L98: Optimierung auf Mannigfaltigkeiten

Lösung (98.1) Zu optimieren ist die Funktion $f(x,y) := xy$ unter der Nebenbedingung $0 = x^2+y^2-1 =: g(x,y)$.

(a) Nach Lagrange gibt es an der gesuchten Stelle (x,y) eine Zahl $\lambda\in\mathbb{R}$ mit $\nabla f = \lambda\,\nabla g$, also

$$(\star)\qquad \begin{bmatrix} y \\ x \end{bmatrix} = \lambda \begin{bmatrix} 2x \\ 2y \end{bmatrix}.$$

Wegen $(x,y)\neq 0$ ist $\lambda\neq 0$; es folgt dann sowohl $x\neq 0$ als auch $y\neq 0$. Aus $(\star)$ ergibt sich daher $x:y=y:x$, also $x^2=y^2$ und damit $y=\pm x$. Einsetzen in die Nebenbedingung $x^2+y^2=1$ liefert $2x^2=1$ und damit $x=\pm 1/\sqrt{2}$. Damit erhalten wir die vier kritischen Punkte $(x,y) = (\pm 1/\sqrt{2},\, \pm 1/\sqrt{2})$. An den beiden Stellen $(1/\sqrt{2}, 1/\sqrt{2})$ und $(-1/\sqrt{2}, -1/\sqrt{2})$ wird dabei von f der Maximalwert $1/2$ angenommen, an den Stellen $(1/\sqrt{2}, -1/\sqrt{2})$ und $(-1/\sqrt{2}, 1/\sqrt{2})$ dagegen der Minimalwert $-1/2$.

(b) Der Kreis hat die Parameterdarstellung $(x,y) = (\cos t, \sin t)$; dann ist $f(x,y) = f(\cos t,\sin t) = \cos t\sin t = \sin(2t)/2$. Dieser Ausdruck wird maximal für $2t = (\pi/2) + 2k\pi$ bzw. $t = (\pi/4)+k\pi$ mit $k\in\mathbb{Z}$, dagegen minimal für $2t = (3\pi/2)+2k\pi$ bzw. $t=(3\pi/4)+k\pi$ mit $k\in\mathbb{Z}$. Im ersten Fall ist $(\cos t,\sin t) = \pm(1/\sqrt{2}, 1/\sqrt{2})$, im zweiten Fall ist $(\cos t,\sin t) = \pm(1/\sqrt{2}, -1/\sqrt{2})$; wir erhalten also (natürlich!) die gleiche Lösung wie in (a).

Lösung (98.2) Zu optimieren ist die Funktion $f(x,y) := xy^2$ unter der Nebenbedingung $g(x,y) = 0$ mit $g(x,y) := 4x^2+9y^2-36 = 0$.

(a) Nach Lagrange gibt es eine Zahl λ mit $\nabla f = \lambda\,\nabla g$, also

$$(\star)\qquad \begin{bmatrix} y^2 \\ 2xy \end{bmatrix} = \lambda \begin{bmatrix} 8x \\ 18y \end{bmatrix}.$$

Ist $\lambda = 0$, so führt $(\star)$ auf $y=0$ und dann $x^2=9$, also $x=\pm 3$. Ist $\lambda\neq 0$, so haben wir $x\neq 0$ und $y\neq 0$ und wegen $(\star)$ dann $2xy : y^2 = (18y):(8x)$, also $2x/y = (9y)/(4x)$ bzw. $8x^2 = 9y^2$. Setzt man dies in die Nebenbedingung $4x^2+9y^2=36$ ein, so ergibt sich $12x^2 = 36$ bzw. $x^2=3$, also $x=\pm\sqrt{3}$. Hieraus folgt dann $y^2 = 8x^2/9 = 8/3$ und damit $y = 2\sqrt{2/3}$. Wir erhalten also die sechs kritischen Punkte $(\pm\sqrt{3}, \pm 2\sqrt{2/3})$ und $(\pm 3, 0)$. An den beiden Punkten $P_{1,2} = (\sqrt{3}, \pm 2\sqrt{2/3})$ nimmt f den Wert $8/\sqrt{3} > 0$ an, an den Punkten $P_{3,4} = (-\sqrt{3}, \pm 2\sqrt{2/3})$ den Wert $-8/\sqrt{3} < 0$, an den Punkten $P_{5,6} = (\pm 3, 0)$ dagegen den Wert 0. Damit ist klar, daß bei $P_{1,2}$ jeweils ein (sogar globales) Maximum vorliegt, bei $P_{3,4}$ dagegen ein (sogar globales) Minimum. Bei $P_5 = (3,0)$ liegt ein lokales Minimum vor, bei $P_6 = (-3,0)$ in lokales Maximum; das folgt beispielsweise daraus, daß entlang einer Kurve (hier der Kreislinie $x^2+y^2=1$) eine Funktion zwischen zwei Minima stets ein Maximum besitzen muß, zwischen zwei Maxima stets ein Minimum.

(b) Wegen $(x/3)^2+(y/2)^2=1$ haben wir $x = 3\cos\varphi$ und $y = 2\sin\varphi$; zu optimieren ist also der Ausdruck $xy^2 = 12\cos\varphi\sin^2\varphi$ bzw. die Funktion $g(\varphi) := \sin^2\varphi\cos\varphi$. Nullsetzen von

$$\begin{aligned} g'(\varphi) &= 2\sin\varphi\cos^2\varphi - \sin^3\varphi \\ = \sin\varphi\,(2\cos^2\varphi - \sin^2\varphi) &= \sin\varphi\,(2-3\sin^2\varphi) \end{aligned}$$

führt auf die Lösungen $\varphi_1 = \arcsin(\sqrt{2/3})$, $\varphi_2 = 2\pi-\varphi_1$, $\varphi_{3,4} = \pi\pm\varphi_1$ sowie $\varphi_5 = 0$ und $\varphi_6 = 0$. Die Winkel φ_i entsprechen genau den in (a) gefundenen Punkten $P_i = (\cos\varphi_i, \sin\varphi_i)$. Ob jeweils ein Minimum oder ein Maximum vorliegt, kann man anhand der Funktion $g''(\varphi) = \cos\varphi(2-9\sin^2\varphi)$ ermitteln.

Lösung (98.3) Wir geben zwei Lösungen an.

• **Erste Lösung.** Nach Lagrange machen wir den Ansatz $\nabla f = \lambda\,(\nabla g)$, also

$$\begin{bmatrix} yz^2 \\ xz^2 \\ 2xyz \end{bmatrix} = 2\lambda \begin{bmatrix} x \\ y \\ z \end{bmatrix}.$$

Wir unterscheiden die Fälle $z=0$ und $z\neq 0$.

Erster Fall: Ist $z=0$, so folgt $x^2+y^2=4$, so daß nicht x und y beide gleich Null sein können; dies erzwingt $\lambda=0$. Wegen $x^2+y^2=4$ gibt es ein $\varphi\in[0,2\pi)$ mit

$$(x,y,z) = (2\cos\varphi,\, 2\sin\varphi,\, 0).$$

Zweiter Fall: Ist $z\neq 0$, so folgt $xy=\lambda$. Genau dann gilt $x=0$, wenn auch $y=0$ gilt; in diesem Fall haben wir $z^2=4$, was auf die Punkte $(0,0,\pm 2)$ führt. An diesen Punkten ist $f=0$. Gilt $xy\neq 0$, so erhalten wir $2x^2 = z^2 = 2y^2$ und damit $4 = x^2+y^2+z^2 = 4x^2$, also $x=\pm 1$. Dies führt auf die Lösungen $(\pm1, \pm1, \pm\sqrt{2})$; wir erhalten $f(1,1,\pm\sqrt{2}) = f(-1,-1,\pm\sqrt{2}) = 2$ und $f(1,-1,\pm\sqrt{2}) = f(-1,1,\pm\sqrt{2}) = -2$.

• **Zweite Lösung.** Die Menge aller Punkte (x,y,z) mit $x^2+y^2+z^2=4$ läßt sich mittels Kugelkoordinaten parametrisieren durch

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = 2\begin{bmatrix} \cos(\varphi)\cos(\theta) \\ \sin(\varphi)\cos(\theta) \\ \sin(\theta) \end{bmatrix} \qquad \begin{matrix} (0\le\varphi<2\pi, \\ -\pi/2\le\theta\le\pi/2). \end{matrix}$$

Wir erhalten dann

$$\begin{aligned} F(\varphi,\theta) &:= f\big(x(\varphi,\theta),\, y(\varphi,\theta),\, z(\varphi,\theta)\big) \\ &= 2\cos(\varphi)\cos(\theta)\cdot 2\sin(\varphi)\cos(\theta)\cdot 4\sin^2(\theta) \\ &= 16\sin(\varphi)\cos(\varphi)\sin^2(\theta)\cos^2(\theta) \\ &= 2\sin(2\varphi)\sin(2\theta)^2. \end{aligned}$$

Da die Sinusfunktion nur Werte zwischen -1 und $+1$ annimmt, können wir sofort das Maximum 2 (angenommen für $\varphi \in \{\pi/4, 5\pi/4\}$ und $\theta \in \{\pm\pi/4\}$) und das Minimum -2 (angenommen für $\varphi \in \{3\pi/4, 7\pi/4\}$ und $\theta \in \{\pm\pi/4\}$) ablesen.

Lösung (98.4) Zu optimieren ist die Funktion $f(x,y) = xy^2$ unter der Nebenbedingung $g(x,y) = 0$ mit $g(x,y) := 4x^2 + 9y^2 - 36 = 0$.

Erste Lösung. Nach Lagrange gibt es eine Zahl λ mit $\nabla f = \lambda \nabla g$, also

$$(\star) \qquad \begin{bmatrix} y^2 \\ 2xy \end{bmatrix} = \lambda \begin{bmatrix} 8x \\ 18y \end{bmatrix}.$$

Ist $\lambda = 0$, so führt $(\star)$ auf $y = 0$ und dann $x^2 = 9$, also $x = \pm 3$. Ist $\lambda \neq 0$, so haben wir $x \neq 0$ und $y \neq 0$ und wegen $(\star)$ dann $2xy : y^2 = (18y) : (8x)$, also $2x/y = (9y)/(4x)$ bzw. $8x^2 = 9y^2$. Setzt man dies in die Nebenbedingung $4x^2 + 9y^2 = 36$ ein, so ergibt sich $12x^2 = 36$ bzw. $x^2 = 3$, also $x = \pm\sqrt{3}$. Hieraus folgt dann $y^2 = 8x^2/9 = 8/3$ und damit $y = 2\sqrt{2/3}$. Wir erhalten also die sechs kritischen Punkte

$$(\pm\sqrt{3}, \pm 2\sqrt{2/3}) \quad \text{und} \quad (\pm 3, 0).$$

An den beiden Punkten $P_{1,2} = (\sqrt{3}, \pm 2\sqrt{2/3})$ nimmt f den Wert $8/\sqrt{3} > 0$ an, an den Punkten $P_{3,4} = (-\sqrt{3}, \pm 2\sqrt{2/3})$ den Wert $-8/\sqrt{3} < 0$, an den Punkten $P_{5,6} = (\pm 3, 0)$ dagegen den Wert 0. Damit ist klar, daß bei $P_{1,2}$ jeweils ein (sogar globales) Maximum vorliegt, bei $P_{3,4}$ dagegen ein (sogar globales) Minimum. Bei $P_5 = (3,0)$ liegt ein lokales Minimum vor, bei $P_6 = (-3,0)$ in lokales Maximum; das folgt beispielsweise daraus, daß entlang einer Kurve (hier der Kreislinie $x^2 + y^2 = 1$) eine Funktion zwischen zwei Minima stets ein Maximum besitzen muß, zwischen zwei Maxima stets ein Minimum.

Zweite Lösung. Wegen $(x/3)^2 + (y/2)^2 = 1$ haben wir $x = 3\cos\varphi$ und $y = 2\sin\varphi$; zu optimieren ist also der Ausdruck $xy^2 = 12\cos\varphi\sin^2\varphi$ bzw. die Funktion $g(\varphi) := \sin^2\varphi\cos\varphi$. Nullsetzen von

$$\begin{aligned} g'(\varphi) &= 2\sin\varphi\cos^2\varphi - \sin^3\varphi \\ &= \sin\varphi\,(2\cos^2\varphi - \sin^2\varphi) = \sin\varphi\,(2 - 3\sin^2\varphi) \end{aligned}$$

führt auf die Lösungen $\varphi_1 = \arcsin(\sqrt{2/3})$, $\varphi_2 = 2\pi - \varphi_1$, $\varphi_{3,4} = \pi \pm \varphi_1$ sowie $\varphi_5 = 0$ und $\varphi_6 = 0$. Die Winkel φ_i entsprechen genau den in in der ersten Lösung gefundenen Punkten $P_i = (\cos\varphi_i, \sin\varphi_i)$. Die Feststellung des Minimums und Maximums erfolgt dann wie in der ersten Lösung. (Ob jeweils ein lokales Minimum oder Maximum vorliegt, kann man auch anhand der Funktion $g''(\varphi) = \cos\varphi(2 - 9\sin^2\varphi)$ ermitteln.)

Lösung (98.5) Der Abstand zweier Punkte (x,y) und (ξ,η) ist gegeben durch $\sqrt{(x-\xi)^2 + (y-\eta)^2}$. Statt dieses Abstands können wir genauso gut die Funktion

$$f(x,y,\xi,\eta) := (x-\xi)^2 + (y-\eta)^2$$

minimieren, die etwas bequemer zu behandeln ist; diese Funktion ist zu minimieren unter den Nebenbedingungen $g_1(x,y,\xi,\eta) = g_2(x,y,\xi,\eta) = 0$ mit $g_1(x,y,\xi,\eta) := 3x^2 + 2xy + 3y^2 - 1$ und $g_2(x,y,\xi,\eta) := \xi^2 + 4\eta^2 - 4$.

