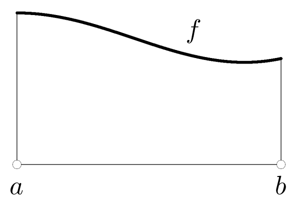
Kapitel III
Ableitung und Integral

Ph: Vous entendez cela, et vous savez le latin, sans doute.
J: Oui; mais faites comme si je ne le savais pas. Expliquez-moi ce que cela veut dire.
Le Bourgeois Gentilhomme

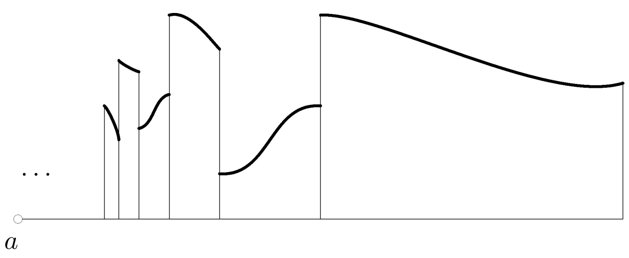
Wir erklären den Kalkül der Differential- und Integralrechnung, Berechnung der Steigung einer Kurve, der Geschwindigkeit einer ungleichförmigen Bewegung, und Berechnung des Flächeninhalts. Auch erklären und diskutieren wir die klassischen Funktionen, Sinus, Kosinus, Logarithmus und Exponentialfunktion.

§ 1. Das Riemann-Integral

Wir setzen uns die Aufgabe, die Größe der Fläche zu bestimmen, die durch eine Funktion $f : [a, b] \to \mathbb{R}$ über dem kompakten Intervall $[a, b]$ begrenzt wird. (In diesem Abschnitt sind alle Intervallgrenzen endlich.)



Für stetige Funktionen ist diese Aufgabe nicht schwer zu lösen. Es genügt dann natürlich auch, wenn f auf den Teilintervallen einer geeigneten Zerlegung des Intervalls stetig ist. Auch wenn f etwa nur an den Punkten $a + \frac{1}{n}$, $n \in \mathbb{N}$, springt, wüßten wir uns wohl zu helfen, solange f wenigstens beschränkt bleibt.



Definieren wir aber eine Funktion $f : [0,1] \to \mathbb{R}$, durch

$$f(x) = \begin{cases} 1 & \text{falls } x \in \mathbb{Q} = \text{Körper der rationalen Zahlen,} \\ 0 & \text{falls } x \notin \mathbb{Q}, \text{ d.h. } x \text{ ist irrational,} \end{cases}$$

so versagt jedenfalls die Anschauung. Um nicht an eine etwas willkürliche Konstruktion gebunden zu sein, die vielleicht mit anderen möglichen Konstruktionen in undurchschaubarer Konkurrenz steht, werden wir die Klasse von Funktionen, die man integrieren kann, sowie die Eigenschaften, die das Integral — also der gesuchte Flächeninhalt — dann haben soll, zunächst axiomatisch beschreiben. Dann werden wir für eine große Klasse von Funktionen zeigen, daß die Axiome auf genau eine Weise zu erfüllen sind. Lernen wir später dann allgemeiner anwendbare Integrale kennen, so bleibt das hier Bewiesene doch immer bestehen.

Definition des Integrals

(1.1) Integrable Funktionen. *Zu jedem kompakten Intervall* $[a, b]$ *sei eine Menge* F_a^b *von Funktionen* $f : [a, b] \to \mathbb{R}$ *ausgezeichnet. Sie heißen auf* $[a, b]$ **integrabel**. *Es gelte:*

(i) *Ist* $f|(a,b)$ *konstant, so ist* $f \in F_a^b$.
(ii) *Ist* $a \leq b \leq c$, *so ist eine Funktion* $f : [a,c] \to \mathbb{R}$ *genau dann in* F_a^c, *wenn* $f \,|\, [a,b] \in F_a^b$ *und* $f \,|\, [b,c] \in F_b^c$.

Insbesondere sind also alle Treppenfunktionen stets integrabel.

(1.2) Das Integral. *Das Integral für die demgemäß gegebene Familie integrabler Funktionen ist eine Familie von Abbildungen*

$$\int_a^b : F_a^b \to \mathbb{R}, \quad a \leq b \in \mathbb{R},$$

die also jeder Funktion $f \in F_a^b$ *eine reelle Zahl, ihr* **Integral**

$$\int_a^b f =: \int_a^b f(x)\,dx \in \mathbb{R}$$

so zuordnet, daß gilt:
(i) **Intervall-Additivität:** *Ist* $f \in F_a^c$ *und* $a \leq b \leq c$, *so ist*

$$\int_a^b f + \int_b^c f = \int_a^c f.$$

(ii) **Monotonie:** *Sind* $f, g \in F_a^b$ *und ist* $f \leq g$, *so ist*

$$\int_a^b f \leq \int_a^b g.$$

(iii) **Normierung:** *Ist* $f = c$ *konstant, so ist* $\int_a^b f = c \cdot (b-a)$.

Kurz: Ein Integral ist ein (endlich) additives, monotones, normiertes Funktional (d.h. eine Abbildung einer Menge von Funktionen nach $\mathbb{R}$). Daß der Flächeninhalt unter einer Funktion f jedenfalls diese Eigenschaften haben sollte, ist eine plausible Forderung.

Das "endlich" steht in Klammern, weil man auch stärkere Integralbegriffe betrachtet, wo das entsprechende Axiom auch für unendliche Zerlegungen gilt. Aber für eine große Klasse von Funktionen, zum Beispiel die stückweise stetigen oder monotonen, ist das Integral durch diese Axiome schon festgelegt. Eine weitere naheliegende Forderung, die aber für die von uns betrachteten Funktionen aus den aufgestellten Axiomen schon folgt, ist die der

Linearität: *Sind* f, g *auf* $[a, b]$ *integrabel und* $\lambda, \mu \in \mathbb{R}$, *so ist auch* $\lambda f + \mu g$ *auf* $[a, b]$ *integrabel, und es gilt:*

$$\int_a^b (\lambda f + \mu g) = \lambda \int_a^b f + \mu \int_a^b g.$$

Übrigens ist die Variable x in der Bezeichnung des Integrals unwesentlich, also $\int_a^b f(x)\,dx = \int_a^b f(t)\,dt \ldots$.

Wir beginnen mit einfachen Folgerungen aus den Axiomen: Nach (1.2, i) ist $\int_a^a f = \int_a^a f + \int_a^a f$, also

$$\int_a^a f = 0. \tag{1.3}$$ □

(1.4) Satz. *Ist* $f : [a, b] \to \mathbb{R}$ *integrabel und* $|f| \le M$, *so ist die Funktion*

$$F : [a, b] \to \mathbb{R}, \quad x \mapsto \int_a^x f$$

stetig, sogar **Lipschitz-stetig** *mit Konstante* M, *das heißt*

$$|F(x) - F(p)| \le M \cdot |x - p|.$$

Beweis: Weil die Behauptung in x und p symmetrisch ist, dürfen wir $p \leq x$ annehmen und setzen $h = x - p$, dann ist

$$|F(x) - F(p)| = \Big| \int\limits_a^{p+h} f - \int\limits_a^{p} f \Big| = \Big| \int\limits_p^{p+h} f \Big|,$$

wegen Additivität. Aus $-M \leq f \leq M$ folgt wegen der Monotonie

$$-Mh = \int\limits_p^{p+h} (-M) \leq \int\limits_p^{p+h} f \leq Mh, \quad \text{also} \quad \Big| \int\limits_p^{p+h} f \Big| \leq Mh.$$

In der Stetigkeitsdefinition kann man zu $\varepsilon > 0$ setzen $\delta \leq \frac{\varepsilon}{M}$. □

Man nennt a die **untere** und b die **obere Grenze** des Integrals $\int_a^b$. Das Integral beschränkter Funktionen ist also stetig als Funktion der oberen Grenze. Der entsprechende Satz gilt natürlich auch für das Integral als Funktion der unteren Grenze. So eignet sich das Integral — wenn man es einmal besitzt — zur Konstruktion neuer bedeutsamer Funktionen, zum Beispiel

$$\log(x) := \int\limits_1^x t^{-1}\, dt,$$

$$\arctan(x) := \int\limits_0^x \frac{1}{1+t^2}\, dt.$$

Der Satz erlaubt eine kleine Verschärfung des Normierungsaxioms.

(1.5) Normierung. *Ist $f|(a,b) = c$ konstant, so ist $\int_a^b f = c\cdot(b-a)$.*

Also auf die Werte an den Endpunkten kommt es nicht an.

Beweis: Es ist f beschränkt, und für $a < b$ und $\varepsilon < \frac{b-a}{2}$ ist

$$\int\limits_a^b f = \int\limits_a^{a+\varepsilon} f + \int\limits_{a+\varepsilon}^{b-\varepsilon} c + \int\limits_{b-\varepsilon}^{b} f.$$

Der mittlere Summand ist $(b-a-2\varepsilon)\cdot c$. Grenzübergang $\varepsilon \to 0$ liefert nach (1.4) $\int_a^b f = 0 + (b-a)\cdot c + 0$. $\square$

Jetzt sehen wir sofort, wie man Treppenfunktionen, die ja nach unseren Forderungen jedenfalls integrabel sind, zu integrieren hat.

(1.6) Satz. *Für Treppenfunktionen sind die Integralaxiome auf genau eine Weise zu erfüllen, nämlich: Ist* $a = z_0 \le z_1 \le \cdots \le z_n = b$ *eine Zerlegung, und* $f(x) = c_k$ *für* $z_{k-1} < x < z_k$, *so ist*

(i) $\int\limits_a^b f = \sum\limits_{k=1}^n c_k(z_k - z_{k-1})$.

Dieses Integral ist auch linear: Für $\lambda, \mu \in \mathbb{R}$ *ist*

(ii) $\int\limits_a^b (\lambda f + \mu g) = \lambda \int\limits_a^b f + \mu \int\limits_a^b g$.

Beweis: Die Formel (i) folgt unmittelbar aus der Additivität und verbesserten Normierung (1.5) des Integrals. Aus dieser Formel folgt dann (ii) unmittelbar. Um zu zeigen, daß die Formel (i) umgekehrt ein Integral definiert, das alle unsere Forderungen erfüllt, muß man nur klarmachen, daß die rechte Seite von (i) unabhängig von der Zerlegung ist, mit der f beschrieben ist. Weil zwei Zerlegungen stets eine gemeinsame Verfeinerung haben, muß man sich nur davon überzeugen, daß die Formel keinen anderen Wert liefert, wenn man einen neuen Zerlegungspunkt $z_k < z < z_{k+1}$ einfügt. Nun freilich, dann wird nur der Summand

$$c_{k+1}(z_{k+1} - z_k) \quad \text{durch} \quad c_{k+1}(z - z_k) + c_{k+1}(z_{k+1} - z)$$

ersetzt. Zum Beweis der Additivität darf man dann annehmen, daß der mittlere Punkt b ein Zerlegungspunkt ist; alle anderen Axiome folgen direkt aus (i). $\square$

Treppenfunktionen können wir jetzt integrieren, aber das ist noch nichts rechtes, wir denken nicht wie ein Computer. Sei nun wieder ein $f \in F_a^b$ gegeben und die F_a^b dabei wie immer so erklärt, daß die Axiome erfüllt sind. Sind dann φ, ψ Treppenfunktionen, so wissen wir:

$$\varphi \leq f \leq \psi \implies \int_a^b \varphi \leq \int_a^b f \leq \int_a^b \psi.$$

Daher erklären wir für beliebige beschränkte Funktionen f auf Intervallen $[a,b]$ das

Oberintegral: $\int_a^{*b} f := \inf\{\int_a^b \psi \mid f \leq \psi \text{ und } \psi \in T_a^b\}$,

Unterintegral: $\int_{*a}^{b} f := \sup\{\int_a^b \varphi \mid \varphi \leq f \text{ und } \varphi \in T_a^b\}$.

