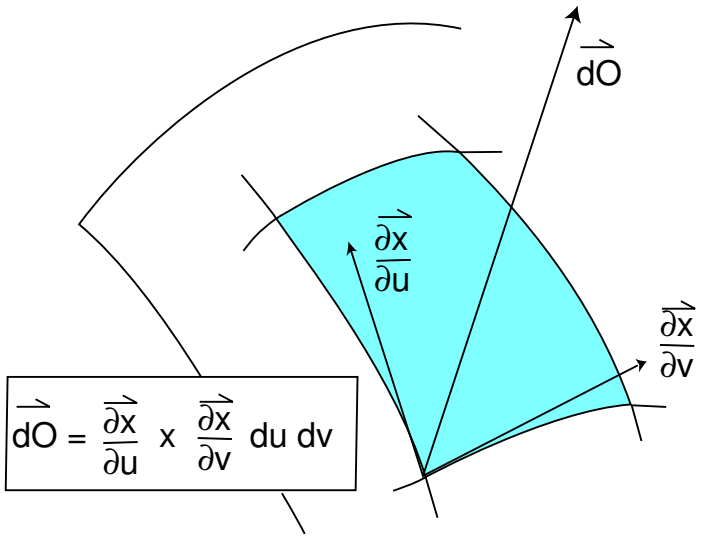
Technische Universität Berlin
Fakultät ii • Institut für Mathematik
Prof. Dr. Dirk Ferus

Analysis II für Ingenieure

(Mit Beispielmaterial aus der Elektrotechnik)



Information. Für die erfolgreiche Teilnahme an diesem Modul erhalten Sie

8 Leistungspunkte nach ECTS.

Entsprechend erwarten wir von durchschnittlich begabten und vorgebildeten Studierenden folgenden Arbeitsaufwand:

Vorlesung	4h/Woche
Übung	2h/Woche
Häusliche Nacharbeit und Hausaufgaben	8h/Woche
Klausurvorbereitung	30h

Version vom 16.05.2007

Inhaltsverzeichnis

Literatur

Als Lehrbücher zu dieser Veranstaltung werden empfohlen:

- **G. Bärwolff, G. Seifert: Höhere Mathematik für Naturwissenschaftler und Ingenieure, Spektrum Akademischer Verlag**
- **Meyberg, Vachenauer: Höhere Mathematik 1, Springer Verlag**

Farbig unterlegt finden Sie Beispiele aus den Ingenieuranwendungen, oft mit expliziten Hinweisen auf Ingenieurskripten des Grundstudiums:

Müller: Mechanik II
W.H. Müller: Mechanik II,
Skript TUB SS 2002

Müller: Mechanik III
W.H. Müller: Mechanik III,
Skript TUB, WS 2002/3

Popov: Mechanik III
V. Popov: Mechanik III,
Skript TUB WS 2002/3

Grundlagen der Elektrotechnik, Teil A
Naunin: Grundlagen der Elektrotechnik,
LaborSkript TUB

Theoretische Elektrotechnik
Henke: Theoretische Elektrotechnik,
Skript TUB 1990/91

Werkstoffe II
Werkstoffe und Bauelemente der Elektrotechnik II,
Skript TUB, Institut für Werkstoffe der Elektrotechnik

Energie-, Impuls- und Stofftransport
Lehrbuch *Baehr/Stephan: Wärme- und Stoffübertragung*
zur VL Auracher: Energie-, Impuls- und Stofftransport

Verfahrentechnik I
Kraume: Verfahrenstechnik I,
Skript TUB

Zu einzelnen Veranstaltungen (wie der Mechanik) existieren mehrere und vielleicht auch neuere Skripten und Lehrbücher. Wir begnügen uns meistens mit *einem* Zitat, auch wenn sich der betreffende Sachverhalt in der Regel an mehreren Stellen findet.

Im Wintersemester beginnt die Veranstaltung mit einem Abschnitt über Unendliche Reihen (vgl. Anhang 0).

1 Mehrdimensionale Differentialrechnung

1.1 Vorbemerkungen

Funktionen, die physikalisch-technische Größen beschreiben, hängen in der Regel von sehr vielen Parametern ab: Die Temperatur im Innern eines Autos hängt ab von der Sonneneinstrahlung, von der Temperatur des Motors, von der Außentemperatur, von der Dicke und spezifischen Wärmeleitfähigkeit der Isolierung, von der Wärmekonvektion durch Lüftung und vielen anderen Parametern. Der Spannungsabfall einer Wechselstromleitung hängt ab von der Länge, dem spezifischen Materialwiderstand, der Temperatur, der Frequenz, dem Widerstand des umgebenden Mediums, ... Richtet man das Augenmerk nur auf die Veränderung der Funktion bei Variation eines Parameters, so hat man eine Funktion einer Variablen. Häufig ändern sich aber gleichzeitig mehrere Parameter, und man muss also Funktionen mehrerer Variablen untersuchen. Insbesondere möchte man die Differential- und Integralrechnung auf solche Funktionen ausdehnen. Als typisches Beispiel können Sie ortsabhängige Funktionen betrachten, da haben Sie schon 3 Variable für die Raumkoordinaten. Aber Funktionen, die von den Raumkoordinaten *zweier* Punkte abhängen, haben schon 6 Variable.

Oft hängen die Funktionen nicht nur von mehreren Variablen ab, sondern ihre Werte sind auch mehrdimensional: sie haben mehrere Komponenten. Je nach geometrischer Interpretation spricht man dann von Abbildungen, Transformationen oder Vektorfeldern.

- Die Bahn eines Massenpunktes ist eine Abbildung der Zeitachse in den 3-dimensionalen Raum (1 Variable, 3 Komponenten).
- Die Abbildung $(x, y) \mapsto (\sqrt{x^2 + y^2}, \arctan \frac{y}{x})$ nennt man eine Transformation (auf Polarkoordinaten, 2 Variable, 2 Komponenten).
- Das (zeitinvariante) Strömungsfeld einer Flüssigkeit oder das elektrische Feld eines Dipols sind Beispiele für Vektorfelder (3 Variable, 3 Komponenten).

1.2 Topologie im $\mathbb{R}^n$

- Der n-dimensionale Raum
- Beschreibung von Teilmengen darin
- Offene und abgeschlossene Mengen

Der *n-dimensionale Euklidische Raum* $\mathbb{R}^n$ besteht aus den n-tupeln reeller Zahlen

$$\vec{x} = (x_1, \ldots, x_n),$$

die wir je nach Zusammenhang auch als „Punkte“ des $\mathbb{R}^n$ oder als „Vektoren“ im $\mathbb{R}^n$ bezeichnen.

Bemerkung. Neben Vektoren des $\mathbb{R}^n$ kommen in der Analysis wie in der linearen Algebra häufig auch Matrizen vor. Aus rechentechnisch-formalen Gründen ist es dann vorzuziehen, Vektoren als *Spalten*

$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

zu schreiben. Im geschriebenen Text und besonders in der Analysis, wo es um Ausdrücke der Form $f(\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix})$ geht, ist diese Notation unhandlich. Manchmal behilft man sich dann mit dem Symbol $(\ldots)^T$. Dabei steht der Buchstabe „T“ für *Transposition* = Vertauschung von Zeilen und Spalten:

$$(x_1, \ldots, x_n)^T = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}.$$

Es ist allerdings durchaus gebräuchlich, einfach $(x_1, \ldots, x_n)$ statt $(x_1, \ldots, x_n)^T$ zu schreiben. Wir werden das auch tun, wenn die gerade betrachtete Situation nicht mehr Sorgfalt erfordert. Letzteres ist zum Beispiel der Fall, wenn auch lineare Abbildungen und Matrizen ins Spiel kommen.

Der $\mathbb{R}^2$ ist die Ebene, der $\mathbb{R}^3$ der „gewöhnliche“ 3-dimensionale Raum. Ihre Elemente bezeichnen wir oft mit (x, y) bzw. (x, y, z).

Für Vektoren im $\mathbb{R}^n$ haben wir ein Skalarprodukt und eine Länge definiert:

$$\vec{x} \cdot \vec{y} = \sum_{k=1}^{n} x_k y_k = x_1 y_1 + \ldots + x_n y_n,$$
$$|\vec{x}\,| = \sqrt{\vec{x} \cdot \vec{x}} = \sqrt{x_1^2 + \ldots + x_n^2}.$$

Dafür gelten folgende Fomeln:

$$|c\vec{x}\,| = |c|\,|\vec{x}\,|$$
$$|\vec{x} + \vec{y}\,| \leq |\vec{x}\,| + |\vec{y}\,| \qquad \text{(Dreiecksungleichung)}$$
$$|\vec{x} \cdot \vec{y}\,| \leq |\vec{x}\,||\vec{y}\,| \qquad \text{(Cauchy-Schwarzsche Ungleichung)}$$

Den Abstand von zwei Punkten im $\mathbb{R}^n$ definieren wir als

$$|\vec{x} - \vec{y}\,| = \sqrt{\sum_{k=1}^{n} (x_k - y_k)^2}.$$

In der Ebene wird diese Definition durch den Satz des Pythagoras motiviert. Aus der vorstehenden Dreiecksungleichung für Vektoren ergibt sich die Dreiecksungleichung für ein Dreieck mit den Eckpunkten $\vec{x}, \vec{y}, \vec{z}$:

$$|\vec{x} - \vec{z}| \leq |\vec{x} - \vec{y}| + |\vec{y} - \vec{z}|.$$

Wir betrachten nun Beispiele von Teilmengen des $\mathbb{R}^n$, die mit diesen Begriffen definiert werden:

Beispiel 1. Durch

$$\{\vec{x} \in \mathbb{R}^3 \mid |\vec{x} - \vec{a}| < r\}$$

wird das Innere einer Kugel vom Radius r um den Punkt $\vec{a}$ beschrieben. Man nennt die entsprechende Menge auch im $\mathbb{R}^n$ eine (offene) Kugel. Im Zweidimensionalen ist eine Kugel im Sinne dieser Definition also ein Kreis.

Die Menge

$$\{\vec{x} \in \mathbb{R}^3 \mid |\vec{x} - \vec{a}| = r\}$$

ist die Oberfläche dieser Kugel, auch Sphäre genannt. □

Beispiel 2. Seien $\vec{a}, \vec{b} \in \mathbb{R}^n$ mit $a_k < b_k$ für alle k. Dann heißt

$$\{\vec{x} \in \mathbb{R}^n \mid a_k \leq x_k \leq b_k\}$$

ein n-dimensionales *Quader*. Für $n = 2$ ist das ein Rechteck mit den diagonal gegenüberliegenden Ecken $\vec{a}, \vec{b}$. □

Beispiel 3.

$$\{(x, y, z) \in \mathbb{R}^3 \mid z > 0\}$$

ist der obere Halbraum des $\mathbb{R}^3$. □

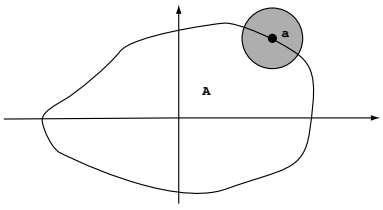
Beispiel 4. Oft beschreibt man eine Teilmenge des $\mathbb{R}^n$, indem man einfach die definierende Gleichung oder Ungleichung angibt, die bei den obigen Beispielen im zweiten Teil der Mengenklammer steht. Man sagt etwa, dass

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} \leq 1, \quad a, b, c > 0$$

ein Ellipsoid (Inneres und Oberfläche) mit den Halbachsen a, b, c „ist“. □

Eine Kugel $\{\vec{x} \in \mathbb{R}^n \mid |\vec{x} - \vec{a}| < r\}$ mit $\vec{a} \in \mathbb{R}^n$ als Mittelpunkt ist eine *Umgebung* des Punktes $\vec{a}$. Allgemeiner nennt man jede Menge $U \subset \mathbb{R}^n$, die eine solche Kugel mit Mittelpunkt $\vec{a}$ enthält auch eine *Umgebung* von $\vec{a}$.

Sei nun $A \subset \mathbb{R}^n$. Man nennt $\vec{a}$ einen *Randpunkt* von A, wenn in *jeder* - noch so kleinen - Kugel um $\vec{a}$ sowohl Punkte von A liegen, wie auch Punkte, die nicht in A (sondern in $\mathbb{R}^n \setminus A$) liegen, wenn also gilt:



$$\text{Für jede Kugel } U \text{ von } \vec{a} \text{ ist } U \cap A \neq \emptyset \text{ und } U \cap (\mathbb{R}^n \setminus A) \neq \emptyset. \tag{1}$$

Ein Randpunkt $\vec{a}$ selber kann zu A gehören oder auch nicht.

Definition 5. Wenn alle Randpunkte von A zu A gehören, heißt A *abgeschlossen*. Wenn keiner dazugehört, heißt A *offen*. Gehören manche dazu, andere nicht, so ist A weder offen noch abgeschlossen.

Denken Sie an offene und abgeschlossene Intervalle: $[0,1]$ enthält alle seine Randpunkte, nämlich 0 und 1, während $]0,1[$ keinen der beiden enthält. Das erste Intervall ist nach alter wie neuer Terminologie abgeschlossen, das zweite offen. Halboffene Intervalle wie $]0,1]$ liefern Beispiele von Mengen, die weder offen noch abgeschlossen sind. Aber: Die Intervalle $[0,\infty[$ und $]-\infty,\infty[$ sind abgeschlossene Mengen, das zweite ist auch offen!

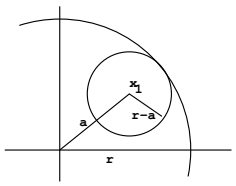
Kriterium für offene Mengen. Eine Menge ist offen, wenn sie keinen ihrer Randpunkte enthält, d.h. wenn keiner ihrer Punkte ein Randpunkt ist. Dann gibt es also um jeden ihrer Punkte eine Kugel U, die nicht die Bedingung (1) erfüllt. Weil der Punkt selber aber in A ist, müssen dann alle Punkte von U in A sein, die Kugel U liegt ganz in A. Die Umkehrung ist auch richtig: Wenn ein Punkt von A Mittelpunkt einer Kugel ist, die ganz in A liegt, ist er natürlich kein Randpunkt. Wir haben also:

Satz 6 (Offenheitskriterium). *Eine Menge A ist offen, wenn es um jeden ihrer Punkte eine Kugel gibt, die ganz in A liegt. Punkte mit dieser Eigenschaft heißen* innere Punkte *von A. Also ist A genau dann offen, wenn es nur innere Punkte enthält.*

Wir werden bei der Betrachtung von Funktionen häufig voraussetzen, dass ihr Definitionsbereich offen ist, und zwar aus folgendem Grund: Will man untersuchen, was mit den Werten der Funktion passiert, wenn man ihr Argument etwas verändert, so bleibt das Argument wenigstens bei kleinen Änderungen immer im Definitionsbereich der Funktion.

Beispiel 7. Die offene Kugel $x^2+y^2+z^2<r^2$ mit $r>0$ ist wirklich offen. Ist nämlich $\vec{x}_1$ ein Punkt daraus, so ist sein Abstand a vom Mittelpunkt der Kugel, also von 0, echt kleiner als r.

Die (offene) Kugel vom Radius $\frac{1}{2}(r-a)$ um den Punkt $\vec{x}_1$ liegt dann ganz in der ursprünglichen Kugel. Auf den Faktor $\frac{1}{2}$ kann man sogar verzichten, wenn man die offene kleine Kugel wählt. Das ist anschaulich völlig klar. (Oder?) Wir rechnen es aber zur Übung noch einmal formal nach: Sei K die gegebene Kugel. Nach Definition ist dann für $\vec{x}_1 \in K$



$$a := |\vec{x}_1 - \vec{0}| < r.$$

Ist nun $\vec{x}_2$ ein Punkt der offenen Kugel vom Radius $r-a$ um $\vec{x}_1$, so gilt $|\vec{x}_2 - \vec{x}_1| < r-a$. Aus der Dreiecksungleichung folgt

$$|\vec{x}_2 - \vec{0}| \le |\vec{x}_2 - \vec{x}_1| + |\vec{x}_1 - \vec{0}| < r - a + a = r.$$

□

Beispiel 8. Die Teilmenge $A = \mathbb{R}^n$ von $\mathbb{R}^n$ hat eine sonderbare Eigenschaft. Keiner ihrer Randpunkte gehört zu der Menge - weil sie keine Randpunkte hat. Andrerseits gehören alle ihre Randpunkte zu der Menge - weil sie keine Randpunkte hat. Der $\mathbb{R}^n$ ist also gleichzeitig offen und abgeschlossen! Dasselbe Schicksal teilt die leere Menge, und zwar aus denselben Gründen. Das sind aber auch die einzigen Mengen im $\mathbb{R}^n$ mit dieser „Pathologie“. □

1.3 Konvergenz im $\mathbb{R}^n$

- Wir verallgemeinern den Begriff der Konvergenz von Folgen reeller Zahlen auf Folgen von Vektoren im $\mathbb{R}^n$.

Konvergenz

Wenn wir im mehrdimensionalen Raum differenzieren und integrieren wollen, müssen wir in diesem Raum Grenzwerte bilden. Wir müssen also zum Beispiel erklären, was es heissen soll, dass eine Folge von Punkten im $\mathbb{R}^n$ konvergiert. Eine Folge im $\mathbb{R}^n$ sieht so aus: $(\vec{x}_k)_{k\in\mathbb{N}}$. Für jedes $k \in \mathbb{N}$ ist $\vec{x}_k$ also ein Punkt im $\mathbb{R}^n$.

Wenn wir von so einem Punkt die Komponenten brauchen, wird die Sache kompliziert, wir benötigen einen zweiten Index:

$$\vec{x}_k = (x_{k1}, \ldots, x_{kn}).$$

Der erste Index soll also der Folgenindex sein, der zweite die Komponente bezeichnen.

Für die folgende Definition kommen wir aber ohne Komponenten aus:

Definition 9. Eine Folge $(\vec{x}_k)_{k\in\mathbb{N}}$ im $\mathbb{R}^n$ heißt *konvergent gegen* $\vec{a} \in \mathbb{R}^n$, wenn

$$\lim_{k\to\infty} |\vec{x}_k - \vec{a}\,| = 0.$$

Beachten Sie, dass $(|\vec{x}_k - \vec{a}\,|)_{k\in\mathbb{N}}$ einfach eine Folge reeller Zahlen ist, und dafür wissen wir, was Konvergenz ist. Man schreibt dann auch

$$\vec{a} = \lim_{k\to\infty} \vec{x}_k.$$

Anschauliche Interpretation

Die Folge $(\vec{x}_k)$ konvergiert genau dann gegen $\vec{a}$, wenn folgendes gilt: In jeder (noch so kleinen) Kugel um $\vec{a}$ liegt der ganze Rest der Folge von einem bestimmten Index k_0 an. Formal heißt das: Wenn man eine beliebige Zahl $r > 0$ vorgibt, so gibt es dazu ein k_0, so dass

$$|\vec{x}_k - \vec{a}| < r \text{ für alle } k \geq k_0.$$

Beispiel 10. Sei $n = 2$ und $\vec{x}_k = (\frac{1}{k}\cos k\theta, \frac{1}{k}\sin k\theta)$ für ein festes θ. Dann ist

$$|\vec{x}_k - \vec{0}| = \sqrt{(\frac{1}{k}\cos k\theta)^2 + (\frac{1}{k}\sin k\theta)^2} = \frac{1}{k} \to 0 \text{ für } k \to \infty.$$

Also ist $\lim_{k\to\infty} \vec{x}_k = \vec{0}$. □

Komponentenweise Konvergenz

Aus

$$|\vec{x}_k - \vec{a}\,| = \sqrt{\sum_{m=1}^{n} (x_{km} - a_m)^2}$$

sieht man sofort, dass

$$\vec{a} = \lim_{k\to\infty} \vec{x}_k \iff \lim_{k\to\infty} x_{km} = a_m \text{ für alle } m.$$

Satz 11 (Komponentenweise Konvergenz). *Eine Folge von Vektoren konvergiert genau dann, wenn alle Komponentenfolgen konvergieren.*

Beispiel 12. Für die Folge $\vec{x}_k = ((-1)^k, \frac{1}{k})$ ist die Folge der ersten Komponenten, also $((-1)^k)_{k\geq 1}$, nicht konvergent. Daher ist die ganze Folge nicht konvergent. □

Im Fall $n = 2$ kann man statt Vektoren auch einfach **komplexe Zahlen** verwenden: Der Vektor (x, y) in der Ebene entspricht $z = x + iy$. So entspricht $\vec{x}_k = (\frac{1}{k}\cos k\theta, \frac{1}{k}\sin k\theta)$ aus dem obigen Beispiel die Folge der $z_k = \frac{1}{k}(\cos k\theta + i \sin k\theta) = \frac{1}{k}e^{ik\theta}$.

Beispiel 13. Wir betrachten für eine komplexe Zahl $c = a+ib$ und ein beliebig vorgegebenes $z_0 = x_0 + iy_0 \neq 0$ die *rekursiv definierte* Folge

$$z_{k+1} = \frac{1}{2}\left(z_k + \frac{c}{z_k}\right).$$

Wenn $c = a > 0$ und $z_0 = x_0 > 0$ reell sind, bleibt die ganze Folge reell, und wir haben früher gesehen, dass sie gegen $\sqrt{a}$ konvergiert. Die komplexe Folge z_k kann man natürlich auch als Folge im $\mathbb{R}^2$ schreiben, aber wir verzichten hier darauf. Die Folge (z_k) konvergiert gegen diejenige der beiden komplexen Wurzeln aus $a + ib$, die näher an z_0 liegt. Für z_0 auf der Mittelsenkrechten zwischen den beiden Werten divergiert die Folge, oder sie bricht ab. □

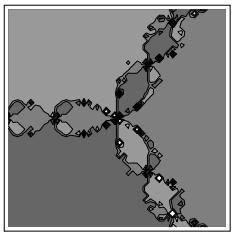
Beispiel 14. Die Folge

$$z_{k+1} = f(z_k) = \frac{2}{3}\left(z_k + \frac{c}{2z_k^2}\right)$$

konvergiert zwar für positive reelle c und z_0 gegen $\sqrt[3]{c}$. Aber es gibt drei komplexe Wurzeln aus c, und die komplexe Ebene besitzt entsprechend drei Bereiche, deren Punkte unter dem Rekursionsverfahren jeweils gegen eine andere dieser drei Wurzeln konvergieren. Anders als im Beispiel 13 bestehen aber diese Bereiche jeweils aus verschiedenen Komponenten, deren Grenzen überaus komplizierte Mengen (Juliamengen) sind. In dem folgenden Mathematica-Beispiel ist $c = 1$. Es werden für jeden Ausgangspunkt z_0 zehn Iterationsschritte durchgeführt und z_0 dann gefärbt mit einer Farbe, die die Position von $f^{10}(z_0)$ zu den drei Fixpunkten $1, -0.5 + 0.866025i$ und $-0.5 - 0.866025i$ angibt. Dazu werden die Niveaus der Funktion $\mathrm{Im}(f^{10}(z))$ benutzt.

$$f[z\_] := \frac{2}{3}\left(z + \frac{1}{2z^2}\right)$$

```
g[x_, y_] := Nest[f, x + iy, 10]
b = 1.4;
ContourPlot[
    Evaluate[Im[g[x, y]]],
    {x, -b, b}, {y, -b, b}, PlotPoints-> 50,
    FrameTicks-> False];
```



Die mit nur geringem Rechenaufwand produzierte Abbildung zeigt (mindestens so gut wie die handelsüblichen Kunstgewerbsprodukte) das sogenannte *chaotische* Verhalten dieses Rekursionsverfahrens. □

Grenzwerte in Teilmengen

Wenn man eine Menge $A \subset \mathbb{R}^n$ und in dieser eine Folge von $\vec{x}_k \in A$ gegeben hat, die gegen ein $\vec{x}$ konvergiert, liegt dann $\vec{x}$ auch in A? Das muss nicht so sein. Man weiß ja nur, dass in jeder Kugel um $\vec{x}$ fast alle Folgenglieder liegen, also sicher Punkte aus A. Aber es können

durchaus darin auch (für jeden Radius) Punkte vom Komplement $\mathbb{R}^n \setminus A$ liegen. Dann ist $\vec{x}$ ein Randpunkt von A. Wenn A abgeschlossen ist, liegt der Grenzwert also auch in A (vorausgesetzt die Folge konvergiert!), andernfalls kann er auch ein Randpunkt sein, der nicht zu A gehört.

So konvergiert die Folge $(1 - \frac{1}{k}, 0)$ von Punkten in der offenen Kreisscheibe

$$U = \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 < 1\}$$

gegen $(1, 0) \notin U$. Betrachtet man sie dagegen als Folge in der abgeschlossenen Kreisscheibe, so liegt der Grenzwert auch in dieser.

Für das folgende erinnern wir daran, dass nach Analysis I jede Folge reeller Zahlen eine monotone Teilfolge enthält und dass monotone beschränkte Folgen konvergent sind. Also enthält jede beschränkte Folge reeller Zahlen eine konvergente Teilfolge. Sei nun $(\vec{x_k})$ eine beschränkte Folge im $\mathbb{R}^n$, d.h. eine, die in einer Kugel vom (vielleicht großen) Radius $R < \infty$ liegt. Dann ist die Folge der ersten Komponenten (x_{k1}) beschränkt. Nach der Vorbemerkung hat also $(\vec{x_k})$ eine Teilfolge, deren erste Komponenten konvergieren. Aus dieser Teilfolge können wir nach demselben Argument eine aussuchen, für die die Folge der zweiten Komponenten konvergiert; die Folge der ersten konvergiert dann immer noch! Nachdem wir das n-mal wiederholt haben, erhalten wir eine Teilfolge von $(\vec{x_k})$, deren sämtliche Komponentenfolgen konvergieren, die also selbst konvergiert: Jede beschränkte Folge im $\mathbb{R}^n$ enthält also eine konvergente Teilfolge.

Haben wir nun eine Folge $(\vec{x}_k)$ in einer *beschränkten und abgeschlossenen* Menge $A \subset \mathbb{R}^n$ (solche Mengen nennen wir auch *kompakt*), so enthält sie eine Teilfolge, die konvergiert. Der Limes liegt wegen der Abgeschlossenheit von A dann wieder in A. Wir haben:

Satz 15 (Folgenkompaktheit). *Jede Folge in einer kompakten Teilmenge $A \subset \mathbb{R}^n$ enthält eine in A konvergente Teilfolge.*

1.4 Abbildungen, Funktionen, Stetigkeit

- An verschiedenen Beispielen trainieren wir unsere Vorstellung von mehrdimensionalen Abbildungen.
- Wir erklären den Begriff der Stetigkeit für solche Abbildungen.
- Der Existenzsatz für Extremwerte ist - wie im 1-Dimensionalen - eine wichtige Ergänzung zu den Extremalkriterien mittels Differentialrechnung.

Wir betrachten Abbildungen $\vec{f} : \mathbb{R}^n \supset G \to \mathbb{R}^m$, die jedem Punkt einer Menge $G \subset \mathbb{R}^n$ einen Punkt im $\mathbb{R}^m$ zuordnet. Die Menge G heißt der Definitionsbereich von $\vec{f}$. Abbildungen mit Werten in $\mathbb{R} = \mathbb{R}^1$, also mit $m = 1$, nennt man auch Funktionen.

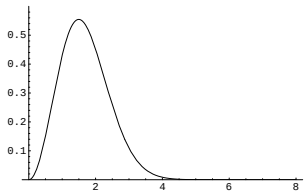
Veranschaulichung von mehrdimensionalen Abbildungen

1. **Veranschaulichung als Graph.**

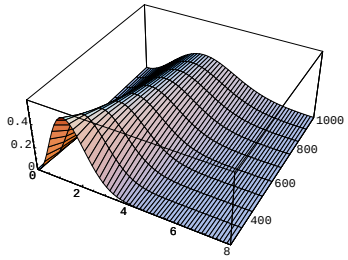
   $f : \mathbb{R} \to \mathbb{R}$
   Graph der Boltzmann-Maxwell-Verteilung aus der kinetischen Gastheorie

   $$v \mapsto 4\pi \left(\frac{M}{2\pi RT}\right)^{3/2} v^2 e^{-\frac{Mv^2}{2RT}}$$

   bei fester Temperatur T.



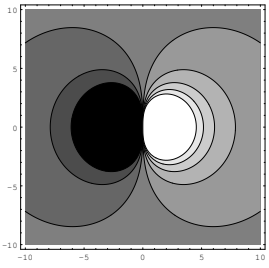
   $f : \mathbb{R}^2 \to \mathbb{R}$
   Der Graph derselben Funktion in Abhängigkeit von 2 Variablen v und T.



2. **Veranschaulichung als Niveaulinien und -flächen.**

   $f : \mathbb{R}^2 \to \mathbb{R}$
   Die Abbildung zeigt ein Dipolpotential der Form $u(x, y) = \frac{Q\vec{a}\cdot\vec{x}}{\sqrt{x^2+y^2}^3}$. Sie wurde erzeugt mit dem Mathematica-Befehl

   ContourPlot[x/$\sqrt{x^2 + y^2}^3$,
   {x,-10,10},{y,-10,10},
   PlotPoints− > 100]

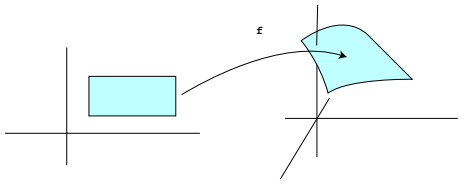


Dabei ist also $Q = 1, \vec{a} = (1, 0)$. Die Option „PlotPoints− >100“ erzwingt eine feinere Darstellung (mehr berechnete Punkte als automatisch), weil sonst Verzerrungen durch Rechenfehler sichtbar werden.

Die Menge $\{\vec{x} \mid f(\vec{x}) = c\}$ nennt man allgemein die *Niveaumenge* von f zum Niveau c. Für Funktionen $f : \mathbb{R}^2 \to \mathbb{R}$ sind die Niveaumengen typischerweise Kurven, für $f : \mathbb{R}^3 \to \mathbb{R}$ Flächen, die sogenannten *Niveaukurven* bzw. *Niveauflächen*. Eine graphische Darstellung von Niveauflächen ist problematisch, aber im Kopf kann man sich sehr gut vorstellen, dass der 3-dimensionale Raum durch die Niveauflächen einer solchen Funktion „geblättert“ wird.

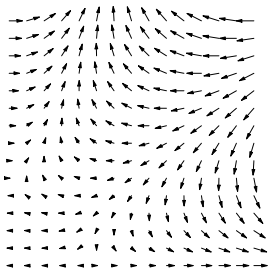
3. **Veranschaulichung als Bildmenge.**

$\vec{f} : \mathbb{R}^2 \to \mathbb{R}^3$



4. **Veranschaulichung als Vektorfeld.**

$\vec{f} : \mathbb{R}^2 \to \mathbb{R}^2$

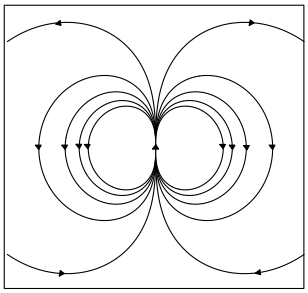


Wiederum kann man sich Vektorfelder im 3-dimensionalen Raum ganz gut vorstellen, aber nicht sehr gut graphisch darstellen.

5. **Veranschaulichung durch Flusslinien.**

Bei der Darstellung von Vektorfeldern überlagern sich schon im $\mathbb{R}^2$ oft Bilder der Pfeile so, dass man nichts mehr erkennen kann.

Dann kann man sich oft übersichtlichere Verhältnisse schaffen, indem man nicht die Pfeile darstellt, sondern die sogenannten *Flusslinien*, die man erhält, wenn ein Partikel den Pfeilen folgt. Allerdings hat man damit Information verloren: Es ist nicht mehr klar, wie schnell die einzelnen Flusslinien durchlaufen werden. (Die Ähnlichkeit mit den Niveaukurven des Dipolpotentials hat einen tieferen Grund, auf den wir hier nicht eingehen.)

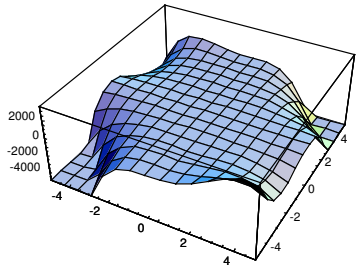


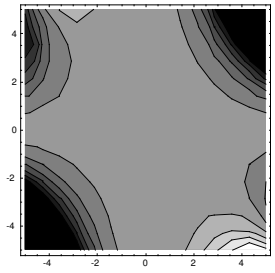
Das folgende Beispiel soll demonstrieren, dass man die „Veranschaulichung“ von Funktionen nicht einfach dem Rechner überlassen kann.

Beispiel 16. Wir versuchen, eine anschauliche Vorstellung vom Verlauf der Funktion

$$f(x,y) = x^3y^2(1-x-y)$$

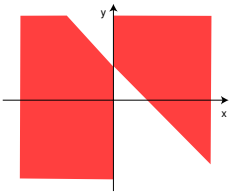
zu gewinnen.





So ganz überzeugend ist das nicht. Deshalb bemühen wir außer Mathematica auch noch unseren Kopf.

Der Faktor x^3y^2 ist ≤ 0 in der Halbebene links der y-Achse, der Faktor $(1-x-y)$ ist ≤ 0 oberhalb der Geraden $y = 1-x$. Das Produkt der beiden Faktoren ist deshalb in dem gefärbten Bereich ≤ 0. Das 0-Niveau ist die Vereinigung der beiden Koordinatenachsen mit der Geraden $y = 1-x$.



Diese Überlegung liefert Einsichten über die Funktion, die die obigen Bildern nicht vermitteln. □

Wir übertragen nun den Grenzwertbegriff von Folgen auf Funktionen.

Definition 17 (Grenzwerte einer Abbildung). Seien $\vec{f}: \mathbb{R}^n \supset G \to \mathbb{R}^m$ eine Abbildung, $\vec{a} \in \mathbb{R}^n$ und $\vec{b} \in \mathbb{R}^m$. Wir sagen, dass $\vec{f}(\vec{x})$ gegen $\vec{b}$ geht, wenn $\vec{x}$ gegen $\vec{a}$ geht, falls gilt:

$$\text{Für jede Folge } (\vec{x}_k) \text{ in } G\backslash\{\vec{a}\} \text{ mit } \lim_{k\to\infty} \vec{x}_k = \vec{a} \text{ gilt } \lim_{k\to\infty} \vec{f}(\vec{x}_k) = \vec{b}.$$

Um offensichtlichen Unsinn zu verhindern, wollen wir außerdem fordern, dass es wenigstens eine Folge $(\vec{x}_k)$ in $G \setminus \{\vec{a}\}$ gibt, die gegen $\vec{a}$ konvergiert. Notation: $\lim_{\vec{x}\to\vec{a}} \vec{f}(\vec{x}) = \vec{b}$.

Kurz gesagt:

$$\boxed{\lim_{\vec{x}\to\vec{a}} \vec{f}(\vec{x}) = \vec{b} : \Longleftrightarrow \lim_{k\to\infty} \vec{f}(\vec{x_k}) = \vec{b} \text{ für jede gegen } \vec{a} \text{ konvergierende Folge.}}$$

Beispiel 18. Die Funktion $f(x,y) := \frac{x^2}{\sqrt{x^2+y^2}}$ ist definiert auf $G = \mathbb{R}^2\backslash\{0\}$. Es ist

$$0 \leq f(x,y) = |x| \frac{|x|}{\sqrt{x^2+y^2}} \leq |x|.$$

Ist daher $\lim_{k\to\infty}(x_k, y_k) = (0,0)$, so folgt $\lim_{k\to\infty} f(x_k, y_k) = 0$. Also

$$\lim_{(x,y)\to(0,0)} \frac{x^2}{\sqrt{x^2+y^2}} = 0.$$

□

Definition 19 (Stetigkeit). Seien $\vec{f}\colon \mathbb{R}^n \supset G \to \mathbb{R}^m$ eine Abbildung und $\vec{a} \in G$. Dann heißt $\vec{f}$ *stetig* in $\vec{a}$, wenn

$$\boxed{\lim_{\vec{x}\to\vec{a}} \vec{f}(\vec{x}) = \vec{f}(\vec{a}).}$$

Die Abbildung $\vec{f}$ heißt *stetig (auf G)*, wenn sie in jedem Punkt $\vec{a} \in G$ stetig ist.

Beispiel 20. Typisches Beispiel einer vektorwertigen Funktion ist das elektrische Feld. In homogenen Medien ist es stetig, in der Regel sogar differenzierbar, während es an Grenzschichten oder dort, wo Punktladungen sitzen, in der Regel unstetig ist.

Beispiel 21. Wir betrachten für $\vec{u} \in \mathbb{R}^3$ die Vektorprodukt-Funktion $\vec{f}\colon \mathbb{R}^3 \to \mathbb{R}^3$ mit

$$\vec{f}(\vec{x}) = \vec{x} \times \vec{u}.$$

Dann ist $|\vec{x} \times \vec{u}| = |\vec{x}||\vec{u}\,|\sin\phi \le |\vec{x}||\vec{u}\,|$, also

$$\begin{aligned} |\vec{f}(\vec{x}) - \vec{f}(\vec{a})| &= |\vec{x} \times \vec{u} - \vec{a} \times \vec{u}\,| = |(\vec{x} - \vec{a}) \times \vec{u}\,| \\ &\le |\vec{x} - \vec{a}||\vec{u}\,| \to 0 \text{ für } \vec{x} \to \vec{a}. \end{aligned}$$

Daher ist $\vec{f}$ stetig. □

Stetigkeit komponentenweise. Weil Konvergenz im $\mathbb{R}^m$ äquivalent zu komponentenweiser Konvergenz ist, gilt $\lim_{k\to\infty} \vec{f}(\vec{x}_k) = \vec{b}$ genau dann, wenn $\lim_{k\to\infty} f_i(\vec{x}_k) = b_i$ für alle i.

Satz 22 (Komponentenweise Stetigkeit). *$\vec{f} = (f_1, \ldots, f_m)$ ist genau dann stetig, wenn alle Komponentenfunktionen f_i stetig sind.*

Beispiel 23. Wir betrachten noch einmal die Abbildung

$$\vec{f}(\vec{x}) = \vec{x} \times \vec{u}$$

aus dem letzten Beispiel und zeigen die komponentenweise Stetigkeit. Die erste Komponentenfunktion ist

$$f_1(x_1, x_2, x_3) = x_2u_3 - x_3u_2.$$

Für $\vec{a} = (a_1, a_2, a_3) \in \mathbb{R}^3$ ist

$$\begin{aligned} |f_1(x_1, x_2, x_3) - f_1(a_1, a_2, a_3)| &= |(x_2 - a_2)u_3 - (x_3 - a_3)u_2| \\ &\le |x_2 - a_2||u_3| + |x_3 - a_3||u_2| \\ &\le |\vec{x} - \vec{a}|(|u_2| + |u_3|). \end{aligned}$$

Für $\vec{x} \to \vec{a}$ geht das gegen Null, daher ist die erste Komponentenfunktion von $\vec{f}$ stetig. Ebenso zeigt man die Stetigkeit für die beiden anderen, und daher ist auch $\vec{f}$ selbst stetig. □

Partielle Stetigkeit. Wir haben gesehen, dass man bei einer Abbildung $\vec{f}\colon \mathbb{R}^n \to \mathbb{R}^m$ die Stetigkeit einfach an den (reellwertigen) Komponentenfunktionen untersuchen kann. $\mathbb{R}^m$ oder $\mathbb{R}$ auf der rechten Seite macht also „keinen großen Unterschied“. Wir betrachten der einfacheren Notation wegen jetzt eine Funktion $f : \mathbb{R}^n \supset G \to \mathbb{R}$. Dann können wir $f(\vec{x}) = f(x_1, \ldots, x_n)$ als Funktion jeder einzelnen Variablen betrachten, indem wir uns

vorstellen, dass die anderen festbleiben. Es stellt sich die naheliegende Frage, ob f in $\vec{a}$ stetig ist, wenn alle die Funktionen

$$\begin{aligned} x_1 &\mapsto f(x_1, a_2, a_3, \ldots, a_n) \\ x_2 &\mapsto f(a_1, x_2, a_3, \ldots, a_n) \\ x_3 &\mapsto f(a_1, a_2, x_3, \ldots, a_n) \\ &\ldots \\ x_n &\mapsto f(a_1, a_2, a_3, \ldots, x_n) \end{aligned}$$

stetig sind. Man nennt das partielle Stetigkeit, weil man immer nur einen Teil der Variablen - nämlich eine - als variabel betrachtet. Folgt aus partieller Stetigkeit die Stetigkeit? Das ist nicht so:

Partielle Stetigkeit impliziert NICHT Stetigkeit.

Beispiel 24. Sei $f : \mathbb{R}^2 \to \mathbb{R}$ gegeben durch $f(0,0) := 0$ und

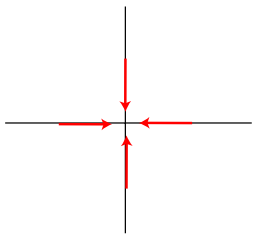
$$f(x,y) := \frac{xy}{x^2+y^2} \text{ für } (x,y) \neq (0,0).$$

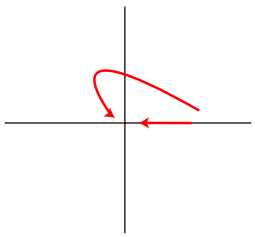
Für $\lambda \in \mathbb{R}$ geht nämlich die Folge $(\frac{1}{k}, \frac{\lambda}{k})$ gegen $(0,0)$, aber es ist

$$f(\frac{1}{k}, \frac{\lambda}{k}) = \frac{\lambda}{k^2(\frac{1}{k^2} + \frac{\lambda^2}{k^2})} = \frac{\lambda}{1+\lambda^2}.$$

Für $\lambda \neq 0$ und $k \to \infty$ geht das also nicht gegen $0 = f(0,0)$. Andererseits ist f in $(0,0)$ wegen $f(x,0) = 0 = f(0,y)$ aber partiell stetig.

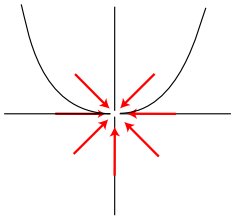
Dieses Beispiel zeigt genauer, warum partielle Stetigkeit viel schwächer ist als „totale“ Stetigkeit: Die Variable $\vec{x}$ muss sich der Stelle $\vec{a}$ auf beliebige Weise nähern dürfen. Bei der partiellen Stetigkeit schränkt man sich aber auf achsenparallele Annäherung ein.





In unserem Beispiel ist die Funktion auf allen Geraden durch den Nullpunkt jeweils konstant (Wert $\lambda/(1+\lambda^2)$), nur im Nullpunkt hat sie definitionsgemäß den Wert 0. Der kommt heraus, wenn man auf der x-Achse ($\lambda = 0$) oder auf der y-Achse ($\lambda = \infty$) an den Nullpunkt heranläuft, aber eben nur dann.

Selbst wenn $f(\vec{x}) \to f(\vec{a})$ bei Annäherung auf allen Geraden durch $\vec{a}$ gilt, folgt daraus nicht die Stetigkeit in $\vec{a}$. Ein Gegenbeispiel liefert die Funktion g mit $g(x,y) = 1$, falls $y = x^2 \neq 0$, und $g(x,y) = 0$ sonst. Wie sieht der Graph dieser Funktion aus?



□

RECHENREGELN. Mit stetigen Abbildungen mehrerer Variabler kann man „rechnen“ wie mit solchen in einer Variablen: Summen und Hintereinanderschaltung stetiger Abbildungen sind stetig. Auch Produkte stetiger Abbildungen sind „meistens“ stetig. Man muss nur spezifizieren, was „Produkt“ bedeuten soll. Für reellwertige Funktionen ist das klar, für vektorwertige kann man das Skalarprodukt darunter verstehen, für $\mathbb{R}^3$-wertige das Vektorprodukt, siehe oben. In allen Fällen bleibt die Stetigkeit erhalten.

EXTREMWERTE STETIGER FUNKTIONEN. Alles in allem sind stetige Funktionen für Ingenieure nicht so wichtig wie zum Beispiel differenzierbare, weil eben Differentialrechnung ein so starkes Hilfsmittel ist. Wir wollen diesen Abschnitt aber mit einem Resultat über stetige Funktionen beschließen, das bei der Suche nach Extremwerten von grundlegender Bedeutung ist. Wenn man über Maximum und Minimum von Abbildungen reden will, betrachtet man sinnvollerweise nur reellwertige Abbildungen, also Funktionen f. Mit Hilfe der Differentialrechnung findet man, wie wir später sehen werden, „extremal-verdächtige“ Punkte; allerdings nur im Inneren des Definitionsbereichs. Dann muss man noch den Rand untersuchen. Bei einer Variablen waren das die zwei Endpunkte des Intervalls, in mehreren Variablen ist die Sache viel schwieriger. Ehe man diesen Aufwand betreibt, wüsste man natürlich gern, ob man nach einem Phantom sucht oder ob das Maximum bzw. Minimum wirklich irgendwo angenommen wird. Darüber gibt der folgenden Satz Auskunft:

Satz 25 (Extremwerte auf Kompakta). *Eine stetige Funktion $f : \mathbb{R}^n \supset A \to \mathbb{R}$ nimmt auf einer nicht-leeren kompakten (= abgeschlossenen und beschränkten) Menge A ihr Maximum und Minimum an.*

Eine typische Anwendung dieses Satzes werden wir später genauer untersuchen (Beispiel 69): Die differenzierbare, also stetige, Funktion f auf der (kompakten!) Kreisscheibe

$$D = \{(x, y) \mid x^2 + y^2 \leq 1\}$$

hat auf dem Rand Werte ≥ 1 und im Inneren nur an der Stelle $(x, y) = (0, 0)$ (mehrdimensionale) Ableitung $= 0$. Weiter ist $f(0, 0) = 0$. Dann weiß man ohne weitere Untersuchung (zum Beispiel der zweiten Ableitungen), dass die Funktion in $(0, 0)$ ihr Minimum und auf dem Rand (irgendwo) ihr Maximum annimmt.

Beweis des Satzes. Sei $M := \sup_{\vec{x} \in A} f(\vec{x})$. Dann ist $M \in \mathbb{R} \cup \{+\infty\}$. Nach Definition des Supremums gibt es eine Folge $(\vec{x}_k)$ in A mit $\lim_{k\to\infty} f(\vec{x}_k) = M$. Weil A kompakt ist, hat die Folge $(\vec{x}_k)$ eine konvergente Teilfolge mit Grenzwert $\vec{a} \in A$. Für diese Teilfolge ist der Grenzwert der Funktionswerte natürlich immer noch M. Also können wir einfach annehmen, dass $\lim_{k\to\infty} \vec{x}_k = \vec{a}$. Weil f stetig ist, ist dann $M = \lim_{k\to\infty} f(\vec{x}_k) = f(\vec{a})$. □

Beispiel 26. Die Kugel

$$B := \{\vec{x} \in \mathbb{R}^3 \,|\, |x| \leq 1\}$$

ist abgeschlossen (warum?) und beschränkt, also kompakt. Für $\vec{a} \in \mathbb{R}^3$ nimmt daher die stetige Funktion

$$f(\vec{x}) := |\vec{x} - \vec{a}|,\ \vec{x} \in B,$$

auf B ihr Maximum und ihr Minimum an. Beschreiben Sie, wo. Beachten Sie dabei die verschiedenen Möglichkeiten für $\vec{a}$ relativ zu B. □

1.5 Lineare Abbildungen

- "Differenzieren" heißt "linear approximieren". Deshalb wiederholen wir vor der Differentialrechung noch einmal die aus der Linearen Algebra bekannten linearen Abbildungen.

Eine sehr wichtige Familie von Abbildungen des $\mathbb{R}^n$ in den $\mathbb{R}^m$ sind die *linearen Abildungen*. Sie werden durch $(m \times n)$-Matrizen wie folgt gegeben: Ist

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ \vdots & \vdots & \dots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}$$

eine solche Matrix, so ist die (ebenfalls mit A bezeichnete) Abbildung gegeben durch das Matrixprodukt

$$A\vec{x} = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ \vdots & \vdots & \dots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n \end{pmatrix}.$$

In diesem Zusammenhang ist es üblich, Vektoren als *Spalten*vektoren zu schreiben. $A\vec{x}$ entsteht dann einfach durch Matrixmultiplikation. Die i-te Komponentenfunktion ist gegeben durch

$$A_i\vec{x} = (A\vec{x})_i = \sum_{j=1}^{n} a_{ij}x_j.$$

Die linearen Abbildungen haben die – einfach nachzurechnenden – charakteristischen Eigenschaften

$$A(c\vec{x}) = cA\vec{x},$$
$$A(\vec{x} + \vec{y}\,) = A\vec{x} + A\vec{y}.$$

Das hat zur Folge, dass zum Beispiel die Gerade $\vec{a} + t\vec{b}$ durch A abgebildet wird auf

$$A(\vec{a} + t\vec{b}) = A\vec{a} + tA\vec{b}$$

und das ist wieder eine Gerade, falls $A\vec{b} \neq \vec{0}$, andernfalls ein Punkt. Lineare Gebilde wie Geraden oder auch Ebenen gehen also unter linearen Abbildungen wieder in lineare Gebilde über: ein Grund für das Adjektiv „linear". Möglicherweise verringert sich allerdings die Dimension dieser Gebilde, wenn der $\text{Rang}(A) < n$ ist.

Abgesehen von den konstanten Abbildungen sind die linearen Abbildungen die einfachsten Abbildungen von $\mathbb{R}^n$ in den $\mathbb{R}^m$ überhaupt. Das sollen die beiden folgenden Beispiele belegen.

Beispiel 27. Für $m = n = 1$ sieht eine lineare Abbildung einfach so aus: Man hat $A = (a)$ und

$$A : \mathbb{R} \to \mathbb{R},\ x \mapsto ax.$$

□

Beispiel 28. Für $m = 1$ ist $A = (a_1 \dots a_n)$ und

$$A\vec{x} = a_1x_1 + \ldots + a_nx_n = \vec{a} \cdot \vec{x}.$$

So ist für $n = 3$ etwa $A\begin{pmatrix} x \\ y \\ z \end{pmatrix} = 2x - y + 7z$ eine solche Abbildung. □

Satz 29. *Lineare Abbildungen sind stetig.*

Das gilt, weil ihre Komponentenfunktionen stetig sind:

$$|A_i\vec{x} - A_i\vec{y}| \leq \sum_{j=1}^{n} |a_{ij}||x_j - y_j| \to 0 \text{ für } \vec{x} \to \vec{y}.$$

1.6 Differentiation

- Wir erklären die Ableitung einer Abbildung in einem Punkt als deren lineare Approximation, also als eine Matrix.
- Die Ableitung kommt unter vielen verschiedenen Namen vor, von denen manche (wie das *totale Differential*) gelegentlich zu Unrecht zu besonders geheimnisvollen Objekten stilisiert werden.

