Ralitsa Bozhanova Jonas Kindervater

Ferienkurs im Anschluss an das Wintersemester 2008

Höhere Mathematik III für Physiker
Analysis 2

16. bis 20. Februar 2009

Inhaltsverzeichnis

1 Der $\mathbb{R}^n$ als metrischer Raum

1.1 Norm

Der mathematische Begriff der Norm ist die Verallgemeinerung des geometrischen Begriffs der Länge oder des Abstandes. Die Norm ist eine Funktion, die jedem Element eines Vektorraums eine nicht negative reelle Zahl zuordnet.

Definition 1.1 (Norm) *Sei V ein $\mathbb{R}$ Vektorraum. Die Funktion $\|.\| : V \to \mathbb{R}$ heißt* ***Norm****, wenn für alle $x, y \in V$ und $\alpha \in \mathbb{R}$ folgende Eigenschaften erfüllt sind:*

$$\|x\| > 0 \text{ mit } \|x\| = 0 \leftrightarrow x = 0 \qquad \text{(Definitheit)}$$
$$\|\alpha x\| = |\alpha| \, \|x\| \qquad \text{(Homogenität)}$$
$$\|x + y\| \leq \|x\| + \|y\| \qquad \text{(Dreiecksungleichung)}$$

Beispiel 1.1 (p-Norm)

$$\|x\|_p = \sqrt[p]{\sum_{k=0}^{n} |x_k|^p} \tag{1}$$

Der Spezialfall $p = 2$ beschreibt die euklidsche Norm: $\|x\|_2 = \sqrt{x_1^2 + \ldots + x_n^2}$

Allgemein wird durch jedes Skalarprodukt $< .,. >$ eine Norm in der Form $\|x\| = \sqrt{< x, x >}$ induziert. Die Umkehrung gilt jedoch nicht, d.h. nicht jede Norm wird durch ein Skalarprodukt induziert.

1.2 Metrik

Jede Norm induziert durch die Funktion $d(x, y) := \|x - y\|$ eine Metrik. Die Metrik ist eine Funktion, die je zwei Elementen eines Raumes einen nicht negativen Wert zuordnet. Dieser nicht negative Wert kann als Abstand zwischen den Punkten aufgefasst werden. Ein **metrischer Raum** ist ein Raum auf dem eine Metrik und damit ein Abstand definiert ist.

Definition 1.2 (Metrik) *Sei X eine beliebige Menge.*
Die Funktion $d : (x, y) \in X^2 \to \mathbb{R}$ heißt Metrik falls für alle $x, y, z \in X$ gilt:

$$d(x, x) = 0 \qquad \text{(Identische Punkte haben den Abstand Null)}$$
$$d(x, y) > 0 \qquad \text{(Unterschiedliche Punkte haben nicht Abstand Null)}$$
$$d(x, y) = d(y, x) \qquad \text{(Symmetrie)}$$
$$d(x, y) \leq d(x, z) + d(z, y) \qquad \text{(Dreiecksungleichung)}$$

2 Funktionen in mehreren Variablen: Differentiation

Bei der zentralen Aufgabe der Physik, der Beschreibung von Naturvorgängen, werden oft Funktionen benutzt, die von mehreren Variablen x_1, ..., x_n abhängen. Dies können beliebige Größen wie die Ortskoordinaten x, y, z, die Zeit t, die elektrische Spannung V oder andere Größen sein. Im Folgenden fassen wir diese Variablen in einem Spaltenvektor $\boldsymbol{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{R}^n$ zusammen. Das Ergebnis der von uns untersuchten Funktionen kann ebenfalls mehrdimensional sein und lässt sich ebenfalls als Spaltenvektor $\boldsymbol{f}(\boldsymbol{x}) = \boldsymbol{f}(x_1, ..., x_n) = \begin{pmatrix} f_1(\boldsymbol{x}) \\ \vdots \\ f_m(\boldsymbol{x}) \end{pmatrix} \in \mathbb{R}^m$ darstellen. Wir untersuchen also Funktionen der Art:

$$\boldsymbol{f} : \mathbb{R}^n \supseteq D \to \mathbb{R}^m$$

$$\boldsymbol{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \mapsto \boldsymbol{f}(x_1, ..., x_n) = \begin{pmatrix} f_1(\boldsymbol{x}) \\ \vdots \\ f_m(\boldsymbol{x}) \end{pmatrix}$$