(a) Nach Lagrange gibt es Zahlen $\lambda, \mu \in \mathbb{R}$ mit $\nabla f = \lambda \nabla g_1 + \mu \nabla g_2$, d. h.

$$\begin{bmatrix} 2(x-\xi) \\ 2(y-\eta) \\ 2(\xi-x) \\ 2(\eta-y) \end{bmatrix} = \lambda \begin{bmatrix} 6x+2y \\ 2x+6y \\ 0 \\ 0 \end{bmatrix} + \mu \begin{bmatrix} 0 \\ 0 \\ 2\xi \\ 8\eta \end{bmatrix},$$

nach Division durch 2 also

$$\begin{bmatrix} x-\xi \\ y-\eta \\ \xi-x \\ \eta-y \end{bmatrix} = \lambda \begin{bmatrix} 3x+y \\ x+3y \\ 0 \\ 0 \end{bmatrix} + \mu \begin{bmatrix} 0 \\ 0 \\ \xi \\ 4\eta \end{bmatrix}.$$

Bilden des Quotienten $(y-\eta)/(x-\xi)$ liefert

$$\frac{y-\eta}{x-\xi} = \frac{x+3y}{3x+y} = \frac{4\eta}{\xi}.$$

Hieraus erhalten wir die Gleichungen $(x-\xi)(x+3y) = (y-\eta)(3x+y)$ und $\xi(x+3y) = 4\eta(3x+y)$, die sich zu der Gleichung

$$\begin{bmatrix} x+3y & -(3x+y) \\ x+3y & -4(3x+y) \end{bmatrix} \begin{bmatrix} \xi \\ \eta \end{bmatrix} = \begin{bmatrix} x^2 - y^2 \\ 0 \end{bmatrix}$$

zusammenfassen lassen. Auflösen nach ξ und η liefert

$$\begin{aligned} \begin{bmatrix} \xi \\ \eta \end{bmatrix} &= \frac{1}{3(x+3y)(3x+y)} \begin{bmatrix} 4(3x+y) & -(3x+y) \\ x+3y & -(x+3y) \end{bmatrix} \begin{bmatrix} x^2-y^2 \\ 0 \end{bmatrix} \\ &= \frac{x^2-y^2}{3(x+3y)(3x+y)} \begin{bmatrix} 4(3x+y) \\ x+3y \end{bmatrix} \end{aligned}$$

bzw.

$$(\star) \qquad \xi = \frac{4(x^2-y^2)}{3(x+3y)}, \quad \eta = \frac{x^2-y^2}{3(3x+y)}.$$

Einsetzen in die Nebenbedingung $\xi^2 + 4\eta^2 = 4$ führt nun auf

$$\frac{(x^2-y^2)^2}{9(x+3y)^2(3x+y)^2}\big(16(3x+y)^2 + 4(x+3y)^2\big) = 4$$

bzw.

$$\frac{(x^2-y^2)^2\big(4(3x+y)^2 + (x+3y)^2\big)}{9(x+3y)^2(3x+y)^2} = 1.$$

Durch Gleichsetzen mit der anderen Nebenbedingung $3x^2 + 2xy + 3y^2 = 1$ ergibt sich

$$\begin{aligned} &(x^2-y^2)^2\big(4(3x+y)^2 + (x+3y)^2\big) \\ &= 9(3x^2+2xy+3y^2)(x+3y)^2(3x+y)^2. \end{aligned}$$

Dividieren wir diese Gleichung durch x^6 und setzen wir $q := y/x$, so erhalten wir

$$\begin{aligned}&(1-q^2)^2\big(4(3+q)^2+(1+3q)^2\big)\\ &= 9(3+2q+3q^2)(1+3q)^2(3+q)^2.\end{aligned}$$

(Wir haben es jetzt geschafft, statt eines Systems von sechs Gleichungen für die sechs Unbekannten $x, y, \xi, \eta, \lambda, \mu$ nur noch eine Gleichung in einer Unbekannten untersuchen zu müssen.) Durch Ausmultiplizieren, Sortieren und Zusammenfassen geht diese Gleichung über in

$$115q^6+876q^5+2249q^4+2712q^3+2285q^2+876q+103 \;=\; 0.$$

Diese Gleichung hat vier reelle Lösungen:

$$\begin{aligned} q_1 &\approx -3.70678,\\ q_2 &\approx -2.55523,\\ q_3 &\approx -0.442293,\\ q_4 &\approx -0.203805.\end{aligned}$$

Für jede dieser Lösungen q_i mit $1 \le i \le 4$ setzen wir $y = q_i x$ in die Gleichung $3x^2 + 2xy + 3y^2 = 1$ ein und erhalten $x^2(3 + 2q_i + 3q_i^2) = 1$, also die beiden Lösungen $x_i = \pm 1/\sqrt{3 + 2q_i + 3q_i^2}$. Dann ist $y_i = q_i x_i$; der Punkt (ξ, η) ergibt sich dann aus Gleichung $(\star)$. Dies liefert für $P = (x, y)$ und $Q = (\xi, \eta)$ die folgenden Möglichkeiten:

- $P_1 = (0.164829, -0.610986)$,
  $Q_1 = (0.276665, 0.990388)$;
- $P_2 = (-0.164829, 0.610986)$,
  $Q_2 = (-0.276665, -0.990388)$;
- $P_3 = (0.239202, -0.611216)$,
  $Q_3 = (0.264558, -0.991221)$;
- $P_4 = (-0.239202, 0.611216)$,
  $Q_4 = (-0.264558, 0.991221)$;
- $P_5 = (0.608323, -0.269057)$,
  $Q_5 = (-1.99593, 0.0637708)$;
- $P_6 = (-0.608323, 0.269057)$,
  $Q_5 = (1.99593, -0.0637708)$;
- $P_7 = (0.606674, -0.123643)$,
  $Q_7 = (1.99519, 0.0693174)$;
- $P_8 = (-0.606674, 0.123643)$,
  $Q_8 = (-1.99519, -0.0693174)$.

Einsetzen in die Funktion f zeigt, daß der minimale Abstand zwischen den Punkten P_3 und Q_3 bzw. P_4 und Q_4 angenommen wird (in der Abbildung grün), der maximale Abstand dagegen zwischen den Punkten P_7 und Q_7 bzw. P_8 und Q_8 (in der Zeichnung violett).

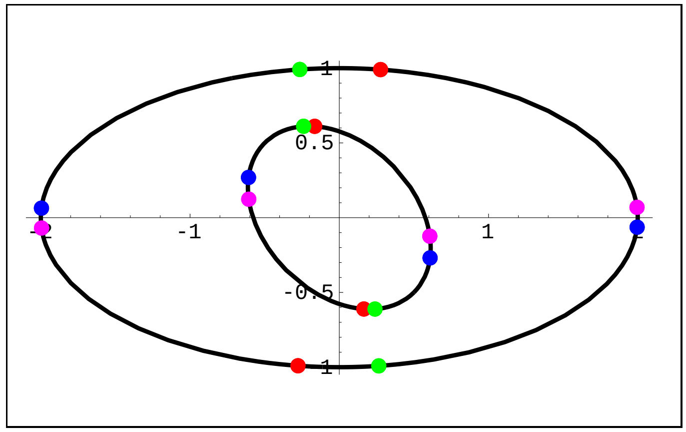


Abb. 98.5: Kritische Punkte der betrachteten Abstandsfunktion.

(b) Unter Benutzung des Hinweises können wir die Gleichung $3x^2 + 2xy + 3y^2 = 1$ auch in der Form

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

schreiben und erhalten dadurch die Parametrisierung $x - y = \cos u$, $\sqrt{2}(x + y) = \sin u$; wegen $(\xi/2)^2 + \eta^2 = 1$ können wir ferner $\xi/2 = \cos v$ und $\eta = \sin v$ schreiben. Wir erhalten also die Darstellungen

$$\begin{bmatrix} x\\ y\end{bmatrix} = \frac{1}{2\sqrt{2}}\begin{bmatrix}\sqrt{2}\cos u + \sin u\\ \sin u - \sqrt{2}\cos u\end{bmatrix} \quad \text{und} \quad \begin{bmatrix}\xi\\ \eta\end{bmatrix} = \begin{bmatrix} 2\cos v\\ \sin v\end{bmatrix}.$$

Um die Sinus- und Kosinusfunktionen zu vermeiden und nur mit rationalen Ausdrücken zu rechnen, führen wir $a := \tan(u/2)$ und $b := \tan(v/2)$ ein und erhalten dann

$$\begin{bmatrix}\cos u\\ \sin u\end{bmatrix} = \frac{1}{1+a^2}\begin{bmatrix}1-a^2\\ 2a\end{bmatrix} \text{ und } \begin{bmatrix}\cos v\\ \sin v\end{bmatrix} = \frac{1}{1+b^2}\begin{bmatrix}1-b^2\\ 2b\end{bmatrix}$$

und folglich

$$x \;=\; \frac{1-a^2+a\sqrt{2}}{2(1+a^2)} \quad \text{und} \quad y \;=\; \frac{a\sqrt{2}-1+a^2}{2(1+a^2)}$$

sowie

$$\xi \;=\; \frac{2(1-b^2)}{1+b^2} \quad \text{und} \quad \eta \;=\; \frac{2b}{1+b^2}.$$

Die zu minimierende Funktion $(x-\xi)^2 + (y-\eta)^2$ ist dann (bis auf einen Faktor 4) gegeben durch

$$\left(\frac{1-a^2+a\sqrt{2}}{1+a^2} - \frac{4(1-b^2)}{1+b^2}\right)^2 + \left(\frac{a\sqrt{2}-1+a^2}{1+a^2} - \frac{4b}{1+b^2}\right)^2$$

und damit (nach Ausmultiplizieren und Sortieren) durch die Funktion

$$\begin{aligned} F(a,b) \;=\;& \frac{2(a^4+1)}{(a^2+1)^2} + \frac{16(b^4-b^2+1)}{(b^2+1)^2}\\ &- \frac{8(a^2-a\sqrt{2}-1)(b^2-1)}{(a^2+1)(b^2+1)} - \frac{8(a^2+a\sqrt{2}-1)b}{(a^2+1)(b^2+1)}.\end{aligned}$$

Die partiellen Ableitungen sind

$$\begin{aligned}\frac{\partial F}{\partial a} &= \frac{8a^3-8a}{(a^2+1)^3} - \frac{8\sqrt{2}(b^2-1)(a^2+2\sqrt{2}\,a-1)}{(b^2+1)(a^2+1)^2}\\ &\quad + \frac{8\sqrt{2}\,b(a^2-2\sqrt{2}\,a-1)}{(b^2+1)(a^2+1)^2}\end{aligned}$$

und

$$\begin{aligned}\frac{\partial F}{\partial b} &= \frac{96b^3-96b}{(b^2+1)^3} - \frac{32(a^2-\sqrt{2}\,a-1)\,b}{(a^2+1)(b^2+1)^2}\\ &\quad + \frac{8(a^2+\sqrt{2}\,a-1)(b^2-1)}{(a^2+1)(b^2+1)^2}.\end{aligned}$$

Nullsetzen und Durchmultiplizieren mit $(a^2+1)^3(b^2+1)/8$ bzw. $(a^2+1)(b^2+1)^3/8$ und anschließendes Sortieren nach Potenzen von a führt auf das Gleichungssystem

$$(\star)\qquad \begin{aligned} 0 &= p_0(b)+p_1(b)\,a+p_2(b)\,a^2+p_3(b)\,a^3,\\ 0 &= q_0(b)+q_1(b)\,a+q_2(b)\,a^2\end{aligned}$$

mit

$$\begin{aligned} p_0(b) &:= \sqrt{2}\cdot(b^2-b-1),\\ p_1(b) &:= -3b^2-4b+3,\\ p_2(b) &:= \sqrt{2}\cdot(-b^2+b+1),\\ p_3(b) &:= b^2+1\end{aligned}$$

und

$$\begin{aligned} q_0(b) &:= -b^4+16b^3-8b+1,\\ q_1(b) &:= \sqrt{2}\cdot(b^4+4b^3+4b-1),\\ q_2(b) &:= b^4+8b^3-16b-1.\end{aligned}$$

Es folgt nun ein Standardtrick, ein System zweier Polynomgleichungen in zwei Variablen in eine einzige Polynomgleichung in einer der beiden Variablen zu überführen. Aus $(\star)$ folgt

$$\begin{bmatrix} p_0(b) & p_1(b) & p_2(b) & p_3(b) & 0\\ 0 & p_0(b) & p_1(b) & p_2(b) & p_3(b)\\ q_0(b) & q_1(b) & q_2(b) & 0 & 0\\ 0 & q_0(b) & q_1(b) & q_2(b) & 0\\ 0 & 0 & q_0(b) & q_1(b) & q_2(b)\end{bmatrix}\begin{bmatrix}1\\a\\a^2\\a^3\\a^4\end{bmatrix} = \begin{bmatrix}0\\0\\0\\0\\0\end{bmatrix}.$$

Dieses Gleichungssystem kann nur dann eine Lösung besitzen, wenn die Koeffizientendeterminante verschwindet; die einzigen möglichen Werte von b sind also die Lösungen der Polynomgleichung

$$\det\begin{bmatrix} p_0(b) & p_1(b) & p_2(b) & p_3(b) & 0\\ 0 & p_0(b) & p_1(b) & p_2(b) & p_3(b)\\ q_0(b) & q_1(b) & q_2(b) & 0 & 0\\ 0 & q_0(b) & q_1(b) & q_2(b) & 0\\ 0 & 0 & q_0(b) & q_1(b) & q_2(b)\end{bmatrix} = 0,$$

und für jede dieser Lösungen müssen wir dann die zugehörigen möglichen Lösungen für a finden (was am bequemsten durch Betrachten der zweiten Gleichung in $(\star)$ erfolgt, die quadratisch in a ist). Die erhaltenen Lösungen (a,b) sind dann in die Funktion F einzusetzen, um zu sehen, wo F sein Minimum bzw. Maximum annimmt. Damit ist klar, wie man prinzipiell die gesuchte Lösung (a,b) findet. Auf die Durchführung der mühseligen Rechnungen verzichten wir hier, zumal aus der vorigen Aufgabe bereits klar ist, was als Ergebnis herauskommen muß.

Lösung (98.6) Geometrisch bedeutet die Aufgabenstellung, den kürzesten Abstand eines Punktes (x,y) auf der Parabel $x=-y^2$ und einem Punkt (ξ,η) auf der Hyperbel $\xi\eta=1$ zu finden (wobei wir nur den rechten Hyperbelast $\xi>0$ betrachten; der andere schneidet den Parabelbogen).

(a) Wir definieren Funktionen $g_1,g_2:\mathbb{R}^4\to\mathbb{R}$ durch $g_1(x,y,\xi,\eta):=x+y^2$ und $g_2(x,y,\xi,\eta):=\xi\eta-1$. Es gibt dann Zahlen $\lambda,\mu\in\mathbb{R}$ mit $\nabla f=\lambda\nabla g_1+\mu\nabla g_2$, also

$$\begin{bmatrix}x-\xi\\y-\eta\\\xi-x\\\eta-y\end{bmatrix} = \lambda\begin{bmatrix}1\\2y\\0\\0\end{bmatrix}+\mu\begin{bmatrix}0\\0\\\eta\\\xi\end{bmatrix}.$$

Fall 1: Es gilt $x-\xi=0$. In diesem Fall folgt $\lambda=0$ und damit $y-\eta=0$. Wir haben dann $x=-y^2$ und $xy=1$, folglich $x=1/y$ und damit $y^3=-1$. Dies liefert die Lösung $x=y=\xi=\eta=-1$ und führt auf den Schnittpunkt $(-1,-1)$ der Parabel $x=-y^2$ mit der Hyperbel $xy=1$.

Fall 2: Es gilt $x-\xi\neq 0$. In diesem Fall haben wir

$$2y = \frac{y-\eta}{x-\xi} = \frac{\xi}{\eta},$$

folglich $\xi-2y\eta=0$ und $2y(x-\xi)=y-\eta$ bzw.

$$\begin{bmatrix}2y & -1\\1 & -2y\end{bmatrix}\begin{bmatrix}\xi\\\eta\end{bmatrix} = \begin{bmatrix}2xy-y\\0\end{bmatrix}$$

und damit

$$\begin{bmatrix}\xi\\\eta\end{bmatrix} = \frac{1}{1-4y^2}\begin{bmatrix}-2y & 1\\-1 & 2y\end{bmatrix}\begin{bmatrix}2xy-y\\0\end{bmatrix} = \frac{y(1-2x)}{1-4y^2}\begin{bmatrix}2y\\1\end{bmatrix}.$$

Folglich ist

$$1 = \xi\eta = \frac{y^2(1-2x)^2}{(1-4y^2)^2}\cdot 2y = \frac{2y^3(1+2y^2)^2}{(1-4y^2)^2}$$

und damit $(1-4y^2)^2=2y^3(1+2y^2)^2$. Ausmultiplizieren und Sortieren führt auf die Gleichung

$$8y^7+8y^5-16y^4+2y^3+8y^2-1 = 0.$$

Diese Gleichung hat eine einzige reelle Lösung, nämlich $y\approx 0.379129$. Hieraus erhalten wir dann $x=-y^2\approx -0.143739$ sowie $\xi\approx 0.870780$ und $\eta\approx 1.148396$. Die beiden Punkte (x,y) und (ξ,η) mit dem kleinstmöglichen Abstand sind in der folgenden Skizze eingezeichnet.

