Kapitel I

Analysis 1

§1 Topologie im $\mathbf{R}^n$

Es sei (X, d) ein *metrischer Raum*. Unter diesen Begriff fallen alle normierten Vektorräume, aber auch beliebige Teilmengen solcher Räume. Wichtigstes Beispiel wird immer der $\mathbb{R}^n$ sein.

Definition:

Eine Folge $(x_n)_{n\in\mathbb{N}}$ von Punkten aus X *konvergiert* gegen ein $x_0 \in X$, wenn gilt:

$$\forall\, \varepsilon > 0\ \exists\, n_0 \in \mathbb{N},\ \text{so daß}\ \forall\, n \geq n_0\ \text{gilt: } d(x_n, x_0) < \varepsilon.$$

Man schreibt dann: $\lim\limits_{n\to\infty} x_n = x_0$.

Die Definition entspricht w ö r t l i c h der Definition der Konvergenz von Zahlenfolgen, wenn man $d(x_n, x_0)$ statt $|\, x_n - x_0 \,|$ schreibt.

Beispiel :

$\mathbf{x}_n := (\frac{1}{n}, \frac{n}{n+1})$ ist eine Punktfolge im $\mathbb{R}^2$. Wir vermuten natürlich, daß sie gegen $\mathbf{x}_0 := (0, 1)$ konvergiert. Sei also $\varepsilon > 0$ vorgegeben. Dann gibt es ein n_0, so daß gilt:

$$\frac{1}{n} < \frac{\varepsilon}{2} \text{ und } \left|\, \frac{n}{n+1} - 1 \,\right| < \frac{\varepsilon}{2} \text{ für } n \geq n_0.$$

Aber dann ist

$$d(\mathbf{x}_n, \mathbf{x}_0) = \sqrt{(\frac{1}{n})^2 + (\frac{n}{n+1} - 1)^2} \leq \frac{1}{n} + \left|\, \frac{n}{n+1} - 1 \,\right| < \varepsilon,$$

für $n \geq n_0$.

Dabei haben wir folgende allgemeine Abschätzung benutzt:

$$\sqrt{a^2 + b^2} = \sqrt{(|\, a \,| + |\, b \,|)^2 - 2|\, a \,| \cdot |\, b \,|} \leq |\, a \,| + |\, b \,|.$$

Man kann das Beweisschema aus dem Beispiel verallgemeinern und zeigen:

Komponentenweise Konvergenz von Vektorfolgen

Eine Folge $\mathbf{x}_\nu = (x_1^{(\nu)}, \ldots, x_n^{(\nu)})$ *im* $\mathbb{R}^n$ *konvergiert genau dann gegen ein* $\mathbf{x}_0 = (x_1^{(0)}, \ldots, x_n^{(0)})$, *wenn jede der* n *Folgen* $(x_i^{(\nu)})$ *im gewöhnlichen Sinne gegen* $x_i^{(0)}$ *konvergiert.*

Damit lassen sich die meisten Konvergenzprobleme in mehreren Veränderlichen auf solche in einer Veränderlichen zurückführen.

Nun soll für den $\mathbb{R}^n$ – und allgemeiner sogar für jeden metrischen Raum X – erklärt werden, was es bedeutet, daß Punkte mehr oder weniger benachbart sind. Dazu brauchen wir den Begriff der „offenen Menge".

Definition:

Sei (X, d) ein metrischer Raum, $x_0 \in X$ und $r > 0$. Dann heißt

$$B_r(x_0) := \{x \in X \mid d(x, x_0) < r\}.$$

die *offene Kugel* um x_0 mit Radius r.

Eine Teilmenge $U \subset X$ heißt *offen*, wenn es zu jedem $x \in U$ ein $\varepsilon > 0$ gibt, so daß $B_\varepsilon(x)$ ganz in U liegt.

Offene Kugeln im $\mathbb{R}^n$ sind das, was man sich darunter vorstellt. „Offen" bedeutet hierbei, daß der Rand nicht dazu gehört. Entsprechend soll man sich offene Teilmengen des $\mathbb{R}^n$ so vorstellen, daß die „Randpunkte" nicht mehr dazu gehören. Das wird später noch weiter präzisiert werden.

Beispiele :

1. Offene Kugeln $B_r(x_0)$ sind offen:

   Ist nämlich $x \in B_r(x_0)$, so ist $\delta := d(x, x_0) < r$, also $r - \delta > 0$. Man kann ein ε mit $0 < \varepsilon < r - \delta$ finden. Für $y \in B_\varepsilon(x)$ gilt dann:

   $$\begin{aligned} d(y, x_0) &\leq d(y, x) + d(x, x_0) \text{ (Dreiecks-Ungleichung)} \\ &< \varepsilon + \delta < (r - \delta) + \delta = r. \end{aligned}$$

   Also liegt $B_\varepsilon(x)$ in $B_r(x_0)$.

2. In $\mathbb{R}$ ist jedes offene Intervall offen. Ist nämlich $a < b$ und $x_0 := \frac{a+b}{2}$ und $r := \frac{b-a}{2}$, so ist $(a, b) = B_r(x_0)$.

Das System aller offenen Mengen eines metrischen Raumes nennt man die „Topologie" dieses Raumes.

Die Topologie hat gewisse charakteristische Eigenschaften:

Eigenschaften offener Mengen

Die offenen Mengen in einem metrischen Raum (X, d) *(und damit insbesondere im* $\mathbb{R}^n$*) haben folgende Eigenschaften:*

1. *Die leere Menge ist offen.*
2. *Der ganze Raum* X *ist offen.*
3. *Sind* $U_1, U_2 \subset X$ *offen, so ist auch* $U_1 \cap U_2$ *offen.*
4. *Ist* $(U_\lambda)_{\lambda \in L}$ *eine beliebige Familie von offenen Mengen in* X*, so ist auch* $\bigcup_{\lambda \in L} U_\lambda$ *offen.*

BEWEIS: 1) Die leere Menge hat keine Elemente, also ist auch nichts zu zeigen.

2) Ist $x \in X$ und $r > 0$, so liegt natürlich $B_r(x)$ in X. Also ist X offen.

3) Seien $U_1, U_2 \subset X$ offen, $x_0 \in U_1 \cap U_2$ beliebig vorgegeben. Dann gibt es $\varepsilon_1, \varepsilon_2 > 0$, so daß $B_{\varepsilon_1}(x_0) \subset U_1$ und $B_{\varepsilon_2}(x_0) \subset U_2$ ist. Setzt man $\varepsilon := \min(\varepsilon_1, \varepsilon_2)$, so ist $B_\varepsilon(x_0) \subset U_1 \cap U_2$. Also ist der Durchschnitt von zwei offenen Mengen wieder offen.

4) Sei $(U_\lambda)_{\lambda \in L}$ eine Familie von offenen Mengen in X, $x_0 \in \bigcup_{\lambda \in L} U_\lambda$. Das bedeutet:

$$\begin{aligned} & \exists\, \lambda \in L \text{ mit } x_0 \in U_\lambda \\ \Longrightarrow \quad & \exists\, \lambda \in L,\ \exists\, \varepsilon > 0, \text{ so daß } B_\varepsilon(x_0) \subset U_\lambda \\ \Longrightarrow \quad & \exists\, \varepsilon > 0 \text{ mit } B_\varepsilon(x_0) \subset \bigcup_{\lambda \in L} U_\lambda \end{aligned}$$

Also ist die Vereinigung der U_λ offen. □

Der Durchschnitt von endlich vielen offenen Mengen ist wieder offen (Induktion!), aber bei unendlichen Durchschnitten kann es schiefgehen. Dafür werden wir später noch ein Beispiel angeben.

Definition:

Eine Teilmenge $A \subset X$ heißt *abgeschlossen*, wenn $\complement_X(A) = X \setminus A$ offen ist.

Beispiel:

In $\mathbb{R}$ ist auch jedes Intervall der Gestalt $(-\infty, b)$ oder (a, ∞) offen, als abzählbare Vereinigung offener Intervalle.[1] Also ist

$$\mathbb{R} \setminus [a, b] = (-\infty, a) \cup (b, \infty)$$

offen und $[a, b]$ damit abgeschlossen.

[1] Es ist z.B. $(a, \infty) = (a, a+2) \cup (a+1, a+3) \cup (a+2, a+4) \cup \ldots$.

Eigenschaften abgeschlossener Mengen

1. *Die leere Menge ist abgeschlossen.*
2. *Der ganze Raum X ist abgeschlossen.*
3. *Sind $A_1, A_2 \subset X$ abgeschlossen, so ist auch $A_1 \cup A_2$ abgeschlossen.*
4. *Ist $(A_\lambda)_{\lambda \in L}$ eine Familie von abgeschlossenen Mengen in X, so ist auch $\bigcap_{\lambda \in L} A_\lambda$ abgeschlossen.*

Die Beweise sind denkbar einfach. Es ist $X \setminus \varnothing = X$, $X \setminus X = \varnothing$, $\complement_X(A_1 \cup A_2) = \complement_X(A_1) \cap \complement_X(A_2)$ und $\complement_X(\bigcap_{\lambda \in L} A_\lambda) = \bigcup_{\lambda \in L} \complement_X(A_\lambda)$.

Wir lernen hier, daß eine Menge durchaus gleichzeitig offen und abgeschlossen sein kann. Das logische Gegenteil zu „offen“ ist nicht „abgeschlossen“, sondern „nicht offen“. Es kann auch passieren, daß eine Menge weder offen noch abgeschlossen ist, wie z.B. das Intervall $[a, b)$.

Mit Hilfe von Folgen läßt sich der Begriff der abgeschlossenen Menge etwas handlicher charakterisieren.

Abgeschlossenheits-Kriterium

Eine Menge A in einem metrischen Raum X ist genau dann abgeschlossen, wenn gilt:

Ist $x_n \in A$ eine gegen ein $x_0 \in X$ konvergente Folge, so liegt auch x_0 in A.

Beweis: 1) A sei abgeschlossen in X, (x_n) eine Folge in A, $x_0 = \lim_{n \to \infty} x_n$.

Annahme, x_0 liegt in $X \setminus A$. Da diese Menge offen ist, gibt es ein $\varepsilon > 0$, so daß sogar $B_\varepsilon(x_0)$ in $X \setminus A$ liegt. Wegen der Konvergenz der Folge muß es aber ein n_0 geben, so daß x_n in $B_\varepsilon(x_0)$ liegt, für $n \geq n_0$. Das ist ein Widerspruch.

2) Sei umgekehrt das Kriterium erfüllt. Wir müssen zeigen, daß $X \setminus A$ offen ist, und wir nehmen an, das wäre nicht der Fall.

Dann gibt es ein $x_0 \in X \setminus A$, so daß gilt:

$$\forall n \in \mathbb{N} \text{ ist } B_{1/n}(x_0) \not\subset X \setminus A.$$

Das bedeutet, daß wir für alle $n \in \mathbb{N}$ ein $x_n \in B_{1/n}(x_0)$ finden können, das nicht in $X \setminus A$ liegt.

(x_n) ist also eine Folge in A. Ist $\varepsilon > 0$ vorgegeben, so gibt es ein n_0, so daß $\frac{1}{n} < \varepsilon$ für $n \geq n_0$ ist. Aber dann ist $d(x_n, x_0) < \frac{1}{n} < \varepsilon$, d.h. (x_n) konvergiert gegen x_0, und nach dem Kriterium liegt x_0 in A. Das ist ein Widerspruch, die Annahme war falsch. $\square$

Beispiel:

Jede 1-punktige Menge $A := \{x_0\}$ in einem metrischen Raum X ist abgeschlossen.

Ist nämlich (x_n) eine Folge in A, so muß $x_n = x_0$ für alle n gelten. Aber dann konvergiert (x_n) auch gegen x_0, und dieser Grenzwert liegt in A.

Definition:

Sei x_0 ein Punkt in einem metrischen Raum X. Eine Teilmenge $U \subset X$ heißt *Umgebung* von x_0, wenn gilt:

Es gibt ein $\varepsilon > 0$, so daß $B_\varepsilon(x_0)$ in U liegt.

Wir sagen, x_0 liegt ganz im Inneren von U, wenn U eine Umgebung für x_0 ist. Ein Randpunkt von U liegt natürlich nicht im Inneren von U.

Die Umgebung U braucht weder offen noch abgeschlossen zu sein. Ist sie aber offen, so sprechen wir auch von einer offenen Umgebung. Ist U abgeschlossen, so sprechen wir von einer abgeschlossenen Umgebung.

Definition:

Sei $M \subset X$ eine beliebige Teilmenge. Ein Punkt $x_0 \in X$ (der nicht in M liegen muß) heißt *Randpunkt* von M, wenn jede Umgebung von x_0 wenigstens einen Punkt aus M und einen Punkt aus $X \setminus M$ enthält.

Die Menge aller Randpunkte von M bezeichnet man mit ∂M („Rand von M").

Beispiele:

1. In $\mathbb{R}$ ist $\partial[a,b) = \{a,b\}$.
2. Im $\mathbb{R}^n$ ist $\partial B_r(x_0) = \{x \in X \mid d(x,x_0) = r\}$.
3. In $\mathbb{R}$ ist $\partial\{\frac{1}{n} \mid n \in \mathbb{N}\} = \{\frac{1}{n} \mid n \in \mathbb{N}\} \cup \{0\}$.

Offene Mengen enthalten keine Randpunkte

Eine Menge $M \subset X$ ist genau dann offen, wenn $M \cap \partial M = \varnothing$ ist.

Der Beweis ist klar: Jeder Punkt einer offenen Menge M liegt ganz im Inneren von M. Ein Randpunkt einer Menge kann aber nie im Inneren der Menge liegen.

Abgeschlossene Mengen enthalten alle Randpunkte

Eine Menge $M \subset X$ ist genau dann abgeschlossen, wenn $\partial M \subset M$ ist.

BEWEIS: Sei zunächst M als abgeschlossen vorausgesetzt, $x_0 \in \partial M$. Wir wollen zeigen, daß x_0 in M liegt. Angenommen, $x_0 \in X \setminus M$. Da $X \setminus M$ offen ist, gibt es eine Umgebung U von x_0, die noch ganz in $X \setminus M$ liegt. Also enthält U keinen Punkt von M, und x_0 kann demnach kein Randpunkt von M sein. Widerspruch!

Ist umgekehrt $\partial M \subset M$, so müssen wir zeigen, daß M abgeschlossen ist. Ein Punkt $x_0 \in X \setminus M$ kann nicht Randpunkt von M sein. Also gibt es eine Umgebung U von x_0 mit $U \cap M = \varnothing$. Also liegt x_0 ganz im Inneren von $X \setminus M$, und da der Punkt beliebig war, bedeutet das, daß $X \setminus M$ offen ist. Damit ist alles gezeigt. □

Ist $M \subset X$ eine beliebige Menge, so nennt man die Menge $\overline{M} := M \cup \partial M$ die *abgeschlossene Hülle* von M.

Offensichtlich ist M genau dann abgeschlossen, wenn $\overline{M} = M$ ist.

Beispiel:

$\overline{B_r(x_0)} = \{x \in X \mid d(x, x_0) \leq r\}$ nennt man auch die *abgeschlossene Kugel* mit Radius r um x_0.

Wir wollen nun die Stetigkeit von Abbildungen zwischen metrischen Räumen definieren. Da wir die Konvergenz von Folgen schon eingeführt haben, können wir die Stetigkeit genauso wie bei den reellen Funktionen formulieren.

Definition:

Eine Abbildung $f : X \to Y$ zwischen metrischen Räumen heißt *stetig* in $x_0 \in X$, wenn gilt:

Für **jede** Folge (x_n) in X mit $\lim_{n\to\infty} x_n = x_0$ ist $\lim_{n\to\infty} f(x_n) = f(x_0)$.

Das bedeutet anschaulich:

„Ist x nahe genug bei x_0, so liegt auch $f(x)$ nahe bei $f(x_0)$."

Man kann also zu jeder vorgegebenen Genauigkeit $\varepsilon > 0$ eine so kleine Umgebung $B_\delta(x_0)$ finden, daß $d(f(x), f(x_0)) < \varepsilon$ für alle $x \in B_\delta(x_0)$ ist:

ε-δ-Kriterium für die Stetigkeit

Eine Abbildung $f : X \to Y$ ist genau dann stetig in x_0, wenn folgende Bedingung erfüllt ist:

$$\forall\, \varepsilon > 0\ \exists\, \delta > 0, \text{ so daß gilt: } d(x, x_0) < \delta \implies d(f(x), f(x_0)) < \varepsilon.$$

BEWEIS:

1) Sei f stetig in $\mathbf{x}_0$. Wir nehmen an, das Kriterium sei nicht erfüllt:

$$\exists\, \varepsilon > 0, \text{ so daß } \forall\, \delta > 0 \text{ gilt: } \exists\, \mathbf{x} \text{ mit } d(\mathbf{x}, \mathbf{x}_0) < \delta \text{ und } d(f(\mathbf{x}), f(\mathbf{x}_0)) \geq \varepsilon.$$

Wir wählen dann zu jedem $n \in \mathbb{N}$ ein $\mathbf{x}_n$ mit

$$d(\mathbf{x}_n, \mathbf{x}_0) < \frac{1}{n} \text{ und } d(f(\mathbf{x}_n), f(\mathbf{x}_0)) \geq \varepsilon.$$

Das bedeutet, daß die Folge $(\mathbf{x}_n)$ gegen $\mathbf{x}_0$ konvergiert, aber $(f(\mathbf{x}_n))$ nicht gegen $f(\mathbf{x}_0)$. Das ist ein Widerspruch.

2) Ist umgekehrt das ε-δ-Kriterium erfüllt, so betrachten wir eine beliebige Folge $(\mathbf{x}_n)$, die gegen $\mathbf{x}_0$ konvergiert. Zu vorgegebenem $\varepsilon > 0$ können wir ein $\delta > 0$ gemäß dem Kriterium wählen. Nun gibt es ein n_0, so daß $d(\mathbf{x}_n, \mathbf{x}_0) < \delta$ für $n \geq n_0$ ist. Aber dann ist auch $d(f(\mathbf{x}_n), f(\mathbf{x}_0)) < \varepsilon$ für $n \geq n_0$. Das bedeutet, daß $(f(\mathbf{x}_n))$ gegen $f(\mathbf{x}_0)$ konvergiert. □

Die Stetigkeit der Metrik

Sei $x_0 \in M$ ein fester Punkt. Dann ist die durch $f(x) := d(x, x_0)$ gegebene Abbildung $f : X \to \mathbb{R}$ stetig.

BEWEIS: Wir verwenden folgende Hilfs-Formel:

$$\forall\, x, y \in X \text{ gilt: } |\, d(x, x_0) - d(y, x_0)\,| \leq d(x, y).$$

Der Beweis dieser Formel ergibt sich aus einer zweimaligen Anwendung der Dreiecksungleichung:

$$\begin{array}{rrcl} & d(x, x_0) & \leq & d(x, y) + d(y, x_0) \\ \text{und} & d(y, x_0) & \leq & d(y, x) + d(x, x_0) \;=\; d(x, y) + d(x, x_0), \end{array}$$

$$\text{also} \qquad -\,d(x, y) \;\leq\; d(x, x_0) - d(y, x_0) \;\leq\; d(x, y).$$

Ist nun x_0^* ein beliebiger Punkt von X und x_n eine gegen x_0^* konvergente Folge, so gilt:

$$\begin{aligned} |\, f(x_n) - f(x_0^*)\,| &= |\, d(x_n, x_0) - d(x_0^*, x_0)\,| \\ &\leq d(x_n, x_0^*) \longrightarrow 0, \end{aligned}$$

für $n \to \infty$.

Das heißt, daß f in x_0^* stetig ist. □

Verknüpfung stetiger Funktionen

Ist $f : X \to Y$ stetig in $x_0 \in X$ und $g : Y \to Z$ stetig in $y_0 := f(x_0)$, so ist auch $g \circ f : X \to Z$ stetig in x_0.

Der BEWEIS ist sehr einfach:

Strebt x_n gegen x_0, so strebt $y_n := f(x_n)$ gegen $y_0 = f(x_0)$. Aber dann strebt auch $(g \circ f)(x_n) = g(y_n)$ gegen $(g \circ f)(x_0) = g(y_0)$. □

Schreibt man eine Abbildung $f : \mathbb{R}^n \to \mathbb{R}^k$ in der Form

$$f = (f_1, \dots, f_k),$$

mit Funktionen $f_i : \mathbb{R}^n \to \mathbb{R}$, so ist f genau dann stetig in $\mathbf{x}_0$, wenn alle Funktionen f_i stetig in $\mathbf{x}_0$ sind.

Beispiel:

$$\text{Sei } f(x_1, x_2) := \begin{cases} \dfrac{2x_1x_2}{x_1^2 + x_2^2} & \text{für } (x_1, x_2) \neq (0,0), \\ 0 & \text{für } (x_1, x_2) = (0,0). \end{cases} .$$

Dann ist $t \mapsto f(t, c)$ und $t \mapsto f(c, t)$ für jede Konstante c eine stetige Funktion von t. Aber dennoch ist f nicht stetig im Nullpunkt, denn es ist z.B.

$$f(t,t) \equiv 1 \text{ für } t \neq 0, \quad \text{und } f(0,0) = 0.$$

Aus den Sätzen über Folgen kann man solche über Stetigkeit herleiten. Wir zitieren hier nur die Ergebnisse:

Regeln für die Stetigkeit

$f, g : \mathbb{R}^n \to \mathbb{R}^k$ seien in $\mathbf{x}_0$ stetige Abbildungen.

1. *$f \pm g : \mathbb{R}^n \to \mathbb{R}^k$ mit $(f \pm g)(\mathbf{x}) := f(\mathbf{x}) \pm g(\mathbf{x})$ ist stetig in $\mathbf{x}_0$.*
2. *Für $\alpha \in \mathbb{R}$ ist $\alpha \cdot f$ mit $(\alpha \cdot f)(\mathbf{x}) := \alpha \cdot f(\mathbf{x})$ stetig in $\mathbf{x}_0$.*
3. *$f \bullet g : \mathbb{R}^n \to \mathbb{R}$ mit $(f \bullet g)(\mathbf{x}) := f(\mathbf{x}) \bullet g(\mathbf{x})$ ist stetig in $\mathbf{x}_0$.*
4. *$(f, g) : \mathbb{R}^n \to \mathbb{R}^k \times \mathbb{R}^k$ mit $(f, g)(\mathbf{x}) := (f(\mathbf{x}), g(\mathbf{x}))$ ist stetig in $\mathbf{x}_0$.*
5. *Ist g auch noch stetig in $\mathbf{y}_0$, so ist $f \times g : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^k \times \mathbb{R}^k$ mit $(f \times g)(\mathbf{x}, \mathbf{y}) := (f(\mathbf{x}), g(\mathbf{y}))$ stetig in $(\mathbf{x}_0, \mathbf{y}_0)$.*

Bemerkung: Die Aussagen des Satzes bleiben richtig, wenn man den Definitionsbereich $\mathbb{R}^n$ durch eine beliebige offene Teilmenge M des $\mathbb{R}^n$ ersetzt. Verzichtet man auch noch auf die Offenheit von M, so muß man die Definition der Stetigkeit geringfügig anpassen, es sind dann nur noch Folgen in M zu betrachten.

Lineare Abbildungen sind stetig

Eine lineare Abbildung $f : \mathbb{R}^n \to \mathbb{R}^k$ ist in jedem Punkt $\mathbf{x} \in \mathbb{R}^n$ stetig.

Beweis: Es reicht, die Behauptung für Linearformen zu zeigen.

a) Die Funktionen $(x_1, \dots, x_n) \mapsto x_i$ sind sicherlich stetig:

Ist nämlich $(\mathbf{x}_\nu) = (x_1^{(\nu)}, \ldots, x_n^{(\nu)})$ eine Folge, die gegen $\mathbf{x}_0 = (x_1^{(0)}, \ldots, x_n^{(0)})$ konvergiert, so konvergiert auch $(x_i^{(\nu)})$ gegen $x_i^{(0)}$.

b) Nach den oben angegebenen Regeln über die Stetigkeit ist dann auch

$$(x_1, \ldots, x_n) \mapsto m_1 x_1 + \cdots + m_n x_n$$

eine stetige Funktion. □

Bemerkung: Da wir die komplexen Zahlen geometrisch mit dem Raum $\mathbb{R}^2$ identifizieren können, ist nun auch klar, was eine offene Teilmenge $U \subset \mathbb{C}$ und eine stetige komplexwertige Funktion auf U ist. Insbesondere sind die komplexen Polynome stetig.

Durch Gleichungen und Ungleichungen definierte Mengen

Seien $f, g : \mathbb{R}^n \to \mathbb{R}$ zwei steige Funktionen. Dann gilt:

1. $\{\mathbf{x} \in \mathbb{R}^n \mid f(\mathbf{x}) = g(\mathbf{x})\}$ *ist abgeschlossen.*
2. $\{\mathbf{x} \in \mathbb{R}^n \mid f(\mathbf{x}) < g(\mathbf{x})\}$ *ist offen.*
3. $\{\mathbf{x} \in \mathbb{R}^n \mid f(\mathbf{x}) \le g(\mathbf{x})\}$ *ist abgeschlossen.*

Beweis: Sei $U := \{\mathbf{x} \in \mathbb{R}^n \mid f(\mathbf{x}) \ne g(\mathbf{x})\}$. Wir müssen zeigen, daß U offen ist.

Sei $\mathbf{x}_0 \in U$. Dann ist $r := |\, f(\mathbf{x}) - g(\mathbf{x})\,| > 0$. Ist $0 < \varepsilon < \frac{r}{2}$, so gibt es ein $\delta > 0$ mit

$$|\, f(\mathbf{x}) - f(\mathbf{x}_0)\,| < \varepsilon \text{ und } |\, g(\mathbf{x}) - g(\mathbf{x}_0)\,| < \varepsilon, \text{ für } d(\mathbf{x}, \mathbf{x}_0) < \delta.$$

Annahme, es gibt ein $\mathbf{x} \in B_\delta(\mathbf{x}_0)$ mit $f(\mathbf{x}) = g(\mathbf{x})$. Dann ist

$$\begin{aligned} |\, f(\mathbf{x}_0) - g(\mathbf{x}_0)\,| &\le |\, f(\mathbf{x}_0) - f(\mathbf{x})\,| + |\, f(\mathbf{x}) - g(\mathbf{x})\,| + |\, g(\mathbf{x}) - g(\mathbf{x}_0)\,| \\ &< \varepsilon + 0 + \varepsilon < r. \end{aligned}$$

Aber das ist ein Widerspruch! Also liegt ganz $B_\delta(\mathbf{x}_0)$ in U, U ist offen.