Dabei ist T_a^b der reelle Vektorraum der Treppenfunktionen auf $[a,b]$. Eigentlich sollte es "Riemann-Oberintegral" ... heißen, aber weil in diesem Band kein anderes Integral auftritt, lassen wir es so. Nach dem Gesagten gilt nun ersichtlich:

(1.7) Satz. *Ist eine beschränkte Funktion $f : [a,b] \to \mathbb{R}$ integrabel, so ist*

$$\int_* f \leq \int f \leq \int^* f. \qquad \square$$

Allgemein ist für eine beschränkte Funktion f jedenfalls

$$\int_* f \leq \int^* f,$$

denn sind $\varphi \leq f \leq \psi$ Treppenfunktionen, so ist $\int \varphi \leq \int \psi$. Da das für alle Treppenfunktionen $\varphi \leq f$ gilt, folgt $\int_* f \leq \int \psi$, und da das für alle Treppenfunktionen $\psi \geq f$ gilt, folgt $\int_* f \leq \int^* f$.

Definition. *Eine beschränkte Funktion $f : [a,b] \to \mathbb{R}$ auf einem kompakten Intervall heißt* **Riemann-integrabel**, *falls*

$$\int_{*a}^{b} f = \int_a^{*b} f.$$

Die Definition verlangt mit anderen Worten: Zu jedem $\varepsilon > 0$ gibt es Treppenfunktionen $\varphi \le f \le \psi$ mit $\int_a^b \psi - \int_a^b \varphi < \varepsilon$, oder noch anders gesagt: Es gibt Treppenfunktionen $\varphi_n \le f \le \psi_n$ mit $\lim_{n\to\infty} \int_a^b (\psi_n - \varphi_n) = 0$. Die Definition ist nach dem Satz (1.7) gemodelt und führt unsere Bemühung um die Konstruktion eines Integrals zum (vorläufigen) Ziel:

(1.8) Satz. *Sei F_a^b die Menge der Riemann-integrablen Funktionen auf $[a, b]$. Dann erfüllen die F_a^b die Axiome für integrable Funktionen, und ein Integral läßt sich für diese Funktionenfamilie auf genau eine Weise erklären, nämlich durch*

$$\int_a^b f := \int_a^{b\,*} f = \int_{a\,*}^b f.$$

Dieses Integral ist auch linear.

Wir werden künftig nur lineare Integrale betrachten und statt "Riemann-integrabel" einfach **integrabel** sagen. Die Eindeutigkeit des Integrals ergibt sich aus (1.7). Daß die Axiome so tatsächlich erfüllt sind, werden wir gleich sehen. Zunächst bemerken wir, daß etwas gewonnen ist:

(1.9) Satz. *Stetige Funktionen und monotone Funktionen sind integrabel.*

Beweis: Ist $f : [a, b] \to \mathbb{R}$ stetig und $\varepsilon > 0$ gegeben, so wähle Treppenfunktionen φ, ψ mit

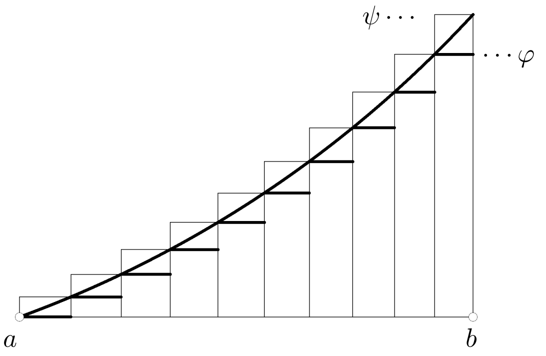
$$\varphi \le f \le \psi, \quad \psi - \varphi < \varepsilon/(b - a),$$

siehe (II, 5.2). Dann ist $\int_a^b (\psi - \varphi) \le \int_a^b \varepsilon/(b - a) = \varepsilon$.

Ist f etwa monoton wachsend und nicht konstant, so wähle eine Zerlegung $a = z_0 \le \cdots \le z_n = b$ mit $z_{k+1} - z_k < \varepsilon/\bigl(f(b) - f(a)\bigr)$.

Dann bestimme φ und ψ durch $\varphi(x) = f(z_k)$, $\psi(x) = f(z_{k+1})$ für $z_k \leq x < z_{k+1}$. Damit ist $\varphi \leq f \leq \psi$, und wir haben:

$$\begin{aligned}\int_a^b (\psi - \varphi) &= \sum_k f(z_{k+1})(z_{k+1} - z_k) - \sum_k f(z_k)(z_{k+1} - z_k) \\ &= \sum_k \big(f(z_{k+1}) - f(z_k)\big)(z_{k+1} - z_k) \\ &< \frac{\varepsilon}{f(b) - f(a)} \cdot \sum_k f(z_{k+1}) - f(z_k) = \varepsilon. \qquad \square\end{aligned}$$



Beweis von (1.8): Offenbar sind Treppenfunktionen integrabel, daher gilt (1.1, i) und (1.2, iii), die Normierung. Wir zeigen (1.1, ii) und (1.2, i), die Additivität: Sei also $a \leq b \leq c$, sei $f : [a, c] \to \mathbb{R}$ gegeben, und es seien $f \mid [a, b]$ und $f \mid [b, c]$ integrabel. Wähle Treppenfunktionen φ_{1n}, ψ_{1n} auf $[a, b]$, sowie φ_{2n}, ψ_{2n} auf $[b, c]$, mit

$$\varphi_{1n} \leq f|[a,b] \leq \psi_{1n}, \quad \int_a^b (\psi_{1n} - \varphi_{1n}) \to 0,$$

$$\varphi_{2n} \leq f|[b,c] \leq \psi_{2n}, \quad \int_b^c (\psi_{2n} - \varphi_{2n}) \to 0,$$

für $n \to \infty$, und $\varphi_{1n}(b) = \varphi_{2n}(b)$, $\psi_{1n}(b) = \psi_{2n}(b)$. Dann setzen sich φ_{1n} mit φ_{2n} und ψ_{1n} mit ψ_{2n} zu Treppenfunktionen φ_n und ψ_n auf $[a,c]$ zusammen, und es gilt

$$\varphi_n \le f \le \psi_n, \quad \int_a^c (\psi_n - \varphi_n) = \int_a^b (\psi_{1n} - \varphi_{1n}) + \int_b^c (\psi_{2n} - \varphi_{2n}) \to 0$$

für $n \to \infty$. Folglich ist f integrabel auf $[a,c]$ und für $n \to \infty$ gilt:

$$\begin{array}{ccccc} \int_a^c \varphi_n & = & \int_a^b \varphi_{1n} & + & \int_b^c \varphi_{2n} \\ \downarrow & & \downarrow & & \downarrow \qquad n \to \infty. \\ \int_a^c f & & \int_a^b f & & \int_b^c f \end{array}$$

Also $\int_a^c f = \int_a^b f + \int_b^c f$. Ist umgekehrt f integrabel, so kann man $\varphi_n \le f \le \psi_n$, $\int_a^c (\psi_n - \varphi_n) \to 0$ wie eben vorgeben, gewinnt daraus φ_{1n}, ψ_{1n} und φ_{2n}, ψ_{2n} durch Einschränkung auf die Teilintervalle $[a,b]$ und $[b,c]$, und weil die Summanden in

$$\int_a^b (\psi_{1n} - \varphi_{1n}) + \int_b^c (\psi_{2n} - \varphi_{2n}) = \int_a^c (\psi_n - \varphi_n) \to 0$$

nicht negativ sind, kann man schließen, daß sie für $n \to \infty$ einzeln gegen 0 gehen. Das zeigt die Additivität. Die Monotonie ist trivial: Ist $f \le g$, so ist $\int_a^b f = \int_{a^*}^b f = \sup\{\int_a^b \varphi \mid \varphi \in T_a^b \text{ und } \varphi \le f\}$ $\le \sup\{\cdots \text{und } \varphi \le g\} = \int_{a^*}^b g = \int_a^b g$.

Schließlich bleibt die Linearität des Riemann-Integrals zu zeigen. Das braucht keine neuen Gedanken. Wie immer seien die φ, ψ und σ, τ Treppenfunktionen. Wir gehen aus von integrablen Funktionen $f, g : [a,b] \to \mathbb{R}$ und Folgen von Treppenfunktionen

$$\varphi_n \le f \le \psi_n, \quad \int_a^b (\psi_n - \varphi_n) \to 0,$$

$$\sigma_n \le g \le \tau_n, \quad \int_a^b (\tau_n - \sigma_n) \to 0,$$

für $n \to \infty$. Insbesondere $\lim_n \int \varphi_n = \int f$, $\lim_n \int \sigma_n = \int g$. Dann folgt $\varphi_n + \sigma_n \le f + g \le \psi_n + \tau_n$, und:

$$\int_a^b \big((\psi_n + \tau_n) - (\varphi_n + \sigma_n)\big) = \int_a^b (\psi_n - \varphi_n) + \int_a^b (\tau_n - \sigma_n) \to 0.$$

Also existiert $\int_a^b (f+g)$ und ist gleich $\lim_n \int (\varphi_n + \sigma_n) = \int f + \int g$.

Ganz ähnlich schließt man für λf, wenn $\lambda \ge 0$. Ist $\lambda < 0$, so beachte man die Umkehrung der Ungleichungen: Aus $\varphi_n \le f \le \psi_n$ und $\int(\psi_n - \varphi_n) \to 0$ folgt $\lambda\psi_n \le \lambda f \le \lambda\varphi_n$ und $\int(\lambda\varphi_n - \lambda\psi_n) \to 0$. $\square$

Aus der Monotonie folgt die

(1.10) Dreiecksungleichung *für das Integral. Für $a \le b$ gilt:*

$$\left|\int_a^b f\right| \le \int_a^b |f|.$$

Beweis: Es ist ja $-|f| \le f \le |f|$, also $-\int |f| \le \int f \le \int |f|$, und weil $\int |f| \ge 0$, folgt die Behauptung. $\square$

Man sollte sich jedoch vorher vergewissern, daß $|f|$ überhaupt integrabel ist.

(1.11) Bemerkung. *Sind die Funktionen f, g auf $[a, b]$ integrabel, so auch die Funktionen*

$$f_+ := \max(f, 0), \quad f_- := \max(-f, 0),$$

sowie $\max(f, g)$, $\min(f, g)$ *und* $|f|$.

