lokale Umkehrung von der Klasse $C^r$ ist, haben wir bereits vor (11.1) begründet. ■

Eine injektive Abbildung $F : M \to \mathbb{R}^n$ (mit offenem $M \subset \mathbb{R}^n$) der Klasse $C^r$ (für ein $r \in \mathbb{N}$) mit überall regulärem Differential nennt man auch einen *Diffeomorphismus* der Klasse $C^r$ von $M$ auf $F(M)$. Ist nun $F : M \to \mathbb{R}^n$ eine Abbildung der Klasse $C^r$ und $DF_{X_0}$ regulär, so gibt es nach (11.2) eine offene Umgebung $U$ von $X_0$ derart, daß $F|U$ ein Diffeomorphismus der Klasse $C^r$ von $U$ auf $F(U)$ ist. Für diesen Sachverhalt sagt man auch, $F$ sei bei $X_0$ lokal ein $C^r$-Diffeomorphismus.

Als eine Folgerung aus dem Satz über die Umkehrabbildung wollen wir nun eine Aussage über sogenannte „implizite Funktionen" gewinnen. Dieses für manche Anwendungen wichtige Thema wollen wir zunächst anschaulich erläutern.

Verschiedenartige Aufgabenstellungen erfordern die Untersuchung von Gebilden im euklidischen Raum, wie „Kurven" oder „Flächen", die beschrieben werden durch Gleichungen, die zwischen den Koordinaten ihrer Punkte bestehen. Als einfaches Beispiel betrachten wir im $\mathbb{R}^3$ die „Einheitssphäre"

$$S^2 := \{X \in \mathbb{R}^3 \mid x_1^2 + x_2^2 + x_3^2 - 1 = 0\}.$$

Die genauere Untersuchung dieser Punktmenge kann dadurch erleichtert werden, daß man die definierende Gleichung $x_1^2 + x_2^2 + x_3^2 = 1$ nach einer Variablen auflöst, indem man etwa

$$x_3 = \sqrt{1 - x_1^2 - x_2^2}$$

schreibt. Man betrachtet dann die Punktmenge

$$S_+ := \{X \in \mathbb{R}^3 \mid x_3 = \sqrt{1 - x_1^2 - x_2^2},\ x_1^2 + x_2^2 < 1\}.$$

Allerdings ist $S_+$ nur ein echter Teil von $S^2$, die „obere Halbsphäre". Offenbar kann die ganze Menge $S^2$ nicht einheitlich dadurch dargestellt werden, daß man eine Koordinate als Funktion der beiden anderen schreibt. Aber zur lokalen Untersuchung von $S^2$ ist eine derartige Darstellung gut geeignet. Es stört nicht, daß man zur Untersuchung von $S^2$ nicht mit nur einer solchen Darstellung auskommt. Will man $S^2$ in einer Umgebung des Punktes $(0, 0, -1)$ untersuchen, so benutzt man die „Auflösung" $x_3 = -\sqrt{1 - x_1^2 - x_2^2}$, in einer Umgebung von $(1, 0, 0)$ benutzt man die Auflösung $x_1 = \sqrt{1 - x_2^2 - x_3^2}$, usw.

Analog kann man Punktmengen untersuchen, die durch mehrere Gleichungen zwischen den Koordinaten definiert sind. Als Beispiel betrachten wir die Punktmenge

$$K := \left\{X \in \mathbb{R}^3 \mid x_1^2 + x_2^2 + x_3^2 - 1 = 0 \text{ und } \left(x_1 - \frac{1}{2}\right)^2 + x_2^2 - \frac{1}{4} = 0\right\},$$

also den Durchschnitt der Sphäre $S^2$ mit einem gewissen Kreiszylinder mit Erzeugenden parallel zur $x_3$-Achse. Eine „lokale Auflösung" lautet

$$x_1 = 1 - x_3^2,\ \ x_2 = x_3\sqrt{1 - x_3^2},\ \ |x_3| < 1,$$

eine andere

$$x_1 = 1 - x_3^2,\ \ x_2 = -x_3\sqrt{1 - x_3^2},\ \ |x_3| < 1.$$

Diese Darstellungen können zur lokalen Untersuchung der Schnittkurve $K$ benutzt werden.

