Thema14—Der Satz über inverse Funktionen und der Satz über implizite Funktionen

In diesem Kapitel betrachten wir die Invertierbarkeit von glatten Abbildungen bzw. die Auflösbarkeit von impliziten Gleichungen. Zunächst einige einfache Beispiele.

1. (Polarkoordinaten.) Die Abbildung $(r, \theta) \mapsto (r\cos\theta, r\sin\theta)$ der Übergangstransformation von Polarkoordinaten in Kartesische Koordinaten. Wir invertieren diese Transformation, indem wir die Gleichungen $x = r\cos\theta$, $y = r\sin\theta$ lösen. Rein formal bekommen wir $r = \sqrt{x^2+y^2}$, $\tan\theta = \frac{y}{x}$.

   Bekanntlich gibt es bei der 2-ten Gleichnung Komplikationen, da der Winkel θ nicht eindeutig durch $\tan\theta$ bestimmt wird.

2. Der Einheitskreis wird implizit durch die Gleichung $x^2 + y^2 = 1$ beschrieben. Wir können diese Gleichung für y auflösen, und bekommen $y = \sqrt{1-x^2}$ oder $y = -\sqrt{1-x^2}$. Wieder gibt es Probleme (etwa in der Nähe von $(\pm 1, 0)$). Insbesondere ist der Kreis “lokal”, aber nicht “global” der Graph einer Funktion $y = f(x)$ von x.

Um die folgende Darstellung nicht mit einer formalen Terminologie zu belasten, werden wir öfters das Wort “glatt” verwenden, ohne genau zu sagen, was wir damit meinen (etwa stetig differenzierbar, n-mal stetig differenzierbar, endlich oft differenzierbar…).

Definition 1 *Eine glatte Abbildung $f : U \to \mathbf{R}^m$, U offen in $\mathbf{R}^n$, heißt Diffeomorphismus von U auf $f(U)$, wenn $f(U)$ wieder offen in $\mathbf{R}^m$ ist und eine glatte Abbildung $g : f(U) \to \mathbf{R}^n$ existiert mit $g \circ f = Id$ und $f \circ g = Id$.*

Bemerkung. f ist dann eine bijektive glatte Abbildung von U auf $f(U)$. Die Abbildung g ist die zu f inverse Abbildung, f^{-1}. Es wird nun verlangt, daß sowohl f als auch g glatt sind. Es gibt glatte Abbildungen f, die invertierbar sind, wobei aber $g = f^{-1}$ nicht glatt ist. Ein Beispiel ist die Abbildung f auf $]-1, 1[$ gegeben durch $f(x) = x^3$. Hier ist $g(x) = \sqrt[3]{x}$. g ist in $x = 0$ nicht differenzierbar.

Satz 2 *Sei f ein Diffeomorphismus zwischen den offenen Mengen $U \subseteq \mathbf{R}^n$ und $f(U) \subseteq \mathbf{R}^m$. Dann für jedes $x \in U$ ist $(Df)_x$ eine invertierbare lineare Abbildung des $\mathbf{R}^n$ in sich.*

Beweis. Nach Voraussetzung gibt es $g : f(U) \to \mathbf{R}^n$ mit $g \circ f = \mathrm{Id}$ und $f \circ g = \mathrm{Id}$. Daher ist nach der Kettenregel für jedes $x \in U$

$$(Dg)_{f(x)}(Df)_x = \mathrm{Id}$$

d.h.

$$(Df)_x(Dg)_{f(x)} = \mathrm{Id}.$$

Die lineare Abbildung $(Df)_x$ besitzt also eine inverse Abbildung und das ist nur möglich, wenn $m = n$ ist.

