# 54 Lokale Extrema mit Nebenbedingungen

Lokale Extrema reellwertiger Funktionen auf *offenen Teilmengen* des $\mathbb{R}^n$ wurden bereits in Abschnitt 52 untersucht. Für zahlreiche Anwendungen ist es wichtig, als Definitionsbereiche auch *Kurven* im $\mathbb{R}^2$, *Flächen* im $\mathbb{R}^3$ oder entsprechende höherdimensionale Mengen im $\mathbb{R}^n$ zuzulassen.

**54.1 Beispiel.** a) Die *Zielfunktion*

$$Z : \mathbb{R}^2 \mapsto \mathbb{R}^2\,, \quad Z(x,y) := \; xy\,, \tag{1}$$

besitzt auf $\mathbb{R}^2$ *kein lokales Extremum,* ihre *Einschränkung* auf die kompakte *Kreislinie* $S^1 = \{(x,y) \in \mathbb{R}^2 \mid x^2+y^2=1\}$ besitzt aber nach Satz 48.17 ein Maximum und ein Minimum.

b) Diese kann man durch *Auflösung* der *Nebenbedingung*

$$f(x,y) := \; x^2+y^2-1 \; = \; 0 \tag{2}$$

bestimmen. Für $y > 0$ ist (2) äquivalent zu

$$y \; = \; g(x) := +\sqrt{1-x^2}\,,$$

und mit (1) ergibt sich

$$z(x) := \; Z(x,g(x)) \; = \; x\sqrt{1-x^2}\,.$$

Man hat

$$z'(x) \; = \; \sqrt{1-x^2} - \frac{x^2}{\sqrt{1-x^2}} \; = \; \frac{1-2\,x^2}{\sqrt{1-x^2}}\,;$$

die kritischen Punkte von $z$ in $(-1,1)$ sind also $x = \pm\frac{\sqrt{2}}{2}$. Eventuelle lokale Extrema der Zielfunktion $Z$ auf der „offenen“ oberen Kreislinie liegen also in den Punkten $P_1 = \frac{1}{2}(\sqrt{2},\sqrt{2})$ oder $P_2 = \frac{1}{2}(-\sqrt{2},\sqrt{2})$.

Die entsprechende Rechnung für $y < 0$ liefert die Punkte $P_3 = \frac{1}{2}(\sqrt{2},-\sqrt{2})$ und $P_4 = \frac{1}{2}(-\sqrt{2},-\sqrt{2})$; lokale Extrema könnten auch noch in den Punkten $P_5 = (1,0)$ oder $P_6 = (-1,0)$ liegen. Wegen $Z(P_1) = Z(P_4) = \frac{1}{2}$, $Z(P_2) = Z(P_3) = -\frac{1}{2}$, und $Z(P_5) = Z(P_6) = 0$ ist $M = \frac{1}{2}$ das Maximum und $m = -\frac{1}{2}$ das Minimum der Zielfunktion $Z$ auf $S^1$.

c) Man kann auch die Kreislinie parametrisieren:

$$S^1 \; = \; \psi(\mathbb{R}) \quad \text{mit} \quad \psi(t) \; = \; (\cos t, \sin t)\,, \quad t \in \mathbb{R}\,. \tag{3}$$

Man betrachtet dann die auf $\mathbb{R}$ definierte Funktion

$$\zeta(t) := \; Z(\psi(t)) \; = \; \cos t \cdot \sin t \; = \; \tfrac{1}{2}\sin 2t\,.$$

Wegen $\dot{\zeta}(t) = \cos 2t$ erhält man die kritischen Parameter $t = \frac{\pi}{4} + k\,\frac{\pi}{2}$, $k \in \mathbb{Z}$, und daraus wieder die Punkte $P_1$, $P_2$ $P_3$, $P_4$ aus b) als Kandidaten für lokale Extrema der Zielfunktion $Z$.