Die Menge, wo $f(\mathbf{x}) = g(\mathbf{x})$ ist, ist demnach abgeschlossen. Genauso wie oben zeigt man, daß die Menge, wo $f(\mathbf{x}) < g(\mathbf{x})$ ist, offen ist. Vertauscht man f und g und geht zum Komplement über, so erhält man schließlich, daß die Menge, wo $f(\mathbf{x}) \le g(\mathbf{x})$ ist, abgeschlossen sein muß. □

Nun sieht man z.B., daß jede Gerade und jede Ebene im $\mathbb{R}^3$ dort eine abgeschlossene Teilmenge ist.

Es sei $M \subset \mathbb{R}^n$ eine Teilmenge und (f_n) eine Folge von (reell– oder komplexwertigen) Funktionen auf M.

Definition:

Die Funktionenfolge (f_n) *konvergiert auf* M *punktweise* gegen eine Funktion f, falls gilt:

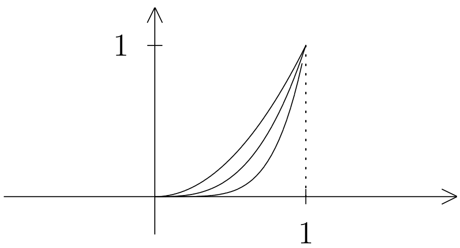
$$\forall\,\mathbf{x} \in M \;:\quad \lim_{n\to\infty} f_n(\mathbf{x}) = f(\mathbf{x})\,.$$

Das Konvergenzverhalten wird also für jeden einzelnen Punkt $\mathbf{x} \in M$ gesondert behandelt. Das globale Verhalten der beteiligten Funktionen spielt dabei keine Rolle.

Beispiel:

Sei $I := [0,1] \subset \mathbb{R}$ und $f_n : I \to \mathbb{R}$ definiert durch $f_n(x) := x^n$. Dann konvergiert diese Funktionenfolge punktweise gegen die Funktion

$$f(x) := \begin{cases} 0 & \text{für } 0 \le x < 1\,, \\ 1 & \text{für } x = 1\,. \end{cases}$$



Obwohl alle Funktionen f_n stetig sind, ist die Grenzfunktion f nicht stetig. Das ist ein wenig wünschenswertes Verhalten.

Wir brauchen also einen besseren Konvergenzbegriff für Funktionenfolgen!

Wechseln wir den Standpunkt! Sei $K \in \{\mathbb{R}, \mathbb{C}\}$. Wir betrachten den Raum

$$\mathcal{B}(M,K) := \{f : M \to K \;:\; \sup\{|\,f(\mathbf{x})\,| \;:\; \mathbf{x} \in M\} < \infty\}.$$

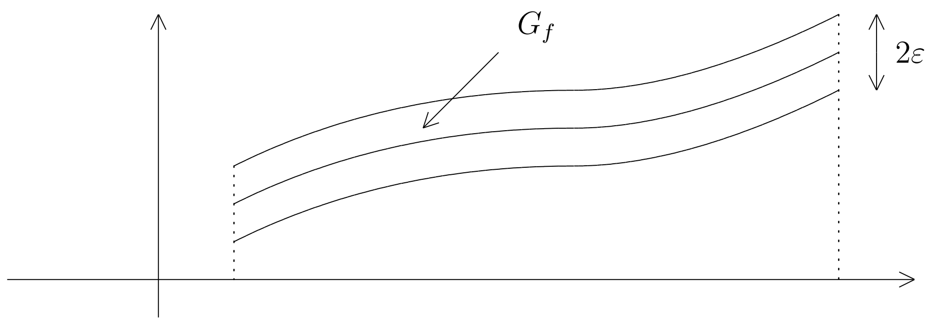
Das ist ein K-Vektorraum mit der Norm $\|\,f\,\| := \sup\limits_M |\,f\,|$, und damit auch ein metrischer Raum (mit $d(f,g) = \|\,f - g\,\|$). Die beschränkten Funktionen auf M fassen wir nun als Punkte eines metrischen Raumes auf.

Eine Folge (f_n) von Elementen von $\mathcal{B}(M,K)$ konvergiert gegen ein $f \in \mathcal{B}(M,K)$, wenn in jeder ε–Umgebung von f fast alle Folgeglieder f_n liegen:

$$\forall\,\varepsilon > 0\;\exists\,n_0 \in \mathbb{N},\text{ so daß } \forall\,n \ge n_0 \text{ gilt: } \|\,f_n - f\,\| < \varepsilon.$$

Wie kann man sich eine ε–Umgebung einer Funktion vorstellen? Wir beschränken uns auf den Fall reeller Funktionen auf einem Intervall I und betrachten die zugehörigen Graphen. Wir definieren einen ε–Schlauch um G_f :

$$S_\varepsilon(f) := \{(x,y) \in \mathbb{R}^2 \mid x \in I \text{ und } f(x) - \varepsilon < y < f(x) + \varepsilon\}.$$



Die ε–Kugel um f besteht nun aus allen Funktionen $g : I \to \mathbb{R}$, deren Graph ganz in $S_\varepsilon(f)$ liegt:

$$\begin{aligned} B_\varepsilon(f) &= \{g : I \to \mathbb{R} \mid \| g - f \| < \varepsilon\} \\ &= \{g : I \to \mathbb{R} \mid \forall x \in I : |g(x) - f(x)| < \varepsilon\} \\ &= \{g : I \to \mathbb{R} \mid \forall x \in I : f(x) - \varepsilon < g(x) < f(x) + \varepsilon\} \\ &= \{g : I \to \mathbb{R} \mid \forall x \in I : (x, g(x)) \in S_\varepsilon(f)\} \\ &= \{g : I \to \mathbb{R} \mid G_g \subset S_\varepsilon(f)\}. \end{aligned}$$

Definition:

Die Folge (f_n) *konvergiert gleichmäßig auf M gegen eine Funktion f*, wenn gilt:

$$\forall \varepsilon > 0 \; \exists n_0 \quad \text{(unabhängig von } \mathbf{x} \in M\text{)},$$

so daß für $n \geq n_0$ und **alle** $\mathbf{x} \in M$ gilt: $\quad | f_n(\mathbf{x}) - f(\mathbf{x}) | < \varepsilon$.

(f_n) konvergiert also genau dann gleichmäßig gegen f, wenn in jedem ε–Schlauch um f fast alle f_n liegen.

Daß wir nun den richtigen Konvergenzbegriff gefunden haben, zeigt der folgende Satz:

Gleichmäßiger Limes stetiger Funktionen

(f_n) konvergiere auf M gleichmäßig gegen f, alle f_n seien stetig. Dann ist auch die Grenzfunktion f auf M stetig.

BEWEIS: Stetigkeit ist eine lokale Eigenschaft, die nur in der Nähe eines (beliebigen) Punktes $\mathbf{x}_0 \in M$ gezeigt werden muß.

Sei $\mathbf{x}_k \in M$ eine Folge von Punkten mit $\lim\limits_{k\to\infty} \mathbf{x}_k = \mathbf{x}_0$. Wir müssen zeigen, daß dann auch $\lim\limits_{k\to\infty} f(\mathbf{x}_k) = f(\mathbf{x}_0)$ ist.

Sei $\varepsilon > 0$. Da (f_n) gleichmäßig auf M gegen f konvergiert, gibt es ein n_0, so daß

$$| f_n(\mathbf{x}) - f(\mathbf{x}) | < \frac{\varepsilon}{3} \text{ für } n \geq n_0 \text{ und alle } \mathbf{x} \in M$$

ist. Wir wählen **ein** solches $n \geq n_0$. Da f_n stetig ist, gibt es ein k_0, so daß

$$| f_n(\mathbf{x}_k) - f_n(\mathbf{x}_0) | < \frac{\varepsilon}{3} \text{ für } k \geq k_0$$

ist. Für solche k gilt dann:

$$\begin{aligned} |\, f(\mathbf{x}_k) - f(\mathbf{x}_0)\,| &\leq |\, f(\mathbf{x}_k) - f_n(\mathbf{x}_k)\,| + |\, f_n(\mathbf{x}_k) - f_n(\mathbf{x}_0)\,| + |\, f_n(\mathbf{x}_0) - f(\mathbf{x}_0)\,| \\ &< \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon. \end{aligned}$$

Das zeigt, daß $(f(\mathbf{x}_k))$ gegen $f(\mathbf{x}_0)$ konvergiert. □

Beispiele :

1. Die Folge $f_n(x) := x^n$ kann auf $[0,1]$ nicht gleichmäßig konvergent sein, weil die Grenzfunktion nicht stetig ist. Aber wie steht es mit der Konvergenz auf $I_r := [0,r]$, mit $0 < r < 1$?

   Die Folge r^n konvergiert gegen Null, und für $0 \leq x \leq r$ ist $0 \leq x^n \leq r^n$. Zu gegebenem $\varepsilon > 0$ gibt es daher ein n_0, so daß für $n \geq n_0$ gilt:

   $$|\, f_n(x) - 0\,| = x^n \leq r^n \leq r^{n_0} < \varepsilon.$$

   Also konvergiert (f_n) auf dem kleineren Intervall $[0,r]$ gleichmäßig gegen die Nullfunktion.

2. Sei $f_n(x) := \frac{1}{n}\sin(nx)$ auf $\mathbb{R}$.

   Ist $\varepsilon > 0$ und $n_0 > \frac{1}{\varepsilon}$, so gilt für $n \geq n_0$:

   $$|\, f_n(x) - 0\,| = \frac{1}{n}|\sin(nx)\,| \leq \frac{1}{n} \leq \frac{1}{n_0} < \varepsilon.$$

   Also konvergiert (f_n) auf ganz $\mathbb{R}$ gleichmäßig.

Die Konvergenz von $f_n \pm g_n$ und $c \cdot f_n$

Wenn (f_n) auf M gleichmäßig gegen f und (g_n) auf M gleichmäßig gegen g konvergiert, dann konvergiert auch $f_n \pm g_n$ gleichmäßig auf M gegen $f \pm g$ und $(c \cdot f_n)$ gleichmäßig auf M gegen $c \cdot f$.

Es scheint schwierig zu sein, Beispiele gleichmäßig konvergenter Folgen von Funktionen zu konstruieren. Dennoch gilt der folgende berühmte Satz:

Satz von Stone/Weierstraß

Ist $f : [a,b] \to \mathbb{R}$ eine stetige Funktion, so gibt es eine Folge von Polynomen, die auf $[a,b]$ gleichmäßig gegen f konvergiert.

Auf den nicht ganz einfachen BEWEIS müssen wir hier verzichten.

§2 Differenzierbare Funktionen

Sei $I \subset \mathbb{R}$ ein Intervall. Wir betrachten „vektorwertige Funktionen" auf I, d.h. Abbildungen $f : I \to \mathbb{R}^n$ für beliebiges $n \in \mathbb{N}$.

Im Falle $n = 1$ liegt eine reelle Funktion vor, die wir uns üblicherweise mit Hilfe ihres Graphen G_f veranschaulichen.

Ist $n = 2$, so nennt man f eine parametrisierte ebene Kurve, die man sich an Hand ihres Bildes $f(I) \subset \mathbb{R}^2$ gut vorstellen kann. Typische Beispiele sind etwa Geraden $t \mapsto \mathbf{x}_0 + t \cdot \mathbf{v}$ oder Kreise $t \mapsto \mathbf{x}_0 + r\cos(t) \cdot \mathbf{e}_1 + r\sin(t) \cdot \mathbf{e}_2$.

Ist $n \geq 3$, so liegt der Graph in einem mindestens 4-dimensionalen Raum und entzieht sich damit endgültig unserer Vorstellung. Aber das Bild $f(I)$ ist zumindest im Falle $n = 3$ noch eine vorstellbare „Raumkurve".

Den Begriff des Grenzwertes von Funktionen können wir mühelos auf vektorwertige Funktionen ausdehnen. Wir schreiben $f : I \to \mathbb{R}^n$ in der Form $f = (f_1, \ldots, f_n)$. Dann ist

$$\lim_{t \to t_0} f(t) = (\lim_{t \to t_0} f_1(t), \ldots, \lim_{t \to t_0} f_n(t)).$$

Nun sind wir gerüstet, um den Begriff der Differenzierbarkeit einzuführen.

Definition:

Eine (vektorwertige) Funktion $f : I \to \mathbb{R}^n$ heißt in $t_0 \in I$ *differenzierbar*, falls

$$f'(t_0) := \lim_{t \to t_0} \frac{f(t) - f(t_0)}{t - t_0}$$

existiert. Der Grenzwert $f'(t_0)$ heißt die *Ableitung* von f in t_0 oder (im Falle $n \geq 2$) der *Tangentialvektor* an f in t_0.

Die Gerade $T_f(t) := f(t_0) + (t - t_0) \cdot f'(t_0)$ nennt man die *Tangente* an f in t_0.

Daß $f : I \to \mathbb{R}^n$ in t_0 differenzierbar ist, kann man auf verschiedene Weise interpretieren:

Zunächst ist der „Differenzenquotient" $\Delta f(t_0, t) := \dfrac{f(t) - f(t_0)}{t - t_0}$ die Richtung der Sekante durch $(t_0, f(t_0))$ und $(t, f(t))$. Läßt man t gegen t_0 gehen, so strebt bei einer in t_0 differenzierbaren Funktion die Richtung der Sekante gegen die Richtung der Tangente. Insbesondere besitzt f in t_0 eine Tangente, d.h. der Graph von f ist dort „glatt".

Andererseits ist jede Sekante eine affin-lineare Approximation von f in der Nähe von t_0. Unter all diesen Approximationen ist die Tangente die beste. Wir wollen sehen, wie gut diese Approximation ist. Dazu definieren wir die lineare Abbildung $L : \mathbb{R} \to \mathbb{R}^n$ durch $L(h) := h \cdot f'(t_0)$ und eine Abbildung r nahe 0 durch

$$r(h) := f(t_0 + h) - f(t_0) - L(h) = f(t_0 + h) - T_f(t_0 + h).$$

Dann gilt:

1. $f(t_0 + h) = f(t_0) + L(h) + r(h)$, falls $t_0 + h \in I$ ist.
2. $\lim\limits_{h \to 0} \dfrac{r(h)}{h} = 0$.

BEWEIS für die zweite Aussage:

$$\frac{r(h)}{h} = \frac{f(t_0 + h) - f(t_0)}{h} - f'(t_0),$$

und dieser Ausdruck strebt offensichtlich gegen Null. $\square$

Damit haben wir ein Kriterium für die Differenzierbarkeit gewonnen. f ist genau dann in t_0 differenzierbar, wenn f so durch eine affin-lineare Abbildung approximiert werden kann, daß die beiden Abbildungen in t_0 übereinstimmen und ihre Differenz mindestens so schnell wie t^2 gegen Null geht.

Eine Abbildung $f = (f_1, \ldots, f_n) : I \to \mathbb{R}^n$ ist genau dann in t_0 differenzierbar, wenn jede der Funktionen f_i in t_0 differenzierbar ist, und es gilt:

$$f'(t_0) = (f_1'(t_0), \ldots, f_n'(t_0)).$$

Wir sagen, f ist auf dem Intervall I differenzierbar, wenn f in jedem $t \in I$ differenzierbar ist. Dann wird durch $t \mapsto f'(t)$ eine neue Funktion $f' : I \to \mathbb{R}^n$ definiert, die *Ableitung* von f.

Wir setzen die Kenntnis der folgenden elementaren Regeln als bekannt voraus:

$$\boxed{\begin{array}{rcll} (x^n)' &=& n \cdot x^{n-1} & \text{für } n \in \mathbb{N},\ x \text{ beliebig}, \\ \sin'(x) &=& \cos(x) & \text{für } x \in \mathbb{R}, \\ \cos'(x) &=& -\sin(x) & \text{für } x \in \mathbb{R}, \\ \exp'(x) &=& \exp(x) & \text{für } x \in \mathbb{R}. \end{array}}$$

Beispiele:

1. Ist $f(t) \equiv \mathbf{c}$ eine konstante Abbildung, so ist $f'(t) \equiv \mathbf{0}$.
2. Ist $f(t) := \mathbf{x}_0 + t \cdot \mathbf{v}$ eine Gerade, so ist

   $$\frac{f(t) - f(t_0)}{t - t_0} = \mathbf{v} \text{ für beliebige Parameter } t,\ t_0,$$

   also $f'(t) \equiv \mathbf{v}$.
3. $f(t) := |\,t\,|$ ist in $t = 0$ nicht differenzierbar, denn

   $$\frac{f(t) - f(0)}{t} = \frac{|\,t\,|}{t} = \begin{cases} 1 & \text{für } t > 0 \\ -1 & \text{für } t < 0 \end{cases}$$

   hat für $t \to 0$ keinen eindeutigen Grenzwert.

Bemerkung: Ist f in t_0 differenzierbar, so ist f in t_0 erst recht stetig. Das folgt sofort aus der Darstellung $f(t) = f(t_0) + L(t - t_0) + r(t - t_0)$.

$(u \circ f)'$ für lineares u

Sei $f : I \to \mathbb{R}^n$ differenzierbar in t_0 und $u : \mathbb{R}^n \to \mathbb{R}^k$ linear. Dann ist auch $u \circ f : I \to \mathbb{R}^k$ in t_0 differenzierbar, und es ist

$$(u \circ f)'(t_0) = u(f'(t_0)).$$

BEWEIS:

$$\frac{u \circ f(t) - u \circ f(t_0)}{t - t_0} = u\left(\frac{f(t) - f(t_0)}{t - t_0}\right),$$

und da u stetig ist, konvergiert der rechte Ausdruck für $t \to t_0$ gegen $u(f'(t_0))$. $\square$

Beispiele:

1. Sind $f, g : I \to \mathbb{R}$ differenzierbar, so auch $f \pm g$, und es ist $(f \pm g)' = f' \pm g'$.

   BEWEIS: $(f, g) : I \to \mathbb{R}^2$ ist differenzierbar, und $u : \mathbb{R}^2 \to \mathbb{R}$ mit $u(x_1, x_2) := x_1 \pm x_2$ ist linear. $\square$

2. Genauso folgt: Mit f ist auch $c \cdot f$ differenzierbar, mit $(c \cdot f)' = c \cdot f'$.

3. Ist $f(t) := (x_0 + r\cos(t), y_0 + r\sin(t))$ die Parametrisierung eines ebenen Kreises, so ist $f'(t) = (-r\sin(t), r\cos(t))$ der zugehörige Tangentialvektor.

Verallgemeinerte Produktregel

Sind $f, g : I \to \mathbb{R}^n$ beide in t_0 differenzierbar, so ist auch $f \bullet g : I \to \mathbb{R}$ in t_0 differenzierbar, und es gilt:

$$(f \bullet g)'(t_0) = f'(t_0) \bullet g(t_0) + f(t_0) \bullet g'(t_0).$$

BEWEIS: Man benutzt einen kleinen Trick:

$$\begin{aligned}
\frac{(f \bullet g)(t) - (f \bullet g)(t_0)}{t - t_0} &= \frac{(f(t) - f(t_0)) \bullet g(t) + f(t_0) \bullet (g(t) - g(t_0))}{t - t_0} \\
&= \frac{f(t) - f(t_0}{t - t_0} \bullet g(t) + f(t_0) \bullet \frac{g(t) - g(t_0)}{t - t_0}.
\end{aligned}$$

strebt für $t \to t_0$ gegen $f'(t_0) \bullet g(t_0) + f(t_0) \bullet g'(t_0)$. $\square$

Hierin ist die gewöhnliche Produktregel $(f \cdot g)' = f' \cdot g + f \cdot g'$ enthalten.

Es gibt eine nette Anwendung der verallgemeinerten Produktregel:

Sei $f : I \to \mathbb{R}^n$ differenzierbar und $\| f(t) \|$ konstant (d.h., die Werte von f befinden sich auf einem Kreis um den Nullpunkt). Dann ist auch $f(t) \bullet f(t) = \| f(t) \|^2$ konstant, und es folgt:

$$0 \equiv (f \bullet f)' = f' \bullet f + f \bullet f' = 2 \cdot f \bullet f'.$$

Also steht $f'(t)$ für jedes t senkrecht auf $f(t)$. Das kann man speziell bei der Parametrisierung des Kreises $f(t) := (r\cos(t), r\sin(t))$ beobachten. Der Tangentialvektor $f'(t)$ steht senkrecht auf dem „Ortsvektor“ $f(t)$.

Aus der Produktregel folgt sogleich auch die bekannte „Quotientenregel“

$$(\frac{f}{g})' = \frac{f'g - fg'}{g^2},$$

die man erhält, indem man $f' = (g \cdot \frac{f}{g})'$ weiter ausrechnet.

Als Anwendung berechnen wir die Ableitung des Tangens:

$$\tan'(t) = \left(\frac{\sin}{\cos}\right)'(t) = \frac{\cos^2(t) - (-\sin^2(t))}{\cos^2(t)} = 1 + \tan^2(t).$$

Kettenregel

Sei $f : I \to \mathbb{R}^n$ differenzierbar in t_0, J ein weiteres Intervall und $g : J \to I$ differenzierbar in $s_0 \in J$, $g(s_0) = t_0$. Dann ist auch $f \circ g : J \to \mathbb{R}^n$ differenzierbar in s_0, und es gilt:

$$(f \circ g)'(s_0) = f'(g(s_0)) \cdot g'(s_0).$$

Den Beweis führen wir nicht aus.

Beispiele:

1. Ist f differenzierbar, so ist auch e^f differenzierbar, und es gilt:

   $$(e^f)'(t) = f'(t) \cdot e^{f(t)}.$$

   Insbesondere ist $a^t = e^{\ln(a)\cdot t}$, also

   $$\boxed{(a^t)' = \ln(a) \cdot a^t, \text{ für } a > 0,\ t \in \mathbb{R}.}$$

2. Sei $f : [0, 2\pi] \to \mathbb{R}^2$ mit $f(t) := (r\cos(t), r\sin(t))$ die Parametrisierung eines Kreises um den Nullpunkt. Setzt man $\varphi(s) := 2s$, so parametrisiert $f \circ \varphi : [0, \pi] \to \mathbb{R}^2$ den gleichen Kreis. Es ist nun $(f \circ \varphi)'(s_0) = 2 \cdot f'(2s_0)$. Das bedeutet, daß der Kreis diesmal mit der doppelten Geschwindigkeit durchlaufen wird.

Definition:

$f : I \to \mathbb{R}^n$ sei in jedem Punkt $t \in I$ differenzierbar. Ist die Ableitung $f' : I \to \mathbb{R}^n$ in $t_0 \in I$ noch ein weiteres Mal differenzierbar, so sagt man, f ist in t_0 zweimal differenzierbar, und man schreibt:

$$f''(t_0) := (f')'(t_0).$$

Induktiv definiert man die n-te Ableitung $f^{(n)}$:

Ist f auf I $(n-1)$-mal differenzierbar und die $(n-1)$-te Ableitung $f^{(n-1)}$ in t_0 noch ein weiteres mal differenzierbar, so sagt man, f ist in t_0 n-mal differenzierbar, und die n-te Ableitung in t_0 wird definiert durch

$$f^{(n)}(t_0) := (f^{(n-1)})'(t_0).$$

Bemerkung : Manchmal benutzt man auch die Leibnizsche Schreibweise:

$$\frac{df}{dt} \text{ statt } f' \quad \text{und} \quad \frac{d^n f}{dt^n} \text{ statt } f^{(n)}.$$

Beispiel :

Sei $f(t) := e^{t^2}$. Dann gilt:

$$\begin{aligned} f'(t) &= 2t \cdot e^{t^2}, \\ f''(t) &= 2 \cdot e^{t^2} + 2t \cdot (2t \cdot e^{t^2}) = (2 + 4t^2) \cdot e^{t^2}, \\ f^{(3)}(t) &= 8t \cdot e^{t^2} + (2 + 4t^2) \cdot (2t \cdot e^{t^2}) = (12t + 8t^3) \cdot e^{t^2}. \end{aligned}$$

Die Versuche lassen folgendes vermuten:

$$f^{(n)}(t) = p(t) \cdot e^{t^2},$$

mit einem Polynom $p(t)$ vom Grad n, das nur gerade bzw nur ungerade Potenzen von t enthält, je nachdem, ob n gerade oder ungerade ist. Für kleine n haben wir das verifiziert. Also können wir einen Induktionsbeweis führen. Ist die Formel für $n \geq 1$ richtig, so gilt:

$$f^{(n+1)}(t) = (p'(t) + 2t \cdot p(t)) \cdot e^{t^2}.$$

Ist etwa $n = 2k$, so enthält $p(t)$ nur gerade Potenzen von t. Aber dann ist $p'(t)$ ein Polynom vom Grad $n - 1$ und $2t \cdot p(t)$ ein Polynom vom Grad $n + 1$, und beide enthalten nur ungerade Potenzen von t. Ähnlich funktioniert es im Falle $n = 2k + 1$.

Also gilt die Formel auch für $n + 1$ und damit für alle n.

Im Rest des Paragraphen betrachten wir nur noch reellwertige Funktionen.

Ableitung der Umkehrfunktion

Ist $f : I \to J \subset \mathbb{R}$ eine bijektive differenzierbare Abbildung und $f'(x_0) \neq 0$, so ist $f^{-1} : J \to I$ in $f(x_0)$ differenzierbar, und es gilt:

$$(f^{-1})'(f(x_0)) = \frac{1}{f'(x_0)}.$$

Auch hier lassen wir den Beweis weg.

Beispiele:

1. $\exp : \mathbb{R} \to \mathbb{R}_+$ ist bijektiv und differenzierbar, und $\exp'(x) = \exp(x)$ ist stets > 0. Also ist auch die Umkehrfunktion $\ln : \mathbb{R}_+ \to \mathbb{R}$ überall differenzierbar, mit $(\ln)'(e^x) = \dfrac{1}{e^x}$, also

$$\boxed{\ln'(x) = \frac{1}{x}.}$$

2. Die Funktion x^a ist für festes $a \neq 0$ und $x > 0$ definiert durch $x^a := e^{a\ln(x)}$. Dann ist $(x^a)' = a \cdot x^{-1} \cdot e^{a\ln(x)}$, also

$$\boxed{(x^a)' = a \cdot x^{a-1}, \text{ für } x > 0 \text{ und } a \neq 0.}$$

Die Funktion $x^x = e^{x\ln(x)}$ ergibt nun beim Differenzieren weder $x \cdot x^{x-1} = x^x$ noch $\ln(x) \cdot x^x$, sondern vielmehr gilt:

$$(x^x)' = (\ln x + 1) \cdot x^x.$$

3. Sei $I \subset \mathbb{R}$ ein Intervall und $f : I \to \mathbb{R}_+$ eine differenzierbare Funktion. Dann ist auch $g := \ln \circ f : I \to \mathbb{R}$ differenzierbar, und es gilt:

$$g'(x) = (\ln \circ f)'(x) = \ln'(f(x)) \cdot f'(x) = \frac{f'(x)}{f(x)}.$$

Man nennt das auch die „logarithmische Ableitung“ von f.