■

Das obige Beispiel motiviert folgende Definition:

Definition 3 *Eine Abbildung f heißt* **lokaler Diffeomorphismus** *in einem Punkt x, wenn eine offene Umgebung U von x existiert, so daß die Restriktion von f auf U ein Diffeomorphismus von U auf $f(U)$ ist.*

Es ist klar, daß die Zusammensetzung von 2 lokalen Diffeomorphismen wieder ein solcher ist. Der Hauptsatz (ohne Beweis) dieses Kapitels ist:

Satz 4 *(Satz über inverse Funktionen): Sei U eine offene Teilmenge des $\mathbf{R}^n$ und $f : U \to \mathbf{R}^n$ glatt. Sei $x_0 \in U$ und $(Df)_{x_0} : \mathbf{R}^n \to \mathbf{R}^n$ invertierbar.*
Dann ist f ein lokaler Diffeomorphismus in x_0, d.h. f bildet eine hinreichend kleine offene Menge U um x_0 diffeomorph auf eine offene Menge $f(U)$ ab.

Beispiel. (Inversion in einer Sphäre.)

1. Die Abbildung $f : \mathbf{R}^n \setminus \{0\} \to \mathbf{R}^n \setminus \{0\}$, definiert durch $f(x) = \dfrac{x}{|x|^2}$ ist ein Diffeomorphismus.

   Man sieht leicht, dass $f \circ f =$ id, womit klar ist, dass f eine Bijektion ist.

   In diesem Fall läßt sich das Differential leicht direkt berechnen:

$$(Df)_x(h) = \lim_{t \to 0} \frac{1}{t}\left(\frac{x + th}{|x + th|^2} - \frac{x}{|x|^2}\right) =$$

$$\begin{aligned} &= \lim_{t \to 0} \frac{x|x|^2 + th|x|^2 - x|x|^2 - 2t(x|h)x - xt^2|h|^2}{t|x + th|^2|x|^2} \\ &= \frac{h|x|^2 - 2(x|h)x}{|x|^4} = \frac{1}{|x|^4}(|x|^2\mathrm{Id} - 2xx^t)(h). \end{aligned}$$

   Im Fall $n = 2$ ist z.B.

$$(f_1(x, y), f_2(x, y)) = \left(\frac{x}{x^2 + y^2}, \frac{y}{x^2 + y^2}\right)$$

   und

$$(Df)_{(x,y)} = \frac{1}{(x^2 + y^2)^2}\begin{bmatrix} y^2 - x^2, & -2xy \\ -2yx, & x^2 - y^2 \end{bmatrix}$$

2. Für die Abbildung $g : \mathbf{R}^2 \to \mathbf{R}^2$, wobei $g(r, \theta) = (r\cos\theta, r\sin\theta)$ gilt

$$\det\ (Dg)_{(r,\theta)} = \det \begin{bmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{bmatrix} = r.$$

   Sie ist also für $r \neq 0$ ein lokaler Diffeomorphismus.

   Für $r = 0$ bildet g jeden Punkt $(0, \theta)$ in $(0, 0)$ ab, ist also nicht mehr injektiv.

3. Ein weiteres wichtiges Beispiel eines lokalen Diffeomorphismus, der aber kein globaler Diffeomorphismus ist, stellt die Abbildung $z \mapsto e^z$ der komplexen Zahlenebene dar, d.h. die Abbildung

$$(x, y) \mapsto (e^x \cos y, e^x \sin y) \text{ des } \mathbf{R}^2 \text{ in sich.}$$

   Hier ist

$$(Df)_{(x,y)} = \begin{bmatrix} e^x\cos y & -e^x\sin y \\ e^x\sin y & e^x\cos y \end{bmatrix}$$

   und $\det(Df)_{(x,y)} = e^{2x} \neq 0$. Trotzdem ist f nicht injektiv, da $f(x, y + 2\pi) = f(x, y)$.

Ist $f : U \to f(U)$ ein glatter Diffeomorphismus, dann kann man f auf zwei Arten interpretieren, entweder, wie bisher als Abbildung, welche jedem $x \in U$ einen eindeutig bestimmten Punkt $f(x) \in f(U)$ zuordnet oder aber als eine **Koordinatentransformation**. Betrachten wir etwa die ebenen Polarkoordinaten.

$$\begin{aligned} x &= r\cos\theta \quad \text{bzw.} \quad r = \sqrt{x^2+y^2} \\ y &= r\sin\theta \quad \text{bzw.} \quad \theta = \operatorname{Arctg}\frac{y}{x}. \end{aligned}$$

Halten wir $\theta = c$ konstant, so ergeben sich die Geraden Arctg $\frac{y}{x} = c$, d.h. $\frac{y}{x} = \tan c$ oder $y = (\tan c)x$, falls $c \neq \pm\frac{\pi}{2}$. Hält man dagegen $r = R$ konstant, so ergeben sich die Kreise $x^2 + y^2 = R^2$, falls $R \neq 0$.

