

Kapitel 4

Integralrechnung

Die Integralrechnung ist neben der Differentialrechnung die zweite wichtige Säule der Analysis. Wir betrachten auch sie in diesem Semester nur im eindimensionalen Fall.

§ 1: Das Riemann-Integral

a) Heuristische Vorüberlegungen

Seiner großen Bedeutung entsprechend, wollen wir uns dem Integralbegriff zunächst heuristisch nähern, bevor wir ihn – mit beträchtlichem technischem Aufwand – in Abschnitt c) 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 sec kennen, können wir $s(5 \text{ sec})$ rekursiv berechnen als

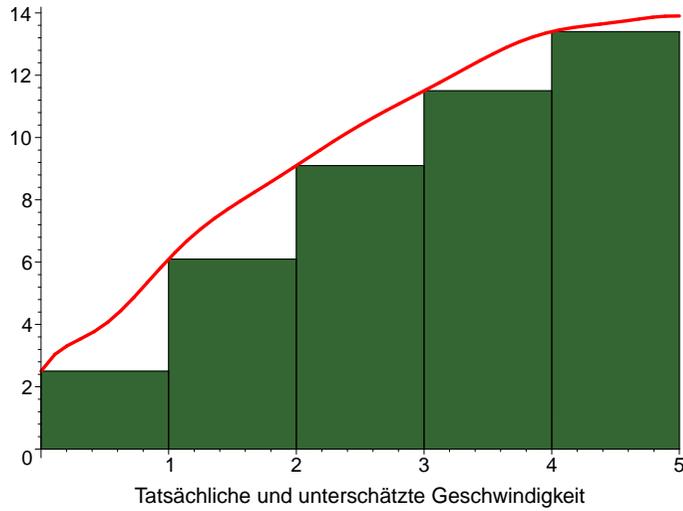
$$\begin{aligned} s(5 \text{ sec}) &\approx v(0 \text{ sec}) \cdot 1 \text{ sec} + v(1 \text{ sec}) \cdot 1 \text{ sec} + v(2 \text{ sec}) \cdot 1 \text{ sec} \\ &\quad + v(3 \text{ sec}) \cdot 1 \text{ sec} + v(4 \text{ sec}) \cdot 1 \text{ sec} \\ &= 2,5 \text{ m} + 6,1 \text{ m} + 9,1 \text{ m} +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

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

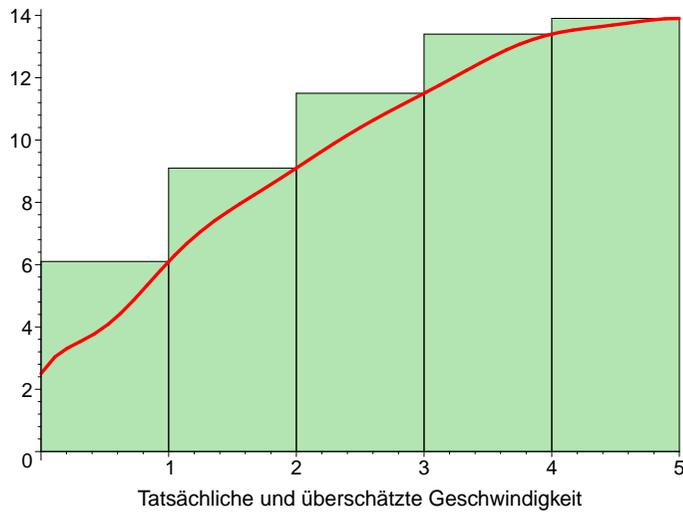
Zumindest mit dieser Genauigkeit ist das sicherlich falsch, denn bei der Berechnung des Wegs sind wir davon ausgegangen, daß das Fahrzeug während der ersten Sekunde konstant mit einer Geschwindigkeit von 2,5 m/sec gefahren ist, exakt ab deren Ende dann aber plötzlich eine Sekunde lang mit 6,1 m/sec, usw. Das ist natürlich unrealistisch; tatsächlich dürfte die Geschwindigkeit etwa so verlaufen sein, wie es die Kurve in den Abbildungen 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.

Alternativ hätten wir auch einen *zu großen* Weg schätzen können, wenn wir die Geschwindigkeit während eines Sekundenintervalls jeweils auf den bekannten 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 42,6m und 54 m liegt.



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

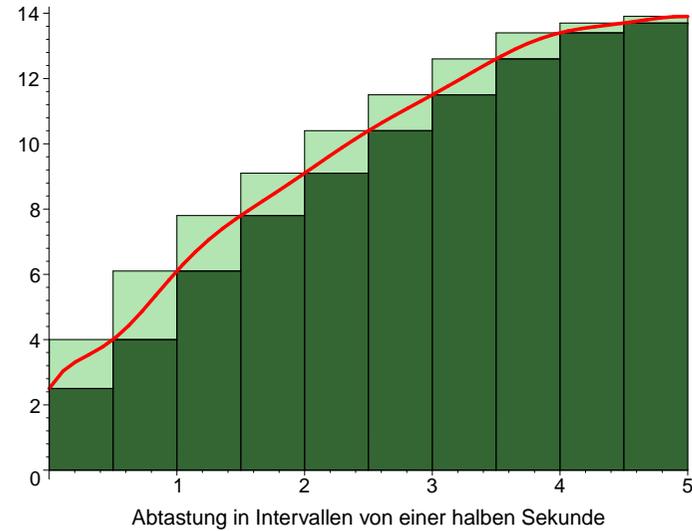
Jetzt ist die „unterschätzte“ Wegstrecke

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

und die „überschätzte“ ist

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

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



Beide Werte wie auch ihre Differenz lassen sich leicht geometrisch veranschaulichen: In der nächsten Abbildung ist die unterschätzte Geschwindigkeit 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

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

und als obere Schranke

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

Mit zehn Geschwindigkeitswerten pro Sekunde verbessert sich die untere Schranke auf

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

und die obere auf

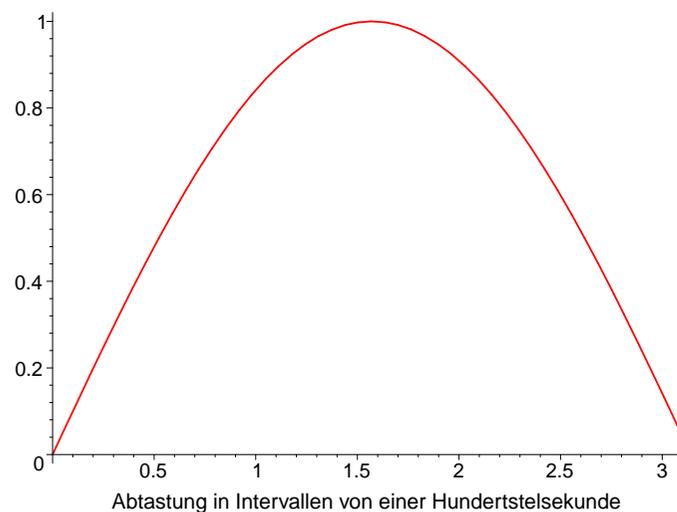
$$\sum_{k=1}^{50} v\left(\frac{k}{10} \text{ sec}\right) \cdot \frac{1}{10} \text{ sec} \approx 49,16 \text{ m}.$$

Wie die Abbildung 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.

Gehen wir von den Zehntelsekunden zu den Hundertsteln, sind die Rechtecke zumindest visuell nicht mehr voneinander und von der Fläche



unter der Kurve zu unterscheiden; zu sehen ist nur noch eine durchweg dunkelgrüne Fläche.



Rechnerisch gibt es noch einen kleinen Unterschied im Dezimeterbereich:

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

und

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

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

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

für die untere Schranke und

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

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 10000 Mal pro Sekunde abgetastet werden, und wir erhalten die Schranken

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

und

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

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

2) Integration als Flächenbestimmung: Die beiden letzten Abbildungen 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

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

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 π , 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 Durchschnittsbestimmung: 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

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

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

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

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

Wir *definieren* daher

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

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

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

Wie wir aus den Übungen wissen (oder noch einmal durch vollständige Induktion beweisen können) ist

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

also können wir dies vereinfachen zu

$$F(x) = x^3 \lim_{n \rightarrow \infty} \frac{1}{n^3} \cdot \frac{n(2n-1)(n-1)}{6} = x^3 \lim_{n \rightarrow \infty} \frac{n(2n-1)(n-1)}{6n^3}$$

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

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

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

$$= x \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=0}^{n-1} \left(e^{\frac{x}{n}}\right)^k.$$

Die Summe ganz rechts ist eine geometrische Reihe mit $q = e^{x/n}$; da wir uns nur für Werte $x \neq 0$ interessieren, können wir die Summenformel

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

anwenden und erhalten

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

$$= x(1 - e^x) \lim_{n \rightarrow \infty} \frac{\frac{1}{n}}{1 - e^{\frac{x}{n}}}.$$

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'HÔPITAL 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

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

gegen eben diesen Grenzwert.