4. Die Funktion $\tan : (-\frac{\pi}{2}, +\frac{\pi}{2}) \to \mathbb{R}$ ist differenzierbar und bijektiv, und $\tan'(x) = 1 + \tan^2(x)$ hat dort keine Nullstelle. Also ist auch die Umkehrfunktion arctan auf $\mathbb{R}$ differenzierbar, und es gilt:

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

Dieses Ergebnis sollte man sich für später merken: Durch Differenzieren der Umkehrfunktion der Winkelfunktion Tangens erhält man eine rationale Funktion.

Definition:

$f : I \to \mathbb{R}$ hat in $x_0 \in I$ ein *lokales Maximum* (bzw. *lokales Minimum*), falls gilt:
$\exists\, \varepsilon > 0$, so daß $f(x) \leq f(x_0)$ (bzw. $f(x) \geq f(x_0)$) für $x \in I$ und $|\, x - x_0\,| < \varepsilon$ ist.
In beiden Fällen sagt man, f hat in x_0 einen (lokalen) Extremwert.

Man beachte: Ist f in der Nähe von x_0 konstant, so hat f dort nach unserer Definition auch einen Extremwert! Wir führen deshalb noch einen zusätzlichen Begriff ein:

Definition:

$f : I \to \mathbb{R}$ hat in $x_0 \in I$ ein *isoliertes Maximum* (bzw. ein *isoliertes Minimum*), falls gilt:

$$\exists\, \varepsilon > 0, \text{ so daß } f(x) < f(x_0) \text{ für } |\, x - x_0\,| < \varepsilon \text{ und } x \neq x_0 \text{ ist}$$

$$(\text{bzw. } f(x) > f(x_0) \text{ im Falle des Minimums}).$$

„Notwendiges Kriterium“ für Extremwerte

Sei I ein Intervall, x_0 ein innerer Punkt von I und $f : I \to \mathbb{R}$ in x_0 differenzierbar. Wenn f in x_0 ein lokales Extremum besitzt, dann ist $f'(x_0) = 0$.

BEWEIS: Wir betrachten den Differenzenquotienten $\Delta f(x_0, x) := \dfrac{f(x) - f(x_0)}{x - x_0}$, und behandeln nur den Fall des lokalen Maximums, beim Minimum geht es analog.

Hat f in x_0 ein lokales Maximum, so ist $f(x) \leq f(x_0)$ für x nahe bei x_0. Ist $x < x_0$, so ist $x - x_0 < 0$ und daher $\Delta f(x_0, x) \geq 0$. Ist jedoch $x > x_0$, so ist $\Delta f(x_0, x) \leq 0$. Aber dann muß $f'(x_0) = \lim_{x \to x_0} \Delta f(x_0, x) = 0$ sein. $\square$

Hinreichende Kriterien behandeln wir später.

Beispiele :

1. Sei $I = [-1, 1]$, $f : I \to \mathbb{R}$ definiert durch $f(x) := x^2$. Dann gilt für alle $x \in I$: $f(x) \geq 0 = f(0)$. Also hat f in $x_0 := 0$ ein (sogar isoliertes) lokales Minimum. Und tatsächlich besitzt $f'(x) = 2x$ in x_0 eine Nullstelle.

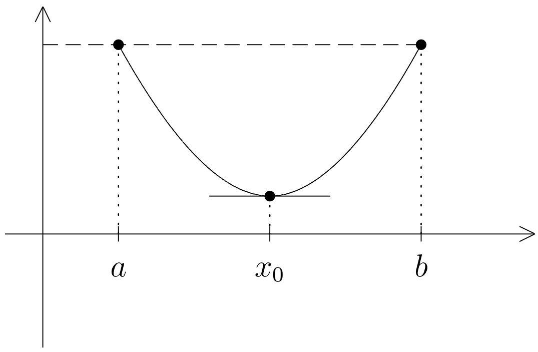
2. $g : I \to \mathbb{R}$ mit $g(x) := |\, x\,|$ hat ebenfalls in $x_0 = 0$ ein lokales Minimum. Aber weil $|\, x\,|$ dort nicht differenzierbar ist, kann man das Kriterium nicht anwenden.

   Für $x \in I$ ist $|\, x\,| \leq 1 = g(-1) = g(1)$. Also hat g in den Punkten $x = -1$ und $x = +1$ jeweils ein lokales Maximum. Aber auch hier kann man das notwendige Kriterium nicht anwenden, denn die Punkte liegen nicht im Innern von I.

§3 Mittelwertsatz und Taylorsche Formel

Der Satz von Rolle

Sei $f : I := [a, b] \to \mathbb{R}$ stetig und im Innern von I differenzierbar. Ist $f(a) = f(b)$, so gibt es einen inneren Punkt x_0 von I mit $f'(x_0) = 0$.



Beweis: Sei $c := f(a) = f(b)$. Ist $f(x) \equiv c$ auf ganz I, so ist auch $f'(x) \equiv 0$.

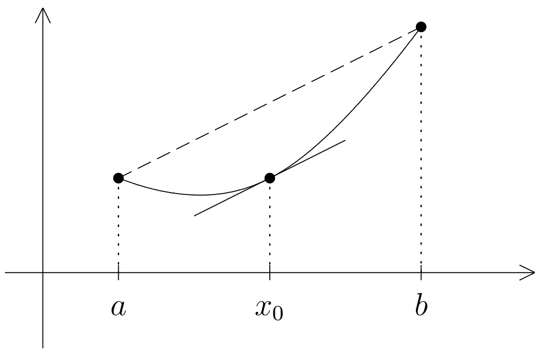
Ist f auf I nicht konstant, so muß entweder das Minimum oder das Maximum von f im Innern von I liegen. Und dort muß dann f' verschwinden. □

Der folgende wichtige Satz ist eine einfache Folgerung:

Der 1. Mittelwertsatz der Differentialrechnung

Sei $f : I := [a, b] \to \mathbb{R}$ stetig und im Innern von I differenzierbar. Dann gibt es einen Punkt x_0 im Innern von I mit

$$f'(x_0) = \frac{f(b) - f(a)}{b - a}.$$



Beweis: Sei $L : \mathbb{R} \to \mathbb{R}$ die eindeutig bestimmte affin-lineare Funktion mit $L(a) = f(a)$ und $L(b) = f(b)$ (also die Sekante durch $(a, f(a))$ und $(b, f(b))$). Dann ist

$$L(x) = L(0) + mx, \text{ mit } m := \frac{f(b) - f(a)}{b - a}.$$

Setzen wir $g := f - L$ auf I, so ist $g(a) = g(b) = 0$ und $g'(x) = f'(x) - m$. Nach dem Satz von Rolle, angewandt auf die Funktion g, existiert ein Punkt x_0 im Innern von I mit $g'(x_0) = 0$, also $f'(x_0) = m$. □

So einfach der Beweis, so mächtig die Konsequenzen:

Funktionen mit verschwindender Ableitung

Sei $f : I \to \mathbb{R}$ stetig und im Inneren von I differenzierbar.

Ist $f'(x) \equiv 0$, so ist f konstant.

Beweis: Sei $I = [a, b]$, $a \le x_1 < x_2 \le b$. Nach dem Mittelwertsatz existiert ein x_0 mit $x_1 < x_0 < x_2$ und

$$0 = f'(x_0) = \frac{f(x_2) - f(x_1)}{x_2 - x_1}.$$

Das ist nur möglich, wenn $f(x_1) = f(x_2)$ ist. Und da die Punkte x_1 und x_2 beliebig gewählt werden können, ist f konstant. □

Ableitung und Monotonie

Sei $f : I := [a, b] \to \mathbb{R}$ stetig und im Inneren von I differenzierbar. f ist genau dann auf I monoton wachsend (bzw. fallend), wenn $f'(x) \ge 0$ (bzw. $f'(x) \le 0$) für alle $x \in (a, b)$ ist.

Ist sogar $f'(x) > 0$ für alle $x \in I$ (bzw. $f'(x) < 0$ für alle $x \in I$), so ist f streng monoton wachsend (bzw. fallend).

Beweis: Wir beschränken uns auf den Fall der wachsenden Funktion.

1) Ist f monoton wachsend, so sind alle Differenzenquotienten ≥ 0, und daher ist auch überall $f'(x) \ge 0$.

2) Nun sei $f'(x) \ge 0$ für alle $x \in (a, b)$. Ist $x_1 < x_2$, so gibt es nach dem Mittelwertsatz ein x_0 mit $x_1 < x_0 < x_2$, so daß gilt:

$$0 \le f'(x_0) = \frac{f(x_2) - f(x_1)}{x_2 - x_1}\,, \text{ also } f(x_1) \le f(x_2).$$

3) Ist f monoton wachsend, aber nicht streng monoton wachsend, so muß f auf einem Teilintervall positiver Länge konstant sein, und dort verschwindet die Ableitung von f. Ist also f' überall positiv, so muß f streng monoton wachsend sein. □

Beispiele:

1. Sei $f(x) := x^3$. Dann ist $f'(x) = 3x^2$ und $f''(x) = 6x$. Da überall $f'(x) \ge 0$ ist, wächst f auf ganz $\mathbb{R}$ monoton. Außerhalb des Nullpunktes ist $f'(x)$ sogar positiv, also wächst f dort streng monoton. Und für kleines positives h ist $f(-h) = -h^3 <$

$0 < h^3 = f(h)$. Daraus folgt, daß f sogar überall streng monoton steigt. Trotzdem ist $f'(0) = 0$.

2. Die Funktionen $\sinh(x) = \dfrac{1}{2}(e^x - e^{-x})$ und $\cosh(x) = \dfrac{1}{2}(e^x + e^{-x})$ sind offensichtlich überall differenzierbar, und es gilt:

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

Da $\sinh'(x) > 0$ für alle x ist, ist sinh streng monoton wachsend und somit umkehrbar. Die Umkehrfunktion wird mit arsinh (*Area-Sinus hyperbolicus*) bezeichnet.

Die Beziehung $y = \sinh(x) = \dfrac{1}{2}(e^x - e^{-x})$ liefert eine quadratische Gleichung für e^x,

$$2y = \frac{(e^x)^2 - 1}{e^x}, \quad \text{also} \quad (e^x)^2 - 2y \cdot e^x - 1 = 0,$$

und damit

$$\operatorname{arsinh}(y) = x = \ln(y + \sqrt{1 + y^2}).$$

Das positive Vorzeichen vor der Wurzel muß gewählt werden, weil $e^x > 0$ ist.

Daraus folgt:

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

Der Cosinus hyperbolicus läßt sich nur für $x \geq 0$ oder für $x \leq 0$ umkehren. Die Umkehrfunktion arcosh (*Area-Cosinus hyperbolicus*) ist jeweils für $y \geq 1$ erklärt. Sie ist gegeben durch

$$\operatorname{arcosh}(y) = \ln(y \pm \sqrt{y^2 - 1}).$$

Schließlich kann man den Mittelwertsatz noch weiter verallgemeinern:

Der 2. Mittelwertsatz der Differentialrechnung

Es seien f und g auf $I := [a, b]$ stetig und im Innern von I differenzierbar. Außerdem sei $g'(x) \neq 0$ im Innern von I.

Dann gibt es einen Punkt c im Innern von I mit

$$\frac{f'(c)}{g'(c)} = \frac{f(b) - f(a)}{g(b) - g(a)}.$$

Für $g(x) = x$ erhält man den 1. Mittelwertsatz zurück.

BEWEIS: Wäre $g(b) - g(a) = 0$, so wäre $g'(x_0) = 0$ für ein x_0 im Innern von I. Das hatten wir aber gerade ausgeschlossen.

Wir benutzen die Hilfsfunktion

$$F(x) := f(x) - \frac{f(b) - f(a)}{g(b) - g(a)} \cdot (g(x) - g(a)).$$

Es ist $F(a) = f(a) = F(b)$. Nach dem Satz von Rolle gibt es ein c im Innern des Intervalls, so daß $F'(c) = 0$ ist. Aber offensichtlich ist

$$F'(c) = f'(c) - \frac{f(b) - f(a)}{g(b) - g(a)} \cdot g'(c).$$

Daraus folgt die gewünschte Gleichung. □

Als Anwendung ergibt sich die

1. Regel von de l'Hospital (der Grenzwert $\frac{0}{0}$)

Die Funktionen f und g seien auf dem offenen Intervall $I := (a, b)$ differenzierbar, und es sei $g'(x) \neq 0$ für $x \in I$.

Außerdem sei $c \in I$ und $f(c) = g(c) = 0$.

Wenn $\displaystyle\lim_{x \to c} \frac{f'(x)}{g'(x)}$ *existiert, dann existiert auch* $\displaystyle\lim_{x \to c} \frac{f(x)}{g(x)}$, *und die beiden Grenzwerte sind gleich.*

Beweis: Wir arbeiten mit einseitigen Grenzwerten. Der Satz gilt dann dementsprechend auch etwas allgemeiner.

Es sei (x_ν) eine Folge von Zahlen mit $c < x_\nu < b$ und $\lim_{\nu \to \infty} x_\nu = c$ (nicht notwendig monoton). Nach dem 2. Mittelwertsatz gibt es Zahlen c_ν mit $c < c_\nu < x_\nu$ und

$$\frac{f(x_\nu)}{g(x_\nu)} = \frac{f(x_\nu) - f(c)}{g(x_\nu) - g(c)} = \frac{f'(c_\nu)}{g'(c_\nu)}.$$

Da auch $\lim_{\nu \to \infty} c_\nu = c$ ist, strebt der letzte Quotient nach Voraussetzung gegen $\frac{f'(c)}{g'(c)}$. Aber das bedeutet, daß

$$\lim_{x \to c+} \frac{f(x)}{g(x)} = \frac{f'(c)}{g'(c)}$$

ist, und analog schließt man für den linksseitigen Grenzwert. □

An Stelle der Annäherung an eine endliche Zahl c kann man auch den Fall $x \to \pm\infty$ betrachten, es gelten analoge Aussagen.

Beispiele:

1. Sei $f(x) := \sin x$ und $g(x) := x$.

   Da $f(0) = g(0) = 0$ ist und $\dfrac{f'(x)}{g'(x)} = \dfrac{\cos x}{1}$ für $x \to 0$ gegen 1 strebt, ist auch

   $$\lim_{x \to 0} \frac{\sin(x)}{x} = \lim_{x \to 0} \frac{\cos(x)}{1} = 1.$$

2. Sei $f(x) := \ln(1-x)$ und $g(x) := x + \cos x$. Dann gilt:
$$\frac{f'(x)}{g'(x)} = \frac{1}{(x-1)(1-\sin x)} \to -1, \text{ für } x \to 0.$$
Aber man darf l'Hospital gar nicht anwenden, denn es ist zwar $f(0) = 0$, aber $g(0) = 1$.

   Tatsächlich ist $\lim\limits_{x\to 0} \dfrac{\ln(1-x)}{x+\cos x} = 0$.

2. Regel von de l'Hospital (der Grenzwert $\frac{\infty}{\infty}$)

Die Funktionen f und g seien auf dem offenen Intervall $I := (a, b)$ differenzierbar, und es sei $g'(x) \neq 0$ für $x \in I$.

Es sei $\lim\limits_{x\to a+} f(x) = \lim\limits_{x\to a+} g(x) = +\infty$.

Wenn $\lim\limits_{x\to a+} \dfrac{f'(x)}{g'(x)}$ existiert, dann existiert auch $\lim\limits_{x\to a+} \dfrac{f(x)}{g(x)}$, und die beiden Grenzwerte sind gleich.

Die gleiche Aussage gilt für die Annäherung an a von links.

Der Beweis benutzt ebenfalls den verallgemeinerten Mittelwertsatz, ist aber etwas komplizierter. Ich lasse ihn hier weg.

Beispiele:

1. Es ist $$\lim_{x\to 0+} x\cdot\ln(x) = \lim_{x\to 0+} \frac{\ln(x)}{x^{-1}} = \lim_{x\to 0+} \frac{x^{-1}}{-x^{-2}} = \lim_{x\to 0+}(-x) = 0.$$

2. Sei $p(x)$ ein Polynom. Mehrfache Anwendung von l'Hospital ergibt:
$$\lim_{x\to\infty} \frac{e^x}{p(x)} = \lim_{x\to\infty} \frac{e^x}{p'(x)} = \ldots = \lim_{x\to\infty} \frac{e^x}{\text{Konstante}} = +\infty.$$
Die Exponentialfunktion wächst stärker als jedes Polynom.

3. Dagegen gilt:
$$\lim_{x\to\infty} \frac{\ln(x)}{p(x)} = \lim_{x\to\infty} \frac{1}{x\cdot p'(x)} = 0.$$
Die Logarithmusfunktion wächst also schwächer als jedes Polynom.

Definition:

Ist $f : I \to \mathbb{R}$ überall k-mal differenzierbar und $f^{(k)}$ auf I noch stetig, so nennt man f auf I *k-mal stetig differenzierbar*.

Die Menge der k-mal stetig differenzierbaren Funktionen auf I bezeichnet man mit $\mathcal{C}^k(I)$.

Bemerkung: $\mathcal{C}^k(I)$ ist ein $\mathbb{R}$-Vektorraum, und $D : \mathcal{C}^k(I) \to \mathcal{C}^{k-1}(I)$ mit $D(f) := f'$ ist eine lineare Abbildung. $\mathrm{Ker}\,(D)$ besteht aus den konstanten Funktionen.

Sei $I \subset \mathbb{R}$ ein Intervall, $f : I \to \mathbb{R}$ differenzierbar, $a \in I$ ein fest gewählter Punkt. Wir wollen versuchen, f in der Nähe von a möglichst gut durch ein Polynom zu approximieren.

Dazu untersuchen wir erst einmal ein Polynom selbst. Da

$$x^n = ((x-a)+a)^n = \sum_{k=0}^{n} \binom{n}{k} a^{n-k}(x-a)^k$$

ist, können wir annehmen, daß wir ein Polynom $p(x)$ vom Grad n in der Form

$$p(x) = \sum_{k=0}^{n} c_k (x-a)^k$$

gegeben haben.

Es ist $((x-a)^k)^{(i)} = \begin{cases} k(k-1)\cdot \ldots \cdot (k-i+1)\cdot (x-a)^{k-i} & \text{für } i<k, \\ k! & \text{für } i=k, \\ 0 & \text{sonst.} \end{cases}$

Daraus folgt, daß $p^{(i)}(a) = i! \cdot b_i$ ist, also

$$p(x) = \sum_{k=0}^{n} \frac{p^{(k)}(a)}{k!}(x-a)^k.$$

Wir nennen das auch die „Entwicklung von p im Punkte a“.

Wir suchen jetzt zu einer Funktion $f \in \mathcal{C}^n(I)$ ein Polynom p mit $f^{(k)}(a) = p^{(k)}(a)$ für $k = 0, 1, \ldots, n$.

Definition:

Sei $f \in \mathcal{C}^n(I)$, $a \in I$ ein fester Punkt. Das Polynom

$$T_n f(x) = T_n f(x; a) := \sum_{k=0}^{n} \frac{f^{(k)}(a)}{k!}(x-a)^k$$

heißt *n–tes Taylorpolynom von f in a*.

Offensichtlich ist

$$T_1 f(x; a) = f(a) + f'(a)(x-a) \text{ die Tangente an } f \text{ in } a \text{, und}$$

$$(T_n f)^{(k)}(a) = f^{(k)}(a), \text{ für } k = 0, 1, 2, \ldots, n.$$

Nun geht es um das Verhalten des *Restgliedes* $R_n(x) := f(x) - T_n f(x)$ in der Nähe von a:

Satz von der Taylorentwicklung

Es sei $f \in \mathcal{C}^n(I)$ und $R_n(x) := f(x) - T_n f(x)$.

1. *Es gibt eine Funktion $\eta(x)$ mit $\lim_{x \to a} \eta(x) = 0$, so daß gilt:*

$$R_n(x) = \eta(x) \cdot (x-a)^n.$$

2. *Ist sogar $f \in \mathcal{C}^{n+1}(I)$, so gibt es zu jedem $x \neq a$ ein $c = c(x)$ zwischen a und x, so daß gilt:*

$$R_n f(x) = \frac{f^{(n+1)}(c)}{(n+1)!} \cdot (x-a)^{n+1}.$$

Man spricht dann auch von der „Lagrangeschen Form“ des Restgliedes.

Die Darstellung $f = T_n f + R_n$ nennt man die „Taylorentwicklung“ der Ordnung n von f im Punkte a.

Beweis: Wir beginnen mit dem zweiten Teil und betrachten

$$\varphi(x) := \frac{R_n(x)}{(x-a)^{n+1}}$$

für $x \neq a$. Die $(n+1)$–te Ableitung von $(x-a)^{n+1}$ ergibt $(n+1)!$. Deshalb führt eine $(n+1)$–fache Anwendung des 2. Mittelwertsatzes zu

$$\begin{aligned} \varphi(x) &= \frac{R_n'(c_1)}{(n+1)(c_1-a)^n} \\ &= \frac{R_n''(c_2)}{n(n+1)(c_2-a)^{n-1}} \\ &= \cdots \\ &= \frac{R_n^{(n+1)}(c)}{(n+1)!} = \frac{f^{(n+1)}(c)}{(n+1)!}, \end{aligned}$$

mit geeigneten Punkten $c_i = c_i(x)$ zwischen a und x, sowie $c := c_{n+1}$, also

$$R_n(x) = \frac{f^{(n+1)}(c)}{(n+1)!}(x-a)^{n+1}.$$

Ist f nur n–mal stetig differenzierbar, so erhält man auf die gleiche Weise die Beziehung

$$f(x) - T_{n-1}f(x) = \frac{f^{(n)}(c)}{n!}(x-a)^n,$$

mit einem geeigneten c zwischen a und x. Wir setzen

$$\eta(x) := \frac{1}{n!} \cdot (f^{(n)}(c) - f^{(n)}(a)).$$

Da c stets zwischen a und x liegt und $f^{(n)}$ stetig ist, ist $\lim_{x \to a} \eta(x) = 0$. Außerdem gilt:

$$\frac{f^{(n)}(c)}{n!}(x-a)^n = \frac{f^{(n)}(a)}{n!}(x-a)^n + \eta(x) \cdot (x-a)^n.$$

Das ergibt die erste Behauptung. □

Bemerkung: Die Beziehung $\lim_{x\to a} \dfrac{f(x)}{g(x)} = 0$ wird gerne mit Hilfe des *Landauschen Symbols* o abgekürzt:

$$f(x) = \mathrm{o}(g(x)) \qquad (\text{für } x \to a\,).$$

Zum Beispiel ist

$$(1+x)^n = 1 + nx + \frac{n(n-1)}{2}x^2 + \ldots = 1 + nx + \mathrm{o}(x),\ (\text{für } x \to 0)$$

oder

$$x \cdot \sin(\frac{1}{x}) = \mathrm{o}(1).$$

Im Falle der Taylorentwicklung kann man nun schreiben: Ist $f \in \mathcal{C}^n(I)$, so ist

$$f(x) = T_n f(x) + \mathrm{o}((x-a)^n),\ (\text{für } x \to a).$$

Wir wollen einige **Beispiele** zur Taylorentwicklung betrachten:

1. Sei $a = 0$ und $f(x) = \sin(x)$. Es ist

   $$\sin'(x) = \cos(x),\ \sin''(x) = -\sin(x),\ \sin^{(3)}(x) = -\cos(x),\ \sin^{(4)}(x) = \sin(x),$$

   und dann wiederholt sich das wieder. Daraus folgt:

   $$\begin{aligned} \sin^{(2n)}(0) &= 0 \\ \text{und } \sin^{(2n+1)}(0) &= (-1)^n. \end{aligned}$$

   Das ergibt für die Entwicklung der Ordnung $2n+2$:

   $$\begin{aligned} \sin(x) &= \sum_{k=0}^{n} \frac{(-1)^k}{(2k+1)!} \cdot x^{2k+1} + \mathrm{o}(x^{2n+2}) \\ &= x - \frac{1}{6}x^3 + \frac{1}{120}x^5 \pm \cdots + \frac{(-1)^n}{(2n+1)!}x^{2n+1} + \mathrm{o}(x^{2n+2}). \end{aligned}$$

2. Analog geht es beim Cosinus. Es ist

   $$\begin{aligned} \cos^{(2n+1)}(0) &= 0 \\ \text{und} \quad \cos^{(2n)}(0) &= (-1)^n. \end{aligned}$$

   Daraus folgt:

   $$\begin{aligned} \cos(x) &= \sum_{k=0}^{n} \frac{(-1)^k}{(2k)!} \cdot x^{2k} + \mathrm{o}(x^{2n+1}) \\ &= 1 - \frac{1}{2}x^2 + \frac{1}{24}x^4 - \frac{1}{720}x^6 \pm \cdots + \frac{(-1)^n}{(2n)!}x^{2n} + \mathrm{o}(x^{2n+1}). \end{aligned}$$

   Wir wollen diese Entwicklung benutzen, um den Cosinus numerisch zu berechnen. Zunächst sind einige Reduktionsschritte nötig.

1) Da $\cos(-x) = \cos(x)$ ist, kann man sich auf den Fall $x \geq 0$ beschränken.

2) Wegen der Periodizität interessiert nur $0 \leq x < 2\pi$.

3) Wegen $\cos(\pi + x) = \cos(\pi - x)$ kann man sich auf $0 \leq x \leq \pi$ beschränken.

4) Zwischen 0 und π ist $\cos(\frac{\pi}{2} + x) = -\cos(\frac{\pi}{2} - x)$. Also braucht man sogar nur den Bereich $0 \leq x \leq \frac{\pi}{2}$ zu betrachten.

5) Für $0 \leq x \leq \frac{\pi}{2} \approx 1.570795$ verwenden wir die Taylorentwicklung

$$\cos(x) = 1 - \frac{x^2}{2} + \frac{x^4}{24} - \frac{x^6}{720} + R_7(x),$$

wobei

$$R_7(x) = \cos(x) - T_7\cos(x) = \frac{\cos^{(8)}(c)}{8!} \cdot x^8 = \frac{\cos(c)}{40320}x^8 \text{ ist, mit } 0 \leq c \leq \frac{\pi}{2},$$

also

$$|\, R_8(x)\,| \leq \frac{1}{40320} \cdot (1.570795\ldots)^8 \approx 0.001.$$

Damit ist

$$\cos(x) \approx 1 - \frac{x^2}{2}\left(1 - \frac{x^2}{12}\left(1 - \frac{x^2}{30}\right)\right), \text{ für } 0 \leq x \leq \frac{\pi}{2}.$$

Setzt man $x = \frac{\pi}{4} = 0.7853981\ldots$ ein, so erhält man $\cos(x) \approx 0.7071$, was recht gut dem tatsächlichen Wert $\frac{1}{2}\sqrt{2}$ entspricht.