Diese Geraden und Kreise entsprechen also den Parallelen zur x- bzw. y-Achse bei den kartesischen Koordinaten. Jeder Punkt ist als Schnitt solcher Koordinatenkurven darstellbar. Diese sind überdies eindeutig bestimmt, wenn man vom Nullpunkt absieht, der eine Sonderrolle spielt, weil dort die Abbildung kein lokaler Diffeomorphismus ist.

Anders ausgedrückt: Das Koordinatennetz für Polarkoordinaten sind die Urbilder in der xy-Ebene der Gleichung $r = c, \theta = d$ in der $r - \theta$-Ebene.

Implizite Funktionen

Die problematischen Fragen bei dem Auflösen der Gleichungen $x^2 + y^2 = 1$ werden im folgenden Satz thematisiert. Wir beginnen mit Funktionen von 2 Variablen.

Satz 5 *(Satz über implizite Funktionen.) Sei F eine glatte Abbildung von einer offenen Menge $U \subseteq \mathbf{R}^2$ in die reellen Zahlen. Sei $(x_0, y_0) \in U$. Setzt man voraus, daß $F_y(x_0, y_0) \neq 0$ ist, so existiert eine eindeutig bestimmte glatte Funktion f auf einer geeigneten offenen Umgebung $U(x_0)$ des Punktes x_0, so daß $f(x_0) = y_0 \iff F(x, f(x)) = 0$ $(x \in U(x_0))$.*

Beweis. Wir betrachten die Abbildung $\phi : U \to \mathbf{R}^2$, definiert durch

$$\phi(x, y) = (x, F(x, y)).$$

Dann ist

$$\det(D\phi)_{(x,y)} = \det \begin{bmatrix} 1 & 0 \\ F_x(x,y) & F_y(x,y) \end{bmatrix} = F_y(x, y),$$

wobei F_x etwa die partielle Ableitung von F bzgl. x bezeichnet. Da $F_y(x_0, y_0) \neq 0$ ist, ist ϕ ein lokaler Diffeomorphismus in (x_0, y_0). Es existiert also eine offene Umgebung $V(x_0, y_0)$, deren Bild $\phi(V(x_0, y_0)) = W(x_0, 0)$ eine offene Umgebung des Bildpunktes

$$(x_0, y_0) = (x_0, 0)$$

ist, auf welcher ϕ eine glatte inverse Abbildung ψ besitzt. Diese hat die Gestalt

$$\psi(x, y) = (x, g(x, y))$$

mit einer glatten Funktion $g(x, y)$.

Wegen $\phi\psi(x, y) = (x, y)$ auf $W(x_0, 0)$ und $\phi\psi(x, y) = \phi(x, g(x, y)) = (x, F(x, g(x, y)))$ ist also $F(x, g(x, y)) = y$ auf $W(x_0, 0)$. Da $W(x_0, 0)$ ein offenes Intervall der Gestalt $U(x_0) \times W(0)$ enthält, ist auf $U(x_0)$ die Funktion $f(x) = g(x, 0)$ stetig differenzierbar und erfüllt dort $F(x, f(x)) = F(x, g(x, 0)) = 0$. Damit ist der Satz bewiesen.

■

Beispiele.

1. Sei $F(x,y) = x^2 + y^2 - 1$. Hier ist $F_y = 2y \neq 0$ außer für die Punkte $(\pm 1, 0)$. In diesen Punkten existiert auch keine Funktion der Gestalt $y = f(x)$, die auf einer offenen Umgebung von ± 1 definiert ist und dort $x^2 + (f(x))^2 - 1 = 0$ erfüllt.