**54.2 Nebenbedingungen.** a) Es seien nun $D \subseteq \mathbb{R}^n$ offen, $Z \in \mathcal{C}^1(D,\mathbb{R})$ und $f \in \mathcal{C}^1(D,\mathbb{R}^m)$ mit $1 \le m < n$. Gesucht sind *lokale Extrema* der Zielfunktion $Z$ unter der *Nebenbedingung* $f(x) = 0$, ausführlich

$$\begin{aligned} f_1(x_1,\ldots,x_n) &= 0\,, \\ f_2(x_1,\ldots,x_n) &= 0\,, \\ &\ldots \\ f_m(x_1,\ldots,x_n) &= 0\,, \end{aligned} \qquad (4)$$

d. h. lokale Extrema von $Z|_S$ auf der Niveaumenge $S = N_0(f)$ von $f$.

b) I. a. kann man weder (lokale) Auflösungen der Nebenbedingungen noch Parametrisierungen von $S$ *explizit* berechnen. Daher ist es interessant, Kriterien für lokale Extrema zu finden, die nur die Kenntnis von $Z$ und $f$ benötigen.

Für eine beliebige Menge $S \subseteq \mathbb{R}^n$ und $q \in S$ wird mit $N_q(S) := T_q(S)^\perp$ der *Normalenraum* an $S$ in $q$ bezeichnet, d. h. der Raum aller Vektoren, die zu allen Tangentenvektoren an $S$ in $q$ orthogonal sind.

**54.3 Satz.** *Es seien $D \subseteq \mathbb{R}^n$ offen, $S \subseteq D$, und die Einschränkung $Z|_S$ der Zielfunktion $Z \in \mathcal{C}^1(D,\mathbb{R})$ besitze ein lokales Extremum in $q \in S$. Dann gilt* $\operatorname{grad} Z(q) \in N_q(S)$.

Beweis. Es seien $\mathfrak{t} \in T_q(S)$ und $\gamma : (-\delta,\delta) \mapsto \mathbb{R}^n$ ein Weg mit $(\gamma) \subseteq S$, $\gamma(0) = q$ und $\dot\gamma(0) = \mathfrak{t}$. Dann besitzt $Z \circ \gamma$ ein lokales Extremum in $0$, und daher ist $\frac{d}{dt}(Z\circ\gamma)(0) = 0$. Die Kettenregel liefert sofort

$$\langle \operatorname{grad} Z(q), \mathfrak{t} \rangle \;=\; \tfrac{d}{dt}(Z\circ\gamma)(0) \;=\; 0\,,$$

also $\operatorname{grad} Z(q) \perp \mathfrak{t}$ und somit $\operatorname{grad} Z(q) \in N_q(S)$.

**54.4 Lagrange-Multiplikatoren.** a) Wie in 54.2 seien nun $1 \le m < n$, $D \subseteq \mathbb{R}^n$ offen, $Z \in \mathcal{C}^1(D,\mathbb{R})$, $f \in \mathcal{C}^1(D,\mathbb{R}^m)$ und $S = N_0(f)$. Die Einschränkung $Z|_S$ von $Z$ besitze ein lokales Extremum in einem *regulären* Punkt $q \in S$, d. h. es gelte

$$\operatorname{rg} Df(q) \;=\; m\,. \qquad (5)$$

b) Wegen (5) ist $S$ nahe $q$ nach dem Satz über implizite Abbildungen ein *Graph,* und nach 50.9 ist der Tangentialkegel an $S$ in $q$ ein $(n-m)$-dimensionaler *Vektorraum* $T_q(S) \subseteq \mathbb{R}^n$. Folglich gilt $\dim N_q(S) = \dim T_q(S)^\perp = m$.