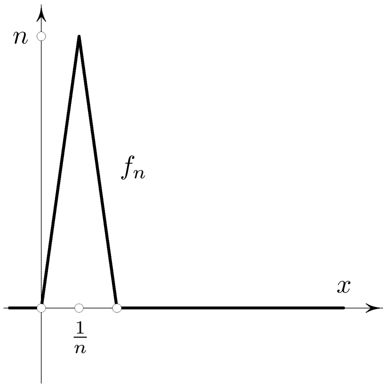
Beweis: Es ist $f_- = f_+ - f$, $|f| = f_+ + f_-$ und $\max(f, g) = \frac{1}{2}(f+g) + \frac{1}{2}|f-g|$, $\min(f, g) = \frac{1}{2}(f+g) - \frac{1}{2}|f-g|$. Wir müssen also nur zeigen, daß f_+ integrabel ist. Hat man Treppenfunktionen φ, ψ mit

$$\varphi \le f \le \psi, \quad \int (\psi - \varphi) < \varepsilon,$$

so folgt $\varphi_+ \le f_+ \le \psi_+$, $\int(\psi_+ - \varphi_+) \le \int(\psi - \varphi) < \varepsilon$, weil $\psi_+ - \varphi_+ \le \psi - \varphi$. □

Das Beispiel der Funktionenfolge (II, 4.1) zeigt, daß für eine konvergente Folge $(f_n) \to f$ von Funktionen auf $[a, b]$ im allgemeinen

$$\lim_{n\to\infty} \int_a^b f_n(x)\,dx \ \neq \int_a^b \big(\lim_{n\to\infty} f_n(x)\big)\,dx.$$



In diesem Fall ist ja $\lim_{n\to\infty} f_n = 0$ und $\int_a^b f_n(x)\,dx = 1$. Gut geht es aber bei gleichmäßiger Konvergenz.

(1.12) Satz *(Stetigkeit des Integrals als Funktional): Konvergiert die Funktionenfolge (f_n) auf $[a, b]$ gleichmäßig gegen f und sind alle f_n integrabel, so ist auch f integrabel und*

$$\int_a^b f \ = \int_a^b \lim_{n\to\infty} f_n \ = \ \lim_{n\to\infty} \int_a^b f_n.$$

Beweis: Ist $|f_n - f| < \varepsilon$ und sind φ, ψ Treppenfunktionen mit $\varphi \le f_n \le \psi$ und $\int(\psi - \varphi) < \varepsilon$, so folgt $f_n - \varepsilon < f < f_n + \varepsilon$, also

$$\varphi - \varepsilon < f < \psi + \varepsilon, \qquad \int \big((\psi + \varepsilon) - (\varphi - \varepsilon)\big) < \varepsilon\big(2(b-a) + 1\big).$$

Da letzteres mit ε beliebig klein wird, ist f integrabel.
Aus $|f_n - f| < \varepsilon$ folgt nach (1.10)

$$\Big|\int f_n - \int f\Big| = \Big|\int (f_n - f)\Big| \le \int |f_n - f| \le \int \varepsilon = \varepsilon(b-a).$$

Das zeigt $(\int f_n) \to \int f$. □

In der Tat, die letze Abschätzung ist der springende Punkt des Beweises. Sie zeigt etwas mehr als das Behauptete: Versieht man den Raum F_a^b mit der Supremumsnorm, so ist die Abbildung

$$\int : F_a^b \to \mathbb{R}, \quad f \mapsto \int_a^b f(x)\,dx$$

Lipschitz-stetig mit Konstante $b - a$. Also

$$\Big|\int_a^b f - \int_a^b g\Big| \le (b-a)\,\|f - g\|. \tag{1.13}$$

Die Stetigkeit, um die es hier geht, ist die des Integrals als Abbildung $\int : F_a^b \to \mathbb{R}$, oder wie man sagt, als Funktional. Sie ist wohl zu unterscheiden von der Stetigkeit des Integrals als Funktion der oberen Grenze $F(x) = \int_a^x f(t)\,dt$.

Bisher haben wir $a \le b$ vorausgesetzt. Für $a > b$ setzen wir jetzt:

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

Dann gilt die Additivitätsformel

$$\int_a^c f = \int_a^b f + \int_b^c f$$

für beliebige Lage der Grenzen a, b, c, falls f auf dem Intervall zwischen $\min\{a, b, c\}$ und $\max\{a, b, c\}$ integrabel ist.

Wir schließen mit einem wichtigen Hilfsmittel, das uns im nächsten Abschnitt dienen wird zu zeigen, wie die Integralrechnung mit der Differentialrechnung zusammenhängt.

(1.14) Mittelwertsatz der Integralrechnung. *Sei f eine stetige und p eine integrable Funktion auf $[a, b]$, und sei $p \geq 0$ (oder $p \leq 0$). Dann existiert ein $\xi \in [a, b]$, sodaß*

$$\int_a^b f \cdot p \;=\; f(\xi) \cdot \int_a^b p.$$

Für $p = 1$ folgt insbesondere

$$\int_a^b f \;=\; f(\xi) \cdot (b - a).$$

Beweis: Sei m das Minimum und M das Maximum von f auf $[a, b]$, dann ist $mp \leq f \cdot p \leq Mp$, also $m \int_a^b p \leq \int_a^b f \cdot p \leq M \int_a^b p$. Jetzt gibt es nach dem Zwischenwertsatz, angewendet auf die Funktion $x \mapsto f(x) \cdot \int_a^b p$, ein $\xi \in [a, b]$ mit $f(\xi) \int_a^b p = \int_a^b f \cdot p$. □

Oft schreibt man $(b - a) =: h$ und $\xi = a + \vartheta h$, $0 \leq \vartheta \leq 1$, dann steht da:

$$\int_a^{a+h} f \cdot p \;=\; f(a + \vartheta h) \int_a^{a+h} p.$$

Diese Version bleibt für $b < a$, also $h < 0$, gültig und ist daher oft vorzuziehen.

Man kann leicht zeigen, daß ein Produkt integrabler Funktionen stets integrabel ist. Wir verweisen auf die Übungen, um nicht durch

Wiederholung zu ermüden. Man beginnt am besten mit der Zerlegung

$$f = f_+ - f_-.$$

§ 2. Die Ableitung

Wir betrachten in diesem Abschnitt Funktionen $f : D \to \mathbb{R}$, deren Definitionsgebiet D ein Intervall ist, das nicht nur einen Punkt enthält. Die **lineare** Funktion mit Steigung a ist die Abbildung $\mathbb{R} \to \mathbb{R}$, $x \mapsto ax$. Addiert man noch eine Konstante, so erhält man eine **affine** Funktion $f(x) = ax + b$. Dann ist für jedes p

$$f(p+h) - f(p) = ah.$$

Also die Abbildung $h \mapsto f(p+h) - f(p)$ ist linear. Eine Funktion ist bei p differenzierbar, wenn sie sich dort durch eine affine Funktion gut approximieren läßt:

Definition. *Eine Funktion* $f : D \to \mathbb{R}$ *heißt bei* $p \in D$ **differenzierbar** *mit* **Ableitung** *(oder* **Differentialquotient***)*

$$a = f'(p) = \frac{df}{dx}(p),$$

wenn gilt:

$$f(p+h) - f(p) = \Phi(h) \cdot h,$$

und Φ *ist stetig am Nullpunkt, mit* $\Phi(0) = a$.

Natürlich besagt die Definition einfach

$$\Phi(h) = \frac{f(p+h) - f(p)}{h}$$

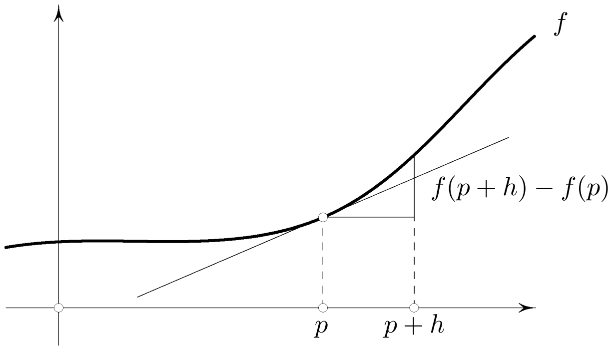
für $h \neq 0$, dies ist der **Differenzenquotient** bei p, und

$$f'(p) = \lim_{h \to 0} \Phi(h).$$

Setzen wir $\varphi(h) := \Phi(h) - a$, so steht da:

$$f(p+h) - f(p) = ah + \varphi(h) \cdot h, \qquad \lim_{h\to 0} \varphi(h) = 0,$$

also $h \mapsto ah$ ist eine lineare Approximation von $h \mapsto f(p+h) - f(p)$, und der Rest $\varphi(h) \cdot h$ verschwindet **von höherer Ordnung** für $h \to 0$, das heißt, wenn man den Rest durch h dividiert, geht der Quotient für $h \to 0$ immer noch gegen 0. Die Funktionen Φ und φ sind auf $\{h \,|\, p+h \in D\}$ definiert.



In etwas pauschalen Betrachtungen schreibt man auch einfach $y = f(x)$ und $f' = \frac{dy}{dx}$, eine von Leibniz eingeführte suggestive Bezeichnung, die uns beim Rechnen leiten kann. Der Differentialquotient $\lim_{h\to 0} \Phi(h)$ ist offenbar eindeutig bestimmt, wenn er existiert. Auch hängt er nur von $f(p+h)$ für kleine h ab, also von f lokal um p. Wir beginnen mit einigen Beispielen:

Eine affine Funktion $f(x) = ax + b$ erfüllt $f(p+h) - f(p) = ah$, hat also an jeder Stelle konstant die Ableitung a. Aus diesem Musterexemplar erhalten wir viele weitere Beispiele durch den

(2.1) Satz *(rationale Operationen). Seien $f, g : D \to \mathbb{R}$ differenzierbar an der Stelle $p \in D$ und $\lambda, \mu \in \mathbb{R}$ Konstanten. Dann sind auch $\lambda f + \mu g$ und $f \cdot g$ und für $f(p) \neq 0$ auch $1/f$ bei p differenzierbar, und es gilt:*

(i) **Linearität**: $(\lambda f + \mu g)' = \lambda \cdot f' + \mu \cdot g'$.
(ii) **Produktregel**: $(f \cdot g)' = f' \cdot g + f \cdot g'$.
(iii) **Quotientenregel**: $(1/f)' = -f'/f^2$.

Beweis: Wir setzen dies gerade so an, wie es in der Definition steht:
(i): $f(p+h) = f(p) + \Phi(h) \cdot h$, $g(p+h) = g(p) + \Psi(h) \cdot h$, und Φ, Ψ seien stetig am Nullpunkt mit den respektiven Ableitungen als Wert. Dann ist folglich

$$(\lambda f + \mu g)(p+h) - (\lambda f + \mu g)(p) = \big(\lambda\Phi(h) + \mu\Psi(h)\big) \cdot h,$$

und $\lambda\Phi + \mu\Psi$ ist stetig am Nullpunkt mit Wert $\lambda f'(p) + \mu g'(p)$.

(ii) $(f \cdot g)(p+h) = (f \cdot g)(p) + \big(f(p)\Psi(h) + \Phi(h)g(p) + \Phi(h)\Psi(h)h\big) \cdot h$

und der lange Term in Klammern geht für $h \to 0$ gegen

$$f(p)g'(p) + f'(p)g(p).$$

(iii) $$\frac{1}{f(p+h)} - \frac{1}{f(p)} = -\frac{f(p+h) - f(p)}{f(p) \cdot f(p+h)} = -\frac{\Phi(h)}{f(p) \cdot f(p+h)} \cdot h,$$

und für $h \to 0$ geht letzterer Bruch gegen $-f'(p)/f^2(p)$. □

Hier haben wir noch ein bißchen gemogelt, wir benutzen die Bemerkung, daß eine bei p differenzierbare Funktion dort insbesondere stetig ist, also demnach $f(p+h)$ in unserem Fall für kleine h nicht verschwindet. In der Tat: $f(p+h) = f(p) + \Phi(h) \cdot h$, und die rechte Seite ist nach Voraussetzung stetig am Nullpunkt. Die Umkehrung gilt nicht, denn die Funktion $x \mapsto |x|$ ist am Nullpunkt stetig, aber sie ist dort nicht differenzierbar.