Im allgemeinen wird es nicht möglich sein, eine „lokale Auflösung" explizit (formelmäßig) zu bewerkstelligen. Hier sind dann Aussagen von Interesse, die die grundsätzliche Möglichkeit einer solchen Auflösung (und die Differenzierbarkeit der darstellenden Funktionen) gewährleisten. Zunächst soll aber die Fragestellung allgemeiner und genauer gefaßt werden.

Eine Punktmenge $N$ des $\mathbb{R}^n$ sei definiert und beschrieben durch ein Gleichungssystem

$$(*) \qquad \begin{cases} f_1(x_1, \dots, x_n) = 0, \\ f_2(x_1, \dots, x_n) = 0, \\ \vdots \\ f_k(x_1, \dots, x_n) = 0. \end{cases}$$

Dabei sei $M \subset \mathbb{R}^n$ eine offene Teilmenge, und

$$f_i : M \to \mathbb{R}, \qquad i = 1, \dots, k < n,$$

seien reellwertige Funktionen der Klasse $C^1$. Sodann sei

$$N := \{X \in M \mid X = (x_1, \dots, x_n) \ \ \text{erfüllt } (*)\}.$$

(Der Buchstabe $N$ soll an „Nullstellenmenge" erinnern.) Das Problem besteht nun darin, die Punktmenge $N$ in einer Umgebung eines gegebenen Punktes $X_0 \in N$ zu beschreiben, indem $k$ Koordinaten als Funktionen der übrigen $n-k$ Koordinaten dargestellt werden, etwa

$$\begin{aligned} x_{n-k+1} &= g_1(x_1, \dots, x_{n-k}), \\ &\ \vdots \\ x_n &= g_k(x_1, \dots, x_{n-k}) \end{aligned}$$

für $(x_1, \dots, x_{n-k})$ aus einer passenden Menge $V \subset \mathbb{R}^{n-k}$. Das ist so zu verstehen: Setzt man die so berechneten Koordinaten in $(*)$ ein, so soll dieses System erfüllt sein, also

$$f_i(x_1, \dots, x_{n-k}, g_1(x_1, \dots, x_{n-k}), \dots, g_k(x_1, \dots, x_{n-k})) = 0$$

für $(x_1, \ldots, x_{n-k}) \in V$ und $i = 1, \ldots, k$. Genauer gesagt: Jeder Punkt

$$(x_1, \ldots, x_{n-k}, g_1(x_1, \ldots, x_{n-k}), \ldots, g_k(x_1, \ldots, x_{n-k}))$$

mit $(x_1, \ldots, x_{n-k}) \in V$ soll zu $N$ gehören; umgekehrt soll aber auch der Durchschnitt von $N$ mit einer ganzen Umgebung von $X_0$ auf diese Weise erhalten werden. Man sagt dann, die Funktionen $g_1, \ldots, g_k$ seien „implizit definiert" durch die Gleichungen

$$f_1(X) = 0, \ldots, f_k(X) = 0.$$

Nützlich ist die hierdurch gegebene lokale Auflösung i.a. nur, wenn die Funktionen $g_1, \ldots, g_k$ ihrerseits so oft differenzierbar sind wie die Funktionen $f_1, \ldots, f_k$. Für die Möglichkeit einer solchen Auflösung gibt Satz (11.3) eine hinreichende Bedingung.

Es ist im folgenden zweckmäßig, die Funktionen $f_1, \ldots, f_k$ als Koordinatenfunktionen einer Abbildung $F : M \to \mathbb{R}^k$ und ebenso die Funktionen $g_1, \ldots, g_k$ als Koordinatenfunktionen einer Abbildung $G : V \to \mathbb{R}^k$ aufzufassen. Unter dem *Graphen* von $G$ versteht man die Menge

$$\text{Graph } G := \{(X, G(X)) \in \mathbb{R}^{n-k} \times \mathbb{R}^k \mid X \in V\}.$$

Der zu beweisende Satz läßt sich dann so formulieren, daß man eine Nullstellenmenge unter passenden Voraussetzungen lokal als Graph darstellen kann.