Nach DE L'HÔPITAL ist

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

$$= \lim_{u \rightarrow \infty} \frac{-1}{xe^{\frac{x}{u}}} = \frac{-1}{x}.$$

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

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

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

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

für rationales x . Für irrationales x dagegen sind alle kx/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: Plötzliche Sprünge können bei den wenigsten Anwendungen ausgeschlossen werden; im Gegensatz zu DIRICHLETS Beispiel ist dort allerdings die Menge der Sprungstellen diskret.

Wirklich kompliziert wird die Situation beispielsweise bei der mathematischen Modellierung von Börsenkurses oder auch des Rauschens in einer elektronischen Schaltung, wo Sprünge mit Hilfe von Zufallsprozessen beschrieben werden. Für solche Anwendungen wird der in diesem Paragraphen definierte Integralbegriff nicht ausreichen; er ist aber eine unabdingbare Voraussetzung, um die dort benötigten Techniken zu verstehen.

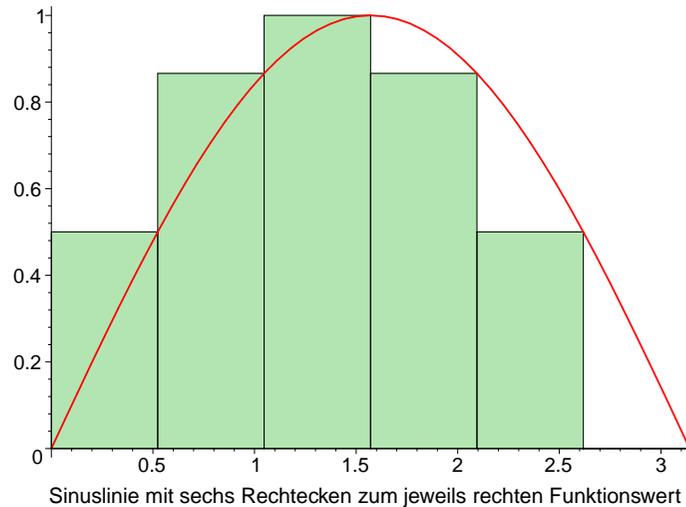
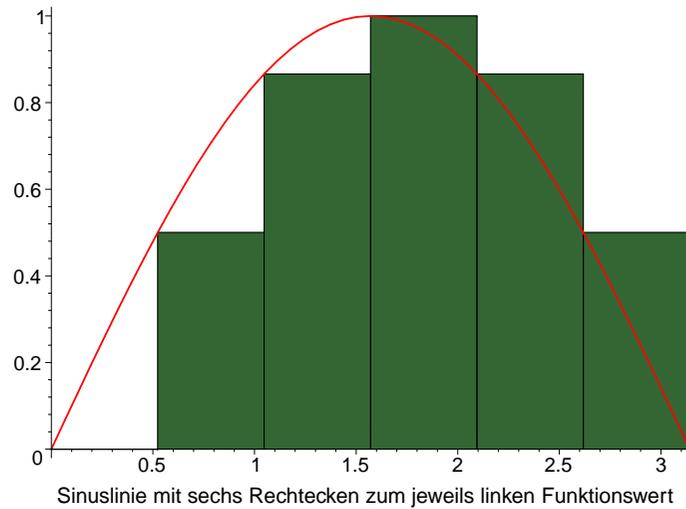
Obwohl wir im folgenden fast jeden der wichtigeren Sätze aus der bisherigen Vorlesung noch einmal in die Erinnerung zurückrufen und anwenden müssen, ist die nun folgende Konstruktion also kein Luxus, sondern zumindest langfristig notwendig auch für praktische Anwendungen.

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 π mit einer Annäherung der Fläche durch sechs Rechtecke.

In beiden Abbildungen 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 π 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 versucht sein, den Wert 1,954 als eine „Taschenrechnerzwei“ zu interpretieren; da aber die Sinuswerte an den Stellen $0, \frac{\pi}{6}, \frac{\pi}{3}, \frac{\pi}{2}, \frac{4\pi}{3}, \frac{5\pi}{6}$ und 1 wohlbekannt sind (im Winkelmaß ausgedrückt sind diese Stellen gerade die Vielfachen von 30°), kann man die Summe der Rechteckflächen exakt berechnen mit dem Ergebnis

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

das schon wegen der Transzendenz von π 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 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 Integralbegriff 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 uns ein alter Bekannter aus Kapitel 3, der *Mittelwertsatz der Differen-*

tialrechnung: Ist F stetig in einem Intervall $[u, v]$ und differenzierbar in (u, v) , so gibt es einen Punkt $\xi \in (u, v)$, so daß

$$F'(ξ) = \frac{f(v) - f(u)}{v - u}$$

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 ξ_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.

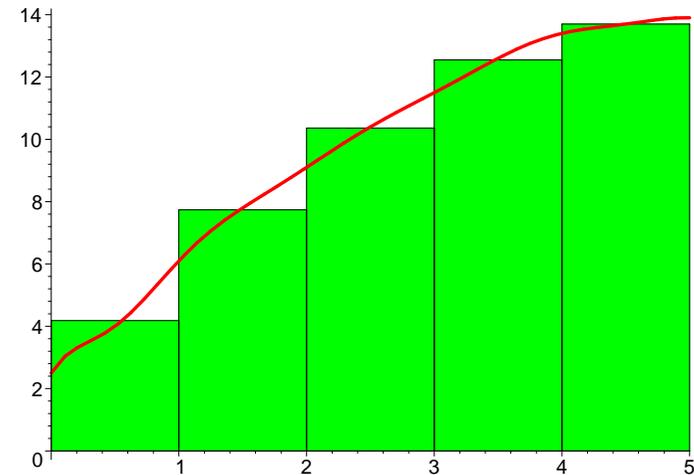
$$F(x_{i+1}) = F(x_i) + f(\xi_i)(x_{i+1} - x_i).$$

Rekursiv folgt, daß

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

Damit ist also die Fläche unter der Kurve $y = f(x)$ *exakt* gleich der Gesamtfläche endlich vieler geeignet gewählter Rechtecke; die Abbildung unten 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.

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



Die Fläche unter der Kurve ist gleich der Fläche der fünf Rechtecke

Die Idee zur Definition des RIEMANN-Integrals ist nun, daß man einfach *beliebige* ξ_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 ξ_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 ξ_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 δ 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

$$\left| \frac{1}{x \pm \delta} - \frac{1}{x} \right| = \left| \frac{\mp \delta}{x(x \pm \delta)} \right| = \frac{\delta}{x(x \pm \delta)} < \varepsilon$$

$$\Leftrightarrow \delta < (x^2 \pm x\delta)\varepsilon \Leftrightarrow \delta \mp x\delta\varepsilon < x^2\varepsilon \Leftrightarrow \delta < \frac{x^2\varepsilon}{1 \mp x\varepsilon}.$$

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

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

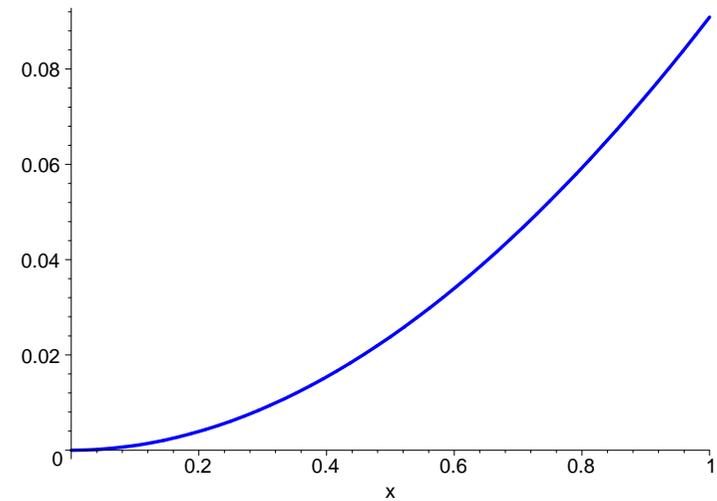
Damit ist die Stetigkeit der Funktion bewiesen. Sie ist aber nicht gleichmäßig stetig, denn es gibt aber kein von x unabhängiges $\delta > 0$ mit dieser Eigenschaft: Der angegebene Ausdruck für das größtmögliche δ geht wegen des Faktors x^2 im Zähler für $x \rightarrow 0$ selbst gegen Null, fällt also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

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.