3. Sei wieder $a = 0$ und $f(x) = e^x$. Da $(e^x)' = e^x$ und $e^0 = 1$ ist, folgt:

$$\begin{aligned} e^x &= \sum_{k=0}^{n} \frac{1}{k!}x^k + \mathrm{o}(x^n) \\ &= 1 + x + \frac{1}{2}x^2 + \frac{1}{6}x^3 + \cdots + \frac{1}{n!}x^n + \mathrm{o}(x^n). \end{aligned}$$

4. Die Funktion $f(x) = \ln(x)$ ist nur für $x > 0$ definiert. Hier nehmen wir $a = 1$ als Entwicklungspunkt. Es ist

$$\ln(1) = 0,\ \ln'(x) = \frac{1}{x},\ \ln''(x) = -\frac{1}{x^2},\ \ln^{(3)}(x) = \frac{2}{x^3},\ \ln^{(4)}(x) = -\frac{6}{x^4},$$

und allgemein

$$\ln^{(k)}(x) = (-1)^{k+1} \cdot \frac{(k-1)!}{x^k}, \text{ für } k \geq 1.$$

Das bedeutet, daß

$$\frac{\ln^{(k)}(1)}{k!} = (-1)^{k+1}\frac{(k-1)!}{k!} = \frac{(-1)^{k+1}}{k} \text{ für } k \geq 1$$

ist, also

$$\begin{aligned} \ln(x) &= \sum_{k=1}^{n} \frac{(-1)^{k+1}}{k}(x-1)^k + \mathrm{o}((x-a)^n) \\ &= (x-1) - \frac{1}{2}(x-1)^2 + \frac{1}{3}(x-1)^3 \pm \cdots + \frac{(-1)^{n+1}}{n}(x-1)^n + \mathrm{o}((x-1)^n). \end{aligned}$$

Diese Entwicklung kann natürlich nur für $x > 0$ gelten, und das asymptotische Verhalten bezieht sich sowieso stets nur auf die Annäherung an den Entwicklungspunkt.

Als Anwendung der Taylorformel können wir jetzt das Problem der lokalen Extrema erledigen:

Hinreichendes Kriterium für Extremwerte

Die Funktion f sei in der Nähe von x_0 n–mal stetig differenzierbar. Es sei

$$\begin{aligned} & f^{(k)}(x_0) = 0 \quad \textit{für } k = 1, \ldots, n-1 \\ \textit{und} \quad & f^{(n)}(x_0) \neq 0. \end{aligned}$$

Ist n ungerade, so besitzt f in x_0 kein lokales Extremum.

Ist n gerade, so liegt ein lokales Extremum in x_0 vor, und zwar

$$\begin{aligned} \textit{ein Maximum, falls } f^{(n)}(x_0) &< 0 \quad \textit{ist,} \\ \textit{und ein Minimum, falls } f^{(n)}(x_0) &> 0 \quad \textit{ist.} \end{aligned}$$

Beweis: Wir verwenden die Lagrangesche Form des Restgliedes bei der Taylorentwicklung. Da $f'(x_0) = f''(x_0) = \ldots = f^{(n-1)}(x_0)$ ist, folgt mit $h := x - x_0$:

$$f(x) = f(x_0 + h) = f(x_0) + \frac{f^{(n)}(c)}{n!}h^n,$$

mit einem geeigneten c zwischen x_0 und x.

Ist $\varepsilon > 0$ klein genug gewählt, so ist $f^{(n)}(x) \neq 0$ für $|\,x - x_0\,| < \varepsilon$, und dann hat $f^{(n)}(c)$ das gleiche Vorzeichen wie $f^{(n)}(x_0)$.

Wir betrachten nur den Fall $f^{(n)}(x_0) > 0$, der andere geht analog. Da c von x (und damit von h) abhängt, können wir schreiben:

$$f(x_0 + h) - f(x_0) = \varphi(h) \cdot h^n,$$

mit einer positiven Funktion φ.

Ist n ungerade, so wechselt h^n bei $h = 0$ sein Vorzeichen, und es kann kein Extremwert vorliegen. Ist n gerade, so bleibt h^n immer ≥ 0 und verschwindet bei $h = 0$. Dann besitzt f in x_0 ein Minimum. □

Gelegentlich interessiert man sich für die folgende Situation:

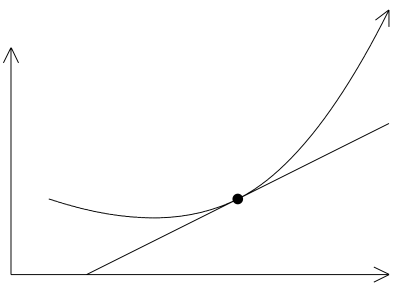
$$f'(x_0) = f''(x_0) = 0 \text{ und } f'''(x_0) \neq 0.$$

Dann besitzt f in x_0 kein lokales Extremum. f'' kann nicht in einer ganzen Umgebung von x_0 verschwinden, denn dann müßte ja auch $f'''(x_0) = 0$ sein. Aber auch wenn eine Folge (x_ν) mit $x_\nu \to x_0$ existieren würde, so daß $f''(x_\nu) = 0$ für alle ν ist, so würde das bedeuten, daß

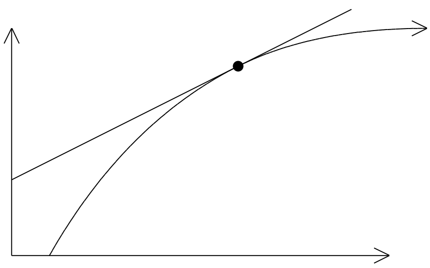
$$f'''(x_0) = \lim_{\nu \to \infty} \frac{f''(x_\nu) - f''(x_0)}{x_\nu - x_0} = 0$$

ist, im Widerspruch zur Voraussetzung.

Also muß es ein $\varepsilon > 0$ geben, so daß $f''(x) \neq 0$ für $|\, x - x_0 \,| < \varepsilon$ und $x \neq x_0$ ist. Da $f''(x_0) = 0$ ist und f'' in x_0 kein Extremum besitzt, muß f'' bei x_0 sein Vorzeichen wechseln, und demnach f' bei x_0 sein Monotonieverhalten. Der Graph einer Funktion mit monoton wachsender Ableitung ist nach links gekrümmt, der einer Funktion mit monoton fallender Ableitung ist nach rechts gekrümmt.



Linkskrümmung



Rechtskrümmung

Definition:

Sei $f \in \mathcal{C}^2(I)$, $x_0 \in I$ und $f''(x_0) = 0$.

f hat in x_0 einen *Wendepunkt*, falls f'' dort sein Vorzeichen ändert (also f von einer Linkskrümmung in eine Rechtskrümmung übergeht, oder umgekehrt).

Bemerkung: Wir brauchen für einen Wendepunkt nicht die Voraussetzung, daß $f'(x_0) = 0$ ist. Wenn das zusätzlich erfüllt ist, sprechen wir von einem *Sattelpunkt*.

Wir haben gezeigt:

Hinreichendes Kriterium für Wendepunkte

Sei I ein offenes Intervall, $f \in \mathcal{C}^3(I)$ und $x_0 \in I$.

Ist $f''(x_0) = 0$ und $f'''(x_0) \neq 0$, so besitzt f in x_0 einen Wendepunkt.

Beispiele:

1. Sei $f(x) := x^3$.

   Es ist $f''(x) = 6x$, also $f''(0) = 0$. Für $x < 0$ ist $f''(x) < 0$, und für $x > 0$ ist $f''(x) > 0$. Also wechselt f von einer Rechtskrümmung zu einer Linkskrümmung und hat damit in 0 einen Wendepunkt.

Tatsächlich ist $f'''(0) = 6 \neq 0$.

2. Sei $f(x) := x^4$, also $f''(x) = 12x^2$, $f'''(x) = 24x$ und $f^{(4)}(x) = 24$.

   Es ist $f'(0) = f''(0) = f'''(0) = 0$, aber $f^{(4)}(0) > 0$. Damit muß f im Nullpunkt ein Minimum besitzen, es kann dort kein Wendepunkt vorliegen!

3. Jetzt betrachten wir noch $f(x) := x^5$.

   Es ist $f'(x) = 5x^4$, $f''(x) = 20x^3$ und $f'''(x) = 60x^2$, also
   $f''(0) = 0$ **und** $f'''(0) = 0$.

   Aber offensichtlich ist $f''(x) < 0$ für $x < 0$ und $f''(x) > 0$ für $x > 0$. Damit besitzt f im Nullpunkt einen Wendepunkt.

§4 Integrale und Stammfunktionen

Definition:

Ist $M \subset \mathbb{R}$ eine beliebige Teilmenge, so nennt man die durch

$$\chi_M(t) := \begin{cases} 1 & \text{für } t \in M, \\ 0 & \text{sonst,} \end{cases}$$

definierte Funktion $\chi_M : \mathbb{R} \to \mathbb{R}$ die *charakteristische Funktion* von M.

Wir wollen im Folgenden charakteristische Funktionen von beschränkten Intervallen betrachten, also von Intervallen I der Form (a, b), $[a, b)$, $(a, b]$ oder $[a, b]$. In jedem dieser Fälle nennt man $l(I) := b - a$ die Länge des Intervalls, und mit $\mathring{I}$ bezeichnet man das offene Intervall (a, b). Ein Punkt liegt genau dann ganz im Innern von I, wenn er in $\mathring{I}$ liegt.

Definition:

Eine *Treppenfunktion* auf $\mathbb{R}$ ist eine endliche Linearkombination

$$\varphi = c_1\chi_{I_1} + \cdots + c_r\chi_{I_r},$$

wobei $I_1, \ldots, I_r$ beschränkte Intervalle und $c_1, \ldots, c_r$ reelle Konstanten sind.

Man nennt φ eine *Treppenfunktion auf dem (abgeschlossenen) Intervall* $[a, b]$, falls $\varphi(t) = 0$ für $t < a$ und $t > b$ ist.

$\mathcal{T}$ sei die Menge aller Treppenfunktionen auf $\mathbb{R}$, und $\mathcal{T}_a^b$ die Teilmenge aller Treppenfunktionen auf $[a, b]$.

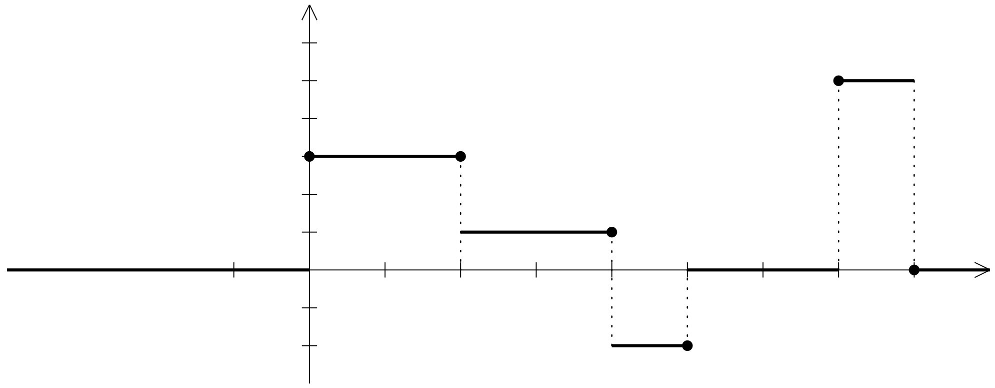
Es ist offensichtlich, daß $\mathcal{T}$ ein Vektorraum über $\mathbb{R}$ und $\mathcal{T}_a^b$ ein Unterraum von $\mathcal{T}$ ist. $\mathcal{T}$ selbst ist übrigens ein Unterraum des Vektorraumes $\mathcal{B}(\mathbb{R}, \mathbb{R})$ aller beschränkten Funktionen auf $\mathbb{R}$, trägt also auf natürliche Weise eine Norm.

Beispiel:

Sei $I_1 := [0, 4]$, $I_2 := (2, 5]$, $I_3 := [7, 8)$ und $\varphi(t) := 3 \cdot \chi_{I_1} - 2 \cdot \chi_{I_2} + 5 \cdot \chi_{I_3}$. Dann gilt:

$$\varphi(t) = \begin{cases} 3 & \text{für } 0 \le t \le 2 \\ 1 & \text{für } 2 < t \le 4 \\ -2 & \text{für } 4 < t \le 5 \\ 0 & \text{für } 5 < t < 7 \\ 5 & \text{für } 7 \le t < 8 \end{cases} \qquad \text{und } \varphi(t) = 0 \text{ für } t < 0 \text{ und } t \ge 8.$$

Dann ergibt sich folgendes Bild:



Es sieht also so aus, als ob eine Treppenfunktion immer stückweise konstant ist. Wir müssen nur $\mathbb{R}$ in geeigneter Weise zerlegen.

Unter einer *Zerlegung* von $\mathbb{R}$ verstehen wir eine aufsteigende endliche Folge von reellen Zahlen $t_0 < t_1 < \ldots < t_n$. Eine weitere Zerlegung nennen wir eine *Verfeinerung* der ersten, wenn sie aus ihr durch Hinzunahme weiterer Teilungspunkte hervorgegangen ist.

Zulässige Zerlegung zu einer Treppenfunktion

Zu einer Treppenfunktion φ auf $\mathbb{R}$ gibt es eine Zerlegung $t_0 < t_1 < \ldots < t_n$ und Zahlen $k_1, \ldots, k_n$, so daß gilt:

$$\begin{array}{rcl} \varphi|_{(t_{i-1},t_i)} & \equiv & k_i \quad \textit{für } i = 1, \ldots, n \\ \textit{und} \quad \varphi(t) & = & 0 \quad \textit{für } t < t_0 \textit{ und } t > t_n. \end{array}$$

Jede Zerlegung mit den obigen Eigenschaften nennt man eine *zulässige Zerlegung* für φ. Man beachte, daß nichts über die Werte von φ in den Punkten t_i ausgesagt wird!

Beweis: Sei $\varphi = \sum_{\varrho=1}^{r} c_\varrho \chi_{I_\varrho}$ eine Treppenfunktion, und $\mathring{I}_\varrho = (a_\varrho, b_\varrho)$, für $\varrho = 1, \ldots, r$.

Man kann offensichtlich eine Zerlegung $t_0 < t_1 < \ldots < t_n$ finden, so daß gilt:

$$\{a_1, b_1, a_2, b_2, \ldots, a_r, b_r\} = \{t_0, t_1, \ldots, t_n\}.$$

Wir setzen

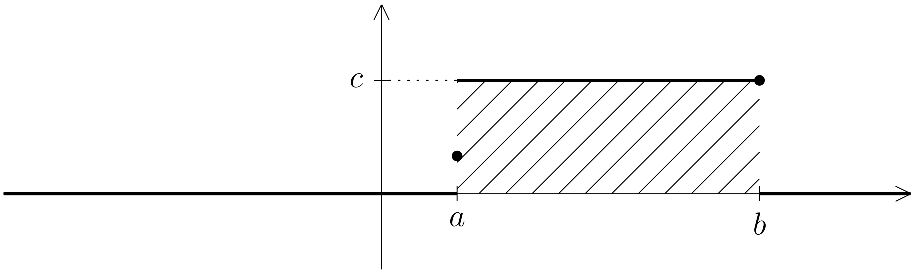
$$\begin{array}{rrcl} & M & := & \{1, \ldots, r\} \\ \text{und} & M_i & := & \{\varrho \in M \mid (t_{i-1}, t_i) \subset (a_\varrho, b_\varrho)\}. \end{array}$$

Für $t \in (t_{i-1}, t_i)$ ist dann

$$\begin{aligned} \varphi(t) &= \sum_{\varrho \in M} c_\varrho \cdot \chi_{I_\varrho}(t) \\ &= \sum_{\varrho \in M_i} c_\varrho \cdot 1 + \sum_{\varrho \in M \setminus M_i} c_\varrho \cdot 0 \\ &= \sum_{\varrho \in M_i} c_\varrho =: k_i = \text{konstant}. \end{aligned}$$

Und liegt t außerhalb $[t_0, t_n]$, so ist natürlich $\varphi(t) = 0$. $\square$

Ist I ein beschränktes Intervall (mit den Endpunkten a und b) und c eine positive reelle Zahl, so ist $\varphi := c \cdot \chi_I$ eine besonders einfache Treppenfunktion und $a = t_0 < t_1 = b$ eine zulässige Zerlegung für φ. Die Fläche unter dem Graphen von φ wird durch die Zahl $\Sigma(\varphi) := c \cdot (b - a) = c \cdot l(I)$ gegeben.



Ist $c < 0$, so kann man höchstens von der Fläche zwischen dem Graphen und der x-Achse sprechen. Der Flächeninhalt ist gleich geblieben, aber die Zahl $\Sigma(\varphi)$ ist nun negativ.

Definition:

Ist $\varphi = c_1\chi_{I_1} + \cdots + c_r\chi_{I_r}$ eine Treppenfunktion, so versteht man unter dem *Integral* von φ die Zahl $\quad \Sigma(\varphi) := \sum_{\varrho=1}^{r} c_\varrho \cdot l(I_\varrho)$.

Diese Definition ist einfach, aber nicht sehr anschaulich. Und da die Darstellung einer Treppenfunktion als Linearkombination von charakteristischen Funktionen nicht eindeutig ist, muß man auch noch zeigen, daß die Definition von $\Sigma(\varphi)$ unabhängig von der gewählten Darstellung von φ ist.

$\Sigma(\varphi)$ ist unabhängig von der Darstellung von φ

Sei $t_0 < t_1 < \ldots < t_n$ eine zulässige Zerlegung für die Treppenfunktion φ, also $\varphi|_{(t_{i-1},t_i)} \equiv k_i$ und $\varphi(t) = 0$ für $t < t_0$ und $t > t_n$, so ist

$$\Sigma(\varphi) = \sum_{i=1}^{n} k_i \cdot (t_i - t_{i-1}).$$

Beweis: Sei $\varphi = \sum_{\varrho=1}^{r} c_\varrho \cdot \chi_{I_\varrho}$ und $\mathring{I}_\varrho = (a_\varrho, b_\varrho)$, für $\varrho = 1, \ldots, r$.

Geht man von der gegebenen zulässigen Zerlegung zu einer Verfeinerung über, so erhält man wieder eine zulässige Zerlegung. Aber für $t_{i-1} < \xi < t_i$ ist $k_i \cdot (\xi - t_{i-1}) + k_i \cdot (t_i - \xi) = k_i \cdot (t_i - t_{i-1})$. Wir können also o.B.d.A. annehmen, daß die gegebene zulässige Zerlegung schon fein genug ist, nämlich so fein, daß gilt:

$$\{a_1, b_1, a_2, b_2, \ldots, a_r, b_r\} \subset \{t_0, \ldots, t_n\}.$$

Die Intervalle (t_{i-1}, t_i) enthalten dann keinen der Punkte a_ϱ oder b_ϱ. Wenn sie also zu einem I_ϱ nicht disjunkt sind, müssen sie schon ganz darin enthalten sein. Wie oben setzen wir

$$\begin{aligned} M &:= \{1, \ldots, r\} \\ \text{und} \quad M_i &:= \{\varrho \in M \mid (t_{i-1}, t_i) \subset I_\varrho\}. \end{aligned}$$

Außerdem sei

$$f(i, \varrho) := \begin{cases} c_\varrho & \text{falls } \varrho \in M_i, \\ 0 & \text{sonst.} \end{cases}$$

Dann ist $\sum_{\varrho=1}^{r} f(i, \varrho) = \sum_{\varrho \in M_i} c_\varrho = k_i$.

Da sich I_ϱ aus allen Intervallen (t_{i-1}, t_i) mit $\varrho \in M_i$ zusammensetzt, ist

$$l(I_\varrho) = \sum_{i \text{ mit } \varrho \in M_i} (t_i - t_{i-1}), \qquad \text{also } c_\varrho \cdot l(I_\varrho) = \sum_{i=1}^{n} f(i, \varrho)(t_i - t_{i-1}),$$

und

$$\begin{aligned} \Sigma(\varphi) &= \sum_{\varrho=1}^{r} c_\varrho \cdot l(I_\varrho) \\ &= \sum_{\varrho=1}^{r} \sum_{i=1}^{n} f(i, \varrho)(t_i - t_{i-1}) \\ &= \sum_{i=1}^{n} \left(\sum_{\varrho=1}^{r} f(i, \varrho) \right) (t_i - t_{i-1}) \\ &= \sum_{i=1}^{n} k_i \cdot (t_i - t_{i-1}). \end{aligned}$$

Das war zu zeigen. □

Eigenschaften des Integrals von Treppenfunktionen

1. *$\varphi \mapsto \Sigma(\varphi)$ ist eine Linearform auf $\mathcal{T}$.*
2. *Es ist $\Sigma(\chi_{[0,1]}) = 1$.*
3. *Ist $\varphi \le \psi$, so ist $\Sigma(\varphi) \le \Sigma(\psi)$*
4. *Ist $\varphi \in \mathcal{T}_a^b$, so ist $|\,\Sigma(\varphi)\,| \le (b-a) \cdot \|\,\varphi\,\|$.*

   Dabei ist $\|\,\varphi\,\| := \sup\{|\,\varphi(t)\,| \,:\, t \in [a, b]\}$.

BEWEIS:

1) Die Linearität folgt sehr einfach aus der Gleichung

$$\Sigma(c_1\chi_{I_1} + \cdots + c_r\chi_{I_r}) = c_1 \cdot \Sigma(\chi_{I_1}) + \cdots + c_r \cdot \Sigma(\chi_{I_r}).$$

2) ist trivial.

3) Ist $\varphi \geq 0$, so ist offensichtlich auch $\Sigma(\varphi) \geq 0$. Das sieht man sofort, wenn man eine zulässige Zerlegung für φ benutzt.

Ist $\psi \geq \varphi$, so ist $\psi - \varphi \geq 0$ und daher

$$\Sigma(\psi) - \Sigma(\varphi) = \Sigma(\psi - \varphi) \geq 0, \quad \text{also } \Sigma(\psi) \geq \Sigma(\varphi).$$

4) Ist $a = t_0 < t_1 < \ldots < t_n = b$ eine zulässige Zerlegung für φ und $\varphi|_{(t_i,t_{i-1})} \equiv k_i$, so ist

$$|\,\Sigma(\varphi)\,| \leq \sum_{i=1}^{n} |\,k_i\,| \cdot (t_i - t_{i-1}) \leq \|\,\varphi\,\| \cdot (b-a).$$

$\square$

Definition:

Sei I ein beschränktes Intervall mit unterer Grenze a und oberer Grenze b. Eine Funktion $f : I \to \mathbb{R}$ heißt *Regelfunktion* auf I, wenn f in jedem Punkt $x \in (a,b)$ sowohl einen linksseitigen als auch einen rechtsseitigen Grenzwert, in a einen rechtsseitigen und in b einen linksseitigen Grenzwert besitzt.

$R(I)$ sei der Vektorraum der Regelfunktionen auf I.

Alle stetigen Funktionen, alle monotonen Funktionen und alle Treppenfunktionen auf I sind Regelfunktionen.

Mit f und g sind auch $|\,f\,|$, $f \cdot g$, $\max(f,g)$ und $\min(f,g)$ Regelfunktionen.

Approximationssatz

Eine Funktion $f : [a,b] \to \mathbb{R}$ ist genau dann eine Regelfunktion, wenn es eine Folge von Treppenfunktionen $\varphi_n \in \mathcal{T}_a^b$ gibt, die auf $[a,b]$ gleichmäßig gegen f konvergiert.

Der Beweis ist nicht ganz einfach und wird deshalb hier weggelassen.

Da jede Treppenfunktion höchstens endlich viele Unstetigkeitsstellen besitzt, ist das folgende Ergebnis nicht allzu verwunderlich:

Unstetigkeitsstellen von Regelfunktionen

Jede Regelfunktion $f : [a,b] \to \mathbb{R}$ ist mit Ausnahme von höchstens abzählbar vielen Stellen stetig.

Jetzt können wir den Integralbegriff vom Raum der Treppenfunktionen auf den Raum der Regelfunktionen ausdehnen:

Definition:

Sei $f : [a,b] \to \mathbb{R}$ eine Regelfunktion und (φ_n) eine Folge von Treppenfunktionen, die auf $[a,b]$ gleichmäßig gegen f konvergiert. Dann nennt man

$$\int_a^b f(x)\,dx := \lim_{n\to\infty} \Sigma(\varphi_n)$$

das *Integral* von f über $[a,b]$

Bemerkung : Es ist nicht allzu schwer, folgendes zu zeigen:

1. Der Grenzwert $\lim\limits_{n\to\infty} \Sigma(\varphi_n)$ existiert für jede Folge von Treppenfunktionen (φ_n), die gleichmäßig gegen f konvergiert.
2. Der Grenzwert ist unabhängig von der Auswahl der Folge (φ_n).

Diese beiden Aussagen machen die Integral-Definition erst sinnvoll.

Eigenschaften des Integrals von Regelfunktionen

$f, g : [a,b] \to \mathbb{R}$ seien Regelfunktionen. Dann gilt:

1. *Linearität: Für $\alpha, \beta \in \mathbb{R}$ ist*

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

2. *Normierung: $\int_0^1 dx = 1$.*
3. *Monotonie:*

$$\text{Ist } f(x) \le g(x) \text{ auf } [a,b], \text{ so ist auch } \int_a^b f(x)\,dx \le \int_a^b g(x)\,dx.$$

4. *Beschränktheit:* $\left| \int_a^b f(x)\,dx \right| \le \int_a^b |f(x)|\,dx \le (b-a) \cdot \| f \|$.

Beweis:

1) Konvergiert (φ_n) gleichmäßig gegen f und (ψ_n) gleichmäßig gegen g, so konvergiert auch $(\alpha \cdot \varphi_n + \beta \cdot \psi_n)$ gleichmäßig gegen $\alpha \cdot f + \beta \cdot g$. Da die Treppenfunktionen einen Vektorraum bilden, gilt:

$$\begin{aligned}
\int_a^b (\alpha \cdot f(x) + \beta \cdot g(x))\,dx &= \lim_{n\to\infty} \Sigma(\alpha \cdot \varphi_n + \beta \cdot \psi_n) \\
&= \lim_{n\to\infty} (\alpha \cdot \Sigma(\varphi_n) + \beta \cdot \Sigma(\psi_n)) \\
&= \alpha \cdot \lim_{n\to\infty} \Sigma(\varphi_n) + \beta \cdot \lim_{n\to\infty} \Sigma(\psi_n)
\end{aligned}$$

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

2) ist trivial.

3) Es reicht zu zeigen: $f \geq 0 \implies \int_a^b f(x)\,dx \geq 0$.

