

2.2 Datensätze und Folgen

2.2.0 Der Weg ins Unendliche: Vom Datensatz zur Folge

(2.0.1) Die menschliche Erfahrung spielt sich im Bereich endlich vieler Operationen ab, faktisch sogar sehr weniger. Denken Sie an die Addition von Zahlen oder an Zeitmessungen. Man addiert 3 oder vielleicht 30 Zahlen, mehr nicht. Man mißt die Zeitdauer eines bestimmten Vorganges, sagen wir einer Pendelschwingung. Wie oft? Einige wenige Male. Aber das ist immer noch weit entfernt von *unendlich*.

(2.0.2) Was ist, wenn man sich vorstellt, man könne Erfahrungen zu Vorgängen mit unendlich vielen Operationen ausführen? Sagen wir die Addition von unendlich vielen Zahlen. Was käme dabei heraus? Man könnte das operativ so realisieren: Für die erste Addition benötigt man 1s (Sekunde). Für das Hinzuaddieren der nächsten Zahl benötigt nur noch 1/2 s. Für das der nächsten 1/4 s, dann 1/8 s usw. Kann man einen Zeitpunkt angeben, an dem man fertig ist? Und was wäre das Ergebnis? Klappt es auch, wenn man für den 3. Schritt 1/3s statt 1/4s benötigt, für den 4. Schritt 1/4s statt 1/8s usw.? Wir werden sehen, dass man im ersten Fall fertig wird, obwohl es keine "letzte Addition" gibt, im zweiten nicht.

(2.0.3) Die Notwendigkeit, zumindest gedanklich unendlich viele Operationen auszuführen, ist seit alterher bekannt. Denken Sie an die Bestimmung des Wertes der Kreiszahl π mit Hilfe von approximierenden Vielecken.

(2.0.4) Bemerkenswerterweise liefert die Mathematik in vielen Fällen eine Antwort auf die gestellte Frage *Was käme heraus, wenn man Erfahrungen mit unendlich vielen Operationen sammeln könnte?* Wenn man also mit übernatürlichen Kräften seine rechnerischen Fähigkeiten beschleunigen könnte, so wie oben angedeutet! Die Mathematik liefert in vielen Fällen ein Resultat der folgenden Art: **Es kommt einerseits ein wohlbestimmtes Ergebnis heraus und andererseits läßt sich dieses auch angeben.**

(2.0.5) Für die Physik ist dieser Sachverhalt von größter Bedeutung, weil viele (für die praktische Arbeit) wichtige und erfolgreiche **Idealisierungen** auf derartigen **"Übergängen ins Unendliche"** beruhen. Dazu gehört etw die Wahrscheinlichkeitsrechnung, die Beschreibung von Materie durch eine kontinuierliche Dichte oder umgekehrt durch punktförmige Atome. Unstetige Einschaltprozesse usw, usw.

(2.0.6) Den Ausgang für die Behandlung von Fragen dieser Art bilden die Abbildungen $(I_n, k \mapsto a_k, M)$ vom Datensatztyp. Durch vorgestelltes Sammeln von "immer mehr" Daten erhält man schließlich Abbildung vom Folgentyp $(\mathbb{N}, k \mapsto a_k, M)$. **Die Verhaltensuntersuchung von Folgen ist daher unser erstes Etappenziel.**

(2.0.7) Und eine solche Verhaltensuntersuchung beginnt mit der Erstellung eines zugehörigen Begriffssystems.

Einige Beispiele für das Auftreten von Folgen.

(2.0.8) **Beispiel:** Die Vorgabe einer **rekursive Formel** zur Berechnung des jeweiligen Nachfolgers gibt einen ersten Beispieltyp. Sagen wir $x_{n+1} = \frac{1}{2}(x_n + \frac{2}{x_n})$. Legt man dann noch einen **Startwert** x_0 fest, dann ist jedes Folgenglied x_n im Prinzip eindeutig festgelegt und man kann es mit endlich vielen Rechenschritten erreichen. Was geschieht bei "unendlicher Wiederholung" dieser Rechnung? Genauer: Was kann man über das Verhalten der Folge $n \mapsto x_n$ für wachsendes n aussagen? Nehmen wir als Beispiel $x_0 = 2$. Dann scheinen mit jedem neuen Folgenglied weitere Nachkommastellen festgelegt zu werden.

n=	0	1	2	3	4	5
x_n =	2	1.5	1.4166..	1.414215686..	1.414213562374	1.414213562373...

Zwei Möglichkeiten des Verhaltens für immer weiter laufendes n bieten sich an, die wir mit (K) und (D) bezeichnen wollen:

(K) Falls ein stabiler Grenzwert existiert, falls also x_n sich mit zunehmendem n einem festen Wert beliebig nähert, dann sollte dieser Wert $x_{n+1} = x_n$ erfüllen und unsere Formel liefert (über die entstehende quadratische Gleichung) dafür nur die beiden Kandidaten $\pm\sqrt{2}$. Ist $x_0 > 0$ gewählt, dann sind alle $x_n > 0$ und es muss $+\sqrt{2}$ sein. Experimente mit dem Taschenrechner oder Computer deuten sofort stark darauf hin, dass dies der Fall ist. Zur verdeutlichung bilden

wir eine neue Folge, nämlich $n \mapsto x_n - \sqrt{2}$. D.h. wir subtrahieren den erhofften Grenzwert. Der Beginn der neuen Folge sieht so aus:

n	0	1	2	3	4
$x_n - \sqrt{2}$	0.085	2.4×10^{-3}	2.1×10^{-6}	1.6×10^{-12}	9×10^{-25}

Also stimmen hier bereits für $n=4$ etwa 24 Nachkommastellen von x_4 mit denen von $\sqrt{2}$ überein. Aber "geht das wirklich immer so weiter"?

□ Versuchen Sie aber einmal die rekursive in eine explizite Formel umzuwandeln, etwa x_3 durch x_0 auszudrücken! (Computeralgebrasystem verwenden!)

(D) Falls kein solch stabiler Grenzwert existiert, was ist dann? Hier wäre zu überlegen, was dann für ein Verhalten überhaupt möglich wäre. Eine einfache Möglichkeit wäre, dass x_n immer weiter wächst. Aber es gibt natürlich noch viele weitere denkbare Verhaltensmuster. Die Experimente mit endlich vielen Operationen deuten jedoch darauf hin, dass zumindest die Folgen unseres Beispiels nicht hierher gehören. Aber können wir sicher sein?

(2.0.9) **Beispiel:** Die *harmonische Reihe*. Auch diese Folge definieren wir zunächst rekursiv

$$s_1 = 1, s_{n+1} = s_n + \frac{1}{n+1}.$$

Wie verhält sie sich? Veranschaulichung des Folgenbeginns mit dem Computer (Graphenstandpunkt) zeigt Wachstum. Trägt man die Daten logarithmisch auf, erhält man den Eindruck linearen Wachstums in $\ln(n)$. Daher wird man als nächstes die Folge $n \mapsto (s_n - \ln(n))$ betrachten und deren Verhalten untersuchen. Es zeigt sich, dass diese Folge beschränkt bleibt. Hier - für s_n - scheint also der Fall (D) vorzuliegen.

□ Es ist in diesem Fall leicht, die rekursive in eine explizite Formel umzuwandeln. Tun Sie das.

(2.0.10) **Beispiel:** Jetzt betrachten wir $n \mapsto t_{n+1} = t_n + \frac{1}{(n+1)^2}$. Das ist eine leichte Abwandlung des zweiten Beispiels. Es zeigt völlig anderes Verhalten, nämlich wieder eines vom Typ (D).

Offen bleibt die Frage, wie man derartige Vermutungen beweist, sichert.

(2.0.11) **Beispiel:** Jetzt betrachten wir $n \mapsto q_{n+1} = q_n + x^n$. Hierbei sei $x \in \mathbb{R}$ ein äußerer Parameter und $q_0 = 1..$ Die zugehörige explizite Formel ist

$$q_n = 1 + x + x^2 + \dots + x^n = \sum_{k=0}^n x^k$$

wie man sofort sieht. Hierfür gibt es nun aber die Formel für die *endliche geometrische Reihe*

$$q_n = 1 + x + x^2 + \dots + x^n = \frac{1 - x^{n+1}}{1 - x}.$$

mit deren Hilfe man das Verhalten der Folge in den Griff bekommt, genauer: Es auf bekannte mathematische Resultate zurückführen kann. Dazu schreiben wir

$$\underbrace{1 + x + x^2 + \dots + x^n}_{\text{interessierende Größe}} = \underbrace{\frac{1}{1-x}}_{\text{fest}} + \underbrace{\frac{-x^{n+1}}{1-x}}_{\text{Fehler}}$$

und schreiben die interessierende Größe als Summe einer festen (von n unabhängigen) Größe und einer Größe, die wir Abweichung oder Fehler nennen.

Ist $|x| < 1$, dann wird dieser Fehler immer kleiner, geht nach Null, da dann x^{n+1} dies tut. Der feste Term sagt voraus, was sich nach unendlich vielen Schritten ergeben sollte. Hat man dagegen $|x| > 1$, dann wird auch die Abweichung betragsmäßig immer größer und der feste Term bildet eine immer schlechtere Voraussage.

(2.0.12) **Beispiel.** Was ist die Dezimalbruchdarstellung einer reellen Zahl? Sagen wir $\pi = 3.1415926535898\dots$? Offenbar können wir das wie folgt schreiben

$$\pi = 3 + \frac{1}{10^1} + \frac{4}{10^2} + \frac{1}{10^3} + \frac{5}{10^4} + \dots = \sum_k \frac{z_k}{10^k}$$

wobei z_k immer eine *Ziffer* ist, eine natürliche Zahl zwischen 0 und 9. Oder rekursiv $\pi_{n+1} = \pi_n + \frac{z_n}{10^n}$.

(2.0.13) **Beispiel:** Die Cantormenge. Vgl. Kap 1....

2.2.1 Das Begriffssystem für Zahlfolgen $\boxed{\mathbb{N} \rightarrow \mathbb{C}}$

(2.1.1) Wir betrachten jetzt Zahlfolgen, also Abbildungen des Typs $a = (\mathbb{N}, n \mapsto a_n, \mathbb{C})$. Wir wollen das zugehörige Begriffssystem ausbauen. Wir werden solche Folgen vielfach vom Feldstandpunkt aus darstellen, also schreiben a_0, a_1, a_2, \dots . Die Werte a_n der Folgenabbildung a nennt man auch die *Folenglieder*. Halten Sie beiden auseinander: Die gesamte Abbildung und einzelne Glieder. Hat man eine Berechnungsformel wie $a_n = \sum_{k=1}^n \frac{1}{k^2}$, so kann man durch Rollenwechsel zwischen Abbildung und Folenglied wechseln.

(2.1.2) Der Feldstandpunkt ist Ausgangspunkt der verbreiteten Schreibweise $(a_k)_{k \in \mathbb{N}}$ für Folgen, die zeigt, dass eine Folge in gewisser Weise ein unendliches Tupel ist.

(2.1.3) Was ist eine $\boxed{\text{Teilfolge von } a}$? a_0, a_2, a_4, \dots soll eine sein. Dagegen $a_0, a_2, a_1, a_4, a_3, \dots$ keine. Hierzu definieren wir:

Definition: Sei $\sigma : \mathbb{N} \rightarrow \mathbb{N}$ eine injektive monoton wachsende Abbildung.

Dann bezeichnen wir $b : \mathbb{N} \rightarrow \mathbb{C}$ mit $b_k = a_{\sigma(k)}$ als *Teilfolge* von a .

Das erste Beispiel mit $b_k = a_{2k}$ ist eine Teilfolge mit $\sigma(k) = 2k$. Im zweiten Fall ist die Monotonie (also $i < j \Rightarrow \sigma(i) < \sigma(j)$) verletzt. Etwa durch $(\sigma(1) = 2) > (\sigma(2) = 1)$.

□ Weitere wichtige Teilfolgen: $a_N, a_{N+1}, a_{N+2}, \dots$ oder a_1, a_3, a_5, \dots oder $a_0, a_2, a_4, a_8, \dots$. Wie sieht das zugehörige σ jeweils aus?

(2.1.4) Natürlich kann man aus einer gegebenen Folge neben den Teilfolgen noch weitere $\boxed{\text{abgeleitete Folgen}}$ bilden. Einige wichtige geben wir an:

Es sei $\pi : \mathbb{N} \rightarrow \mathbb{N}$ bijektiv, also eine Permutation von \mathbb{N} . dann entsteht $b_n = a_{\pi(n)}$ aus a durch *Vertauschung der Folenglieder* ($a_0, a_4, a_3, a_8, a_7, a_{16}, a_{15}, \dots$ etwa entsteht durch Vertauschen der Glieder einer Teilfolge!)

Eine zweite wichtige abgeleitete Folge ist die *Folge der Beträge*. Zu $a = (\mathbb{N}, n \mapsto a_n, \mathbb{C})$ bilde

$$|a| = (\mathbb{N}, n \mapsto |a_n|, [0, \infty[).$$

Für $a = (\mathbb{N}, n \mapsto (-1)^n, \mathbb{R})$ etwa hat man $|a| = (\mathbb{N}, n \mapsto 1, \mathbb{R})$. Das ist eine *konstante Folge*.

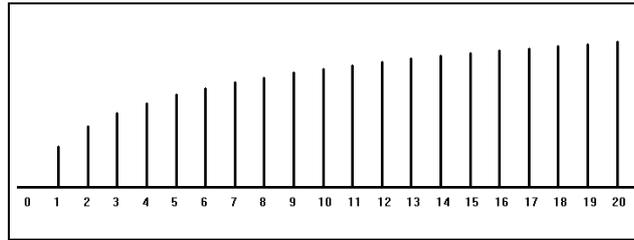
□ Etwa zu $f = (\mathbb{N}, n \mapsto e^{in}, \mathbb{C})$ bilden wir $|f| = (\mathbb{N}, n \mapsto |e^{in}| = \dots?)$.

(2.1.5) Bei der Zordnung $n \mapsto \frac{1}{n}$ ist $n=0$ nicht zugelassen. Man hilft sich, indem man umnummeriert und $k \mapsto \frac{1}{k+1}$ betrachtet. Teilweise ist es jedoch besser, die alte Numerierung zu behalten und stattdessen den Beginn zu verschieben. Für die Verhaltensdiskussion für große n sind eventuelle Sonderfälle und Abnormitäten zu Beginn ja bedeutungslos. Wir vereinbaren, dass wir solche abweichenden Anfangsglieder bei Bedarf stillschweigend fortlassen, wenn es nur um das Verhalten für $n \rightarrow \infty$ geht. Folgende Sprechweise ist üblich: Gilt eine Eigenschaft der Glieder einer Folge erst ab einem Wert N_0 , dann sagt man, die Eigenschaft gelte *für fast alle n* . Also: "fast alle" bedeutet "bis auf endlich viele Ausnahmen".

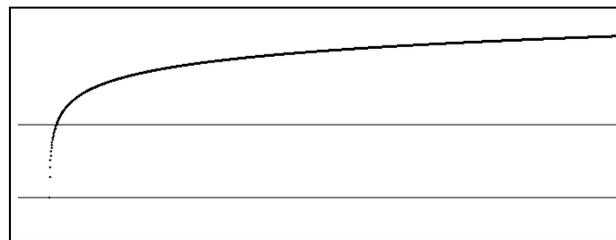
2.2.2 Verhaltensveranschaulichung und einige charakteristische Eigenschaften von Folgen

(2.2.1) Die günstigste Veranschaulichung ist die vom Graphenstandpunkt, eventuell mit gewissen spezifischen Modifikationen. Horizontal trägt man den Anfang der Urbildmenge auf und heftet dann jeweils als vertikale Höhe den zugehörigen Feldwert an. Die Frage ist häufig, wie weit man mit n gehen kann. Trägt man ausreichend viele n in kurzem Abstand auf, gehen die diskreten Punkte in einen Kurvenzug über. Manchmal ist es günstig, horizontal nicht n , sondern den Logarithmus von n aufzutragen. Damit wird der Weg " n nach unendlich" komprimiert. Ebenso kann man vertikal eine logarithmische Auftragung versuchen, eventuell des Betrages des Abstandes zu einem Grenzwert.

