4 Lagrangeformalismus

4.1 Lagrange– und Sattelfunktionen

Eine typische Aufgabe ist, daß eine Funktion zweier Veränderlicher $L(x,y)$ gegeben ist und diese Funktion bezüglich einer Variabler (hier x) maximiert und bezüglich der anderen Variablen (hier y) minimiert wird. Dazu kann man die beiden Varianten

$$s^* = \inf_{y\in Y}\sup_{x\in X} L(x,y)\ ,\quad s_* = \sup_{x\in X}\inf_{y\in Y} L(x,y)$$

betrachten und die Frage stellen, wie sich beide Größen s^* und s_* zueinander verhalten.
Es stellt sich heraus, daß stets $s_* \le s^*$ (genannt: **schwaches Dualitätsprinzip**, siehe Punkt 4.1.1 auf Seite 70) und unter speziellen Bedingungen $s_* = s^*$ (genannt: **starkes Dualitätsprinzip**) gilt.
Beide Aufgaben – inf-sup und sup-inf – sind physikalisch verschiedene Aufgaben. Im Fall $s_* < s^*$ haben beide Aufgaben nichts miteinander zu tun. Der Fall $s_* = s^*$ beschreibt den Zusammenhang dieser beiden Aufgaben und ist daher von besonderem Interesse.
Eine zentrale Aufgabe in der Physik ist es, zu einem Problem eine Funktion $L(x,y)$ zu finden, für die $s_* = s^*$ gilt. Die beiden Funktionen

$$f(x) = \inf_{y\in Y} L(x,y)\ ,\quad g(y) = \sup_{x\in X} L(x,y)$$

spielen dann die Rolle von zueinander dualen Funktionen, von denen die eine maximiert und die andere minimiert wird. Im einführenden Beispiel, der Suche nach den Gleichgewichtslagen eines Graphen, gebildet aus Massen und Federn, bestanden die beiden zueinander dualen Aufgaben in der Minimierung zweier potentieller Energien $Q(y)$ und $P(x)$. Der Zusammenhang mit den eben eingeführten Funktionen ist $f(y) = Q(y)$ und $g(x) = -P(x)$. Der Minimierung von $P(x)$ entspricht die Maximierung von $g(x) = -P(x)$.

Im weiteren seien $\mathcal{X}$ und $\mathcal{Y}$ beliebige Mengen und $L : \mathcal{X}\times\mathcal{Y} \longrightarrow \mathbb{R}$ eine beliebige reellwertige Funktion.

Definition: L heißt **Lagrangefunktion** oder **Lagrangian**, falls $s_* = s^*$ gilt. Man sagt auch, daß eine Lagrangefunktion ein Minimax-Theorem erfüllt oder daß für L das starke Dualitätsprinzip gilt.

Es stellt sich heraus, daß das Finden von nichttrivialen Lagrangefunktionen nicht einfach ist. Man kann sagen, daß Lagrangefunktionen selten sind. Das führt dazu, daß jede gefundene Lagrangefunktion interessante physikalische Interpretationen zuläßt. Heute wird eine neue Idee in der theoretischen Physik eigentlich nur dann akzeptiert, wenn ihr eine Lagrangefunktion zugrunde liegt. Viele Artikel in physikalischen Journalen, insbesondere zur Quantenfeldtheorie, beginnen mit dem Satz: “Wir postulieren folgenden Lagrangian.”

Definition: Man sagt, eine Funktion L hat einen **Sattelpunkt** (x_0,y_0), falls

$$L(x,y_0) \le L(x_0,y_0) \le L(x_0,y),\ \forall\ (x,y)\in\mathcal{X}\times\mathcal{Y}$$

gilt. Dabei ist $L(x_0,y_0)$ der **Sattelwert**. Eine Funktion L, die einen Sattelpunkt hat, wird **Sattelfunktion** genannt.

Eine Sattelfunktion kann mehrere Sattelpunkte haben. Es gilt aber der

Satz: Eine Sattelfunktion hat genau einen Sattelwert.

Beweis: Angenommen es sei (x_1, y_1) ein weiterer Sattelpunkt. Dann gelten nach Definition für alle $(x, y) \in \mathcal{X} \times \mathcal{Y}$ die beiden Ungleichungen

$$
\begin{array}{lll}
L(x_0, y) \leq & L(x_0, y_0) & \leq L(x, y_0) \\
L(x_1, y) \leq & L(x_1, y_1) & \leq L(x, y_1)
\end{array}
$$

Sei jetzt o.B.d.A. $L(x_0, y_0) < L(x_1, y_1)$, dann können wir in der ersten Ungleichung $x = x_1$ und in der zweiten $y = y_0$ setzen und erhalten $L(x_0, y_0) \geq L(x_1, y_0) \geq L(x_1, y_1)$ als Widerspruch zur Annahme. □

Satz: Die Sattelpunkte bilden ein Kreuzprodukt. D.h., wenn (x_0, y_0) und (x_1, y_1) Sattelpunkte sind, dann sind auch (x_0, y_1) und (x_1, y_0) Sattelpunkte.

Beweis: Ekland, Temam S.174

4.1.1 Schwaches Dualitätsprinzip

Wir beweisen hier die schon erwähnte, stets geltende Ungleichung $s_* \leq s^*$, die schwaches Dualitätsprinzip genannt wird.

Satz: Es seien $\mathcal{X}$ und $\mathcal{Y}$ beliebige Mengen und $L : \mathcal{X} \times \mathcal{Y} \to \mathbb{R}$ eine beliebige Funktion. dann gilt

$$
\sup_{x \in \mathcal{X}} \inf_{y \in \mathcal{Y}} L(x, y) \leq \inf_{y \in \mathcal{Y}} \sup_{x \in \mathcal{X}} L(x, y)
$$

Beweis: Es seien

$$
f(x) = \inf_{y \in Y} L(x, y) \ , \quad g(y) = \sup_{x \in X} L(x, y)
$$

dann gilt offensichtlich

$$
f(x) \leq L(x, y) \leq g(y), \quad (x, y) \in \mathcal{X} \times \mathcal{Y}
$$

Aus $f(x) \leq g(y)$ folgt natürlich auch

$$
\sup_{x \in X} f(x) \leq \inf_{y \in Y} g(y)
$$

oder

$$
\sup\nolimits_{x \in \mathcal{X}} \ \inf\nolimits_{y \in \mathcal{Y}} L(x, y) \ \leq \ \inf\nolimits_{y \in \mathcal{Y}} \ \sup\nolimits_{x \in \mathcal{X}} L(x, y) \qquad \square
$$

Bemerkung: Entscheident für das Relationszeichen ist die innere Extremwertbildung: Wenn man von einer Größe min und max bildet ist natürlich min $\leq$ max. Die äußere Extremwertbildung verändert diese Eigenschaft dann nicht mehr.

4.1.2 Äquivalenz von Sattelfunktionen und Lagrangefunktionen

Satz: L ist eine Sattelfunktion mit Sattelpunkt (x_0, y_0) und Sattelwert $L(x_0, y_0)$ genau dann, wenn gilt

$$L(x_0, y_0) = \max_{x\in X} \inf_{y\in Y} L(x, y) = \min_{y\in Y} \sup_{x\in X} L(x, y)$$

Beweis: Wir setzen

$$\begin{aligned} g(y) &= \sup_{x\in X} L(x, y) \\ f(x) &= \inf_{y\in Y} L(x, y) \end{aligned}$$

$\Longrightarrow$ Es sei L eine Sattelfunktion, wir beweisen, daß L ein Minimax-Theorem erfüllt.
Es gilt

$$L(x, y_0) \le L(x_0, y_0) \le L(x_0, y), \ \forall\ x, y$$

Hieraus folgt

$$\sup_{x\in X} L(x, y_0) \le L(x_0, y_0) \le \inf_{y\in Y} L(x_0, y)$$

Außerdem gilt (weil das inf nicht größer als ein spezieller Wert sein kann)

$$\begin{aligned} \inf_{y\in Y} \sup_{x\in X} L(x, y) &\le \sup_{x\in X} L(x, y_0) \\ \inf_{y\in Y} L(x_0, y) &\le \sup_{x\in X} \inf_{y\in Y} L(x, y) \end{aligned}$$

was zusammengefaßt

$$\inf_{y\in Y} \sup_{x\in X} L(x, y) \le L(x_0, y_0) \le \sup_{x\in X} \inf_{y\in Y} L(x, y)$$

ergibt. Zusammen mit der schwachen Dualität (umgekehrte Relation) folgt Gleicheit aller Größen

$$\begin{aligned} \inf_{y\in Y} \sup_{x\in X} L(x, y) = \inf_{y\in Y} g(y) = g(y_0) = \sup_{x\in X} L(x, y_0) &= L(x_0, y_0) \\ \sup_{x\in X} \inf_{y\in Y} L(x, y) = \sup_{x\in X} f(x) = f(x_0) = \inf_{y\in Y} L(x_0, y) &= L(x_0, y_0) \end{aligned}$$

Da es Punkte $x_0 \in X$ und $y_0 \in Y$ gibt, bei denen die sup und inf in

$$\inf_{y\in Y} g(y) = g(y_0)\ , \quad \sup_{x\in X} f(x) = f(x_0)$$

angenommen werden, kann an ihre Stelle max bzw. min geschrieben werden, was

$$\min_{y\in Y} \sup_{x\in X} L(x, y) = g(y_0) = L(x_0, y_0) = f(x_0) = \max_{x\in X} \inf_{y\in Y} L(x, y)$$

ergibt.