Sei also $f \geq 0$. Ist (φ_n) eine Folge von Treppenfunktionen, die gleichmäßig gegen f konvergiert, so ist jeweils auch $\varphi_n^* := \varphi_n + \| f - \varphi_n \|$ eine Treppenfunktion, und es gilt:

$$\| f - \varphi_n^* \| = \| f - \varphi_n - (\| f - \varphi_n \|) \| \leq 2 \cdot \| f - \varphi_n \|.$$

Also konvergiert auch (φ_n^*) gleichmäßig gegen f. Und schließlich ist

$$\varphi_n^* = f + (\varphi_n - f) + \| \varphi_n - f \| \geq f - \| \varphi_n - f \| + \| \varphi_n - f \| = f \geq 0,$$

also auch $\Sigma(\varphi_n^*) \geq 0$, und damit

$$\int_a^b f(x)\,dx = \lim_{n\to\infty} \Sigma(\varphi_n^*) \geq 0.$$

4) Es ist

$$-| f | \leq f \leq | f |,$$

also auch

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

Aber das bedeutet:

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

Konvergiert nun (φ_n) gleichmäßig gegen f, so konvergiert auch $(| \varphi_n |)$ gegen $| f |$ und $\| \varphi_n \|$ gegen $\| f \|$, und da

$$\Sigma(| \varphi_n |) \leq (b-a) \cdot \| \varphi_n \|$$

ist, für jedes n, so ist auch

$$\int_a^b | f(x) |\,dx = \lim_{n\to\infty} \Sigma(| \varphi_n |) \leq (b-a) \cdot \| f \|.$$

□

Außerdem zeigt man leicht:

Intervall-Additivität

Ist $a < c < b$ und $f : [a,b] \to \mathbb{R}$ eine Regelfunktion, so ist

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

Und der Vollständigkeit halber definiert man:

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

$$\text{und} \quad \int_b^a f(x)\,dx := -\int_a^b f(x)\,dx, \qquad \text{für } a < b.$$

Ab sofort können wir die Definition des Integrals mit Hilfe von approximierenden Treppenfunktionen eigentlich wieder vergessen, denn wir benutzen nur noch die **Eigenschaften** des Integrals, die wir gerade bewiesen haben. Allerdings sollte man im Auge behalten, daß das Integral bei positiven Funktionen die Fläche unter dem Graphen wiedergibt.

Hilfssatz:

Ist $f : [a, b] \to \mathbb{R}$ *eine Regelfunktion und* $m \le f(x) \le M$ *für alle* $x \in [a, b]$, *so ist*

$$m \cdot (b - a) \le \int_a^b f(x)\,dx \le M \cdot (b - a).$$

BEWEIS: Dies ist eine triviale Anwendung der Monotonie–Eigenschaft. □

Mittelwertsatz der Integralrechnung

Ist $f : [a, b] \to \mathbb{R}$ *stetig und* $p : [a, b] \to \mathbb{R}$ *eine nirgends negative Regelfunktion, so gibt es ein* $c \in [a, b]$ *mit*

$$\int_a^b f(x)p(x)\,dx = f(c) \cdot \int_a^b p(x)\,dx.$$

BEWEIS: Es gibt reelle Zahlen m, M, so daß $m \le f(x) \le M$ auf $[a, b]$ ist. Wir wählen m als das globale Minimum und M als das globale Maximum von f auf $[a, b]$. Diese Werte werden sogar in gewissen Punkten des Intervalls von f angenommen.

Dann ist

$$m \cdot p(x) \le f(x)p(x) \le M \cdot p(x) \text{ auf } [a, b],$$

und wegen der Linearität und der Monotonie des Integrals ist dann auch

$$m \cdot \int_a^b p(x)\,dx \le \int_a^b f(x)p(x)\,dx \le M \cdot \int_a^b p(x)\,dx.$$

Durch $F(x) := f(x) \cdot \int_a^b p(t)\,dt$ wird nun eine stetige Funktion $F : [a, b] \to \mathbb{R}$ definiert, die die Werte $m \cdot \int_a^b p(t)\,dt$ und $M \cdot \int_a^b p(t)\,dt$ in $[a, b]$ annimmt. Nach dem Zwischenwertsatz und wegen der obigen Ungleichung muß F dann in einem geeigneten Punkt $c \in [a, b]$ auch den Wert $\int_a^b f(x)p(x)\,dx$ annehmen. Also ist

$$f(c) \cdot \int_a^b p(t)\,dt = F(c) = \int_a^b f(x)p(x)\,dx.$$

□

Bemerkung: Wendet man den Mittelwertsatz auf $p(x) \equiv 1$ an, so erhält man die Formel

$$\boxed{\exists\, c \in [a, b] \text{ mit } \int_a^b f(x)\,dx = f(c) \cdot (b - a).}$$

Bis jetzt sieht es so aus, als sei es sehr mühsam, Integrale konkret auszurechnen. Isaac Newton und Gottfried Wilhelm Leibniz verdanken wir jedoch eine viel effektivere Methode.

Wir betrachten ein Integral mit **variabler Obergrenze** über eine stetige Funktion f und erhalten so eine neue Funktion:

$$F(x) := \int_a^x f(u)\, du.$$

Aufgepaßt! F hängt von der Obergrenze x ab, nicht von der Integrationsvariablen u.

Es sei nun $a < x_0 < x < b$. Dann gilt:

$$F(x) - F(x_0) = \int_a^x f(u)\, du - \int_a^{x_0} f(u)\, du = \int_{x_0}^x f(u)\, du.$$

Als stetige Funktion nimmt f auf dem Intervall $[x_0, x]$ ein Minimum $m(f, x)$ und ein Maximum $M(f, x)$ an. Und die Monotonie des Integrals liefert:

$$m(f, x) \cdot (x - x_0) \leq F(x) - F(x_0) \leq M(f, x) \cdot (x - x_0).$$

Da $x - x_0 > 0$ ist, folgt:

$$m(f, x) \leq \frac{F(x) - F(x_0)}{x - x_0} \leq M(f, x).$$

In der Mitte steht ein Differenzenquotient von F, links und rechts jeweils eine von x abhängige Größe. Läßt man nun x gegen x_0 wandern, so streben $m(f, x)$ und $M(f, x)$ beide gegen $f(x_0)$. Also muß auch der Differenzenquotient in der Mitte einen Limes besitzen, und man erhält:

$$f(x_0) = \lim_{x \to x_0} \frac{F(x) - F(x_0)}{x - x_0} = F'(x_0).$$

Diese Beziehung bleibt richtig, wenn man auch Argumente $x < x_0$ zuläßt.

Definition:

Sei I ein Intervall, $f : I \to \mathbb{R}$ stetig, $F : I \to \mathbb{R}$ differenzierbar und $F' = f$.
Dann heißt F eine *Stammfunktion* von f.

Wir haben gerade gesehen, daß $F(x) = \displaystyle\int_a^x f(u)\, du$ eine Stammfunktion von f ist.

Sind F_1, F_2 zwei Stammfunktionen von f, so ist $(F_1 - F_2)' = 0$, also $F_1 - F_2 = \mathit{konst}$. Damit gilt für eine beliebige Stammfunktion F von f:

$$\int_a^x f(u)\, du = F(x) + c, \text{ mit einer geeigneten Konstante } c.$$

Setzt man $x = a$, so erhält man die Gleichung $0 = F(a) + c$, also $c = -F(a)$. Setzt man $x = b$ ein, so erhält man:

$$\int_a^b f(u)\, du = F(x) \Big|_a^b := F(b) - F(a).$$

Zusammengefaßt haben wir bewiesen:

Fundamentalsatz der Differential- und Integralrechnung

Sei $f : [a, b] \to \mathbb{R}$ eine stetige Funktion. Dann gilt:

1. $F(x) := \int_a^x f(u)\,du$ *ist eine Stammfunktion von* f.
2. *Sind* F_1, F_2 *zwei Stammfunktionen von* f, *so ist* $F_1 - F_2$ *konstant.*
3. *Ist* F *eine Stammfunktion von* f, *so ist*

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

Bemerkung : Die Menge aller Stammfunktionen von f wird oft mit dem Symbol $\int f(x)\,dx$ bezeichnet, und man spricht dann von *dem unbestimmten Integral.* Diese Bezeichnungsweise korrekt durchzuhalten, ist mühsam. Wir sprechen deshalb lieber von *einem unbestimmten Integral* und verstehen darunter eine Funktion der Gestalt

$$F(x) := \int_a^x f(u)\,du\,,$$

wobei die Untergrenze a beliebig im Definitionsbereich von f gewählt werden kann. Aus Bequemlichkeit werden wir dabei manchmal auch die Integralgrenzen weglassen.

Beispiele :

1. Da $(x^{n+1})' = (n+1)\cdot x^n$ ist, also $\left(\dfrac{x^{n+1}}{n+1}\right)' = x^n$, folgt:

$$\int_a^b x^n\,dx = \frac{1}{n+1}(b^{n+1} - a^{n+1}).$$

2. Es soll $\int_{-2}^{+2} |\,x-1\,|\,dx$ berechnet werden. Dabei stört zunächst die Betragsfunktion. Es ist aber halb so schlimm, wir müssen nur die Nullstellen ermitteln und stückweise integrieren:

$$\begin{aligned}
\int_{-2}^{+2} |\,x-1\,|\,dx &= \int_{-2}^{1}(1-x)\,dx + \int_1^2 (x-1)\,dx \\
&= (x - \frac{x^2}{2})\,\Big|_{-2}^{1} + (\frac{x^2}{2} - x)\,\Big|_1^2 \\
&= (\frac{1}{2} - (-4)) + (0 - (-\frac{1}{2})) \\
&= 5.
\end{aligned}$$

3. Da $\cos'(x) = -\sin(x)$ ist, folgt:

$$\int_0^\pi \sin(x)\,dx = -(\cos(\pi) - \cos(0)) = 2$$

und

$$\int_0^{2\pi} \sin(x)\,dx = -(\cos(2\pi) - \cos(0)) = 0.$$

Im zweiten Fall heben sich positiv und negativ zu rechnende Flächenteile gegenseitig weg.

4. Bekanntlich ist $\arctan'(x) = \dfrac{1}{x^2+1}$ und $\arctan(0) = 0$. Daraus folgt:

$$\arctan(x) = \int_0^x \frac{1}{1+u^2}\,du.$$

Da $\tan(\frac{\pi}{4}) = 1$ ist, ist $\arctan(1) = \dfrac{\pi}{4}$, also

$$\boxed{\frac{\pi}{4} = \int_0^1 \frac{1}{1+u^2}\,du}$$

Das liefert uns eine interessante Definition für die Zahl π .

Wir fassen noch die gängigen Stammfunktionen zu einer Tabelle zusammen:

Funktion	Stammfunktion	
x^n	$\frac{1}{n+1} \cdot x^{n+1}$	$n \in \mathbb{N}$
$\frac{1}{x^n}$	$\frac{-1}{(n-1)x^{n-1}}$	$n \in \mathbb{N}, n \geq 2, x \neq 0$
$\frac{1}{\sqrt{x}}$	$2\sqrt{x}$	$x > 0$
$\frac{1}{x}$	$\ln(\lvert x \rvert)$	$x \neq 0$
$\frac{1}{1+x^2}$	$\arctan(x)$	
$\sin(x)$	$\cos(x)$	
$\cos(x)$	$\sin(x)$	
$\frac{1}{\cos^2(x)} = 1 + \tan^2(x)$	$\tan(x)$	$x \neq (n + \frac{1}{2})\pi,\ n \in \mathbb{Z}$
$\frac{1}{\sin^2(x)}$	$-\cot(x)$	$x \neq n\pi,\ n \in \mathbb{Z}$
a^x	$\frac{1}{\ln(a)} \cdot a^x$	
e^x	e^x	
$\sinh(x)$	$\cosh(x)$	
$\cosh(x)$	$\sinh(x)$	
$\frac{1}{\sqrt{1+x^2}}$	$\operatorname{arsinh}(x)$	

§5 Integrationstechniken

Wir wollen Differentiations-Regeln neu betrachten und in die Sprache der Integrale und Stammfunktionen übersetzen.

Wir beginnen mit der *Produktregel*:

$$(f \cdot g)' = f' \cdot g + f \cdot g'.$$

Sie besagt, daß $f \cdot g$ eine Stammfunktion von $f' \cdot g + f \cdot g'$ ist. Nun kommt es selten vor, daß der Integrand deutlich sichtbar diese Form besitzt, aber umso häufiger ist er von der Form $f \cdot g'$. Wenn wir die Gleichung nach $f \cdot g'$ auflösen und dann integrieren, erhalten wir:

Regel der Partiellen Integration

Sind f und g über $[a, b]$ stetig differenzierbar, so ist

$$\int_a^b f(x)g'(x)\,dx = (f(x) \cdot g(x))\Big|_a^b - \int_a^b f'(x)g(x)\,dx.$$

Auf den ersten Blick ist der Nutzen dieser Formel noch nicht zu sehen, denn statt des Integrals über $f \cdot g'$ muß man nun das Integral über $f' \cdot g$ berechnen. Den Vorteil erkennt man am besten an Beispielen:

Beispiele :

1. Sei $0 < a < b$. Dann gilt:

$$\begin{aligned} \int_a^b \ln(x)\,dx &= \int_a^b \ln(x) \cdot x'\,dx \quad \text{(denn es ist } x' = 1) \\ &= (\ln(x) \cdot x)\Big|_a^b - \int_a^b \ln'(x) \cdot x\,dx \quad \text{(Partielle Integration)} \\ &= (x \cdot \ln(x))\Big|_a^b - (x)\Big|_a^b \quad \text{(denn es ist } \ln'(x) \cdot x = 1). \end{aligned}$$

   Also ist $x \cdot \ln(x) - x$ eine Stammfunktion von $\ln(x)$. Zur Probe kann man ja differenzieren.

   Die Regel der Partiellen Integration hat hier weitergeholfen, weil $f' \cdot g$ viel einfacher als $f \cdot g'$ ist. Wie man in solchen Fällen jeweils f und g wählen muß, sagt einem aber keiner. Hier fängt die Kunst des Integrierens an.

2. Das folgende Beispiel ist besonders typisch und hat schon fast den Charakter eines Kochrezeptes:

$$\begin{aligned} \int_a^b \sin^2(x)\,dx &= \int_a^b (-\cos'(x)) \cdot \sin(x)\,dx \\ &= (-\cos(x) \cdot \sin(x))\Big|_a^b - \int_a^b (-\cos(x)) \cdot \cos(x)\,dx \end{aligned}$$

$$\begin{aligned} &= -(\cos(x)\cdot\sin(x))\Big|_a^b + \int_a^b \cos^2(x)\,dx \\ &= -(\cos(x)\cdot\sin(x))\Big|_a^b + (x)\Big|_a^b - \int_a^b \sin^2(x)\,dx. \end{aligned}$$

Jetzt ist das Integral, das wir ausrechnen wollten, wieder aufgetaucht! Trotzdem hilft uns das weiter: Das Integral über $\sin^2(x)$ auf der rechten Seite hat nämlich das richtige Vorzeichen. Wir können es auf die andere Seite der Gleichung bringen und erhalten:

$$\int_a^b \sin^2(x)\,dx = \frac{1}{2}\cdot\left((x-\cos(x)\cdot\sin(x))\Big|_a^b\right).$$

3. Von ähnlicher Bauart ist das folgende Beispiel:

$$\begin{aligned} \int_a^b e^x\cdot\sin(x)\,dx &= \int_a^b e^x\cdot(-\cos'(x))\,dx \\ &= -(e^x\cdot\cos(x))\Big|_a^b - \int_a^b (e^x)'\cdot(-\cos(x))\,dx \\ &= -(e^x\cdot\cos(x))\Big|_a^b + \int_a^b e^x\cdot\cos(x)\,dx. \end{aligned}$$

Eine zweite Rechnung liefert:

$$\int_a^b e^x\cdot\cos(x)\,dx = (e^x\cdot\sin(x))\Big|_a^b - \int_a^b e^x\cdot\sin(x)\,dx.$$

Setzt man das in das erste Ergebnis ein, so erhält man:

$$\int_a^b e^x\cdot\sin(x)\,dx = \frac{1}{2}\cdot\left((e^x\cdot(\sin(x)-\cos(x)))\Big|_a^b\right).$$

Die zweite Technik, die wir kennenlernen wollen, ist noch etwas schwieriger.

Durch Probieren kann man rasch herausfinden, daß nicht nur $-\cos(x)$ Stammfunktion von $\sin(x)$ ist, sondern auch $-\dfrac{1}{a}\cos(ax)$ Stammfunktion von $\sin(ax)$. Wenn man zur Probe differenziert, benötigt man die Kettenregel. Daher liegt es nahe, diese Regel auf ihre Verwendbarkeit in der Integrationstheorie hin zu untersuchen:

Sei $f: I \to \mathbb{R}$ stetig, $F: I \to \mathbb{R}$ eine Stammfunktion von f. Weiter sei $\varphi: [\alpha,\beta] \to \mathbb{R}$ eine stetig differenzierbare Funktion, mit $\varphi([\alpha,\beta]) \subset I = [a,b]$.

Dann ist die Verknüpfung $F\circ\varphi: [\alpha,\beta]\to\mathbb{R}$ definiert und differenzierbar, es ist

$$(F\circ\varphi)'(t) = F'(\varphi(t))\cdot\varphi'(t) = (f\circ\varphi)(t)\cdot\varphi'(t).$$

Also ist $F\circ\varphi$ eine Stammfunktion von $(f\circ\varphi)\cdot\varphi'$.

Für das bestimmte Integral von $(f\circ\varphi)\cdot\varphi'$ über $[\alpha,\beta]$ ergibt das:

$$\int_\alpha^\beta f(\varphi(t))\cdot\varphi'(t)\,dt = \int_\alpha^\beta (F\circ\varphi)'(t)\,dt = F(\varphi(\beta)) - F(\varphi(\alpha)).$$

Andererseits ist

$$F(\varphi(\beta)) - F(\varphi(\alpha)) = \int_{\varphi(\alpha)}^{\varphi(\beta)} F'(x)\,dx.$$

Da $F' = f$ ist, haben wir bewiesen:

Substitutionsregel

Sei $\varphi : [\alpha, \beta] \to \mathbb{R}$ stetig differenzierbar, $\varphi([\alpha, \beta]) \subset I$ und $f : I \to \mathbb{R}$ stetig. Dann gilt:

$$\int_{\varphi(\alpha)}^{\varphi(\beta)} f(x)\,dx = \int_{\alpha}^{\beta} f(\varphi(t)) \cdot \varphi'(t)\,dt.$$

Merken kann man sich diese Regel mit folgender Eselsbrücke:
Ist $x = x(t)$, so ist

$$\int f(x)\,dx = \int f(x(t))\frac{dt}{dt}\,dx = \int f(x(t))\frac{dx}{dt}\,dt.$$

Auch hier zeigt sich der Nutzen am besten an Hand von Beispielen. Dabei beginnen wir mit dem einfacheren Fall, der Anwendung der Substitutionsregel **von rechts nach links**. Der Integrand ist dann in der Form $f(\varphi(t))\varphi'(t)$ gegeben.

Beispiele:

1. Häufig möchte man eine Funktion der Form $x \mapsto f(x + c)$ integrieren. Hier wird in f die Funktion $\varphi(t) := t + c$ eingesetzt. Da $\varphi'(t) \equiv 1$ ist, folgt:

$$\int_a^b f(t + c)\,dt = \int_{a+c}^{b+c} f(x)\,dx.$$

2. Nun untersuchen wir Funktionen der Form $x \mapsto f(x \cdot c)$. Hier wird $\varphi(t) := t \cdot c$ eingesetzt, mit $\varphi'(t) \equiv c$. Die Substitutionsregel liefert eine Formel für das Integral über $c \cdot f(t \cdot c)$. Wir können aber die Konstante c auf die andere Seite der Gleichung bringen und erhalten dann:

$$\int_a^b f(t \cdot c)\,dt = \frac{1}{c} \cdot \int_{a \cdot c}^{b \cdot c} f(x)\,dx.$$

3. Auch das folgende Beispiel ist eigentlich ein alter Bekannter:

   Es sei $f(x) := \dfrac{1}{x}$ und $\varphi(t)$ eine stetig differenzierbare Funktion ohne Nullstellen über $[\alpha, \beta]$. Dann gilt:

$$f(\varphi(t)) \cdot \varphi'(t) = \frac{\varphi'(t)}{\varphi(t)}.$$

   Also ist

$$\int_{\alpha}^{\beta} \frac{\varphi'(t)}{\varphi(t)}\,dt = \int_{\varphi(\alpha)}^{\varphi(\beta)} \frac{1}{x}\,dx = (\ln|\,x\,|)\Big|_{\varphi(\alpha)}^{\varphi(\beta)} = (\ln|\,\varphi(t)\,|)\Big|_{\alpha}^{\beta}.$$

   Das hätte man aber auch über die logarithmische Ableitung erhalten:

Da $\dfrac{\varphi'(t)}{\varphi(t)} = (\ln\circ|\varphi|)'(t)$ ist, ist $\ln\circ|\varphi|$ eine Stammfunktion von $\dfrac{\varphi'}{\varphi}$.

Z.B. ist

$$\int_a^b \tan(t)\,dt = \int_a^b \frac{-\cos'(t)}{\cos(t)}\,dt = -(\ln|\cos(t)|)\Big|_a^b .$$

4. Sei $f(x) := x^n$, φ beliebig. Dann gilt:

$$\begin{aligned}\int_\alpha^\beta \varphi(t)^n\cdot\varphi'(t)\,dt &= \int_{\varphi(\alpha)}^{\varphi(\beta)} f(x)\,dx \\ &= \frac{1}{n+1}x^{n+1}\Big|_{\varphi(\alpha)}^{\varphi(\beta)} \\ &= \frac{1}{n+1}\cdot\left(\varphi(\beta)^{n+1}-\varphi(\alpha)^{n+1}\right).\end{aligned}$$

Speziell ist $\dfrac{1}{2}\varphi(t)^2$ Stammfunktion von $\varphi(t)\varphi'(t)$, also etwa

$$\int_\alpha^\beta \frac{\ln(t)}{t}\,dt = \frac{1}{2}\ln(t)^2\Big|_\alpha^\beta .$$

Wenn die rechte Seite der Substitutionsregel der Ausgangspunkt ist, kann man die Substitution φ leicht erkennen. Schwieriger wird es, wenn man mit der linken Seite beginnt. Dann die geeignete Substitution zu finden, erfordert Kreativität und viel Routine. Wir beginnen mit einem noch relativ einfachen Beispiel:

Es soll das Integral $\displaystyle\int_a^b \sqrt{1-x^2}\,dx$ berechnet werden, für $-1 < a < b < +1$.

Für $a \to -1$ und $b \to +1$ wird damit die Fläche des halben Einheitskreises berechnet.

Wir suchen nach einer Substitution, durch die der Integrand einfacher wird. Die Gleichung $y = \sqrt{1-x^2}$ erinnert an die Gleichung $\cos(x) = \sqrt{1-\sin^2(x)}$, deshalb kann man es ja einmal mit der Substitution $\varphi(t) := \sin(t)$ versuchen. Da die Sinus-Funktion das Intervall $[-\frac{\pi}{2}, +\frac{\pi}{2}]$ bijektiv auf das Intervall $[-1,+1]$ abbildet, können wir die Substitutionsregel in folgender Form anwenden:

$$\int_a^b f(x)\,dx = \int_{\varphi^{-1}(a)}^{\varphi^{-1}(b)} f(\varphi(t))\varphi'(t)\,dt.$$

Setzen wir $\alpha := \arcsin(a)$ und $\beta := \arcsin(b)$, so erhalten wir:

$$\begin{aligned}\int_a^b \sqrt{1-x^2}\,dx &= \int_\alpha^\beta \sqrt{1-\sin^2(t)}\cdot\cos(t)\,dt \\ &= \int_\alpha^\beta \cos^2(t)\,dt.\end{aligned}$$

Tatsächlich hat sich die Situation vereinfacht, das neue Integral kann in der bekannten Weise mit Hilfe partieller Integration berechnet werden. Es ist

$$\int_\alpha^\beta \cos^2(t)\,dt = \int_\alpha^\beta \cos(t)\sin'(t)\,dt$$

$$\begin{aligned}
&= \cos(t)\sin(t)\Big|_\alpha^\beta + \int_\alpha^\beta \sin^2(t)\,dt \\
&= \cos(t)\sin(t)\Big|_\alpha^\beta + \int_\alpha^\beta (1-\cos^2(t))\,dt \\
&= (\cos(t)\sin(t)+t)\Big|_\alpha^\beta - \int_\alpha^\beta \cos^2(t)\,dt.
\end{aligned}$$

Also ist

$$\int_\alpha^\beta \cos^2(t)\,dt = \frac{1}{2}\cdot(\cos(t)\cdot\sin(t)+t)\Big|_\alpha^\beta .$$

Das Ergebnis kann nun noch etwas umformuliert werden. Zunächst können wir $\cos(t)$ durch $\sqrt{1-\sin^2(t)}$ ersetzen, so daß nur noch die Substitutions-Funktion $\sin(t)$ vorkommt. Dann ist es natürlich wünschenswert, die Hilfsgrößen α und β loszuwerden, damit das Ergebnis von a und b abhängt. Setzen wir die Gleichungen

$$\alpha = \arcsin(a) \quad \text{und} \quad \beta = \arcsin(b)$$

ein, so erhalten wir:

$$\int_a^b \sqrt{1-x^2}\,dx = \frac{1}{2}\cdot(t\cdot\sqrt{1-t^2}+\arcsin(t))\Big|_a^b .$$

Läßt man hier $a \to -1$ und $b \to +1$ gehen, so konvergiert auch die rechte Seite gegen $\frac{1}{2}\cdot(\arcsin(+1)-\arcsin(-1)) = \frac{\pi}{2}$. Also ist

$$\int_{-1}^{+1} \sqrt{1-x^2}\,dx = \frac{\pi}{2}.$$

Das sollte für die Fläche des halben Einheitskreises auch herauskommen!