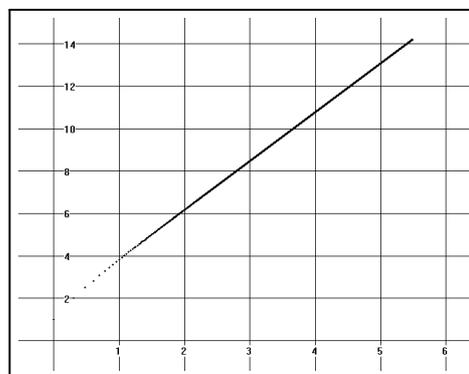
(2.2.2) Wir betrachten das Beispiel der harmonischen Reihe mit $a_n = \sum_{k=1}^n \frac{1}{k}$. Im ersten Bild sind die ersten 20 Glieder aufgetragen. a_{20} ist etwa 3.6. Beachten Sie $a_1 = 1$, $a_2 = \frac{3}{2}$ usw.



Einige Beschreibungsmerkmale: Alle Folgenglieder (=Folgewerte) sind positiv. Man sagt, die Folge habe *positive Glieder* oder manchmal verkürzt: Die Folge sei *positiv*. Außerdem wachsen die Werte mit zunehmendem n . Man spricht von einer *monoton wachsenden Folge*. Aber sind die Werte auch beschränkt? Das ist im Bild noch nicht zu erkennen, auch nicht abzuschätzen. Ebensov wenig ist vom Bild sicher, ob die Monotonie immer weiter erhalten ist. Die rekursive Form des Rechenausdrucks zeigt das allerdings sofort: Der Nachfolger ist immer um das Stück $\frac{1}{n+1}$ größer. Aber dieser Zuwachs wird immer kleiner, geht nach Null. Im nächsten Bild sind die ersten 2000 Glieder aufgetragen, nur noch die Endpunkte, nicht mehr die gesamte Feldhöhe. Der obere horizontale Strich hat die Höhe von a_{20} . Man sieht also deutlich weitere Zunahme. Aber irgendwie wirkt die Form ähnlich zur Form der ersten Figur.



(2.2.3) Als nächstes tragen wir a_n gegen $\log(n)$ auf, stauchen die horizontale Achse also stark zusammen. Ein Horizontalwert 4 steht daher für 10^4 . Insgesamt geht n in diesem Bild bis 500000. Jetzt liegt die Vermutung ziemlich nahe, dass a_n mit $\log(n)$ linear anwächst: $a_n \approx \alpha \log(n) + \beta$.



Aber die Frage bleibt, ob sich dieses Verhalten nicht irgendwann für viel größeres n doch ändert. Wir haben noch keinen Beweis unserer Vermutung.

(2.2.4) Projiziert man in derartigen Bildern für $\text{Graph}(a)$ die Figur auf die vertikale Achse, dann erhält man die Menge $\text{Bild}(a)$. Eine Folge heißt *nach oben beschränkt*, wenn $\text{Bild}(a)$ als Menge nach oben beschränkt ist. In unserem Fall erhebt sich die Frage: Ist die Folge der harmonischen Reihe wirklich nach oben unbeschränkt?

2.2.3 Vorgabemethoden für Folgen

Wie werden Folgen in der Regel vorgegeben? Das ist häufig der erste Schritt einer Problemdiskussion. Wir nennen die wichtigsten Methoden.

(2.3.1) Vorgabe über eine explizite Berechnungsformel

- Durch Angabe einer **expliziten Berechnungsformel**. Etwa $a_n = \frac{1}{n}$ oder $b_n = \sqrt{n}e^{i\frac{n}{100}}$ Ein spezielles Beispiel bilden die konstanten Folgen $a_n = c$. Wie steht es mit $c_n = \frac{1}{n-2}$? Hier ist c_2 nicht zulässig. In den meisten Fällen sind die ersten Folgenglieder für die Verhaltensdiskussion bei n nach unendlich irrelevant. Dann betrachtet man stillschweigend die Teilfolge der Glieder ab $n=3$. Numeriert diese auch nicht um.
- Die Berechnungsformel gehört vielfach zu einer vertrauten reellen Funktion. Oder: Die Folge ist einfach die **Restriktion einer Funktion auf die natürlichen Zahlen**. So ist $a_n = \frac{1}{n}$ die Restriktion der Hyperbelfunktion $x \mapsto \frac{1}{x}$ auf die Zahlen $1, 2, 3, \dots$. Kennt man dann das Verhalten der reellen Funktion, so kann man oft bequem Schlüsse auf das Folgenverhalten ziehen!

(2.3.2) Vorgabe über eine **rekursive Berechnungsformel**

- Durch Angabe einer **rekursiven Berechnungsformel** sowie eventueller Anfangswerte. Siehe dazu die Eingangsbeispiele. Aber auch $a_{n+1} = a_n + na_{n-1}$ ist ein Beispiel. Hier allerdings muss man a_0 **und** a_1 als Anfangswerte vorgeben.

□ Berechnen Sie für das letzte Beispiel a_2, \dots, a_5 wenn $a_0 = a_1 = 1$ ist und wenn $a_0 = 1$ und $a_1 = x$ ist.

- Vorgabe durch *Iteration einer Funktion*. Ist $f: \mathbb{R} \rightarrow \mathbb{R}$ oder $f: \mathbb{C} \rightarrow \mathbb{C}$ eine Funktion, dann gibt $x_{n+1} = f(x_n)$ eine Rekursion. Natürlich ist zusätzlich ein Startwert x_0 festzulegen.

(2.3.4) Durch **verbale Festlegung der Zuordnung** wie "Sei p_n die n -te Primzahl" und darauf aufbauende Konstruktionen, etwa " $\pi_n =$ Anzahl der Primzahlen bis n ". Die erste Folge beginnt wie folgt:

1	2	3	4	5	6	7	8	9	10	11	12	13
2	3	5	7	11	13	17	19	23	29	31	37	41

und die zweite, die man auch *Verteilungsfunktion* nennt

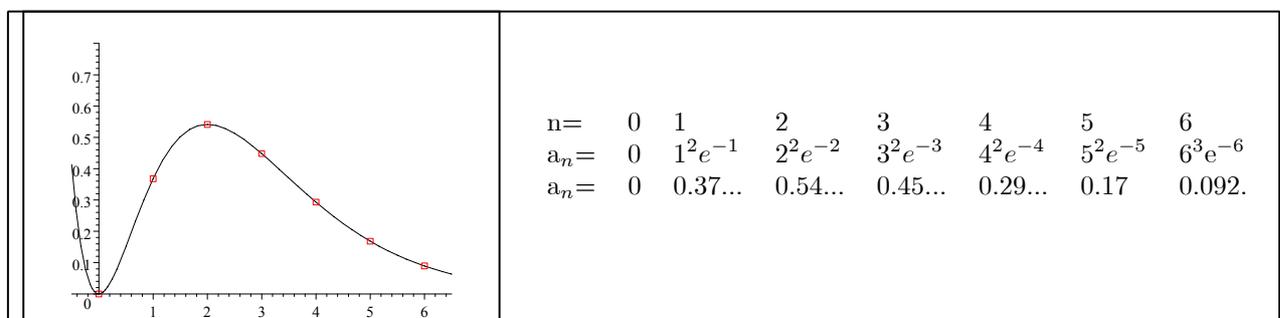
1	2	3	4	5	6	7	8	9	10	11	12	13
0	1	2	2	3	3	4	4	4	4	5	5	6

Solche Verteilungsfunktionen erweisen sich vielfach als nützlich.

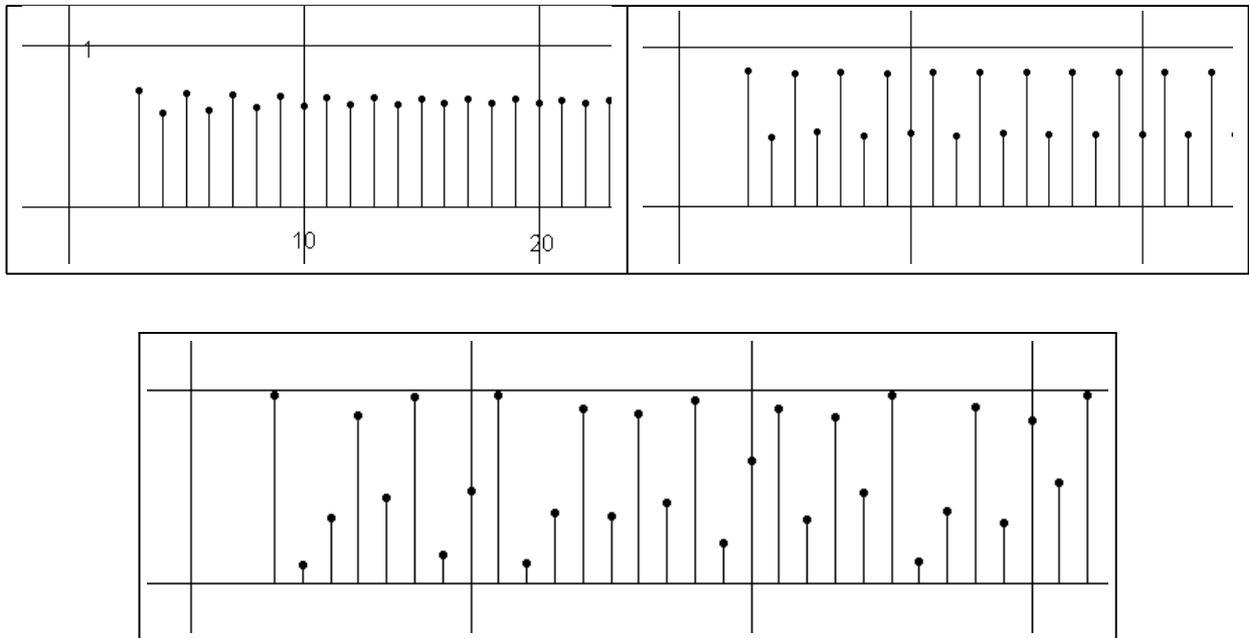
(2.3.5) Hat man eine Folge gegeben, so erhält man weitere Folgen durch Bildung von Teilfolgen und abgeleiteten Folgen. Vgl.

- "Auswahlfolgen" Siehe etwa (2.6.8). Ein allgemeines mathematisches Resultat sichert die Existenz einer Folge mit bestimmten Eigenschaften. Dann kann man mit dieser Folge argumentieren, daraus Konsequenzen herleiten, auch wenn man die Folgeglieder nicht kennt, nur weiß, dass sie existieren.

(2.3.6) Die Restriktionsmethode aus (2.3.1) ist besonders nützlich. Wir geben ein Beispiel. Dazu betrachten wir $f = (\mathbb{R}, x \mapsto x^2 e^{-x}, \mathbb{R})$ und restringieren auf \mathbb{N} . Also $f|_{\mathbb{N}}$ in der Schreibweise von Kapitel 1. Das gibt für $a = f|_{\mathbb{N}}$ die ersten Folgenglieder (links die Graphik, rechts die Tabelle)



(2.3.7) Jetzt noch ein Beispiel für die Funktionsiteration aus (2.3.2). Wir wählen $f(x)=ax(1-x)$. Hierbei sei a äußerer Parameter. $2 < a < 4$ ist der interessante Bereich. Für x_0 sollte man eine Zahl zwischen 0 und 1 wählen. Die Bilder zeigen das Anfangsverhalten der entstehenden Iterationsfolgen für $a=2.9$ und $a=3.4$ und $a=3.9$ dargestellt vom Feldstandpunkt. In den ersten beiden Bildern sind die ersten Glieder von $n=3$ bis $n=23$ aufgetragen. Alle Werte liegen zwischen 0 und 1. Im ersten Fall scheinen sich die Werte rasch einem Grenzwert zu nähern. Im zweiten dagegen wechseln sie zwischen zwei Werten. Oder die beiden Teilfolgen $n \mapsto x_{2n}$ und $n \mapsto x_{2n+1}$ scheinen unterschiedlich Grenzwerte zu haben. Im dritten Bild für $a=3.9$ dagegen ergibt sich keine Regelmäßigkeit. Zur Verdeutlichung sind über 30 Glieder aufgetragen. Allerdings liegen alle Werte immer noch zwischen 0 und 1.



2.2.4 Die Algebraische Struktur des Folgenraumes (Kap.3,4)

Sei $\mathcal{F}(\mathbb{N}, \mathbb{C})$ die Menge aller Folgen. Dazu kann man zunächst Teilmengen bilden. Etwa $\mathcal{F}(\mathbb{N}, \mathbb{R})$ oder $\mathcal{F}(\mathbb{N}, \mathbb{Z})$. Letzteres ist die Menge aller Folgen mit ganzzahligen Werten. Jede dieser Mengen trägt eine algebraische Struktur. Die ersten beiden sind Vektorräume über \mathbb{C} bzw \mathbb{R} , die dritte ist \mathbb{Z} -Modul. Aber man hat auch eine Ringstruktur, die durch Wertemultiplikation entsteht. Sagen wir $a = \left(\frac{1}{(n+1)^2}\right)_{n \in \mathbb{N}}$ und $b = ((-1)^n n)_{n \in \mathbb{N}}$.

Dann bilden wir $ab = \left(\frac{(-1)^n n}{(n+1)^2}\right)_{n \in \mathbb{N}}$ als neue Folge. und natürlich $a+b = \left((-1)^n n + \frac{1}{(n+1)^2}\right)_{n \in \mathbb{N}}$.

- Weisen Sie die Ringstruktur für die drei gegebenen Fälle nach.
- Wann kann man die Folge $\frac{1}{a}$ bilden? Was ist, wenn $a_n = 0$ nur für endlich viele n gilt.
- Wieso ist "Nur endlich viele a_n sind Null" nicht dasselbe wie "Nur für endliche viele n ist $a_n = 0$ "? Halten Sie beides sorgfältig auseinander!

2.2.5 Folgenverhalten (I): Nullfolgen und konvergente Folgen

(2.5.1) Von besonderer Bedeutung ist die Teilmenge der *Nullfolgen*. Das sind - intuitiv gesehen - Folgen, deren Folgenglieder sich beliebig genau der Null nähern. Also ein Verhalten zeigen wie es der Beginn der Folge $n \mapsto x_n - \sqrt{2}$ aus (2.0.8) zeigt. Auch die Folge aus (2.3.6) scheint eine Nullfolge zu sein. Die Folgen aus (2.3.7) zeigen die geforderte Eigenschaft (noch??) nicht. Um zu verstehen, was *Nullfolge* genauer bedeutet, stellen wir uns vor, dass alle Folgenglieder a_n als Dezimalbrüche vorgegeben sind. In einem Sichtfenster werden - Feldstandpunkt - nacheinander alle a_n in Dezimaldarstellung angegeben. Dann müssen

sämtliche Kommastellen irgendwann Null werden, falls sich die Folgenglieder der Null beliebig annähern. Das Sichtfenster sieht wie folgt aus, wobei für den gesamten Rest der nicht mehr gezeigten Stellen steht.

n=123	a ₁₂₃ = 0.000312474
-------	--------------------------------	------

Oder auch: Betrachtet man die 7. Nachkommastelle. Dann muss es ein $N_0 \in \mathbb{N}$ geben, von dem ab ($n \geq N_0$) alle Stellen bis mindestens zur 7. immer Null sind.

Oder auch: Wird die 7. Stelle von a_n immer wieder einmal ungleich Null, **wie groß man n auch wählt**, dann kann a_n keine Nullfolge bilden, kann sich der Null nicht beliebig annähern.

(2.5.2) Man muss bei dieser Konstruktion sehr genau auf das Szenenbild, Reihenfolge und Art der Vorgabe der einzelnen Objekte achten.

- Zunächst wird die Folge $a: \mathbb{N} \rightarrow \mathbb{R}$ vorgegeben
- dann wird eine Genauigkeitsgrenze ("Stellenzahl") in Form einer (beliebigen) Zahl $\varepsilon > 0$ vorgegeben.
- Nun muss man **dazu** eine Zahl $N_0 = N_0(\varepsilon)$ aus \mathbb{N} bestimmen, von der ab (d.h. für $n \geq N_0$) sich a_n um weniger als ε von der Null unterscheidet. Als Formel $|a_n - 0| < \varepsilon$.
- Und das muss für jedes $\varepsilon > 0$ möglich sein.

(2.5.3) Unter diesen Umständen wird man das (im Sichtfenster dargestellte) Folgenverhalten beschreiben als: "Die Folgenglieder nähern sich beliebig der Null". Und eine solche Folge wollen wir eine *Nullfolge* nennen.

(2.5.4) Zur Verdeutlichung dieses für viele zunächst schwierigen Sachverhaltes formulieren wir das Szenenbild in Form eines Zweipersonenspieles. Die beiden Personen nennen wir *Herausforderer* und *Verteidiger*.

