

Für Extremwertbetrachtungen bei implizit definierten Funktionen braucht man, wie wir im nächsten Semester sehen werden, die zweite Ableitung vor allem in den Punkten, in denen die erste verschwindet; dort vereinfacht sich die Formel zu

$$f''(x) = -\frac{F_{xx}(x, f(x))}{F_y(x, f(x))} \quad \text{falls } f'(x) = 0.$$

Genau wie Funktionen einer Veränderlichen können auch Funktionen mehrerer Veränderlicher implizit definiert sein; die entsprechende Verallgemeinerung des Satzes über implizite Funktionen folgt fast vollständig aus der koordinatenweisen Anwendung des obigen Satzes, lediglich für die Stetigkeit der Funktion muß man das entsprechende Argument aus dem gerade beendeten Beweis noch einmal anwenden. Die Aussage ist

**Satz:**  $D \subseteq \mathbb{R}^n$  sei eine offene Menge,  $F \in C^1(D, \mathbb{R})$  eine stetig differenzierbare Funktion auf  $D$ , und  $\mathbf{a} = (a_1, \dots, a_n) \in D$  sei ein Punkt mit  $f(\mathbf{a}) = 0$ . Falls  $F_{x_n}(\mathbf{a}) \neq 0$  ist, gibt es eine offene Umgebung  $U$  von  $(a_1, \dots, a_{n-1})$  in  $\mathbb{R}^{n-1}$  und eine Funktion  $f \in C^1(U, \mathbb{R})$ , so daß für alle Punkte  $\mathbf{y} = (y_1, \dots, y_{n-1}) \in U$  gilt:

$$F(\mathbf{y}, f(\mathbf{y})) = 0.$$

Für alle  $i \leq n - 1$  ist

$$f_{x_i}(\mathbf{y}) = -\frac{F_{x_i}(\mathbf{y}, f(\mathbf{y}))}{F_{x_n}(\mathbf{y}, f(\mathbf{y}))}.$$

■

### §2: Vektorfelder

Eine Funktion  $f: D \rightarrow \mathbb{R}^m$  auf einer Teilmenge  $D \subseteq \mathbb{R}^n$  läßt sich für  $n, m \geq 2$  nicht mehr durch einen Graphen veranschaulichen, denn der Graph

$$\Gamma_f = \{(\mathbf{x}, \mathbf{y}) \in \mathbb{R}^n \times \mathbb{R}^m \mid \mathbf{y} = f(\mathbf{x})\}$$

liegt dann im mindestens vierdimensionalen  $\mathbb{R}^{n+m}$ .

Im für uns wichtigsten Fall  $n = m = 2$  oder  $3$  gibt es aber eine andere Möglichkeit der Veranschaulichung: Für  $\mathbf{x} \in D$  läßt sich  $f(\mathbf{x})$  als Vektor auffassen, und  $f$  kann veranschaulicht werden, indem man für hinreichend viele Punkte  $\mathbf{x} \in D$  diesen Vektor (oder ein geeignetes konstantes Vielfaches davon) im Punkt  $\mathbf{x}$  anträgt; siehe dazu die Abbildungen 38 und 39.

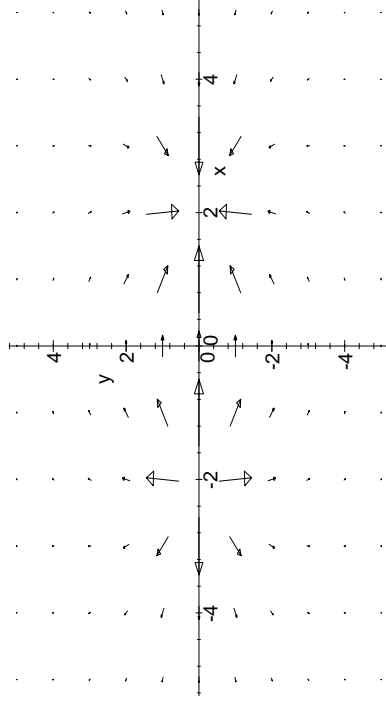


Abb. 38: Das elektrische Feld zweier entgegengesetzt gleicher Punktladungen

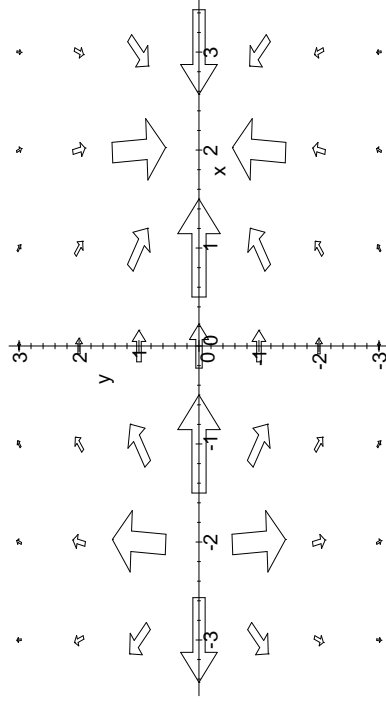


Abb. 39: Dasselbe Feld in einem kleineren Bereich mit „flächigeren“ Vektoren

Wie Abbildung 38 zeigt, kann es dabei bei betragsmäßig stark variierenden Funktionswerten zu Problemen kommen: Wie man die Konstante auch wählt, mit der alle Vektoren multipliziert werden, hat man immer entweder einige extrem lange Vektoren oder aber Vektoren, die so kurz sind, daß man nichts mehr erkennen kann: Ab einer gewissen Entfernung vom Zentrum sind die Richtungen des Felds praktisch nicht mehr wahrnehmbar. Eine leichte Verbesserung ergibt sich, wenn man wie in Abbildung 39 die Vektoren „flächig“ zeichnet, wobei die Länge des Vektors proportional zur *Fläche* des eingezeichneten Pfeils ist, aber auch das hilft nicht sehr viel.

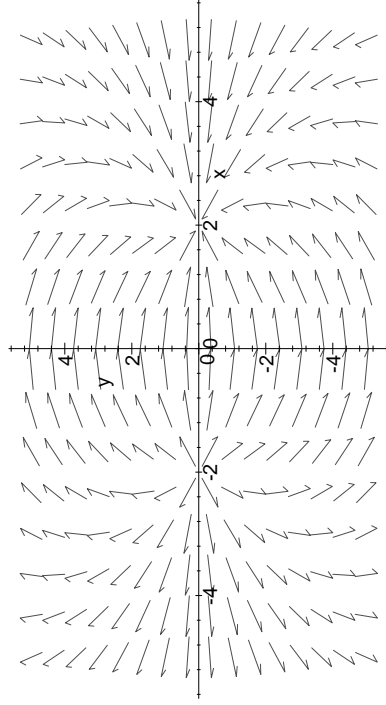


Abb. 40: Das Richtungsfeld zum obigen Feld

Da gerade die Richtung sehr wesentlich für die Konstruktion von Feldlinien ist, begnügt man sich oft damit, nur die Richtungen darzustellen, ansonsten aber sämtlichen Vektoren dieselbe Länge zu geben; siehe dazu etwa Abbildung 40. Dadurch geht zwar ein Teil der Information verloren, aber das entsprechende Bild kann trotzdem oft mehr aussagen, als eines, bei dem fast alle Pfeile visuell zu Punkten degeneriert sind. Auch bei der Darstellung als Richtungsfeld müssen allerdings natürlich alle Punkte  $\mathbf{x}$ , für die  $f(\mathbf{x}) = (0, \dots, 0)$  ist, durch Punkte oder überhaupt nicht dargestellt werden, denn dort gibt es keine Richtung.

### a) Der Begriff des Vektorfelds

Nicht nur zur Veranschaulichung kann es nützlich sein, die Werte einer Funktion als Vektoren zu betrachten: Die für uns wichtigsten Beispiele von Funktionen mehrerer Veränderlicher, deren Wert nicht in  $\mathbb{R}$  sondern in einem  $\mathbb{R}^n$  liegt, sind die aus der Physik bekannten Felder wie etwa ein Kraftfeld, ein elektrisches Feld oder ein Magnetfeld. Da solche Felder typischerweise als Vektoren geschrieben werden, schreiben wir deshalb in diesem Paragraphen die *Funktionswerte* vektoriell, wohingegen die Argumente weiterhin Punkte sind, die als Tupel geschrieben werden. Wir betrachten somit Funktionen

$$\vec{V}: D \rightarrow \mathbb{R}^n; \quad \mathbf{x} \mapsto \vec{V}(\mathbf{x}) = \begin{pmatrix} V_1(\mathbf{x}) \\ \vdots \\ V_n(\mathbf{x}) \end{pmatrix},$$

wobei die Komponentenfunktionen  $V_i: D \rightarrow \mathbb{R}$  genau die im vorigen Paragraphen betrachteten skalarwertigen Funktionen mehrerer Veränderlicher sind.

Der mathematische Begriff, der physikalische Felder beschreibt, ist der des Vektorfelds:

**Definition:** Ein Vektorfeld auf der Teilmenge  $D \subseteq \mathbb{R}^n$  ist eine Funktion  $\vec{V}: D \rightarrow \mathbb{R}^n$ .

(Da die typischen elektromagnetischen Felder wie  $\vec{E}$ ,  $\vec{D}$ ,  $\vec{B}$ ,  $\vec{H}$  oder auch das Kraftfeld  $\vec{F}$  üblicherweise durch Großbuchstaben bezeichnet werden, verwenden wir auch für Vektorfelder meist Großbuchstaben wie  $\vec{V}$  oder  $\vec{W}$ .)

### b) Die Jacobi-Matrix

Wir werden uns in diesem Paragraphen zwar hauptsächlich mit Vektorfeldern im  $\mathbb{R}^n$  beschäftigen, und uns selbst da teilweise auf den Fall  $n = 3$  beschränken; zumindest die grundlegenden Definitionen sollen aber für beliebige Funktionen auf einer Teilmenge  $D$  von  $\mathbb{R}^n$  mit Werten in irgendeinem  $\mathbb{R}^m$  betrachtet werden.

Fast von selbst versteht sich die

**Definition:** Eine Funktion  $\vec{f}: D \rightarrow \mathbb{R}^m$  auf  $D \subseteq \mathbb{R}^n$  heißt stetig  $b_{zw}$ . differenzierbar, wenn alle Komponentenfunktionen  $f_i$  stetig  $b_{zw}$ . differenzierbar sind. Entsprechend definieren wir auch die Menge  $\mathcal{C}^k(D, \mathbb{R}^m)$  aller  $k$ -fach stetig differenzierbarer Funktionen von  $D$  nach  $\mathbb{R}^m$  als Menge aller Funktionen, deren Komponenten  $f_i \in \mathcal{C}^k(D, \mathbb{R})$   $k$ -mal stetig differenzierbar sind.

Wie bei skalarwertigen Funktionen überlegt sich leicht, daß  $\mathcal{C}^k(D, \mathbb{R}^m)$  ein  $\mathbb{R}$ -Vektorraum ist. Damit wissen wir also, was Differenzierbarkeit einer Funktion  $\vec{f}: D \rightarrow \mathbb{R}^n$  bedeutet; was aber ist ihre Ableitung?

Bei skalarwertigen Funktionen einer oder mehrerer Veränderlicher hatten wir die Ableitung jeweils als Linearisierung der Funktion in der Umgebung des betrachteten Punktes definiert; genauso gehen wir auch hier vor: Für jede der Komponentenfunktion  $f_i$  ist

$$f_i(\mathbf{x} + \vec{h}) = f_i(\mathbf{x}) + \text{grad } f_i(\mathbf{x}) \cdot \vec{h} + o(\|\vec{h}\|),$$

wobei der Punkt hier für das Skalarprodukt von Vektoren steht. Im Hinblick auf vektorwertige Funktionen ist es aber besser, anstelle des Skalarprodukts ein Matrixprodukt zu verwenden; dazu müssen wir den Gradienten als Zeilenvektor schreiben, d.h. wenn  $\cdot$  das Matrixprodukt bezeichnet, ist

$$f_i(\mathbf{x} + \vec{h}) = f_i(\mathbf{x}) + {}^t \text{grad } f_i(\mathbf{x}) \cdot \vec{h} + o(\|\vec{h}\|).$$

In dieser Notation können wir die als Zeilenvektoren aufgefaßten Gradienten zu einer  $n \times n$ -Matrix zusammenfassen, der JACOBI-Matrix

$$J_{\vec{f}}(\mathbf{x}) = \begin{pmatrix} {}^t \text{grad } f_1(\mathbf{x}) \\ \vdots \\ {}^t \text{grad } f_m(\mathbf{x}) \end{pmatrix} = \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial x_1} & \cdots & \frac{\partial f_m}{\partial x_n} \end{pmatrix}.$$

Mit dieser Matrix ist dann

$$\vec{f}(\mathbf{x} + \vec{h}) = \vec{f}(\mathbf{x}) + J_{\vec{f}}(\mathbf{x}) \cdot \vec{h} + o(\|\vec{h}\|),$$

wobei hier das Produkt als Matrixprodukt zu verstehen wird und das Symbol  $o(\|\vec{h}\|)$  schlampigerweise auch für einen Vektor benutzt wird, dessen sämtliche Komponenten  $o(\|\vec{h}\|)$  sind.

### c) Die Divergenz eines Vektorfelds

Die JACOBI-Matrix eines Vektorfelds hängt stark ab vom gewählten Koordinatensystem. Wir wollen aus dieser Matrix Größen ableiten, von denen wir später sehen werden, daß sie *intrinsic*, d.h. also vom Koordinatensystem unabhängige Eigenschaften des Vektorfelds beschreiben.

**Definition:** a) Die *Spur* einer  $n \times n$ -Matrix  $A = (a_{ij})$  ist die Summe

$$\text{Spur } A = a_{11} + a_{22} + \cdots + a_{nn}$$

der Diagonalelemente von  $A$ .

b) Die Spur der JACOBI-Matrix  $J_{\vec{v}}(\mathbf{x})$  eines Vektorfelds heißt *Divergenz* oder *Quelldichte* von  $\vec{v}$ :

$$\text{div } \vec{v}(\mathbf{x}) = \text{Spur } J_{\vec{v}}(\mathbf{x}).$$

Um ein anschauliches Verständnis von der Divergenz zu bekommen, betrachten wir den (sehr speziellen) Fall, daß die JACOBI-Matrix von  $\vec{v}$  eine Diagonalmatrix ist:

$$J_{\vec{v}}(\mathbf{x}) = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix}$$

Dann ist für jeden Einheitsvektor  $\vec{e}_i$

$$\vec{v}(\mathbf{x} + h\vec{e}_i) = \vec{v}(\mathbf{x}) + h \cdot J_{\vec{v}}(\mathbf{x})\vec{e}_i + o(h) = \vec{v}(\mathbf{x}) + h\lambda_i\vec{e}_i + o(h),$$

und für einen beliebigen Vektor  $\vec{h} = h_1\vec{e}_1 + \cdots + h_n\vec{e}_n$  ist

$$\vec{v}(\mathbf{x} + \vec{h}) = \vec{v}(\mathbf{x}) + J_{\vec{v}}(\mathbf{x})\vec{h} + o(\|\vec{h}\|) = \vec{v}(\mathbf{x}) + \begin{pmatrix} \lambda_1 h_1 \\ \vdots \\ \lambda_n h_n \end{pmatrix} + o(\|\vec{h}\|).$$

Falls alle  $\lambda_i$  positiv sind, hat hier jede Komponente des Vektors  $\vec{v}(\mathbf{x}) + J_{\vec{v}}(\mathbf{x})\vec{h}$  dasselbe Vorzeichen wie die entsprechende Komponente von  $\vec{h}$ ; entfernt man sich also vom Punkt  $\mathbf{x}$ , so zeigt auch die Veränderung des Vektorfelds  $\vec{v}$  weg vom Punkt  $\mathbf{x}$  und umgekehrt; die Situation

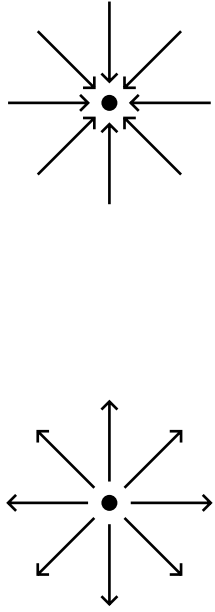


Abb. 41: Quellen und Senken

ist also ungefähr so wie auf der linken Seite von Abbildung 41. Man bezeichnet  $\boldsymbol{x}$  in einem solchen Fall als eine *Quelle* des Vektorfelds  $\vec{V}$ .

Falls alle  $\lambda_i$  negativ sind, gehen die Änderungen jeweils in Gegenrichtung, es sieht also ungefähr so aus wie im rechten Teil von Abbildung 41, und man spricht von einer *Senke*.

Im allgemeinen werden allerdings weder *alle*  $\lambda_i$  positiv noch *alle*  $\lambda_i$  negativ sein; die Divergenz als Summe der  $\lambda_i$  sagt uns dann, ob sich der Punkt eher wie eine Quelle oder eher wie eine Senke verhält.

Nun wird es natürlich nur selten vorkommen, daß die JACOBI-Matrix eine Diagonalmatrix ist; wie wir am Ende dieses Kapitels sehen werden, ist die Divergenz trotzdem auch im allgemeinen Fall eine Maßzahl dafür, ob ein Punkt Quelle oder Senke ist, und zwar in einer viel präziseren Weise, als es die obige Diskussion für Diagonalmatrizen zeigte.

Im übrigen sind Diagonalmatrizen doch nicht so speziell, wie es zunächst den Anschein hat: Im nächsten Semester werden wir sehen, daß es zu jeder symmetrischen Matrix ein Koordinatensystem gibt, bezüglich dessen die durch die Matrix definierte lineare Abbildung von  $\mathbb{R}^n$  nach  $\mathbb{R}^n$  eine Diagonalmatrix als Abbildungsmatrix hat; zumindest für den symmetrischen Anteil der JACOBI-Matrix ist die obige Diskussion also richtig. Mit dem antisymmetrischen Anteil beschäftigen wir uns im nächsten Abschnitt.

In diesem Abschnitt wollen wir uns nur noch überlegen, wie man die Divergenz eines Vektorfelds tatsächlich ausrechnet. Die Divergenz ist

nach Definition gleich der Spur der JACOBI-Matrix, also ist

$$\begin{aligned} \operatorname{div} \vec{V} &= \operatorname{Spur} J_{\vec{V}}(x) = \operatorname{Spur} \begin{pmatrix} \frac{\partial V_1}{\partial x_1} & \frac{\partial V_1}{\partial x_2} & \cdots & \frac{\partial V_1}{\partial x_n} \\ \frac{\partial V_2}{\partial x_1} & \frac{\partial V_2}{\partial x_2} & \cdots & \frac{\partial V_2}{\partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial V_n}{\partial x_1} & \frac{\partial V_n}{\partial x_2} & \cdots & \frac{\partial V_n}{\partial x_n} \end{pmatrix} \\ &= \frac{\partial V_1}{\partial x_1} + \frac{\partial V_2}{\partial x_2} + \cdots + \frac{\partial V_n}{\partial x_n}. \end{aligned}$$

Genauso wie wir den Gradienten einer skalarwertigen Funktion  $f$  auch als  $\nabla f$  schreiben, schreiben wir die Divergenz eines Vektorfelds  $\vec{V}$  auch als  $\nabla \cdot \vec{V}$ . Beides kann formal auch als Produkt bzw. Skalarprodukt mit einem vektorwertigen Operator interpretiert werden:

$$\nabla f = \begin{pmatrix} \frac{\partial}{\partial x_1} \\ \frac{\partial}{\partial x_2} \\ \vdots \\ \frac{\partial}{\partial x_n} \end{pmatrix} \cdot f = \begin{pmatrix} \frac{\partial f}{\partial x_1} \\ \frac{\partial f}{\partial x_2} \\ \vdots \\ \frac{\partial f}{\partial x_n} \end{pmatrix}$$

und

$$\nabla \cdot \vec{V} = \begin{pmatrix} \frac{\partial}{\partial x_1} \\ \frac{\partial}{\partial x_2} \\ \vdots \\ \frac{\partial}{\partial x_n} \end{pmatrix} \cdot \begin{pmatrix} V_1 \\ V_2 \\ \vdots \\ V_n \end{pmatrix} = \frac{\partial V_1}{\partial x_1} + \frac{\partial V_2}{\partial x_2} + \cdots + \frac{\partial V_n}{\partial x_n}.$$

Diese Interpretation darf allerdings nicht zu ernst genommen werden, denn obwohl Skalarprodukt und Produkt mit einem Skalar beide kommutativ sind, ist weder die Gleichung  $\nabla \cdot \vec{V} = \vec{V} \cdot \nabla$  noch die Gleichung  $\nabla f = f \nabla$  auch nur annähernd sinnvoll, geschweige denn richtig.

#### d) Die Rotation eines dreidimensionalen Vektorfelds

Erinnern wir uns zunächst an symmetrische und antisymmetrische Matrizen: Eine Matrix ist *symmetrisch*, wenn  $a_{ij} = a_{ji}$  ist für alle Indexpaare  $(i, j)$ ; sie ist *antisymmetrisch* im Falle  $a_{ij} = -a_{ji}$ . Jede reelle Matrix  $A$  kann als Summe einer symmetrischen Matrix  $S$  und einer

antisymmetrischen Matrix  $T$  geschrieben werden; die Einträge von  $S$  und  $T$  sind jeweils

$$\frac{a_{ij} + a_{ji}}{2} \quad \text{und} \quad \frac{a_{ij} - a_{ji}}{2}.$$

Für  $n = 1$  ist jede  $n \times n$ -„Matrix“ symmetrisch, es gibt als keinen antisymmetrischen Anteil.

Für  $n = 2$  hat jede antisymmetrische Matrix  $A$  die Form

$$\begin{pmatrix} 0 & a \\ -a & 0 \end{pmatrix},$$

für den antisymmetrischen Anteil der JACOBI-Matrix ist

$$a = \frac{1}{2} \left( \frac{\partial f_1}{\partial y} - \frac{\partial f_2}{\partial x} \right).$$

Damit ist der antisymmetrische Anteil durch diese eine Zahl bestimmt.

Für  $n > 3$  gibt es keine einfachere Darstellung des antisymmetrischen Anteils als die schiefsymmetrische Matrix selbst, bleibt also noch der für viele Anwendungen besonders wichtige Fall  $n = 3$ .

Eine antisymmetrische  $3 \times 3$ -Matrix  $A$  hat die Form

$$A = \begin{pmatrix} 0 & \alpha & \beta \\ -\alpha & 0 & \gamma \\ -\beta & -\gamma & 0 \end{pmatrix}.$$

Für einen Vektor  $\vec{x} \in \mathbb{R}^3$  ist daher

$$A\vec{x} = \begin{pmatrix} 0 & \alpha & \beta \\ -\alpha & 0 & \gamma \\ -\beta & -\gamma & 0 \end{pmatrix} \cdot \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} \alpha y + \beta z \\ -\alpha x + \gamma z \\ -\beta x - \gamma y \end{pmatrix}.$$

Diese spezielle Form des Matrix-Vektor-Produkts sollte Erinnerungen an das Vektorprodukt im  $\mathbb{R}^3$  wecken:

$$\begin{pmatrix} a \\ b \\ c \end{pmatrix} \times \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} bz - cy \\ cx - az \\ ay - bx \end{pmatrix},$$

d.h.

$$\begin{pmatrix} a \\ b \\ c \end{pmatrix} \times \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 & -c & b \\ c & 0 & -a \\ -b & a & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

und damit

$$\begin{pmatrix} 0 & \alpha & \beta \\ -\alpha & 0 & \gamma \\ -\beta & -\gamma & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} -\gamma \\ \beta \\ -\alpha \end{pmatrix} \times \begin{pmatrix} x \\ y \\ z \end{pmatrix}.$$

Somit entspricht im Dreidimensionalen die Multiplikation mit einer schiefsymmetrischen Matrix genau dem Vektorprodukt (das im übrigen auch nur im Dreidimensionalen definierbar ist).

Der Vektor  $\vec{a}$  läßt sich im vorliegenden Fall leicht ausrechnen: Die JACOBI-Matrix

$$J_{\vec{V}}(\mathbf{x}) = \begin{pmatrix} \frac{\partial V_1}{\partial x} & \frac{\partial V_1}{\partial y} & \frac{\partial V_1}{\partial z} \\ \frac{\partial V_2}{\partial x} & \frac{\partial V_2}{\partial y} & \frac{\partial V_2}{\partial z} \\ \frac{\partial V_3}{\partial x} & \frac{\partial V_3}{\partial y} & \frac{\partial V_3}{\partial z} \end{pmatrix}$$

eines dreidimensionalen Vektorfelds  $\vec{V}$  hat antisymmetrischen Anteil

$$A = \frac{1}{2} (A - {}^t A) = \frac{1}{2} \begin{pmatrix} 0 & \frac{\partial V_1}{\partial y} - \frac{\partial V_2}{\partial x} & \frac{\partial V_1}{\partial z} - \frac{\partial V_3}{\partial x} \\ \frac{\partial V_2}{\partial x} - \frac{\partial V_1}{\partial y} & 0 & \frac{\partial V_2}{\partial z} - \frac{\partial V_3}{\partial y} \\ \frac{\partial V_3}{\partial x} - \frac{\partial V_1}{\partial z} & \frac{\partial V_3}{\partial y} - \frac{\partial V_2}{\partial z} & 0 \end{pmatrix},$$

und wie wir uns gerade überlegt haben, ist für jeden Vektor  $\vec{v}$

$$A\vec{v} = \vec{a} \times \vec{v} \quad \text{mit} \quad \vec{a} = \frac{1}{2} \begin{pmatrix} \frac{\partial V_1}{\partial y} - \frac{\partial V_2}{\partial x} \\ \frac{\partial V_2}{\partial x} - \frac{\partial V_1}{\partial y} \\ \frac{\partial V_3}{\partial x} - \frac{\partial V_1}{\partial z} \end{pmatrix}.$$

Da der Faktor  $\frac{1}{2}$  in konkreten Rechnungen nur hinderlich wäre, betrachten wir den Vektor  $2\vec{a}$  und definieren:

**Definition:**  $\vec{V}: D \rightarrow \mathbb{R}^3$  sei ein Vektorfeld auf der offenen Teilmenge  $D \subseteq \mathbb{R}^3$ . Dann bezeichnen wir das Vektorfeld

$$\text{rot } \vec{V}: D \rightarrow \mathbb{R}^3; \quad \mathbf{x} \mapsto \begin{pmatrix} \frac{\partial V_1}{\partial y} - \frac{\partial V_2}{\partial x} \\ \frac{\partial V_2}{\partial z} - \frac{\partial V_3}{\partial x} \\ \frac{\partial V_3}{\partial x} - \frac{\partial V_1}{\partial z} \end{pmatrix}$$

als Rotation von  $\vec{V}$ .

Für ein Vektorfeld  $\vec{V} \in \mathcal{C}^1(D, \mathbb{R}^3)$ , dessen JACOBI-Matrix symmetrischen Anteil  $S$  hat, ist also

$$\vec{V}(\vec{x} + \vec{h}) = \vec{V}(\vec{x}) + S\vec{h} + \frac{1}{2}(\text{rot } \vec{V}) \times \vec{h} + o(\|\vec{h}\|).$$

Der Name *Rotation* wird im folgenden (etwas umfangreichen) Beispiel verständlich: Wir gehen aus vom Vektor

$$\vec{a} = 2 \text{ rot } \vec{V}(\vec{x}).$$

Nach Definition des Vektorprodukts steht  $\vec{a} \times \vec{h}$  senkrecht sowohl auf  $\vec{a}$  als auch auf  $\vec{h}$ . Bedenkt man noch, daß Vektoren Äquivalenzklassen von Pfeilen sind, die man sich nicht unbedingt so vorstellen muß, daß sie alle im selben Punkt beginnen, bietet sich folgende Interpretation an: Wenn wir den Vektor  $\vec{h}$  um die von  $\vec{a}$  aufgespannte Achse drehen, zeigt  $\vec{a} \times \vec{h}$  in Richtung der Tangente der Bahn des Endpunkts von  $\vec{h}$ . (Damit dies stimmt, muß die Achse so orientiert werden, daß sich  $\vec{h}$  im mathematisch positiven, d.h. Gegenurzeigersinn, um  $\vec{a}$  dreht. Von der Ebene der Uhr aus gesehen zeigt dann  $\vec{a}$  nach oben,  $\vec{h}$  zu dem sich drehenden Punkt, und  $\vec{a} \times \vec{h}$  in Richtung des Gegenurzeigersinns, wie man sich z.B. mit der Dreifingerregel leicht veranschaulicht.)

Der antisymmetrische Anteil der JACOBI-Matrix sollte also etwas mit einer Drehbewegung zu tun haben, genauer gesagt mit deren Geschwindigkeitsvektor, denn dieser zeigt in Richtung der Tangente der Bahnkurve.

Um dies genauer zu untersuchen, betrachten wir zunächst eine einfache Drehung eines Vektors um die  $z$ -Achse mit Drehwinkel  $\varphi$ . Diese ist aufgrund der Definition von Sinus und Cosinus ausdrückbar durch die lineare Abbildung

$$\begin{aligned} x &\mapsto x \cos \varphi - y \sin \varphi \\ y &\mapsto x \sin \varphi + y \cos \varphi \\ z &\mapsto z \end{aligned}$$

Als nächstes wollen wir einen Punkt  $\vec{x} \in \mathbb{R}^3$  um eine beliebige Achse drehen; der Drehwinkel sei weiterhin  $\varphi$ . Dazu sei  $\vec{a}$  der Einheitsvektor in Richtung der Achse; er sei so orientiert, daß der Drehsinn (bei positivem Winkel  $\varphi$ ) gleich dem Gegenurzeigersinn ist. Dann gelten in den auf  $\vec{a}$  senkrecht stehenden Ebenen im wesentlichen dieselben Formeln, falls wir nur die Koordinateneinheitsvektoren der  $x$ - und der  $y$ -Achse ersetzen durch zwei aufeinander und auf  $\vec{a}$  senkrecht stehende gleichlange Vektoren. Als ein solcher Vektor bietet sich der Verbindungsvektor  $\vec{r}$  von der Achse zum Punkt  $\vec{x}$  an, d.h. jener Vektor, der den Abstand zwischen  $\vec{x}$  und der Achse beschreibt. Ein sowohl auf  $\vec{r}$  als auch auf  $\vec{a}$  senkrecht stehender Vektor ist  $\vec{s} = \vec{a} \times \vec{r}$ ; da  $\vec{a}$  die Länge eins hat, ist  $\vec{s}$  genauso lang wie  $\vec{r}$ . In der von  $\vec{r}$  und  $\vec{s}$  aufgespannten Ebene bildet die Drehung um den Winkel  $\varphi$  somit den Vektor  $\alpha \vec{r} + \beta \vec{s}$  ab auf

$$(\alpha \cos \varphi - \beta \sin \varphi) \vec{r} + (\alpha \sin \varphi + \beta \cos \varphi) \vec{s}.$$

Nun zerlegen wir den Ortsvektor  $\vec{x}$  eines Punktes  $\vec{x}$  in einen Vektor  $\vec{x}_a$  auf der Achse und den radialen, d.h. senkrecht auf der Achse stehenden, Anteil  $\vec{r}$ .

$\vec{x}_a$  ist die Projektion von  $\vec{x}$  auf die Drehachse; da wir  $\vec{a}$  als Einheitsvektor vorausgesetzt haben, ist dies nach den allgemeinen Regeln der Vektorrechnung einfach

$$\vec{x}_a = (\vec{a} \cdot \vec{x}) \vec{a},$$

denn die Länge des Vektors wird bei der Projektion mit dem Cosinus des Winkels zur Achse multipliziert, und die Richtung des auf die Achse projizierten Vektors ist natürlich gleich der Richtung der Achse.