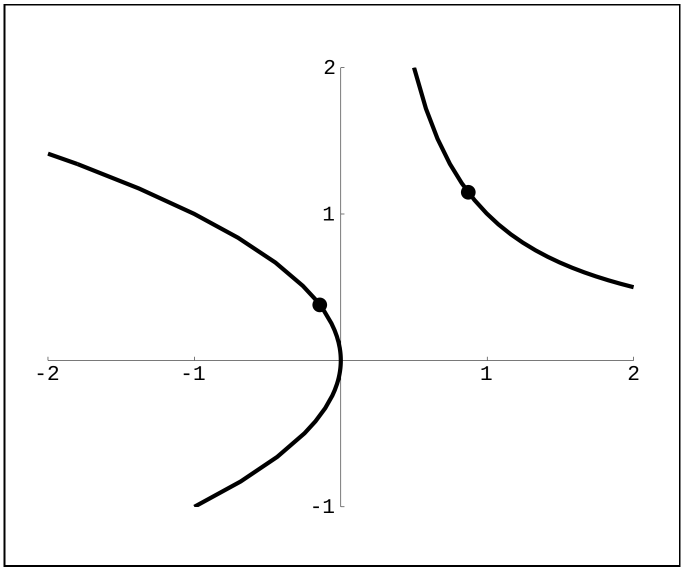


Abb. 98.6: Kürzester Abstand zwischen der Parabel $x = -y^2$ und dem Hyperbelast $y = 1/x$ mit $x > 0$.

Bemerkung: Die Abbildung läßt vermuten, daß die Verbindungslinie der beiden Punkte die jeweiligen Kurven jeweils senkrecht trifft. In Aufgabe (98.12) werden wir sehen, daß dies kein Zufall ist!

(b) Wir müssen die Funktion

$$F(y,\eta) \;:=\; f(-y^2,\, y,\, 1/\eta,\, \eta) \;=\; \frac{(y^2+1/\eta)^2+(y-\eta)^2}{2}$$

minimieren und damit die partiellen Ableitungen von F gleich Null setzen. Wir erhalten

$$\begin{aligned}\frac{\partial F}{\partial y} &= 2y(y^2+1/\eta)+(y-\eta) \;=\; 0 \quad \text{und}\\ \frac{\partial F}{\partial \eta} &= -(y^2+1/\eta)/\eta^2+(\eta-y) \;=\; 0,\end{aligned}$$

folglich entweder $y^2+1/\eta = y-\eta = 0$ und damit $y=\eta=-1$ oder aber

$$\frac{y-\eta}{y^2+1/\eta} \;=\; -2y \;=\; -1/\eta^2$$

und damit $y = 1/(2\eta^2)$. Im ersten Fall erhalten wir $(x,y) = (\xi,\eta) = (-1,-1)$. Im zweiten Fall ergeben sich die Gleichungen $x = -y^2 = -1/(4\eta^4)$ sowie $\xi = 1/\eta$; die Gleichung $0 = \partial F/\partial y$ geht damit über in

$$\begin{aligned}0 &= \frac{1}{\eta^2}\left(\frac{1}{4\eta^4}+\frac{1}{\eta}\right)+\frac{1}{2\eta^2}-\eta\\ &= \frac{-4\eta^7+2\eta^4+4\eta^3+1}{4\eta^6}.\end{aligned}$$

Also erfüllt η die Gleichung $4\eta^7-2\eta^4-4\eta^3-1=0$; diese Gleichung hat eine einzige reelle Nullstelle, nämlich $\eta \approx 1.148396$. Wir erhalten dann $\xi = 1/\eta \approx 0.870780$, weiter $x = -1/(4\eta^4) \approx -0.143739$ und $y = 1/(2\eta^2) \approx 0.379129$. Dies ist natürlich das gleiche Ergebnis wie das in (a) erhaltene.

Lösung (98.7) Da das Rotationsparaboloid die Parametrisierung $(x,y) \mapsto (x,y,x^2+y^2)$ besitzt, können wir die Aufgabe direkt (ohne Anwendung der Methode von Lagrange) lösen. Statt des Abstandes können wir genausogut die Funktion

$$F(x,y,z) \;:=\; \frac{1}{2}\big((x-1)^2+(y-3)^2+(x^2+y^2-2)^2\big)$$

minimieren, die etwas bequemer zu handhaben ist. Wegen $F(x,y,z) \to \infty$ für $(x,y,z) \to \infty$ ist dabei die Existenz eines globalen Minimums von vornherin klar. Nullsetzen der partiellen Ableitungen führt auf

$$\begin{aligned}0 &= F_x \;=\; x-1+(x^2+y^2-2)\cdot 2x,\\ 0 &= F_y \;=\; y-3+(x^2+y^2-2)\cdot 2y.\end{aligned}$$

Durchmultiplizieren der ersten Gleichung mit y und der zweiten Gleichung mit x liefert

$$y(x-1) \;=\; -2xy(x^2+y^2-2) \;=\; x(y-3)$$

und damit $xy-y = xy-3x$, also $y = 3x$. Setzen wir dies in die erste Gleichung ein, so ergibt sich $0 = x-1+2x(10x^2-2) = 20x^3-3x-1$. Ausprobieren zeigt, daß $x = 1/2$ eine Lösung dieser Gleichung ist; anschließende Polynomdivision liefert $(20x^3-3x-1):(2x-1) = 10x^2+5x+1 = 10\big(x+(1/4)\big)^2+(3/8)$, so daß keine weitere Lösung mehr existiert. Wir erhalten also $x = 1/2$, $y = 3x = 3/2$ und $z = x^2+y^2 = 5/2$. Das Rotationsparaboloid kommt also im Punkt $(1/2,3/2,5/2)$ dem Punkt $(1,3,2)$ am nächsten.

Wir lösen die Aufgabe auch noch mit der Methode von Lagrange. Zu minimieren ist die Funktion

$$f(x,y,z) \;:=\; (x-1)^2+(y-3)^2+(z-1)^2$$

unter der Nebenbedingung $g(x,y,z) = 0$ mit $g(x,y,z) := x^2+y^2-z$. Wir machen nach Lagrange den Ansatz $\nabla f = \lambda \nabla g$, also

$$\begin{bmatrix}2(x-1)\\2(y-3)\\2(z-2)\end{bmatrix} \;=\; \lambda \begin{bmatrix}2x\\2y\\-1\end{bmatrix}.$$

Aus den beiden ersten Komponenten dieser Vektorgleichung ergibt sich $2y(x-1) = 2\lambda xy = 2x(y-3)$, damit $xy-y = xy-3x$ und folglich $y = 3x$. Die erste Komponente liefert $x \neq 0$ und dann $\lambda = (x-1)/x$. Die letzte Komponente liefert $\lambda = 4-2z$. Vergleich der beiden letzten Gleichungen ergibt $4-2z = (x-1)/x$ bzw. $z = (3x+1)/(2x)$. Die Gleichung $z = x^2+y^2$ geht damit über in

$$\frac{3x+1}{2x} \;=\; x^2+(3x)^2 \;=\; 10x^2$$

bzw. $20x^3-3x-1=0$. Wieder ergibt sich $x = 1/2$, damit $y = 3x = 3/2$ und folglich $z = x^2+y^2 = 5/2$.

Lösung (98.8) Der Mittelpunkt der Ellipse E ist der Nullpunkt $(0,0,0)$; die kleine bzw. große Halbachse von E ist daher der minimale bzw. maximale Abstand eines Punktes $(x,y,z) \in E$ zum Nullpunkt. Wir suchen daher das Minimum und das Maximum der Funktion

$$f(x,y,z) \;:=\; x^2+y^2+z^2$$

unter den Nebenbedingungen $g_1 = g_2 = 0$ mit $g_1(x,y,z) := (x/2)^2+(y/\sqrt{5})^2+(z/5)^2-1$ und $g_2(x,y,z) := x+y-z$. (Wegen der Stetigkeit von f und der Kompaktheit von E ist die Existenz eines Minimums bzw. Maximums von vornherein klar.) Nach Lagrange machen wir den Ansatz $\nabla f = \lambda(\nabla g_1) + \mu(\nabla g_2)$, also

$$2\begin{bmatrix} x \\ y \\ z \end{bmatrix} \;=\; 2\lambda \begin{bmatrix} x/4 \\ y/5 \\ z/25 \end{bmatrix} + \mu \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix}.$$

Damit diese Gleichung gelten kann, müssen die in ihr auftretenden Spaltenvektoren linear abhängig sein; es muß also

$$0 \;=\; \det \begin{bmatrix} x & x/4 & 1 \\ y & y/5 & 1 \\ z & z/25 & -1 \end{bmatrix} \;=\; \frac{xy}{20} - \frac{4yz}{25} + \frac{21xz}{100}$$

bzw. $0 = 5xy - 16yz + 21xz$ gelten. (Die Lagrange-Multiplikatoren λ und μ, deren Werte uns ohnehin nicht interessieren, wurden durch die Herleitung der letzten Gleichung eliminiert.) Setzen wir die Nebenbedingung $z = x+y$ in die letzte Gleichung ein, so erhalten wir $0 = 5xy-16y(x+y)+21x(x+y) = 21x^2+10xy-16y^2$. Die andere Nebenbedingung $g_1(x,y,z) = 0$ läßt sich schreiben als $100 = 25x^2+20y^2+4z^2 = 25x^2+20y^2+4(x+y)^2 = 29x^2+8xy+24y^2$. Dividieren wir die Gleichung $0 = 21x^2+10xy-16y^2$ durch $-16x^2$ und schreiben wir $q := y/x$, so erhalten wir $q^2-(5/8)q-(21/16)$ und damit $q = (5/16) \pm (19/16)$, also $q = 3/2$ oder $q = -7/8$. Wir haben also entweder $y = 3x/2$ oder aber $y = -7x/8$. Diese Bedingung wird nun jeweils in die Gleichung $29x^2+8xy+24y^2 = 100$ eingesetzt.

- **Erster Fall:** $y = 3x/2$. Dies liefert $100 = 29x^2 + 12x^2+54x^2 = 95x^2$ bzw. $x^2 = 29/19$. Wegen $y = 3x/2$ und $z = x+y = 5x/2$ liefert dies die beiden Lösungen

  $$(x,y,z) \;=\; \pm\left(\sqrt{\frac{29}{19}}, \frac{3}{2}\sqrt{\frac{29}{19}}, \frac{5}{2}\sqrt{\frac{29}{19}}\right)$$

  mit

  $$x^2+y^2+z^2 \;=\; \frac{29}{19}\left(\frac{4}{4}+\frac{9}{4}+\frac{25}{4}\right) \;=\; \frac{29}{2}.$$

- **Zweiter Fall:** $y = -7x/8$. Dies liefert $100 = 29x^2 - 7x^2 + (147/8)x^2$ bzw. $x^2 = 800/323$. Wegen $y = -7x/8$ und $z = x+y = x/8$ liefert dies die beiden Lösungen

  $$(x,y,z) \;=\; \pm\left(\sqrt{\frac{800}{323}}, -\frac{7}{8}\sqrt{\frac{800}{323}}, \frac{1}{8}\sqrt{\frac{800}{323}}\right) \;=\; \frac{75}{17}.$$

Die beiden Halbachsen sind also $\sqrt{29/2} \approx 3.80789$ und $\sqrt{75/17} \approx 2.10042$.

Lösung (98.9) Die Punkte auf P sind genau die Punkte der Form (a, a^2+1), die Punkte auf G die Punkte der Form $(2b, b)$; wir müssen also $a, b \in \mathbb{R}$ so finden, daß die Funktion

$$f(a,b) \;:=\; (a-2b)^2 + (a^2+1-b)^2$$

(die das Quadrat des Abstandes zwischen den Punkten (a, a^2+1) und $(2b, b)$ angibt) minimal wird. Wir erhalten die partiellen Ableitungen

$$\begin{aligned} \frac{\partial f}{\partial a} &= 2(a-2b) + 2(a^2+1-b)\cdot 2a, \\ \frac{\partial f}{\partial b} &= 2(a-2b)\cdot(-2) + 2(a^2+1-b)\cdot(-1); \end{aligned}$$

Nullsetzen führt auf

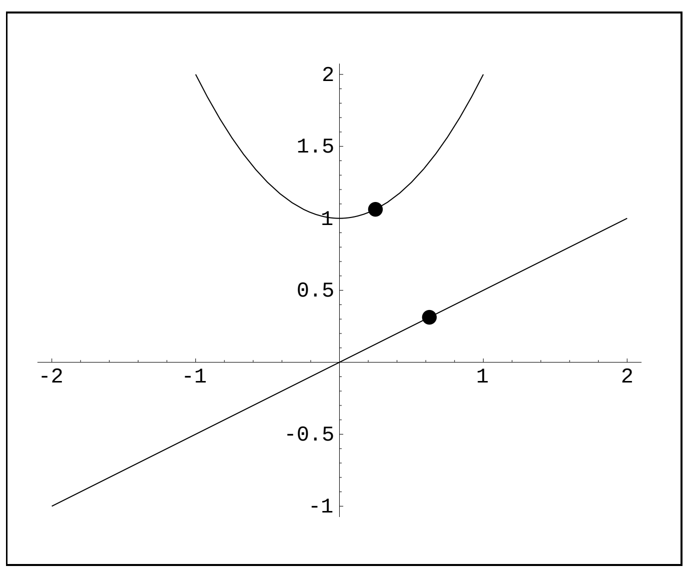
(1) $\quad a - 2b + 2a(a^2+1-b) \;=\; 0,$

(2) $\quad 2(a-2b) + (a^2+1-b) \;=\; 0.$

Aus (2) folgt

(3) $\quad a^2+1-b \;=\; -2(a-2b);$

Einsetzen in (1) liefert dann $0 = a-2b-4a(a-2b) = (a-2b)(1-4a)$ und damit $a = 2b$ oder $a = 1/4$. Einsetzen der ersten Möglichkeit in (3) liefert $4b^2-b+1 = 0$; da diese Gleichung keine reelle Lösung hat, bleibt nur die zweite Möglichkeit $a = 1/4$. Einsetzen in (3) liefert $(17/16) - b = 4b - (1/2)$ und damit $5b = 25/16$, also $b = 5/16$. Die Funktion f hat also nur einen einzigen kritischen Punkt, nämlich $(a,b) = (1/4,\, 5/16)$. Wegen $f(a,b) \to \infty$ für $(a,b) \to \infty$ (genaue Begründung?) besitzt f zwangsläufig ein (sogar globales) Minimum (genaue Begründung?); also liegt an der Stelle $(a,b) = (1/4,\, 5/16)$ ein globales Minimum vor. Der kürzeste Abstand wird also zwischen den Punkten $(a, a^2+1) = (1/4,\, 17/16)$ und $(2b, b) = (5/8,\, 5/16)$ angenommen.