Das Geheimnis der Differentialrechnung ist die Approximation von Funktionen durch einfache (nämlich lineare) Funktionen. Wir beschreiben das zunächst noch einmal für den Fall einer Variablen. Man hat

$$f(x) \approx f(x_0) + f'(x_0) \cdot (x - x_0) \tag{2}$$

Links steht eine beliebige differenzierbare Funktion $f(x)$. In der Nähe der Stelle x_0 kann man sie approximieren durch die Funktion auf der rechten Seite: eine lineare Funktion plus Konstante, d.h. eine Funktion der Form $mx+b$. Man hat den (im allgemeinen gekrümmten) Graphen der Funktion durch eine Gerade (nämlich die Tangente) approximiert. Den

$$Fehler = f(x) - (f(x_0) + f'(x_0) \cdot (x - x_0))$$

kann man ohne weiteren Aufwand jedenfalls qualitativ beschreiben: Für $x \to x_0$ geht er schneller als linear gegen Null:

$$\lim_{x \to x_0} \frac{Fehler}{x - x_0} = 0.$$

Wir wollen nun die Differentiation für Funktionen *mehrerer* Variablen behandeln. Zunächst stellen wir aber noch fest, dass man (2) auch so schreiben kann

$$f(x + \Delta x) \approx f(x) + f'(x)\Delta x \quad \text{mit} \lim_{\Delta x \to 0} \frac{Fehler}{\Delta x} = 0.$$

Definition 30. Seien $G \subset \mathbb{R}^n$ offen und $\vec{f} : G \to \mathbb{R}^m$ eine Abbildung. Dann heißt $\vec{f}$ in $\vec{x} \in G$ *differenzierbar*, wenn es eine lineare Abbildung, also eine $(m \times n)$-Matrix A gibt, so dass

$$\boxed{\vec{f}(\vec{x} + \vec{\Delta x}) \approx \vec{f}(\vec{x}) + A\vec{\Delta x} \quad \text{und} \lim_{\vec{\Delta x} \to 0} \frac{Fehler}{|\vec{\Delta x}|} = 0.} \tag{3}$$

Im Nenner rechts muss man den Betrag schreiben, weil man durch Vektoren nicht dividieren kann. Wenn $\vec{f}$ in $\vec{x}$ differenzierbar ist, gibt es auch nur *eine* solche Matrix - wir erklären im nächsten Abschnitt, wie man sie finden kann -, und die nennt man dann *die Ableitung* oder *das Differential* oder *die Funktionalmatrix* von $\vec{f}$ an der Stelle $\vec{x}$. Man schreibt dafür

$$\vec{f}\,'(\vec{x}) := A$$
$$\textit{oder}$$
$$d_{\vec{x}}\vec{f} := A.$$

Die Notation df (für reellwertige Funktionen) ist zum Beispiel in der *Thermodynamik* ausgesprochen üblich. Man nennt die Ableitung dort das *totale Differential* im Gegensatz zu den *partiellen* Ableitungen. Wir kommen darauf zurück.

$$\left.\begin{array}{l} Totales\,Differential \\ Differential \\ Ableitung\,und \\ Funktionalmatrix \end{array}\right\} \text{sind also nur verschiedene Namen für dasselbe Ding.}$$

Schreibt man $\vec{\Delta f} := \vec{f}(\vec{x} + \vec{\Delta x}) - \vec{f}(\vec{x})$, so erhält man also die Grundidee der Differentialrechnung, die lineare Approximation, in der ebenso komprimierten wie suggestiven Formel

$$\boxed{\vec{\Delta f} \approx \vec{f}\,'(\vec{x})\vec{\Delta x}.} \tag{4}$$

Der Fehler (linke Seite minus rechte Seite) geht für $\vec{\Delta x} \to 0$ schneller als linear gegen 0. Dabei sind $\vec{\Delta f}$ ein m-reihiger Vektor, $\vec{\Delta x}$ ein n-reihiger Vektor und $\vec{f}\,'(\vec{x})$ eine $(m \times n)$-Matrix.

Bemerkungen.

1. Die Ableitung von $\vec{f}$ an der Stelle $\vec{x}$ ist also eine von $\vec{x}$ abhängige Matrix oder lineare Abbildung. Man kann sie daher auch als abhängig von *zwei* Argumenten verstehen, nämlich von der Stellle $\vec{x}$, an der differenziert wird, und von $\vec{\Delta x}$ als Argument der linearen Abbildung A.

2. Die Definition der Differenzierbarkeit liefert erst einmal kein Rezept, wie man die Ableitung denn ausrechnen kann.

Beispiel 31. Für den Fall $m = n = 1$ ist die neue Ableitung eine $(1,1)$-Matrix, in der einfach die alte Ableitung als einzige Komponente steht. □

Beispiel 32. Wir betrachten $\vec{f}\colon \mathbb{R}^2 \to \mathbb{R}^2$ mit

$$\vec{f}(x,y) = \begin{pmatrix} xy \\ x^2 - y^2 \end{pmatrix}.$$

Dafür ist

$$\begin{aligned}
\vec{f}(x+\Delta x, y+\Delta y) &= \begin{pmatrix} (x+\Delta x)(y+\Delta y) \\ (x+\Delta x)^2 - (y+\Delta y)^2 \end{pmatrix} \\
&= \begin{pmatrix} xy + y\Delta x + x\Delta y + \Delta x\Delta y \\ x^2 - y^2 + 2x\Delta x - 2y\Delta y + (\Delta x)^2 - (\Delta y)^2 \end{pmatrix} \\
&= \begin{pmatrix} xy \\ x^2 - y^2 \end{pmatrix} + \begin{pmatrix} y\Delta x + x\Delta y \\ 2x\Delta x - 2y\Delta y \end{pmatrix} + \begin{pmatrix} \Delta x\Delta y \\ (\Delta x)^2 - (\Delta y)^2 \end{pmatrix} \\
&= \begin{pmatrix} xy \\ x^2 - y^2 \end{pmatrix} + \begin{pmatrix} y & x \\ 2x & -2y \end{pmatrix} \begin{pmatrix} \Delta x \\ \Delta y \end{pmatrix} + \begin{pmatrix} \Delta x\Delta y \\ (\Delta x)^2 - (\Delta y)^2 \end{pmatrix}.
\end{aligned}$$

Es bleibt der Fehler abzuschätzen:

$$\left| \frac{\Delta x \Delta y}{|(\Delta x, \Delta y)|} \right| = \underbrace{\frac{|\Delta x|}{\sqrt{x^2+y^2}}}_{\leq 1} |\Delta y| \to 0 \text{ für } (\Delta x, \Delta y) \to (0,0)).$$

Ebenso schließt man für die 2. Komponenten $(\Delta x)^2 - (\Delta y)^2$. Damit ist $\vec{f}$ überall differenzierbar und die Ableitung an der Stelle (x,y) ist

$$\vec{f}\,'(x,y) = d_{(x,y)}\vec{f} = \begin{pmatrix} y & x \\ 2x & -2y \end{pmatrix}.$$

□

Beispiel 33 (Geometrische Interpretation). Für den Fall $m = 1$ und $n = 2$, den unsere geometrische Anschauung gerade noch bewältigt, schreiben wir für den Augenblick wieder (x_0, y_0) statt $\vec{x}$ und (x,y) statt $\vec{x} + \vec{\Delta x}$. Weiter schreiben wir die $(1,2)$-Matrix A in der Form $A = (a\,b)$. Dann wird aus der Approximation (3)

$$f(x,y) \approx f(x_0,y_0) + (a\,b)\begin{pmatrix} x - x_0 \\ y - y_0 \end{pmatrix} = f(x_0,y_0) + a(x-x_0) + b(y-y_0). \tag{5}$$

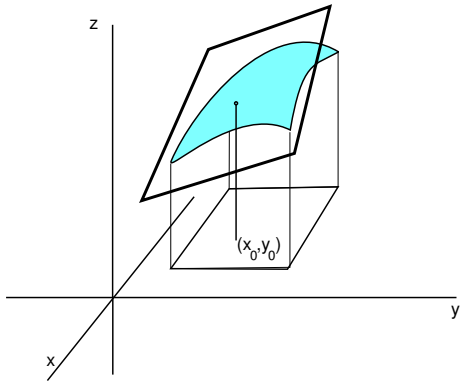
Die linke Seite gibt die Gleichung

$$z = f(x, y),$$

also die Gleichung für den Graphen von f. Die rechte Seite gibt

$$z = f(x_0, y_0) + a(x - x_0) + b(y - y_0),$$

und das ist die Gleichung für die Tangentialebene an den Graphen im Punkt $(x_0, y_0, f(x_0, y_0))$. Die Koeffizienten a und b sind die Steigungen der Tangentialebene in der x- bzw. in der y-Richtung.



□

Beispiel 34 („Physikalische“ Interpretation). Sei $\vec{\Delta x} = t\vec{v} \in \mathbb{R}^n$. Die Abbildung

$$t \mapsto \vec{x} + t\vec{v}, \quad 0 \leq t,$$

beschreibt eine gerade Bahnkurve, die mit Geschwindigkeit $\vec{v}$ in $\vec{x}$ startet. Die Gleichung

$$\vec{f}(\vec{x} + t\vec{v}) \approx \vec{f}(\vec{x}) + t\vec{f}\,'(\vec{x})\vec{v}$$

besagt entsprechend, dass sich die Funktion $\vec{f}$ längs der Bahn mit der Geschwindigkeit

$$\vec{f}\,'(\vec{x})\vec{v}$$

ändert. Beachten Sie, dass die Ableitung gewissermaßen *zwei* Argumente hat: $\vec{x}$ gibt an, an welcher Stelle man sich gerade befindet und $\vec{v}$, in welcher Richtung und wie schnell man den Punkt $\vec{x}$ verändert.

Komponentenweise Differentiation. Wir schreiben

$$\vec{f} = \begin{pmatrix} f_1 \\ \vdots \\ f_m \end{pmatrix} \text{ und } A = \begin{pmatrix} A_1 \\ \vdots \\ A_m \end{pmatrix}$$

und (3) komponenentenweise. Wir finden

$$f_i(\vec{x} + \vec{\Delta x}) \approx f_i(\vec{x}) + A_i\vec{\Delta x} \quad \text{und} \quad \lim_{\vec{\Delta x} \to 0} \frac{Fehler_i}{|\vec{\Delta x}|} = 0.$$

Daraus sieht man sofort, dass $\vec{f}$ genau dann in $\vec{x}$ differenzierbar ist, wenn alle seine Komponentenfunktionen f_i in $\vec{x}$ differenzierbar sind. Die einzeiligen Matrizen $A_i = f_i'(\vec{x})$ bilden dann gerade die Zeilen der Matrix $A = \vec{f}\,'(\vec{x})$.

Nach dieser Feststellung kann man sich bei der Differentialrechnung oft auf *reellwertige* Funktionen beschränken. Vektorwertige differenziert man einfach komponentenweise.

RICHTUNGSABLEITUNG. Für $\vec{\Delta x} = t\vec{v}$ gilt, weil $\vec{f}'(\vec{x})$ linear ist,

$$\vec{f}(\vec{x} + t\vec{v}) = \vec{f}(\vec{x}) + t\vec{f}'(\vec{x})\vec{v} + Fehler(t\vec{v}),$$

wobei

$$\lim_{t\to 0} \frac{Fehler(t\vec{v})}{t} = \lim_{t\to 0} \frac{Fehler(t\vec{v})}{\pm|t\vec{v}|}|\vec{v}| = \pm \lim_{\vec{\Delta x}\to 0} \frac{Fehler(\vec{\Delta x})}{|\vec{\Delta x}|}|\vec{v}| = 0.$$

Daraus folgt

$$\vec{f}'(\vec{x})\vec{v} = \lim_{t\to 0} \frac{\vec{f}(\vec{x} + t\vec{v}) - \vec{f}(\vec{x})}{t}. \tag{6}$$

Wenn $\vec{v}$ ein *Einheitsvektor* oder eine sogenannte *Richtung* ist, d.h. wenn $|\vec{v}| = 1$ ist, nennt man diesen Ausdruck auch die *Richtungsableitung von $\vec{f}$ an der Stelle $\vec{x}$ in Richtung $\vec{v}$* und schreibt dafür

$$\boxed{\frac{\partial \vec{f}}{\partial \vec{v}}(\vec{x}) := \vec{f}'(\vec{x})\vec{v}.} \tag{7}$$

Die Gleichung (6) kann man zur Berechnung der Ableitung $\vec{f}'(\vec{x})$ benutzen. Wir gehen darauf im nächsten Abschnittt genauer ein.

1.7 Partielle Ableitungen und totales Differential

- Die Ableitung=das totale Differential ist eine Matrix, die partiellen Ableitungen sind deren Einträge.
- Partielle Differentiation ist also eine Methode zur Berechnung der Ableitung.

Wie berechnet man die Ableitungsmatrix einer differenzierbaren Funktion?

Sei $\vec{f} : \mathbb{R}^n \supset G \to \mathbb{R}^m$ in $\vec{x}$ differenzierbar. Wir betrachten die Richtungsableitung in Richtung des j-ten Standard-Basisvectors $\vec{e}_j = (0, \ldots, 0, 1, 0, \ldots, 0)$ mit einer 1 in der j-ten Komponente. Dafür ist dann

$$\frac{\partial \vec{f}}{\partial \vec{e}_j}(\vec{x}) = \vec{f}\,'(\vec{x})\vec{e}_j = \lim_{t\to 0} \frac{\vec{f}(\vec{x} + t\vec{e}_j) - \vec{f}(\vec{x})}{t} = \lim_{t\to 0} \frac{\vec{f}(x_1, \ldots, x_j + t, \ldots, x_n) - \vec{f}(x_1, \ldots, x_n)}{t}.$$

Aber die rechte Seite ist einfach die Ableitung $g'(x)$ der Funktion

$$g(x) := \vec{f}(x_1, \ldots, x_{j-1}, x, x_{j+1}, \ldots, x_n)$$

von einer Variablen an der Stelle $x = x_j$. Wir nennen das die *j-te partielle Ableitung* von $\vec{f}$ an der Stelle $\vec{x}$ und bezeichnen sie mit

$$\frac{\partial \vec{f}}{\partial x_j}(\vec{x}), \text{gelegentlich auch mit } \vec{f}_{x_j}(\vec{x}).$$

Definition 35. Die *partielle Ableitung von $f : \mathbb{R}^n \supset G \to \mathbb{R}^n$ an der Stelle $\vec{x}$* ist definiert als

$$\boxed{\frac{\partial \vec{f}}{\partial x_j}(\vec{x}) := \frac{\partial \vec{f}}{\partial \vec{e}_j}(\vec{x}) = \lim_{t\to 0} \frac{\vec{f}(x_1, \ldots, x_j + t, \ldots, x_n) - \vec{f}(x_1, \ldots, x_n)}{t}.}$$

Ist $\vec{f} = \begin{pmatrix} f_1 \\ \vdots \\ f_m \end{pmatrix}$, so ist $\frac{\partial \vec{f}}{\partial x_j} = \begin{pmatrix} \frac{\partial f_1}{\partial x_j} \\ \vdots \\ \frac{\partial f_m}{\partial x_j} \end{pmatrix}$. Und für eine $(m \times n)$-Matrix A liefert $A\vec{e}_j$ bekanntlich gerade die j-te Spalte. Damit erhalten wir das gewünschte Rezept zur Berechnung der Ableitungsmatrix mittels "Analysis 1-Differentiation":

Satz 36 (Ableitungsmatrix, Funktionalmatrix). *Die Ableitung einer differenzierbaren Abbildung vom $\mathbb{R}^n$ in den $\mathbb{R}^m$ an der Stelle $\vec{x}$ ist eine $(m \times n)$-Matrix gebildet aus den partiellen Ableitungen:*

$$\vec{f}\,'(\vec{x}) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(\vec{x}) & \ldots & \frac{\partial f_1}{\partial x_n}(\vec{x}) \\ & \ldots & \\ \frac{\partial f_m}{\partial x_1}(\vec{x}) & \ldots & \frac{\partial f_m}{\partial x_n}(\vec{x}) \end{pmatrix}.$$

Man nennt sie auch die Ableitungsmatrix *oder* Funktionalmatrix.

Dann ist

$$\boxed{\vec{f}\,'(\vec{x})\vec{\Delta x} = \vec{f}\,'(\vec{x}) \begin{pmatrix} \Delta x_1 \\ \vdots \\ \Delta x_n \end{pmatrix} = \frac{\partial \vec{f}}{\partial x_1}(\vec{x})\Delta x_1 + \ldots + \frac{\partial \vec{f}}{\partial x_n}(\vec{x})\Delta x_n,} \tag{8}$$

und die Fehlerapproximation aus (4) ist damit gegeben durch

$$\boxed{\Delta\vec{f} = \vec{f}(\vec{x} + \vec{\Delta x}) - \vec{f}(\vec{x}) \approx \vec{f}\,'(\vec{x})\vec{\Delta x} = \frac{\partial \vec{f}}{\partial x_1}(\vec{x})\Delta x_1 + \ldots + \frac{\partial \vec{f}}{\partial x_n}(\vec{x})\Delta x_n.} \tag{9}$$

Beispiel 37 (Totales Differential, Pfaffsche Formen).
Thermodynamik I
Einen ganz einfachen Spezialfall des totalen Differentials haben wir für die erste Komponentenfunktion

$$f(x_1, \ldots, x_n) = x_1.$$

Dafür ist

$$dx_1 = (\frac{\partial x_1}{\partial x_1}, \ldots, \frac{\partial x_1}{\partial x_n}) = (1, 0, \ldots, 0).$$

Es folgt

$$\frac{\partial f}{\partial x_1}\, dx_1 = (\frac{\partial f}{\partial x_1}, 0, \ldots, 0)$$

und entsprechend für die anderen Komponentenfunktionen. Das bedeutet aber

$$df = (\frac{\partial f}{\partial x_1}, \ldots, \frac{\partial f}{\partial x_n}) = \frac{\partial f}{\partial x_1} dx_1 + \ldots + \frac{\partial f}{\partial x_n} dx_n. \tag{10}$$

Eine anschauliche Interpretation dieser Formel ist die Gleichung (9), und so wird das totale Differential oft verwendet: als eine Fehlerapproximation. Die diversen partiellen Ableitungen geben an, „wie empfindlich“ f auf Änderungen der entsprechenden Variablen ragiert, und das totale Differential kombiniert diese Informationen für den Fall, dass sich alle Variablen ändern.
Allgemeiner finden Sie (zum Beispiel in der *Thermodynamik*) Ausdrücke der Form

$$g_1 dx_1 + \ldots + g_n dx_n, \tag{11}$$

wobei die Funktionen $g_1, \ldots, g_n$ nicht unbedingt die sämtlichen partiellen Ableitungen einer Funktion f sind, zum Beispiel $x\,dx + xy\,dy$. Ein solcher Ausdruck (11) ist also nichts anderes, als eine lineare Abbildung oder Matrix

$$(g_1, \ldots, g_n)$$

mit Funktionen als Komponenten, die man auch ein *(unvollständiges) Differential* oder eine *Pfaffsche Form* nennt.
Wenn (11) das totale Differential einer Funktion f ist, d.h. wenn es eine Funktion f gibt, so dass

$$\frac{\partial f}{\partial x_i} = g_i,$$

heißt das Differential auch *vollständig*.
Ein Beispiel für eine Pfaffsche Form ist

$$x_2 dx_1$$

oder, in der (V, P)-Ebene der Thermodynamik,

$$PdV.$$

Überlegen Sie, dass PdV unvollständig, aber $PdV + VdP$ vollständig ist.

Partielle und totale Differenzierbarkeit. Wenn für eine Funktion alle partiellen Ableitungen existieren, so nennt man sie *partiell differenzierbar*. Wie bei der partiellen Stetigkeit folgt auch hier aus der partiellen Differenzierbarkeit NICHT die Differenzierbarkeit. Die Funktion mit

$$f(x,y) := \frac{xy}{x^2+y^2} \text{ und } f(0,0) := 0$$

liefert auch hier ein Gegenbeispiel.

Aber es gilt der folgende wunderbare Satz:

Satz 38. *Existieren für $\vec{f}\colon \mathbb{R}^n \supset G \to \mathbb{R}^m$ alle partiellen Ableitungen und sind diese stetig, so ist $\vec{f}$ differenzierbar.*

Damit ist die mehrdimensionale Differentiation in der Regel endgültig auf die 1-dimensionale zurückgeführt:

Verfahren. Will man eine Funktion mehrerer Variablen differenzieren, so berechnet man die partiellen Ableitungen.

- Sind diese *stetig*, so ist die Funktion (total) differenzierbar und man hat die Ableitungsmatrix bereits berechnet.
- Existiert eine der partiellen Ableitungen nicht, so ist die Funktion nicht differenzierbar.
- Existieren die partiellen Ableitungen, aber sind sie nicht alle stetig, so kann die Funktion differenzierbar sein oder auch nicht. In diesem Fall – der glücklicherweise selten ist – muss man die Existenz des Fehler-Grenzwertes in der Definition (3) nachprüfen, wobei A die Matrix der partiellen Ableitungen ist.
- Ist f nicht stetig, so ist f nicht differenzierbar.

Beispiel 39. Dies ist ein einfaches numerisches Beispiel für die Berechnung der Funktionalmatrix und der dadurch gegebenen linearen Abbildung.

Sei $\vec{f}\colon \mathbb{R}^2 \to \mathbb{R}^3$ gegeben durch

$$\vec{f}(x,y) = \begin{pmatrix} xy \\ x+y \\ x \end{pmatrix}.$$

Dann sind die partiellen Ableitungen offensichtlich stetig und

$$\vec{f}\,'(x,y) = \begin{pmatrix} y & x \\ 1 & 1 \\ 1 & 0 \end{pmatrix}.$$

Weiter hat man zum Beispiel

$$\vec{f}\,'(7,5) = \begin{pmatrix} 5 & 7 \\ 1 & 1 \\ 1 & 0 \end{pmatrix}, \quad \vec{f}\,'(7,5)\begin{pmatrix} 2 \\ 3 \end{pmatrix} = \begin{pmatrix} 31 \\ 5 \\ 2 \end{pmatrix}.$$

□

1.8 Der Gradient

- Wir lernen den Gradienten kennen als eine andere Schreibweise aber auch als eine andere Interpretation der Ableitung einer reellwertigen Funktion.

Der Fall reellwertiger Funktionen ($m = 1$) ist vermutlich der wichtigste in der Differentialrechnung. In diesem Fall kann man die $(1 \times n)$-Matrix $f'(\vec{x})$ auch durch die transponierte $(n \times 1)$-Matrix $f'(\vec{x})^T$, d.h. als Vektor im $\mathbb{R}^n$ auffassen und das Produkt $f'(\vec{x})\vec{\Delta x}$ als Skalarprodukt lesen. Man nennt $f'(\vec{x})^T$ auch *den Gradienten* von f an der Stelle $\vec{x}$.

Definition 40 (Gradient). Der *Gradient* einer differenzierbaren Funktion $f : \mathbb{R}^n \supset G \to \mathbb{R}$ ist die Ableitung $f'(\vec{x})$ interpretiert als Vektor oder, genauer, als Vektorfeld:

$$\boxed{\operatorname{grad}_{\vec{x}} f := \vec{\nabla}_{\vec{x}} f := \begin{pmatrix} \frac{\partial f}{\partial x_1}(\vec{x}) \\ \vdots \\ \frac{\partial f}{\partial x_n}(\vec{x}) \end{pmatrix} = \left(\frac{\partial f}{\partial x_1}(\vec{x}), \ldots, \frac{\partial f}{\partial x_n}(\vec{x}) \right)^T.}$$

Gemäß der zu Anfang von Abschnitt 1.2 gemachten Bemerkung über Spalten- und Zeilennotation nennen wir aber auch $(\frac{\partial f}{\partial x_1}(\vec{x}), \ldots, \frac{\partial f}{\partial x_n}(\vec{x}))$ den Gradienten, wenn wir diesen Ausdruck als Vektor und nicht als Matrix oder lineare Abbildung verstehen.

Damit ergibt sich

$$f(\vec{x} + \vec{\Delta x}) \approx f(\vec{x}) + \operatorname{grad}_{\vec{x}} f \cdot \vec{\Delta x} = f(\vec{x}) + \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(\vec{x}) \Delta x_i.$$

oder

$$\Delta f \approx \operatorname{grad}_{\vec{x}} f \cdot \vec{\Delta x} = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i} \Delta x_i = f'(\vec{x})\vec{\Delta x}. \tag{12}$$

Das ist also die Änderung von f in Richtung von $\vec{\Delta x}$, vgl. (7). Für $|\vec{u}| = 1$ ist

$$\operatorname{grad}_{\vec{x}} f \cdot \vec{u} = \frac{\partial f}{\partial \vec{u}}.$$

Der Gradient ermöglicht auch bei beliebiger Dimension n noch eine „anschauliche“ Interpretation der Ableitung. Nach (12) wird die Änderung von f an der Stelle $\vec{x}$ in der Richtung $\vec{\Delta x}$ gegeben durch

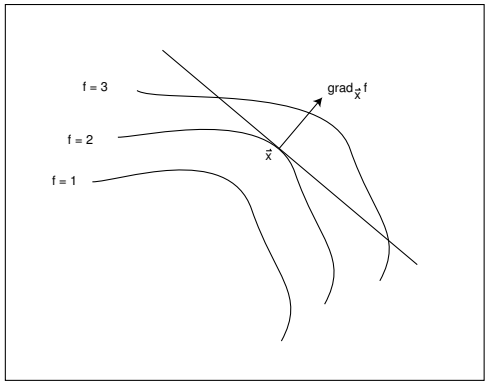
$$\Delta f \approx |\operatorname{grad}_{\vec{x}} f| \cdot |\vec{\Delta x}| \cos \alpha,$$

wobei α den Winkel zwischen dem Gradienten und $\vec{\Delta x}$ bezeichnet. Die Änderung ist

- am größten, wenn $\vec{\Delta x}$ in die Richtung des Gradienten weist,
- am kleinsten (negativ), wenn $\vec{\Delta x}$ entgegengesetzt zur Richtung des Gradienten weist,
- und 0, wenn $\vec{\Delta x}$ orthogonal zum Gradienten weist.

Satz 41 (Bedeutung des Gradienten). *Der Gradient einer Funktion (eines skalaren Feldes) f weist in die Richtung des stärksten Wachstums von f. Er steht senkrecht auf dem Niveau von f.*

Die letztere Aussage ergibt sich so: Bewegt man sich von $\vec{x}$ aus auf der Niveaulinie oder -fläche, so bleibt f natürlich konstant, d.h. in dieser Richtung ist $\Delta f = 0$ und deshalb $\operatorname{grad}_{\vec{x}} f \cdot \vec{\Delta} x = 0$.



Beispiel 42 (Elektrisches Feld). Das elektische Feld $\vec{E}$ ist der negative Gradient der elektrischen Potentials U:

$$\vec{E} = -\operatorname{grad} U.$$

1.9 Anwendungsbeispiele für die Ableitung

- Wir lernen verschiedene Anwendungen der Differentialrechnung auf Probleme aus der Geometrie, der Mechanik und der Strömungsmechanik kennen.

In den Beispielen dieses Abschnitts berechnen wir partielle Ableitungen und unterstellen kommentarlos deren Stetigkeit.

Beispiel 43 (Geometrisches Beispiel). Das Dach eines Gebäudes ist gegeben als Graph der Funktion

$$f(x,y) = \sin 2x \, \cos y, \qquad -.4 \le x \le 1, \; -1 \le y \le 0.$$

An der Kanten $x = 1$ stößt es gegen eine vertikale Mauer. Entsteht dabei eine Rinne oder gar eine Mulde, in der das Regenwasser stehen bleibt?

Wir betrachten die Funktion

$$h(x,y,z) = z - f(x,y).$$

Das Dach ist gerade das 0-Niveau von h, und der Gradient von h

$$\operatorname{grad} h = (-2\cos 2x \, \cos y, \sin 2x \, \sin y, 1)$$

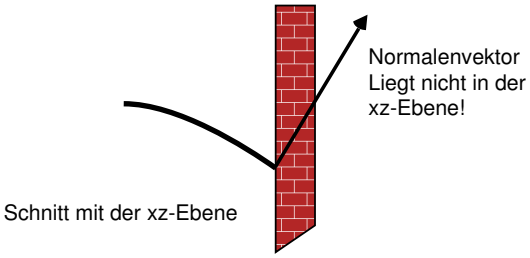
ausgewertet an der Stelle $(x, y, f(x,y))$ liefert deshalb einen Normalenvektor senkrecht auf der Dachfläche.

Seine z-Komponente ist 1, er zeigt also nach oben. Seine x-Komponente auf der Schnittkante mit der Wand ist

$$-2\cos 2 \cos y, \quad -1 \le y \le 0.$$

Weil $\cos 2 < 0$, aber $\cos t > 0$ für $-1 \le t \le 0$, ist das > 0. Also entsteht zwischen Dach und Wand eine Rinne.
Die y-Komponente des Normalenvektors längs der Schnittlinie ist

$$\sin 2 \, \sin y \approx 0.91 \, \sin y$$

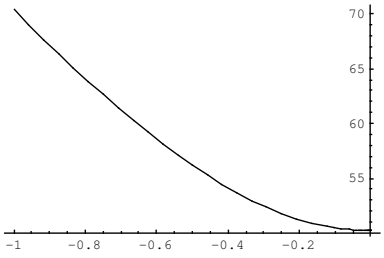


Für $-1 \le y \le 0$ ist der Sinus $\sin y$ negativ, die Komponente also überall < 0. Daher gibt es keine Mulde.

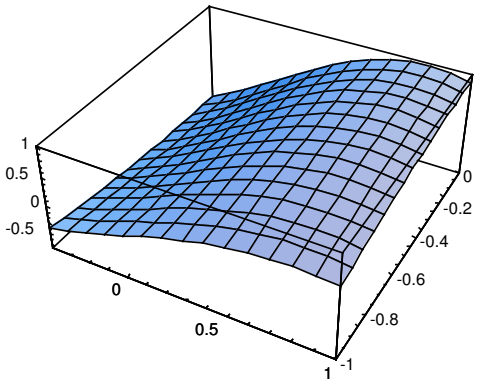
Will man genauer den Anschlußwinkel zwischen Dach und Wand in Abhängigkeit von y ermitteln, so muss man den Winkel δ zwischen der oben berechneten Dach-Normalen und der Mauer-Normalen $(1,0,0)$ ermitteln:

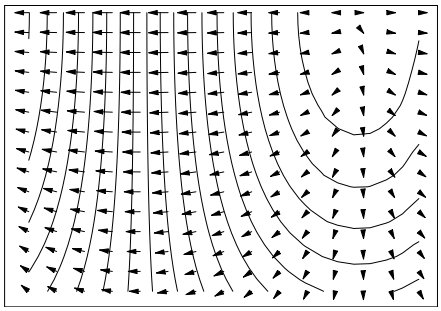
$$\delta = \arccos \frac{-2\cos 2 \cos y}{\sqrt{(-2\cos 2 \cos y)^2 + (\sin 2 \sin y)^2 + 1^2}}$$

Ein Plot dieser Funktion zeigt, dass der Winkel zwischen 70^0 und etwas über 50^0 liegt.



Die folgenden Abbildungen zeigen das Dach und das Feld $-\operatorname{grad} f$, also das Geschwindigkeitsfeld abfließenden Regenwassers zusammen mit den Niveaulinien.





□

Beispiel 44 (Feldeffekt-Transistor FET).
Werkstoffe II, Abschnitt 7
Der Drainstrom I_D eines Transistors ist eine Funktion der Steuerspannung U_{GS} und der Drain-Source-Spannung U_{DS}. Deshalb gilt

$$dI_D = \frac{\partial I_D}{\partial U_{GS}}\, dU_{GS} + \frac{\partial I_D}{\partial U_{DS}}\, dU_{DS}.$$

Dieselbe Information ist enthalten in der Formel

$$\operatorname{grad} I_D = \left(\frac{\partial I_D}{\partial U_{GS}}, \frac{\partial I_D}{\partial U_{DS}}\right)$$

oder in

$$I_D'(U_{GS}, U_{DS}) = \left(\frac{\partial I_D}{\partial U_{GS}}\ \frac{\partial I_D}{\partial U_{DS}}\right).$$

Die erste Formel drückt aber deutlicher aus, dass Änderungen von I_D sich ergeben als Überlagerung der Änderungen dU_{GS} und dU_{DS} gewichtet jeweils mit der partiellen Ableitung, d.h. mit der Sensitivität von I_D gegenüber dieser Variablen.

Beispiel 45 (Wechselstromwiderstand).
Grundlagen der Elektrotechnik, Teil A
Der Wechselstromwiderstand einer RL-Reihenschaltung ist $W = \sqrt{R^2 + \omega^2 L^2}$. Wir betrachten das als Funktion von drei Variablen und erhalten

$$\frac{\partial W}{\partial R} = \frac{R}{\sqrt{R^2 + \omega^2 L^2}}, \frac{\partial W}{\partial \omega} = \frac{\omega L^2}{\sqrt{R^2 + \omega^2 L^2}}, \frac{\partial W}{\partial L} = \frac{\omega^2 L}{\sqrt{R^2 + \omega^2 L^2}}.$$

Im Bereich $R > 0$ sind das stetige Funktionen, und deshalb ist W dort differenzierbar. Man hat

$$W(R + \Delta R, \omega + \Delta\omega, L + \Delta L) \approx \sqrt{R^2 + \omega^2 L^2} + \frac{R}{\sqrt{\cdots}}\Delta R + \frac{\omega L^2}{\sqrt{\cdots}}\Delta\omega + \frac{\omega^2 L}{\sqrt{\cdots}}\Delta L$$

oder

$$\operatorname{grad} W = \left(\frac{R}{\sqrt{R^2+\omega^2L^2}}, \frac{\omega L^2}{\sqrt{R^2+\omega^2L^2}}, \frac{\omega^2 L}{\sqrt{R^2+\omega^2L^2}}\right).$$

Beispiel 46 (Stationäre Strömungen, inkompressible Flüssigkeiten).
Verfahrentechnik I, Abschnitt 1.5.1
Eine stationäre Strömung im $\mathbb{R}^n$ läßt sich mathematisch beschreiben durch eine Abbildung

$$\vec{\Phi} : \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}^n, (t, \vec{x}) \mapsto \vec{\Phi}_t(\vec{x}).$$

Hält man $\vec{x}$ fest, so beschreibt $t \mapsto \vec{\Phi}_t(\vec{x})$ eine Kurve, die *Flusslinie*, auf der sich das strömende Teilchen bewegt, welches sich zur Zeit $t = 0$ an der Stelle $\vec{x}$ befindet. Hält man hingegen t fest, so ist

$$\vec{\Phi}_t : \vec{x} \mapsto \vec{\Phi}_t(\vec{x})$$

eine Abbildung des ganzen Raumes, die alle Teilchen um die Zeit t weiterrückt.
Man hat deshalb

$$\vec{\Phi}_0(\vec{x}) = \vec{x} \tag{13}$$

und

$$\vec{\Phi}_{s+t}(\vec{x}) = \vec{\Phi}_s(\vec{\Phi}_t(\vec{x})).$$

Aus der Linearen Algebra wissen wir, dass die Determinante einer linearen Abbildung (Matrix) die *Volumenverzerrung* durch diese Abbildung beschreibt. Für die differenzierbare Abbildung

$$\vec{\Phi}_t : \mathbb{R}^n \to \mathbb{R}^n$$

wird die *infinitesimale Volumenverzerrung* durch die Determinante der Ableitung

$$\det \vec{\Phi}'_t = \det\left(\frac{\partial(\Phi_t)_i}{\partial x_j}\right)$$

gegeben (vgl. auch Abschnitt 3.4). Das heißt, ein kleiner Bereich an der Stelle $\vec{x}$ wird durch die Strömung nach der Zeit t in einen kleinen Bereich an der Stelle $\vec{\Phi}_t(\vec{x})$ transportiert, dessen Volumen sich zum ursprünglichen verhält wie $\det \vec{\Phi}'_t(\vec{x}) : 1$.
Wenn uns die Volumenverzerrung entlang einer Flußlinie interessiert, müssen wir

$$\frac{\partial}{\partial t} \det \vec{\Phi}'_t(\vec{x})$$

berechnen. Wir machen das für den Fall $n = 2$ (zweidimensionales „Volumen“ = Fläche) und schreiben

$$\vec{\Phi}_t(x, y) = \begin{pmatrix}\phi_1(t, x, y)\\ \phi_2(t, x, y)\end{pmatrix}.$$

Dann erhalten wir

$$\det \vec{\Phi}'_t(x, y) = \begin{vmatrix}\frac{\partial\phi_1}{\partial x} & \frac{\partial\phi_1}{\partial y}\\ \frac{\partial\phi_2}{\partial x} & \frac{\partial\phi_2}{\partial y}\end{vmatrix} = \frac{\partial\phi_1}{\partial x}\frac{\partial\phi_2}{\partial y} - \frac{\partial\phi_2}{\partial x}\frac{\partial\phi_1}{\partial y}.$$

Das differenzieren wir nach der Zeit und benutzen dabei die Produktregel und – im Vorgriff auf Satz von Schwarz aus Abschnitt 1.13 – die Formel

$$\frac{\partial}{\partial t}\frac{\partial\phi_1}{\partial x} = \frac{\partial}{\partial x}\frac{\partial\phi_1}{\partial t} = \frac{\partial\dot{\phi}_1}{\partial x}$$

Wir erhalten

$$\frac{\partial}{\partial t}\det \vec{\Phi}_t'(x,y) = \frac{\partial \dot{\phi}_1}{\partial x}\frac{\partial \phi_2}{\partial y} + \frac{\partial \phi_1}{\partial x}\frac{\partial \dot{\phi}_2}{\partial y} - \left(\frac{\partial \dot{\phi}_2}{\partial x}\frac{\partial \phi_1}{\partial y} + \frac{\partial \phi_2}{\partial x}\frac{\partial \dot{\phi}_1}{\partial y}\right).$$

Wir betrachten das für $t = 0$. Wegen (13) ist

$$\vec{\Phi}_0(\vec{x}) = \begin{pmatrix}\phi_1(0,x,y)\\ \phi_2(0,x,y)\end{pmatrix} = \begin{pmatrix}x\\ y\end{pmatrix}$$

und deshalb

$$\frac{\partial \phi_1}{\partial x} = 1 = \frac{\partial \phi_2}{\partial y}, \quad \frac{\partial \phi_2}{\partial x} = 0 = \frac{\partial \phi_1}{\partial y}.$$

Also finden wir

$$\frac{\partial}{\partial t}\det \vec{\Phi}_t'(x,y)\Big|_{t=0} = \frac{\partial \dot{\phi}_1}{\partial x} + \frac{\partial \dot{\phi}_2}{\partial y}.$$

Das Vektorfeld

$$\vec{v}(\vec{x}) := \frac{\partial}{\partial t}\vec{\Phi}_t(x)\Big|_{t=0} = \begin{pmatrix}\dot{\phi}_1(0,x,y)\\ \dot{\phi}_2(0,x,y)\end{pmatrix}$$

ist aber gerade das Geschwindigkeitsfeld zu unserer Strömung $\vec{\Phi}$, und die Flächenverzerrung durch die Strömung an der Stelle (x, y) ist gegeben durch

$$\frac{\partial v_1}{\partial x}(x,y) + \frac{\partial v_2}{\partial y}(x,y).$$

Für Dimension $n = 3$ wird die Rechnung etwas komplizierter, ergibt aber die analoge Formel.

Wir halten das Ergebnis in einem Satz fest. Vgl. auch Abschnitt 2.1.

Satz 47. *Die Volumenverzerrung einer stationären Strömung im $\mathbb{R}^3$ berechnet sich aus ihrem Geschwindigkeitsfeld $\vec{v}$ mit der Formel*

$$\operatorname{div}\vec{v} := \frac{\partial v_1}{\partial x_1} + \frac{\partial v_2}{\partial x_2} + \frac{\partial v_3}{\partial x_3}.$$

Diesen Ausdruck nennt man die Divergenz *von $\vec{v}$.*
Divengenzfreie Strömungen ($\operatorname{div}\vec{v} = 0$*) sind also volumentreu (inkompressibel).*

1.10 Rechenregeln für die Differentiation

- Wir lernen insbesondere verschiedene Variationen der Produktregel und der Kettenregel kennen.

Die Rechenregeln für die Differentiation haben – bei einer wie bei mehreren Variablen – eigentlich *zwei* Aspekte. Sie besagen, dass

1. Summen, Vielfache, Produkte oder Kompositionen von differenzierbaren Funktionen wieder differenzierbar sind. Das erspart es uns, bei jeder neuen Funktion auf die Definition der Differenzierbarkeit zurückzugreifen und Grenzwerte zu berechnen. Weiter liefern sie
2. Formeln für die Berechnung der Ableitung von Summen etc. aus den Ableitungen der einzelnen Bausteine.

Die erste Aussage trifft unverändert für Funktionen mehrerer Variablen zu, die Formeln werden gelegentlich etwas komplizierter, weil die Ableitung nicht mehr ein Skalar, sondern eine Matrix ist.

Einfach ist es mit der Summe oder Vielfachen von differenzierbaren Funktionen

$$\vec{f}, \vec{g} : \mathbb{R}^n \supset G \to \mathbb{R}^m.$$

Dafür gilt

$$(\vec{f} + \vec{g})' = \vec{f}' + \vec{g}', (\lambda \vec{f})' = \lambda \vec{f}'.$$

Hier stehen also die Summe oder das skalare Vielfache von Matrizen.

Die **Produktregel** setzt voraus, dass für die Abbildungen ein Produkt definiert ist. Wir geben folgende Fälle an, in denen die aus der Differenzierbarkeit der beiden Faktoren auch die des Produktes folgt:

1. Wenigstens eine der beiden Abbildungen ist reellwertig. Dann ist das Produkt

   $$(f\vec{g})(\vec{x}) := f(\vec{x})\vec{g}(\vec{x})$$

   definiert.
2. $\vec{f}$ und $\vec{g}$ sind $\mathbb{R}^m$-wertig und „Produkt“ bedeutet „Skalarprodukt“:

   $$(\vec{f} \cdot \vec{g})(\vec{x}) := \vec{f}(\vec{x}) \cdot \vec{g}(\vec{x}).$$

3. $\vec{f}$ und $\vec{g}$ sind $\mathbb{R}^3$-wertig und „Produkt“ bedeutet „Vektorprodukt“:

   $$(\vec{f} \times \vec{g})(\vec{x}) := \vec{f}(\vec{x}) \times \vec{g}(\vec{x}).$$

Wir erinnern an die Definitionen des Skalarproduktes und des Vektorproduktes:

$$\begin{pmatrix} f_1 \\ \vdots \\ f_n \end{pmatrix} \cdot \begin{pmatrix} g_1 \\ \vdots \\ g_n \end{pmatrix} = f_1 g_1 + \ldots + f_n g_n$$

$$\begin{pmatrix} f_1 \\ \vdots \\ f_3 \end{pmatrix} \times \begin{pmatrix} g_1 \\ \vdots \\ g_3 \end{pmatrix} = \begin{pmatrix} f_2 g_3 - f_3 g_2 \\ f_3 g_1 - f_1 g_3 \\ f_1 g_2 - f_2 g_1 \end{pmatrix}.$$

Die Formel für die Differentiation der Produkte in Matrixschreibweise, also z.B. für $(\vec{f}\times\vec{g})'$ ist etwas kompliziert, und wir beschränken uns darauf, sie für die partiellen Ableitungen anzugeben:

Produkt mit reeller Funktion:

$$\boxed{\frac{\partial(f\vec{g})}{\partial x_j} = \frac{\partial f}{\partial x_j}\vec{g} + f\frac{\partial \vec{g}}{\partial x_j}}$$

Skalarprodukt:

$$\boxed{\frac{\partial(\vec{f}\cdot\vec{g})}{\partial x_j} = \frac{\partial \vec{f}}{\partial x_j}\cdot\vec{g} + \vec{f}\cdot\frac{\partial \vec{g}}{\partial x_j}}$$

Vektorprodukt:

$$\boxed{\frac{\partial(\vec{f}\times\vec{g})}{\partial x_j} = \frac{\partial \vec{f}}{\partial x_j}\times\vec{g} + \vec{f}\times\frac{\partial \vec{g}}{\partial x_j}.}$$

Beispiel 48.

$$\frac{\partial}{\partial y}\left(e^{xy}\begin{pmatrix}x+2y\\ xy\end{pmatrix}\right) = xe^{xy}\begin{pmatrix}x+2y\\ xy\end{pmatrix} + e^{xy}\begin{pmatrix}2\\ x\end{pmatrix} = e^{xy}\begin{pmatrix}x^2+2xy+2\\ x^2y+x\end{pmatrix}.$$

□

Rechenregeln für den Gradienten. Aus der Produktregel für die (partiellen) Ableitungen folgt

$$\operatorname{grad}_{\vec{x}}(fg) = g(\vec{x})\operatorname{grad}_{\vec{x}} f + f(\vec{x})\operatorname{grad}_{\vec{x}} g,$$

Sind $f:\mathbb{R}\to\mathbb{R}$ und $g:\mathbb{R}^n\to\mathbb{R}$, so ist nach der Kettenregel für Funktionen *einer* Veränderlichen $\frac{\partial(f\circ g)}{\partial x_j} = f'(g)\frac{\partial g}{\partial x_j}$, also

$$\operatorname{grad}_{\vec{x}}(f\circ g) = f'(g(\vec{x}))\operatorname{grad}_{\vec{x}} g.$$

Beispiel 49 (Coulombpotential).
Theoretische Elektrotechnik, Abschnitt 1.4
Das Potential einer Punktladung Q ist gegeben durch

$$U = \frac{Q}{4\pi\epsilon}\frac{1}{r}$$

mit

$$r = \sqrt{x^2+y^2+z^2}.$$

Es gilt

$$\operatorname{grad} r = \frac{1}{2\sqrt{\sum x_i^2}}\operatorname{grad}(\sum x_i^2) = \frac{1}{2r}(2x_1,\ldots,2x_n) = (\frac{x_1}{r},\ldots,\frac{x_n}{r}) = \frac{\vec{x}}{r},$$

und daher ist das elektrische Feld von U, also das negative Gradientenfeld, gegeben durch

$$\vec{E} = -\operatorname{grad} U = -\frac{Q}{4\pi\epsilon}\operatorname{grad}\frac{1}{r} = \frac{Q}{4\pi\epsilon}r^{-2}\operatorname{grad} r == \frac{Q}{4\pi\epsilon}\frac{\vec{x}}{r^3}.$$

Beispiel 50 (Dipolpotential).
Theoretische Elektrotechnik, Abschnitt 1.4
Das Potential eines Dipols im Nullpunkt ist gegeben durch

$$V(\vec{x}) = -\frac{1}{4\pi\epsilon}\frac{\vec{p}\cdot\vec{x}}{|\vec{x}|^3}.$$

Das Dipolmoment $\vec{p}$ gibt die Stärke und Ausrichtung des Dipols an. Nun ist

$$\mathrm{grad}_{\vec{x}}(\vec{p}\cdot\vec{x}) = \mathrm{grad}_{\vec{x}}(\sum p_i x_i) = \vec{p} \qquad \text{und} \qquad \mathrm{grad}_{\vec{x}}(\vec{x}\cdot\vec{x}) = 2\vec{x}.$$

Damit erhält man das elektrische Feld des Dipols als

$$\vec{E}(\vec{x}) = \mathrm{grad}_{\vec{x}}(-V) = \frac{1}{4\pi\epsilon}\left(\frac{\vec{p}}{|\vec{x}|^3} - \frac{3}{2}\frac{2(\vec{p}\cdot\vec{x})\vec{x}}{|\vec{x}|^5}\right) = \frac{1}{4\pi\epsilon}\left(\frac{\vec{p}}{|\vec{x}|^3} - 3\frac{(\vec{p}\cdot\vec{x})}{|\vec{x}|^5}\vec{x}\right).$$

Mit Mathematica berechnet sich das Feld eines Dipolpotentials folgendermaßen:

```
<< Calculus`VectorAnalysis`
In:=Grad[(p1 x + p2 y + p3 z)/(x^2 + y^2 + z^2)^(3/2), Cartesian[x,y,z]]
Out = {-(3 x (p1 x + p2 y + p3 z))/(x^2 + y^2 + z^2)^(5/2) + p1/(x^2 + y^2 + z^2)^(3/2),
  -(3 y (p1 x + p2 y + p3 z))/(x^2 + y^2 + z^2)^(5/2) + p2/(x^2 + y^2 + z^2)^(3/2),
  -(3 z (p1 x + p2 y + p3 z))/(x^2 + y^2 + z^2)^(5/2) + p3/(x^2 + y^2 + z^2)^(3/2)}
```

Dabei wurde zuerst das für den Gradienten benötigte Package *Calculus`VectorAnalysis`* eingelesen.

Die Kettenregel schreibt sich genauso wie bei Funktionen einer Variablen, wenn man die Matrixmultiplikation verwendet:

Satz 51 (Kettenregel mit Ableitungsmatrizen). *Hat man differenzierbare Abbildungen $\vec{g}:\mathbb{R}^n\to\mathbb{R}^m$ und $\vec{f}:\mathbb{R}^m\to\mathbb{R}^p$, so ist $\vec{f}\circ\vec{g} = \vec{f}(\vec{g}):\mathbb{R}^n\to\mathbb{R}^p$ auch differenzierbar, und es gilt*

$$(\vec{f}(\vec{g}))'(\vec{x}) = \vec{f}\,'(\vec{g}(\vec{x}))\,\vec{g}\,'(\vec{x}).$$

Dabei steht auf der rechten Seite das Produkt von zwei Matrizen,

Wir bezeichnen die Variablen im $\mathbb{R}^m$ mit $\vec{y} = (y_1,\ldots,y_m)$ und schreiben die Matrizen mit partiellen Ableitungen aus: $\vec{f}' = (\frac{\partial f_k}{\partial y_i})_{k,i}$ und $\vec{g}' = (\frac{\partial g_i}{\partial x_j})_{i,j}$. Dann hat man

Satz 52 (Kettenregel mit partiellen Ableitungen).

$$\frac{\partial(f_k\circ\vec{g})}{\partial x_j}(\vec{x}) = \sum_{i=1}^{m}\frac{\partial f_k}{\partial y_i}(\vec{g}(\vec{x}))\frac{\partial g_i}{\partial x_j}(\vec{x}). \tag{14}$$

Beispiel 53. Die Ladung q eines Kondensators der Kapazität C mit der Spannung u ist $q = Cu$. Im Fall, dass Spannung und Kapazität zeitabhängig sind, ist der Strom gegeben durch

$$i(t) = \frac{dq}{dt} = \frac{dC}{dt}u(t) + C(t)\frac{du}{dt} \quad \text{(Produktregel)}.$$

Ist aber die Kapazität auch noch von der Spannung abhängig, also $C = C(u,t)$, so berechnet sich der Strom mit der Kettenregel:

$$i(t) = \frac{dq}{dt} = \frac{dC}{dt}u(t) + C(t)\frac{du}{dt} = \left(\frac{\partial C}{\partial u}\frac{du}{dt} + \frac{\partial C}{\partial t}\right)u(t) + C(t)\frac{du}{dt}.$$

Beispiel 54. Seien $f(x,y) := \frac{-1}{\sqrt{x^2+y^2}}$ und $\vec{g}(\rho,\phi) = (\rho\cos\phi, \rho\sin\phi)$. Dann ist

$$\begin{aligned}\frac{\partial f(\vec{g})}{\partial \rho} &= \frac{\rho\cos\phi}{\sqrt{(\rho\cos\phi)^2 + (\rho\sin\phi)^2}^3}\cos\phi + \frac{\rho\sin\phi}{\sqrt{(\rho\cos\phi)^2 + (\rho\sin\phi)^2}^3}\sin\phi \\ &= \frac{\rho\cos^2\phi + \rho\sin^2\phi}{\sqrt{\rho^2}^3} = \frac{1}{\rho^2}.\end{aligned}$$

Natürlich kann man auch explizit ausrechnen, dass $f(\vec{g})(\rho,\phi) = \frac{-1}{\rho}$ und dann direkt differenzieren. □

Beispiel 55. Nach dem Cosinussatz gilt im Dreieck

$$c = \sqrt{a^2 + b^2 - 2ab\cos\gamma}.$$

Wie groß ist der Fehler Δc, wenn a, b und γ mit Fehlern $\Delta a, \Delta b$ und $\Delta\gamma$ gemessen sind? Es gilt

$$\begin{aligned}\Delta c &\approx \frac{1}{2\sqrt{a^2 + b^2 - 2ab\cos\gamma}}\left((2a - 2b\cos\gamma)\Delta a + (2b - 2a\cos\gamma)\Delta b + 2ab\sin\gamma\,\Delta\gamma\right) \\ &= \frac{1}{\sqrt{a^2 + b^2 - 2ab\cos\gamma}}\left((a - b\cos\gamma)\Delta a + (b - a\cos\gamma)\Delta b + ab\sin\gamma\,\Delta\gamma\right).\end{aligned}$$

Vergleiche hierzu auch den Abschnitt 1.12 über Fehlerapproximation und Fehlerschranken. □

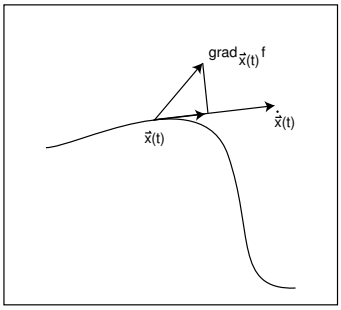
Beispiel 56. Seien $f : \mathbb{R}^n \to \mathbb{R}$ ein ortsabhängiges Potential und $\vec{x} : \mathbb{R} \to \mathbb{R}^n$, $t \mapsto \vec{x}(t)$ die Bahn eines Teilchens im Raum $\mathbb{R}^n$. Dann erhalten wir

$$\frac{d(f\circ\vec{x})}{dt}(t) = \frac{\partial f}{\partial x_1}(\vec{x}(t))\dot{x}_1(t) + \ldots + \frac{\partial f}{\partial x_n}(\vec{x}(t))\dot{x}_n(t) = (\operatorname{grad}_{\vec{x}(t)} f)\cdot\dot{\vec{x}}(t).$$

Dabei haben wir wie üblich die Ableitung nach t mit einem Punkt bezeichnet. Interpretiert man den Gradienten des Potentials f als Kraftfeld, so ist also die Änderung des Potentials längs der Bahnkurve gleich der Geschwindigkeit des Teilchens multipliziert mit der Komponente der Kraft in Richtung der Bahn. Beachten Sie

$$(\operatorname{grad} f)\cdot\dot{\vec{x}} = |\operatorname{grad} f||\dot{\vec{x}}|\cos\phi.$$

Verläuft die Bahnkurve $\vec{x}(t)$ in einem Niveau von f, so ist also $f(\vec{x}(t)) = const.$ und deshalb $(\operatorname{grad} f)\cdot\dot{\vec{x}} = 0$. Das ist der „wahre“ Beweis dafür, dass der Gradient auf den Niveaus senkrecht steht.


grad $_{\vec{x}(t)}$ f
$\dot{\vec{x}}$(t)
$\vec{x}$(t)

□

1.11 Koordinatensysteme

- Symmetrische Probleme lassen sich oft bequemer in anderen als den kartesischen Koordinaten beschreiben. Dann spielt bei der Differentialtion die Kettenregel eine wichtige Rolle.
- Wir untersuchen häufig gebrauchte Koordinatensysteme.