$\Longleftarrow$ Es sei

$$L := \max_{x\in X} \inf_{y\in Y} L(x, y) = \min_{y\in Y} \sup_{x\in X} L(x, y)$$

Die max und min in der Gleichheit bedeuten, daß es Punkte x_0 und y_0 gibt mit

$$L \ = \ f(x_0) = \max_{x\in X} f(x) = g(y_0) = \min_{y\in Y} g(y)$$

Wir beweisen, daß (x_0, y_0) Sattelpunkt von $L(x, y)$ ist und $L(x_0, y_0) = L$ gilt.

Wir berechnen $L(x_0, y_0)$. Es gilt

$$\begin{aligned} L &= f(x_0) = \inf_{y\in Y} L(x_0, y) \leq L(x_0, y_0) \\ L &= g(y_0) = \sup_{x\in X} L(x, y_0) \geq L(x_0, y_0) \end{aligned}$$

Da Gleichheit gilt, gilt $L = L(x_0, y_0)$ und

$$\inf_{y\in Y} L(x_0, y) = L(x_0, y_0) = \sup_{x\in X} L(x, y_0)$$

oder

$$\sup_{x\in X} L(x, y_0) = L(x_0, y_0) = \inf_{y\in Y} L(x_0, y)$$

Zusammen mit

$$\begin{aligned} L(x, y_0) &\leq \sup_{x\in X} L(x, y_0) \;, && \forall\, x \\ \inf_{y\in Y} L(x_0, y) &\leq L(x_0, y) && , \forall\, y \end{aligned}$$

folgt $L(x, y_0) \leq L(x_0, y_0) \leq L(x_0, y), \;\forall\, x, y$. □

4.1.3 Zusammenfassung der Ungleichungen und Gleichungen

Aus Sicht der Lagrangefunktion haben wir

$$\begin{array}{rcccccl} f(x) &=& \inf_y L(x,y) &\leq L(x,y) \leq& \sup_x L(x,y) &=& g(y) \\ f(x_0) = \sup_x f(x) &=& \sup_x \inf_y L(x,y) &= L(x_0,y_0) =& \inf_y \sup_x L(x,y) &=& \inf_y g(y) = g(y_0) \end{array}$$

Aus Sicht der Sattelfunktion haben wir

$$f(x) \leq L(x, y_0) \leq f(x_0) = L(x_0, y_0) = g(y_0) \leq L(x_0, y) \leq g(y)$$

4.1.4 Beweisvarianten für Sattelprobleme

Zum Beweis, daß eine Funktion L Sattel- bzw. Lagrangefunktion ist bietet der letzte Satz zwei Möglichkeiten:

- Beweis des Minimax-Theorems:

  Man berechnet zwei Funktionen

  $$g(y) = \sup_x L(x, y) \;, \quad f(x) = \inf_y L(x, y)$$

  und beweist

  $$\inf_y g(y) = \sup_x f(x)$$

  Sollten die Extremalwerte hier in Punkten x_0 und y_0 angenommen werden, dann ist (x_0, y_0) ein Sattelpunkt von L.

- Beweis der Satteleigenschaft:

  Man "errät" den Sattelpunkt (x_0, y_0) (z.B. durch Lösung der Gleichungen $\partial_x L = 0$, $\partial_y L = 0$) und beweist

  $$L(x, y_0) \leq L(x_0, y_0) \leq L(x_0, y)$$

  Diese Methode ist oft viel einfacher, erfordert aber die Kenntnis des Sattelpunktes.

4.2 Existenzbedingungen für Sattelpunkte und Extremwerte

Wie immer ist die Existenz eines Extremwertes eine viel schächere Bedingung als die Existenz eines Punktes in dem der Extremwert angenommen wird. Man will vom Verhalten eines Funktionals auf die Existenz von Punkten schließen.
Wir werden im weiteren voraussetzen, daß für eine uns interessierende Lagrangefunktion ein Sattelpunkt existiert, falls wir ihn brauchen.
Der Vollständigkeit halber führen wir hier ohne Beweis einige solcher Existenzsätze an. Die Idee der Beweise besteht im allgemeinen im folgenden: Man konstruiert eine Folge, die gegen den Extremwert konvergiert. Hier reicht meist die schwache Konvergenz. Die konvergente Folge wird aus einer beschränkten Folge konstruiert. Hierzu ist Kompaktheit und Reflexivität Voraussetzung.

4.2.1 Existenzsätze für Extremalaufgaben

Im Allgemeinen nehmen stetige Funktionen auf nicht-kompakten Mengen kein Minimum oder Maximum an. Da wir hier häufig Funktionen auf Banachräumen (die natürlich nicht kompakt sind) betrachten, sind Aussagen sinnvoll, die die Existenz von Extremwerten in nicht kompakten Situationen implizieren. Intuitiv ist klar, daß man für die Existenz eines Minimums die Nicht-Kompaktheit durch schnelles Wachstum außerhalb kompakter Mengen kompensieren kann. Dazu dient folgende

Definition: Eine Funktion $F : \mathcal{X} \to \mathbb{R}$ heißt **koerzitiv**, falls für jede Folge x_n mit $\lim_{n\to\infty} \|x_n\| = \infty$ gilt: $\lim_{n\to\infty} F(x_n) = \infty$.

Dazu gelten folgende Sätze:

Satz: Es sei $\mathcal{X}$ reflexiv und $F : \mathcal{X} \to \mathbb{R}$ stetig, konvex und koerzitiv. Dann nimmt F auf $\mathcal{X}$ sein Minimum an.

Satz: Es sei $\mathcal{X}$ reflexiv und $F : \mathcal{X} \to \mathbb{R}$ schwach unterhalbstetig und koerzitiv. Dann nimmt F auf $\mathcal{X}$ sein Minimum an.
Hier heißt schwach unterhalbstetig, daß für alle schwach konvergenten Folgen $x_n \to x$
$\liminf_{n\to\infty} f(x_n) \geq f(x)$ gilt.

4.2.2 Existenzsätze für Sattelpunkte

Die grundlegenden Voraussetzungen für eine Funktion L Sattelfunktion zu sein sind folgende:
Es seien $\mathcal{X}$ und $\mathcal{Y}$ reflexive Banachräume, $A \subset \mathcal{X}$ und $B \subset \mathcal{Y}$ Teilmengen und $L : A \times B \to \mathbb{R}$.
Im weiteren gelten folgende Voraussetzungen:

- A nichtleer, konvex, abgeschlossen
- B nichtleer, konvex, abgeschlossen
- Für feste $y \in B$ sei $-L(\cdot, y)$ konvex und unterhalbstetig
- Für feste $x \in A$ sei $L(x, \cdot)$ konvex und unterhalbstetig

Satz: Es seien A und B beschränkt. Dann besitzt L einen Sattelpunkt.

Satz: Es sei L als Funktion jeder Veränderlicher koerzitiv. Dann besitzt L einen Sattelpunkt.

Der klassische Satz zur Existenz von Sattelpunkten ist der folgende
Satz (vonNeumann): Es seien $\mathcal{X}$ und $\mathcal{Y}$ reflexive Banachräume, $A \subset \mathcal{X}$ und $B \subset \mathcal{Y}$ abgeschlossen, beschränkt und konvex. Es sei

$$\begin{array}{lll} x & \to & L(x,y) \quad \text{konkav und stetig } \forall\, y \in B \\ y & \to & L(x,y) \quad \text{konvex und stetig } \forall\, x \in A \end{array}$$

dann existiert ein Punkt (x_0, y_0) mit

$$\sup_{x\in A} \inf_{y\in B} L(x,y) = \inf_{y\in B} \sup_{x\in A} L(x,y) = L(x_0,y_0)$$

Die Bedingungen an die Räume und die Funktion L lassen sich weiter abschwächen. Hierzu gibt es viele alternative Sätze.