Damit läßt sich auch die radiale Komponente problemlos berechnen: Da beide Komponenten zusammen den Vektor  $\vec{x}$  ergeben müssen, ist

$$\vec{r} = \vec{x} - \vec{x}_a = \vec{x} - (\vec{a} \cdot \vec{x}) \vec{a}.$$

Nun ist  $\vec{x}_a$  parallel zu  $\vec{a}$ , und das Vektorprodukt zweier paralleler Vektoren verschwindet. Deshalb können wir auch schreiben

$$\vec{s} = \vec{a} \times \vec{r} = \vec{a} \times (\vec{x} - (\vec{a} \cdot \vec{x}) \vec{a}) = \vec{a} \times \vec{x}.$$

Da die Axialkomponente eines Vektors bei Drehungen erhalten bleibt, wird speziell der Vektor  $\vec{x} = \vec{x}_a + \vec{r}$  also abgebildet auf

$$\begin{aligned} \vec{x}_a + \cos \varphi \vec{r} + \sin \varphi \vec{s} \\ = (\vec{a} \cdot \vec{x}) \vec{a} + \cos \varphi (\vec{x} - (\vec{a} \cdot \vec{x}) \vec{a}) + \sin \varphi \vec{a} \times \vec{x} \\ = (1 - \cos \varphi) (\vec{a} \cdot \vec{x}) \vec{a} + \cos \varphi \vec{x} + \sin \varphi \vec{a} \times \vec{x}. \end{aligned}$$

Um den Geschwindigkeitsvektor einer Drehung auszurechnen, müssen wir diese zunächst zu einem dynamischen Prozeß machen, d.h. anstelle einer Drehung mit konstantem Drehwinkel  $\varphi$  betrachten wir eine Drehung mit Winkelgeschwindigkeit  $\omega$ . Wir gehen wieder aus von einem festen Punkt  $\vec{x}$  und betrachten dessen Position zum Zeitpunkt  $t$ , d.h. wir betrachten die Funktion

$$\vec{f}: \begin{cases} \mathbb{R} & \rightarrow \mathbb{R}^3 \\ t & \mapsto (1 - \cos \omega t) (\vec{a} \cdot \vec{x}) \vec{a} + \cos \omega t \vec{x} + \sin \omega t \vec{a} \times \vec{x}. \end{cases}$$

Der Geschwindigkeitsvektor im Punkt  $\vec{x}$  ist die Ableitung dieser Funktion im Nullpunkt, also der Vektor

$$\frac{d\vec{f}}{dt}(0) = \left( \omega \sin \omega t (\vec{a} \cdot \vec{x}) \vec{a} - \omega \sin \omega t \vec{x} + \omega \cos \omega t \vec{a} \times \vec{x} \right) \Big|_{t=0} = \omega \vec{a} \times \vec{x}.$$

Mit anderen Worten: Bei der Drehung mit Winkelgeschwindigkeit  $\omega$  um die von  $\vec{a}$  aufgespannte Achse ist der Geschwindigkeitsvektor im Punkt  $\vec{x}$  gleich  $\omega \vec{a} \times \vec{x}$ .

Der Vektor  $\omega \vec{a}$  läßt sich auffassen als eine vektorielle Winkelgeschwindigkeit (die in der Physik in der Tat genau so definiert wird), und für einen beliebigen Vektor  $\vec{w} \in \mathbb{R}^3$  ist das Vektorfeld

$$\vec{V}: \mathbb{R}^3 \rightarrow \mathbb{R}^3; \quad \vec{x} \mapsto \vec{w} \times \vec{x}$$

das Feld der Geschwindigkeitsvektoren zu einer Rotation mit vektorieller Winkelgeschwindigkeit  $\vec{\omega}$ .

Es sei noch darauf hingewiesen, daß  $\text{rot } \vec{V}$  in der englischsprachigen Literatur meist als  $\text{curl } \vec{V}$  geschrieben wird nach dem englischen Wort *curl* = Locke.

Wie Gradient und Divergenz können wir auch die Rotation formal mit Hilfe des Operators  $\nabla$  ausdrücken: Für ein dreidimensionales Vektorfeld  $\vec{V}$  ist

$$\nabla \times \vec{V} = \begin{pmatrix} \frac{\partial}{\partial x} \\ \frac{\partial}{\partial y} \\ \frac{\partial}{\partial z} \end{pmatrix} \times \begin{pmatrix} V_1 \\ V_2 \\ V_3 \end{pmatrix} = \begin{pmatrix} \frac{\partial V_3}{\partial y} - \frac{\partial V_2}{\partial z} \\ \frac{\partial V_1}{\partial z} - \frac{\partial V_3}{\partial x} \\ \frac{\partial V_2}{\partial x} - \frac{\partial V_1}{\partial y} \end{pmatrix} = \text{rot } \vec{V}.$$

**e) Erste Beispiele**

Zum besseren Verständnis von Divergenz und Rotation und um den Umgang mit den beiden Begriffen zu üben, wollen wir sie für zwei einfache Felder ausrechnen: Für das elektrische Feld einer Punktladung und für das Magnetfeld eines stromdurchflossenen Leiters.

**1) Das elektrische Feld einer Punktladung:** Eine Punktladung  $q$  im Nullpunkt erzeugt nach dem COULOMBSchen Gesetz ein elektrisches Feld  $\vec{E}$  der Stärke

$$\vec{E}(\mathbf{x}) = \frac{q}{4\pi\epsilon_0} \frac{\vec{x}}{|\vec{x}|^3},$$

denn die Feldstärke nimmt mit der Entfernung vom Nullpunkt quadratisch ab und hat die Richtung (oder – je nach Vorzeichen von  $q$  – auch Gegenrichtung) des Ortsvektors, d.h. der Einheitsvektor der Richtung ist  $\pm \vec{x}/|\vec{x}|$ , und er muß noch durch  $|\vec{x}|^2$  dividiert werden.

Das Feld  $\vec{E}$  ist nur definiert auf  $\mathbb{R}^3 \setminus \{(0, 0, 0)\}$ ; im Nullpunkt haben wir einen Nenner Null, und auch physikalisch ist es nicht sinnvoll, nach dem Feld *im Innern* einer Punktladung zu fragen: Punktladungen sind nun einmal mathematische Idealisierungen, und wirklich anwendbar ist das COULOMBSche Gesetz ohnehin erst ab einem gewissen Mindestabstand: Bei zu kleinen Abständen spielen andere Wechselwirkungen eine wesentlich größere Rolle.

In Koordinaten ist also

$$\vec{E}(\mathbf{x}) = \begin{pmatrix} E_1(\mathbf{x}) \\ E_2(\mathbf{x}) \\ E_3(\mathbf{x}) \end{pmatrix} = \frac{q}{4\pi\epsilon_0} \frac{1}{\|\vec{x}\|^3} \begin{pmatrix} x \\ y \\ z \end{pmatrix}.$$

Zur Berechnung der Divergenz müssen wir die erste Komponente nach  $x$ , die zweite nach  $y$  und die dritte nach  $z$  ableiten und dann die drei Ableitungen addieren. Beachten wir, daß

$$\|\vec{x}\|^3 = (x^2 + y^2 + z^2)^{3/2}$$

ist, folgt sofort aus der Quotientenregel, daß

$$\begin{aligned} \frac{\partial E_1}{\partial x}(\mathbf{x}) &= \frac{q}{4\pi\epsilon_0} \frac{\partial}{\partial x} \frac{x}{\|\vec{x}\|^3} = \frac{q}{4\pi\epsilon_0} \frac{\|\vec{x}\|^3 - x \cdot \frac{3}{2} \cdot 2x \|\vec{x}\|^{-1}}{\|\vec{x}\|^6} \\ &= \frac{q}{4\pi\epsilon_0} \frac{\|\vec{x}\|^2 - 3x^2}{\|\vec{x}\|^5} = \frac{q}{4\pi\epsilon_0} \frac{1}{\|\vec{x}\|^5} (-2x^2 + y^2 + z^2). \end{aligned}$$

Entsprechend ist

$$\frac{\partial E_2}{\partial y}(\mathbf{x}) = \frac{q}{4\pi\epsilon_0} \frac{1}{\|\vec{x}\|^5} (x^2 - 2y^2 + z^2)$$

und

$$\frac{\partial E_3}{\partial z}(\mathbf{x}) = \frac{q}{4\pi\epsilon_0} \frac{1}{\|\vec{x}\|^5} (x^2 + y^2 - 2z^2),$$

also

$$\begin{aligned} \text{div } \vec{E}(\mathbf{x}) &= \frac{q}{4\pi\epsilon_0} \frac{1}{\|\vec{x}\|^5} (-2x^2 + y^2 + z^2 + x^2 - 2y^2 + z^2 + y^2 - 2z^2) \\ &= 0. \end{aligned}$$

Zur Berechnung der Rotation benötigen wir die Einträge der JACOBI-Matrix, die nicht in der Diagonale stehen; dabei werden die Komponenten von  $\vec{E}$  nach jenen Variablen abgeleitet, die *nicht* im Zähler stehen. Die Situation ist also stets dieselbe, und es genügt, wenn wir beispielsweise die Ableitung von  $E_1$  nach  $y$  ausrechnen.

Aus den üblichen Ableitungsregeln folgt sofort, daß

$$\frac{\partial E_1}{\partial y} = \frac{q}{4\pi\epsilon_0} \left( -\frac{3}{2} \right) \frac{x \cdot 2y}{\|\vec{x}\|^5} = -\frac{3q}{4\pi\epsilon_0} \frac{xy}{\|\vec{x}\|^5}$$

ist, und analog ist auch

$$\frac{\partial E_2}{\partial x} = \frac{q}{4\pi\epsilon_0} \left( -\frac{3}{2} \right) \frac{y \cdot 2x}{\|\vec{x}\|^5} = -\frac{3q}{4\pi\epsilon_0} \frac{xy}{\|\vec{x}\|^5}.$$

Die  $x$ -Komponente von  $\text{rot } \vec{E}$  ist somit

$$\frac{\partial E_2}{\partial x} - \frac{\partial E_1}{\partial y} = 0,$$

und genauso überzeugt man sich auch vom Verschwinden der anderen Komponenten.

Somit ist also das Feld einer Punktladung außerhalb des Punktes selbst sowohl quellen- als auch wirbelfrei, genau wie es nach den MAXWELLSchen Gleichungen auch sein muß.

**2) Das Magnetfeld eines stromdurchflossenen Leiters:** Das Magnetfeld eines von einem Strom der Stärke  $I$  durchflossenen geradlinigen Drahts ist nach dem BIOT-SAVARTSchen Gesetz

$$\vec{H}(\mathbf{x}) = \frac{I}{4\pi\|\vec{r}\|^2} \vec{a} \times \vec{x},$$

wobei  $\vec{a}$  der Einheitsvektor in Richtung des Stroms und  $\vec{r}$  der Abstandsvektor zwischen dem Draht und dem Punkt  $\vec{x}$  ist.  $\vec{H}$  ist überall im  $\mathbb{R}^3$  erklärt, außer auf der von  $\vec{a}$  aufgespannten Geraden, wo  $\vec{r}$  der Nullvektor ist.

Zum einfacheren Rechnen wählen wir das Koordinatensystem so, daß  $\vec{a}$  der Einheitsvektor in Richtung der positiven  $z$ -Achse ist; für den Punkt mit Koordinaten  $(x, y, z)$  ist dann

$$\vec{r} = \begin{pmatrix} x \\ y \\ 0 \end{pmatrix} \quad \text{und} \quad \vec{a} \times \vec{x} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \times \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} -y \\ x \\ 0 \end{pmatrix},$$

d.h.

$$\vec{H}(\mathbf{x}) = \frac{I}{4\pi(x^2 + y^2)} \begin{pmatrix} -y \\ x \\ 0 \end{pmatrix}.$$

Dieses Vektorfeld ist erklärt auf  $\mathbb{R}^3$  minus der  $z$ -Achse; auf der  $z$ -Achse, d.h. im Innern des Leiters, bekommen wir eine Null in den Nenner.

Die drei Komponenten dieses Felds könnten wir nun wie oben nach den drei Variablen  $x, y, z$  ableiten, um so Rotation und Divergenz zu berechnen; es gibt allerdings eine einfachere Möglichkeit, die JACOBI-Matrix eines solchen Felds zu bestimmen:

Wie auch das  $\vec{E}$ -Feld im obigen Beispiel läßt sich  $\vec{H}$  schreiben als Produkt einer skalaren Funktion mit einer linearen Abbildung  $\mathbb{R}^3 \rightarrow \mathbb{R}^3$ . Solche Funktionen, bei denen alle Nichtlinearität in einer skalaren Funktion steckt, bezeichnen wir als *quasilinear*:

**Definition:** Eine Funktion  $\vec{f}: D \rightarrow \mathbb{R}^m$  auf einer Teilmenge  $D \subseteq \mathbb{R}^n$  heißt *quasilinear*, wenn es eine Funktion  $\varphi: D \rightarrow \mathbb{R}$  und eine Matrix  $A \in \mathbb{R}^{n \times m}$  gibt, so daß  $\vec{f}(\mathbf{x}) = \varphi(\mathbf{x}) \cdot A\vec{x}$  ist.

In diesem Sinne ist  $\vec{H}(\mathbf{x})$  quasilinear mit

$$\varphi(\mathbf{x}) = \frac{I}{4\pi(x^2 + y^2)} \quad \text{und} \quad A = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

und das obige Feld  $\vec{E}$  ist quasilinear mit

$$\varphi(\mathbf{x}) = \frac{1}{4\pi\epsilon_0\|\vec{r}\|^3} \quad \text{und} \quad A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

**Lemma:** Für eine quasilineare Funktion  $\vec{f}: D \rightarrow \mathbb{R}^m$  auf einer Teilmenge  $D \subseteq \mathbb{R}^n$  ist

$$J_{\vec{f}}(\mathbf{x}) = \varphi(\mathbf{x}) \cdot A + (A\vec{x}) \cdot {}^t \text{grad } \varphi(\mathbf{x}).$$

In dieser Formel steht der erste Malpunkt für die Multiplikation eines Skalars mit einer Matrix, der zweite für das Matrixmultiplikation des



Spaltenvektors  $A\vec{x}$  mit dem zum Zeilenvektor transponierten Gradienten. Es geht hier also nicht um ein Skalarprodukt, sondern um eine Matrixmultiplikation, deren Ergebnis eine  $n \times m$ -Matrix ist.

Ausgeschrieben wird die Formel zu

$$\frac{\partial f_i}{\partial x_j} = \varphi(\mathbf{x}) \cdot a_{ij} + \left( \sum_{\nu=1}^n a_{i\nu} x_\nu \right) \cdot \frac{\partial \varphi}{\partial x_j}.$$

Zum Beweis leiten wir  $f_i(\mathbf{x}) = \varphi(\mathbf{x}) \sum_{\nu=1}^n a_{i\nu} x_\nu$ , nach  $x_j$  ab; nach der Produktregel ist, wie gewünscht,

$$\begin{aligned} \frac{\partial f_i}{\partial x_j}(\mathbf{x}) &= \varphi(\mathbf{x}) \frac{\partial}{\partial x_j} \left( \sum_{\nu=1}^n a_{i\nu} x_\nu \right) + \frac{\partial \varphi}{\partial x_j}(\mathbf{x}) \sum_{\nu=1}^n a_{i\nu} x_\nu \\ &= \varphi(\mathbf{x}) \cdot a_{ij} + \frac{\partial \varphi}{\partial x_j}(\mathbf{x}) \sum_{\nu=1}^n a_{i\nu} x_\nu. \end{aligned}$$

■

Dieses Lemma wenden wir gleich an auf das Feld  $\vec{H}(\mathbf{x})$ ; hier ist

$$\frac{\partial \varphi}{\partial x}(\mathbf{x}) = \frac{-2x \cdot I}{4\pi(x^2 + y^2)^2} \quad \text{und} \quad \frac{\partial \varphi}{\partial y}(\mathbf{x}) = \frac{-2y \cdot I}{4\pi(x^2 + y^2)^2},$$

d.h.

$$\text{grad } \varphi(\mathbf{x}) = \frac{-2I}{4\pi(x^2 + y^2)^2} \begin{pmatrix} x \\ y \\ 0 \end{pmatrix}.$$

Somit ist

$$\begin{aligned} J_{\vec{H}}(\mathbf{x}) &= \frac{I}{4\pi(x^2 + y^2)} \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \\ &\quad + \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} \cdot \frac{-2I}{4\pi(x^2 + y^2)^2} \begin{pmatrix} x & y & 0 \end{pmatrix} \\ &= \frac{I}{4\pi(x^2 + y^2)^2} \begin{pmatrix} 0 & -x^2 - y^2 & 0 \\ x^2 + y^2 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \\ &\quad + \begin{pmatrix} -y \\ x \\ 0 \end{pmatrix} \cdot \begin{pmatrix} -2x & -2y & 0 \end{pmatrix} \\ &= \frac{I}{4\pi(x^2 + y^2)^2} \begin{pmatrix} 0 & -x^2 - y^2 & 0 \\ x^2 + y^2 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \\ &\quad + \begin{pmatrix} 2xy & 2y^2 & 0 \\ -2x^2 & -2xy & 0 \\ 0 & 0 & 0 \end{pmatrix} \\ &= \frac{I}{4\pi(x^2 + y^2)^2} \begin{pmatrix} 2xy & y^2 - x^2 & 0 \\ y^2 - x^2 & -2xy & 0 \\ 0 & 0 & 0 \end{pmatrix}. \end{aligned}$$

Diese Matrix ist symmetrisch und hat Spur Null, d.h. auch hier ist

$$\text{div } \vec{H}(\mathbf{x}) = 0 \quad \text{und} \quad \text{rot } \vec{H}(\mathbf{x}) = \vec{0},$$

und auch dies natürlich wieder in Übereinstimmung mit den MAXWELLSchen Gleichungen.

Da die Feldlinien des hier betrachteten Felds Kreise sind, mag es auf den ersten Blick seltsam erscheinen, daß die Rotation von  $\vec{H}$  verschwindet. Der Grund liegt darin, daß alle Kreise ihren Mittelpunkt auf der  $z$ -Achse haben, die wir in der obigen Rechnung ausschließen mußten: Die Rotation in einem Punkt  $\mathbf{x}$  mißt die Rotation um  $\mathbf{x}$ , und beim betrachteten Feld rotiert eben alles um die  $z$ -Achse, so daß die Rotation überall sonst verschwindet.

**f) Allgemeine Rechenregeln**

Nachdem wir gerade mit relativ großem Aufwand für zwei Beispiele Divergenz und Rotation berechnet haben, wollen wir nun einige Rechenregeln herleiten, die uns unter anderem zeigen werden, daß zumindest ein Teil der obigen Rechnungen hätte vermieden werden können.

Wir wollen uns als nächstes überlegen, was passiert, wenn wir zwei der drei Operatoren grad, div und rot hintereinanderausführen – sofern dies möglich ist.

Der Gradient einer skalaren Funktion  $f: D \rightarrow \mathbb{R}$  auf einer offenen Teilmenge  $D \subseteq \mathbb{R}^n$  ist ein Vektorfeld; es ist also möglich, dessen Divergenz und Rotation zu berechnen. Um keine Probleme mit der Existenz und der Reihenfolge von Ableitungen zu haben, setzen wir voraus, daß  $f$  in  $C^2(D, \mathbb{R})$  liegt; dann ist

$$\operatorname{div} \operatorname{grad} f = \operatorname{div} \begin{pmatrix} \frac{\partial f}{\partial x_1} \\ \vdots \\ \frac{\partial f}{\partial x_n} \end{pmatrix} = \frac{\partial^2 f}{\partial x_1^2} + \dots + \frac{\partial^2 f}{\partial x_n^2}.$$

Die rechte Seite schreiben wir auch kurz als

$$\Delta f \stackrel{\text{def}}{=} \frac{\partial^2 f}{\partial x_1^2} + \dots + \frac{\partial^2 f}{\partial x_n^2}$$

und bezeichnen

$$\Delta \stackrel{\text{def}}{=} \frac{\partial^2}{\partial x_1^2} + \dots + \frac{\partial^2}{\partial x_n^2}$$

als LAPLACE-Operator.

(Der LAPLACE-Operator ist älter als Divergenz, Gradient, Rotation und so weiter; er spielt eine große Rolle sowohl bei Wellen als auch bei Potentialfeldern. HAMILTON wählte das Zeichen  $\nabla$  in Anlehnung an diesen Operator; formal kann man  $\Delta = \nabla \cdot \nabla$  schreiben.)

Auch die Rotation eines Gradientenfelds läßt sich leicht berechnen, allerdings natürlich nur im  $\mathbb{R}^3$ , da wir für andere Dimensionen keine

Rotation definiert haben.

$$\operatorname{rot} \operatorname{grad} f = \operatorname{rot} \begin{pmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \\ \frac{\partial f}{\partial z} \end{pmatrix} = \begin{pmatrix} \frac{\partial^2 f}{\partial z \partial y} - \frac{\partial^2 f}{\partial y \partial z} \\ \frac{\partial^2 f}{\partial x \partial z} - \frac{\partial^2 f}{\partial z \partial x} \\ \frac{\partial^2 f}{\partial y \partial x} - \frac{\partial^2 f}{\partial x \partial y} \end{pmatrix} = \vec{0}.$$

Die Rotation eines stetig differenzierbaren Gradientenfelds ist also stets Null; da das elektrische Feld der negative Gradient des elektrischen Potentials ist, hätten wir oben also für

$$\vec{E}(\mathbf{x}) = \frac{1}{4\pi\epsilon_0} \frac{\vec{x}}{\|\vec{x}\|^3} = -\operatorname{grad} \frac{1}{4\pi\epsilon_0 \|\vec{x}\|}$$

auf die Berechnung der Rotation verzichten können.

Die Divergenz eines Vektorfelds ist eine skalare Funktion; der einzige Operator, der darauf angewandt werden kann, ist daher der Gradient, und das Ergebnis ist völlig uninteressant: siehe Frage 5) auf dem zehnten Übungsblatt.

Die Rotation eines Vektorfelds im Dreidimensionalen schließlich ist wieder ein Vektorfeld, wir können also Divergenz und Rotation davon berechnen:

$$\begin{aligned} \operatorname{div} \operatorname{rot} \vec{V} &= \operatorname{div} \begin{pmatrix} \frac{\partial V_3}{\partial y} - \frac{\partial V_2}{\partial z} \\ \frac{\partial V_1}{\partial z} - \frac{\partial V_3}{\partial x} \\ \frac{\partial V_2}{\partial x} - \frac{\partial V_1}{\partial y} \end{pmatrix} \\ &= \frac{\partial^2 V_3}{\partial x \partial y} - \frac{\partial^2 V_2}{\partial x \partial z} + \frac{\partial^2 V_1}{\partial y \partial z} - \frac{\partial^2 V_3}{\partial y \partial x} + \frac{\partial^2 V_2}{\partial z \partial x} - \frac{\partial^2 V_1}{\partial z \partial y} = 0 \end{aligned}$$

für ein zweimal stetig differenzierbares Vektorfeld. Da ein Magnetfeld als Rotation eines Vektorpotentials geschrieben werden kann, erklärt dies, warum im obigen Beispiel  $\operatorname{div} \vec{H} = 0$  war.

Bleibt noch die Rotation der Rotation; für diese ist nach Aufgabe 2 des zehnten Übungsblatts

$$\operatorname{rot} \operatorname{rot} \vec{V} = \operatorname{grad} \operatorname{div} \vec{V} - \begin{pmatrix} \Delta V_1 \\ \Delta V_2 \\ \Delta V_3 \end{pmatrix}.$$

Ebenfalls dort sind auch die beiden „Produktregeln“

$$\operatorname{div}(f\vec{V}) = f \operatorname{div} \vec{V} + (\operatorname{grad} f) \cdot \vec{V}$$

und

$$\operatorname{rot}(f\vec{V}) = f \operatorname{rot} \vec{V} + (\operatorname{grad} f) \times \vec{V}$$

zu finden. Mit dem Operator  $\nabla$  geschrieben werden sie völlig analog zur klassischen LEIBNIZ-Regel:

$$\nabla \cdot (f\vec{V}) = f \nabla \cdot \vec{V} + \nabla f \cdot \vec{V}$$

und

$$\nabla \times (f\vec{V}) = f \nabla \times \vec{V} + \nabla f \times \vec{V}$$

Speziell für ein lineares Vektorfeld

$$\vec{V}: \mathbb{R}^n \rightarrow \mathbb{R}^n; \quad \mathbf{x} \mapsto A\vec{x} \quad \text{mit} \quad A \in \mathbb{R}^{n \times n}$$

ist

$$\operatorname{div} \vec{V} = \sum_{i=1}^n \frac{\partial V_i}{\partial x_i} = \sum_{i=1}^n a_{ii} = \operatorname{Spur} A$$

und, im Falle  $n = 3$ ,

$$\operatorname{rot} \vec{V} = \begin{pmatrix} a_{32} - a_{23} \\ a_{13} - a_{31} \\ a_{21} - a_{12} \end{pmatrix};$$

für eine quasilineare Funktion  $\vec{f}(\mathbf{x}) = \varphi(\mathbf{x})A\vec{x}$  ist also nach den beiden Produktregeln

$$\operatorname{div} \vec{f} = \varphi \operatorname{Spur} A + (\operatorname{grad} \varphi) \cdot A\vec{x}$$

und

$$\operatorname{rot} \vec{f} = \varphi \cdot \begin{pmatrix} a_{32} - a_{23} \\ a_{13} - a_{31} \\ a_{21} - a_{12} \end{pmatrix} + (\operatorname{grad} \varphi) \times A\vec{x},$$

was uns in den Beispielen des vorigen Abschnitts manche Rechnung erspart hätte.

### g) Nichtkartesische Koordinatensysteme

Auch in diesem letzten Abschnitt von §2 soll es um Möglichkeiten gehen, wie wir bei den oben betrachteten und anderen Beispielen effizienter rechnen können. Das Feld einer Punktladung beispielsweise ist kugelsymmetrisch, jedoch haben wir bei der Berechnung von Divergenz und Rotation dieses Feldes die Kugelsymmetrie mutwillig zerstört, indem wir in einem kartesischen Koordinatensystem rechneten. Hier wollen wir uns beschäftigen, wie man Koordinatensysteme finden kann, die einer solchen Symmetrie angepaßt sind.

Entsprechend der Vielzahl möglicher Symmetrien kennt die Mathematik auch eine Vielzahl solcher Koordinatensysteme; wir wollen uns auf die beiden wichtigsten und am häufigsten angewandten beschränken: die Kugel- und die Zylinderkoordinaten. Zum Einstieg in das Thema betrachten wir aber zunächst die etwas einfachere Situation der Polarkoordinaten in der Ebene.

**1) Polarkoordinaten in  $\mathbb{R}^2$ :** Ein Punkt  $P \in \mathbb{R}^2$  wird charakterisiert durch seinen Abstand  $r$  vom Nullpunkt  $O$  sowie den Winkel  $\varphi$  zwischen dem Vektor  $\vec{OP}$  und dem Einheitsvektor der (positiven)  $x$ -Achse, wobei dieser Winkel für  $P = O$  nicht definiert ist. Ansonsten messen wir ihn im Bogenmaß und wählen den Drehsinn so, die Richtung von der positiven  $x$ -Achse zur positiven  $y$ -Achse positiv gerechnet wird.

Für einen Punkt mit kartesischen Koordinaten  $(x, y)$  ist dann

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

und für  $(x, y) \neq (0, 0)$  ist  $\varphi$  so bestimmt, daß gilt

$$x = r \cos \varphi \quad \text{und} \quad y = r \sin \varphi.$$

$\varphi$  ist dabei, wie üblich, natürlich nur modulo  $2\pi$  bestimmt. Da Sinus und Cosinus nur auf Intervallen der Länge  $\pi$  injektiv sind, reicht eine der beiden Umkehrfunktionen allein nicht aus, um  $\varphi$  direkt anzugeben, sondern man braucht noch mindestens eine Fallunterscheidung. Eine mögliche Formel für  $\varphi$ , auf Basis des Arkuscosinus, ist etwa

$$\varphi = \begin{cases} \arccos \frac{x}{r} & \text{falls } y \geq 0 \\ -\arccos \frac{x}{r} & \text{falls } y < 0. \end{cases}$$

Natürlich kann man entsprechende Formeln auch mit dem Arkussinus oder Arkustangens aufstellen.

Anstelle des Gitternetzes eines kartesischen Koordinatensystems hat man bei den Polarkoordinaten ein Netz aus Kreisen um den Nullpunkt, auf denen der Radius  $r$  konstant ist, und aus vom Nullpunkt ausgehenden Strahlen, auf denen  $\varphi$  konstant ist.

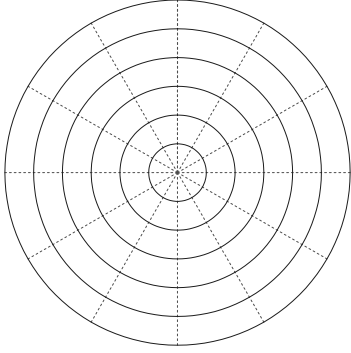


Abb. 42: Das Netz der Polarkoordinaten

Für eine Funktion  $f(x, y)$  auf einer Teilmenge des  $\mathbb{R}^2$  können wir die Polarkoordinatendarstellung

$$F(r, \varphi) \stackrel{\text{def}}{=} f(r \cos \varphi, r \sin \varphi)$$

betrachten; wir wollen sehen, wie man mit  $F$  allein immer noch den Gradienten oder beispielsweise auch die LAPLACESche  $\Delta f$  berechnen kann.

Dazu müssen wir zunächst die partiellen Ableitungen von  $f$  nach  $x$  und  $y$  durch die von  $F$  nach  $r$  und  $\varphi$  ausdrücken. Dies ist auf direkte Weise etwas schwierig, da wir keine gute Formel für die Berechnung von  $\varphi$  aus  $x$  und  $y$  haben. Die umgekehrte Aufgabe dagegen ist umso

einfacher: Nach der Kettenregel ist

$$\frac{\partial F}{\partial r} = \frac{\partial}{\partial r} f(r \cos \varphi, r \sin \varphi) = \frac{\partial f}{\partial x} \cos \varphi + \frac{\partial f}{\partial y} \sin \varphi$$

und

$$\frac{\partial F}{\partial \varphi} = \frac{\partial}{\partial \varphi} f(r \cos \varphi, r \sin \varphi) = -\frac{\partial f}{\partial x} r \sin \varphi + \frac{\partial f}{\partial y} r \cos \varphi.$$

Da  $r > 0$  ist, können wir dies auch in der Form

$$F_r = f_x \cos \varphi + f_y \sin \varphi$$

$$\frac{F_\varphi}{r} = -f_x \sin \varphi + f_y \cos \varphi$$

schreiben. Dies ist ein lineares Gleichungssystem für  $f_x$  und  $f_y$ ; wenn wir von der mit  $\cos \varphi$  multiplizierten ersten Gleichung die mit  $\sin \varphi$  multiplizierte zweite subtrahieren, erhalten wir

$$f_x = F_r \cos \varphi - \frac{F_\varphi}{r} \sin \varphi;$$

entsprechend erhalten wir

$$f_y = F_r \sin \varphi + \frac{F_\varphi}{r} \cos \varphi,$$

wenn wir die mit  $\sin \varphi$  multiplizierte erste Gleichung zur mit  $\cos \varphi$  multiplizierten zweiten addieren.

Entsprechend können wir auch die zweiten partiellen Ableitungen von  $f$  durch partielle Ableitungen von  $F$  ausdrücken, allerdings werden hier die Ausdrücke schon deutlich komplexer, so daß hier auf die detaillierte Rechnung verzichtet sei. Wichtig ist vor allem der LAPLACE-Operator, und wenn es auch sehr umständlich wäre, dessen Darstellung in Polarkoordinaten *herzuleiten* können wir die fertige Formel

$$\Delta f = f_{xx} + f_{yy} = F_{rr} + \frac{F_r}{r} + \frac{F_{\varphi\varphi}}{r^2},$$

doch leicht und relativ schnell beweisen kann, indem wir

$$F_{rr} = (f_{xx} \cos \varphi + f_{xy} \sin \varphi) \cos \varphi + (f_{xy} \cos \varphi + f_{yy} \sin \varphi) \sin \varphi$$

und

$$F_{\varphi\varphi} = (f_{xx}r \sin \varphi - f_{xy}r \cos \varphi)r \sin \varphi - f_x r \cos \varphi - (f_{xy}r \sin \varphi - f_{yy}r \cos \varphi)r \cos \varphi - f_y r \sin \varphi$$

nach der Kettenregel berechnen und einsetzen.