Sie finden häufig die Kettenregel (14) in der folgenden Form

$$\frac{\partial f_k}{\partial x_j} = \sum_{i=1}^{m} \frac{\partial f_k}{\partial y_i} \frac{\partial y_i}{\partial x_j}. \tag{15}$$

Das ist viel einfacher zu behalten als (14). Aber es bedarf einer Interpretation. Ich will dabei annehmen, dass $p = 1$ ist, so dass wir uns den Index k sparen können. Weiter nehmen wir der Einfachheit halber noch $m = n = 2$ an. Statt y_1, y_2 wollen wir gern x, y schreiben und statt x_1, x_2 benutzen wir u, v. Dann erhalten wir

$$\frac{\partial f}{\partial u} = \frac{\partial f}{\partial x} \frac{\partial x}{\partial u} + \frac{\partial f}{\partial y} \frac{\partial y}{\partial u} \tag{16}$$

und ebenso für v. In dieser Formel ist nicht gesagt, an welcher Stelle die Ableitungen zu nehmen sind, aber das macht man vermutlich automatisch richtig. Schwieriger ist die Frage, von welchen Variablen f eigentlich abhängt? Jedenfalls wird es nach x, y, u und eventuell auch nach v differenziert. Ist es eine Funktion von vier Variablen? Dieses Problem, das vor allem bei der Verwendung verschiedener Koordinatensysteme vorkommt, will ich jetzt klären.

Der **Ingenieur** stellt sich unter f eine physikalische Größe vor, zum Beispiel ein elektrisches Potential in der Ebene (weil $m = 2$). Das Potential hängt vom Ort (x, y) ab, und deshalb ist $f = f(x, y)$. Vielleicht ist f aber viel einfacher in Polarkoordinaten $(u, v) = (\rho, \phi)$ zu beschreiben. Dann ist für den Ingenieur eben $f = f(\rho, \phi)$. Wenn etwa $f(x, y) = \frac{1}{\sqrt{x^2+y^2}}$, dann ist $f(\rho, \phi) = \frac{1}{\rho}$. Weil das Potential als physikalische Größe unabhängig davon ist, welche Koordinaten wir zu seiner Beschreibung verwenden, benutzt der Ingenieur allemal denselben Buchstaben f, um es zu bezeichnen.

Für den **Mathematiker** ist das verwirrend: Als mathematische Funktion sieht f in beiden Fällen ganz verschieden aus. Was ist denn $f(4, 3)$? Ist es $\frac{1}{5}$ oder $\frac{1}{4}$. Sind $(4, 3)$ kartesische Koordinaten oder Polarkoordinaten? Ist $y = 3$ oder $\phi = 3$? – Der Mathematiker würde für die Funktionen verschiedene Symbole benutzen, zum Beispiel $\tilde{f}(\rho, \phi) = \frac{1}{\rho}$, und dann ist „mathematisch exakt“

$$\tilde{f}(\rho, \phi) = f(\rho \cos \phi, \rho \sin \phi).$$

Die Formel (16) würde er schreiben als

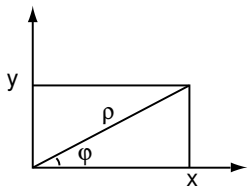
$$\frac{\partial \tilde{f}}{\partial \rho} = \frac{\partial f}{\partial x} \frac{\partial x}{\partial \rho} + \frac{\partial f}{\partial y} \frac{\partial y}{\partial \rho}.$$

Für den Mathematiker geht es also nicht um die physikalische Größe „Potential“, sondern um die mathematischen Modelle dafür, und die sind in Abhängigkeit von den verwendeten Koordinaten ganz verschieden.

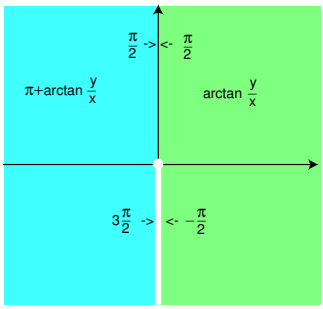
Die Formeln (15), (16) beruhen auf der „physikalischen Interpretation“. f ist eine Funktion der x-Variablen, aber diese sind wiederum Funktionen der u-Variablen - in der mathematischen Version ist $\vec{x} = \vec{g}$. Weil in (15) links eine Ableitung nach u_j steht, ist f als Funktion der u-Variablen zu interpretieren. Außerdem muss man die partiellen Ableitungen an den richtigen Stellen nehmen, die $\frac{\partial f}{\partial x_i}$ eben an der Stelle $\vec{x} = \vec{x}(\vec{u})$.

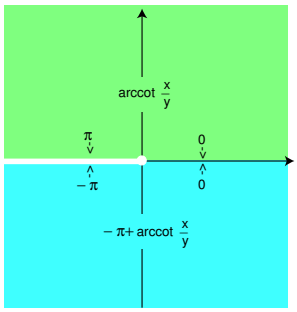
Wir stellen hier die gebräuchlichsten Koordinatensysteme zusammen und geben an, welche partiellen Ableitungen sich aus der Kettenregel ergeben. Dabei benutzen wir im Sinne der gerade gegebenen Erklärungen die „physikalische Interpretation“.

EBENE POLARKOORDINATEN. Die Position eines Punktes $\neq (0,0)$ in der Ebene wird beschrieben durch seinen Abstand ρ vom Ursprung und den Winkel ϕ zwischen dem Ortsvektor und der positiven x-Achse. Beachten Sie aber, dass man keine in der ganzen Ebene ohne den Ursprung stetige Winkelfunktion erklären kann: Läuft man einmal um den Ursprung herum, so ist der Winkel um 2π gewachsen. Der Winkel ist nur bis auf Vielfache von 2π eindeutig bestimmt.



Das beste was man bekommen kann, ist eine auf einem “Schlitzgebiet” stetige Winkelfunktion. Wir geben dafür zwei Beispiele:





Beachten Sie aber: Wenn man es nur mit $\sin\phi$ und $\cos\phi$ zu tun hat, macht die Mehrdeutigkeit nichts aus. Auch wenn es nicht um den Wert von ϕ, sondern, wie beim Differenzieren, um kleine Änderungen von ϕ geht, spielt die Mehrdeutigkeit keine Rolle. Darum sind zum Beispiel $\frac{\partial\phi}{\partial x}$ oder $\frac{\partial x}{\partial\phi}$ wohldefiniert.

Aus

$$x = \rho\cos\phi, \quad y = \rho\sin\phi. \tag{17}$$

erhalten wir

$$\frac{\partial x}{\partial\rho} = \cos\phi, \quad \frac{\partial x}{\partial\phi} = -\rho\sin\phi,$$
$$\frac{\partial y}{\partial\rho} = \sin\phi, \quad \frac{\partial y}{\partial\phi} = \rho\cos\phi,$$

und damit

$$\frac{\partial f}{\partial\rho} = \frac{\partial f}{\partial x}\frac{\partial x}{\partial\rho} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial\rho} = \frac{\partial f}{\partial x}\cos\phi + \frac{\partial f}{\partial y}\sin\phi$$
$$\frac{\partial f}{\partial\phi} = \frac{\partial f}{\partial x}\frac{\partial x}{\partial\phi} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial\phi} = -\rho\frac{\partial f}{\partial x}\sin\phi + \rho\frac{\partial f}{\partial y}\cos\phi$$

Beachten Sie, dass man diese Formeln anschaulich verstehen kann, wenn man sie in der Form

$$\frac{\partial f}{\partial\rho} = \operatorname{grad} f\cdot(\cos\phi,\sin\phi),$$
$$\frac{\partial f}{\partial\phi} = \operatorname{grad} f\cdot(-\rho\sin\phi,\rho\cos\phi),$$

schreibt. Die ρ-Ableitung ist die Richtungsableitung in radialer Richtung, die ϕ-Ableitung die in dazu senkrechter Richtung, wobei der Faktor ρ berücksichtigt, dass eine Änderung des Winkels ϕ en Punkt (x, y) proportional zu seinem Abstand von 0 bewegt.

Für die umgekehrte Richtung findet man aus $\rho = \sqrt{x^2 + y^2}$

$$\frac{\partial \rho}{\partial x} = \frac{x}{\rho}, \quad \frac{\partial \rho}{\partial y} = \frac{y}{\rho}.$$

Die Berechnung von $\frac{\partial \phi}{\partial x}$ kann man, ohne sich auf diverse arctan-Funktionen einzulassen, sehr elegant machen, indem man zum Beispiel die zweiten Gleichung (17) nach x differenziert. Man erhält

$$0 = \frac{\partial \rho}{\partial x} \sin\phi + \rho \cos\phi \frac{\partial \phi}{\partial x} = \frac{x}{\rho} \sin\phi + x \frac{\partial \phi}{\partial x} = \frac{x}{\rho^2} y + x \frac{\partial \phi}{\partial x}.$$

Auflösen nach $\frac{\partial \phi}{\partial x}$ und eine ähnliche Rechnung für die y-Ableitung liefern

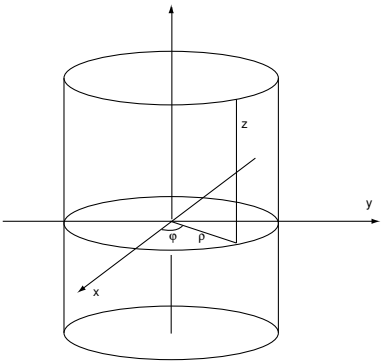
$$\frac{\partial \phi}{\partial x} = \frac{-y}{\rho^2}, \quad \frac{\partial \phi}{\partial y} = \frac{x}{\rho^2}.$$

Damit erhält man

$$\frac{\partial f}{\partial x} = \frac{x}{\rho} \frac{\partial f}{\partial \rho} - \frac{y}{\rho^2} \frac{\partial f}{\partial \phi}, \frac{\partial f}{\partial y} = \frac{y}{\rho} \frac{\partial f}{\partial \rho} + \frac{x}{\rho^2} \frac{\partial f}{\partial \phi}.$$

ZYLINDERKOORDINATEN. Zylinderkoordinaten in $\mathbb{R}^3$ sind ebene Polarkoordinaten, die durch die z-Koordinate ergänzt werden:

$$\begin{aligned} x &= \rho \cos\phi \qquad \rho = \sqrt{x^2 + y^2} \\ y &= \rho \sin\phi \\ z &= z \end{aligned}$$

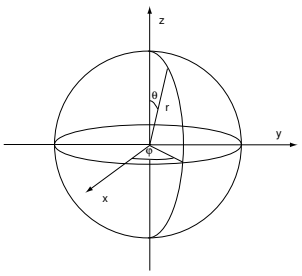


Entsprechend sind die Formeln für die partiellen Ableitungen nach diesen Koordinaten im wesentlichen dieselben wie für ebene Polarkoordinaten:

$$\begin{aligned} \frac{\partial f}{\partial \rho} &= \frac{\partial f}{\partial x} \cos\phi + \frac{\partial f}{\partial y} \sin\phi \\ \frac{\partial f}{\partial \phi} &= -\rho \frac{\partial f}{\partial x} \sin\phi + \rho \frac{\partial f}{\partial y} \cos\phi \\ \frac{\partial f}{\partial z} &= \frac{\partial f}{\partial z}. \end{aligned}$$

KUGELKOORDINATEN. In Kugelkoordinaten beschreibt man Punkte im $\mathbb{R}^3$ durch ihren Abstand r vom Ursprung und durch ihre "geographische Breite" θ und ihre "geographische Länge" ϕ. Dabei ist die Konvention hinsichtlich θ nicht einheitlich. Wir wollen θ als "Azimutwinkel" vom "Nordpol" aus messen. (Alternativ kann man θ auch wie die übliche geographische Breite von Äquator aus messen.)

$$\begin{aligned} x &= r\sin\theta\cos\phi \quad r = \sqrt{x^2+y^2+z^2} \\ y &= r\sin\theta\sin\phi \\ z &= r\cos\theta. \end{aligned}$$



Wie bei den ebenen Polarkoordinaten finden wir

$$\begin{aligned} \frac{\partial f}{\partial r} &= \frac{\partial f}{\partial x}\sin\theta\cos\phi + \frac{\partial f}{\partial y}\sin\theta\sin\phi + \frac{\partial f}{\partial z}\cos\theta \\ \frac{\partial f}{\partial \theta} &= \frac{\partial f}{\partial x} r\cos\theta\cos\phi + \frac{\partial f}{\partial y} r\cos\theta\sin\phi - \frac{\partial f}{\partial z} r\sin\theta \\ \frac{\partial f}{\partial \phi} &= -\frac{\partial f}{\partial x} r\sin\theta\sin\phi + \frac{\partial f}{\partial y} r\sin\theta\cos\phi. \end{aligned}$$

Können Sie das auch wie bei den ebenen Polarkoordinaten anschaulich interpretieren?
Für die partiellen Ableitungen nach x, y, z von in Polarkoordinaten gegebenen Funktionen finden wir sofort

$$\frac{\partial r}{\partial x} = \frac{x}{r}, \quad \frac{\partial r}{\partial y} = \frac{y}{r}, \quad \frac{\partial r}{\partial z} = \frac{z}{r}.$$

Differentiation von $z = r\cos\theta$ nach x liefert

$$0 = \frac{x}{r}\cos\theta - r\sin\theta\frac{\partial\theta}{\partial x} = \sin\theta\cos\phi\cos\theta - r\sin\theta\frac{\partial\theta}{\partial x}.$$

Auflösen nach $\frac{\partial\theta}{\partial x}$ und ähnliche Rechnungen für die y- und z-Ableitung liefern

$$\frac{\partial\theta}{\partial x} = \frac{\cos\theta\cos\phi}{r}, \quad \frac{\partial\theta}{\partial y} = \frac{\cos\theta\sin\phi}{r}, \quad \frac{\partial\theta}{\partial z} = \frac{-\sin\theta}{r}.$$

Nun differenzieren wir $y = r\sin\theta\sin\phi$ nach x:

$$\begin{aligned} 0 &= \frac{x}{r}\sin\theta\sin\phi + r\cos\theta\sin\phi\frac{\partial\theta}{\partial x} + r\sin\theta\cos\phi\frac{\partial\phi}{\partial x} \\ &= \sin^2\theta\cos\phi\sin\phi + \cos^2\theta\sin\phi\cos\phi + r\sin\theta\cos\phi\frac{\partial\phi}{\partial x} \\ &= \cos\phi\sin\phi + r\sin\theta\cos\phi\frac{\partial\phi}{\partial x} \end{aligned}$$

Auflösen nach $\frac{\partial\phi}{\partial x}$ und ähnliche Rechnungen für die y- und z-Ableitung liefern

$$\frac{\partial\phi}{\partial x} = \frac{-\sin\phi}{r\sin\theta}, \quad \frac{\partial\phi}{\partial y} = \frac{\cos\phi}{r\sin\theta}, \quad \frac{\partial\phi}{\partial z} = 0.$$

Beispiel 57. Wenn wir das Coulombpotential $f(x,y,z) = -\frac{1}{r}$ approximieren, so erhalten

wir bei Verwendung kartesischer Koordinaten

$$f(x+\Delta x, y+\Delta y, z+\Delta z) \approx f(x,y,z) + \frac{x}{r^3}\Delta x + \frac{y}{r^3}\Delta y + \frac{z}{r^3}\Delta z.$$

In Kugelkoordinaten erhält man

$$f(r+\Delta r, \theta+\Delta\theta, \phi+\Delta\phi) \approx f(r,\theta,\phi) + \frac{1}{r^2}\Delta r.$$

Nur Änderungen des Arguments in radialer Richtung machen sich bemerkbar, und zwar mit einem Proportionalitätsfaktor $1/r^2$. Aus der kartesichen Formel kann man das auch herauslesen, aber mit größerer Mühe. Können Sie's? □

Wir kommen auf die Differentialrechnung in anderen Koordinatensystemen im Kapitel „Vektoranalysis" noch einmal ausführlich zurück. Dafür stellen wir noch einmal unsere Ergebnisse zusammen. Die Ableitungen der kartesischen Koordinaten, also zum Beispiel $\frac{\partial x}{\partial r}$ sind allerdings so leicht zu finden, dass wir sie nicht noch einmal aufschreiben.

Polarkoordinaten	$x = \rho\cos\phi$ $y = \rho\sin\phi$	$\rho = \sqrt{x^2+y^2}$
$\frac{\partial\rho}{\partial x} = \frac{x}{\rho}$	$\frac{\partial\rho}{\partial y} = \frac{y}{\rho}$	
$\frac{\partial\phi}{\partial x} = \frac{-\sin\phi}{\rho}$	$\frac{\partial\phi}{\partial y} = \frac{\cos\phi}{\rho}$	

Zylinderkoordinaten	$x = \rho\cos\phi$ $y = \rho\sin\phi$ $z = z$	$\rho = \sqrt{x^2+y^2}$
$\frac{\partial\rho}{\partial x} = \frac{x}{\rho}$	$\frac{\partial\rho}{\partial y} = \frac{y}{\rho}$	$\frac{\partial\rho}{\partial z} = 0$
$\frac{\partial\phi}{\partial x} = \frac{-\sin\phi}{\rho}$	$\frac{\partial\phi}{\partial y} = \frac{\cos\phi}{\rho}$	$\frac{\partial\phi}{\partial z} = 0$
$\frac{\partial z}{\partial x} = 0$	$\frac{\partial z}{\partial y} = 0$	$\frac{\partial z}{\partial z} = 1$

Kugelkoordinaten	$x = r\sin\theta\cos\phi$ $y = r\sin\theta\sin\phi$ $z = r\cos\theta$	$r = \sqrt{x^2+y^2+z^2}$
$\frac{\partial r}{\partial x} = \frac{x}{r}$	$\frac{\partial r}{\partial y} = \frac{y}{r}$	$\frac{\partial r}{\partial z} = \frac{z}{r}$
$\frac{\partial\theta}{\partial x} = \frac{\cos\theta\cos\phi}{r}$	$\frac{\partial\theta}{\partial y} = \frac{\cos\theta\sin\phi}{r}$	$\frac{\partial\theta}{\partial z} = \frac{-\sin\theta}{r}$
$\frac{\partial\phi}{\partial x} = \frac{-\sin\phi}{r\sin\theta}$	$\frac{\partial\phi}{\partial y} = \frac{\cos\phi}{r\sin\theta}$	$\frac{\partial\phi}{\partial z} = 0$

1.12 Fehlerapproximation und Fehlerschranken

- Wir lernen eine wichtige Anwendung der Differentialrechnung kennen: Die Behandlung von (Mess-)fehlern mittels der Ableitung.

Wenn wir eine reellwertige Funktion f berechnen, deren Eingangsdaten x_i mit Fehlern (zum Beispiel Messfehlern) Δx_i behaftet sind, und wir wissen wollen, welchen Einfluss diese Fehler auf den Funktionswert haben, dann können wir natürlich die Approximation durch die Ableitung benutzen:

$$\Delta f \approx f'(\vec{x})\vec{\Delta x} = \sum \frac{\partial f}{\partial x_i}(\vec{x})\Delta x_i.$$

Beispiel 58 (Approximation des Fehlers). Wir betrachten die Funktion $W(R, L) := \sqrt{R^2 + \omega^2 L^2}$ mit fester Frequenz $\omega = 6000\,[sec^{-1}]$. Es sei $R = 20 \pm 0.1\,[\Omega]$ und $L = 5 \cdot 10^{-3} \pm 10^{-4}\,[H]$. Dann ist

$$\frac{\partial W}{\partial R} = \frac{R}{W}, \quad \frac{\partial W}{\partial L} = \frac{\omega^2 L}{W}$$

und

$$W = 36.06[\Omega],$$
$$\Delta W \approx \frac{\partial W}{\partial R}\Delta R + \frac{\partial W}{\partial L}\Delta L = \frac{20}{\sqrt{1300}}0.1[\Omega] + \frac{18 \cdot 10^4}{\sqrt{1300}}10^{-4}[\Omega] = 0.55[\Omega].$$

Das gibt als approximative Einschließung

$$35.51 \leq W \pm \Delta W \leq 36.61. \tag{18}$$

Beispiel 59 (Approximation des Relativen Fehlers). Wir setzen das vorangehende Beispiel fort: Für die relativen Fehler $\frac{\Delta W}{W}$, $\frac{\Delta R}{R}$ und $\frac{\Delta L}{L}$ ergibt sich (Division mit W und Erweitern mit R bzw. L)

$$\frac{\Delta W}{W} \approx \frac{R^2}{W^2}\frac{\Delta R}{R} + \frac{\omega^2 L^2}{W^2}\frac{\Delta L}{L} = \frac{4}{13}\frac{\Delta R}{R} + \frac{9}{13}\frac{\Delta L}{L}.$$

Diese Formel gibt Aufschluß darüber, welcher relative Messfehler in welchem Maße für den relativen Fehler im Ergebnis verantwortlich ist. Das kann für die Planung von Versuchsreihen eine wichtige Information sein. Und die bekommt man nicht durch „Rumrechnen“.

Die vorstehenen Fehlerabschätzungen benutzen die Formel $\Delta f \approx f'(x)\Delta\vec{x}$, und sind deshalb *Approximationen* für den Fehler, keine wirklich sicheren *Fehlerschranken*. Letztere kann man im hier betrachteten Beispiel aber relativ leicht direkt bekommen:

Beispiel 60 (Fehlerschranken durch Monotonie). Weil W in R und L offensichtlich monoton wachsend ist, liegt für R und L in den angegebenen Genauigkeitsintervallen der Wert von W im Intervall

$$W(R - \Delta R, L - \Delta L) \leq W + \Delta W \leq W(R + \Delta R, L + \Delta L),$$

also

$$35.50[\Omega] \leq W + \Delta W \leq 36.61[\Omega].$$

Das ist mit dem Rechner bequemer zu erreichen, als erst die partiellen Ableitungen zu bilden und auszuwerten, und es gibt verläßlichere Information: Der Fehler kann, wie man sieht, größer sein, als in der Approximation (18) angegeben.

Wenn man keine Monotonie wie in diesem Beispiel hat, muss natürlich keiner der Werte

$$f(x_1 \pm \Delta x_1, \ldots, x_n \pm \Delta x_n) - f(x_1, \ldots, x_n)$$

(mit allen möglichen Vorzeichenkombinationen) die maximale Abweichung realisieren. Und wenn die Feststellung der Monotonie aufwendig ist, ist die Approximation mit dem totalen Differential vielleicht einfacher. *Schranken* kann man dann mit dem folgenden Satz erhalten. Er präzisiert die sehr einleuchtende Tatsache, dass Schranken für die Ableitung (= Wachstums*rate*) einer Funktion Schranken für deren Wachstum liefern, und er gibt zuverlässige Fehlerschranken.

Satz 61 (Fehlerschrankensatz). *Seien $f : \mathbb{R}^n \supset G \to \mathbb{R}$ differenzierbar und G offen. Seien $\vec{x}, \vec{y} \in G$. Die Verbindungsstrecke S zwischen $\vec{x}$ und $\vec{y}$ liege ganz in G. Seien $M_1, \ldots, M_n$ reelle Zahlen, so dass für alle $\vec{z} \in S$*

$$\left| \frac{\partial f}{\partial x_i}(\vec{z}) \right| \leq M_i.$$

Dann gilt

$$|f(\vec{y}) - f(\vec{x})| \leq \sum_i M_i |y_i - x_i|.$$

Im allgemeinen Fall ist die Voraussetzung über die Verbindungsstrecke von selbst erfüllt, wenn G *konvex* ist:

Definition 62 (Konvex). Eine Teilmenge $G \subset \mathbb{R}^n$ heißt *konvex*, wenn sie mit je zwei Punkten auch deren Verbindungsstrecke enthält: Für $\vec{x}, \vec{y} \in G$ gilt

$$S = \{(1-t)\vec{x} + t\vec{y} \mid 0 \leq t \leq 1\} \subset G.$$

Zum Beispiel sind Kreise, Kugeln, Ellipsen oder Ellipsoide konvex.

Beweis des Fehlerschrankensatzes. Wir setzen $h(t) := f((1-t)\vec{x} + t\vec{y})$ für $0 \leq t \leq 1$. Dann ist h nach der Kettenregel differenzierbar und $\dot{h}(t) = \sum \frac{\partial f}{\partial x_i}((1-t)\vec{x} + t\vec{y})(-x_i + y_i)$. Daher erhalten wir für ein geeignetes t zwischen 0 und 1

$$\begin{aligned} |f(\vec{y}) - f(\vec{x})| &= |h(1) - h(0)| = |\dot{h}(t)| \\ &\leq \sum |\frac{\partial f}{\partial x_i}((1-t)\vec{x} + t\vec{y})||y_i - x_i| \leq \sum M_i |y_i - x_i|. \end{aligned}$$

□

Sind also die Größen x_i mit Messfehlern Δx_i behaftet, so ergibt sich für den resultierenden Fehler von f:

$$\boxed{|\Delta f| := |f(x_1 + \Delta x_1, \ldots, x_n + \Delta x_n) - f(x_1, \ldots, x_n)| \leq \sum M_i |\Delta x_i|.}$$

Beispiel 63 (Fehlerschranken mit dem Satz). Für das oben betrachtete Beispiel ist etwa

$$\left|\frac{\partial W}{\partial R}\right| = \left|\frac{R}{\sqrt{R^2+\omega^2L^2}}\right| \leq \left|\frac{R}{R}\right| = 1, \quad \left|\frac{\partial W}{\partial L}\right| = \left|\frac{\omega^2 L}{\sqrt{R^2+\omega^2L^2}}\right| \leq \omega \left|\frac{\omega L}{\omega L}\right| = \omega.$$

Damit erhalten wir als grobe, aber **zuverlässige** Abschätzung

$$|\Delta W| \leq |\Delta R| + \omega|\Delta L| \leq 0.1\,[\Omega] + 6000 \cdot 10^{-4}\,[Hsec^{-1}] = 0.7\,[\Omega].$$

1.13 Höhere Ableitungen. Extremwerte

- Wenn die lineare Approximation einer Funktion nicht ausreicht, kann man höhere Ableitungen heranziehen. Mittels zweiter Ableitungen kann man den Funktionsgraphen nicht nur durch seine Tangentialebene, sondern durch eine Fläche wie ein Paraboloid oder Hyperboloid approximieren.
- Dazu lernen wir den Satz von Taylor in mehreren Variablen kennen.
- Eine wichtige Anwendung sind Extremalkriterien.

Definition 64. Sei $f : \mathbb{R}^n \supset G \to \mathbb{R}$ eine Funktion, deren sämtliche partielle Ableitungen $\frac{\partial f}{\partial x_j}$ in G existieren und selber differenzierbar sind. Dann nennt man die Funktionen

$$\frac{\partial^2 f}{\partial x_i \partial x_j} := \frac{\partial(\frac{\partial f}{\partial x_j})}{\partial x_i}$$

die *zweiten partiellen Ableitungen* von f. Sind diese Funktionen stetig, so nennt man f zweimal stetig differenzierbar. In diesem Fall gilt der *Satz von Schwarz:*

$$\boxed{\frac{\partial^2 f}{\partial x_i \partial x_j} = \frac{\partial^2 f}{\partial x_j \partial x_i}.}$$

Er vereinfacht die Berechnung der (sämtlichen) zweiten partiellen Ableitungen um fast die Hälfte.

Entsprechend definiert man höhere als zweite partielle Ableitungen. Höhere partielle Ableitungen von vektor*wertigen* Funktionen definiert man komponentenweise, aber wir beschränken uns in diesem Abschnitt auf reellwertige Funktionen, insbesondere deshalb weil wir Extremwerte betrachten wollen, und die machen für vektorwertige Funktionen keinen Sinn. Wie im Fall einer Variablen, kann man die höheren Ableitungen benutzen, um eine (entsprechend oft differenzierbare) Funktion durch ein „Polynom“ besser als durch eine lineare Funktion zu approximieren. Es gibt auch hier eine Taylorformel, die wir aber nur für die zweite Ordnung notieren:

Satz 65 (Taylorformel). *Sei $f : \mathbb{R}^n \supset G \to \mathbb{R}$ zweimal stetig partiell differenzierbar. Die Verbindungssstrecke von $\vec{x}$ und $\vec{x} + \vec{\Delta x}$ liege in G. Dann gilt*

$$f(\vec{x} + \vec{\Delta x}) = f(\vec{x}) + \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(\vec{x})\Delta x_i + \frac{1}{2}\sum_{i,j=1}^{n} \frac{\partial^2 f}{\partial x_i \partial x_j}(\vec{x} + t\vec{\Delta x})\Delta x_i \Delta x_j \qquad (19)$$

für ein geeignetes $t \in]0, 1[$. Weiter gilt

$$f(\vec{x} + \vec{\Delta x}) = f(\vec{x}) + \sum_{i} \frac{\partial f}{\partial x_i}(\vec{x})\Delta x_i + \frac{1}{2}\sum_{i,j} \frac{\partial^2 f}{\partial x_i \partial x_j}(\vec{x})\Delta x_i \Delta x_j + Fehler, \qquad (20)$$

wobei

$$\lim_{\vec{\Delta x}\to 0} \frac{Fehler}{|\vec{\Delta x}|^2} = 0.$$

Man beweist den Satz durch Rückführung auf den eindimensionalen Fall, aber wir gehen darauf nicht ein. Wir gehen auch nicht auf konkrete Approximationen von Funktionen mit

Hilfe des Satzes von Taylor ein. Dass aber die Approximation (20) einem ein „anschauliches“ Bild vom Funktionsverlauf in der Nähe eines Punktes vermittelt, ist sehr wichtig zum Beispiel für Extremwertuntersuchungen.

Sind etwa die *ersten* partiellen Ableitungen an der Stelle $\vec{x}$ alle 0, sind weiter die „gemischten“ zweiten Ableitungen auch $= 0$ und die „doppelten“ zweiten, also die $\frac{\partial^2 f}{(\partial x_i)^2}$, an dieser Stelle alle positiv, so sieht man aus

$$f(\vec{x} + \vec{\Delta x}) \approx f(\vec{x}) + \underbrace{\frac{1}{2} \sum_i \frac{\partial^2 f}{(\partial x_i)^2}(\vec{x})(\Delta x_i)^2}_{\geq 0}, \tag{21}$$

dass f an der Stelle $\vec{x}$ ein *lokales Minimum* hat. Dem gehen wir nun genauer nach.

Sei also $f : \mathbb{R}^n \supset G \to \mathbb{R}$ eine differenzierbare Funktion auf einer OFFENEN Menge G und $\vec{x} \in G$. Ist $\operatorname{grad}_{\vec{x}} f \neq 0$, also etwa $\frac{\partial f}{\partial x_i}(\vec{x}) \neq 0$, und wählt man $\vec{\Delta x} = (0, \ldots, \Delta x_i, \ldots, 0)$, so gilt

$$f(\vec{x} + \vec{\Delta x}) \approx f(\vec{x}) + \frac{\partial f}{\partial x_i}(\vec{x}) \Delta x_i.$$

Indem man Δx_i mit verschiedenem Vorzeichen versieht, kann man also $f(\vec{x} + \vec{\Delta x}) > f(\vec{x})$ und $f(\vec{x} + \vec{\Delta x}) < f(\vec{x})$ für beliebig kleine $|\Delta x_i|$ erzielen: Beliebig nah bei $\vec{x}$ nimmt f größere und kleiner Werte an als an der Stelle selbst. Also hat f in $\vec{x}$ kein Extremum. Das liefert:

Satz 66 (Lokale Extrema: Notwendiges Kriterium). *Notwendig dafür, dass f im inneren Punkt $\vec{x}$ ein lokales Maximum oder Minimum hat, ist*

$$\operatorname{grad}_{\vec{x}} f = \vec{0}.$$

Man nennt Punkte $\vec{x}$, in denen der Gradient verschwindet auch kritische Punkte *von f.*

Für $n = 1$ ist das einfach die bekannte Bedingung $f'(x) = 0$. Aus der Taylorformel folgt weiter sofort:

Satz 67 (Lokale Extrema: Hinreichendes Kriterium). *Sei $f : G \to \mathbb{R}$ zweimal stetig partiell differenzierbar. Sei $\vec{x} \in G$ und*

$$\operatorname{grad}_{\vec{x}} f = \vec{0}.$$

Wir betrachten die sogenannte Hesseform

$$\operatorname{hess}_{\vec{x}} f(\vec{u}) := \sum_{i,j} \frac{\partial^2 f}{\partial x_i \partial x_j}(\vec{x}) u_i u_j \quad \vec{u} \in \mathbb{R}^n.$$

(i) Ist $\operatorname{hess}_{\vec{x}} f(\vec{u}) > 0$ für alle $\vec{u} \neq \vec{0}$, so hat f in $\vec{x}$ ein striktes lokales Minimum.

(ii) Ist $\operatorname{hess}_{\vec{x}} f(\vec{u}) < 0$ für alle $\vec{u} \neq \vec{0}$, so hat f in $\vec{x}$ ein striktes lokales Maximum.

(iii) Wechselt $\operatorname{hess}_{\vec{x}} f(\vec{u})$ für verschiedene $\vec{u}$ das Vorzeichen, so hat f in $\vec{x}$ kein *lokales Extremum.*

(iv) Hat man nur $\operatorname{hess}_{\vec{x}} f(\vec{u}) \geq 0$ oder $\operatorname{hess}_{\vec{x}} f(\vec{u}) \leq 0$ für alle $\vec{u} \neq 0$, so gibt das Kriterium keine Auskunft.

Man nennt die Hesseform im Fall (i) positiv definit*, im Fall (ii)* negativ definit*, im Fall (iii)* indefinit *und in den Fällen (iv)* positiv bzw. negativ semidefinit*,*

In gewissem Sinne ersetzt also die Hesseform $\operatorname{hess}_{\vec{x}} f$ oder die aus allen zweiten partiellen Ableitungen gebildete sogenannte *Hessematrix*

$$f''(\vec{x}) := \begin{pmatrix} \frac{\partial^2 f}{\partial x_1^2} & \cdots & \frac{\partial^2 f}{\partial x_1 x_n} \\ \vdots & & \vdots \\ \frac{\partial^2 f}{\partial x_n x_1} & \cdots & \frac{\partial^2 f}{\partial x_n^2} \end{pmatrix}$$

die zweite Ableitung im Fall einer Variablen, wobei wir dann nach dem Satz von Taylor schreiben dürfen

$$f(\vec{x} + \vec{\Delta x}) = f(\vec{x}) + f'(\vec{x})\vec{\Delta x} + \frac{1}{2}(\Delta x_1, \ldots, \Delta x_n) \begin{pmatrix} \frac{\partial^2 f}{\partial x_1^2} & \cdots & \frac{\partial^2 f}{\partial x_1 x_n} \\ \vdots & & \vdots \\ \frac{\partial^2 f}{\partial x_n x_1} & \cdots & \frac{\partial^2 f}{\partial x_n^2} \end{pmatrix} \begin{pmatrix} \Delta x_1 \\ \vdots \\ \Delta x_n \end{pmatrix} + \text{Fehler}.$$

Die positive oder negative Definitheit der Hesseform ist im allgemeinen nicht leicht nachzuprüfen. In der linearen Algebra wird gezeigt, dass sie gleichbedeutend damit ist, dass alle Eigenwerte der Hessematrix positiv bzw. negativ sind.

Wir untersuchen hier nur den Fall $n = 2$ genauer. Wir schreiben

$$\begin{pmatrix} \frac{\partial^2 f}{\partial x^2}(\vec{x}) & \frac{\partial^2 f}{\partial x \partial y}(\vec{x}) \\ \frac{\partial^2 f}{\partial x \partial y}(\vec{x}) & \frac{\partial^2 f}{\partial y^2}(\vec{x}) \end{pmatrix} = \begin{pmatrix} A & B \\ B & C \end{pmatrix}$$

und $\vec{u} = (v, w)$. Dann ist

$$\operatorname{hess}_{\vec{x}} f(\vec{u}) = Av^2 + 2Bvw + Cw^2.$$

Wir fragen, wann das für alle $\vec{u} = (v, w) \neq (0, 0)$ positiv ist. Setzen wir $w = 0$, so sehen wir, dass $A > 0$ sein muss. Ist andrerseits $w \neq 0$, so schreiben wir

$$Av^2 + 2Bvw + Cw^2 = \left(A(\frac{v}{w})^2 + 2B\frac{v}{w} + C\right) w^2 \underset{t:=\frac{v}{w}}{=} \left(At^2 + 2Bt + C\right) w^2.$$

Die Klammer ganz rechts beschreibt aber eine (wegen $A > 0$) nach oben offene Parabel mit Scheitel an der Stelle $2At + 2B = 0$, also $t = -\frac{B}{A}$. Dort ist der Funktionswert

$$A(\frac{B}{A})^2 + 2B(-\frac{B}{A}) + C = \frac{-B^2 + AC}{A}.$$

Also ist die Hesseform genau dann *positiv definit*, wenn $A > 0$ und $AC - B^2 > 0$. Der Ausdruck

$$AC - B^2 = \frac{\partial^2 f}{\partial x^2}\frac{\partial^2 f}{\partial y^2} - \left(\frac{\partial^2 f}{\partial x \partial y}\right)^2$$

heißt die Determinante der Hessematrix.

Eine analoge Untersuchung ergibt, dass die Hesseform genau dann *negativ definit* ist, wenn $A < 0$, die Determinante aber wieder positiv ist. Ist die Determinante negativ, so ist die Hesseform indefinit.

Wir fassen das zusammen:

Satz 68 (Hinreichendes Kriterium für $n = 2$). *Sei $f : \mathbb{R}^2 \supset G \to \mathbb{R}$ auf der offenen Menge G zweimal stetig differenzierbar. Sei $\vec{x} \in G$ und*

$$\frac{\partial f}{\partial x}(\vec{x}) = \frac{\partial f}{\partial y}(\vec{x}) = 0.$$

Im folgenden sind alle Ableitungen an der Stelle $\vec{x}$ zu nehmen. Es gilt:

$$\frac{\partial^2 f}{\partial x^2}\frac{\partial^2 f}{\partial y^2} - \left(\frac{\partial^2 f}{\partial x \partial y}\right)^2 < 0 \qquad \Longrightarrow \textit{ kein lokales Extremum}$$

$$\frac{\partial^2 f}{\partial x^2}\frac{\partial^2 f}{\partial y^2} - \left(\frac{\partial^2 f}{\partial x \partial y}\right)^2 > 0 \textit{ und } \frac{\partial^2 f}{\partial x^2} > 0 \qquad \Longrightarrow \textit{lokales Minimum}$$

$$\frac{\partial^2 f}{\partial x^2}\frac{\partial^2 f}{\partial y^2} - \left(\frac{\partial^2 f}{\partial x \partial y}\right)^2 > 0 \textit{ und } \frac{\partial^2 f}{\partial x^2} < 0 \qquad \Longrightarrow \textit{lokales Maximum}$$

Ist die Determinante der Hessematrix $= 0$, so gibt das Kriterium keine Auskunft.

Übrigens ist bei $\frac{\partial^2 f}{\partial x^2}\frac{\partial^2 f}{\partial y^2} - \left(\frac{\partial^2 f}{\partial x \partial y}\right)^2 > 0$ die Bedingung $\frac{\partial^2 f}{\partial x^2} > 0$ gleichbedeutend mit $\frac{\partial^2 f}{\partial y^2} > 0$, so dass die letzten beiden Bedingungen des Satzes nicht unsymmetrisch in x und y sind.

Die Untersuchung einer Funktion auf Extremwerte ist keine Routineangelegenheit. Die Tests mittels Differentialrechnung geben oft nur unvollständige Auskunft, und man muss sich im konkreten Einzelfall „etwas einfallen lassen“, wie die folgenden Beispiele zeigen.

Beispiel 69. Berechne das Minimum von

$$f(x, y) := x^2 + (2 - x)^3 y^2$$

auf der Kreisscheibe

$$D: x^2 + y^2 \leq 1.$$

Zunächst ist f eine stetige Funktion auf der kompakten Menge D. Daher nimmt f sein Minimum tatsächlich irgendwo auf D an. Wir müssen das Innere und die Randpunkte separat untersuchen. Wir beginnen mit dem Inneren. Wir finden

$$\frac{\partial f}{\partial x} = 2x - 3(2-x)^2y^2, \quad \frac{\partial f}{\partial y} = 2(2-x)^3y.$$

Aus $\frac{\partial f}{\partial y} = 0$ folgt $x = 2$ und dann $\frac{\partial f}{\partial x} = 4 \neq 0$, oder $y = 0$ und dann $\frac{\partial f}{\partial x} = 2x$. Also ist $(x, y) = (0, 0)$ der einzige Punkt, in dem $\operatorname{grad} f = 0$. Also liegt das Minimum in $(0, 0)$, wo der Funktionswert $f(0, 0) = 0$ ist, oder auf dem Rand. Auf dem Rand ist aber $x^2 + y^2 = 1$ und deshalb

$$f(x, y) = x^2 + \underbrace{(2-x)^3}_{\geq 1} y^2 \geq x^2 + y^2 = 1.$$

Also ist das Minimum 0; es wird (nur) in $(0, 0)$ angenommen. Man kann „sicherheitshalber“ noch die Hessematrix in $(0, 0)$ berechnen. Sie ist

$$\begin{pmatrix} 2 + 6(2-x)y^2 & -6(2-x)^2y \\ -6(2-x)^2y & 2(2-x)^3 \end{pmatrix}_{(0,0)} = \begin{pmatrix} 2 & 0 \\ 0 & 16 \end{pmatrix}.$$

Das ist offenbar positiv definit, und deshalb liegt in $(0, 0)$ ein lokales Minimum vor. Wir wussten aber schon, dass dort das globale Minimum liegt.

Sucht man auch das Maximum auf der Kreisscheibe so muss dieses auf dem Rand $x^2+y^2 = 1$ liegen. Für $x = \pm 1$ ist $y = 0$ und $f(\pm 1, 0) = 1$. Differenziert man $x^2 + (2-x)^3(1-x^2)$ und bestimmt numerisch die Nullstellen im Intervall $[-1, +1]$, so findet man nur eine, und zwar bei $x = -0.48972$. Weil f in y quadratisch ist, liegt das Maximum also bei $(-0.48972, \pm 0.760174)$ und hat den Wert $f_{max} = 11.9716$. □

Beispiel 70. Berechne das Minimum von

$$f(x, y) := x^2 + (2-x)^3y^2,$$

aber diesmal auf ganz $\mathbb{R}^2$. Man findet, dass auch auf $\mathbb{R}^2$ nur in $(0, 0)$ beide partiellen Ableitungen verschwinden. Dort ist ein lokales Minimum. Daraus kann man aber *nicht* schließen, dass dies auch das globale Minimum ist. Die Funktion ist ein ungerades Polynom in x, und deshalb gilt wie oben

$$\sup f = +\infty, \quad \inf f = -\infty.$$

Das Minimum wird also gar nicht angenommen. □

Beispiel 71. Gesucht sind die lokalen und globalen Extrema der Funktion

$$f(x, y) = x^3y^2(1 - x - y)$$

auf $\mathbb{R}^2$. Randpunkte gibt es also nicht. In y ist f ein Polynom dritten Grades. Wählt man $x > 0$ beliebig, so ist

$$\lim_{y \to \pm\infty} f(x, y) = \mp\infty.$$

Insbesondere ist also

$$\sup f(x, y) = \infty, \quad \inf f(x, y) = -\infty$$

und f nimmt kein *globales* Maximum oder Minimum an. Für den Gradient findet man:

$$\operatorname{grad} f(x,y) = \begin{pmatrix} 3x^2y^2(1-x-y) - x^3y^2 \\ 2x^3y(1-x-y) - x^3y^2 \end{pmatrix} = x^2y \begin{pmatrix} 3y - 4xy - 3y^2 \\ 2x - 2x^2 - 3xy \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

genau dann, wenn

$$x^2y = 0,$$

d.h. wenn (x,y) auf einer der beiden Achsen liegt, oder wenn

$$\begin{aligned} 3 - 4x - 3y &= 0 \\ 2 - 2x - 3y &= 0, \end{aligned}$$

d.h. wenn

$$(x,y) = (\frac{1}{2}, \frac{1}{3}).$$

Wir untersuchen nun zunächst, ob im letzteren Punkt wirklich ein lokales Extremum vorliegt.

- Die Hessematrix in diesem Punkt ergibt sich nach kurzer Rechnung als

$$\begin{pmatrix} \frac{\partial^2 f}{\partial x^2} & \frac{\partial^2 f}{\partial x \partial y} \\ \frac{\partial^2 f}{\partial x \partial y} & \frac{\partial^2 f}{\partial y^2} \end{pmatrix} = \begin{pmatrix} -\frac{1}{9} & -\frac{1}{12} \\ -\frac{1}{12} & -\frac{1}{8} \end{pmatrix}.$$

  Die Determinante ist

$$\frac{\partial^2 f}{\partial x^2}\frac{\partial^2 f}{\partial y^2}|_{(\frac{1}{2},\frac{1}{3})} - (\frac{\partial^2 f}{\partial x \partial y})|^2_{(\frac{1}{2},\frac{1}{3})} = \frac{1}{72} - \frac{1}{144} > 0$$

  und deshalb liegt an dieser Stelle ein lokales Extremum vor, und zwar ein lokales Maximum, weil

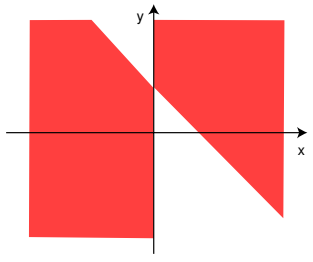
$$\frac{\partial^2 f}{\partial x^2}|_{(\frac{1}{2},\frac{1}{3})} = -\frac{1}{9} < 0.$$

  Der Funktionswert ist $f(\frac{1}{2}, \frac{1}{3}) = \frac{1}{432}$.

- Die Entscheidung, ob bei $(x,y) = (\frac{1}{2}, \frac{1}{3})$ ein Maximum oder Minimum vorliegt, kann man auch ohne Hessematrix treffen. Vergleiche dazu die nachstehende Abbildung und Beispiel 16. Auf dem kompakten Dreieck mit den Endpunkten $(0,0), (1,0)$ und $(0,1)$ nimmt die stetige Funktion f ihr Maximum an. Auf dem Rand des Dreiecks ist sie $= 0$ und im Innern positiv. Also liegt das Maximum irgendwo im Innern, und der einzige Kandidat ist $(x,y) = (\frac{1}{2}, \frac{1}{3})$.

Untersuchung der Achsen. Andere Kandidaten für lokale Extremwerte sind nur noch die Punkte auf den beiden Achsen. Der Funktionswert dort ist 0. Die Determinante der Hessematrix ist auf den Achsen $= 0$, so dass das Kriterium mit der zweiten Ableitung versagt.

Nach Beispiel 16 hat man nebenstehende Verteilung der Funktionswerte von f. In den farbigen Bereichen ist f negativ. In der Nähe eines jeden Punktes auf der y-Achse gibt es sowohl Punkte mit negativem wie solche mit positivem Funktionswert, also liegen dort keine lokalen Extremalpunkte. Ein Gleiches gilt für die beiden Punkte $(0,0)$ und $(0,1)$ auf der x-Achse.



In allen anderen Punkten der x-Achse liegen hingegen lokale Extrema. Wie die obige Skizze zeigt, hat f in den Punkten

$$\{(x,0) \mid 0 < x < 1\}$$

ein (schwaches) lokales Minimum (in allen Nachbarpunkten ist $f \geq 0$), und in den Punkten

$$\{(x,0) \,|\, x > 1 \text{ oder } x < 0\}$$

liegen (schwache) lokale Maxima. □

1.14 Extrema mit Nebenbedingungen

- Wir suchen Extremwerte, wenn sich das Argument der Funktion nicht frei im Raum bewegen kann, sondern auf eine Fläche eingeschränkt ist.

Wo nimmt die Funktion $f(x,y,z) = xyz$ auf der abgeschlossenen Kugel

$$B = \{(x,y,z) \,|\, x^2 + y^2 + z^2 \leq 1\}$$

ihr Maximum an? Weil f stetig und B kompakt ist, wird dieses Maximum wirklich irgendwo angenommen, die Frage macht also Sinn.

- Wir untersuchen zunächst die Punkte im Inneren von B, wo $\operatorname{grad} f = 0$:

$$\frac{\partial f}{\partial x} = yz = 0, \quad \frac{\partial f}{\partial y} = xz = 0, \quad \frac{\partial f}{\partial z} = xy = 0.$$

  Diese Gleichungen sind genau dort erfüllt, wo zwei der drei Koordinaten $= 0$ sind, und das sind genau die Koordinatenachsen. Dort ist dann aber $f = 0$, und da f offenbar in B auch positive Werte annimmt, wird hier gewiß nicht das Maximum angenommen. (Auch nicht das Minimum.)