($\mathbb{R}^n$ heißt *Urbildraum*, D *Definitionsbereich*, $\mathbb{R}^m$ *Bildraum* und $\boldsymbol{f}(D) = \{\boldsymbol{f}(\boldsymbol{x}); \boldsymbol{x} \in D\}$ heißt *Bild*)

Im Folgenden wollen wir die durch diese Abbildungsvorschrift gegebenen Funktionen untersuchen. Dafür unterscheiden wir folgende Fälle:

Kurven - n=1, m beliebig d.h. $\boldsymbol{f} : \mathbb{R} \supseteq I \to \mathbb{R}^m$ (siehe Abschnitt 2.1)

Beispiele: Ortskuve eines Massenpunktes,...

Skalarfelder - m=1, n beliebig d.h. $f : \mathbb{R}^n \supseteq D \to \mathbb{R}$ (siehe Abschnitt 2.2)

Beispiele: Temperatur, elektrisches Potential einer Ladungsverteilung,...

Vektorfelder - n, m beliebig d.h. $\boldsymbol{f} : \mathbb{R}^n \supseteq D \to \mathbb{R}^m$ (siehe Abschnitt 5.1.2)

Beispiele: Strömungsfelder, magnetisches Feld, ...

2.1 Kurven im $\mathbb{R}^n$

2.1.1 Parameterdarstellung

Für die allgemeine Beschreibung einer Kurve eignet sich am besten die Parameterdarstellung. Diese kann man sich so vorstellen, dass die Kurve die Bahn eines Teilchen beschreibt und der Parameter t der Zeit entspricht. Damit handelt es sich bei einer Kurve um eine stetige Abbildung:

$$\boldsymbol{x} : \mathbb{R} \supseteq I \to \mathbb{R}^n$$

Definition 2.1 (Kurvenstück) *Sei $G \subseteq \mathbb{R}^n$ und $[a,b] \subseteq \mathbb{R}$ ein abgeschlossenes Intervall. Dann ist jede stetig differenzierbare Funktion $\boldsymbol{x} : [a,b] \to G$ ein* ***Kurvenstück*** *in G mit Anfangspunkt $\boldsymbol{x}(a)$ und Endpunkt $\boldsymbol{x}(b)$*

Eine Kurve besteht aus einer Aneinanderreihung von Kurvenstücken.

Definition 2.2 (Bogenlänge)

$$s(t) = \int_a^t |\dot{\boldsymbol{x}}(t)|\, dt \tag{2}$$

Die Zahl s(t) heißt Bogenlänge über [a,t]. Für n=2, 3 entspricht sie der tatsächlichen Länge des Bogens. Achtung: Hier wird auch der mehrfacher Durchlauf berücksichtigt.

Beispiel 2.1 (Parametrisierung des Einheitskreises) *Wähle Parametrisierung des Einheitskreises $\boldsymbol{x}(t) = \begin{pmatrix} cos(2\pi t) \\ sin(2\pi t) \end{pmatrix}$ und berechne die Bogenlänge für die Intervalle $[0,1]$ und $[0,2]$:*

$$s(1) = \int_0^1 \left| \begin{pmatrix} cos(2\pi t) \\ sin(2\pi t) \end{pmatrix} \right| dt = 2\pi \int_0^1 dt = 2\pi$$

$$s(2) = 4\pi$$

2.1.2 Das begleitende Dreibein

Die Bewegung eines Massenpunktes längs einer durch $\boldsymbol{x} : [a,b] \to \mathbb{R}^3$ parametrisierten Spur und die Auswirkung von Kräften und Momenten in den einzelnen Punkten der Trajektorie wird am besten durch ein begleitendes kartesisches Koordinatensystem $(\boldsymbol{x}(t); \boldsymbol{T}, \boldsymbol{N}, \boldsymbol{B})$ beschrieben. In diesem Koordinatensystem verändern sich ebenfalls die Vektoren $\boldsymbol{T}$, $\boldsymbol{N}$ und $\boldsymbol{B}$ mit dem Parameter t. Das begleitende Dreibein besteht aus folgenden Vektoren:

Definition 2.3 (Der Tangentenvektor)

$$\boldsymbol{T}(t) := \frac{1}{|\dot{\boldsymbol{x}}(t)|} \dot{\boldsymbol{x}}(t) \tag{3}$$

Definition 2.4 (Der Hauptnormalenvektor)

$$\boldsymbol{N}(t) := \frac{1}{\left|\dot{\boldsymbol{T}}(t)\right|} \dot{\boldsymbol{T}}(t) \tag{4}$$

Definition 2.5 (Der Binormalenvektor)

$$\boldsymbol{T}(t) := \boldsymbol{T}(t) \times \boldsymbol{N(t)} \tag{5}$$

Definition 2.6 (Die Krümmung)

$$\kappa(t) := \frac{1}{\dot{s}(t)} \left|\dot{\boldsymbol{T}}(t)\right| = \frac{\left|\dot{\boldsymbol{T}}(t)\right|}{|\dot{\boldsymbol{x}}(t)|} \tag{6}$$

2.2 Skalarfelder

Unter einem Skalarfeld versteht man eine Abbildung $\boldsymbol{f} : \mathbb{R}^n \supseteq D \to \mathbb{R}$. Dabei kann es sich wie oben beschrieben um die Temperatur im Raum oder das elektrische Potential einer Ladungsverteilung handeln. Skalarfelder können auf unterschiedliche Weisen gegeben sein:

durch eine **explizite** Rechenvorschrift
Bsp: $f(u, v, w) = 3uv^2 + 5w^4 + 4$ mit $D = \mathbb{R}^3$

durch eine **implizite** Gleichungen
Bsp: $z^3 - xz - y = 0$

durch komplizierte Vorschrifen, wie **Differentialgleichungen**

2.2.1 Stetigkeit

Genau wie im $\mathbb{R}^1$ will man bei Funktionen mit mehreren Veränderlichen Aussagen über Grenzwerte und Stetigkeit treffen. Hierzu muss die so genannte r-Umgebung eingeführt werden. Das ist die Menge der Punkte, deren Abstand von einem Punkt $\boldsymbol{a} \in \mathbb{R}^n$ kleiner ist als eine Zahl $r > 0$. Diese Menge wird wie folgt beschrieben:

$$U_r(\boldsymbol{a}) := \{\boldsymbol{x} \in \mathbb{R}^n; |\boldsymbol{x} - \boldsymbol{a}| < r\}$$

Für n=1 ist $U_r(a)$ das offene Intervall $a - r < x < a + r$, für n=2 ist $U_r(\boldsymbol{a})$ die Kreisscheibe um den Punkt $\boldsymbol{a}$ mit Radius r allerdings ohne Randpunkte. Für n=3 beschreibt $U_r(\boldsymbol{a})$ eine Kugel um $\boldsymbol{a}$ mit Radius r ohne Randpunkte. Dies führt zu einer genaueren Betrachtung des Randes einer Menge.

Definition 2.7 (offene und abgeschlossene Mengen) *Sei D eine Teilmenge des* $\mathbb{R}^n$

a) *Ein Punkt $\boldsymbol{a} \in \mathbb{R}^n$ heißt **innerer Punkt** von D, wenn es eine r-Umgebung von $\boldsymbol{a}$ gibt, die ganz in D enthalten ist.*

b) *D ist **offen**, wenn jeder Punkt von D ein innerer Punkt ist.*

c) *Ein Punkt $\boldsymbol{b} \in \mathbb{R}^n$ heißt Randpunkt von D, wenn jede r-Umgebung von $\boldsymbol{b}$ sowohl mindestens einen Punkt aus D als auch mindestens einen nicht zu D gehördenden Punkt enthält. Die Menge aller Randpunkte von D heißt **Rand von D** und wird mit ∂D bezeichnet.*

d) *Eine Menge heißt **abgeschlossen**, wenn sie alle ihre Randpunkte enthält.*

Beispiel 2.2 (Kreisscheibe)

1) *Die „offene Kreisscheibe (ohne Rand)“*

$$K = \{(x, y); (x - x_0)^2 + (y - y_0)^2 < r^2$$

ist eine offene Menge in der der Rand ∂K: $(x - x_0)^2 + (y - y_0)^2 = r^2$ nicht enthalten ist.