**2) Zylinderkoordinaten im  $\mathbb{R}^3$ :** Für den  $\mathbb{R}^3$  gibt es zwei naheliegende Verallgemeinerung der Polarkoordinaten: Einmal können wir in der  $(x, y)$ -Ebene Polarkoordinaten einführen, die  $z$ -Koordinate aber unverändert lassen, zum anderen können wir neben dem Abstand vom Nullpunkt zwei Winkelkoordinaten einführen, wie z.B. die geographische Länge und Breite. Hier soll es zunächst um den ersten Fall gehen, wir haben also drei Koordinaten  $(r, \varphi, z)$  mit  $r > 0$ , die über die Gleichungen

$$x = r \cos \varphi, \quad y = r \sin \varphi \quad \text{und} \quad z = z$$

mit den kartesischen Koordinaten verbunden sind. Diese Koordinaten heißen *Zylinderkoordinaten*, da die Flächen  $r = \text{konstant}$  Zylinder sind. Insbesondere ist die Koordinate  $r$  hier nicht mehr gleich dem Abstand eines Punktes vom Nullpunkt, sondern gleich dem Abstand von der  $z$ -Achse.

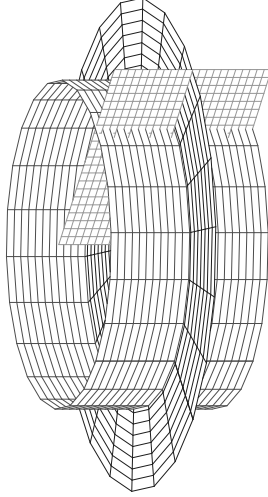


Abb. 43: Flächen mit konstantem  $r, \varphi$  und  $z$  für Zylinderkoordinaten

Da die  $z$ -Koordinate des kartesischen Koordinatensystems übernommen wird, brauchen wir für diese Koordinaten nicht neu zu rechnen, sondern

können die obigen Ergebnisse praktisch wörtlich übernehmen: Mit

$$F(r, \varphi, z) = f(r \cos \varphi, r \sin \varphi, z)$$

ist

$$f_x = F_r \cos \varphi - \frac{F_\varphi}{r} \sin \varphi$$

$$f_y = F_r \sin \varphi + \frac{F_\varphi}{r} \cos \varphi$$

$$f_z = F_z$$

und

$$\Delta f = f_{xx} + f_{yy} + f_{zz} = F_{rr} + \frac{F_r}{r} + \frac{F_{\varphi\varphi}}{r^2} + F_{zz}.$$

**3) Kugelkoordinaten:** Hier führen wir zusätzlich zu den beiden Polarkoordinaten der Ebenen als dritte Koordinate noch eine *Winkelkoordinate*  $\vartheta$  ein, wir haben also drei Koordinaten  $(r, \varphi, \vartheta)$ , wobei  $r$  wieder positiv sein muß und den Abstand vom Nullpunkt bezeichnet.

Im Gegensatz zur Konvention der Geographen, wonach der Äquator, d.h. der Schritzkreis einer Kugel um den Nullpunkt mit der  $(x, y)$ -Ebenen, die Breite Null hat, ist es in der Mathematik, Physik und Technik üblich, dem Äquator die Koordinate  $\vartheta = \pi/2$  zu geben, dem „Nordpol“, d.h. dem Schnittpunkt mit der positiven  $z$ -Achse,  $\vartheta = 0$  und dem Südpol  $\vartheta = \pi$ . In der  $(x, z)$ -Ebene ist also

$$x = r \sin \vartheta \quad \text{und} \quad z = r \cos \vartheta,$$

in der  $(y, z)$ -Ebene

$$y = r \sin \vartheta \quad \text{und} \quad z = r \cos \vartheta,$$

und ganz allgemein ist

$$R = r \sin \vartheta \quad \text{und} \quad z = r \cos \vartheta,$$

wobei  $R$  den Abstand eines Punktes von der  $z$ -Achse bezeichnet.

Für einen Punkt im Abstand  $R$  von der  $z$ -Achse ist

$$x = R \cos \varphi \quad \text{und} \quad y = R \sin \varphi,$$

zusammen mit den obigen Formeln erhalten wir also als Zusammenhang zwischen Kugelkoordinaten und kartesischen Koordinaten

$$x = r \cos \varphi \sin \vartheta$$

$$y = r \sin \varphi \sin \vartheta$$

$$z = r \cos \vartheta.$$

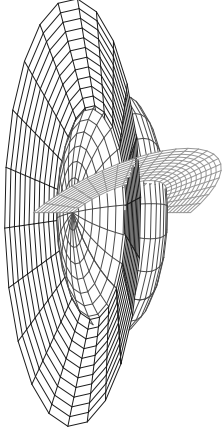


Abb. 44: Flächen mit konstantem  $r$ ,  $\varphi$  und  $\vartheta$  für Kugelkoordinaten

Die Flächen mit konstantem  $r$  sind hier natürlich Kugeln, konstantes  $\varphi$  führt wie bei den Zylinderkoordinaten auf Halbebenen, und die Flächen mit konstantem  $\vartheta$  schließlich sind Kegel um die  $z$ -Achse.

Wieder können wir einer Funktion  $f(x, y, z)$  in kartesischen Koordinaten durch

$$F(r, \varphi, \vartheta) \stackrel{\text{def}}{=} f(r \cos \varphi \sin \vartheta, r \sin \varphi \sin \vartheta, r \cos \vartheta)$$

eine neue Funktion zuordnen und durch Anwendung der Kettenregel so wie Lösen eines linearen Gleichungssystems die partiellen Ableitungen zueinander in Beziehung setzen; die Vorgehensweise ist im wesentlichen wie bei den ebenen Polarkoordinaten, nur daß wir jetzt ein System aus drei linearen Gleichungen bekommen:

$$F_r = f_x \cos \varphi \sin \vartheta + f_y \sin \varphi \sin \vartheta + f_z \cos \vartheta$$

$$\frac{F_\varphi}{r} = -f_x \sin \varphi \sin \vartheta + f_y \cos \varphi \sin \vartheta$$

$$\frac{F_\vartheta}{r} = f_x \cos \varphi \cos \vartheta + f_y \sin \varphi \cos \vartheta - f_z \sin \vartheta.$$

Der GAUSS-Algorithmus führt schnell auf die Lösung

$$f_x = F_r \cos \varphi \sin \vartheta - \frac{F_\varphi \sin \varphi}{r \sin \vartheta} + \frac{F_\vartheta \cos \varphi \cos \vartheta}{r}$$

$$f_y = F_r \sin \varphi \sin \vartheta - \frac{F_\varphi \cos \varphi}{r \sin \vartheta} + \frac{F_\vartheta \sin \varphi \cos \vartheta}{r}$$

$$f_z = F_r \cos \vartheta - \frac{F_\vartheta \sin \vartheta}{r}.$$

Im Prinzip genauso können auch die zweiten Ableitungen berechnet werden; eine grausame, wenn auch nicht sonderlich schwierige Rechnung führt beispielsweise auf die Formel

$$\begin{aligned} \Delta f &= f_{xx} + f_{yy} + f_{zz} \\ &= F_{rr} + \frac{2F_r}{r} + \frac{F_{\varphi\varphi}}{r^2 \sin^2 \vartheta} + \frac{F_{\vartheta\vartheta}}{r^2} + \frac{F_\vartheta}{r^2} \tan \vartheta. \end{aligned}$$

Als Beispiel wollen wir damit noch einmal die Divergenz des elektrischen Felds einer Ladung  $q$  im Nullpunkt berechnen: Diese erzeugt ein Potential

$$U(r, \varphi, \vartheta) = \frac{q}{4\pi\epsilon_0 r},$$

das Feld ist  $\vec{E} = -\text{grad } U$  und

$$\text{div } \vec{E} = -\text{div grad } U = -\Delta U = -\frac{q}{4\pi\epsilon_0} \left( \frac{2}{r^3} + \frac{2}{r} \cdot \frac{-1}{r^2} \right) = 0.$$

### §3: Integralrechnung

Die Integralrechnung ist neben der Differentialrechnung die zweite wichtige Säule der Analysis. Wir betrachten zunächst den eindimensionalen Fall, der sich auch für die mehrdimensionale Theorie an mehreren Stellen als wichtig erweisen wird.

#### a) Heuristische Vorüberlegungen

Seiner großen Bedeutung entsprechend, wollen wir uns dem Integralbegriff zunächst heuristisch nähern, bevor wir ihn – mit beträchtlichem technischem Aufwand – im nächsten Paragraphen formal einführen.

Die Integration dient primär drei Zwecken:

- Sie dient zur Umkehrung der Differentiation.
- Sie dient zur Flächenberechnung.
- Sie dient zur Durchschnittsberechnung.

Wie wollen uns alle drei Aspekte kurz ansehen.

**1) Integration als Umkehrung der Differentiation:** Das klassische Beispiel zur Einführung der Differentiation ist die Geschwindigkeit als Ableitung des Wegs nach der Zeit. Es liegt daher nahe, zur Einführung des Integrals die Frage nach der Berechnung des Wegs aus der Geschwindigkeit zu betrachten.

Wir nehmen also an, ein Fahrzeug sei zur Zeit  $t = 0$  an der Stelle  $s = 0$  und beschleunige in fünf Sekunden auf die im Stadtverkehr zulässige Höchstgeschwindigkeit von 13,9 m/sec. Welchen Weg hat es bis dahin zurückgelegt?

Die Geschwindigkeit sei durch folgende Tabelle gegeben:

|             |     |     |     |      |      |      |
|-------------|-----|-----|-----|------|------|------|
| $t$ [sec]   | 0   | 1   | 2   | 3    | 4    | 5    |
| $v$ [m/sec] | 2,5 | 6,1 | 9,1 | 11,5 | 13,4 | 13,9 |

Ist  $s(t)$  der bis zum Zeitpunkt  $t$  zurückgelegte Weg und  $v(t)$  die dann erreichte Geschwindigkeit, so ist bekanntlich

$$v(t) = \frac{ds}{dt} \approx \frac{\Delta s}{\Delta t} = \frac{s(t+1 \text{ sec}) - s(t)}{(t+1 \text{ sec}) - t} = \frac{s(t+1 \text{ sec}) - s(t)}{1 \text{ sec}}.$$

Also ist

$$s(t+1 \text{ sec}) \approx s(t) + v(t) \cdot 1 \text{ sec}.$$

Da wir sowohl  $s(0 \text{ sec}) = 0 \text{ m}$  als auch die Werte  $v(t)$  für  $t = 0, 1, 2, 3$  und  $4 \text{ sec}$  kennen, können wir  $s(5 \text{ sec})$  rekursiv berechnen als

$$\begin{aligned} s(5 \text{ sec}) &\approx v(0 \text{ sec}) \cdot 1 \text{ sec} + v(1 \text{ sec}) \cdot 1 \text{ sec} + v(2 \text{ sec}) \cdot 1 \text{ sec} \\ &\quad + v(3 \text{ sec}) \cdot 1 \text{ sec} + v(4 \text{ sec}) \cdot 1 \text{ sec} \\ &= 2,5 \text{ m} + 6,1 \text{ m} + 9,1 \text{ m} + 11,5 \text{ m} + 13,4 \text{ m} = 42,6 \text{ m}. \end{aligned}$$

Also hat das Fahrzeug in fünf Sekunden 42,6 m zurückgelegt?

Zumindest mit dieser Genauigkeit ist das sicherlich falsch, denn bei der Berechnung des Wegs sind wir davon ausgegangen, daß das Fahrzeug während der ersten Sekunde konstant mit einer Geschwindigkeit von 2,5 m/sec gefahren ist, exakt ab deren Ende dann aber plötzlich eine Sekunde lang mit 6,1 m/sec, usw. Das ist natürlich unrealistisch; tatsächlich dürfte die Geschwindigkeit etwa so verlaufen sein, wie es die Kurve in Abbildung 45 angibt, wohingegen wir mit dem ebenfalls eingezeichneten stufenförmigen Verlauf gerechnet haben. Wir haben die Geschwindigkeit somit fast durchgängig unterschätzt und damit auch einen zu kleinen Weg berechnet.

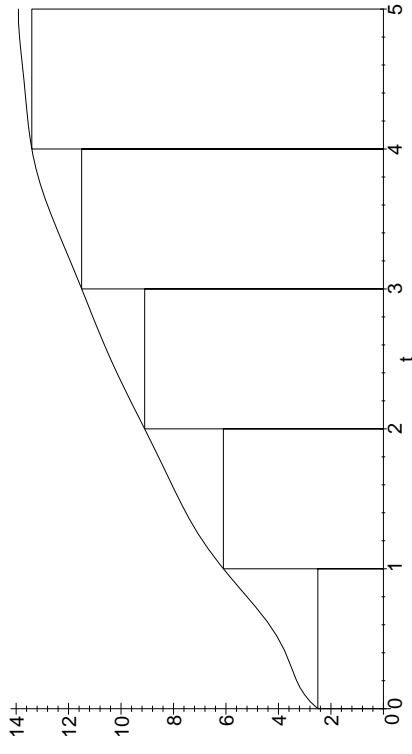


Abb. 45: Tatsächliche und unterschätzte Geschwindigkeit

Alternativ hätten wir auch einen zu großen Weg schätzen können, wenn wir die Geschwindigkeit während eines Sekundenintervalls jeweils auf den bekanntesten Wert am Ende dieses Intervalls gesetzt hätten; das Ergebnis wäre dann gewesen

$$\begin{aligned} s(5 \text{ sec}) &\approx v(1 \text{ sec}) \cdot 1 \text{ sec} + v(2 \text{ sec}) \cdot 1 \text{ sec} + v(3 \text{ sec}) \cdot 1 \text{ sec} \\ &\quad + v(4 \text{ sec}) \cdot 1 \text{ sec} + v(5 \text{ sec}) \cdot 1 \text{ sec} \\ &= 6,1 \text{ m} + 9,1 \text{ m} + 11,5 \text{ m} + 13,4 \text{ m} + 13,9 \text{ m} = 54,0 \text{ m}. \end{aligned}$$

Tatsächlich wissen wir im Augenblick also nur, daß der zurückgelegte Weg irgendwo zwischen 42,6 m und 54 m liegt.

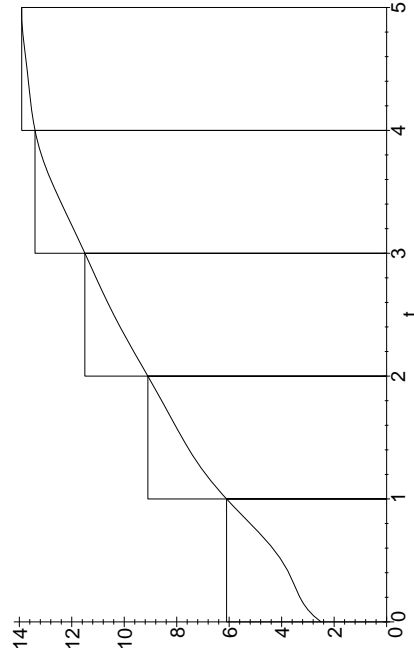


Abb. 46: Tatsächliche und überschätzte Geschwindigkeit

Eine bessere Schätzung bekommen wir, wenn wir anstelle der Sekundenintervalle Intervalle von nur einer halben Sekunde betrachten – vorausgesetzt natürlich, wir kennen die Geschwindigkeit auch für halbzahlige Sekundenwerte. Die entsprechenden Werte seien die in der folgenden Tabelle angegebenen:

|             |     |     |     |     |     |      |      |      |      |      |      |
|-------------|-----|-----|-----|-----|-----|------|------|------|------|------|------|
| $t$ [sec]   | 0   | 0,5 | 1   | 1,5 | 2   | 2,5  | 3    | 3,5  | 4    | 4,5  | 5    |
| $v$ [m/sec] | 2,5 | 4,4 | 6,1 | 7,8 | 9,1 | 10,4 | 11,5 | 12,6 | 13,4 | 13,8 | 13,9 |

Jetzt ist die „unterschätzte“ Wegstrecke

$$\sum_{i=0}^9 v\left(\frac{i}{2} \text{ sec}\right) \cdot \frac{1}{2} \text{ sec} = 45,6 \text{ m},$$

und die „überschätzte“ ist

$$\sum_{i=1}^{10} v\left(\frac{i}{2} \text{ sec}\right) \cdot \frac{1}{2} \text{ sec} = 51,3 \text{ m},$$

die Unsicherheit hat sich also etwa halbiert. Abbildung 47 zeigt den Verlauf von unterschätzter, tatsächlicher und überschätzter Geschwindigkeit für die Halbskundenintervalle.

Beide Werte wie auch ihre Differenz lassen sich anhand von Abbildung 47 leicht geometrisch veranschaulichen: Die unterschätzte Ge-

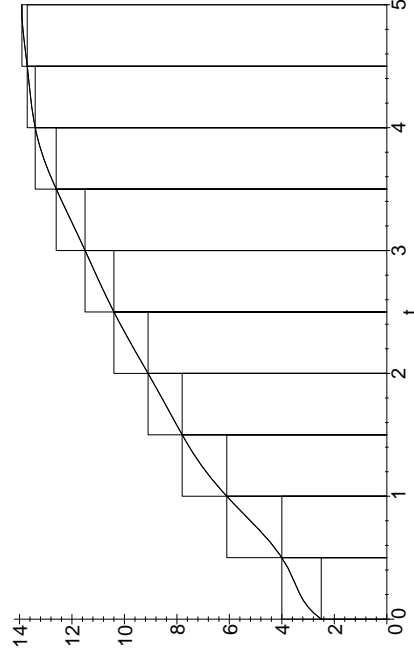


Abb. 47: Abtastung in Intervallen von einer halben Sekunde

schwindigkeit ist die Fläche der Rechtecke unterhalb der unteren Treppenfunktion, die überschätzte entsprechend die Fläche der Rechtecke unterhalb der oberen Treppenfunktion, wobei die Basis der Rechtecke jeweils auf der Zeitachse liegt. Die Differenz ist somit gleich der Fläche der Differenzrechtecke.

Diese können wir durch weitere Verfeinerung der Abtastung verkleinern. Dazu nehmen wir an, die eingezeichnete Kurve (die tatsächlich einfach eine notdürftig angepasste Polynomfunktion darstellt) gebe den tatsächlichen Geschwindigkeitsverlauf wieder, und lassen den Computer rechnen:

Wenn wir die Geschwindigkeit viermal pro Sekunde bestimmen und den zurückgelegten Weg damit schätzen, erhalten wir als untere Schranke

$$\sum_{i=0}^{19} v\left(\frac{i}{4} \text{ sec}\right) \cdot \frac{1}{4} \text{ sec} \approx 47,14 \text{ m}$$

und als obere Schranke

$$\sum_{i=1}^{20} v\left(\frac{i}{4} \text{ sec}\right) \cdot \frac{1}{4} \text{ sec} \approx 49,99 \text{ m}.$$



Mit zehn Geschwindigkeitswerten pro Sekunde verbessert sich die untere Schranke auf

$$\sum_{i=0}^{49} v\left(\frac{i}{10} \text{ sec}\right) \cdot \frac{1}{10} \text{ sec} \approx 48,02 \text{ m}$$

und die obere auf

$$\sum_{i=1}^{50} v\left(\frac{i}{10} \text{ sec}\right) \cdot \frac{1}{10} \text{ sec} \approx 49,16 \text{ m} .$$

Wie Abbildung 48 zeigt, unterscheiden sich nun die Rechtecke der unteren Abschätzung nur noch wenig von denen der oberen, und die Fläche unter der Geschwindigkeitskurve liegt schon recht nahe bei der Fläche der Rechtecke aus jeder der beiden Abschätzungen.

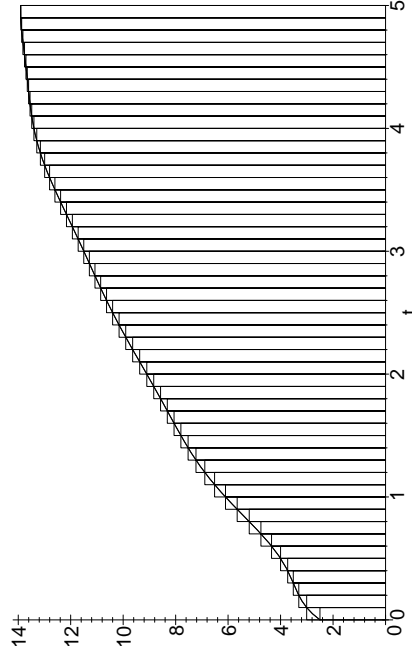


Abb. 48: Abtastung in Intervallen von einer Zehntelsekunde

Gehen wir von den Zehntelsekunden zu den Hundertstel, sind die Rechtecke zumindest visuell in Abbildung 49 nicht mehr voneinander und von der Fläche unter der Kurve zu unterscheiden: Daß man hier keine durchweg schwarze Fläche sieht, liegt ausschließlich an sogenannten *alias*-Effekten, d.h. Diskretisierungsfehlern der Rastergraphik.

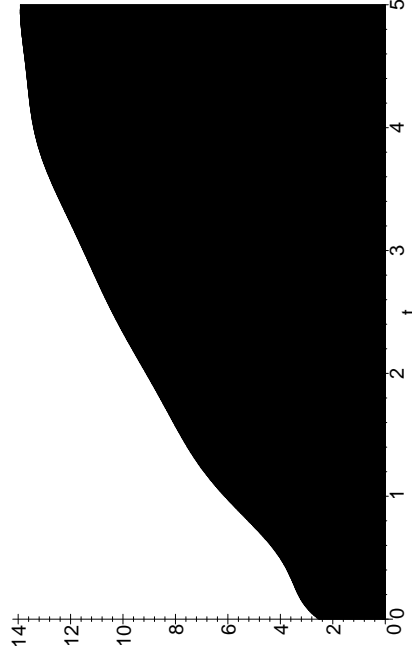


Abb. 49: Abtastung in Intervallen von einer Hundertstelsekunde

Rechnerisch gibt es noch einen kleinen Unterschied im Dezimeterbereich:

$$\sum_{i=0}^{499} v\left(\frac{i}{100} \text{ sec}\right) \cdot \frac{1}{100} \text{ sec} \approx 48,54 \text{ m}$$

und

$$\sum_{i=1}^{500} v\left(\frac{i}{100} \text{ sec}\right) \cdot \frac{1}{100} \text{ sec} \approx 48,66 \text{ m} .$$

Bei einer nochmaligen Verzehnfachung der Intervallanzahl ist natürlich graphisch nichts neues mehr zu sehen; die Schätzwerte verbessern sich auf

$$\sum_{i=0}^{4999} v\left(\frac{i}{1000} \text{ sec}\right) \cdot \frac{1}{1000} \text{ sec} \approx 48,595 \text{ m}$$

für die untere Schranke und

$$\sum_{i=1}^{5000} v\left(\frac{i}{1000} \text{ sec}\right) \cdot \frac{1}{1000} \text{ sec} \approx 48,607 \text{ m}$$

für die obere.

Um die eingangs gestellte Frage nach dem zurückgelegten Weg millimetergenau beantworten zu können (sofern man diese Genauigkeit wirklich als sinnvoll betrachtet), muß die Geschwindigkeitskurve 10 000 Mal pro Sekunde abgetastet werden, und wir erhalten die Schranken

$$\sum_{i=0}^{49999} v \left( \frac{i}{10000} \text{ sec} \right) \cdot \frac{1}{10000} \text{ sec} \approx 48,6006 \text{ m}$$

und

$$\sum_{i=1}^{50000} v \left( \frac{i}{10000} \right) \text{ sec} \cdot \frac{1}{10000} \text{ sec} \approx 48,6017 \text{ m};$$

das Fahrzeug legte in fünf Sekunden also 48 m und 60,1 cm zurück.

**2) Integration als Flächenbestimmung:** Die Abbildungen 48 und vor allem 49 zeigen, daß die immer feiner werdenden Rechtecke, die wir für die obigen Abschätzungen wählten, die Fläche unter der Kurve  $v = v(t)$  immer besser annähern. Wenn wir die Einheiten vergessen und  $t$  wie auch  $v$  als Längen ansehen, können wir die oben betrachteten Ausdrücke

$$\sum_{i=0}^{5n-1} v \left( \frac{i}{n} \right) \cdot \frac{1}{n} \quad \text{und} \quad \sum_{i=1}^{5n} v \left( \frac{i}{n} \right) \cdot \frac{1}{n}$$

also auch als Näherungswerte für die Fläche unter der Kurve ansehen und den Flächeninhalt als ihren gemeinsamen Grenzwert für immer größer werdendes  $n$  berechnen.

Genau genommen handelt es sich hier allerdings nicht um die *Berechnung* dieser Fläche, sondern um die *Definition* des Flächeninhalts, denn es gibt schließlich *a priori* keinen Begriff des Flächeninhalts einer krummlinig begrenzten Fläche. Die hier gewählte Definition über eine Ausschöpfung mit immer feiner werdenden Rechtecken ist zwar sehr natürlich, aber nicht zwangsläufig. Sie ist übrigens weit älter als jeder Begriff eines Integrals oder auch nur Grenzwerts: Bereits vor über zwei Jahrtausenden, um etwa 370 vor Christus, definierte der griechische Mathematiker EUDOXOS VON CNIDOS (ca. 408–ca. 355), ein Schüler und späterer Konkurrent PLATOS, Flächeninhalte auf diese Weise; später hat ARCHIMEDES VON SYRAKUS (287–212) diese Methode perfektioniert

und angewandt auf Kreise und Parabeln; insbesondere berechnete er damit seine Abschätzung  $223/71 < \pi < 22/7$  für  $\pi$ , in Dezimalzahlen ausgedrückt  $3,1408 < \pi < 3,1429$ , wobei diese Abschätzung allerdings nicht auf Rechtecken beruht, sondern auf der Konstruktion des regelmäßigen 96-Ecks.

**3) Integration als Durchschnittbestimmung:** Kehren wir wieder zurück zum Eingangsbeispiel eines sich beschleunigenden Fahrzeugs, und fragen wir uns, wie hoch die *Durchschnittsgeschwindigkeit* war. Bei endlich vielen Werten ist der Durchschnitt natürlich einfach das arithmetische Mittel, also z.B.

$$\begin{aligned} & \frac{1}{5} \cdot (v(0 \text{ sec}) + v(1 \text{ sec}) + v(2 \text{ sec}) + v(3 \text{ sec}) + v(4 \text{ sec})) \\ &= \frac{1}{5} \cdot 42,6 \text{ m/sec} = 8,52 \text{ m/sec}, \end{aligned}$$

wenn wir die Geschwindigkeiten vom *Anfang* jedes Sekundenintervalls nehmen, und

$$\begin{aligned} & \frac{1}{5} \cdot (v(1 \text{ sec}) + v(2 \text{ sec}) + v(3 \text{ sec}) + v(4 \text{ sec}) + v(5 \text{ sec})) \\ &= \frac{1}{5} \cdot 54,0 \text{ m/sec} = 10,8 \text{ m/sec}, \end{aligned}$$

wenn wir die vom Ende nehmen.

Natürlich läßt sich auch hier die Abtastung immer weiter verfeinern; die entstehenden Ausdrücke

$$\frac{1}{5n} \cdot \sum_{i=0}^{5n-1} v \left( \frac{i}{n} \text{ sec} \right) \quad \text{und} \quad \frac{1}{5n} \cdot \sum_{i=1}^{5n} v \left( \frac{i}{n} \text{ sec} \right)$$

entsprechen bis auf einen Faktor von  $\frac{1}{5 \text{ sec}}$  genau denen aus Teil a), wo wir den Weg abgeschätzt haben – wie es ja auch in der Tat der Fall sein muß: Die Durchschnittsgeschwindigkeit ist schließlich nichts anderes als der zurückgelegte Weg dividiert durch die zugrundeliegende Zeitspanne von fünf Sekunden.

Trotzdem ist die Interpretation eines Integrals als Durchschnitt gelegentlich von unabhängigem Interesse, da vor allem bei Anwendungen in der

Quantenphysik oft zwar Durchschnitte existieren, aber keine „Zähler“ und „Nenner“, als deren Quotienten man sie interpretieren könnte.

### b) Integration elementarer Funktionen

Abstrahieren wir vom Beispiel der Wegberechnung anhand einer Geschwindigkeitskurve und betrachten wir das allgemeine Problem, zu einer gegebenen reellwertigen Funktion  $f$  eine neue Funktion  $F$  zu finden derart, daß  $F'(x) = f(x)$  ist.

Da jede konstante Funktion die Ableitung Null hat, ist mit  $F$  für jede reelle Zahl  $c$  auch  $F + c$  eine solche Funktion; um ein konkretes  $F$  hinzuschreiben, müssen wir also einen Funktionswert von  $F$  festlegen; der Einfachheit halber sei dies, in Analogie zum obigen Beispiel, der Wert  $F(0) = 0$ .

Damit übernimmt  $f$  die Rolle der Geschwindigkeit  $v$ , dem Weg entspricht die Funktion  $F$ , die Variable ist nun  $x$  anstelle der Zeit  $t$ , und da uns  $F$  an einer beliebigen Stelle  $x$  interessiert, nicht nur wie im obigen Beispiel an der Stelle 5, empfiehlt es sich,  $n$  jetzt als Gesamtzahl der Intervalle zu nehmen, nicht wie oben als Anzahl der Intervalle pro Einheit (Sekunde). Die Länge eines jeden Intervalls wird dann zu  $x/n$ , und die Analoga zu obigen Summen sind die Ausdrücke

$$\sum_{i=0}^{n-1} f\left(\frac{ix}{n}\right) \cdot \frac{x}{n} \quad \text{und} \quad \sum_{i=1}^n f\left(\frac{ix}{n}\right) \cdot \frac{x}{n},$$

deren gemeinsamer Grenzwert für immer größer werdendes  $n$  der gesuchte Wert  $F(x)$  sein sollte.

Wir definieren daher

$$F(x) = \lim_{\text{def } n \rightarrow \infty} \sum_{i=0}^{n-1} f\left(\frac{ix}{n}\right) \cdot \frac{x}{n} \quad (*)$$

in der Hoffnung, daß dieser Grenzwert existiert und mit dem anderen übereinstimmt – zumindest in hinreichend vielen interessanten Fällen. Bevor wir uns im nächsten Paragraphen dieser Frage zuwenden, wollen wir hier zunächst etwas mit dieser vorläufigen Definition spielen und sehen, was sie bei einfachen Funktionen liefert.

1) Die Funktion  $f(x) = x^2$ : Hier erhalten wir

$$\begin{aligned} F(x) &= \lim_{\text{def } n \rightarrow \infty} \sum_{i=0}^{n-1} f\left(\frac{ix}{n}\right) \cdot \frac{x}{n} = \lim_{n \rightarrow \infty} \sum_{i=0}^{n-1} \left(\frac{ix}{n}\right)^2 \cdot \frac{x}{n} \\ &= x^3 \lim_{n \rightarrow \infty} \frac{1}{n^3} \sum_{i=0}^{n-1} i^2. \end{aligned}$$

Wie man in seiner Formelsammlung nachschlägt oder sich von seinem Computeralgebrasytem sagen läßt, ist

$$\sum_{i=0}^{n-1} i^2 = \frac{n(2n-1)(n-1)}{6},$$

eine Formel, die dann nachträglich leicht durch vollständige Induktion bewiesen werden kann:

Für  $n = 1$  steht links und rechts Null, die Formel ist also korrekt.

Angenommen, wir hätten sie bewiesen für  $n - 1$  anstelle von  $n$ , d.h. wir wüßten, daß für ein spezielles  $n > 1$

$$\sum_{i=0}^{n-2} i^2 = \frac{(n-1)(2n-3)(n-2)}{6}$$

ist. Dann folgt

$$\begin{aligned} \sum_{i=0}^{n-1} i^2 &= \sum_{i=0}^{n-2} i^2 + (n-1)^2 \\ &= \frac{(n-1)(2n-3)(n-2)}{6} + (n-1)^2 \\ &= \frac{n-1}{6} \cdot ((2n-3)(n-2) + 6(n-1)) \\ &= \frac{n-1}{6} \cdot (2n^2 - 7n + 6 + 6n - 6) \\ &= \frac{n-1}{6} \cdot (2n^2 - n) = \frac{(n-1)(2n-1)n}{6} \\ &= \frac{n(2n-1)(n-1)}{6}, \end{aligned}$$

d.h. die Formel für  $n$  selbst.