- Also wird das Maximum auf dem Rand von B angenommen. Es bleibt die Aufgabe, auf dem Rand von B, also auf der Sphäre $x^2 + y^2 + z^2 = 1$, das Maximum von f zu finden. Wir haben damit – allgemeiner formuliert – das folgende Problem: Wir suchen ein Extremum einer Funktion $f = f(x,y,z)$ auf einer Fläche im $\mathbb{R}^3$. Diese Fläche ist gegeben als Niveaufläche einer Funktion g, nämlich in unserem Fall durch die Gleichung

$$g(x,y,z) = x^2 + y^2 + z^2 - 1 = 0.$$

  (Beachten Sie: Wir haben eine gewisse Normierung vorgenommen, indem wir die 1 auf die linke Seite gebracht haben. Die Fläche ist damit das 0-Niveau unserer Funktion g.)

Man sagt auch:

> Man sucht ein Extremum der Funktion f
> *unter der Nebenbedingung* $g = 0$.

Geometrisch bedeutet unser Problem: Man finde ein achsenparalleles Quader größten Inhalts unter der Nebenbedingung, dass die Ecken auf einer Sphäre vom Radius 1 liegen. Ist (x,y,z) eine solche Ecke, so ist $f(x,y,z) = xyz$ gerade ein Achtel des Volumens.

1. Lösungsansatz. Man löst die Gleichung $g(x,y,z) = 0$ zum Beispiel nach z auf, in unserem Fall

$$z = \sqrt{1 - x^2 - y^2},$$

und bestimmt dann die Extremalstellen von $\tilde{f}(x,y) = f(x,y,z(x,y))$, in unserem Fall also von

$$\tilde{f}(x,y) = xy\sqrt{1 - x^2 - y^2}.$$

Man findet

$$\frac{\partial \tilde{f}}{\partial x} = y\sqrt{1 - x^2 - y^2} - \frac{x^2 y}{\sqrt{1 - x^2 - y^2}} = 0,$$

$$\frac{\partial \tilde{f}}{\partial y} = x\sqrt{1 - x^2 - y^2} - \frac{xy^2}{\sqrt{1 - x^2 - y^2}} = 0.$$

Wir können wieder die Fälle $x = 0$ oder $y = 0$ ausschließen, weil hier $\tilde{f} = 0$. Dann folgt aber $(1 - x^2 - y^2) - x^2 = 0 = (1 - x^2 - y^2) - y^2$ und daraus leicht $x^2 = y^2 = \frac{1}{3}$. Dann ist $z = \sqrt{1 - 2/3} = \frac{1}{\sqrt{3}}$ und wir finden die beiden Maximalpunkte $(x, y, z) = (\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}})$ und $(x, y, z) = (-\frac{1}{\sqrt{3}}, -\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}})$. (Die Punkte mit gemischten Vorzeichen in den beiden ersten Komponenten liefern einen negativen Funktionswert, also sicher kein Maximum.)

Probleme bei diesem Lösungsverfahren sind:

- Die Auflösung nach z ist nicht eindeutig: Warum nicht $z = -\sqrt{1 - x^2 - y^2}$? Das hätte mit $(x, y, z) = (\frac{1}{\sqrt{3}}, -\frac{1}{\sqrt{3}}, -\frac{1}{\sqrt{3}})$ und $(x, y, z) = (-\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, -\frac{1}{\sqrt{3}})$ zwei weitere Maximalpunkte geliefert. Sind das nun alle?
- Bei komplizierteren Funktionen g mag die Auflösung nach einer der Variablen gar nicht explizit möglich sein.
- Wichtiger: Die Funktion $z = \sqrt{1 - x^2 - y^2}$ und damit auch $\tilde{f}$ sind nur für die Punkte mit $x^2 + y^2 < 1$ differenzierbar. Deshalb muss man die Punkte mit $x^2 + y^2 = 1$ noch gesondert untersuchen.
- Die willkürliche Auswahl von z zerstört die Symmetrie des Problems und führt deshalb zu mühsamen Rechnungen.

2. Lösungsansatz. Bevor wir diesen diskutieren, wollen wir noch einmal an folgendes erinnern. Bei gewöhnlichen Extremalproblemen ohne Nebenbedingungen ist die notwendige Bedingung für ein Extremum in einem inneren Punkt $\vec{x}$, dass dort die Ableitung (= der Gradient) verschwindet, so dass also alle Richtungsableitungen $\operatorname{grad}_{\vec{x}} f \cdot \vec{u}$ verschwinden. Ist eine Richtungsableitung in $\vec{x}$ etwa > 0, so ist dort sicher kein Maximum, und weil dann die Ableitung in Gegenrichtung < 0 ist, auch kein Minimum.

Wenn wir jetzt f nur auf einer Niveaufläche von g untersuchen, ergibt sich folgendes Bild:

Der Vektor $\operatorname{grad}_{\vec{x}} g$ steht an der Stelle $\vec{x}$ senkrecht auf der g-Niveaufläche. Wenn also die Einschränkung von f auf diese Niveaufläche in $\vec{x}$ ein lokales Extremum hat, müssen alle Richtungsableitungen $\operatorname{grad}_{\vec{x}} f \cdot \vec{u} = 0$ sein, wenn $\vec{u}$ in Richtung der Niveaufläche zeigt. Also steht $\operatorname{grad} f$ senkrecht auf allen Vektoren, die auf $\operatorname{grad} g$ senkrecht stehen. Das bedeutet aber, dass $\operatorname{grad} f$ in $\pm$ dieselbe Richtung zeigt wie $\operatorname{grad} g$: Es gibt ein $\lambda \in \mathbb{R}$ mit

$$\operatorname{grad}_{\vec{x}} f = \lambda \operatorname{grad}_{\vec{x}} g.$$

Schreibt man das in partiellen Ableitungen, so findet man

$$\frac{\partial f}{\partial x_i} = \lambda \frac{\partial g}{\partial x_i} \quad \text{für } i = 1, \ldots, n (= 3 \text{ in unserem Spezialfall}).$$

Das sind n Gleichungen für $n+1$ Unbekannte, nämliche die gesuchten x_i und die Hilfsvariable λ, die auch *Lagrangescher Multiplikator* heißt. Zum Glück haben wir noch eine $(n + 1)$-te Gleichung, nämlich $g(\vec{x}) = 0$.

Bei dieser Überlegung haben wir stillschweigend vorausgesetzt, dass $g = 0$ wirklich in der Nähe von $\vec{x}$ eine glatte Niveaufläche beschreibt, auf der der Gradient senkrecht steht. Das ist aber nur in solchen $\vec{x}$ sicher, in denen $\operatorname{grad}_{\vec{x}} g \neq \vec{0}$. Die Punkte mit $\operatorname{grad}_{\vec{x}} g = \vec{0}$ muss man separat untersuchen.

> Zum Beispiel liefert $g(x, y, z) = yz = 0$ die Vereinigung der xz- mit der xy-Ebene. Entlang der x-Achse schneiden sich diese beiden unter 90^0. Dort hat man also *keine* glatte Niveaufläche. Der Gradient verschwindet dort.
> Der Satz über implizit definierte Funktionen (auf den wir nicht näher eingehen) besagt, dass man

$g(x_1, \ldots, x_i, \ldots, x_n) = 0$ in der Nähe von $\vec{x}$ eindeutig nach x_i auflösen kann, wenn $g(\vec{x}) = 0$ und $\frac{\partial g}{\partial x_i}(\vec{x}) \neq 0$, also insbesondere $\operatorname{grad}_{\vec{x}} g \neq 0$. Und x_i ist dann eine differenzierbare Funktion der anderen Variablen – stetige Differenzierbarkeit von g vorausgesetzt. In diesem Fall kann man unseren 1. Lösungsversuch präzisieren und kommt genau zu dem von uns geometrisch hergeleiteten Kriterium, das wir nun noch einmal als Satz zusammenfassen.

Satz 72 (Extremwerte mit Nebenbedingungen). *Die reellwertigen Funktionen f und g auf der offenen Menge $G \subset \mathbb{R}^n$ mögen stetige partielle Ableitungen haben.*
Gesucht sind die Extremwerte von f unter der Nebenbedingung $g = 0$.
Die einzig möglichen Kandidaten sind dann die Lösungen $\vec{x} \in G$ von einem der beiden folgenden Gleichungssysteme

$$\begin{aligned} \operatorname{grad}_{\vec{x}} f &= \lambda \operatorname{grad}_{\vec{x}} g, \\ g(\vec{x}) &= 0. \end{aligned}$$

oder

$$\begin{aligned} \operatorname{grad}_{\vec{x}} g &= \vec{0}, \\ g(\vec{x}) &= 0, \end{aligned}$$

- *Im ersten Fall hat man eine zusätzliche dummy-Variable λ, den sogenannten* Lagrange-Multiplikator*, der uns eigentlich nicht interessiert. Das entstehende System von $n+1$ Gleichungen für die $n+1$ Unbekannten $x_1, \ldots, x_n, \lambda$ ist im allgemeinen nicht-linear und deshalb schwierig zu lösen. Typischerweise erhalten wir eine Anzahl einzelner Punkte als Lösungen. Die Funktionswerte dort entscheiden über (globale) Maxima und Minima.*
- *Der zweiten Fall, der sogenannte singuläre Fall, hat nichts mit f zu tun. Er liefert die Punkte, wo das g-Niveau keine glatte Fläche oder Kurve ist, und wo unsere Argumentation deshalb versagt. Diese Punkte muss man (ad hoc) mit anderen Methoden untersuchen.*

Bemerkung. Eine gängige Variante für den ersten Fall ist die folgende: Man bildet die Funktion

$$h(\vec{x}, \lambda) := f(\vec{x}) - \lambda g(\vec{x})$$

von $n+1$ Variablen und bestimmt deren Extrema ohne Nebenbedingung. Der Gradient von h ist nämlich

$$\begin{pmatrix} \frac{\partial f}{\partial x_1} - \lambda \frac{\partial g}{\partial x_1} \\ \vdots \\ \frac{\partial f}{\partial x_n} - \lambda \frac{\partial g}{\partial x_n} \\ -g(\vec{x}) \end{pmatrix}$$

Die Bedingung $\operatorname{grad} h = 0$ liefert also dasselbe Gleichungssystem wie im Satz.

Beispiel 73. Wir betrachten wieder das Problem, $f(x, y, z) = xyz$ unter der Nebenbedingung $g(x, y, z) = x^2 + y^2 + z^2 - 1 = 0$ zu maximieren. Es ist $\operatorname{grad} g = (2x, 2y, 2z) \neq 0$ für alle $(x, y, z) \neq (0, 0, 0)$, also insbesondere für alle, die die Nebenbedingung erfüllen. Daher

gibt es keine singulären Punkte. Aus dem zweiten Gleichungssystem erhält man

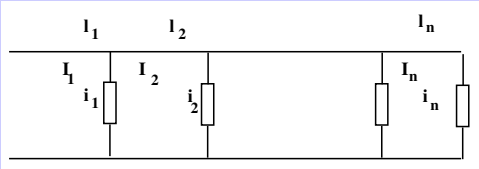
$$\begin{aligned} yz &= \lambda\, 2x, \\ xz &= \lambda\, 2y, \\ xy &= \lambda\, 2z. \end{aligned}$$

Multipliziert man diese Gleichungen mit x, y, z, addiert und verwendet die Nebenbedingung, so hat man $3xyz = 2\lambda$. Einsetzen von λ in die obigen Gleichungen liefert $x^2 = y^2 = z^2 = \frac{1}{3}$ oder zwei Koordinaten sind 0, die dritte dann wegen der Nebenbedingung ± 1. Die letzteren Punkte scheiden als Achsenpunkte wieder aus ($f = 0$). Mögliche Extrema liegen also in den Punkten

$$(\pm\frac{1}{\sqrt{3}}, \pm\frac{1}{\sqrt{3}}, \pm\frac{1}{\sqrt{3}})$$

mit voneinander unabhängigen Vorzeichen. Die entsprechenden Funktionswerte sind $f = \pm\frac{1}{3\sqrt{3}}$. Die positiven sind die Maxima, die negativen die Minima. Diese Rechnung ist deutlich einfacher als die im 1. Lösungsansatz. □

Beispiel 74 (Vgl. Fichtenholz, Bd. 1). Gegeben sei ein Stromkreis mit $n > 1$ Verbrauchern in Parallelschaltung



Der Stromverbrauch der Abnehmer sei konstant $i_1, \ldots, i_n$, so dass in den Leitungsabschnitten die Ströme

$$\begin{aligned} I_1 &= i_1 + i_2 + \ldots + i_n \\ I_2 &= \quad\; i_2 + \ldots + i_n \\ &\vdots \\ I_n &= \qquad\qquad\; i_n \end{aligned}$$

fließen. Die Aufgabe ist die Minimierung des Materialverbrauchs durch geeignete Wahl der Leiterquerschnitte q_i auf den Abschnitten der Länge l_i bei vorgegebenem Spannungsabfall $2U_0$ für den gesamten Leiter. Wir beschränken uns auf den oberen Leiter und bezeichnen mit ρ den spezifischen Widerstand. Dann lautet das mathematische Modell:

$$V(q_1, \ldots, q_n) = l_1 q_1 + \ldots + l_n q_n \stackrel{!}{=} min.$$

unter der Nebenbedingung

$$U(q_1, \ldots, q_n) = \rho\frac{l_1}{q_1} I_1 + \ldots + \rho\frac{l_n}{q_n} I_n - U_0 = 0.$$

Man erhält aus $\operatorname{grad} V = \lambda \operatorname{grad} U$

$$l_i = \lambda\rho\frac{l_i}{q_i^2}I_i. \tag{22}$$

(Dabei haben wir das obige λ durch $-\lambda$ ersetzt. Natürlich darf man das, denn λ war der Name für irgendeine reelle Zahl.) Die hier berechneten partiellen Ableitungen von U sind alle $\neq 0$, so dass es keine singulären Punkte gibt. Wir erhalten

$$q_i = \sqrt{\lambda\rho}\sqrt{I_i}$$

Multipliziert man andrerseits (22) mit q_i und summiert, so folgt aus der Nebenbedingung

$$\sum q_i l_i = \lambda U_0,$$

also

$$\lambda = \frac{1}{U_0}\sum \sqrt{\lambda\rho}\sqrt{I_i}l_i,$$
$$\sqrt{\lambda} = \frac{\sqrt{\rho}}{U_0}\sum \sqrt{I_i}l_i.$$

Mit diesem λ folgt

$$q_i = \sqrt{\lambda\rho}\,\sqrt{I_i}.$$

Dies ist also der einzig mögliche Punkt für Extrema unter der angegebenen Nebenbedingung. Der zugehörige Verbrauch ist

$$V^* = \sqrt{\lambda\rho}\,(l_1\sqrt{I_1} + \ldots + l_n\sqrt{I_n})$$

Warum ist das wirklich das Minimum?
Damit die Nebenbedingung erfüllbar ist, muss für jedes i

$$q_i > \rho\frac{l_i I_i}{U_0}$$

sein. Und ist eines der q_i nah an dieser unteren Schranke und die Nebenbedingung erfüllt, so müssen die anderen q_j sehr groß sein. Daraus schließt man (mit ein bisschen Mathematik), dass für großes M die Menge

$$\{(q_1,\ldots,q_n) \mid \rho\frac{l_i I_i}{U_0} < q_i \leq M\} \cap \{(q_1,\ldots,q_n) \mid U(q_1,\ldots,q_n) = 0\}$$

kompakt ist, so dass die stetige Funktion f darauf ihr Maximum und Minimum annimmt. Ist aber auch nur eines der q_i sehr groß, so wird auch der Wert von V sehr groß und solche Punkte kommen für das Minimum nicht in Frage. Also wird das Minimum in

$$\{(q_1,\ldots,q_n) \mid \rho\frac{l_i I_i}{U_0} < q_i < M\} \cap \{(q_1,\ldots,q_n) \mid U(q_1,\ldots,q_n) = 0\}$$

angenommen und dann notwendig im oben gefundene Punkt.

2 Vektoranalysis

2.1 Klassische Differentialoperatoren

- Die partiellen Ableitungen einer Abbildung kann man auf verschiedene Weise zu neuen Abbildungen kombinieren. Solche Kombinationsvorschriften nennt man dann *Differentialoperatoren*.
- Einige der so konstruierten Differentialoperatoren spielen in der Physik eine wichtig Rolle. Wir wollen ihre Definitionen kennenlernen und eine anschauliche Vorstellung für sie entwickeln.

Wir betrachten in diesem Abschnitt Abbildungen

$$u : \mathbb{R}^n \supset G \to \mathbb{R} \text{ und } \vec{v} : \mathbb{R}^n \supset G \to \mathbb{R}^n.$$

die wir aber als „Felder“ interpretieren: Wir stellen uns vor, dass $\vec{v}(\vec{x})$ ein Vektor ist, der an der Stelle $\vec{x}$ „sitzt“, so dass $\vec{v}$ ein *Vektorfeld* ist. Beispiele liefern elektrische oder magnetische Felder, Geschwindigkeitsfelder stationärer Strömungen oder Kraftfelder wie das der Gravitation. Reellwertige Funktionen u bezeichnen wir dann als *skalare Felder*: an jeder Stelle sitzt ein skalarer Wert. Temperaturverteilungen, Materialdichten oder Potentiale von Kraftfeldern sind Beispiele dafür.

Der Gradient eines skalaren Feldes. Den Gradienten haben wir schon früher ausführlich behandelt. Für $r = \sqrt{x^2 + y^2 + z^2}$ war

$$\operatorname{grad} r = \frac{\vec{x}}{r}, \quad \operatorname{grad} \frac{1}{r} = -\frac{\vec{x}}{r^3}, \quad \operatorname{grad} \frac{-1}{r^2} = \frac{2\vec{x}}{r^4}. \tag{23}$$

Beispiel 75 (Coulombpotential).
Theoretische Elektrotechnik, Abschnitt 1.4
Sei $u = \frac{Q}{4\pi\,\epsilon\,|\vec{x}|}$ das Potential einer Ladung im Ursprung. Dann ist das zugehörige elektrische Feld gegeben als

$$\vec{E}_{\vec{x}} = -\operatorname{grad}_{\vec{x}} u = \frac{Q}{4\pi\,\epsilon|\vec{x}|^3}\,\vec{x}. \tag{24}$$

Die Divergenz eines Vektorfeldes. Die Divergenz eines Vektorfeldes ist ein skalares Feld. Wir haben sie bereits im Beispiel 46 kennengelernt. Wir geben hier ein etwas andere Interpretation: Wieder verstehen wir das Vektorfeld $\vec{v}$ als stationäres Strömungsfeld, d.h. als das Geschwindigkeitsfeld strömender Teilchen, deren Geschwindigkeit an jeder Stelle zeitunabhängig (=stationär) ist. Für einen Augenblick nehmen wir an, $\vec{v}$ sei konstant (= homogenes Strömungsfeld). Wir betrachten eine (ebene durchlässige) Fläche in der Strömung und fragen, wieviel Flüssigkeit in der Zeiteinheit durch die Fläche hindurchströmt.

Man stellt sich vor, dass die Fläche in der Strömung mitschwimmt. Das von ihr überstrichene Volumen ist dann genau die gesuchte Größe.

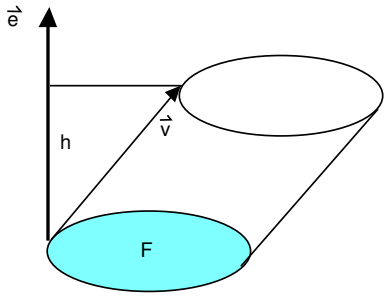
Dieses Volumen ist aber gleich dem Flächeninhalt F multipliziert mit der Komponente h von $\vec{v}$ in der zur Fläche senkrechten Richtung. Ist die Strömung senkrecht zur Fläche, so erhält man

$$F\,|\vec{v}|.$$

Andernfalls ergibt sich

$$F\vec{e}\cdot\vec{v},$$

wobei $\vec{e}$ ein zur Fläche senkrechter Einheitsvektor ist.



Nun betrachten wir eine räumlich inhomogene Strömung. Die Strömung habe Quellen und Senken, die aber nicht in Punkten konzentriert, sondern stetig verteilt seien: Wir nehmen an, dass es eine Quelldichte-Funktion $\rho(\vec{x})$ gibt. Legen wir um $\vec{x}$ einen kleinen Würfel der Kantenlänge $2t$, so ist die davon eingeschlossene Quelle bei kleinem t ungefähr

$$Volumen \times Quelldichte \approx (2t)^n\rho(\vec{x}),$$

wobei Sie sich vorstellen können, dass $n = 3$. (Bei größeren Würfeln müßte man eigentlich ρ integrieren. Darauf kommen wir im Kapitel 3 zurück.) Die Strömungsbilanz durch die Oberfläche des Würfels ist dann gerade gleich der eingeschlossenen Quelle.

$$\begin{aligned}(2t)^n\rho(\vec{x}) &\approx \sum_{i=1}^{2n} \text{Fluß durch die i-te Seite} \\ &= \sum_{i=1}^{n}(2t)^{n-1}\,\vec{e}_i\cdot\vec{v}(\vec{x}+t\vec{e}_i) + \sum_{i=1}^{n}(2t)^{n-1}\,(-\vec{e}_i)\cdot\vec{v}(\vec{x}-t\vec{e}_i).\end{aligned}$$

Es folgt, wenn v_i die i-te Komponente von $\vec{v}$ bezeichnet,

$$\begin{aligned}\rho(\vec{x}) &\approx \frac{1}{2t}\sum\left(v_i(\vec{x}+t\vec{e_i}) - v_i(\vec{x}-t\vec{e_i})\right) \\ &= \frac{1}{2}\sum\left(\frac{v_i(\vec{x}+t\vec{e_i})-v_i(\vec{x})}{t} + \frac{v_i(\vec{x}-t\vec{e_i})-v_i(\vec{x})}{-t}\right).\end{aligned}$$

Für $t\to 0$ gehen *beide* Summanden gegen $\frac{\partial v_i}{\partial x_i}$ und die Quelldichte wird

$$\rho(\vec{x}) = \sum_{i=1}^{n}\frac{\partial v_i}{\partial x_i}(\vec{x}).$$

Man bezeichnet die Quelldichte ρ als die *Divergenz* des Vektorfeldes $\vec{v}$:

$$\boxed{\operatorname{div}_{\vec{x}}\vec{v} = \sum_{i=1}^{n}\frac{\partial v_i}{\partial x_i}(\vec{x}).}$$

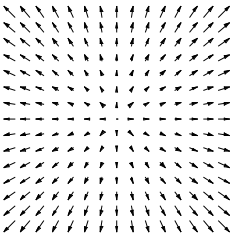
Die Divergenz eines Vektorfeldes ist also ein skalares Feld.

Ein Feld mit $\operatorname{div}\vec{v} = 0$ heißt ein *divergenzfreies* oder *quellenfreies* Vektorfeld.

Beispiel 76. $\vec{v}(\vec{x}) = \vec{v}_0 = const.$ liefert $\operatorname{div}\vec{v} = 0$. □

Beispiel 77. Wir betrachten das Feld $\vec{v}(\vec{x}) = a\vec{x}$ mit $a > 0$.

Die Divergenz ist in diesem Fall konstant $= na$. Das erscheint auf den ersten Blick paradox, weil das Feld doch „aus 0 kommt“. Aber wenn man beachtet, dass $\vec{v}(\vec{x}) = a(\vec{x}-\vec{x}_0)+a\vec{x}_0$, so wird klar: Das Feld kommt nicht nur „aus 0“, sondern, bis auf einen homogenen, also quellenfreien Anteil, ebenso „aus $\vec{x}_0$“.



□

Beispiel 78 (Coulombfeld).
Theoretische Elektrotechnik, Abschnitt 1.4
Für das elektrische Feld einer Punktladung $\vec{E} = \frac{Q}{4\pi\epsilon|\vec{x}|^3}\,\vec{x}$ ergibt sich mit (23) $\operatorname{div}\vec{E} = 0$ für alle $\vec{x} \neq 0$. Beachten Sie, dass physikalisch betrachtet die einzige „Quelle“ dieses Feldes in $\vec{x} = 0$ sitzt.

Für die Berechnung der Divergenz ist die folgende Rechenregel hilfreich, die sich unmittelbar aus der Definition der Divergenz ergibt. Für ein skalares Feld ψ und ein Vektorfeld $\vec{v}$ gilt:

$$\operatorname{div}(\psi\vec{v}) = (\operatorname{grad}\psi)\cdot\vec{v} + \psi\operatorname{div}\vec{v}.$$

Beispiel 79 (Magnetfeld eines geraden Leiters). Man rechnet nach: $\operatorname{div}\vec{H} = 0$ für

$$\vec{H}(x,y,z) := \frac{I}{2\pi}\left(\frac{-y}{x^2+y^2}, \frac{x}{x^2+y^2}, 0\right).$$

H ist das induzierte Magnetfeld im Außenraum eines (unendlich langen) Leiters, der z-Achse, mit Strom I.

Beispiel 80 (Kontinuitätsgleichung).
Theoretische Elektrotechnik, Abschnitte 2, 5 und 6
Die Divergenz kommt in den Anwendungen häufig in sogenannten Kontinuitätsgleichungen vor. Bei der Stromleitung ist zum Beispiel die Divergenz der Stromdichte proportional zur zeitlichen Änderung der Elektronendichte:

$$\operatorname{div}\vec{j} = q\frac{\partial n}{\partial t}.$$

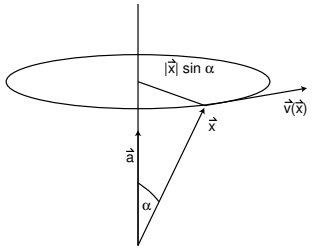
Wir kommen darauf im Rahmen des Gaußschen Integralsatzes zurück.

Die Rotation eines Vektorfeldes. Wir motivieren die Definition der Rotation in drei Schritten:

1.

Wir beginnen mit einer Vorbetrachtung über Rotationen im 3-dimensionalen Raum und betrachten eine Drehung mit konstanter Geschwindigkeit um eine Achse durch den Nullpunkt des $\mathbb{R}^3$. Das Geschwindigkeitsfeld dieser Drehung läßt sich bequem so beschreiben:

$$\vec{v}(\vec{x}) = \vec{a} \times \vec{x}. \tag{25}$$



Dabei zeigt $\vec{a}$ in die Richtung der Achse, und die Länge $|\vec{a}|$ ist gerade die Winkelgeschwindigkeit. Beachten Sie, dass $\vec{v}$ tangential ist an die Kreise in Ebenen senkrecht zu $\vec{a}$.

Nun kann man (25) auch so schreiben:

$$\vec{a} \times \vec{x} = \begin{pmatrix} a_2x_3 - a_3x_2 \\ a_3x_1 - a_1x_3 \\ a_1x_2 - a_2x_1 \end{pmatrix} = \begin{pmatrix} 0 & -a_3 & a_2 \\ a_3 & 0 & -a_1 \\ -a_2 & a_1 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}. \tag{26}$$

Man kann eine Achsendrehung also beschreiben durch einen „Drehvektor“ $\vec{a}$ oder, gleichwertig, durch eine (zur Diagonalen *schiefsymmetrische*) Matrix, deren Koeffizienten gerade die Koeffizienten des Drehvektors in der obigen Anordnung sind.

2. Wir nennen eine (quadratische) Matrix (a_{ij}) *symmetrisch*, wenn $a_{ij} = a_{ji}$, und *schiefsymmetrisch*, wenn $a_{ij} = -a_{ji}$ für alle i, j gilt. Eine schiefsymmetrische 3×3-Matrix entspricht einer Drehung um eine Achse und damit eineindeutig einem Vektor des $\mathbb{R}^3$ unter der Beziehung (26).

   Mit Hilfe der an der Diagonalen gespiegelten Matrix A^T, der sogenannten Transponierten, kann man eine beliebige $n \times n$-Matrix (a_{ij}) immer schreiben als

$$\begin{aligned}(a_{ij}) &= \frac{1}{2}\left((a_{ij}) - (a_{ij})^T\right) + \frac{1}{2}\left((a_{ij}) + (a_{ij})^T\right) \\ &= \frac{1}{2}\left(a_{ij} - a_{ji}\right) + \frac{1}{2}\left(a_{ij} + a_{ji}\right).\end{aligned}$$

   Man hat also eine Zerlegung in den sogenannten *Rotationsanteil* von (a_{ij}) und den *rotationsfreien* Anteil.

3. Nun betrachten wir eine beliebige (differenzierbare, stationäre) Strömung $\vec{v}$ und approximieren diese in der Nähe von $\vec{x}_0$ linear:

$$\vec{v}(\vec{x}) \approx \vec{v}(\vec{x}_0) + \underbrace{\vec{v}\,'(\vec{x})}_{Matrix}\ \underbrace{(\vec{x} - \vec{x}_0)}_{Vektor}.$$

   Der erste Summand ist eine homogene stationäre Strömung. Der zweite Teil ist das Produkt der Ableitung (Funktionalmatrix, Matrix der partiellen Ableitungen) mit $\vec{x} - \vec{x}_0$. Der Rotationsanteil von $\vec{v}\,'(\vec{x})$ ist

$$\frac{1}{2}\left(\vec{v}\,' - (\vec{v}\,')^T\right) = \frac{1}{2}\left(\frac{\partial v_i}{\partial x_j} - \frac{\partial v_j}{\partial x_i}\right)_{i,j=1,\ldots,n}.$$

   Dem entspricht der Vektor

$$\frac{1}{2}\begin{pmatrix} \frac{\partial v_3}{\partial x_2} - \frac{\partial v_2}{\partial x_3} \\ \frac{\partial v_1}{\partial x_3} - \frac{\partial v_3}{\partial x_1} \\ \frac{\partial v_2}{\partial x_1} - \frac{\partial v_1}{\partial x_2} \end{pmatrix}.$$

Dieser ist also der Rotationsanteil der Ableitung (=linearen Approximation) von $\vec{v}$.

Definition 81. Das Vektorfeld

$$\boxed{\operatorname{rot}\vec{v} := \begin{pmatrix} \frac{\partial v_3}{\partial x_2} - \frac{\partial v_2}{\partial x_3} \\ \frac{\partial v_1}{\partial x_3} - \frac{\partial v_3}{\partial x_1} \\ \frac{\partial v_2}{\partial x_1} - \frac{\partial v_1}{\partial x_2} \end{pmatrix}.}$$

heißt die *Rotation* von $\vec{v}$.

Die Rotation eines Vektorfeldes ist wieder ein Vektorfeld.

Ist $\operatorname{rot}\vec{v} = \vec{0}$, so nennt man das Feld *rotations- oder wirbelfrei*.

Man hat also die Strömung $\vec{v}$ lokal approximiert durch einen homogenen Anteil, einen Rotationsanteil und einen rotationsfreien Anteil

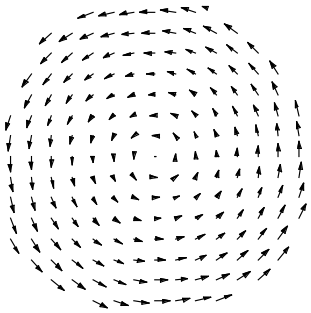
$$\vec{v}(\vec{x}) \approx \underbrace{\vec{v}(\vec{x}_0)}_{homogen} + \underbrace{\frac{1}{2}\operatorname{rot}_{\vec{x}_0}\vec{v} \times (\vec{x}-\vec{x}_0)}_{Rotationsanteil} + \underbrace{\frac{1}{2}\left(\vec{v}\,'(\vec{x}_0) + (\vec{v}\,'(\vec{x}_0))^T\right)(\vec{x}-\vec{x}_0)}_{wirbelfrei}.$$

Beispiel 82. Ähnlich wie im Beispiel 77 zur Divergenz findet man,

dass das Vektorfeld $\vec{v}(\vec{x}) = \frac{1}{2}\,\vec{a} \times \vec{x}$ an jeder Stelle $\vec{x}_0$ die Rotation $\operatorname{rot}_{\vec{x}_0}\vec{v} = \vec{a}$ besitzt. Das ist offenbar, weil

$$\vec{v}(\vec{x}) = \frac{1}{2}\,\vec{a} \times \vec{x} = \frac{1}{2}\,\vec{a} \times \vec{x}_0 + \frac{1}{2}\,\vec{a} \times (\vec{x}-\vec{x}_0).$$

Natürlich kann man auch einfach losrechnen, indem man die partiellen Ableitungen von $\vec{v}(\vec{x}) = \frac{1}{2}(a_2x_3 - a_3x_2, a_3x_1 - a_1x_3, a_1x_2 - a_2x_1)$ bildet.



□

Für die Berechnung der Rotation ist die folgende Rechenregel hilfreich, die sich unmittelbar aus der Definition der Rotation ergibt. Für ein skalares Feld ψ und ein Vektorfeld $\vec{v}$ gilt:

$$\operatorname{rot}(\psi\vec{v}) = (\operatorname{grad}\psi) \times \vec{v} + \psi \operatorname{rot}\vec{v}. \tag{27}$$

Beispiel 83. Für das elektrische Feld der Punktladung $\vec{E} = \frac{Q}{4\pi\epsilon}\frac{\vec{x}}{|\vec{x}|^3}$ rechnet man nach, dass $\operatorname{rot}\vec{E} = 0$.

Beispiel 84. Für

$$H(x,y,z) = \frac{I}{2\pi(x^2+y^2)} \begin{pmatrix} -y \\ x \\ 0 \end{pmatrix}$$

findet man mit (27) und (23)

$$\operatorname{rot} \vec{H} = \operatorname{rot} \frac{I}{2\pi(x^2+y^2)} \begin{pmatrix} -y \\ x \\ 0 \end{pmatrix} = -\frac{I}{2\pi(x^2+y^2)^2} \begin{pmatrix} 2x \\ 2y \\ 0 \end{pmatrix} \times \begin{pmatrix} -y \\ x \\ 0 \end{pmatrix} + \frac{I}{2\pi(x^2+y^2)} \operatorname{rot} \begin{pmatrix} -y \\ x \\ 0 \end{pmatrix}$$

$$= -\frac{I}{2\pi(x^2+y^2)^2} \begin{pmatrix} 0 \\ 0 \\ 2x^2+2y^2 \end{pmatrix} + \frac{I}{2\pi(x^2+y^2)} \begin{pmatrix} 0 \\ 0 \\ 2 \end{pmatrix} = \vec{0}.$$

Beispiel 85 (Biot-Savartsches Gesetz).
Theoretische Elektrotechnik, Abschnit 3.3
Ist $\vec{J}$ ein konstanter Vektor, so ist (vgl. Beispiel 75)

$$\operatorname{rot} \frac{\mu_0}{4\pi} \frac{\vec{J}}{r} = \frac{\mu_0}{4\pi} \operatorname{grad} \frac{1}{r} \times \vec{J} + \frac{\mu_0}{4\pi} \frac{1}{r} \underbrace{\operatorname{rot} \vec{J}}_{=0} = -\frac{\mu_0}{4\pi} \frac{\vec{x}}{r^3} \times \vec{J} = \frac{\mu_0}{4\pi} \vec{J} \times \frac{\vec{x}}{r^3}.$$

Wenn $\vec{J}$ die Stromdichte an der Stelle $\vec{0}$ bezeichnet, so ist dieser Ausdruck die Dichte der magnetischen Induktion an der Stelle $\vec{x}$.

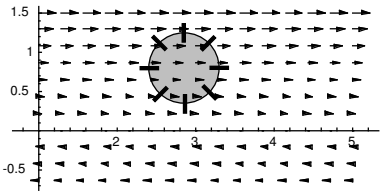
Rechenregeln für die Differentialoperatoren. Alle drei Operatoren sind linear: Auf Summen kann man sie gliedweise anwenden, konstante skalare Faktoren wandern heraus.

Für $\psi, u : \mathbb{R}^n \supset G \to \mathbb{R}$ und $\vec{v}, \vec{w} : \mathbb{R}^n \supset G \to \mathbb{R}^n$ hat man „Produktregeln“, die wir zum Teil oben schon angegeben und benutzt haben

$$\begin{aligned} \operatorname{grad}(\psi u) &= \psi \operatorname{grad} u + u \operatorname{grad} \psi, \\ \operatorname{div}(\psi \vec{v}) &= (\operatorname{grad} \psi) \cdot \vec{v} + \psi \operatorname{div} \vec{v}, \\ \operatorname{rot}(\psi \vec{v}) &= (\operatorname{grad} \psi) \times \vec{v} + \psi \operatorname{rot} \vec{v} && \text{(n=3)}, \\ \operatorname{div}(\vec{v} \times \vec{w}) &= (\operatorname{rot} \vec{v}) \cdot \vec{w} - \vec{v} \cdot \operatorname{rot} \vec{w} && \text{(n=3)}. \end{aligned}$$

Beispiel 86 (Ein Messverfahren für die Rotation). Die Divergenz kann man bestimmen, indem man die Strömungsbilanz eines kleinen durchlässigen Würfels mißt. Für die Rotation betrachtet man ein kleines Schaufelrad, dessen Nabe am Meßpunkt $\vec{x}$ fixiert ist und dessen Achse um diesen Punkt frei drehbar ist. Dreht man die Achse so, dass die Strömung das Rad mit maximaler Geschwindigkeit dreht, so zeigt die Achse in Richtung der Rotation und die geeignet normierte Geschwindigkeit liefert die Größe (den Betrag) der Rotation.

Das Vektorfeld $\vec{v}(x,y,z) = \begin{pmatrix} y \\ 0 \\ 0 \end{pmatrix}$ hat die Rotation $\operatorname{rot}\vec{v} = \begin{pmatrix} 0 \\ 0 \\ -1 \end{pmatrix}$. In der Abbildung erkennt



man, dass das Rad sich im mathematisch negativen Sinn dreht, weil oben eine stärkere Strömung herrscht als unten.

"Beweis" für das Schaufelrad-Verfahren bei der linearisierten Strömung. Sei A eine 3×3-Matrix und

$$A = \frac{1}{2}(A + A^T) + \frac{1}{2}(A - A^T) =: A^+ + A^-$$

die Zerlegung in den symmetrischen und schiefsymmetrischen Teil, und sei $\vec{a} \in \mathbb{R}^3$ mit

$$A^- x = \vec{a} \times \vec{x} \text{ für alle } \vec{x}.$$

Sei weiter $\vec{u}, \vec{v}$ ein Paar orthonormaler Vektoren, so dass der Kreis

$$\vec{x}(t) = \vec{u}\cos t + \vec{v}\sin t$$

das "Schaufelrad" beschreibt. Das Drehmoment, das die Strömung $\vec{v}(\vec{x}) = A\vec{x}$ auf das Rad ausübt, ist gegeben durch die Integration der tangentialen Komponente längs des Rades, die wir durch Skalarmultiplikation mit dem Tangentialvektor $\dot{\vec{x}}(t)$ beschreiben. In der Rechnung benutzen wir, dass wegen der Symmetrie von A^+ gilt:

$$\frac{d}{dt}\langle A^+\vec{x}(t), \vec{x}(t)\rangle = \langle A^+\dot{\vec{x}}(t), \vec{x}(t)\rangle + \langle A^+\vec{x}(t), \dot{\vec{x}}(t)\rangle = 2\langle A^+\vec{x}(t), \dot{\vec{x}}(t)\rangle.$$

Damit erhalten wir

$$\begin{aligned} F &= \int_0^{2\pi} \langle A\vec{x}(t), \dot{\vec{x}}(t)\rangle dt = \int_0^{2\pi} \langle A^+\vec{x}(t), \dot{\vec{x}}(t)\rangle dt + \int_0^{2\pi} \langle A^-\vec{x}(t), \dot{\vec{x}}(t)\rangle dt \\ &= \int_0^{2\pi} \frac{1}{2}\frac{d}{dt}\langle A^+\vec{x}(t), \vec{x}(t)\rangle dt + \int_0^{2\pi} \langle \vec{a} \times \vec{x}(t), \dot{\vec{x}}(t)\rangle dt \\ &= \frac{1}{2}\underbrace{\big(\langle A^+\vec{x}(2\pi), \vec{x}(2\pi)\rangle - \langle A^+\vec{x}(0), \vec{x}(0)\rangle\big)}_{=0} + \int_0^{2\pi} \underbrace{\det(a, \vec{x}(t), \dot{\vec{x}}(t))}_{=konstant} dt \\ &= 2\pi \det(\vec{a}, \vec{u}, \vec{v}) = 2\pi\langle a, \vec{u} \times \vec{v}\rangle. \end{aligned}$$

Das wird maximal, wenn die Achse $\vec{u} \times \vec{v}$ des Schaufelrades in Richtung des Vektors $\vec{a}$ liegt, und $|\vec{a}|$ ist dann das bis auf den Faktor 2π das Drehmoment. □

Wir stellen für die in den Anwendungen wichtigen Felder im $\mathbb{R}^3$ noch einmal die Wirkung der Differentialoperatoren dar.

Physikalische Bedeutung	Feld	grad	div	rot
	r	$\frac{\vec{x}}{r}$	–	–
Coulombpotential	$\frac{-1}{r}$	$\frac{\vec{x}}{r^3}$	–	–
Coulombfeld	$\frac{\vec{x}}{r^3}$	–	0	$\vec{0}$
Magnetfeld des Leiters	$\frac{1}{x^2+y^2}\begin{pmatrix} -y \\ x \\ 0 \end{pmatrix}$	–	0	$\vec{0}$
"Biot-Savart-Feld"	$\frac{1}{r}\vec{J}$	–	$-\frac{\vec{x}\cdot\vec{J}}{r^3}$	$-\frac{\vec{x}\times\vec{J}}{r^3}$

2.2 Mehrfach-Anwendung der Differentialoperatoren

- Häufig kommen Kombinationen der klassischen Differentialoperatoren vor.
- Der Laplaceoperator div grad $= \Delta$ ist ein Differentialoperator von besonderer Wichtigkeit für mehrdimensionale Schwingungsprobleme oder für Diffusionsprozesse.
- Andere Kombinationen sind vor allem wichtig im Zusammenhang mit der Frage nach der Existenz sogenannter Potentiale.

In physikalischen Anwendungen kommen sehr häufig Kombinationen von zwei der im letzten Abschnitt besprochenen Differentialoperatoren vor. Sinnvoll sind

$$\operatorname{div}\operatorname{grad} u,\ \operatorname{rot}\operatorname{grad} u,\ \operatorname{div}\operatorname{rot}\vec{v},\ \operatorname{rot}\operatorname{rot}\vec{v},\ \operatorname{grad}\operatorname{div}\vec{v}.$$

Diese Kombinationen wollen wir in diesem Abschnitt betrachten.

Bemerkung. Physikalische Felder hängen sehr häufig von Ort und Zeit ab, wie die folgenden Beispiele unterstreichen. Sie sind also Funktionen auf einem 4-dimensionalen Raum. In den Differentialgleichungen für diese Größen behandelt man in der Regel die Zeitvariable gesondert und wendet Differentialoperatoren nur auf die Ortsvariablen an.

2.2.1 Der Laplaceoperator

Auf ein zweimal differenzierbares skalares Feld u kann man den *Laplaceoperator* anwenden:

$$\Delta u := \operatorname{div}(\operatorname{grad} u) = \frac{\partial^2 u}{\partial x_1^2} + \ldots + \frac{\partial^2 u}{\partial x_n^2}$$

Für $n = 1$ ist $\Delta u = u''$. Tatsächlich spielt Δ in höheren Dimensionen oft die Rolle der „zweiten Ableitung“. Auch auf zweimal differenzierbare Vektorfelder kann man den Laplaceoperator anwenden: einfach komponentenweise:

$$\Delta \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} = \begin{pmatrix} \Delta v_1 \\ \vdots \\ \Delta v_n \end{pmatrix}.$$

Beispiel 87 (Diffusion und Wärmeleitung).
Werkstoffe II, Abschnitt 8.3, Verfahrentechnik I, Kapitel 2 oder Energie-, Impuls- und Stofftransport, Abschnitt 2.5
Die raumzeitliche Temperaturverteilung T in einem (wärmeisotropen) Medium wird beschrieben durch die sogenannte *Wärmeleitungsgleichung*

$$\frac{\partial T}{\partial t} = \lambda\, \Delta T.$$

Diese Gleichung beschreibt gleichermaßen auch die Konzentration von Stoffen bei einem Diffusionsvorgang in (isotropen) Medien. Man bezeichnet sie deshalb auch als *Diffusionsgleichung*.
Wir erklären den physikalischen Hintergrund für diese Formel später in Verbindung mit dem Gaußschen Integralsatz, vgl. Beispiel 146.

Beispiel 88.

Energie-, Impuls- und Stofftransport, Abschnitt 2.5

Die explizite Lösung der Wärmeleitungsgleichung ist nur bei speziellen zusätzlichen Rand- und Anfangswerten möglich (vgl. Beispiel 129), im allgemeinen ist man auf numerische Lösungen angewiesen. Dazu benutzt man Diskretisierungsverfahren, die wir hier an einem einfachen Beispiel für den eindimensionalen Fall

$$\frac{\partial T}{\partial t} = \lambda \frac{\partial^2 T}{\partial x^2}. \tag{28}$$

skizzieren. Dies Modell wird zum Beispiel benutzt bei der Wärmeleitung oder Dampfdiffusion durch eine Wand. Man wählt kleine Δx und Δt und setzt

$$T_{n,k} := T(x_0 + n\Delta x, t_0 + k\Delta t).$$

Aus der (eindimensionalen) Taylorformel ergibt sich

$$T(x + \Delta x, t) \approx T(x,t) + \frac{\partial T}{\partial x}(x,t)\Delta x + \frac{1}{2}\frac{\partial^2 T}{\partial x^2}(x,t)(\Delta x)^2$$
$$T(x - \Delta x, t) \approx T(x,t) - \frac{\partial T}{\partial x}(x,t)\Delta x + \frac{1}{2}\frac{\partial^2 T}{\partial x^2}(x,t)(\Delta x)^2$$

und daraus durch Addition für $x = x_0 + n\Delta x$ und $t = t_0 + k\Delta t$

$$T_{n+1,k} + T_{n-1,k} - 2T_{n,k} \approx \frac{\partial^2 T}{\partial x^2}(x,t)(\Delta x)^2.$$

Einfacher erhält man ebenso

$$T_{n,k+1} - T_{n,k} \approx \frac{\partial T}{\partial t}(x,t)\Delta t.$$

Daher ersetzt man die Gleichung (28) durch die Differenzengleichung

$$T_{n,k+1} - T_{n,k} = \lambda \frac{\Delta t}{(\Delta x)^2}(T_{n+1,k} + T_{n-1,k} - 2T_{n,k}).$$

Kennt man die anfängliche Temperaturverteilung, also $T_{n,0}$ für alle n, so ergeben sich rekursiv daraus die $T_{n,k}$ für alle k.

Beispiel 89 (Schwingung, Wellen).

Müller: Mechanik III, Abschnitt 17; Popov: Mechanik III; Theoretische Elektrotechnik, Abschnitt 5/6

Die Schwingung einer Saite wird beschrieben durch ihre Auslenkung $u(x,t)$ an der Stelle x zur Zeit t. Dafür leitet die Physik folgende Gleichung her:

$$\frac{\partial^2 u}{\partial t^2} = \lambda \frac{\partial^2 u}{\partial x^2}.$$

Für die Schwingungen einer ebenen Membran oder für räumliche Schwingungsprobleme ergibt sich die mehrdimensionale *Wellengleichung*

$$\frac{\partial^2 u}{\partial t^2} = \lambda \Delta u.$$

Beispiel 90 (Harmonische Funktionen). Funktionen mit $\Delta u = 0$, die sogenannten *harmonischen Funktionen* spielen als Potentiale quellfreier Felder in der Potentialtheorie eine große Rolle. Wir berechnen ein Beispiel.
Für $k \in \mathbb{Z}$ und $r = \sqrt{x_1^2 + \ldots + x_n^2}$ ist

$$\frac{\partial r^k}{\partial x_i} = kr^{k-2}x_i, \quad \frac{\partial^2 r^k}{\partial x_i^2} = kr^{k-2} + k(k-2)r^{k-4}x_i^2.$$

Daher bekommt man

$$\Delta r^k = nkr^{k-2} + k(k-2)r^{k-4}\sum x_i^2 = k(n+k-2)r^{k-2}$$

und $r^{2-n} = \frac{1}{r^{n-2}}$ ist harmonisch:

$$\Delta \frac{1}{r^{n-2}} = 0.$$

2.2.2 Das Potential

Durch simples Nachrechnen unter Ausnutzung der Vertauschbarkeit zweiter partieller Ableitungen beweist man für Funktionen $u : \mathbb{R}^3 \supset G \to \mathbb{R}$ mit stetigen 2. partiellen Ableitungen:

$$\boxed{\operatorname{rot} \operatorname{grad} u = \vec{0}.} \tag{29}$$

Definition 91 (Potential). Ist $\vec{v}$ ein Vektorfeld, so heißt eine Funktion u ein *Potential* für $\vec{v}$, wenn

$$\operatorname{grad} u = -\vec{v}.$$

Das Vektorfeld $\vec{v}$ nennt man in diesem Fall ein *Potentialfeld* oder ein *konservatives* Feld.

Offenbar ist ein Potential nur bis auf eine Konstante bestimmt. Im Fall $n = 1$ ist $-u$ also eine Stammfunktion von $v = v_1$. Aber im Gegensatz zu skalaren Funktionen, die - wenn sie nur stetig sind - nach dem Hauptsatz der Differential-und Integralrechnung immer Stammfunktionen haben, besitzt nicht jedes Vektorfeld ein Potential. Wegen (29) gilt:

Satz 92 (Notwendige Potentialbedingung). *Notwendig für die Existenz eines Potentials zum stetig differenzierbaren Vektorfeld $\vec{v} : \mathbb{R}^3 \supset G \to \mathbb{R}^3$ ist die Bedingung*

$$\operatorname{rot} \vec{v} = \vec{0}.$$

Nur wirbelfreie Vektorfelder können *ein Potential haben.*

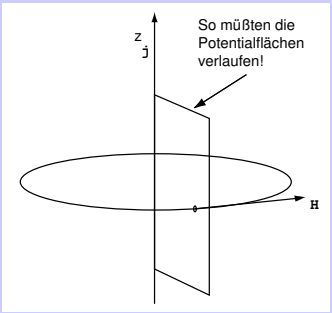
Sie *müssen* allerdings nicht, auch wenn das häufig behauptet wird.

Beispiel 93. Das elektrische Feld einer Punktladung und das Magnetfeld eines unendlichen Leiters sind beide wirbelfrei. Das erste hat ein Potential, vgl. Beispiel 75. Das zweite hingegen, also

$$\vec{H}(x,y,z) := \frac{I}{2\pi(x^2+y^2)} \begin{pmatrix} -y \\ x \\ 0 \end{pmatrix},$$

hat kein Potential.

Wir wollen das plausibel machen: Das Feld ist tangential an Kreise um die z-Achse senkrecht zu dieser. Diese Kreise nennt man auch die „Feldlinien“. Gäbe es ein Potential u, so zeigte der Gradient von $-u$ in Richtung dieser Feldlinien. Folgt man einer solchen, so bewegt man sich immer in Richtung des Gradienten von $-u$. Das heißt, längs der Feldlinie nimmt der Wert von $-u$ ständig zu. Aber wenn man einmal herum ist, muss $-u$ wieder beim selben Wert angekommen sein. Widerspruch!



Immerhin gilt:

Satz 94 (Hinreichende Potentialbedingung). *Ist $\vec{v}$ stetig differenzierbar und wirbelfrei auf einem offenen Definitionsbereich und ist G eine offene konvexe Teilmenge des Definitionsbereichs von $\vec{v}$, so existiert in G ein Potential für $\vec{v}$. Um jeden Punkt gibt es eine offene Kugel, die im Definitionsbereich liegt. Deshalb haben wirbelfreie Felder immer* lokale Potentiale.