- Die Folge gehört zum Verteidiger. Er behauptet, die Folge sei eine Nullfolge, "verteidigt" diese Aussage.
- Der Herausforderer nennt eine Genauigkeitsschranke in Form einer Zahl $\varepsilon > 0$. Das Sichtfenster ist so eingestellt, dass Abweichungen von Null unter dieser Genauigkeit nicht mehr sichtbar werden.
- Der Verteidiger muss dazu eine $N_0 = N_0(\varepsilon)$ angeben, so dass oberhalb von N_0 (also $n \geq N_0$ "für alle n mit $n \geq N_0$ ") nur noch Nullen im Sichtfenster erscheinen. Das Sichtfenster sieht also ab N_0 **immer** wie folgt aus

n	0.0000000000000000
---	--------------------	-------

Die Zahl der im Fenster gezeigten Nullen ist durch das ε des Herausforderers festgelegt.

(2.5.5) $\varepsilon = 0$ ist **nicht** zulässig. Hierfür müsste man ein unendlich langes Sichtfenster für unendlich vielen Stellen zeigen. Aber $\varepsilon > 0$ darf beliebig klein gewählt werden und ergibt immer ein **endliches** Sichtfenster. Natürlich gibt es kein kleinstes ε , keine kleinste echt positive Zahl. Daher das Spiel: Der Herausforderer **darf** sein ε beliebig klein wählen und der Verteidiger **muss** parieren können.

□ Im Beispiel $n \mapsto x_n - \sqrt{2}$ aus (2.0.8) sieht es so aus, als hätte man $N(10^{-20}) = 4$. Was erwarten Sie für $N(10^{-40})$?

(2.5.6) Der Verteidiger braucht daher eine Zuordnung $(]0, \infty[, \varepsilon \mapsto N_0(\varepsilon), \mathbb{N})$, die für ihn das Gewünschte leistet. Wir nennen eine solche Funktion eine *Strategiefunktion*. Sie setzt den Verteidiger in die Lage, jeden Zug des Herausforderers parieren zu können. Bitte beachten Sie: Das Spiel verlangt keineswegs, dass der Verteidiger einen möglichst kleinen N_0 -Wert angibt. Das wird nirgends verlangt. Nur **nach** diesem Wert **muss** es im Sichtfenster nur noch Nullen geben, vorher **darf** es ruhig solche geben. D.h., wenn es bequem ist, kann man N_0 ruhig vergrößern.

(2.5.7) Wie weit kann man N_0 **verkleinern**? Dazu betrachten wir die folgenden Mengen

$$\begin{aligned} Z_\varepsilon &= \{n \in \mathbb{N} \mid |a_n| < \varepsilon\} \subset \mathbb{N} \\ Z &= \{(\varepsilon, n) \mid \varepsilon > 0, |a_n| < \varepsilon\} \subset]0, \infty[\times \mathbb{N} . \\ Z_S &= \{(\varepsilon, n) \mid \varepsilon > 0, (|a_k| < \varepsilon \text{ für alle } k \geq n)\} \end{aligned}$$

□ Analysieren Sie die Eigenschaften von Z_S und beantworten Sie damit die Frage, wie weit man N_0 bestenfalls verkleinern kann. Wann kann man keine Strategiefunktion finden? D.h. wie drückt sich das in der Menge Z_S aus?

(2.5.8) Wir werden unten genauer diskutieren, wie man Strategiefunktionen konkret findet.

(2.5.9) Warum sind Nullfolgen so nützlich? Angenommen eine Folge geht für n nach unendlich nicht gegen Null, sondern nähert sich einem anderen Wert $A \in \mathbb{R}$. Wie ist das zu präzisieren? Nun dann bilden wir die neue abgeleitete Folge $n \mapsto d_n = a_n - A$. Und wenn a_n gegen A geht, dann sollte d_n gegen Null gehen, einen Nullfolge bilden. Wir können also stets zu dem besprochenen Fall zurückkehren.

Eine derartige Folge nennt man *konvergent (mit Grenzwert A)* und schreibt $\lim_{n \rightarrow \infty} a_n = A$ (Gelesen: Limes von a_n für n nach unendlich gleich A) Die Eingangsfolge aus (2.0.8) hat den Grenzwert $\sqrt{2}$. Wir schreiben auch kürzer $\lim(a)$ oder *lima* oder $\lim_n a_n$ statt $\lim_{n \rightarrow \infty} a_n$.

(2.5.10) Es liegt jetzt nahe, einige allgemeine Resultate algebraischer Art einzuschieben: Sei dazu $\mathcal{F}(\mathbb{N}, \mathbb{R})$ der Vektorraum aller Zahlfolgen und $\mathcal{N}(\mathbb{N}, \mathbb{R})$ die Teilmenge aller Nullfolgen und $\mathcal{K}(\mathbb{N}, \mathbb{R})$ die Teilmenge aller konvergenten Folgen.

□ Zeigen Sie: $\mathcal{K}(\mathbb{N}, \mathbb{R})$ und $\mathcal{N}(\mathbb{N}, \mathbb{R})$ sind Teilräume.

(2.5.11) Eine zweite Strukturierung ergibt sich wie folgt: Sei $\mathcal{K}(\mathbb{N}, \mathbb{R})$ der Raum aller konvergenten Folgen. Wir bilden für $a, b \in \mathcal{K}(\mathbb{N}, \mathbb{R})$

$$a \sim b \Leftrightarrow (a - b \text{ ist Nullfolge})$$

□ Zeigen Sie, dass das eine Äquivalenzrelation ergibt.

Die Klassen bestehen aus genau solchen Folgen, die ein und denselben Grenzwert haben.

(2.5.12) Bisher haben wir den Begriff der Nullfolge für reelle Folgen $\mathbb{N} \rightarrow \mathbb{R}$ eingeführt. **Wie steht es mit komplexen Folgen** $c: \mathbb{N} \rightarrow \mathbb{C}$? Nun jedes c_n ist eine komplexe Zahl, die die eindeutige kartesische Darstellung $c_n = u_n + iv_n$ besitzt. Und $c_n \rightarrow 0$ sollte offensichtlich bedeuten, dass die beiden reellen Folgen $n \rightarrow u_n$ und $n \rightarrow v_n$ Nullfolgen sind. Eine komplexe Zahl geht nach Null, wenn beide - Real- und Imaginärteil - nach Null gehen. Damit haben wir das Problem komplexwertiger Folgen auf das reellwertiger zurückgeführt.

(2.5.13) Neben der kartesischen Darstellung der komplexen Zahlen gibt es noch die Polardarstellung. Für unsere Folgenglieder sei das $c_n = r_n e^{i\varphi_n}$. Hierbei ist nur r_n eindeutig. Der geometrische Pfeil c_n geht nach Null, wenn r_n nach Null geht. Dagegen ist der Wert von φ_n ohne Bedeutung. Nun ist aber $n \rightarrow r_n$ erneut eine Zahlfolge, sogar mit nicht negativen Gliedern. ebenso ist klar, dass " (r_n) ist Nullfolge" gleichwertig ist zu " (u_n) und (v_n) sind Nullfolgen".

Fassen wir zusammen: Hat man es mit einer Folge $c: \mathbb{N} \rightarrow \mathbb{C}$ zu tun, dann bedeutet " c ist Nullfolge", dass entweder die Betragsfolge (r_n) Nullfolge ist oder aber die Realteilfolge (u_n) und die Imaginärteilfolge (v_n) beides Nullfolgen sind. Auf diese Weise führen wir alle Konvergenzprobleme für Folgen $\mathbb{N} \rightarrow \mathbb{C}$ auf Probleme für reelle Nullfolgen zurück. Und letztere behandeln wir im Augenblick vorzugsweise.

2.2.5a Die Bestimmung von Strategiefunktionen

(2.5.14) Im Zusammenhang mit der Bestimmung von Strategiefunktionen treten meist zwei Problemtypen auf:

- 1. Aufgabentyp:** Eine Folge a ist gegeben, von der man zeigen möchte, dass eine Nullfunktion vorliegt. Man muß konkret für diesen Fall eine Strategiefunktion konstruieren.
- 2. Aufgabentyp:** Man hat ein oder zwei Folgen, von denen man weiß, dass sie konvergent sind. Dazu eine abgeleitete Folge c . Ist dann auch c konvergent? Und kann man den Grenzwert von c durch die Grenzwerte der Eingabefolgen ausdrücken. Diese Fragestellung führt zu einer Reihe von *Grenzwertsätzen*.

(2.5.15) Wir beginnen mit dem ersten Problemkreis: Die **explizite Konstruktion einer Strategiefunktion**. Eine solche erfolgt in der Regel in mehreren Schritten, die zusammen eine Ungleichungskette der folgenden Art ergeben:

$$\boxed{|a_n| \stackrel{(1)}{\leq} |\bar{a}_n| \stackrel{(2)}{\leq} |\bar{a}_N| \stackrel{(3)}{<} \varepsilon.}$$

Dabei sind die einzelnen Schritte wie folgt zu interpretieren

(1) steht für eine **vorsichtige Vergrößerung** meist mit dem Zweck, den Rechenausdruck zu vereinfachen. "Vorsichtig" besagt: Dieser Schritt ist so auszuführen, dass die beiden nachfolgenden Schritte erfolgreich realisiert werden können.

(2) Steht für eine **Monotonieabschätzung**. Dabei wird Gültigkeit für alle n mit $n \geq N$ verlangt.

(3) schließlich ist als **Bestimmungsgleichung für N** zu interpretieren. Dabei ist $\varepsilon > 0$ als äußerer Parameter anzusetzen. Die gesuchte Strategiefunktion ist eine Lösung dieser Bestimmungsgleichung. Hat man eine Lösung $\varepsilon \mapsto N(\varepsilon)$, dann ist auch jede Vergrößerung $M(\varepsilon)$ mit $M(\varepsilon) \geq N(\varepsilon)$ eine solche.

(2.5.16) Beachten Sie: Hat man die drei Schritte erfolgreich realisiert, dann hat man insgesamt die von der Konvergenzbedingung geforderte Ungleichung $|a_n| < \varepsilon$ für alle $n \geq N(\varepsilon)$. Hat a den Grenzwert A , so ist überall a_n durch $a_n - A$ zu ersetzen. Das Vorgehen ist dasselbe. Die zu erfüllende Endbedingung lautet $|a_n - A| < \varepsilon$ für $n \geq N(\varepsilon)$. Vielfach kommt dann noch ein Schritt (0) hinzu, der in der Explikation von $|a_n - A|$ für den betrachteten Fall besteht.

(2.5.17) Das Ganze kann man wie folgt zusammenfassen - **und das ist dann die übliche Formulierung und Forderung**

Zu jedem $\varepsilon > 0$ gibt es ein $N=N(\varepsilon)$, so dass $|a_n-A|<\varepsilon$ für alle $n \geq N(\varepsilon)$ gilt.

(2.5.18) 1. **Beispiel:**

$a_n = \frac{n^3+2}{3n^5+2n+4}$ Zu beweisen: a ist Nullfolge.

$ a_n = a_n = \frac{n^3+2}{3n^5+2n+4}$	$\leq \frac{n^3+2}{3n^5}$ $\leq \frac{3n^3}{3n^5} = \frac{1}{n^2}$ $\leq \frac{1}{N^2} < \varepsilon$	a_n ist positiv. Betrag kann entfallen Nenner verkleinern. Gehört zu (1) Gilt ab $n=1$. Zähler vergrößern. Noch (1). Für $n \geq N$. Monotonie von $\frac{1}{n^2}$. Schritte (2) und (3).
<div style="border: 1px solid black; padding: 2px; display: inline-block;"> $\frac{2}{3} \frac{1}{N^2} \stackrel{?}{\leq} \varepsilon$ </div>	$N \geq \sqrt{\frac{2}{\varepsilon}}$	Löst die Bestimmungsgleichung. Die 2 sichert $<$ statt \leq .
<p>Ergebnis: Mit $N(\varepsilon) = \left\lceil \sqrt{\frac{1}{\varepsilon}} \right\rceil + 1$ gilt $0 < a_n = a_n < \varepsilon$ für $n \geq N(\varepsilon)$</p>		

In diesem Fall haben wir am Ende noch die kleinste zulässige Lösung angegeben. Zur Erinnerung $[x]$ bezeichnet die größte ganze Zahl unterhalb x .

Mit etwas Übung werden solche Abschätzungen Routine. Natürlich kann und wird man die drei Schritte fallspezifisch modifizieren. Und wenn man keine Konvergenz hat, wird man mit der Negation der Aussage oder einer geeigneten Modifikation arbeiten.

(2.5.19) **2. Beispiel.** $a_n = q^n$, wobei $q \in \mathbb{C}$ äußerer Parameter ist. Für $|q| < 1$ ist das eine Nullfolge. Wie sieht in diesem Fall die Abschätzung aus?

$|a_n-0| \stackrel{(0)}{=} |q|^n \stackrel{(2)}{\leq} |q|^N \stackrel{(3)}{<} \varepsilon$ für $N(\varepsilon) = \left\lceil \frac{|\ln \varepsilon|}{|\ln |q||} \right\rceil + 1$ für $\varepsilon, |q| < 1$

$N \ln |q| < \ln \varepsilon \quad |\ln \varepsilon| < N |\ln |q|| \quad N > \frac{|\ln \varepsilon|}{|\ln |q||} = \frac{\ln \varepsilon}{\ln |q|}$ für $\varepsilon, |q| < 1$

Die obere Zeile: Im Schritt (0) wird hier nur $|a_n - 0|$ für den betrachteten Fall expliziert. Der Schritt (1) ist unnötig. Schritt (2) benutzt die Monotonie der Funktion $x \mapsto |q|^x$. Wichtig: Für $|q| < 1$ ist das eine monoton fallende Funktion. Wegen $n \geq N$ wird genau das benötigt. Bei Schritt (2) wird meist auf bekannte Monotonieeigenschaften von Funktionen zurückgegriffen und die dazu benötigten Rechenausdrücke werden im ersten Schritt angesteuert. Für (3) ist erneut eine optimale Abschätzung angegeben. Diese wird in der zweiten Zeile hergeleitet. Der Logarithmus wächst monoton. Das gibt die erste Ungleichung. Für $\varepsilon < 1$ sind beide Logarithmen negativ! Daher gibt der Übergang zum Betrag die 2. Ungleichung (Wichtig!). Daraus

folgt dann die Auflösung nach N, die man als Strategiegleichung bezeichnen kann. Verlangt man schließlich, dass N ganzzahlig ist, erhält man die gegebene Abschätzung.

(2.5.20) Bedeutet die Bedingung $\varepsilon < 1$ eine Einschränkung? Nein! Denn ist $\varepsilon \mapsto N_0(\varepsilon)$ eine Strategiefunktion für $0 < \varepsilon < \varepsilon_0$, dann ist

$$N(\varepsilon) = \begin{cases} N_0(\varepsilon) & \text{für } 0 < \varepsilon < \varepsilon_0 \\ N_0(\varepsilon_0) & \text{für } \varepsilon_0 \leq \varepsilon \end{cases}$$

offensichtlich Strategiefunktion für alle $\varepsilon > 0$.

(2.5.21) **3. Beispiel:** Die Abschätzung für die Folge der geometrischen Reihe folgt jetzt leicht. Erneut müssen wir zunächst den Explikationsschritt (0) ausführen ("D.h. hier..."). Dafür steht uns in diesem Fall die gegebene Summenformel zur Verfügung:

$$\begin{aligned} |q_n - A| &\stackrel{(0)}{=} \left| \sum_{k=0}^n q^k - \frac{1}{1-q} \right| \stackrel{(0)}{=} \left| \frac{-q^{n+1}}{1-q} \right| \stackrel{(0)}{=} \frac{|q|^{n+1}}{|1-q|} \stackrel{(1)}{\leq} \frac{|q|^n}{1-|q|} \\ &\stackrel{(2)}{<} \frac{|q|^N}{1-|q|} \stackrel{(3)}{<} \varepsilon \quad \text{mit} \quad N(\varepsilon) = \left\lceil \frac{|\ln(\varepsilon(1-|q|))|}{|\ln|q||} \right\rceil + 1 \quad \varepsilon, |q| < 1. \end{aligned}$$

Die Gültigkeit von Schritt (1) gerade auch für komplexes q überlegt man sich leicht geometrisch: Der gegebene Wert gibt den kleinstmöglichen Nenner. Die Ungleichung läßt sich für $\varepsilon < 1$ leicht lösen durch $N > \frac{|\ln(\varepsilon(1-|q|))|}{|\ln|q||}$

(2.5.22) **4. Beispiel:** Jetzt wandeln wir unser Rechenschema fallspezifisch ab, um zu zeigen, dass die zur harmonischen Reihe gehörige Folge tatsächlich über alle Grenzen wächst. Vgl. (2.2.3-4). D.h. wir zeigen, dass es zu jeder Zahl $K > 0$ ein N gibt, so dass $|a_n| = \left| \sum_{k=1}^n \frac{1}{k} \right| > K$ wird, sobald nur $n \geq N$ wird. In unserer Veranschaulichung: Man zieht eine Parallele zur n-Achse in der Höhe K. Oben in (2.2.4) war $K=3.6$. Dann gibt es einen N-Wert, von dem ab der Graph sicher immer oberhalb dieser Geraden verläuft.