Nach dem Prinzip der vollständigen Induktion ist diese damit für *alle* Werte von  $n$  bewiesen; wir können sie in obige Rechnung einsetzen und erhalten

$$\begin{aligned} F(x) &= x^3 \lim_{n \rightarrow \infty} \frac{1}{n^3} \cdot \frac{n(2n-1)(n-1)}{6} = x^3 \lim_{n \rightarrow \infty} \frac{n(2n-1)(n-1)}{6n^3} \\ &= \frac{x^3}{6} \lim_{n \rightarrow \infty} \left(\frac{n}{n}\right) \left(\frac{2n-1}{n}\right) \left(\frac{n-1}{n}\right) = \frac{x^3}{6} \cdot 2 = \frac{x^3}{3}. \end{aligned}$$

Diese Funktion hat tatsächlich die Ableitung  $F'(x) = x^2$ , zumindest in diesem Beispiel funktioniert also die Definition (\*).

2) **Die Exponentialfunktion:** Versuchen wir dasselbe nochmal für die Funktion  $f(x) = e^x$ . Hier führt (\*) (wieder mit der Festlegung  $F(0) = 0$ ) auf den Ansatz

$$\begin{aligned} F(x) &= \lim_{n \rightarrow \infty} \sum_{i=0}^{n-1} f\left(\frac{ix}{n}\right) \cdot \frac{x}{n} = \lim_{n \rightarrow \infty} \sum_{i=0}^{n-1} e^{\frac{ix}{n}} \cdot \frac{x}{n} \\ &= x \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=0}^{n-1} \left(e^{\frac{x}{n}}\right)^i. \end{aligned}$$

Die Summe ganz rechts ist eine geometrische Reihe und kann leicht ausgerechnet werden: Setzen wir zur Abkürzung  $q = e^{x/n}$ , so ist

$$\begin{aligned} \sum_{i=0}^{n-1} q^i &= 1 + q + q^2 + q^3 + \dots + q^{n-1} \\ q \sum_{i=0}^{n-1} q^i &= q + q^2 + q^3 + \dots + q^{n-1} + q^n \\ \hline \Rightarrow (1-q) \sum_{i=0}^{n-1} q^i &= 1 - q^n. \end{aligned}$$

Für  $q \neq 1$  können wir durch  $1 - q$  dividieren und erhalten die Formel

$$\sum_{i=0}^{n-1} q^i = \frac{1 - q^n}{1 - q}.$$

Da uns nur der Fall  $x \neq 0$  interessiert, ist  $q = e^{\frac{x}{n}} \neq 1$ ; wir können die Formel also anwenden und erhalten

$$\begin{aligned} F(x) &= x \lim_{n \rightarrow \infty} \frac{1}{n} \cdot \frac{1 - \left(e^{\frac{x}{n}}\right)^n}{1 - e^{\frac{x}{n}}} = x \lim_{n \rightarrow \infty} \frac{1}{n} \cdot \frac{1 - e^x}{1 - e^{\frac{x}{n}}} \\ &= x(1 - e^x) \lim_{n \rightarrow \infty} \frac{1}{n} \cdot \frac{x}{1 - e^{\frac{x}{n}}}. \end{aligned}$$

Im Grenzwert ganz rechts gehen für  $n \rightarrow \infty$  leider sowohl der Zähler als auch der Nenner gegen Null, wir müssen also die Regel von DE L'HOSPITAL anwenden. Diese gilt zwar nur für Grenzwerte von Werten stetiger Funktionen, während wir hier den Grenzwert einer diskreten Folge suchen, aber wenn für eine reelle Variable  $u$  der Grenzwert

$$\lim_{u \rightarrow \infty} \frac{\frac{1}{u}}{1 - e^{\frac{x}{u}}}$$

existiert, konvergiert natürlich auch die Folge der Zahlen

$$\frac{1}{1 - e^{\frac{x}{n}}} \quad \text{für } n = 1, 2, 3, \dots$$

gegen eben diesen Grenzwert.

Nach DE L'HOSPITAL ist

$$\begin{aligned} \lim_{u \rightarrow \infty} \frac{\frac{1}{u}}{1 - e^{\frac{x}{u}}} &= \lim_{u \rightarrow \infty} \frac{\frac{d}{du} \frac{1}{u}}{\frac{d}{du} (1 - e^{\frac{x}{u}})} = \lim_{u \rightarrow \infty} \frac{-\frac{1}{u^2}}{-e^{\frac{x}{u}} \left(\frac{-x}{u^2}\right)} \\ &= \lim_{u \rightarrow \infty} \frac{-1}{x e^{\frac{x}{u}}} = \frac{-1}{x}. \end{aligned}$$

Eingesetzt in den Ansatz für  $F(x)$  führt dies auf

$$F(x) = x(1 - e^x) \cdot \frac{-1}{x} = e^x - 1,$$

eine Funktion, deren Ableitung in der Tat  $e^x$  ist.

3) **Die Dirichletsche Sprungfunktion:** Bevor wir zu übermütig werden, möchte ich als letztes Beispiel noch eine Funktion betrachten, die sich

im Gegensatz zu Quadrat und Exponentialfunktion alles andere als gut verhält, die DIRICHLETSche Sprungfunktion

$$f(x) = \begin{cases} 1 & \text{für } x \in \mathbb{Q} \\ 0 & \text{für } x \notin \mathbb{Q} \end{cases}$$

Mit  $F(0) = 0$  ist wieder

$$F(x) = \lim_{\text{def } n \rightarrow \infty} \sum_{i=0}^{n-1} f\left(\frac{ix}{n}\right) \cdot \frac{x}{n}.$$

Dabei ist die Zahl  $ix/n$  für  $i = 0$  rational, da Null; für  $i \neq 0$  ist sie genau dann rational, wenn auch  $x$  rational ist. In diesem Fall ist also  $f(ix/n) = 1$  für alle  $i$ , d.h.

$$F(x) = \lim_{n \rightarrow \infty} \sum_{i=0}^{n-1} \frac{x}{n} = \lim_{n \rightarrow \infty} n \cdot \frac{x}{n} = x$$

für rationales  $x$ . Für irrationales  $x$  dagegen sind alle  $ix/n$  außer der Null irrational, also ist  $f(ix/n) = 0$  für alle  $i \neq 0$ . Damit bleibt von der Summe nur der erste Summand stehen, d.h.

$$F(x) = \lim_{n \rightarrow \infty} \frac{x}{n} = 0.$$

Also ist

$$F(x) = \begin{cases} x & \text{für } x \in \mathbb{Q} \\ 0 & \text{für } x \notin \mathbb{Q} \end{cases},$$

eine offensichtlich nicht differenzierbare Funktion, und wenn  $F'(x)$  nicht einmal existiert, kann es natürlich auch nicht gleich  $f(x)$  sein.

Auch unter dem Gesichtspunkt der Interpretation von  $F(x)$  als Fläche unterhalb der Kurve  $y = f(x)$  ist das Ergebnis unbefriedigend, denn da  $f(x) \geq 0$  für alle  $x$ , sollte  $F(x)$  eine monoton wachsende (oder zumindest nicht fallende) Funktion von  $x$  sein, wohingegen das gerade berechnete  $F$  bei jeder rationalen Zahl auf Null zurückfällt.

Der Ansatz (\*) hat also auch seine Tücken, und es wird Zeit, diese heuristische Definition durch eine bessere zu ersetzen.

### c) Definition des Riemann-Integrals

Tatsächlich ist der Ansatz (\*) nicht so schlecht: Für die beiden gutartigen Funktionen und für das Eingangsbeispiel der Streckenbestimmung führte er schließlich zu vernünftigen Ergebnissen, und wie wir bald sehen werden, kann man Integrale über stetige Funktionen *immer* nach diesem Ansatz bestimmen; zumindest als Veranschaulichung des Integralbegriffs sollte man ihn daher durchaus im Hinterkopf behalten. Problematisch wird er erst bei „schlechten“ Funktionen wie der im letzten Beispiel.

**1) Warum lohnt sich ein allgemeinerer Ansatz?:** Dies legt natürlich die Frage nahe, ob man solche „schlechten“ Funktionen für Anwendungen wirklich braucht, und wenn ja, ob man sie so sehr braucht, daß sich der erhebliche technische Aufwand, den wir in diesem Paragraphen treiben müssen, lohnt.

Die DIRICHLETSche Sprungfunktion ist ein rein theoretisch konstruiertes Gegenbeispiel, das wohl keinerlei praktische Anwendung haben dürfte. Eine ihrer charakteristischen Eigenschaften taucht allerdings auch bei praktisch relevanten Funktionen auf: Auch ein periodisch wiederholter Rechteckimpuls hat abzählbar unendlich viele Sprungstellen, und zumindest als Idealisierung eines realen Signals spielt diese Funktion eine Rolle – sie ist allerdings erheblich harmloser als DIRICHLETS Beispiel, da die Menge ihrer Sprungstellen diskret ist.

Wirklich kompliziert wird die Situation dagegen beispielsweise bei der mathematischen Modellierung des Rauschens in einer elektronischen Schaltung: Hier liegt eine noch einmal deutlich schwierigere Situation vor als im obigen Beispiel, und in der Tat wird auch der in diesem Paragraphen definierte Integralbegriff nicht ausreichen, um mit diesem Problem fertig zu werden. Er ist aber eine unabdingbare Voraussetzung, um die dazu benötigten Techniken zu verstehen.

Obwohl wir im folgenden fast jeden der wichtigeren Sätze aus der Analysis I noch einmal in die Erinnerung zurückrufen und anwenden müssen, ist die nun folgende Konstruktion also kein Luxus, sondern zumindest langfristig notwendig auch für praktische Anwendungen.

2) **Wo sollte der bisherige Ansatz modifiziert werden?**: Ein Punkt, bei dem der bisherige Ansatz nur aufgrund der Einfachheit der betrachteten Beispiele so erfolgreich war, ist die *Abschätzung* des Grenzwerts durch die endlichen Approximationen. Beim Beispiel der Geschwindigkeit eines beschleunigenden Fahrzeugs war das problemlos, denn die Geschwindigkeit war eine monoton wachsende Funktion, die für jedes Teilintervall am linken Ende ihr Minimum und am rechten ihr Maximum annimmt.

Auch  $f(x) = x^2$  ist für  $x \geq 0$  monoton wachsend,  $f(x) = e^x$  sogar für alle  $x \in \mathbb{R}$ , so daß wir auch hier leicht untere und obere Schranken für  $F(x)$  berechnen können.

Für nichtmonotone Funktionen dagegen ist die Situation völlig anders: Betrachten wir als Beispiel etwa die Sinuslinie zwischen Null und  $\pi$  mit einer Annäherung der Fläche durch sechs Rechtecke. In den Abbildungen 50 und 51 sind diese Rechtecke eingezeichnet, einmal bezogen auf den Funktionswert am linken Ende des Teilintervalls und einmal bezogen auf den am rechten. Wie man sieht, liegen die Rechtecke exakt spiegelsymmetrisch zueinander und haben damit insbesondere die gleiche Summe der Flächeninhalte, nämlich jeweils ungefähr 1,954. Die Fläche unter der Sinuslinie zwischen Null und  $\pi$  ist allerdings, wie wir bald sehen werden,

$$-\cos \pi - (-\cos 0) = -(-1) - (-1) = 1 + 1 = 2.$$

Wenn man ein (jedem Wissenschaftler zu empfehlendes) gesundes Mißtrauen gegen Computer und Taschenrechner mitbringt, könnte man versuchen, den Wert 1,954 als eine „Taschenrechnerzwei“ zu interpretieren; da aber die Sinuswerte an den Stellen  $0, \frac{\pi}{6}, \frac{\pi}{3}, \frac{\pi}{2}, \frac{2\pi}{3}, \frac{5\pi}{6}$  und 1 wohlbekannt sind (im Winkelmaß ausgedrückt sind diese Stellen gerade die Vielfachen von  $30^\circ$ ), kann man die Summe der Rechteckflächen exakt berechnen mit dem Ergebnis

$$\frac{\pi}{6}(2 + \sqrt{3}),$$

das schon wegen der Transzendenz von  $\pi$  nicht gleich zwei sein kann. Somit haben wir definitiv keine obere Schranke für den korrekten Wert, und zumindest *a priori* haben wir auch keine untere, denn daß die

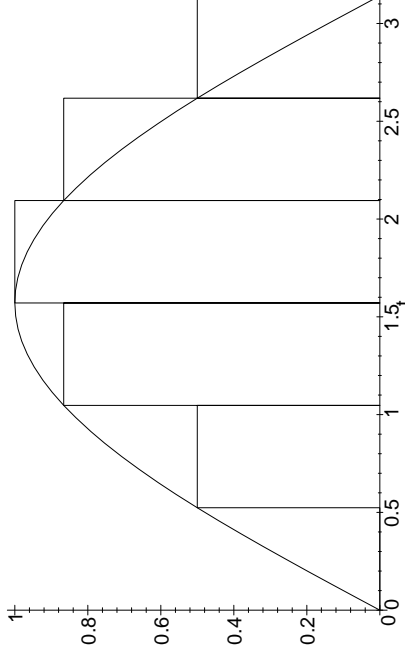


Abb. 50: Sinuslinie mit sechs Rechtecken zum jeweils linken Funktionswert

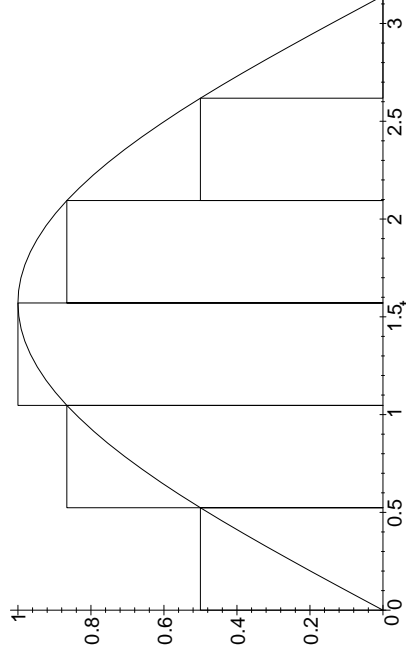


Abb. 51: Sinuslinie mit sechs Rechtecken zum jeweils rechten Funktionswert

Summe der Rechteckflächen kleiner ist als die Fläche unter der Kurve folgte je erst nachträglich durch Vergleich der beiden Werte.

Um wirklich eine untere Schranke für die Fläche zu bekommen, müßte man hier im Beispiel für die ersten drei Rechtecke den Funktionswert am linken Ende des Teilintervalls nehmen und für die letzten drei den vom rechten; für eine obere Schranke müßte man genau umgekehrt vorgehen.

Für eine beliebige Funktion gibt es aber offensichtlich keinen Grund, warum das Minimum oder Maximum überhaupt an einem Intervallende angenommen werden sollte; für einen möglichst allgemeinen Integrallbegriff sollte man also auch Punkte im Intervallinnern zur Ermittlung der Höhe des Rechtecks heranziehen: Anstelle der bislang betrachteten Unterteilung

$$a < a + \delta < a + 2\delta < \dots < a + (n-1)\delta < a + n\delta = b \quad \text{mit} \quad \delta = \frac{b-a}{n}$$

sollte also eine beliebige Unterteilung

$$a = x_0 < x_1 < x_2 < \dots < x_{n-1} < x_n = b$$

treten.

**3) Anwendung des Mittelwertsatzes:** Einen weiteren Grund dafür liefert der wohl wichtigste Satz aus der Analysis I, der *Mittelwertsatz der Differentialrechnung*: Ist  $F$  stetig in einem Intervall  $[u, v]$  und differenzierbar in  $(u, v)$ , so gibt es einen Punkt  $\xi \in (u, v)$ , so daß

$$F'(\xi) = \frac{f(v) - f(u)}{v - u}$$

ist.

Wir suchen zu einer gegebenen Funktion  $f$  eine differenzierbare Funktion  $F$  mit  $F' = f$ . Falls wir annehmen, daß wir diese Funktion schon hätten, so wäre  $F' = f$ , und nach dem Mittelwertsatz gäbe es in jedem Intervall  $(x_i, x_{i+1})$  ein Element  $\xi_i$ , so daß

$$\frac{F(x_{i+1}) - F(x_i)}{x_{i+1} - x_i} = F'(\xi_i) = f(\xi_i)$$

wäre, d.h.

$$F(x_{i+1}) = F(x_i) + f(\xi_i)(x_{i+1} - x_i).$$

Rekursiv folgt, daß

$$\begin{aligned} F(x) &= F(x_n) = F(x_{n-1}) + f(\xi_{n-1})(x_n - x_{n-1}) \\ &= F(x_{n-2}) + f(\xi_{n-1})(x_{n-1} - x_{n-2}) + f(\xi_{n-1})(x_n - x_{n-1}) \\ &= \vdots \\ &= F(x_0) + \sum_{i=0}^{n-1} f(\xi_i)(x_{i+1} - x_i). \end{aligned}$$

Damit ist also die Fläche unter der Kurve  $y = f(x)$  *exakt* gleich der Gesamfläche endlich vieler geeignet gewählter Rechtecke; Abbildung 52 veranschaulicht dies anhand der ganz zu Beginn betrachteten Geschwindigkeitsfunktion. Hier ist nicht nur die Fläche unter der Kurve exakt gleich der Fläche der fünf Rechtecke, sondern auch die Fläche eines jeden Rechtecks gleich der Fläche zwischen Kurve und Grundlinie des Rechtecks, wir kommen also mit endlich vielen Rechtecken aus und brauchen nicht einmal einen Grenzwert zu berechnen.

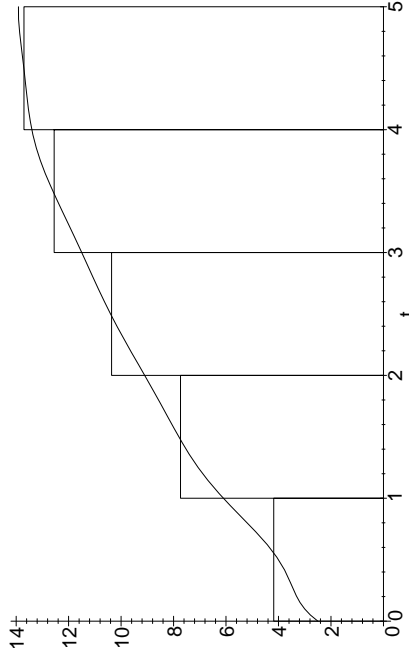


Abb. 52: Die Fläche unter der Kurve ist gleich der Fläche der fünf Rechtecke

Der Haken bei der Sache ist natürlich, daß wir die „geeignet gewählten“ Rechtecke nicht kennen: Im obigen Ausdruck ist alles bekannt *außer* den Werten  $\xi_i$ , an denen die Funktion  $f$  ausgewertet wird.

Die Idee zur Definition des RIEMANN-Integrals ist nun, daß man einfach *beliebige*  $\xi_i$  zwischen  $x_i$  und  $x_{i+1}$  wählt in der Hoffnung, daß bei immer kleiner werdendem Abstand zwischen zwei aufeinanderfolgenden  $x_i$  der Grenzwert nicht mehr von der Wahl der  $\xi_i$  abhängt.

Diese Hoffnung hat natürlich nur eine Chance auf Erfüllung, wenn die Funktion  $f$  hinreichend stetig ist; ansonsten können sich die Werte von  $f$  an zwei beliebig nahe beieinanderliegenden Stellen immer noch beliebig stark unterscheiden, indem sie beispielsweise wie oben davon abhängen, ob  $\xi_i$  rational ist oder nicht.

**4) Gleichmäßige Stetigkeit:** Die Stetigkeit allein reicht allerdings immer noch nicht ganz aus:

Erinnern wir uns:  $f$  heißt *stetig* auf einer Teilmenge  $D \subseteq \mathbb{R}$ , wenn es zu jedem  $\varepsilon > 0$  und zu jedem  $x \in (a, b)$  ein  $\delta > 0$  gibt, so daß gilt: Ist  $y \in D$  und  $|y - x| < \delta$ , so folgt, daß  $|f(y) - f(x)| < \varepsilon$  ist.

Wir möchten mehr: Wir wollen, daß sich in *jedem* Rechteck der Wert von  $f(\xi)$  höchstens um  $\varepsilon > 0$  ändert, falls nur die Breite des Rechtecks unter einer gewissen Schranke liegt, d.h. wir möchten, daß  $\delta$  nicht von  $x$  abhängt

Dies führt auf folgende Verschärfung des Stetigkeitsbegriffs:

**Definition:**  $D \subseteq \mathbb{R}$  sei eine Teilmenge von  $\mathbb{R}$  und  $f: D \rightarrow \mathbb{R}$  eine Funktion.  $f$  heißt *gleichmäßig stetig* in  $D$ , wenn es zu jedem  $\varepsilon > 0$  ein  $\delta > 0$  gibt, so daß für alle  $x, y \in D$  gilt:

$$|y - x| < \delta \implies |f(y) - f(x)| < \varepsilon.$$

Um zu sehen, daß dies mehr ist als die bloße Stetigkeit, betrachten wir die Funktion  $f(x) = 1/x$ . Diese Funktion ist stetig auf der Menge aller positiver reeller Zahlen, denn für  $0 < \delta < x$  und  $\varepsilon > 0$  ist

$$\begin{aligned} \left| \frac{1}{x \pm \delta} - \frac{1}{x} \right| &= \left| \frac{\mp \delta}{x(x \pm \delta)} \right| = \frac{\delta}{x(x \pm \delta)} < \varepsilon \\ \Leftrightarrow \delta < (x^2 \pm x\delta)\varepsilon &\Leftrightarrow \delta \mp x\delta\varepsilon < x^2\varepsilon \Leftrightarrow \delta < \frac{x^2\varepsilon}{1 \mp x\varepsilon}. \end{aligned}$$

Der Ausdruck ganz rechts ist offensichtlich kleiner, wenn im Nenner das Pluszeichen steht, also gilt: Für gegebenes  $x > 0$  und  $\varepsilon > 0$  gilt für jede positive reelle Zahl  $y$

$$|y - x| < \delta \stackrel{\text{def}}{=} \frac{x^2\varepsilon}{1 + x\varepsilon} \implies \left| \frac{1}{y} - \frac{1}{x} \right| < \varepsilon.$$

Damit ist die Stetigkeit der Funktion bewiesen. Sie ist aber nicht gleichmäßig stetig, denn es gibt aber kein von  $x$  unabhängiges  $\delta > 0$  mit dieser Eigenschaft: Der angegebene Ausdruck für das größtmögliche  $\delta$  geht wegen des Faktors  $x^2$  im Zähler für  $x \rightarrow 0$  selbst gegen Null, fällt also unter jede vorgegebene positive Schranke. Abbildung 53 zeigt  $\delta$  in Abhängigkeit von  $x$  für den speziellen Wert  $\varepsilon = \frac{1}{10}$ .

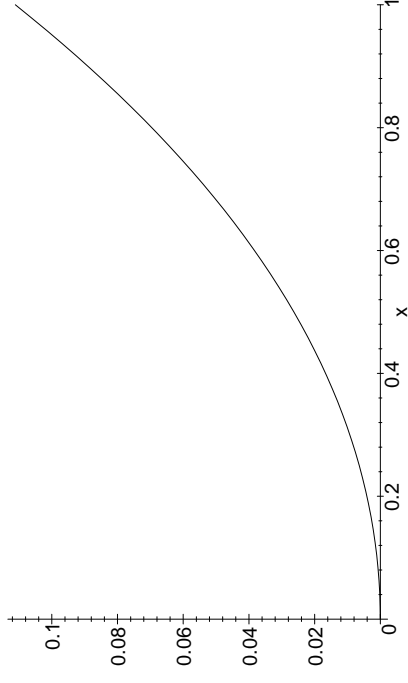


Abb. 53:  $\delta$  in Abhängigkeit von  $x$  für  $f(x) = 1/x$  und  $\varepsilon = 0,1$

In einem abgeschlossenen Intervall sollte so etwas nicht möglich sein, und in der Tat gilt der (durchaus nichttriviale)

**Satz:**  $f: [a, b] \rightarrow \mathbb{R}$  sei eine stetige Funktion auf dem abgeschlossenen Intervall  $[a, b]$  mit  $a, b \in \mathbb{R}$ . Dann ist  $f$  auf  $[a, b]$  gleichmäßig stetig.

Der Beweis ist technisch und indirekt:



Angenommen,  $f$  wäre *nicht* gleichmäßig stetig. Dann gäbe es für mindestens ein  $\varepsilon > 0$  zu jedem  $\delta > 0$  Punkte  $x, y \in [a, b]$ , so daß  $|y - x| < \delta$ , aber  $|f(y) - f(x)| \geq \varepsilon$  wäre.

Aufgrund der Annahme, der Satz sei falsch, können wir ein solches  $\varepsilon$  fixieren und betrachten die speziellen Werte  $\delta = \frac{1}{n}$ : Nach dem gerade gesagten gibt es dazu Punkte  $x_n, y_n \in [a, b]$ , so daß  $|y_n - x_n| < \frac{1}{n}$ , aber  $|f(y_n) - f(x_n)| \geq \varepsilon$  ist.

Nun kommt der nichttriviale Teil: Wir benötigen aus der Analysis I den

**Satz von Bolzano-Weierstraß:** Jede Folge  $(x_n)$  von Punkten aus dem abgeschlossenen Intervall  $[a, b]$  hat (mindestens) eine konvergente Teilfolge  $x_{n_\nu}$ .

Insbesondere muß also die hier betrachtete Folge  $(x_n)$  eine konvergente Teilfolge  $(x_{n_\nu})_{\nu \in \mathbb{N}}$  haben; deren Grenzwert sei  $c \in [a, b]$ .

Da  $|x_{n_\nu} - y_{n_\nu}| < 1/n_\nu$  ist, muß dann auch die Folge der  $y_{n_\nu}$  gegen  $c$  konvergieren, und da  $f$  nach Voraussetzung eine stetige Funktion ist, gilt

$$\lim_{\nu \rightarrow \infty} f(x_{n_\nu}) = \lim_{\nu \rightarrow \infty} f(y_{n_\nu}) = f(c).$$

Andererseits ist aber für jedes  $\nu$  nach Konstruktion der Folgen  $(x_n)$  und  $(y_n)$  der Abstand zwischen  $f(x_{n_\nu})$  und  $f(y_{n_\nu})$  größer oder gleich der festgewählten Zahl  $\varepsilon$ , so daß die beiden Folgen unmöglich denselben Grenzwert haben können.

Mithin führt die Annahme,  $f$  sei *nicht* gleichmäßig stetig, auf einen Widerspruch, und damit ist der Satz bewiesen. ■

**5) Definition einer Approximation für das Integral:** Nach diesen Vorarbeiten können wir ernsthaft darangehen, das Integral einer Funktion zu definieren. Es gibt in der Mathematik verschiedene Integralbegriffe; für uns reicht die Definition des deutschen Mathematikers BERNHARD RIEMANN, das inzwischen nach ihm benannte RIEMANN-Integral.



GEORG FRIEDRICH BERNHARD RIEMANN (1826-1866) war Sohn eines lutherischen Pastors und schrieb sich 1946 auf Anraten seines Vaters an der Universität Göttingen für das Studium der Theologie ein. Schon bald wechselte an die Philosophische Fakultät, um dort unter anderem bei GAUSS Mathematikvorlesungen zu hören. Nach Promotion 1851 und Habilitation 1854 erhielt er dort 1857 einen Lehrstuhl. Trotz seines frühen Todes initiierte er grundlegende auch noch heute fundamentale Entwicklungen in der Geometrie, der Zahlentheorie und über abelsche Funktionen. Seine Vermutung über die Nullstellen der (heute als RIEMANNNSCH bezeichneten)  $\zeta$ -Funktion ist die berühmteste offene Vermutung der heutigen Mathematik.

Wir gehen aus von einer Funktion  $f: [a, b] \rightarrow \mathbb{R}$  auf einem abgeschlossenen Intervall  $[a, b]$ ; gesucht ist eine Funktion  $F: [a, b] \rightarrow \mathbb{R}$ , für die (idealerweise)  $F'(x) = f(x)$  sein sollte für alle  $x$  aus dem offenen Intervall  $(a, b)$ .

Wir wählen eine Unterteilung

$$a = x_0 < x_1 < x_2 < \dots < x_{n-1} < x_n = b$$

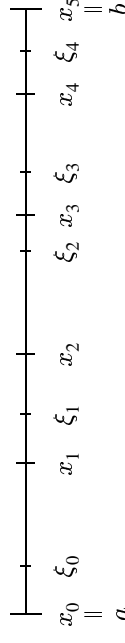
des Intervalls von  $a$  bis  $b$ . Da wir im folgenden viel mit dieser Unterteilung rechnen werden, führen wir die Abkürzung

$$\mathbf{x} = (x_0, x_1, \dots, x_n)$$

dafür ein und verabreden, daß die  $x_i$  immer, wenn von  $\mathbf{x}$  die Rede ist, die obigen Gleichungen und Ungleichungen erfüllen sollen, daß sie also tatsächlich eine Unterteilung von  $[a, b]$  definieren.

Falls die gesuchte differenzierbare Funktion  $F: [a, b] \rightarrow \mathbb{R}$  existiert, sagt uns – wie wir in Teil c) gesehen haben – der Mittelwertsatz der Differentialrechnung, daß für geeignete Elemente  $\xi_i$  mit

$$a = x_0 < \xi_0 < x_1 < \xi_1 < x_2 < \dots < \xi_{n-2} < x_{n-1} < \xi_{n-1} < x_n = b$$



folgt

$$F(x) = F(a) + \sum_{i=0}^{n-1} f(\xi_i)(x_{i+1} - x_i).$$

Natürlich kennen wir die Zahlen  $\xi_i$  nicht – selbst wenn wir wissen, daß  $F$  und damit die  $\xi_i$  überhaupt existieren. Deshalb betrachten wir einfach *beliebige* Zahlen  $\xi_i$  mit

$$a = x_0 < \xi_0 < x_1 < \xi_1 < x_2 < \dots < \xi_{n-2} < x_{n-1} < \xi_{n-1} < x_n = x$$

und führen auch hier wieder die Abkürzung

$$\underline{\xi} = (\xi_0, \xi_1, \dots, \xi_{n-1})$$

ein mit der Verabredung, daß immer, wenn ein  $\underline{\xi}$  in Zusammenhang mit einem  $x$  auftritt, diese Ungleichungen erfüllt sein sollen.

**Definition:** Die RIEMANNSCHE SUMME zu einem Paar  $(x, \underline{\xi})$  ist

$$I(x, \underline{\xi}) = \sum_{i=0}^{n-1} f(\xi_i)(x_{i+1} - x_i).$$

Das RIEMANN-Integral soll der Grenzwert einer Folge RIEMANNSCHER SUMMEN sein, wobei die Unterteilungen  $x$  immer feiner werden. Dieses „Feinerwerden“ müssen wir natürlich auch noch erklären: Seien  $x$  und  $y$  zwei Unterteilungen des Intervalls  $[a, x]$ ; konkret sei

$$a = x_0 < x_1 < \dots < x_{n-1} < x_n = x$$

und

$$a = y_0 < y_1 < \dots < y_{m-1} < y_m = x.$$

Dann heißt  $y$  eine *Verfeinerung* von  $x$ , wenn jedes Teilintervall  $(y_i, y_{i+1})$  von  $y$  ganz in einem der Teilintervalle  $(x_j, x_{j+1})$  von  $x$  liegt;  $y$  entsteht also aus  $x$ , indem einige der Intervalle von  $x$  noch weiter unterteilt werden. Insbesondere muß dann  $m \geq n$  sein.