c) Nach Satz 54.3 gilt $\operatorname{grad} f_i(q) \in N_q(S)$ für $i = 1,\ldots,m$, und diese Vektoren sind wegen (5) linear unabhängig. Folglich ist $\{\operatorname{grad} f_1(q),\ldots,\operatorname{grad} f_m(q)\}$ eine *Basis* von $N_q(S)$.

d) Nach Satz 54.3 gilt auch $\operatorname{grad} Z(q) \in N_q(S)$, und daher muß

$$\operatorname{grad} Z(q) \;=\; \sum_{i=1}^m \mu_i \operatorname{grad} f_i(q) \qquad (6)$$

mit geeigneten $\mu_1,\ldots,\mu_m \in \mathbb{R}$ gelten. Die Zahlen $\mu_1,\ldots,\mu_m \in \mathbb{R}$ heißen *Lagrange-Multiplikatoren.*

e) Für die *Hilfsfunktion* $h \in \mathcal{C}^1(D \times \mathbb{R}^m, \mathbb{R})$,

$$h(x,\lambda) := Z(x) - \sum_{i=1}^{m} \lambda_i f_i(x)\,, \tag{7}$$

gilt also $\operatorname{grad}_x h(q,\mu) = 0$ und auch $\operatorname{grad}_\lambda h(q,\mu) = -f(q) = 0$ ! Die $(n+m)$ Gleichungen

$$\operatorname{grad} h(x,\lambda) = 0 \tag{8}$$

in den $(n+m)$ Unbekannten $(x_1,\ldots,x_n,\lambda_1,\ldots,\lambda_m) \in D \times \mathbb{R}^m$ stellen also eine *notwendige Bedingung* für das Vorliegen eines lokalen Extremums unter der Nebenbedingung „$f(x) = 0$" in *regulären* Punkten von $S = N_0(f)$ dar.

Die folgende Formulierung eines notwendigen Kriteriums für lokale Extrema gilt auch in *singulären* Punkten von $S = N_0(f)$:

**54.5 Satz.** *Es seien $D \subseteq \mathbb{R}^n$ eine offene Menge, $Z \in \mathcal{C}^1(D,\mathbb{R})$, $f \in \mathcal{C}^1(D,\mathbb{R}^m)$ und $S = \{x \in D \mid f(x) = 0\}$. Besitzt dann $Z|_S$ ein lokales Extremum in $q \in S$, so folgt*

$$\operatorname{rg} D(f,Z)(q) := \operatorname{rg} \begin{pmatrix} Df(q) \\ DZ(q) \end{pmatrix} < m+1\,. \tag{9}$$

Beweis. Ist $\operatorname{rg} Df(q) = m$, so gilt (6), d.h. $DZ(q)$ ist eine Linearkombination der Zeilen von $Df(q)$. Ist aber $\operatorname{rg} Df(q) < m$, so ist (9) offensichtlich auch erfüllt.

**54.6 Beispiele.** a) Im Fall $m = n-1$ ist $D(f,Z)(q)$ eine *quadratische* Matrix, (9) also äquivalent zu

$$\det D(f,Z)(q) = 0\,. \tag{10}$$

b) Die Funktion $Z(x,y) := xy$ auf der Kreislinie $S^1$ aus 54.1 wird mittels (10) untersucht. Es gilt $D(f,Z)(x,y) = \begin{pmatrix} 2x & 2y \\ y & x \end{pmatrix}$, also $\det D(f,Z)(x,y) = 2x^2 - 2y^2$. Aus $\det D(f,Z)(x,y) = 0$ und $x^2+y^2-1 = 0$ ergibt sich daher sofort $x^2 = y^2 = \frac{1}{2}$.

c) Man kann auch (7) verwenden. Man hat

$$\begin{aligned} h(x,y,\lambda) &= x\,y - \lambda\,(x^2+y^2-1)\,, \quad \text{also} \\ \operatorname{grad} h(x,y,\lambda) &= (y - 2\,\lambda\,x\,,\, x - 2\,\lambda\,y\,,\, -(x^2+y^2-1))^\top\,. \end{aligned}$$

Aus $\operatorname{grad} h(x,y,\lambda) = 0$ folgt wieder $x^2 = 2\lambda xy = y^2$, also $x^2 = y^2 = \frac{1}{2}$ sowie $\lambda = \pm\frac{1}{2}$.