(11.3) **Satz** (über implizit definierte Funktionen). *Sei $k < n$, $M \subset \mathbb{R}^n$ offen, $F : M \to \mathbb{R}^k$ eine Abbildung der Klasse $C^r$ (für ein $r \in \mathbb{N}$), sei*

$$N := \{X \in M \mid F(X) = 0\}.$$

*Sei $X_0 \in N$ und $DF_{X_0}$ vom Rang $k$. Dann gibt es (nach passender Identifizierung von $\mathbb{R}^n$ mit $\mathbb{R}^{n-k} \times \mathbb{R}^k$) eine offene Umgebung $U$ von $X_0$ in $M$, eine offene Menge $V$ in $\mathbb{R}^{n-k}$ und eine Abbildung $G : V \to \mathbb{R}^k$ der Klasse $C^r$ mit*

$$N \cap U = \text{ Graph } G.$$

*Beweis.* Das Differential $DF_{X_0}$ ist nach Voraussetzung vom Rang $k$. Die Funktionalmatrix von $F$ in $X_0$ hat also $k$ linear unabhängige Spalten. Wir dürfen o.B.d.A. (d.h. nach passender Umnumerierung der Basisvektoren) annehmen, daß dies die letzten $k$ Spalten sind. Wir identifizieren dann $\mathbb{R}^{n-k}$ mit dem von den ersten $n-k$ Basisvektoren und $\mathbb{R}^k$ mit dem von den letzten $k$ Basisvektoren des $\mathbb{R}^n$ aufgespannten Unterraum, und wir identifizieren $\mathbb{R}^n$ mit $\mathbb{R}^{n-k} \times \mathbb{R}^k$. Jeder Punkt $X$ aus $\mathbb{R}^n$ läßt sich also eindeutig in der Form $(X', X'')$ mit $X' \in \mathbb{R}^{n-k}$ und $X'' \in \mathbb{R}^k$ schreiben. Wir definieren die Abbildung

$$\begin{array}{rccc} \Phi : & M & \to & \mathbb{R}^{n-k} \times \mathbb{R}^k \\ & (X', X'') & \mapsto & (X', F(X)). \end{array}$$

Dann ist $\Phi$ eine Abbildung der Klasse $C^r$. Wegen

$$J\Phi = \begin{pmatrix} 1 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & & \ddots & \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 1 & 0 & \cdots & 0 \\ \partial_1 f_1 & & \cdots & \partial_{n-k} f_1 & \partial_{n-k+1} f_1 & \cdots & \partial_n f_1 \\ \vdots & & \vdots & \vdots & \vdots & \vdots & \vdots \\ \partial_1 f_k & & \cdots & \partial_{n-k} f_k & \partial_{n-k+1} f_k & \cdots & \partial_n f_k \end{pmatrix}$$

ist $J\Phi$ an der Stelle $X_0$ vom Rang $n$. Nach Satz (11.2) ist daher $\Phi$ in einer Umgebung $U_0$ von $X_0$ ein Diffeomorphismus der Klasse $C^r$. Sei $\Psi$ seine Umkehrabbildung. $\Psi$ ist definiert auf einer offenen Umgebung $W$ von $\Phi(X_0) = (X_0', F(X_0)) = (X_0', 0)$. Wir setzen

$$V := \{X' \in \mathbb{R}^{n-k} \mid (X', 0) \in W\},$$

dann ist $V \subset \mathbb{R}^{n-k}$ eine offene Umgebung von $X_0'$. Wegen $\Phi(X', X'') = (X', F(X', X''))$ ist

$$\Psi(X', Z) = (X', \Psi_2(X', Z)),$$

wodurch eine Abbildung $\Psi_2 : W \to \mathbb{R}^k$ definiert wird. Wir setzen nun