Die Wirbelfreiheit ist auch hinreichend für die Existenz eines Potentials, wenn G *sternförmig* oder *einfach zusammenhängend* ist. *Sternförmig* bedeutet, dass es in G einen Punkt gibt, so dass die Verbindungsstrecke zwischen diesem Punkt und jedem anderen Punkt von G ganz in G liegt. *Einfach zusammenhängend* bedeutet, dass “G keine Löcher hat”, genauer, dass man jede geschlossene Kurve in G auf einen Punkt zusammenziehen kann.

Beispiel 95. Die Funktion

$$u = -\frac{I}{2\pi} \arctan \frac{y}{x}$$

ist auf der (konvexen) Menge $\{(x,y,z) \mid x > 0\}$ wirklich ein Potential für das Magnetfeld $\vec{H}(x,y,z) := \frac{I}{2\pi(x^2+y^2)}(-y,x,0)$ des unendlich langen Leiters, vgl. Beispiel 79. Sie ist sogar ein lokales Potential auf einer Umgebung eines jeden Punktes mit $x \neq 0$. Aber sie ist eben nicht auf dem *ganzen* Komplement der $z-$Achse definiert.

Wie findet man zu einem gegebenen wirbelfreien Vektorfeld ein (lokales) Potential bzw. eine Stammfunktion? Wir zeigen das an einem

Beispiel 96 (Potentialbestimmung). Gegeben sei das Vektorfeld

$$\vec{v}(x,y,z) = \begin{pmatrix} 3x^2y \\ x^3+z \\ y+1 \end{pmatrix}.$$

Rechnen Sie nach, dass $\operatorname{rot} v = \vec{0}$ ist. Gesucht ist eine Stammfunktion, d.h. eine Funktion $\phi : \mathbb{R}^3 \to \mathbb{R}$ mit

$$\begin{aligned} \frac{\partial \phi}{\partial x} &= 3x^2y \\ \frac{\partial \phi}{\partial y} &= x^3+z \\ \frac{\partial \phi}{\partial z} &= y+1 \end{aligned}$$

Aus der ersten Gleichung ergibt sich $\phi = x^3y$, nämlich durch Integration nach x. Offenbar sind aber die y- und z-Ableitung von ϕ nicht richtig. Der Grund ist, dass wir die Integrationskonstante unterschlagen haben, und weil wir bezüglich x integriert haben, kann diese „Konstante“ von y und z abhängen:

$$\phi = x^3y + c(y,z).$$

Einsetzen in die zweite Gleichung liefert

$$x^3 + \frac{\partial c}{\partial y} = x^3 + z,$$

also, und jetzt sind wir vorsichtiger,

$$c = yz + h(z).$$

Daraus folgt mit der dritten Gleichung schließlich

$$y + \frac{\partial h}{\partial z} = y + 1,$$

also $h(z) = z + const.$ Damit haben wir die Stammfunktion

$$\phi(x,y,z) = x^3y + yz + z + const.$$

und das Potential

$$u(x,y,z) = -(x^3y + yz + z + const.)$$

gefunden.

Dieses Verfahren funktioniert nur, wenn wirklich ein lokales Potential existiert. Probieren Sie

$$\vec{v}(x,y,z) = \begin{pmatrix} 3x^2y \\ x^2+z \\ y+1 \end{pmatrix}.$$

□

Bemerkung. Die Frage nach der Existenz eines Potentials für ein Vektorfeld macht auch in höheren Dimensionen Sinn. Die Rotation gibt es allerdings nur in Dimension 3. In höheren Dimensionen tritt an die Stelle der Wirbelfreiheit die sogenannnte Integrabilitätsbedingung

$$\boxed{\frac{\partial v_i}{\partial x_j} = \frac{\partial v_j}{\partial x_i}.}$$

Das ist nämlich, falls ein Potential existiert, gerade die Bedingung für die Vertauschbarkeit der zweiten Ableitungen, und im 3-dimensionalen Fall ergibt sich wieder die Wirbelfreiheit.

2.2.3 Das Vektorpotential

Durch simples Nachrechnen unter Ausnutzung der Vertauschbarkeit zweiter partieller Ableitungen beweist man für Vektorfelder mit stetigen zweiten partiellen Ableitungen:

$$\boxed{\operatorname{div}\operatorname{rot}\vec{w} = 0.}. \tag{30}$$

Wir sagen, das Vektorfeld $\vec{v}$ hat ein *Vektorpotential*, wenn es sich in der Form.

$$\vec{v} = \operatorname{rot}\vec{w}$$

schreiben lässt. $\vec{w}$ heißt dann ein Vektorpotential. Aus (30) sieht man sofort, dass

$$\operatorname{div}\vec{v} = 0,$$

also die Quellfreiheit von $\vec{v}$, eine notwendige Bedingung für die Existenz eines Vektorpotentials ist. Diese Bedingung ist wiederum lokal, nicht aber global, auch hinreichend.

Satz 97 (Existenz eines Vektorpotentials). *Eine notwendige Bedingung für die Existenz eines Vektorpotentials zu einem stetig differenzierbaren Vektorfeld* $\vec{v} : \mathbb{R}^3 \supset G \to \mathbb{R}^3$ *ist*

$$\operatorname{div}\vec{v} = 0.$$

Auf konvexen Mengen G ist sie auch hinreichend.
Quellfreie Felder haben also lokale Vektorpotentiale.

Beispiel 98 (Coulombfeld).
Theoretische Elektrotechnik, Abschnitt 1.4
Ein divergenzfreies Feld ohne globales Vektorpotential ist das elektrische Feld einer Punktladung, vgl. Beispiel 75. Aus Symmetriegründen müßte ein Vektorpotential nämlich von der Form

$$\vec{w} = f(r)\vec{x}$$

mit einer vom Radius abhängigen Funktion f sein. Dafür ist aber

$$\operatorname{rot} \vec{w} = \operatorname{grad}(f(r)) \times \vec{x} + f(r) \underbrace{\operatorname{rot} \vec{x}}_{=0} = \frac{f'(r)}{r}\, \vec{x} \times \vec{x} = 0.$$

Sie können aber mit einiger Mühe nachrechnen, dass

$$\vec{w}(x,y,z) = \frac{Q}{4\pi\epsilon_0 (x^2+y^2)\sqrt{x^2+y^2+z^2}} \begin{pmatrix} yz \\ -xz \\ 0 \end{pmatrix}$$

ein lokales Vektorpotential für alle Punkte liefert, die *nicht* auf der z-Achse liegen.

2.2.4 Die Wellengleichung

Durch einfaches Nachrechnen bestätigt man

$$\boxed{\operatorname{rot} \operatorname{rot} \vec{v} = \operatorname{grad} \operatorname{div} \vec{v} - \Delta \vec{v}.}$$

Beispiel 99. Wir leiten als Anwendung die Wellengleichung aus den Maxwellschen Gleichungen her. Die Maxwellschen Gleichungen im Vakuum

$$\operatorname{rot} \vec{E} = -\frac{\partial(\mu_0 \vec{H})}{\partial t} \quad \text{(Induktionsgesetz)} \tag{31}$$

$$\operatorname{rot} \vec{H} = \frac{\partial(\epsilon_0 \vec{E})}{\partial t} \quad \text{(Durchflutungsgesetz)} \tag{32}$$

$$\operatorname{div} \vec{E} = 0 \quad \text{(el. Quellenfreiheit)} \tag{33}$$

$$\operatorname{div} \vec{H} = 0 \quad \text{(magn. Quellenfreiheit)} \tag{34}$$

liefern

$$\begin{aligned} \epsilon_0 \mu_0 \frac{\partial^2 \vec{E}}{\partial t^2} &\underset{(32)}{=} \mu_0 \frac{\partial}{\partial t} \operatorname{rot} \vec{H} = \operatorname{rot} \frac{\partial(\mu_0 \vec{H})}{\partial t} \underset{(31)}{=} -\operatorname{rot} \operatorname{rot} \vec{E} \\ &= -\operatorname{grad} \operatorname{div} \vec{E} + \Delta \vec{E} \underset{(33)}{=} \Delta \vec{E} \end{aligned}$$

Die Gleichung

$$\boxed{\Delta \vec{E} = \epsilon_0 \mu_0 \frac{\partial^2 \vec{E}}{\partial t^2}}$$

heißt die *Wellengleichung für das elektrische Feld.*
Vergleichen Sie etwa Müller: Mechanik III, Popov: Mechanik III oder Theoretische Elektrotechnik, Abschnitt 5/6.

2.3 Nablakalkül. Andere Koordinaten

- Wir lernen eine andere Notation für die klassischen Differentiationsoperatoren kennen.
- Wir berechnen diese Operatoren in anderen Koordinatensystemen.

NABLAOPERATOR. Wir definieren den „Operator“

$$\nabla := \begin{pmatrix} \frac{\partial}{\partial x_1} \\ \vdots \\ \frac{\partial}{\partial x_n} \end{pmatrix}$$

und rechnen mit diesem in der „offensichtlichen“ Weise, d.h. als ob es ein ordinärer Vektor wäre. Dann finden wir

$$\begin{aligned} \operatorname{grad} u &= \nabla u \\ \operatorname{div} \vec{v} &= \nabla \cdot \vec{v} \\ \operatorname{rot} \vec{v} &= \nabla \times \vec{v}. \end{aligned}$$

Das liefert einfache Merkregeln für die obigen Differentialoperatoren. Weiter kann man sich angewöhnen, wie folgt zu rechnen:

$$\begin{aligned} \operatorname{div}(u\vec{v}) &= \nabla \cdot (u\vec{v}) = (\nabla u)\vec{v} + u\nabla \cdot \vec{v} = (\operatorname{grad} u) \cdot \vec{v} + u \operatorname{div} \vec{v}, \\ \operatorname{rot}(u\vec{v}) &= \nabla \times (u\vec{v}) = (\nabla u) \times \vec{v} + u(\nabla \times \vec{v}) = \operatorname{grad} u \times \vec{v} + u \operatorname{rot} \vec{v}. \end{aligned}$$

Hier rechnet man also mit ∇ wie mit einer vektoriellen Ableitung. Man verwendet eine „Produktregel“, deren Produkte man aber geeignet interpretieren muss. Berücksichtigt man $\vec{a} \cdot (\vec{b} \times \vec{c}) = \vec{c} \cdot (\vec{a} \times \vec{b})$ (Spatprodukt), so ergibt sich weiter:

$$\operatorname{div}(\vec{v} \times \vec{w}) = \nabla \cdot (\vec{v} \times \vec{w}) = \vec{w} \cdot (\nabla \times \vec{v}) = \vec{w} \cdot \operatorname{rot} \vec{v}$$

aber **das ist nur die halbe Wahrheit!** Die richtige Formel haben wir früher angegeben:

$$\operatorname{div}(\vec{v} \times \vec{w}) = \vec{w} \cdot \operatorname{rot} \vec{v} - \vec{v} \cdot \operatorname{rot} \vec{w}.$$

Vermutlich ist der Nablaoperator doch keine so große Hilfe: Mit ihm lässt sich nur dann einfach rechnen, wenn man vorher weiß, was herauskommen soll.

ANDERE KOORDINATEN. Bei Problemen mit gewissen Symmetrien verwendet man häufig andere Koordinatensysteme. Zum Beispiel lässt sich das zentrale Potential in Kugelkoordinaten viel einfacher schreiben, als in Kartesischen Koordinaten:

$$\frac{1}{r} \text{ ist einfacher als } \frac{1}{\sqrt{x^2 + y^2 + z^2}}.$$

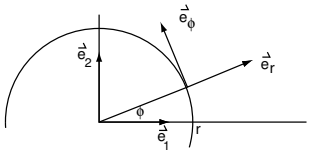
Es ist deshalb wünschenswert, Formeln für grad, div und rot in diesen Koordinatensystemen zu haben. Allerdings ist es dabei angemessen, auch die Vektorfeldern nicht durch ihre Komponenten $(v_1, \ldots, v_n)$ bezüglich der Kartesischen Basisvektoren $\vec{e}_1, \ldots, \vec{e}_n$ zu beschreiben, sondern bezüglich einer Basis, die an das jeweilige Koordinatensystem angepasst ist. Dabei wählt man gern orthogonale Koordinatensysteme. Ein wesentlicher Unterschied ist aber, dass diese Basen nicht mehr konstant sind, sondern vom jeweiligen Punkt abhängen, in dem man das Vektorfeld betrachtet. Das wird deutlicher an einem Beispiel: In ebenen Polarkoordinaten[1] verwendet man

$$\vec{e}_\rho := \begin{pmatrix} \cos\phi \\ \sin\phi \end{pmatrix} = \begin{pmatrix} \frac{x}{\rho} \\ \frac{y}{\rho} \end{pmatrix}, \quad \vec{e}_\phi := \begin{pmatrix} -\sin\phi \\ \cos\phi \end{pmatrix} = \begin{pmatrix} -\frac{y}{\rho} \\ \frac{x}{\rho} \end{pmatrix}.$$

[1] Vergleiche Abschnitt 1.11

Dann ist

$$\vec{e}_1 = \cos\phi\vec{e}_\rho - \sin\phi\vec{e}_\phi = \frac{x}{\rho}\vec{e}_\rho - \frac{y}{\rho}\vec{e}_\phi,$$
$$\vec{e}_2 = \sin\phi\vec{e}_\rho + \cos\phi\vec{e}_\phi = \frac{y}{\rho}\vec{e}_\rho + \frac{x}{\rho}\vec{e}_\phi.$$



Beispiel 100 (Magnetfeld des unendlich langen geraden Leiters). Das Vektorfeld

$$\vec{H} = \frac{I}{2\pi}\left(\frac{-y}{x^2+y^2}, \frac{x}{x^2+y^2}\right) = \frac{I}{2\pi}\left(\frac{-y}{x^2+y^2}\,\vec{e}_1 + \frac{x}{x^2+y^2}\,\vec{e}_2\right)$$

schreibt sich als

$$\vec{H} = \frac{I}{2\pi}\frac{-y}{x^2+y^2}\left(\frac{x}{\rho}\vec{e}_\rho - \frac{y}{\rho}\vec{e}_\phi\right) + \frac{I}{2\pi}\frac{x}{x^2+y^2}\left(\frac{y}{\rho}\vec{e}_\rho + \frac{x}{\rho}\vec{e}_\phi\right) = \frac{I}{2\pi\rho}\,\vec{e}_\phi.$$

Im dreidimensionalen Raum mit Zylinderkoordinaten hat das (wahre) Magnetfeld des geraden Leiters dann noch eine Komponente $= 0$ in z-Richtung.

Wir führen nun beispielhaft die Berechnung von Gradient und Divergenz in ebenen Polarkoordinaten vor und geben anschließend die Resultate für Zylinder- und Kugelkoordinaten ohne Rechnung an.

Berechnung des Gradienten in ebenen Polarkoordinaten. Wegen

$$\frac{\partial\rho}{\partial x} = \frac{x}{\rho}, \quad \frac{\partial\phi}{\partial x} = \frac{-y}{\rho^2},$$
$$\frac{\partial\rho}{\partial y} = \frac{y}{\rho}, \quad \frac{\partial\phi}{\partial y} = \frac{x}{\rho^2},$$

ergibt sich der Gradient einer Funktion $u(\rho,\phi)$ mit der Kettenregel so:

$$\begin{aligned}
\operatorname{grad} u &= \frac{\partial u}{\partial x}\begin{pmatrix}1\\0\end{pmatrix} + \frac{\partial u}{\partial y}\begin{pmatrix}0\\1\end{pmatrix} \\
&= \left(\frac{\partial u}{\partial\rho}\frac{\partial\rho}{\partial x} + \frac{\partial u}{\partial\phi}\frac{\partial\phi}{\partial x}\right)\begin{pmatrix}1\\0\end{pmatrix} + \left(\frac{\partial u}{\partial\rho}\frac{\partial\rho}{\partial y} + \frac{\partial u}{\partial\phi}\frac{\partial\phi}{\partial y}\right)\begin{pmatrix}0\\1\end{pmatrix} \\
&= \frac{\partial u}{\partial\rho}\left(\begin{pmatrix}\frac{x}{\rho}\\0\end{pmatrix} + \begin{pmatrix}0\\\frac{y}{\rho}\end{pmatrix}\right) + \frac{\partial u}{\partial\phi}\left(\begin{pmatrix}\frac{-y}{\rho^2}\\0\end{pmatrix} + \begin{pmatrix}0\\\frac{x}{\rho^2}\end{pmatrix}\right) \\
&= \frac{\partial u}{\partial\rho}\vec{e}_\rho + \frac{1}{\rho}\frac{\partial u}{\partial\phi}\vec{e}_\phi.
\end{aligned}$$

Berechnung der Divergenz in ebenen Polarkoordinaten. Sei andrerseits ein Vektorfeld in der Form

$$\vec{v}(\rho,\phi) = v_\rho\vec{e}_\rho + v_\phi\vec{e}_\phi$$

gegeben und $\operatorname{div}\vec{v}$ gesucht. Dann ist

$$\begin{aligned}
\vec{v} &= v_\rho(\cos\phi\vec{e}_1 + \sin\phi\vec{e}_2) + v_\phi(-\sin\phi\vec{e}_1 + \cos\phi\vec{e}_2) \\
&= (v_\rho\cos\phi - v_\phi\sin\phi)\vec{e}_1 + (v_\rho\sin\phi + v_\phi\cos\phi)\vec{e}_2.
\end{aligned}$$

Es folgt

$$\begin{aligned}
\operatorname{div}\vec{v} =& \frac{\partial}{\partial x}(v_\rho\cos\phi - v_\phi\sin\phi) + \frac{\partial}{\partial y}(v_\rho\sin\phi + v_\phi\cos\phi)\\
=& \left(\frac{\partial\rho}{\partial x}\frac{\partial}{\partial\rho} + \frac{\partial\phi}{\partial x}\frac{\partial}{\partial\phi}\right)(v_\rho\cos\phi - v_\phi\sin\phi) + \left(\frac{\partial\rho}{\partial y}\frac{\partial}{\partial\rho} + \frac{\partial\phi}{\partial y}\frac{\partial}{\partial\phi}\right)(v_\rho\sin\phi + v_\phi\cos\phi)\\
=& \frac{x}{\rho}\frac{\partial v_\rho}{\partial\rho}\cos\phi - \frac{x}{\rho}\frac{\partial v_\phi}{\partial\rho}\sin\phi + \frac{-y}{\rho^2}\left(\frac{\partial v_\rho}{\partial\phi}\cos\phi - v_\rho\sin\phi - \frac{\partial v_\phi}{\partial\phi}\sin\phi - v_\phi\cos\phi\right)\\
&+ \frac{y}{\rho}\frac{\partial v_\rho}{\partial\rho}\sin\phi + \frac{y}{\rho}\frac{\partial v_\phi}{\partial\rho}\cos\phi + \frac{x}{\rho^2}\left(\frac{\partial v_\rho}{\partial\phi}\sin\phi + v_\rho\cos\phi + \frac{\partial v_\phi}{\partial\phi}\cos\phi - v_\phi\sin\phi\right)\\
=& \frac{\partial v_\rho}{\partial\rho} + \frac{1}{\rho}\frac{\partial v_\phi}{\partial\phi} + \frac{1}{\rho}v_\rho.
\end{aligned}$$

Beim $\vec{v}$ aus dem vorstehenden Beispiel ist $v_\rho = 0, v_\phi = \frac{1}{\rho}$, und deshalb $\operatorname{div}\vec{v} = 0$ ohne jede Rechnung.

Differentialoperatoren in Zylinderkoordinaten. Bei Zylinderkoordinaten (ρ,ϕ,z) im $\mathbb{R}^3$ benutzt man die Basis $(\vec{e}_\rho,\vec{e}_\phi,\vec{e}_z)$ mit $\vec{e}_z = \vec{e}_3$. Es ergeben sich die folgenden Formeln im $\mathbb{R}^3$:

$$\begin{aligned}
\operatorname{grad} u &= \frac{\partial u}{\partial\rho}\vec{e}_\rho + \frac{1}{\rho}\frac{\partial u}{\partial\phi}\vec{e}_\phi + \frac{\partial u}{\partial z}\vec{e}_z,\\
\operatorname{div}\vec{v} &= \frac{\partial v_\rho}{\partial\rho} + \frac{1}{\rho}\frac{\partial v_\phi}{\partial\phi} + \frac{\partial v_z}{\partial z} + \frac{1}{\rho}v_\rho,\\
\operatorname{rot}\vec{v} &= \left(\frac{1}{\rho}\frac{\partial v_z}{\partial\phi} - \frac{\partial v_\phi}{\partial z}\right)\vec{e}_\rho + \left(\frac{\partial v_\rho}{\partial z} - \frac{\partial v_z}{\partial\rho}\right)\vec{e}_\phi + \left(\frac{\partial v_\phi}{\partial\rho} - \frac{1}{\rho}\frac{\partial v_\rho}{\partial\phi} + \frac{1}{\rho}v_\phi\right)\vec{e}_z,\\
\Delta u &= \frac{\partial^2 u}{\partial\rho^2} + \frac{1}{\rho}\frac{\partial u}{\partial\rho} + \frac{1}{\rho^2}\frac{\partial^2 u}{\partial\phi^2} + \frac{\partial^2 u}{\partial z^2}.
\end{aligned}$$

Das folgende Beispiel beschäftigt sich noch einmal mit dem Laplaceoperator und diskutiert ein Standardverfahren zur Gewinnung von Lösungen der Schwingungsgleichung im räumlich eindimensionalen und zweidimensionalen Fall, also etwa die Schwingungen einer Saite oder einer Membran. Für rechteckige Membrane sind die Kartesischen Koordinaten angemessen, für kreisförmige Membrane sind Polarkoordinaten vorzuziehen.

Beispiel 101 (Separation in Polarkoordinaten).
Müller: Mechanik III, Abschnitt 17 Popov: Mechanik III, Theoretische Elektrotechnik, Abschnitt 5/6
Sucht man Lösungen der eindimensionalen Schwingungsgleichung

$$\frac{\partial^2 u}{\partial t^2} = \frac{\partial^2 u}{\partial x^2}$$

(bei der wir zur Vereinfachung $\lambda = 1$ gesetzt haben), so kann man den auf Euler und Bernoulli zurückgehenden *Separationsansatz* benutzen: Man nimmt an, dass die Lösung von der Form

$$u(x,t) = g(x)h(t)$$

ist.
Beachten Sie: Nicht jede Lösung ist von dieser Form, aber sehr viele, und durch *Kombination* verschiedener solcher Lösungen erhält man „alle“.
Setzt man den Ansatz in die Differentialgleichung ein, so findet man

$$g(x)h''(t) = g''(x)h(t).$$

Division mit $g(x)h(t)$ liefert

$$\frac{h''(t)}{h(t)} = \frac{g''(x)}{g(x)}.$$

Die linke Seite hängt nur von t ab, die rechte nur von x. Ändert man t, so ändert sich die rechte Seite nicht, aber dann ändert sich auch die linke Seite nicht! Beide Seiten sind also konstant:

$$\frac{h''(t)}{h(t)} = -\omega^2 = \frac{g''(x)}{g(x)}.$$

Wir haben der Einfachheit halber angenommen, dass die Konstante negativ ist, dann kann man sie als $-\omega^2$ schreiben. Positive Konstanten liefern *weitere* Lösungen. Der Separationsansatz vereinfacht die partielle Differentialgleichung zu einem nicht gekoppelten System gewöhnlicher Differentialgleichungen

$$h'' + \omega^2 h = 0,$$
$$g'' + \omega^2 g = 0.$$

Diese haben Sinus-Cosinus Lösungen, aus denen man zum Beispiel

$$u(x,t) = A\cos\omega(t-t_0)\cos\omega(x-x_0)$$

mit beliebigen Konstanten A, t_0, x_0 erhält.
Nun betrachten wir den Fall einer schwingenden kreisförmigen Platte. Die Schwingungsgleichung ist (wieder mit zu 1 normierten Parametern)

$$\frac{\partial^2 u}{\partial t^2} = \Delta u.$$

Sucht man rotationssymmetrische, d.h. von ϕ unabhängige Lösungen, so bietet sich die Verwendung von Polarkoordinaten und der Ansatz

$$u(\rho,t) = g(\rho)h(t)$$

an. Nun benutzt man den Laplaceoperator in Polarkoordinaten, d.h. in „Zylinderkoordinaten ohne z“:

$$g(\rho)h''(t) = g''(\rho)h(t) + \frac{1}{\rho}g'(\rho)h(t).$$

Dasselbe Argument wie oben führt zu

$$\frac{h''(t)}{h(t)} = -\omega^2 = \frac{g''(\rho)}{g(\rho)} + \frac{1}{\rho}\frac{g'(\rho)}{g(\rho)}.$$

Wieder ergibt sich für die Zeitfunktion eine Sinus-Cosinus-Gleichung

$$h'' + \omega^2 h = 0,$$

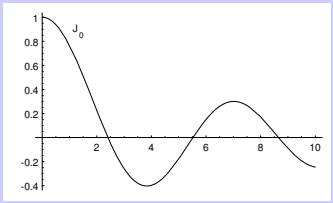
aber für die Radialkomponente ergibt sich eine kompliziertere Differentialgleichung:

$$\rho g'' + g' + \omega^2 \rho g = 0.$$

Die beschränkten Lösungen davon (die unbeschränkten sind aus physikalischen Gründen uninteressant) sind

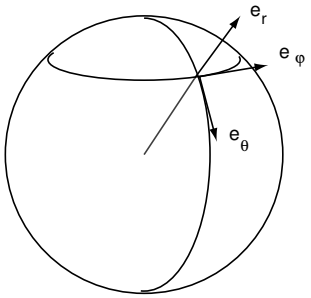
$$g(\rho) = AJ_0(\omega\rho),$$

wobei A eine beliebige Konstante und J_0 die Besselfunktion der Ordnung 0 ist.



DIFFERENTIALOPERATOREN IN KUGELKOORDINATEN. Bei Kugelkoordinaten verwendet man die Basis

$$\begin{aligned}
\vec{e}_r &= \frac{1}{r}(x\vec{e}_1 + y\vec{e}_2 + z\vec{e}_3)\\
\vec{e}_\theta &= \cos\theta\cos\phi\vec{e}_1 + \cos\theta\sin\phi\vec{e}_2 - \sin\theta\vec{e}_3\\
\vec{e}_\phi &= \frac{1}{r\sin\theta}(-y\vec{e}_1 + x\vec{e}_2).
\end{aligned}$$



Man findet

$$\begin{aligned}
\operatorname{grad} u =& \frac{\partial u}{\partial r}\vec{e}_r + \frac{1}{r\sin\theta}\frac{\partial u}{\partial\phi}\vec{e}_\phi + \frac{1}{r}\frac{\partial u}{\partial\theta}\vec{e}_\theta,\\
\operatorname{div}\vec{v} =& \frac{1}{r^2}\frac{\partial(r^2 v_r)}{\partial r} + \frac{1}{r\sin\theta}\frac{\partial v_\phi}{\partial\phi} + \frac{1}{r\sin\theta}\frac{\partial(v_\theta\sin\theta)}{\partial\theta},\\
\operatorname{rot}\vec{v} =& \frac{1}{r\sin\theta}\left(\frac{\partial(v_\phi\sin\theta)}{\partial\theta} - \frac{\partial v_\theta}{\partial\phi}\right)\vec{e}_r + \frac{1}{r}\left(\frac{1}{\sin\theta}\frac{\partial v_r}{\partial\phi} - \frac{\partial(rv_\phi)}{\partial r}\right)\vec{e}_\theta\\
&+ \frac{1}{r}\left(\frac{\partial(rv_\theta)}{\partial r} - \frac{\partial v_r}{\partial\theta}\right)\vec{e}_\phi,\\
\Delta u =& \frac{\partial^2 u}{\partial r^2} + \frac{2}{r}\frac{\partial u}{\partial r} + \frac{1}{r^2\sin^2\theta}\frac{\partial^2 u}{\partial\phi^2} + \frac{1}{r^2\sin\theta}\frac{\partial(\sin\theta\frac{\partial u}{\partial\theta})}{\partial\theta}.
\end{aligned}$$

Beachten Sie, dass diese Formeln zwar einigermaßen kompliziert aussehen, dass aber andrerseits häufig die Vektorfelder in diesen Koordinaten besonders einfach aussehen, so dass

die Berechnung der Differentialoperatoren dann doch sehr einfach ist. Das zeigen die beiden nachstehenden Beispiele:

Beispiel 102. Für $u(r, \theta, \phi) = r^k$ erhält man

$$\Delta u = k(k-1)r^{k-2} + \frac{2}{r}kr^{k-1} = k(k+1)r^{k-2}.$$

(Beachte: $n = 3$.) □

Beispiel 103 (Coulombfeld).
Theoretische Elektrotechnik, Abschnitt 1.4
Zum Potential $U = \frac{Q}{4\pi\epsilon r}$ der Punktladung Q erhält man das elektrische Feld

$$\vec{E} = -\operatorname{grad} U = \frac{Q}{4\pi\epsilon r^2}\vec{e}_r.$$

Ohne große Rechnung sieht man dann sofort

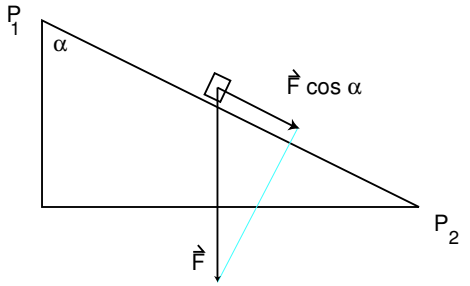
$$\operatorname{div} \vec{E} = \frac{1}{r^2}\frac{\partial(\frac{Q}{4\pi\epsilon})}{\partial r} = 0,$$
$$\operatorname{rot} \vec{E} = 0.$$

2.4 Kurvenintegrale

- Wir lernen, wie man Vektorfelder über eine Kurve integriert und wie man die Energie eines Massenpunktes berechnet, der sich in einem Kraftfeld bewegt.
- Wir lernen, wie man skalare Funktionen über Kurven integriert und insbesondere, wie man die Länge einer Kurve berechnet.
- Kurvenintegrale liefern auch das Modell für die Beschreibung thermodynamischer Prozesse.

In diesem Abschnitt werden Vektorfelder vorwiegend als Kraftfelder interpretiert. Wir bezeichnen sie daher mit dem gebräuchlichen Buchstaben $\vec{F}$ (für *force*).

Bewegt sich ein Massenpunkt in einem Kraftfeld $\vec{F}$ vom Punkt $\vec{P}_1$ zum Punkt $\vec{P}_2$, so gewinnt er bei dieser Bewegung (kinetische) Energie. Im homogenen Kraftfeld ist die Energie gegeben durch die Streckenlänge mal die Komponente der Kraft in Richtung der Strecke: $A = \vec{F} \cdot (\vec{P}_2 - \vec{P}_1)$.



Ist das Kraftfeld aber räumlich veränderlich und bewegt sich der Massenpunkt auf einer Kurve von $\vec{P}_1 = \vec{x_0}$ über die Zwischenpunkte $\vec{x_i}$ nach $\vec{P}_2 = \vec{x_n}$, so ist die Energie näherungsweise

$$A = \sum_{i=1}^{n} \vec{F}(\vec{x_i}) \cdot (\vec{x_i} - \vec{x}_{i-1}).$$

Ist die Kurve gegeben durch eine *Parametrisierung*

$$\vec{x} : [a, b] \to \mathbb{R}^3,$$

und wählt man die Zwischenpunkte $\vec{x}_i = \vec{x}(t_i)$ immer feiner, so ergibt sich im Grenzwert

$$\begin{aligned} A &= \lim_{n\to\infty} \sum_{i=1}^{n} \vec{F}(\vec{x}(t_i)) \cdot \frac{\vec{x}(t_i) - \vec{x}(t_{i-1})}{t_i - t_{i-1}} (t_i - t_{i-1}) \\ &= \int_a^b \vec{F}(\vec{x}(t)) \cdot \dot{\vec{x}}(t) dt. \end{aligned}$$

Das führt zu folgender

Definition 104. Das Integral des Vektorfeldes $\vec{F}$ über die Kurve $\vec{x} : [a, b] \to \mathbb{R}^n$ ist definiert durch

$$\boxed{\int_{\vec{x}} \vec{F} \cdot d\vec{s} := \int_a^b \vec{F}(\vec{x}(t)) \cdot \dot{\vec{x}}(t) dt.}$$

Dabei setzen wir offenbar voraus, dass $\vec{x}(t)$ differenzierbar ist. Wenn wir überdies annehmen, dass $\vec{F}$ und $\dot{\vec{x}}$ stetig sind, so existiert das Integral. $\dot{\vec{x}}$ ist der Geschwindigkeitsvektor, und weil „Geschwindigkeit x Zeit = zurückgelegte Strecke“, ist $\dot{\vec{x}}dt$ das vektorielle Streckenelement. Daher die Bezeichnung $d\vec{s}$.

In der Praxis kommen oft Kurven vor, die nur stückweise eine stetige Ableitung haben, die also durch Aneinanderhängen von endlich vielen Kurven mit stetiger Ableitung entstehen.

Ein Beispiel ist die Integration über den Rand eines Dreiecks. Dafür addiert man einfach die entsprechenden Kurvenintegrale.

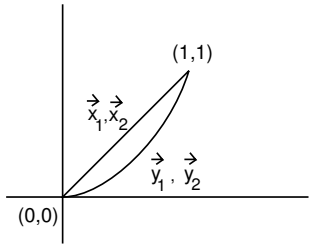
Beispiel 105.
Theoretische Elektrotechnik, Abschnitt 1.4
Eine Punktladung Q erfährt im elektrischen Feld $\vec{E}$ die Kraft $Q\vec{E}$, und $Q\int_{\vec{x}} \vec{E}\cdot\vec{ds}$ ist die auf dem Weg $\vec{x}$ gewonnene Energie.

Beispiel 106. Sei $\vec{F}(x,y) = (x^2y, x+y)$.

Wir betrachten vier Kurven zwischen $(0,0)$ und $(1,1)$, nämlich

$$\vec{x}_1(t) := (t,t), \quad \vec{x}_2(t) := (1-t, 1-t),$$
$$\vec{y}_1(t) := (t,t^2), \quad \vec{y}_2(t) := (\sqrt{t}, t)$$

für $0 \le t \le 1$.



Dann erhalten wir

$$\int_{\vec{x}_1} \vec{F}\cdot\vec{ds} = \int_0^1 (t^3, 2t)\cdot(1,1)dt = \int_0^1 (t^3+2t)dt = \frac{5}{4},$$
$$\int_{\vec{x}_2} \vec{F}\cdot\vec{ds} = \int_0^1 ((1-t)^3, 2(1-t))\cdot(-1,-1)dt$$
$$\underset{1-t=u}{=} \int_1^0 (u^3, 2u)\cdot(-1,-1)(-du) = -\int_0^1 (u^3+2u)du = -\frac{5}{4},$$
$$\int_{\vec{y}_1} \vec{F}\cdot\vec{ds} = \int_0^1 (t^4, t+t^2)\cdot(1,2t)dt = \int_0^1 (t^4+2t^2+2t^3)dt = \frac{41}{30},$$
$$\int_{\vec{y}_2} \vec{F}\cdot\vec{ds} = \int_0^1 (t^2, \sqrt{t}+t)\cdot(\frac{1}{2\sqrt{t}}, 1)dt = \int_0^1 (\frac{1}{2}t^{3/2} + t + \sqrt{t})dt = \frac{41}{30}.$$

□

In diesem Beispiel sehen wir:

- Das Kurvenintegral eines Vektorfeldes ist im allgemeinen nicht nur vom Anfangs- und Endpunkt der Kurve abhängig, sondern davon, wie die Kurve dazwischen verläuft.
- Integration über „dieselbe“ Kurve in entgegengesetzter Richtung liefert den negativen Wert für das Integral.
- Integration über „dieselbe“ Kurve in verschiedenen gleichgerichteten Parametrisierungen liefert denselben Wert für das Integral.

Die beiden letzten Tatsachen gelten allgemein: Das Kurvenintegral hängt nicht von der Parametrisierung ab, es ist „parameterinvariant“, solange man die Orientierung erhält. Wir beweisen das zur (Rechen-)Übung. Sei

$$\vec{x}_2(t) = \vec{x}_1(h(t)), \quad \alpha \le t \le \beta$$

und

$$a = h(\alpha) \le h(t) \le h(\beta) = b.$$

Dann ist

$$\begin{aligned}\int_{\vec{x}_2} \vec{F}d\vec{s} &= \int_\alpha^\beta \vec{F}(\vec{x}_2(t))\cdot\dot{\vec{x}}_2(t)dt = \int_\alpha^\beta \vec{F}(\vec{x}_1(h(t)))\cdot\dot{\vec{x}}_1(h(t))\dot{h}(t)dt\\ &= \int_{h(\alpha)}^{h(\beta)} \vec{F}(\vec{x}_1(\tau))\cdot\dot{\vec{x}}_1(\tau)d\tau \underset{(*)}{=} \int_a^b \vec{F}(\vec{x}_1(\tau))\cdot\dot{\vec{x}}_1(\tau)d\tau\\ &= \int_{\vec{x}_1} \vec{F}d\vec{s}.\end{aligned}$$

Kehrt h hingegen die Orientierung um, d.h. ist

$$b = h(\alpha) \geq h(t) \geq h(\beta) = a,$$

so erbt man der Stelle $(*)$ ein Minuszeichen. □

Bei Potentialfeldern (d.h. wenn es eine Stammfunktion gibt!) hängt das Integral nicht mehr davon ab, wie man vom Anfangspunkt zum Endpunkt kommt, es ist „wegunabhängig“. (Aber nur bei Wegen mit gleichem Anfangspunkt und gleichem Endpunkt!)

Beispiel 107. Sei $\vec{F} = -\operatorname{grad} u$ ein Vektorfeld mit Potential u. Dann gilt

$$\begin{aligned}\int_{\vec{x}} \vec{F}\cdot\vec{ds} &= \int_a^b \vec{F}(\vec{x}(t))\cdot\dot{\vec{x}}(t)dt = \int_a^b \left(\sum_i F_i(\vec{x}(t))\dot{x}_i(t)\right)dt\\ &= \int_a^b \left(\sum_i -\frac{\partial u}{\partial x_i}(\vec{x}(t))\dot{x}_i(t)\right)dt = -\int_a^b \left(\sum_i \frac{\partial u}{\partial x_i}(\vec{x}(t))\frac{dx_i}{dt}\right)dt\\ &\underset{(Kettenregel)}{=} -\int_a^b \left(\frac{du(\vec{x}(t))}{dt}\right)dt = u(\vec{x}(a)) - u(\vec{x}(b)).\end{aligned}$$

Also

$$\int_{\vec{x}} \vec{F}\cdot\vec{ds} = u(Anfangspunkt) - u(Endpunkt).$$

Die freiwerdende Energie ist gleich der Potentialdifferenz an den Endpunkten der Kurve.

Man sagt:

Satz 108. *Das Kurvenintegral von Potentialfeldern ist wegunabhängig, nämlich gleich der Potentialdifferenz zwischen Anfangs- und Endpunkt des Weges.*

Hiervon gilt auch die Umkehrung: Ist $\vec{F}$ ein auf einer offenen und wegzusammenhängenden Teilmenge $G \subset \mathbb{R}^n$ definiertes Vektorfeld und ist das Kurvenintegral wegunabhängig für alle Kurven in G, so besitzt $\vec{F}$ in G ein Potential, nämlich

$$u(\vec{x_1}) := -\int_{\vec{x}} \vec{F}\cdot\vec{ds},$$

wobei $\vec{x_0}$ ein beliebig gewählter fester Punkt in G und $\vec{x}$ eine Kurve von $\vec{x}_0$ nach $\vec{x}_1$ in G ist. Nach Voraussetzung ist es egal, welche Kurve man dafür wählt. Ändert man $\vec{x}_0$, so ändert sich das Potential um eine Konstante, bleibt also ein Potential. Natürlich müsste man nachrechnen, dass u die richtigen partiellen Ableitungen hat. Das ist relativ einfach, wenn man für $\frac{\partial u}{\partial x_i}$ die Kurve $\vec{x}$ so wählt, dass sie auf dem letzten Stück einfach parallel zur x_i-Achse verläuft; aber wir führen den Beweis nicht vor.

Satz 109. *Vektorfelder mit wegunabhängigem Integral besitzen ein Potential.*

INTEGRATION SKALARER FUNKTIONEN ÜBER KURVEN. Wir definieren noch einen anderen Begriff von „Kurvenintegral“. Wenn wir eine Kurve $\vec{x} : [a, b] \to \mathbb{R}^n$ gegeben haben, und weiter eine Funktion $\rho : \vec{x}([a, b]) \to \mathbb{R}$, die wir uns etwa als Massendichte vorstellen, so definieren wir

$$\int_{\vec{x}} \rho \, ds := \int_a^b \rho(\vec{x}(t))|\dot{\vec{x}}(t)| \, dt.$$

Für den Fall $\rho = 1$ erhält man

$$\text{Länge}(\vec{x}) := \int_a^b ds = \int_a^b |\dot{\vec{x}}| \, dt.$$

Wenn Sie t als Zeitparameter interpretieren, dann ist das also die Formel

$$\text{„Strecke} = \text{Geschwindigkeit} \times \text{Zeit“}$$

in integraler Form.

Beispiel 110.
Müller: Mechanik I, Abschnitt 5
Wir berechnen den Schwerpunkt des Halbkreisbogens

$$C := \{(r \cos t, r \sin t) \mid 0 \leq t \leq \pi\}.$$

Die x-Koordinate ist aus Symmetriegründen offenbar $x_S = 0$. Die y-Koordinate ist gegeben durch

$$y_S = \frac{\int_C y ds}{\int_C ds} = \frac{\int_0^\pi r \sin t \sqrt{r^2 \sin^2 t + r^2 \cos^2 t} dt}{\int_0^\pi \sqrt{r^2 \sin^2 t + r^2 \cos^2 t} dt} = \frac{2r^2}{\pi r} = \frac{2}{\pi} r.$$

Wir schließen diesen Abschnitt mit einer anderen Formulierung des Kurvenintegrals über Felder, die inhaltlich aber nichts Neues bringt:

Beispiel 111 (Felder und Integrale in der Thermodynamik).
Thermodynamik I
In der Thermodynamik (und anderenorts) verwendet man statt der Vektorfelder

$$\vec{v} = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} \text{ die Pfaffschen Formen } v_1\,dx_1 + \ldots + v_n\,dx_n,$$

vergleichen Sie Beispiel 37.
Das Vektorfeld $\vec{v}$ hat ein Potential, wenn es eine Funktion u gibt, so dass

$$\vec{v} = -\operatorname{grad} u,$$

d.h.

$$-\frac{\partial u}{\partial x_i} = v_i$$

Das bedeutet für die Pfaffsche Form aber gerade, dass sie ein vollständiges Differential ist:

$$d(-u) = v_1\,dx_1 + \ldots + v_n\,dx_n.$$

Das Kurvenintegral einer Pfaffschen Form ist definiert als

$$\int_{\vec{x}} \sum_{i=1}^{n} v_i\,dx_i := \int_a^b \left(\sum_{i=1}^{n} v_i(\vec{x}(t))\,\frac{dx_i}{dt}(t) \right) dt.$$

Es ist also gerade das Kurvenintegral des entsprechenden Vektorfeldes.
Vektorfelder oder Pfaffsche Formen sind nur zwei verschiedene Schreibweisen für dasselbe mathematische Objekt, es gibt gewissermaßen ein Lexikon, um eine Notation in die andere zu übersetzen.

Elektrodynamik/Mechanik	Thermodynamik
Vektorfeld	Pfaffsche Form
Potentialfeld	Vollständiges Differential
wirbelfrei	$\frac{\partial v_i}{\partial x_j} = \frac{\partial v_j}{\partial x_i}$ (geschlossen)
Potential(=wegunabhängiges Integral)	Zustandsvariable
wegabhängiges Integral	Prozessvariable

3 Mehrdimensionale Integration

3.1 Integration von Funktionen in mehreren Variablen

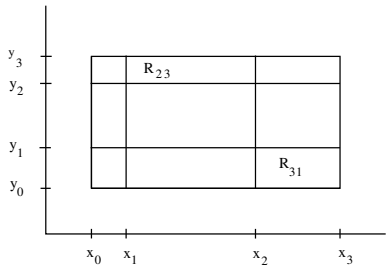
- Wie das eindimensionale Integral als Mittel zur Flächenberechnung unter einem Funktionsgraphen eingeführt wurde, kann man das zweidimensionale Integral verstehen als Mittel zur Volumenberechnung. Aber was soll dann ein dreidimensionales Integral bedeuten?
- Dazu interpretieren wir das zweidimensionale Integral einer Funktion δ als Gesamtladung einer Fläche, auf der eine Ladungsdichte δ verteilt ist. Entsprechend kann man sich ein dreidimensionales Integral dann vorstellen als Gesamtladung einer durch den Integranden gegebenen räumlichen Ladungsverteilung.

Das Integral von (positiven reellwertigen) Funktionen einer Variablen dient der Berechnung von Flächeninhalten unter dem Graphen dieser Funktionen. Entsprechend kann man Volumina unter dem Graphen von (positiven reellwertigen) Funktionen zweier Variabler mit einem erweiterten Integralbegriff berechnen. Wir beginnen mit einer „Treppenfunktion“ auf einem Rechteck R im $\mathbb{R}^2$. Wir nehmen an, dass dieses Rechteck durch achsenparalle Streifen bei x_i bzw. y_j in kleine Rechtecke R_{ij} zerlegt ist, vgl. die Abbildung.

Wir schreiben

$$\Delta x_i := x_i - x_{i-1},\ \Delta y_j := y_j - y_{j-1}.$$

Sei nun $f : R \to \mathbb{R}$ eine Treppenfunktion auf R, d.h. auf dem Innern eines jeden kleinen Rechtecks R_{ij} konstant. Die Werte von f auf den Rechteck-Rändern interessieren uns nicht weiter, aber wir wollen voraussetzen, dass f beschränkt ist.



Ist überdies $f \geq 0$, so ist das Volumen unter dem Graphen von f wie folgt zu berechnen: Sei (ξ_{ij}, η_{ij}) ein beliebiger Punkt im Innern des Rechtecks R_{ij}. Dann ist $f(\xi_{ij}, \eta_{ij})\Delta x_i \Delta y_j$ das Volumen über dem Rechteck R_{ij}, und das Gesamtvolumen ist

$$\iint_R f(x,y)dxdy := \sum_j \sum_i f(\xi_{ij}, \eta_{ij})\Delta x_i \Delta y_j. \tag{35}$$

Wir benutzen (35) als Definition für das Integral einer Treppenfunktion f, auch wenn diese wechselndes Vorzeichen hat.

Wir erweitern nun den Integralbegriff auf solche Funktionen – einstweilen definiert auf dem Rechteck R –, die sich “gut durch Treppenfunktionen einschließen lassen”: Zu jedem $k \in \mathbb{N}$ soll es Treppenfunktionen f_k und f^k geben, so dass

$$f_k(x,y) \leq f(x,y) \leq f^k(x,y) \quad \text{für alle } (x,y) \in R$$

ist, und der Bereich zwischen f_k und f^k ein Volumen $< \frac{1}{k}$ hat:

$$\iint_R f^k(x,y)dxdy - \iint_R f_k(x,y)dxdy < \frac{1}{k}. \tag{36}$$

Wir sagen dann, dass die Folge $(f_k)_{k>0}$ die Funktion f gut approximiert und setzen

$$\iint\limits_R f(x,y)dxdy = \lim_{k\to\infty} \iint\limits_R f_k(x,y)dxdy \left(= \lim_{k\to\infty} \iint\limits_R f^k(x,y)dxdy \right).$$

Natürlich ist das nur sinnvoll, weil man beweisen kann, dass dieser Limes tatsächlich immer existiert und dass er unabhängig davon ist, welche Folge (f_k) von gut approximierenden Treppenfunktionen man für die Approximation wählt.

Sei schließlich $f : B \to \mathbb{R}$ definiert auf einer kompakten Menge $B \subset \mathbb{R}^2$. „Kompakt“ hieß: abgeschlossen und beschränkt, vergleichen Sie Abschnitt 1.2. Wenn Sie irgendwo der Mandelbrod-Menge (Apfelmännchen) begegnet sind oder sich an das Beispiel 14 erinnern, wissen Sie vielleicht, dass der Rand von kompakten Mengen sehr „wild“ aussehen kann. Wir wollen deshalb für das Kapitel 3 folgende Vereinbarung treffen:

Generalvoraussetzung. Wir betrachten nur kompakte Mengen im $\mathbb{R}^2$ bzw. $\mathbb{R}^3$, deren Rand stückweise glatt ist, d.h. aus endlich vielen glatten Kurven bzw. Flächen besteht.

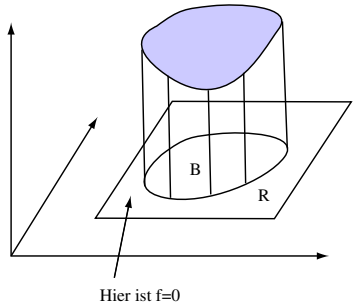
Beispiele sind Kreise, Rechtecke, Dreiecke bzw. Kugeln, Quader, Würfel,...

Diese Generalvorausetzung hat den Vorteil, dass man sie ruhig vergessen kann: Einerseits kommen in Ihrer Praxis kaum andere kompakte Mengen vor, andrerseits bleiben die im weiteren über kompakte Mengen gemachten Aussagen im wesentlichen richtig ohne die Generalvoraussetzung, wenn man den allgemeineren Lebesgueschen Integralbegriff verwendet.

Die Beschränktheit von B bedeutet, dass B enthalten ist in einem Rechteck R, und wir definieren einfach $f(x,y) := 0$, falls $(x,y) \in R\backslash B$. Wir setzen dann

$$\iint\limits_B f(x,y)dxdy := \iint\limits_R f(x,y)dxdy,$$

falls das rechte Integral existiert.



Integrierbare Funktionen. Auf diese Weise ist das Integral über die kompakte Menge B erklärt für eine gewisse Klasse von Funktionen, die wir *integrierbare Funktionen* nennen wollen.

Alle beschränkten Funktionen mit *nicht zuviel Unstetigkeit* sind integrierbar.

Stetige Funktionen haben gar keine Unstetigkeit und lassen sich deshalb integrieren. Aber auch manche unstetigen Funktionen lassen sich integrieren, Treppenfunktionen zum Beispiel. Bei den letzteren liegen die Unstetigkeitsstellen in den Ränder der kleinen Rechtecke. Allgemein soll *nicht zuviel Unstetigkeit* genau bedeuten:

Die Unstetigkeitsmenge ist enthalten in der Vereinigung von endlich vielen glatten Kurven.

Gelegentlich muss man vektorwertige Funktionen integrieren:

Definition 112 (Integration vektorwertige Funktionen). Vektor*wertige* Funktionen integriert man komponenetenweise:

$$\int_B \begin{pmatrix} f_1(x,y) \\ f_2(x,y) \end{pmatrix} dxdy := \begin{pmatrix} \int_B f_1(x,y)dxdy \\ \int_B f_2(x,y)dxdy \end{pmatrix}.$$

Beispiel 113.
Müller: Mechanik I, Abschnitt 5
Die Masse M einer Platte $B \subset \mathbb{R}^2$ von variabler Dichte $\rho(x,y)$ ist gegeben durch

$$M = \iint_B \rho(x,y)dxdy.$$

Der Schwerpunkt der Platte ist

$$\vec{s} = \frac{1}{M} \iint_B \rho(x,y) \begin{pmatrix} x \\ y \end{pmatrix} dxdy.$$

Wie soll man sich die „zweidimensionalen“ Integrale vorstellen? Für das Integral $\iint_B f(x,y)dxdy$ gibt es mindestens zwei häufig benutzte Anwendungen und Interpretationen.