4.2.3 Differenzierbare Sattelfunktionen

Es sei L für feste Variable als Funktion der anderen Variablen differenzierbar (es existiere die Gateauxableitung).
Satz: Ein Punkt (x_0, y_0) ist Sattelpunkt von L genau dann, wenn

$$\begin{array}{rcl} \left\langle \dfrac{\partial}{\partial x} L(x,y_0)\Big|_{x=x_0}, \quad x - x_0 \right\rangle & \leq & 0, \ \forall x \in A \\ \left\langle \dfrac{\partial}{\partial y} L(x_0,y)\Big|_{y=y_0}, \quad y - y_0 \right\rangle & \geq & 0, \ \forall y \in B \end{array}$$

4.2.4 Notwendige Sattelbedingungen

Analog zum Verschwinden der Ableitungen läßt sich zeigen, daß eine notwendige Bedingung für einen Sattelpunkt das Verschwinden beider Ableitungen

$$\frac{\partial}{\partial x} L(x,y_0)\Big|_{x=x_0} = 0, \quad \frac{\partial}{\partial y} L(x_0,y)\Big|_{y=y_0} = 0$$

ist. Diese Eigenschaft beschreibt eine horizontale Tangentialebene. Es ist natürlich nur eine notwendige Bedingung, denn eine Funktion zweier Varabler hat eine horizontale Tangentialebene auch in einem Maximum oder Minimum.
Es folgt eine Übersicht über Minima, Maxima und Sattelpunkte. Es sei $x \in \mathcal{X}$, $y \in \mathcal{Y}$

	global	lokal
Minimum	$g(y) \geq g(y_0)$	$\frac{d}{dy} g(y)\Big\|_{y=y_0} = 0$
Maximum	$f(x_0) \geq f(x)$	$\frac{d}{dx} f(x)\Big\|_{x=x_0} = 0$
Sattelpunkt	$L(x,y_0) \leq L(x_0,y_0) \leq L(x_0,y)$	$\frac{\partial}{\partial x} L(x,y_0)\Big\|_{x=x_0} = 0, \quad \frac{\partial}{\partial y} L(x_0,y)\Big\|_{y=y_0} = 0$

4.3 Beispiele für Lagrangefunktionen

4.3.1 Ein quadratisches Funktional

Es sei

$$L(x,y) = \frac{1}{2}\langle y, \mathbf{M}y\rangle + \langle y, \mathbf{D}x\rangle - \frac{1}{2}\langle x, \mathbf{B}x\rangle$$

mit positiven Operatoren $\mathbf{M}$ und $\mathbf{B}$. Wir untersuchen, ob es sich bei L um eine Lagrange- oder Sattelfunktion handelt.
Falls es eine Sattelfunktion ist, kann man den Sattelpunkt mit dem notwendigen Kriterium – beide partiellen Abletungen verschwinden – berechnen:

$$\begin{aligned}
\frac{\partial}{\partial x} L(x,y) &= -\mathbf{B}x + \mathbf{D}^* y = 0 \\
\frac{\partial}{\partial y} L(x,y) &= +\mathbf{M}x + \mathbf{D}x = 0
\end{aligned}$$

Offensichtlich ist $(x_0, y_0) = (0,0)$ Lösung des Systems. Dieser Punkt ist Sattelpunkt, falls die Ungleichungskette

$$\begin{array}{ccccc}
L(x,0) & \leq & L(0,0) & \leq & L(0,y) \\
-\frac{1}{2}\langle x, \mathbf{B}x\rangle & \leq & 0 & \leq & \frac{1}{2}\langle y, \mathbf{M}y\rangle
\end{array}$$

erfüllt ist, was offensichtlich der Fall ist (weil $\mathbf{B}, \mathbf{M} > 0$).
Für die Frage, ob es sich um eine Lagrangefunktion handelt, berechnen wir f und g:

$$\begin{aligned}
f(x) &= \inf_{y \in Y} L(x,y) = \inf_{y \in Y} \left(\frac{1}{2}\langle y, \mathbf{M}y\rangle + \langle y, \mathbf{D}x\rangle - \frac{1}{2}\langle x, \mathbf{B}x\rangle \right) = \\
&= -\frac{1}{2}\langle x, \mathbf{B}x\rangle - \sup_{y \in Y} \left(\langle y, (-\mathbf{D}x)\rangle - \frac{1}{2}\langle y, \mathbf{M}y\rangle \right) = \\
&= -\frac{1}{2}\langle x, \mathbf{B}x\rangle - \frac{1}{2}\langle \mathbf{M}^{-1}\mathbf{D}x, \mathbf{D}x\rangle = -\frac{1}{2}\big\langle x, \big(\mathbf{B} + \mathbf{D}^*\mathbf{M}^{-1}\mathbf{D}\big)x\big\rangle \\
g(y) &= \sup_{x \in X} L(x,y) = \sup_{x \in X} \left(\frac{1}{2}\langle y, \mathbf{M}y\rangle + \langle y, \mathbf{D}x\rangle - \frac{1}{2}\langle x, \mathbf{B}x\rangle \right) = \\
&= \frac{1}{2}\langle y, \mathbf{M}y\rangle + \sup_{x \in X} \left(\langle \mathbf{D}^* y, x\rangle - \frac{1}{2}\langle x, \mathbf{B}x\rangle \right) = \\
&= \frac{1}{2}\langle y, \mathbf{M}y\rangle + \frac{1}{2}\langle \mathbf{D}^* y, \mathbf{B}^{-1}\mathbf{D}^* y\rangle = \frac{1}{2}\big\langle y, \big(\mathbf{M} + \mathbf{D}\mathbf{B}^{-1}\mathbf{D}^*\big)y\big\rangle
\end{aligned}$$

Da die Operatoren $\mathbf{B} + \mathbf{D}^*\mathbf{M}^{-1}\mathbf{D}$ und $\mathbf{M} + \mathbf{D}\mathbf{B}^{-1}\mathbf{D}^*$ positiv sind, folgt $\sup_{x \in X} f(x) = 0$ und $\inf_{y \in Y} g(y) = 0$. L ist also auch Lagrangefunktion.

4.3.2 Kinetische minus potentielle Energie

Beispiel 1a: Es seien $F : \mathcal{X} = \mathbb{R} \to \mathbb{R}$ und $G : \mathcal{Y} = \mathbb{R} \to \mathbb{R}$ Funktionen, die ein globale Minima in x_0 bzw. y_0 haben, Es sei also $F(x_0) = \min_x F(x)$ und $G(y_0) = \min_y G(y)$. Dann ist

$$L(x,y) = G(y) - F(x)$$

Lagrangefunktion, denn es gilt

$$\begin{aligned}
f(x) &= \inf_{y\in Y} L(x,y) = \inf_{y\in Y}\big(G(y)-F(x)\big) = -F(x) + \inf_{y\in Y} G(y) = G(y_0) - F(x) \\
s_* &= \sup_{x\in X} f(x) = \sup_{x\in X}\big(G(y_0)-F(x)\big) = G(y_0) - \inf_{x\in X} F(x) = G(y_0) - F(x_0) \\
g(y) &= \sup_{x\in X} L(x,y) = \sup_{x\in X}\big(G(y)-F(x)\big) G(y) - \inf_{x\in X} F(x) = G(y) - F(x_0) \\
s^* &= \inf_{y\in Y} g(y) = \inf_{y\in Y}\big(G(y)-F(x_0)\big) = -F(x_0) + \inf_{y\in Y} G(y) = G(y_0) - F(x_0) = s_*
\end{aligned}$$

Lagrangefunktionen vom Typ $L(x,y) = G(y) - F(x)$ spielen in der Physik eine wichtige Rolle. Sie beschreiben die Dynamik von klassischen Systemen. Dabei hängt $G(y)$ mit der kinetischen und $F(x)$ mit der potentiellen Energie zusammen.

Beispiel 1b: Es seien $F : \mathcal{X} = \mathbb{R} \longrightarrow \mathbb{R}$ und $G : \mathcal{Y} = \mathbb{R} \longrightarrow \mathbb{R}$ Funktionen, die ein globale Minima in x_0 bzw. y_0 haben, Es sei also $F(x_0) = \min_x F(x)$ und $G(y_0) = \min_y G(y)$. Dann ist (x_0, y_0) ein Sattelpunkt von L und $L(x_0,y_0) = G(y_0) - F(x_0)$, denn aus $F(x_0) \le F(x)$ und $G(y_0) \le G(y)$ folgt offensichtlich

$$G(y_0) - F(x) \le F(x_0) - G(y_0) \le G(y) - F(x_0)$$

In diesem Beispiel gilt

$$\begin{aligned}
f(x) &= \inf_{y\in Y} L(x,y) = G(y_0) - F(x) = L(x,y_0) \\
g(y) &= \sup_{x\in X} L(x,y) = G(y) - F(x_0) = L(x_0,y)
\end{aligned}$$

Im allgemeinen gilt nur

$$\begin{aligned}
f(x) &= \inf_{y\in Y} L(x,y) \le L(x,y_0) \le L(x_0,y_0) \\
g(y) &= \sup_{x\in X} L(x,y) \ge L(x_0,y) \ge L(x_0,y_0)
\end{aligned}$$

wogegen

$$\begin{aligned}
\sup_{x\in X} f(x) &= \sup_{x\in X}\inf_{y\in Y} L(x,y) = L(x_0,y_0) \\
\inf_{y\in Y} g(y) &= \inf_{y\in Y}\sup_{x\in X} L(x,y) = L(x_0,y_0)
\end{aligned}$$

gilt, falls die entsprechenden Werte x_0 und y_0 existieren.