$$G(X') := \Psi_2(X', 0) \qquad \text{für } X' \in V.$$

Dann ist $G : V \to \mathbb{R}^k$ eine Abbildung der Klasse $C^r$. Anwendung von $\Psi$ auf die Gleichung

$$\Phi(X_0', X_0'') = (X_0', F(X_0)) = (X_0', 0)$$

ergibt

$$(X_0', X_0'') = \Psi(X_0', 0) = (X_0', \Psi_2(X', 0)) = (X_0', G(X_0')),$$

also $G(X_0') = X_0''$. Für $X' \in V$ gilt $\Psi(X', 0) = (X', G(X'))$ (nach Definition von $\Psi_2$ und $G$). Anwendung von $\Phi$ ergibt

$$(X', 0) = \Phi(X', G(X')) = (X', F(X', G(X'))),$$

also $F(X', G(X')) = 0$ und damit

$$(X', G(X')) \in N.$$

Wir können o.B.d.A. $V \times G(V) \subset U_0$ annehmen, denn dies läßt sich durch Verkleinerung von $V$ erreichen, da $G$ in $X_0'$ stetig und $G(X_0') = X_0''$ ist. Setze $U := (V \times \mathbb{R}^k) \cap U_0$. Dann gilt

$$N \cap U = \text{ Graph } G.$$

Zum Beweis sei $(X', Y) \in$ Graph $G$, also $X' \in V$ und $Y = G(X')$. Dann ist $(X', Y) = (X', G(X')) \in V \times G(V) \subset U$ und, wie eben gezeigt, $(X', G(X')) \in N$.

Ist umgekehrt $(X', Y) \in N \cap U$, so ist

$$(X', F(X', Y)) = (X', 0) = (X', F(X', G(X'))),$$

also nach Definition von $\Phi$

$$\Phi(X', Y) = \Phi(X', G(X')).$$

Da $\Phi$ injektiv ist, folgt $Y = G(X')$, also $(X', Y) \in$ Graph $G$. ■

Satz (11.2) gibt die Möglichkeit, auch für $k \neq n$ differenzierbare Abbildungen $F : M \to \mathbb{R}^k$ mit nichtausgeartetem Differential in einer Weise zu beschreiben, die zeigt, daß solche Abbildungen sich lokal qualitativ so wie ihr Differential verhalten. Für $k < n$ ist diese Beschreibung im wesentlichen schon in (11.3) enthalten; für $k > n$ wird sie auf analoge Weise erhalten.

(11.4) **Definition.** Sei $F : M \to \mathbb{R}^k$ differenzierbar in $X_0 \in M$. Dann wird der Rang des Differentials $DF_{X_0}$ (= Rang der Funktionalmatrix $JF(X_0)$) als *Rang* der Abbildung $F$ in $X_0$ bezeichnet.

Sei nun $M \subset \mathbb{R}^n$ offen und $F : M \to \mathbb{R}^k$ eine Abbildung der Klasse $C^1$. Der Rang von $F$ in $X \in M$ kann dann höchstens gleich $\min\{n, k\}$ sein. Ist er gleich dieser Zahl, so sagt man, $F$ sei in $X$ *von maximalem Rang.* Eine Teilaussage von Satz (11.2) kann man auch so formulieren: Im Fall $k = n$ ist jede $C^1$-Abbildung, die in einem Punkt von maximalem Rang ist, in einer Umgebung dieses Punktes ein Diffeomorphismus. Wir wollen analoge Aussagen für $k \neq n$ gewinnen. Sie besagen, daß man differenzierbare Abbildungen maximalen Ranges lokal in sehr übersichtlicher Weise durch Diffeomorphismen und lineare Abbildungen beschreiben kann.

Bezeichnung. Sind $f : A \to B$, $g : B \to C$, $h : A \to B$ ($A, B, C$ Mengen) Abbildungen mit $g \circ f = h$, so drückt man dies auch aus durch die Sprechweise: „Das Diagramm

$$\begin{tikzcd} A \arrow[r, "f"] \arrow[dr, "h"'] & B \arrow[d, "g"] \\ & C \end{tikzcd}$$

ist kommutativ.“ [Anschaulich: Man kommt, ausgehend von einem $a \in A$, zu demselben Element von $C$, gleichgültig auf welchem Wege man den Pfeilen folgt.]