2) *Die Kugel (mit Rand)*

$$K = \{(x, y, z); (x - x_0)^2 + (y - y_0)^2 + (z - z_0)^2 \leq r^2$$

ist eine Menge die den Rand mit enthält.

Definition 2.8 (beschränkte und kompakte Menge)

a) *Eine Menge $D \in \mathbb{R}^n$ heißt **beschränkt**, wenn es eine Konstante $K > 0$ gibt mit*

$$|\boldsymbol{x}| < K \quad \forall \boldsymbol{x} \in D.$$

b) *Die abgeschlossenen und beschränkten Mengen des $\mathbb{R}^n$ nennt man **kompakt**.*

Definition 2.9 (Grenzwert und Stetigkeit) *Sei $f : \mathbb{R}^n \supseteq D \to \mathbb{R}$ und $\boldsymbol{a} \in D \cup \partial D$.*

a) f hat in $\boldsymbol{a}$ den ***Grenzwert*** *$c \in \mathbb{R}$, in Zeichen*

$$\lim_{\boldsymbol{x} \to \boldsymbol{a}} f(\boldsymbol{x}) = c$$

oder $f(x) \to c$ für $\boldsymbol{x} \to \boldsymbol{a}$, wenn es zu jeder (beliebig kleinen Schranke $\epsilon > 0$ eine r-Umgebung $U_r(\boldsymbol{a})$ gibt, so dass $|f(\boldsymbol{x}) - c| < \epsilon$ für alle $\boldsymbol{x} \in D \cap U_r(\boldsymbol{a})$ gilt.

b) f heißt in $\boldsymbol{a} \in D$ ***stetig****, wenn $\lim_{\boldsymbol{x} \to \boldsymbol{a}} f(\boldsymbol{x}) = f(\boldsymbol{a})$ gilt.*

c) f heißt auf D ***stetig****, wenn f in allen $\boldsymbol{a} \in D$* ***stetig*** *ist.*

Des Weiteren gilt, dass Summen, Produkte, Quotienten stetiger Funktionen stetig sind. Allerdings ergibt sich die Stetigkeit von $f(x, y)$ in (x_0, y_0) noch nicht aus der Stetigkeit von $f(x, y_0)$ und $f(x_0, y)$.

Beispiel 2.3 (Parabelfalte) *Man untersuche*

$$f(x, y) = \begin{cases} \frac{xy^2}{x^2+y^4} & , \text{ für } (x, y) \neq (0, 0) \\ 0 & , \text{ für } (x, y) = (0, 0) \end{cases}$$

Es sind sowohl $f(x, 0) = 0$ ($x \in \mathbb{R}$) als auch $f(0, y) = 0$ ($y \in \mathbb{R}$) in x=0 bzw. y=0 stetig. Jedoch ist f (als Funktion von zwei Variablen) in $(0, 0)$ nicht stetig. Bei Annäherung an $(0, 0)$ längs der Geraden (tu, tv) mit $u \neq 0$ erhält man:

$$\lim_{t \to 0} f(tu, tv) = \lim_{t \to 0} \frac{t^3uv^2}{t^2(u^2 + t^2v^4)} = \lim_{t \to 0} \frac{tuv^2}{u^2 + t^2v^4} \leq \lim_{t \to 0} \frac{uv^2}{tv^4)} = 0$$

Bei Annäherung längs der (t^2, t) erhält man hingegen:

$$\lim_{t \to 0} f(t^2, t) = \lim_{t \to 0} \frac{t^4}{2t^4} = \frac{1}{2}$$

Da die beiden Grenzwerte nicht gleich sind, ist die Funktion f in $(0, 0)$ nicht stetig.