4.3.3 Beispiel 2

Beispiel 2: Es sei

$$A(y) \le B(x,y) \le C(x)$$

mit den Eigenschaften, daß ein Punkt (x_0, y_0) existiert, für den

$$\begin{aligned}
A(y_0) - A(y) &\ge B(x_0,y_0) - B(x_0,y) \\
C(x) - C(x_0) &\ge B(x,y_0) - B(x_0,y_0)
\end{aligned}$$

gilt. Dann ist

$$L(x,y) = A(y) + C(x) - B(x,y)$$

Lagrangefunktion mit dem Sattelpunkt (x_0, y_0) denn es gilt

$$\begin{aligned} L(x,y_0) - L(x_0,y_0) &= A(y_0) + C(x) - B(x,y_0) - \big(A(y_0) + C(x_0) - B(x_0,y_0)\big) = \\ &= C(x) - C(x_0) + B(x_0,y_0) - B(x,y_0) \geq 0 \end{aligned}$$

und analog $L(x_0, y) \leq L(x_0, y_0)$.

4.3.4 Beispiel 3. Störungssatz

Beispiel 3: Es sei $L_0(x, y)$ Lagrangefunktion mit dem Sattelpunkt (x_0, y_0) und $f(x)$ und $g(y)$ definiert wie oben. Dann ist

$$\begin{aligned} L(x,y) &= f(x) + g(y) - L_0(x,y) = \\ &= \sup_{y\in Y} L_0(x,y) + \inf_{x\in X} L_0(x,y) - L_0(x,y) \end{aligned}$$

ebenfalls Lagrangefunktion mit dem Sattelpunkt (x_0, y_0) denn es gilt $g(y) \leq L(x_0, y)$ und $L(x, y_0) \leq f(x)$, woraus

$$\begin{aligned} g(y_0) - g(y) &\geq L(x_0,y_0) - L(x_0,y) \\ f(x) - f(x_0) &\geq L(x,y_0) - L(x_0,y_0) \end{aligned}$$

folgt. Damit ist dieses Beispiel auf Beispiel 2 zurückgeführt.
Im allgemeinen gilt $L(x, y) \not\equiv L_0(x, y)$. Gleichheit gilt nur für $L(x, y) = \frac{1}{2}f(x) + \frac{1}{2}g(y)$.
ÜA: Finde $L(x, y)$, wenn

$$L_0(x,y) = \frac{1}{2}\langle y, \mathbf{M}y\rangle + \langle y, \mathbf{D}x\rangle - \frac{1}{2}\langle x, \mathbf{B}x\rangle$$

4.3.5 Beispiel 4. Produkte

Unter einer Sattelfunktion stellt man sich als erstes $L(x, y) = y^2 - x^2$ vor. Das Bild ist ein typischer Sattel. Der Übergang in eine anderes Koordinatensystem führt zu $L(x, y) = y^2 - x^2 = (y - x)(y + x) = st =: L(s, t)$ läßt natürlich das Bild unverändert. $L(x, y) = xy$ ist also auch eine Sattelfunktion. Das läßt sich verallgemeinern.
Beispiel 4: Es seien $F(x)$ und $G(y)$ Funktionen mit Nullstellen und den Wachstumsbedingungen $\sup_x F(x) = +\infty$, $\inf_x F(x) = -\infty$, $\sup_y G(y) = +\infty$, $\inf_y G(y) = -\infty$. (Alternativ kann man Stetigkeit voraussetzen, dann folgt aus den Wachstumsbedingungen die Existenz von Nullstellen.)
Dann ist $L(x, y) = F(x) \cdot G(y)$ eine Lagrangefunktion.
Beweis: Wir beweisen das minimax-Theorem. Nach Voraussetzung sind die Mengen $G^{-1}(0)$ und $F^{-1}(0)$ nicht leer. Dann folgt

$$\begin{aligned} \sup_x F(x)\cdot G(y) &= \chi_{G^{-1}(0)} \\ \inf_y \sup_x F(x)\cdot G(y) &= \inf_y \chi_{G^{-1}(0)} = 0 \\ \inf_y F(x)\cdot G(y) &= -\chi_{F^{-1}(0)} \\ \sup_x \inf_x F(x)\cdot G(y) &= \sup_x(-\chi_{F^{-1}(0)}) = -\inf_x \chi_{F^{-1}(0)} = 0 \end{aligned}$$

Für die Sattelpunkte gilt

$$(x_0, y_0) \in F^{-1}(0) \times G^{-1}(0)$$

4.4 Spieltheorie

4.4.1 Lagrangefunktionen in der Spieltheorie

Besonders deutlich wird die Rolle von Lagrangefunktionen in der Spieltheorie. Hier treffen zwei – ähnlich wie in vielen ökonomischen Problemen – zwei Parteien mit unterschiedlichen Interessen aufeinander. In der Physik ist das genauso, nur kann sich der Mensch das Wirken "unterschiedlicher Interessen" schlecht vorstellen. Aber hier gilt wie immer in der Naturwissenschaft: Unsere Wahrnehmung verschafft uns zwar ein objektives Bild der Natur, aber was für Fragen wir ihr stellen ist subjektiv.

Es sei $\mathcal{X} = \{1, ..., n\}$ und $\mathcal{Y} = \{1, ..., m\}$. Dann läßt sich eine allgemeine Funktion $L : \mathcal{X} \times \mathcal{Y} \to \mathbb{R}$ als Matrix $\mathbf{A} = (a_{ij})$, $i = 1, ..., m$, $j = 1, ..., n$ darstellen.
Gegeben sei eine Matrix $L(i, j) = \mathbf{A} = (a_{ij})$, $i = 1, ..., m$, $j = 1, ..., n$, die die möglichen Züge zweier Spieler beschreibt. Spieler A kann einen Wert in der i-ten Zeile bestimmen, Spieler B einen Wert in der j-ten Spalte. Die auf diese Weise festgelegte Zahl a_{ij} beschreibt den Betrag, den B an A zahlen muß. Diese scheinbare Asymmetrie legt nur fest, bezüglich welcher Variablen min oder max genommen wird.
Die Zahlen können negativ sein, dann bezeichne sie "Schulden", die B an A zahlen, d.h. ihm aufbürden, muß, was gleichbedeutend mit einem Betrag ist, den A an B zahlen muß.

$$L(i, j) = \mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} & \cdots & a_{1n} \\ a_{21} & a_{22} & a_{23} & a_{24} & \cdots & a_{2n} \\ a_{31} & a_{32} & a_{33} & a_{34} & \cdots & a_{3n} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & a_{m3} & a_{m4} & \cdots & a_{mn} \end{pmatrix}$$

Klar, daß B wenig zahlen und A viel haben will. Wir untersuchen, ob das Spiel "fair" ist. Dieser Begriff hat nichts mit Gerechtigkeit oder ausgelichenem Gewinn zu tun sondern bedeutet, daß es gleichgültig ist, wer das Spiel beginnt. Will man "Gerechtigkeit" oder "Gewinnausgleich" erziehlen, muß man das über die Zahlen in der Matrix regeln. So könnte das Spiel häufig stattfinden und die Werte a_{ij} um den Mittelwert 0 herum gestreut sein. Das interessiert uns hier nicht. Es geht nur darum, ob das Spiel fair ist.
Jedes endliche Spiel kann man durch so eine Matrix darstellen. Ein typisches unfaires Spiel ist das Spiel "Schere, Stein, Papier". Hier verliert der, der anfängt. Deshalb fangen in der Realität auch beide Spieler gleichzeitig an.

1) Angenommen A beginnt:
Wenn er eine Zeile i, also $(a_{i1}, ..., a_{in})$ festlegt, wird B in dieser Zeile das Minimum wählen, also $\min_j a_{ij}$. Da A das ahnt, wird es also jenes i wählen, bei dem $\min_j a_{ij}$ maximal wird. Das ist

$$J_A = \max_i \min_j a_{ij} \ .$$

2) Angenommen B beginnt:
Wenn er eine Spalte j, also $(a_{1j}, ..., a_{mj})^*$ festlegt, wird A in dieser Spalte das Maximum wählen, also $\max_i a_{ij}$. Da B das ahnt, wird es also jenes j wählen, bei dem $\max_i a_{ij}$ maximal wird. Das ist

$$J_B = \min_j \max_i a_{ij} \ .$$

Übersichtlich dargestellt, werden folgende Operationen ausgeführt:

$$\begin{array}{ccc}
\begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} & \cdots & a_{1n} \\ a_{21} & a_{22} & a_{23} & a_{24} & \cdots & a_{2n} \\ a_{31} & a_{32} & a_{33} & a_{34} & \cdots & a_{3n} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & a_{m3} & a_{m4} & \cdots & a_{mn} \end{pmatrix} & \overset{\max}{\Longrightarrow} & \begin{pmatrix} M_1 \\ M_2 \\ M_3 \\ \vdots \\ M_m \end{pmatrix} \\
\Big\Downarrow \min & & \Big\Downarrow \min \\
\begin{pmatrix} m_1 & m_2 & m_3 & m_4 & \cdots & m_n \end{pmatrix} & \overset{\max}{\Longrightarrow} & J_A \overset{?}{=} J_B
\end{array}$$

Wie verhalten sich J_A und J_B zueinander? Wie immer gilt schwache Dualität, also $J_B \geq J_A$. Fair ist das Spiel, wenn $J_A = J_B$, dann ist egal, wer anfängt. Ob das möglich ist, hängt von der Matrix $\mathbf{A}$ ab. Der mit einem fairen Spiel zusammenhängende Sattelwert heißt Nash-Gleichgewicht (nach John Nash, Nobel- und Abelpreisträger).