Nachdem das Prinzip klargeworden ist, können wir weitere Beispiele behandeln:

Beispiele :

1. Es soll $\int e^{\sqrt{x}}\,dx$ berechnet werden:

   Wir verwenden die Substitution $x(t) = t^2$, also $x'(t) = 2t$. Dann ist $t = \sqrt{x}$ und

   $$\begin{aligned}
   \int e^{\sqrt{x}}\,dx &= \int e^t\cdot 2t\,dt \\
   &= 2(t-1)e^t \;=\; 2(\sqrt{x}-1)\cdot e^{\sqrt{x}}.
   \end{aligned}$$

2. Wir werden weiter unten zeigen, daß man jede rationale Funktion (außerhalb ihrer Polstellen) integrieren kann. Deshalb möchte man Integranden gerne mit Hilfe einer geeigneten Substitution rational machen:

   In $\int \frac{dx}{\sqrt[3]{x}+\sqrt{x}}$ benutzen wir die Substitution $x(t) = t^6$, mit $x'(t) = 6t^5$ und $t = \sqrt[6]{x}$. Dann gilt:

   $$\int \frac{dx}{\sqrt[3]{x}+\sqrt{x}} = \int \frac{6t^5}{t^2+t^3}\,dt = 6\int \frac{t^3}{1+t}\,dt.$$

3. Ist F rational, so integriert man $f(x) := F(e^x)$ mit Hilfe der Substitution $x(t) = \ln(t)$, also $x'(t) = \frac{1}{t}$ und $t = e^x$. Dann ist

$$\int F(e^x)\,dx = \int \frac{F(t)}{t}\,dt.$$

4. Jetzt sollte das Schema deutlich geworden sein.

   Sei etwa $F(x, y)$ eine rationale Funktion in x und y. Um $f(x) := F(x, \sqrt[m]{ax+b})$ zu integrieren, setzen wir $\sqrt[m]{ax+b} = t$, also $x(t) = \dfrac{1}{a}(t^m - b)$, mit $x'(t) = \dfrac{m}{a}t^{m-1}$. Das ergibt:

$$\int F(x, \sqrt[m]{ax+b})\,dx = \int F(\frac{1}{a}(t^m - b), t)\frac{m}{a}t^{m-1}\,dt.$$

5. Zum Schluß noch ein schwierigeres Beispiel:

   Es soll $\displaystyle\int F(\sin x, \cos x)\,dx$ für eine rationale Funktion $F = F(x, y)$ berechnet werden:

   In der Klasse der Winkelfunktionen, ihrer Umkehrungen und deren Ableitungen kennen wir nur einen einzigen Fall, wo eine rationale Funktion auftaucht:

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

   Daher versuchen wir, Sinus und Cosinus durch den Tangens auszudrücken: Es ist

$$\tan^2(x) = \frac{\sin^2(x)}{\cos^2(x)} = \frac{1-\cos^2(x)}{\cos^2(x)} \text{ und damit } 1 + \tan^2(x) = \frac{1}{\cos^2(x)},$$

   also

$$\begin{aligned} \cos^2(x) &= \frac{1}{1+\tan^2(x)} \\ \text{und} \quad \sin^2(x) &= 1 - \cos^2(x) = \frac{\tan^2(x)}{1+\tan^2(x)}. \end{aligned}$$

   Das ist noch nicht ganz befriedigend, weil jetzt für die Darstellung von Sinus und Cosinus durch den Tangens noch Wurzeln benötigt werden. Es gilt aber:

$$\begin{aligned} \sin(x) &= 2\sin(\frac{x}{2})\cos(\frac{x}{2}) = \frac{2\tan(\frac{x}{2})}{1+\tan^2(\frac{x}{2})}, \\ \text{und} \quad \cos(x) &= \cos^2(\frac{x}{2}) - \sin^2(\frac{x}{2}) = \frac{1-\tan^2(\frac{x}{2})}{1+\tan^2(\frac{x}{2})}. \end{aligned}$$

   Jetzt verwenden wir die Substitution $t = \tan(\dfrac{x}{2})$, also $x(t) = 2\arctan(t)$ und $x'(t) = \dfrac{2}{1+t^2}$. Dann folgt:

$$\int F(\sin x, \cos x)\,dx = \int F(\frac{2t}{1+t^2}, \frac{1-t^2}{1+t^2}) \cdot \frac{2}{1+t^2}\,dt.$$

   Damit ist der ganze Integrand rational geworden.

6. Manchmal kann es auch sinnvoll sein, Winkelfunktionen einzusetzen, wie wir es schon beim ersten Beispiel zur Substitutionsregel gesehen haben. Die Substitution $x(t) = \sin(t)$ ergibt z.B.:

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

Wir wollen nun – wie versprochen – zeigen, daß jede rationale Funktion integrierbar ist:

$$\text{Es sei} \quad f(x) = \frac{P(x)}{Q(x)}, \text{ mit Polynomen } P \text{ und } Q.$$

Wir gehen in mehreren Schritten vor:

1. Schritt:

Ist $\deg(P) \geq \deg(Q)$, so führt man eine Polynomdivision durch:

$$f(x) = P_0(x) + \frac{R(x)}{Q(x)}, \text{ mit Polynomen } P_0 \text{ und } R,$$

sowie $\deg(R) < \deg(Q)$.

Da Polynome problemlos integriert werden können, braucht man nur den Fall $f(x) = \dfrac{P(x)}{Q(x)}$ mit $\deg(P) < \deg(Q)$ zu betrachten.

2. Schritt:

Bestimme alle Nullstellen von $Q(x)$! Da einige Nullstellen komplex sein können, ergibt sich eine Zerlegung

$$Q(x) = (x-c_1)^{k_1} \cdot \ldots \cdot (x-c_r)^{k_r} \cdot q_1(x)^{s_1} \cdot \ldots \cdot q_l(x)^{s_l},$$

mit reellen Zahlen $c_1, \ldots, c_r$ und unzerlegbaren quadratischen Polynomen $q_1(x), \ldots, q_l(x)$.

Dieser Schritt kann natürlich in der Praxis ein unüberwindbares Hindernis darstellen, denn für $\deg(Q) > 4$ gibt es kein konstruktives Verfahren zur Bestimmung der Nullstellen. Theoretisch existiert die Zerlegung aber!

3. Schritt:

Wenn die Zerlegung von $Q(x)$ in „irreduzible Faktoren“ vom Typ $x - c$ bzw. $q(x)$ gegeben ist, dann kann man $\dfrac{P(x)}{Q(x)}$ als Summe von *Partialbrüchen* schreiben:

$$f(x) = \frac{P(x)}{Q(x)} = \sum_{\varrho=1}^{r} \sum_{i=1}^{k_\varrho} \frac{A_{\varrho,i}}{(x-c_\varrho)^i} + \sum_{\lambda=1}^{l} \sum_{j=1}^{s_\lambda} \frac{B_{\lambda,j}x + C_{\lambda,j}}{q_\lambda(x)^j}.$$

Hier ist eine Andeutung des Beweises für den ersten Teil der Summe:
Es sei $Q(x) = (x-c)^k \cdot g(x)$, mit $g(c) \neq 0$. Setzt man $b := \dfrac{P(c)}{g(c)}$, so ist

$$\begin{aligned} \frac{P(x)}{Q(x)} &= \frac{b \cdot g(x) + (P(x) - b \cdot g(x))}{(x-c)^k \cdot g(x)} \\ &= \frac{b}{(x-c)^k} + \frac{P(x) - b \cdot g(x)}{(x-c)^k \cdot g(x)}. \end{aligned}$$

Da $P(c) - b \cdot g(c) = 0$ ist, muß der Zähler des zweiten Bruches den Faktor $x - c$ enthalten. Es gibt also ein $k' < k$, so daß $P(x) - b \cdot g(x) = (x-c)^{k'} \cdot h(x)$ mit einem geeigneten Polynom $h(x)$ mit $h(c) \neq 0$ ist. Also ist

$$\frac{P(x)}{Q(x)} = \frac{b}{(x-c)^k} + \frac{h(x)}{(x-c)^{k-k'} \cdot g(x)}.$$

Den zweiten Summanden behandelt man nun wieder genauso, und nach endlich vielen Schritten ist man am Ziel.

In der Praxis nimmt man den Satz von der Partialbruchzerlegung einfach zur Kenntnis, setzt die Zähler der Partialbrüche mit unbestimmten Koeffizienten an, multipliziert aus und versucht, über einen Koeffizientenvergleich zu Gleichungen für die gesuchten Koeffizienten zu kommen.

4. Schritt:

Schließlich sucht man nach Stammfunktionen für Funktionen der Art

$$\frac{A}{(x-c)^k} \quad \text{und} \quad \frac{Bx + C}{q(x)^s}.$$

a) $\displaystyle\int \frac{1}{x-c}\, dx = \ln|\, x - c\,|.$

b) $\displaystyle\int \frac{1}{(x-c)^k}\, dx = \frac{1}{1-k}(x-c)^{1-k}$, für $k \geq 2$.

c) Bei Integralen der Form $\displaystyle\int \frac{h(x)}{(x^2 + ax + b)^n}\, dx$ mit affin-linearem $h(x)$ müssen wir zunächst eine Substitution vornehmen. Da wir nur den Fall betrachten, wo der Nenner **keine** reelle Nullstelle besitzt, ist $a^2 - 4b < 0$.

Wir machen den Ansatz $\quad x(t)^2 + a \cdot x(t) + b = c^2(t^2 + 1)$. Dann muß gelten:

$$x(t) = \frac{1}{2} \cdot (-a \pm \sqrt{a^2 - 4b + 4c^2(t^2+1)}).$$

Setzen wir $c := \frac{1}{2}\sqrt{4b - a^2}$, also $a^2 - 4b = -4c^2$, so ergibt sich die Substitution

$$x(t) := ct - \frac{a}{2}, \quad x'(t) = c.$$

Dann ist

$$\int \frac{h(x)}{(x^2 + ax + b)^n}\, dx = \int \frac{c \cdot h(x(t))}{(c^2(t^2+1))^n}\, dt = \frac{1}{c^{2n-1}} \int \frac{h(x(t))}{(t^2+1)^n}\, dt.$$

Der Zähler $h(x(t))$ ist wieder eine affin-lineare Funktion.

Wir müssen also noch die folgenden Integraltypen behandeln:

d) $\displaystyle\int \frac{1}{t^2+1}\,dt = \arctan(t)$.

e) $\displaystyle\int \frac{t}{t^2+1}\,dt = \frac{1}{2}\ln(t^2+1)$.

f) $\displaystyle\int \frac{t}{(t^2+1)^n}\,dt = \frac{1}{2}\int \frac{x'(t)}{x(t)^n}\,dt = \frac{1}{2}\int \frac{1}{x^n}\,dx$, für $x(t) = t^2+1$ (und $n > 1$).

Also ist $\displaystyle\int \frac{t}{(t^2+1)^n}\,dt = -\frac{1}{2(n-1)}\cdot\frac{1}{(t^2+1)^{n-1}}$.

g) Es bleibt der schwierigste Fall, nämlich $\displaystyle\int \frac{1}{(t^2+1)^n}\,dt$.

Behauptung: Es gibt für jedes $n \ge 1$ eine aus elementaren Funktionen zusammengesetzte Funktion F_n mit $F_n(0) = 0$ und $F_n' = \dfrac{1}{(t^2+1)^n}$.

Beweis: Wir führen Induktion nach n.

Natürlich setzen wir $F_1(t) := \arctan(t)$. Aber wie weiter? Wenn es die gesuchten Funktionen gibt, dann gilt:

$$\begin{aligned}
[2n\cdot F_{n+1}(t) - (2n-1)\cdot F_n(t)]' &= \frac{2n}{(t^2+1)^{n+1}} - \frac{2n-1}{(t^2+1)^n} \\
&= \frac{1}{(t^2+1)^n} - \frac{2n(t^2+1-1)}{(t^2+1)^{n+1}} \\
&= \frac{1}{(t^2+1)^n} + t\cdot\frac{(-n)\cdot 2t}{(t^2+1)^{n+1}} \\
&= \left(\frac{t}{(t^2+1)^n}\right)',
\end{aligned}$$

also

$$F_{n+1}(t) = \frac{2n-1}{2n}F_n(t) + \frac{t}{2n(t^2+1)^n} + C,$$

mit einer Integrations-Konstanten C, die wegen $F_n(0) = F_{n+1}(0) = 0$ ebenfalls $= 0$ gesetzt werden sollte. Diese rekursive Definition tut's tatsächlich!

In der Theorie läßt sich also jede rationale Funktion elementar integrieren. In der Praxis dürfte es oft an der Nullstellenbestimmung im Nenner scheitern.

Beispiele:

1. Sei $f(x) := \dfrac{1}{x^3-x^2+x-1} = \dfrac{1}{(x-1)(x^2+1)}$.

   Wir machen den Ansatz $\dfrac{1}{x^3-x^2+x-1} = \dfrac{a}{x-1} + \dfrac{b+cx}{x^2+1}$. Dann muß $(a+c)x^2 + (b-c)x + a - b = 1$ sein, und das führt zu dem

Gleichungssystem $\quad a + c = 0, \qquad b - c = 0 \quad$ und $\quad a - b = 1.$

Also muß $a = \dfrac{1}{2}$ und $b = c = -\dfrac{1}{2}$ sein, d.h. $f(x) = \dfrac{1}{2} \cdot \left[\dfrac{1}{x-1} - \dfrac{x+1}{x^2+1}\right]$. Daraus folgt:

$$\begin{aligned} \int f(x)\,dx &= \frac{1}{2} \cdot \left[\int \frac{1}{x-1}\,dx - \int \frac{1}{x^2+1}\,dx - \int \frac{x}{x^2+1}\,dx\right] \\ &= \frac{1}{2} \cdot \left[\ln|\,x-1\,| - \arctan(x) - \frac{1}{2}\ln(x^2+1)\right]. \end{aligned}$$

2. Sei $f(x) := \dfrac{x^4}{x^3-1} = \dfrac{x(x^3-1)+x}{x^3-1} = x + \dfrac{x}{x^3-1} = x + \dfrac{x}{(x-1)(x^2+x+1)}$.

$$\text{Der Ansatz} \quad \frac{x}{x^3-1} = \frac{a}{x-1} + \frac{b+cx}{x^2+x+1}$$

liefert $a + c = 0, \quad a + b - c = 1$ und $a - b = 0$, also $a = b = \dfrac{1}{3}$ und $c = -\dfrac{1}{3}$.

$$\text{Damit ist} \quad f(x) = x + \frac{1}{3} \cdot \left[\frac{1}{x-1} + \frac{1-x}{x^2+x+1}\right].$$

Das Integral $\displaystyle\int \frac{1-x}{x^2+x+1}\,dx$ behandeln wir mit der Substitution

$$x(t)^2 + x(t) + 1 = c^2(t^2+1), \text{ mit } c = \frac{1}{2}\sqrt{4-1} = \frac{1}{2}\sqrt{3}.$$

Das ergibt $x(t) = \frac{1}{2}(\sqrt{3}t - 1)$. Dann gilt:

$$\begin{aligned} \int \frac{1-x}{x^2+x+1}\,dx &= \frac{\sqrt{3}}{2}\int \frac{1-x(t)}{x(t)^2+x(t)+1}\,dt \\ &= \frac{\sqrt{3}}{4}\int \frac{3-\sqrt{3}t}{\frac{3}{4}(t^2+1)}\,dt \quad = \quad \int \frac{\sqrt{3}-t}{t^2+1}\,dt, \end{aligned}$$

mit $t = \dfrac{2x+1}{\sqrt{3}}$ und $t^2 = \dfrac{1}{3}(4x^2+4x+1)$. Also ist

$$\begin{aligned} \int f(x)\,dx &= \int x\,dx + \frac{1}{3}\int \frac{1}{x-1}\,dx + \frac{1}{\sqrt{3}}\int \frac{1}{t^2+1}\,dt - \frac{1}{3}\int \frac{t}{t^2+1}\,dt \\ &= \frac{1}{2}x^2 + \frac{1}{3}\ln|\,x-1\,| + \frac{1}{\sqrt{3}}\arctan\frac{2x+1}{\sqrt{3}} - \frac{1}{6}\ln\left(\frac{4}{3}(x^2+x+1)\right). \end{aligned}$$

Zum Schluß soll noch eine abstraktere Integrationsmethode erwähnt werden, deren Nutzen erst später offenbar werden wird.

Vertauschung von Limes und Integral

Die Folge von stetigen Funktionen $f_n : [a,b] \to \mathbb{R}$ konvergiere gleichmäßig gegen die Grenzfunktion $f : [a,b] \to \mathbb{R}$. Dann ist

$$\lim_{n\to\infty} \int_a^b f_n(t)\,dt = \int_a^b f(t)\,dt.$$

BEWEIS: Sei $\varepsilon > 0$ vorgegeben.
Dann gibt es ein N, so daß für $n \geq N$ und alle $x \in [a,b]$ gilt:

$$|\, f(x) - f_n(x)\,| < \frac{\varepsilon}{b-a}.$$

Also ist

$$\begin{aligned} |\int_a^b f(t)\,dt - \int_a^b f_n(t)\,dt\,| &= |\int_a^b (f(t) - f_n(t))\,dt\,| \\ &\leq \int_a^b |\, f(t) - f_n(t)\,|\,dt \\ &< \frac{\varepsilon}{b-a} \cdot (b-a) = \varepsilon. \end{aligned}$$

Das ergibt die gewünschte Formel. □

Bemerkung: Die Funktion f ist als Grenzwert einer gleichmäßig konvergenten Folge von stetigen Funktionen wieder stetig. Setzt man nur voraus, daß die f_n Regelfunktionen sind, so kann man zeigen, daß f zumindest auch wieder eine Regelfunktion ist, und der Satz bleibt richtig.

Vertauschung von Limes und Ableitung

Die Folge von stetig differenzierbaren Funktionen $f_n : [a,b] \to \mathbb{R}$ sei punktweise konvergent gegen eine Funktion $f : [a,b] \to \mathbb{R}$. Außerdem sei f_n' gleichmäßig konvergent. Dann ist f stetig differenzierbar und

$$\lim_{n\to\infty} f_n'(x) = f'(x).$$

BEWEIS: Sei $f^* := \lim_{n\to\infty} f_n'$.

Ist $x_0 \in I$ fest gewählt und $x \in I$ ein weiterer Punkt, so ist

$$\begin{aligned} f(x) - f(x_0) &= \lim_{n\to\infty} (f_n(x) - f_n(x_0)) \\ &= \lim_{n\to\infty} \int_{x_0}^x f_n'(t)\,dt \\ &= \int_{x_0}^x f^*(t)\,dt. \end{aligned}$$

Also ist

$$f(x) = f(x_0) + \int_{x_0}^x f^*(t)\,dt$$

eine differenzierbare Funktion mit Ableitung $f'(x) = f^*(x)$. □

Beispiel:

Sei $f_n : [0, 2\pi] \to \mathbb{R}$ definiert durch

$$f_n(x) := \frac{1}{n}\sin(n^2 x).$$

Dann ist $\| f_n - 0 \| = \sup\{| f_n(x) | \ :\in [0, 2\pi] \} \leq \frac{1}{n}$, also (f_n) gleichmäßig konvergent gegen 0.

$f_n'(x) = n \cos(n^2 x)$ oszilliert mit zunehmendem n immer stärker, und mit wachsender Amplitude. Z.B. ist

$$f_1'(\pi) = -1,\ f_2'(\pi) = 2,\ f_3'(\pi) = -3,\ f_4'(\pi) = 4, \ldots .$$

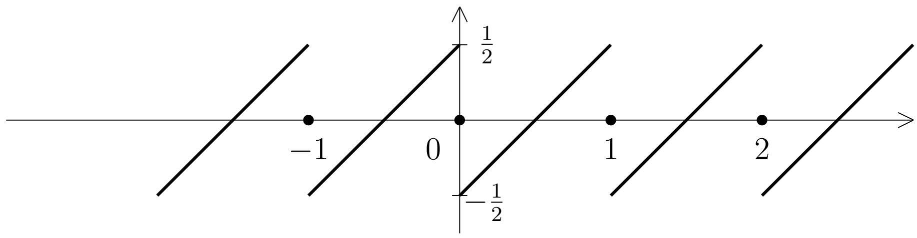
Das bedeutet, daß (f_n') nicht konvergiert. Der Satz über die Vertauschbarkeit von Limes und Ableitung ist hier nicht anwendbar.

§6 Numerische Integration

Manche Integrale können nicht exakt berechnet werden. Dann bleibt nur noch die numerische Auswertung. Wir wollen zwei solche Methoden betrachten.

Sei $H : \mathbb{R} \to \mathbb{R}$ definiert durch

$$H(x) := \begin{cases} x - n - \frac{1}{2} & \text{für } n < x < n+1 \\ 0 & \text{für } x \in \mathbb{Z} \end{cases}$$



Eulersche Summationsformel

Sei $n \in \mathbb{N}$. Ist $f : [0,n] \to \mathbb{R}$ stetig differenzierbar, so ist

$$\sum_{k=0}^{n} f(k) = \int_0^n f(x)\,dx + \frac{1}{2}(f(0) + f(n)) + \int_0^n H(x) f'(x)\,dx.$$

Beweis: Es ist

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

also

$$\begin{aligned} \int_0^n f(x)\,dx &= \sum_{k=0}^{n-1} \int_k^{k+1} f(x)\,dx \\ &= \frac{1}{2}(f(0) + f(n)) + \sum_{k=1}^{n-1} f(k) - \int_0^n H(x) f'(x)\,dx. \end{aligned}$$

Daraus folgt die Behauptung. $\square$

Jetzt wollen wir das Integral $\int_0^n H(x) f'(x)\,dx$ weiter ausrechnen. Allerdings müssen wir voraussetzen, daß f zweimal stetig differenzierbar ist.

Sei $F(x) := \frac{1}{2}((x-n)^2 - (x-n)) = \frac{1}{2}(x^2 - x(2n+1) + (n^2+n))$, für $n \le x < n+1$.

Dann ist F periodisch mit Periode 1, also $F(x+1) = F(x)$, denn:

Ist $n \leq x < n+1$, so ist $n+1 \leq x+1 < n+2$, also

$$F(x+1) = \frac{1}{2}(((x+1)-(n+1))^2 - ((x+1)-(n+1))) = \frac{1}{2}((x-n)^2 - (x-n)) = F(x).$$

Weiter ist $F(n) = 0$ für $n \in \mathbb{Z}$ und $F'(x) = x - n - \frac{1}{2} = H(x)$ für $n < x < n+1$. Also gilt – wenn f zweimal stetig differenzierbar ist – :

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

Summiert man über k, so erhält man:

$$\int_0^n H(x)f'(x)\,dx = -\int_0^n F(x)f''(x)\,dx.$$

Nach dem Mittelwertsatz der Integralrechnung gibt es ein $c \in [0,n]$ mit

$$\begin{aligned}\int_0^n F(x)f''(x)\,dx &= f''(c)\cdot\int_0^n F(x)\,dx \\ &= n\cdot f''(c)\cdot\int_0^1 \frac{x^2-x}{2}\,dx \\ &= n\cdot f''(c)\cdot\left[\frac{x^3}{6}-\frac{x^2}{4}\right]\Big|_0^1 \\ &= n\cdot f''(c)\cdot\left(\frac{1}{6}-\frac{1}{4}\right) \\ &= -\frac{n}{12}\cdot f''(c).\end{aligned}$$

Damit ist

$$\int_0^n H(x)f'(x)\,dx = \frac{n}{12}\cdot f''(c),$$

also

$$\begin{aligned}\int_0^n f(x)\,dx &= \sum_{k=0}^{n} f(k) - \frac{1}{2}(f(0)+f(n)) - \int_0^n H(x)f'(x)\,dx \\ &= \frac{1}{2}(f(0)+f(n)) + \sum_{k=1}^{n-1} f(k) - \frac{n}{12}\cdot f''(c).\end{aligned}$$

Diese Formel wollen wir benutzen, um für eine zweimal stetige Funktion f das Integral $\int_a^b f(x)\,dx$ durch eine Summe von Flächen geeigneter Trapeze zu approximieren:

Dazu unterteilen wir $[a,b]$ in n Teilintervalle der Länge $h := \dfrac{b-a}{n}$. Es sei $x_i := a + i\cdot h$, für $i = 0,1,\ldots,n$. Dann ist

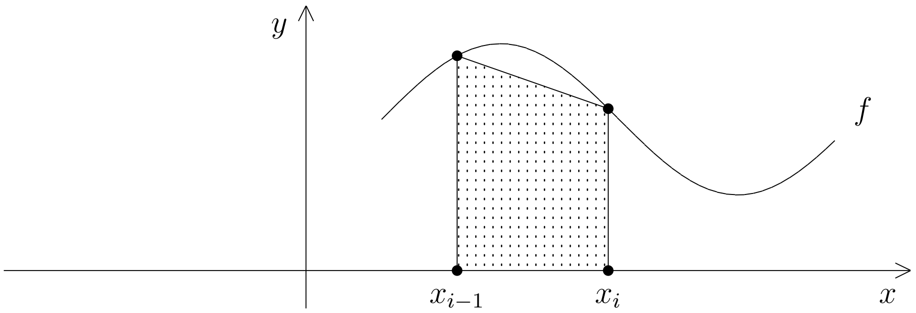
$$a = x_0 < x_1 < x_2 < \ldots < x_n = b,$$

und $x_i - x_{i-1} = h$ für $i = 1, \ldots, n$.

Das Trapez mit den Ecken $(x_{i-1}, 0)$, $(x_i, 0)$, $(x_i, f(x_i))$ und $(x_{i-1}, f(x_{i-1}))$ hat die Fläche

$$F_i := \frac{h}{2}(f(x_{i-1}) + f(x_i)),$$

und $T := \sum_{i=1}^{n} F_i$ kann als Approximation für $\int_a^b f(x)\,dx$ benutzt werden.



Es ist

$$\begin{aligned}
T = T(h) &= F_1 + F_2 + \cdots + F_n \\
&= h \cdot \left[\frac{f(x_0) + f(x_1)}{2} + \frac{f(x_1) + f(x_2)}{2} + \cdots + \frac{f(x_{n-1}) + f(x_n)}{2} \right] \\
&= h \cdot \left[\frac{f(a) + f(b)}{2} + \sum_{k=1}^{n-1} f(a + h \cdot k) \right].
\end{aligned}$$

Trapezregel

Sei $f : [a, b] \to \mathbb{R}$ $2\times$ *stetig differenzierbar,* $|\, f''(x) \,| \le K$ *auf* $[a, b]$, $h := \frac{b-a}{n}$.
Dann ist

$$\left| \int_a^b f(x)\,dx - T(h) \right| \le \frac{nh^3}{12} \cdot K.$$

Beweis: Sei $\varphi : [a, b] \to \mathbb{R}$ definiert durch $\varphi(t) := \dfrac{1}{h}(t - a)$. Dann ist $\varphi'(t) = \frac{1}{h}$, und φ bildet $[a, b]$ bijektiv auf $[0, n]$ ab.