2.2.2 Partielle Differentiation und Gradient

Die partielle Differentiation oder partielle Ableitung einer Funktion beschreibt die Ableitung nach einer Variablen. Alle übrigen Variablen werden als Konstanten angenommen. Es handelt sich also anschaulich um eine Richtungsableitung entlang der Koordinatenachse. Die partiellen Differentiation einer Funktion kann wie folgt geschrieben werden:

Definition 2.10 (Partielle Differentiation)

$$f_{x_i}(\boldsymbol{x}) = \frac{\partial f(\boldsymbol{x})}{\partial x_i} := \lim_{t \to 0} \frac{f(x_1, \ldots, x_i + t, \ldots, x_n) - f(x_1, \ldots, x_i, \ldots, x_n))}{t}$$

Beispiel 2.4 ()

$$f(x,y,z) = 3xy^2 + e^{2yz^2}$$

$$f_y(x,y,z) = 6xy + 2z^2 e^{2yz^2} \qquad f_x(x,y,z) = 3y^2$$

$$f_{yy}(x,y,z) = 6x + 4z^4 e^{2yz^2} \qquad f_{xx}(x,y,z) = 0$$

Eine Funktion f heißt k-mal partiell differenzierbar, wenn alle ihre k-ten partiellen Ableitungen

$$f\underbrace{x_i, \ldots, x_j}_{k-mal} = \underbrace{\frac{\partial}{\partial x_i} \cdots \frac{\partial f}{\partial x_j}}_{k-mal}$$

existieren. Sind alle Ableitungen stetig, dann heißt f k-mal stetig partiell differenzierbar. Wir bezeichnen die Menge aller auf einer offenen Menge $D \subseteq \mathbb{R}^n$ k-mal stetig differenzierbaren Funktionen $f : D \to \mathbb{R}$ mit $\mathcal{C}^k(D, \mathbb{R})$ und nennen $f \in \mathcal{C}^k(D, \mathbb{R})$ eine $\mathcal{C}^k$-Funktion.

$$\boxed{\begin{aligned} \mathcal{C}^k(D,\mathbb{R}) &:= \{f : D \to \mathbb{R};\ f \text{ k-mal stetig partiell differenzierbar}\} \\ \mathcal{C}^0(D,\mathbb{R}) &:= \{f : D \to \mathbb{R};\ f \text{ ist stetig}\} \end{aligned}}$$

Satz 2.1 (Satz von Schwarz - Vertauschbarkeit der partiellen Ableitung)
Für jede $\mathcal{C}^2$-Funktion $f : D \to \mathbb{R}, D \subseteq \mathbb{R}^n$ offen, gilt:

$$\frac{\partial}{\partial x_i}\left(\frac{\partial f}{\partial x_j}\right) = \frac{\partial}{\partial x_j}\left(\frac{\partial}{\partial x_i}\right) \tag{7}$$

mit $1 \leq i, j \leq n$

Beispiel 2.5 (Vertauschbarkeit der partiellen Ableitung)

1)

$$f(x,y) = x^2y^3 + y\ln x(x>0)$$

$$f_x(x,y) = 2y^3x + \frac{y}{x}\ , \qquad f_y(x,y) = 3x^2y^2 + \ln x$$

$$f_{xy}(x,y) = 6y^2x + \frac{1}{x}\ , \qquad f_{yx}(x,y) = 6y^2x + \frac{1}{x}$$

2) Gegeben sei die Funktion

$$f(x,y) = \begin{cases} xy\frac{x^2-y^2}{x^2+y^2}\,, & (x,y) \neq (0,0) \\ 0\,, & (x,y) = (0,0) \end{cases}$$

Da für alle $x, y \in \mathbb{R}$ *stets* $f(x,0) = f(0,y) = 0$*, gilt* $f_x(0,0) = f_y(0,0) = 0$*. Für* $(x,y) \neq (0,0)$ *erhält man durch Differentitation*

$$f_x = y\left(\frac{x^2-y^2}{x^2+y^2} + \frac{4x^2y^2}{(x^2+y^2)^2}\right)$$

$$f_y = x\left(\frac{x^2-y^2}{x^2+y^2} - \frac{4x^2y^2}{(x^2+y^2)^2}\right)$$

Damit folgt für $f_{xy} = \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right)$

$$f_{xy}(0,0) = \lim_{y\to 0}\frac{f_x(x,0)-f_x(0,0)}{y} = \lim_{y\to 0}\frac{-y^2}{y^2} = -1$$

und andererseits

$$f_{yx}(0,0) = \lim_{x\to 0}\frac{f_y(x,0)-f_y(0,0)}{x} = \lim_{x\to 0}\frac{x^2}{x^2} = 1$$