Das einfachste nichttriviale Beispiel: Wenn $n = 1$ oder $m = 1$, ist das Spiel immer fair. Es gibt keine Wahlmöglichkeit.
Es sei $\mathcal{X} = \mathcal{Y} = \{1, 2\}$. Dann läßt sich eine allgemeine Funktion $L : \mathcal{X} \times \mathcal{Y} \longrightarrow \mathbb{R}$ als Matrix $L = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ darstellen. Wir vergleichen

$$\min_{\text{Spalte}} \max_{\text{Zeile}} \begin{pmatrix} a & b \\ c & d \end{pmatrix} \leq \max_{\text{Zeile}} \min_{\text{Spalte}} \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

ÜA: Wann ist L eine Lagrangefunktion? **Lösung:** Es sei

$$\begin{aligned} J_B = s^* &= \min\big\{\max\{a, b\}, \max\{c, d\}\big\} \\ J_A = s_* &= \max\big\{\min\{a, c\}, \min\{b, d\}\big\} \end{aligned}$$

Max und Min sind invariant bezüglich der Permutation der Zeilen bzw. Spalten. Wir können deshalb die Zeilen und Spalten derart umordnen, daß a das kleinste Element der Matrix ist. Das führt auf folgende Möglichkeiten:

$$\begin{aligned}
a < b < c < d &\implies s^* = c,\ s_* = c \quad ! \\
a < b < d < c &\implies s^* = d,\ s_* = d \quad ! \\
a < c < b < d &\implies s^* = c,\ s_* = c \quad ! \\
a < c < d < b &\implies s^* = c,\ s_* = c \quad ! \\
a < d < b < c &\implies s^* = d,\ s_* = b \\
a < d < c < b &\implies s^* = d,\ s_* = c
\end{aligned}$$

Damit definieren 2/3 aller willkürlich gewählten 2×2-Matrizen ein faires Spiel.
Bemerkung: Die Eigenschaft, eine Lagrangefunktion zu sein, ändert sich für eine Matrix nicht, wenn die Spalten und/oder Zeilen permutiert werden. Das sieht man daran, daß die max- bzw min-Bildung permutationsinvariant ist.
Ein andere Grund für diese Einsicht ist, daß eine Permutation der Spalten und/oder Zeilen eine Umnummerirung der Zustände bedeutet.
Das ist der Grund, warum man einen Eintrag in der Matrix festsetzen kann. Im Beispiel eben, wurde a als kleinstes Element gesetzt. Deshalb reichte es, 6 Ungleichungsketten zu betrachten (anstelle von 24).

ÜA: Ist folgende Aussage richtig? Es sei L eine $n \times m$ Matrix mit Einträgen aus $[0,1]$. Der Anteil der Matrizen, die das Minimax-Theorem erfüllen ist $m!n!/(n+m-1)!$.

- Es gibt einfache Fälle, in denen $L = (a_{ij})$ Lagr.funktion ist:

  Wenn es eine Zeile gibt, in der alle Einträge größer sind als alle Einträge in der Matrix, dann ist das Max in jeder Spalte aus dieser Zeile und der Sattelwert das kleinste Element in dieser Zeile.

  Umgekehrt ist das kleinste Element in dieser Zeile größer als alle anderen kleinsten Zeilenelemente und damit das größte aller kleinsten Elemente.

  Analog ist es, wenn es eine Spalte gibt, in der alle Einträge kleiner sind als alle Einträge in der Matrix.

- Wenn es dagegen eine Zeile gibt, in der alle Einträge kleiner sind als alle Einträge in der Matrix, dann kann man daraus nichts schließen.

4.4.2 Gemischte Spielstrategien (stochastisches Spiel)

Gegeben sei wieder eine Matrix $\mathbf{A} = (a_{ij})$, $i = 1, ..., m$, $j = 1, ..., n$, die die möglichen Züge zweier Spieler beschreibt. Diesmal muß Spieler A aber nicht einen definierten Wert in der i-ten Zeile bestimmen, sondern kann eine konvexe Kombination der Werte der i-ten Zeile bestimmen ($=$ Wahrscheinlichkeitsmaß auf $\mathcal{X}$). Spieler A gibt also einen Vektor $(x_1, ..., x_m)$ aus dem Simplex $\mathcal{P}_m \subset \mathbb{R}^*_m$ an (das sind Vektoren mit nichtnegativen Komponenten deren Summe 1 ist). Das gleiche gilt für Spieler B und der j-ten Spalte. Er muß einen Vektor $(y_1, ..., y_n)$ aus dem Simplex $\mathcal{P}_n \subset \mathbb{R}^*_n$ bestimmen. Auf diese Weise wird ein Wert $\sum_{i,j=1}^{m,n} a_{ij} x_i y_j$ festgelegt. Das ist der Betrag, den A an B zahlen muß. Das führt auf die beiden Ergebnisse (in Abhängigkeit davon, wer anfängt)

$$
\begin{aligned}
J_A &= \min_{x \in \mathcal{P}_m} \max_{y \in \mathcal{P}_n} \sum_{i,j=1}^{m,n} a_{ij} x_i y_j = \min_{x \in \mathcal{P}(\mathcal{Y})} \max_{y \in \mathcal{P}(\mathcal{X})} \langle x, \mathbf{A} y \rangle \\
J_B &= \max_{y \in \mathcal{P}_n} \min_{x \in \mathcal{P}_m} \sum_{i,j=1}^{m,n} a_{ij} x_i y_j = \max_{y \in \mathcal{P}(\mathcal{X})} \min_{x \in \mathcal{P}(\mathcal{Y})} \langle x, \mathbf{A} y \rangle
\end{aligned}
$$

Interessanterweise gilt für beliebige Matrizen $\mathbf{A}$ stets $J_A = J_B$. Die Funktion $L(x,y) = \langle x, \mathbf{A} y \rangle$ enthält stets einen Sattelpunkt. Das folgt aus dem klassischen Satz von vonNeumann.
Das deterministische Spiel ist ein Spezialfall eines stochastischen Spiels. Es gilt

$$
\min_{x \in \partial_e \mathcal{P}_n} \max_{y \in \partial_e \mathcal{P}_m} \langle x, \mathbf{A} y \rangle = \min_i \max_j \langle e^*_i, \mathbf{A} e^*_j \rangle = \min_i \max_j \langle e_i, \mathbf{A} e^*_j \rangle = \min_i \max_j a_{ij}
$$

Die Möglichkeit, nicht nur reine Zustände, sondern auch gemischte Zustände auszuwählen, regularisiert das Spiel.
Der Übergang von einem deterministischen zu einem stochastischen Problem macht also jedes Problem zu einem Sattelproblem. Im letzten Semester hatten wir festgestellt, daß derselbe Übergang jedes allgemeine Problem zu einem konvexen, kompakten und stetigen Problem macht.

Bemerkung: Der häufig verwendet Begriff “stochstisches Spiel” ist irreführend. Im Spiel findet kein Zufall statt. Es besteht nur die Möglichkeit, eine konvexe Kombination der Zustände auszuwählen, was im allgemeinen gerade durch ein Wahrscheinlichkeitsmaß realisiert wird.

4.4.3 Sattelpunkt bei entarteter Matrix

Angenommen, die Matrix $\mathbf{A}$ hat einen Eigenwert 0, zu dem zwei Eigenvektoren $x_0 \in \mathcal{P}_m$ und $y_0 \in \mathcal{P}_n$ gehören. Es sei also $\mathbf{A}y_0 = 0$ und $\mathbf{A}^* x_0 = 0$. Dann ist (x_0, y_0) ein Sattelpunkt von $L(x,y) = \langle x, \mathbf{A}y \rangle$.
Beweis: Aus $\mathbf{A}y_0 = 0$ folgt $\langle x, \mathbf{A}y_0 \rangle = 0$ für alle x und analog gilt $\langle \mathbf{A}^* x_0, y \rangle = \langle x_0, \mathbf{A}y \rangle = 0$ für alle y. Also gilt auch $\langle x_0, \mathbf{A}y_0 \rangle = 0$ und damit $\langle x_0, \mathbf{A}y \rangle \leq \langle x_0, \mathbf{A}y_0 \rangle \leq \langle x, \mathbf{A}y_0 \rangle$.