δ in Abhängigkeit von x für $f(x) = 1/x$ und $\varepsilon = 0,1$

Aufgrund der Annahme, der Satz sei falsch, können wir ein solches ε 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 Kapitel 2 den

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

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_ν} 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 ν 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 ε , 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 uns bereits aus Kapitel 2, §6 bekannten deutschen Mathematikers BERNHARD RIEMANN, das inzwischen nach ihm benannte RIEMANN-Integral.

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

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

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

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

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

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

folgt

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

Natürlich kennen wir die Zahlen ξ_i nicht – selbst wenn wir wissen, daß F und damit die ξ_i überhaupt existieren. Deshalb betrachten wir

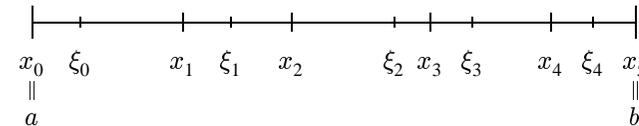
einfach *beliebige* Zahlen ξ_i mit

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

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 \underline{x} auftritt, diese Ungleichungen erfüllt sein sollen.



Definition: Die RIEMANNsche Summe zu einem Paar $(\underline{x}, \underline{\xi})$ ist

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

Das RIEMANN-Integral soll der Grenzwert einer Folge RIEMANNscher Summen sein, wobei die Unterteilungen \underline{x} immer feiner werden. Dieses „Feinerwerden“ müssen wir natürlich auch noch erklären: Seien \underline{x} und \underline{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 \underline{y} eine *Verfeinerung* von \underline{x} , wenn jedes Teilintervall (y_i, y_{i+1}) von \underline{y} ganz in einem der Teilintervalle (x_j, x_{j+1}) von \underline{x} liegt; \underline{y} entsteht also aus \underline{x} , indem einige der Intervalle von \underline{x} noch weiter unterteilt werden. Insbesondere muß dann $m \geq n$ sein.

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

- $\underline{x}^{(\nu+1)}$ stets eine Verfeinerung von $\underline{x}^{(\nu)}$ ist und
- die maximale Länge der Teilintervalle von $\underline{x}^{(\nu)}$ mit wachsendem ν 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 $\underline{x}^{(\nu)}$ 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

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

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 $(\underline{x}^{(\nu)})$ 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

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

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

ξ 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 ξ, 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 $(\underline{x}^{(\nu)})$ von immer feiner werdenden Unterteilungen des Intervalls $[a, x]$; dazu entsprechende Folgen $(\underline{\xi}^{(\nu)})$ von Zwischenwerten und die Grenzwerte der RIEMANN-Summen $I(\underline{x}^{(\nu)}, \underline{\xi}^{(\nu)})$.

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

Um das einzusehen, betrachten wir für jede Unterteilung $\underline{x}^{(\nu)}$ 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

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

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

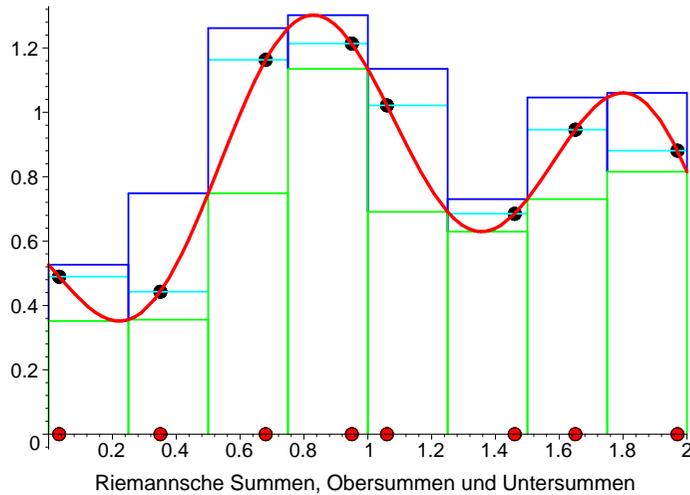
$$S^{(\nu)} \stackrel{\text{def}}{=} \sum_{i=0}^{n_\nu-1} M_i^{(\nu)}(x_{i+1}^{(\nu)} - x_i^{(\nu)})$$

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

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

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

ein Zusammenhang, der unten 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 $I(\underline{x}^{(\nu)}, \underline{\xi}^{(\nu)})$ 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

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

auch $I(\underline{x}^{(\nu)}, \underline{\xi}^{(\nu)})$ 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 $\underline{x}^{(\nu+1)}$ entsteht aus $\underline{x}^{(\nu)}$ dadurch, daß einige der Intervalle weiter unterteilt werden. Ist aber $(x_i^{(\nu+1)}, x_{i+1}^{(\nu+1)})$ Teilintervall von $(x_j^{(\nu)}, x_{j+1}^{(\nu)})$, so kann das Minimum im Teilintervall natürlich nicht kleiner sein als 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. Da wir aus Kapitel 2 wissen, daß jede monotone und beschränkte Folge reeller Zahlen konvergent ist, 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

$$\Delta^{(\nu)} = \max_{\text{def } i} (M_i^{(\nu)} - m_i^{(\nu)})$$

ist also

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

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 $\underline{x}^{(\nu)}$ jede der Differenzen $x_{i+1}^{(\nu)} - x_i^{(\nu)}$ kleiner als δ , 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 $\underline{x}^{(\nu)}$ immer feiner werden, d.h. zu jedem $\delta > 0$ gibt es in der Tat ein ν_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 ε , 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

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

auch die Gleichung

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

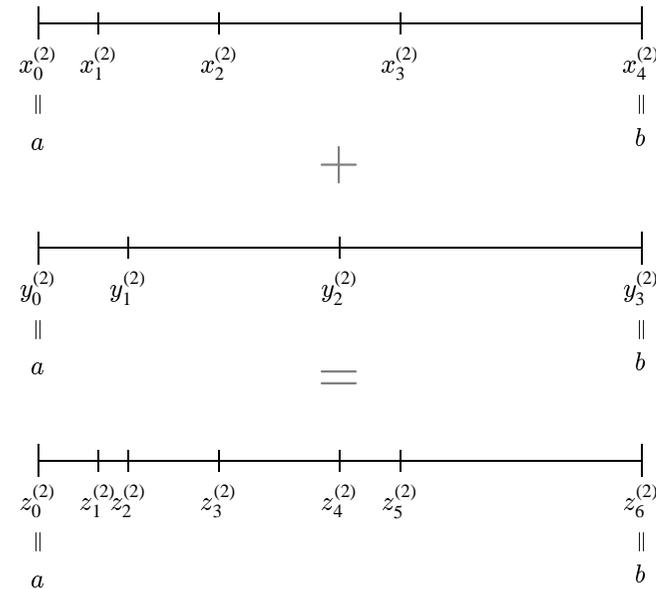
Da die linke und die rechte Seite obiger Ungleichung nicht von den $\xi_i^{(\nu)}$ abhängen, ist auch $\lim I(\underline{x}^{(\nu)}, \underline{\xi}^{(\nu)})$ davon unabhängig, und damit ist der erste Schritt des Beweises beendet.