2. Sei $F(x,y) = x^2y^5 - \sin(x-y) + x = 0$ und $(x_0, y_0) = (0,0)$. Hier ist $F_y = 5x^2y^4 + \cos(x-y)$ und daher ist $F_y(0,0) = 1 \neq 0$. Es existiert also eine Umgebung von $x = 0$, auf welcher die Gleichung $F(x,y) = 0$ in der Form $y = y(x)$ auflösbar ist.

3. (Niveaulinien.) Sei $f(x,y)$ eine reellwertige glatte Funktion auf einer offenen Menge U. Ist für jedes $(x,y) \in U$ mit $f(x,y) = c$ das Differential $(Df)_{(x,y)} \neq 0$, dann ist die Menge aller (x,y) mit $f(x,y) = c$ eine glatte Kurve oder besteht aus mehreren glatten Kurven. Man nennt diese dann Niveaulinien. Denn für jeden Punkt $(a,b) \in U$ mit $f(a,b) = c$ eine offene Umgebung existiert, auf welcher die Menge aller Punkte (x,y) mit $f(x,y) = c$ eine glatte orientierte Kurve durch (a,b) bilden.

   Nach Voraussetzung ist $Df = (f_x, f_y) \neq 0$, d.h. entweder $f_x \neq 0$ oder $f_y \neq 0$. Im ersten Fall ist die durch $f(x,y) = c$ definierte implizite Funktion von der Gestalt $x = \phi(y)$ und im zweiten Fall von der Gestalt $y = \psi(x)$.

Wir wissen bereits, daß die implizit definierte Funktion f aus dem obigen Satz glatt ist. Man kann die Ableitung wie folgt berechnen.
Wir haben in einer Umgebung von x_0 die Identität $F(x, f(x)) = 0$. Wenden wir darauf die Kettenregel an, so ergibt sich

$$0 = \frac{d}{dx} F(x, f(x)) = F_x + F_y f'(x).$$

Da $F_y(x_0, y_0) \neq 0$ ist, ergibt sich daraus

$$f'(x_0) = -\frac{F_x(x_0, y_0)}{F_y(x_0, y_0)}.$$

Beispiel. Aus $x^2 + y^2 = 1$ folgt $2x + 2y\frac{dy}{dx} = 0$ – also $\frac{dy}{dx} = -\frac{x}{y}$.
Ist in einem Punkt (x_0, y_0) das Differential $(DF)_{(x_0,y)0} = 0$, d.h. $F_x(x_0, y_0) = F_y(x_0, y_0) = 0$, so können verschiedene Situationen vorliegen.
Es kann sein, daß (x_0, y_0) ein isolierter Punkt ist, d.h. daß in einer Ugebung von (x_0, y_0) keine weitere Lösung von $F(x,y) = 0$ existiert, wie etwa im Fall $F(x,y) = x^2 + y^2 = 0$. Oder aber es können sich mehrere glatte Kurven in diesem Punkt schneiden, wie etwa für $F(x,y) = x^2 - y^2 = 0$. Hier fallen also zwei Lösungen zusammen.

Berechnung der höheren Ableitung von f Man geht von der Gleichung $F_x + F_y f'(x) = 0$ aus und differenziert diese nach x:

$$\begin{aligned} 0 &= \frac{d}{dx}(F_x(x, f(x)) + F_y(x, f(x)) f'(x)) = F_{xx}(x, f(x)) \\ &+ F_{xy}(x, f(x)) f'(x) + F_{yx}(x, f(x)) f'(x) + F_{yy}(x, f(x))(f'(x))^2 \\ &+ F_y(x, f(x)) f''(x), \end{aligned}$$