Zur Ausführung betrachten wir folgende Abschätzungen:

$$\begin{aligned} \frac{1}{3} + \frac{1}{4} &> \frac{1}{4} + \frac{1}{4} = \frac{2}{4} = \frac{1}{2} \\ \frac{1}{5} + \frac{1}{6} + \frac{1}{7} + \frac{1}{8} &> \frac{4}{8} = \frac{1}{2} \quad \text{usw.} \quad \left(\sum_{s=2^{i-1}+1}^{2^i} \frac{1}{s} \right) \end{aligned}$$

Damit folgt mit $2^N > n$, also $N > \frac{\ln n}{\ln 2}$:

$$|a_n| \stackrel{(0)}{=} \left| \sum_{k=1}^n \frac{1}{k} \right| \stackrel{(1)}{\geq} \sum_{k=1}^{2^N} \frac{1}{k} = 1 + \sum_{r=1}^N \left(\sum_{s=2^{r-1}+1}^{2^r} \frac{1}{s} \right) \stackrel{(1)}{\geq} 1 + \sum_{r=1}^N \frac{1}{2} = 1 + \frac{N}{2} \stackrel{(2)}{>} K.$$

Die zu erfüllende Ungleichung ergibt $\boxed{N > 2K-1}$. Also sollte man $\frac{\ln n}{\ln 2} > 2K-1$ wählen. und das heißt $\boxed{n > 2^{2K-1}}$. Für $K=10$ etwa ergibt sich $2^{19} = 524288$ und $\sum_{k=1}^{2^{19}} \frac{1}{k} \approx 13.747$. Natürlich ist in Schritt (1) echt vergrößert worden.

2.2.5b Die Grenzwertsätze

(2.5.23) Damit beschließen wir die Beispiele für die explizite Konstruktion von Strategiefunktionen oder Strategiegleichungen und kommen zum zweiten Problemkreis aus (2.5.14), dem der Grenzwertsätze.

(2.5.24) Beispielsweise gilt der folgende **Satz**, wobei wir daran erinnern, dass abgeleitete Folgen wie $\alpha a + \beta b$ oder ab durch komponentenweise Verknüpfung definiert sind. Etwa $(ab)_n = a_n b_n$. Vgl. Kap3.

Sei $\mathcal{K}(\mathbb{N}, \mathbb{C})$	die Menge aller konvergenten Folgen und $a, b \in \mathcal{K}(\mathbb{N}, \mathbb{C})$ mit $A = \lim(a)$ und $B = \lim(b)$. Weiter sei $\alpha, \beta \in \mathbb{C}$.
Dann	ist auch $\alpha a + \beta b \in \mathcal{K}(\mathbb{N}, \mathbb{C})$ mit $\lim(\alpha a + \beta b) = \alpha A + \beta B$ und $ab \in \mathcal{K}(\mathbb{N}, \mathbb{C})$ mit $\lim(ab) = AB$. Oder $\lim(\alpha a + \beta b) = \alpha \lim a + \beta \lim b$ und $\lim(ab) = (\lim a)(\lim b)$.

D.h. der Grenzwert der abgeleiteten Folge existiert und Verknüpfung und Grenzwertbildung vertauschen. Etwa $\lim(ab) = (\lim a)(\lim b)$.

Wir **beweisen** den ersten Teil des Satzes:

▼ Laut Voraussetzung gibt es zu a und b Strategiefunktionen N_a und N_b . O.b.d.A dürfen wir $\alpha, \beta \neq 0$ annehmen. Wir benötigen eine Strategiefunktion für $c = \alpha a + \beta b$. Dazu argumentieren wir wie folgt

$$|c_n - C| \stackrel{(0)}{=} |(\alpha a_n + \beta b_n) - (\alpha A + \beta B)| \stackrel{(0)}{=} |\alpha(a_n - A) + \beta(b_n - B)|$$

$$\stackrel{(1)}{\leq} |\alpha|(a_n - A)| + |\beta|(b_n - B)| \stackrel{(V)}{<} |\alpha| \cdot \frac{\varepsilon}{2|\alpha|} + |\beta| \cdot \frac{\varepsilon}{2|\alpha|} = \varepsilon.$$

In (1) wird die Dreiecksungleichung benutzt. Die Voraussetzungen gehen in (V) ein: N_a ist die vorhandene Strategiefunktion für a. Wählt man daher $n \geq N_a(\frac{\varepsilon}{2|\alpha|})$ wird $|(a_n - A)| < \frac{\varepsilon}{2|\alpha|}$ wie in der Rechnung benutzt. Im 2. Summanden ist entsprechend $n \geq N_b(\frac{\varepsilon}{2|\alpha|})$ zu fordern.

Beide Bedingungen sind erfüllt für $N_c(\varepsilon) = \max(N_a(\frac{\varepsilon}{2|\alpha|}), N_b(\frac{\varepsilon}{2|\alpha|}))$. Das ist unsere gesuchte Strategiefunktion.

▼ Für die Produktfolge kann man entsprechend folgende Abschätzung durchführen:

$$|c_n - C| = |a_n b_n - AB| = |(a_n - A)b_n + A(b_n - B)| \leq |a_n - A||b_n| + |A|(b_n - B)|$$

Jetzt benötigt man noch einen Hilfssatz, der besagt:

(2.5.25) **Hat** b einen Grenzwert, **dann** ist Bild(b) beschränkt. D.h. es gibt $C > 0$ mit $|b_n| \leq C$.

- Begründen Sie die einzelnen Schritte und führen Sie den Beweis unter Verwendung des Hilfssatzes zu Ende.
- Beweisen Sie den Hilfssatz. (Hinweis: Es gibt eine Strategiefunktion N_b , wählen Sie $n_0 = N_b(1)$ und nutzen Sie für $n > n_0$ die folgende Abschätzung

$$|b_n| = |b_n - b_{n_0} + b_{n_0}| \leq |b_n - b_{n_0}| + |b_{n_0}|.$$

- Welche weiteren Grenzwertsätze dieser Art erwarten Sie? Schauen Sie in einem Analysisbuch nach, welche es gibt und welche zusätzlichen Bedingungen jeweils noch erforderlich sind.

(2.5.26) **Zusammenstellung der Grenzwertsätze:** Sofern nichts anderes gesagt wird, seien in der nachfolgenden Liste seien a,b,c... Folgen mit Grenzwerten $A = \lim a_n, B = \lim b_n \dots$ Und \rightarrow steht für "(die Folge) konvergiert gegen". **Dann gilt:**

$a+b \rightarrow A+B$	oder	$\lim(a+b) = \lim a + \lim b$	(G1)
$\alpha a \rightarrow \alpha A$ für $\alpha \in \mathbb{C}$	oder	$\lim(\alpha a) = \alpha \lim a$	(G2)
$ab \rightarrow AB$	oder	$\lim(ab) = \lim a \cdot \lim b$	(G3)
$a \rightarrow 0$ und b beschränkt	dann:	$ab \rightarrow 0$	(G4)
t Teilfolge von a mit $a \rightarrow A$	dann:	$t \rightarrow A$	(G5)
$a \rightarrow A \neq 0$ und $a_n \neq 0$ ab n_0	dann:	$\frac{1}{a} \rightarrow \frac{1}{A}$	(G6)
$a \rightarrow A$ und f stetige Funktion	dann:	$f(a) \rightarrow f(A)$	(G7)
$a \rightarrow A$	dann:	$ a \rightarrow A $	(G8)
$a_n \leq b_n$ oder $a_n < b_n$ ab n_0	dann:	$A \leq B$ (nicht $A < B!$)	(G9)

(2.5.27) Eine einfache nützliche Folgerung aus (G.9) formulieren wir noch.

Sei $a = (\mathbb{N}, n \rightarrow a_n, \mathbb{R})$ eine konvergente Folge mit Grenzwert A, die folgende Eigenschaft besitzt:
Es gibt ein $N_0 \in \mathbb{N}$, so dass $a_n < K$ für alle $n \geq n_0$.

Dann gilt $A = \lim_n a_n \leq K$.

Zum Beweis wähle in (G.9) $b_n = K$, d.h. b ist konstante Folge.

(2.5.28) Beachten Sie unbedingt, dass in (G.9) beim Grenzübergang aus $<$ ein \leq werden kann (nicht muss!). Wählen Sie als Beispiel $a_n = -\frac{1}{n}$. Jedes Einzelglied ist echt negativ, der Grenzwert jedoch Null. Der Satz gilt entsprechend für $>$ und \geq . Dazu genügt es, von $n \rightarrow a_n$ zur Folge $n \rightarrow -a_n$ überzugehen.

Aus (G.9) folgt ein nützliches Konvergenzkriterium, das **Sandwichkriterium:**

Es sei (a_n) ein Folge, deren Konvergenzverhalten interessiert. b und c seien konvergente Folgen.
Man habe zwei Abschätzungen $b_n \leq a_n$ und $a_n \leq c_n$ für alle $n \geq N_0$ für ein N_0 –
Weiter sei $\lim_n b_n = A$ und $\lim_n c_n = A$,
Dann ist auch a konvergent mit demselben Grenzwert. Also $\lim_n a_n = A$.

Häufig liegen b und c in der Form $b_n = a_n + p_n$ und $c_n = a_n + q_n$ vor, wobei p und q Nullfolgen sind!

(2.5.29) **Zusammenfassung:** Grenzwertdiskussionen wird man meist so führen, dass man beide Typen von Argumentationen zusammen benutzt: Mit Hilfe der Grenzwertsätze führt man die Diskussion auf das Verhalten einfacherer Folgen zurück. Und diese sind entweder bekannt oder aber sie müssen explizit behandelt werden.

(2.5.30) Als nächstes geben wir eine Zusammenstellung wichtiger konvergenter Folgen, deren Verhalten man kennen sollte, die häufig als Folgen oder als elementare Bestandteile anderer Folgen auftauchen.

	a_n	Folgenbeginn	Grenzwert	Funktion ($x \rightarrow \infty$)	Reihe		
(F.1)	$\frac{1}{n^s}$ $s > 0$	$\frac{1}{1^s}, \frac{1}{2^s}, \frac{1}{3^s}, \dots$	0	$x \mapsto x^{-s}$	$\sum \frac{1}{n^s}$		
(F.2)	$\sqrt[n]{a^p}$ $a > 1$	$a^p, \sqrt[2]{a^p}, \sqrt[3]{a^p}, \sqrt[4]{a^p}$	1	$x \mapsto a^{\frac{x}{n}}$			
(F.3)	$\sqrt[n]{n^r}$	$1^r, \sqrt[2]{2^r}, \sqrt[3]{3^r}, \sqrt[4]{4^r}$	1	$x \mapsto e^{\frac{x}{n} \ln x}$			
(F.4)	q^n $ q < 1$	$1, q, q^2, q^3, q^4, \dots$	0	$x \mapsto q^x = e^{x \ln q}$	$\sum q^n$		
(F.5)	$\frac{n^k}{z^n} = n^k q^n \rightarrow 0$ $k \in \mathbb{N}, q < 1, z > 1$	$0, q, 2^k q^2, 3^k q^3, \dots$	0	$x \mapsto e^{k \ln x + x \ln q}$			
(F.6)	$\frac{a^n}{n!}$	$1, \frac{a}{1}, \frac{a^2}{2}, \frac{a^3}{6}, \frac{a^4}{24}, \dots$	0	??	$\sum \frac{a^n}{n!}$		
(F.7)	$(1 + \frac{x}{n})^n$		e^x	$x \mapsto (1 + \frac{x}{n})^n$			

(2.5.31) Bei einem Teil der Folgen ist die aus den Gliedern rekursiv gebildete Summe - die *unendliche Reihe* - von größerer Bedeutung als die Folge selbst. Diese Fälle sind in der letzten Spalte angedeutet. Sie werden später ausführlich besprochen.

Diskussion der Beispiele:

▼(2.5.32) (F1)

$$|a_n - 0| = \frac{1}{n^s} \leq \frac{1}{N^s} < \varepsilon \quad \left(\frac{1}{\varepsilon} \right)^{\frac{1}{s}} < N$$

□ Kommentieren Sie die einzelnen Schritte. Nutzen Sie die Kenntnis des Verhaltens der Funktion ($[1, \infty[, x \mapsto x^{-s}, \mathbb{R}$). Vgl. (2.3.1).

▼(2.5.33) (F2) Hier gewinnen wir zunächst zwei Abschätzungen. Sei zuerst $a \geq 1$.

$$\text{Aus } 1 \leq a \text{ folgt mit Monotonie } \left(\frac{p}{n} < 0 \right) \quad 1 \leq \sqrt[n]{a^p} = a_n$$

Und mit Hilfe der AGM-Ungleichung erhält man

$$a_n = \sqrt[n]{a^p} = \sqrt[n]{1^{n-p} a^p} < \frac{(n-p)1 + pa}{n} = 1 + \frac{p}{n}(a-1) \quad a > 1.$$

Für $a=x+1$ ergibt das übrigens die Bernoullische Ungleichung. Zusammen hat man

$$1 \leq a_n < 1 + \frac{p}{n}(a-1)$$

Das Sandwichkriterium liefert $\lim_n a_n = 1$. Oder $\sqrt[n]{a^p} \xrightarrow{n \rightarrow \infty} 1$.

Und wie steht es mit $0 < a < 1$? Betrachte dazu $b_n = \frac{1}{a_n} = \sqrt[n]{\left(\frac{1}{a}\right)^p}$ mit $\frac{1}{a} > 1$. Das erste Resultat und der Grenzwertsatz (G6) ergibt das Resultat auch für $0 < a < 1$. Nur dass man es für $a > 1$ mit einer monoton fallenden und bei $a < 1$ mit einer monoton wachsenden Folge zu tun hat.

□ Begründen Sie das Resultat über die Restriktionsmethode, also über die Diskussion der Funktion $x \mapsto a^{\frac{x}{n}}$.

▼ (2.5.34) (F3) $a_n = \sqrt[n]{n^r} = n^{\frac{r}{n}}$. Erneut arbeiten wir mit dem Sandwichkriterium. Sicherlich ist $1 \leq a_n$ für $(n \geq 1)$. Aus der **AGM-Ungleichung** folgt für $n > 2r$

$$\sqrt[n]{n^r} = \sqrt[n]{(\sqrt{n})^{2r} 1^{n-2r}} < \frac{1}{n} (2r\sqrt{n} + n - 2r) < 1 + \frac{2r}{\sqrt{n}}$$

Damit folgt der behauptete Grenzwert. $\lim_{n \rightarrow \infty} \sqrt[n]{n^r} = 1$.

▼ (2.5.35) (F4) $a_n = q^n$ mit $|q| < 1$ und $q \in \mathbb{C}$. Es sei $q = |q|e^{i\alpha}$ in Polardarstellung. Also $c_n = |q|^n e^{in\alpha}$. Wir müssen zeigen, dass $n \mapsto |q|^n$ eine Nullfolge ist. Dann gibt das in (2.5.13) beschriebene Vorgehen:

$$r^n \leq r^N < \varepsilon \quad N \ln r < \ln \varepsilon \quad (\text{falls } 0 < r < 1) : \quad N |\ln r| > |\ln \varepsilon|$$

$$N > \frac{|\ln \varepsilon|}{|\ln r|} = \frac{\ln \frac{1}{\varepsilon}}{\ln \frac{1}{r}} \quad \mathbf{0 < r < 1 \text{ zu fordern}} \quad 0 < \varepsilon < 1 \text{ wählbar}$$

Für $r > 1$ ist die Bedingung nicht erfüllbar!