Aus den Regeln folgt durch Induktion sofort, wie man rationale Funktionen ableitet; für $f(x) = x^n$ ergibt sich nämlich

$$(x^n)' = nx^{n-1},$$

also für ein Polynom erhält man:

$$f(x) = \sum_{k=0}^{n} a_k x^k \Longrightarrow f'(x) = \sum_{k=1}^{n} k a_k x^{k-1}. \tag{2.2}$$

Auch ergibt sich aus (i), (ii) für einen Quotienten $\frac{f}{g} = (\frac{1}{g}) \cdot f$ die allgemeinere

(2.3) Quotientenregel. $\left(\dfrac{f}{g}\right)' = \dfrac{gf' - fg'}{g^2}.$

Wie sich die Ableitung bei Zusammensetzungen von Funktionen verhält, sagt die

(2.4) Kettenregel. *Gegeben seien Funktionen*

$$D \underset{f}{\longrightarrow} B \underset{g}{\longrightarrow} \mathbb{R}.$$

Sei f differenzierbar bei p und g differenzierbar bei $q = f(p)$. Dann ist die Zusammensetzung $g \circ f$ bei p differenzierbar mit der Ableitung

$$(g \circ f)'(p) = g'(q) \cdot f'(p).$$

In salopper Notation schreibt man: Ist $y = f(x)$ und $z = g(y)$, so ist

$$\frac{dz}{dx} = \frac{dz}{dy} \cdot \frac{dy}{dx}.$$

Beweis: Nach bewährtem Muster schreiben wir

$$f(p+h) = f(p) + \Phi(h) \cdot h, \qquad \Phi(0) = f'(p),$$
$$g(q+k) = g(q) + \Psi(k) \cdot k, \qquad \Psi(0) = g'(q),$$

also $gf(p+h) = g\big(f(p)+\Phi(h)\cdot h\big) = g\big(f(p)\big) + \Psi\big(\Phi(h)\cdot h\big)\cdot\Phi(h)\cdot h$, setze für die zweite Gleichung $\Phi(h)\cdot h$ für k ein.

Nun, für $h \to 0$ geht auch $\Phi(h)\cdot h \to 0$, also $\Psi(\Phi(h)\cdot h) \to \Psi(0) = g'(q)$, und $\Phi(h) \to \Phi(0) = f'(p)$. Der ganze Faktor bei h im letzten Term geht also gegen $g'(q)\cdot f'(p)$, und das ist nach Definition der Ableitung die Behauptung. □

Man stellt sich unter dx gern eine verschwindend kleine Verschiebung von p aus in x-Richtung, unter dy die entsprechende Änderung der Funktion, und unter $\frac{dy}{dx}$ dann den Quotienten vor. Das rechtfertigt die Notation, macht die Rechenregeln plausibel und hilft auch, das mathematische Wesen anderswo, zum Beispiel in der Physik, zu interpretieren. Statt "verschwindend klein" sagt man auch wohl "infinitesimal", in mittelalterlichem Vertrauen auf die wissenschaftszeugende Kraft der lateinischen Sprache. Die Differentiale dx, $df\ldots$ selbst werden dabei oft nur etwas schüchtern vorgeführt, als sei es eigentlich gemogelt. Doch haben auch sie eine klare Bedeutung:

Betrachten wir eine an jeder Stelle des Intervalls differenzierbare Funktion $f : D \to \mathbb{R}$. Wir haben dann jedem Punkt $p \in D$ eine lineare Abbildung zugeordnet, nämlich die Abbildung

$$d_pf = df(p) : \mathbb{R} \to \mathbb{R}, \quad h \mapsto f'(p)\cdot h.$$

Dies ist ja eben an der Stelle p die lineare Approximation der Abbildung $h \mapsto f(p+h) - f(p)$. Das **Differential** df von f ordnet so jedem $p \in D$ die lineare Abbildung $\mathbb{R} \to \mathbb{R}$, $h \mapsto f'(p)\cdot h$ zu. Insbesondere die Identität x hat das Differential dx, das natürlich jedem Punkt die identische Abbildung zuordnet:

$$d_px : \mathbb{R} \to \mathbb{R}, \quad h \mapsto h.$$

Und wie jede lineare Abbildung $h \mapsto ah$ von $\mathbb{R}$ nach $\mathbb{R}$ ein Vielfaches der Identität ist (nämlich $a\cdot id$), so ist insbesondere

$$df = f'\cdot dx.$$

Das ist an jeder Stelle p die Gleichung $d_p f : h \mapsto f'(p) \cdot h$ von oben. Im Eindimensionalen sieht das künstlich aus, weil eine lineare Abbildung $\mathbb{R} \to \mathbb{R}$ dasselbe ist wie eine Zahl; sie ist bestimmt durch das Bild der Eins. Aber im Höherdimensionalen wird $d_p f$ auch eine lineare Abbildung höherdimensionaler Räume sein, gegeben durch eine Matrix. Darauf werden wir im dritten Band zurückkommen und die Bedeutung der Differentiale erklären.

Ist $f : D \to \mathbb{R}$ an jedem Punkt $p \in D$ differenzierbar, so liefert f die neue Funktion

$$f' : D \to \mathbb{R}, \quad p \mapsto f'(p).$$

Ist auch sie differenzierbar, so kann man fortfahren und $f'' = (f')'$ bilden, und so fort und induktiv die n**-te Ableitung**

$$f^{[n]} = \left(f^{[n-1]}\right)', \quad f^{[0]} = f,$$

solange die Funktionen eben noch differenzierbar sind. In der Notation von Leibniz schreibt man

$$f^{[n]} = \frac{d^n f}{dx^n}.$$

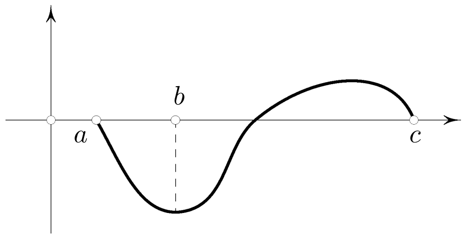
Existiert $f^{[n]}$ und ist stetig, so heißt f entsprechend n**-mal stetig differenzierbar**.

Die Ableitung einer differenzierbaren Funktion muß allerdings nicht differenzierbar sein, zum Beispiel $x \cdot |x|$ ist überall differenzierbar, mit Ableitung $2|x|$, was im Nullpunkt nicht mehr differenzierbar ist.

§ 3. Das lokale Verhalten von Funktionen

Da die Ableitung einer Funktion die Steigung der Tangente an den Graphen ist, ist es anschaulich plausibel, daß eine Funktion auf einem Intervall genau dann konstant ist, wenn ihre Ableitung verschwindet, daß sie genau dann monoton wächst, wenn ihre Ableitung nie negativ ist, und dergleichen mehr. Der Angelpunkt zum wirklichen Beweis solcher Aussagen ist der Mittelwertsatz der Differentialrechnung. Man beginnt nach gefestigter Tradition mit folgendem Spezialfall:

(3.1) Satz von Rolle. *Sei f eine stetige Funktion auf dem kompakten Intervall $[a, c]$, und sie sei im Innern, also auf dem offenen Intervall (a, c) differenzierbar. Ist dann $f(a) = f(c) = 0$, so existiert ein $b \in (a, c)$ mit $f'(b) = 0$.*



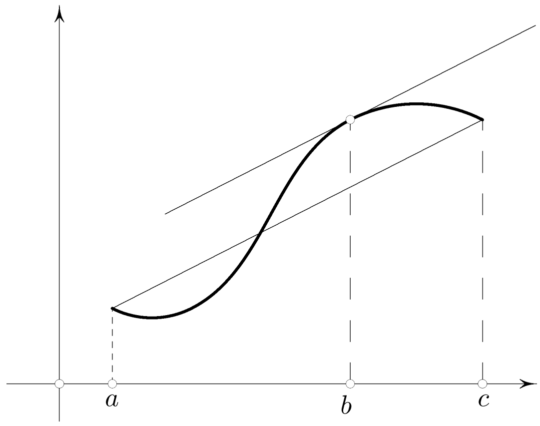
Beweis: Ist f konstant, so ist jedes $b \in (a, c)$ recht. Sonst aber nimmt f sein Maximum oder sein Minimum nicht auf den Randpunkten, also auf einem inneren Punkt b an. Sei also $f(b)$ etwa der minimale Wert von f, der andere Fall ist ebenso zu behandeln. Wir betrachten die Definition der Differenzierbarkeit am Punkt b:

$$f(b + h) - f(b) = \Phi(h) \cdot h, \quad \Phi(0) = f'(b).$$

Die linke Seite ist für kleine $|h|$ definiert und stets nicht negativ. Also muß $\Phi(h)$, wo es nicht verschwindet, dasselbe Vorzeichen wie h haben: $\Phi(h) \geq 0$ für $h > 0$ und $\Phi(h) \leq 0$ für $h < 0$. Aus Stetigkeit muß dann $\Phi(0) = 0$ sein, sonst bliebe es doch lokal um 0 positiv oder negativ. $\square$

(3.2) Mittelwertsatz der Differentialrechnung. *Sei f eine stetige Funktion auf dem kompakten Intervall $[a,c]$, die auf dem Inneren (a,c) differenzierbar ist. Dann existiert ein innerer Punkt $b \in (a,c)$ mit*

$$f(c) - f(a) = f'(b) \cdot (c-a).$$



Beweis: Subtrahiere die lineare Verbindung der Endpunkte, also definiere eine differenzierbare Funktion g durch

$$g(x) := f(x) - \Big(f(a) + (x-a)\frac{f(c)-f(a)}{c-a}\Big).$$

Dann ist $g(a) = g(c) = 0$, also nach Rolle $g'(b) = 0$ für ein b im Innern, und das heißt

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

Wohlgemerkt, die Behauptungen glaubt man sofort, aber wir wollen nach und nach so viel darauf bauen, daß hier alles darauf ankommt, formale Beweise zu führen.

(3.3) Folgerung. *Unter den Voraussetzungen des Mittelwertsatzes gilt:*

(i) f *ist konstant genau wenn* $f' = 0$ *auf* (a, c).

(ii) f *wächst monoton genau wenn* $f' \geq 0$ *auf* (a, c).

(iii) *Ist* $f'(x) > 0$ *für alle* $x \in (a, c)$, *so wächst* f *streng monoton. Entsprechendes gilt für monotones Fallen.*

Beweis: (i): Ist $x > a$ und $f' = 0$, so ist $f(x) - f(a) = f'(\xi) \cdot (x - a) = 0$, $a < \xi < x$.
(ii): Für $a \leq x < y$ ist $f(y) - f(x) = f'(\xi) \cdot (y - x)$ mit $x < \xi < y$. Weil $f'(\xi) \geq 0$ nach Voraussetzung, ist $f(x) \leq f(y)$. Ist umgekehrt kein Differenzenquotient $\frac{f(y)-f(x)}{y-x}$ negativ, so auch kein Differentialquotient.
(iii): Ebenso wie (ii) mit $<$ statt $\leq$. □

Die Funktion $y = x^3$ wächst streng monoton, obwohl $(x^3)' = 3x^2$ im Ursprung verschwindet.