- **Geometrische Interpretation.** Unsere anfängliche Motivation zeigt das Integral als Volumen unter dem Graphen von f und damit als Lösung eines *dreidimensionalen* Problems.
- **Physikalische Interpretation.** Im ersten Teil des vorstehenden Beispiels ergab das Integral hingegen die Masse der (dünnen) Platte B und damit die Lösung eines *zweidimensionalen* Problems. Auch die Schwerpunktberechnung muss man *so* sehen.

„Dreidimensionale“ Integration. Integration von Funktionen auf kompakten Mengen des dreidimensionalen Raumes definiert man ganz analog. Für sie spielt allerdings nur das Analogon der „Physikalischen Interpretation“ eine Rolle: vierdimensionale Volumina sind für den Ingenieur kein Thema.

Wieder sind stetige Funktionen integrierbar, und es sind bei beschränkten Funktionen auch Unstetigkeitsstellen erlaubt, wenn diese enthalten sind in der Vereinigung von endlich vielen glatten *Flächen*. So ist zum Beispiel die Funktion

$$f(x,y,z) := \begin{cases} 1 & \text{für } x^2+y^2+z^2 \leq 1 \text{ und } x \geq 0 \\ 0 & \text{sonst} \end{cases}$$

auf einer abgeschlossenen Halbkugel 1 und außerhalb 0. Die Menge der Unstetigkeitsstellen ist enthalten in der Kugelfläche $x^2 + y^2 + z^2 = 1$ vereinigt mit der Ebene $x = 0$. Also ist f integrierbar, und nach Definition ist das Integral genau das Volumen der Halbkugel.

Beispiel 114.
Theoretische Elektrotechnik, Abschnitt 1.4
Das elektrische Feld einer Punktladung Q an der Stelle $\vec{x} = (x, y, z)$ hat an der Stelle $\vec{x^*} = (x^*, y^*, z^*)$ den Wert

$$\vec{E}(\vec{x^*}) = \frac{Q}{4\pi\epsilon} \frac{\vec{x^*} - \vec{x}}{|\vec{x^*} - \vec{x}|^3} = \frac{Q}{4\pi\epsilon} \frac{1}{\sqrt{(x^* - x)^2 + (y^* - y)^2 + (z^* - z)^2}^3} \begin{pmatrix} x^* - x \\ y^* - y \\ z^* - z \end{pmatrix}.$$

Die x-Komponente davon ist

$$E_1(\vec{x^*}) = \frac{Q}{4\pi\epsilon} \frac{1}{\sqrt{\cdots}^3} (x^* - x).$$

Ist die Ladung nun stetig über den kompakten Bereich B verteilt und $\rho(x, y, z)$ die Ladungsdichte, so ist die x-Komponente von $\vec{E}$ an der Stelle (x^*, y^*, z^*) gegeben durch

$$E_1(\vec{x^*}) = \frac{1}{4\pi\epsilon} \iiint_B \frac{\rho(x, y, z)}{\sqrt{(x^* - x)^2 + (y^* - y)^2 + (z^* - z)^2}^3} (x^* - x) dx dy dz.$$

Entsprechend findet man die anderen Komponenten.

RECHENREGELN. Wir geben nun eine Zusammenstellung der Rechenregeln für das mehrdimensionale Integral.

Die integrierbaren Funktionen bilden einen Vektorraum, d.h. Linearkombinationen von integrierbaren Funktionen sind wieder integrierbar. Das Integral ist ein lineares Funktional auf diesem Vektorraum, was ganz bescheiden bedeutet, dass man (endliche) Summen von integrierbaren Funktionen gliedweise integrieren und konstante Faktoren aus dem Integral herausziehen darf. Wir schreiben das für zweidimensionale Integrale, es gilt aber allgemein.

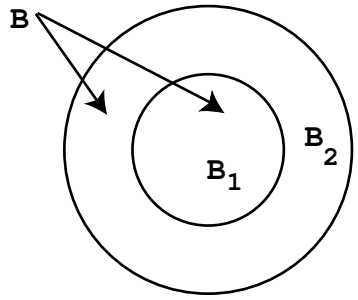
$$\iint_B (f(x, y) + g(x, y)) dx dy = \iint_B f(x, y) dx dy + \iint_B g(x, y) dx dy,$$

$$\iint_B \lambda f(x, y) dx dy = \lambda \iint_B f(x, y) dx dy,$$

$$\iint_B f(x, y) dx dy \leq \iint_B g(x, y) dx dy, \quad \text{falls } f \leq g,$$

$$\left| \iint_B f(x, y) dx dy \right| \leq \iint_B |f(x, y)| dx dy.$$

Ist schließlich der Integrationsbereich B durch eine oder mehrere glatte Kurven (im 3-dimensionalen: *glatte Flächen*) in zwei Teile B_1, B_2 zerlegt,



so gilt

$$\iint\limits_{B=B_1\cup B_2} f(x,y)dxdy = \iint\limits_{B_1} f(x,y)dxdy + \iint\limits_{B_2} f(x,y)dxdy.$$

Zum Beispiel kann man das Integral über den Kreisring B_2 als Differenz zweier Integrale (über den großen und den kleinen Kreis) berechnen.

3.2 Berechnung durch Riemannsche Summen

- Beim Übergang von der Mechanik einzelner Massenpunkte zur Mechanik starrer Körper und zur Kontinuumsmechanik entstehen mehrdimensionale Integrale als Grenzwerte von Summen. Unsere Integraldefinition erweist sich als wunderbar passend zu dieser Methode, viel passender als etwa eine Definition durch iterierte eindimensionale Integrale, wie sie im nächsten Abschnitt vorkommen.

Für die Berechnung mehrdimensionaler Integrale gibt es verschiedene Methoden. Zunächst kann man einfach auf die Definition zurückgreifen, aber der folgende Sachverhalt vereinfacht das:

Für die stetigen (oder nicht zu unstetigen beschränkten Funktionen im oben erwähnten Sinne) kann man das Integral berechnen, indem man (im 2- dimensionalen Fall) ein Rechteck R um B in feine Rechtecke unterteilt, aus jedem der feinen Rechtecke einen Funktionswert auswählt und die so entstehende Treppenfunktion integriert (d.h. „summiert“):

$$S = \sum_{i,j} f(\xi_{ij}, \eta_{ij}) \Delta x_i \Delta x_j.$$

Die dabei auftretenden Summen nennt man auch Riemannsche Summen. Für eine Folge von Unterteilungen, die „auf richtige Weise“ immer feiner werden, konvergiert dann die Folge der Treppenfunktions-Integrale gegen das gesuchte Integral. „Auf richtige Weise“ bedeutet, dass man die Teilrechtecke in x- und y-Richtung gleichermaßen kleiner machen muss. Ein gutes Maß dafür ist die maximale Diagonale der Teilrechtecke. Geht sie gegen null, so konvergieren die Riemannschen Summen gegen das Integral. Wie bei Integralen in einer Variablen kann man dies zur *numerischen Näherung des Integrals* benutzen.

Beispiel 115. Wir wollen die Funktion $f(x, y) = \cos x \cos y$ über den Viertelkreis

$$B := \{(x, y) \mid x^2 + y^2 \leq 1, x \geq 0, y \geq 0\}$$

integrieren. Wir setzen $f(x, y) = 0$ für $x^2 + y^2 > 1$ und integrieren über das Rechteck

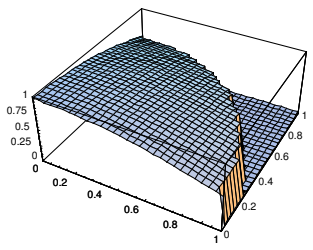
$$R = \{(x, y) \mid 0 \leq x \leq 1, 0 \leq y \leq 1\},$$

indem wir dieses in Quadrate der Kantenlänge $1/n$ zerlegen.

Wir betrachten die Riemannschen Summen mit Funktionswerten in der rechten oberen bzw. linken unteren Ecke der Teilquadrate, d.h.

$$F(n) = \sum_{i,j=1}^{n} \frac{f(\frac{i}{n}, \frac{j}{n})}{n^2},$$

$$G(n) = \sum_{i,j=1}^{n} \frac{f(\frac{i-1}{n}, \frac{j-1}{n})}{n^2}.$$



Die Abbildung macht deutlich, dass $F(n)$ eine untere und $G(n)$ eine obere Schranke für das Integral liefert. Mit Mathematica erhält man

n	$F(n)$	$G(n)$
10	0.5246	0.6873
20	0.5629	0.6456
50	0.5874	0.6209
200	0.6004	0.6088

und mit der Integrationssoftware von Mathematica:

```
NIntegrate[f[x, y], {x, 0, 1}, {y, 0, 1}]
0.604740
```

□

3.3 Berechnung durch Mehrfach-Integration

- Wir lernen, wie man mehrdimensionale Integrale über einfache Integrationsbereiche zurückführt auf mehrere eindimensionale Integrale.

Neben numerischen Methoden hätte man natürlich gern einfache Methoden für die exakte Berechnung von mehrdimensionalen Integralen, am liebsten so etwas wie die Geschichte mit der Stammfunktion. Dafür gibt es keinen wirklichen Ersatz, aber es gibt andere wichtige Hilfen, auf die wir in diesem und dem nächsten Abschnitten eingehen.

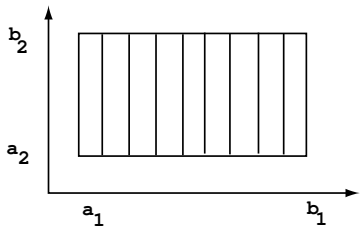
Wir betrachten zunächst eine stetige Funktion $f : R \to \mathbb{R}$ auf einem Rechteck

$$R = [a_1, b_1\,] \times [a_2, b_2\,].$$

Man kann beweisen, dass dann gilt:

$$\boxed{\iint\limits_R f(x,y)dxdy = \int_{a_1}^{b_1} \left(\int_{a_2}^{b_2} f(x,y)dy \right) dx.}$$

Die rechte Seite ist so zu verstehen: Für jedes feste $x \in [a_1, b_1\,]$ ist $f(x,.) : [a_2, b_2\,] \to \mathbb{R}$ eine (stetige) Funktion. Diese integriert man über das Intervall $[a_2, b_2\,]$ (inneres Integral der rechten Seite) und erhält für das gewählte x eine reelle Zahl. Auf diese Weise entsteht eine Funktion von x, die man anschließend über $[a_1, b_1\,]$ integriert (äußeres Integral der rechten Seite). Die Behauptung ist, dass dies gerade das gesuchte Integral über das Rechteck R liefert. Das zweidimensionale Integral ist also auf zwei eindimensionale reduziert worden. Anschaulich kann man das so symbolisieren:

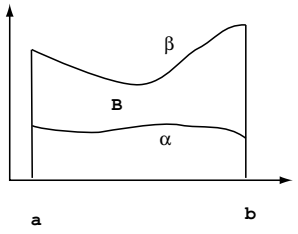


Beispiel 116. Sei $R = [0,1] \times [0,1]$. Dann ist

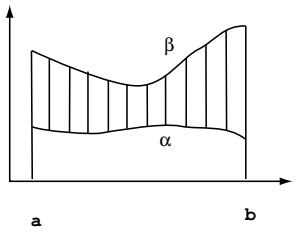
$$\iint\limits_R (xy+1)dxdy = \int_0^1 \left(\int_0^1 (xy+1)dy \right) dx = \int_0^1 (xy^2/2 + y)|_0^1 dx$$
$$= \int_0^1 (x/2+1)dx = (x^2/4 + x)|_0^1 = \frac{5}{4}.$$

□

Was aber tut man, wenn der Integrationsbereich kein Rechteck ist? Ein typischer Fall ist dieser:



Der untere Rand des Integrationsbereichs B ist der Graph einer Funktion $\alpha : [a, b] \to \mathbb{R}$, der obere der einer Funktion $\beta : [a, b] \to \mathbb{R}$ mit $\alpha \leq \beta$. (Was bedeutet es, wenn an manchen Stellen $\alpha(x) = \beta(x)$?) Auch in diesem Fall kann man die Integration über R auf zwei eindimensionale Integrale reduzieren: Symbolisch sieht das so aus:



Genauer formulieren wir das Ergebnis nun in einem Satz.

Satz 117 (Mehrfachintegration, $n = 2$). *Seien $\alpha, \beta : [a, b] \to \mathbb{R}$ differenzierbar und $\alpha \leq \beta$. Sei*

$$B = \{(x, y) \mid a \leq x \leq b \text{ und } \alpha(x) \leq y \leq \beta(x)\}.$$

Dann ist B kompakt. Ist $f : B \to \mathbb{R}$ stetig, so gilt:

$$\iint_B f(x, y)dxdy = \int_a^b \left(\int_{\alpha(x)}^{\beta(x)} f(x, y)dy \right) dx.$$

Beispiel 118. Sei $B = \{(x, y) \mid 0 \leq x, 0 \leq y, x^2 + y^2 \leq 1\}$ der Viertelkreis vom Radius 1. Dann ist

$$\begin{aligned}
\iint_B (xy + 1)dxdy &= \int_0^1 \left(\int_0^{\sqrt{1-x^2}} (xy + 1)dy \right) dx = \int_0^1 (xy^2/2 + y)|_0^{\sqrt{1-x^2}} dx \\
&= \int_0^1 (x/2(1 - x^2) + \sqrt{1 - x^2})dx \\
&= (x^2/4 - x^4/8)|_0^1 + \int_0^1 \sqrt{1 - x^2}dx \\
&= \frac{1}{8} + \frac{1}{4}(\text{Fläche des Einheitskreises}) = \frac{1}{8} + \frac{\pi}{4}.
\end{aligned}$$

□

Beispiel 119.
Müller: Mechanik I, Abschnitt 5
Für die y-Koordinate des Schwerpunktes des Halbkreises

$$B = \{(x, y) \mid x^2 + y^2 \leq r^2\}$$

erhalten wir mit Beispiel 113 den Wert

$$y_S = \frac{1}{\pi r^2/2} \int_{-r}^{+r} \int_0^{\sqrt{r^2-x^2}} ydydx = \frac{1}{\pi r^2/2} \frac{2r^3}{3} = \frac{2}{3} \frac{2r}{\pi}.$$

Vergleichen Sie das mit dem Schwerpunkt des Halbkreisbogens in Beispiel 110.

Entsprechendes gilt auch in drei Dimensionen. An die Stelle des „äußeren Integrationsintervalls“ tritt dann ein zweidimensionaler Integrationsbereich, auf den man dann hoffentlich den vorstehenden Satz anwenden kann, so dass man das dreidimensionale Integral auf drei eindimensionale reduziert.

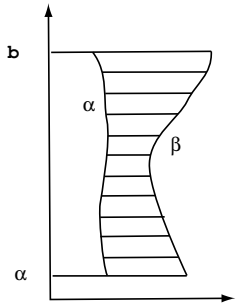
Satz 120 (Mehrfachintegration, $n = 3$). *Seien $B^* \subset \mathbb{R}^2$ kompakt und $\alpha, \beta : B^* \to \mathbb{R}$ differenzierbar mit $\alpha \leq \beta$. Sei*

$$B = \{(x, y, z) \,|\, (x, y) \in B^* \text{ und } \alpha(x, y) \leq z \leq \beta(x, y)\}.$$

Dann ist B kompakt. Ist $f : B \to \mathbb{R}$ stetig, so gilt:

$$\iiint_B f(x, y, z)dxdydz = \iint_{B^*} \left(\int_{\alpha(x,y)}^{\beta(x,y)} f(x, y, z)dz \right) dxdy.$$

Bemerkungen. 1) Natürlich lassen sich die Rollen der Variablen in den vorstehenden Sätzen vertauschen. Man kann eventuell auch *erst* über x und dann über y integrieren:



2) Häufig kann man ein gegebenes B nicht so darstellen, wie es in den Sätzen vorausgesetzt ist, aber man kann es in verschiedene Bereiche zerlegen, die so aussehen. Dann integriert man über die einzelnen Bereiche und addiert die Ergebnisse, vgl. Beispiel 123.

3) Ein Hauptproblem bei der Mehrfachintegration ist, den Integrationsbereich entsprechend zu beschreiben. Man braucht dazu geometrische Vorstellungskraft (und da hapert es leider oft).

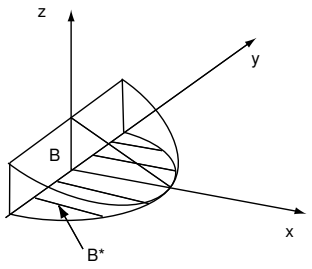
Beispiel 121. Berechne

$$\iiint_B 2xzdxdydz,$$

wobei B der kompakte Bereich zwischen den vier Flächen

$$\begin{aligned} x &= 0, \\ x &= 1 - y^2, \\ z &= 0, \\ x + 2z &= 1 \end{aligned}$$

ist.



Wir bezeichnen mit B^* den „Boden“ des Bereiches B, also das Gebiet der xy-Ebene zwischen der y-Achse und der Parabel $x = 1 - y^2$. Dann ist

$$B = \{(x, y, z) \mid (x, y) \in B^* \text{ und } 0 \leq z \leq \frac{1}{2}(1 - x)\}.$$

Daher folgt

$$\begin{aligned}
\iiint\limits_B 2xzdxdydz &= \iint\limits_{B^*} \left(\int_{z=0}^{\frac{1-x}{2}} 2xzdz \right) dxdy = \iint\limits_{B^*} \left(xz^2|_{z=0}^{z=\frac{1-x}{2}} \right) dxdy \\
&= \iint\limits_{B^*} \frac{1}{4}(x^3 - 2x^2 + x)dxdy = \int_{y=-1}^{1} \left(\int_{x=0}^{1-y^2} \frac{1}{4}(x^3 - 2x^2 + x)dx \right) dy \\
&= \int_{y=-1}^{1} \frac{1}{4}(\frac{1}{4}x^4 - \frac{2}{3}x^3 + \frac{1}{2}x^2)|_{x=0}^{1-y^2} dy \\
&= \frac{1}{4} \int_{y=-1}^{1} \left(\frac{1}{4}(1 - y^2)^4 - \frac{2}{3}(1 - y^2)^3 + \frac{1}{2}(1 - y^2)^2 \right) dy \\
&= \frac{2}{63}.
\end{aligned}$$

Stattdessen hätte man auch das Dreieck in der xz-Ebene als B^* wählen können. Dann ergeben sich die folgenden Integrale. Beachten Sie, dass in der Notation die „d-Terme“ von innen nach außen die Reihenfolge der Integrationsvariablen angeben.

$$\iiint\limits_B 2xzdxdydz = \int_{z=0}^{1/2} \int_{x=0}^{1-2z} \int_{y=-\sqrt{1-x}}^{\sqrt{1-x}} 2xz\, dydxdz.$$

□

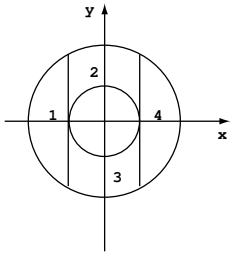
Beispiel 122. Sie können auch ein Programm für numerische Integration benutzen. In Mathematica sieht das vorstehende Beispiel so aus:

```
In[1]:=NIntegrate[ 2 x z,{y,-1,1},{x,0,1-y^2},{z,0,(1-x)/2}]
Out[1]=0.031746
```

Eine Zeile Eingabe, ein Klick, schon hat man das Ergebnis. Wenn Sie allerdings die geschweiften Klammern ansehen, die den Laufbereich der einzelnen Variablen beschreiben, dann erkennen Sie, dass Ihnen das Programm die Umsetzung des geometrisch gegebenen Integrationsbereichs in eine algebraische Beschreibung, eben die Entscheidung über die Reihenfolge der Integrationen und die Bestimmung der zugehörigen Integrationsintervalle, leider nicht abnimmt. □

Beispiel 123. Sei $B := \{(x, y) \mid r^2 \leq x^2 + y^2 \leq R^2\}$ der Kreisring mit innerem und äußerem Radius r bzw. R. Dann ist

$$
\begin{aligned}
&\iint\limits_B f(x,y)dxdy \\
&= \iint\limits_{B_1} f(x,y)dxdy + \ldots + \iint\limits_{B_4} f(x,y)dxdy \\
&= \int_{x=-R}^{-r}\int_{y=-\sqrt{R^2-x^2}}^{+\sqrt{R^2-x^2}} f(x,y)dxdy + \int_{x=-r}^{+r}\int_{y=+\sqrt{r^2-x^2}}^{+\sqrt{R^2-x^2}} f(x,y)dxdy \\
&\quad + \int_{x=-r}^{+r}\int_{y=-\sqrt{R^2-x^2}}^{-\sqrt{r^2-x^2}} f(x,y)dxdy + \int_{x=r}^{R}\int_{y=-\sqrt{R^2-x^2}}^{+\sqrt{R^2-x^2}} f(x,y)dxdy.
\end{aligned}
$$



Oft empfiehlt es sich allerdings, stattdessen die „Bereichsadditivität“ des Integrals zu benutzen und so zu rechnen:

$$
\iint\limits_B f(x,y)dxdy = \underbrace{\int_{x=-R}^{R}\int_{y=-\sqrt{R^2-x^2}}^{+\sqrt{R^2-x^2}} f(x,y)dxdy}_{\text{Integral über den großen Kreis}} - \underbrace{\int_{x=-r}^{r}\int_{y=-\sqrt{r^2-x^2}}^{+\sqrt{r^2-x^2}} f(x,y)dxdy}_{\text{Integral über den kleinen Kreis}}.
$$

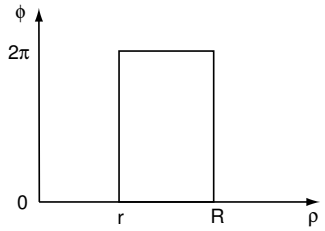
Natürlich klappt das nur, wenn f auch auf dem kleinen Kreis definiert ist.

Ist aber z.B. $f(x,y) = \sqrt{x^2+y^2-r^2}$, so ist das nicht der Fall und die Kreisscheiben-Zerlegung macht gar keinen Sinn. Man muss dann doch auf die obige komplizierte Zerlegung zurückgreifen. Oder man muss sich etwas anderes einfallen lassen, vgl. nächsten Abschnitt. □

3.4 Berechnung durch Koordinatentransformation

- Bei kreisförmigen Integrationsbereichen liegt die Verwendung von Polarkoordinaten nahe. Allgemeiner kann man vielleicht einem Integrationsbereich besser angepasste Koordinaten verwenden, als die kartesischen. Bei der Integration muss man dann aber auch das Flächen- bzw. Volumenelement den neuen Koordinaten anpassen. Wir lernen, wie man das macht.

Manche Bereiche lassen sich am bequemsten nicht in kartesischen, sondern in anderen Koordinaten beschreiben. Der Kreisring aus dem Beispiel im letzten Abschnitt war mühevoll zu behandeln. In Polarkoordinaten entspricht diesem Kreisring einfach ein Rechteck



PROBLEMSTELLUNG. Ist nun

$$\iint_B f(x,y)dxdy$$

gesucht, kann man dann einfach f in Polarkoordinaten ausdrücken und

$$\iint_R f(\rho,\phi)d\rho d\phi$$

berechnen? NEIN, das gibt im allgemeinen nicht das gewünschte Ergebnis. Der Grund dafür wird deutlich, wenn man eine konstante Funktion, etwa $f = 1$ betrachtet. Das erste Integral liefert das Volumen eines Körpers der konstanten Höhe 1 über dem Kreisring B, also einfach die Fläche des Kreisrings

$$\pi(R^2 - r^2) = \pi(R-r)(R+r).$$

Aber in Polarkoordinaten ist ebenfalls $f = 1$. Das zweite Integral liefert daher die Fläche des Rechtecks

$$2\pi(R-r),$$

also im allgemeinen etwas anderes.

Denkt man an das Volumen (lokal) als „Grundfläche×Höhe“, so sieht man: Die Transformation der (nun wieder beliebigen) Funktion f passt zwar die Höhe den neuen Koordinaten an, aber man muss die Grundfläche korrigieren.

Zwischenbetrachtung. Eine lineare Abbildung

$$\begin{pmatrix} u \\ v \end{pmatrix} \mapsto \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix}$$

bildet das von den beiden Einheitsvektoren $\vec{e}_1, \vec{e}_2$ aufgespannte Quadrat der Fläche 1 auf ein Parallelogramm mit den Seiten $\begin{pmatrix} a_{11} \\ a_{21} \end{pmatrix}$ und $\begin{pmatrix} a_{12} \\ a_{22} \end{pmatrix}$ ab, und das hat Flächeninhalt $|a_{11}a_{22} - a_{12}a_{21}| =$

$|\det A|$. Die Flächenverzerrung durch eine lineare Abbildung von $\mathbb{R}^2$ in sich ist gegeben durch den Absolutbetrag ihrer Determinante.

Eine differenzierbare Koordinatentransformation approximiert man lokal durch ihre Ableitung. Das ist eine lineare Abbildung mit der Funktionalmatrix (aus den 1. partiellen Ableitungen) als Matrix. Die Flächenverzerrung ist dann gegeben durch den Absolutbetrag der Determinante der Funktionalmatrix.

Wir fassen das Ergebnis in einem Satz zusammen:

Satz 124 (Transformationsformel für $n = 2$). *Seien $B, R \subset \mathbb{R}^2$ kompakte Bereiche und sei $f : B \to \mathbb{R}$ stetig. Sei*

$$\vec{x} : \begin{pmatrix} u \\ v \end{pmatrix} \mapsto \begin{pmatrix} x(u,v) \\ y(u,v) \end{pmatrix}$$

eine Transformation mit stetigen partiellen Ableitungen, welche R surjektiv und (mit möglichen Ausnahmen in Randpunkten von R) injektiv auf B abbildet. Dann gilt

$$\iint\limits_B f(x,y)dxdy = \iint\limits_R f(x(u,v), y(u,v)) \left| \det \begin{pmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{pmatrix} \right| dudv.$$

Für die hier auftretende Funktionaldeterminante finden sie in der Literatur auch diese Notation

$$\frac{\partial(x,y)}{\partial(u,v)} := \det \begin{pmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{pmatrix}.$$

Beispiel 125. Für Polarkoordinaten

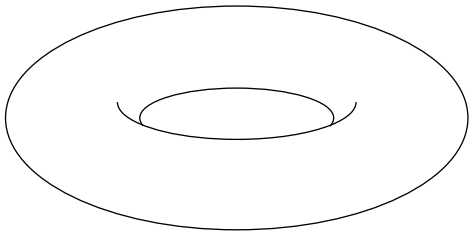
$$x = \rho\cos\phi, y = \rho\sin\phi$$

spielen ρ und ϕ spielen also die Rolle von u und v. Die Funktionaldeterminante ist

$$\det \begin{pmatrix} \frac{\partial x}{\partial \rho} & \frac{\partial x}{\partial \phi} \\ \frac{\partial y}{\partial \rho} & \frac{\partial y}{\partial \phi} \end{pmatrix} = \det \begin{pmatrix} \cos\phi & -\rho\sin\phi \\ \sin\phi & \rho\cos\phi \end{pmatrix} = \rho\cos^2\phi + \rho\sin^2\phi = \rho.$$

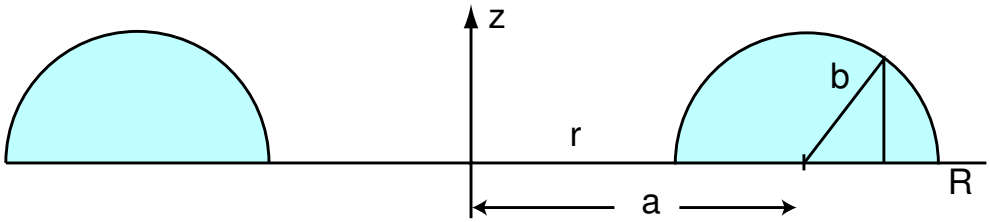
Weil $\rho \geq 0$ ist, kann man sich die Absolutstriche sparen. □

Beispiel 126. Wir wollen das Volumen eines Torus



durch Integration der Höhenfunktion über dem Kreisring mit den Radien $0 < r < R$ berechnen. Dann ist $a = \frac{r+R}{2}$ der Radius der "Seele" des Torus und $b = \frac{R-r}{2}$ der Radius des erzeugenden Kreises. In Polarkoordinaten ist die Höhenfunktion dann gegeben durch

$$f(\rho, \phi) = \sqrt{b^2 - (a - \sqrt{\rho^2})^2}.$$



Damit ergibt sich

$$\frac{1}{2}V \qquad = \int_{\phi=0}^{2\pi} \int_{\rho=a-b}^{a+b} \sqrt{b^2 - (a-\rho)^2}\, \rho d\rho d\phi = \int_{\phi=0}^{2\pi} \frac{\pi}{2} ab^2\, d\phi = \pi^2 ab^2.$$

Das eindimensionale Integral über ρ erfordert eine Anwendung der Substitutionsregel oder ähnliches, das wir unterschlagen haben. Das Ergebnis kann man auch so schreiben: Das Volumen des ganzen Torus ist $2\pi^2 ab^2 = (2\pi a)(\pi b^2)$, also gerade die Länge der „Seele" multipliziert mit der Fläche des Querschnitts.

> Dies ist ein Spezialfall der *Guldinschen Regel*: Entsteht ein Körper durch Rotation einer ganz in einer Halbebene durch die Rotationsachse gelegenen Fläche, so ist sein Volumen gleich dem Flächeninhalt dieser Fläche multipliziert mit der Länge des Kreises, den der Flächenschwerpunkt beschreibt.

□

Bemerkungen. Praktischerweise erlaubt der Satz 124, dass $\vec{x}$ auf dem Rand von R nicht injektiv ist. Gerade bei den Polarkoordinaten bekäme man sonst Probleme: Die offensichtliche Beschreibung des Kreises ist für $\rho = 0$ und für $\phi = 0 = 2\pi$ nicht injektiv.
Die Transformationsformel gilt für beliebige krummlinige Koordinatensysteme, nicht nur für Polarkoordinaten; aber in der Praxis des Ingenieurs kommen wohl vor allem diese vor. Wir notieren deshalb noch einmal in Kurzfassung

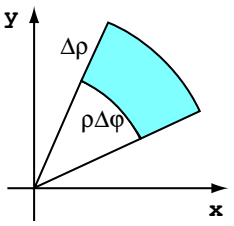
Satz 127 (Integration in Polarkoordinaten).

$$\iint_B f(x,y)dxdy = \iint_R f(\rho,\phi)\rho\, d\rho d\phi.$$

Dabei ist R der B entsprechende Bereich in der (ρ, ϕ)-Ebene, und $f(\rho, \phi)$ bedeutet genauer $f(x(\rho, \phi), y(\rho, \phi))$, vgl. die Bemerkung über Koordinatensysteme im Abschnitt 1.11.

FLÄCHENELEMENTE. In Anbetracht der Definition des mehrdimensionalen Integrals bezeichnet man $dA = dxdy$ auch als das (infinitesimale) Flächenelement. Bei der Koordinatentransformation muss man dies nach dem obigen Satz ersetzen durch das Flächenelement in Polarkoordinaten

$$\boxed{dA = dxdy \to \rho\, d\rho d\phi.}$$



Wir wollen das noch anschaulich verdeutlichen. Dem Rechteck mit den diagonalen Ecken $(\rho, \phi), (\rho + \Delta\rho, \phi + \Delta\phi)$ entspricht in der (x, y)-Ebene annähernd ein Rechteck mit den Seiten $\Delta\rho$ und $\rho\Delta\phi$, also mit dem Flächeninhalt $\rho\Delta\rho\Delta\phi$.

Beispiel 128 (Gaußsche Fehlerfunktion). Sei $B(r)$ ein Kreis vom Radius r um den Nullpunkt und

$$f(x, y) = e^{-\frac{1}{2}(x^2+y^2)}.$$

Dann erhalten wir mit der Transformationsformel bei Verwendung von Polarkoordinaten

$$\iint\limits_{B(r)} e^{-\frac{1}{2}(x^2+y^2)} dxdy = \iint\limits_{R(r)} e^{-\frac{\rho^2}{2}} \rho d\rho d\phi.$$

Dabei ist $R(r)$ das Rechteck $0 \le \phi \le 2\pi$, $0 \le \rho \le r$, und der Faktor ρ ist ein Geschenk des Himmels, denn $e^{-\rho^2/2}\rho$ lässt sich im Gegensatz zu $e^{-\rho^2/2}$ integrieren:

$$\iint\limits_{B(r)} e^{-\frac{1}{2}(x^2+y^2)} dxdy = \int_0^{2\pi} (\int_0^r e^{-\rho^2/2} \rho d\rho) d\phi = \int_0^{2\pi} (-e^{-\rho^2/2})_0^r d\phi$$

$$= \int_0^{2\pi} (1 - e^{-r^2/2}) d\phi = 2\pi(1 - e^{-r^2/2}).$$

Lässt man hier $r \to \infty$ gehen, so erhält man ein zweidimensionales „uneigentliches“ Integral

$$\iint\limits_{\mathbb{R}^2} e^{-\frac{1}{2}(x^2+y^2)} dxdy = \lim_{r\to\infty} 2\pi(1 - e^{-r^2/2}) = 2\pi.$$

Wir betrachten nun andrerseits das Quadrat $Q(r)$ mit der Kantenlänge $2r$ um den Nullpunkt. Dafür ist

$$\iint\limits_{Q(r)} e^{-\frac{1}{2}(x^2+y^2)} dxdy = \int_{-r}^{r} (\int_{-r}^{r} e^{-\frac{1}{2}(x^2+y^2)} dx) dy = \int_{-r}^{r} (e^{-\frac{1}{2}y^2} \underbrace{\int_{-r}^{r} e^{-\frac{1}{2}x^2} dx}_{=:I(r)}) dy$$

$$= I(r) \int_{-r}^{r} e^{-\frac{1}{2}y^2} dy = I(r)^2.$$

Lässt man hier $r \to \infty$ gehen, so strebt die linke Seite wieder gegen $\iint\limits_{\mathbb{R}^2} e^{-\frac{1}{2}(x^2+y^2)} dxdy = 2\pi$, die rechte Seite aber gegen das Quadrat von

$$I(\infty) := \lim_{r\to\infty} I(r) = \int_{-\infty}^{\infty} e^{-\frac{1}{2}x^2} dx.$$

Wir erhalten $\int_{-\infty}^{+\infty} e^{-x^2/2} dx = \sqrt{2\pi}$. Dieses Integral haben wir früher nicht herausbekommen, weil $e^{-\frac{1}{2}x^2}$ keine elementare Stammfunktion besitzt, und auch Substitutionsregel und partielle Integration nicht zum Ziel führten. Aber dies ist nicht *irgendein* Integral, wie Sie

daran erkennen können, dass die Funktion $e^{-\frac{1}{2}x^2}$ sogar den Weg auf den alten deutschen Zehn-Mark-Schein gefunden hatte.

Mit der Substitutionsregel finden Sie leicht

$$\frac{2}{\sqrt{\pi}} \int_0^{+\infty} e^{-t^2} dt = 1.$$

Die Funktion

$$\operatorname{erf}(x) := \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt$$

heißt *Gaußsche Fehlerfunktion*. Sie spielt als *Normalverteilung* in der Wahrscheinlichkeitsrechnung eine wichtige Rolle, und der eben berechnete Grenzwert

$$\boxed{\operatorname{erf}(\infty) := \lim_{x\to\infty} \operatorname{erf}(x) = 1}$$

ist dafür essentiell. □

Beispiel 129 (Diffusion und Wärmeleitung). Die Fehlerfunktion kommt aber auch in ganz anderem Zusammenhang vor. Um das zu erklären, rechnen wir für ein $\lambda > 0$ die Ableitungen der Funktion

$$u(x,t) := \operatorname{erf}(\frac{x}{2\sqrt{\lambda t}}) \tag{37}$$

aus. Wir finden

$$\frac{\partial u}{\partial t} = \frac{2}{\sqrt{\pi}} e^{-\frac{x^2}{4\lambda t}} \frac{-x}{4\sqrt{\lambda}\, t^{3/2}}, \quad \frac{\partial^2 u}{\partial x^2} = \frac{2}{\sqrt{\pi}} e^{-\frac{x^2}{4\lambda t}} \frac{1}{2\sqrt{\lambda t}} \frac{-2x}{4\lambda t}.$$

Daher gilt

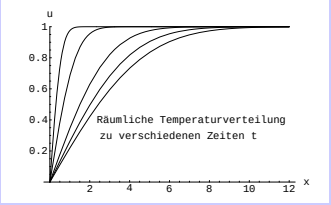
$$\boxed{\frac{\partial u}{\partial t} = \lambda \frac{\partial^2 u}{\partial x^2}.}$$

Die Funktion $u(x,t)$ ist also eine (aber keineswegs die einzige!) Lösung der sogenannten (1+1)-dimensionalen *Wärmeleitungs-* oder *Diffusionsgleichung*. Sie begegnet Ihnen in dieser Funktion in ungezählten Anwendungen, etwa

- bei der Herstellung von Fremdstoffkonzentrationen in einem Halbleiter durch Diffusion, vgl. [Werkstoffe II, Abschnitt 8.3],
- bei Diffusionsprozessen in einer Vielzahl von Varianten, vgl. [Verfahrentechnik I, Kapitel 2] oder [Energie-, Impuls- und Stofftransport, Abschnitt 2.5],

  und gleichermaßen
- bei Wärmeleitungsphänomenen, etwa bei der Temperaturentwicklung in einer einseitig erhitzten feuerdämmenden Wand, vgl. [Energie-, Impuls- und Stofftransport, Abschnitt 2.3].

Offenbar gilt für u wie in (37), dass $u(0,t) = 0$ und $u(x,0) := \lim_{t\searrow 0} u(x,t) = 1$. Deshalb kann man u deuten als die Temperatur eines (nach rechts unendlich langen) Metallstabes, der zur Zeit $t = 0$ überall die Temperatur 1^0 hat, und dann am Ende $x = 0$ ständig auf 0^0 gekühlt wird.



DER 3-DIMENSIONALE FALL. Hier muss man die Volumenverzerrung durch die Koordinatentransformation berücksichtigen. Man erhält ganz analog zum obigen 2-dimensionalen Fall

Satz 130 (Transformationsformel für $n = 3$). *Seien $B, R \subset \mathbb{R}^3$ kompakte Bereiche und sei $f : B \to \mathbb{R}$ stetig. Sei*

$$\vec{x} : \begin{pmatrix} u \\ v \\ w \end{pmatrix} \mapsto \begin{pmatrix} x(u,v,w) \\ y(u,v,w) \\ z(u,v,w) \end{pmatrix}$$

eine Transformation mit stetigen partiellen Ableitungen, welche R surjektiv und (mit möglichen Ausnahmen in Randpunkten von R) injektiv auf B abbildet. Dann gilt

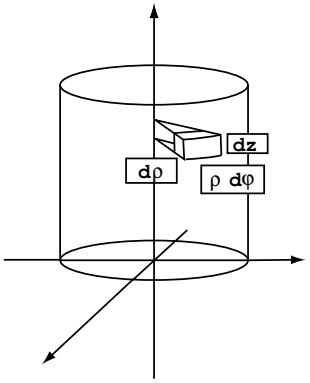
$$\iiint_B f(x,y,z)dxdydz =$$

$$\iiint_R f(x(u,v,w), y(u,v,w), z(u,v,w)) \left| \det \begin{pmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} & \frac{\partial x}{\partial w} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} & \frac{\partial y}{\partial w} \\ \frac{\partial z}{\partial u} & \frac{\partial z}{\partial v} & \frac{\partial z}{\partial w} \end{pmatrix} \right| dudvdw.$$

Wichtige Anwendungen betreffen die

Zylinderkoordinaten

$$\begin{aligned} x &= \rho \cos \phi, \\ y &= \rho \sin \phi, \\ z &= z, \end{aligned}$$



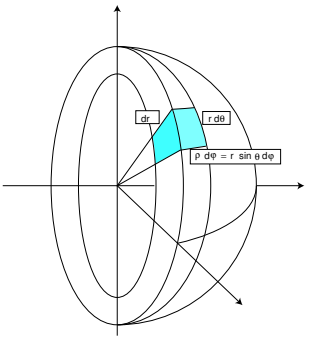
wo sich das Volumenelement

$$\boxed{dV = \rho d\rho d\phi dz}$$

ergibt, und die sphärischen Polarkoordinaten oder

Kugelkoordinaten

$$\begin{aligned} x &= r \sin\theta \cos \phi, \\ y &= r \sin\theta \sin \phi, \\ z &= r \cos\theta. \end{aligned}$$



In diesem Fall berechnen wir die Funktionaldeterminante

$$\det\begin{pmatrix} \frac{\partial x}{\partial r} & \frac{\partial x}{\partial \theta} & \frac{\partial x}{\partial \phi} \\ \frac{\partial y}{\partial r} & \frac{\partial y}{\partial \theta} & \frac{\partial y}{\partial \phi} \\ \frac{\partial z}{\partial r} & \frac{\partial z}{\partial \theta} & \frac{\partial z}{\partial \phi} \end{pmatrix} = \det\begin{pmatrix} \sin\theta\cos\phi & r\cos\theta\cos\phi & -r\sin\theta\sin\phi \\ \sin\theta\sin\phi & r\cos\theta\sin\phi & r\sin\theta\cos\phi \\ \cos\theta & -r\sin\theta & 0 \end{pmatrix} = r^2\sin\theta.$$

Also ist das Volumenelement in Kugelkoordinaten gegeben durch

$$\boxed{dV = r\rho\, dr\, d\theta\, d\phi = r^2\sin\theta dr d\theta d\phi.}$$

Beispiel 131.
Müller: Mechanik II, Abschnitt 14.2
Das Trägheitsmoment einer Hohlkugel H vom inneren Radius R_1 und äußeren Radius R_2 und von der konstanten Dichte δ_0 bezüglich einer Achse durch den Mittelpunkt ist gegeben durch

$$\begin{aligned}\Theta &= \iiint_H (x^2+y^2)\delta_0 dxdydz = \delta_0\int_0^{2\pi}\int_0^{\pi}\int_{R_1}^{R_2}(r^2\sin^2\theta)r^2\sin\theta dr d\theta d\phi \\ &= \delta_0\int_0^{2\pi}\int_0^{\pi}\int_{R_1}^{R_2} r^4\sin^3\theta dr d\theta d\phi = \delta_0\frac{1}{5}(R_2^5-R_1^5)\cdot 2\pi\cdot \underbrace{\int_0^{\pi}\sin^3\theta d\theta}_{=\frac{4}{3}} \\ &= \frac{8\pi}{15}\delta_0(R_2^5-R_1^5).\end{aligned}$$

Die Masse der Hohlkugel ist $\delta_0\cdot\frac{4}{3}\pi(R_2^3-R_1^3)$, und damit schreibt sich

$$\Theta = \frac{2}{5}M\frac{R_2^5-R_1^5}{R_2^3-R_1^3}.$$

z.B. kann man durch Messung von Θ (vgl. Mechanik) sowie von M und R_2 den inneren Radius R_1 bestimmen.
Das Trägheitsmoment der Vollkugel ($R_1 = 0, R_2 =: R$) ergibt sich zu $\Theta_V = \frac{2}{5}MR^2$, und für $R_1 \nearrow R_2 =: R$ bekommt man (etwa mit der Regel von l'Hospital) das Trägheitsmoment einer sehr dünnen Kugelschale als $\Theta_S = \frac{2}{3}MR^2$. Bei gleicher Masse und gleichem äußerem Radius (folglich bei geringerer Dichte!) hat also die Vollkugel kleineres Trägheitsmoment als die Kugelschale, sie rollt schneller eine schiefe Ebene hinunter.
Das Trägheitsmoment eines Hohlzylinders spielt etwa bei Drehspulinstrumenten eine Rolle. Mit Hilfe von Zylinderkoordinaten berechnet man wie oben (aber etwas einfacher) für einen Hohlzylinder der Höhe h mit Radien $R_1 < R_2$

$$\Theta = \frac{M}{2}(R_1^2+R_2^2).$$

Zum Abschluss betrachten wir ein Beispiel für Nicht-Standard-Koordinaten.

Beispiel 132. Seien $B^* \subset \mathbb{R}^2$ ein kompakter Bereich und $(a,b,h) \in \mathbb{R}^3$ ein Punkt mit $h > 0$. Die Verbindungsstrecke zwischen einem Punkt $(x,y,0)$ und (a,b,h) ist dann

$$\begin{pmatrix} (1-t)x + ta \\ (1-t)y + tb \\ th \end{pmatrix}, \quad 0 \le t \le 1,$$

und deshalb ist

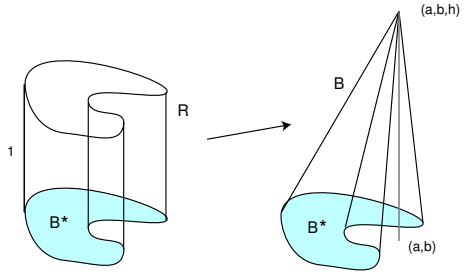
$$B := \{((1-t)x + ta, (1-t)y + tb, th) \mid (x,y) \in B^*, 0 \le t \le 1\}$$

der (schiefe) Kegel über der Grundfläche B^* mit der Spitze (a,b,h). Wir wollen sein Volumen $V(B)$ berechnen.

Beachten Sie, dass

$$R := \{(u,v,w) \mid (u,v) \in B^*, 0 \le w \le 1\}$$

der Zylinder der Höhe 1 über der Grundfläche B^* ist.



Die Abbildung

$$\begin{pmatrix} u \\ v \\ w \end{pmatrix} \quad \mapsto \quad \begin{pmatrix} (1-w)u + wa \\ (1-w)v + wb \\ wh \end{pmatrix}$$

bildet R bijektiv und stetig differenzierbar auf B ab. Ihre Funktionaldeterminante ist

$$\det \begin{pmatrix} 1-w & 0 & -u+a \\ 0 & 1-w & -v+b \\ 0 & 0 & h \end{pmatrix} = (1-w)^2 h.$$

Daher liefert die Transformationsformel

$$\begin{aligned} V(B) &= \iiint_B dxdydz = \iiint_R (1-w)^2 h \, dudvdw \\ &= h \iint_{B^*} \left(\int_0^1 (1-w)^2 dw \right) dudv = h \iint_{B^*} -\frac{1}{3}(1-w)^3 \Big|_0^1 dudv \\ &= \frac{h}{3} \iint_{B^*} dudv = \frac{h}{3} F(B^*), \end{aligned}$$

wobei $F(B^*)$ die Fläche von B^* bezeichnet. Das Volumen ist "ein Drittel Grundfläche mal Höhe". Diese Formel kennen Sie vielleicht für einfache Figuren wie den geraden Kreiskegel oder die Pyramide aus der Schule. □

3.5 Flächen im Raum. Skalare Oberflächenintegrale

- Wir lernen, wie man Flächen im Raum mit Koordinaten versieht (parametrisiert) und wie man über so parametrisierte Flächen integriert.
- Insbesondere lernen wir den Begriff des Oberflächenelementes einer Fläche kennnen.

Wir betrachten Flächen im dreidimensionalen Raum $\mathbb{R}^3$, zum Beispiel die Kugelfläche, den Graphen einer Funktion $f(x, y)$ von zwei Variablen, die Oberfläche eines chemischen Reaktors oder eines mechanischen Körpers. Um mit einer solchen Fläche F rechnen zu können, zum Beispiel ihren Flächeninhalt oder den Fluß einer physikalischen Größe durch sie berechnen zu können, braucht man Koordinaten (u, v) auf der Fläche, die die Punkte möglichst eindeutig festlegen.

Die Wahl solcher Koordinaten, kann Ihnen in einfachen Fällen vielleicht eine mathematische Software abnehmen, in der Regel ist sie aber in zweifacher Weise eine Herausforderung an Ihre Fähigkeiten: Sie brauchen dazu nämlich erstens eine geometrische Vorstellung von der Fläche und zweitens die Fähigkeit, Koordinatenfunktionen dafür wirklich aufs Papier (in den Rechner) zu bringen. Beides erfordert Übung. Die Beispiele in diesem Abschnitt repräsentieren Standardsituationen.

Beispiel 133 (Parametrisierung der Sphäre). Auf der Kugelfläche oder Sphäre vom Radius R bieten sich dafür die „geographische Breite“ $u = \theta$ und die „geographische Länge“ $v = \phi$ an, die den Punkt

$$\vec{x} = \vec{x}(\theta, \phi) = \begin{pmatrix} R\sin\theta\cos\phi \\ R\sin\theta\sin\phi \\ R\cos\theta \end{pmatrix} \tag{38}$$

auf der Kugel festlegen. Die Koordinaten durchlaufen den Bereich

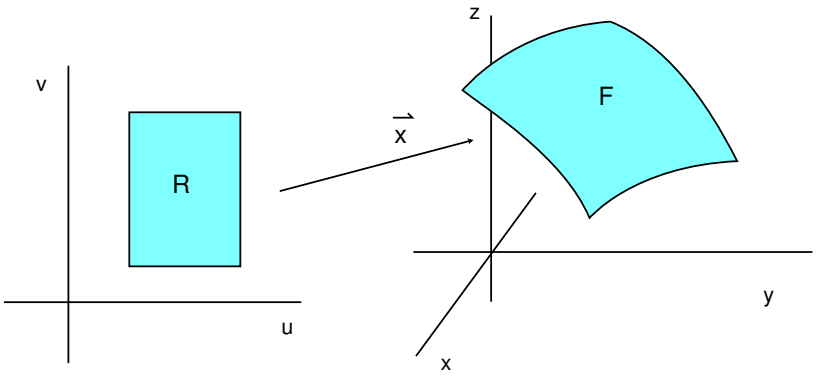
$$B:\ 0 \le \theta \le \pi, \quad 0 \le \phi \le 2\pi,$$

also ein Rechteck. □

Man hat dann zwei Abbildungen, die zueinander invers sind: Die Abbildung, die dem Punkt $\vec{x} \in F$ auf der Fläche den Punkt (u, v) im *Koordinaten-* oder *Parameterbereich* B zuordnet, und deren Umkehrabbildung

$$\vec{x}(u, v) := \begin{pmatrix} x(u, v) \\ y(u, v) \\ z(u, v) \end{pmatrix},$$

die jedem Parameterpaar $(u, v) \in B$ den entsprechenden Punkt auf der Fläche zuordnet. Diese zweite Abbildung ist für unsere Zwecke brauchbarer, und wir nennen sie eine *Parametrisierung der Fläche* F.



Technische Anforderungen an eine Parametrisierung. Wir wollen von unseren Parametrisierungen folgendes verlangen:

1. Der Parameterbereich B ist kompakt, und der Rand ∂B besteht aus höchstens endlich vielen glatten Kurven (Beispiele: Kreis oder Rechteck).
2. $\vec{x} : B \to F$ ist surjektiv und auf $B \backslash \partial B$ injektiv. (∂B bezeichnet den Rand von B. Die Bedingung bedeutet, dass die Koordinaten eines Punktes „meistens“ eindeutig festliegen, dass aber zum Beispiel bei der Kugel $(\theta, 0)$ und $(\theta, 2\pi)$ zum selben Punkt gehören.)
3. $\vec{x} : B \to \mathbb{R}^3$ hat stetige partielle Ableitungen und
4. $\frac{\partial \vec{x}}{\partial u}$ und $\frac{\partial \vec{x}}{\partial v}$ sind überall auf $B \backslash \partial B$ linear unabhängig.

Definition 134 (Fläche). Wir nennen nun eine Teilmenge $F \subset \mathbb{R}^3$ eine (glatte) Fläche, wenn es für F eine solche Parametrisierung gibt. Im allgemeinen gibt es dann sogar mehrere verschiedene Parametrisierungen.