**54.7 Satz.** *Es seien $D \subseteq \mathbb{R}^n$ eine offene Menge, $Z \in \mathcal{C}^2(D,\mathbb{R})$, $f \in \mathcal{C}^2(D,\mathbb{R}^m)$ und $S = \{x \in D \mid f(x) = 0\}$. Es seien $q \in S$ ein regulärer Punkt von $S$ mit* $\operatorname{grad} Z(q) \in N_q(S)$ *und $A = Hh(q)$ die Hesse-Matrix der Hilfsfunktion $h$ aus (7). Dann gilt:*

*a) $A > 0$ auf $T_q(S) \Rightarrow Z|_S$ hat ein isoliertes lokales Minimum in $q$,*

*b) $A < 0$ auf $T_q(S) \Rightarrow Z|_S$ hat ein isoliertes lokales Maximum in $q$,*

*c) $A$ indefinit auf $T_q(S) \Rightarrow Z|_S$ hat kein lokales Extremum in $q$.*

**54.8 Beispiel.** Für $c > 0$ bestimmen wir alle lokalen Extrema der Funktion $Z(x,y,z) = xy + xz + yz$ unter der Nebenbedingung $f(x,y,z) = xyz - c^3 = 0$.

a) Wegen $Df = (yz, xz, xy) \neq 0$ auf $S := N_0(f)$ sind alle Punkte von $S$ regulär. Für die gemäß (7) gebildete Hilfsfunktion $h(x,y,z,\lambda) := Z(x,y,z) - \lambda f(x,y,z)$ ist

$$\operatorname{grad} h \;=\; (y+z-\lambda yz, x+z-\lambda xz, x+y-\lambda xy, xyz-c^3)^\top,$$

und $\operatorname{grad} h(x,y,z,\lambda) = 0$ liefert

$$\begin{aligned} x(y+z) &= \lambda xyz = \lambda c^3, \\ y(x+z) &= \lambda xyz = \lambda c^3, \\ z(x+y) &= \lambda xyz = \lambda c^3. \end{aligned}$$

Es folgt $xz - yz = 0$, $yx - zx = 0$ und somit $x = y = z = c$, $\lambda = \frac{2}{c}$.

b) Die Hesse-Matrix $A = Hh(q)$ von $h$ im Punkt $q = (c,c,c)$ ist

$$A \;=\; \begin{pmatrix} 0 & 1-\lambda z & 1-\lambda y \\ 1-\lambda z & 0 & 1-\lambda x \\ 1-\lambda y & 1-\lambda x & 0 \end{pmatrix} \;=\; \begin{pmatrix} 0 & -1 & -1 \\ -1 & 0 & -1 \\ -1 & -1 & 0 \end{pmatrix}.$$

Die Eigenwerte von $A$ sind $1$, $1$ und $-2$, $A$ ist also *indefinit.*

c) Nach Satz 54.7 ist nur die *Einschränkung* der Hesse-Form auf den *Tangentialraum* $T_q(S)$ interessant. Nach 54.4 c) ist aber

$$\begin{aligned} T_q(S) &= N_q(S)^\perp = [\operatorname{grad} f(q)]^\perp = [(c^2,c^2,c^2)^\top]^\perp \\ &= \{(s+t,-s,-t)^\top \mid s,t \in \mathbb{R}\} \end{aligned}$$

der *Eigenraum* von $A$ zum doppelten Eigenwert $1$. Folglich ist $Q_A$ auf $T_q(S)$ *positiv definit,* und $Z|_S$ besitzt ein lokales Minimum in $q$.