Bezeichnung. Sind $A, B$ Mengen, so sind auf dem kartesischen Produkt $A \times B$ die *Projektionen* $\pi_1, \pi_2$ definiert durch

$$\begin{array}{rrcl} \pi_1: & A \times B & \to & A, \\ & (a,b) & \mapsto & a \end{array} \qquad \begin{array}{rrcl} \pi_2: & A \times B & \to & B \\ & (a,b) & \mapsto & b. \end{array}$$

Sind $A, B$ Vektorräume, so wird die *kanonische Injektion* $i$ von $A$ in $A \times B$ definiert durch

$$\begin{array}{rrcl} i: & A & \to & A \times B \\ & a & \mapsto & (a,0). \end{array}$$

(11.5) **Satz** (über lokal surjektive Abbildungen). *Sei $k < n$. Sei $M \subset \mathbb{R}^n$ offen, sei $F : M \to \mathbb{R}^k$ eine Abbildung der Klasse $C^r$ $(r \geq 1)$. Sei $X_0 \in M$ und $F$ in $X_0$ vom Rang $k$ (also $DF_{X_0}$ surjektiv). Dann gibt es eine offene Umgebung $U$ von $X_0$ in $M$, eine offene Menge $V$ in $\mathbb{R}^{n-k}$ und einen $C^r$-Diffeomorphismus $h : U \to V \times F(U)$, so daß das Diagramm*

$$\begin{tikzcd} U \arrow[r, "h"] \arrow[dr, "F"'] & V \times F(U) \arrow[d, "\pi_2"] \\ & F(U) \end{tikzcd}$$

*kommutativ ist.*

Bemerkung. Die Abbildung $F$ ist also in einer Umgebung des Punktes $X_0$, in dem das Differential surjektiv ist, dargestellt als Komposition eines $C^r$-Diffeomorphismus und einer surjektiven linearen Abbildung.

*Beweis von Satz* (11.5). Nach Voraussetzung hat die Funktionaldeterminante $JF(X_0)$ $k$ linear unabhängige Spalten. Wir können o.B.d.A. annehmen, daß dies die letzten $k$ Spalten sind. Wir fassen jetzt wieder $\mathbb{R}^n$ als das kartesische Produkt

$\mathbb{R}^{n-k} \times \mathbb{R}^k$ auf und bezeichnen mit $\pi_1, \pi_2$ die zugehörigen Projektionen. Aus dem Beweis von Satz (11.3) folgt, daß die Abbildung

$$\begin{array}{rccc} \varphi: & M & \to & \mathbb{R}^{n-k} \times \mathbb{R}^k \\ & X & \mapsto & (\pi_1(X), F(X)) \end{array}$$

auf einer offenen Umgebung $U'$ von $X_0$ ein $C^r$-Diffeomorphismus ist. Das Bild $\varphi(U')$ enthält eine offene Umgebung von $\varphi(X_0) = (\pi_1(X_0), F(X_0))$ der Form $V \times W$, dabei ist also $V$ offen in $\mathbb{R}^{n-k}$. Setze $U := \varphi^{-1}(V \times W)$ und $h := \varphi|U$. Dann ist $F(U) = W$, und für $X \in U$ gilt $h(X) = (\pi_1(X), F(X))$, also $\pi_2 \circ h = F$. ■

Der folgende Satz ist das Gegenstück zu (11.5) für den Fall $k > n$.