Wichtig ist hier, daß die Eigenvektoren in den Simplizes liegen. Das ist für Matrizen (und entsprechende beschränkte Operatoren im unendlichdim. Fall) stets erfüllt, wenn die Nebendiagonalelemente alle nichtnegativ (oder nichtpositiv) sind. Es lassen sich dann nämlich offensichtlich für jede Matrix $\mathbf{A}$ positive reelle Zahlen a und b derart finden, daß die Matrix $\mathbf{M} = b\mathbf{I} - a\mathbf{A}$ aus nichtnegativen Elementen besteht. Für solche Matrizen ist bekannt (Satz von Frobenius-Peron), daß sie einen positiven Eigenvektor haben, wenn b ihr Spektralradius ist. Dieser positiven Eigenvektor ist – geeignet skaliert – der gesuchte Eigenvektor zum Eigenwert 0 von $\mathbf{A}$.

Entartete Matrizen sind typisch für Nullsummenspiele, bei denen alle Zeilen- und Spaltensummen 0 sind. Dann ist der konstante Vektor Eigenvektor zur 0.

4.4.4 "Schere, Stein, Papier"

Eine geeignete Matrix $\mathbf{A}$ für dieses Spiel ist

$$\mathbf{A} = \begin{pmatrix} 0 & -1 & 1 \\ 1 & 0 & -1 \\ -1 & 1 & 0 \end{pmatrix}$$

Offensichtlich ist das Spiel unfair: $\max_i \min_j a_{ij} = -1 < 1 = \min_j \max_i a_{ij}$
Wir finden eine gemischte Strategie: Dazu betrachten wir

$$L(x,y) = \langle x, \mathbf{A}y \rangle = x_1 y_2 - x_2 y_1 + x_2 y_3 - x_3 y_2 + x_3 y_1 - x_1 y_3 = \begin{vmatrix} 1 & 1 & 1 \\ x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \end{vmatrix}$$

Man sieht leicht, daß $x_0 = y_0 = (1/3, 1/3, 1/3)$ Sattelpunkt zum Sattelwert $L(x_0, y_0) = 0$ ist. Dieser Sattelpunkt ist einzig. Das sieht man daran, daß $L(x,y)$ "zufällig" der doppelte Flächeninhalt eines Dreiecks mit den Punkten $P_i = (x_i, y_i)$ ist. Der ist genau dann $= 0$ (der Sattelwert ist eindeutig!), wenn die Punkte zusammenfallen.

4.5 Bedingte Extremalaufgaben (Lagrangemethode)

4.5.1 Lagrangefunktionen und Extremalaufgaben

Aus der Analysis ist gut bekannt, daß man bedingte Extremalaufgaben mit der Lagrangemethode – der Einführung von Lagrange-Multiplikatoren – lösen kann. Wir beweisen hier diese Methode mit Mitteln der konvexen Analysis für allgemeine nicht-glatte Funktionen.
Die Idee der Lagrangemethode besteht darin, anstelle eines Extremalproblems auf einer Teilmenge, ein spezielles Sattelproblem auf dem ganzen Raum zu betrachten.
Die Richtigkeit dieses Zusammenhangs erklärt die Merkwürdigkeit, daß anstelle eines n-dimensionalen Extremalproblems mit $k < n$ Nebenbedingungen ein $(n+k)$-dimensionales Problem betrachtet wird, obwohl man natürlicherweise annimmt, daß sich die Dimension – also die Anzahl der Freiheitsgrade – durch die Nebenbedingungen auf $(n-k)$ reduzieren sollte.

Wir betrachten ein Minimumproblem für ein Funktional $F : \mathcal{Y} \longrightarrow \mathbb{R}$ mit einer allgemeinen Nebenbedingung. Es sei $G : \mathcal{Y} \to \mathcal{X}^*$ eine Abbildung, $c \in \mathcal{X}^*$ und

$$A = \{y \in \mathcal{Y} \mid G(y) = c\}$$

die zulässige Menge. Wir wollen das Problem

$$F(y_0) = \min_{y \in A} F(y)$$

lösen. Die Lagrangemethode besteht darin, nach dem Sattelpunkt der Funktion

$$L(x, y) = F(y) - \langle x, G(y) - c\rangle$$

zu suchen. x heißt **Lagrangemultiplikator**.

Es stellt sich heraus, daß L einen Sattelpunkt (x_0, y_0) hat, gdw. F einen Minimumspunkt y_0 hat. Dabei ist die Hinrichtung (wenn Sattel, dann Minimum) unter allgemeinen Bedingungen gültig, wogegen für die Rückrichtung – wegen der geforderten Existenz des Lagrange-Multiplikators x_0 – zusätzliche Annahmen an F und G gestellt werden müssen.
Diese zusätzlichen Bedingungen sind:

- $\mathcal{Y}$ sei reflexiv.
- Es sei $\inf_y F(y) > -\infty$.

  Bemerkung: Wenn wir ein Minimum von F suchen, sollte F nach unten beschränkt sein.
- Es sei $\sup_y \|G(y)\|_{\mathcal{X}^*} < \infty$.

  Bemerkung: Die Abbildung G interessiert uns eigentlich nur auf der Menge A, d.h., dort, wo G den Wert c annimmt. Außerhalb von A kann G beliebig fortgesetzt werden. Wir nehmen an, daß das auf beschränkte Weise geschiet.

Mit diesen Voraussetzungen gilt folgender
Satz: Der Punkt (x_0, y_0) ist ein Sattelpunkt von L genau dann, wenn $y_0 \in A$ und $F(y_0)$ das Minimum von $F(y)$ auf A bildet.
Beweis $\Longrightarrow$: Es sei (x_0, y_0) Sattelpunkt der Funktion $L(x, y)$. Dann gilt nach Definition des Sattelpunktes

$$\begin{array}{ccccc} L(x, y_0) & \leq_1 & L(x_0, y_0) & \leq_2 & L(x_0, y) \\ F(y_0) - \langle x, G(y_0) - c\rangle & \leq_1 & F(y_0) - \langle x_0, G(y_0) - c\rangle & \leq_2 & F(y) - \langle x_0, G(y) - c\rangle \end{array}$$

Aus Ungleichung $\leq_1$ folgt $0 \leq \langle x - x_0, G(y_0) - c\rangle$ für alle x und wegen der Linearität der dualen Paarung $0 \geq \langle x - x_0, G(y_0) - c\rangle$ also $0 = \langle x - x_0, G(y_0) - c\rangle$ für alle $x \in \mathcal{X}$ und somit $G(y_0) - c = 0$ also $y_0 \in A$.
Aus Ungleichung $\leq_2$ folgt dann

$$F(y_0) \leq F(y) - \langle x_0, G(y) - c\rangle$$

für alle $y \in \mathcal{Y}$. Wir betrachten $y \in A$. Dann folgt $F(y_0) \leq F(y)$ für alle $y \in A$. Aber das ist gerade die Definition des Minimumpunktes für y_0.
Beweis $\Longleftarrow$: Hier ist zu zeigen, daß unter der Annahme, daß $F(y)$ in y_0 ein Minimum auf A hat, es ein x_0 gibt, sodaß (x_0, y_0) ein Sattelpunkt für L ist. Das heißt, aus einer Ungleichung auf A folgt eine Ungleichung auf ganz $\mathcal{Y}$.
Wir zeigen, daß die Funktion $L(x, y)$ ein minimax-Theorem erfüllt.
1): Wir betrachten die Funktion

$$g(y) = \sup_x \big(F(y) - \langle x, G(y) - c\rangle\big)$$

Für $y \notin A$, ist $G(y) - c \neq 0$. Dann läßt sich stets eine Folge x_n finden, sodaß $\langle x_n, G(y) - c\rangle$ beliebig klein, also $-\langle x_n, G(y) - c\rangle$ beliebig groß wird. Da $F(y) > -\infty$, ist $g(y) = +\infty$ für $y \notin A$. Für $y \in A$ ist $g(y) = F(y)$. Das heißt,

$$F(y_0) = \min_{y\in A} F(y) = \min_{y\in A} g(y) = \min_y \sup_x \big(F(y) - \langle x, G(y) - c\rangle\big)$$

2) Wir zeigen jetzt, daß

$$F(y_0) \leq \max_x \inf_y \big(F(y) - \langle x, G(y) - c\rangle\big)$$

gilt. Ist das der Fall, so folgt die Behauptung aus der schwachen Dualität.
Es sei

$$f(x) = \inf_y \big(F(y) - \langle x, G(y) - c\rangle\big)$$

Diese Funktion ist als Infimum affiner Funktionen konkav und oberhalbstetig.
Den Fall $f(x) \equiv -\infty$ können wir wegen der Voraussetungen an das Wachstum von F und G ausschließen. Wir untersuchen, wie sich $f(x)$ für $\|x\| \to \infty$ verhält.
Die Funktion $G(x) - c$ ist außerhalb der Menge A positiv oder negativ. Für $x \longrightarrow +\infty$ sind die positiven Werte von $G(y)$ die dominierenden in $F(y) - xG(y)$, für $x \to -\infty$ die negativen Werte von $G(y)$. In beiden Fällen geht also $f(x) \longrightarrow -\infty$. Damit ist $-f(x)$ konvex, koerzitiv und unterhalbstetig. $f(x)$ nimmt ihr Maximum im Inneren von $\mathcal{X}$ an. Dieser Punkt sei x_0. Es gilt also

$$f(x_0) = \max_x f(x) = \max_x \inf_y \big(F(y) - \langle x, G(y) - c\rangle\big)$$