Der zweite ist zum Glück erheblich einfacher:

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

Betrachten wir zwei Folgen $(\underline{x}^{(\nu)})$ und $(\underline{y}^{(\nu)})$ von Unterteilungen. Für jeden Index ν können wir zu den Unterteilungen $\underline{x}^{(\nu)}$ und $\underline{y}^{(\nu)}$ eine gemeinsame Verfeinerung $\underline{z}^{(\nu)}$ 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 $\underline{x}^{(\nu)}$; entsprechend seien $\tilde{s}^{(\nu)}$ und $\tilde{S}^{(\nu)}$ die zur Unterteilung $\underline{z}^{(\nu)}$. Da $\underline{z}^{(\nu)}$ eine Verfeinerung von $\underline{x}^{(\nu)}$ 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 $I(\underline{z}^{(\nu)}, \tilde{\xi}^{(\nu)})$, 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

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

Genauso folgt, daß auch

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

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

zeigt schließlich die Behauptung

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

Damit ist der Satz vollständig bewiesen. ■

7) Stückweise stetige Funktionen: Gelegentlich möchte man die Bedingung der Stetigkeit wenigstens ein bißchen lockern, um beispielsweise auch einen Funktionen mit gelegentlichen Sprüngen 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 \underline{x} , 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 \underline{x} eine Unterteilung des Intervalls $[0, 1]$. Unabhängig von der Wahl dieser Unterteilung können wir stets eine Folge $\underline{\xi}$ von rationalen Zwischenwerten finden, aber auch eine Folge $\underline{\tilde{\xi}}$ von irrationalen. Dann ist, unabhängig von der Unterteilung \underline{x} ,

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

und

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

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 $\int_a^a f(x) dx = 0$.

§2: Wichtige Eigenschaften des Integrals

a) Erste Integrationsregeln

Das RIEMANN-Integral stellt die heuristischen Überlegungen vom Beginn dieses Kapitels 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 \underline{x} des Intervalls $[a, b]$ und eine dazu kompatible Sequenz $\underline{\xi}$ von Zwischenwerten; $I(\underline{x}, \underline{\xi})$ sei die zugehörige RIEMANN-Summe für f und $J(\underline{x}, \underline{\xi})$ die für g . Da für jedes ξ_i nach Voraussetzung $f(\xi_i) \leq g(\xi_i)$ ist, folgt unmittelbar aus der Definition der RIEMANN-Summen, daß $I(\underline{x}, \underline{\xi}) \leq J(\underline{x}, \underline{\xi})$ ist. Da sich eine solche Kleinerleichbeziehung 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 α, β 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 \underline{x} und jede damit kompatible Zwischenwertsequenz $\underline{\xi}$ 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: Wie der Mittelwertsatz der Differentialrechnung die Ableitung mit der Steigung einer geeigneten Sehne in Verbindung bringt, sagt uns der Mittelwertsatz der Integralrechnung, daß die Fläche unterhalb einer Kurve gleich der eines gewissen Rechtecks ist. Das haben wir bei der Konstruktion des RIEMANN-Integrals bereit zur Motivation benutzt; hier soll es nun im Hinblick auf spätere Anwendungen in etwas größerer Allgemeinheit bewiesen

werden. Insbesondere wird durch dieses Satz auch die zu Beginn dieses Kapitels postulierte Interpretation von Integralen als Mittelwerte auf eine solide Grundlage stellt.

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 Monotonieregel

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

Es gibt daher eine reelle Zahl μ 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 wegen der Stetigkeit von f gibt es nach dem Zwischenwertsatz mindestens ein $\xi \in [a, b]$ mit $f(\xi) = \mu$.

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 dieser Satz für nur *stückweise* stetige Funktion f im allgemeinen *nicht* gilt: 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.

b) 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

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

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

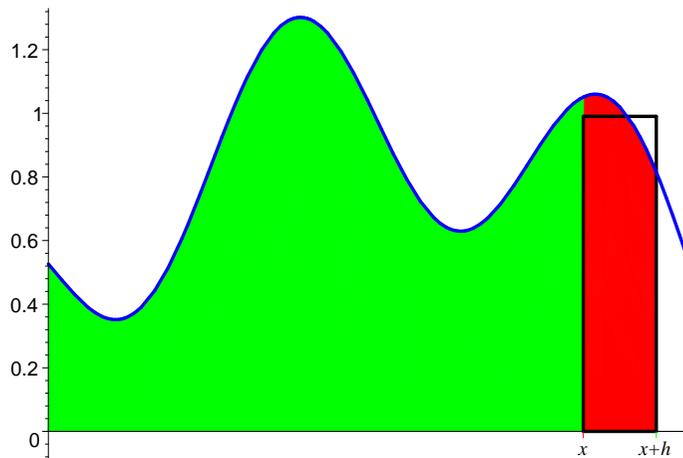
$$F_a(x+h) - F_a(x) = \int_a^{x+h} f(x) dx - \int_a^x f(x) dx = \int_x^{x+h} f(x) dx$$

(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 ξ zwischen x und $x+h$, so daß das Integral gleich $|h| \cdot f(\xi)$ ist; in der Abbildung ist dies die Fläche des schwarz umrandeten Rechtecks. Somit ist

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



$F(x)$ = grüne Fläche, $F(x+h) - F(x)$ = rote Fläche

Geht nun h gegen Null, so muß die zwischen x und $x+h$ liegende Zahl ξ

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 Kapitel 3: Wenn die Ableitung einer Funktion $h(x)$ verschwindet, muß die Funktion konstant sein. Also gibt es eine Konstante C , so daß $F(x) - F_a(x) = C$ ist und damit auch $F(x) = F_a(x) + C$, wie behauptet. ■

Für die Berechnung bestimmter Integrale bedeutet dies

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 auf eine Konstante:

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: Die Ableitung von x^n beispielsweise ist nx^{n-1} , also ist wegen der Linearitätsregel

$$\int x^n dx = \frac{x^{n+1}}{n+1} + C \quad \text{falls } n \neq -1.$$

Für $n = -1$ wissen wir, daß $x^{-1} = 1/x$ die Ableitung des Logarithmus ist; somit haben wir in diesem Fall die Stammfunktion

$$\int \frac{dx}{x} = \log x + C.$$

Die Funktion e^{ax} hat Ableitung ae^{ax} , also ist

$$\int e^{ax} dx = \frac{e^{ax}}{a} + C \quad \text{falls } a \neq 0.$$

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.

c) Partielle Integration

Eine wichtige Regel der Differentialrechnung ist die *Produktregel*

$$(uv)'(x) = u'(x)v(x) + u(x)v'(x).$$

Für die Zwecke der Integralrechnung schreiben wir sie besser als

$$u'(x)v(x) = (uv)'(x) - u(x)v'(x),$$

was integriert auf

$$\int u'(x)v(x) dx = u(x)v(x) - \int u(x)v'(x) dx$$

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 $u(x)v'(x)$ leichter integrierbar als $u'(x)v(x)$.

Als Beispiel dazu betrachten wir das Integral

$$\int x \sin x dx.$$

Hier könnten wir entweder $u'(x) = x$ und $v(x) = \sin x$ setzen; dann wäre aber $u(x)v'(x) = -\frac{1}{2}x^2 \cos x$ komplizierter als unser gegebener Integrand. Setzen wir aber $v(x) = x$ und $u'(x) = \sin x$, so können wir $u(x) = -\cos x$ setzen und $v'(x) = 1$. Somit ist

$$\int x \sin x dx = -x \cos x + \int \cos x dx = \sin x - x \cos x + C.$$

Auch in einigen gänzlich unerwarteten Fällen führt partielle Integration manchmal zum Erfolg, beispielsweise können wir damit die Stammfunktion des Logarithmus bestimmen. Der Logarithmus ist zwar nicht als Produkt zweier Funktionen definiert, aber wir können ihn natürlich schreiben als

$$\log x = u'(x)v(x) \quad \text{mit} \quad u'(x) = 1 \quad \text{und} \quad v(x) = \log x.$$

Dann ist

$$u(x) = x \quad \text{und} \quad v'(x) = \frac{1}{x};$$

nach der Regel zur partiellen Integration ist daher

$$\begin{aligned} \int \log x dx &= x \log x - \int x \cdot \frac{1}{x} dx = x \log x - \int 1 dx \\ &= x \log x - x + C = x(\log x - 1) + C. \end{aligned}$$

d) Substitutionsregel

Auch die Kettenregel $(f \circ g)'(x) = f'(g(x)) \cdot g'(x)$ 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 $f(x) = 1/x$ und damit $F(x) = \log|x|$ führt die Substitutionsregel

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

zur Integrationsregel

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

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

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

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

Auch Integrale wie

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

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

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

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

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

Als etwas umfangreicheres Beispiel wollen wir

$$\int \frac{dx}{x^2 + bx + c}$$

berechnen. Da wir bislang (außer im Fall $b = c = 0$) noch keine Funktionen kennen, die Ableitungen dieser Form haben könnten, müssen wir unsere bekannten Funktionen durchforsten und schauen, welche auf einen Ausdruck dieser Art führen könnten.