(11.6) **Satz** (über lokal injektive Abbildungen). *Sei $k > n$. Sei $M \subset \mathbb{R}^n$ offen, sei $F : M \to \mathbb{R}^k$ eine Abbildung der Klasse $C^r$ $(r \geq 1)$. Sei $X_0 \in M$ und $F$ in $X_0$ vom Rang $n$ (also $DF_{X_0}$ injektiv). Dann gibt es eine offene Umgebung $U$ von $X_0$ in $M$, eine offene Umgebung $V$ von 0 in $\mathbb{R}^{k-n}$, eine offene Umgebung $W$ von $F(X_0)$ in $\mathbb{R}^k$ und einen $C^r$-Diffeomorphismus $h : U \times V \to W$, so daß das Diagramm*

$$\begin{tikzcd} U \arrow[r, "i"] \arrow[dr, "F"'] & U \times V \arrow[d, "h"] \\ & W \end{tikzcd}$$

(*wo $i$ die kanonische Injektion bezeichnet*) *kommutativ ist.*

Bemerkung. Die Abbildung $F$ ist also in einer Umgebung des Punktes $X_0$, in dem das Differential injektiv ist, dargestellt als Komposition einer injektiven linearen Abbildung und eines $C^r$-Diffeomorphismus.

*Beweis von Satz* (11.6). Nach Voraussetzung hat die Funktionalmatrix $JF(X_0)$ $n$ linear unabhängige Zeilen. Wir können o.B.d.A. annehmen, daß dies die ersten $n$ Zeilen sind. Wir fassen den Raum $\mathbb{R}^k$ als das kartesische Produkt $\mathbb{R}^n \times \mathbb{R}^{k-n}$ auf und definieren

$$\begin{array}{rccc} \psi: & M \times \mathbb{R}^{k-n} & \to & \mathbb{R}^k \\ & (X, Y) & \mapsto & F(X) + (0, Y). \end{array}$$

Dann ist $\psi$ von der Klasse $C^r$, und die Funktionalmatrix von $\psi$ ist von der Form

$$\left(\begin{array}{c|c} A & 0 \\ \hline B & E \end{array}\right),$$

wo $A$ aus den ersten $n$ Zeilen der Funktionalmatrix von $F$ gebildet ist und $E$ die $(k-n)$-reihige Einheitsmatrix ist. Im Punkt $(X_0, 0)$ ist also die Funktionaldeterminante von $\psi$ ungleich Null. Nach (11.2) gibt es daher offene Umgebungen $U \times V$ (o.B.d.A. in dieser Produktform) von $(X_0, 0)$ und $W$ von $\psi(X_0, 0) = F(X_0)$, so daß $h := \psi|U \times V$ ein $C^r$-Diffeomorphismus von $U \times V$ auf $W$ ist. Für $X \in U$ gilt $h \circ i(X) = h(X, 0) = F(X)$. ■

Es ist nun Gelegenheit, die Untersuchung von Extremwerten reeller Funktionen fortzuführen. Wir wollen eine notwendige Bedingung für das Vorliegen eines lokalen Extremums unter Nebenbedingungen aufstellen, die sogenannte *Multiplikatorenregel von Lagrange*. Zunächst betrachten wir als Beispiel eine typische Aufgabe dieser Art. Es seien die Extrema der Funktion $f(x, y) = x + y$ zu finden unter der Nebenbedingung $5x^2 + 6xy + 2y^2 = 1$. Mit anderen Worten, es sei

$$N := \{(x, y) \in \mathbb{R}^2 \mid 5x^2 + 6xy + 2y^2 = 1\},$$

und es sollen die Extrema der Einschränkung $f|N$ gefunden werden. In diesem einfachen Fall kann man natürlich so vorgehen, daß man die Gleichung $5x^2 + 6xy + 2y^2 = 1$ nach $x$ oder $y$ auflöst, zum Beispiel

$$y = -\frac{3}{2}x \pm \sqrt{\frac{1}{2} - \frac{1}{4}x^2}.$$

Wir setzen daher

$$N^\pm := \left\{(x, y) \in \mathbb{R}^2 \mid |x| < \sqrt{2},\ y = -\frac{3}{2}x \pm \sqrt{\frac{1}{2} - \frac{1}{4}x^2}\right\},$$

$$g^\pm(x) := x + \left(-\frac{3}{2}x \pm \sqrt{\frac{1}{2} - \frac{1}{4}x^2}\right) \qquad \text{für } |x| < \sqrt{2}.$$