Fasst man alle partiellen Ableitungen einer Funktion zu einem Vektor zusammen, so erhält man den **Gradienten** von f an der Stelle $\boldsymbol{x}$:

Definition 2.11 (Gradient)

$$grad\ f(\boldsymbol{x}) := \begin{pmatrix} \frac{\partial f}{\partial x_1}(\boldsymbol{x}) \\ \frac{\partial f}{\partial x_2}(\boldsymbol{x}) \\ \vdots \\ \frac{\partial f}{\partial x_n}(\boldsymbol{x}) \end{pmatrix} \in \mathbb{R}^n$$

$\boldsymbol{x} \mapsto grad\, f(\boldsymbol{x})$ ist ein Vektorfeld (Siehe Abschnitt 5.1.2).

Beispiel 2.6 ()

$$f(x,y,z) = 3xy^2 + 4x^2z^2 + e^{2xz}$$

$$grad\ f(x,y,z) := \begin{pmatrix} 3y^2 + 8xz^2 + 2ze^{2xz} \\ 6xy \\ 8x^2z + 2xe^{2xz} \end{pmatrix}$$

2.2.3 Die Kettenregel

Hängt der Vektor $\boldsymbol{x}$ von einem kontinuierlichem Parameter t ab, also $\boldsymbol{x} = \boldsymbol{x}(t)$, so beschreibt dies eine Kurve im $\mathbb{R}^n$. Die Änderung der Funktion f entlang dieser Kurve lässt sich wie folgt bestimmen:

Satz 2.2 (Die Kettenregel) *Für jede $\mathcal{C}^1$-Funktion $f : D \to \mathbb{R}$, $D \subseteq \mathbb{R}^n$ offen, und für jedes Kurvenstück $\boldsymbol{x} : \mathbb{R} \supseteq [a,b] \to D$ gilt*

$$\begin{aligned} \frac{d}{dt} f(\boldsymbol{x}(t)) &= \frac{d}{dt} f(x_1(t), \ldots, x_n(t)) \\ &= f_{x_1}(\boldsymbol{x}(t))\dot{x_1}(t) + \ldots + f_{x_n}(\boldsymbol{x}(t))\dot{x_n}(t) \\ &= \nabla f(\boldsymbol{x}(t)) \cdot \dot{\boldsymbol{x}}(t) \end{aligned} \tag{8}$$

Mit der Kettenregel kann man zeigen, dass der Gradient immer senkrecht auf Niveauflächen steht. Die Niveauflächen einer Funktion f sind definiert durch:

$$N_c := \{\mathbf{x} \in D \subseteq \mathbb{R}^n; f(\mathbf{x}) = c\}\ ;\ c \in \mathbb{R}$$

Für **alle** Kurven $\mathbf{x} = \mathbf{x}(t)$, die in der Niveaufläche N_c liegen gilt:

$$\frac{d}{dt} f(\mathbf{x}(t)) = \frac{d}{dt} c = 0 = \nabla f \cdot \dot{\mathbf{x}}(t)$$

Das Skalarprodukt zwischen dem Gradienten und der Kurventangente ist immer Null. Das bedeutet der Gradient von f ist orthogonal zu allen Kurventangenten und somit orthogonal zur Niveaufläche N_c

2.2.4 Die Richtungsableitung

Wie in Abschnitt 2.2.2 dargestellt, beschreibt die partielle Differentiation einer Funktion $\frac{\partial f}{\partial x_i}(\boldsymbol{x})$ die Ableitung der Funktion f in Richtung der durch x_i beschriebenen Koordinatenachse. Es gibt keinen Grund nur die Ableitungen entlang der Koordinatenachsen zu berücksichtigen. Deshalb bezeichnet man den Grenzwert:

$$\partial_{\boldsymbol{v}} f(\boldsymbol{x}) := \lim_{t \to 0} \frac{1}{t} \left[f(\boldsymbol{x} + t\boldsymbol{v}) - f(\boldsymbol{x}) \right]$$

als Ableitung von f an der Stelle $\boldsymbol{x}$ längs dem beliebigen Vektor $\boldsymbol{v}$. Ist $\boldsymbol{v}$ ein Einheitsvektor ($|\boldsymbol{v}| = 1$), dann heißt $\partial_{\boldsymbol{v}} f(\boldsymbol{x})$ die Richtungsableitung von f an der Stelle $\boldsymbol{x}$ in Richtung $\boldsymbol{v}$.