Die oben angegebene Parametrisierung der Sphäre erfüllt die gestellten Bedingungen. Wir prüfen die 4. Bedingung und erinnern daran, dass $\vec{a}, \vec{b} \in \mathbb{R}^3$ genau dann linear unabhängig sind, wenn

$$\vec{a} \times \vec{b} \neq \vec{0}.$$

Die partiellen Ableitungen von $\vec{x}$ sind

$$\frac{\partial \vec{x}}{\partial \theta} = \begin{pmatrix} R\cos\theta\cos\phi \\ R\cos\theta\sin\phi \\ -R\sin\theta \end{pmatrix}, \quad \frac{\partial \vec{x}}{\partial \phi} = \begin{pmatrix} -R\sin\theta\sin\phi \\ R\sin\theta\cos\phi \\ 0 \end{pmatrix}.$$

Sie haben das Vektorprodukt

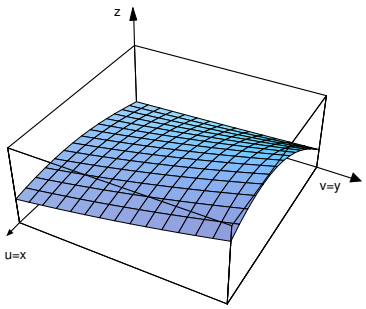
$$\frac{\partial \vec{x}}{\partial \theta} \times \frac{\partial \vec{x}}{\partial \phi} = (R\sin\theta)\vec{x} = (R^2\sin\theta)\vec{e_r}, \tag{39}$$

wobei $\vec{e_r} = \frac{\vec{x}}{|\vec{x}|}$ der radiale Einheitsvektor ist (Nachrechnen). Für $0 < \theta < \pi$ ist das $\neq 0$, und deshalb sind die partiellen Ableitungen dort linear unabhängig. Auch die Bedingung der Injektivität ist erfüllt: Die Kugelkoordinaten sind eindeutig bis auf die „Randfälle“ $\theta = 0, \theta = \pi, \phi = 0, \phi = 2\pi$.

Wir geben zwei weitere Beispiele für die Parametrisierung geometrisch gegebener Flächen.

Beispiel 135 (Parametrisierung von Funktionsgraphen). Der Graph einer differenzierbaren Funktion $f : \mathbb{R}^2 \supset B \to \mathbb{R}$ ist eine glatte Fläche, auf der der „Fußpunkt“ natürliche Koordinaten $(u, v) = (x, y)$ liefert. Die Parametrisierung ist dann gegeben durch

$$\vec{x}(u, v) := \begin{pmatrix} u \\ v \\ f(u, v) \end{pmatrix}. \tag{40}$$

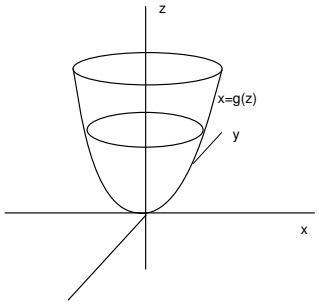


Konkret ist der Graph von $f(x, y) = \sqrt{R^2 - x^2 - y^2}$ eine Funktion, deren Graph die obere Halbkugelfläche vom Radius R ist. In diesem Fall liefert also (40) eine weitere Parametrisierung der Halbkugel neben der durch (38) mit $0 \leq \theta \leq \frac{\pi}{2}$ gegebenen. □

Beispiel 136 (Parametrisierung von Rotationsflächen). Ist $x = g(z)$ eine positive Funktion von z und rotiert man deren (in der (x, z)-Ebene liegenden) Graphen um die z-Achse, so erhält man eine Rotationsfläche, die man wie folgt parametrisieren kann:

$$\vec{x}(u, \phi) = \begin{pmatrix} g(u) \cos \phi \\ g(u) \sin \phi \\ u \end{pmatrix}. \tag{41}$$

Für $x = \sqrt{z}$ erhält man ein Rotationsparaboloid:



Dagegen ergibt $x = \sqrt{R^2 - z^2}$ eine weitere Parametrisierung der Kugelfläche. □

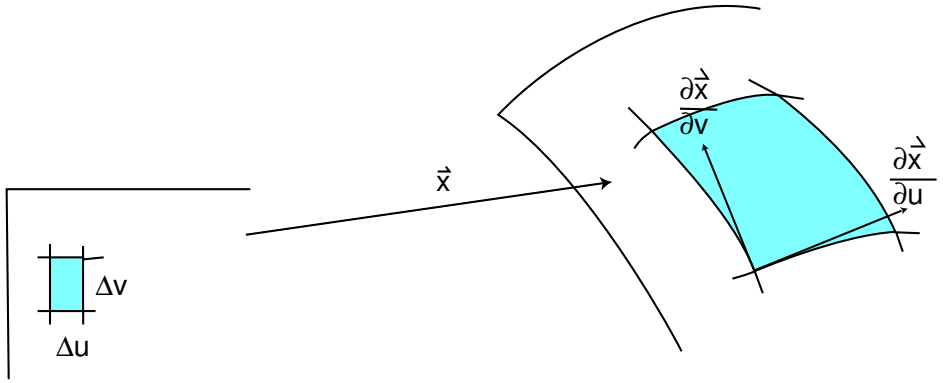
Das Oberflächenelement. Gegeben sei ein glattes Flächenstück

$$\vec{x} : B \to \mathbb{R}^3, \begin{pmatrix} u \\ v \end{pmatrix} \mapsto \begin{pmatrix} x(u, v) \\ y(u, v) \\ z(u, v) \end{pmatrix}.$$

Dann spannen die partiellen Ableitungen $\frac{\partial \vec{x}}{\partial u}, \frac{\partial \vec{x}}{\partial v}$ die Tangentialebene an die Fläche auf. Einem kleinen Rechteck mit diagonalen Ecken (u, v) und $(u+\Delta u, v+\Delta v)$ entspricht auf der Fläche annähernd ein Parallelogramm, das von den Vektoren $\Delta u \frac{\partial \vec{x}}{\partial u}$ und $\Delta v \frac{\partial \vec{x}}{\partial v}$ aufgespannt wird, und das die Fläche

$$\left| \frac{\partial \vec{x}}{\partial u} \times \frac{\partial \vec{x}}{\partial v} \right| \Delta u \Delta v$$

besitzt.



Wir bezeichnen deshalb den Ausdruck

$$dO = \left| \frac{\partial \vec{x}}{\partial u} \times \frac{\partial \vec{x}}{\partial v} \right| du dv$$

als das *(infinitesimale) Oberflächenelement*[2]. Zum Beispiel berechnet sich das Oberflächenelement der **Kugelkoordinatenparametrisierung** (38) wegen (39) als

$$\boxed{dO = R^2 \sin\theta d\theta d\phi.}$$

Wir definieren nun –ähnlich wie bei den Kurvenintegralen in Abschnitt 2.4– zwei Arten von Oberflächenintegralen: Wir integrieren skalare Funktionen und wir integrieren (im nächsten Abschnitt) Vektorfelder über eine Fläche.

Integration skalarer Funktionen.

Definition 137. Sei $\vec{x} : B \to \mathbb{R}^3$ eine glatte Fläche, $F := \vec{x}(B)$ die Bildmenge und $\delta : F \to \mathbb{R}$ eine skalare Funktion. Dann definiert man

$$\boxed{\iint\limits_F \delta dO := \iint\limits_{\vec{x}} \delta dO := \iint\limits_B \delta(\vec{x}) \left| \frac{\partial \vec{x}}{\partial u} \times \frac{\partial \vec{x}}{\partial v} \right| du dv.} \tag{42}$$

Für $\delta = 1$ erhält man die *Oberfläche* von F:

$$\boxed{O(F) = \iint\limits_B \left| \frac{\partial \vec{x}}{\partial u} \times \frac{\partial \vec{x}}{\partial v} \right| du dv.}$$

Ist δ eine Massen- oder Ladungsverteilung auf der Fläche, so liefert das Integral die Gesamtmasse oder -ladung.

[2] Vergleichen Sie die Situation bei der Transformationsformel. Eine ebene Koordinatentransformation kann man auch als Fläche im $\mathbb{R}^3$ ansehen, die „zufällig“ in $\mathbb{R}^2$ liegt. Dann ist $dO = dA$.

Beispiel 138 (Flächenträgheitsmoment).
Müller: Mechanik II, Abschnitt 14.2
Wir betrachten die Funktion $\delta(x,y,z) = \delta_0(x^2+y^2)$ für eine positive Konstante δ_0 und wollen diese über die Kugelfläche $S(R)$ vom Radius R um 0 integrieren. Wir benutzen die Parametrisierung (38) und finden:

$$\begin{aligned}\iint\limits_{S(R)} \delta_0(x^2+y^2)dO &= \delta_0 \int_0^{2\pi}\int_0^{\pi} (R^2\sin^2\theta\cos^2\phi + R^2\sin^2\theta\sin^2\phi)R^2\sin\theta d\theta d\phi \\ &= \delta_0 \int_0^{2\pi}\int_0^{\pi} R^4\sin^3\theta d\theta d\phi = \delta_0\cdot R^4\cdot 2\pi\int_0^{\pi}\sin^3\theta d\theta \\ &= \delta_0\cdot R^4\cdot 2\pi\cdot\frac{4}{3} = \frac{8\pi}{3}\delta_0 R^4.\end{aligned}$$

Ist δ_0 die homogene Massendichte der Kugelschale, so ist δ die Dichte des Trägheitsmoments bezüglich der z-Achse, und das obige Integral liefert das gesamte Trägheitmoment Θ. Die Gesamtmasse ist nach der Formel für die Kugelfläche $M = 4\pi R^2\delta_0$, so dass $\Theta = \frac{2}{3}MR^2$ in übereinstimmung mit Beispiel 131.

Beispiel 139 (Oberfläche von Rotationsflächen). Für die Parametrisierung einer Rotationsfläche wie im Beispiel 136 finden wir:

$$\frac{\partial\vec{x}}{\partial u}\times\frac{\partial\vec{x}}{\partial\phi} = \begin{pmatrix} g'(u)\cos\phi \\ g'(u)\sin\phi \\ 1\end{pmatrix}\times\begin{pmatrix} -g(u)\sin\phi \\ g(u)\cos\phi \\ 0\end{pmatrix} = \begin{pmatrix} -g(u)\cos\phi \\ -g(u)\sin\phi \\ g(u)g'(u)\end{pmatrix}.$$

Die Oberfläche wird daher

$$O(F) = \int_0^{2\pi}\int_a^b g(u)\sqrt{1+g'(u)^2}dud\phi = \int_a^b \underbrace{2\pi g(u)}_{Kreisumfang}\ \underbrace{\sqrt{1+g'(u)^2}du}_{Streifenbreite}.$$

□

Beispiel 140. Eine Ringkernspule besitze einen torusförmigen Kern mit einer Seele vom Radius R und einem Querschnitt vom Radius r. Hat die Wicklung eine Dicke h, die klein im Vergleich zu r ist, so ist das Volumen der Wicklung annähernd $h\times$Oberfläche des Torus. Wir berechnen diese Oberfläche. Eine Parametrisierung ist gegeben durch

$$\vec{x}(\phi,\theta) = R\begin{pmatrix}\cos\phi\\ \sin\phi\\ 0\end{pmatrix} + r\left(\begin{pmatrix}\cos\phi\\ \sin\phi\\ 0\end{pmatrix}\cos\theta + \begin{pmatrix}0\\ 0\\ 1\end{pmatrix}\sin\theta\right) = \begin{pmatrix}(R+r\cos\theta)\cos\phi\\ (R+r\cos\theta)\sin\phi\\ r\sin\theta\end{pmatrix}.$$

Wir erledigen die Rechenarbeit mittels Mathematica

```
In[1]:= u=D[{(R+ r Cos[Theta])Cos[Phi],(R+ r Cos[Theta])Sin[Phi],
          r Sin[Theta]},Theta];
        v=D[{(R+ r Cos[Theta])Cos[Phi],(R+ r Cos[Theta])Sin[Phi],
          r Sin[Theta]},Phi];
        w=Sqrt[Simplify[Cross[u,v].Cross[u,v]]]
Out[2]= Sqrt[r^2 (R + r Cos[Theta])^2]
        Integrate[w,{Theta,0,2 Pi},{Phi,0,2 Pi}]
Out[3]= (4 Pi^2 R Sqrt[r^2 (r+R)^2])/(r+R)
```

Mit u und v berechnen wir die partiellen Ableitungen nach θ,ϕ. Das Semikolon am Zeilenende unterdrückt den Ausdruck der Ergebnisse. Mit w berechnen wir dann das Skalarprodukt vom Kreuzprodukt der partiellen Ableitungen mit sich selbst, also $(\frac{\partial x}{\partial\theta}\times\frac{\partial x}{\partial\phi})\cdot(\frac{\partial x}{\partial\theta}\times\frac{\partial x}{\partial\phi})$ und ziehen daraus die Wurzel. Das Ergebnis findet sich in Out[2], das Ergebnis der Integration in Out[3]. Beachten Sie, dass das Programm nicht weiß, dass $0<r<R$. Folglich sind die Vorzeichen von $R+r\cos\theta$ oder $r+R$ nicht klar. Wir wissen darüber aber Bescheid und können das letzte Ergebnis zusammenfassen: Die Oberfläche ist

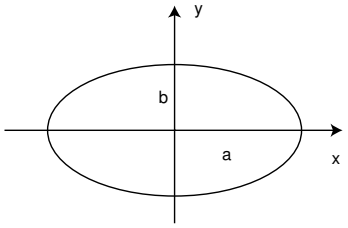
$$O = 4\pi^2 rR = (2\pi R)(2\pi r).$$

Dahinter steht wieder eine „Guldinsche Regel“ wie im Beispiel 126 des vorangehenden Abschnitts.

Beispiele für die Parametrisierung von Kurven und Flächen. Wir beschließen diesen Abschnitt mit einigen Beispielen für die Parametrisierung von Kurven und Flächen.

Ellipse

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1, \quad a, b > 0.$$



Parametrisierung der oberen Hälfte als Graph:

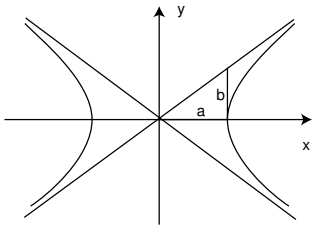
$$\vec{x}(u) = \left(u, b\sqrt{1 - \frac{u^2}{a^2}} \right), \quad -a \leq u \leq a.$$

Parametrisierung der ganzen Ellipse

$$\vec{x}(t) = (a\cos t, b\sin t), \quad 0 \leq t \leq 2\pi.$$

Hyperbel

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1, \quad a, b > 0.$$



Parametrisierung des rechten Astes als Graph über der y-Achse:

$$\vec{x}(u) = \left(a\sqrt{1 + \frac{u^2}{a^2}}, u \right), \quad u \in \mathbb{R}.$$

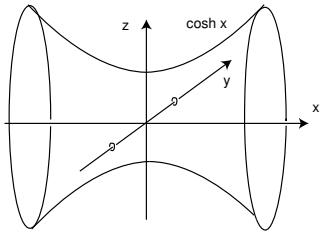
Andere Parametrisierung:

$$\vec{x}(t) = (a\cosh t, b\sinh t), \quad t \in \mathbb{R}$$

Katenoid

Rotationsfläche der Kettenlinie

$$z = \cosh x.$$

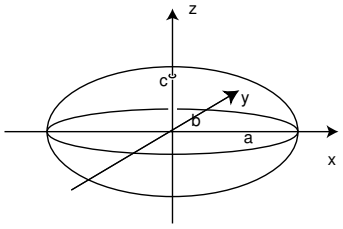


Parametrisierung als Rotationsfläche y-Achse:

$$\vec{x}(t, \phi) = (t, \cosh t\, \cos\phi, \cosh t \sin\phi), \quad y \in \mathbb{R}.$$

Ellipsoid
Gleichung

$$\frac{x^2}{a^2}+\frac{y^2}{b^2}+\frac{z^2}{c^2}=1, \quad a,b,c>0.$$



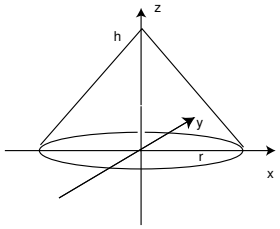
Parametrisierung der oberen Hälfte als Graph:

$$\vec{x}(u,v)=\left(u,v,c\sqrt{1-(\frac{u^2}{a^2}+\frac{v^2}{b^2})}\right), \quad \frac{u^2}{a^2}+\frac{v^2}{b^2}\leq 1.$$

Parametrisierung mit „elliptischen“ Koordinaten:

$$\vec{x}(\theta,\phi)=(a\sin\theta\cos\phi, b\sin\theta\sin\phi, c\cos\theta), \quad 0\leq\theta\leq\pi/2, 0\leq\phi\leq 2\pi.$$

Kreiskegel
Höhe h, Basisradius r.



Parametrisierung als Rotationsfläche um die z-Achse:

$$\vec{x}(u,\phi)=\left(a(1-\frac{u}{h})\cos\phi, a(1-\frac{u}{h})\sin\phi, u\right) \quad 0\leq u\leq h,\ 0\leq\phi\leq 2\pi.$$

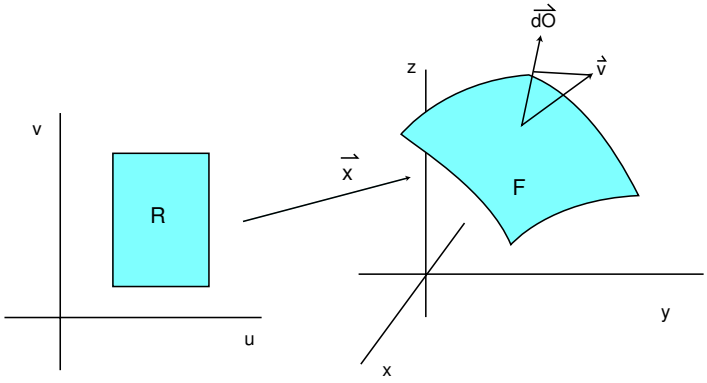
3.6 Integration von Vektorfeldern: Flussintegrale

- Den Fluss einer Strömung durch eine virtuelle Fläche berechnet man durch Integration des Geschwindigkeitsfeldes der Strömung über diese Fläche.
- Dazu erklären wir das vektorielle Oberflächenelement und das Flussintegral.

Viel wichtiger als die Integration skalarer Funktionen über Flächen ist in der Elektrotechnik oder der Dynamik von Gasen oder Flüssigkeiten die Integration von Vektorfeldern. Dabei deutet man die Vektorfelder als Geschwindigkeitsfelder stationärer Strömungen, und das Integral soll den Fluss durch die Fläche in der Zeiteinheit liefern. Dazu multipliziert man die Komponente des Vektorfeldes *senkrecht* zur Fläche mit dem Oberflächenelement und integriert dies. Es ist nützlich zu erinnern, dass $\frac{\partial \vec{x}}{\partial u} \times \frac{\partial \vec{x}}{\partial v}$ ein Vektor senkrecht zur Fläche ist, dessen Länge gerade der Betrag des Oberflächenelements ist. Man nennt

$$\boxed{d\vec{O} := \frac{\partial \vec{x}}{\partial u} \times \frac{\partial \vec{x}}{\partial v} dudv}$$

deshalb auch das *vektorielle Oberflächenelement*. Der gesuchte Fluss pro Zeiteinheit durch das Oberflächenelement ist also gerade $\vec{v} \cdot d\vec{O}$.



Definition 141 (Flußintegral). Sei $\vec{x} : B \to \mathbb{R}^3$ eine glatte Fläche und $F := \vec{x}(B)$ die Bildmenge. Sei weiter $\vec{v}$ ein stetiges Vektorfeld auf einer offenen Menge, die F enthält. Dann definieren wir das *(Fluss)integral* von $\vec{v}$ über F durch

$$\boxed{\iint\limits_F \vec{v} \cdot d\vec{O} := \iint\limits_{\vec{x}} \vec{v} \cdot d\vec{O} := \iint\limits_B \vec{v}(\vec{x}(u,v)) \cdot \left(\frac{\partial \vec{x}}{\partial u} \times \frac{\partial \vec{x}}{\partial v} \right) dudv.}$$

Der Integrand ist ein sogenanntes *Spatprodukt*, nämlich von der Form $\vec{a} \cdot (\vec{b} \times \vec{c})$. Aber dafür gibt es die Formel $\vec{a} \cdot (\vec{b} \times \vec{c}) = \det(\vec{a}, \vec{b}, \vec{c})$, bei der man kein Vektorprodukt mehr ausrechnen muss. Also kurz

$$\boxed{\iint\limits_F \vec{v} \cdot d\vec{O} := \iint\limits_B \det \left(\vec{v}(\vec{x}), \frac{\partial \vec{x}}{\partial u}, \frac{\partial \vec{x}}{\partial v} \right) dudv.}$$

Diese Formel ist allerdings kein Fortschritt, wenn man $d\vec{O}$ bereits ausgerechnet hat, wie wir z.B. in (39) für die Kugelkoordinatenparametrisierung.

Beispiel 142. Wie groß ist der Fluss des homogenen Feldes $\vec{E} = \begin{pmatrix} 0 \\ 0 \\ E_3 \end{pmatrix}$ durch den Graphen der Funktion $f(u,v) = uv + v^2$ über der Kreisscheibe $B : u^2 + v^2 \leq 9$? Wir benutzen die Parametrisierung

$$\vec{x}(u,v) = \begin{pmatrix} u \\ v \\ uv + v^2 \end{pmatrix}.$$

Dann ist

$$\iint_F \vec{E} \cdot d\vec{O} = \iint_B \det \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ E_3 & v & u+2v \end{pmatrix} dudv = \iint_B E_3 dudv = 9\pi E_3.$$

Der Fluss hängt also gar nicht von der Funktion f ab, er ist für alle Graphen über der Kreisscheibe B gleich. Machen Sie sich klar, dass man das auch erwarten sollte.

Beispiel 143 (Fluss des Coulombfeldes).
Theoretische Elektrotechnik, Abschnitt 1.4
Wie groß ist der Fluss des elektrischen Feldes einer Punktladung Q im Ursprung durch eine Kugelfläche $S(R)$ vom Radius R um den Ursprung? In diesem Fall ist also

$$\vec{E} = \frac{Q}{4\pi\epsilon} \frac{\vec{x}}{|\vec{x}|^3}.$$

Wir benutzen die Kugelkoordinatenparametrisierung (39) und erhalten

$$\begin{aligned} \iint_{S(R)} \vec{E} \cdot d\vec{O} &= \int_0^{2\pi} \int_0^{\pi} \frac{Q}{4\pi\epsilon} \frac{1}{R^3} \vec{x} \cdot (R \sin\theta \vec{x})\, d\theta d\phi \\ &= \frac{Q}{4\pi\epsilon} \frac{R^3}{R^3} \int_0^{2\pi} \int_0^{\pi} \sin\theta d\theta d\phi = \frac{Q}{\epsilon}. \end{aligned}$$

Der Fluss ist insbesondere unabhängig vom Radius der Kugel. Das sollte man auch erwarten, weil die Stärke der Strömung quadratisch mit dem Radius abnimmt, während die Fläche der Sphäre quadratisch mit dem Radius wächst.

Parameterinvarianz und Orientierung. Wir haben uns bei der Integration über Kugeln in den vorstehenden Beispielen immer für die Parametrisierung durch Kugelkoordinaten entschieden. Wäre das Ergebnis bei einer Parametrisierung z.B. als Rotationsfläche von $x = \sqrt{R^2 - z^2}$ dasselbe gewesen? Wenn wir $\iint_F \ldots$ schreiben, betrachten wir entgegen unserer anfänglichen Definition die Fläche F eben doch als Punktmenge im $\mathbb{R}^3$ und nicht als Abbildung. Tatsächlich kann man zeigen, dass – wie es in den physikalischen Beispielen natürlich nicht anders sein darf – das Integral skalarer Funktionen über eine Fläche unabhängig ist von der Parametrisierung. Beim Flussintegral ergibt sich allerdings mögli-

cherweise ein Vorzeichenwechsel, der auch physikalisch verständlich ist: Man muss spezifizieren, in welcher Richtung der Durchfluss positiv gerechnet werden soll. Bei parametrisierten Flächen geschieht das so, dass man die Richtung von $d\vec{O}$ als positiv definiert. Dazu braucht man aber eben eine Parametrisierung. Verwendet man für die Kugel statt $(u, v) = (\theta, \phi)$ die Koordinaten $(u, v) = (\phi, \theta)$ in anderer Reihenfolge, so ändert $d\vec{O}$ und damit auch das Flussintegral sein Vorzeichen.

Wir wollen bei der Kugel und allgemeiner bei Oberflächen kompakter Bereiche vereinbaren, dass das vektorielle Oberflächenelement nach außen weisen soll.

Falls Sie eine Parametrisierung gefunden haben, bei der das Oberflächenelement ins Innere des kompakten Bereiches weist, können Sie am einfachsten damit weiterarbeiten, müssen aber das Vorzeichen des Flussintegrals umkehren.

Warum vereinbaren wir nicht gleich für alle Flächen, wohin das Oberflächenelement zeigen soll? Weil es dafür keine sinnvolle Methode gibt: Wenn die Fläche z.B. ein Quadrat ist, das irgendwie im 3-dimensionalen Raum liegt, dann macht es keinen Sinn von innen und außen zu sprechen. Das geht nur bei Oberflächen von Körpern. (Vergleichen Sie auch die *Konvention* zum Satz von Stokes im nächsten Abschnitt.)

FLÄCHENSTÜCKE UND FLÄCHEN. Typische Flächen, über die man im Ingenieurbereich integrieren möchte, sind neben den Kugeln die Oberflächen von Würfeln (allgemeiner von Quadern) und von Kreiszylindern. Die beiden letzteren bestehen allerdings nicht aus einem, sondern aus mehreren glatten Flächenstücken im Sinne unserer Definition. Jede „Seite“ ist ein solches glattes Flächenstück. Wir definieren deshalb einfach eine Fläche F als eine endliche Familie von glatten Flächenstücken $F_1, \ldots, F_k$ und definieren

$$\iint\limits_F \ldots = \iint\limits_{F_1} \ldots + \ldots + \iint\limits_{F_k} \ldots$$

Das Integral über einen Würfel ist also die Summe über die Integrale über die sechs Seiten, wobei wieder alle vektoriellen Oberflächenelemente nach außen weisen sollen. Das Integral über einen Kreiszylinder ist das Integral über den Mantel plus die Integrale über Boden und Deckel, wieder mit der Orientierungsvereinbarung wie oben.

3.7 Der Integralsatz von Gauß

- Wir lernen den Integralsatz von Gauß kennen, der in der Sprache der Mathematik einen anschaulichen physikalischen Sachverhalt beschreibt: Die Strömungsbilanz durch eine *geschlossene* virtuelle Fläche gibt Aufschluss über die von ihr eingeschlossenen Quellen.

Wir haben früher die Divergenz eines Vektorfeldes als seine Quelldichte eingeführt. Betrachtet man die Fluss-Bilanz über die Oberfläche eines dreidimensionalen Bereiches B, so ist aus physikalischen Gründen einleuchtend, dass der Überschuss des austretenden Flusses über den eintretenden gerade die gesamte Quelle im Innern des Bereiches wiedergibt. Dies lässt sich mathematisch beweisen und ist der Inhalt des folgenden Satzes.

Satz 144 (Integralsatz von GAUSS = Divergenzsatz). *Sei B ein kompakter Bereich, dessen Rand ∂B eine Fläche im Sinne des letzten Abschnitts ist. Die Parametrisierung dieser Fläche sei so gewählt, dass die Normale überall aus dem Bereich B heraus weist.*
Sei $\vec{v}$ ein Vektorfeld auf einer offenen Umgebung von B mit stetigen partiellen Ableitungen. Dann gilt

$$\iiint_B \operatorname{div} \vec{v}\, dxdydz = \iint_{\partial B} \vec{v} \cdot d\vec{O}.$$

Typische Beispiele für B sind die Vollkugel vom Radius R oder ein Würfel der Kantenlänge $2R$. Im ersten Fall kann man die Kugelfläche $\partial B = S(R)$ dann durch die Kugelkoordinaten parametrisieren. Im zweiten Fall besteht ∂B aus den sechs Seiten des Würfels, die einzeln so zu parametrisieren sind, dass das vektorielle Flächenelement nach außen weist, vgl. den folgenden

Beweis. Wir beschränken uns auf den Spezialfall, dass B ein Würfel W der Kantenlänge $2R$ um den Nullpunkt ist. Zunächst wählen wir auch ein sehr spezielles Vektorfeld, nämlich eines der Form

$$\vec{v}(x,y,z) = \begin{pmatrix} f(x,y,z) \\ 0 \\ 0 \end{pmatrix}.$$

Dann ist $\operatorname{div} \vec{v} = \frac{\partial f}{\partial x}(x,y,z)$ und

$$\begin{aligned} \iiint_W \operatorname{div} \vec{v}\, dxdydz &= \int_{-R}^{R}\int_{-R}^{R}\left(\int_{-R}^{R} \frac{\partial f}{\partial x} dx\right) dydz \\ &= \int_{-R}^{R}\int_{-R}^{R} (f(R,y,z) - f(-R,y,z))\, dydz. \end{aligned} \tag{43}$$

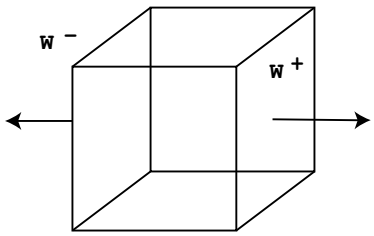
Wir berechnen nun andrerseits das Flussintegral von $\vec{v}$ durch die Randfläche des Würfels. Weil $\vec{v}$ nur eine x-Komponente hat, ist der Fluß durch die obere und untere sowie durch die vordere und hintere Würfelseite $= 0$, und wir müssen nur die rechte und linke Seite

$$W^+ = \{(x,y,z) \in \partial W \,|\, x = +R\}$$

und

$$W^- = \{(x,y,z) \in \partial W \,|\, x = -R\}$$

des Würfels betrachten:



Wir wählen folgende Parametrisierungen:

$$W^+: \quad \vec{x}^+(u,v) = (R,u,v),$$
$$W^-: \quad \vec{x}^-(u,v) = (-R,u,v),$$

wo $-R \le u,v \le R$. Das vektorielle Oberflächenelement ist für beide Parametrisierungen

$$d\vec{O} = \begin{pmatrix}0\\1\\0\end{pmatrix} \times \begin{pmatrix}0\\0\\1\end{pmatrix} dudv = \begin{pmatrix}1\\0\\0\end{pmatrix} dudv.$$

Für W^- weist es also in den Würfel hinein, und wir müssen das korrigieren, indem wir das Integral abziehen. Der Fluss ist damit

$$\begin{aligned}\iint\limits_{\partial W} \vec{v}\cdot d\vec{O} &= \int_{-R}^{R}\int_{-R}^{R} \begin{pmatrix}f(R,u,v)\\0\\0\end{pmatrix} \cdot \begin{pmatrix}1\\0\\0\end{pmatrix} dudv - \int_{-R}^{R}\int_{-R}^{R} \begin{pmatrix}f(-R,u,v)\\0\\0\end{pmatrix} \cdot \begin{pmatrix}1\\0\\0\end{pmatrix} dudv \\ &= \int_{-R}^{R}\int_{-R}^{R} f(R,u,v)\,dudv - \int_{-R}^{R}\int_{-R}^{R} f(-R,u,v)\,dudv \\ &= \int_{-R}^{R}\int_{-R}^{R} (f(R,u,v) \ - f(-R,u,v))dudv.\end{aligned}$$

Vergleich mit (43) zeigt die Behauptung.

Natürlich klappt der Beweis auch, wenn $\vec{v}$ von der Form $\begin{pmatrix}0\\g(x,y,z)\\0\end{pmatrix}$ oder $\begin{pmatrix}0\\0\\h(x,y,z)\end{pmatrix}$ ist. Deshalb gilt er aber auch für beliebiges

$$\vec{v} = \begin{pmatrix}f(x,y,z)\\g(x,y,z)\\h(x,y,z)\end{pmatrix} = \begin{pmatrix}f(x,y,z)\\0\\0\end{pmatrix} + \begin{pmatrix}0\\g(x,y,z)\\0\end{pmatrix} + \begin{pmatrix}0\\0\\h(x,y,z)\end{pmatrix}.$$

Schließlich muss man zeigen, dass der Satz auch für „deformierte“ Würfel richtig bleibt und dass man „jeden“ Bereich B in solche zerlegen kann. Das ist aber technisch sehr kompliziert. □

Der entscheidende Schritt im obigen Beweis, bei dem nämlich die Ableitung von $\vec{v}$ (Divergenz) verschwindet und aus dem Dreifachintegral ein Doppelintegral wird, ist die Stelle (43). Hier benutzt man den Fundamentalsatz der Differentialrechnung $\int_a^b f'(x)dx = f(b) - f(a)$, den man als eine eindimensionale Version des Satzes von Gauß auffassen kann: Es ist ja $f'(x) = \frac{\partial f}{\partial x} = \operatorname{div} f$ für eindimensionale Vektorfelder. Rechts steht dann gewissermaßen als 0-faches Integral die Differenz der Funktionswerte auf dem Rand des Integrationsbereichs $[a,b]$.

Beispiel 145 (Ein Messverfahren für die Divergenz). Betrachten wir einen Würfel $W(R)$ der Kantenlänge $2R$ um einen Punkt $\vec{x}$ im Strömungsfeld $\vec{v}$, so ist, wenn die partiellen

Ableitungen von $\vec{v}$ stetig sind und R klein ist, die Divergenz von $\vec{v}$ auf dem Würfel annähernd konstant, also ist

$$\operatorname{div}_{\vec{x}} \vec{v} \approx \frac{1}{(2R)^3} \iiint\limits_{W(R)} \operatorname{div} \vec{v}\, dxdydz = \frac{1}{(2R)^3} \iint\limits_{\partial W(R)} \vec{v} \cdot d\vec{O}.$$

Der mit dem Volumen normierte Fluß/Zeiteinheit durch die Würfelfläche misst also die Quelldichte $\operatorname{div} \vec{v}$. □

Beispiel 146 (Wärmeleitungsgleichung).
Energie-, Impuls- und Stofftransport, Abschnitt 2.1
Wir betrachten einen kompakten Bereich B in einem (wärmeisotropen) Medium und wollen darin die Veränderung der Temperatur in Abhängigkeit von Ort und Zeit beschreiben. Die Gesamtenergie U in B ergibt sich durch Integration der spezifischen inneren Energie u multipliziert mit der Dichte ρ:

$$U = \iiint\limits_B \rho\, u\, dxdydz.$$

Die zeitliche Änderung $\frac{dU}{dt}$ der Gesamtenergie ist gleich dem gesamten Wärmefluß durch die Oberfläche in B *hinein, daher das Minuszeichen*:

$$\frac{dU}{dt} = \frac{d}{dt} \iiint\limits_B \rho\, u\, dxdydz = \iiint\limits_B \rho \frac{\partial u}{\partial t}\, dxdydz = -\iint\limits_{\partial B} \dot{\vec{q}} \cdot d\vec{O}.$$

Mit dem Gaußschen Satz folgt

$$\iiint\limits_B \rho\, \frac{\partial u}{\partial t}\, dxdydz = -\iiint\limits_B \operatorname{div} \dot{\vec{q}}\, dxdydz. \tag{44}$$

Daraus ergibt sich wie im vorstehenden Beispiel

$$\rho\, \frac{\partial u}{\partial t} = -\operatorname{div} \dot{\vec{q}}. \tag{45}$$

Nun ist die zeitliche Änderung der spezifischen Energie proportional zur Änderung der Temperatur

$$\frac{\partial u}{\partial t} = c \frac{\partial T}{\partial t}, \qquad c = \text{spezifische Wärmekapazität},$$

und nach dem Fourierschen Gesetz ist der Wärmefluß proportional zum Gradienten der Temperatur: Die Wärme fließt in Richtung des steilsten Temperatur*abfalls*:

$$\dot{\vec{q}} = -\lambda \operatorname{grad} T, \qquad \lambda = \text{Wärmeleitfähigkeit}.$$

Einsetzen der beiden letzten Identitäten in (45) liefert wegen $\Delta = \operatorname{div} \operatorname{grad}$ die Wärmeleitungsgleichung.

$$\boxed{c\rho \frac{\partial T}{\partial t} = \lambda \Delta T.}$$

Mit ähnlicher Argumentation ergibt sich, dass die Gleichung auch Diffusionsprozesse beschreibt, vgl. z.B. Energie-, Impuls- und Stofftransport, Abschnitt 2.1, VTI, Kapitel 2 oder Werkstoffe II, Abschnitt 8.3.

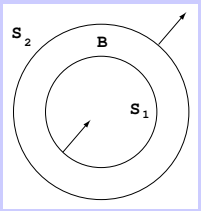
Beispiel 147 (Fluss des Coulombfeldes).
Theoretische Elektrotechnik, Abschnitt 1.4
Als Vektorfeld betrachten wir das elektrische Feld der Punktladung

$$\vec{E}(x,y,z) = \frac{Q}{4\pi\epsilon}\frac{\vec{x}}{\sqrt{x^2+y^2+z^2}^3},$$

vgl. auch Beispiel 143. Auf dieses Vektorfeld und auf B=Vollkugel um 0 mit Radius R lässt sich der Satz von Gauß NICHT anwenden, denn $\vec{E}$ ist auf B nicht differenzierbar: in 0 ist es nicht einmal stetig. Überall sonst in $\mathbb{R}^3$ ist allerdings $\operatorname{div}\vec{E} = 0$, wie wir früher nachgerechnet haben. Wir betrachten nun als Bereich B eine *Hohlkugel* um 0 mit innerem Radius R_1 und äußerem Radius R_2. Wir bezeichnen mit S_1 und S_2 die beiden Randkomponenten, die innere und äußere Kugelfläche. Nach dem Satz von Gauß ist dann wegen des Verschwindens der Divergenz

$$0 = \iiint_B \operatorname{div}\vec{E}\,dxdydz = \iint_{S_1}\vec{E}\cdot d\vec{O}_1 + \iint_{S_2}\vec{E}\cdot d\vec{O}_2$$

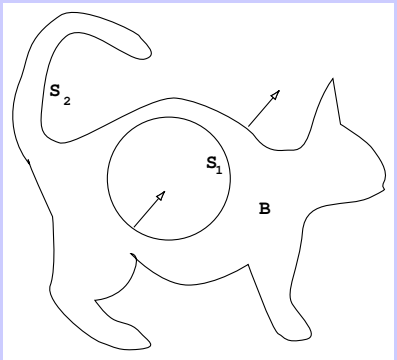
Beachten Sie die Orientierung der beiden Randkomponenten: Auf S_2 zählt der Fluss weg vom Nullpunkt positiv, auf S_1 aber negativ, weil das aus B herausweisende vektorielle Oberflächenelement *in* die kleine Kugel *hinein* zeigt:



Bleibt man hingegen bei der früheren Konvention über die Parametrisierung von Kugelflächen (nach außen ist positiv), so muss man schreiben

$$0 = -\iint_{S_1}\vec{E}\cdot d\vec{O}_1 + \iint_{S_2}\vec{E}\cdot d\vec{O}_2.$$

Der Fluss durch die kleine Kugelfläche ist also gleich dem durch die große. Das hatten wir früher schon einmal explizit ausgerechnet. Aber der Satz von Gauß liefert erheblich viel mehr: Man kann um die innere Kugel einen beliebigen Körper herumlegen, und der Fluss durch die äußere Fläche bleibt gleich dem durch die kleine Kugel.

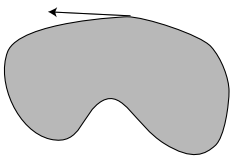


Natürlich könnte man die kleine Kugel auch noch deformieren ...

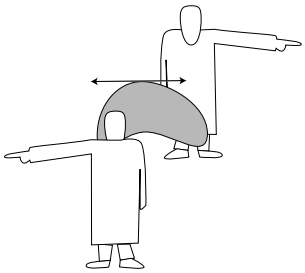
3.8 Der Integralsatz von Stokes

- Wir überlegen, wie man den Rand eines Flächenstücks im dreidimensionalen Raum orientieren kann.
- Der Satz von Stokes, den wir dann betrachten wollen, ist weniger anschaulich als der Satz von Gauß. Er spielt aber zum Beispiel bei Modellen für die elektromagnetische Induktion ein wichtige Rolle.
- Wir betrachten als Anwendungen die Gaußsche Formel zur Flächenberechnung in der Geodäsie und den Carnotschen Kreisprozess.

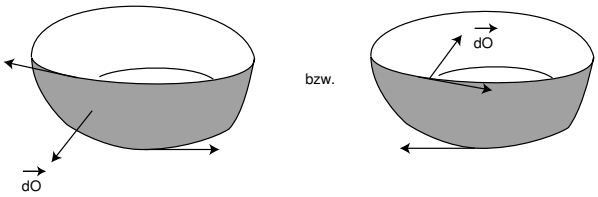
Eine Konvention. Im folgenden Integralsatz von Stokes kommt ein Kurvenintegral über die Randkurve einer Fläche vor, das davon abhängt, in welcher Richtung man die Kurve durchläuft. Bei einem Gebiet in der Ebene kann man die Randkurve so durchlaufen, dass das Gebiet zur Linken liegt: Man nennt das auch den mathematisch positiven Umlaufsinn (entgegen dem Uhrzeiger).



Bei einer Fläche im Raum ist aber kein solcher Umlaufsinn auf natürliche Weise ausgezeichnet:

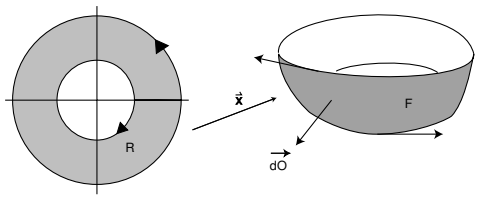


Ist die Fläche im Raum hingegen parametrisiert, so zeichnet das vektorielle Oberflächenelement eine Seite der Fläche als „Oberseite“ aus. Von *der* Seite aus soll man dann die mathematisch positive Umlaufrichtung für die Randkurve wählen:



(Bei dieser Fläche hat die Randkurve zwei *Komponenten*.)

Kann man die Fläche in einem Stück parametrisieren, so hat ihr Parameterbereich in der uv-Ebene eine kanonische Randorientierung mit dem Parameterbereich zur Linken. Diese Parametrisierung der Randkurve entspricht unter der Parameterabbildung $\vec{x}$ gerade der richtigen Randorientierung der Fläche.



Satz 148 (Integralsatz von Stokes). *Sei F eine (parametrisierte) Fläche im Raum $\mathbb{R}^3$ mit stückweise glatter Randkurve ∂F, die durchlaufen wird, wie oben beschrieben. Sei $\vec{v}$ ein Vektorfeld mit stetigen partiellen Ableitungen auf einer offenen Umgebung von F. Dann gilt*

$$\iint_F \operatorname{rot} \vec{v} \cdot d\vec{O} = \int_{\partial F} \vec{v} \cdot d\vec{s}.$$

Ist die Fläche geschlossen (wie die Kugel- oder Würfelfläche), also $\partial F = \emptyset$, so gilt

$$\iint_F \operatorname{rot} \vec{v} \cdot d\vec{O} = 0.$$

Auf den Beweis gehen wir nicht ein.

Bemerkungen. 1. Beide Integrale in dem Satz hängen von der „Orientierung“ ab. Beim Flächenintegral kommt es darauf an, in welche Richtung die Normale zeigt, beim Randintegral auf den Umlaufsinn. Eine Änderung ergibt jeweils eine Änderung des Vorzeichens. Durch unsere Vereinbarung über die Randorientierung haben wir die beiden so gekoppelt, dass der Satz richtig ist.

2. Manche Leute können sich nicht merken, auf welcher Seite $\vec{v}$ und auf welcher $\operatorname{rot} \vec{v}$ steht. Merkregel: „Ein Integral weniger = eine Differentiation weniger“. Dasselbe gilt auch für den Satz von Gauß.

3. Wir erinnern noch einmal daran, dass Flächen für uns aus endlich vielen glatten Flächenstücken bestehen, und dass glatte Flächenstücke eine Parametrisierung mit den zu Beginn von Abschnitt 3.5 gemachten Eigenschaften besitzen. Die Randkurve ∂F besteht deshalb immer aus einer oder mehreren geschlossenen Kurve. Das wird gelegentlich unterstrichen durch die folgende Notation

$$\oint_{\partial F} \vec{v} \cdot d\vec{s}.$$

Man nennt das auch ein Kontourintegral.

Beispiel 149. Dieses Beispiel hat keine physikalische Bedeutung, es soll nur noch einmal die Integraldefinitionen demonstrieren, die in den Stokesschen Satz eingehen. Sei

$$\vec{v}(x, y, z) = \begin{pmatrix} y^3 \\ x^2 \\ z \end{pmatrix}$$

und F die obere Halbsphäre vom Radius R mit der Kugelkoordinatenparametrisierung. Wir berechnen zunächst das Randintegral. Die Randkurve ist der Kreis vom Radius R in der xy-Ebene mit dem positivem Umlaufsinn:

$$\partial F: \quad \vec{x}(t) = R(\cos t, \sin t, 0), \quad 0 \le t \le 2\pi.$$

Damit wird

$$\begin{aligned}\int_{\partial F} \vec{v}\cdot\vec{ds} &= \int_0^{2\pi} \vec{v}(\vec{x}(t))\cdot\dot{\vec{x}}(t)dt \\ &= \int_0^{2\pi} \begin{pmatrix} R^3\sin^3 t \\ R^2\cos^2 t \\ 0 \end{pmatrix}\cdot\begin{pmatrix} -R\sin t \\ R\cos t \\ 0 \end{pmatrix} dt = \int_0^{2\pi}(-R^4\sin^4 t + R^3\cos^3 t)dt \\ &= -R^4\int_0^{2\pi}\sin^4 t dt = -\frac{3}{4}\pi R^4.\end{aligned}$$

Nun berechnen wir das Flächenintegral. Zunächst ist $\operatorname{rot}\vec{v} = \begin{pmatrix} 0 \\ 0 \\ 2x-3y^2 \end{pmatrix}$. Das vektorielle Oberflächenelement der Kugelkoordinaten war $R\sin\theta\vec{x}d\theta d\phi$. Damit erhalten wir

$$\begin{aligned}\iint_F \operatorname{rot}\vec{v}\cdot d\vec{O} &= \int_0^{2\pi}\int_0^{\pi/2} \begin{pmatrix} 0 \\ 0 \\ 2R\sin\theta\cos\phi - 3R^2\sin^2\theta\sin^2\phi \end{pmatrix}\cdot(R\sin\theta\begin{pmatrix} R\sin\theta\cos\phi \\ R\sin\theta\sin\phi \\ R\cos\theta \end{pmatrix})d\theta d\phi \\ &= \int_0^{2\pi}\int_0^{\pi/2} R^2\sin\theta\cos\theta(2R\sin\theta\cos\phi - 3R^2\sin^2\theta\sin^2\phi)d\theta d\phi \\ &= 2R^3\int_0^{\pi/2}\sin^2\theta\cos\theta d\theta \underbrace{\int_0^{2\pi}\cos\phi d\phi}_{=0} - 3R^4\int_0^{\pi/2}\sin^3\theta\cos\theta d\theta\int_0^{2\pi}\sin^2\phi d\phi \\ &= -3R^4(\frac{1}{4}\sin\frac{\pi}{2})\pi = -\frac{3}{4}\pi R^4.\end{aligned}$$

Stimmt also. □

Beispiel 150.
Theoretische Elektrotechnik, Abschnitt 6
Aus der 2. Maxwellschen Gleichung

$$\operatorname{rot} \vec{E} = -\frac{\partial \vec{B}}{\partial t}$$

erhält man durch Integration über eine Fläche F:

$$\begin{aligned} -\frac{\partial}{\partial t} \iint_F \vec{B} \cdot d\vec{O} &= \iint_F -\frac{\partial \vec{B}}{\partial t} \cdot d\vec{O} \\ &= \iint_F \operatorname{rot} \vec{E} \cdot d\vec{O} \\ &= \int_{\partial F} \vec{E} \cdot d\vec{s}. \end{aligned}$$

Dies ist das *Faradaysche Induktionsgesetz*: Die Änderung des magnetischen Flusses durch die Fläche ist bis aufs Vorzeichen gleich der elektrischen Randspannung.
Ebenso folgt aus der 1. Maxwellschen Gleichung $\operatorname{rot} \vec{H} = \vec{j} + \frac{\partial \vec{D}}{\partial t}$ das *Durchflutungsgesetz*

$$\iint_F \vec{j} \cdot d\vec{O} + \frac{\partial}{\partial t} \iint_F \vec{D} \cdot d\vec{O} = \int_{\partial F} \vec{H} \cdot d\vec{s}.$$

Beispiel 151. Sei $\vec{v} = -\operatorname{grad} u$ ein Vektorfeld, welches ein *Potential* u besitzt. Dann gilt für jede Kurve $\vec{x} : [a, b] \to \mathbb{R}^3$, dass $\int_{\vec{x}} \vec{v} \cdot d\vec{s} = u(\vec{x}(a)) - u(\vec{x}(b))$. Also gilt

$$\int_{\vec{x}} \vec{v} \cdot d\vec{s} = 0 \quad \text{für jede geschlossene Kurve } \vec{x}. \tag{46}$$

Ist die geschlossene Kurve der Rand einer Fläche F, so folgt das auch aus dem Satz von Stokes, denn dann ist $\int_{\vec{x}} \vec{v} \cdot d\vec{s} = -\iint_F \operatorname{rot} \operatorname{grad} u \cdot d\vec{O}$, und $\operatorname{rot} \operatorname{grad} u = 0$.

Sei nun $\vec{v} = \operatorname{rot} \vec{w}$ ein Vektorfeld, welches ein *Vektor*potential $\vec{w}$ besitzt. Dann gilt nach dem Satz von Stokes

$$\iint_F \vec{v} \cdot d\vec{O} = \iint_F \operatorname{rot} \vec{w} \cdot d\vec{O} = 0 \quad \text{für jede } \textit{geschlossene} \text{ Fläche } F. \tag{47}$$

Das ist also eine notwendige Bedingung für die Existenz eines Vektorpotentials, so wie (46) eine notwendige (und hinreichende) Bedingung für die Existenz eines gewöhnlichen Potentials ist.