Wir zeigen $f(x_0) \geq F(y_0)$. Damit wäre zusammen mit der stets geltenden schwachen Dualität $f(x_0) = F(y_0)$ gezeigt.
Angenommen, $f(x_0) < F(y_0)$, d.h., es gilt

$$\inf_y \big(F(y) - \langle x_0, G(y) - c\rangle\big) < F(y_0) = \inf_{y\in A} F(y)$$

dann kann das Infimum links nicht auf A angenommen werden. Dann gibt es eine Folge $(y_n) \not\subset A$ mit $F(y_n) - \langle x_0, G(y_n) - c\rangle < F(y_0)$ und $G(y_n) > 0$ (oder $G(y_n) < 0$). Jetzt läßt sich x_0 verkleinern (vergrößern) und daß Supremum in $\sup_x f(x)$ wird größer. Das ist ein Widerspruch dazu, daß $f(x_0)$ das Supremum war. □

Bemerkung 1: Es sei $G(x) = c$. Dann ist $L(x,y) = F(y) - x(G(y) - c)$ zu wählen. Wenn L glatt ist, dann ist $\frac{\partial}{\partial c} L = x$ und x läßt sich interpretieren als die Größe, die die Änderung von L in Abhängigkeit von c angibt.

Bemerkung 2: Die Idee hinter der Methode besteht darin, daß man außerhalb der Menge M den Ausdruck $xG(y)$ durch Wahl des x beliebig groß oder klein machen kann, da dort $G(y) \neq 0$. D.h., eigentlich betrachtet man eine Funktion, die außerhalb von M den Wert ∞ annimmt.

Bemerkung 3:
Ein bedingtes Extremalproblem ist eigentlich ein Sattelproblem für eine Funktion, die in einer Variablen zu einer linearen Funktion entartet ist.

4.5.2 Die Euler-Lagrange-Gleichung

Angenommen F und G sind ausreichend glatt, dann erfüllt der Sattelpunkt (x_0, y_0) die notwendige Bedingung $\frac{\partial}{\partial x} L = 0$ und $\frac{\partial}{\partial y} L = 0$. Die erste Gleichung ist einfach die Nebenbedingung. Die zweite Gleichung ist formal soetwas wie $F'(y) - \langle x, G'(y)\rangle = 0$. Um diesen Ausdruck zu verstehen, führen wir folgende Definition ein:
Definition: Eine Abbildung $G : \mathcal{Y} \to \mathcal{X}$ heißt im Punkt y **Frechet**-differenzierbar, wenn ein linearer Operator $G'(y) : \mathcal{Y} \to \mathcal{X}$ existiert und für alle $h \in \mathcal{Y}$ gilt

$$G(y+h) = G(y) + G'(y)h + \omega(x,h)$$

mit $\|\omega\|_{\mathcal{X}} \to 0$ für $\|h\|_{\mathcal{Y}} \to 0$.
Definition: Ein Punkt $y \in M$ heißt regulärer Punkt der Menge M, falls G in diesem Punkt Frechet-differenzierbar ist und $R(G'(y)) = \mathcal{X}$.

Eine notwendige Bedingung für das genannte Minimumproblem gibt folgender
Satz: Es sei y_0 Lösung des Minimumproblems

$$F(y_0) = \min_{y \in M} F(y) \ .$$

F sei Gateauxdifferenzierbar, G Frechet-differenzierbar in y_0 und y_0 regulärer Punkt von M, dann existiert ein Punkt $x_0 \in \mathcal{X}$, sodaß gilt

$$\langle h, F'(y_0)\rangle - \langle G'(y_0)h, x\rangle = 0, \ \ h \in \mathcal{Y}$$

Diese Gleichung heißt vage (oder schwache) Euler-Lagrange-Gleichung des Minimumproblems und x_0 heißt **Lagrangemultiplikator**.
Falls $G'(y_0)$ darüber hinaus dicht definiert ist, existiert sein adjungierter $[G'(y_0)]^*$ und x_0 erfüllt die starke Euler-Lagrange-Gleichung

$$F'(y_0) - [G'(y_0)]^* x : 0 = 0$$

Zusammen mit der Gleichung der Nebenbedingung kann auf diese Weise der Sattelpunkt (x_0, y_0) berechnet werden.

4.5.3 Zusatzbedingungen als Ungleichungen (Variablen umbenennen!)

Die Lagrangemethode läßt sich auf Nebenbedingungen in Ungleichungsform verallgemeinern. Zu beachten ist hierbei, daß das entsprechende Problem für den Lagrangemultiplikator nicht im ganzen Raum betrachtet werden darf sondern nur dort (z.B. im positiven Kegel), daß die Nebenbedingung wie im Fall einer Gleichung als Nebenbedingung in eine charakteristische Funktion übergeht.
Wir betrachten den allgemeineren Fall mehrerer Zusatzbedingungen, die durch eine Ungleichung gegeben seien. Es seien $F : \mathcal{Y} \longrightarrow \mathbb{R}$, $G : \mathcal{Y} \longrightarrow \mathcal{X}$, $x^* \in \mathcal{X}^*$. Wir betrachten das bedingte Minimumproblem

$$\begin{aligned} F(y_0) &= \inf_{y\in\mathcal{Y}} F(y) \\ M &= \{y \in \mathcal{Y} \mid G(y) \leq 0\} \end{aligned}$$

Zu diesem Problem betrachten wir die Funktion

$$L(y, x^*) = F(y) + \langle G(y), x^* \rangle \ , (y, x^*) \in \mathcal{Y} \times \mathcal{X}_+$$

Es gilt folgender
Satz: (x_0^*, y_0) ist genau dann ein Sattelpunkt von $L(y, x^*)$ mit $y^* \in \mathcal{Y}_+^*$. wenn F in y_0 ein Minimum unter der angegebenen Nebenbedingung annimmt.
Beweis: Da L Lagrangefunktion mit Sattelpunkt (x_0, y_0^*) ist, gilt

$$\begin{array}{ccccccc} L(x_0, y^*) & \leq & L(x_0, y_0^*) & \leq & L(x, y_0^*), & & x \in \mathcal{X},\ y_0^* \in \mathcal{Y}_+^* \\ F(x_0) + \langle G(x_0), y^* \rangle & \leq_1 & F(x_0) + \langle G(x_0), y_0^* \rangle & \leq_2 & F(x) + \langle G(x), y_0^* \rangle, & & x \in \mathcal{X},\ y_0^* \in \mathcal{Y}_+^* \end{array}$$

Ungleichung $\leq_1$ gilt für beliebige $y^* \geq 0$, also auch für beliebige $y^* = y_0^* + h^*$ mit $h^* \geq 0$. Daraus folgt $\langle G(x_0), y_0^* + h^* \rangle \leq_1 \langle G(x_0), y_0^* \rangle$ also $\langle G(x_0), h^* \rangle \leq 0$ für beliebige $h^* \geq 0$. Also $G(x_0) \leq 0$. Setzt man in Ungleichung $\leq_1$ $y^* = 0$ erhält man $0 \leq \langle G(x_0), y_0^* \rangle$. Da $\langle G(x_0), y^* \rangle \leq 0$ für alle $y^* \geq 0$ gilt, folgt $\langle G(x_0), y_0^* \rangle = 0$. Damit folgt aus Ungleichung $\leq_2$

$$F(x_0) \leq_2 F(x) + \langle G(x), y_0^* \rangle \ .$$

Das heißt, für alle $x \in X$, also die x mit $G(x) \leq 0$ gilt $F(x_0) \leq F(x)$.
Satz: Es sei (x_0, y_0^*) ein Sattelpunkt von $L(x, y^*)$ mit $y \in \mathcal{Y}^*$. Dann ist x_0 Lösungspunkt der Aufgabe

$$\begin{aligned} F(x_0) &= \inf_{x\in X} F(x) \\ X &= \{x \in \mathcal{X} \mid G(x) = 0\} \end{aligned}$$

Der Unterschied zum Satz eben besteht darin, daß im aktuellen Fall die Sattelpunktbedingung für alle $y \in \mathcal{Y}^*$ gelten muß. Entsprechend kleiner ist die Menge X der zulässigen x.
Beweis: Da L Lagrangefunktion mit Sattelpunkt (x_0, y_0^*) ist, gilt

$$\begin{array}{ccccccc} L(x_0, y^*) & \leq & L(x_0, y_0^*) & \leq & L(x, y_0^*), & & x \in \mathcal{X},\ y^* \in \mathcal{Y}^* \\ F(x_0) + \langle G(x_0), y^* \rangle & \leq_1 & F(x_0) + \langle G(x_0), y_0^* \rangle & \leq_2 & F(x) + \langle G(x), y_0^* \rangle, & & x \in \mathcal{X},\ y^* \in \mathcal{Y}^* \end{array}$$

Ungleichung $\leq_1$ gilt für beliebige y^*, also auch für beliebige $y^* = y_0^* + h^*$ mit beliebigem h^*. Daraus folgt $\langle G(x_0), y_0^* + h^* \rangle \leq_1 \langle G(x_0), y_0^* \rangle$ also $\langle G(x_0), h^* \rangle \leq 0$ für beliebige h^*. Das ist nur für $G(x_0) = 0$ möglich. Damit folgt aus Ungleichung $\leq_2$

$$F(x_0) \leq_2 F(x) + \langle G(x), y_0^* \rangle \ .$$

Das heißt, für alle $x \in X$ gilt $F(x_0) \leq F(x)$.