▼(F5) (2.5.36) $a_n = n^p q^n$ mit $|q| < 1$ und $p \geq 1$. Hier haben wir es mit der Restriktion der reellen Funktion $x \mapsto x^p |q|^x = x^p e^{x \ln |q|} = e^{p \ln x + x \ln |q|}$ zu tun. Da $\ln |q| < 0$ ist, geht die Funktion für $x \rightarrow \infty$ und damit die Folge nach Null. wie angegeben. $\lim_{n \rightarrow \infty} n^p q^n = 0$

Man kann das etwas aufwendiger auch **direkt beweisen**: Zunächst zeigen wir $|n^r q^n| \leq 1$ oder $\sqrt[n]{n^r} \leq \frac{1}{|q|} = 1 + h$ für große n und $r \geq 1$ ganz. (F3) zeigt, dass für $n > 2r$

$$\sqrt[n]{n^r} < 1 + \frac{2r}{\sqrt{n}} \quad \text{Also } \frac{2r}{\sqrt{n}} \leq \frac{2r}{\sqrt{N}} \leq h$$

$$\text{oder } N \geq \frac{4r^2}{h^2} = \frac{4r^2}{\left(\frac{1}{|q|} - 1\right)^2}.$$

Nun ist für $n > N_q$

$$\sqrt[n]{n^r} < 1 + \frac{2r}{\sqrt{n}} < 1 + h = q^{-1}.$$

$$n^r < q^{-n} \quad \text{und} \quad n^r q^n < 1 \quad \text{für } n > N_q$$

Wähle $r = p + k$. Also $n^{p+k} q^n < 1$ und damit $n^p q^n < \frac{1}{n^k}$.

▼(2.5.37) Noch (F.5). Die Abschätzung läßt sich verbessern und ergibt dann ein verbessertes Resultat.:

$$\sqrt[n]{n} = \left(n^{\frac{1}{4}} n^{\frac{1}{4}} n^{\frac{1}{4}} n^{\frac{1}{4}} 1^{n-4} \right)^{\frac{1}{n}} \leq \frac{1}{n} \left(4n^{\frac{1}{4}} + n - 4 \right) = \frac{4}{n^{\frac{3}{4}}} + 1 - \frac{4}{n} < \frac{4}{n^{\frac{3}{4}}} + 1$$

$$0 < \sqrt[n]{n} (\sqrt[n]{n} - 1) < \frac{4}{n^{\frac{3}{4}}} \quad \text{Also } \lim_{n \rightarrow \infty} \sqrt[n]{n} (\sqrt[n]{n} - 1) = 0$$

Man sieht das auch über die Entwicklung

$$n^{\frac{1}{n}} = e^{\frac{1}{n} \ln n} = 1 + \frac{1}{n} \ln n + ..$$

$$\left(n^{\frac{1}{n}} - 1 \right) \sqrt[n]{n} = \frac{1}{\sqrt[n]{n}} \ln n + ..$$

oder allgemeiner für $0 \leq \varepsilon < 1$

$$\lim_{n \rightarrow \infty} n^\varepsilon (\sqrt[n]{n} - 1) = 0 \quad \text{für } 0 \leq \varepsilon < 1$$

▼ (F.6) (2.5.38) $c_n = \frac{a^n}{n!}$ mit $\lim_{n \rightarrow \infty} c_n = 0$ Zunächst einmal gibt es für jedes a eine $n_0 \in \mathbb{N}$, so dass $|a| < n_0$ ist. Wähle jetzt $n > n_0$. Dann schreiben wir c_n wie folgt

$$c_n = \frac{a}{1} \cdot \frac{a}{2} \cdot \frac{a}{3} \cdots \frac{a}{n_0} \cdot \frac{a}{n_0 + 1} \cdots \frac{a}{n}$$

Die ersten n_0 Faktoren ergeben eine feste Zahl K . Alle weiteren Faktoren sind ≤ 1 . Daher können wir wie folgt vergrößern

$$|c_n| \stackrel{(1)}{\leq} |K| \cdot \frac{a}{n} \stackrel{(1)}{\leq} |K| \cdot \frac{a}{N} < \varepsilon.$$

Das reicht offensichtlich

(2.5.39) (F.7) wird in Kap.2.3.6 ausführlich behandelt.

□ Wogegen konvergieren $\frac{1}{\sqrt[n]{n}}$ und $\sqrt[n]{n+1}$?

2.2.6 Folgenverhalten (II): Divergente Folgen und Häufungswerte von Folgen

Bisher haben wir uns mit Folgen befaßt, die ein besonders gutartiges Verhalten zeigten, bei denen sich die Glieder einem bestimmten Grenzwert immer mehr annäherten, den konvergenten Folgen. Was ist mit den anderen, den nicht konvergenten Folgen? Läßt sich über deren Verhalten auch etwas aussagen?

(2.6.1) Zunächst einmal gibt es Folgen wie $n \mapsto n$ oder $n \mapsto -n^2$, die in mancherlei Hinsicht den konvergenten Folgen ähnlich sind. Beide Folgen nähern sich einem "Grenzwert", nur dass dieser Wert keine reelle Zahl ist, sondern (plus oder minus) "unendlich". Folgen dieses Typs nennt man "bestimmt divergent". Also

Definition: Eine Folge $a = (\mathbb{N}, n \mapsto a_n, \mathbb{R})$ heißt *bestimmt divergent (nach $+\infty$)*, wenn es zu jeder Zahl $K > 0$ ein $N = N(K)$ gibt, so dass $a_n > K$ gilt. Analog definiert man *bestimmt divergent nach $-\infty$* .

Etwas lax kann man sagen: Bestimmt divergente Folgen konvergieren gegen $\pm\infty$.

(2.6.2) Welche Folgen werden noch immer nicht erfaßt? Beispielsweise $n \mapsto (-1)^n$ und $n \mapsto (-1)^n n$.

(2.6.3) Ein ganz wesentliches Merkmal dieser verbleibenden Folgen besteht darin, dass sie sich vielfach nicht **einem** Wert nähern, sondern dass ihre Glieder sich regelmäßig oder auch unregelmäßig abwechselnd **mehreren Werten** nähern können. Diese Eigenschaft wird durch folgende Definition erfaßt. Zuvor sei an die Definition von *Teilfolge* in (2.1.3) erinnert.

Definition: Sei $a = (\mathbb{N}, n \mapsto a_n, \mathbb{C})$ Folge. Dann heißt $H \in \mathbb{C}$ ein *Häufungswert von a* , wenn es eine Teilfolge h von a gibt, die gegen H konvergiert. Weiter sei $H(a)$ die Menge aller Häufungswerte von a .

(2.6.4) Eine bestimmt divergente Folge wie $n \mapsto n^2$ hat **keinen** Häufungswert (denn jede Teilfolge ist erneut bestimmt divergent). Eine konvergente Folge mit $\text{lima} = A$ hat ein **einelementiges** $H(a)$. Nämlich $H(a) = \{A\}$. Der Grenzwert ist der einzige Häufungswert. Wie steht es mit $n \mapsto (-1)^n$? Hier hat man die beiden konvergenten Teilfolgen $n \mapsto a_{2n} = +1$ und $n \mapsto a_{2n+1} = -1$ und damit die beiden Häufungswerte ± 1 . Also $H(a) = \{+1, -1\}$ **zweielementig**.

(2.6.5) Wie steht es mit $n \mapsto \sin(n)$? Man kann zeigen, dass hier alle Zahlen zwischen -1 und $+1$ Häufungswert sind. D.h. man kann für jedes solche r eine Indexfolge $k \mapsto n_k$ finden, so dass $k \mapsto \sin(n_k)$ gegen r konvergiert. Damit ist in diesem Fall $H(a) = [-1, 1]$. (Nehmen Sie etwa $r=1$. Da n nie ein Vielfaches von π ist, wird der Wert 1 nie angenommen. Schreibt man jedoch $n = (2k+1)\pi + r_n$ mit natürlichem k und möglichst kleinem $|r_n|$, dann findet man n -Werte, für die $|r_n|$ beliebig klein wird.)

□ Schreiben Sie für den Computer ein entsprechendes Suchprogramm (zur Bestimmung der ersten glieder einer derartigen Teilfolge).

(2.6.6) Häufungswerte müssen nicht immer (als Werte) angenommen werden. Sie können es wie im Fall $n \mapsto (-1)^n$, müssen es aber nicht. Betrachten Sie etwa $n \mapsto (-1)^n (1 + \frac{1}{n+1})^n$.

(2.6.7) Neben dem Sandwichkriterium ist auch das folgende Kriterium nützlich:

Monotoniekriterium:	Es SEI $a: \mathbb{N} \rightarrow \mathbb{R}$ nach oben beschränkt und monoton wachsend. (Oder nach unten beschränkt und monoton fallend.)
----------------------------	---

Dann

ist a konvergent mit Grenzwert $A = \sup(\text{Bilda})$ im wachsenden und $A = \inf(\text{Bilda})$ im fallenden Fall.

Beweis: Es genügt den wachsenden Fall zu betrachten. Andernfalls geht man zur Folge $-a$ über. Nach dem Vollständigkeitsaxiom existiert das Supremum S von $Bild a$, die kleinste obere Schranke in der Menge aller Werte. Setze $\varepsilon_k = \frac{1}{k}$. Dann gibt es mindestens ein Folgenglied a_{n_k} mit $\boxed{\frac{1}{k} > S - a_{n_k} \geq 0}$. Andernfalls könnte man S sicher um $\frac{1}{2}\varepsilon_k$ verkleinern. Da a monoton wächst, erfüllen auch alle a_n mit $n \geq n_k$ die Ungleichung. Daher darf man $n_{k+1} > n_k$ wählen. Die Teilfolge $k \mapsto a_{n_k}$ konvergiert daher gegen S , etwa mit der Strategiefunktion $N(\varepsilon) = \lceil \frac{1}{\varepsilon} \rceil + 1$. Da a monoton wächst, liegen auch alle a_n zwischen $S - \frac{1}{k}$ und S und die Folge gesamte ist konvergent.

(2.6.8) Der folgende allgemeine und wichtige Satz macht Aussagen über die Existenz von Häufungspunkten

Satz von Bolzano-Weierstraß
 Sei $a: \mathbb{N} \rightarrow \mathbb{C}$ eine beschränkte Folge.
 Dann existiert eine konvergente Teilfolge.

D.h. eine **beschränkte** Folge hat mindestens einen Häufungswert. $n \mapsto n^2$ hat keinen Häufungspunkt, ist aber auch unbeschränkt.

Beweis: Wir beginnen mit dem Fall $\mathbb{N} \rightarrow \mathbb{R}$.

- Sei $(\mathbb{N}, n \mapsto a_n, \mathbb{R})$ eine beschränkte Zahlfolge. D.h. es gibt $S > 0$ mit $-S \leq a_n \leq S$ für alle n .
- Sei weiter \mathcal{G} die folgende Menge (*Menge der Gipfelpunkte*)

$$\mathcal{G} = \{(n, a_n) \mid a_n \geq a_k \text{ für alle } k > n\} \subset Graph(a).$$

Immer gilt: Wenn (n, a_n) und $(m, a_m) \in \mathcal{G}$ mit $n < m$, dann ist $a_n \geq a_m$. Die durch \mathcal{G} definierte (eventuell endliche) Teilfolge g von a ist **monoton fallend**.

- Es gibt zwei Möglichkeiten: \mathcal{G} ist **entweder endlich oder unendlich**.

Beispiele für unendliches \mathcal{G} :

$$a = (\boxed{1}, -1, -1, \frac{1}{2}, \boxed{\frac{1}{2}}, -\frac{1}{2}, -\frac{1}{2}, -\frac{1}{2}, \frac{1}{3}, \frac{1}{3}, \boxed{\frac{1}{3}}, -\frac{1}{3}, -\frac{1}{3}, -\frac{1}{3}, -\frac{1}{3}, \frac{1}{4}, \dots)$$

$$\mathcal{G}(a) = ((0, 1), (3, \frac{1}{2}), (8, \frac{1}{3}), (15, \frac{1}{4}), \dots)$$

Oder b mit $b_n = \frac{1}{n+1}$. Dann ist $Graph(b) = \mathcal{G}(b)$. Ist \mathcal{G} unendlich, dann liefert der Monotoniesatz damit sofort eine konvergente Teilfolge.

- **Sei umgekehrt \mathcal{G} endlich.** D.h. es gibt N_0 mit folgender Eigenschaft: Zu jedem $n_0 \geq N_0$ gibt es $n_1 > n_0$ mit $a_{n_1} > a_{n_0}$. (Etwa in der Folge c mit $c_n = n$.)

Das gibt folgende Mengen:

$$\mathcal{H}_{n_0} = \{n \mid n \in \mathbb{N}, n > n_0, a_n > a_{n_0}\}$$

Damit bilden wir die Teilfolge $t_0 = a_{N_0}$. Danach wählen wir $n_1 \in \mathcal{H}_{N_0}$ und setzen $t_1 = a_{n_1}$. Dann wählen wir $n_2 \in \mathcal{H}_{n_1}$ und setzen $t_2 = a_{n_2} > a_{n_1}$. Usw. Da all diese Mengen Elemente enthalten, die Wahl also möglich ist und stets $t_{i+1} > t_i$ gilt, ist das eine monoton wachsende Folge.

- Damit gibt es in jedem Fall eine monotone beschränkte und damit nach dem Monotoniesatz konvergente Teilfolge.
- Jetzt sei $a: \mathbb{N} \rightarrow \mathbb{C}$. Es gilt $a_n = x_n + iy_n$ mit $x_n, y_n \in \mathbb{R}$. Es gilt $|x_k| \leq |x_k + iy_k|$. Die so abgeleitete beschränkte **reelle** Folge $n \mapsto x_n$ hat dann eine konvergente Teilfolge. $k \mapsto t_k = x_{n_k}$. Bilde die zugehörige Folge $k \mapsto y_{n_k}$ der Imaginärteile, die erneut beschränkt ist. Sie besitzt auch eine konvergente Teilfolge $r \mapsto s_r = y_{n_{k_r}}$. Die entstehende Folge $r \mapsto x_{n_{k_r}} + iy_{n_{k_r}}$ ist in \mathbb{C} konvergent. Und das heißt nach Satz , dass der Grenzwert Häufungspunkt der Ausgangsfolge ist.

(2.6.9) Wir haben noch die Folgerung:

Korollar: Jede reelle beschränkte Zahlfolge hat einen größten und einen kleinsten Häufungspunkt. ($\overline{\lim}_n a_n$ und $\underline{\lim}_n a_n$)

□ Bestimmen Sie $\overline{\lim}_n a_n$ und $\underline{\lim}_n a_n$ für $n \mapsto \sin(n)$.

2.2.7 Cauchyfolgen

(2.7.1) Die wichtigsten Folgen sind die konvergenten Folgen. Bei ihnen nähern sich die Folgenglieder einem Grenzwert an, der dann im Sinne unserer einleitenden Bemerkungen als das Resultat von abzählbar unendlich vielen Operationen anzusehen ist. Bisher haben wir so getan, als sei uns dieser Grenzwert bekannt. Wir sind etwa zur neuen Folge $n \mapsto a_n - A$ übergegangen und haben so getan, als könnten wir über $A = \lim(a)$ verfügen. **Was ist aber, wenn wir A nicht kennen, ja wenn wir nicht einmal wissen, ob a überhaupt konvergiert?** Dann scheinen alle unsere Überlegungen, insbesondere die zu den Strategiefunktionen zusammenzustürzen.

(2.7.2) Manchmal kann man dann noch A als äußeren Parameter ansetzen und schauen, ob man für einen bestimmten Wert von A eine Strategiegleichung herleiten kann. Aber in vielen Fällen geht das nicht oder reicht nicht.

□ Beweisen Sie als Zwischenüberlegung: Eine Folge $\mathbb{N} \rightarrow \mathbb{C}$ kann höchstens einen Grenzwert haben.

(2.8.3) Wie kann man also bereits mit endlich vielen (wenn auch beliebig vielen) Schritten vorhersagen, was für Verhältnisse bei unendlich vielen Schritten entstehen? Zum Glück läßt sich das entwickelte Schema so abändern, dass es diese anspruchsvolle Frage beantwortet. Hierzu definieren wir (vgl. (2.5.17)):

Definition Eine Folge $a: \mathbb{N} \rightarrow \mathbb{C}$ heißt *Cauchyfolge*, wenn es zu jedem $\varepsilon > 0$ eine $N = N(\varepsilon)$ gibt, für das $|a_n - a_m| < \varepsilon$ gilt für alle $n, m \geq N(\varepsilon)$.

Wir sehen: Das ist die alte Bedingung, nur dass die Ungleichung $|a_n - A| < \varepsilon$ durch $|a_n - a_m| < \varepsilon$ ersetzt ist. In erstere geht der Grenzwert A ein, in der letzteren benötigt man nur Folgenglieder von a, wenn auch alle oberhalb von N.

(2.8.4) Ansonsten können wir alles übernehmen: Den Begriff der Strategiefunktion, aber auch die Methode zur Herleitung einer solchen. Beachten Sie: Zu einer konvergenten Cauchyfolge gehören jetzt **zwei** Strategiefunktionen. Eine, sagen wir N_K , für die Konvergenzabschätzung von $|a_n - A|$ und eine, sagen wir N_{CF} , für die Cauchyfolgenabschätzung von $|a_n - a_m|$.