Lösung (98.10) Es sei (a,b,c,d,e) ein Punkt, an dem f unter den angegebenen Nebenbedingungen ein lokales Minimum oder Maximum annimmt. Die Methode von Lagrange liefert die Existenz von Zahlen $u,v,w\in\mathbb{R}$ mit

$$\begin{bmatrix}4a^3\\4b^3\\4c^3\\4d^3\\4e^3\end{bmatrix} = u\begin{bmatrix}1\\1\\1\\1\\1\end{bmatrix} + v\begin{bmatrix}2a\\2b\\2c\\2d\\2e\end{bmatrix} + w\begin{bmatrix}3a^2\\3b^2\\3c^2\\3d^2\\3e^2\end{bmatrix}.$$

Alle fünf Zahlen a,b,c,d,e sind dann Nullstellen des Polynoms $p(x):=4x^3-3wx^2-2vx-u$. Ein Polynom dritten Grades hat aber höchstens drei verschiedene Nullstellen; daher treten unter den fünf Zahlen a,b,c,d,e mindestens drei gleiche Zahlen oder zwei Paare jeweils gleicher Zahlen auf. Da sowohl die Funktion f als auch die Nebenbedingungen vollkommen symmetrisch in a,b,c,d,e sind, dürfen wir also annehmen, daß die Gleichung $a=b=c$ gilt oder daß die Gleichungen $a=b$ und $c=d$ gelten. (Andere Lösungen ergeben sich dann durch Umbenennen der Variablen.)

Erster Fall: $a=b=c$. In diesem Fall lauten die Nebenbedingungen folgendermaßen:

$$\begin{aligned}0 &= 3a+d+e,\\ 4 &= 3a^2+d^2+e^2,\\ 0 &= 3a^3+d^3+e^3.\end{aligned}$$

Die erste Gleichung liefert $a=-(d+e)/3$; Einsetzen in die zweite und die dritte Gleichung liefert $2d^2+de+2e^2=6$ sowie $(d+e)^3=9(d^3+e^3)=9(d+e)(d^2-de+e^2)$. Aufgrund der letzten Bedingung gilt $d+e=0$ oder $(d+e)^2=9(d^2-de+e^2)$ bzw. $8d^2-11de+8e^2=0$. Wir betrachten die beiden Möglichkeiten separat.

- Gilt $d+e=0$, also $e=-d$, so geht die Gleichung $2d^2+de+2e^2=6$ über in $d^2=2$; ferner ist $a=-(d+e)/3=0$. Dies liefert die zwei Lösungen

$$(1)\qquad (a,b,c,d,e) = \pm(0,0,0,\sqrt{2},-\sqrt{2}).$$

- Die Gleichung $0=8d^2-11de+8e^2=8\big(d^2-(11/8)de+e^2\big)=8\big((d-11e/16)^2+(135e^2/258)\big)$ hat $(d,e)=(0,0)$ als einzige reelle Lösung. Diese Möglichkeit kann also nicht eintreten.

Zweiter Fall: $a=b$ und $c=d$. In diesem Fall lauten die Nebenbedingungen folgendermaßen:

$$\begin{aligned}0 &= 2a+2c+e,\\ 4 &= 2a^2+2c^2+e^2,\\ 0 &= 2a^3+2c^3+e^3.\end{aligned}$$

Die erste Gleichung liefert $e=-2(a+c)$; Einsetzen in die zweite und die dritte Gleichung liefert $3a^2+4ac+3c^2=2$ und $0=a^3+c^3-4(a+c)^3=-3(a+c)(a^2+3ac+c^2)$. Aufgrund der letzten Bedingung gilt $a+c=0$ oder $a^2+3ac+c^2=0$. Wir betrachten die beiden Möglichkeiten separat.

- Gilt $a+c=0$, also $c=-a$, so geht die Gleichung $3a^2+4ac+3c^2=2$ über in $a^2=1$ und damit $a=\pm1$. Ferner gilt $e=-2(a+c)=0$; dies liefert die beiden Lösungen

$$(2)\qquad (a,b,c,d,e) = \pm(1,1,-1,-1,0).$$

- Gelten die Gleichungen $3a^2+4ac+3c^2=2$ und $a^2+3ac+c^2=0$, so subtrahieren wir von der ersten das Dreifache der zweiten Gleichung und erhalten $ac=-2/5$, also $c=-2/(5a)$. Einsetzen in die Gleichung $a^2+3ac+c^2=0$ liefert dann $0=a^2-(6/5)+4/(25a^2)$ bzw. $0=a^4-(6/5)a^2+(4/25)$ und damit $a^2=(3\pm\sqrt{5})/5$, also $a=\pm\sqrt{(3\pm\sqrt{5})/5}$. Hieraus folgt $c=-2/(5a)=\mp\sqrt{(3\mp\sqrt{5})/5}$, wobei sowohl vor der Wurzel als auch innerhalb der Wurzel jeweils dasjenige Vorzeichen zu wählen ist, das dem im Ausdruck für a auftretenden Vorzeichen entgegengesetzt ist. Mit $\alpha:=\sqrt{(3+\sqrt{5})/5}$ und $\beta:=\sqrt{(3-\sqrt{5})/5}$ erhalten wir also die Lösungen

$$(3)\qquad \begin{aligned}(a,b,c,d,e) &= \pm(\alpha,\alpha,-\beta,-\beta,2\beta-2\alpha) \text{ und}\\ (a,b,c,d,e) &= \pm(\beta,\beta,-\alpha,-\alpha,2\alpha-2\beta).\end{aligned}$$

Damit haben wir alle kritischen Punkte bestimmt. Wegen $f(0,0,0,\sqrt{2},-\sqrt{2})=8$, $f(1,1,-1,-1,0)=4$ und $f(\alpha,\alpha,-\beta,-\beta,2\beta-2\alpha)=f(\beta,\beta,-\alpha,-\alpha,2\alpha-2\beta)=24/5=4.8$ nimmt also f unter den angegebenen Nebenbedingungen an den Punkten der Form (2) das (globale) Minimum 4 an, an den Punkten der Form (1) das (globale) Maximum 8.

Lösung (98.11) Die Menge M aller 2×2-Matrizen, deren erste Spalte die Länge 1 und deren zweite Spalte die Länge 2 hat, ist eine 2-dimensionale Untermannigfaltigkeit des 4-dimensionalen Vektorraums $\mathbb{R}^{2\times2}$ aller reellen (2×2)-Matrizen. Wir können M sowohl in Parameterform als auch in Gleichungsform darstellen:

$$\begin{aligned}M &= \{\begin{bmatrix}\cos u & 2\cos v\\ \sin u & 2\sin v\end{bmatrix} \mid u,v\in\mathbb{R}\}\\ &= \{\begin{bmatrix}a & b\\ c & d\end{bmatrix} \mid a^2+c^2=1,\, b^2+d^2=4\}.\end{aligned}$$

Im ersten Fall ist die zu optimierende Funktion gegeben durch

$$F(u,v) := \det\begin{bmatrix}\cos u & 2\cos v\\ \sin u & 2\sin v\end{bmatrix} = 2\sin(v-u).$$

Diese Funktion nimmt ihren maximalen Wert (nämlich 2) genau dann an, wenn $\sin(v-u)=1$ gilt, also $v=$

$u+(\pi/2)+k\cdot 2\pi$ mit $k\in\mathbb{Z}$ und damit $\cos v=-\sin u$ und $\sin v=\cos u$, also auf allen Matrizen der Form

$$\begin{bmatrix}\cos u & -2\sin u\\ \sin u & 2\cos u\end{bmatrix}.$$

Analog nimmt die Funktion F ihren minimalen Wert (nämlich -2) genau dann an, wenn $\sin(v-u)=-1$ gilt, also $v=u-\pi/2+k\cdot 2\pi$ mit $k\in\mathbb{Z}$ und damit $\cos v=\sin u$ und $\sin u=-\cos v$, also auf allen Matrizen der Form

$$\begin{bmatrix}\cos u & 2\sin u\\ \sin u & -2\cos u\end{bmatrix}.$$

Benutzen wir die Gleichungsdarstellung für M, so haben wir

$$A\ =\ \begin{bmatrix}a & b\\ c & d\end{bmatrix}\quad \text{mit } a^2+c^2=1 \text{ und } b^2+d^2=4.$$

Zu minimieren ist dann die Funktion $f(a,b,c,d):=ad-bc$ unter den Nebenbedingungen $g_1(a,b,c,d):=a^2+c^2-1=0$ und $g_2(a,b,c,d):=b^2+d^2-4=0$ zu minimieren bzw. zu maximieren. Der Lagrange-Ansatz $\nabla f=\lambda(\nabla g_1)+\mu(\nabla g_2)$ lautet

$$\begin{bmatrix}d\\ -c\\ -b\\ a\end{bmatrix}\ =\ 2\lambda\begin{bmatrix}a\\ 0\\ c\\ 0\end{bmatrix}+2\mu\begin{bmatrix}0\\ b\\ 0\\ d\end{bmatrix}.$$

Hieraus folgen die Bedingungen

$$\begin{bmatrix}d\\ -b\end{bmatrix}=2\lambda\begin{bmatrix}a\\ c\end{bmatrix}\quad\text{und}\quad\begin{bmatrix}-c\\ a\end{bmatrix}=2\mu\begin{bmatrix}b\\ d\end{bmatrix},$$

nach Normbildung also $2=2|\lambda|$ und $1=4|\mu|$ bzw. $\lambda=\pm 1$ und $\mu=\pm 1/4$. Wir erhalten dann entweder

$$A\ =\ \begin{bmatrix}a & b\\ c & d\end{bmatrix}=\begin{bmatrix}a & -2c\\ c & 2a\end{bmatrix}\quad\text{mit}\quad \det(A)\ =\ 2$$

(in diesem Fall haben wir $\lambda=1$ und $\mu=1/4$) oder

$$A\ =\ \begin{bmatrix}a & b\\ c & d\end{bmatrix}=\begin{bmatrix}a & 2c\\ c & -2a\end{bmatrix}\quad\text{mit}\quad \det(A)\ =\ -2$$

(in diesem Fall haben wir $\lambda=-1$ und $\mu=-1/4$). Da die Vektoren der Länge 1 genau die Vektoren der Form $(a,c)^T=(\cos u,\sin u)^T$ sind, ist dies das gleiche Ergebnis wie das, das wir unter Benutzung der Parameterdarstellung erhalten haben.

Lösung (98.12) Die Funktion $f:M\times N\to\mathbb{R}$ mit

$$f(x,y)\ :=\ (1/2)\,\|x-y\|^2$$

nehme ein lokales Minimum in $x_0\in M$ und $y_0\in N$ an. Sind dann $I,J\subseteq\mathbb{R}$ Intervalle mit $0\in I$ und $0\in J$ und sind $\alpha:I\to\mathbb{R}^n$ und $\beta:J\to\mathbb{R}^n$ irgendwelche ganz in M bzw. N verlaufenden Kurven mit $\alpha(0)=x_0$ und $\beta(0)=y_0$, so nimmt die durch

$$F(s,t)\ :=\ (1/2)\,\|\alpha(s)-\beta(t)\|^2$$

definierte Funktion $F:I\times J\to\mathbb{R}$ ein lokales Minimum in $(s,t)=(0,0)$ an. Eine dafür notwendige Bedingung ist das Verschwinden der partiellen Ableitungen

$$\begin{aligned}\frac{\partial F}{\partial s}(0,0)\ &=\ \langle\alpha(0)-\beta(0),\alpha'(0)\rangle\ =\ \langle x_0-y_0,\alpha'(0)\rangle,\\ \frac{\partial F}{\partial t}(0,0)\ &=\ -\langle\alpha(0)-\beta(0),\beta'(0)\rangle\ =\ -\langle x_0-y_0,\beta'(0)\rangle.\end{aligned}$$

Da $\alpha'(0)$ und $\beta'(0)$ beliebige Elemente von $T_{x_0}M$ bzw. $T_{y_0}N$ sind, bedeutet dies $\langle x_0-y_0,v\rangle=\langle x_0-y_0,w\rangle=0$ für alle $v\in T_{x_0}M$ und alle $w\in T_{y_0}N$. Das ist aber die Behauptung.

Lösung (98.13) Wir definieren $f:\mathbb{R}^n\times\mathbb{R}^n\to\mathbb{R}$ und $g:\mathbb{R}^n\to\mathbb{R}$ durch

$$f(x,y):=\|x-y\|^2\quad\text{und}\quad g(x):=c_1x_1^4+\cdots+c_nx_n^4;$$

zu maximieren ist dann der Ausdruck $f(x,y)$ unter den Nebenbedingungen $g(x)=0$ und $g(y)=0$. Wird das Maximum an einer Stelle (x,y) angenommen, so gibt es nach Lagrange Konstanten $\lambda,\mu\in\mathbb{R}$ mit

$$(\star)\qquad 2\begin{bmatrix}x-y\\ y-x\end{bmatrix}\ =\ \lambda\begin{bmatrix}(\nabla g)(x)\\ 0\end{bmatrix}+\mu\begin{bmatrix}0\\ (\nabla g)(y)\end{bmatrix}.$$

Dabei sind λ und μ beide von Null verschieden, denn aus $\lambda=0$ oder $\mu=0$ würde $x=y$ und damit $f(x,y)=0$ folgen, was offensichtlich der *minimal* mögliche Wert für f ist. Also gibt es eine Konstante $\alpha\neq 0$ mit $(\nabla g)(y)=\alpha\,(\nabla g)(x)$ bzw. $y_i^3=\alpha x_i^3$ für alle i; mit $a:=\sqrt[3]{\alpha}$ gilt dann $y_i=ax_i$ für alle i. Wegen $g(x)=g(y)=0$ folgt hieraus $a^4=1$ und damit $a=\pm 1$. Die Bedingung $a=1$ würde wieder auf $y=x$ führen; also gilt $a=-1$ und damit $y=-x$. (Die beiden gesuchten Punkte gehen also durch Punktspiegelung am Nullpunkt ineinander über.) Bedingung $(\star)$ reduziert sich dann auf $4x=\lambda(\nabla g)(x)$ bzw. $x_i=\lambda\cdot c_ix_i^3$ für $1\le i\le n$. Hieraus folgt entweder $x_i=0$ oder aber $1=\lambda\cdot c_ix_i^2$. (Da nicht $x_i=0$ für alle Indices i gelten kann, folgt hieraus $\lambda>0$.) Es sei nun J die Menge derjenigen Indices i mit $x_i\neq 0$. Wir haben dann $x_i^2=1/(\lambda c_i)$ für $i\in J$ und $x_i=0$ für $i\notin J$, folglich $0=g(x)=\sum_{i\in J}1/(\lambda^2c_i)-1$ und damit $\lambda=\sqrt{\sum_{i\in J}1/c_i}$. Es folgt $f(x,y)=f(x,-x)=4\|x\|^2=4\sum_{i\in J}x_i^2=4\sum_{i\in J}1/(c_i\lambda)=(4/\lambda)\sum_{i\in J}(1/c_i)=4\sqrt{\sum_{i\in J}1/c_i}$. Da der Wert $f(x,y)$ so groß wie möglich sein soll, muß $J=\{1,\ldots,n\}$ gelten; d.h., für die optimale Lösung gilt $x_i\neq 0$ für alle i. Der größtmögliche Abstand ist also

$$4\cdot\sqrt{\frac{1}{c_1}+\frac{1}{c_2}+\cdots+\frac{1}{c_n}};$$

er wird jeweils angenommen zwischen den Punkten x und $-x$, wobei x von der Form

$$x = \frac{1}{\sqrt[4]{(1/c_1)+\cdots+(1/c_n)}} \begin{bmatrix} \pm 1/\sqrt{c_1} \\ \pm 1/\sqrt{c_2} \\ \vdots \\ \pm 1/\sqrt{c_n} \end{bmatrix}$$

(mit beliebigen Vorzeichenauswahlen) ist. **Beispiel:** Der maximale Abstand zwischen zwei Punkten der Menge $\{(x,y)\in\mathbb{R}^2 \mid 2x^4+3y^4=1\}$ wird angenommen zwischen den Punkten $\pm(\sqrt[4]{3/10}, \sqrt[4]{2/15})$ bzw. $\pm(-\sqrt[4]{3/10}, \sqrt[4]{2/15})$.