Satz 2.3 (Richtungsableitung) *Für jede auf der offenen Menge $D \subseteq \mathbb{R}^n$ total differenzierbaren Funktion f (insbesondere für $f \in \mathcal{C}^1(D, \mathbb{R})$ und für jeden Vektor $\boldsymbol{v} \in \mathbb{R}^n$, $\boldsymbol{v} \neq 0$, gilt:*

$$\boxed{\partial_{\boldsymbol{v}} f(\boldsymbol{x}) = \nabla f(\boldsymbol{x}) \cdot \boldsymbol{v} = \sum_{i=1}^{n} f_{x_i}(\boldsymbol{x}) v_i} \tag{9}$$

Mit $|\boldsymbol{v}| = 1$ stellt (9) die Richtungsableitung (bzw. den Anstieg) von f an der Stelle $\boldsymbol{x}$ in Richtung $\boldsymbol{v}$ dar.

2.2.5 Totale Ableitung

Oft ist es notwendig die Funktion $f(\boldsymbol{x})$ an einem Punkt $P = (x_0, y_0)$ zu approximieren. Im Fall $\mathbb{R}^2$, d.h. $z = f(x, y)$ kann man sich diese Approximation als Tangentialebene an den Punkt P vorstellen. Im Folgenden werden wir das totale Differential für diesen Fall herleiten. Es lässt sich jedoch ohne weiteres auf den $\mathbb{R}^n$ erweitern.
Eine Tangentialeben $t(x, y)$ hat die Form:

$$z = ax + by + c = t(x, y)$$

Da die Ebene durch den Punkt P gehen soll und außerdem die Steigung an diesem Punkt approximieren soll, muss gelten:

$$z_0 = ax_0 + by_0 + c; \quad f_x(x_0, y_0) = a; \quad f_y(x_0, y_0) = b$$

und man erhält

$$(z - z_0) = f_x(x_0, y_0)\,(x - x_0) + f_y(x_0, y_0)\,(y - y_0)$$

Geht man nun in die Infinitesimalschreibweise über ($(x - x_0) \to \Delta x \to dx$) erhält man:

Satz 2.4 (Totales Differential) *Die Funktion $f(\boldsymbol{x})$ sei in $\boldsymbol{x}_0$ nach allen Variablen x_i differenzierbar. Dann nennt man*

$$\boxed{dz = f_{x_1}(\boldsymbol{x}_0)\,dx_1 + \ldots + f_{x_n}(\boldsymbol{x}_0)\,dx_n} \tag{10}$$

totales Differential der Funktion $f(\boldsymbol{x})$ in $\boldsymbol{x}_0$ und $f(\boldsymbol{x})$ ist in $\boldsymbol{x}_0$ total differenzierbar.

Mit $z_0 = f(\mathbf{x}_0)$ und $z \approx f(\mathbf{x})$ erhält man außerdem die lineare Approximation von $f(\mathbf{x})$ an der Stelle $\mathbf{x}_0$

$$\boxed{f(\mathbf{x}) \approx f(\mathbf{x}_0) + \nabla f(\mathbf{x}_0) \cdot (\mathbf{x} - \mathbf{x}_0)} \tag{11}$$

2.2.6 Taylorentwicklung von Funktionen in mehreren Veränderlichen

Im Folgenden soll die schon bekannte Taylorentwicklung von Funktionen mit einer Variablen erweitert werden, auf Funktionen mit mehreren Veränderlichen. Hierzu betrachten wir eine Funktion einer reellen Variablen $h(t) := f(\boldsymbol{x}_0 + t\boldsymbol{v})$ mit $\boldsymbol{x}_0 \in D \subseteq \mathbb{R}^n$, $\boldsymbol{v} \in \mathbb{R}^n$