Nun betrachten wir eine Folge  $(x^{(\nu)})$  von Unterteilungen derart, daß

- $x^{(\nu+1)}$  stets eine Verfeinerung von  $x^{(\nu)}$  ist und

- die maximale Länge der Teilintervalle von  $x^{(\nu)}$  mit wachsendem  $\nu$  gegen Null geht, d.h.

$$\lim_{\nu \rightarrow \infty} \max_{i=0}^{n_{\nu}-1} (x_{i+1}^{(\nu)} - x_i^{(\nu)}) = 0.$$

Zu jeder Unterteilung  $x^{(\nu)}$  wählen wir willkürlich eine (im Sinne obiger Konvention dazu passende) Folge  $\underline{\xi}^{(\nu)}$  von Zwischenpunkten und fragen nach Existenz und gegebenenfalls Wert des Grenzwerts

$$\lim_{\nu \rightarrow \infty} I(x^{(\nu)}, \underline{\xi}^{(\nu)}).$$

Dieser Grenzwert muß natürlich im allgemeinen nicht existieren, und wenn er existiert, kann er von der Wahl der Zwischenpunkte  $\xi_i^{(\nu)}$  und von der Wahl einer Folge  $(x^{(\nu)})$  von Unterteilungen abhängen. Wenn er unabhängig von all diesen Wahlen existiert und immer denselben Wert hat, bezeichnen wir diesen gemeinsamen Wert als das RIEMANN-Integral von  $f$  zwischen  $a$  und  $x$ , in Zeichen

$$\int_a^x f(\xi) d\xi = \lim_{\nu \rightarrow \infty} I(x^{(\nu)}, \underline{\xi}^{(\nu)}).$$

Falls dieses Integral für jedes  $x \in [a, b]$  existiert, sagen wir,  $f$  sei RIEMANN-integrierbar auf  $[a, b]$ .

$\xi$  heißt *Integrationsvariable* und übernimmt die Rolle des Summationsindex in einer Summe: Genau wie beispielsweise

$$\sum_{i=0}^{n-1} i^2 = \sum_{j=0}^{n-1} j^2 = \sum_{k=0}^{n-1} k^2 = \frac{n(2n-1)(n-1)}{6}$$

davon unabhängig ist, ob der Summationsindex  $i, j$  oder  $k$  heißt, ist auch

$$\int_a^x f(\xi) d\xi = \int_a^x f(u) du = \int_a^x f(t) dt$$

unabhängig davon, ob die Integrationsvariable mit  $\xi, u$  oder  $t$  bezeichnet wird.

**6) Existenz des Riemann-Integrals für stetige Funktionen:** Damit haben wir das RIEMANN-Integral definiert; wir wissen allerdings bisher für keine einzige Funktion, daß es existiert. In diesem Abschnitt wollen wir uns überlegen, daß es zumindest für *stetige* Funktionen immer existiert:

**Satz:** Jede stetige Funktion  $f: [a, b] \rightarrow \mathbb{R}$  ist RIEMANN-integrierbar auf  $[a, b]$ .

Zum *Beweis* müssen wir zeigen, daß das RIEMANN-Integral

$$\int_a^x f(\xi) d\xi$$

für jedes  $x \in [a, b]$  existiert.

Dazu fixieren wir ein beliebig vorgegebenes  $x \in [a, b]$  und betrachten Folgen  $(x^{(\nu)})$  von immer feiner werdenden Unterteilungen des Intervalls  $[a, x]$ ; dazu entsprechende Folgen  $(\xi^{(\nu)})$  von Zwischenwerten und die Grenzwerte der RIEMANN-Summen  $I(x^{(\nu)}, \xi^{(\nu)})$ .

**1. Schritt:** Für eine feste Folge  $(x^{(\nu)})$  von Unterteilungen existiert der Grenzwert und ist unabhängig von der Wahl der Zwischenwerte  $\xi_i^{(\nu)}$ :

Um das einzusehen, betrachten wir für jede Unterteilung  $x^{(\nu)}$  die RIEMANNschen Obersummen und Untersummen: Da  $f$  nach Voraussetzung stetig ist, nimmt es in jedem abgeschlossenen Intervall sowohl sein Maximum als auch sein Minimum an; insbesondere werden also im Intervall  $[x_i^{(\nu)}, x_{i+1}^{(\nu)}]$  ein Maximum  $M_i^{(\nu)}$  und ein Minimum  $m_i^{(\nu)}$  angenommen. Damit gilt unabhängig von der Wahl des Zwischenwerts  $\xi_i^{(\nu)}$

$$m_i^{(\nu)} \leq f(\xi_i^{(\nu)}) \leq M_i^{(\nu)}.$$

Nach Definition ist

$$I(x^{(\nu)}, \xi^{(\nu)}) = \sum_{i=0}^{n_\nu-1} f(\xi_i^{(\nu)})(x_{i+1}^{(\nu)} - x_i^{(\nu)});$$

ersetzen wir hierin  $\xi_i^{(\nu)}$  jeweils durch  $m_i^{(\nu)}$ , so erhalten wir eine untere Abschätzung

$$s^{(\nu)} \stackrel{\text{def}}{=} \sum_{i=0}^{n_\nu-1} m_i^{(\nu)}(x_{i+1}^{(\nu)} - x_i^{(\nu)})$$

für  $I(x^{(\nu)}, \xi^{(\nu)})$ ; diese bezeichnen wir als **RIEMANNsche Untersumme** der Unterteilung  $x^{(\nu)}$ . Entsprechend liefert die Ersetzung durch  $M_i^{(\nu)}$  eine obere Abschätzung

$$S^{(\nu)} \stackrel{\text{def}}{=} \sum_{i=0}^{n_\nu-1} M_i^{(\nu)}(x_{i+1}^{(\nu)} - x_i^{(\nu)})$$

für  $I(x^{(\nu)}, \xi^{(\nu)})$ , die **RIEMANNsche Obersumme** der Unterteilung  $x^{(\nu)}$ .

Unabhängig von der Wahl der Zwischenwerte  $\xi_i^{(\nu)}$  gilt somit

$$s^{(\nu)} \leq I(x^{(\nu)}, \xi^{(\nu)}) \leq S^{(\nu)},$$

ein Zusammenhang, der in Abbildung 54 noch einmal graphisch dargestellt ist: Die RIEMANNsche Untersumme ist gleich der Fläche der durchgezogenen Rechtecke, für die Obersumme kommt noch der gestrichelte Anteil dazu, und für die RIEMANNsche Summe  $I(x^{(\nu)}, \xi^{(\nu)})$  muß man bis zu den punktierten Linien gehen. Die Werte  $\xi_i^{(\nu)}$  sind auf der  $x$ -Achse durch kleine Kreise gekennzeichnet, ebenso die Punkte  $(\xi_i^{(\nu)}, f(\xi_i^{(\nu)}))$  auf der Kurve.

Wenn wir jetzt noch zeigen können, daß die Folge der  $s^{(\nu)}$  und die der  $S^{(\nu)}$  beide gegen denselben Grenzwert konvergieren, muß wegen der Einschließung

$$s^{(\nu)} \leq I(x^{(\nu)}, \xi^{(\nu)}) \leq S^{(\nu)}$$

auch  $I(x^{(\nu)}, \xi^{(\nu)})$  gegen diesen Grenzwert konvergieren, und die Behauptung im ersten Schritt ist bewiesen.

Vergleichen wir dazu zunächst die Untersummen  $s^{(\nu)}$  und  $s^{(\nu+1)}$ : Die Unterteilung  $x^{(\nu+1)}$  entsteht aus  $x^{(\nu)}$  dadurch, daß einige der Intervalle weiter unterteilt werden. Ist aber  $(x_j^{(\nu+1)}, x_{j+1}^{(\nu+1)})$  Teilintervall von  $(x_j^{(\nu)}, x_{j+1}^{(\nu)})$ , so kann das Minimum im Teilintervall natürlich nicht kleiner sein als

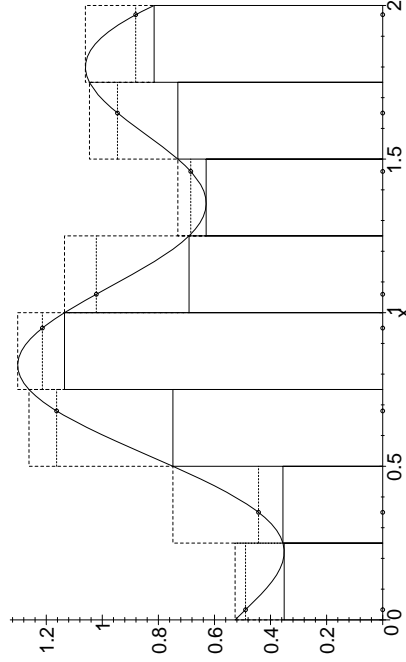


Abb. 54: Riemannsche Summen, Obersummen und Untersummen

im größeren Intervall, d.h.  $s^{(\nu+1)} \geq s^{(\nu)}$ . Somit ist die Folge der  $s^{(\nu)}$  monoton wachsend.

Genauso folgt, daß die Folge der  $S^{(\nu)}$  monoton fallend ist, und da stets  $s^{(\nu)} \leq S^{(\nu)}$  ist, folgt weiter, daß

$$s^{(1)} \leq s^{(2)} \leq s^{(3)} \leq \dots \leq S^{(3)} \leq S^{(2)} \leq S^{(1)}.$$

Insbesondere ist also  $S^{(1)}$  eine obere Schranke für die Folge der  $s^{(\nu)}$  und  $s^{(1)}$  eine untere Schranke für die Folge der  $S^{(\nu)}$ .

Damit ist also die Folge  $(s^{(\nu)})$  monoton wachsend und nach oben beschränkt, während  $(S^{(\nu)})$  monoton fallend und nach unten beschränkt ist. Nach einem Satz aus der Analysis I (*Jede monotone und beschränkte Folge reeller Zahlen ist konvergent.*) folgt daraus die Konvergenz beider Folgen.

Nun fehlt nur noch, daß beide denselben Grenzwert haben.

Dazu betrachten wir die Differenzen

$$\begin{aligned} S^{(\nu)} - s^{(\nu)} &= \sum_{i=0}^{n_\nu-1} M_i^{(\nu)}(x_{i+1}^{(\nu)} - x_i^{(\nu)}) - \sum_{i=0}^{n_\nu-1} m_i^{(\nu)}(x_{i+1}^{(\nu)} - x_i^{(\nu)}) \\ &= \sum_{i=0}^{n_\nu-1} (M_i^{(\nu)} - m_i^{(\nu)})(x_{i+1}^{(\nu)} - x_i^{(\nu)}). \end{aligned}$$

Mit

$$\Delta_i^{(\nu)} = \max_i (M_i^{(\nu)} - m_i^{(\nu)})$$

ist also

$$\begin{aligned} S^{(\nu)} - s^{(\nu)} &\leq \sum_{i=0}^{n_\nu-1} \Delta_i^{(\nu)}(x_{i+1}^{(\nu)} - x_i^{(\nu)}) = \Delta^{(\nu)} \sum_{i=0}^{n_\nu-1} (x_{i+1}^{(\nu)} - x_i^{(\nu)}) \\ &= \Delta^{(\nu)}(x_n^{(\nu)} - x_0^{(\nu)}) = \Delta^{(\nu)}(x - a), \end{aligned}$$

und das ist eine Nullfolge, falls wir zeigen können, daß die Folge der  $\Delta^{(\nu)}$  eine ist.

Hier kommt nun die in Abschnitt c) eingeführte gleichmäßige Stetigkeit von  $f$  zum Tragen: Danach gibt es zu jedem  $\varepsilon > 0$  ein  $\delta > 0$ , so daß für alle Punkte  $\xi_1, \xi_2 \in [a, b]$  gilt: Ist  $|\xi_1 - \xi_2| < \delta$ , so folgt  $|f(\xi_1) - f(\xi_2)| < \varepsilon$ .

Ist dabei für eine Unterteilung  $\mathbf{x}^{(\nu)}$  jede der Differenzen  $x_{i+1}^{(\nu)} - x_i^{(\nu)}$  kleiner als  $\delta$ , so ist auch  $M_i^{(\nu)} - m_i^{(\nu)} < \varepsilon$ , denn sowohl  $M_i^{(\nu)}$  als auch  $m_i^{(\nu)}$  sind Funktionswerte, die irgendwo im Intervall  $[x_i^{(\nu)}, x_{i+1}^{(\nu)}]$  angenommen werden.

Nun haben wir aber vorausgesetzt, daß die Unterteilungen  $\mathbf{x}^{(\nu)}$  immer feiner werden, d.h. zu jedem  $\delta > 0$  gibt es in der Tat ein  $\nu_0$ , so daß  $x_{i+1}^{(\nu)} - x_i^{(\nu)} < \delta$  für alle  $\nu > \nu_0$  und alle  $i$ . Somit sind für  $\nu > \nu_0$  alle Differenzen  $M_i^{(\nu)} - m_i^{(\nu)}$  kleiner als  $\varepsilon$ , und damit ist auch  $\Delta^{(\nu)} < \varepsilon$  für  $\nu > \nu_0$ . Also sind sowohl  $(\Delta^{(\nu)})$  als auch  $(S^{(\nu)} - s^{(\nu)})$  Nullfolgen, d.h.

$$\lim_{\nu \rightarrow \infty} s^{(\nu)} = \lim_{\nu \rightarrow \infty} S^{(\nu)}.$$

Wie schon erwähnt, folgt daraus aufgrund der Einschließung

$$s^{(\nu)} \leq I(\mathbf{x}^{(\nu)}, \underline{\xi}^{(\nu)}) \leq S^{(\nu)},$$

auch die Gleichung

$$\lim_{\nu \rightarrow \infty} S^{(\nu)} = \lim_{\nu \rightarrow \infty} I(\mathbf{x}^{(\nu)}, \underline{\xi}^{(\nu)}) = \lim_{\nu \rightarrow \infty} S^{(\nu)}.$$

Da die linke und die rechte Seite obiger Ungleichung nicht von den  $\xi_i^{(\nu)}$  abhängen, ist auch  $\lim_{\nu \rightarrow \infty} I(\mathbf{x}^{(\nu)}, \underline{\xi}^{(\nu)})$  davon unabhängig, und damit ist der erste Schritt des Beweises beendet.

Der zweite ist zum Glück erheblich einfacher:

**2. Schritt:**  $\lim_{\nu \rightarrow \infty} I(\mathbf{x}^{(\nu)}, \underline{\xi}^{(\nu)})$  ist auch unabhängig von der Folge  $(\mathbf{x}^{(\nu)})$ :

Betrachten wir zwei Folgen  $(\mathbf{x}^{(\nu)})$  und  $(\mathbf{y}^{(\nu)})$  von Unterteilungen. Für jeden Index  $\nu$  können wir zu den Unterteilungen  $\mathbf{x}^{(\nu)}$  und  $\mathbf{y}^{(\nu)}$  eine gemeinsame Verfeinerung  $\mathbf{z}^{(\nu)}$  konstruieren, indem wir einfach die sämtlichen Zahlen  $x_i^{(\nu)}$  und  $y_j^{(\nu)}$  der Größe nach ordnen als  $z_0^{(\nu)} = a, \dots, z_r^{(\nu)} = x$ .

Wie oben seien  $s^{(\nu)}$  und  $S^{(\nu)}$  die RIEMANNSCHE Unter- und Obersumme zur Unterteilung  $\mathbf{x}^{(\nu)}$ ; entsprechend seien  $\tilde{s}^{(\nu)}$  und  $\tilde{S}^{(\nu)}$  die zur Unterteilung  $\mathbf{z}^{(\nu)}$ . Da  $\mathbf{z}^{(\nu)}$  eine Verfeinerung von  $\mathbf{x}^{(\nu)}$  ist, wissen wir aus dem ersten Schritt, daß

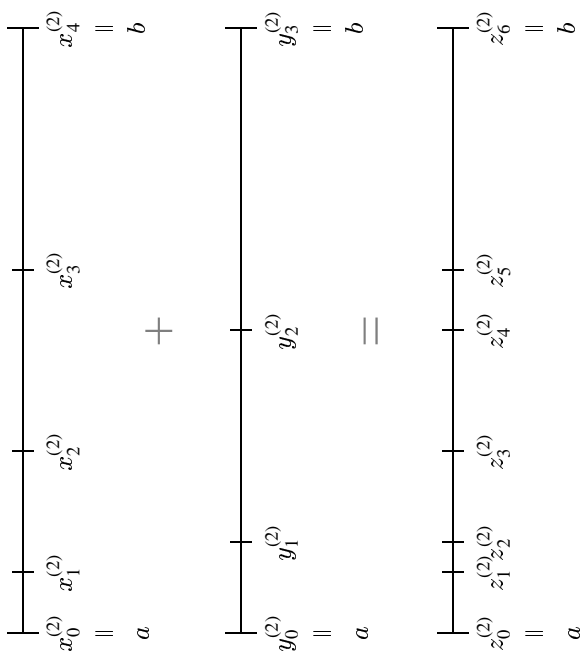
$$s^{(\nu)} \leq \tilde{s}^{(\nu)} \leq \tilde{S}^{(\nu)} \leq S^{(\nu)}$$

ist; da die Folgen  $(s^{(\nu)})$  und  $(S^{(\nu)})$  denselben Grenzwert haben, müssen auch  $(\tilde{s}^{(\nu)})$  und  $(\tilde{S}^{(\nu)})$  gegen diesen Wert konvergieren und damit auch die Folge der  $I(\mathbf{z}^{(\nu)}, \tilde{\xi}^{(\nu)})$ , wobei  $(\tilde{\xi}^{(\nu)})$  irgendeine Folge von Zwischenwerten bezeichnet; wie wir bereits aus dem ersten Schritt wissen, hängt der Grenzwert nicht davon ab. Also ist

$$\lim_{\nu \rightarrow \infty} I(\mathbf{x}^{(\nu)}, \underline{\xi}^{(\nu)}) = \lim_{\nu \rightarrow \infty} I(\mathbf{z}^{(\nu)}, \tilde{\xi}^{(\nu)}).$$

Genauso folgt, daß auch

$$\lim_{\nu \rightarrow \infty} I(\mathbf{y}^{(\nu)}, \tilde{\xi}^{(\nu)}) = \lim_{\nu \rightarrow \infty} I(\mathbf{z}^{(\nu)}, \tilde{\xi}^{(\nu)}),$$



wobei  $(\tilde{\xi}^{(\nu)})$  die Folge der Zwischenwerte zu  $\mathbf{y}^{(\nu)}$  bezeichnet, und dies zeigt schließlich die Behauptung

$$\lim_{\nu \rightarrow \infty} I(\mathbf{x}^{(\nu)}, \underline{\xi}^{(\nu)}) = \lim_{\nu \rightarrow \infty} I(\mathbf{y}^{(\nu)}, \tilde{\xi}^{(\nu)}).$$

Damit ist der Satz vollständig bewiesen. ■

**7) Stückweise stetige Funktionen:** Gelegentlich möchte man die Bedingung der Stetigkeit wenigstens ein bißchen lockern, um beispielweise auch einen Rechteckimpuls, der periodisch zwischen 0 und 1 (oder -1 und 1) wechselt, behandeln zu können. Die Existenz des RIEMANN-Integral wird durch solche kleineren Abweichungen nicht beeinträchtigt:

**Definition:** Eine Funktion  $f: [a, b] \rightarrow \mathbb{R}$  heißt *stückweise stetig*, wenn

es eine Unterteilung

$$a = a_0 < a_1 < \dots < a_{r-1} < a_r = b$$

des Intervalls  $[a, b]$  gibt, so daß  $f$  auf jedem der offenen Intervalle  $(a_i, a_{i+1})$  stetig ist.

$f$  ist also überall stetig außer eventuell in endlich vielen Punkten  $a_0, \dots, a_r$ . Der Wert in diesen Punkten kann, aber muß nicht mit dem linksseitigen oder rechtsseitigen Grenzwert der Funktion übereinstimmen; beispielsweise definiert man Rechteckimpulse gelegentlich auch durch die Funktion

$$f(x) = \begin{cases} 1 & \text{falls } 2n - 1 < x < 2n \text{ für ein } n \in \mathbb{Z} \\ 0 & \text{falls } x \in \mathbb{Z} \\ -1 & \text{falls } 2n < x < 2n + 1 \text{ für ein } n \in \mathbb{Z} \end{cases}.$$

**Satz:** Eine stückweise stetige Funktion  $f: [a, b] \rightarrow \mathbb{R}$  ist RIEMANN-integrierbar.

*Beweis:*  $f$  sei stetig in den offenen Intervallen  $(a_i, a_{i+1})$  mit

$$a = a_0 < a_1 < \dots < a_{r-1} < a_r = b.$$

Wiederholt man den Beweis des entsprechenden Satzes für Funktionen, die auf ganz  $[a, b]$  stetig sind, so funktioniert fast alles problemlos auch für  $f$ . Schwierigkeiten gibt es nur mit den Intervallen einer Unterteilung  $\alpha$ , die einen der Punkte  $a_i$  enthalten. In diesen Intervallen ist  $f$  nicht notwendigerweise stetig, so daß die Differenz zwischen dem Supremum und dem Infimum von  $f$  dort nicht mit Verkleinerung des Intervalls gegen Null gehen muß, sondern auch gegen einen von Null verschiedenen Wert, nämlich die „Sprunghöhe“ bei  $a_i$ , konvergieren kann.

Nun gibt es aber in jeder Unterteilung nur endlich viele Intervalle, die ein  $a_i$  enthalten, und auch die Sprunghöhen sind begrenzt; gehen also alle Intervalllängen gegen Null, so geht auch wie im Beweis des Satzes die Differenz zwischen RIEMANNschen Ober- und Untersummen gegen Null. ■

8) **Noch einmal die Dirichletsche Sprungfunktion:** In Abschnitt b3) waren wir nicht zufrieden mit dem Verhalten des dort heuristisch eingeführten Integrals für die DIRICHLETSche Sprungfunktion; schauen wir, was das nun eingeführte RIEMANN-Integral daraus macht. Sei also

$$f(x) = \begin{cases} 1 & \text{für } x \in \mathbb{Q} \\ 0 & \text{für } x \notin \mathbb{Q} \end{cases};$$

wir interessieren uns für  $\int_0^1 f(x) dx$ .

Dazu sei  $\alpha$  eine Unterteilung des Intervalls  $[0, 1]$ . Unabhängig von der Wahl dieser Unterteilung können wir stets eine Folge  $\xi$  von rationalen Zwischenwerten finden, aber auch eine Folge  $\tilde{\xi}$  von irrationalen. Dann ist, unabhängig von der Unterteilung  $\alpha$ ,

$$I(\alpha, \xi) = \sum_{i=0}^{n-1} f(\xi_i)(x_{i+1} - x_i) = \sum_{i=0}^{n-1} (x_{i+1} - x_i) = 1 - 0 = 1$$

und

$$I(\alpha, \tilde{\xi}) = \sum_{i=0}^{n-1} f(\tilde{\xi}_i)(x_{i+1} - x_i) = 0.$$

Damit kann kein gemeinsamer Grenzwert existieren, und somit ist die DIRICHLETSche Sprungfunktion nicht RIEMANN-integrierbar – ein sehr viel überzeugenderes Ergebnis als das aus Abschnitt b3).

9) **Ausblick: Das Lebesgue-Integral:** Wir könnten allerdings auch argumentieren, daß es „nur“ abzählbar unendlich viele rationale Zahlen, aber überabzählbar viele irrationale zwischen null und eins gibt; daher sollten letztere das Geschehen dominieren, und  $\int_0^1 f(x) dx$  sollte verschwinden.

In der Tat kann man eine Verallgemeinerung des RIEMANN-Integrals definieren, das LEBESGUE-Integral, das für stückweise stetige Funktionen mit dem RIEMANN-Integral übereinstimmt und für die DIRICHLETSche Sprungfunktion den Wert Null liefert. Dazu geht man nach dem französischen Mathematiker HENRI LÉON LEBESGUE (1875-1941) bei den Unter- und Obersummen nicht wie bei RIEMANN von endlich vielen

Rechtecken aus, sondern von *abzählbar unendlich* vielen. (Wie man dies im einzelnen macht, braucht uns hier nicht zu interessieren; wir werden uns im folgenden stets auf das RIEMANN-Integral beschränken.) Dieses LEBESGUE-Integral existiert *fast* immer: Man kann zwar die *Existenz* von Funktionen, für die es nicht existiert, *beweisen*, explizite Beispiele solcher Funktionen sind aber nicht bekannt.

**10) Anwendung auf Flächeninhalte:** Nachdem wir nun mit einem exakten Integralbegriff haben, bietet sich an, Flächen über dieses Integral zu *definieren*:

**Definition:** Ist  $f: [a, b] \rightarrow \mathbb{R}$  eine nichtnegative Funktion, so bezeichnen wir die Zahl

$$\int_a^b f(x) dx$$

als *Fläche* zwischen der Kurve  $y = f(x)$  und der  $x$ -Achse zwischen den Geraden  $x = a$  und  $x = b$ .

Was passiert, wenn  $f$  auch negative Werte annimmt? Für  $f(x) = x$  etwa ist  $\int_{-1}^1 f(x) dx = 0$ , wie man sich leicht überlegt anhand von Unterteilungen, die symmetrisch zur Null liegen. Auch dies läßt sich als Aussage über Flächeninhalte interpretieren – wenn man davon ausgeht, daß die Formel

$$\text{Fläche} = \text{Länge} \times \text{Breite}$$

für die Rechteckfläche auch bei *negativer* Länge und/oder Breite gelten soll. Da Integrale nicht nur zur Flächenbestimmung, sondern auch etwa zur Bestimmung der Ladung in einem Kondensator verwendet werden, ist dies die Interpretation, die man in der Mathematik in den meisten Fällen vorzieht; für die klassische Fläche im Sinne des Tapezierens muß man mit

$$\int_a^b |f(x)| dx$$

arbeiten.

Im Sinne dieser negativen Längen und Breiten ist es auch sinnvoll, Integrale für  $b < a$  zu definieren durch

$$\int_a^b f(x) dx \stackrel{\text{def}}{=} - \int_b^a f(x) dx.$$

Insbesondere ist dann natürlich  $\int_a^a f(x) dx = 0$ .

### d) Erste Integrationsregeln

Das RIEMANN-Integral stellt die heuristischen Überlegungen vom Beginn dieses Paragraphen insofern auf eine exaktere mathematische Grundlage, als wir nun einen Integralbegriff haben, der auch bei extrem ungewöhnlichen Funktionen wie der DIRICHLETSchen Sprungfunktion keine unerwarteten Ergebnisse liefert. Was allerdings den Ausgangspunkt betrifft, die Umkehrung der Differentiation, wissen wir über das neue Integral noch gar nichts, und auch sonst kennen wir noch nicht viele Regeln über den Umgang damit und insbesondere auch über seine Berechnung *ohne* die umständlichen und sehr langsam konvergierenden RIEMANN-Summen.

In diesem Abschnitt sollen die ersten (und einfachsten) solchen Regeln zusammengestellt werden; danach erweitern wir zunächst unser Instrumentarium, um dann damit auch kompliziertere und interessantere Regeln zu beweisen.

**1) Monotonieregel:** Eine der einfachsten, aber trotzdem oft nützlichen Regeln für den Umgang mit Integralen übersetzt Größerbeziehungen zwischen Integranden in Größenbeziehungen zwischen Integralen:

**Satz:**  $f$  und  $g$  seien stückweise stetige Funktionen auf dem Intervall  $[a, b]$ , und für alle  $x \in [a, b]$  sei  $f(x) \leq g(x)$ . Dann ist auch

$$\int_a^b f(x) dx \leq \int_a^b g(x) dx.$$

Insbesondere ist

$$\left| \int_a^b f(x) dx \right| \leq \int_a^b |f(x)| dx$$

*Beweis:* Wir betrachten eine Unterteilung  $\alpha$  des Intervalls  $[a, b]$  und eine dazu kompatible Sequenz  $\xi$  von Zwischenwerten;  $I(\alpha, \xi)$  sei die zugehörige RIEMANN-Summe für  $f$  und  $J(\alpha, \xi)$  die für  $g$ . Da für jedes  $\xi_i$  nach Voraussetzung  $f(\xi_i) \leq g(\xi_i)$  ist, folgt unmittelbar aus der Definition der RIEMANN-Summen, daß

$$I(\alpha, \xi) \leq J(\alpha, \xi)$$

ist. Da sich eine solche Kleinergleichbeziehung auch auf Grenzwerte überträgt, ist daher

$$\int_a^b f(x) dx \leq \int_a^b g(x) dx,$$

wie behauptet.

Insbesondere können wir dies auch anwenden auf die Ungleichungskette

$$-|f(x)| \leq f(x) \leq |f(x)|$$

und erhalten

$$-\int_a^b |f(x)| dx \leq \int_a^b f(x) dx \leq \int_a^b |f(x)| dx,$$

d.h.

$$\left| \int_a^b f(x) dx \right| \leq \int_a^b |f(x)| dx.$$

2) **Linearität und Zusammensetzung:** Genauso einfach ist die Linearitätsregel: ■

**Satz:** Für zwei stückweise stetige Funktionen  $f, g: [a, b] \rightarrow \mathbb{R}$  und zwei reelle Zahlen  $\alpha, \beta$  ist

$$\int_a^b (\alpha f(x) + \beta g(x)) dx = \alpha \int_a^b f(x) dx + \beta \int_a^b g(x) dx.$$

*Beweis:* Wie beim letzten Satz: Für jede RIEMANN-Summe zu einer festen Unterteilung  $\alpha$  und jede damit kompatible Zwischenwertsequenz  $\xi$  gilt die zur Behauptung des Satzes analoge Gleichung, und damit gilt im Limes auch der Satz selbst. ■

Für die Zusammensetzung von Integrationsintervallen gilt erwartungsgemäß

**Satz:** Für  $a \leq b \leq c$  und eine stückweise stetige Funktion  $f: [a, c] \rightarrow \mathbb{R}$  ist

$$\int_a^c f(x) dx = \int_a^b f(x) dx + \int_b^c f(x) dx.$$

*Beweis:*  $\int_a^c f(x) dx$  kann als Grenzwert von RIEMANN-Summen zu einer beliebigen Folge sich verfeinernder Unterteilungen von  $[a, c]$  berechnet werden; insbesondere können also Unterteilungen gewählt werden, die den Zwischenpunkt  $b$  enthalten, und dafür folgt die Behauptung unmittelbar. ■

3) **Der Mittelwertsatz der Integralrechnung:** Der Mittelwertsatz der Differentialrechnung ist der wohl wichtigste Satz aus der Analysis I. Mit dieser überragenden Bedeutung kann der Mittelwertsatz der Integralrechnung nicht konkurrieren; trotzdem ist auch er eine sehr nützliche Aussage, die insbesondere auch die zu Beginn dieses Paragraphen postulierte Interpretation von Integralen als Mittelwerten für das RIEMANN-Integral auf eine solide Grundlage stellt.



Im Hinblick auf spätere Anwendungen sei der Satz zunächst etwas allgemeiner formuliert, als wir ihn im Augenblick benötigen; an der Schwierigkeit des Beweises ändert diese Verallgemeinerung nichts.

**Satz:**  $f: [a, b] \rightarrow \mathbb{R}$  sei eine stetige Funktion;  $g: [a, b] \rightarrow \mathbb{R}$  sei stückweise stetig und nirgends negativ. Dann gibt es ein  $\xi \in [a, b]$ , so daß

$$\int_a^b f(x)g(x) dx = f(\xi) \int_a^b g(x) dx.$$

Die Funktion  $g$  sollte man sich dabei als eine „Gewichtsfunktion“ vorstellen, die die verschiedenen  $x$ -Werte verschieden stark gewichtet. Am anschaulichsten ist der Spezialfall  $g \equiv 1$ , der deshalb noch einmal getrennt angegeben sei:

**Korollar:** Für eine stetige Funktion  $f: [a, b] \rightarrow \mathbb{R}$  gibt es ein  $\xi \in [a, b]$  mit

$$f(\xi) = \frac{1}{b-a} \int_a^b f(x) dx.$$

*Mit anderen Worten:* Der „Mittelwert“ der Funktion  $f$  im Intervall  $[a, b]$  existiert nicht nur, sondern wird auch an (mindestens) einer Stelle im Intervall angenommen.