(3.4) Satz über die Umkehrfunktion. *Sei* D *ein Intervall mit mehr als einem Punkt, sei* $p \in D$ *und* $f : D \to \mathbb{R}$ *stetig, streng monoton (d.h. injektiv) und differenzierbar bei* p *mit* $f'(p) \neq 0$. *Dann ist* $f^{-1} : fD \to D$ *bei* $q = f(p)$ *differenzierbar, und*

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

In salopper Notation schreibt man dafür

$$\frac{dx}{dy}(q) = \frac{1}{\frac{dy}{dx}(p)}.$$

Beweis: Für $x \in D$ ist nach Definition der Differenzierbarkeit

$$f(x) = f(p) + \Phi(x)(x - p), \quad \Phi(p) = f'(p) \neq 0,$$

und Φ ist bei p und damit übrigens auf ganz D stetig. Für $x \neq p$ ist auch $\Phi(x) = (f(x) - f(p))/(x-p) \neq 0$. Sei nun $y = f(x)$, $q = f(p)$, dann gilt demnach $y - q = \Phi\bigl(f^{-1}(y)\bigr)\bigl(f^{-1}(y) - f^{-1}(q)\bigr)$, also

$$f^{-1}(y) - f^{-1}(q) = \frac{1}{\Phi\bigl(f^{-1}(y)\bigr)} \cdot (y - q).$$

Wir wissen schon (II, 3.9), daß f^{-1}, also $1/\Phi \circ f^{-1}$ bei q stetig ist, und der Wert dort ist $1/\Phi(p) = 1/f'(p)$. □

Die Voraussetzungen des Satzes sind insbesondere für jedes $p \in D$ erfüllt, wenn f auf ganz D differenzierbar und stets $f' \neq 0$ ist.

Beispiel. Die Funktion $y = x^n$ liefert $y' = nx^{n-1} > 0$ für $x > 0$. Die inverse Funktion ist $x = \sqrt[n]{y} =: y^{\frac{1}{n}}$. Leiten wir die Gleichung $(x^{\frac{1}{n}})^n = x$ nach der Kettenregel ab, so steht da $n(x^{\frac{1}{n}})' \cdot (x^{\frac{1}{n}})^{n-1} = 1$, also

$$(x^{\frac{1}{n}})' = \tfrac{1}{n}\, x^{\frac{1}{n}-1}.$$

Allgemeiner erhält man so für rationales $\alpha = \frac{m}{n}$, daß x^α auf $\{x > 0\}$ eine wohldefinierte streng monotone differenzierbare Funktion ist, mit

$$(x^\alpha)' = \alpha\, x^{\alpha-1}.$$

Wir werden dieses Ergebnis jedoch bald auf ganz anderem Wege für beliebige reelle Zahlen α erhalten.

Jetzt wollen wir das lokale Verhalten einer differenzierbaren Funktion um einen Punkt genauer beschreiben. Sei dazu f auf einem offenen Intervall D gegeben, und $p \in D$. Man nennt p ein **lokales Maximum** von f, wenn p eine Umgebung U in D besitzt, sodaß $f|U$ in p ein Maximum annimmt, also $f(x) \leq f(p)$ für alle $x \in U$. Ist hier $f(x) < f(p)$ für alle $x \neq p$, so heißt das lokale Maximum **isoliert**. Hat $-f$ ein (isoliertes) lokales Maximum, so hat f ein (isoliertes) lokales **Minimum** an der Stelle p. Ein **Extremum** ist ein Maximum oder Minimum. Ähnlich nennen wir f um p **lokal streng monoton** wachsend oder fallend, wenn $f|U$ für eine Umgebung U

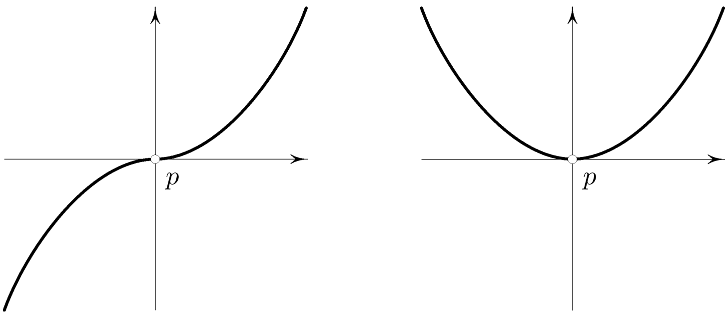
von p diese Eigenschaft hat. Eine beliebig oft differenzierbare Funktion hat stets eine dieser vier Verhaltensweisen, es sei denn, daß alle ihre Ableitungen im Punkt p verschwinden.

(3.5) Lemma. *Sei f differenzierbar auf D und $f'(p) = 0$. Dann gilt:*

(i) *Ist $f'(p+h) \cdot h > 0$ für kleine $h \neq 0$, so ist p ein isoliertes lokales Minimum von f. Dies ist insbesondere der Fall, wenn f' lokal um p streng monoton wächst.*

(ii) *Ist p ein isoliertes lokales Minimum von f', so wächst f lokal um p streng monoton.*

Sieht also f' aus, wie die eine Figur, so f wie die andere.



Beweis: (i): Hier hat $f'(p{+}h)$ gleiches Vorzeichen wie h, also es ist positiv für $h > 0$ und negativ für $h < 0$, also wächst f streng monoton für $h > 0$ und fällt streng monoton für $h < 0$, was insbesondere die Behauptung zeigt.
(ii): Hier ist $f'(p+h) > 0$ für $|h| > 0$, also wächst f lokal streng monoton. $\square$

Der erste Fall liegt insbesondere vor, wenn $f''(p)$ existiert und positiv ist, denn dann ist

$$f'(p+h) = \Phi(h) \cdot h, \quad \Phi(0) > 0,$$

also $f'(p+h) \cdot h = \Phi(h) \cdot h^2 > 0$ für kleine $h \neq 0$.

(3.6) Satz *über das lokale Verhalten. Sei f auf dem offenen Intervall D definiert und $(n-1)$-mal differenzierbar, $n \geq 2$. Für ein $p \in D$ existiere auch $f^{[n]}(p)$, und es sei*

$$f^{[n]}(p) \neq 0, \quad f^{[k]}(p) = 0 \quad \text{für} \quad k < n.$$

Dann hat man einen der folgenden vier Fälle:

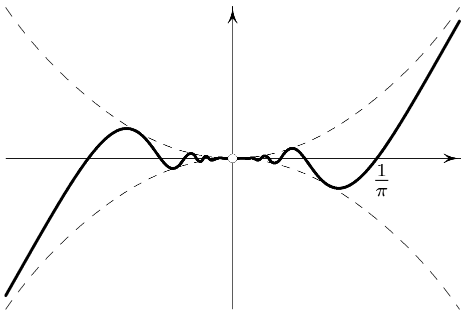
n gerade,	$f^{[n]}(p) > 0$	$\Longrightarrow$	*p ist ein isoliertes lokales Minimum.*
n gerade,	$f^{[n]}(p) < 0$	$\Longrightarrow$	*p ist ein isoliertes lokales Maximum.*
n ungerade,	$f^{[n]}(p) > 0$	$\Longrightarrow$	*f wächst lokal um p streng monoton.*
n ungerade,	$f^{[n]}(p) < 0$	$\Longrightarrow$	*f fällt lokal um p streng monoton.*

Beweis: Den Fall $f^{[n]}(p) < 0$ führt man durch Multiplikation mit -1 auf den Fall $f^{[n]}(p) > 0$ zurück, den wir jetzt betrachten. Nach Voraussetzung ist $f^{[n-2]''}(p) > 0$, und demnach fällt $f^{[n-2]}$ unter Fall (i) des Lemmas. Damit fällt aber nach dem Lemma $f^{[n-2k]}$ unter Fall (i) und $f^{[n-2k-1]}$ unter Fall (ii). Je nachdem ob n gerade oder ungerade ist, fällt also $f = f^{[0]}$ unter (i) bzw. (ii). □

Der Satz sagt, daß sich $f(p+h)$ für kleine h ebenso verhält, wie die Funktion

$$h \mapsto f^{[n]}(p) \cdot h^n.$$

Das werden wir später noch besser verstehen, wenn von der Taylorentwicklung die Rede ist (vergl. auch Bd. 2, V, 3.3). Verschwinden allerdings alle Ableitungen von f bei p, so kann f lokal um p sehr pathologisch werden, weder extremal noch monoton.



Ein Beispiel ist die Funktion

$$f(x) = x^2 \sin\left(\tfrac{1}{x}\right), \quad f(0) = 0.$$

Wir werden auch noch Beispiele kennenlernen, die beliebig oft differenzierbar sind. Aber solche Beispiele sind doch untypisch.

Um zu zeigen, daß ein Punkt nun wirklich ein absolutes Extremum ist, sind natürlich weitere Überlegungen nötig. Manchmal hilft der Satz, daß eine stetige Funktion auf einem kompakten Intervall jedenfalls ein Extremum hat. Kann man die Randpunkte ausschließen, so bleiben meist nur noch wenige Nullstellen der Ableitung von f als Kandidaten.

Übrigens ist in unseren Definitionen nicht ausgeschlossen, daß der Punkt p ein Randpunkt des Intervalls D ist. Man spricht dann auch von **einseitiger** Differenzierbarkeit von f bei p.

§ 4. Der Hauptsatz

Dieser Satz lehrt, daß Ableiten und Integrieren zueinander inverse Operationen sind. Wir betrachten wie immer ein Intervall D mit mehr als einem Punkt und eine darauf definierte Funktion f. Eine Funktion $F : D \to \mathbb{R}$ heißt eine **Stammfunktion** von f, wenn

$F' = f$. Zwei Stammfunktionen F und F_1 unterscheiden sich um eine Konstante, denn $(F - F_1)' = F' - F_1' = f - f = 0$.

(4.1) Hauptsatz. *Ist $a \in D$ und f stetig auf D, so ist*

$$x \mapsto \int_a^x f(t)\,dt$$

eine Stammfunktion von f auf D. Ist also F irgendeine Stammfunktion von f, so ist

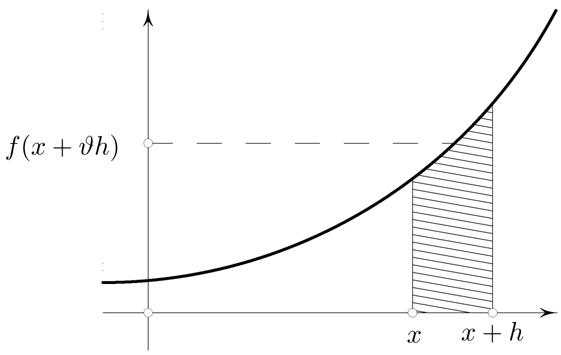
$$\int_a^x f(t)\,dt = F(x) - F(a) =: \big[F\big]_a^x.$$

Beweis: Sei $g(x) := \int_a^x f(t)\,dt$, dann ist nach dem Mittelwertsatz der Integralrechnung

$$g(x+h) - g(x) = \int_x^{x+h} f = f(x + \vartheta_h \cdot h) \cdot h, \quad 0 \le \vartheta_h \le 1.$$

Und $\lim_{h\to 0} f(x + \vartheta_h \cdot h) = f(x)$, das zeigt $g'(x) = f(x)$.

Eine Stammfunktion ist bis auf eine Konstante, also durch ihren Wert an einem Punkt bestimmt, und die beiden Stammfunktionen $\int_a^x f$ und $F(x) - F(a)$ stimmen für $x = a$ überein. $\square$



Wegen des Hauptsatzes nennt man eine Stammfunktion von f auch ein **unbestimmtes Integral** und bezeichnet sie durch

$$\int f \;=\; \int f(t)\,dt.$$

Zum Beispiel ein Polynom $f(x) = \sum_{k=0}^{n} a_k x^k$ hat das unbestimmte Integral

$$\int f(t)\,dt \;=\; a + \sum_{k=0}^{n} \frac{a_k}{k+1} x^{k+1}.$$

So gewinnen wir durch Differenzieren von Funktionen mühelos viele Formeln für Integrale, die aus der Definition des Integrals direkt nur schwer zu finden wären, denn für das Rechnen mit elementaren Funktionen ist das Differenzieren meist leichter als das Integrieren. Auf die Dauer und prinzipiell erweist sich dagegen doch das Integral als der Operator mit den besseren Eigenschaften: Man beweist zum Beispiel die Lösbarkeit von Differentialgleichungen, indem man daraus eine Integralgleichung macht. Jede stetige Funktion läßt sich integrieren und liefert als Integral nach dem Hauptsatz eine differenzierbare Funktion: Integrieren macht die Funktionen besser. Ableiten dagegen macht sie schlechter, die Ableitung einer differenzierbaren Funktion braucht nichteinmal stetig zu sein.