Wenn man die Ableitungen der wichtigsten elementaren Funktionen kennt, weiß man, daß hier die Umkehrfunktion des Tangens in Frage kommt: Der Tangens ist definiert als Quotient von Sinus und Kosinus und damit an den Nullstellen des Kosinus nicht definiert; wir müssen also alle $x \in \mathbb{R}$ ausschließen, die sich in der Form $(2k+1)\pi/2$ schreiben lassen mit $k \in \mathbb{Z}$. Nach der Quotientenregel ist

$$\tan' x = \frac{\cos^2 x + \sin^2 x}{\cos^2 x} = 1 + \tan^2 x,$$

der Tangens ist also in allen Intervallen, in denen er definiert ist, monoton wachsend. Speziell können wir das offene Intervall $(-\pi/2, \pi/2)$ betrachten. Da Sinus und Kosinus in $(0, \pi/2)$ positiv sind, wächst die Funktion bei linksseitiger Annäherung an $\pi/2$ unbeschränkt; bei rechtsseitiger Annäherung an $-\pi/2$ nimmt sie immer (betrags)größer werdende negative Werte an, denn in $(-\pi/2, 0)$ ist der Sinus negativ, der Kosinus aber positiv. Somit bildet der Tangens das Intervall $(-\pi/2, \pi/2)$ monoton steigend ab auf \mathbb{R} und wir haben eine differenzierbare Umkehrfunktion

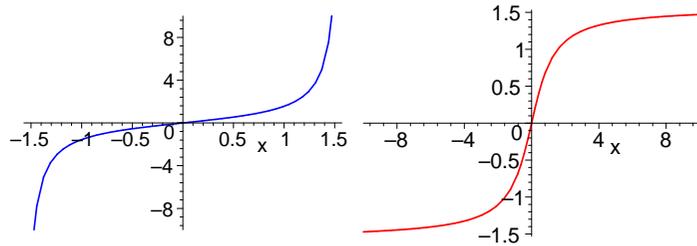
$$\arctan: \mathbb{R} \rightarrow (-\pi/2, \pi/2),$$

deren Ableitung im Punkt $y = \tan x$ wir nach der Regel zur Ableitung

der Umkehrfunktion berechnen können:

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

Somit ist $\int \frac{dy}{y^2 + 1} = \arctan y + C$.



Tangens und Arkustangens

Durch quadratische Ergänzung können wir versuchen, den Nenner unseres Integranden dieser Form anzunähern:

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

setzen wir

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

ist daher nach der Regel über lineare Substitutionen

$$\int \frac{dx}{x^2 + bx + c} = \int \frac{dy}{y^2 + d},$$

und damit sind wir schon recht nahe am obigen Integral.

Falls d positiv ist, gibt es ein $a > 0$ mit $d = a^2$ und

$$\int \frac{dy}{y^2 + d} = \int \frac{dy}{y^2 + a^2} = \int \frac{\frac{dy}{a}}{\left(\frac{y}{a}\right)^2 + 1}.$$

Hierauf können wir die Substitutionsregel anwenden mit

$$f(y) = \frac{1/a}{y^2 + 1}, \quad F(y) = \frac{1}{a} \arctan y \quad \text{und} \quad g(y) = \frac{y}{a};$$

wir erhalten

$$\int \frac{dy}{y^2 + d} = \int \frac{\frac{dy}{a}}{\left(\frac{y}{a}\right)^2 + 1} \cdot \frac{1}{a} = \frac{1}{a} \arctan \left(\frac{y}{a}\right) + C.$$

Für negative d können wir $d = -a^2$ schreiben und uns dann entweder ähnlich wie oben überlegen, daß die Umkehrfunktion des Tangens hyperbolicus $\tanh x = \frac{\sinh x}{\cosh x}$ eine Stammfunktion von $1/(x^2 - 1)$ ist, oder aber wir schreiben

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

und haben dann

$$\begin{aligned} \int \frac{dy}{y^2 + d} &= \int \frac{dy}{y^2 - a^2} = \frac{1}{2a} \left(\int \frac{dy}{y - a} - \int \frac{dy}{y + a} \right) \\ &= \frac{1}{2a} (\log(y - a) - \log(y + a)) + C = \frac{1}{2a} \log \frac{y - a}{y + a} + C. \end{aligned}$$

Setzen wir zur Abkürzung $\operatorname{artanh} x = \frac{1}{2} \log \frac{x - 1}{x + 1}$ (gesprochen *Area-tangens hyperbolicus*), so läßt sich dies auch schreiben als $\frac{1}{2a} \operatorname{artanh} \frac{y}{a}$. Somit ist mit $\Delta = b^2 - 4c = -4\Delta$

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

(Man kann sich leicht überlegen, daß der so definierte Areatangens hyperbolicus tatsächlich eine Umkehrfunktion von $\tanh x = \frac{\sinh x}{\cosh x}$ ist.)

e) Wie findet man eine Stammfunktion?

Wie das letzte Beispiel zeigt, kann die Suche nach einer Stammfunktion sehr aufwendig sein; während wir die Ableitung einer differenzierbaren Funktion im allgemeinen ohne großes Nachdenken durch schematisches Anwenden einfacher Regeln finden können, müssen wir uns bei der Suche nach einer Stammfunktion immer neue Tricks einfallen lassen, und manchmal kann es auch passieren, daß keiner unserer Ansätze zum

Erfolg führt. Für die Funktionen

$$\frac{\sin x}{x}, \quad \frac{1}{\log x}, \quad e^{-x^2} \quad \text{und} \quad \frac{1}{\sqrt{x^3 - 1}}$$

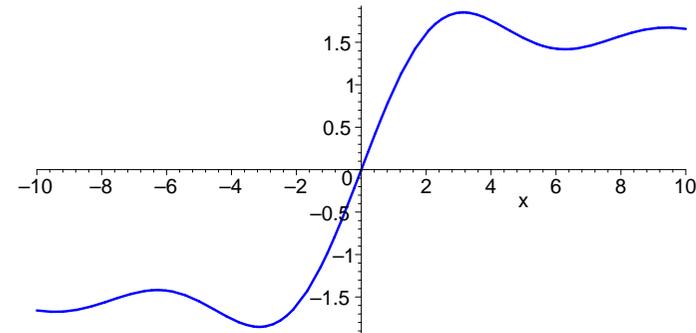
etwa kann man zeigen, daß es keine aus Grundrechenarten, Wurzeln, Exponentialfunktionen, Logarithmen, trigonometrischen und Hyperbelfunktionen sowie deren Umkehrfunktionen zusammengesetzte Funktion gibt, deren Ableitung eine der genannten Funktionen ist. Trotzdem existieren diese Stammfunktionen natürlich, und sie sind auch wichtig: Die Funktion $\frac{\sin x}{x}$ und ihre Stammfunktion spielen eine wesentliche Rolle zum Beispiel bei der Vermeidung sogenannter *alias*-Effekte sowohl bei Rastergraphiken als auch bei digitalen Audioaufzeichnungen, die Stammfunktion von $\frac{\log x}{x}$ ist eng verbunden mit der Verteilung der Primzahlen, die von e^{-x^2} wird benötigt für die wohl wichtigste Verteilung in der Statistik, die GAUSSsche Normalverteilung, und Integrale, bei denen im Nenner die Wurzel aus einem kubischen Polynom ohne mehrfache Nullstellen steht, treten z.B. auf bei der Berechnung der Bogenlänge der Ellipse und damit auch bei den GAUSS-KRÜGER-Koordinaten, die sowohl unseren amtlichen topographischen Karten als auch dem Katasterwesen zugrunde liegen. Via RIEMANN-Summen können diese Stammfunktionen auch problemlos numerisch berechnet werden, und mehr können wir mit der Wurzelfunktion oder den trigonometrischen Funktionen auch nicht machen. Daher führt man einfach neue Funktionen ein wie

den *Integralsinus*
$$\text{Si}(x) = \int_0^x \frac{\sin \xi}{\xi} d\xi$$

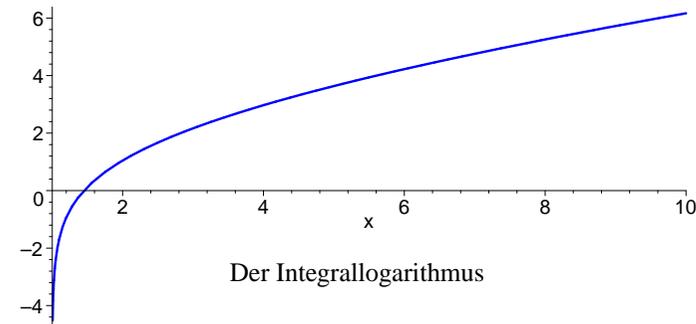
den *Integrallogarithmus*
$$\text{Li}(x) = \int_0^x \frac{d\xi}{\log \xi}$$

die *Fehlerfunktion*
$$\text{Erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-\xi^2} d\xi$$