Zum Beweis des Satzes betrachten wir das Minimum  $m$  und das Maximum  $M$  von  $f$  auf  $[a, b]$ . Da  $g$  nirgends negativ wird, gilt dann auch für jedes  $x \in [a, b]$

$$m g(x) \leq f(x)g(x) \leq M g(x)$$

und damit nach der Monotonieergel

$$m \int_a^b g(x) dx \leq \int_a^b f(x)g(x) dx \leq M \int_a^b g(x) dx.$$

Es gibt daher eine reelle Zahl  $\mu$  zwischen  $m$  und  $M$ , den Mittelwert, so daß

$$\int_a^b f(x)g(x) dx = \mu \int_a^b g(x) dx$$

ist, und da  $f$  stetig ist, gibt es nach dem Zwischenwertsatz mindestens ein  $\xi \in [a, b]$ , für das  $f(\xi) = \mu$  ist.

Damit ist der Satz bewiesen, und das Korollar ist natürlich einfach der Spezialfall  $g \equiv 1$ , für den das Integral über  $g$  gleich  $b - a$  ist. ■

Man beachte, daß es hier nicht genügt, daß  $f$  nur eine *stückweise* stetige Funktion ist: Für

$$f(x) = \begin{cases} 0 & \text{für } x < 0 \\ 1 & \text{für } x \geq 0 \end{cases}$$

ist

$$\frac{1}{1 - (-1)} \int_{-1}^1 f(x) dx = \frac{1}{2} \int_0^1 1 dx = \frac{1}{2}$$

nicht in der Form  $f(\xi)$  darstellbar.

### e) Der Hauptsatz der Differential- und Integralrechnung

Eine Motivation zur Einführung eines Integrals war die Suche nach einer Umkehroperation zur Differentiation: Genau wie die Differentiation aus einem zeitabhängigen Weg eine Geschwindigkeit macht, wollten wir ausgehend von einer zeitabhängigen Geschwindigkeit den Weg zurückbekommen. Der Hauptsatz der Differential- und Integralrechnung sagt uns nun endlich, daß wir dieses Ziel erreicht haben:

**Hauptsatz der Differential- und Integralrechnung:**  $f: [a, b] \rightarrow \mathbb{R}$  sei eine stetige Funktion und

$$F_a(x) = \int_a^x f(\xi) d\xi.$$

Dann ist  $F'_a(x) = f(x)$  für alle  $x \in (a, b)$ .

Ist umgekehrt  $F$  eine differenzierbare Funktion, deren Ableitung auf dem offenen Intervall  $(a, b)$  mit  $f$  übereinstimmt, so gibt es eine reelle Zahl  $C$  derart, daß  $F(x) = F_a(x) + C$  ist.

*Beweis:* Die Ableitung von  $F_a$  ist

$$F'_a(x) = \lim_{h \rightarrow 0} \frac{F_a(x+h) - F_a(x)}{h};$$

für positives  $h$  ist

$$F_a(x+h) - F_a(x) = \int_a^{x+h} f(x) dx - \int_a^x f(x) dx = \int_x^{x+h} f(x) dx$$

(s. Abschnitt b)); für negatives  $h$  ist entsprechend

$$F_a(x+h) - F_a(x) = - \int_{x+h}^x f(x) dx.$$

Nach dem Mittelwertsatz der Integralrechnung aus dem vorigen Abschnitt gibt es in jedem der beiden Fälle ein  $\xi$  zwischen  $x$  und  $x+h$ , so daß das Integral gleich  $|h| \cdot f(\xi)$  ist, d.h.

$$\frac{F_a(x+h) - F_a(x)}{h} = f(\xi).$$

Geht nun  $h$  gegen Null, so muß die zwischen  $x$  und  $x+h$  liegende Zahl  $\xi$  gegen  $x$  gehen; wegen der Stetigkeit von  $f$  ist also  $F'_a(x) = f(x)$

Ist  $F$  eine weitere Funktion mit Ableitung  $f$ , so hat die Differenz  $h$  von  $F$  und  $F_a$  die Ableitung Null. Nun erinnern wir uns an die Analysis I: Wenn die Ableitung einer Funktion  $h(x)$  verschwindet, muß die Funktion konstant sein. In der Tat: Gäbe es im Definitionsbereich von  $h$  zwei Werte  $x_1 \neq x_2$  mit  $h(x_1) \neq h(x_2)$ , so gäbe es nach dem Mittelwertsatz der Differentialrechnung ein  $\xi \in (x_1, x_2)$ , so daß

$$\frac{h(x_2) - h(x_1)}{x_2 - x_1} = h'(\xi)$$

wäre. Hier steht links eine von Null verschiedene Zahl, rechts aber Null, ein Widerspruch.

Also gibt es eine Konstante  $C$ , so daß  $F(x) - F_a(x) = C$  oder

$$F(x) = F_a(x) + C$$

ist, wie behauptet. ■

Als Anwendung für die Berechnung bestimmter Integrale ergibt sich das folgende

**Korollar:** Ist  $F$  irgendeine differenzierbare Funktion mit der Eigenschaft, daß  $F'(x) = f(x)$  ist im Intervall  $[a, b]$ , so gilt

$$\int_a^b f(x) dx = F(b) - F(a).$$

*Beweis:* Mit obigen Bezeichnungen ist das Integral gleich  $F_a(b)$ . Zur Funktion  $F$  gibt es nach dem Hauptsatz eine Konstante  $C$ , so daß  $F(x) = F_a(x) + C$  ist. Damit ist  $F(b) - F(a) = F_a(b)$ , wie gewünscht. ■

Für „einfache“ Integranden ist dies im allgemeinen die beste Möglichkeit zur Berechnung eines Integrals; die Funktionen  $F$  haben daher einen Namen verdient:

**Definition:** Eine differenzierbare Funktion  $F$  heißt *Stammfunktion* von  $f$ , wenn  $F'(x) = f(x)$  ist; wir schreiben

$$F(x) = \int f(\xi) d\xi + C,$$

wobei die *Integrationskonstante*  $C$  die Nichteindeutigkeit der Stammfunktion ausdrückt.

Diese Nichteindeutigkeit sorgt auch dafür, daß zwar nach obigem Hauptsatz die Differentiation die Integration rückgängig macht, die Integration umgekehrt aber die Differentiation nur bis auf eine Konstante:

$$\textbf{Korollar:} \int f'(x) dx = f(x) + C$$

*Beweis:* Klar, denn  $f$  ist eine Stammfunktion von  $f'$ . ■

Falls wir eine Stammfunktion von  $f$  kennen, ist die Berechnung des Integrals

$$\int_a^b f(x) dx = F(b) - F(a)$$

also ganz einfach. Eine erste Auswahl von Stammfunktionen erhalten wir durch Rückwärtslesen von Differentiationsregeln, zum Beispiel

$$\begin{aligned} \frac{d}{dx} x^n &= nx^{n-1} \Rightarrow \int x^n dx = \frac{x^{n+1}}{n+1} && \text{falls } n \neq -1 \\ \frac{d}{dx} e^{ax} &= ae^{ax} \Rightarrow \int e^{ax} dx = \frac{e^{ax}}{a} && \text{falls } a \neq 0, \end{aligned}$$

und auch Formeln wie

$$\int \sin \omega x dx = -\frac{\cos \omega x}{\omega} + C \quad \text{und} \quad \int \cos \omega x dx = \frac{\sin \omega x}{\omega} + C$$

sind nun problemlos. Für weitere Stammfunktionen müssen wir allerdings zunächst noch einige Funktionen kennenlernen.

### f) Trigonometrische Funktionen, Hyperbelfunktionen und ihre Umkehrfunktionen

Die Ableitung einer rationalen Funktion ist wieder eine rationale Funktion; für Stammfunktionen gilt allerdings nicht dergleichen, denn wie wir gerade gesehen haben, ist schon die Stammfunktion von  $x^{-1}$  nicht mehr rational, sondern der natürliche Logarithmus. Wie wir bald sehen werden, reichen (im Reellen) auch Logarithmen noch nicht aus, um alle rationalen Funktionen integrieren zu können; wir brauchen zusätzlich noch die sogenannten Arkus- und Areefunktionen.

Die Arkusfunktionen sind die Umkehrungen der trigonometrischen Funktionen: Letztere ordnen einem Winkel eine reelle Zahl zu; die Umkehrfunktionen liefern also als Ergebnis einen Winkel oder auch Bogen, auf lateinisch *arcus* genannt.

Unter den trigonometrischen Funktionen sind zumindest der Sinus und der Cosinus aus der Analysis I bekannt; die meisten kennen wohl auch

bereits deren Quotienten

$$\tan x = \frac{\sin x}{\cos x},$$

den Tangens. Im rechtwinkligen Dreieck gibt er das Verhältnis zwischen Gegenkathete und Ankathete eines Winkels an und damit das, was man üblicherweise als *Steigung* bezeichnet: das Verhältnis der Höhendifferenz zur horizontalen Entfernung.

Da  $\cos x$  für  $x = \pi/2 + k\pi$  mit  $k \in \mathbb{Z}$  verschwindet, ist  $\tan x$  an diesen Stellen nicht definiert; in ihrer Umgebung wird er links, wo Sinus und Cosinus dasselbe Vorzeichen haben (+ für gerades  $k$ , – für ungerades) beliebig groß, rechts beliebig negativ; siehe dazu auch Abbildung 55.

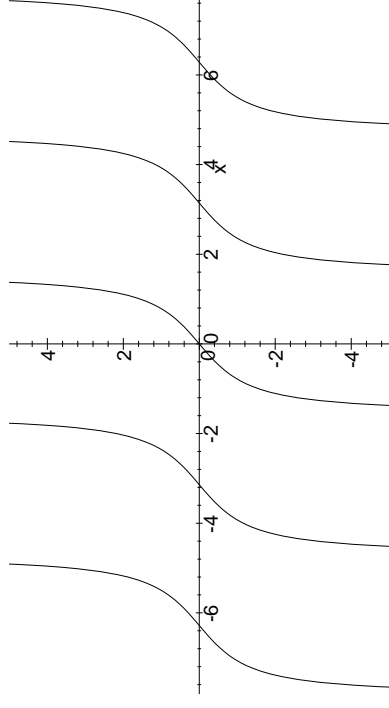


Abb. 55: Der Tangens

Aus den speziellen Werten von Sinus und Cosinus berechnet man leicht die entsprechenden Werte für Tangens:

|          |     |                      |                 |                 |                 |          |
|----------|-----|----------------------|-----------------|-----------------|-----------------|----------|
| $x$      | $0$ | $\frac{\pi}{6}$      | $\frac{\pi}{4}$ | $\frac{\pi}{3}$ | $\frac{\pi}{2}$ | $\infty$ |
| $\tan x$ | $0$ | $\frac{\sqrt{3}}{3}$ | $1$             | $\sqrt{3}$      | $\infty$        |          |

Da alle trigonometrischen Funktionen periodisch sind, können sie keine global definierten Umkehrfunktionen haben; wir müssen sie also jeweils

auf ein Intervall einschränken, auf dem sie injektiv oder – besser noch – monoton sind.

Im Fall des Sinus bietet sich das Intervall von  $-\pi/2$  bis  $\pi/2$  an, in dem er monoton von  $-1$  auf  $1$  ansteigt; wir definieren daher *den Arkussinus* oder – genauer ausgedrückt – den *Hauptwert* des Arkussinus für dieses Intervall:

$$\arcsin: [-1, 1] \rightarrow \left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$$

ist die Umkehrfunktion von

$$\sin: \left[-\frac{\pi}{2}, \frac{\pi}{2}\right] \rightarrow [-1, 1].$$

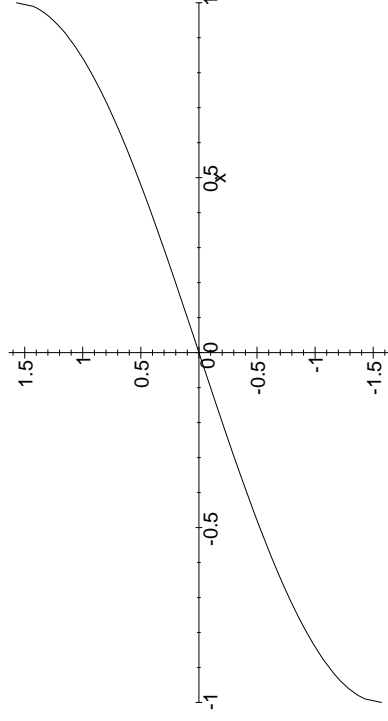


Abb. 56: Der Arkussinus

Der Cosinus fällt von seinem Maximalwert  $1$  an der Stelle Null monoton ab, bis er bei  $\pi$  den Wert  $-1$  erreicht hat. Wir definieren den *Hauptwert* der Umkehrfunktion *Arkuscossinus* oder kurz *den* Arkuscossinus daher für dieses Intervall:

$$\arccos: [-1, 1] \rightarrow [0, \pi]$$

ist die Umkehrfunktion von

$$\cos: [0, \pi] \rightarrow [-1, 1].$$

Aufgrund der Beziehung  $\cos \varphi = \sin(\frac{\pi}{2} - \varphi)$  und der Wahl der Wertebereiche ist

$$\arccos x = \frac{\pi}{2} - \arcsin x.$$

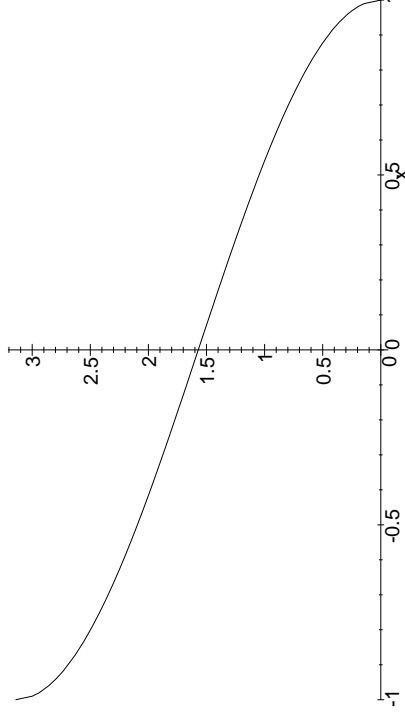


Abb. 57: Der Arkuscossinus

Der Tangens schließlich ist monoton ansteigend im offenen Intervall  $(-\pi/2, \pi/2)$  und nimmt dort jeden reellen Wert an; wir definieren den *Hauptwert* der Umkehrfunktion *Arkustangens* oder kurz *den* Arkustangens daher für dieses Intervall:

$$\arctan: \mathbb{R} \rightarrow \left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$$

ist die Umkehrfunktion von

$$\tan: \left(-\frac{\pi}{2}, \frac{\pi}{2}\right) \rightarrow \mathbb{R}.$$

Gelegentlich ganz nützlich ist der Wert an der Stelle  $1$ : Da nach obiger Tabelle  $\tan \frac{\pi}{4}$  den Wert  $1$  hat, ist  $\arctan(1) = \frac{\pi}{4}$  oder

$$\pi = 4 \arctan(1).$$

Falls man in einem Programm den Wert  $\pi$  benötigt und ihn über diese Formel als Programmkonstante definiert, kann man – eine saubere Implementierung der Arkusfunktionen vorausgesetzt – sicher sein, daß  $\pi$

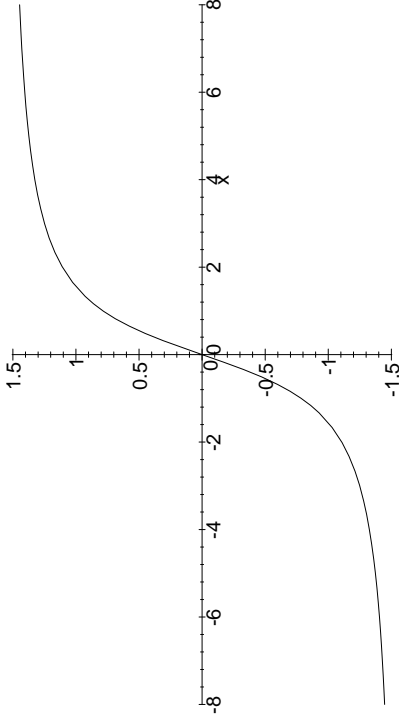


Abb. 58: Der Arkustangens

mit der maximal möglichen Stellenzahl der jeweiligen Gleitkommaziffermetik dargestellt wird – ohne daß man diese Stellenzahl oder gar  $\pi$  mit der entsprechenden Genauigkeit zu kennen braucht.

Abgesehen vom gerade definierten Hauptwert lassen sich natürlich noch in offensichtlicher Weise weitere Zweige definieren; wir wollen hier auf Einzelheiten verzichten. Der Leser sollte sich allerdings klarmachen, daß es beispielsweise für den Arkussinus sowohl monoton wachsende als auch monoton fallende Zweige gibt.

Mit Hilfe der Arkusfunktionen lassen sich die Polarkoordinatendarstellung und die kartesische Darstellung der komplexen Zahlen ineinander umrechnen: Da  $z = r e^{i\varphi}$  Realteil  $x = r \cos \varphi$  und Imaginärteil  $y = r \sin \varphi$  hat, sollte außer der bereits bekannten Formel

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

für  $z \neq 0$  auch so etwas wie

$$\varphi = \arctan \frac{y}{x} = \arccos \frac{x}{r} = \arcsin \frac{y}{r}$$

gelten. Dies im Prinzip auch richtig, allerdings nicht immer für den Hauptwert: Da alle drei Umkehrfunktionen nur Werte in einem Intervall der Länge  $\pi$  annehmen, während das Argument (genauer: der

Hauptwert des Arguments) einer komplexen Zahl im doppelt so langen Intervall  $(-\pi, \pi]$  liegt, kann obige Gleichung bei keiner der drei Funktionen für alle  $z$  gelten. Eine korrekte Formel mit dem Hauptwert des Arkuscossinus ist beispielsweise

$$\arg z = \begin{cases} \arccos \frac{\Re z}{r} & \text{falls } \Im z \geq 0 \\ -\arccos \frac{\Re z}{r} & \text{falls } \Im z < 0 \end{cases}.$$

Der für uns interessanteste Aspekt der Arkusfunktionen sind ihre *Ableitungen*, die uns neue, bislang unbekannte Stammfunktionen liefern sollen.

Wie in der Analysis I bei der Ableitung des Logarithmus gehen wir aus von der Formel

$$g'(x) = \frac{1}{f'(g(x))}$$

für die Umkehrfunktion  $g$  einer Funktion  $f$ . Da die Ableitung des Sinus der Cosinus ist, erhalten wir also

$$\arcsin'(x) = \frac{1}{\cos(\arcsin x)}.$$

Wegen  $\sin^2 y + \cos^2 y = 1$  ist

$$\cos^2(\arcsin x) = 1 - \sin^2(\arcsin x) = 1 - x^2;$$

da der Cosinus im Intervall  $[-\pi/2, \pi/2]$ , in dem der Arkussinus seine Werte annimmt, größer oder gleich Null ist, folgt

$$\cos(\arcsin x) = \sqrt{1 - x^2} \quad \text{und} \quad \arcsin'(x) = \frac{1}{\sqrt{1 - x^2}}.$$

Genauso ist

$$\arccos'(x) = \frac{-1}{\sin(\arccos x)} = \frac{-1}{\sqrt{1 - x^2}},$$

da der Sinus zwischen 0 und  $\pi$  keine negativen Werte annimmt.

Für den Arkustangens schließlich müssen wir zunächst Tangens selbst ableiten; nach der Quotientenregel ist

$$\tan'(x) = \frac{d}{dx} \frac{\sin x}{\cos x} = \frac{\cos^2 x + \sin^2 x}{\cos^2 x} = 1 + \tan^2 x$$

und dementsprechend

$$\arctan'(x) = \frac{1}{1 + \tan^2(\arctan x)} = \frac{1}{1 + x^2}.$$

Da sich die Ableitungen von Arkussinus und Arkuscossinus nur im Vorzeichen unterscheiden (was natürlich von vornherein klar war, da sich die Werte der beiden Funktionen stets zu  $\pi/2$  ergänzen) sind als Stammfunktionen vor allem der Arkussinus und der Arkustangens interessant; wir haben die beiden neuen Formeln

$$\int \frac{dx}{1+x^2} dx = \arctan x + C \quad \text{und} \quad \int \frac{dx}{\sqrt{1-x^2}} dx = \arcsin x + C.$$

Als nächstes wollen wir uns überlegen, wie wir die Integrale

$$\int \frac{dx}{1-x^2} dx \quad \text{und} \quad \int \frac{dx}{\sqrt{1+x^2}} dx$$

mit dem jeweils anderen Vorzeichen bekommen können. Eine offensichtliche Lösung besteht darin, einfach  $x$  durch  $ix$  zu ersetzen, und das legt es nahe, in Analogie zu den EULERSchen Formeln

$$\cos x = \frac{e^{ix} + e^{-ix}}{2} \quad \text{und} \quad \sin x = \frac{e^{ix} - e^{-ix}}{2i}$$

neue Funktionen

$$\cosh x = \frac{e^x + e^{-x}}{2} \quad \text{und} \quad \sinh x = \frac{e^x - e^{-x}}{2}$$

zu definieren, den *Cosinus hyperbolicus* und den *Sinus hyperbolicus*.

Der Zusatz *hyperbolicus* läßt sich leicht verstehen: Genau wie der gewöhnliche Sinus und Cosinus als *Kreisfunktionen* bezeichnet werden, weil die Punkte  $(\sin t, \cos t)$  auf dem Kreis  $x^2 + y^2 = 1$  liegen, haben die Hyperbelfunktionen ihren Namen daher, daß

$$\begin{aligned} \cosh^2 t - \sinh^2 t &= \left( \frac{e^t + e^{-t}}{2} \right)^2 - \left( \frac{e^t - e^{-t}}{2} \right)^2 \\ &= \frac{e^{2t} + 2 + e^{-2t}}{4} - \frac{e^{2t} - 2 + e^{-2t}}{4} = 1 \end{aligned}$$

ist, so daß die Punkte  $(\cosh t, \sinh t)$  auf der Hyperbel  $x^2 - y^2 = 1$  liegen, d.h.

$$\cosh^2 x - \sinh^2 x = 1 \quad \text{für alle } x \in \mathbb{R}.$$

Damit ist insbesondere  $\cosh^2 x$  stets  $\geq 1$ ; da der Cosinus hyperbolicus genau wie die beiden Exponentialfunktionen, aus denen er zusammengesetzt ist, keine negativen Werte annehmen kann, folgt also

$$\cosh x \geq 1 \quad \text{für alle } x \in \mathbb{R}.$$

Sofort aus der Definition folgen die Ableitungsregeln

$$\cosh'(x) = \sinh x \quad \text{und} \quad \sinh' x = \cosh x;$$

auf Grund obiger Ungleichung ist daher der Sinus hyperbolicus monoton steigend auf ganz  $\mathbb{R}$ . Man sieht leicht, daß sein Vorzeichen jeweils das von  $x$  ist (Insbesondere ist also auch  $\sinh 0 = 0$ ), d.h. der Cosinus hyperbolicus ist monoton fallend für negative und monoton steigend für positive  $x$ .

Für große Werte von  $x$  wird  $e^{-x}$  beliebig klein, sowohl  $\sinh x$  als auch  $\cosh x$  unterscheiden sich für solche Werte also beliebig wenig von  $\frac{1}{2}e^x$  und gehen insbesondere gegen unendlich. Für stark negative Werte wird  $e^x$  beliebig klein und das Verhalten beider Funktionen wird dominiert durch den Term  $\pm e^{-x}$ . Daher geht  $\sinh x$  für  $x \rightarrow -\infty$  gegen  $-\infty$  und  $\cosh x$  gegen  $+\infty$ . Offensichtlich ist  $\sinh x$  eine ungerade,  $\cosh x$  aber eine gerade Funktion. In Abbildung 59 sind gestrichelt auch noch die Funktionen  $\frac{1}{2}e^x$  im positiven und  $\pm \frac{1}{2}e^{-x}$  im negativen Bereich eingezeichnet; wie man sieht, ist die Übereinstimmung mit  $\sinh x$  bzw.  $\cosh x$  schon ab etwa  $|x| \geq 2$  recht gut.

Der Graph des Cosinus hyperbolicus erinnert an eine durchhängende Kette, und tatsächlich kann die Variationsrechnung zeigen, daß eine Kette mit reibungsfrei gegeneinander beweglichen Gliedern genau diese Form hat.

In völliger Analogie zum Tangens definieren wir auch noch einen *Tangens hyperbolicus* durch die Vorschrift

$$\tanh x \stackrel{\text{def}}{=} \frac{\sinh x}{\cosh x} = \frac{e^x + e^{-x}}{e^x - e^{-x}};$$

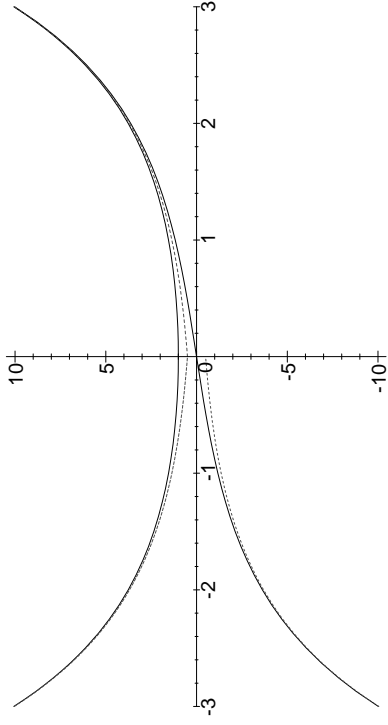


Abb. 59: Sinus hyperbolicus und Cosinus hyperbolicus

da der Cosinus hyperbolicus stets größer oder gleich eins ist, kann der Nenner nie Null werden, so daß die Funktion auf ganz  $\mathbb{R}$  definiert ist. Dabei ist

$$\begin{aligned}\lim_{x \rightarrow \infty} \tanh x &= \lim_{x \rightarrow \infty} \frac{e^x + e^{-x}}{e^x - e^{-x}} = \lim_{x \rightarrow \infty} \frac{e^x(1 + e^{-2x})}{e^x(1 - e^{-2x})} \\ &= \lim_{x \rightarrow \infty} \frac{1 + e^{-2x}}{1 - e^{-2x}} = 1\end{aligned}$$

und entsprechend

$$\lim_{x \rightarrow -\infty} \tanh x = \lim_{x \rightarrow -\infty} \frac{e^x + e^{-x}}{e^x - e^{-x}} = \lim_{x \rightarrow -\infty} \frac{e^{2x} + 1}{e^{2x} - 1} = -1.$$

Die Ableitung des Tangens hyperbolicus ist nach der Quotientenregel

$$\tanh' x = \frac{\cosh^2 x - \sinh^2 x}{\cosh^2 x} = \begin{cases} \frac{1}{\cosh^2 x} & \text{nach obiger Formel} \\ 1 - \tanh^2 x & \text{durch Ausdividieren} \end{cases},$$

wobei wie beim Tangens je nach Anwendung mal die eine, mal die andere Form des Ergebnisses nützlicher ist.

Aus beiden Ausdrücken sieht man sofort, daß die Ableitung stets positiv ist, der Tangens hyperbolicus steigt also monoton von  $-1$  nach  $+1$ , wobei beide Werte nur asymptotisch angenommen werden.

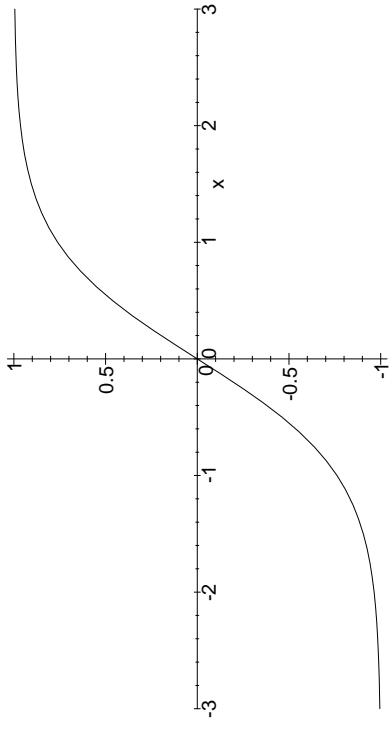


Abb. 60: Der Tangens hyperbolicus

Die Umkehrfunktionen der Hyperbelfunktionen werden als *Areafunktionen* bezeichnet und heißen *Areasinus hyperbolicus*, *Areacosinus hyperbolicus* und *Areatangens hyperbolicus*.

Aus obiger Diskussion folgt, daß der Areasinus hyperbolicus

$$\operatorname{arsinh}: \mathbb{R} \rightarrow \mathbb{R}$$

auf ganz  $\mathbb{R}$  definiert ist und der Areatangens hyperbolicus

$$\operatorname{artanh}: (-1, 1) \rightarrow \mathbb{R}$$

nur auf dem offenen Einheitsintervall. Der Cosinus hyperbolicus hat keine eindeutig bestimmte Umkehrfunktion, da er für positive und für negative Argumente jeweils denselben Wert annimmt. Wir definieren den Areacosinus hyperbolicus

$$\operatorname{arcosh}: \mathbb{R}_{\geq 1} \rightarrow \mathbb{R}_{\geq 0}$$

als die Umkehrfunktion des positiven Zweigs.

Da Sinus hyperbolicus und Cosinus hyperbolicus asymptotisch wie eine Exponentialfunktion ansteigend, steigen Areasinus hyperbolicus und Areacosinus hyperbolicus wie Logarithmen, also sehr langsam.

Auch diese Funktionen sind wieder vor allem wegen ihrer Ableitungen interessant; eine weitgehend zum Fall der Arkusfunktionen analoge

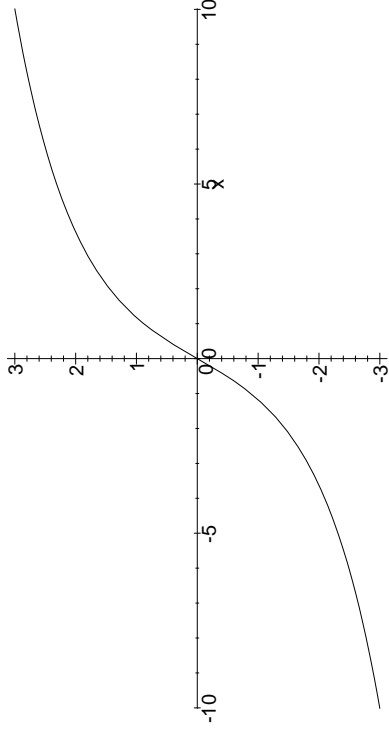


Abb. 61: Der Areasinus hyperbolicus

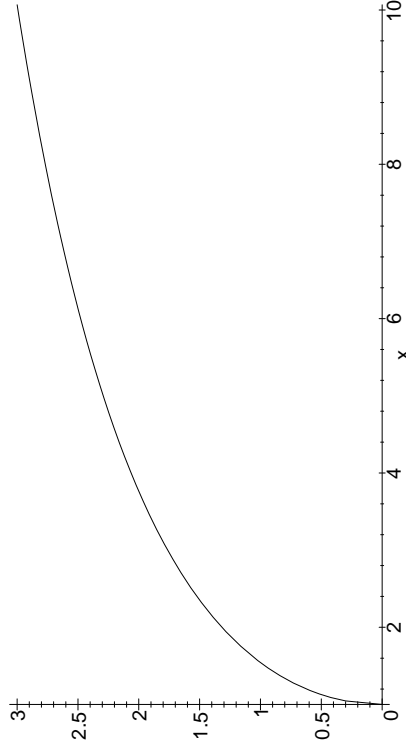


Abb. 62: Der Areacosinus hyperbolicus

Rechnung zeigt, daß

$$\operatorname{arsinh}'(x) = \frac{1}{\cosh(\operatorname{arsinh} x)} = \frac{1}{\sqrt{1+x^2}}$$

und

$$\operatorname{arcosh}'(x) = \frac{1}{\sinh(\operatorname{arcosh} x)} = \frac{1}{\sqrt{x^2-1}}$$

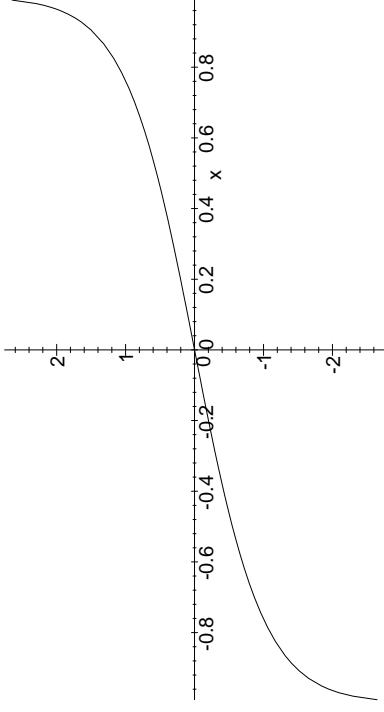


Abb. 63: Der Areatangens hyperbolicus

ist. Für den Areatangens hyperbolicus schließlich ist

$$\operatorname{artanh}'(x) = \frac{1}{1 - \tanh^2(\operatorname{artanh} x)} = \frac{1}{1 - x^2}.$$

Dies führt also auf die neuen Stammfunktionen

$$\int \frac{dx}{\sqrt{1+x^2}} = \operatorname{arsinh}(x) + C, \quad \int \frac{dx}{\sqrt{x^2-1}} = \operatorname{arcosh}(x) + C$$

und

$$\int \frac{dx}{1-x^2} = \operatorname{artanh}(x) + C.$$

Die letztere wird allerdings nicht unbedingt gebraucht, denn da

$$\frac{1}{1-x^2} = \frac{1}{2} \cdot \left( \frac{1}{1-x} + \frac{1}{1+x} \right)$$

ist, können wir dieses Integral auch ausrechnen als

$$\int \frac{dx}{1-x^2} = \frac{\ln|1-x| + \ln|1+x|}{2} + C,$$

wobei diese Formel noch den Vorteil eines größeren Definitionsbereichs hat.