Jede Regel der Differentialrechnung läßt sich — jedenfalls für stetige Funktionen — nach dem Hauptsatz als Regel der Integralrechnung deuten. Aus der Kettenregel wird so die

(4.2) Transformationsformel. *Ist $f : D \to \mathbb{R}$ eine stetige Funktion und $\varphi : [a,b] \to D$ stetig differenzierbar, so ist*

$$\int_a^b f \circ \varphi(t) \cdot \varphi'(t)\,dt \;=\; \int_{\varphi(a)}^{\varphi(b)} f(u)\,du.$$

In sinnfälliger Notation schreiben wir: Ist $u = \varphi(t)$, so ersetze zur Transformation des Integrals $du = \varphi'(t)\,dt$.

Beweis: Für jedes $x \in [a,b]$ zeigen wir:

$$F(x) := \int_a^x f \circ \varphi \cdot \varphi' = \int_{\varphi(a)}^{\varphi(x)} f =: G(x).$$

Für $x = a$ verschwinden beide Funktionen, und es ist $F' = G'$ zu zeigen. Nach dem Hauptsatz ist $F'(x) = f\big(\varphi(x)\big) \cdot \varphi'(x)$, und G muß man, entsprechend der Kettenregel, erst nach der oberen Grenze, und dann diese nach x ableiten. Das ergibt auch $f\big(\varphi(x)\big) \cdot \varphi'(x)$. □

Beispiel. Gesucht ist ein Integral $\int x\sqrt{1+x}\,dx$. Setze $\sqrt{1+x} = u$, also $x = u^2 - 1$, $dx = 2u\,du$, und es ergibt sich

$$2\int (u^2-1)u^2\,du = 2\int (u^4 - u^2)\,du = 2\Big(\frac{u^5}{5} - \frac{u^3}{3}\Big),$$

was das Problem im wesentlichen löst.

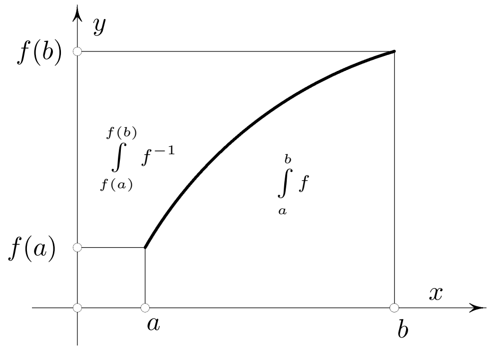
Leichte Anwendungen der Transformationsformel sind die Formeln:

$$\begin{aligned} \int_a^b f(t+c)\,dt &= \int_{a+c}^{b+c} f(x)\,dx, && x = t+c. \\ c\int_a^b f(ct)\,dt &= \int_{ac}^{bc} f(x)\,dx, && x = ct. \\ \int_a^b t^{n-1} f(t^n)\,dt &= \frac{1}{n}\int_{a^n}^{b^n} f(x)\,dx, && x = t^n. \end{aligned}$$

Ist f eine streng monotone Funktion auf $[a,b]$ mit Bild $[f(a), f(b)]$ oder $[f(b), f(a)]$, so hat man auf dem Bildintervall die inverse Funktion, und es gilt:

$$\int_a^b f + \int_{f(a)}^{f(b)} f^{-1} = bf(b) - af(a). \tag{4.3}$$

Wissen wir also eine Funktion zu integrieren, so können wir auch das Integral der Umkehrfunktion hinschreiben, wenn die Funktion umkehrbar ist.



Beweis: Ist f differenzierbar, so setze x für b und differenziere nach x. Es kommt beidseits $xf'(x) + f(x)$ heraus. Im allgemeinen hilft die Figur weiter: Eine Treppenfunktion unter f liefert eine Treppenfunktion ber f^{-1} und umgekehrt. □

Schließlich hat man als oft geschicktes Hilfsmittel die

(4.4) Partielle Integration. *Sind f und g auf dem Intervall $[a, b]$ stetig differenzierbare Funktionen, so gilt:*

$$\int_a^b f(t)\, g'(t)\, dt \;=\; \big[f \cdot g\big]_a^b - \int_a^b f'(t)\, g(t)\, dt.$$

Kurz notiert: $\int f\, dg \;=\; fg - \int g\, df$.

Beweis: Ersetze b durch x und differenziere, dann steht da:
$f \cdot g' = (f \cdot g)' - g \cdot f'$, also die Produktregel. □

Wählt man speziell $g(x) = x$, so erhält man

$$\int f = xf - \int xf'.$$

Als Anwendung betrachten wir die Funktion $\log(x) := \int_1^x t^{-1}\, dt$, die wir im nächsten Abschnitt genauer studieren werden. Es ist also

$$\log'(x) = 1/x.$$

Mit der letzten Formel erhalten wir daher

$$\int \log(x) = x\log(x) - \int 1 = x\bigl(\log(x) - 1\bigr).$$

Ähnlich gelingt die Berechnung der Integrale

$$A_m := \int \frac{dx}{(1+x^2)^m}.$$

$A_m = \frac{x}{(1+x^2)^m} + 2m\int \frac{x^2\,dx}{(1+x^2)^{m+1}} = \frac{x}{(1+x^2)^m} + 2mA_m - 2mA_{m+1}$,

wegen $x^2 = (1+x^2) - 1$. Damit haben wir die Rekursionsformel

$$A_{m+1} = \frac{1}{2m}\cdot\frac{x}{(1+x^2)^m} + \frac{2m-1}{2m}A_m. \tag{4.5}$$

Wir werden bald lernen, für A_1 auch

$$A_1(x) = \arctan(x)$$

zu schreiben. Daraus lassen sich dann rekursiv alle A_n berechnen.

Eine Transformation zu finden, die ein gegebenes Integral in ein bekanntes umformt, verlangt Übung und Geschicklichkeit, und oft auch psychologische Einfühlung in die Neigungen des Aufgabenstellers. Man darf jedoch nicht erwarten, daß jedes Integral schon bekannter Funktionen sich wieder durch bekannte Funktionen algebraisch oder durch Zusammensetzung ausdrücken läßt. Auch muß man bedenken, daß eine Funktion wie

$$\log(\arctan\sqrt{1+\cos^2 x}) - \sin(\arctan(\log|x|))$$

auch nicht ernstlich bekannt ist. Daher ist es gerade in Anwendungen, aber auch in theoretischen Betrachtungen, oft vernünftiger, eine durch ein Integral definierte Funktion direkt zu untersuchen und durch Näherungsverfahren das Integral zu berechnen, als sich in Lösungsformeln zu verstricken. Immerhin soll der Student auch einiges Geschick im Rechnen erwerben. Nun ist allerdings doch schon offenbar, daß es uns überhaupt an Funktionen fehlt. Spezielle Funktionen sind nicht nur Beispiele zur Anwendung der allgemeinen Sätze, sondern sie bilden den eigentlich konkreten Inhalt der Analysis.

§ 5. Logarithmus und Exponentialfunktion

Der natürliche **Logarithmus** ist die Funktion

$$\log : \mathbb{R}_+ \to \mathbb{R}, \quad x \mapsto \int_1^x t^{-1}\, dt.$$

Er ist also durch die Gleichungen

$$\log(1) = 0, \quad \log'(x) = \frac{1}{x}$$

bestimmt. Weil auf dem Definitionsgebiet der positiven reellen Zahlen $\log' > 0$ ist, wächst der Logarithmus streng monoton und ist, wie die Funktion $1/x$, beliebig oft differenzierbar.

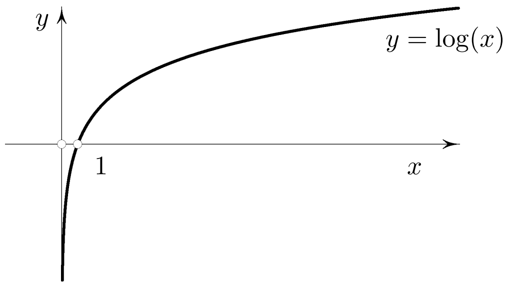
(5.1) $$\log(ax) = \log(a) + \log(x).$$

Beweis: Beide Seiten sind gleich für $x = 1$ und haben gleiche Ableitung. □

Insbesondere folgt $\log(x) + \log(x^{-1}) = \log(1) = 0$, also $\log(x^{-1}) = -\log(x)$. Für $n \in \mathbb{N}$ folgt induktiv $\log(x^n) = n\log(x)$, also $m\log(x^{\frac{1}{m}}) = \log(x)$, das heißt $\log(x^{\frac{1}{m}}) = \frac{1}{m}\log(x)$. Aus beidem zusammen folgt

$$\log(x^\alpha) = \alpha \log(x)$$

für rationale Zahlen α. Weil nun zum Beispiel $\log(2) > \log(1) = 0$, folgt $\log(2^n) = n\log(2) \to \infty$ für $n \to \infty$, und $\log(2^{-n}) = -n\log(2) \to -\infty$ für $n \to \infty$. Das Bild von log ist also die ganze reelle Gerade, und die Steigung $1/x$ fällt monoton und ist 1 im Punkt 1. Sie geht gegen ∞ für $x \to 0$ und gegen 0 für $x \to \infty$. Alles zusammengenommen haben wir schon ein ziemlich klares Bild des Graphen der Logarithmenfunktion:



Für $x < 0$ hat man nach der Kettenregel $\log(-x)' = 1/x$, also

$$\int \frac{1}{x} = \log|x| \quad \text{auf} \quad \mathbb{R} \setminus \{0\}.$$

Der Logarithmus hat eine Umkehrfunktion, die **Exponentialfunktion**

$$\exp : \mathbb{R} \to \mathbb{R}_+,$$

und nach dem Satz über die Umkehrfunktion ist diese ebenfalls differenzierbar und wächst streng monoton. Weil für $y = \log(x)$ gilt $y' = x^{-1}$ gilt für die Umkehrfunktion $x = \exp(y)$, $dx/dy = 1/x^{-1} = x$, also haben wir

$$\exp(0) = 1, \quad \exp'(x) = \exp(x). \tag{5.2}$$

Diese Gleichungen charakterisieren die Exponentialfunktion, genauer:

(5.3) Satz: *Sei* $u : \mathbb{R} \to \mathbb{R}$ *eine differenzierbare Funktion, sodaß* $u' = \alpha u$, $u(0) = \beta$ *für Konstanten* $\alpha, \beta \in \mathbb{R}$. *Dann ist*

$$u(x) = \beta \exp(\alpha x).$$

Beweis: Nach dem Gesagten erfüllt die Funktion $\beta \exp(\alpha x)$ jedenfalls die geforderten Gleichungen. Gilt nun dasselbe von u, so setze $v(x) = u(x)/\exp(\alpha x)$. Dann folgt

$$v'(x) = \frac{\exp(\alpha x)u'(x) - u(x)\alpha\exp(\alpha x)}{\exp(\alpha x)^2} = 0$$

wegen $u' = \alpha u$. Also $v = v(0) = u(0) = \beta$. □

Die Exponentialfunktion $\exp(\alpha x)$ beschreibt also das Wachstum oder die Abnahme mit konstanter Rate, ein Vorgang, der uns überall, segensreich oder verhängnisvoll, vor Augen steht. Die Differentialgleichung $y' = \alpha y$, die wir jetzt vollständig gelöst haben, ist eine der wichtigsten überhaupt.