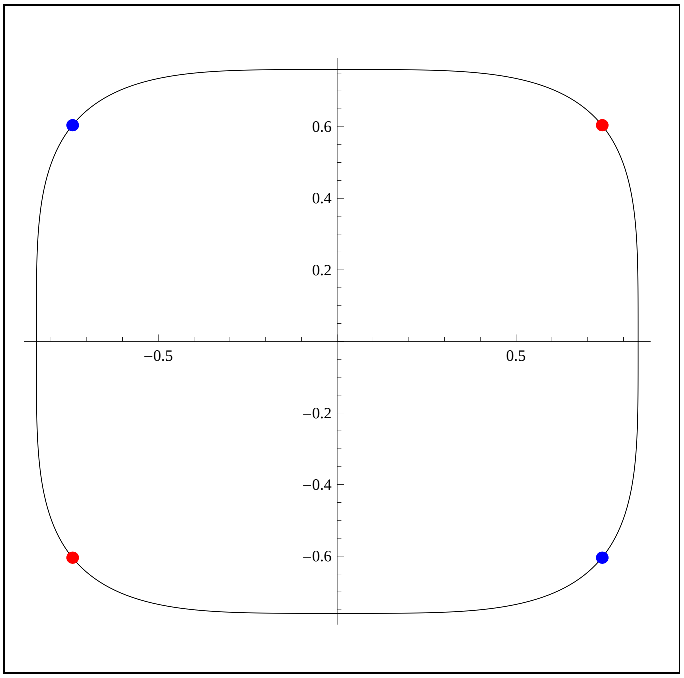


Abb. 98.13: Maximaler Abstand zwischen zwei Punkten der Menge $\{(x,y)\in\mathbb{R}^2 \mid 2x^4+3x^4=1\}$.

Lösung (98.14) Wir lösen die Aufgabe zunächst für eine beliebige, aber fest gewählte Zahl $n\in\mathbb{N}$. Zu minimieren ist die Funktion $f(a_1,\ldots,a_n) := a_1a_2\cdots a_n$ unter der Nebenbedingung $g(a_1,\ldots,a_n)=0$ mit $g(a_1,\ldots,a_n) := a_1+a_2+\cdots+a_n-a$. Der Lagrange-Ansatz $(\nabla f)=\lambda(\nabla g)$ führt auf

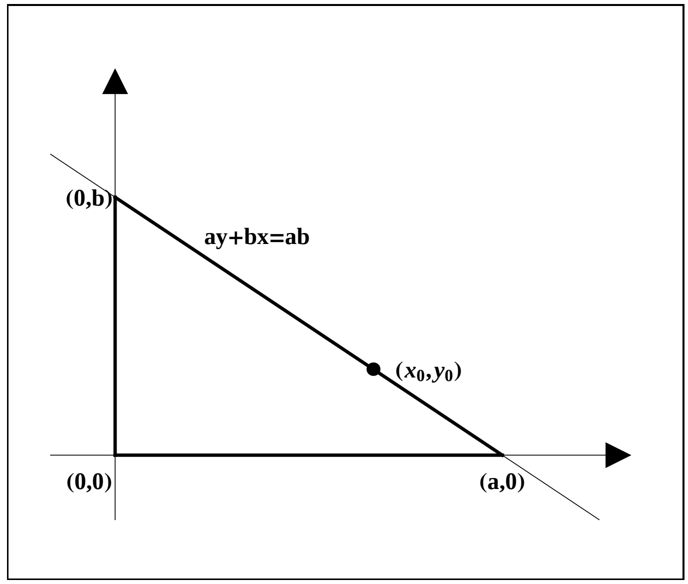
$$\begin{bmatrix} f(a)/a_1 \\ \vdots \\ f(a)/a_n \end{bmatrix} = \lambda \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix};$$

hieraus folgt sofort $a_1=a_2=\cdots=a_n$, aufgrund der Nebenbedingung also $a_i=a/n$ für alle i. Dies liefert (für fest gewähltes n) das maximale Produkt $(a/n)^n$. Wir müssen nun noch klären, für welchen Wert $n\in\mathbb{N}$ dieser Ausdruck maximal wird. Dazu betrachten wir die Funktion $\varphi:(0,\infty)\to(0,\infty)$ mit

$$\varphi(x) := \left(\frac{a}{x}\right)^x = e^{x\ln(a/x)}.$$

Ableiten liefert $\varphi'(x)=\varphi(x)\cdot\big(\ln(a/x)-1\big)$; die einzige Nullstelle von φ' ist $x=a/e$. Der Wert $\varphi(a/e)=e^{a/e}>1$ ist wegen $\varphi(x)\to 1$ für $x\to 0+$ und $\varphi(x)\to 0$ für $x\to\infty$ das globale Maximum der Funktion φ. Die Funktion φ wächst daher streng monoton für $0<x<a/e$ und fällt streng monoton für $x>a/e$. Ist also $n:=[a/e]$ (so daß $n\le a/e<n+1$ gilt), so ist das gesuchte Maximum der Einschränkung von φ auf $\mathbb{N}$ (und damit die Lösung des Optimierungsproblems) gegeben durch $\max\{(a/n)^n, \big(a/(n+1)\big)^{n+1}\}$.

Lösung (98.15) Es seien $(a,0)$ und $(0,b)$ die Endpunkte der Katheten des gesuchten Dreiecks; die Gerade durch diese beiden Punkte hat die Gleichung $ay+bx=ab$.



Zu minimieren ist also die Funktion

$$f(a,b) := a+b+\sqrt{a^2+b^2}$$

unter der Nebenbedingung $g(a,b)=0$ mit

$$g(a,b) := ab-ay_0-bx_0.$$

Die Nebenbedingung kann in der Form $(a-x_0)(b-y_0)=x_0y_0$ geschrieben werden, also $b-y_0=x_0y_0/(a-x_0)$ bzw.

$$(1) \qquad b = \frac{x_0y_0}{a-x_0}+y_0 = \frac{ay_0}{a-x_0}.$$

Nach Lagrange gibt es zu der optimalen Lösung (a,b) eine Zahl $\lambda\in\mathbb{R}$ mit $(\nabla f)(a,b)=\lambda(\nabla g)(a,b)$, also

$$\begin{bmatrix} 1+a/\sqrt{a^2+b^2} \\ 1+b/\sqrt{a^2+b^2} \end{bmatrix} = \lambda \begin{bmatrix} b-y_0 \\ a-x_0 \end{bmatrix}.$$

Um den (uns eigentlich gar nicht interessierenden) Parameter λ zu eliminieren, dividieren wir die x- durch die y-Komponente dieser Vektorgleichung und erhalten

$$\frac{b-y_0}{a-x_0} = \frac{a+\sqrt{a^2+b^2}}{b+\sqrt{a^2+b^2}} = \frac{1+\sqrt{1+(b/a)^2}}{(b/a)+\sqrt{1+(b/a)^2}}.$$

Setzen wir (1) in diese Gleichung ein, so erhalten wir

$$\frac{\frac{ay_0}{a-x_0} - y_0}{a-x_0} = \frac{1+\sqrt{1+\left(\frac{y_0}{a-x_0}\right)^2}}{\frac{y_0}{a-x_0} + \sqrt{1+\left(\frac{y_0}{a-x_0}\right)^2}}.$$

Bringen wir die linke Seite auf den Hauptnenner und erweitern wir die rechte Seite mit $a-x_0$, so geht diese Gleichung über in

$$\frac{x_0y_0}{(a-x_0)^2} = \frac{a-x_0+\sqrt{(a-x_0)^2+y_0^2}}{y_0+\sqrt{(a-x_0)^2+y_0^2}},$$

mit der Abkürzung $u := a-x_0$ also in

$$\begin{aligned}\frac{x_0y_0}{u^2} &= \frac{u+\sqrt{u^2+y_0^2}}{y_0+\sqrt{u^2+y_0^2}} \cdot \frac{y_0-\sqrt{u^2+y_0^2}}{y_0-\sqrt{u^2+y_0^2}} \\ &= \frac{(u+\sqrt{u^2+y_0^2})(y_0-\sqrt{u^2+y_0^2})}{-u^2}\end{aligned}$$

bzw.

$$\begin{aligned}x_0y_0 &= (\sqrt{u^2+y_0^2}+u)(\sqrt{u^2+y_0^2}-y_0) \\ &= u^2+y_0^2+(u-y_0)\sqrt{u^2+y_0^2}-uy_0\,.\end{aligned}$$

Isolieren der Wurzel führt auf

$$(y_0-u)\sqrt{u^2+y_0^2} = u^2+y_0^2-uy_0-x_0y_0,$$

anschließendes Quadrieren auf

$$(y_0-u)^2(u^2+y_0^2) = (u^2+y_0^2-uy_0-x_0y_0)^2\,.$$

Multiplizieren wir beide Seite aus, so heben sich alle Terme dritter und vierter Ordnung in u weg, und es bleibt die quadratische Gleichung

$$(y_0-2x_0)u^2 + 2x_0y_0u + x_0y_0(x_0-2y_0) = 0.$$

In dem Sonderfall $y_0 = 2x_0$ führt dies auf $u = 3x_0/2$, andernfalls auf die quadratische Gleichung

$$u^2 + \frac{2x_0y_0}{y_0-2x_0}u + x_0y_0\frac{x_0-2y_0}{y_0-2x_0} = 0$$

mit den Lösungen

$$u = \frac{-x_0y_0 \pm \sqrt{2x_0y_0}(x_0-y_0)}{y_0-2x_0}$$

und wegen $a = u+x_0$ dann

$$(2) \qquad a = \frac{-2x_0^2 \pm \sqrt{2x_0y_0}(x_0-y_0)}{y_0-2x_0}.$$

Das Vorzeichen der Wurzel wird durch die Bedingung bestimmt, daß $a > x_0$ gelten muß; man überzeugt sich schnell davon, daß das positive Vorzeichen für $y_0 \le x_0$ und das negative Vorzeichen für $y_0 \ge x_0$ gewählt werden muß. Einsetzen von (2) in (1) liefert dann

$$b = \frac{2y_0\big((x_0-y_0)^2-2x_0^2\big) \pm \sqrt{2x_0y_0}(y_0-2x_0)(x_0-y_0)}{(x_0-2x_0)(2x_0-y_0)}.$$

Lösung (98.16) Das Volumen V und die Oberfläche F eines Quaders mit den Kantenlängen x, y, z sind gegeben durch die Formeln

$$V(x,y,z) = xyz \quad \text{und} \quad F(x,y,z) = 2(xy+yz+zx).$$

(a) Zu minimieren ist F unter der Nebenbedingung $V = V_0$. Die Existenz eines Minimums ist dabei von vornherein klar, weil die stetige Funktion F auf der Menge aller $(x,y,z) \in (0,\infty)^3$ mit $V(x,y,z) = V_0$ (also $z = V_0/(xy)$) nach unten durch Null beschränkt ist und die Werte dieser Funktion für $x, y \to 0$ bzw. $x, y \to \infty$ gegen Unendlich gehen. Nach Lagrange gibt es an einer Stelle (x,y,z), an der das Minimum angenommen wird, eine Zahl λ mit $\nabla F = \lambda \cdot (\nabla V)$, also

$$(\star) \qquad 2\begin{bmatrix} y+z \\ z+x \\ x+y \end{bmatrix} = \lambda \begin{bmatrix} yz \\ zx \\ xy \end{bmatrix}.$$

Da aufgrund der Fragestellung nur positive Werte für x, y, z in Frage kommen, gilt $\lambda \neq 0$. Bilden wir den Quotienten der beiden ersten bzw. der beiden letzten Komponenten in $(\star)$, so erhalten wir die Gleichungen

$$\frac{x}{y} = \frac{x+z}{y+z} \quad \text{und} \quad \frac{y}{z} = \frac{y+x}{z+x},$$

aus denen sich die Gleichungen $xy + xz = xy + yz$ sowie $yz + xy = yz + xz$ bzw. $x = y$ und $y = z$ ergeben. Die minimale Oberfläche wird also für einen Würfel (mit der Kantenlänge $\sqrt[3]{V_0}$) angenommen.

(b) Zu maximieren ist V unter der Nebenbedingung $F = F_0$, wobei die Existenz eines Maximums wieder von vornherein klar ist. Nach Lagrange gibt es an einer Stelle (x,y,z), an der das Maximum angenommen wird, eine Zahl μ mit $\nabla V = \mu \cdot (\nabla F)$, wobei wie in (a) sofort klar ist, daß $\mu \neq 0$ gelten muß. Dann gilt aber $\nabla F = (1/\mu)\nabla V$, und das ist (mit $\lambda = 1/\mu$) die gleiche Bedingung wie in (a). Das maximale Volumen wird also für einen Würfel (mit der Kantenlänge $\sqrt{F_0/6}$) angenommen.

Lösung (98.17) Wir bezeichnen mit x, y, z die Länge, Breite und Höhe der Kiste. Deren Volumen sowie der Materialverbrauch sind dann gegeben durch

$$V(x,y,z) = xyz \quad \text{und} \quad F(x,y,z) = xy + 2xz + 2yz.$$

(a) Zu minimieren ist F unter der Nebenbedingung $V = V_0$. Ähnlich wie in der vorigen Aufgabe ist dabei die

Existenz eines Minimums von vornherein klar. Wird das Minimum für die Abmessungen x, y, z angenommen, so gibt es nach Lagrange eine Zahl $\lambda \in \mathbb{R}$ mit $(\nabla F)(x,y,z) = \lambda\,(\nabla V)(x,y,z)$, also

$$(\star) \qquad \begin{bmatrix} y+2z \\ x+2z \\ 2x+2y \end{bmatrix} = \lambda \begin{bmatrix} yz \\ xz \\ xy \end{bmatrix}.$$

Weil nur Zahlen $x, y, z > 0$ in Frage kommen, gilt $\lambda \neq 0$. Bilden wir den Quotienten der beiden ersten bzw. der beiden letzten Komponenten in $(\star)$, so erhalten wir die Gleichungen

$$\frac{y}{x} = \frac{y+2z}{x+2z} \quad \text{und} \quad \frac{z}{y} = \frac{x+2z}{2x+2y},$$

aus denen sich die Gleichungen $xy+2xz = xy+2yz$ sowie $xy+2yz = 2yz+2yz$ bzw. $x = y$ und $x = 2z$ ergeben. Die minimale Oberfläche wird also für einen Quader angenommen, für den $x = y = 2z$ gilt, der also eine quadratische Grundfläche besitzt und dessen Höhe halb so groß wie Grundseite ist.

(b) Zu maximieren ist V unter der Nebenbedingung $F = F_0$. Wieder ist die Existenz einer Lösung von vornherein klar. Der Ansatz nach Lagrange führt auf eine Gleichung $\nabla V = \mu(\nabla F)$. Man sieht wieder schnell, daß $\mu \neq 0$ gelten muß; wir erhalten daher $\nabla F = (1/\mu)\nabla V$, mit $\lambda := 1/\mu$ also die gleiche Bedingung wie in Teil (a). Also liefert das in (a) gefundene Ergebnis $x : y : z = 2 : 2 : 1$ auch die Lösung für die Fragestellung in (b).