### g) Partielle Integration

Eine wichtige Regel der Differentialrechnung ist die *Produktregel*

$$\frac{d}{dx}(f(x)g(x)) = f'(x)g(x) + f(x)g'(x).$$

Für die Zwecke der Integralrechnung schreiben wir sie besser als

$$f'(x)g(x) = \frac{d}{dx}(f(x)g(x)) - f(x)g'(x),$$

was integriert auf

$$\int f'(x)g(x) dx = f(x)g(x) - \int f(x)g'(x) dx$$

führt – die Regel der partiellen Integration. Sie ist dann nützlich, wenn sich der Integrand als Produkt schreiben läßt, wobei einer der Faktoren eine bekannte Stammfunktion hat; gelegentlich ist dann das Produkt  $f(x)g'(x)$  leichter integrierbar als  $f'(x)g(x)$ .

Mit den wenigen Funktionen, die wir bislang kennen, lassen sich nur wenige interessante Beispiele konstruieren; die volle Kraft dieser Regel werden wir erst später kennenlernen, wenn wir unser Repertoire an Funktionen erweitert haben. Hier sei nur  $\int xe^x dx$  betrachtet: Da die Exponentialfunktion ihre eigene Ableitung und Stammfunktion ist, empfiehlt sie sich für die Rolle der Funktion  $f$ , während  $g(x) = x$  mit  $g'(x) = 1$  auf Vereinfachungen auf der rechten Seite hoffen läßt. In der Tat führt partielle Integration mit

$$\int xe^x dx = xe^x - \int e^x dx = xe^x - e^x + C = (x - 1)e^x + C$$

auf eine Stammfunktion. Genauso lassen sich rekursiv auch die Integrale  $\int x^n e^x dx$  berechnen:

$$\int x^2 e^x dx = x^2 e^x - \int 2xe^x dx = (x^2 - 2x + 2)e^x + C$$

und

$$\int x^n e^x dx = x^n e^x - n \int x^{n-1} e^x dx.$$

### h) Substitutionsregel

Auch die Kettenregel

$$\frac{d}{dx}f(g(x)) = f'(g(x)) \cdot g'(x)$$

läßt sich umschreiben zu einer Integrationsregel, der *Substitutionsregel*

$$\int f(g(x)) \cdot g'(x) dx = F(g(x)) + C,$$

wobei  $F$  eine Stammfunktion von  $f$  ist.

Diese Regel dürfte wohl meist der erfolgversprechendste Versuch zum Auffinden einer Stammfunktion sein – vor allem, wenn man sie von rechts nach links liest, um für eine bekannte Funktion  $f$  die unbekannte Stammfunktion  $F$  zu berechnen. Als Substitution  $g$  wählt man hier eine „geeignete“ bijektive Funktion, die zu einer Vereinfachung auf der linken Seite führt.

**1) Der Spezialfall logarithmischer Ableitungen:** Für  $f(x) = 1/x$  führt die Substitutionsregel

$$\int f(g(x))g'(x) dx = F(g(x)) + C \quad \text{mit} \quad F'(x) = f(x)$$

zur Integrationsregel

$$\int \frac{g'(x)}{g(x)} dx = \ln |g(x)| + C.$$

Als erste Anwendung betrachten wir

$$\tan x = \frac{\sin x}{\cos x} = -\frac{g'(x)}{g(x)}$$

mit  $g(x) = \cos x$ . Nach der gerade bewiesenen Regel ist daher

$$\int \tan x dx = -\ln |\cos(x)| + C,$$

eine Funktion, die genau wie der Tangens selbst an den Nullstellen der Cosinusfunktion nicht definiert ist.

Beim Tangens hyperbolicus gibt es keine Probleme mit dem Vorzeichen; hier ist ganz entsprechend

$$\int \tanh x \, dx = \ln |\cosh(x)| + C = \ln \cosh(x) + C,$$

da der Cosinus hyperbolicus nur positive Werte annimmt.

Auch Integrale wie

$$\int \frac{x}{1+x^2} dx$$

lassen sich nach dieser Regel ausrechnen: Da die Ableitung des Nenners  $2x$  ist, folgt

$$\int \frac{x}{1+x^2} dx = \frac{1}{2} \ln |1+x^2| + C = \frac{1}{2} \ln(1+x^2) + C.$$

Damit können wir dann beispielsweise den Arkustangens integrieren: Partielle Integration zeigt, daß

$$\begin{aligned} \int 1 \cdot \arctan x \, dx &= x \arctan x - \int \frac{x}{1+x^2} dx + C \\ &= x \arctan x - \frac{1}{2} \ln(1+x^2) + C \end{aligned}$$

ist.

**2) Substitutionen mit linearen Funktionen:** Eine der elementarsten Anwendungen der Substitutionsregel, auf die man üblicherweise auch ohne diese Regel kommt, ist die Substitution mit linearen Funktionen  $g(x) = ax + b$ ; hier besagt die Substitutionsregel, daß

$$\int f(ax+b) \cdot a \, dx = F(ax+b) \quad \text{mit} \quad F'(x) = f(x)$$

ist, oder besser

$$\int f(ax+b) \, dx = \frac{F(ax+b)}{a} \quad \text{mit} \quad F'(x) = f(x).$$

Somit ist beispielsweise

$$\int \sin(\omega t + \varphi) \, dt = \frac{-\cos(\omega t + \varphi)}{\omega} + C$$

oder

$$\int \frac{dx}{x^2+a^2} = \int \frac{\frac{dx}{a^2}}{\left(\frac{x}{a}\right)^2+1} = \frac{1}{a} \arctan\left(\frac{x}{a}\right) + C.$$

Entsprechend berechnet man auch

$$\int \frac{dx}{x^2-a^2} = -\frac{1}{a} \operatorname{artanh}\left(\frac{x}{a}\right) + C$$

und sogar ganz allgemein  $\int \frac{dx}{ax^2+bx+c}$ :

Da der Fall  $a = 0$  uninteressant ist (bzw. eine unbedingt notwendige Übungsaufgabe für alle, die nicht sofort einsehen warum), schließen wir dies aus und schreiben

$$ax^2+bx+c = a \left(x + \frac{b}{2a}\right)^2 + c - \frac{b^2}{4a}.$$

Mit

$$y = x + \frac{b}{2a} \quad \text{und} \quad d = \frac{c}{a} - \frac{b^2}{4a^2}$$

ist dann

$$\begin{aligned} \int \frac{dx}{ax^2+bx+c} &= \frac{1}{a} \int \frac{dy}{y^2+d} \\ &= \frac{1}{a} \begin{cases} \frac{1}{\sqrt{d}} \arctan\left(\frac{y}{\sqrt{d}}\right) + C & \text{falls } d > 0 \\ \frac{1}{y} & \text{falls } d = 0 \\ \frac{1}{\sqrt{-d}} \operatorname{artanh}\left(\frac{y}{\sqrt{-d}}\right) + C & \text{falls } d < 0, \end{cases} \end{aligned}$$

und je nach Vorzeichen von  $d$  kann dies entweder über eine der obigen Formeln oder (für  $d = 0$ ) elementar ausgerechnet werden. Insgesamt erhalten wir mit  $\Delta = b^2 - 4ac$  (und damit  $d = -\Delta/4a^2$ )

$$\int \frac{dx}{ax^2+bx+c} = \begin{cases} \frac{2}{\sqrt{-\Delta}} \arctan\left(\frac{2ax+b}{\sqrt{-\Delta}}\right) + C & \text{falls } \Delta < 0 \\ \frac{2}{2ax+b} + C & \text{falls } \Delta = 0 \\ \frac{2}{\sqrt{\Delta}} \operatorname{artanh}\left(\frac{2ax+b}{\sqrt{\Delta}}\right) + C & \text{falls } \Delta > 0. \end{cases}$$

### 3) Substitutionen mit trigonometrischen und Hyperbelfunktionen:

Wegen der Beziehungen

$$\sin^2 x + \cos^2 x = 1 \quad \text{und} \quad \cosh^2 x - \sinh^2 x = 1$$

bieten sich Substitutionen mit trigonometrischen Funktionen und Hyperbelfunktionen an bei Integralen, in denen Ausdrücke der Form  $\sqrt{1 \pm x^2}$  und ähnliche vorkommen.

Betrachten wir als einfachstes Beispiel die Stammfunktion von  $\sqrt{1-x^2}$  selbst. Mit der Substitution  $x = \sin t$  mit  $-\frac{\pi}{2} \leq t \leq \frac{\pi}{2}$  erhalten wir (durch Rückwärtslesen der Substitutionsregel)

$$\int \sqrt{1-x^2} dx = \int \sqrt{1-\sin^2 t} \cos t dt,$$

und dies ist  $\int \cos^2 t dt$ , da Cosinus zwischen  $-\frac{\pi}{2}$  und  $\frac{\pi}{2}$  nur nichtnegative Werte annimmt.

Dieses Integral kennen wir von Aufgabe 1b) des dritten Übungsblatts:

$$\int \cos^2 t dt = \frac{1}{2}(\sin t \cos t + t) + C.$$

Mit der Rücksubstitution  $t = \arcsin x$  erhalten wir schließlich nach der kurzen Nebenrechnung

$$\cos(\arcsin x) = \sqrt{1-\sin^2(\arcsin x)} = \sqrt{1-x^2}$$

das Ergebnis

$$\int \sqrt{1-x^2} dx = \frac{1}{2}(x\sqrt{1-x^2} + \arcsin x) + C.$$

Aus ähnliche Weise läßt sich auch die Stammfunktion von  $\sqrt{1+x^2}$  bestimmen: Hier bietet sich die Substitution  $x = \sinh t$  an und wir erhalten

$$\int \sqrt{1+x^2} dx = \int \sqrt{1-\sinh^2 t} \cosh t dt = \int \cosh^2 t dt,$$

da  $\cosh t$  nur positive Werte annimmt.

Leider kennen wir das rechtsstehende Integral noch nicht; angesichts der großen Ähnlichkeiten zwischen trigonometrischen Funktionen und

Hyperbelfunktionen lohnt es sich aber sicherlich, als ersten Ansatz vor einer partiellen Integration zu schauen, ob vielleicht etwas analoges gilt wie oben. In der Tat ist nach der Produktregel

$$\frac{d}{dt} \frac{1}{2}(\sinh t \cosh t + t) = \cosh^2 t + C,$$

also

$$\int \cosh^2 t dt = \frac{1}{2}(\sinh t \cosh t + t) + C.$$

Somit ist

$$\int \sqrt{1+x^2} dx = \frac{1}{2}(x\sqrt{1+x^2} + \operatorname{arsinh} x) + C.$$

Als letztes Beispiel, in dem  $x$  einmal nicht quadratisch vorkommt, betrachten wir noch

$$\int \sqrt{\frac{1-x}{1+x}} dx.$$

Da es in diesem Abschnitt um Substitutionen mit trigonometrischen Funktionen geht, versuchen wir es wieder mit dem Ansatz  $x = \sin t$  für  $-\frac{\pi}{2} \leq t \leq \frac{\pi}{2}$  und erhalten

$$\int \sqrt{\frac{1-x}{1+x}} dx = \int \sqrt{\frac{1-\sin t}{1+\sin t}} \cos t dt.$$

Hier hilft nun ein Trick weiter: Erweitern wir den Bruch unter der Quadratwurzel mit  $1-\sin t$ , so wird das Integral zu

$$\begin{aligned} \int \sqrt{\frac{(1-\sin t)^2}{1-\sin^2 t}} \cos t dt &= \int \frac{1-\sin t}{\cos t} \cos t dt = \int (1-\sin t) dt \\ &= t + \cos t + C. \end{aligned}$$

Damit ist

$$\begin{aligned} \int \sqrt{\frac{1-x}{1+x}} dx &= \arcsin x + \cos(\arcsin x) + C \\ &= \arcsin x + \sqrt{1-x^2} + C. \end{aligned}$$

Bei einem so komplizierten Integral empfiehlt es sich, das Ergebnis durch Differentiation zu überprüfen; wir erhalten

$$\frac{d}{dx} \left( \arcsin x + \sqrt{1-x^2} \right) = \frac{1}{\sqrt{1-x^2}} - \frac{x}{\sqrt{1-x^2}},$$

was zunächst so aussieht, als sei irgend etwas falsch gelaufen. Beachtet man aber, daß

$$\sqrt{1+x} \cdot \sqrt{1-x} = \sqrt{1-x^2}$$

ist, so ist

$$\frac{1}{\sqrt{1-x^2}} - \frac{x}{\sqrt{1-x^2}} = \frac{1-x}{\sqrt{1-x^2}} = \frac{\sqrt{1-x}}{\sqrt{1+x}},$$

was erheblich vertrauenserweckender aussieht.

**4) Integrale der Form  $\int h(e^{ax}) dx$ :** Bei solchen Integralen führt oft die Substitution  $u = e^{ax}$  oder, anders ausgedrückt,  $x = g(u) = \frac{1}{a} \ln u$  zu einer Vereinfachung. Die Substitutionsregel

$$\int f(g(u)) g'(u) du = \int f(x) dx$$

wird wegen  $g'(u) = \frac{dx}{du} = \frac{1}{au}$  zu

$$\int h(u) \frac{du}{au} = \int h(e^{ax}) dx;$$

insbesondere wird das Integral also bei rationalem  $h$  auf ein Integral einer rationalen Funktion zurückgeführt, und dieses kann nach dem im nächsten Abschnitt skizzierten Verfahren der Partialbruchzerlegung berechnet werden.

Im allgemeinen schreibt man eine Substitution wie die obige kurz in der Form

$$u = e^{ax}, \quad dx = \frac{du}{au},$$

wobei man letztere Beziehung auch aus

$$\frac{du}{dx} = ae^{ax} = au \quad \text{oder} \quad du = audx$$

herleiten kann.

Betrachten wir als Beispiel das Integral  $\int \frac{e^{3x}}{e^{2x}-1} dx$ : Da sowohl  $e^{2x}$  als auch  $e^{3x}$  im Integranden vorkommen, empfiehlt sich die Substitution  $u = e^x$  mit  $du = u dx$ ; damit wird

$$\int \frac{e^{3x}}{e^{2x}-1} dx = \int \frac{u^3}{u^2-1} \frac{du}{u} = \int \frac{u^2}{u^2-1} du,$$

und wir sind beim Integral einer rationalen Funktion angelangt. Elementare Bruchrechnung zeigt, daß

$$\frac{u^2}{u^2-1} = 1 + \frac{1}{u^2-1} = 1 + \frac{\frac{1}{2}}{u-1} - \frac{\frac{1}{2}}{u+1},$$

d.h.

$$\frac{u^2}{u^2-1} du = \int du + \frac{1}{2} \left( \int \frac{du}{u-1} - \int \frac{du}{u+1} \right) = u + \ln \left| \frac{u-1}{u+1} \right| + C.$$

(Für komplizierte Funktionen werden wir im nächsten Abschnitt ein Verfahren kennenlernen, daß rationale Funktionen in Summen aus einfachen Summanden zerlegt.)

Nun muß nur noch die Substitution  $u = e^x$  rückgängig gemacht werden, und wir erhalten als Ergebnis

$$\int \frac{e^{3x}}{e^{2x}-1} dx = e^x + \frac{1}{2} \ln \left| \frac{e^x-1}{e^x+1} \right| + C.$$

Dieselbe Technik funktioniert natürlich auch bei Integralen über Hyperbelfunktionen, da man diese genauso gut als Ausdrücke in Exponentialfunktionen schreiben kann. Beispielsweise ist

$$\begin{aligned} \int \frac{dx}{\sinh x} &= 2 \int \frac{dx}{e^x - e^{-x}} = 2 \int \frac{du}{u \left( u - \frac{1}{u} \right)} = 2 \int \frac{du}{u^2 - 1} \\ &= \ln \left| \frac{u-1}{u+1} \right| + C = \ln \left| \frac{e^x-1}{e^x+1} \right| + C. \end{aligned}$$

5) **Integrale der Form**  $\int h(\sin x, \cos x) dx$ : Als letzte Anwendung der Substitutionsregel in diesem Abschnitt wollen wir noch Integrale betrachten, die von trigonometrischen Funktionen abhängen. Hier führt oft die Substitution

$$x = 2 \arctan t, \quad dx = \frac{2 dt}{t^2 + 1}$$

zum Erfolg; sie macht  $\sin x$  zu  $\sin(2 \arctan t)$  und  $\cos x$  zu  $\cos(2 \arctan t)$ , was wir wie folgt ausrechnen können: Aus

$$\tan^2 x = \frac{\sin^2 x}{\cos^2 x} = \frac{1 - \cos^2 x}{\cos^2 x} = \frac{1}{\cos^2 x} - 1$$

folgt, daß

$$\cos(2x) = 2 \cos^2(x) - 1 = \frac{2}{1 + \tan^2 x} - 1 = \frac{1 - \tan^2 x}{1 + \tan^2 x}$$

ist und somit

$$\cos(2 \arctan t) = \frac{1 - t^2}{1 + t^2}.$$

Da der Arkustangens nur Werte zwischen  $-\frac{\pi}{2}$  und  $\frac{\pi}{2}$  annimmt, liegt  $2 \arctan t$  zwischen  $-\pi$  und  $\pi$ , ein Intervall, in dem der Sinus dasselbe Vorzeichen hat wie sein Argument. Daher ist

$$\sin(2 \arctan t) = \sqrt{1 - \cos^2(2 \arctan t)} = \sqrt{1 - \frac{(1 - t^2)^2}{(1 + t^2)^2}} = \frac{2t}{1 + t^2}.$$

Als Beispiel betrachten wir

$$\int \frac{dx}{\cos x} = \int \frac{1+t^2}{1-t^2} \frac{2 dt}{1+t^2} = -2 \int \frac{dt}{t^2 - 1}.$$

Das letzte Integral kennen wir:

$$\int \frac{dt}{t^2 - 1} = \frac{1}{2} \ln \left| \frac{1-t}{1+t} \right| + C.$$

Also ist

$$\int \frac{dx}{\cos x} = -\frac{2}{2} \ln \left| \frac{1-t}{1+t} \right| + C = \ln \left| \frac{1+t}{1-t} \right| + C = \ln \left| \frac{1 - \tan \frac{x}{2}}{1 + \tan \frac{x}{2}} \right| + C.$$

### i) Integration rationaler Funktionen

Bekanntlich ist ein Polynom in  $x$  ein Ausdruck der Form

$$f(x) = \sum_{\nu=0}^n a_{\nu} x^{\nu} = a_n x^n + \dots + a_1 x + a_0;$$

ist  $a_n \neq 0$ , so bezeichnet man  $n$  als den *Grad* des Polynoms, in Zeichen

$$n = \deg f.$$

Polynome vom Grad eins, zwei, drei bzw. vier heißen linear, quadratisch, kubisch bzw. biquadratisch.

Unter einer *rationalen Funktion* versteht man eine Funktion der Form

$$x \mapsto \frac{f(x)}{g(x)},$$

wobei  $f(x)$  und  $g(x)$  Polynome sind. Es ist klar, daß diese Funktion nur in jenen Punkten  $x \in \mathbb{R}$  definiert ist, in denen  $g(x) \neq 0$  ist.

Die Integrationstheorie rationaler Funktionen ist seit langem bekannt; im Prinzip kann man zu jeder rationalen Funktion eine Stammfunktion angeben, indem man wie folgt vorgeht: (Auf Einzelheiten und Beweise muß hier weitgehend verzichtet werden, da wir nicht über das notwendige mathematische Instrumentarium verfügen.)

**1. Schritt:** Man zerlege den Nenner in ein Produkt aus linearen und quadratischen Polynomen. Die ist zumindest grundsätzlich möglich, denn nach dem *Fundamentalsatz der Algebra* läßt sich ein Polynom vom Grad  $n$  (mir reellen oder komplexen Koeffizienten) stets in der Form

$$g(x) = a_n (x - z_1)(x - z_2) \cdots (x - z_n)$$

schreiben, wobei die *komplexen* Zahlen  $z_i$  die Nullstellen von  $g$  sind. Wie man sich leicht überlegt, folgt für ein reelles Polynom  $g$  aus  $g(z) = 0$  sofort, daß auch  $g(\bar{z}) = 0$  ist, d.h. die nichtreellen Nullstellen treten als Paare konjugiert komplexer Zahlen auf, die auch jeweils die gleiche Vielfachheit haben, da letztere über das Verschwinden von Ableitungen definiert werden kann.

Das Produkt der Linearfaktoren zu zwei konjugiert komplexen Zahlen ist

$$(x - z)(x - \bar{z}) = x^2 - (z + \bar{z})x + z\bar{z} = x^2 - 2\Re z \cdot x + |z|^2,$$

ein quadratisches Polynom mit reellen Koeffizienten. Also läßt sich  $g(x)$  als Produkt linearer und quadratischer Polynome mit reellen Koeffizienten schreiben. Faßt man gleiche Faktoren zusammen, so erhält man demnach eine Darstellung

$$g(x) = a_n \ell_1(x)^{e_1} \cdot \dots \cdot \ell_r(x)^{e_r} \cdot q_1(x)^{f_1} \cdot \dots \cdot q_s(x)^{f_s}$$

mit linearen Polynomen  $\ell_i$ , quadratischen Polynomen  $q_j$  und natürlichen Zahlen  $e_i, f_j$ . Dabei sind sowohl die  $\ell_i$  als auch die  $q_j$  paarweise voneinander verschieden.

Diese Zerlegung kann algorithmisch selbst bei Polynomen mit rationalen Koeffizienten algorithmisch sehr aufwendig sein; daher kommt der obige Zusatz „im Prinzip“ für die Durchführbarkeit der Integration rationaler Funktionen.

**2. Schritt:** Man schreibe die Funktion  $f(x)/g(x)$  als Summe eines Polynoms und von *Partialbrüchen* der Form

$$\frac{\alpha_{jk}x + \gamma_{jk}}{\ell_j(x)^k} \quad \text{und} \quad \frac{\beta_{jk}x + \gamma_{jk}}{q_j(x)^k},$$

wobei  $k$  von eins bis  $e_i$  bzw.  $f_j$  läuft und  $\alpha_{ik}, \beta_{jk}$  und  $\gamma_{jk}$  reelle Zahlen sind, d.h.

$$\frac{f(x)}{g(x)} = Q(x) + \sum_{i=1}^r \sum_{k=1}^{e_i} \frac{\alpha_{ik}}{\ell_i(x)^k} + \sum_{j=1}^s \sum_{k=1}^{f_j} \frac{\beta_{jk}x + \gamma_{jk}}{q_j(x)^k}.$$

Wir müssen uns zunächst überlegen, warum das möglich ist:

Beginnen wir mit zwei teilerfremden Polynomen  $p(x)$  und  $q(x)$ . Der erweiterte EUKLIDISCHE Algorithmus aus Kapitel I, §2e) liefert dazu neue Polynome  $\alpha(x), \beta(x)$ , so daß

$$\alpha(x)p(x) + \beta(x)q(x) = 1$$

ist und damit

$$\frac{1}{p(x)q(x)} = \frac{\alpha(x)p(x) + \beta(x)q(x)}{p(x)q(x)} = \frac{\alpha(x)}{q(x)} + \frac{\beta(x)}{p(x)}.$$

Ist ein drittes Polynom  $r(x)$  teilerfremd sowohl zu  $p(x)$  als auch zu  $q(x)$ , sind insbesondere auch  $p(x)q(x)$  und  $r(x)$  teilerfremd; es gibt also eine Darstellung

$$\frac{1}{p(x)q(x)r(x)} = \frac{\phi(x)}{p(x)q(x)} + \frac{\psi(x)}{r(x)} = \frac{\phi(x)\alpha(x)}{q(x)} + \frac{\phi(x)\beta(x)}{p(x)} + \frac{\psi(x)}{r(x)}$$

mit zwei neuen Polynomen  $\phi(x)$  und  $\psi(x)$ , insgesamt also wieder eine Summe dreier Brüche mit den drei Faktoren  $p(x), q(x)$  und  $r(x)$  als Nennern.

Induktiv folgt auf diese Weise, daß es zu  $n$  paarweise teilerfremden Polynomen  $p_1(x), \dots, p_n(x)$  stets Polynome  $\alpha_1(x), \dots, \alpha_n(x)$  gibt, so daß

$$\frac{1}{p_1(x) \cdots p_n(x)} = \frac{\alpha_1(x)}{p_1(x)} + \dots + \frac{\alpha_n(x)}{p_n(x)}$$

ist. Für ein beliebiges weiteres Polynom ist daher auch

$$\frac{f(x)}{p_1(x) \cdots p_n(x)} = \frac{\alpha_1(x)f(x)}{p_1(x)} + \dots + \frac{\alpha_n(x)f(x)}{p_n(x)}.$$

Dies wenden wir an auf die obige Zerlegung des Nenners  $g$ : Da die  $\ell_i(x)$  und die  $q_j(x)$  paarweise teilerfremd sind, gilt dasselbe auch für ihre Potenzprodukte, es gibt also Polynome  $\varphi_i(x), \psi_j(x)$ , so daß gilt

$$\frac{f(x)}{g(x)} = \sum_{i=1}^r \frac{\varphi_i(x)}{\ell_i(x)^{e_i}} + \sum_{j=1}^s \frac{\psi_j(x)}{q_j(x)^{f_j}}.$$

Um den Grad dieser Polynome zu reduzieren, dividieren wir für jeden Bruch den Zähler mit Rest durch den Nenner und fassen die Quotientenpolynome zusammen zu einem einzigen Polynom  $Q(x)$ . Damit erhalten wir eine neue Darstellung

$$\frac{f(x)}{g(x)} = Q(x) + \sum_{i=1}^r \frac{\tilde{\varphi}_i(x)}{\ell_i(x)^{e_i}} + \sum_{j=1}^s \frac{\tilde{\psi}_j(x)}{q_j(x)^{f_j}},$$

in der jeder Zähler  $\tilde{\varphi}_i(x)$  und  $\tilde{\psi}_j(x)$  kleineren Grad hat als der zugehörige Nenner.

Zur weiteren Zerlegung der Zähler überlegen wir uns zunächst, daß es zu je zwei Polynomen  $\varphi(x)$  und  $g(x)$  eine Darstellung

$$\varphi(x) = \alpha_0(x) + \alpha_1(x)g(x) + \alpha_2g(x)^2 + \dots$$

gibt mit  $\deg \alpha_i(x) < \deg g(x)$ . (Für ein lineares Polynom  $g(x) = x - a$  ist das gerade die TAYLOR-Entwicklung von  $\varphi$  im Punkt  $a$ .) Für  $\alpha_0(x)$  nehmen wir dazu den Divisionsrest von  $\varphi(x)$  durch  $g(x)$ ; ist  $q_0(x)$  der Quotient, gilt dann

$$g(x) = \alpha_0(x) + g(x)q_0(x).$$

Ist nun  $q_1$  der Quotient und  $\alpha_1$  der Rest bei der Division von  $q_0$  durch  $g$ , so erhalten wir die neue Darstellung

$$g(x) = \alpha_0(x) + g(x)(\alpha_1(x) + g(x)q_1(x)) = \alpha_0(x) + \alpha_1(x)g(x) + q_1(x)g(x)^2.$$

$\alpha_2(x)$  ist nun entsprechend der Divisionsrest von  $q_1(x)$  durch  $g(x)$  und so weiter, bis ein Quotient  $q_r(x)$  kleineren Grad als  $g(x)$  hat und damit der letzte Koeffizient  $\alpha_{r+1}(x)$  ist.

Ist  $g(x)$  ein lineares Polynom, haben alle  $\alpha_{r_i}(x)$  kleineren Grad als eins, sind also Konstanten. Da  $\deg \tilde{\varphi}(x) < \deg \ell_i^{e_i} = e_i$  ist, können wir also schreiben

$$\tilde{\varphi}(x) = \alpha_0 + \alpha_1 \ell_i(x) + \dots + \alpha_{e_i-1} \ell_i^{e_i-1}(x)$$

mit reellen Zahlen  $\alpha_{r_i}$ , und

$$\begin{aligned} \frac{\tilde{\varphi}(x)}{\ell_i(x)^{e_i}} &= \frac{\alpha_0 + \alpha_1 \ell_i(x) + \dots + \alpha_{e_i-1} \ell_i^{e_i-1}(x)}{\ell_i(x)^{e_i}} \\ &= \frac{\alpha_0}{\ell_i(x)^{e_i}} + \frac{\alpha_1}{\ell_i(x)^{e_i-1}} + \dots + \frac{\alpha_{e_i-1}}{\ell_i(x)}. \end{aligned}$$

Entsprechend erhalten wir die Darstellungen

$$\frac{\tilde{\psi}(x)}{q_k(x)^{f_i}} = \frac{\beta_0 x + \gamma_0}{q_i(x)^{f_i}} + \frac{\beta_1 x}{q_i(x)^{f_i-1}} + \dots + \frac{\beta_{f_i-1} x + \gamma_{f_i-1}}{q_i(x)}.$$

Faßt man alles zusammen und gibt den Zählern die richtigen Indizes, erhält man die oben angegebene Zerlegung, die sogenannte *Partialbruchzerlegung*.

Der gerade durchgeführte Beweis ist zwar (bei einer gegebenen Zerlegung des Nenners) konstruktiv, das verwendete Verfahren ist allerdings nicht besonders effizient. Nachdem wir wissen, daß eine Zerlegung wie oben existiert und auch alle Nenner kennen, ist es beispielsweise erheblich einfacher, von der obigen Zerlegung auszugehen (mit unbestimmten Zählern) und die Partialbrüche nach den üblichen Regeln der Bruchrechnung zu addieren. Der Nenner des entstehenden Bruchs ist  $g(x)$ , der Zähler  $f(x) - Q(x)g(x)$ , wobei  $Q(x)$  der Quotient bei der Division von  $f(x)$  durch  $g(x)$  ist. Der durch Addition der Partialbrüche berechnete Zähler ist eine lineare Funktion in den Koeffizienten  $\alpha, \beta, \gamma$ ; Koeffizientenvergleich mit  $f(x) - Q(x)g(x)$  liefert ein lineares Gleichungssystem für die Koeffizienten, das wir schnell und einfach nach GAUSS lösen können.