Aus (5.1) folgt durch Umkehrung unmittelbar

$$\exp(a + x) = \exp(a) \cdot \exp(x), \tag{5.4}$$

denn beide haben gleichen Logarithmus $a + x$. Wir setzen

$$\exp(1) := e.$$

(Es ist $e = 2{,}718281\dots$). Dann gilt jedenfalls für rationale Exponenten

$$\exp(\alpha) = e^{\alpha},$$

denn beide Seiten haben gleichen Logarithmus α. Für andere Exponenten hatten wir bisher keine Potenz erklärt, und definieren jetzt

$$e^x := \exp(x).$$

Will man für eine beliebige Zahl $a > 0$ die **allgemeine Potenz** a^x erklären, so sollte doch $\log(a^x) = x \log(a)$ sein, und so setzen wir jetzt übereinstimmend mit dem früheren

$$a^x := e^{x \log(a)}. \tag{5.5}$$

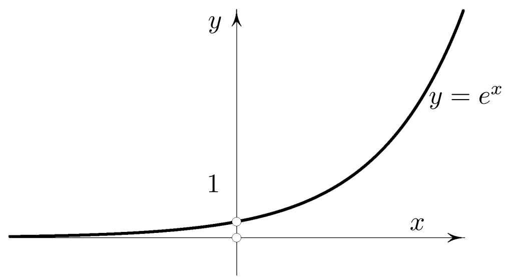
Es folgt dann $a^{x+y} = e^{(x+y)\log(a)} = e^{x \log(a)} \cdot e^{y \log(a)} = a^x \cdot a^y$, und

$$(a^x)' = \log(a) \cdot a^x. \tag{5.6}$$

Auch erhalten wir jetzt für beliebige Exponenten:
$(x^{\alpha})' = (e^{\alpha \log(x)})' = \frac{\alpha}{x} e^{\alpha \log(x)} = \frac{\alpha}{x} x^{\alpha} = \alpha x^{\alpha - 1}$, also

$$\begin{aligned} (x^{\alpha})' &= \alpha\, x^{\alpha-1} \quad \text{für} \quad x > 0. \\ \int x^{\alpha} &= \frac{x^{\alpha+1}}{\alpha + 1} \quad \text{für} \quad x > 0,\ \alpha \neq -1. \end{aligned} \tag{5.7}$$

Der Graph der Exponentialfunktion ergibt sich aus dem des Logarithmus.



Daß es die Exponentialfunktion als Bijektion $\mathbb{R} \to \mathbb{R}_+$ mit der Eigenschaft

$$\exp(a + b) = \exp(a) \cdot \exp(b)$$

gibt, ist ein grundlegender Satz über die algebraische Struktur des Körpers der reellen Zahlen. Wir betrachten einerseits die Gruppe $(\mathbb{R}, +)$ der beliebigen reellen Zahlen mit der Addition als Verknüpfung, und andererseits die Gruppe $(\mathbb{R}_+, \cdot)$ der positiven reellen Zahlen mit der Multiplikation als Verknüpfung. Dann definiert die Exponentialfunktion einen Isomorphismus

$$\exp : (\mathbb{R}, +) \to (\mathbb{R}_+, \cdot)$$

dieser Gruppen, mit inversem Isomorphismus log. Die beiden algebraischen Strukturen sind gleichsam dieselben mit unterschiedlicher Benennung. Was sich in der einen mit der Addition sagen läßt, gilt genauso in der anderen, wenn man die Addition durch die Multiplikation ersetzt. Das ist eine Besonderheit des reellen Zahlkörpers. Zum Beispiel beim Körper $\mathbb{Q}$ der rationalen Zahlen ist die additive Struktur geheimnisvoll, die multiplikative Struktur dagegen aufgrund der eindeutigen Primfaktorzerlegung sehr einfach.

§ 6. Winkelfunktionen

Wie schon erwähnt, definieren wir die Funktion $\arctan : \mathbb{R} \to \mathbb{R}$, den **Arkustangens**, durch

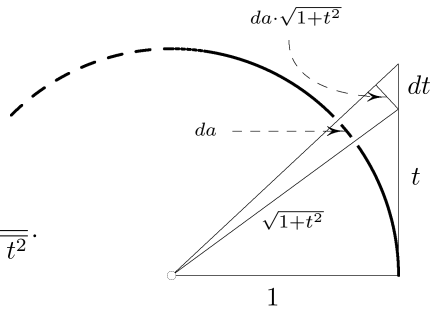
$$\arctan(t) = \int\limits_0^t \frac{1}{1+s^2}\, ds.$$

Schreiben wir der Kürze halber $a(t) := \arctan(t)$, so ist diese Funktion demnach durch die Gleichungen bestimmt:

$$a'(t) = \frac{1}{1+t^2}, \quad a(0) = 0. \tag{6.1}$$

Natürlich soll diese Funktion den im Bogenmaß gemessenen Winkel mit Tangens t angeben. Wir werden unsere Definitionen von Winkel, Tangens und Bogenmaß schon dementsprechend einrichten. Den Anschluß an die Anschauung liefert eine Physikerbetrachtung an folgender Figur:

$$\frac{da \cdot \sqrt{1+t^2}}{dt} = \frac{1}{\sqrt{1+t^2}}.$$



Man nennt diese Funktion auch den **Hauptzweig** des Arkustangens. Wir legen dem Folgenden nur die obige formale Definition (6.1) zugrunde. Weil $a'(t) = \frac{1}{1+t^2} > 0$, wächst die Funktion streng monoton mit abnehmender Steigung und ist beliebig oft differenzierbar. Der Arkustangens ist eine **ungerade** (**antisymmetrische**) Funktion, das heißt:

$$a(-t) = -a(t). \tag{6.2}$$

Wir definieren die Zahl π durch

$$\pi := 4 \cdot a(1).$$

Weil $\frac{1}{2} \le \frac{1}{1+t^2} \le 1$ auf $[0,1]$, folgt $\frac{1}{2} \le a(1) \le 1$, also $2 \le \pi \le 4$. Natürlich kann man π aufgrund der Definition ohne Mühe auf viele Stellen genau berechnen. Die numerische Integration von $\frac{1}{1+t^2}$, also z.B. das Einschließen dieser Funktion zwischen Treppenfunktionen, wäre ein mögliches, aber kein geschicktes Verfahren.

(6.3) $$a(t) + a(t^{-1}) = \pi/2 \quad \text{für} \quad t > 0.$$

Beweis: Die Formel gilt für $t = 1$, und die linke Seite hängt nicht von t ab, weil ihre Ableitung verschwindet:

$$\frac{1}{1+t^2} - \frac{t^{-2}}{1+t^{-2}} = 0. \qquad \square$$

Also für $t \to \infty$ geht $a(t) \to \pi/2$, weil $a(t^{-1}) \to a(0) = 0$. Aus (6.2) folgt somit, daß der Arkustangens eine umkehrbare Abbildung

$$a : \mathbb{R} \to (-\pi/2, \pi/2)$$

definiert. Die Umkehrabbildung heißt der **Tangens**

$$\tan : (-\pi/2, \pi/2) \to \mathbb{R},$$

und wir setzen diese Funktion periodisch fort:

$$\tan(x + \pi) := \tan(x).$$

Auf dem Intervall $(-\pi/2, \pi/2)$ wächst der Tangens streng monoton, und aus (6.2) folgt

(6.4) $$\tan(-x) = -\tan(x),$$

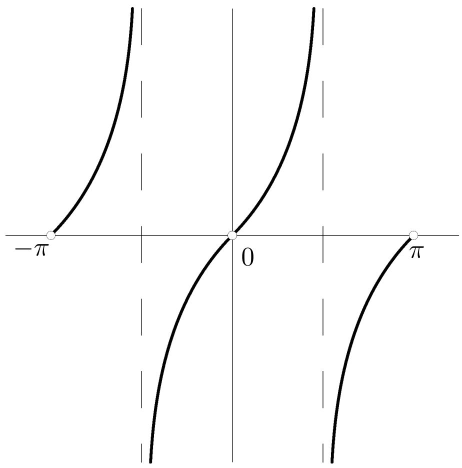
weil beide Seiten gleichen Arkus $-x$ liefern. Aus (6.3) folgt ganz entsprechend:

(6.5) $$\tan(x) \cdot \tan\left(\frac{\pi}{2} - x\right) = 1 \quad \text{für} \quad x \notin \left\{\frac{k\pi}{2} \mid k \in \mathbb{N}\right\}.$$

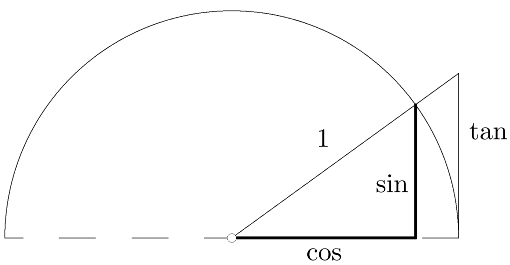
Der Satz über die Ableitung der Umkehrfunktion liefert

$$\frac{d}{dx}\tan(x) \;=\; 1+\tan^2(x). \tag{6.6}$$

Also außerhalb $\{\pi/2+k\pi \mid k\in\mathbb{Z}\}$ ist der Tangens beliebig oft differenzierbar. Dabei ist $\tan'(0)=1$ und $\tan'(x)\to\infty$ für $x\to\pm\pi/2$ auf dem Intervall $(-\pi/2,\pi/2)$. Die Steigung $\tan'$ fällt monoton auf $(-\pi/2,0]$ und steigt monoton auf $[0,\pi/2)$. Mit der Periodizität des Tangens ergibt sich somit als Graph die Figur:



Aus der Funktion Tangens erhält man **Sinus** und **Kosinus**, wenn man aufgrund elementargeometrischer Betrachtungen von den Formeln $\sin/\cos=\tan$, $\sin^2+\cos^2=1$ ausgeht:



$\cos(x) := \frac{1}{\sqrt{1+\tan^2(x)}}$, $\sin(x) := \frac{\tan(x)}{\sqrt{1+\tan^2(x)}}$ für $|x| < \pi/2$.
$\cos(\pi/2) := 0$, $\sin(\pi/2) := 1$.
$\cos(x+\pi) := -\cos(x)$, $\sin(x+\pi) := -\sin(x)$.