(2.8.5) Zunächst müssen wir klären, welche Beziehung zwischen *konvergenter Folge* und *Cauchyfolge* besteht. Das erfordert zwei Richtungen:

Satz: Jede konvergente Folge ist Cauchyfolge.

Beweis: Sei a konvergent mit Grenzwert A und Strategiefunktion N_K . Dann gilt

$$|a_n - a_m| = |(a_n - A) - (a_m - A)| \leq |a_n - A| + |a_m - A| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon \quad \text{für } n, m \geq N_K\left(\frac{\varepsilon}{2}\right)$$

Das ist die gewünschte Abschätzung. Oder: Als Strategiefunktion für die Cauchyabschätzung dürfen wir wählen $N_{CF}(\varepsilon) = N_K\left(\frac{\varepsilon}{2}\right)$.

(2.8.6) Wie sieht es mit der anderen Richtung aus? Hier liefert uns das neue Konzept die Möglichkeit, in unsere Problem Diskussion einen wichtigen weiteren Freiheitsgrad einzubeziehen. Bisher haben wir Folgen $\mathbb{N} \rightarrow \mathbb{R}$ betrachtet. In vielen Fällen hat man aber Folgen $\mathbb{N} \rightarrow \mathbb{Q}$, also Folgen mit rationalen Gliedern. Etwa rationale Approximationen von $\sqrt{2}$ oder π . Sei etwa π_n die Zahl π nur bis zur n-ten Nachkommastelle. Etwa $\pi_2 = 3.14$ und $\pi_3 = 3.141$ usw. Dann haben wir $(\mathbb{N}, n \mapsto \pi_n, \mathbb{Q})$.

□ Zeigen Sie: Das ist eine Cauchyfolge und eine konvergente Folge mit Grenzwert π .

Wir haben damit eine Cauchyfolge $\mathbb{N} \rightarrow \mathbb{Q}$ die zwar konvergent ist. **Aber dieser Grenzwert liegt nicht mehr in der Wertemenge \mathbb{Q} dieser Folge!**

(2.8.7) Das gibt Anlass zu der folgenden Präzisierung unseres Begriffsystems:

Es sei $G \subset \mathbb{C}$ und $a: \mathbb{N} \rightarrow G$ eine Folge mit Werten in G .

Dann sollte man bei einem eventuellen Grenzwert unterscheiden, ob dieser auch in G liegt oder eventuell nur in \mathbb{C} . Wir haben bereits gesehen: Eine Cauchyfolge $\mathbb{N} \rightarrow \mathbb{Q}$ muss keineswegs einen Grenzwert in \mathbb{Q} haben.

(2.8.9) Aber es gilt:

Satz: Eine Cauchyfolge $\mathbb{N} \rightarrow \mathbb{R}$ hat immer einen Grenzwert in \mathbb{R}
 Eine Cauchyfolge $\mathbb{N} \rightarrow \mathbb{C}$ hat immer einen Grenzwert in \mathbb{C}

Beachten Sie Eine Folge $\mathbb{N} \rightarrow \mathbb{Q}$ oder auch $\mathbb{N} \rightarrow \mathbb{N}$ kann über Wertemengenerweiterung stets als Folge $\mathbb{N} \rightarrow \mathbb{R}$ interpretiert werden. Anders herum geht das nur unter Zusatzannahmen. So ist $n \mapsto \frac{1}{\pi^n}$ keine Folge $\mathbb{N} \rightarrow \mathbb{Q}$ oder gar $\mathbb{N} \rightarrow \mathbb{N}$.

(2.8.10) Zum Beweis benötigen wir das Vollständigkeitsaxiom der reellen Zahlen über den Satz von Bolzano-Weierstraß.

□ Zeigen Sie zunächst: Eine Cauchyfolge ist immer beschränkt.

Beweis: Jetzt argumentieren wir wie folgt: Da a beschränkt ist, gibt es nach Bolzano-Weierstraß eine konvergente Teilfolge $k \mapsto a_{n_k}$. Der Grenzwert sei A . Wir behaupten: A ist auch Grenzwert von a selbst. Dazu schätzen wir ab

$$|a_n - A| = |(a_n - a_{n_k}) + (a_{n_k} - A)| \leq |a_n - a_{n_k}| + |a_{n_k} - A|.$$

Jetzt wählen wir K so groß, dass für $k > K$ gilt $|a_{n_k} - A| < \frac{\varepsilon}{2}$. Da die Teilfolge konvergiert, geht das. Dann wählen wir N so groß, dass für $n, m \geq N$ die Cauchyabschätzung $|a_n - a_m| < \frac{\varepsilon}{2}$ gilt. Nach eventueller Vergrößerung von K , so dass $n_k > N$ wird, hat man die gesuchte Abschätzung.

(2.8.11) Beispiel: $s_n = \sum_{k=1}^n \frac{1}{k^2}$. Die Folge wächst monoton, aber ist sie auch beschränkt? Numerische Experimente mit dem Computer deuten das an, aber kann man sicher sein? Und was ist der Grenzwert? Dieser wird sich später als $\frac{\pi^2}{6}$ erweisen. Aber wie soll man das herausbekommen. Aber der Nachweis, dass eine Cauchyfolge vorliegt, ist ziemlich leicht. Die folgende Abschätzung für $n > m$ zeigt das:

$$\begin{aligned} |s_n - s_m| &= |a_{m+1} + \dots + a_n| = \frac{1}{(m+1)^2} + \dots + \frac{1}{n^2} && \text{(Explication, Betrag entfällt)} \\ &\leq \frac{1}{m(m+1)} + \frac{1}{(m+1)(m+2)} + \dots + \frac{1}{(n-1)n} && \text{(Vorsichtig vergrößert)} \\ &= \left(\frac{1}{m} - \frac{1}{m+1}\right) + \left(\frac{1}{m+1} - \frac{1}{m+2}\right) + \dots + \left(\frac{1}{n-1} - \frac{1}{n}\right) && \text{(Partialbruch!)} \\ &= \frac{1}{m} - \frac{1}{n} \leq \frac{1}{m} \leq \frac{1}{N} < \varepsilon && \text{(Problemlos erfüllbar)} \end{aligned}$$

Natürlich benötigt diese Rechnung eine Idee: Das Wegheben der Beiträge in der vorletzten Zeile. Die vorherigen Zeilen bereiten das vor. Und beachten Sie. Es ist etwas anderes, auf so eine Idee zu kommen, sie zu finden, als sie nachzuvollziehen und zu verstehen. Letzteres wird hier erwartet!

Fazit: Die Folge (mit Werten in \mathbb{Q}) ist eine Cauchyfolge und hat damit sicherlich einen Grenzwert in \mathbb{R} mit noch unbekanntem Wert.

2.2.9 Die Konstruktion der reellen Zahlen

(2.9.1) Wir haben in Kap.2.1 bereits besprochen, wie man konstruktiv von der leeren Menge zu \mathbb{N} , von dort nach \mathbb{Z} und von dort weiter zum Körper \mathbb{Q} gelangt. Aber wie kommt man von \mathbb{Q} nach \mathbb{R} ? Also von den Brüchen zur Gesamtheit aller reellen Zahlen? Es geht nicht mehr durch "Lösbarmachen aller Gleichungen eines bestimmten Typs" wie in den meisten anderen Fällen. Man benötigt eine neue Idee.

(2.9.2) Wie kann man von den rationalen Zahlen zu den reellen gelangen? Wir erwarten, dass wir jede reelle Zahl als unendlichen Dezimalbruch darstellen können. Etwa $\pi = 3.141\,592\,653\dots$ Mit dieser Darstellung ist es leicht, eine solche Zahl rational zu approximieren. D.h. eine Folge $\mathbb{N} \rightarrow \mathbb{Q}$ bilden, deren

Grenzwert die gegebene Zahl ist. Im Beispiel: $\pi_1 = 3.1$, $\pi_2 = 3.14$, $\pi_3 = 3.141$, usw. Offensichtlich gilt $\lim_n \pi_n = \pi$. Denn es ist $|\pi_n - \pi| < 10^{-n}$. Und eine Cauchyfolge ist das auch. Jedes π_n ist natürlich eine rationale Zahl. Beachten Sie, dass wir einen Rollenkonflikt haben: Es liegt nahe, π einerseits als Bezeichnung für die Folge $n \mapsto \pi_n$ einzuführen und andererseits als Bezeichnung für den Grenzwert dieser Folge. Die Dezimalbruchdarstellung ist dann auch gleich der Folge der zugehörigen Approximationen.

□ Schätzen Sie $|\pi_n - \pi_m|$ ab.

(2.9.3) Jetzt stellen wir uns auf den Standpunkt, dass wir \mathbb{R} noch nicht zur Verfügung hätten. Mit \mathbb{Q} können wir die Cauchyfolgen in \mathbb{Q} bilden, die aber teilweise keinen Grenzwert (in \mathbb{Q}) haben.

Die Idee ist dann, die Folge selbst als neues Objekt, eben eine reelle Zahl anzusehen. Denn all deren Eigenschaften kann man ja mit mehr oder weniger Mühe aus der Folge herausholen. Aber natürlich gibt es nicht nur die eine von uns gewählte Folge, sondern es gibt viele gleichwertige Folgen. Und wir können innerhalb von \mathbb{Q} entscheiden, ob zwei Folgen dieselbe reelle Zahl erfassen oder nicht. **Tun sie es, ist die Differenz eine Nullfolge** und das ist eine in \mathbb{Q} entscheidbare Aussage. D.h. letztlich, wir müssen zu einer Äquivalenzrelation für die Cauchyfolgen) samt zugehöriger Partition übergehen.

(2.9.4) Der Einstieg in das recht umfangreiche **Programm zur Konstruktion der reellen Zahlen** ist damit klar:

- Sei $C = C(\mathbb{N}, \mathbb{Q})$ die Menge aller Cauchyfolgen $\mathbb{N} \rightarrow \mathbb{Q}$.

– Darin definieren wir die Relation für $a, b \in C$

$$a \sim b \iff (a - b \text{ ist Nullfolge})$$

- Das **ist** eine Äquivalenzrelation. Die Klassenmenge bezeichnen wir mit R . Die von der Folge $a \in C$ erzeugte Klasse bezeichnen wir mit $[a]$. Diese Klassen $[a]$ von R sind die Kandidaten für die reellen Zahlen. Oder "Eine reelle Zahl ist eine Äquivalenzklasse von Cauchyfolgen $\mathbb{N} \rightarrow \mathbb{Q}$ ". $[a] = [b]$ bedeutet: $a - b$ ist Nullfolge!
- Für jede rationale Zahl $r \in \mathbb{Q}$ ist die konstante Folge $\underline{r} = (\mathbb{N}, n \mapsto r, \mathbb{Q})$ Element aus C und Vertreter einer Klasse $[\underline{r}]$. Das gibt uns eine injektive Abbildung $(\mathbb{N}, r \mapsto \underline{r}, R)$. Oder auch: Die rationalen Zahlen sind in R eingebettet.

□ Weisen Sie nach, dass eine Äquivalenzrelation vorliegt.

□ Zeigen Sie, dass jede Klasse $[a]$ abzählbar unendlich viele Elemente enthält.

(2.9.5) Jetzt muss man sukzessive die Gültigkeit aller in Kap. 2.1.2 angegebenen Axiome für die reellen Zahlen über geeignete Konstruktionen nachweisen. Dies Programm führen wir hier nur teilweise durch. Beginnen wir mit der Addition:

Seien $a, b \in C$ und $[a], [b]$ die davon erzeugten Klassen. Wir definieren $[a] + [b] = [a + b]$, wobei $a + b = (\mathbb{N}, n \mapsto (a + b)_n = a_n + b_n, \mathbb{Q})$ sein soll. Das liefert das übliche Wohldefiniertheitsproblem: Angenommen a' und b' sind zwei weitere Cauchyfolgen mit $[a] = [a']$ und $[b] = [b']$. Gilt dann $[a + b] = [a' + b']$? Tatsächlich folgt das leicht mit Hilfe der Grenzwertsätze für die Nullfolgen! Im Sinne der Abbildungen gilt ja $(a + b) - (a' + b') = (a - a') + (b - b')$. Und nach den Grenzwertsätzen ist das erneut eine Nullfolge. Also sind $a + b$ und $a' + b'$ äquivalent und erzeugen die gleiche Klasse.

Als nächstes ist zu zeigen, dass $(R, +)$ die Gruppeneigenschaften erfüllt. Etwa $[a] + [b] = [b] + [a]$. Das ist weitgehend trivial. Und für rationale Zahlen ergibt sich gerade die alte Addition.

Entsprechend ist die Multiplikation einzuführen und der gesamte zugehörige Satz der Körperaxiome zu überprüfen.

(2.9.6) Neben den algebraischen Rechenregeln hatten wir noch die Ordnungsrelation und die Vollständigkeit.

Im ersteren Fall müssen wir die Menge $P \subset R$ der später positiven reellen Zahlen einführen und die zugehörigen Axiome realisieren. Wir definieren:

$$P = \{a \mid a \in C, \text{ es gibt } N \in \mathbb{N}, \text{ so dass } a_n > 0 \text{ für alle } n \geq N\} \subset C.$$

Damit folgen (R.11-13) ohne größere Probleme.

(2.9.7) Es bleibt das Vollständigkeitsaxiom (R.14):

Auch dessen Gültigkeit ist beweisbar (was wir hier nicht tun) und damit haben wir die reellen Zahlen konstruiert. In jedem Fall ist noch zusätzlich zu prüfen, ob die Konstruktion in \mathbb{R} für Elemente aus \mathbb{Q} das alte Resultat liefert

(2.9.8) Jetzt kann man \mathbb{R} mit R identifizieren: Also die approximierende Cauchyfolge mit ihrem Grenzwert. Alle Folgerungen der Axiome gelten auch für das so konstruierte \mathbb{R} , insbesondere der erste Teil von (2.8.9): **Jede Cauchyfolge in R ist konvergent.**

2.3 Unendliche Reihen

2.3.0 Vorbemerkung

(3.0.1) Die große Bedeutung der unendlichen Reihen beruht darauf, dass es mit ihrer Hilfe möglich ist, viele wichtige und benötigte irrationale Zahlen als Ergebnis einer unendlichen Reihe darzustellen bzw. sie durch rationale Zahlen zu approximieren. Sagen wir π oder $\sin(1)$ oder auch nur $\sqrt{3}$.

Nehmen wir $\sin(1)$ mit im Bogenmaß gemessenem Winkel. Ein Computeralgebrasystem liefert nach Fixierung einer bestimmten Stellenzahl - hier 45 :

$$\sin 1 = 0.841\ 470\ 984\ 807\ 896\ 506\ 652\ 502\ 321\ 630\ 298\ 999\ 622\ 563\ 061\dots$$

Wie wird so etwas erhalten? Sicher nicht durch geometrisches Vermessen eines Dreiecks. Auch nicht über eine einfache rekursive Formel für die einzelnen Stellen. Man benötigt einen rationalen Rechenausdruck, der nur die elementaren Rechenoperationen enthält. Aber den gibt es nicht. Erst wenn man die Zahl der Summanden nach unendlich gehen läßt, erhält man eine immer bessere Approximation, der Unterschied zwischen wahrem und approximativem Wert wird zu einer Nullfolge. Man muss eine allgemeine Theorie entwickeln, die es erlaubt, solche Approximationen beliebiger Genauigkeit zu finden. Die Theorie der Reihen erweist sich hierfür als wichtiges Werkzeug.

(3.0.2) Noch eine kleine Überlegung am Rande zu den Ziffern von $\sin(1)$. Es liegt nahe, einmal auszuzählen, wie oft die einzelnen Ziffern auftreten, die 0, die 1 usw. Wir nehmen nur die oben gegebenen Stellen und finden (oben die Ziffern, in der Mitte die absoluten Häufigkeiten - etwa 7 mal die Ziffer 6 und unten die relativen Häufigkeiten $7/45 = 0.15$ für die 6).

0	1	2	3	4	5	6	7	8	9
6	3	6	3	3	4	7	2	5	6
0.13	0.07	0.13	0.07	0.07	0.08	0.15	0.04	0.11	0.13

- Wie werden sich die Einträge in diese Tabelle ändern, wenn man immer mehr Stellen mitnimmt? Vermuten Sie einen Grenzwert?