Für $(x, y) \in N^+$ gilt dann $f(x, y) = g^+(x)$. Für $(x, y) \in N^+$ hat also $f$ in $(x, y)$ genau dann ein lokales Extremum, wenn $g^+$ in $x$ ein lokales Extremum hat. Nun ist $(g^+)'(x) = 0$ nur für $x = -1$. Hat also $f$ in $N^+$ ein lokales Extremum, so im Punkt $(-1, 2)$. Analog gilt: Hat $f$ in $N^-$ ein lokales Extremum, so im Punkt $(1, -2)$. Man muß allerdings noch nachweisen, z.B. durch Auflösung nach $x$, daß $f$ in den nicht erfaßten Punkten mit $|x| = \sqrt{2}$ keine lokalen Extrema hat. Da $f$ ein Maximum und ein Minimum annimmt, hat $f$ also genau in den beiden Punkten $(-1, 2)$ und $(1, -2)$ lokale Extrema.

Im allgemeinen wird eine solche explizite Auflösung nicht möglich sein. Sie ist aber auch nicht erforderlich, da man folgendermaßen argumentieren kann. Wir können (wie später allgemeiner gezeigt und verwendet wird), die Menge $N$ lokal parametrisieren, das heißt die Punkte $(x, y) \in N$ darstellen in der Form $x = \varphi(t)$, $y = \psi(t)$ mit stetig differenzierbaren Funktionen $\varphi, \psi$. Dann setzen wir $h(t) = f(\varphi(t), \psi(t))$ und bestimmen die kritischen Stellen von $h$. Eine notwendige Bedingung ist $h'(t) = 0$, also

$$\partial_1 f(\varphi(t), \psi(t))\varphi'(t) + \partial_2 f(\varphi(t), \psi(t))\psi'(t) = 0.$$

Ist $g(x, y) = 0$ die Nebenbedingung, so gilt $g(\varphi(t), \psi(t)) = 0$, also

$$\partial_1 g(\varphi(t), \psi(t))\varphi'(t) + \partial_2 g(\varphi(t), \psi(t))\psi'(t) = 0.$$

An einer kritischen Stelle von $h$ sind also, wenn dort nicht gerade $\varphi' = \psi' = 0$ ist, die Vektoren $(\partial_1 f, \partial_2 f)$ und $(\partial_1 g, \partial_2 g)$ linear abhängig. Es gibt also, wenn an der kritischen Stelle $(\partial_1 g, \partial_2 g) \neq (0, 0)$ ist, eine Zahl $\lambda$ mit

$$\begin{aligned} \partial_1 f - \lambda \partial_1 g &= 0, \\ \partial_2 f - \lambda \partial_2 g &= 0. \end{aligned}$$

Zusammen mit $g = 0$ sind das drei Gleichungen, aus denen man im Prinzip die Koordinaten des kritischen Punktes und den Wert von $\lambda$ bestimmen kann. Dabei ist die Zahl $\lambda$, der Multiplikator von Lagrange, in der Regel nicht von Interesse, sondern hat nur eine Hilfsfunktion.

Im obigen Beispiel ist $f(x, y) = x + y$, $g(x, y) = 5x^2 + 6xy + 2y^2 - 1$. Man erhält also die Gleichungen

$$\begin{aligned} &1 - \lambda(10x + 6y) = 0, \\ &1 - \lambda(6x + 4y) = 0, \\ &5x^2 + 6xy + 2y^2 - 1 = 0 \end{aligned}$$

mit den Lösungen $(x, y, \lambda) = (-1, 2, \frac{1}{2})$ und $(1, -2, -\frac{1}{2})$.

Wir wollen dieses Verfahren nun auf mehrere Nebenbedingungen ausdehnen und exakt begründen.