Damit sind Sinus und Kosinus zunächst für alle x so definiert, daß es mit unseren anschaulichen Vorstellungen verträglich ist. Aus dem, was wir über den Tangens schon wissen, folgt offenbar:

$$\begin{aligned} &\cos(-x) = \cos(x), && \sin(-x) = -\sin(x), \\ &\cos(x+2\pi) = \cos(x), && \sin(x+2\pi) = \sin(x). \end{aligned} \tag{6.7}$$

Sinus und Kosinus sind 2π-periodisch. Es ist leicht zu sehen, daß die beiden Funktionen an den Stellen $\pi/2 + k\pi$, wo sie zusammengestückelt sind, stetig bleiben, nämlich

$$\lim_{x\to\pm\pi/2} \cos(x) = 0, \qquad \lim_{x\to\pm\pi/2} \sin(x) = \pm 1.$$

Die zweite Gleichung zum Beispiel folgt für $t = \tan(x)$ aus

$$\lim_{t\to\infty} \frac{\pm t}{\sqrt{1+t^2}} = \pm 1.$$

Jetzt findet man die grundlegenden Differentialgleichungen

$$\sin'(x) = \cos(x), \quad \cos'(x) = -\sin(x). \tag{6.8}$$

Dies bestätigt man zunächst für das Intervall $|x| < \pi/2$ aus (6.6) und der Definition von Sinus und Kosinus durch Nachrechnen. Dann überträgt sich die Gleichung wegen der Periodizität (6.7) auf alle Punkte bis auf die $\pi/2 + k\pi$. Für diese hilft dann das bei solchem Basteln mit Funktionen oft nützliche allgemeine

(6.9) Lemma. *Sei D ein reelles Intervall mit mehr als einem Punkt, f eine auf D stetige Funktion, und $p \in D$. Angenommen f ist auf $D \setminus \{p\}$ differenzierbar und*

$$\lim_{x\to p} f'(x) = a, \quad x \neq p,$$

dann ist f *auch bei* p *differenzierbar und* $f'(p) = a$.

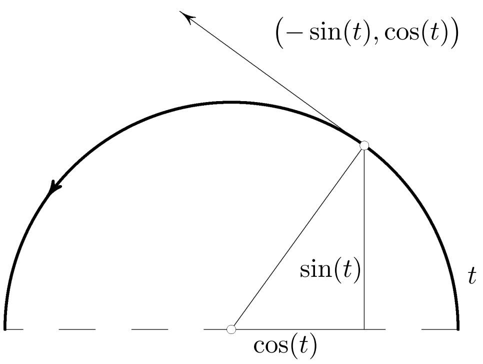
Beweis: Nach dem Mittelwertsatz der Differentialrechnung ist

$$\frac{f(p+h) - f(p)}{h} = f'(p + \vartheta_h h), \quad 0 < \vartheta_h < 1,$$

und für $h \to 0$ konvergiert dies gegen a. □

Findet man also, wie in unserem Fall, eine stetige Funktion, die auf einem Intervall außer bei p die Ableitung einer anderen stetigen Funktion ist, so ist sie es auch an der Stelle p.

Die Differentialgleichungen (6.8) haben eine sehr anschauliche Bedeutung: Wir denken uns, daß ein Punkt mit Geschwindigkeit 1 auf dem Einheitskreis in positiver Richtung herumläuft. Er beginne zur Zeit $t = 0$ seinen Lauf im Punkt $(1, 0)$. Nach der Zeit t befindet er sich an der Stelle $\big(\cos(t), \sin(t)\big)$.



Der Geschwindigkeitsvektor des auf dem Kreis herumlaufenden Punktes hat Länge 1, steht senkrecht auf dem Ortsvektor $\big(\cos(t), \sin(t)\big)$, und bildet diesem folgend ein Rechtssystem, und damit muß es sich um den Vektor $\big(-\sin(t), \cos(t)\big)$ handeln. Andererseits erhält man die Geschwindigkeit durch Ableiten der Ortskoordinaten nach der Zeit, also

$$d/dt\cos(t) = -\sin(t), \quad d/dt\sin(t) = \cos(t).$$

Nach dieser kinematischen Abschweifung ins Zweidimensionale zurück zu unseren Funktionen. Sie sind durch diese beiden Differentialgleichungen mit den Anfangsbedingungen

$$\sin(0) = 0, \quad \cos(0) = 1$$

vollkommen bestimmt, und wir hätten sie auch dadurch definieren können. Nur, daß es solche Funktionen gibt, haben wir aus dem Vorhergehenden erschlossen.

(6.10) Satz *(Schwingungsgleichung). Sei $u : \mathbb{R} \to \mathbb{R}$ eine zweimal differenzierbare Funktion mit*

$$u'' + u = 0.$$

Dann ist

$$u(x) = \alpha\cos(x) + \beta\sin(x), \quad \alpha = u(0), \quad \beta = u'(0).$$

Beweis: Setze $v := u'$, dann gilt

$$u' = v, \quad v' = -u, \quad u(0) := \alpha, \quad v(0) := \beta. \tag{6.11}$$

Diese Bedingungen erfüllt auch das Paar von Funktionen

$$\alpha\cos(x) + \beta\sin(x), \quad \beta\cos(x) - \alpha\sin(x),$$

und wir wollen zeigen, daß nur dieses Paar die Gleichungen (6.11) erfüllt. Nun, bilden wir die Differenzen

$$U := u - (\alpha\cos(x) + \beta\sin(x)), \quad V := v - (\beta\cos(x) - \alpha\sin(x)),$$

so gilt offenbar:

$$U' = V, \quad V' = -U, \quad U(0) = V(0) = 0,$$

und wir haben zu zeigen, daß dies nur vom Paar der Nullfunktionen erfüllt wird. Denken wir noch einmal an die kinematische Deutung, so besagen diese Gleichungen eigentlich, daß der Punkt (U, V) so über die Ebene läuft, daß seine Bewegungsrichtung stets senkrecht zum Ortsvektor ist, seine Anfangsposition aber der Nullpunkt ist. Freilich kann er dann da nicht wegkommen. Betrachten wir also demgemäß das Quadrat des Abstands, also ohne alle Deutung die Funktion $E(x) = U^2(x) + V^2(x)$. Es ist

$$E' = (U^2 + V^2)' = 2UU' + 2VV' = 2UV - 2UV = 0,$$

und $E(0) = 0$, also $E = 0$, also $U = V = 0$. □

Auch die Differentialgleichung

$$u'' + \omega^2 u = 0$$

können wir jetzt vollständig lösen. Ist $\omega \neq 0$, so ist die allgemeine Lösung

$$u = \alpha \cos(\omega t) + \beta \sin(\omega t), \quad \alpha, \beta \in \mathbb{R}.$$

Das erhält man aus dem Vorigen, indem man setzt:

$$w(t) := u(t/\omega).$$

Dann gilt $w'' + w = 0$, also $w(t) = \alpha \cos(t) + \beta \sin(t)$, also die Behauptung. (Und für $\omega = 0$?)

Die Behauptung läßt sich auch so aussprechen: Die zweimal differenzierbaren Funktionen $u : \mathbb{R} \to \mathbb{R}$, welche die Differentialgleichung

$$u'' + \omega^2 u = 0$$

erfüllen, bilden einen 2-dimensionalen Vektorraum. Eine Basis bilden die Funktionen $\sin(\omega t)$, $\cos(\omega t)$. Übrigens ist die Abbildung

$$u \mapsto u'' + \omega^2 u$$

eine lineare Abbildung des Raumes der zweimal differenzierbaren Funktionen. Wir haben den Kern berechnet.

Beim Beweis des Satzes haben wir als gleichwertiges Resultat gefunden, daß die Bedingung

$$u' = v, \quad v' = -u, \quad u(0) = \alpha, \quad v(0) = \beta$$

ein Paar von Funktionen u, v festlegt, nämlich

$$u = \alpha\cos(x) + \beta\sin(x), \quad v = \beta\cos(x) - \alpha\sin(x).$$

Als Anwendung des Satzes zeigen wir das

(6.12) Additionstheorem.

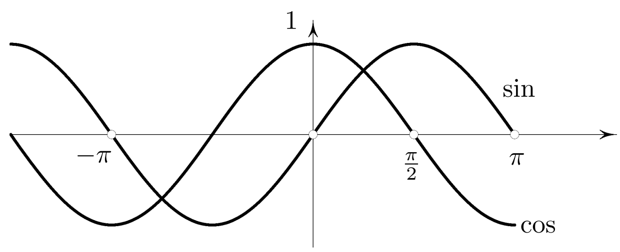
$$\sin(x+y) = \sin(x)\cos(y) + \cos(x)\sin(y),$$
$$\cos(x+y) = \cos(x)\cos(y) - \sin(x)\sin(y).$$

Beweis: Für festes y sei $u(x) = \sin(x+y)$, dann erfüllt u offenbar die Voraussetzung des Satzes mit

$$\alpha = u(0) = \sin(y), \quad \beta = u'(0) = \cos(y).$$

Also ist $u(x) = \alpha\cos(x) + \beta\sin(x)$, und das ist die erste Formel. Die zweite folgt ähnlich. □

Die Graphen von Sinus und Kosinus sind aus dem, was wir wissen, leicht anzugeben:



Die Relation

$$\sin^2(x) + \cos^2(x) = 1 \tag{6.13}$$

stimmt für $x = 0$, also allgemein, weil die Ableitung der linken Seite verschwindet. Schließlich zeigen wir das, wovon wir in der Motivation ausgegangen sind:

(6.14) Satz. *Sind $a, b \in \mathbb{R}$ und ist $a^2 + b^2 = 1$, so gibt es genau ein $t \in [0, 2\pi)$ mit*

$$a = \cos(t), \quad b = \sin(t).$$

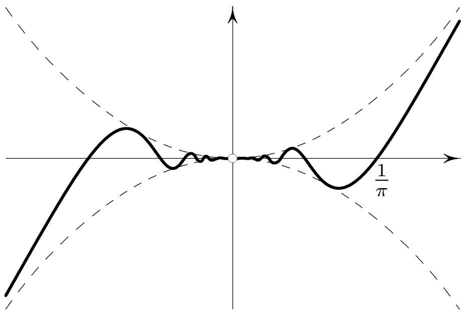
Beweis: Man macht eine Fallunterscheidung nach dem Quadranten der Ebene, in dem der Punkt (a, b) liegt. Sei etwa $0 \le a, b$. Weil $\cos(0) = 1$, $\cos(\pi/2) = 0$, und weil der Kosinus im Intervall $[0, \pi/2]$ streng monoton fällt (die Ableitung $-\sin$ ist auf dem offenen Intervall negativ), gibt es nach dem Zwischenwertsatz genau ein $t \in [0, \frac{\pi}{2}]$ mit $\cos(t) = a$, und dann ist $\sin(t) = \sqrt{1 - \cos^2(t)} = \sqrt{1 - a^2} = b$. Bleibt zu prüfen, daß $\big(\cos(t), \sin(t)\big)$ für $t \in (\pi/2, 2\pi)$ nicht in den ersten Quadranten fällt. Für die anderen Quadranten schließt man analog. □

Wir wollen die Erklärung der Winkelfunktionen nun als gesichert ansehen. Sie werden uns noch oft begegnen, und wir werden bald zu Aussagen kommen, die aus anschaulicher Betrachtung mit Schulargumenten nicht mehr zu erhalten sind.

Mit dem Sinus konstruieren wir eine Funktion, die überall differenzierbar, deren Ableitung jedoch nicht überall stetig ist, nämlich

$$f(x) = x^2 \sin(x^{-1}) \quad \text{für} \quad x \neq 0, \quad f(0) = 0.$$

Für $x \neq 0$ ist $f'(x) = 2x \cdot \sin(x^{-1}) - \cos(x^{-1})$. Für $x = 0$ zeigt die Definition der Ableitung mit $\Phi(h) = h \cdot \sin(h^{-1})$ unmittelbar $f'(0) = 0$. Dagegen konvergiert $f'(x)$, $x \neq 0$, für $x \to 0$ nicht.



Die Folge von Funktionen

$$f_n(x) = \tfrac{1}{n}\sin(n^2 x)$$

konvergiert gleichmäßig gegen die Nullfunktion, aber die Folge der Ableitungen

$$f_n'(x) = n\cos(n^2 x)$$

divergiert. Die Nähe der Funktionswerte sagt nichts über die Nähe der Ableitungen.