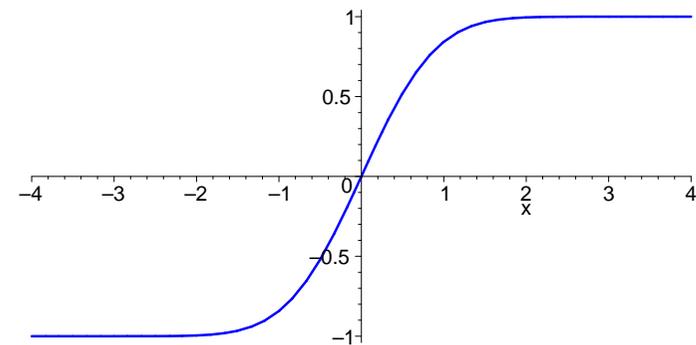
sowie eine ganze Reihe sogenannter *elliptischer Funktionen*. (Der Vorfaktor bei der Fehlerfunktion sorgt dafür, daß $\lim_{x \rightarrow \infty} \text{Erf}(x) = 1$ ist; das ist nützlich für Anwendungen in der Statistik.)



Der Integralsinus



Der Integrallogarithmus



Die Fehlerfunktion

Wir können also auch bei harmlos aussehenden Funktionen nicht sicher sein, daß wir ihre Stammfunktion durch bekannte Funktionen ausdrücken können. Schlimmer noch: Wir können das Problem auch nicht immer in Teilprobleme zerlegen. So ist etwa für $f(x) = \log x \cdot \sin x$

$$f'(x) = \frac{\sin x}{x} + \log x \cdot \cos x,$$

also

$$\int \left(\frac{\sin x}{x} + \log x \cdot \cos x \right) dx = \log x \cdot \sin x + C.$$

Nach der Linearitätsregel ist auch

$$\int \left(\frac{\sin x}{x} + \log x \cdot \cos x \right) dx = \int \frac{\sin x}{x} dx + \int \log x \cdot \cos x dx,$$

aber da das erste Integral rechts auf den nicht elementar ausdrückbaren Integralsinus führt, sind beide Integrale auf der rechten Seite nicht durch elementare Funktionen ausdrückbar.

Früher sagte man daher häufig, die Differentiation sei ein Handwerk, die Integration aber eine Kunst. In den Sechzigerjahren galt Integration auch als Paradebeispiel für Anwendungen der Künstlichen Intelligenz; wie bei vielen der damals populären KI-Anwendungen gab es jedoch auch hier keine nennenswerten Erfolge.

Wer allerdings eines der gängigen Computeralgebrasysteme nach Stammfunktionen suchen läßt, bekommt im allgemeinen erstaunlich schnell eine Antwort, die zwar gelegentlich für den Laien schwer verständlich ist, in einfachen Fällen aber meist etwas liefert, von dem man sich leicht überzeugen kann, daß es eine Stammfunktion ist. Was hier im Hintergrund abläuft hat nichts mit Künstlicher Intelligenz zu tun, sondern mit dem was Computeralgebrasystemen ihren Namen gegeben hat, mit Algebra also. Ausgehend von Sätzen aus dem neunzehnten Jahrhundert kann man für große Klassen von Funktionen nicht nur eine eventuell existierende elementar darstellbare Stammfunktion konstruieren, sondern gegebenenfalls auch beweisen, daß keine existiert. Die dabei verwendeten Algorithmen aus der Differentialalgebra und die für dem Beweis von deren Korrektheit benötigten Sätze aus der Funktionentheorie liegen deutlich jenseits des Niveaus einer Vorlesung Analysis I und haben insbesondere nichts zu tun mit den vielen Integrationstricks, die man

in Analysis-Lehrbüchern findet. Wegen der Allgegenwart auch freier Computeralgebrasysteme sei deshalb hier auf die Behandlung weiterer solcher Tricks verzichtet. Wer wissen möchte, wie ein Computeralgebrasystem integriert, findet eine Kurzdarstellung zum Beispiel bei

J. H. DAVENPORT, Y. SIRET, E. TOURNIER: Computer algebra: systems and algorithms for algebraic computations, *Academic Press*, 1993;

für eine ausführliche Darstellung mit Beweisen sei verwiesen auf

MANUEL BRONSTEIN: Symbolic Integration I: Transcendental Functions, *Springer*, 2005

Der deutlich schwierigere Fall von Integranden mit Wurzelausdrücken wird behandelt in

JAMES HAROLD DAVENPORT: On the integration of algebraic functions, *Lecture notes in computer science* **102**, 1981

f) 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 zuordnen 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.

Wir definieren daher für eine Funktion, die in einem halboffenen Intervall $[a, b)$ mit $b \in \mathbb{R} \cup \{\infty\}$ definiert und dort stückweise stetig ist, das rechtsseitig uneigentliche Integral als

$$\int_a^b f(x) dx = \lim_{c \nearrow b} \int_a^c f(x) dx,$$

falls dieser Grenzwert existiert; andernfalls sagen wir, das Integral sei *divergent*.

Völlig analog erklären wir linksseitige uneigentliche Integrale: f sei definiert und stückweise stetig auf dem halboffenen Intervall $(a, b]$ mit $a \in \mathbb{R} \cup \{-\infty\}$; dann ist

$$\int_a^b f(x) dx = \lim_{c \searrow a} \int_c^b f(x) dx,$$

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 $(a, c]$ und $[c, b)$. 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 uneigentlichen 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

$$\begin{aligned} \int_{-\infty}^{\infty} \frac{dx}{1+x^2} &= \int_{-\infty}^0 \frac{dx}{1+x^2} + \int_0^{\infty} \frac{dx}{1+x^2} \\ &= \lim_{c \rightarrow -\infty} \int_c^0 \frac{dx}{1+x^2} + \lim_{d \rightarrow \infty} \int_0^d \frac{dx}{1+x^2} \\ &= - \lim_{c \rightarrow -\infty} \arctan c + \lim_{d \rightarrow \infty} \arctan d = \frac{\pi}{2} + \frac{\pi}{2} = \pi. \end{aligned}$$

Die Forderung, daß *jedes* der Teilintegrale einzeln konvergieren soll, ist gelegentlich zu 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

$$\int_{-2}^4 \frac{dx}{x^3} = \int_{-2}^2 \frac{dx}{x^3} + \int_2^4 \frac{dx}{x^3} = 0 + \frac{3}{32} = \frac{3}{32}$$

ist. 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 gravierenden Konsequenzen, als daß man einfach durch diesen Punkt hindurch weitergehen könnte.

Andererseits beschreiben aber in Anwendungen der Mathematik viele Funktionen nur näherungsweise die Größe, die sie darstellen sollen: Mathematische Modelle sind praktisch immer *vereinfachende* Modelle der Wirklichkeit. So gilt beispielsweise das OHMSche Gesetz sicherlich nicht mehr, wenn man einen 5Ω -Widerstand aus einer auf 5 V Spannung ausgelegten Schaltung im Hochspannungslabor mit 100 kV belastet, und es gilt auch nicht mehr ohne Korrekturterme, wenn man einen Wechselspannung mit 500 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 $\int_a^b f(x) dx$. 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 $\frac{3}{32}$ der CAUCHYSche Hauptwert des Integrals, wohingegen das Integral selbst nicht existiert.