Das zentrale Kraftfeld

$$\vec{v} = -\frac{\vec{x}}{|\vec{x}|^3} = \operatorname{grad} \frac{1}{|\vec{x}|}$$

erfüllt zwar die notwendige Bedingung $\operatorname{div} \vec{v} = 0$, aber nicht die Bedingung (47) und hat deshalb kein Vektorpotential. □

Beispiel 152 (Ein Messverfahren für die Rotation). Vergleiche Beispiel 86. Wir betrachten in einem Punkt $\vec{x}$ der Strömung $\vec{v}$ ein Rad vom (kleinen) Radius R, das die Strömung um eine zur z-Achse parallele Achse dreht. Dann ist die Drehgeschwindigkeit proportional zu

$$\int_{\partial K} \vec{v} \cdot d\vec{s}.$$

Dabei ist das Integral über den Rand des Rades zu nehmen. Betrachten wir nun die Kreisfläche K des Rades mit einer Parametrisierung, so dass

$$d\vec{O} = \vec{e_3} dO$$

ist, und wählen wir R so klein, dass $\operatorname{rot} \vec{v}$ auf dem Rad annähernd konstant ist, so folgt aus dem Satz von Stokes

$$\operatorname{rot} \vec{v} \cdot \vec{e_3} \approx \frac{1}{\pi R^2} \iint\limits_K \operatorname{rot} \vec{v} \cdot \vec{e_3} dO = \frac{1}{\pi R^2} \int_{\partial K} \vec{v} \cdot \vec{ds}.$$

Die z-Komponente der Rotation ist also die normierte Drehgeschwindigkeit des Rades. Entsprechend kann man mit anderer Achsenstellung die x- und y-Komponenten bestimmen. Stellt man die Achse so, dass die Drehgeschwindigkeit (im positiven Sinne) maximal ist, so zeigt die Achse in Richtung der Rotation. □

Beispiel 153 (Greensche Formel). In diesem Beispiel betrachten wir ebene Vektorfelder $\vec{v} = \begin{pmatrix} p(x,y) \\ q(x,y) \end{pmatrix}$ auf die natürliche Weise als Vektorfelder im $\mathbb{R}^3$, nämlich vermöge

$$\vec{v}(x,y,z) = \begin{pmatrix} p(x,y) \\ q(x,y) \\ 0 \end{pmatrix}.$$

Dann ist

$$\operatorname{rot} \vec{v} = \begin{pmatrix} 0 \\ 0 \\ \frac{\partial q}{\partial x} - \frac{\partial p}{\partial y} \end{pmatrix}.$$

Ist nun $F \subset \mathbb{R}^2 \subset \mathbb{R}^3$ ein ebener Bereich mit Randkurve $\vec{x}$, so ist das vektorielle Flächenelement von F einfach $\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} dxdy$ und der Satz von Stokes liefert

$$\iint\limits_F \left(\frac{\partial q}{\partial x} - \frac{\partial p}{\partial y} \right) dxdy = \int_{\vec{x}} \begin{pmatrix} p \\ q \end{pmatrix} \cdot \vec{ds}$$

oder, wenn $\vec{x} = \begin{pmatrix} x \\ y \end{pmatrix} : [a,b] \to \mathbb{R}^2$,

$$\iint\limits_F \left(\frac{\partial q}{\partial x} - \frac{\partial p}{\partial y} \right) dxdy = \int_a^b \left(p(\vec{x}(t)) \frac{dx}{dt} + q(\vec{x}(t)) \frac{dy}{dt} \right) dt. \tag{48}$$

Man schreibt das auch kurz als sogenannte GREENsche Formel

$$\boxed{\iint\limits_F \left(\frac{\partial q}{\partial x} - \frac{\partial p}{\partial y} \right) dxdy = \int_{\vec{x}} (pdx + qdy).}$$

Das ist hübsch, aber wenn Sie die rechte Seite berechnen müssen, ist es gut, sich an (48) zu erinnern. □

Beispiel 154. Betrachtet man im vorstehenden Beispiel den Spezialfall $p(x,y) = 0$, $q(x,y) = x$, so wird $\frac{\partial q}{\partial x} - \frac{\partial p}{\partial y} = 1$. Dann folgt aber

$$\iint_F 1 dx dy = \int_a^b (p\dot{x} + q\dot{y}) dt = \int_a^b x(t)\dot{y}(t) dt.$$

Das linke Integral ist gerade der Flächeninhalt von F. Mit $p = -y, q = 0$ erhält man ein analoges Resultat. Es ist

$$\boxed{\text{Fläche von } F = \int_a^b x(t)\dot{y}(t) dt = -\int_a^b y(t)\dot{x}(t) dt.} \tag{49}$$

Man kann also die Fläche von F durch ein Randintegral ausrechnen. Das wurde früher bei den Planarimetern zur graphischen Flächenberechnung benutzt. □

Beispiel 155 (Gaußsche Flächenformel). Ist F speziell ein Polygon mit Eckpunkten (x_i, y_i), $i = 0, \ldots, n$ und $(x_n, y_n) = (x_0, y_0)$, $(x_{n+1}, y_{n+1}) := (x_1, y_1)$, so ist

$$\vec{x}(t) = (x_i + t(x_{i+1} - x_i), y_i + t(y_{i+1} - y_i))$$

eine Parametrisierung der i-ten Randkante, und der Integralbeitrag dieser Kante wird

$$\begin{aligned}\int_{\vec{x}} x(t)\dot{y}(t) dt &= \int_0^1 (x_i + t(x_{i+1} - x_i))(y_{i+1} - y_i) dt \\ &= x_i(y_{i+1} - y_i) + \frac{1}{2}(x_{i+1} - x_i)(y_{i+1} - y_i) \\ &= \frac{1}{2} x_i(y_{i+1} - y_i) + \frac{1}{2} x_{i+1}(y_{i+1} - y_i).\end{aligned}$$

Daraus ergibt sich die in der Landvermessung heute noch wichtige Flächenformel von Gauß:

$$\begin{aligned}\text{Fläche von } F &= \frac{1}{2}\sum_{i=0}^{n-1} x_i(y_{i+1} - y_i) + \frac{1}{2}\sum_{i=0}^{n-1} x_{i+1}(y_{i+1} - y_i) \\ &= \frac{1}{2}\sum_{i=0}^{n-1} x_i(y_{i+1} - y_i) + \frac{1}{2}\sum_{i=1}^{n} x_i(y_i - y_{i-1}) \\ &= \frac{1}{2}\sum_{i=1}^{n} x_i(y_{i+1} - y_{i-1}).\end{aligned}$$

Beispiel 156.

Thermodynamik I, Abschnitt 7.5

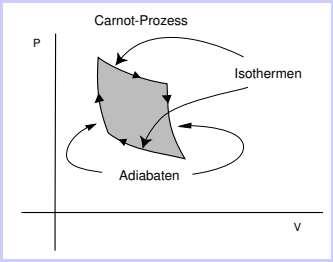
Die Formel (49) spielt eine Rolle in der Thermodynamik. Die Zustände eines idealen Gases kann man sich als Punkte auf einer Fläche im dreidimensionalen (V, p, T)-Raum vorstellen, die durch die Gasgleichung $pV = \nu RT$ gegeben ist, also als der Graph der Funktion $T = \frac{pV}{\nu R}$ (R =universelle Gaskonstante, ν =Molzahl.)

Auf dieser Fläche kann man offenbar zwei der drei Variablen als Koordinaten benutzen, üblicherweise (V, p). Ein thermodynamischer Prozess ist eine Kurve auf der Zustandsfläche, die man also beschreiben kann in der Form $t \mapsto (V(t), p(t))$. Dann ist

$$A = \int_{t_1}^{t_2} p(t)\dot{V}(t)dt$$

die bei der Überführung vom Zustand zur Zeit t_1 in den zur Zeit t_2 vom Prozess geleistete geleistete Arbeit. ($p\Delta V = \nu R\Delta T = \Delta Q$ ist gerade die Wärmemenge, welche die Volumenänderung um ΔV beim Druck p freisetzt.) Bei geschlossenen Kurven, sogenannten *Kreisprozessen*, ist das nach (49) also das Negative der eingeschlossenen Fläche, wenn diese vom Prozess gegen den Uhrzeigersinn umlaufen wird.

Beim berühmten Carnotschen Kreisprozess wird ein krummliniges Viereck im Uhrzeigersinn umlaufen. Die Seiten sind abwechselnd *Isothermen*, die durch Hyperbelgleichungen $pV = const$ (also $T = const$) charakterisiert sind, und sogenannte *Adiabaten*, die durch die Poissongleichung $pV^\kappa = const.$ charakterisiert sind $(1 < \kappa < 2)$. Weil der Umlauf gegen den Uhrzeiger erfolgt, gibt der Flächeninhalt die geleistete Arbeit.



Durch

$$V(x,T) := \frac{x}{(\nu RT)^{\frac{1}{\kappa-1}}}, \qquad p(x,T) := \frac{(\nu RT)^{\frac{\kappa}{\kappa-1}}}{x} \tag{50}$$

wird ein Rechteck $x_1 \le x \le x_2, T_1 \le T \le T_2$ bei geeigneter Wahl von x_i, T_i gerade auf die obige Fläche abgebildet, weil

$$\begin{aligned} p(x,T)V(x,T) &= (\nu RT)^{\frac{\kappa}{\kappa-1}-\frac{1}{\kappa-1}} = \nu RT, \\ p(x,T)V(x,T)^\kappa &= x^{\kappa-1}. \end{aligned}$$

Die (x,T)-Koordinatenlinien entsprechen also gerade den Isothermen und Adiabaten.
Nach der Transformationsformel ist die Fläche dann

$$\begin{aligned} -\int_{\partial F} pdV &= \int_{T_1}^{T_2}\int_{x_1}^{x_2} \left|\frac{\partial(V,p)}{\partial(x,T)}\right| dxdT \\ &= \int_{T_1}^{T_2}\int_{x_1}^{x_2} \left|\det \begin{pmatrix} (\nu RT)^{-\frac{1}{\kappa-1}} & x\,\frac{-1}{\kappa-1}(\nu RT)^{-\frac{1}{\kappa-1}-1}\nu R \\ \frac{-1}{x^2}(\nu RT)^{\frac{\kappa}{\kappa-1}} & \frac{1}{x}\frac{\kappa}{\kappa-1}(\nu RT)^{\frac{1}{\kappa-1}}\nu R \end{pmatrix}\right| dxdT \\ &= \int_{T_1}^{T_2}\int_{x_1}^{x_2} \left|\frac{1}{x}\frac{\kappa}{\kappa-1}\nu R - \frac{1}{x}\frac{1}{\kappa-1}\nu R\right| dxdT \\ &= \int_{T_1}^{T_2}\int_{x_1}^{x_2} \frac{\nu R}{x} dxdT = \nu R(T_2-T_1)\ln\frac{x_2}{x_1}. \end{aligned}$$

Nach (50) ist schließlich $\frac{x_2}{x_1} = \frac{V(x_2,T)}{V(x_1,T)} =: \frac{V_2}{V_1}$. Daher ist die gesamte (mechanische) Arbeit

$$A = \nu R(T_2 - T_1) \ln \frac{V_2}{V_1}.$$

Die auf der T_2-Isotherme vom Prozess (aus einem externen Speicher) absorbierte Wärmemenge ist

$$Q_2 = \int_{V_1}^{V_2} p dV = \int_{V_1}^{V_2} \frac{\nu R T_2}{V} dV = \nu R T_2 \ln \frac{V_2}{V_1},$$

auf der T_1-Isotherme gibt der Prozess die Wärmemenge $\nu R T_1 \ln \frac{V_2}{V_1}$ ab.

0 Anhang: Unendliche Reihen

Unendliche Reihen sind Verallgemeinerungen endlicher Summen auf den Fall unendlich vieler Glieder. Von besonderer Bedeutung sind „Polynome von unendlichem Grad“, die sogenannten *Potenzreihen*

$$\sum_{k=0}^{\infty} a_k (x - x_0)^k$$

und „trigonometrische Polynome von unendlicher Ordnung“, die sogenannten *Fourierreihen*

$$\frac{a_0}{2} + \sum_{k=1}^{\infty} (a_k \cos(k\omega t) + b_k \sin(k\omega t)).$$

Zunächst aber betrachten wir die einfachere Situation unendlicher Reihen, deren Glieder Konstanten sind.

0.1 Reihen mit konstanten Gliedern. Konvergenzkriterien.

- Was sind unendliche Reihen, was sind Partialsummen?
- Wann ist eine unendliche Reihe konvergent?
- Die geometrische Reihe als “allerwichtigste” konvergente Reihe
- Rechenregeln für Reihen

Eine *unendliche Reihe* hat die Form

$$a_0 + a_1 + a_2 + \ldots = \sum_{k=0}^{\infty} a_k,$$

wobei die a_k reelle Zahlen sind. Die endlichen Summen

$$s_n = a_0 + a_1 + a_2 + \ldots + a_n = \sum_{k=0}^{n} a_k$$

heißen die *Partialsummen* der Reihe.

Definition 157 (Reihenkonvergenz). Wenn die Folge $(s_n)_{n\in\mathbb{N}}$ der Partialsummen gegen A konvergiert, nennt man die unendliche Reihe gegen A konvergent und schreibt

$$\sum_{k=0}^{\infty} a_k = A.$$

Man nennt A dann auch *die Summe* der unendlichen Reihe.

Statt $\mathbb{N}$ betrachtet man auch andere Summationsbereiche, z.B. $\sum_{k=1}^{\infty} \frac{1}{k}$.

Beispiel 158 (Die geometrische Reihe). Sei $q \in \mathbb{R}$ und sei $a_k = q^k$. Die zugehörige Reihe

$$\sum_{k=0}^{\infty} q^k = 1 + q + q^2 + q^3 + \ldots$$

heißt die **geometrische Reihe**. Sie ist die wichtigste Reihe überhaupt. Die Partialsummen sind gegeben durch

$$s_n = \sum_{k=0}^{n} q^k = \begin{cases} n+1 & \text{falls } q = 1 \\ \frac{1-q^{n+1}}{1-q} & \text{sonst.} \end{cases}$$

Für $q = 1$ ist die Reihe also divergent. Für $|q| > 1$ ist die Folge $(q^{n+1})_{n\in\mathbb{N}}$ und daher auch die Partialsummenfolge $s_n = \frac{1-q^{n+1}}{1-q}$ divergent. Dasselbe gilt für $q = -1$. Die geometrische Reihe ist also für alle q mit $|q| \geq 1$ divergent. Andrerseits ist $\lim_{n\to\infty} q^{n+1} = 0$ für $|q| < 1$, und deshalb erhalten wir

$$\sum_{k=0}^{\infty} q^k = \frac{1}{1-q} \quad \text{für } |q| < 1.$$

Insbesondere folgt daraus z.B.

$$0,999\ldots = \sum_{k=1}^{\infty} 9 \cdot (\frac{1}{10})^k = \frac{9}{10}\sum_{k=0}^{\infty}(\frac{1}{10})^k = \frac{9}{10}\frac{1}{1-\frac{1}{10}} = 1.$$

□

Beispiel 159 (Dopplereffekt). Eine ruhende Schallquelle sendet Wellen mit der Frequenz ν, der Schwingungsdauer $T = 1/\nu$ und der Ausbreitungsgeschwindigkeit c aus. Die Wellenlänge ist dann $\lambda = cT = c/\nu$.

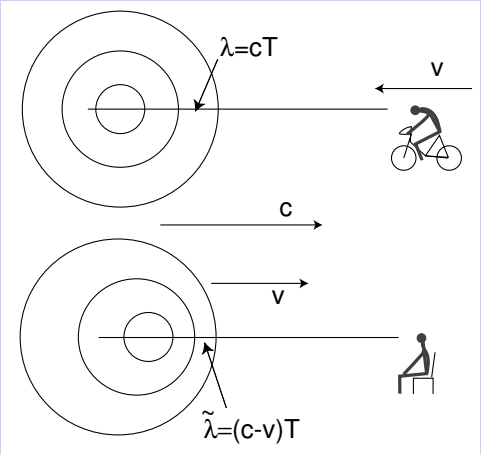
Bewegt sich der Empfänger mit der Geschwindigkeit $v < c$ auf die ruhende Quelle zu, so verkürzt sich die empfangene Frequenz auf

$$\hat{\nu} = \frac{c+v}{\lambda} = \nu\left(1 + \frac{v}{c}\right).$$

Bewegt sich andrereits die Quelle mit der Geschwindigkeit $v < c$ auf den Empfänger zu, so verkürzt sich die Wellenlänge auf $\tilde{\lambda} = (c-v)/\nu = c/\tilde{\nu}$, die empfangene Frequenz ist

$$\tilde{\nu} = \nu\frac{1}{1-v/c} = \nu\left(1 + \frac{v}{c} + (\frac{v}{c})^2 + \ldots\right).$$

In diesem Fall ist die Frequenz also höher als bei bewegtem Empfänger.



Rechenregeln für konvergente unendliche Reihen. Sind $\sum a_k = A$ und $\sum b_k = B$ konvergente unendliche Reihen, so konvergiert $\sum(a_k + b_k)$ gegen $A + B$:

$$\sum(a_k + b_k) = \sum a_k + \sum b_k. \tag{51}$$

Ist weiter $c \in \mathbb{R}$ so folgt

$$\sum(ca_k) = c\sum a_k. \tag{52}$$

Mit Produkten unendlicher Reihen ist es komplizierter. Man muss jedes Glied der einen Reihe mit jedem Glied der anderen Reihe malnehmen, was bei unendlich vielen Gliedern Probleme

macht. Eine einleuchtende Anordnung der Produkte gibt die sogenannte Produktformel von Cauchy:

$$(\sum_{i=0}^{\infty} a_i)(\sum_{j=0}^{\infty} b_j) = \sum_{k=0}^{\infty} c_k, \text{ wobei } c_k := a_0 b_k + a_1 b_{k-1} + \ldots + a_k b_0. \tag{53}$$

Die Gleichung stimmt, wenn alle drei Reihen konvergent sind. Aber aus der Konvergenz der Reihen $\sum_{i=0}^{\infty} a_i$ und $\sum_{j=0}^{\infty} b_j$ folgt im allgemeinen nicht die Konvergenz der Reihe $\sum_{k=0}^{\infty} c_k$. Vergleiche dazu aber die Bemerkungen über *absolut konvergente* Reihen weiter unten.

KONVERGENZKRITERIEN. Wie prüft man, ob eine Reihe konvergiert? Ein einfaches notwendiges Kriterium ist das folgende: Wenn die Reihe $\sum a_k$ konvergiert, wenn also die Partialsummenfolge (s_n) gegen einen Wert A konvergiert, dann konvergiert auch die Folge (s_{n-1}) gegen A, sie hat ja dieselben Glieder, nur in der Numerierung um eins verschoben. Aber dann gilt $0 = \lim s_n - \lim s_{n-1} = \lim(s_n - s_{n-1}) = \lim a_n$. Nur wenn $\lim a_k = 0$ hat die Reihe also eine Chance, konvergent zu sein:

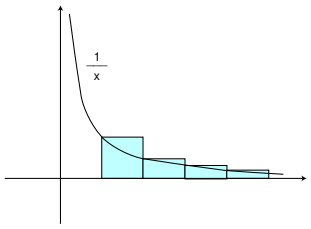
Satz 160 (Notwendiges Konvergenzkriterium). $\sum a_k$ *konvergent* $\implies \lim a_k = 0$.

Dass aber eine Reihe mit dieser Eigenschaft nicht konvergent sein muss, zeigt das sehr berühmte „Gegenstück“ zur geometrischen Reihe:

Beispiel 161 (Divergenz der harmonischen Reihe). Die harmonische Reihe ist die Reihe

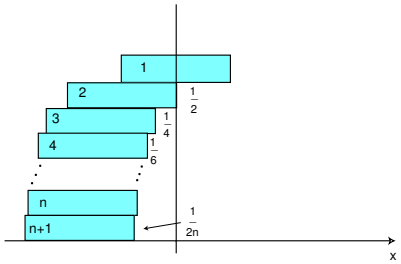
$$\sum_{k=1}^{\infty} \frac{1}{k} = 1 + \frac{1}{2} + \frac{1}{3} + \ldots$$

Wir erinnern daran, dass das uneigentliche Integral $\int_1^{\infty} \frac{1}{x}\,dx$ nicht existiert.
Aus dem Vergleich von Integral und Summe in der Abbildung folgt daher die Divergenz der harmonischen Reihe.



□

Die Divergenz der harmonischen Reihe hat folgende „praktische Anwendung“:



Wir bauen einen Turm aus Ziegelsteinen der Länge 1 „von oben nach unten“, indem wir den bereits gebauten Turm so auf den nächsten Stein setzen, dass sein Schwerpunkt gerade über der Kante des neuen untersten Steins liegt, der Turm also gerade eben nicht umkippt. Ist S_{n-1} die x-Koordinate des Schwerpunktes der ersten $n-1$ Steine, so liegt der Schwerpunkt des n-ten Steins also bei $S_{n-1}+\frac{1}{2}$, und der Schwerpunkt des erweiterten Turms bei

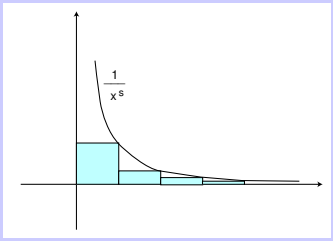
$$S_n = \frac{1}{n}((n-1)S_{n-1} + (S_{n-1} + \frac{1}{2})) = S_{n-1} + \frac{1}{2n} = \frac{1}{2}\sum_{k=1}^{n} \frac{1}{k}.$$

Wegen der Divergenz der harmonischen Reihe kann man also den überhang beliebig groß machen. Zum Beispiel ist $\frac{1}{2}\sum_{k=1}^{4} \frac{1}{k} = 1.04$, mit fünf Steinen kann man einen überhang von mehr als einem Stein und mit 32 Steinen einen von mehr als zwei Steinen realisieren.

Beispiel 162 (Dipolketten).
Werkstoffe I, Abschnitt 2.5

Für $s > 1$ existiert das Integral

$$\int_1^\infty \frac{dx}{x^s}.$$



Am nebenstehenden Bild sieht man deshalb, dass

$$\sum_{k=1}^{\infty} \frac{1}{k^s}$$

für $s > 1$ konvergiert.

Die für $s > 1$ definierte Funktion

$$\zeta(s) := \sum_{k=1}^{\infty} 1/k^s$$

heißt die *Riemannsche Zetafunktion.*
In der Theorie der Dielektrika untersucht man eindimensionale Dipolketten. Für jedes $k \in \mathbb{Z}\backslash\{0\}$ sitzt an der Stelle $\vec{x} = k\vec{a}$ ein Dipol mit Dipolmoment $\vec{p} = p\frac{\vec{a}}{|\vec{a}|}$, der im Nullpunkt ein elektrisches Feld der Stärke

$$\vec{E}_k = \frac{1}{4\pi\epsilon}\left(3\frac{\vec{p}\cdot\vec{x}}{|\vec{x}|^5}\,\vec{x} - \frac{1}{|\vec{x}|^3}\,\vec{p}\right) = \frac{1}{4\pi\epsilon}\left(3\frac{p}{|k|^3|\vec{a}|^3}\,\frac{\vec{a}}{|\vec{a}|} - \frac{1}{|k|^3|\vec{a}|^3}\,p\frac{\vec{a}}{|\vec{a}|}\right) = \frac{p}{2\pi\epsilon|\vec{a}|^3}\,\frac{\vec{a}}{|\vec{a}|}\,\frac{1}{|k|^3}$$

erzeugt. Das gesamte Feld ist dann

$$\vec{E} = \frac{p}{2\pi\epsilon|\vec{a}|^3}\,\frac{\vec{a}}{|\vec{a}|}\sum_{k\in\mathbb{Z},\,k\neq 0}\frac{1}{|k|^3} = \frac{p}{\pi\epsilon|\vec{a}|^3}\sum_{k=1}^{\infty}\frac{1}{k^3} = \frac{\zeta(3)p}{\pi\epsilon|\vec{a}|^3}\,\frac{\vec{a}}{|\vec{a}|} = \frac{1.202p}{\pi\epsilon|\vec{a}|^3}\,\frac{\vec{a}}{|\vec{a}|}.$$

Die Überlegungen der vorangehenden Beispiele kann man ausbauen zu einem Satz, der eine Beziehung zwischen uneigentlichen Integralen und unendlichen Reihen herstellt:

Satz 163 (Reihen-Integral-Kriterium). *Sei $f\colon [m,\infty[\to\mathbb{R}$, $m \in \mathbb{N}$ eine monoton fallende Funktion mit $f(x) \geq 0$ für alle $x \in [m,\infty[$. Dann existiert das uneigentliche Integral $\int_m^\infty f(x)\,dx$ genau dann, wenn die Reihe $\sum_{k=m}^\infty f(k)$ konvergiert.*

0.2 Weitere Konvergenzkriterien

- Wir lernen die wichtigsten Konvergenzkriterien für unendliche Reihen.

Zurück zur Frage: **Wie prüft man, ob eine unendliche Reihe konvergent ist?** Wir wollen dafür jetzt zwei hinreichende Kriterien geben. „Hinreichend“ bedeutet: Wenn dies und das erfüllt ist, *dann* ist die Reihe konvergent. Aber sie kann auch konvergent sein, wenn diese Voraussetzungen nicht vorliegen.

Zunächst vergleichen wir zwei Reihen $\sum a_k$ und $\sum b_k$. Von der zweiten Reihe sei schon bekannt, dass sie konvergiert, und wir wollen annehmen, dass

$$0 \le a_k \le b_k \quad \text{für alle } k \tag{54}$$

gilt. Weil die Summanden ≥ 0 sind, sind die Folgen der Partialsummen monoton wachsend und

$$\sum_{k=0}^{n} a_k \le \sum_{k=0}^{n} b_k \le \sum_{k=0}^{\infty} b_k =: M.$$

Die Partialsummenfolge der Reihe $\sum a_k$ ist also beschränkt, und daher nach dem Monotoniekriterium konvergent.

Dieses Argument kann man nur auf Reihen ohne negative Glieder anwenden. Aber mit einem Trick läßt es sich verallgemeinern: Hat man statt (54) die allgemeinere Bedingung

$$|a_k| \le b_k \quad \text{für alle } k,$$

so kann man die Reihe $\sum c_k = \sum(a_k + |a_k|)$ betrachten und mit der Reihe $2\sum b_k = \sum 2b_k$ vergleichen. Die zweite Reihe ist natürlich wie die Reihe $\sum b_k$ konvergent, und weil

$$0 \le a_k + |a_k| \le 2b_k,$$

ist die Reihe $\sum c_k$ nach unserer gerade gemachten Überlegung konvergent. Natürlich konvergiert danach auch die Reihe $\sum |a_k|$ und damit schließlich die Reihe

$$\sum a_k = \sum((a_k + |a_k|) - |a_k|) = \sum(a_k + |a_k|) - \sum |a_k|.$$

Wir fassen das zusammen:

Satz 164 (Vergleichskriterium=Majorantenkriterium). *Die Reihe $\sum b_k$ sei konvergent, und es gelte*

$$|a_k| \le b_k \text{ für alle } k.$$

(Hier genügt: Für alle k von einem gewissen $k_0 \in \mathbb{N}$ an.) Dann ist Reihe

$$\sum_{k=0}^{\infty} a_k$$

konvergent.

Beispiel 165. Wir vergleichen die Reihen

$$\sum_{k=1}^{\infty} \frac{1}{k^2} \quad \text{und} \quad \sum_{k=2}^{\infty}\left(\frac{1}{k-1} - \frac{1}{k}\right)$$

Für $k > 1$ ist

$$\left|\frac{1}{k^2}\right| = \frac{1}{k^2} < \frac{1}{k(k-1)} = \frac{1}{k-1} - \frac{1}{k}.$$

Die Konvergenz der zweiten Reihe kann man aber leicht einsehen, weil man ihre Partialsummen explizit bestimmen kann:

$$\sum_{k=2}^{n}(\frac{1}{k-1} - \frac{1}{k}) = (\frac{1}{1} - \frac{1}{2}) + (\frac{1}{2} - \frac{1}{3}) + \ldots (\frac{1}{n-1} - \frac{1}{n}) = 1 - \frac{1}{n} \to 1.$$

Damit haben wir noch einmal die Konvergenz von $\sum_{k=1}^{\infty} \frac{1}{k^2}$ bewiesen. Schwieriger ist, den Grenzwert der Reihe zu bestimmen. Er ist $\zeta(2) = \sum \frac{1}{k^2} = \pi^2/6$. □

Der Nachteil beim Vergleichskriterium ist, dass man schon eine konvergente Majorante $\sum b_k$ haben muss. Das folgende Kriterium benutzt nur die zu untersuchende Reihe und ist deshalb meistens die **erste Wahl**, wenn man eine Reihe auf Konvergenz untersuchen will. Erst wenn es keine Auskunft über die Konvergenz gibt, versucht man andere Kriterien.

Satz 166 (Quotientenkriterium). *Ist die Reihe $\sum_{k=0}^{\infty} a_k$ konvergent? Berechne dazu*

$$\lim_{k\to\infty} \left|\frac{a_{k+1}}{a_k}\right|.$$

- *Ist der Grenzwert < 1, so ist die Reihe konvergent.*
- *Ist der Grenzwert > 1, so ist die Reihe divergent.*
- *Ist der Grenzwert $= 1$ oder existiert er gar nicht, so gibt dieses Kriterium keine Auskunft.*

Das ist so zu verstehen, dass die a_k alle (oder wenigstens alle von einer Stelle k_0 an) ungleich null sein müssen und der Limes auch existiert.

Beweis. Zum Beweis wählen wir ein q echt zwischen $\lim|\frac{a_{k+1}}{a_k}|$ und 1:

$$\lim|\frac{a_{k+1}}{a_k}| < q < 1.$$

Weil $\lim|\frac{a_{k+1}}{a_k}| < q$ ist dann (nach der Definition der Konvergenz)

$$\left|\frac{a_{k+1}}{a_k}\right| < q \tag{55}$$

für alle k von einem gewissen k_0 an. Daraus folgt für $k \geq k_0$:

$$|a_k| < q|a_{k-1}| < q^2|a_{k-2}| < \ldots < q^{k-k_0}|a_{k_0}| = q^k \frac{|a_{k_0}|}{q^{k_0}}.$$

Aber nach dem Resultat über die geometrische Reihe ist

$$\sum_{k=0}^{\infty} q^k \frac{|a_{k_0}|}{q^{k_0}} = \frac{|a_{k_0}|}{q^{k_0}} \sum_{k=0}^{\infty} q^k = \frac{|a_{k_0}|}{q^{k_0}} \frac{1}{1-q}.$$

Die erste Behauptung folgt daher aus dem Vergleichssatz angewendet auf die Reihe $\sum b_k$ mit $b_k = q^k \frac{|a_{k_0}|}{q^{k_0}}$.

Die zweite Behauptung ergibt sich wie folgt: Ist der Limes > 1, so bedeutet das für große k, dass $|a_{k+1}| > |a_k|$. Also konvergieren die a_k sicher nicht gegen null. □

Bemerkung: Beim Beweis haben wir nicht wirklich die Existenz des Limes benutzt, sondern nur gebraucht, dass es eine Zahl q **kleiner als** 1 gibt, so dass (55) für alle $k \geq k_0$ gilt. Das Quotientenkriterium gilt also auch mit dieser schwächeren Voraussetzung:

Gibt es $q < 1$ und k_0, so dass $|\frac{a_{k+1}}{a_k}| \leq q$ für alle $k \geq k_0$, so ist die Reihe $\sum a_k$ konvergent.

Beispiel 167. Die Reihe

$$\sum_{k=0}^{\infty} \frac{k}{2^k}$$

ist konvergent nach dem Quotientenkriterium: Es ist

$$\left|\frac{a_{k+1}}{a_k}\right| = \left|\frac{\frac{k+1}{2^{k+1}}}{\frac{k}{2^k}}\right| = \frac{2^k(k+1)}{2^{k+1}k} = \frac{1}{2}(1+\frac{1}{k}) \to \frac{1}{2} < 1.$$

□

Beispiel 168 (Wo das Quotientenkriterium versagt). Für die beiden Reihen

$$\sum_{k=1}^{\infty} \frac{1}{k} \text{ und } \sum_{k=1}^{\infty} \frac{1}{k^2}$$

ist der Limes des Quotienten aufeinanderfolgender Glieder beide Male $= 1$. Das Quotientenkriterium macht in diesen Fällen keine Aussage. Aber wie wir gesehen haben, ist die erste Reihe divergent, die zweite (zum Beispiel nach dem Vergleichskriterium) konvergent. □

Zum Abschluss betrachten wir noch eine Reihe, für welche die vorstehenden Kriterien ebenfalls versagen, nämlich die sogenannte **alternierende harmonische Reihe**:

$$\sum_{k=1}^{\infty} (-1)^{k+1} \frac{1}{k} = 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} \pm \dots$$

Diese Reihe ist konvergent. Die Tatsache der Konvergenz kann man relativ leicht einsehen. Bei jedem Schritt wird abwechselnd die Partialsumme erhöht oder erniedrigt, wobei die Differenz monoton kleiner wird. Die ungeraden Partialsummen $1, 1 - \frac{1}{2} + \frac{1}{3}, \dots$ sind daher eine monoton fallende, die geraden $1 - \frac{1}{2}, 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4}, \dots$ eine monoton wachsende Folge. Beide sind beschränkt und daher konvergent. Und weil sich die benachbarten s_n und s_{n+1} nur um $\pm\frac{1}{n+1}$ unterscheiden, konvergieren sie gegen denselben Grenzwert. Also konvergiert die gesamte Folge der Partialsummen. Diese Argumentation liefert ein weiteres Konvergenzkriterium:

Satz 169 (Leibniz-Kriterium). *Eine alternierende Reihe deren Glieder absolut genommen eine monoton gegen null konvergierende Folgen bilden, ist konvergent.*

Der Grenzwert der obigen Reihe ist wiederum schwierig zu bestimmen. Er ist $\ln 2$.

An dieser Reihe kann man noch ein wichtiges Phänomen deutlich machen: Bei unendlichen Reihen ist die Reihenfolge der Glieder im allgemeinen nicht mehr gleichgültig, es gilt **kein „Kommutativgesetz“**. Das liegt daran, dass eine Umordnung der Glieder die Partialsummenfolge völlig durcheinander bringt. Bei der alternierenden harmonischen Reihe sind nämlich die Reihen der positiven bzw. negativen Glieder

$$1 + \frac{1}{3} + \frac{1}{5} + \dots \quad \frac{1}{2} + \frac{1}{4} + \frac{1}{6} + \dots$$

beide divergent. (Die zweite ist einfach die halbe harmonischen Reihe und deshalb divergent; die Glieder und darum die Partialsummen der ersten Reihe sind aber offensichtlich größer als die der zweiten, weshalb auch die erste Reihe divergent ist.) Darum kann man von der alternierenden Reihe zunächst so viele positive Glieder (der Reihe nach) addieren, bis man z.B. über 27 ist. Dann addiert man negative Glieder solange, bis man unter 27 kommt. Dann addiert man weiter positive Glieder, bis man wieder über 27 ist. Auf diese Weise erwischt man schließlich alle Glieder der Reihe und hat sie so umgeordnet, dass die neue Reihe nun gegen 27 konvergiert.

Diese Zauberei funktionierte, weil man Glieder verschiedenen Vorzeichens hat, die für sich genommen divergente Reihen bilden. Wenn nicht nur die Reihe $\sum a_k$, sondern auch die Reihe $\sum |a_k|$ konvergiert, kann das wohl nicht mehr passieren.

Definition 170 (Absolute Konvergenz). Wenn die Reihe $\sum |a_k|$ konvergiert, nennt man die Reihe $\sum a_k$ *absolut konvergent*. (Nach dem Vergleichskriterium ist sie dann auch im gewöhnlichen Sinne konvergent.)

Vorteile absolut konvergenter Reihen. Vergleichskriterium und Quotientenkriterium liefern *absolut* konvergente Reihen. Diese haben den Vorteil, dass man mit ihnen „ziemlich“ so wie mit endlichen Summen rechnen darf. Genauer:

- Absolut konvergente Reihen darf man beliebig umordnen, die Konvergenz und der Grenzwert ändern sich nicht.
- Das Cauchyprodukt zweier absolut konvergenter Reihen ist wieder (absolut) konvergent, vgl. (53).[3]

Komplexe Reihen. Für Reihen mit komplexen Gliedern definiert man die Konvergenz genauso wie im reellen Fall. Alle die obigen Regeln – mit Ausnahme des Leibniz-Kriteriums – gelten auch im Komplexen. Zum Beispiel gilt

$$\boxed{\sum_{k=0}^{\infty} z^k = \frac{1}{1-z} \quad \text{für alle } z \in \mathbb{C} \text{ mit } |z| < 1.}$$

[3] Es genügt, wenn nur eine der beiden Reihen absolut konvergent und die andere konvergent ist.

0.3 Funktionenreihen

- Reihen, deren Glieder nicht konstant, sondern Funktionen sind, nennt man Funktionenreihen.
- Wir lernen die beiden wichtigsten Typen solcher Reihen kennen, die Potenzreihen und die Fourierreihen.

POTENZ- UND TAYLORREIHEN. Die wichtigsten Funktionen erhält man durch eine Verallgemeinerung der Polynome auf unendliche Summen:

Definition 171. Eine *Potenzreihe* ist eine unendliche Reihe der Form

$$\sum_{k=0}^{\infty} a_k (z - z_0)^k. \tag{56}$$

Wir wollen diese Reihen gleich im Komplexen betrachten. Das ist nicht komplizierter, aber dei Begriffe „Konvergenzradius“ und „Konvergenzkreis“ werden viel anschaulicher. Wir nehmen deshalb an, dass die Koeffizienten a_k und der Entwicklungspunkt z_0 komplexe Zahlen sind. Weiter ist z eine komplexe Variable.

Auf der Menge $D \subset \mathbb{C}$ aller z, für die die Reihe konvergiert, liefert $f(z) := \sum_{k=0}^{\infty} a_k (z - z_0)^k$ also eine Funktion

$$f : D \to \mathbb{C}.$$

Wir untersuchen nun die Frage, für welche Werte von z die Reihe konvergiert. Es ist einleuchtend, dass sie konvergiert, wenn $|z - z_0|$ „klein“ ist, und divergiert, wenn diese Zahl „groß“ ist. Dabei hängt die Bedeutung von „klein“ und „groß“ vermutlich von den a_k ab. Wir versuchen das Quotientenkriterium. Beachten Sie aber, dass jetzt $a_k(z - z_0)^k$ die Rolle übernimmt, die im Satz 166 der Term a_k spielte. Wir müssen also

$$\left| \frac{a_{k+1}(z - z_0)^{k+1}}{a_k (z - z_0)^k} \right| = \left| \frac{a_{k+1}}{a_k} \right| \cdot |z - z_0|$$

betrachten.

Wir nehmen an, dass der Grenzwert $A := \lim_{k\to\infty} \left| \frac{a_{k+1}}{a_k} \right|$ existiert.[4] Dann ist die Reihe für $A|z - z_0| < 1$ nach dem Quotientenkriterium (absolut) konvergent, für $A|z - z_0| > 1$ divergent. Bei $A = 0$ ist die Reihe also für alle $z \in \mathbb{C}$ konvergent, bei $0 < A < \infty$ ist die Reihe für alle z mit

$$|z - z_0| < R := \frac{1}{A} \quad \text{konvergent}, \tag{57}$$

$$|z - z_0| > R := \frac{1}{A} \quad \text{divergent}. \tag{58}$$

Ist $A = \infty$, so ist die Reihe nur für $z = z_0$ konvergent.

Es gilt:

[4] Das muss nicht so sein. Dann kommen wir mit diesem Kriterium nicht weiter.

Satz 172 (Konvergenz von Potenzreihen). *Die Potenzreihe*

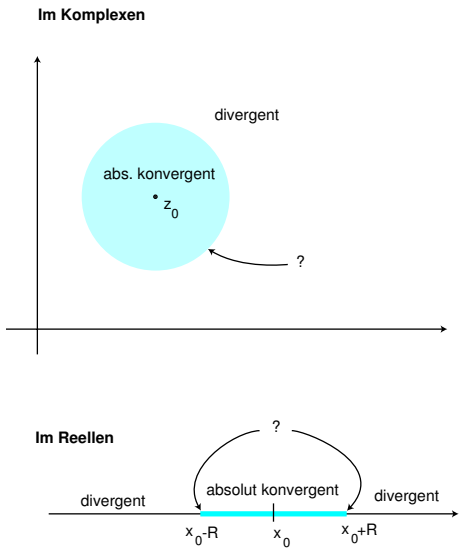
$$\sum_{k=0}^{\infty} a_k (z - z_0)^k$$

ist für alle z in einem offenen Kreis (dem Konvergenzkreis*) vom Radius R um den Mittelpunkt z_0 (absolut) konvergent und für alle z außerhalb des abgeschlossenen Kreises divergent. Das Konvergenzverhalten auf dem Rand des Kreises muss man bei jeder speziellen Reihe in jedem Punkt einzeln untersuchen.*
ür den sogenannten Konvergenzradius R *gilt*

$$R = \lim_{k \to \infty} \frac{1}{\left| \frac{a_{k+1}}{a_k} \right|},$$

falls der Limes exitiert. Dabei ist auch der Wert $R = \infty$ zugelassen, ein Kreis mit unendlichem Radius ist die ganze Ebene $\mathbb{C}$.

Ein „offener Kreis“ ist ein Kreis ohne die berandende Kreislinie, beim „abgeschlossenen Kreis“ ist diese einbegriffen. Beachten Sie, dass der Quotient der a_k hier das Reziproke von der Formel im Quotientenkriterium ist. Beachten Sie weiter, dass wir die obige Aussage nur unter der Annahme bewiesen haben, dass der Grenzwert existiert. Sie bleibt aber richtig, auch wenn der Grenzwert nicht existiert, nur hat man dann keine so einfache Formel mehr für R.
Ein Wort zur Sprache: Sagen Sie nicht, die Reihe sei innerhalb des Konvergenzradius konvergent. Der Konvergenzradius ist eine Zahl, z.B. 7. Was soll es bedeuten, dass die Reihe innerhalb von 7 konvergiert?



Beispiel 173. Die geometrische Reihe $\sum z^k$ ist eine Potenzreihe mit Konvergenzradius $R = 1$. Es gilt

$$f(z) = \sum_{k=0}^{\infty} z^k = \frac{1}{1 - z} \quad \text{für } |z| < 1.$$

□

Beispiel 174. Die Potenzreihe

$$\sum_{k=0}^{\infty} \frac{5^k}{k+1} z^{2k} \tag{59}$$

enthält nur gerade Potenzen, es ist $a_k = 0$ für alle ungeraden k. Darum sind die Quotienten aufeinanderfolgender Koeffizienten abwechselnd null oder nicht definiert. Der obige Grenzwert existiert also nicht. Aber man kann natürlich das Quotientenkriterium direkt versuchen:

$$\lim_{k \to \infty} \left| \frac{\frac{5^{k+1}}{k+2} z^{2k+2}}{\frac{5^k}{k+1} z^{2k}} \right| = \lim_{k \to \infty} \left| \frac{5(k+1)z^2}{k+2} \right| = 5|z|^2.$$

Also hat man Konvergenz für $|z| < \frac{1}{\sqrt{5}}$ und Divergenz für $|z| > \frac{1}{\sqrt{5}}$. Der Konvergenzradius ist $R = \frac{1}{\sqrt{5}}$. □

Im folgenden Satz beschränken wir uns wieder auf reelle Reihen.

Satz 175 (Differentiation von Potenzreihen). *Die reelle Potenzreihe*

$$\sum_{k=0}^{\infty} a_k (x - x_0)^k$$

habe Konvergenzradius $R > 0$. Sie konvergiert dann also auf einem Intervall $]x_0 - R, x_0 + R[$ und definiert dort eine Funktion $f(x)$. Diese Funktion ist differenzierbar und es gilt

$$f'(x) = \sum_{k=1}^{\infty} k a_k (x - x_0)^{k-1} = \sum_{k=0}^{\infty} (k+1) a_{k+1} (x - x_0)^k.$$

Der Konvergenzradius der abgeleiteten Reihe ist wieder R.
Kurz: *Potenzreihen darf man gliedweise differenzieren. Der Konvergenzradius ändert sich dabei nicht. Ebenso darf man Potenzreihen gliedweise integrieren.*

Beispiel 176. Die Reihe $\sum_{k=0}^{\infty} \frac{z^{k+1}}{k+1}$ hat Konvergenzradius1. Sie ist

- für $z = 1$ divergent (harmonische Reihe)
- für $z = -1$ konvergent (alternierende harmonische Reihe).

Für reelles $x \in [-1, 1[$ gilt

$$\sum_{k=0}^{\infty} \frac{x^{k+1}}{k+1} = -\ln(1 - x). \tag{60}$$

Die Ableitungen der beiden Seiten sind nämlich $\sum_{k=0}^{\infty} x^k$ bzw. $\frac{1}{1-x}$, also gleich. Deshalb sind die beiden Seiten von (60) nach dem Konstanzkriterium gleich bis auf eine additive Konstante c. Einsetzen von $x = 0$ liefert $c = 0$. Dieses liefert die Gleichheit (60) nur für $|x| < 1$. Sie ist aber auch für $x = -1$ wahr und liefert, wenn man mit (-1) multiplziert, den Wert der alternierenden harmonischen Reihe:

$$\sum_{k=0}^{\infty} \frac{(-1)^k}{k+1} = 1 - \frac{1}{2} + \frac{1}{3} - + \ldots = \ln(2).$$

□

Bei der Untersuchung der trigonometrischen Funktionen und der Exponentialfunktion in Analysis I haben wir die Taylorpolynome dieser Funktionen ausgerechnet und diese Funktionen dadurch approximiert. Die Vermutung liegt nahe, dass man sie durch unendliche Taylorpolynome, d.h. durch die bei $n \to \infty$ entstehenden Potenzreihen exakt darstellen kann.

Beispiel 177 (Exponentialreihe). Die Potenzreihe

$$\sum_{k=0}^{\infty} \frac{x^k}{k!}$$

hat den Konvergenzradius $R = \infty$. (Nachrechnen!) Also definiert sie eine differezierbare Funktion $y : \mathbb{R} \to \mathbb{R}$, für die gilt

$$y(0) = 1 + \frac{0^1}{1!} + \frac{0^2}{2!} + \ldots = 1$$

und

$$y'(x) = \sum_{k=0}^{\infty} k\,\frac{x^{k-1}}{k!} = \sum_{k=1}^{\infty} \frac{x^{k-1}}{(k-1)!} = \sum_{k=0}^{\infty} \frac{x^k}{k!} = y(x).$$

Damit ist der noch offene Existenzbeweis für die Exponentialfunktion geführt:

$$\boxed{\exp x = \sum_{k=0}^{\infty} \frac{x^k}{k!}.}$$

□

Ebenso schließt man für die trigonometrischen Funktionen und erhält die Reihendarstellungen:

Beispiel 178 (Sinus- und Cosinusreihe).

$$\sin x = \sum_{m=0}^{\infty} \frac{(-1)^m}{(2m+1)!}\,x^{2m+1}$$
$$\cos x = \sum_{m=0}^{\infty} \frac{(-1)^m}{(2m)!}\,x^{2m}.$$

□

Die eben betrachteten Reihen, die Exponentialreihe und die trigonometrischen Reihen, sind Beispiele für sogenannte *Taylorreihen*:

Ist die Funktion $f(x)$ auf einem Intervall um x_0 unendlich oft differenzierbar, so kann man die Potenzreihe

$$\sum_{k=0}^{\infty} \frac{f^{(k)}(x_0)}{k!}\,(x - x_0)^k$$

bilden, die Taylorreihe zu $f(x)$ im Punkt x_0.

Ist $f(x)$ durch eine Potenzreihe um x_0 gegeben, so ist diese auch gerade die Taylorreihe im Punkt x_0.

Aber im allgemeinen ist es nicht klar, ob das Restglied für $n \to \infty$ wirklich gegen 0 geht, d.h. ob die Taylorreihe von f wirklich gegen f konvergiert. Und es gibt Funktionen, bei denen das falsch ist.

FOURIERREIHEN. Wir betrachten eine Funktion $f : \mathbb{R} \to \mathbb{R}$ mit Periode $T = \frac{2\pi}{\omega}$ mit Fourierkoeffizienten a_k und b_k. Dann heißt

$$\frac{a_0}{2} + \sum_{k=1}^{\infty} (a_k \cos(k\omega t) + b_k \sin(k\omega t))$$

die *Fourierreihe* von $f(t)$.

Satz 179 (Konvergenz von Fourierreihen[5]). *Die Funktion $f : \mathbb{R} \to \mathbb{R}$ sei $T = \frac{2\pi}{\omega}$-periodisch und stückweise monoton. Dann gilt:*

- *An allen Stetigkeitsstellen t von f konvergiert die Fourierreihe gegen $f(t)$.*
- *An den Unstetigkeitsstellen existieren wegen der stückweisen Monotonie der links- und rechtsseitige Grenzwert $f(t-)$ und $f(t+)$ von f, und die Fourierreihe konvergiert gegen das daraus gebildete arithmetische Mittel.*

Man hat also für alle t:

$$\frac{f(t-) + f(t+)}{2} = \frac{a_0}{2} + \sum_{k=1}^{\infty} (a_k \cos k\omega t + b_k \sin k\omega t).$$

Beispiel 180. Die 2-periodische Funktion

$$f(t) = \begin{cases} 1 - t & \text{für } 0 < t < 2 \\ 0 & \text{für } t = 0 \end{cases}$$

hat nach Definition an den Sprungstellen gerade den Mittelwert 0 als Funktionswert. Daher wird sie überall durch ihre Fourierreihe dargestellt:

$$f(t) = \sum_{k=1}^{\infty} \frac{2}{\pi k} \sin k\pi t.$$

□

Beispiel 181. Ähnlich erhält man für die 2π-periodische Funktion mit

$$g(t) := \begin{cases} -1 & \text{für } -\pi < t < 0 \\ 1 & \text{für } 0 < t < \pi. \\ 0 & \text{für } t = 0, \pi, \end{cases}$$

dass

$$g(t) = \frac{4}{\pi} \sum_{k=0}^{\infty} \frac{1}{2k+1} \sin(2k+1)t.$$

□

Aus der letzten Behauptung im Satz über die Approximation im quadratischen Mittel folgt

[5]Einen Beweis für den Satz in dieser bequemen, aber nicht sehr verbreiteten Version findet man etwa in Mangoldt-Knopp, Einführung in die Höhere Mathematik, Band III.

Satz 182 (Parsevalsche Gleichung). *Die Funktion $f : \mathbb{R} \to \mathbb{R}$ sei $T = \frac{2\pi}{\omega}$-periodisch und stückweise monoton mit Fourierkoeffizienten a_k und b_k. Dann gilt:*

$$\frac{2}{T}\int_0^T f(t)^2 dt = \frac{a_0^2}{2} + \sum_{k=1}^{\infty}(a_k^2 + b_k^2).$$

Beispiel 183. Wir betrachten die beiden Funktionen aus den vorangehenden Beispielen 180 und 181. Wir erhalten

$$\sum_{k=1}^{\infty} \frac{4}{\pi^2}\,\frac{1}{k^2} = \frac{2}{2}\int_0^2 (1-t)^2 dt = -\left.\frac{(1-t)^3}{3}\right|_0^2 = \frac{2}{3}.$$

Also ist

$$\boxed{\sum_{k=1}^{\infty} \frac{1}{k^2} = \frac{\pi^2}{6}.}$$

Die Konvergenz der Reihe $\sum \frac{1}{k^2}$ haben wir früher mit dem Reihen-Integral-Kriterium bzw. mit dem Majorantenkriterium nachgewiesen. Jetzt haben wir sozusagen zufällig auch den Reihenwert bestimmt.

Ebenso ergibt sich

$$\frac{16}{\pi^2}\sum_{k=0}^{\infty}\left(\frac{1}{2k+1}\right)^2 = \frac{2}{2\pi}\int_{-\pi}^{+\pi} g(t)^2 dt = 2,$$

und damit

$$\boxed{\sum_{k=0}^{\infty}\left(\frac{1}{2k+1}\right)^2 = \frac{\pi^2}{8}.}$$

□

Wie Potenzreihen, so darf man auch Fourierreihen gliedweise integrieren, aber man darf sie i.a. nicht gliedweise differenzieren: Nicht nur die Sprungstellen machen Probleme, wie man denken könnte. Nach dem Beispiel 180 ist

$$1 - t = 2\sum_{k=1}^{\infty}\frac{\sin k\pi t}{k\pi} \quad \text{für } 0 < t < 2.$$

Durch gliedweises Differenzieren der Reihe bekommt man die Reihe $2\sum_{k=1}^{\infty}\cos k\pi t$, und an der Stelle $t = 1$, an der f keinen Sprung hat, ist die Reihe $2\sum_{k=1}^{\infty}\cos k\pi = 2\sum_{k=1}^{\infty}(-1)^k$ gar nicht konvergent. Dasselbe gilt zum Beispiel für alle irrationalen Werte von t.

Konvergente Fourierreihen darf man im allgemeinen NICHT gliedweise differenzieren.