(11.7) **Satz** (Multiplikatorenregel von Lagrange). *Sei $M \subset \mathbb{R}^n$ offen, seien $f : M \to \mathbb{R}$ und $G : M \to \mathbb{R}^k$ mit $1 \leq k < n$ Abbildungen der Klasse $C^1$, sei*

$$N := \{X \in M \mid G(X) = 0\}.$$

*Die Einschränkung $f|N$ habe an der Stelle $X_0 \in N$ ein lokales Extremum, und die Abbildung $G$ mit Koordinatenfunktionen $g_1, \ldots, g_k$ sei an der Stelle $X_0$ vom*

*Rang $k$. Dann gibt es reelle Zahlen $\lambda_1, \dots, \lambda_k$ mit*

$$\partial_i \left( f - \sum_{j=1}^{k} \lambda_j g_j \right) (X_0) = 0 \qquad \text{für } i = 1, \dots, n.$$

*Beweis.* Nach Satz (11.3) gibt es eine Umgebung $U$ von $X_0$, eine offene Menge $V \subset \mathbb{R}^{n-k}$ und eine injektive Abbildung $\varphi : V \to \mathbb{R}^n$ der Klasse $C^1$ vom Rang $n-k$ mit $U \cap N = \varphi(V)$ (nämlich $\varphi(Z) = (Z, G(Z))$ mit $G$ wie in (11.3)). Sei $Z_0$ das Urbild von $X_0$. Die Funktion $h = f \circ \varphi$ hat an der Stelle $Z_0$ ein lokales Extremum. Daher gilt $Dh_{Z_0} = 0$, nach der Kettenregel also

$$Df_{X_0}(D\varphi_{Z_0}(Y)) = 0 \qquad \text{für alle } Y \in \mathbb{R}^{n-k}.$$

Wegen $G \circ \varphi = 0$ gilt auch $DG_{X_0}(D\varphi_{Z_0}(Y)) = 0$, also

$$(Dg_i)_{X_0}(D\varphi_{Z_0}(Y)) = 0 \qquad \text{für alle } Y \in \mathbb{R}^{n-k}.$$

Die im folgenden auftretenden Differentiale von $f$ und $g_i$ sind sämtlich an der Stelle $X_0$ genommen. Wir haben gezeigt, daß

$$Df(H) = 0 \quad \text{und} \quad Dg_i(H) = 0 \quad \text{für} \quad i = 1, \dots, k \quad \text{und} \quad H \in D\varphi_{Z_0}(\mathbb{R}^{n-k})$$

gilt. Sei $L := D\varphi_{Z_0}(\mathbb{R}^{n-k})$ und $L^\perp$ der Orthogonalraum in $\mathbb{R}^n$. Dann ist $\dim L^\perp = k$. Sei $(W_1, \dots, W_k)$ eine Basis von $L^\perp$. Die Zeilen der Matrix

$$\begin{pmatrix} Dg_1(W_1) & \cdots & Dg_k(W_1) \\ \vdots & & \vdots \\ Dg_1(W_k) & \cdots & Dg_k(W_k) \end{pmatrix}$$

sind die Koordinatenvektoren der Bildvektoren $DG_{X_0}(W_1), \dots, DG_{X_0}(W_k)$, sind also linear unabhängig. Daher sind auch die Spalten der Matrix linear unabhängig, und es gibt reelle Zahlen $\lambda_1, \dots, \lambda_k$ mit

$$\begin{pmatrix} Df(W_1) \\ \vdots \\ Df(W_k) \end{pmatrix} = \lambda_1 \begin{pmatrix} Dg_1(W_1) \\ \vdots \\ Dg_1(W_k) \end{pmatrix} + \dots + \lambda_k \begin{pmatrix} Dg_k(W_1) \\ \vdots \\ Dg_k(W_k) \end{pmatrix}.$$

Da jeder Vektor $E \in \mathbb{R}^n$ in der Form $E = H + c_1 W_1 + \dots + c_k W_k$ mit $H \in L$ und $c_1, \dots, c_k \in \mathbb{R}$ dargestellt werden kann, folgt

$$Df(E) = \lambda_1 Dg_1(E) + \dots + \lambda_k Dg_k(E).$$

Mit $E = E_i$, $i = 1, \dots, n$, folgt die Behauptung. ■