Nun sei $g := f \circ \varphi^{-1} : [0, n] \to \mathbb{R}$. Dann ist

$$\begin{aligned}
g(t) &= f(a + ht), \\
g'(t) &= h \cdot f'(a + ht) \\
\text{und} \quad g''(t) &= h^2 \cdot f''(a + ht), \\
\text{sowie} \quad f(x) &= g\left(\frac{x - a}{h}\right).
\end{aligned}$$

Daraus folgt.

$$\int_a^b f(x)\,dx = h \cdot \int_a^b g(\varphi(x))\varphi'(x)\,dx$$

$$
\begin{aligned}
&= h \cdot \int_0^n g(t)\,dt \\
&= h \cdot \left[\frac{1}{2}(g(0)+g(n)) + \sum_{k=1}^{n-1} g(k) - \frac{n}{12} g''(c)\right] \\
&= h \cdot \left[\frac{1}{2}(f(a)+f(b)) + \sum_{k=1}^{n-1} f(a+hk) - \frac{nh^2}{12} f''(a+hc)\right] \\
&= T(h) - \frac{nh^3}{12} f''(a+hc),
\end{aligned}
$$

mit $0 < c < n$, also $a < a + hc < a + nh = b$. Das liefert die gewünschte Abschätzung. □

Beispiel:

Sei $f(x) := \dfrac{1}{x}$, $a = 1$ und $b = 2$. Dann ist

$$\int_a^b f(x)\,dx = \int_1^2 \frac{1}{x}\,dx = \ln(2) - \ln(1) = \ln(2).$$

Wir wollen diesen Wert mit Hilfe der Trapezregel näherungsweise berechnen. Dazu sei $n = 5$, also $h = \dfrac{b-a}{n} = \dfrac{1}{5}$. Das führt zu der Unterteilung

$$x_0 = 1,\ x_1 = \frac{6}{5},\ x_2 = \frac{7}{5},\ x_3 = \frac{8}{5},\ x_4 = \frac{9}{5} \text{ und } x_5 = 2.$$

Für den Näherungswert ergibt sich somit:

$$
\begin{aligned}
T(h) &= \frac{b-a}{2n}\left[f(x_0) + f(x_n) + 2 \cdot \sum_{k=1}^{n-1} f(x_k)\right] \\
&= \frac{1}{10}\left[1 + \frac{1}{2} + 2 \cdot (\frac{5}{6} + \frac{5}{7} + \frac{5}{8} + \frac{5}{9})\right] \\
&\approx \frac{1}{10}[1.5 + 2 \cdot (0.8333 + 0.7143 + 0.6250 + 0.5556)] \\
&= \frac{1}{10}[1.5000 + 5.4564] \;=\; 0.69564.
\end{aligned}
$$

Für die Fehlerabschätzung brauchen wir $f''(x) = \dfrac{2}{x^3}$. Auf dem Intervall $[1,2]$ ist $1 \le x^3 \le 8$, also $|\,f''(x)\,| \le 2$. Das ergibt:

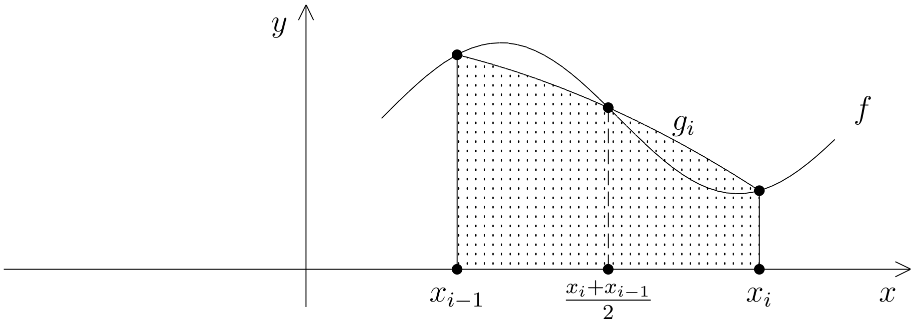
$$|\ln(2) - T(h)\,| \le \frac{nh^3}{12} \cdot 2 = \frac{1}{6 \cdot 25} = 0.0066\ldots < 0.01,$$

also $0.685 < \ln(2) < 0.706$.

Der Taschenrechner liefert $\ln(2) = 0.6931471\ldots$. Unsere Fehlerabschätzung ist also in Ordnung, aber das Ergebnis der Approximation ist noch relativ ungenau.

Wir versuchen es nun mit einer besseren Approximation.

Dazu konstruieren wir zu je zwei Punkten x_{i-1}, x_i ein quadratisches Polynom g_i, das in den drei Punkten x_{i-1}, x_i und $\dfrac{x_{i-1}+x_i}{2}$ jeweils mit f übereinstimmt[2] und dann ersetzen wir das Integral über f durch das viel leichter zu berechnende Integral über g_i.



Die quadratische Interpolation von f wird in der Literatur meist durch Rückgriff auf allgemeine Interpolationssätze erledigt. Es geht aber viel einfacher. Wir brauchen nämlich nicht die Funktion g_i, sondern ihr Integral. Dabei kann man sich mit Hilfe der Substitutionsregel auf den Fall zurückziehen, daß das fragliche Intervall symmetrisch zum Nullpunkt liegt.

Gegeben seien also zunächst ein $\varepsilon > 0$ und drei Werte y_1, y_2, y_3. Gesucht ist ein Polynom $g(x) = Ax^2 + Bx + C$ mit

$$g(-\varepsilon) = y_1, \quad g(0) = y_2 \quad \text{und} \quad g(\varepsilon) = y_3,$$

bzw. das Integral

$$\int_{-\varepsilon}^{\varepsilon} g(x)\,dx = \left(\frac{A}{3}x^3 + \frac{B}{2}x^2 + Cx\right)\Big|_{-\varepsilon}^{\varepsilon} = \frac{2A}{3}\varepsilon^3 + 2C\varepsilon.$$

Wie man sieht, wird B garnicht benötigt.

Aus den Bedingungsgleichungen für g erhält man:

$$y_2 = C \quad \text{und} \quad y_1 + y_3 = 2(A\varepsilon^2 + C).$$

Also ist

$$\int_{-\varepsilon}^{\varepsilon} g(x)\,dx = \frac{2\varepsilon}{3}[A\varepsilon^2 + 3C] = \frac{\varepsilon}{3}[2(A\varepsilon^2 + C) + 4C] = \frac{\varepsilon}{3}(y_1 + 4y_2 + y_3).$$

Nun müssen wir dieses Ergebnis auf ein beliebiges Intervall übertragen. Die Abbildung $\varphi(t) := t - \dfrac{a+b}{2}$ bildet $[a, b]$ bijektiv auf $[-\varepsilon, \varepsilon]$ ab, mit $\varepsilon := \dfrac{b-a}{2}$. Also ist

$$\begin{aligned}
\frac{b-a}{6}[g(-\varepsilon) + 4g(0) + g(\varepsilon)] &= \frac{\varepsilon}{3}(y_1 + 4y_2 + y_3) \\
&= \int_{\varphi(a)}^{\varphi(b)} g(x)\,dx \\
&= \int_a^b g(\varphi(t))\varphi'(t)\,dt \\
&= \int_a^b g(t - \frac{a+b}{2})\,dt \\
&= \int_a^b \widetilde{g}(t)\,dt,
\end{aligned}$$

[2]Dieses quadratische Polynom ist eindeutig bestimmt!

wobei $\widetilde{g}(t) := g(t - \dfrac{a+b}{2})$ wieder ein quadratisches Polynom ist, diesmal aber mit

$$\widetilde{g}(a) = g(-\varepsilon) = y_1, \quad \widetilde{g}(\frac{a+b}{2}) = g(0) = y_2 \quad \text{und} \quad \widetilde{g}(b) = g(\varepsilon) = y_3.$$

Wenn nun $y_1 = f(a)$, $y_2 = f(\frac{a+b}{2})$ und $y_3 = f(b)$ ist, so folgt:

$$\int_a^b \widetilde{g}(t)\,dt = \frac{b-a}{6}[f(a) + 4 \cdot f(\frac{a+b}{2}) + f(b)].$$

Das liefert uns die gewünschte Näherungssumme.

Simpsonsche Regel

Sei $f : [a,b] \to \mathbb{R}$ $4\times$ stetig differenzierbar, $|\, f^{(4)}(x)\,| \le K$ auf $[a,b]$, $h := \frac{b-a}{n}$, $x_k := a + kh$ für $k = 0,1,2,\dots,n$ und

$$S(h) := \frac{h}{6} \cdot \left[f(x_0) + f(x_n) + 2 \cdot \sum_{k=1}^{n-1} f(x_k) + 4 \cdot \sum_{k=1}^{n} f(\frac{x_{k-1}+x_k}{2}) \right].$$

Dann ist

$$|\int_a^b f(x)\,dx - S(h)\,| \le \frac{nh^5}{2880} \cdot K.$$

Den Beweis für die Fehlerabschätzung können wir hier nicht erbringen. Die Größe nh^5 stimmt mit $(b-a)h^4$ überein. Und die Simpsonsche Näherungssumme ergibt sich durch

$$\begin{aligned} S(h) &= \sum_{k=1}^{n} \frac{h}{6}[f(x_{k-1} + f(x_k) + 4 \cdot f(\frac{x_{k-1}+x_k}{2})] \\ &= \frac{h}{6} \cdot \left[f(x_0) + f(x_n) + 2 \cdot \sum_{k=1}^{n-1} f(x_k) + 4 \cdot \sum_{k=1}^{n} f(\frac{x_{k-1}+x_k}{2}) \right]. \end{aligned}$$

Wir wenden nun das Simpson-Verfahren auf die Berechnung von $\ln(2)$ an:

Es sei wieder $n = 5$. Dann ist

$$\begin{aligned} S(h) &= \frac{1}{30}\Big[1 + \frac{1}{2} + 2 \cdot (\frac{5}{6} + \frac{5}{7} + \frac{5}{8} + \frac{5}{9}) \\ &\quad + 4 \cdot (\frac{10}{11} + \frac{10}{13} + \frac{10}{15} + \frac{10}{17} + \frac{10}{19})\Big] \\ &\approx \frac{1}{30}[6.9564 + 13.8380] \approx 0.6931466. \end{aligned}$$

Wie gut dieses Ergebnis ist, zeigt die Fehlerabschätzung. Es ist $f^{(4)}(x) = \dfrac{24}{x^5}$, also $|\, f^{(4)}\,| \le 24$. Damit ist

$$|\ln(2) - S(h)\,| \le \left(\frac{1}{5}\right)^4 \cdot \frac{24}{2880} = \frac{24}{1\,800\,000} \approx 1.333\ldots \cdot 10^{-5} < 10^{-4} = 0.0001.$$

Das sagt uns, daß $0.69304 < \ln(2) < 0.69325$ ist. Die ersten drei Stellen nach dem Komma sind gesichert, die 4. Stelle muß zwischen 0 und 2 liegen. Das ist schon ein sehr gutes Ergebnis. Benutzt man eine feinere Unterteilung, so kann man das Ergebnis weiter verbessern.

§7 Lineare Differentialgleichungen mit konstanten Koeffizienten

V und W seien zwei K-Vektorräume, $K = \mathbb{R}$ oder $K = \mathbb{C}$.

Ist $L : V \to W$ eine lineare Abbildung und $w \in W$ ein fester Vektor, so nennt man die Gleichung

$$(S_i) \qquad L(x) = w$$

ein *(inhomogenes) lineares Gleichungssystem* für x.

Die Gleichung

$$(S_h) \qquad L(x) = 0$$

nennt man das *zugehörige homogene Gleichungssystem.*

Die Lösungsmenge von (S_h) ist der Vektorraum $\text{Ker}\,(L)$. Ist dieser Raum endlich-dimensional, so versucht man, eine Basis $\{a_1, \ldots, a_n\}$ des Lösungsraumes zu finden. Der Ausdruck

$$c_1a_1 + c_2a_2 + \cdots + c_na_n$$

mit variablen Koeffizienten $c_1, \ldots, c_n \in K$ wird dann als die *allgemeine Lösung* von (S_h) bezeichnet.

Für $w \in W$ sei nun $\mathcal{L}(w)$ die Lösungsmenge des inhomogenen Systems $L(x) = w$. Sie bildet (für $w \neq 0$) keinen Vektorraum. Ist x_0 eine spezielle Lösung von (S_i) und x eine weitere Lösung, so gilt:

$$L(x - x_0) = L(x) - L(x_0) = w - w = 0, \quad \text{ also } x - x_0 \in \text{Ker}\,(L).$$

Das bedeutet:

$$\mathcal{L}(w) = \{x \in V \mid x = x_0 + v \text{ mit } v \in \text{Ker}\,(L)\,\} = x_0 + \text{Ker}\,(L)$$

ist ein affiner Unterraum von V.

Ist $\text{Ker}\,(L)$ endlich-dimensional mit Basis $\{a_1, \ldots, a_n\}$, so bezeichnet man den Ausdruck

$$x_0 + c_1a_1 + c_2a_2 + \cdots + c_na_n$$

als *allgemeine Lösung* von (S_i).

Beispiele :

1. V endlich-dimensional. Dann ist natürlich auch $\text{Ker}\,(L)$ endlich-dimensional.

2. Sei $\subset \mathbb{R}$ ein Intervall. Betrachte $D : \mathcal{C}^k(I) \to \mathcal{C}^{k-1}(I)$ mit $D(f) := f'$.

   Die Lösungsmenge $\text{Ker}\,(D) = \{f \in \mathcal{C}^k(I) \mid f' = 0\}$ besteht offensichtlich gerade aus den konstanten Funktionen, ist also ein 1-dimensionaler Unterraum, mit Basis $\{1\}$.

   Eine spezielle Lösung f_0 des inhomogenen Systems $D(f) = g$ ist eine Stammfunktion von g. Die allgemeine Lösung hat also die Gestalt

   $$f = f_0 + c = \int g(x)\,dx + c.$$

Ist $0 \le r \le k$, so definiert man $D^r : \mathcal{C}^k(I) \to \mathcal{C}^{k-r}(I)$ durch

$$D^r(f) := \underbrace{D \circ D \circ \ldots \circ D}_{r\text{-mal}}(f) = f^{(r)}.$$

Speziell ist also $D^0(f) = f$ und $D^1(f) = D(f) = f'$. Ist $r \ge 2$, so setzt sich der „Differentialoperator“ D^r aus r linearen Abbildungen zusammen:

$$\mathcal{C}^k(I) \xrightarrow{D} \mathcal{C}^{k-1}(I) \xrightarrow{D} \mathcal{C}^{k-2}(I) \xrightarrow{D} \ldots \xrightarrow{D} \mathcal{C}^{k-r}(I).$$

Da jeder der Räume $\mathcal{C}^{k-r}(I)$ in dem Raum $\mathcal{C}^0(I)$ aller stetigen Funktionen enthalten ist, kann man D^r stets als Abbildung $D^r : \mathcal{C}^k(I) \to \mathcal{C}^0(I)$ auffassen, solange nur $r \le k$ ist.

Ist nun $p(x) = x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0$ ein (normiertes) Polynom vom Grad n, so setzt man

$$p(D) = D^n + a_{n-1}D^{n-1} + \cdots + a_1 D + a_0.$$

$p(D)$ ist eine lineare Abbildung von $\mathcal{C}^n(I)$ nach $\mathcal{C}^0(I)$. Man nennt $p(D)$ einen *linearen Differentialoperator* und $p(x)$ das zugehörige *charakteristische Polynom.*

Aus traditionellen Gründen nennt man generell eine lineare Abbildung $L : V \to W$ auch einen *linearen Operator*, wenn V ein Vektorraum von Funktionen ist. Statt $L(f)$ schreibt man häufig $L[f]$. Auch wir werden im Folgenden diese Schreibweise verwenden.

Definition:

Eine *lineare Differentialgleichung n-ter Ordnung mit konstanten Koeffizienten* ist eine Gleichung der Gestalt

$$y^{(n)} + a_{n-1}y^{(n-1)} + \cdots + a_1 y' + a_0 y = q(x),$$

mit reellen (oder komplexen) Koeffizienten a_i und einer stetigen Funktion q.

Eine *Lösung der Differentialgleichung* über einem (offenen) Intervall I ist eine reell- oder komplexwertige Funktion $f \in \mathcal{C}^n(I)$ mit

$$(D^n + a_{n-1}D^{n-1} + \cdots + a_1 D + a_0)[f] = q.$$

Beispiel:

Wir betrachten eine Differentialgleichung 1. Ordnung:

$$y' + ay = q(x).$$

Den Fall $a = 0$ haben wir schon behandelt, die Lösungsgesamtheit ist dann durch das unbestimmte Integral gegeben:

$$f(x) = \int q(t)\,dt + c.$$

Wir können also voraussetzen, daß $a \ne 0$ ist. Aus den Überlegungen zu Anfang dieses Abschnittes wissen wir, daß wir zunächst die allgemeine Lösung der zugehörigen homogenen Gleichung $y' + ay = 0$ finden müssen.

Auf jeden Fall ist $f(x) \equiv 0$ eine Lösung der homogenen Gleichung. Zu jeder Lösung $f \neq 0$ über einem offenen Intervall $I = (a, b)$ gibt es wenigstens einen Punkt $x_0 \in I$ mit $f(x_0) \neq 0$. Und wegen der Stetigkeit von f gibt es dann sogar ein $\varepsilon > 0$, so daß $U_\varepsilon(x_0) \subset I$ und $f(x) \neq 0$ für $x \in J := U_\varepsilon(x_0)$ ist. Auf dem (ebenfalls offenen) Intervall J ist

$$-a = \frac{f'(x)}{f(x)} = \frac{|\, f\,|'(x)}{|\, f\,|(x)} = (\ln \circ |\, f\,|)'(x).$$

Also muß $(\ln \circ |\, f\,|)(x) = -ax + c$ sein, mit einer geeigneten Konstanten c. Und daraus folgt:

$$|\, f(x)\,| = e^c \cdot e^{-ax}, \qquad \text{also } f(x) = C \cdot e^{-ax},\ C \in \mathbb{R}.$$

Tatsächlich ist jede der Funktionen $f_C(x) := C \cdot e^{-ax}$ sogar eine Lösung der homogenen Gleichung über ganz $\mathbb{R}$. Wir wollen nun zeigen, daß $f(x) = f_C(x)$ auf ganz I gilt.

Zu diesem Zweck setzen wir

$$x_+ := \sup\{x \in I \mid f(x) = f_C(x)\}.$$

Dann ist offensichtlich $x_0 + \varepsilon \leq x_+ \leq b$. Wir nehmen an, es sei $x_+ < b$ und versuchen, aus dieser Aussage einen Widerspruch herzuleiten.

Da aus Stetigkeitsgründen $f(x_+) = f_C(x_+) \neq 0$ ist, gibt es ein $\delta > 0$, so daß $U_\delta(x_+) \subset I$ und $f(x) \neq 0$ für $x \in U_\delta(x_+)$ ist. Wie oben folgt, daß es ein $C^* \in \mathbb{R}$ gibt, so daß $f(x) = f_{C^*}(x)$ für $x \in U_\delta(x_+)$ ist. Für $x < x_+$ und nahe bei x_+ ist nun $f_C(x) = f_{C^*}(x)$. Das ist nur möglich, wenn $C = C^*$ ist. Aber dann ist $f(x) = f_C(x)$ auch noch rechts von x_+, und das ist ein Widerspruch zur Definition von x_+. Also war die Annahme falsch, und es muß $x_+ = b$ sein. Definiert man jetzt $x_- := \inf\{x \in I \mid f(x) = f_C(x)\}$, so folgt analog, daß $x_- = a$ ist. Aber das bedeutet, daß $f(x) = f_C(x)$ auf $I = (a, b)$ ist.

Nun ist klar, daß $\{e^{-ax}\}$ eine Basis des Lösungsraumes der homogenen Gleichung über jedem Intervall $I \subset \mathbb{R}$ ist.

Wir müssen jetzt nur noch eine spezielle („partikuläre") Lösung der inhomogenen Gleichung $y' + ay = q(x)$ finden. Ein bewährtes Verfahren ist die „Variation der Konstanten". Man setzt die zu findende Lösung y_p in der Form

$$y_p(x) = C(x) \cdot e^{-ax}$$

an, also in der Gestalt der allgemeinen Lösung der homogenen Gleichung, aber mit einem von x abhängigen Koeffizienten. Dann erhält man:

$$\begin{aligned} q(x) - a \cdot C(x)e^{-ax} &= y_p'(x) \\ &= (C'(x) - a \cdot C(x)) \cdot e^{-ax}. \end{aligned}$$

Diese Gleichungen sind nur zu erfüllen, wenn $C'(x) = q(x) \cdot e^{ax}$ ist, also

$$C(x) = \int q(t) e^{at}\, dt.$$

Ist etwa $q(x)$ überall definiert, so erhalten wir als Lösungsvorschlag

$$y_p(x) = \left(\int_0^x q(t)e^{at}\,dt \right) \cdot e^{-ax}.$$

Differenzieren zeigt, daß y_p tatsächlich eine Lösung ist.

Wir testen das Verfahren am Beispiel der DGL $y' + 5y = x^2$.

Die allgemeine Lösung der homogenen Gleichung ist $y_h(x) = C \cdot e^{-5x}$. Eine partikuläre Lösung der inhomogenen Gleichung ist

$$y_p(x) = \left(\int_0^x t^2 e^{5t}\,dt \right) \cdot e^{-5x}.$$

Das Integral kann man durch partielle Integration berechnen, oder durch einen Ansatz für die Stammfunktion in der Form $(ax^2 + bx + c)e^{5x}$. Man erhält dann

$$y_p(x) = \frac{1}{5}(x^2 - \frac{2}{5}x + \frac{2}{25}).$$

Bevor wir nun den allgemeinen Fall behandeln, müssen wir uns mit der Frage der Eindeutigkeit der Lösung befassen.

Nullstellen von Funktionen mit beschränktem Wachstum

Sei $I \subset \mathbb{R}$ ein Intervall, $f : I \to \mathbb{C}$ differenzierbar und $|\, f'(t)\,| \le C \cdot |\, f(t)\,|$, mit einer Konstanten $C > 0$.

Ist $f(t_0) = 0$ für ein $t_0 \in I$, so ist $f(t) \equiv 0$.

Beweis: a) Wir betrachten zunächst den Fall: f reell und ≥ 0, also auch f' reell und $f' \le C \cdot f$.

Sei $g(t) := f(t)e^{-Ct}$. Dann ist

$$g'(t) = (f'(t) - C \cdot f(t))e^{-Ct} \le 0,$$

also g monoton fallend.

Außerdem ist $g(t_0) = 0$, denn es ist ja schon $f(t_0) = 0$.

Für $t \ge t_0$ muß also $g(t) \le 0$ sein. Daraus folgt, daß auch $f(t) \le 0$ für $t \ge t_0$ ist, und wegen der Zusatzannahme $f \ge 0$ bedeutet das, daß $f(t) \equiv 0$ für $t \ge t_0$ ist. Und mit einem ähnlichen Argument folgert man, daß das auch für $t \le t_0$ gilt.

b) Ist f beliebig, so setze man $h := f\overline{f}$. Das ist eine reellwertige Funktion ≥ 0, und es ist

$$|\, h'\,| = |\, f'\overline{f} + f\overline{f}'\,| \le 2 \cdot |\, f'\overline{f}\,| \le 2C \cdot |\, f\overline{f}\,| = \text{Konstante} \cdot |\, h\,|.$$

Mit (a) folgt nun, daß $h(x) \equiv 0$ ist, also auch $f(x) \equiv 0$. □

Eindeutigkeit der Lösung

$f_1, f_2 \in \mathcal{C}^n(I)$ seien Lösungen einer DGL

$$y^{(n)} + a_{n-1}y^{(n-1)} + \cdots + a_1 y' + a_0 y = q(x)$$

auf dem Intervall I, und es sei $f_1^{(k)}(t_0) = f_2^{(k)}(t_0)$ für $k = 0, 1, 2, \ldots, n-1$ und ein spezielles $t_0 \in I$.

Dann ist $f_1 = f_2$.

Beweis: Wir verwenden die Funktion $g := \sum_{k=0}^{n-1} |\, f^{(k)}\,|^2 = \sum_{k=0}^{n-1} f^{(k)}\overline{f^{(k)}}$, mit $f := f_2 - f_1$.

Dann ist g differenzierbar, und es gilt:

$$\begin{aligned} g' &= \sum_{k=0}^{n-2}(f^{(k)}\overline{f^{(k+1)}} + f^{(k+1)}\overline{f^{(k)}}) \\ &\quad + f^{(n-1)}\overline{f^{(n)}} + f^{(n)}\overline{f^{(n-1)}}, \end{aligned}$$

wobei $f^{(n)} = -a_{n-1}f^{(n-1)} - \ldots - a_1 f' - a_0 f$ ist.

Da $|\, f^{(k)}\,| \le \sqrt{g}$ ist, für $k = 0, \ldots, n-1$, ist

$$|\, g'\,| \le 2((n-1)g + \sum_{i=0}^{n-1} |\, a_i\,| g) = C \cdot |\, g\,|,$$

mit der Konstanten $C := 2((n-1) + \sum_{i=0}^{n-1} |\, a_i\,|) > 0$.

Außerdem ist $g(t_0) = 0$. Also muß $g(t) \equiv 0$ und damit auch $f(t) \equiv 0$ sein. Aber das bedeutet, daß $f_1 = f_2$ ist. $\square$

Nun folgt:

Die Dimension des Lösungsraumes

Ist $p(x)$ ein normiertes Polynom n-ten Grades und $L := p(D) : \mathcal{C}^n(I) \to \mathcal{C}^0(I)$ der zugehörige Differentialoperator über einem (offenen) Intervall I, so ist

$$\dim_{\mathbb{C}} \mathrm{Ker}\,(L) \le n,$$

d.h., es gibt höchstens n über $\mathbb{C}$ linear unabhängige Lösungen der homogenen Differentialgleichung $L[f] = 0$.

Beweis: Sei $t_0 \in I$ beliebig und $\mathcal{L}(0) := \mathrm{Ker}\,(L)$ der Lösungsraum der homogenen DGL. Die Abbildung $A : \mathcal{L}(0) \to \mathbb{C}^n$ mit

$$A(f) := (f(t_0), f'(t_0), \ldots, f^{(n-1)}(t_0)),$$

die jeder Lösungs-Funktion einen vollständigen Satz von „Anfangsbedingungen“ im Punkt t_0 zuordnet, ist linear, und nach dem Satz über die Eindeutigkeit der Lösungen ist sie auch injektiv. Also ist $\text{Ker}\,(A) = \{0\}$.

Wir nehmen nun an, es gibt $n+1$ linear unabhängige Lösungen $f_1, \ldots, f_{n+1}$ in $\mathcal{L}(0)$. Dann kann man eine lineare Abbildung

$$B : \mathbb{C}^{n+1} \to \mathbb{C}^n$$

definieren, durch

$$B(c_1, \ldots, c_{n+1}) := A(c_1 f_1 + \cdots c_{n+1} f_{n+1}).$$

Ist $B(\mathbf{c}) = \mathbf{0}$, so ist $A(c_1 f_1 + \cdots c_{n+1} f_{n+1}) = \mathbf{0}$. Da $\text{Ker}\,(A) = \{0\}$ ist, folgt:

$$c_1 f_1 + \cdots c_{n+1} f_{n+1} = 0.$$

Und da die $n+1$ Funktionen linear unabhängig sind, muß $c_1 = \ldots = c_{n+1} = 0$ sein. Das bedeutet, daß $\text{Ker}\,(B) = \{\mathbf{0}\}$ ist, also

$$\dim \text{Im}\,(B) = (n+1) - \dim \text{Ker}\,(B) = n+1.$$

Aber das ist unmöglich, weil $\text{Im}\,(B)$ im $\mathbb{C}^n$ enthalten ist. □

Wir werden später sehen, daß sogar die Gleichheit gilt: $\dim(\mathcal{L}(0)) = n$. Um das zu zeigen, reicht es nun aus, n unabhängige Lösungen zu konstruieren.