Lösung (98.18) Wir identifizieren die Straße mit dem Intervall $[0, \ell]$ und bezeichnen mit $0 \le x_1 \le x_2 \le \cdots \le x_n \le \ell$ die Positionen der Notrufsäulen. Für jedes der Teilintervalle $[0, x_1]$, $[x_i, m_i]$ und $[m_i, x_{i+1}]$ (wobei $m_i := (x_i + x_{i+1})/2$ mit $1 \le i \le n-1$ sowie $[x_n, \ell]$ berechnen wir das Produkt aus der Wahrscheinlichkeit, daß sich eine gegebene Person (unter Annahme der Gleichverteilung entlang der Straße) in dem betrachteten Intervall befindet, und dem durchschnittlichen Weg zur nächsten Notrufsäule in diesem Intervall; dies liefert $(x_1/\ell)\cdot(x_1/2) = x_1^2/(2\ell)$ für das Intervall $[0, x_1]$, weiter $\big((x_{i+1}-x_i)/(2\ell)\big)\cdot(x_{i+1}-x_i)/4 = (x_{i+1}-x_i)^2/(8\ell)$ für jedes der Teilintervalle $[x_i, m_i]$ und $[m_i, x_{i+1}]$ und schließlich $\big((\ell - x_n)/\ell\big)\cdot(\ell - x_n)/2$ für das Intervall $[x_n, \ell]$. Die zu minimierende Funktion ist dann gerade die Summe all dieser Ausdrücke, also

$$F(x_1, \ldots, x_n) := \frac{1}{2\ell}\left(x_1^2 + \sum_{i=1}^{n-1} \frac{(x_{i+1}-x_i)^2}{2} + (\ell - x_n)^2\right).$$

Da der Faktor $1/(2\ell)$ für die Minimierung keine Rolle spielt, können wir genausogut die Funktion

$$f(x_1, \ldots, x_n) := x_1^2 + \sum_{i=1}^{n-1} \frac{(x_{i+1}-x_i)^2}{2} + (\ell - x_n)^2$$

minimieren. Nullsetzen der partiellen Ableitungen

$$\begin{aligned}
(\partial_1 f)(x_1, \ldots, x_n) &= 2x_1 - (x_2 - x_1), \\
(\partial_i f)(x_1, \ldots, x_n) &= (x_i - x_{i-1}) - (x_{i+1} - x_i), \\
(\partial_n f)(x_1, \ldots, x_n) &= (x_n - x_{n-1}) - 2(\ell - x_n),
\end{aligned}$$

also

$$\begin{aligned}
(\partial_1 f)(x_1, \ldots, x_n) &= 3x_1 - x_2, \\
(\partial_i f)(x_1, \ldots, x_n) &= -x_{i-1} + 2x_i - x_{i+1}, \\
(\partial_n f)(x_1, \ldots, x_n) &= -x_{n-1} + 3x_n - 2\ell
\end{aligned}$$

(wobei $2 \le i \le n-1$) führt auf die Gleichungen

$$\begin{aligned}
x_2 &= 3x_1, \\
x_3 &= -x_1 + 2x_2 = 5x_1, \\
x_4 &= -x_2 + 2x_3 = 7x_1, \\
&\vdots \\
x_n &= -x_{n-2} + 2x_{n-1} = (2n-1)x_1, \\
2\ell &= 3x_n - x_{n-1} = 4n \cdot x_1.
\end{aligned}$$

Setzen wir $a := \ell/n$, so haben wir also $x_1 = a/2$, $x_{i+1} - x_i = 2x_1 = a$ für $1 \le i \le n-1$ sowie $x_n = (2n-1)\cdot a/2 = \ell - a/2$.

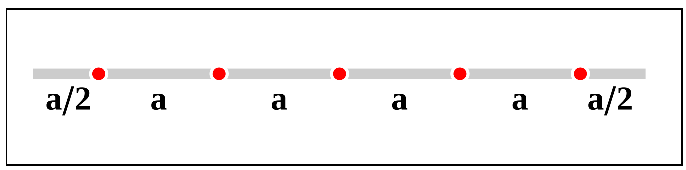
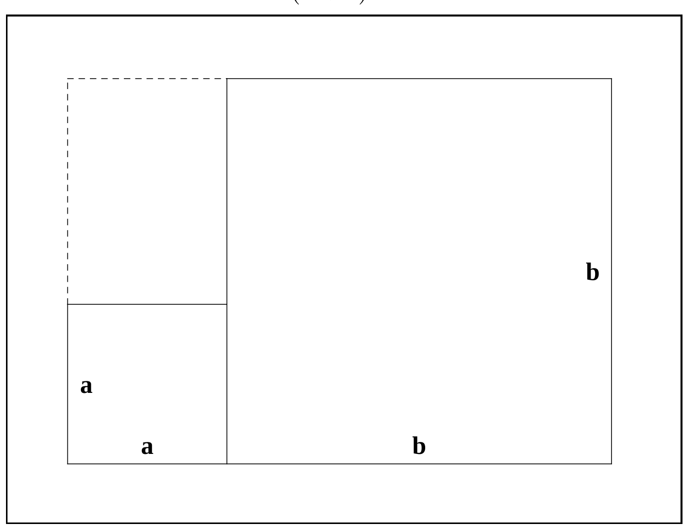


Abb. 98.18: Optimale Verteilung von Notrufsäulen entlang einer Straße.

Lösung (98.19) Sind $0 \le a \le b$ die Kantenlängen zweier Quadrate Q_1 und Q_2, so muß ein Rechteck, in das Q_1 und Q_2 nebeneinander passen sollen, offenbar mindestens den Flächeninhalt $(a+b)b$ haben.



Zu minimieren ist also die Funktion $f(a,b) := (a+b)b = ab + b^2$ unter der Nebenbedingung $g(a,b) = 0$ mit $g(a,b) :=$

a^2+b^2-1, wobei implizit noch $0 \le a \le b$ vorausgesetzt wird. Nach Lagrange machen wir den Ansatz $\nabla f = \lambda \nabla g$, also

$$\begin{bmatrix} b \\ a+2b \end{bmatrix} = 2\lambda \begin{bmatrix} a \\ b \end{bmatrix}.$$

Wegen $(a,b) \neq (0,0)$ gilt $\lambda \neq 0$, und sowohl a als auch b sind von Null verschieden (und damit echt positiv). Elimination von λ führt auf die Gleichung $b : a = (a+2b) : b$, also $b^2 = a^2+2ab$, folglich $2b^2 = a^2+2ab+b^2 = (a+b)^2$ und damit $b\sqrt{2} = a+b$ bzw. $a = b(\sqrt{2}-1)$. Einsetzen in die Nebenbedingung liefert $1 = a^2+b^2 = b^2(\sqrt{2}-1)^2+b^2 = (4-2\sqrt{2})b^2$ und damit

$$b^2 = \frac{1}{4-2\sqrt{2}} \cdot \frac{4+2\sqrt{2}}{4+2\sqrt{2}} = \frac{4+2\sqrt{2}}{8} = \frac{2+\sqrt{2}}{4}.$$

Damit ergibt sich $a^2 = 1-b^2$ zu

$$a^2 = \frac{2-\sqrt{2}}{4}.$$

Wir erhalten also die eindeutige Lösung

$$\begin{bmatrix} a \\ b \end{bmatrix} = \frac{1}{2}\begin{bmatrix} \sqrt{2-\sqrt{2}} \\ \sqrt{2+\sqrt{2}} \end{bmatrix} \approx \begin{bmatrix} 0.382683 \\ 0.923880 \end{bmatrix}$$

und damit den gesuchten Wert $F = f(a,b) = b(a+b) = (1+\sqrt{2})/2 \approx 1.20711$. (Die obige Abbildung zeigt die beiden Quadrate mit Kantenlängen a und b, für das ein Rechteck mit dieser Fläche tatsächlich benötigt wird.)

Lösung (98.20) Gesucht ist

$$\begin{aligned} C &:= \inf\left\{ \frac{\|x\|_p}{\|x\|_q} \;\middle|\; x \in \mathbb{K}^n \setminus \{0\} \right\} \\ &= \inf\left\{ \left\| \frac{x}{\|x\|_q} \right\|_p \;\middle|\; x \in \mathbb{K}^n \setminus \{0\} \right\} \\ &= \inf\{ \|y\|_p \mid \|y\|_q = 1 \} \\ &= \inf\{ \sqrt[p]{|y_1|^p + \cdots + |y_n|^p} \mid |y_1|^q + \cdots + |y_n|^q = 1 \} \\ &= \inf\{ \sqrt[p]{\xi_1^p + \cdots + \xi_n^p} \mid \xi_i \ge 0,\, \xi_1^q + \cdots + \xi_n^q = 1 \}. \end{aligned}$$

Mit den Funktionen $F(\xi_1,\ldots,\xi_n) := (\sum_{i=1}^n \xi_i^p)^{1/p}$ und $g(\xi_1,\ldots,\xi_n) := \xi_1^q + \xi_2^q + \cdots + \xi_n^q - 1 = 0$ läuft die Aufgabe also darauf hinaus, die Funktion F unter der Nebenbedingungen $g(\xi) = 0$ zu minimieren (wobei zusätzlich noch $\xi_i \ge 0$ gelten muß, was wir aber zunächst ignorieren). Statt F können wir natürlich genauso gut die Funktion $f := F^p$ minimieren, also $f(\xi_1,\ldots,\xi_n) = \sum_{i=1}^n \xi_i^p$. Wenn dann an der Stelle ξ das gesuchte Minimum angenommen wird, so gibt es nach dem Satz von Lagrange einen Lagrange-Multiplikator λ mit $(\nabla f)(\xi) = \lambda\,(\nabla g)(\xi)$, also

$$p \begin{bmatrix} \xi_1^{p-1} \\ \xi_2^{p-1} \\ \vdots \\ \xi_n^{p-1} \end{bmatrix} = \lambda \cdot q \begin{bmatrix} \xi_1^{q-1} \\ \xi_2^{q-1} \\ \vdots \\ \xi_n^{q-1} \end{bmatrix}.$$

Für $\xi_i \neq 0$ und $\xi_j \neq 0$ folgt hieraus $(\xi_i/\xi_j)^{p-1} = (\xi_i/\xi_j)^{q-1}$, wegen $p \neq q$ also $\xi_i/\xi_j = 1$ bzw. $\xi_i = \xi_j$. Für die gesuchte Minimalstelle kommen also nur Vektoren in Frage, bei denen eine Anzahl r der Komponenten den gleichen Wert $\xi \neq 0$ hat, während die anderen Komponenten von ξ verschwinden. Da die Funktionen f und g symmetrisch in ihren Argumenten sind, dürfen wir o.B.d.A. $\xi_i = \xi$ für $1 \le i \le r$ und $\xi_i = 0$ für $i > r$ annehmen. Die Nebenbedingung liefert dann $r\xi^q = 1$ und damit $\xi = \sqrt[q]{1/r} = r^{-1/q}$. Der Wert von f ist dann $r\xi^p = r \cdot r^{-p/q} = r^{1-p/q}$; für die eigentlich zu minimierende Funktion $F = f^{1/p}$ ergibt sich der Wert $r^{(1/p)-(1/q)}$. Wir unterscheiden zwei Fälle:

- für $p < q$ wird $r^{(1/p)-(1/q)}$ minimal, wenn r möglichst klein ist, wenn also $r = 1$ gilt;
- für $p > q$ wird $r^{(1/p)-(1/q)}$ minimal, wenn r möglichst groß ist, wenn also $r = n$ gilt.

Die kleinstmögliche Zahl C mit $\|x\|_p \le C\|x\|_q$ für alle x ist also $C = 1$ für $p < q$ und $C = n^{(1/p)-(1/q)}$ für $p > q$.

Lösung (98.21) Der Ansatz $\nabla f = \lambda \nabla g$ führt auf die Gleichung

$$\begin{bmatrix} 2 \\ 3 \end{bmatrix} = \lambda \begin{bmatrix} 1/(2\sqrt{x}) \\ 1/(2\sqrt{y}) \end{bmatrix}$$

und damit $2/3 = \sqrt{y}/\sqrt{x}$ bzw. $\sqrt{y} = 2\sqrt{x}/3$. Dann gilt $5 = \sqrt{x} + \sqrt{y} = 5\sqrt{x}/3$, folglich $\sqrt{x} = 3$ und damit $x = 9$. Es folgt $\sqrt{y} = 2$ und daher $y = 4$. Im Punkt $(9,4)$ nimmt nun die Funktion f den Wert $f(9,4) = 30$ an. Der Punkt $(0,25)$ erfüllt aber auch die Gleichung $g(x,y) = 0$, und es gilt $f(0,25) = 75$ (und dies ist tatsächlich der maximal mögliche Wert). Das Maximum von f unter der Nebenbedingung $g = 0$ kann also nicht mit der Methode von Lagrange gefunden werden. Das ist aber kein Widerspruch, denn die Funktion g ist im Punkt $(0,25)$ gar nicht differentiierbar, der Satz von Lagrange also nicht anwendbar.

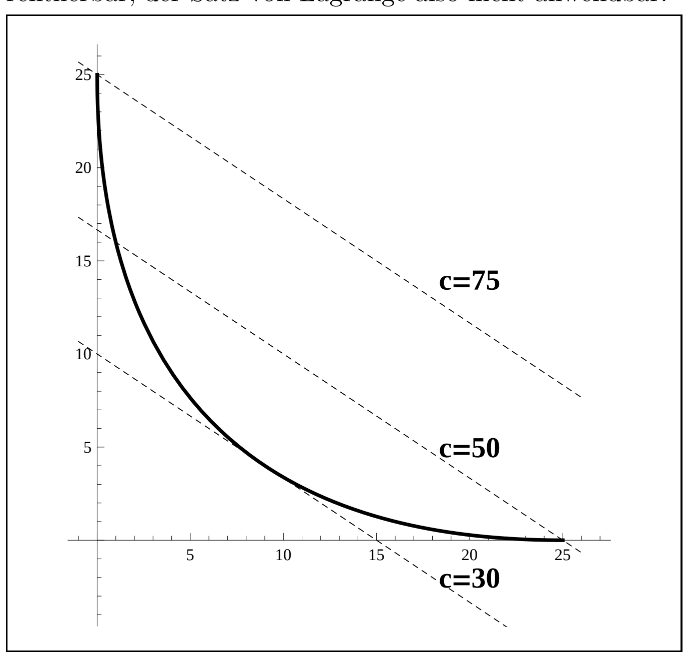


Abb. 98.21: Kurve $g(x,y) = 0$ und Geraden $f(x,y) = 2x+3y = c$ für verschiedene Werte c.

Lösung (98.22) Wir machen nach Lagrange den Ansatz $\nabla f = \lambda \nabla g$; zusammen mit der Nebenbedingung $g = 0$ erhalten wir also die Gleichungen

$$(\star) \qquad \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \lambda \begin{bmatrix} 3x^2 \\ -2y \end{bmatrix} \quad \text{und} \quad x^3 = y^2.$$

Aus der linken Gleichung erhalten wir zunächst $\lambda \neq 0$ und dann $y = 0$. Die rechte Gleichung liefert anschließend $x = 0$. Dann ist aber die linke Gleichung in $(\star)$ nicht mehr erfüllbar. Das Gleichungssystem $(\star)$ hat also keine Lösung. Andererseits gilt $x^3 = y^2 \geq 0$ für alle (x, y), die die Nebenbedingung erfüllen, so daß $f(x, y) = x$ unter dieser Nebenbedingung das (sogar globale) Minimum 0 annimmt $\big($und zwar an der Stelle $(0, 0)\big)$. Daß dieses Minimum nicht mit der Methode von Lagrange gefunden werden kann, widerspricht nicht dem dieser Methode zugrundeliegenden Satz, denn bei diesem wird die Regularität der Nebenbedingungen an dem fraglichen Punkt vorausgesetzt. In unserem Fall ist aber $(0, 0)$ kein regulärer Punkt der Funktion g.