2.3.1 Reihen sind spezielle Folgen

(3.1.1) Reihen sind ein spezieller Typ von Folgen. **D.h. insbesondere, dass alle Resultate zu den Folgen auf die Reihen anwendbar sind.** Wie dieser Sachverhalt entsteht, haben wir bereits mehrfach gesehen. Man hat eine erste Folge $a = (\mathbb{N}, n \mapsto a_n, \mathbb{C})$ und bildet damit eine neue Folge $n \mapsto s_n$, deren Glieder der Rekursion $s_n = s_{n-1} + a_n$ und $s_0 = a_0$ genügen sollen. Das läßt sich sofort in eine explizite Formel umwandeln, nämlich $s_n = a_0 + a_1 + \dots + a_n = \sum_{k=0}^n a_k$. Diese Folge der s_n wird als "unendliche Reihe $\sum a_k$ " bezeichnet, ihre Glieder s_n werden "Partialsummen" genannt.

Nochmals: **Eine unendliche Reihe ist die Folge ihrer Partialsummen und alle einschlägigen Resultate zu den Folgen sind anwendbar.**

(3.1.2) Ist die Folge der Partialsummen konvergent, dann pflegt man auch den Grenzwert mit dem Summensymbol zu bezeichnen. Also $\lim_{n \rightarrow \infty} s_n = \sum_{k=0}^{\infty} a_k$. In diesem Fall kann das Symbol $\sum_{k=0}^{\infty} a_k$ zwei Bedeutungen haben: "(Konvergente) Folge der Partialsummen" und "Grenzwert dieser Folge". Ist die Folge endlich, dann bricht auch die Reihe ab, hat einen letzten Term den man mit der endlichen Summe identifiziert. $s_n = \sum_{k=0}^n a_k$

(3.1.3) In vielen Fällen enthalten die Summanden a_k noch einen äußeren Parameter, den wir meist mit x bezeichnen. Also $n \mapsto a_n(x)$ und $s_n = \sum_{k=0}^n a_k(x)$. D.h.: Für jede zulässige Wahl von x gibt das eine zugehörige eigene Reihe. Das Symbol erfasst eine ganze Schar von Reihen.

(3.1.4) Wir haben bereits einige Reihen behandelt, die wir nun alle in der Summenschreibweise wiedergeben. Zwei davon mit äußerem Parameter.

$\sum \frac{1}{k}$	harmonische Reihe	divergent
$\sum x^n$	geometrische Reihe	konvergent für $ x < 1$
$\sum \frac{x^k}{k!}$	Bisher ?	Bisher ?
$\sum \frac{1}{n^2}$	konvergent!	Bisher ?

(3.1.5) Erneut können wir **die Mengen aller unendlichen Reihen** bilden, darin die Teilmenge der konvergenten Reihen usw. Parametrisiert werden die unendlichen Reihen durch die zugehörige Folge a . Und natürlich tragen diese Mengen jeweils gewisse algebraische Strukturen. Zunächst einmal per Wertemengenübertragung eine Vektorraumstruktur.

Typischerweise sieht das dann so aus. Seien a, b zwei Folgen $\mathbb{N} \rightarrow \mathbb{C}$ mit den zugeordneten Reihen Σa_k und Σb_k . Weiter sei $\alpha, \beta \in \mathbb{C}$. Dann haben wir

$$\Sigma_{k=0}^n (\alpha a_k + \beta b_k) = \alpha \Sigma_{k=0}^n a_k + \beta \Sigma_{k=0}^n b_k.$$

Und das definiert die Partialsummen der neuen Reihe $\alpha a + \beta b$.

(3.1.6) Sei $I = \{A, A+1, \dots, E\}$ ein Intervall aus \mathbb{N} . I darf auch nach oben unbeschränkt sein. Also $I = \{A, A+1, A+2, \dots\}$. In diesem Fall setzen wir symbolisch $E = \infty$. Weiter sei $\pi : I \rightarrow J$ eine bijektive Parametrisierung einer zweiten Indexmenge J . J sei Indexmenge einer Familie komplexer Zahlen. Also $(a_j)_{j \in J}$. Dann definieren wir rekursiv $\Sigma_{i=A}^E a_{\pi(i)}$ bzw. $\Sigma_{i=A}^{\infty} a_{\pi(i)}$ über $s_A = a_{\pi(A)}$ und $s_k = s_{k-1} + a_{\pi(k)}$. Ist I endlich, bricht diese Folge ab und wir haben $s_E = \Sigma_{i=A}^E a_{\pi(i)}$. Ist $I = \mathbb{N}$, dann bezeichnet $\Sigma_{i=A}^{\infty} a_{\pi(i)}$ stets die Folge der Partialsummen und im konvergenten Fall bei Bedarf auch den Grenzwert. Sollten Zweifel über die benutzte Parametrisierung π bestehen, schreiben wir auch $\Sigma_{\pi} a$ oder $\Sigma_{\pi, i=A}^E a_{\pi(i)}$.

(3.1.7) Einige Beispiele:

- $\pi : k \mapsto 2k$. Dann ist $\Sigma_{k=0}^N a_{\pi(k)} = \Sigma_{k=0}^N a_{2k}$.
- Es sei $\pi : \mathbb{N} \rightarrow \mathbb{N}$ eine Permutation von \mathbb{N} . Dann liefert $\Sigma_{i=0}^{\infty} a_{\pi(i)}$ eine Umordnung der Reihenfolge der Summanden.
- Sei a eine Folge und Σa die zugehörige Reihe. Weiter sei $J \subset \mathbb{N}$ eine (endliche oder unendliche) nicht leere Teilmenge von \mathbb{N} . Dann können wir die Elemente von J der Größe nach anordnen: $J = \{j_0, j_1, j_2, \dots, j_E\}$ mit $j_i < j_{i+1}$. Das definiert eine bijektive Parametrisierung $i \mapsto j_i$ die wir mit ω bezeichnen wollen. also $\omega(i) = j_i$. Dafür können wir $\Sigma_{i=0}^E a_{j_i}$ definieren. Ist z.B. $J = \{1, 7, 3, 4\}$, dann ist $\Sigma_{i=0}^4 a_{j_i} = ((a_1 + a_3) + a_4) + a_7$. Die so definierte (endliche oder unendliche) Folge der Partialsummen bezeichnen wir meist mit $\Sigma_{i \in J}^{\omega} a_i$. Beachten Sie: Das Symbol ω deutet an, dass die Elemente aus J in der natürlichen Reihenfolge zu nehmen sind, nicht in einer beliebigen!

□ Sei $I = \{(1,1), (1,2), (2,1), (2,2)\}$. Wir setzen $\pi(0) = (1,1)$, $\pi(1) = (1,2)$, $\pi(2) = (2,1)$ und $\pi(3) = (2,2)$. Dann ist $\Sigma_{\pi} a = ((a_{11} + a_{12}) + a_{21}) + a_{22}$. Was ist $\Sigma_{\psi} a$ für $\psi(0) = (1,1)$, $\psi(1) = (2,1)$, $\psi(2) = (2,2)$ und $\psi(3) = (1,2)$?

(3.1.7) **Ergebnis:** Sei $(a_j)_{j \in J}$ eine Familie komplexer Zahlen mit irgendeiner Indexmenge J . Und $\pi : I \rightarrow J$ sei eine bijektive Parametrisierung durch ein Intervall $I = [A, E]$ von \mathbb{N} . Dann können wir den Term $\Sigma_{\pi, i=A}^E a_{\pi(i)}$ in der beschriebenen Weise definieren. Und ist $\mathbb{N} \subset \mathbb{N}$ irgendeine nicht leere Teilmenge der natürlichen Zahlen, die zugleich als Indexmenge $(a_k)_{k \in \mathbb{N}}$ einer Familie komplexer Zahlen auftritt, dann können wir diese Zahlen ihrer Größe nach ordnen und die Reihe $\Sigma_{n \in \mathbb{N}}^{\omega} a_n$ bilden.

2.3.2 Absolute Konvergenz

(3.2.1) **Gibt es überhaupt etwas, was bei Reihen anders ist als bei Folgen?** Ja - und zwar durchaus Wichtiges! Die Summen erlauben mehr umformende Rechenoperationen als die Folgen und die resultierenden Beziehungen lassen sich keineswegs problemlos ins Unendliche übertragen. Insbesondere gehören hierzu gewisse Konsequenzen des Assoziativgesetzes, des Kommutativgesetzes und der Distributivgesetze. Betrachten wir als Beispiel einmal folgende Termumformung, die für endlich viele Summanden infolge von Assoziativ- und Kommutativgesetz zulässig ist: (Fehlende Klammern sind immer im Sinne der Reihenrekursion zu interpretieren. D.h. $a_2 + a_4 + a_6$ steht für $(a_2 + a_4) + a_6$.)

$$a_1 + a_2 + a_3 + a_4 + a_5 + a_6 = (a_2 + a_4 + a_6) + (a_3 + a_5) + a_1$$

Darf man derartiges auf eine unendliche Summe ausdehnen? Gilt also auch dort notwendig beispielsweise $\Sigma a_k = (\Sigma a_{2k}) + (\Sigma a_{2k+1})$?

Wir illustrieren das Problem an einem einfachen Beispiel, das zwar wenig relevant ist, aber die auftretenden Schwierigkeiten aufzeigt. Sei $a_n = (-1)^n$. Also $a = (1, -1, 1, -1, 1, \dots)$. Dann folgt $s_{2n} = 1$ und $s_{2n+1} = 0$. Die zugehörige Reihe konvergiert keineswegs. Ausgeschrieben: $\sum a_k = 1 - 1 + 1 - 1 + 1 - \dots$. Nimmt man jetzt mit dem Assoziativgesetz eine Beklammerung vor, so erhält man eine neue Reihe nämlich

$$(1 - 1) + (1 - 1) + (1 - 1) + \dots = 0 + 0 + 0 + \dots = 0$$

D.h. die durch Anwendung des Assoziativgesetzes auf die Summe entstehende neue Folge $b_n = a_{2n} + a_{2n+1}$ ist konvergent mit $\sum b_k = 0$.

□ Jetzt sei $c_0 = a_0 + a_2$ und $c_1 = a_4 - a_1$ und allgemein $c_k = c_{k+3} - c_k$. Welche der üblichen Rechenregeln für Summen werden hier für die unendliche Summe benutzt? Was ergibt sich für $\sum c_k$?

(3.2.2) Wir sehen: Die üblichen Rechenregeln für Summen, gegeben durch Kommutativgesetz, Assoziativgesetz und Distributivgesetz darf man keineswegs ungeprüft und ungerechtfertigt auf unendliche Summen ausdehnen. Und das heißt: Beim Übergang nach unendlich darf man bei Summen reeller Zahlen keineswegs eventuelle Klammern fortlassen und die Reihenfolge der Summanden beliebig vertauschen. Bei endlichen Zahlsummen hängt das Ergebnis von solchen Umformungen nicht ab. Bei unendlichen kann das aber der Fall sein. **Daher haben wir über die Definitionen in (3.1.5) einen ganz bestimmten Rechenweg festgelegt, der sich für den Übergang nach unendlich als besonders vorteilhaft erweist.**

(3.2.3) Allerdings zeigt sich, dass Ausdehnung zulässig ist, falls die beteiligten Reihen eine zusätzliche Bedingung erfüllen, die als "absolute Konvergenz" bezeichnet wird.

Definition: Sei $(\mathbb{N}, n \mapsto a_n, \mathbb{C})$ eine Folge und $\sum a_k$ die zugehörige Reihe.
Dann heißt die Reihe *absolut konvergent*, wenn die Reihe $\sum |a_k|$ der Absolutbeträge konvergent ist. D.h. die Folge der Partialsummen $n \mapsto \sum_{k=0}^n |a_k|$ der Beträge.

Insbesondere sind alle konvergenten Reihen, bei denen fast alle a_n dasselbe Vorzeichen haben, auch absolut konvergent. Unsere Behauptung zur absoluten Konvergenz ist natürlich zu beweisen, was wir weiter unten großenteils tun werden. Dieser Beweis benutzt den in (3.1.6-7) eingeführten Formalismus. Der Leser sollte sich diese überaus wichtigen Resultate von "absolut konvergent" einprägen und immer wieder zur Orientierung heranziehen.

Will man mit unendlichen Summen ebenso rechnen wie mit endlichen, dann muss man absolute Konvergenz verlangen!

(3.2.4) Beachten Sie: Da Beträge nicht negativ sind, ist $n \mapsto \sum_{k=0}^n |a_k|$ eine positive monoton wachsende Folge. (Bei jeder Vergrößerung von n wird etwas nicht Negatives zur Partialsumme hinzugefügt.) Ist die Summe beschränkt, dann ist sie konvergent (Monotoniekriterium). Ist sie unbeschränkt, dann divergiert die Folge der Partialsummen unbestimmt.

(3.2.5) Reihen sind Folgen: Schauen wir uns zunächst an, was das Cauchy Kriterium, das ja generell über die Konvergenz entscheidet, für Reihen besagt. Sei a die Folge, s die Folge der Partialsummen und $\sigma = (\mathbb{N}, n \mapsto \sum_{k=0}^n |a_k|, \mathbb{R})$ die Folge der Partialsummen der Beträge. Dann hat man für das Cauchy Kriterium folgende Explikation (mit $n < m$):

$$|s_n - s_m| = \left| \sum_{k=0}^n a_k - \sum_{k=0}^m a_k \right| = \left| \sum_{k=m+1}^n a_k \right| = |a_{m+1} + \dots + a_n|.$$

Den mittleren Term lassen wir fast immer fort. Für absolute Konvergenz lautet dagegen die Forderung

$$|\sigma_n - \sigma_m| = \left| \sum_{k=0}^n |a_k| - \sum_{k=0}^m |a_k| \right| = \left| \sum_{k=m+1}^n |a_k| \right| = |a_{m+1}| + \dots + |a_n|$$

Mit der Dreiecksungleichung folgt andererseits

$$|s_n - s_m| = |a_{m+1} + \dots + a_n| \leq |a_{m+1}| + \dots + |a_n| = |\sigma_n - \sigma_m|.$$

Und natürlich ist denkbar, dass die linke Seite $< \varepsilon$ ist, ohne dass die rechte das auch ist. Wir erwarten daher, dass aus absoluter Konvergenz die gewöhnliche folgt, aber dass es umgekehrt konvergente Reihen gibt, die nicht absolut konvergent sind.

(3.2.6) Merke: **Die Explikation des Cauchyriteriums für eine Reihe** $\sum a_i$ ergibt ($n > m$)

Zu jedem $\varepsilon > 0$ gibt es ein N , so dass für alle $n, m \geq N(\varepsilon)$
gilt $|s_n - s_m| = |a_{n+1} + a_{n+2} + \dots + a_m| < \varepsilon$

(3.2.7) Die in (3.1.12) bewiesene Ungleichung $|s_n - s_m| \leq |\sigma_n - \sigma_m|$ zeigt:

Eine absolut konvergente Reihe ist konvergent.

(3.2.8) Ein Beispiel einer konvergenten, aber nicht absolut konvergenten Reihe gehört zur Folge $a_n = \frac{(-1)^n}{n+1}$. Die betrachtete Reihe ist daher $\sum \frac{(-1)^n}{n+1} = 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \dots$. Die zugehörige Reihe der Absolutbeträge ist $\sum \frac{1}{n+1} = 1 + \frac{1}{2} + \frac{1}{3} + \dots$. Und das ist die harmonische Reihe, von der wir wissen, dass sie bestimmt divergiert. Die Reihe selbst ist konvergent, wie man mit Hilfe des Cauchyriteriums zeigt. Dabei beschränken wir uns auf eine Abschätzung von $|s_{2m+2k} - s_{2m}|$. Die anderen Indexkombinationen lassen sich leicht ergänzen.

$$\begin{aligned} |s_{2m+2k} - s_{2m}| &= \frac{1}{2m+1} - \frac{1}{2m+2} + \frac{1}{2m+3} - \frac{1}{2m+4} + \dots + \frac{1}{2m+2k-1} - \frac{1}{2m+2k} \\ &\leq \frac{1}{2m+1} - \frac{1}{2m+2} + \frac{1}{2m+2} - \frac{1}{2m+4} + \frac{1}{2m+4} + \dots + \frac{1}{2m+2k-2} - \frac{1}{2m+2k} \\ &= \frac{1}{2m+1} - \frac{1}{2m+2k} = \frac{k}{(2m+1)(m+k)} = \frac{1}{(2m+1)(1+\frac{m}{k})} \leq \frac{1}{2m+1} \leq \frac{1}{2N} < \varepsilon. \end{aligned}$$

Hier wurde jeder 2. Summand vorsichtig durch Verkleinern des Nenners vergrößert, so dass danach dasselbe Fortheben wie in (2.8.11) erfolgt. Die letzte Ungleichung ist leicht für alle k und ε erfüllbar, sofern nur $N \leq m$ ausreichend groß ist.