4.6 Dualität in vier Räumen

4.6.1 Der Übergang von L zu H

Lagrangefunktionen sind ein fundamentales mathematisches Hilfsmittel. Trotzdem gibt es eine Reihe von Nachteilen (die positiven Eigenschaften wurden bereits ausgiebig gewürdigt).

- Wie kommt man drauf?
- Sie haben die Einheit der Energie, sind aber nicht additiv, also keine Energie? Das ist eine andere Konstruktion, die es nur bei Energien gibt, nicht bei anderen extensiven Größen.

Dazu kann man sich folgende Fragen stellen:

- Wo sind die vier Räume aus der Einführungsaufgabe? Die gesternten Räume kamen damals natürlicherweise vor. Unser L lebt aber nur in den ungesternten Räumen.
-
- Gibt es Funktionen, die als Energie interpretiert werden können, und deren Spezialfall $P(x)$ ist?

Wir wollen jetzt die Lagrangefunktionen in die Welt der vier Räume einbetten.
Als erstes stellen wir fest, daß in Spezialfällen (Einführungsbeispiel), sowohl L als auch P und Q die Einheit der Energie hatten. Dabei waren P und Q tatsächlich als Energie zu verstehen. Sie waren additiv und setzten sich aus der Summe der Teilenergien zusammen. L hat auch die Einheit der Energie. Allerdings ist L ganz offensichtlich nicht additiv, sondern eher sowas wie subtraktiv. Das ist physikalisch natürlich Unsinn. D.h., L hat zwar die Einheit der Energie, läßt sich aber nirgends als Energie wiederfinden.
Ausgehend von den Ungleichungen

$$\begin{array}{ccccccccccccc} -P(x) & = & f(x) & = & \inf_y L(x,y) & \le & L(x,y) & \le & \sup_x L(x,y) & = & g(y) & = & Q(y) \\ & & f(x) & \le & L(x,y_0) & \le & L(x_0,y_0) & \le & L(x_0,y) & \le & g(y) & & \\ & & \sup_x f(x) & = & f(x_0) & = & L(x_0,y_0) & = & g(y_0) & = & \inf_y g(y) & & \end{array}$$

können wir uns fragen, ob sich die beiden Extrema $f(x)$ und $g(y)$ ähnlich wie $L(x,y_0)$ und $L(x_0,y)$ als spezielle Funktionswerte von geeigneten Funktionen bestimmen lassen. Es ist

$$f(x) = \inf_y L(x,y) = -\sup_y \big(-L(x,y)\big) = -\sup_y \big(\langle y,y^*\rangle - L(x,y)\big)\Big|_{y^*=0} = -H(x,0)$$

wobei

$$H(x,y^*) = \sup_y \big(\langle y,y^*\rangle - L(x,y)\big)$$

die Legendretransformation von L bezüglich der Variablen y ist. Wir nehmen im weiteren an, daß $L(x,\cdot)$ konvex und uhs ist. dann gilt auch

$$L(x,y) = \sup_{y^*} \big(\langle y,y^*\rangle - H(x,y^*)\big)$$

Zur Funktion $H(x,y^*)$ können wir die Legendretransformation bezüglich beider Variablem betrachten:

$$\begin{aligned} H^*(x^*,y) &= \sup_{x,y^*} \big(\langle x,x^*\rangle + \langle y,y^*\rangle - H(x,y^*)\big) \\ H(x,y^*) &= \sup_{x^*,y} \big(\langle x,x^*\rangle + \langle y,y^*\rangle - H^*(x^*,y)\big) \end{aligned}$$

Wir berechnen $H^*(0,y)$:

$$H^*(0,y) = \sup_{x,y^*}\big(\langle y,y^*\rangle - H(x,y^*)\big) = \sup_x L(x,y) = g(y)$$

Damit erhalten wir

$$\begin{aligned} f(x) &= \inf_y L(x,y) = -H(x,0) \\ g(y) &= \sup_x L(x,y) = H^*(0,y) \end{aligned}$$

Aus $f(x) \le g(y)$ folgt die Ungleichung

$$H^*(0,y) + H(x,0) \ge 0$$

die im Falle des Sattelpunktes zum Satz des Pythagoras $H^*(0,y_0) + H(x_0,0) = 0$ wird.
Man kann einen Zusammenhang zwischen L und H^* finden:

$$\begin{aligned} L(x,y) &= \sup_{y^*}\big(\langle y,y^*\rangle - H(x,y^*)\big) = \\ &= \sup_{y^*}\Big(\langle y,y^*\rangle - \sup_{x^*,y_1}\big(\langle x,x^*\rangle + \langle y_1,y^*\rangle - H^*(x^*,y_1)\big)\Big) = \\ &= \sup_{y^*}\inf_{x^*,y_1}\big(\langle y,y^*\rangle - \langle x,x^*\rangle - \langle y_1,y^*\rangle + H^*(x^*,y_1)\big) = \\ &= \inf_{x^*,y_1}\big(\chi_y(y_1) - \langle x,x^*\rangle + H^*(x^*,y_1)\big) = \\ &= \inf_{x^*}\big(-\langle x,x^*\rangle + H^*(x^*,y)\big) = -\sup_{x^*}\big(\langle x,x^*\rangle - H^*(x^*,y)\big) \\ H^*(x^*,y) &= \sup_{x^*}\big(\langle x,x^*\rangle + L(x,y)\big) \end{aligned}$$

4.6.2 Der Übergang von H zu L^*

Ausgehend von H^* können wir wieder eine duale Lagrangefunktion definieren:

$$\begin{aligned} L^*(x^*,y^*) &= \sup_y\big(\langle y,y^*\rangle - H^*(x^*,y)\big) \\ H^*(x^*,y) &= \sup_{y^*}\big(\langle y,y^*\rangle - L^*(x^*,y^*)\big) \end{aligned}$$

Hieraus ergibt sich

$$\begin{aligned} f^\circ(x^*) &= \inf_{y^*} L^*(x^*,y^*) = -\sup_{y^*}\big(-L^*(x^*,y^*)\big) = -H^*(x^*,0) \\ g^\circ(y^*) &= \sup_{x^*} L^*(x^*,y^*) = \sup_{x^*,y^*}\big(\langle y,y^*\rangle - H^*(x^*,y)\big) = H(0,y^*) \end{aligned}$$

Einfachste Beispiele zeigen $f^* \neq f^\circ$, $g^* \neq g^\circ$.
Einen Zusammenhang zwischen L^* und L erhält man so:

$$\begin{aligned} L^*(x^*,y^*) &= \sup_{y^*}\big(\langle y,y^*\rangle - H^*(x^*,y)\big) = \\ &= \sup_{y^*}\Big(\langle y,y^*\rangle - \big(\sup_{x^*}\big(\langle x,x^*\rangle + L(x,y)\big)\Big) = \\ &= \sup_{y^*}\inf_{x^*}\big(\langle y,y^*\rangle - \langle x,x^*\rangle - L(x,y)\big) \end{aligned}$$

Umkehrung:

$$\begin{aligned}
L(x,y) &= -\sup_{x^*}\big(\langle x,x^*\rangle - H^*(x^*,y)\big) = \inf_{x^*}\big(H^*(x^*,y) - \langle x,x^*\rangle\big) = \\
&= \inf_{x^*}\Big(\big(\sup_{y^*}\big(\langle y,y^*\rangle - L^*(x^*,y^*)\big)\big) - \langle x,x^*\rangle\Big) = \\
&= \inf_{x^*}\sup_{y^*}\big(\langle y,y^*\rangle - \langle x,x^*\rangle - L^*(x^*,y^*)\big)
\end{aligned}$$

4.6.3 Beispiel 1

$$\begin{aligned}
L(x,y) &= T(y) - U(x) \\
H(x,y^*) &= T^*(y^*) + U(x) \\
H^*(x^*,y) &= T(y) + U^*(x^*) \\
f(x) &= -T^*(0) - U(x) \\
g(y) &= T(y) + U^*(0) \\
L^*(x^*,y^*) &= T^*(y^*) - U^*(x^*) \\
f^\circ(x^*) &= \inf_{y^*} L^*(x^*,y^*) = -T(0) - U^*(x^*) \neq f^*(x^*) \\
g^\circ(y^*) &= \sup_{x^*} L^*(x^*,y^*) = T^*(y^*) + U(0) \neq g^*(x^*)
\end{aligned}$$