Wir brauchen noch ein paar Ergebnisse über das Rechnen mit Differentialoperatoren.

Vertauschbarkeit von Differentialoperatoren

Sind p_1, p_2 zwei normierte Polynome, so ist

$$(p_1 \cdot p_2)(D) = p_1(D) \circ p_2(D) = p_2(D) \circ p_1(D).$$

Beweis: Die Formel ist einfach nachzurechnen. Es liegt daran, daß $D^p D^q = D^q D^p = D^{p+q}$ ist, und an der Tatsache, daß die Koeffizienten konstant sind. □

Bei nicht-konstanten Koeffizienten geht's übrigens schief:

Ist $L_1 = \sin(x)D$ und $L_2 = e^{cx}D$, so ist

$$(L_1 \circ L_2)[f] = \sin(x)D[e^{cx}f'] = \sin(x) \cdot (c \cdot e^{cx}f' + e^{cx}f'') = (\sin(x)e^{cx}D^2 + c\sin(x)e^{cx}D)[f]$$

und

$$(L_2 \circ L_1)[f] = e^{cx}D[\sin(x)f'] = e^{cx}(\cos(x)f' + \sin(x)f'') = (\sin(x)e^{cx}D^2 + \cos(x)e^{cx}D)[f].$$

$p(D)[f]$ in speziellen Fällen

1. *Für $\lambda \in \mathbb{C}$ ist $p(D)[e^{\lambda x}] = p(\lambda) \cdot e^{\lambda x}$.*
2. *Ist $f \in \mathcal{C}^k(I)$ beliebig und $\lambda \in \mathbb{C}$, so ist*
$$(D-\lambda)^k[f \cdot e^{\lambda x}] = f^{(k)} \cdot e^{\lambda x}.$$
3. *Ist $f \neq 0$ ein* **Polynom** *vom Grad r und $\lambda \in \mathbb{C}$* **keine** *Nullstelle von $p(x)$, so gibt es ein Polynom g vom Grad r, so daß gilt:*
$$p(D)[f(x) \cdot e^{\lambda x}] = g(x) \cdot e^{\lambda x}.$$
4. *Ist $f = g + \mathbf{j}h$, $p(x)$ ein Polynom mit* **reellen** *Koeffizienten und $p(D)[f] = 0$, so ist auch $p(D)[g] = 0$ und $p(D)[h] = 0$.*

Beweis: 1) Es ist $D^k(e^{\lambda x}) = \lambda^k \cdot e^{\lambda x}$. Diese Regel überträgt sich sofort auf Differentialoperatoren der Gestalt

$$D^n + a_{n-1}D^{n-1} + \cdots + a_1 D + a_0.$$

2) Es ist

$$\begin{aligned} (D-\lambda)[f \cdot e^{\lambda x}] &= f' \cdot e^{\lambda x} + f \cdot \lambda \cdot e^{\lambda x} - \lambda \cdot f \cdot e^{\lambda x} \\ &= f' \cdot e^{\lambda x}. \end{aligned}$$

Ein trivialer Induktionsbeweis liefert die Behauptung.

3) Man kann das Polynom $p(x)$ nach Potenzen von $x - \lambda$ entwickeln:

$$p(x) = b_n(x-\lambda)^n + b_{n-1}(x-\lambda)^{n-1} + \cdots + b_1(x-\lambda) + b_0.$$

Dabei ist $b_n = 1$ und $b_0 = p(\lambda) \neq 0$. Wegen der Vertauschbarkeit von Differentialoperatoren ist dann auch

$$p(D) = b_n(D-\lambda)^n + b_{n-1}(D-\lambda)^{n-1} + \cdots + b_1(D-\lambda) + b_0.$$

Nun folgt:

$$\begin{aligned} p(D)[f(x) \cdot e^{\lambda x}] &= \sum_{k=0}^{n} b_k (D-\lambda)^k [f(x) \cdot e^{\lambda x}] \\ &= \sum_{k=0}^{n} b_k f^{(k)}(x) \cdot e^{\lambda x}. \end{aligned}$$

$f^{(k)}(x)$ ist ein Polynom vom Grad $\leq r - k$. Der Term x^r taucht nur in dem Summanden $b_0 \cdot f^{(0)}(x) = b_0 \cdot f(x)$ auf, und weil $b_0 \neq 0$ ist, hat das Polynom

$$g(x) := \sum_{k=0}^{n} b_k f^{(k)}(x)$$

den Grad r.

4) Ist $L := p(D)$, so ist $L[g + \mathbf{j}h] = L[g] + \mathbf{j}L[h]$, mit Realteil $L[g]$ und Imaginärteil $L[h]$. Ist nun $L[g + \mathbf{j}h] = 0$, so muß $L[g] = 0$ und $L[h] = 0$ sein. □

Die Unabhängigkeit von $e^{\lambda_1 x}, \ldots, e^{\lambda_r x}$.

Die Zahlen $\lambda_1, \ldots, \lambda_r \in \mathbb{C}$ seien paarweise verschieden, und $f_1(x), \ldots, f_r(x)$ seien Polynome mit

$$f_1(x) \cdot e^{\lambda_1 x} + \cdots + f_r(x) \cdot e^{\lambda_r x} \equiv 0.$$

Dann ist $f_1 = \ldots = f_r = 0$.

BEWEIS: Wir führen Induktion nach r.

Ist $f_1(x) \cdot e^{\lambda_1 x} \equiv 0$, so muß $f_1 = 0$ sein, weil $e^{\lambda_1 x} > 0$ für jedes x ist.

Sei nun $r \geq 2$, und die Behauptung für $r - 1$ bewiesen. Es sei

$$f_1(x) \cdot e^{\lambda_1 x} + \cdots + f_r(x) \cdot e^{\lambda_r x} \equiv 0.$$

Wir wenden den Differentialoperator $(D - \lambda_r)^k$ auf die Gleichung an, und zwar mit einem $k > \text{grad}(f_r(x))$. Es ist

$$\begin{array}{rrcl} & (D - \lambda_r)^k[f_r(x) \cdot e^{\lambda_r x}] & = & f_r^{(k)}(x) \cdot e^{\lambda_r x} = 0 \\ \text{und} & (D - \lambda_r)^k[f_i(x) \cdot e^{\lambda_i x}] & = & g_i(x) \cdot e^{\lambda_i x}, \text{ für } i = 1, \ldots, r-1, \end{array}$$

mit Polynomen $g_i(x)$ mit $\text{grad}(g_i) = \text{grad}(f_i)$ für $i = 1, \ldots, r - 1$. Also ist

$$g_1(x) \cdot e^{\lambda_1 x} + \cdots + g_{r-1}(x) \cdot e^{\lambda_{r-1} x} \equiv 0.$$

Nach Induktionsvoraussetzung muß dann $g_1 = \ldots = g_{r-1} = 0$ sein, also auch $f_1 = \ldots = f_{r-1} = 0$. Dann ist selbstverständlich auch $f_r = 0$. □

Jetzt können wir ein „Fundamentalsystem" von Lösungen der DGL $p(D)[f] = 0$ angeben, d.h. eine Basis des Lösungsraumes.

Das Fundamentalsystem von $p(D)[f] = 0$.

Sei $p(x)$ das charakteristische Polynom der homogenen DGL

$$y^{(n)} + a_{n-1}y^{(n-1)} + \cdots + a_1 y' + a_0 = 0.$$

Sind $\lambda_1, \ldots, \lambda_r$ die paarweise verschiedenen Nullstellen von $p(x)$ in $\mathbb{C}$, mit Vielfachheiten $k_1, \ldots, k_r$, so bilden die Funktionen

$$x^\nu \cdot e^{\lambda_i x}, \text{ mit } i = 1, \ldots, r \text{ und } \nu = 0, \ldots, k_i - 1,$$

ein Fundamentalsystem von Lösungen der Differentialgleichung.

Insbesondere hat der Lösungsraum die Dimension $k_1 + \cdots + k_r = n$ über $\mathbb{C}$.

Beweis: 1) Wir zeigen zunächst, daß die angegebenen Funktionen tatsächlich Lösungen sind. Ist λ eine Nullstelle von $p(x)$ mit der Vielfachheit k, so gibt es ein Polynom $q(x)$ vom Grad $n-k$, so daß $p(x) = q(x)\cdot(x-\lambda)^k$ ist. Und wegen der Vertauschbarkeit der Differentialoperatoren ist dann auch

$$p(D) = q(D) \circ (D-\lambda)^k.$$

Aber weil $\lambda = \lambda_i$ für ein $i \in \{1, \ldots, r\}$ ist, folgt

$$\begin{aligned} p(D)[x^\nu e^{\lambda_i x}] &= q(D) \circ (D-\lambda_i)^{k_i}[x^\nu \cdot e^{\lambda_i x}] \\ &= q(D)[(x^\nu)^{(k_i)} \cdot e^{\lambda_i x}] = 0 \end{aligned}$$

für $\nu = 0, \ldots, k_i - 1$.

2) Nun müssen wir noch sehen, daß die Lösungen linear unabhängig sind. Da wir die Funktionen mit zwei Indizes versehen haben, müssen wir das auch bei den Koeffizienten tun:

Sei $\sum_{i=1}^{r} \sum_{\nu=0}^{k_i-1} c_{i,\nu} x^\nu e^{\lambda_i x} = 0$. Dann setzen wir $f_i(x) := \sum_{\nu=0}^{k_i-1} c_{i,\nu} x^\nu$, für $i = 1, \ldots, r$. Das sind alles Polynome, und wir haben

$$f_1(x) \cdot e^{\lambda_1 x} + \cdots + f_r(x) \cdot e^{\lambda_r x} \equiv 0.$$

Also muß $f_1 = \ldots = f_r = 0$ sein. Aber ein Polynom ist nur dann das Nullpolynom, wenn alle seine Koeffizienten verschwinden. Damit ist $c_{i,\nu} = 0$ für alle i und ν, und die Lösungsfunktionen sind linear unabhängig. □

Bemerkung: Wir sind eigentlich an reellen Lösungen interessiert. Sind alle Nullstellen $\lambda_1, \ldots, \lambda_r$ des charakteristischen Polynoms reell, so ist alles in Ordnung. Hat $p(x)$ reelle Koeffizienten und ist $\lambda = \alpha + \mathbf{j}\beta$ eine komplexe Nullstelle, so ist auch $\overline{\lambda}$ eine Nullstelle, und es gilt:

$$\begin{aligned} e^{\lambda x} &= e^{\alpha x}(\cos(\beta x) + \mathbf{j}\sin(\beta x)), \\ e^{\bar{\lambda} x} &= e^{\alpha x}(\cos(\beta x) - \mathbf{j}\sin(\beta x)). \end{aligned}$$

Also können wir die komplexen Lösungen $x^\nu e^{\lambda x}$ und $x^\nu e^{\overline{\lambda} x}$ durch die linear unabhängigen reellen Lösungen

$$x^\nu e^{\alpha x}\cos(\beta x) \text{ und } x^\nu e^{\alpha x}\sin(\beta x)$$

ersetzen.

Beispiel:

Wir betrachten die homogene lineare DGL 2. Ordnung

$$y'' + 2ay' + by = 0,$$

mit reellen Koeffizienten und der Anfangsbedingung $y(0) = 1$ und $y'(0) = 0$.

Das charakteristische Polynom ist $p(x) = x^2 + 2ax + b$. Die Nullstellen der Gleichung $p(x) = 0$ sind gegeben durch

$$x = \frac{-2a \pm \sqrt{4a^2 - 4b}}{2} = -a \pm \sqrt{\Delta},$$

mit $\Delta := a^2 - b$.

1. $\Delta > 0$. (starke Dämpfung)

   Dann ist $p(x) = (x - \lambda_1)(x - \lambda_2)$, mit $\{\lambda_1, \lambda_2\} = \{-a \pm \sqrt{\Delta}\}$. Die allgemeine Lösung hat die Gestalt

   $$y(x) = c_1 e^{\lambda_1 x} + c_2 e^{\lambda_2 x}.$$

   Dann ist $y'(x) = c_1\lambda_1 e^{\lambda_1 x} + c_2\lambda_2 e^{\lambda_2 x}$, und das Einsetzen der Anfangsbedingungen ergibt

   $$c_1 + c_2 = 1 \text{ und } c_1\lambda_1 + c_2\lambda_2 = 0.$$

   Also ist $c_2 = 1 - c_1$ und $c_1\lambda_1 + (1 - c_1)\lambda_2 = 0$ Die gesuchte Lösung hat deshalb die Gestalt

   $$y(x) = \frac{1}{\lambda_2 - \lambda_1}(\lambda_2 e^{\lambda_1 x} - \lambda_1 e^{\lambda_2 x}).$$

2. $\Delta = 0$. (kritische Dämpfung)

   Dann besitzt $p(x)$ die zweifache reelle Nullstelle $-a$. Also hat die allgemeine Lösung die Gestalt

   $$y(x) = (c_1 + c_2 x)e^{-ax}.$$

   Wegen $y'(x) = (c_2 - a(c_1 + c_2 x))e^{-ax}$ ergeben die Anfangsbedingungen:

   $$c_1 = 1 \text{ und } c_2 - ac_1 = 0, \text{ also } c_2 = a.$$

   Die gesuchte Lösung hat die Gestalt

   $$y(x) = (1 + ax)e^{-ax}.$$

3. $\Delta < 0$. (schwache Dämpfung)

   Setzt man $\omega := \sqrt{-\Delta}$ und $\lambda := -a + \mathbf{j}\omega$, so hat $p(x)$ die komplexen Nullstellen λ und $\bar{\lambda}$. In diesem Fall hat die allgemeine Lösung die Gestalt

   $$y(x) = e^{-ax}(c_1 \cos(\omega x) + c_2 \sin(\omega x)).$$

   Dann ist

   $$\begin{aligned} y'(x) &= -ae^{-ax}(c_1\cos(\omega x) + c_2\sin(\omega x)) \\ &\quad + e^{-ax}(-\omega c_1 \sin(\omega x) + \omega c_2 \cos(\omega x)) \\ &= e^{-ax}[(c_2\omega - ac_1)\cos(\omega x) - (ac_2 + \omega c_1)\sin(\omega x)]. \end{aligned}$$

   Die Randbedingungen ergeben

   $$c_1 = 1 \text{ und } c_2\omega - ac_1 = 0, \text{ also } c_2 = \frac{a}{\omega}.$$

Das führt zu der Lösung

$$y(x) = e^{-ax}(\cos(\omega x) + \frac{a}{\omega}\sin(\omega x)).$$

Man kann sie umformen zu einer gedämpften Schwingung

$$y(x) = Ae^{-ax}\sin(\omega x + \varphi).$$

Es bleibt noch das Problem, eine partikuläre Lösung der inhomogenen DGL

$$y^{(n)} + a_{n-1}y^{(n-1)} + \cdots + a_1 y + a_0 = q(x)$$

zu finden.

Man kann auch in diesem allgemeinen Fall mit der Methode der Variation der Konstanten arbeiten, aber in der Praxis ist das oft schwer durchführbar. In gewissen Fällen kann man stattdessen die partikuläre Lösung durch einen *direkten Ansatz* finden. Wir werden uns hier auf solche Fälle beschränken.

Partikuläre Lösung durch Ansatz

Sei $p(x)$ ein normiertes Polynom (mit reellen Koeffizienten), $L := p(D)$. Die DGL $L[y] = f(x)e^{\lambda x}$ mit einem Polynom $f(x)$ vom Grad s besitzt eine partikuläre Lösung $y_p(x) = g(x)e^{\lambda x}$, mit einem Polynom $g(x)$ und folgenden Eigenschaften:

1. *Ist λ* **keine** *Nullstelle des charakteristischen Polynoms $p(x)$, so hat auch $g(x)$ den Grad s.*
2. *Ist λ eine m–fache Nullstelle von $p(x)$, so hat $g(x)$ den Grad $m + s$.*

Beweis: Sei λ_k eine Nullstelle des charakteristischen Polynoms. Dann enthält $p(D)$ einen Faktor der Gestalt $(D - \lambda_k)$. Nun ist

$$(D - \lambda_k)[g(x)e^{\lambda x}] = (g'(x) + \lambda \cdot g(x) - \lambda_k \cdot g(x)) \cdot e^{\lambda x}.$$

Ist $\lambda = \lambda_k$, so wähle man $g(x)$ so, daß $g'(x) = f(x)$ ist. Der Grad von g muß dann $s + 1$ betragen. Ist λ sogar eine m–fache Nullstelle, so iteriert man das Verfahren.

Ist $\lambda \neq \lambda_k$, so muß man $g(x)$ so wählen, daß

$$g'(x) + (\lambda - \lambda_k) \cdot g(x) = f(x)$$

ist. Zur Abkürzung setzen wir $\mu := \lambda - \lambda_k$. Ist $f(x) = \sum_{i=0}^{s} b_i x^i$, so setzen wir $g(x)$ mit unbestimmten Koeffizienten an:

$$g(x) = \sum_{i=0}^{s} a_i x^i.$$

Dann ist

$$g'(x) + \mu g(x) = (a_1 + \mu a_0) + (2a_2 + \mu a_1)x + \cdots + (sa_s + \mu a_{s-1})x^{s-1} + \mu a_s x^s.$$

Damit $g'(x) + \mu \cdot g(x) = f(x)$ ist, muß also gelten:

$$\begin{aligned} \mu \cdot a_s &= b_s, \\ s \cdot a_s + \mu \cdot a_{s-1} &= b_{s-1}, \\ &\vdots \\ 2 \cdot a_2 + \mu \cdot a_1 &= b_1 \\ \text{und} \quad a_1 + \mu \cdot a_0 &= b_0. \end{aligned}$$

Dieses lineare Gleichungssystem kann man sehr leicht nach $a_s, a_{s-1}, \ldots, a_1, a_0$ auflösen.

Ist

$$L = (D - \lambda_1)^{m_1} \circ (D - \lambda_2)^{m_2} \circ \ldots \circ (D - \lambda_k)^{m_k},$$

so wende man das gewonnene Ergebnis mehrfach an. $\square$

Beispiele :

1. Zur DGL $y'' - y = (3 + 2x + x^2)e^x$ gehört das charakteristische Polynom $p(x) = x^2 - 1 = (x-1)(x+1)$. Also ist $\{e^x, e^{-x}\}$ ein Fundamentalsystem für die homogene Gleichung.

   Die Inhomogenität hat die Gestalt $f(x)e^x$, mit dem quadratischen Polynom $f(x) = 3+2x+x^2$. Da 1 eine einfache Nullstelle des charakteristischen Polynoms ist, gibt es eine partikuläre Lösung der Gestalt $y_p(x) = g(x)e^x$, mit einem Polynom $g(x)$ vom Grad $2 + 1 = 3$. Um g zu bestimmen, gehen wir wie im obigen Beweis vor.

   Zunächst suchen wir ein Polynom $h(x)$ mit $h' = f$. Offensichtlich ist

   $$(3x + x^2 + \frac{1}{3}x^3)' = 3 + 2x + x^2.$$

   Setzen wir $h(x) := 3x + x^2 + \frac{1}{3}x^3$, so ist

   $$(D - 1)[h(x)e^x] = f(x)e^x.$$

   Nun müssen wir noch ein Polynom $g(x) = a_0 + a_1x + a_2x^2 + a_3x^3$ finden, so daß $(D+1)[g(x)e^x] = h(x)e^x$ ist, also $g'(x)+\mu\cdot g(x) = 3x+x^2+\frac{1}{3}x^3$, mit $\mu = 1-(-1) = 2$. Das führt auf das Gleichungssystem

   $$\begin{aligned} \frac{1}{3} &= 2 \cdot a_3 \\ 1 &= 2 \cdot a_2 + 3 \cdot a_3 \\ 3 &= 2 \cdot a_1 + 2 \cdot a_2 \\ 0 &= 2 \cdot a_0 + 1 \cdot a_1 \end{aligned}$$

   für die Koeffizienten a_i. Die Auflösung ergibt

   $$g(x) = -\frac{5}{8} + \frac{5}{4}x + \frac{1}{4}x^2 + \frac{1}{6}x^3.$$

   Einsetzen zeigt, daß $g(x) \cdot e^x$ tatsächlich die DGL löst:

Ist nämlich $y_p(x) := g(x) \cdot e^x = (-\frac{5}{8} + \frac{5}{4}x + \frac{1}{4}x^2 + \frac{1}{6}x^3)e^x$, so ist

$$\begin{aligned} y_p'(x) &= y_p(x) + (\frac{5}{4} + \frac{1}{2}x + \frac{1}{2}x^2)e^x \\ \text{und } y_p''(x) &= y_p(x) + 2(\frac{5}{4} + \frac{1}{2}x + \frac{1}{2}x^2)e^x + (\frac{1}{2} + x)e^x \\ &= y_p(x) + (3 + 2x + x^2)e^x, \end{aligned}$$

also $y_p''(x) - y_p(x) = (3 + 2x + x^2)e^x$.

2. In der Praxis wird man wohl nicht wie oben in mehreren Schritten vorgehen, sondern man wird sofort $g(x)$ mit unbestimmten Koeffizienten anzusetzen und in die DGL einsetzen. Ein Koeffizientenvergleich führt dann zum gewünschten Ergebnis. Wir wollen auch diese Methode an einem Beispiel ausprobieren.

   Das charakteristische Polynom der DGL $y''' + 3y'' + 3y' + y = xe^{-x}$ ist

   $$p(x) = x^3 + 3x^2 + 3x + 1 = (x + 1)^3.$$

   $x = -1$ ist dreifache Nullstelle von $p(x)$, das bedeutet, daß die drei Funktionen

   $$f_1(x) = e^{-x}, \quad f_2(x) = xe^{-x} \quad \text{und} \quad f_3(x) = x^2e^{-x}$$

   ein Fundamentalsystem bilden.

   Die Inhomogenität hat die Form $f(x)e^{\lambda x}$ mit $f(x) = x$ und $\lambda = -1$. Da λ dreifache Nullstelle des charakteristischen Polynoms ist, machen wir für die partikuläre Lösung den Ansatz

   $$y_p(x) = g(x)e^{-x}, \qquad \text{mit } g(x) = ax^4 + bx^3 + cx^2 + dx + e.$$

   Dann gilt:

   $$\begin{aligned} y_p'(x) &= (g'(x) - g(x))e^{-x}, \\ y_p''(x) &= (g''(x) - 2g'(x) + g(x))e^{-x} \\ \text{und} \quad y_p'''(x) &= (g'''(x) - 3g''(x) + 3g'(x) - g(x))e^{-x}, \end{aligned}$$

   also $y_p'''(x) + 3y_p''(x) + 3y_p'(x) + y_p(x) = g'''(x)e^{-x}$.

   Nun ist $g'(x) = 4ax^3+3bx^2+2cx+d$, $g''(x) = 12ax^2+6bx+2c$ und $g'''(x) = 24ax+6b$. Setzt man das ein, so erhält man die Bedingungsgleichung

   $$24ax + 6b = x.$$

   Der Koeffizientenvergleich liefert $a = \dfrac{1}{24}$ und $b = 0$. Die Koeffizienten c und d tauchen gar nicht auf und können daher $= 0$ gesetzt werden. Als partikuläre Lösung ergibt sich

   $$y_p(x) = \frac{1}{24}x^4 \cdot e^{-x}.$$

   Die Probe zeigt, daß das richtig ist.

   Die allgemeine Lösung lautet nun

   $$y(x) = (c_1 + c_2x + c_3x^2 + \frac{1}{24}x^4)e^{-x}.$$

Hat die rechte Seite z.B. die Gestalt $f(x)e^{\lambda x}\cos(\omega x)$ oder $f(x)e^{\lambda x}\sin(\omega x)$, so muß man es mit einer anderen Methode versuchen:

Zusammengesetzte Inhomogenitäten

1. *Sind die Funktionen y_i Lösungen der DGLn $p(D)[y] = f_i(x)$, so ist $y := c_1 y_1 + \cdots + c_r y_r$ Lösung der DGL $p(D)[y] = c_1 f_1(x) + \cdots + c_r f_r(x)$.*

2. *Hat $p(x)$ reelle Koeffizienten und ist $Y = y_1 + \mathbf{j}y_2$ Lösung der DGL $p(D)[y] = f_1(x) + \mathbf{j}f_2(x)$, so ist y_1 Lösung der DGL $p(D)[y] = f_1$.*

Beweis: 1) ist trivial.

2) Hat $p(x)$ reelle Koeffizienten, so ist $p(D)[y_1 + \mathbf{j}y_2] = p(D)[y_1] + \mathbf{j}p(D)[y_2]$. Ist dieser Ausdruck $= f_1(x) + \mathbf{j}f_2(x)$, so muß insbesondere $p(D)[y_1] = f_1(x)$ sein. □

Beispiel:

Wir betrachten die DGL $y'' + 4y = x^2 + 5\cos(2x)$. Das charakteristische Polynom zur DGL ist

$$p(x) = x^2 + 4 = (x + 2\mathbf{j})(x - 2\mathbf{j}).$$

Also bilden die beiden Funktionen $f_1(x) = \cos(2x)$ und $f_2(x) = \sin(2x)$ ein Fundamentalsystem für die homogene Gleichung.

Für eine partikuläre Lösung der inhomogenen DGL suchen wir jeweils partikuläre Lösungen für die beiden DGLn $y'' + 4y = x^2$ und $y'' + 4y = 5\cos(2x)$. Die Summe ist dann die gesuchte spezielle Lösung.

Für die DGL $y'' + 4y = x^2$ machen wir den Ansatz $y_p(x) = ax^2 + bx + c$. Einsetzen und anschließender Koeffizientenvergleich liefern $y_p(x) = \frac{1}{4}x^2 - \frac{1}{8}$.

Bei der DGL $y'' + 4y = 5\cos(2x)$ gehen wir folgendermaßen vor:

Statt für die vorliegende DGL suchen wir zunächst nach einer partikulären Lösung für die DGL $y'' + 4y = 5e^{\mathbf{j}2x}$. Deren Realteil ist dann die gesuchte partikuläre Lösung der Ausgangsgleichung. Da $2\mathbf{j}$ Nullstelle des charakteristischen Polynoms ist, machen wir den Ansatz $y_p(x) = cxe^{2\mathbf{j}x}$. Einsetzen in die DGL und anschließender Koeffizientenvergleich liefern $y_p(x) = -\frac{5}{4}\mathbf{j}xe^{2\mathbf{j}x}$. Damit ist $y_p(x) = \operatorname{Re}(y_p(x)) = \frac{5}{4}x\sin(2x)$ eine spezielle Lösung der DGL $y'' + 4y = 5\cos(2x)$.

Die allgemeine Lösung der DGL $y'' + 4y = x^2 + 5\cos(2x)$ ist nun insgesamt

$$y(x) = \frac{1}{4}x^2 - \frac{1}{8} + \frac{5}{4}x\sin(2x) + c_1\cos(2x) + c_2\sin(2x).$$