Lösung (98.23) Sind die Gradienten $(\nabla g_i)(p)$ linear unabhängig, so gilt die Aussage nach dem Satz von Lagrange mit $\lambda_0 := 1$. Sind die Gradienten $(\nabla g_i)(p)$ dagegen linear abhängig, so läßt sich der Nullvektor als nichttriviale Linearkombination dieser Vektoren darstellen; es gibt also einen Vektor $(\lambda_1, \ldots, \lambda_m) \neq (0, \ldots, 0)$ mit $\sum_{i=1}^m \lambda_i (\nabla g_i)(p) = 0$. Die Behauptung gilt dann mit $\lambda_0 := 0$.

Lösung (98.24) Es gilt $(\nabla f)(p) = \sum_{i=1}^m \lambda_i (\nabla g_i)(p) = g'(p)\lambda$. Beiderseitige Multiplikation mit $g'(p)^T$ von links liefert $g'(p)^T (\nabla f)(p) = g'(p)^T g'(p)\lambda$. Durchmultiplizieren mit $\big(g'(p)^T g'(p)\big)^{-1}$ von links liefert nun die Behauptung. Die Invertierbarkeit von $g'(p)^T g'(p)$ ergibt sich daraus, daß nach Voraussetzung $g'(p)$ und damit auch $g'(p)^T g'(p)$ den Rang m hat.

Lösung (98.25) (a) Nach Lagrange machen wir den Ansatz $\nabla L = 0$, also

$$\begin{bmatrix} y \\ x \end{bmatrix} + \lambda \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

Diese Bedingung führt zusammen mit der Nebenbedingung $x + y = 2$ auf die einzige Lösung $x = y = 1 = -\lambda$, also $(x, y, \lambda) = (x_0, y_0, \lambda_0)$. Daß es sich um ein Maximum handelt, wird sofort klar, wenn man die Nebenbedingung eliminiert , also $y = 2 - x$ und dann $f(x, y) = xy = x(2 - x) = 2x - x^2$ schreibt.

(b) Fassen wir L als Funktion von drei Variablen x, y, λ auf, so erhalten wir an einer beliebigen Stelle (x, y, λ) (insbesondere also an der kritischen Stelle (x_0, y_0, λ_0)) die Hesse-Matrix

$$L''(x, y, z) = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix}.$$

Diese hat die Eigenwerte $\lambda_1 = 2$ und $\lambda_{2,3} = -1$, ist also indefinit. Damit kann L keine lokalen Extrema besitzen. Für $\Phi(x, y) := L(x, y, \lambda_0) = L(x, y, -1) = f(x, y) - g(x, y) = xy - x - y + 2$ erhalten wir den Gradienten

$$(\nabla \Phi)(x, y) = \begin{bmatrix} y - 1 \\ x - 1 \end{bmatrix}$$

und die Hesse-Matrix

$$\Phi''(x, y) = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}.$$

Da an jeder Stelle $(x, y) \in \mathbb{R}^2$ die Matrix $\Phi''(x, y)$ die Eigenwerte ± 1 hat und daher indefinit ist, besitzt Φ keine lokalen Extrema.

Lösung (98.26) Die Lagrangefunktion ist

$$L(x, y; \lambda) := x^2 + y^2 + \lambda \left(y - x^3 - \frac{28}{27} \right).$$

Die Bedingung $\nabla L = 0$ führt auf

$$\begin{bmatrix} 2x - 3\lambda x^2 \\ 2y + \lambda \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

Die erste Gleichung liefert $x(2 - 3\lambda x) = 0$. Ist $x = 0$, so folgt zunächst $y = 28/27$ und dann $\lambda = -56/27$. Ist dagegen $x \neq 0$, so folgt $\lambda = 2/(3x)$ aus der ersten Gleichung und $y = -\lambda/2 = -1/(3x)$ aus der zweiten Gleichung; Einsetzen in die Nebenbedingung $y = x^3 + (28/27)$ liefert dann

$$\frac{-1}{3x} = x^3 + \frac{28}{27} \quad \text{bzw.} \quad 27x^4 + 28x + 9 = 0.$$

Diese letzte Gleichung hat zwei reelle Lösungen, nämlich $x = -1/3$ und

$$x_\star = \frac{1}{9} \left(1 - \frac{2}{\sqrt[3]{18\sqrt{418} - 368}} + \sqrt[3]{18\sqrt{418} - 368} \right).$$

Wir haben $x_\star \approx -0.867472$ und $y^\star = x_\star^3 + (28/27) \approx 0.384258$. Für $x = -1/3$ erhalten wir $y = 1$ und $\lambda = -2$. Im Punkt $p = (-1/3, 1)$ ist nun

$$L''(p; \lambda) = L''(-1/3, 1; -2) = \begin{bmatrix} -2 & 0 \\ 0 & 2 \end{bmatrix}$$

indefinit. Die Matrix B aus Aufgabe (98.27) ist wegen $T_p M = \mathbb{R}\,(3, 1)^T$ gegeben durch

$$\begin{bmatrix} 3 & 1 \end{bmatrix} \begin{bmatrix} -2 & 0 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} 3 \\ 1 \end{bmatrix} = [-16]$$

und damit negativ definit; an der Stelle p liegt also ein lokales Maximum von f vor.

Bemerkung. In geometrischer Formulierung fragen wir in dieser Aufgabe, an welchen Punkten der Kurve $y = x^3 + (28/27)$ der Abstand zum Nullpunkt ein lokales Minimum oder Maximum annimmt. Es gibt drei kritische Punkte: ein (sogar globales) Minimum im Punkt $(x_\star, y_\star)$, ein lokales Maximum im Punkt $(-1/3, 1)$ und ein lokales Minimum im Punkt $(0, 28/27)$. Das wird auch aus den folgenden Skizzen klar.

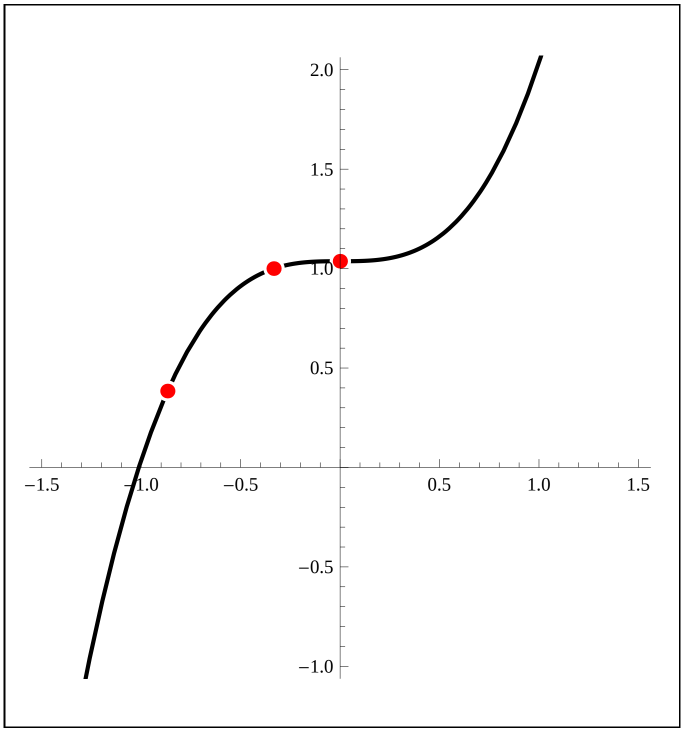


Abb. 98.26a: Kritische Punkte der Funktion $f(x,y) = x^2 + y^2$ auf der Kurve $y = x^3 + (28/27)$.

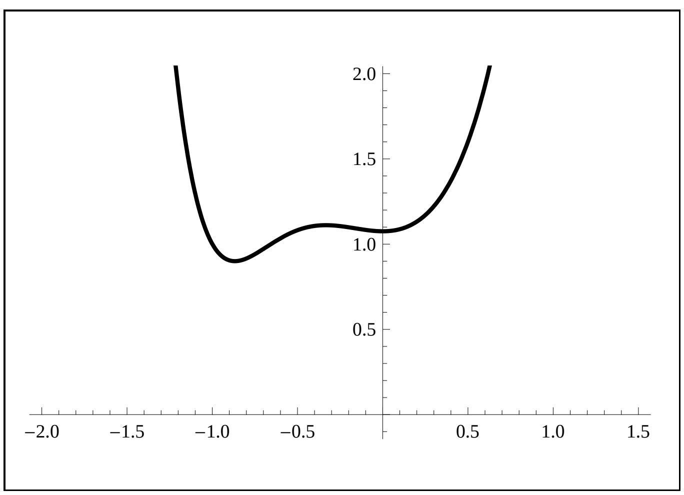


Abb. 98.26b: Verlauf der Funktion $x \mapsto x^2 + (x^3 + 28/27)^2$.

Lösung (98.27) Es sei φ eine Parametrisierung von M in einer Umgebung von $\varphi(u^\star) = p$, ausgeschrieben

$$\varphi(u_1,\ldots,u_d) \;=\; \begin{bmatrix} x_1(u_1,\ldots,u_d) \\ \vdots \\ x_n(u_1,\ldots,u_d) \end{bmatrix}$$

mit $d = \dim(M)$. Dann nimmt f genau dann ein lokales Minimum in p an, wenn $F \circ \varphi$ ein lokales Minimum in $u^\star$ annimmt. Sind nun Lagrange-Multiplikatoren $\lambda_1, \ldots, \lambda_m$ so gewählt, daß für $L := F + \lambda_1 G_1 + \cdots + \lambda_m G_m$ die Bedingung $L'(p) = 0$ erfüllt ist, so stimmen F und L auf M überein; also nimmt $F \circ \varphi$ genau dann ein lokales Minimum in $u^\star$ an, wenn dies auch für $L \circ \varphi$ zutrifft. Dafür ist nun notwendig bzw. hinreichend, daß $(L \circ \varphi)''(u^\star)$ positiv semidefinit bzw. positiv definit ist. Für alle $\xi, \eta \in \mathbb{R}^d$ wollen wir nun

$$(\star)\quad \langle \xi, (L\circ\varphi)''(u^\star)\,\eta\rangle \;=\; \frac{\mathrm{d}}{\mathrm{d}t}\Big|_{t=0} \frac{\mathrm{d}}{\mathrm{d}s}\Big|_{s=0} L\big(\varphi(u^\star + s\xi + t\eta)\big)$$

berechnen. Zunächst ist $(\mathrm{d}/\mathrm{d}s) L\big(\varphi(u^\star + s\xi + t\eta)\big)$ nach der Kettenregel gegeben durch

$$\sum_{i=1}^n \sum_{k=1}^d \frac{\partial L}{\partial x_i}\big(\varphi(u^\star + s\xi + t\eta)\big) \frac{\partial \varphi_i}{\partial u_k}(u^\star + s\xi + t\eta)\,\xi_k;$$

setzen wir dies mit $s = 0$ in $(\star)$ ein, so ergibt sich

$$\begin{aligned} \langle \xi, (L\circ\varphi)''(u^\star)\,\eta\rangle \;&=\; \frac{\mathrm{d}}{\mathrm{d}t}\Big|_{t=0} \frac{\mathrm{d}}{\mathrm{d}s}\Big|_{s=0} L\big(\varphi(u^\star + s\xi + t\eta)\big) \\ &= \frac{\mathrm{d}}{\mathrm{d}t}\Big|_{t=0} \sum_{i=1}^n \sum_{k=1}^d \frac{\partial L}{\partial x_i}\big(\varphi(u^\star + t\eta)\big) \frac{\partial \varphi_i}{\partial u_k}(u^\star + t\eta)\,\xi_k\,. \end{aligned}$$

Unter Benutzung der Ketten- und der Produktregel erhalten wir weiter

$$\begin{aligned} &\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial x_i}\big(\varphi(u^\star + t\eta)\big) \frac{\partial \varphi_i}{\partial u_k}(u^\star + t\eta)\xi_k\right) \;= \\ &\sum_{j=1}^n \sum_{\ell=1}^d \frac{\partial^2 L}{\partial x_i \partial x_j}\big(\varphi(u^\star + t\eta)\big) \frac{\partial \varphi_j}{\partial u_\ell}(u^\star + t\eta) \frac{\partial \varphi_i}{\partial u_k}(u^\star + t\eta)\xi_k \eta_\ell \\ &\qquad + \sum_{\ell=1}^d \frac{\partial L}{\partial x_i}\big(\varphi(u^\star + t\eta)\big) \frac{\partial^2 \varphi}{\partial u_k \partial u_\ell}(u^\star + t\eta)\xi_k \eta_\ell; \end{aligned}$$

setzen wir dies mit $t = 0$ in die obige Gleichung ein, so ergibt sich

$$\begin{aligned} \langle \xi, (L\circ\varphi)''(u^\star)\,\eta\rangle \;&=\; \frac{\mathrm{d}}{\mathrm{d}t}\Big|_{t=0} \frac{\mathrm{d}}{\mathrm{d}s}\Big|_{s=0} L\big(\varphi(u^\star + s\xi + t\eta)\big) \\ &= \sum_{i,j=1}^n \sum_{k,\ell=1}^d \frac{\partial^2 L}{\partial x_i \partial x_j}\big(\varphi(u^\star)\big) \frac{\partial \varphi_j}{\partial u_\ell}(u^\star) \frac{\partial \varphi_i}{\partial u_k}(u^\star)\xi_k \eta_\ell \\ &\qquad + \sum_{i=1}^n \sum_{k,\ell=1}^d \frac{\partial L}{\partial x_i}\big(\varphi(u^\star)\big) \frac{\partial^2 \varphi}{\partial u_k \partial u_\ell}(u^\star)\xi_k \eta_\ell \\ &= \langle \varphi'(u^\star)\xi, L''(p)\varphi'(u^\star)\eta\rangle \;+\; L'(p) \begin{bmatrix} \langle \xi, \varphi_1''(u^\star)\eta\rangle \\ \vdots \\ \langle \xi, \varphi_n''(u^\star)\eta\rangle \end{bmatrix} \\ &= \langle \varphi'(u^\star)\xi, L''(p)\varphi'(u^\star)\eta\rangle\,. \end{aligned}$$

(Der letzte Schritt gilt dabei wegen $L'(p) = 0$, und dies ist genau der Grund, warum wir die Rechnung für L statt für

F durchführten! Für F gilt *nicht* $F'(p) = 0$, denn $F'(p)$ annulliert zwar T_pM, aber nicht notwendigerweise ganz $\mathbb{R}^n$.) Wir erhalten daher die folgende Kette äquivalenter Bedingungen:

- $(L \circ \varphi)''(u^\star)$ ist positiv definit (bzw. semidefinit);
- $\langle \varphi'(u^\star)\xi, L''(p)\varphi'(u^\star)\xi\rangle > 0$ (bzw. ≥ 0) für alle $\xi \in \mathbb{R}^d \setminus \{0\}$;
- $\langle v, L''(p)v\rangle > 0$ (bzw. ≥ 0) für alle $v \in (T_pM) \setminus \{0\}$;
- $\langle \sum_i \lambda_i v_i, L''(p) \sum_i \lambda_i v_i\rangle > 0$ (bzw. ≥ 0) für alle $\lambda \in \mathbb{R}^d \setminus \{0\}$;
- $\lambda^T A^T L''(p) A\lambda > 0$ (bzw. ≥ 0) für alle $\lambda \in \mathbb{R}^d \setminus \{0\}$;
- $\langle \lambda, B\lambda\rangle > 0$ (bzw. ≥ 0) für alle $\lambda \in \mathbb{R}^d \setminus \{0\}$;
- B ist positiv definit (bzw. semidefinit).