Für Linearfaktoren  $\ell_i(x)$ , die nur in der ersten Potenz vorkommen, gibt es sogar ein noch einfacheres Verfahren: Ist  $x_i$  die (einzige) Nullstelle von  $\ell_i(x)$ , so überlegt man sich leicht (durch Addition der Partialbrüche und des polynomialen Anteils), daß

$$\alpha_i = \lim_{x \rightarrow x_i} \frac{(x - x_i) f(x)}{g(x)}$$

ist. Da Zähler und Nenner für  $x = x_i$  verschwinden, muß dieser Grenzwert nach DE L'HOSPITAL berechnet werden; wir erhalten

$$\alpha_i = \lim_{x \rightarrow x_i} \frac{(x - x_i) f'(x) - f(x)}{g'(x)} = - \frac{f(x_i)}{g'(x_i)},$$

wobei hier der Nenner nicht verschwinden kann, da  $x_i$  nur eine einfache Nullstelle von  $g$  ist.

**3. Schritt:** Integration der Partialbrüche: Die Integrale der Funktionen  $\frac{\alpha_{i,k}}{\ell_i(x)^k}$  lassen sich über die Substitution  $u = \ell_i(x)$  sofort auf  $\int \frac{du}{u^k}$  zurückführen und sind somit problemlos. Weiter ist

$$\int \frac{\beta x + \gamma}{ax^2 + bx + c} dx = \frac{\beta}{2a} \int \frac{2ax + b}{ax^2 + bx + c} dx + \int \frac{\gamma - \frac{\beta b}{2a}}{ax^2 + bx + c} dx,$$

wobei das erste Integral nach der Regel über die logarithmische Ableitung aus Abschnitt a) berechnet werden kann, während wir das zweite in Abschnitt b) behandelt haben.

Bleiben schließlich noch die Integrale mit Potenzen eines quadratischen Polynoms im Nenner; hier wird  $k$  durch partielle Integration rekursiv erniedrigt.

Als ganz einfaches Beispiel, in dem alles explizit durchführbar ist, betrachten wir

$$\int \frac{dx}{x^2 - 5x + 6}.$$

Da hier sechs das Produkt und fünf die Summe der Nullstellen des Nenners ist, errät man leicht, daß  $x^2 - 5x + 6 = (x - 2)(x - 3)$  ist. Die Partialbruchzerlegung bestimmen wir nun durch probieren: Da

$$\frac{1}{x - 2} - \frac{1}{x - 3} = \frac{(x - 3) - (x - 2)}{(x - 2)(x - 3)} = \frac{-1}{(x - 2)(x - 3)}$$

ist, folgt

$$\int \frac{dx}{(x - 2)(x - 3)} = - \int \frac{dx}{x - 2} + \int \frac{dx}{x - 3} = \ln \left| \frac{x - 3}{x - 2} \right| + C.$$

## j) Symmetrie

Bei bestimmten Integralen kann man sich gelegentlich einen Großteil der Mühe des Ausrechnens ersparen, indem man Symmetrien berücksichtigt: Ist etwa die Funktion  $f$  spiegelsymmetrisch zur Achse  $x = c$ , d.h.  $f(c + x) = f(c - x)$  für alle  $x$ , für die beide Seiten definiert sind, so ist auch

$$\int_c^{c+a} f(x) dx = \int_{c-a}^c f(x) dx,$$

d.h.

$$\int_c^{c+a} f(x) dx = 2 \int_c^{c+a} f(x) dx.$$

Viel interessanter wird es, wenn  $f$  *punktsymmetrisch* zu einem Punkt  $(c, 0)$  auf der  $x$ -Achse ist, d.h.  $f(c + x) = -f(c - x)$ . Dann ist

$$\int_c^{c+a} f(x) dx = - \int_{c-a}^c f(x) dx \quad \text{und} \quad \int_{c-a}^{c+a} f(x) dx = 0,$$

ohne daß man irgendetwas über die Stammfunktion von  $f$  wissen müßte. So ist beispielsweise

$$\int_{-1}^1 \frac{\sin(x)}{|x|} dx = 0,$$

aber ein Leser, der mit den uns bislang bekannten Methoden nach einer Stammfunktion suchen möchte, dürfte eine sehr harte Zeit und wenig Erfolg haben.

## k) Einige nicht elementar integrierbare Funktionen

Tatsächlich ist die Stammfunktion von  $\frac{\sin x}{x}$  nicht durch elementare transzendenten Funktionen wie die Exponentialfunktion, die trigonometrischen Funktionen sowie deren Umkehrfunktionen und/oder Wurzeln und Grundrechenarten darstellbar, und sie teilt diese Eigenschaft mit einer ganzen Reihe weiterer Stammfunktionen.

Da einige dieser Stammfunktionen trotzdem für Anwendungen wichtig sind, behilft man sich damit, daß man einigen speziellen dieser Funktionen Namen gibt und versucht, alles darauf zurückzuführen. Die so definierten Funktionen können dann, genau wie der Sinus und die Exponentialfunktion, in Tabellenwerke und Unterprogrammbibliotheken aufgenommen werden, so daß man mit den dadurch abgedeckten Integralen genauso rechnen kann wie mit den elementaren transzendenten Funktionen.

Einige wichtige dieser neuen Funktionen sind



1) **Der Integralsinus:** Da  $\lim_{t \rightarrow 0} \frac{\sin t}{t} = 1$  ist, ist das bestimmte Integral

$$\text{Si}(x) = \int_0^x \frac{\sin t}{t} dt$$

für alle  $x \in \mathbb{R}$  wohldefiniert und eine stetige Funktion von  $x$ , der **Integralsinus**. Er ist beispielsweise wichtig für Tiefpaßfilterungen, wo Faltungsglieder mit  $\frac{\sin x}{x}$  berechnet werden können.

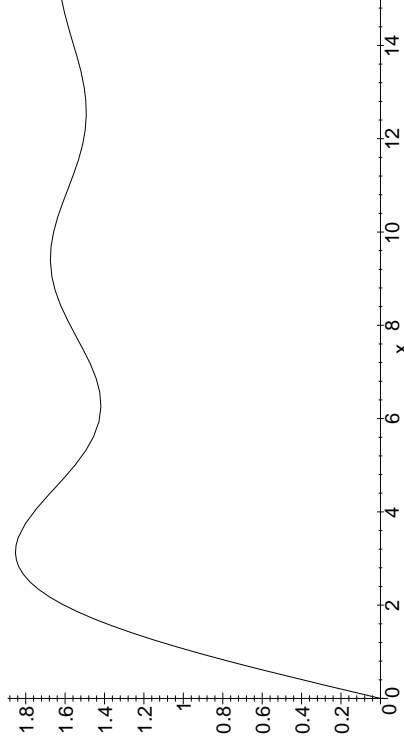


Abb. 64: Der Integralsinus

2) **Die Fehlerfunktion:** Auch die Funktion  $e^{-x^2}$  hat keine elementar angebbare Stammfunktion; hier definiert man als eine Stammfunktion die **Fehlerfunktion** oder **error function** durch

$$\text{Erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt.$$

(Der Vorfaktor sorgt dafür, daß  $\lim_{x \rightarrow \infty} \text{Erf}(x) = 1$  ist, was wir im Augenblick allerdings noch nicht beweisen können.)

Sie ist wichtig für die Statistik, denn für eine normalverteilte Zufallsvariable (dazu gehören die meisten Meßgrößen) mit Mittelwert  $\bar{x}$  und

Standardabweichung  $\sigma$  ist die Wahrscheinlichkeit dafür, daß ein Meßwert zwischen  $a$  und  $b$  liegt, gleich

$$\frac{1}{2\pi} \int_a^b e^{-\frac{1}{2} \left( \frac{x-\bar{x}}{\sigma} \right)^2} dx,$$

ein Integral, von dem man sich leicht überlegt, daß es mit Hilfe von Erf berechnet werden kann als

$$\frac{1}{2} \text{Erf} \left( \frac{\sqrt{2} \bar{x} - a}{\sigma} \right) - \frac{1}{2} \text{Erf} \left( \frac{\sqrt{2} \bar{x} - b}{\sigma} \right).$$

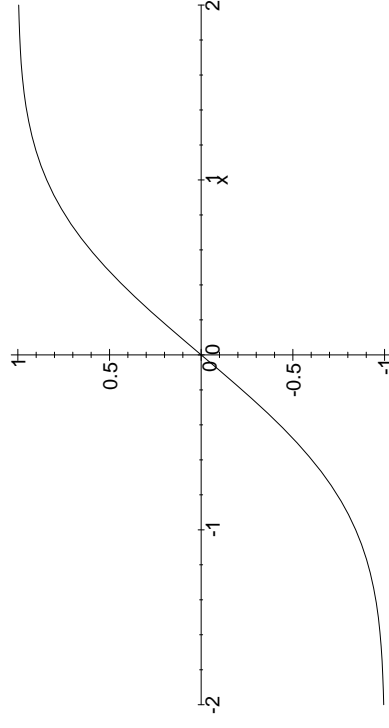


Abb. 65: Die Fehlerfunktion

3) **Elliptische Integrale:** Bei der Berechnung der Bogenlänge eines Ellipsensegments kommt man, je nach Art des Ansatzes und der Substitution, auf Integrale der Form

$$\int \frac{dx}{\sqrt{x^3 + ax^2 + bx + c}},$$

wobei das Polynom im Nenner keine mehrfachen Nullstellen hat, oder auf Integrale der Form

$$\int \frac{dx}{\sqrt{1 - k^2 \sin^2 x}} \quad \text{oder} \quad \int \sqrt{1 - k^2 \sin^2 x} dx$$

mit  $0 < k < 1$ . Alle diese Integrale heißen *elliptische Integrale* und sind nicht elementar ausdrückbar. Eine hauptsächlich auf ADRIEN MARRIE LEGENDRE (1752–1833) und KARL THEODOR WILHELM WEIERSTRASS (1815–1897) zurückgehende Theorie der elliptischen Funktionen stellt das Instrumentarium bereit, mit dem man diese vor allem in der Geodäsie, Kartographie und im Maschinenbau wichtigen Integrale berechnen kann.

**4) Algebraische Integrale:** Die erste Form der elliptischen Integrale ist ein Spezialfall eines sogenannten algebraischen Integrals; das sind Integrale, deren Integranden durch Wurzeln und Grundrechenarten (oder allgemeiner implizit durch Lösungen von Polynomgleichungen) gegeben sind. Die komplizierteren dieser Integrale sind fast alle nicht elementar ausdrückbar; es gibt inzwischen algorithmische Verfahren, die entscheiden, wann ein solches Integral elementar ausdrückbar ist, und die dann auch eine Stammfunktion finden können. Natürlich gibt es auch hier spezielle Funktionen, mit denen sich weitere dieser Integrale ausdrücken lassen.

### 1) Uneigentliche Integrale

Sei  $a > 0$  und  $b > a$ . Dann ist für eine reelle Zahl  $r \neq 1$

$$\int_a^b \frac{dx}{x^r} = \frac{-1}{(r-1)x^{r-1}} \Big|_a^b = \frac{1}{r-1} \left( \frac{1}{a^{r-1}} - \frac{1}{b^{r-1}} \right).$$

Falls  $r > 1$  ist, können wir hiervon den Grenzwert für  $b$  gegen unendlich betrachten und es liegt nahe, diesen als Wert des Integrals von  $a$  bis unendlich zu bezeichnen:

$$\int_a^\infty \frac{dx}{x^r} \stackrel{\text{def}}{=} \frac{1}{r-1} \frac{1}{a^{r-1}} \quad \text{für } r > 1.$$

Auch wenn wir nicht bis unendlich integrieren wollen, gibt es Beispiele von Integralen, denen wir via Grenzwertbetrachtung einen sinnvollen

Wert zuzuordnen können, ohne daß das Integral im Sinne unserer bisherigen Definitionen existieren würde: Beispielsweise ist für  $a \leq c < b$  und eine reelle Zahl  $0 < r < 1$

$$\int_a^c \frac{dx}{(b-x)^r} = \frac{(b-x)^{1-r}}{r-1} \Big|_a^c = \frac{(b-a)^{1-r}}{r-1} - \frac{(b-c)^{1-r}}{r-1},$$

und auch hier liegt es nahe, den Grenzwert für  $c \rightarrow b$  als Wert des Integrals von  $a$  bis  $b$  zu bezeichnen. Wir müssen hier allerdings vorsichtig sein mit dem Grenzübergang, denn die obige Formel gilt natürlich nur für  $c < b$ ; für  $c > b$  ist das Integral undefiniert. In einer solchen Situation können wir daher nicht den gewöhnlichen Limes betrachten, sondern brauchen einen eingeschränkten Grenzwertbegriff, der nur Folgen von einer der beiden Seiten berücksichtigt: Allgemein schreiben wir

$$y = \lim_{x \rightarrow b-} f(x),$$

wenn es für jedes  $\varepsilon > 0$  ein  $\delta > 0$  gibt, so daß  $|f(x) - y| < \varepsilon$  für alle  $x$  mit  $b - \delta < x < b$ . (Zur Erinnerung: Beim gewöhnlichen Grenzwert fordert man dies für alle  $x$  mit  $|x - b| < \delta$ , d.h.  $b - \delta < x < b + \delta$ .)

Völlig analog läßt sich natürlich auch ein rechtsseitiger Grenzwert

$$y = \lim_{x \rightarrow a+} f(x)$$

definieren durch die Bedingung, daß es für jedes  $\varepsilon > 0$  ein  $\delta > 0$  gibt, so daß  $|f(x) - y| < \varepsilon$  für alle  $x$  mit  $b < x < b + \delta$ .

Schließlich sollten wir, wenn wir schon beim Definieren sind, zur Bezeichnungskonsistenz vereinbaren, daß halboffene Intervalle wie

$$[a, b) = \{x \in \mathbb{R} \mid a \leq x < b\}$$

auch für unendliche  $b$  definiert sein sollen durch

$$[a, \infty) \stackrel{\text{def}}{=} \{x \in \mathbb{R} \mid a \leq x\};$$

entsprechend auch

$$(-\infty, b] \stackrel{\text{def}}{=} \{x \in \mathbb{R} \mid x \leq b\} \quad \text{und} \quad (-\infty, \infty) = \mathbb{R}.$$

Für unendliche Intervallgrenzen sind die Definitionen für links- und rechtsseitige Grenzwerte nicht sinnvoll anwendbar; wir vereinbaren daher, daß für  $b = \infty$  der Ausdruck  $\lim_{c \rightarrow b-} \int_a^c f(x) dx$  für den gewöhnlichen Grenzwert  $\lim_{c \rightarrow \infty} \int_a^c f(x) dx$  steht; entsprechend bei  $\lim_{x \rightarrow a+} f(x)$  für  $a = -\infty$ .

Mit all diesen Definitionen können wir dann eine Funktion  $f$  betrachten, die in einem halboffenen Intervall  $[a, b)$  mit  $b \in \mathbb{R} \cup \{\infty\}$  definiert und stückweise stetig ist; für diese definieren wir das rechtsseitig uneigentliche Integral

$$\int_a^b f(x) dx = \lim_{c \rightarrow b-} \int_a^c f(x) dx,$$

falls dieser Grenzwert existiert; andernfalls sagen wir, das Integral sei *divergent*.

Völlig analog definieren wir linksseitige uneigentliche Integrale:  $f$  sei stückweise stetig auf dem halboffenen Intervall  $(a, b]$  mit  $a \in \mathbb{R} \cup \{-\infty\}$ ; dann ist

$$\int_a^b f(x) dx = \lim_{c \rightarrow a+} \int_c^b f(x) dx,$$

falls dieser Grenzwert existiert; andernfalls sagen wir, das Integral sei *divergent*.

Diese Definitionen sind immer noch nicht allgemein genug: Eine Funktion könnte auch an *beiden* Enden eines Intervalls  $(a, b)$  undefiniert sein, wobei wir auch die Sonderfälle  $a = -\infty$  und/oder  $b = \infty$  zulassen wollen, und zusätzlich könnte sie auch noch Undefiniertheitsstellen  $c_1 < \dots < c_r$  im Intervallinnern haben.

In diesem Fall läßt sich das Intervall so in Teilintervalle zerlegen, daß  $f$  in jedem Teilintervall höchstens an *einem* der beiden Intervallenden uneigentlich ist: Falls es keine Undefiniertheitsstellen im Intervallinnern gibt, wählen wir willkürlich ein  $c_0$  zwischen  $a$  und  $b$  und betrachten die beiden Intervalle  $(a, c_0]$  und  $[c_0, b)$ . Im anderen Fall können zwei Zusatzpunkte notwendig sein: ein Punkt  $c_0$  zwischen  $a$  und  $c_1$  sowie ein Punkt  $c_{r+1}$  zwischen  $c_r$  und  $b$ .

Wir sagen dann, das uneigentliche Integral  $\int_a^b f(x) dx$  konvergiere, wenn *jedes* der Integrale

$$\int_a^{c_0} f(x) dx, \int_{c_0}^{c_1} f(x) dx, \dots, \int_{c_{r+1}}^b f(x) dx$$

konvergiert; die Summe ihrer Werte bezeichnen wir als den Wert des Integrals von  $a$  bis  $b$ .

Als Beispiel betrachten wir das an beiden Grenzen uneigentliche Integral

$$\int_{-\infty}^{\infty} \frac{dx}{1+x^2}.$$

Da der Integrand eine gerade Funktion ist, empfiehlt es sich, das Integrationsintervall, das hier aus ganz  $\mathbb{R}$  besteht, bei Null zu unterteilen, und wir erhalten

$$\begin{aligned} \int_{-\infty}^{\infty} \frac{dx}{1+x^2} &= \int_{-\infty}^0 \frac{dx}{1+x^2} + \int_0^{\infty} \frac{dx}{1+x^2} = 2 \int_0^{\infty} \frac{dx}{1+x^2} \\ &= 2 \lim_{c \rightarrow \infty} \int_0^c \frac{dx}{1+x^2} = 2(\lim_{c \rightarrow \infty} \arctan c - \arctan 0) = \pi. \end{aligned}$$

Obige Forderung, daß *jedes* der Teilintegrale einzeln konvergieren soll, ist gelegentlich etwas sehr restriktiv. Für das bei Null uneigentliche Integral

$$\int_{-2}^4 \frac{dx}{x^3}$$

könnte man etwa argumentieren, daß der Integrand eine ungerade Funktion ist, so daß aus Symmetriegründen das Integral von  $-2$  bis  $2$  verschwinden sollte und

$$\int_{-2}^4 \frac{dx}{x^3} = \int_{-2}^2 \frac{dx}{x^3} + \int_2^4 \frac{dx}{x^3} = 0 + \frac{3}{32} = \frac{3}{32}$$

sein sollte. Diese Art der Argumentation ist durch das, was wir bislang gelernt haben, nicht gedeckt, und es gibt auch gute Gründe, sie zu verbieten: Schließlich geht die Stammfunktion  $-\frac{1}{2}x^{-2}$  für  $x \rightarrow 0$  gegen minus unendlich, und wenn eine Größe mit Relevanz in der realen Welt unendlich groß wird, hat dies im Allgemeinen zu gravierende Konsequenzen, als daß man einfach durch diesen Punkt hindurch weitergehen könnte.

Andererseits sind aber viele der mathematischen Formeln, die in den Naturwissenschaften und der Technik angewandt werden, nur näherungsweise gültig: Mathematische Modelle sind praktisch immer *vereinfachte* Modelle der Wirklichkeit, beispielsweise gilt das OHMSche Gesetz sicher nicht mehr, wenn man einen  $5\ \Omega$ -Widerstand aus einer auf  $5\ \text{V}$  Spannung ausgelegten Schaltung im Hochspannungslabor mit  $100\ \text{kV}$  belastet, und es gilt auch nicht mehr ohne Korrekturterme, wenn man einen Wechselspannung mit  $500\ \text{MHz}$  anlegt.

Entsprechend gibt es durchaus Situationen, in denen das mathematische Modell einen unendlich großen Wert vorhersagt, wohingegen in der Realität limitierende Faktoren, die für Werte im „üblichen“ Größenbereich noch keine nennenswerte Rolle spielen, für eine Begrenzung sorgen. Falls man in einer solchen Situation sicher sein kann, daß auch in der realen Situation noch die Symmetrie zum Nullpunkt erhalten bleibt, kann man so wie oben argumentieren; falls allerdings die Symmetrie *nicht* erhalten bleibt, können durch die Begrenzung der Funktion beliebig große Abweichungen erzeugt werden, über die man mit dem vereinfachten mathematischen Modell nichts aussagen kann.

Da somit alles von der Anwendung abhängt, kann die Mathematik hier nicht mehr bieten als eine *Definition*: Falls für die Funktion  $f$ , die auf  $[a, c) \cup (c, b]$  definiert ist, der Grenzwert

$$\lim_{h \rightarrow 0} \int_a^{c-h} f(x) dx + \int_{c+h}^b f(x) dx$$

existiert, bezeichnen wir ihn als CAUCHYSchen Hauptwert von  $\int_a^b f(x) dx$  nach dem aus der Analysis I bekannten französischen Mathematiker

Baron AUGUSTIN LOUIS CAUCHY (1789–1857). Entsprechend reden wir auch in komplizierteren Situationen mit mehreren Unstetigkeitsstellen vom CAUCHYSchen Hauptwert, falls sich eine Aufteilung in Teilintervalle finden läßt, so daß für jedes Teilintervall der CAUCHYSche Hauptwert existiert. Im obigen Beispiel wäre also  $\frac{3}{2}$  der CAUCHYSche Hauptwert des Integrals, wohingegen das Integral selbst undefiniert ist.

Die Frage, wann der CAUCHYSche Hauptwert für ein eigentlich divergentes Integral verwendet werden sollte, ist keine mathematische Frage: Unrein mathematischen Gesichtspunkten gibt es **nie** eine Rechtfertigung für die Verwendung des CAUCHYSchen Hauptwerts. Der CAUCHYSche Hauptwert ist nur dann sinnvoll anwendbar, wenn man davon ausgeht, daß ein mathematisches Modell eine Situation nur für nicht zu große Funktionswerte (ungefähr) korrekt beschreibt, und wenn man gleichzeitig sicher ist, daß die Unendlichkeitsstelle des mathematischen Modells für die Anwendung unproblematisch ist und gleichzeitig die Symmetrie, die der Berechnung des CAUCHYSchen Hauptwerts zugrundeliegt, auch in der realen Anwendung gilt.

Der CAUCHYSche Hauptwert darf auch *nie* als eine Rechtfertigung dafür verstanden werden, daß man unbesonnen

$$\int_a^b f(x) dx = F(b) - F(a)$$

setzt, wobei  $F$  eine zwar in den Punkten  $a$  und  $b$ , nicht aber auch für jeden Zwischenwert  $a \leq x \leq b$  definierte Stammfunktion von  $f$  ist: So etwas kann zu Ergebnissen wie

$$\int_{-2}^2 \frac{dx}{x^2} = -\frac{1}{x} \Big|_{-2}^2 = \frac{-1}{2} - \frac{-1}{-2} = -1$$

führen, und natürlich ist keine Anwendung denkbar, in der eine negative Zahl in sinnvoller Weise als Integral über eine überall positive Funktion angesehen werden kann. In der Tat existiert im obigen Beispiel weder das Integral noch dessen CAUCHYScher Hauptwert, da sich die Unendlichkeiten links und rechts der Null hier nicht wegheben, sondern verstärken.

Zu einer etwas systematischeren Untersuchung uneigentlicher Integrale empfiehlt es sich, zunächst die Potenzen zu betrachten. Für positive Exponenten  $r$  ist  $x^r$  überall definiert, so daß Integrale über einen *endlichen* Bereich unproblematisch sind; für Integrale über einen *unendlichen* Bereich rechnet man leicht nach, daß sie immer divergieren. Für  $r = 0$  ändert sich nichts an dieser Situation; interessant ist also nur der Fall  $r < 0$ , wo sowohl der Wert  $x = 0$  als auch unendliche obere und/oder untere Grenzen zu Problemen führen können. Für negatives  $r$  ist  $x^r = 1/x^{-r}$ , wir interessieren uns also für

$$\int_a^b \frac{dx}{x^r} = \frac{1}{(1-r)x^{r-1}} \Big|_a^b \quad \text{für } r > 0, \quad r \neq 1.$$

Für  $a > 0$  und  $b \rightarrow \infty$  existiert ein Grenzwert genau dann, wenn  $r - 1 > 0$ , also  $r > 1$  ist; alsdann ist

$$\int_a^\infty \frac{dx}{x^r} = \lim_{b \rightarrow \infty} \int_a^b \frac{dx}{x^r} = \frac{1}{(r-1)a^{r-1}}.$$

Für  $b > 0$  und  $a \rightarrow \infty$  existiert ein Grenzwert genau dann, wenn  $r - 1 < 0$ , also  $r < 1$  ist; alsdann ist

$$\int_0^b \frac{dx}{x^r} = \lim_{\alpha \rightarrow 0} \int_\alpha^b \frac{dx}{x^r} = \frac{1}{1-r} b^{1-r}.$$

Zusammen mit der Monotonieeigenschaft des RIEMANN-Integrals aus §3a) ergeben sich hieraus zwei allgemeine Kriterien für die Konvergenz uneigentlicher Integrale:

**Satz:** Die Funktion  $f$  sei stetig für  $x \geq a$  und  $g$  sei stetig für  $0 < x \leq b$ . Dann gilt:

- 1.) Falls es eine reelle Zahl  $K$  und eine reelle Zahl  $r > 1$  gibt, so daß  $|f(x)| \leq \frac{K}{x^r}$  ist, konvergiert  $\int_a^\infty f(x) dx$ .
- 2.) Falls es eine reelle Zahl  $K$  und eine reelle Zahl  $0 < r < 1$  gibt, so daß  $|g(x)| \leq \frac{K}{x^r}$  ist, konvergiert  $\int_0^b g(x) dx$ .

*Beweis:* 1.) Nach der Monotonieregel ist für jedes  $b \geq a$

$$\int_a^b |f(x)| dx \leq K \int_a^b \frac{dx}{x^r},$$

und letzteres Integral konvergiert, wie wir gerade nachgerechnet haben, unter den angenommenen Voraussetzungen. Das linksstehende Integral ist somit beschränkt durch eine von  $b$  unabhängige Konstante; da es wegen der Nichtnegativität des Betrags zusätzlich eine monoton wachsende Funktion von  $b$  ist, existiert daher der Grenzwert nach dem bekannten, schon für die Existenz des RIEMANN-Integrals verwendeten Satz aus der Analysis I, wonach jede monotone und beschränkte Folge reeller Zahlen konvergent ist.

Genau dasselbe gilt für die ebenfalls nichtnegative Funktion  $f(x) + |f(x)|$ , die durch  $\frac{2K}{x^r}$  beschränkt ist; also existieren die uneigentlichen Integrale

$$\int_a^\infty (f(x) + |f(x)|) dx \quad \text{und} \quad \int_a^\infty |f(x)| dx,$$

und damit auch das gesuchte Integral als ihre Differenz.

2.) geht völlig analog. ■

Als Beispiel hierfür betrachten wir die Strahlung eines schwarzen Körpers: Nach dem RAYLEIGH'schen Strahlungsgesetz, wonach alle mit der Geometrie des Körpers verträglichen Wellenzahlvektoren gleichwahrscheinlich sind, wäre die Energiedichte proportional zum Quadrat der Frequenz, die Gesamtenergie also proportional zu

$$\int_0^\infty \nu^2 d\nu,$$

einem offensichtlich divergenten Integral: Das ist die sogenannte „UV-Katastrophe“, die dieses Modell zum Widerspruch führt.

Die Abhilfe besteht bekanntlich darin, daß man die Gleichverteilung der Frequenzen durch eine BOSE-EINSTEIN-Statistik ersetzt und als neue

Energiedichte nun nach dem PLANCKSchen Strahlungsgesetz

$$\frac{8\pi h}{c^2} \frac{\nu^3}{e^{\frac{h\nu}{kT}} - 1}$$

erhält, wobei  $h$  das PLANCKSche Wirkungsquantum,  $k$  die BOLTZMANN-Konstante,  $c$  die Lichtgeschwindigkeit und  $T$  die absolute Temperatur bezeichnet. Hier ist die Gesamtenergie

$$\frac{8\pi h}{c^2} \int_0^\infty \frac{\nu^3}{e^{\frac{h\nu}{kT}} - 1} d\nu,$$

und die UV-Katastrophe wird genau dann vermieden, wenn dieses Integral konvergiert.

Mit der Substitution  $x = \frac{h\nu}{kT}$  erhalten wir

$$\frac{8\pi h}{c^2} \int_0^\infty \frac{\nu^3}{e^{\frac{h\nu}{kT}} - 1} d\nu = \frac{8\pi h}{c^2} \left(\frac{kT}{h}\right)^4 \int_0^\infty \frac{x^3}{e^x - 1} dx.$$

Dieses Integral ist an beiden Grenzen uneigentlich; betrachten wir also eine positive Konstante  $a$  und getrennt die Integrale von 0 bis  $a$  und von  $a$  bis  $\infty$ .

Letzteres Integral existiert nach dem gerade bewiesenen Satz, wenn wir  $K, r$  finden können, so daß

$$\frac{x^3}{e^x - 1} \leq \frac{K}{x^r} \quad \text{für alle } x \geq a.$$

Diese Ungleichung ist äquivalent zu

$$\frac{1}{e^x - 1} \leq \frac{K}{x^{r+3}} \quad \text{oder} \quad e^x \geq 1 + \frac{x^{r+3}}{K},$$

und dies gilt mit geeignetem  $a, K$  und (beispielsweise)  $r = 2$ , da die Exponentialfunktion schneller wächst als jedes Polynom.

Entsprechend gilt für  $0 \leq x \leq 1$  die Ungleichung etwa mit  $r = \frac{1}{2}$ , da dort  $e^x - 1$  stärker wächst als jede  $x$ -Potenz. Also konvergiert das Integral sowohl an seiner unteren als auch seiner oberen Grenze, und

dazwischen ist ohnehin alles unproblematisch, da der Integrand stetig ist. ■

Als (vorerst) letztes Beispiel eines uneigentlichen Integrals sei noch

$$\Gamma(x) \stackrel{\text{def}}{=} \int_0^\infty e^{-t} t^{x-1} dt \quad \text{für } x > 0$$

betrachtet. Dieses Integral ist natürlich immer uneigentlich an der oberen Grenze; für  $x < 1$  zusätzlich auch noch an der unteren.

Diese untere Grenze ist völlig harmlos, denn für  $0 < x < 1$  ist

$$e^{-t} t^{x-1} = \frac{e^{-t}}{t^{1-x}} \leq \frac{1}{t^{1-x}} \quad \text{mit} \quad 0 < 1 - x < 1,$$

so daß obiger Satz sofort die Konvergenz zeigt.

Auch die obere Grenze ist unproblematisch, denn die dazu notwendige Abschätzung

$$e^{-t} t^{x-1} \leq \frac{K}{t^r} \iff e^t \geq t^{r+x-1}$$

folgt wieder, da die Exponentialfunktion stärker wächst als jede Potenz.

Die somit für alle  $x > 0$  definierte Funktion  $x \mapsto \Gamma(x)$  heißt EULERSche *Gamma-Funktion*.



LEONHARD EULER (1707–1783) wurde in Basel geboren und ging auch dort zur Schule und, im Alter von 14 Jahren, zur Universität. Dort legte er zwei Jahre später die Magisterprüfung in Philosophie ab und begann mit dem Studium der Theologie; daneben hatte er sich seit Beginn seines Studiums unter Anleitung von JOHANN BERNOULLI mit Mathematik beschäftigt. 1726 beendete er sein Studium in Basel und bekam eine Stelle an der Petersburger Akademie der Wissenschaften, die er 1727 antrat. Auf Einladung FRIEDRICHS DES GROSSEN wechselte er 1741 an die preußische Akademie der Wissenschaften; nachdem sich das Verhältnis zwischen den beiden dramatisch verschlechtert hatte, kehrte er 1766 nach St. Petersburg zurück. Im gleichen Jahr erblindete er vollständig; trotzdem schrieb er rund die Hälfte seiner zahlreichen Arbeiten (73 Bände) danach. Sie enthalten bedeutende Beiträge zu vielen Gebieten der Mathematik, Physik, Astronomie und Kartographie.