Die Frage, wann der CAUCHYSche Hauptwert für ein divergentes Integral verwendet werden sollte, ist keine mathematische Frage: Unter rein 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

$$\int_0^b \frac{dx}{x^r} = \lim_{a \rightarrow 0} \int_a^b \frac{dx}{x^r} = \frac{1}{1-r} b^{1-r}.$$

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, 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 betrachten wir das Integral

$$\Gamma(x) \stackrel{\text{def}}{=} \int_0^\infty e^{-t} t^{x-1} dt \quad \text{für } x > 0.$$

Es 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

$$e^{-t} t^{x-1} = \frac{e^{-t}}{t^{1-x}} \leq \frac{1}{t^{1-x}} \quad \text{mit } 0 < 1-x < 1,$$

so daß obiger Satz sofort die Konvergenz zeigt.

Auch die obere Grenze ist unproblematisch, denn die dazu notwendige Abschätzung

$$e^{-t} t^{x-1} \leq \frac{K}{t^r} \iff e^t \geq \frac{t^{r+x-1}}{K}$$

folgt, da die Exponentialfunktion stärker wächst als jede Potenz.

Die somit für alle $x > 0$ definierte Funktion $x \mapsto \Gamma(x)$ heißt EULERSche Gamma-Funktion. Ihre wichtigste Eigenschaft der folgt durch partielle

Integration:

$$\Gamma(x) = \int_0^\infty e^{-t} t^{x-1} dt = e^{-t} \frac{t^x}{x} \Big|_0^\infty + \frac{1}{x} \int_0^\infty e^{-t} t^x dt = \frac{\Gamma(x+1)}{x}$$

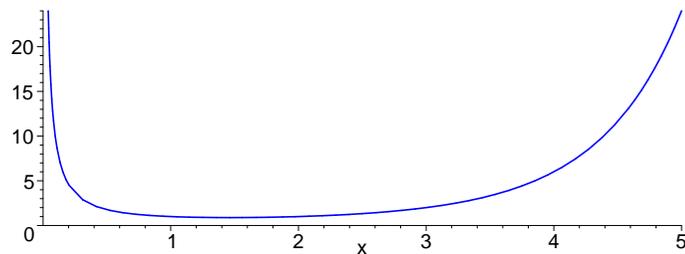
oder $\Gamma(x+1) = x\Gamma(x)$ für alle reellen $x \geq 0$. Mit dem elementar berechenbaren Wert

$$\Gamma(1) = \int_0^\infty e^{-t} dt = -e^{-t} \Big|_0^\infty = 1$$

führt dies auf die Formel

$$\Gamma(n) = (n-1)! \quad \text{für alle } n \in \mathbb{N};$$

die Γ -Funktion ist also eine Art stetig gemachter Fakultätsfunktion.



Die Γ -Funktion

GAUSS definierte auf andere Weise eine stetige Funktion $\Pi(x)$, für die $\Pi(n) = n!$ ist, aber wie sich bald herausstellte, ist $\Pi(x) = \Gamma(x+1)$, so daß nur eine der beiden Funktionen wirklich gebraucht wird. Nach einigen Modewechseln im letzten Jahrhundert entscheidet man sich heute meist für $\Gamma(x)$: Diese Funktionswerte sind in Tafelwerken tabelliert, und numerische Verfahren für ihre Berechnung stehen in den einschlägigen Unterprogramm-bibliotheken und Computeralgebrasystemen zur Verfügung.

g) Summen, Reihen und Integrale

Das Integralzeichen entstand als ein stilisiertes „S“ wie Summe und wir haben bestimmte Integrale auch definiert über RIEMANNSche Summen.

Von daher sollte es nicht verwundern, daß umgekehrt auch Integrale Aussagen Summen und Reihen ermöglichen. Einer der einfachsten Zusammenhänge ist das folgende

Lemma: Die Funktion f sei stetig, monoton fallend und nichtnegativ für alle $x \geq m, m \in \mathbb{Z}$. Dann ist für jedes $n \geq m$

$$\int_m^{n+1} f(x) dx \leq \sum_{k=m}^n f(k) \leq f(m) + \int_m^{n+1} f(x) dx.$$

Beweis: Wegen der Monotonie von f ist für jede ganze Zahl $k \geq m$ und jede reelle Zahl x mit $k \leq x \leq k+1$

$$f(k+1) \leq f(x) \leq f(k).$$

Da das Integral über eine konstante Funktion über ein Intervall der Länge eins einfach diese Konstante ist, folgt daher aus der Monotonieregel die Ungleichungskette

$$f(k+1) = \int_k^{k+1} f(k+1) dx \leq \int_k^{k+1} f(x) dx \leq \int_k^{k+1} f(k) dx = f(k).$$

Summieren wir dies über alle k vom m bis zu einer oberen Schranke n , erhalten wir

$$\sum_{k=m}^n f(k+1) = \sum_{k=m}^{n+1} f(k) - f(m) \leq \int_m^{n+1} f(x) dx \leq \sum_{k=m}^n f(k).$$

Somit ist

$$\int_m^{n+1} f(x) dx \leq \sum_{k=m}^n f(k) \leq f(m) + \int_m^{n+1} f(x) dx.$$

■

Lassen wir hier n gegen unendlich gehen, erhalten wir den

Satz: Die Funktion f sei stetig und monoton fallend für alle $x \geq m, m \in \mathbb{Z}$. Dann gilt: Die Reihe $\sum_{k=m}^\infty f(k)$ konvergiert genau dann, wenn

das uneigentliche Integral $\int_m^\infty f(x) dx$ existiert. In diesem Fall ist

$$\int_m^\infty f(x) dx \leq \sum_{k=m}^\infty f(k) \leq f(m) + \int_m^\infty f(x) dx.$$

Beweis: Wegen der Nichtnegativität von f sind

$$S_n = \sum_{k=m}^n f(k) \quad \text{und} \quad I_n = \int_m^n f(x) dx$$

monoton wachsend mit n . Falls das uneigentliche Integral konvergiert, ist sein Wert I daher eine obere Schranke für alle I_n und damit ist $I + f(m)$ nach dem gerade bewiesenen Lemma eine obere Schranke für alle S_n . Somit ist die Folge $(S_n)_{n \geq m}$ monoton wachsend und beschränkt, also nach Kapitel 2, §3, konvergent. Konvergiert umgekehrt diese Folge gegen einen Grenzwert S , so ist dieser eine obere Schranke für alle S_n , also nach dem Lemma auch für alle I_n , und damit konvergiert das Integral. Die behauptete Ungleichungskette folgt somit aus den Rechenregeln für Grenzwerte in Kapitel 2, §2. ■

Als erste Anwendung können wir noch einmal die harmonische Reihe betrachten: Nach dem Lemma ist

$$\log(n+1) \leq H_n = \sum_{k=1}^n \frac{1}{k} \leq 1 + \log(n+1),$$

die Folge der H_n geht also mit $\log n$ gegen unendlich, was auch einen neuen Beweis für die Divergenz der harmonischen Reihe liefert.

Betrachten wir dagegen für ein $r > 1$ die Summe $\sum_{k=1}^\infty \frac{1}{k^r}$, so konvergiert diese, denn das uneigentliche Integral

$$\int_1^\infty \frac{dx}{x^r} = \frac{1}{(1-r)x^{r-1}} \Big|_1^\infty = \frac{1}{r-1}$$

existiert. Für $r = 2$ etwa erhalten wir die Abschätzung

$$1 \leq \sum_{k=1}^\infty \frac{1}{k^2} \leq 2,$$

die wir in Kapitel 2 auch ohne Intergration gefunden haben. Um etwas Besseres zu bekommen, können wir die ersten Summanden stehen lassen und erst den Rest abschätzen: Beispielsweise ist

$$\sum_{k=1}^4 \frac{1}{k^2} = \frac{205}{144} = 1,4236\bar{1} \quad \text{und} \quad \int_5^\infty \frac{dx}{x^2} \leq \sum_{k=5}^\infty \frac{1}{k^2} \leq \frac{1}{5} + \int_5^\infty \frac{dx}{x^2};$$

da das Integral der Wert $1/5$ hat, liegt die Gesamtsumme somit zwischen $1,6236\bar{1}$ und $1,8236\bar{1}$.

h) Die Eulersche Summenformel

Wir betrachten irgendeine reellwertige differenzierbare Funktion f , deren Definitionsbereich das Intervall $[1, n]$ enthält.

Für eine reelle Zahl x bezeichnen wir wie üblich mit $[x]$ die größte ganze Zahl kleiner oder gleich x und mit $\{x\} \stackrel{\text{def}}{=} x - [x]$ den gebrochenen Anteil von x . Für eine ganze Zahl k ist somit $\{x\} = x - k$ für alle x aus dem Intervall $[k, k+1)$.