Damit ist die Behauptung für lokale Minima bewiesen. Die Behauptung für lokale Maxima folgt in völlig analoger Weise (oder durch Anwendung der Behauptung für Minima auf $-f$ statt auf f). Ist schließlich B indefinit, so ist B weder positiv noch negativ semidefinit, so daß wegen (a) und (c) bei p weder ein lokales Minimum noch ein lokales Maximum vorliegen kann.

Lösung (98.28) Hier wird $(T_pM)^\perp$ aufgespannt von dem Vektor $\nabla g = (g_x, g_y)^T$ an der Stelle p; folglich wird T_pM wird aufgespannt von $v_1 := (-g_y, g_x)^T$. Die Matrix B in Aufgabe (98.27) reduziert sich dann auf die (1×1)-Matrix (D), und die Aussagen (b) und (d) aus Aufgabe (98.27) liefern die Behauptung.

Lösung (98.29) Hier wird $(T_pM)^\perp$ aufgespannt von dem Vektor $\nabla g = (g_x, g_y, g_z)^T$ an der Stelle p. Ist $(g_x, g_y) \neq (0,0)$, so wird eine Basis von T_pM gebildet von den Vektoren $v_1 := (-g_y,\, g_x,\, 0)^T$ und

$$v_2 \;:=\; (\nabla g)\times v_1 \;=\; \begin{bmatrix} g_x \\ g_y \\ g_z \end{bmatrix} \times \begin{bmatrix} -g_y \\ g_x \\ 0 \end{bmatrix} \;=\; \begin{bmatrix} -g_x g_z \\ -g_y g_z \\ g_x^2 + g_y^2 \end{bmatrix}$$

an der Stelle p. Ist dagegen $(g_x, g_y) = (0,0)$ und damit $\nabla g = (0,0,g_z)^T$ mit $g_z \neq 0$, so wird eine Basis von T_pM gebildet von den Vektoren $v_1 := (1,\,0,\,0)^T$ und

$$v_2 \;:=\; (\nabla g)\times v_1 \;=\; \begin{bmatrix} 0 \\ 0 \\ g_z \end{bmatrix} \times \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \;=\; \begin{bmatrix} 0 \\ g_z \\ 0 \end{bmatrix}$$

(wieder an der Stelle p genommen). Die angegebene Matrix M ist also genau die Matrix B in Aufgabe (98.27), und die Behauptung folgt aus den Aussagen (b) und (d) in Aufgabe (98.27).

Lösung (98.30) Hier wird $(T_pM)^\perp$ aufgespannt von den Vektoren $(\nabla g_1)(p)$ und $(\nabla g_2)(p)$, und T_pM wird dann aufgespannt von $(\nabla g_1)(p) \times (\nabla g_2)(p)$. Die Matrix B in Aufgabe (98.27) reduziert sich daher auf die (1×1)-Matrix (D), und die Aussagen (b) und (d) aus Aufgabe (98.27) liefern die Behauptung.

Lösung (98.31) Die Lagrangefunktion lautet

$$L(x,y,z;\lambda) \;:=\; x(y+z) + \lambda(x^2+y^2+z^2-1)\,.$$

Die Bedingung $\nabla L = 0$ führt auf

$$(\star) \qquad \begin{bmatrix} y+z+2\lambda x \\ x+2\lambda y \\ x+2\lambda z \end{bmatrix} \;=\; \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

Subtraktion der beiden letzten Komponenten dieser Gleichung liefert $0 = 2\lambda(z-y)$ und damit $\lambda = 0$ oder $z = y$.

Erster Fall: $\lambda = 0$. In diesem Fall gilt $y+z=0$ und damit $z = -y$. Die Nebenbedingung $x^2+y^2+z^2 = 1$ geht dann über in $2y^2 = 1$ und damit $y^2 = 1/2$. Dies liefert die beiden Lösungen

$$(x,y,z;\lambda) = \left(0, \frac{1}{\sqrt{2}}, \frac{-1}{\sqrt{2}}; 0\right),$$
$$(x,y,z;\lambda) = \left(0, \frac{-1}{\sqrt{2}}, \frac{1}{\sqrt{2}}; 0\right).$$

Zweiter Fall: $z = y$. Die erste Komponente in $(\star)$ liefert dann $y = -\lambda x$; Einsetzen in die zweite Komponente ergibt $x = -2\lambda y = 2\lambda^2 x$. Wäre $x = 0$, dann auch $y = 0$ und folglich $z = 0$ im Widerspruch zur Nebenbedingung $x^2+y^2+z^2 = 1$. Also ist $x \neq 0$, und Division der Gleichung $x = 2\lambda^2 x$ durch x liefert $1 = 2\lambda^2$ bzw. $\lambda = \pm 1/\sqrt{2}$. Die Nebenbedingung $x^2+y^2+z^2 = 1$ geht über in $1 = x^2 + \lambda^2 x^2 + \lambda^2 x^2 = 2x^2$ und damit $x = \pm 1/\sqrt{2}$; zusammen mit den Bedingungen $z = y = -\lambda x$ erhalten wir damit die Lösungen

$$(x,y,z;\lambda) = \left(\frac{1}{\sqrt{2}}, \frac{-1}{2}, \frac{-1}{2}; \frac{1}{\sqrt{2}}\right),$$
$$(x,y,z;\lambda) = \left(\frac{-1}{\sqrt{2}}, \frac{1}{2}, \frac{1}{2}; \frac{1}{\sqrt{2}}\right),$$
$$(x,y,z;\lambda) = \left(\frac{1}{\sqrt{2}}, \frac{1}{2}, \frac{1}{2}; \frac{-1}{\sqrt{2}}\right),$$
$$(x,y,z;\lambda) = \left(\frac{-1}{\sqrt{2}}, \frac{-1}{2}, \frac{-1}{2}; \frac{-1}{\sqrt{2}}\right).$$

Wir erhalten also die sechs kritischen Punkte

$$p_{1,2} \;=\; \pm\left(0, \frac{1}{\sqrt{2}}, \frac{-1}{\sqrt{2}}\right),$$
$$p_{3,4} \;=\; \pm\left(\frac{1}{\sqrt{2}}, \frac{-1}{2}, \frac{-1}{2}\right),$$
$$p_{5,6} \;=\; \pm\left(\frac{1}{\sqrt{2}}, \frac{1}{2}, \frac{1}{2}\right)$$

mit den zugehörigen Lagrange-Multiplikatoren

$$\lambda_{1,2} = 0, \quad \lambda_{3,4} = \frac{1}{\sqrt{2}}, \quad \lambda_{5,6} = \frac{-1}{\sqrt{2}}.$$

Die Hesse-Matrix der Lagrange-Funktion lautet

$$L''(x,y,z;\lambda) \;=\; \begin{bmatrix} 2\lambda & 1 & 1 \\ 1 & 2\lambda & 0 \\ 1 & 0 & 2\lambda \end{bmatrix}.$$

Da die betrachtete Mannigfaltigkeit $M = \{(x,y,z) \in \mathbb{R}^3 \mid x^2+y^2+z^2 = 1\}$ eine Sphäre mit Mittelpunkt im Nullpunkt ist, ist der Tangentialraum eines Punktes $p \in M$ gerade $T_pM = p^\perp$. Wir wählen nun für $1 \le i \le 6$ jeweils eine Matrix $A_i \in \mathbb{R}^{3\times 2}$, deren Spalten eine Basis von $T_{p_i}M = p_i^\perp$ bilden, zum Beispiel

$$\begin{aligned} A_{1,2} &:= \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 1 \end{bmatrix}, \\ A_{3,4} &:= \begin{bmatrix} 1 & 1 \\ \sqrt{2} & 0 \\ 0 & \sqrt{2} \end{bmatrix}, \\ A_{5,6} &:= \begin{bmatrix} 1 & 1 \\ -\sqrt{2} & 0 \\ 0 & -\sqrt{2} \end{bmatrix}. \end{aligned}$$

Dann bilden wir die Matrizen $B_i := A_i^T L''(p_i) A_i$ für $1 \le i \le 6$. Für $i = 1,2$ ist

$$B_i = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 2 \\ 2 & 0 \end{bmatrix}$$

indefinit; also liegen an den beiden Stellen p_1 und p_2 Sattelpunkte vor. Für $i = 3,4$ ist

$$\begin{aligned} B_i &= \begin{bmatrix} 1 & \sqrt{2} & 0 \\ 1 & 0 & \sqrt{2} \end{bmatrix} \begin{bmatrix} \sqrt{2} & 1 & 1 \\ 1 & \sqrt{2} & 0 \\ 1 & 0 & \sqrt{2} \end{bmatrix} \begin{bmatrix} 1 & 1 \\ \sqrt{2} & 0 \\ 0 & \sqrt{2} \end{bmatrix} \\ &= \begin{bmatrix} 5\sqrt{2} & 3\sqrt{2} \\ 3\sqrt{2} & 5\sqrt{2} \end{bmatrix} \end{aligned}$$

positiv definit; also liegen an den beiden Stellen p_3 und p_4 Minima vor. Für $i = 5,6$ ist

$$\begin{aligned} B_i &= \begin{bmatrix} 1 & -\sqrt{2} & 0 \\ 1 & 0 & -\sqrt{2} \end{bmatrix} \begin{bmatrix} -\sqrt{2} & 1 & 1 \\ 1 & -\sqrt{2} & 0 \\ 1 & 0 & -\sqrt{2} \end{bmatrix} \begin{bmatrix} 1 & 1 \\ -\sqrt{2} & 0 \\ 0 & -\sqrt{2} \end{bmatrix} \\ &= \begin{bmatrix} -5\sqrt{2} & -3\sqrt{2} \\ -3\sqrt{2} & -5\sqrt{2} \end{bmatrix} \end{aligned}$$

negativ definit; also liegen an den beiden Stellen p_5 und p_6 Maxima vor.

Wir geben eine Variante der Lösung an, die vielleicht intuitiv klar ist, die aber erst dann präzise gemacht werden kann, wenn uns der Begriff der Differentialform zur Verfügung steht. Mit $G(x,y,z) := x^2+y^2+z^2-1$ wird die Lagrangefunktion

$$L(x,y,z;\lambda) \;:=\; F(x,y,z) + \lambda \cdot G(x,y,z)$$

gebildet. Notwendige Bedingung für ein Extremum ist dann die Bedingung $(\partial_x L, \partial_y L, \partial_z L, \partial_\lambda L) = (0,0,0,0)$; dies führt auf sechs potentielle Lösungen $(x_i, y_i, z_i; \lambda_i)$ mit $1 \le i \le 6$. Ich betrachte beispielhaft eine dieser Lösungen, nämlich

$$(x,y,z) = \left(\frac{1}{\sqrt{2}}, -\frac{1}{2}, -\frac{1}{2}\right) =: p \quad \text{mit} \quad \lambda = \frac{1}{\sqrt{2}}.$$

Mit dem festgehaltenen Wert für λ ergibt sich die Lagrangefunktion

$$L(x,y,z) \;=\; x(y+z) + \frac{1}{\sqrt{2}} \cdot (x^2+y^2+z^2-1).$$

Wir erhalten

$$\begin{aligned} L_x &= y + z + \sqrt{2}x, \\ L_y &= x + \sqrt{2}y, \\ L_z &= x + \sqrt{2}z \end{aligned}$$

und dann

$$\begin{bmatrix} L_{xx} & L_{xy} & L_{xz} \\ L_{yx} & L_{yy} & L_{yz} \\ L_{zx} & L_{zy} & L_{zz} \end{bmatrix} = \begin{bmatrix} \sqrt{2} & 1 & 1 \\ 1 & \sqrt{2} & 0 \\ 1 & 0 & \sqrt{2} \end{bmatrix},$$

insbesondere also

$$L''(p) \;=\; \begin{bmatrix} \sqrt{2} & 1 & 1 \\ 1 & \sqrt{2} & 0 \\ 1 & 0 & \sqrt{2} \end{bmatrix}.$$

Eine Variation $\mathrm{d}p = (\mathrm{d}x,\, \mathrm{d}y,\, \mathrm{d}z)$ liefert daher die quadratische Abweichung

$(\star)$

$$\begin{aligned} \langle \mathrm{d}p, L''(p)\mathrm{d}p \rangle &= \sqrt{2}\,(\mathrm{d}x^2 + \mathrm{d}y^2 + \mathrm{d}z^2) + 2\,\mathrm{d}x\,\mathrm{d}y + 2\,\mathrm{d}x\,\mathrm{d}z \\ &= \frac{1}{\sqrt{2}}(\mathrm{d}x + \sqrt{2}\mathrm{d}y)^2 + \frac{1}{\sqrt{2}}(\mathrm{d}x + \sqrt{2}\mathrm{d}z)^2. \end{aligned}$$

Nun dürfen von p aus nur Variationen $\mathrm{d}p$ betrachtet werden, die tangential zur Mannigfaltigkeit $M := \{(x,y,z) \in \mathbb{R}^3 \mid G(x,y,z) = 0\}$ verlaufen, für die also $0 = \mathrm{d}G = 2(\mathrm{d}x + \mathrm{d}y + \mathrm{d}z)$ gilt. Wir können also $\mathrm{d}x = -\mathrm{d}y - \mathrm{d}z$ schreiben. Setzen wir dies in $(\star)$ ein, so ergibt sich

$$\begin{aligned} \langle \mathrm{d}p, L''(p)\mathrm{d}p \rangle &= \frac{1}{\sqrt{2}} \Big((\sqrt{2}-1)\mathrm{d}y - \mathrm{d}z \Big)^2 \\ &\quad + \frac{1}{\sqrt{2}} \Big(-\mathrm{d}y + (\sqrt{2}-1)\,\mathrm{d}z \Big)^2. \end{aligned}$$

Dieser Ausdruck ist > 0 für alle $(\mathrm{d}y, \mathrm{d}z) \neq (0,0)$; also liegt in p ein lokales Minimum vor.

Lösung (98.32) Die definierende Gleichung $\nabla f(x) = \sum_{i=1}^m \lambda_i (\nabla g_i)(x)$ der Lagrange-Multiplikatoren läßt sich schreiben als $f'(x)^T = g'(x)^T \lambda$ bzw. $f'(x) = \lambda^T g'(x)$, in unserem Fall also

$$(1) \qquad f'\big(x_\star(c)\big) \;=\; \lambda(c)^T g'\big(x_\star(c)\big).$$

Andererseits gilt die Identität $g\big(x_\star(c)\big) = c$, nach Ableiten bezüglich c also

$$(2) \qquad g'\big(x_\star(c)\big) x_\star'(c) \;=\; \mathbf{1}.$$

Unter Benutzung von (1) und (2) erhalten wir dann

$$\begin{aligned} \frac{\partial}{\partial c} f\big(x_\star(c)\big) &= f'\big(x_\star(c)\big) x_\star'(c) \;=\; \lambda(c)^T g'\big(x_\star(c)\big) x_\star'(c) \\ &= \lambda(c)^T\, \mathbf{1} \;=\; \lambda(c)^T. \end{aligned}$$

Die i-te Komponente dieser Gleichung lautet

$$\frac{\partial}{\partial c_i} f\big(x_\star(c)\big) \;=\; \lambda_i(c).$$

Dies kann man so deuten, daß der i-te Lagrange-Multiplikator $\lambda_i(c)$ ein Maß für die Sensitivität der Änderung des Optimums $f\big(x_\star(c)\big)$ bezüglich einer Änderung des Wertes c_i in der Nebenbedingung $g(x) = c_i$ ist.