Daher gilt:

$$f''(x) = -\frac{F_{xx} + 2F_{xy} f'(x) + F_{yy}(f'(x))^2}{F_y}.$$

Setzt man hier $f'(x) = -\dfrac{F_x}{F_y}$ ein, bekommt man die Gleichung

$$f''(x) = \frac{2F_{xy}F_xF_y - F_{xx}F_y^2 + F_{yy}(F_x)^2}{(F_y)^3}.$$

Analog kann man höhere Ableitungen berechnen.
Wir betrachten jetzt den Fall einer glatten Funktion $F(x_1, x_2, \ldots, x_n)$ von n Variablen, insbesonders die Frage: Wann läßt sich x_n als Funktion $x_n = f(x_1, \ldots, x_{n-1})$ der übrigen Variablen darstellen? Man zeigt völlig analog: Sei F auf einer offenen Menge U definiert und $x_0 \in U$ mit $F(x_0) = 0$. Ist $F_{x_n}(x_0) \neq 0$, so existiert eine eindeutig bestimmte stetig differenzierbare Funktion $f(x_1, \ldots, x_{n-1})$ auf einer geeigneten offenen Umgebung $U((x_1^0, \ldots, x_{n-1}^0))$ des Punktes $(x_1^0, \ldots, x_{n-1}^0) \in \mathbf{R}^{n-1}$, so daß $f((x_1^0, \ldots, x_{n-1}^0)) = x_{n_0}$ und $F(x_1, \ldots, x_{n-1}, f(x_1, \ldots, x_{n-1})) = 0$ auf dieser Umgebung erfüllt ist.
Der Beweis verläuft genauso wie oben. Man betrachtet die Abbildung $\phi : U \to \mathbf{R}^n$, definiert durch

$$\phi(x_1, x_2 \ldots, x_{n-1}, x_n) = (x_1, \ldots, x_{n-1}, F(x_1, \ldots, x_{n-1}, x_n)).$$

Dann ist det $(D\phi)_{x_0} = F_x(x_0) \neq 0$ und daher existiert eine glatte inverse Abbildung ψ, usw.
Der Satz über implizite Funktionen kann als Präzisierung des heuristischen Prinzips aufgefaßt werden, daß man aus einer Gleichung in s Unbekannten immer eine Unbekannte ausrechnen kann.
Allgemeiner wird man erwarten, daß man aus m Gleichungen in $m + n$ Unbekannten immer m Unbekannte ausrechnen kann. Diese hängen dann von den restlichen n Parametern ab. Es muß allerdings gewährleistet sein, daß die m Gleichungen wirklich m verschiedene Bedingungen darstellen und nicht auf weniger Gleichungen reduziert werden können. In der linearen Algebra heißt das, daß die entsprechenden linearen Gleichungen linear unabhängig sein müssen, d.h., daß eine $m \times m$-Untermatrix existieren muß, deren Determinante von 0 verschieden ist.
Eine ganz analoge Eigenschaft muß im allgemeinen Fall erfüllt sein, wie der folgende grundlegende Satz zeigt.

Satz 6 *(Allgemeiner Satz über implizite Funktionen.) Sei F eine glatte Abbildung von einer offenen Menge $U \subseteq \mathbf{R}^{m+n}$ in den $\mathbf{R}^m$, die dem Element $(x, y) = (x_1, \ldots, x_m, y_1, \ldots, y_n) \in U$ eine Zahl zuordnet. Sei $(x_0, y_0) \in U$ mit $F(x_0, y_0) = (0, 0, \ldots, 0)$. Setzt man voraus, daß die Funktionaldeterminante*

$$\frac{\partial(F_1, F_2, \ldots, F_m)}{\partial(y_1, y_2, \ldots, y_m)} \neq 0$$

ist, so existiert eine eindeutig bestimmte glatte Abbildung f auf einer geeigneten offenen Umgebung $U(x_0)$ des Punktes $x_0 \in \mathbf{R}^n$ in den $\mathbf{R}^m$, so daß $f(x_0) = y_0$ gilt und $F(x, f(x)) = 0$ auf $U(x_0)$ erfüllt ist.