Partielle Integration führt auf die Gleichung

$$\begin{aligned} \int_k^{k+1} (\{x\} - \tfrac{1}{2}) f'(x) dx &= (x - k - \tfrac{1}{2}) f(x) \Big|_k^{k+1} - \int_k^{k+1} f(x) dx \\ &= \frac{f(k+1) + f(k)}{2} - \int_k^{k+1} f(x) dx. \end{aligned}$$

Addition aller solcher Gleichungen von $k = 1$ bis $k = n - 1$ liefert

$$\int_1^n (\{x\} - \tfrac{1}{2}) f'(x) dx = \frac{f(1)}{2} + \sum_{k=2}^{n-1} f(k) + \frac{f(n)}{2} - \int_1^n f(x) dx,$$

womit man die Summe der $f(k)$ berechnen kann:

Satz (EULERSche Summenformel, einfachster Fall): Für eine differenzierbare Funktion $f: D \rightarrow \mathbb{R}$, deren Definitionsbereich das Intervall $[1, n]$ umfaßt, ist

$$\sum_{k=1}^n f(k) = \int_1^n f(x) dx + \frac{f(1) + f(n)}{2} + \int_1^n (\{x\} - \frac{1}{2}) f'(x) dx.$$

Als Beispiel wollen wir eine Näherungsformel für $n!$ herleiten: Wir schreiben $\log n! = \sum_{k=1}^n \log k$ und haben dann nach der EULERSchen Summenformel

$$\begin{aligned} \log n! &= \int_1^n \log x dx + \frac{\log n}{2} + \int_1^n \frac{\{x\} - \frac{1}{2}}{x} dx \\ &= x(\log x - 1) \Big|_1^n + \frac{\log n}{2} + \int_1^n \frac{\{x\} - \frac{1}{2}}{x} dx \\ &= n(\log n - 1) + 1 + \frac{\log n}{2} + \int_1^n \frac{\{x\} - \frac{1}{2}}{x} dx. \end{aligned}$$

In dieser Formel stört noch das rechte Integral; dieses können wir wie folgt abschätzen: Für eine natürliche Zahl k ist

$$\begin{aligned} \int_k^{k+1} \frac{\{x\} - \frac{1}{2}}{x} dx &= \int_{-\frac{1}{2}}^{\frac{1}{2}} \frac{x}{k + \frac{1}{2} + x} dx \\ &= \int_0^{\frac{1}{2}} \left(\frac{x}{k + \frac{1}{2} + x} - \frac{x}{k + \frac{1}{2} - x} \right) dx = \int_0^{\frac{1}{2}} \frac{-2x^2}{(k + \frac{1}{2})^2 - x^2} dx. \end{aligned}$$

Im Intervall von 0 bis $\frac{1}{2}$ ist der Integrand monoton fallend, d.h.

$$0 \geq \frac{-2x^2}{(k + \frac{1}{2})^2 - x^2} \geq \frac{-\frac{1}{2}}{(k + \frac{1}{2})^2 - \frac{1}{4}} = \frac{-2}{(2k + 1)^2 - 1} \geq -\frac{1}{4k^2},$$

und damit ist

$$0 \geq \int_k^{k+1} \frac{\{x\} - \frac{1}{2}}{x} dx = \int_0^{\frac{1}{2}} \frac{-2x^2}{(k + \frac{1}{2})^2 - x^2} dx \geq -\frac{1}{8k^2},$$

denn wir können das Integral abschätzen durch das Produkt aus der Länge des Integrationsintervalls und dem Minimum des Integranden. Summation von $k = 1$ bis $n - 1$ schließlich gibt die Abschätzung

$$0 \geq \int_1^n \frac{\{x\} - \frac{1}{2}}{x} dx \geq -\sum_{k=1}^{n-1} \frac{1}{4k^2}$$

für das störende Integral aus der obigen Formel.

Von der rechtsstehenden Summe wissen wir aus Kapitel 2, §6, daß sie (egal ob mit oder ohne acht im Nenner) für $n \rightarrow \infty$ konvergiert und können daher folgern, daß gilt

$$\log n! = n(\log n - 1) + \frac{\log n}{2} + R_n$$

mit einem Fehlerterm R_n , der für $n \rightarrow \infty$ beschränkt bleibt. Somit ist

$$n! \approx C \cdot n^n e^{-n} \sqrt{n}$$

mit einer gewissen Konstanten C , deren Wert $\sqrt{2\pi}$ nur auf dem Umweg über ein zweidimensionales Integral und damit nicht mit unseren Mitteln berechnet werden kann. Dieser Näherungsausdruck heißt STIRLINGSche Formeln; genau wie die Summenformel kann er noch beliebig verbessert werden.

Der schottische Mathematiker JAMES STIRLING (1692–1770) war Anhänger des gestürzten Königs Jakob II Stuart und hatte deshalb große politische Probleme bei seinem Studium; unter anderem wurde er deshalb von der Universität Oxford ausgeschlossen. 1717–1722 lebte er in Venedig und hatte auch gute Kontakte zu NICOLAUS BERNOULLI an der Universität von Padua; außerdem brachte er aus Venedig die Produktionsgeheimnisse der dortigen Glasbläser mit. Ab 1724 arbeitete er zehn Jahre lang als Mathematiklehrer in London, wo er viel mit NEWTON zusammentraf; 1735 wurde er Direktor einer schottischen Bergbaugesellschaft. In seine Londoner Zeit fällt die Veröffentlichung seines bedeutendsten Werks *Methodus Differentialis sive Tractatus de Summatione et Interpolatione Serierum Infinitarum* im Jahre 1730, das die obige Formel als Beispiel zwei zu Proposition 28 enthält. Ebenfalls ziemlich bekannt wurde seine 1735 veröffentlichte Arbeit über die Gestalt der Erde.

§3: Zusammenfassung

Im letzten Kapitel dieses Semesters befaßten wir uns mit der Integration von Funktionen einer Veränderlichen. Die Integration ist die Umkehrung der Differentiation; ist $F(x)$ eine differenzierbare Funktion mit Ableitung $f(x)$, so bezeichnen wir $F(x)$ als eine *Stammfunktion* von $f(x)$ und schreiben

$$\int f(x) dx = F(x) + C .$$

Stammfunktionen sind nicht nur bis auf eine additive Konstante $C \in \mathbb{R}$ bestimmt.

Integrale können auch eingeführt werden via Flächenberechnung; wir bezeichnen die Fläche unterhalb des Graphen der Funktion $f(x)$ zwischen den Werten $x = a$ und $x = b$ als das *bestimmte Integral*

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

Flächenteile, die unterhalb der x -Achse liegen, werden dabei als negativ betrachtet. Ist $F(x)$ eine Stammfunktion von $f(x)$, so ist

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

Das bestimmte Integral existiert zumindest dann, wenn die Funktion f im Intervall $[a, b]$ stückweise stetig ist.

Zur Berechnung von Stammfunktionen können wir die Differentiationsregeln umkehren: Aus der Linearität der Ableitung folgt die der Integration:

$$\int (cf(x) + dg(x)) dx = c \int f(x) dx + d \int g(x) dx .$$

Die LEIBNIZ-Regel $(uv)' = u'v + uv'$ führt auf die Regel zur *partiellen Integration*:

$$\int u'(x)v(x) dx = u(x)v(x) - \int u(x)v'(x) dx ;$$

die Kettenregel führt zur Substitutionsregel

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

wobei F eine Stammfunktion von f bezeichnet. Ein wichtiger Spezialfall ergibt sich für $f(x) = 1/x$ und damit $F(x) = \log |x|$:

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

Als Grenzwerte endlicher Integrale lassen sich in manchen Fällen auch bestimmte Integrale mit unterer Grenze $-\infty$ und/oder oberer Grenze ∞ definieren; auch bei Definitionslücken der zu integrierenden Funktion läßt sich gelegentlich (aber bei weitem nicht immer!) durch Grenzwertbetrachtungen ein Integral definieren.

Der Zusammenhang zwischen bestimmten Integralen und Summen läßt sich ausnutzen, um endliche wie auch unendliche Summen abzuschätzen; insbesondere ergibt sich daraus ein Konvergenzkriterium für unendliche Reihen.