Beweis. Wir betrachten die Abbildung $\phi : U \to \mathbf{R}^{m+n}$, definiert durch

$$\phi(x, y) = (x, F(x, y)) \text{ für } x \in \mathbf{R}^n, y \in \mathbf{R}^m.$$

Dann ist

$$\det D\phi_{(x,y)} = \det \begin{bmatrix} 1 & 0 & \dots & 0 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 1 & 0 & \dots & 0 \\ \frac{\partial F_1}{\partial x_1} & \frac{\partial F_1}{\partial x_2} & \dots & \frac{\partial F_1}{\partial x_n} & \frac{\partial F_1}{\partial y_1} & \dots & \frac{\partial F_1}{\partial y_m} \\ \vdots & & & & & & \\ \frac{\partial F_t}{\partial x_1} & \frac{\partial F_1}{\partial x_2} & \dots & \frac{\partial F_m}{\partial x_n} & \frac{\partial F_m}{\partial y_1} & \dots & \frac{\partial F_m}{\partial y_m} \end{bmatrix}.$$

Das ist aber nichts anderes als

$$\det \begin{bmatrix} \frac{\partial F_1}{\partial y_1} & \dots & \frac{\partial F_1}{\partial y_m} \\ \dots & \dots & \dots \\ \dots & \dots & \dots \\ \frac{\partial F_m}{\partial y_1} & \dots & \frac{\partial F_m}{\partial y_m} \end{bmatrix} = \frac{\partial(F_1, \dots, F_m)}{\partial(y_1, \dots, y_m)}.$$

Von hier ab verläuft der Beweis wieder genau wie oben.

■

Maxima und Minima mit Nebenbedingungen (Lagrange Multiplikationen)
Wir betrachten jetzt Extremalprobleme der folgenden Art: Man soll Maxima oder Minima für eine Funktion $f(x, y)$ berechnen, wenn x und y nicht unabhängige Veränderliche sind, sondern durch eine Nebenbedingung $g(x, y) = 0$ miteinander verknüpft sind.
Im Prinzip könnte man stets eine Veränderliche durch die andere aus $g(x, y) = 0$ ausdrücken und hätte das Problem auf Extrema einer Veränderlichen reduziert. In der Praxis ist das aber sehr oft schwierig. Hier erweist sich eine Methode, die auf Lagrange zurückgeht, als besonders nützlich: Nehmen wir an, daß g auf einer offenen Menge U des $\mathbf{R}^2$ definiert ist und daß $M = \{(x, y) : g(x, y) = 0\}$ eine glatte Menge ist. Das ist definitionsgemäß genau dann der Fall, wenn für jeden Punkt $(a, b) \in M$ das Differential $(Dg)_{(a,b)} \neq 0$. Wir suchen einen solchen Punkt $(a, b) \in M$, in welchem die Funktion f ein lokales Extremum annimmt. Dazu betrachten wir eine lokale Parameterdarstellung für M in der Umgebung des Punktes (a, b), etwa $t \mapsto \varphi(t) = (\varphi_1(t), \varphi_2(t))$ mit $\varphi(t_0) = (a, b)$.
Dann nimmt die Funktion $t \mapsto f(\varphi(t))$ im Punkt t_0 ein lokales Extremum an. Daher gilt $f(\varphi(t))' = 0$ im Punkt t_0 oder $(Df)_{(a,b)}(\dot\varphi(t_0)) = 0$. Weiters ist nach Definition von M die Funktion g auf M identisch 0 und daher gilt $g(\varphi(t)) \equiv 0$.
Differenziert man diese Identität im Punkt t_0, so erhält man $(Dg)_{(a,b)}(\dot\varphi(t_0)) = 0$. Also steht auch der Vektor $(\operatorname{grad} g)(a, b)$ orthogonal auf dem Tangentenvektor $\dot\varphi(t_0)$. Da im $\mathbf{R}^2$ die Vektoren, die auf einem festen Vektor senkrecht stehen, einen eindimensionalen Teilraum aufspannen, muß gelten $\operatorname{grad} f(a, b) = \lambda \operatorname{grad} g(a, b)$ mit einer Konstanten λ oder anders ausgedrückt, $(Df)_{(a,b)} = \lambda(Dg)_{(a,b)}$. Im Koordinantenschreibweise heißt das:

$$f_x - \lambda g_x = 0 \text{ und } f_y - \lambda g_y = 0.$$

Fassen wir diese Überlegung zusammen, so erhalten wir die **Lagrangesche Regel**: Um die kritischen Punkte der Funktion $f(x, y)$ unter der Nebenbedingung $g(x, y) = 0$ zu bestimmen, betrachte man die neue Funktion $F(x, y) = f(x, y) - \lambda g(x, y)$ mit einem

Multiplikator λ und betrachte die drei Gleichungen

$$\begin{aligned} f_x - \lambda g_x &= 0 \\ f_y - \lambda g_y &= 0 \\ g(x, y) &= 0. \end{aligned}$$

Die allgemeine Situation ist wie folgt: Sei f eine Funktion von s Veränderlichen und seien $1 \leq p \leq s$ Bedingungen an die x_i gegeben, etwa $g_1 = g_2 = \cdots = g_p = 0$. Man bestimme lokale Extrema von f unter diesen Nebenbedingungen.

Satz 7 *Sei $g : U \to \mathbf{R}^p$, (U offen in $\mathbf{R}^n$, $1 \leq p \leq n$), eine glatte Abbildung und $M = \{x \in U : g(x) = 0\}$. Für jedes $x \in M$ sei das Differential $(Dg)_x : \mathbf{R}^n \to \mathbf{R}^p$ surjektiv. Sei weiters f eine glatte Funktion, die auf einer offenen $U \subset M$ definiert ist.*
Dann gilt: Nimmt f in einem Punkt $a \in M$ ein lokales Extremum an, so existieren p reelle Zahlen $\lambda_1, \ldots, \lambda_p$ mit

$$(Df)_a = \sum_{i=1}^{p} \lambda_i (Dg_i)_a.$$

Eine benutzerfreundlichere Formulierung wäre:
Sind in einer differenzierbaren Funktion $y = f(x_1, x_2, \ldots, x_n)$ die n Veränderlichen $x_1, \ldots, x_n$ nicht unabhängig, sondern durch p unabhängige Nebenbedingungen

$$g_1(x_1, \ldots, x_n) = 0, \ldots, g_p(x_1, \ldots, x_n) = 0, \quad 1 \leq p < n,$$

miteinander verknüpft, so führt man p zusätzliche Zahlen $\lambda_1, \ldots, \lambda_p$, die sogenannten **Lagrangeschen Multiplikatoren**, ein und betrachtet die $n + p$ Gleichungen

$$\frac{\partial F}{\partial x_1} = 0, \ldots, \frac{\partial F}{\partial x_n} = 0, \quad g_1 = 0, \ldots, g_p = 0$$

für die $n+p$-Unbekannten $x_1, \ldots, x_n, \lambda_1, \ldots, \lambda_p$, wobei $F = f - (\lambda_1 g_1 + \cdots + \lambda_p g_p)$ gesetzt wurde. Unter den im Satz angegebenen Voraussetzungen muß jedes lokale Extremum eine Lösung dieser Gleichungen darstellen.

Beispiel. Man berechne das Minimum der Funktion

$$f(x_1, \ldots, x_n) = x_1 + x_2 + \cdots + x_n \qquad (x_i > 0)$$

unter der Nebenbedingung $g(x_1, \ldots, x_n) = x_1 x_2 \ldots x_n - 1 = 0$.
Hier führt man $F(x_1, \ldots, x_n) = x_1 + \cdots + x_n - \lambda(x_1 \ldots x_n - 1)$ ein. Es muß dann gelten:

$$\frac{\partial F}{\partial x_i} = 1 - \lambda \frac{x_1 \ldots x_n}{x_i} = 0$$

und $x_1 \ldots x_n = 1$.
Daraus erhält man $x_i = \lambda$ für alle i und $\lambda^n = 1$. Es ist also das Minimum gegeben durch $f(1, \ldots, 1) = n$. Anders ausgedrückt:

$$\frac{x_1}{\sqrt[n]{x_1 \ldots x_n}} + \cdots + \frac{x_n}{\sqrt[n]{x_1 \ldots x_n}} \geq n \text{ oder } \frac{x_1 + x_2 + \cdots + x_n}{n} \geq \sqrt[n]{x_1 x_2 \ldots x_n}$$

für alle $x_i > 0$. Das ist die Ungleichung zwischen dem arithmetischen und geometrischen Mittel.