□ Beweisen Sie: Eine notwendige, aber keineswegs hinreichende Bedingung für die Konvergenz einer Reihe $\sum a_k$ ist, dass $k \rightarrow a_k$ eine Nullfolge ist. (Hinweis: Cauchyriterium und Gegenbeispiel)

(3.2.9) Wir formulieren nachfolgend einen Satz, der die harmlose Gutartigkeit der absolut konvergenten Reihen belegt und einen zweiten, der die interessante Boshaftigkeit der konvergenten, aber nicht absolut konvergenten Reihen aufzeigt. Das nächste Beispiel zeigt, um welche Termumformungen es dabei geht:

(3.1.18) Wie kann man eine Umformung wie

$$a_1 + a_2 + a_3 + a_4 + a_5 + a_6 = ((a_2 + a_4 + a_6) + (a_3 + a_5)) + a_1$$

formal beschreiben? Jede Summe ohne Klammern steht im Zweifelsfall für die in (3.1.5) festgelegte Standardbeklammerung. Wir haben zweierlei: Eine **Permutation** π der Indexmenge $I = \{1, 2, \dots, 6\}$, mit deren Hilfe wir die neue Folge b mit $b_i = a_{\pi(i)}$ einführen. Im Beispiel ist $b = (a_2, a_4, a_6, a_3, a_5, a_1)$. Und π ist natürlich eine bijektive Parametrisierung der Indexmenge durch ein Intervall. Und dann eine Partition P der Indexmenge. Im Beispiel $P = \{\{1, 2, 3\}, \{3, 5\}, \{1\}\}$. Für diese brauchen wir zur Festlegung der Reihenfolge auch eine **Intervallparametrisierung** ρ (der Partition). Sagen wir $\rho(1) = \{1, 2, 3\}$, $\rho(2) = \{3, 5\}$ und $\rho(3) = \{1\}$. Dann können wir den rechts stehenden Term schreiben als $\sum_{\rho, k=1}^3 \left(\sum_{i \in p(k)}^{\omega} a_{\pi(i)} \right)$.

(3.2.10) Für endlich viele Summanden gilt als Konsequenz der Rechenregeln für Zahlen $\sum_{i=1}^6 a_i = \sum_{\rho, k=1}^3 \left(\sum_{i \in p(k)}^{\omega} a_{\pi(i)} \right)$. Aber wie steht es mit dem Übergang zu unendlich vielen Summanden? D.h. genauer, wenn für die Partition P abzählbar viele Klassen zugelassen werden und wenn die Klassen selbst abzählbar viele Elemente enthalten dürfen?

(3.2.11) **Satz:**

Es sei $a = (\mathbb{N}, n \mapsto a_n, \mathbb{C})$ eine Folge mit unendlicher Reihe $\sum a_k$. Weiter sei $\pi : \mathbb{N} \rightarrow \mathbb{N}$ eine bijektive Abbildung, also eine Permutation der Indexmenge. Und T eine unendliche Teilmenge von \mathbb{N} .

Dann ist die Reihe $\sum_{n \in T}^{\omega} a_{\pi(n)}$ auch absolut konvergent. Insbesondere ist jede Umordnung der Reihe erneut absolut konvergent.

Zur Konstruktion: Zuerst permutiert man die Elemente mit π . Aus (a_0, a_1, \dots) wird $(a_{\pi(0)}, a_{\pi(1)}, a_{\pi(2)}, \dots)$. Dann streicht man alle nicht zu T gehörigen Feldplätze. Aus dem was bleibt wird in der gegebenen Reihenfolge die neue Reihe gebildet.

(3.2.12) Beweis: Die absolute Konvergenz besagt $\sum_{i=0}^{\infty} |a_i| < \infty$. Nun ist $T \subset \mathbb{N}$. Sei $N_1 = \max_{i \leq N} (\pi(i))$. Klar ist $N \leq N_1$. Dann folgt für die endlichen Summen $\sum_{i \in T, i \leq N_1} |a_{\pi(i)}| \leq \sum_{i=0}^{N_1} |a_i| \leq \sum_{i=0}^{\infty} |a_i| < \infty$. Denn jeder Summand links tritt auch in der mittleren Summe auf. Grenzübergang liefert das Gewünschte (G9) aus (2.5.26). **Die permutierten Teilfolgen liefern daher absolut konvergente Reihen.**

(3.2.13) **Umordnungssatz für absolut konvergente Reihen**

Es sei	$a = (\mathbb{N}, n \mapsto a_n, \mathbb{C})$ eine Folge mit unendlicher Reihe $\sum a_k$. Weiter sei $\pi : \mathbb{N} \rightarrow \mathbb{N}$ eine bijektive Abbildung, also eine Permutation der Indexmenge. Schließlich sei $P = (I, i \mapsto p(i), \mathfrak{P}(\mathbb{N}))$ eine bijektive Intervallparametrisierung einer Partition der natürlichen Zahlen mit $I = [A, E] \subset \mathbb{N}$. Die Reihe $\sum a_k$ sei absolut konvergent .
Dann gilt	$\sum a_k = \sum_{i=A}^E \left(\sum_{j \in p(i)}^{\omega} a_{\pi(j)} \right)$

Kurz: Bei absoluter Konvergenz sind Umformungen der in (3.1.18) beschriebenen Art auch für abzählbar unendlich viele Summanden zulässig.

(3.2.14) Illustrationsbeispiele für die Anwendung des Satzes:

- Wählt man P einelementig, also $P = \{\mathbb{N}\}$, dann folgt wie in (3.1.20), da π beliebig ist:

Bei einer absolut konvergenten Reihe darf man die Reihenfolge der Summanden beliebig verändern.

- Vielfach Wählt man P zweielementig: $P = \{G, U\}$ mit $G = \{0, 2, 4, \dots\}$ und $U = \{1, 3, 5, \dots\}$. Ist dann $\sum a_k$ absolut konvergent, hat man:

$$\sum_i a_i = \sum_i a_{2i} + \sum_i a_{2i+1}$$

□ Zerlegen Sie nach dieser Methode $\sum_{n=0}^{\infty} \frac{(ix)^n}{n!}$ in Real- und Imaginärteil. Sie dürfen absolute Konvergenz voraussetzen.

(3.2.15) **Beweis** (von (3.2.13)):

$\sum a_n$ ist absolut konvergent, also auch konvergent. Sei $\sum_{n=1}^{\infty} a_n = S$. Gebe $\varepsilon > 0$ vor. Zu zeigen: Man kann erreichen, dass $|\sum_{i=A}^{*E} \left(\sum_{k \in p(i)}^{\omega*} a_{\pi(k)} \right) - S| < \varepsilon$ wird. Dabei soll Σ^* bedeuten, dass von der Summe jeweils nur eine **endliche Partialsumme** genommen werden soll. Gefordert wird, dass S durch eine endliche Teilsommen approximiert werden kann. Allerdings sollen alle eventuellen Vorgängersummanden auch mitgenommen werden.

Sei jetzt N so groß gewählt, dass $\sum_{k=N}^{\infty} |a_k| < \varepsilon$ ist. D.h. aber: $|a_1 + \dots + a_N - S| < \varepsilon$. Oder auch $|\boxed{a_1 + \dots + a_N + \sum_{k=N}^{**} a_k - S}| < \varepsilon$. (*) Dabei sei $\sum_{k=N}^{**} a_k$ eine endliche Summe weiterer a_k mit $k > N$. Denn das Herausziehen weiterer Summanden aus der Reihe verkleinert den Restbetrag (Dreiecksungleichung). Jetzt restringieren wir die Partition P auf ein endliches Anfangsintervall I_0 von I, derart, dass alle $a_{\pi(i)}$ mit $\pi(i) \leq N$ in einem $p(i)$ mit $i \in I_0$ vorkommen. Da N endlich ist, ist das möglich. Und in jedem $p(i)$ mit $i \in I_0$ nehmen wir in der Partialsumme nur endlich viele Glieder, mindestens so viele, dass alle $a_{\pi(k)}$ mit $\pi(k) \in p(i)$ und $\pi(k) < N$ enthalten sind. Das gibt wieder endlich viele Terme und definiert die Summen $\sum_{i \in p}^{\omega*}$. In diesen Summen tauchen eventuell weitere $a_{\pi(i)}$ mit $\pi(i) > N$ auf, aber das sind erneut höchstens endlich viele. In der insgesamt so konstruierten endlichen Summe sind a_1, \dots, a_N sicher enthalten, so dass nach (*) wie gefordert $|\sum_{p \in P}^* \left(\sum_{i \in p}^{\omega*} a_{\pi(i)} \right) - S| < \varepsilon$ gilt! Und nach Hinzunahme weiterer a_i zur Summe bleibt das so.

Geht man, wie von der Formel verlangt, in den inneren Summen zur Grenze $s_i = \sum_{j \in p(i)}^{\omega} a_{\pi(j)}$ über (Mit Grenzübergang, wenn $p(i)$ unendlich ist, sonst nur alle $j \in p(i)$; diese Grenzwerte existieren nach (3.1.20)), dann bleibt die Ungleichung erhalten und man hat $|\sum_{p \in P}^* s_p - S| \leq \varepsilon$. Ist die Indexmenge I der Partition endlich, genügt das. Andernfalls muss man auch hier in der durch P vorgegebenen Reihenfolge zu Grenze übergehen, was keine Verletzung der Ungleichung bewirken

kann, da ja immer nur weitere Terme a_i mit $i > N$ hinzugenommen werden. Also bleibt auch jetzt die Ungleichung erhalten.

Insgesamt ergibt sich, da $\varepsilon > 0$ beliebig war das gewünschte Resultat

$$\lim_i \sum_{i \in I}^\omega s_i = S = \sum_n a_n.$$

(3.2.16) **Fazit:** Für endliche Indexmengen I und zugehörige Familien $(a_i)_{i \in I}$ komplexer Zahlen haben wir in Kap (1.2) das Symbol $\boxed{\sum_{i \in I} a_i}$ eingeführt. Es besagt, dass über alle Elemente von I zu summieren war, wobei es nicht auf die Reihenfolge und die Art der Beklammerung ankam. Denn jede zulässige Interpretation der Summe ergab dasselbe Resultat. Dieses Konzept erwies sich für das Rechnen als ausgesprochen nützlich. Jetzt sind wir in der Lage, das Symbol $\sum_{i \in I} a_i$ von endlichen auf abzählbare Mengen zu verallgemeinern.

Sei also I eine höchstens abzählbare Indexmenge und $(a_i)_{i \in I}$ eine zugehörige Familie komplexer Zahlen. Wir nennen diese Familie *absolut summabel*, wenn es eine bijektive Parametrisierung $\pi : \mathbb{N} \rightarrow I$ gibt, derart dass $\sum_{i=0}^\infty a_{\pi(i)}$ absolut konvergent ist. Dann ist jede Umordnung und auch jede Zusammenfassung konvergent und wir bezeichnen das Ergebnis mit $\boxed{\sum_{i \in I} a_i}$. Hierin geht also wieder nur die Menge I ein, nicht aber die Art der Abzählung ihrer Elemente wie im Falle von $\sum_{\pi, i \in I}$.

Der Hauptnutzen dieses Symbols besteht darin, dass man damit ausgesprochen effektiv rechnen kann. Die jetzt folgende Behandlung der Doppelsummen liefert hierfür ein gutes Beispiel.

(3.2.17) **Doppelsummen** treten als Folge der Distributivgesetze auf. Typischerweise gilt für endlich viele Summanden unter Verwendung des Summenzeichens

$$(\sum_{j \in J} a_j)(\sum_{k \in K} b_k) = \sum_{(j,k) \in J \times K} a_j b_k.$$

Dabei darf man die Summen wegen Assoziativ- und Kommutativgesetz beliebig interpretieren. In den links stehenden Summen können wir etwa über (3.1.6) zu unendlichen Reihen übergehen. Die Doppelsumme $\sum_{(j,k) \in J \times K} a_j b_k$ der rechten Seite dagegen macht Probleme, da $J \times K$ keine Teilmenge von \mathbb{N} ist. Es gibt viele Wege, daraus Folgen oder Konstruktionen nach (3.2.18) zu machen. Drei davon sind besonders wichtig. Stellt man $\mathbb{N} \times \mathbb{N}$ als (nach rechts unten offene) Matrix dar, dann lassen sich diese drei Reihungen wie folgt beschreiben. Es liegt natürlich jeweils eine der im Hauptsatz beschriebenen Strukturen vor.

- Zerlege die Matrix in ihre (unendlichen) Zeilen. Das gibt die Partition. Dann zuerst die Zeilen aufsummieren und anschließend über die Zeilen summieren. Als Formel:

$$\sum_{i=0}^\infty (\sum_{k=0}^\infty a_{ik})$$

- Analog kann man in die Spalten zerlegen:

$$\sum_{k=0}^\infty (\sum_{j=0}^\infty a_{jk})$$

- Und man kann nach Diagonalen ($i+k=d$) zerlegen, diese aufsummieren und dann über die Diagonalsummen summieren

$$\sum_{d=0}^\infty (\sum_{i=0}^d a_{i,d-i})$$

In alle drei Fällen gibt es für die Partition bereits eine natürliche vorgegebene Numerierung, also Intervallparametrisierung. Die dritte Summation erweist sich als besonders wichtig. Insbesondere sind alle inneren Summen endlich. Es bleibt nur die eine unendliche Summe.

Als schematische Matrix sieht das so aus. Gleiche Symbol wie \blacktriangle deuten eine Klasse der dritten Zerlegung an.

Doppelsumme $\sum_{ik} S_{ik}$	$\sigma_{.1} =$ $\sum_k S_{k1}$	$\sigma_{.2} =$ $\sum_k S_{k2}$	$\sigma_{.3} =$ $\sum_k S_{k3}$	$\sigma_{.4} =$ $\sum_k S_{k4}$	$\sigma_{.5} =$ $\sum_k S_{k5}$	$\sigma_{.6} \dots$
$\sigma_{1.} = \sum_k S_{1k}$	S ₁₁ ♠	S ₁₂ ∇	S ₁₃ ▲	S ₁₄ #	S ₁₅ ♠	S ₁₆
$\sigma_{2.} = \sum_k S_{2k}$	S ₂₁ ∇	S ₂₂ ▲	S ₂₃ #	S ₂₄ ♠	S ₂₅	
$\sigma_{3.} = \sum_k S_{3k}$	S ₃₁ ▲	S ₃₂ #	S ₃₃ ♠	S ₃₄	S ₃₅	
$\sigma_{4.} = \sum_k S_{4k}$	S ₄₁ #	S ₄₂ ♠	S ₄₃	S ₄₄	S ₄₅	
$\sigma_{5.} \dots \dots$	S ₅₁ ♠					

(3.2.18) Der Hauptsatz zeigt: Ist $(a_{ij})_{i,j \in \mathbb{N}}$ summabel, d.h. gibt es irgendeine Abzählung, für die $\sum_{i=0}^{\infty} |a_{\pi(i)}|$ konvergiert, dann liefern alle Abzählungen dasselbe Resultat $\sum_{(i,j) \in \mathbb{N} \times \mathbb{N}} a_{ij}$. Das gilt insbesondere für die drei beschriebenen Berechnungsweisen.

(3.2.19) **Satz:**

Ist $\sum_{i=1}^{\infty} (\sum_{k=0}^{\infty} |a_{ik}|) < \infty$ oder $\sum_{k=1}^{\infty} (\sum_{i=0}^{\infty} |a_{ik}|) < \infty$,
dann ist die Familie $(a_{ik})_{i,k \in \mathbb{N}}$ absolut summabel.

Beweis: Es genügt, absolute Konvergenz für **eine** Abzählung nachzuweisen. Wir nehmen die Diagonalzählung. Dann folgt die Behauptung etwa für die Spaltenzerlegung über die folgende Ungleichungskette, wenn man wie üblich fadenartig nacheinander die einzelnen Diagonalen durchgeht:

$$\sum_{d=0}^N (\sum_{k=0}^d |a_{k,d-k}|) \leq \sum_{i,k=0}^N |a_{ik}| = \sum_{k=1}^N (\sum_{i=k}^N |a_{ik}|) \leq \sum_{k=1}^{\infty} (\sum_{i=k}^{\infty} |a_{ik}|)$$

□ Begründen Sie die einzelnen Schritte genauer!

(3.2.20) Im Fall des Produktes zweier unendlicher Reihen $\sum a_j$ und $\sum b_k$ nennt man die dritte Konstruktion *das Cauchyprodukt der beiden Reihen*. Das ist gegeben durch die Reihe $\sum_{d=0}^{\infty} (\sum_{j=0}^d a_j b_{d-j})$.

(3.2.21) **Satz** (über das Cauchyprodukt absolut konvergenter Reihen)

Sind $\sum a$ und $\sum b$ absolut konvergent,
dann ist das Cauchyprodukt der beiden Reihen absolut konvergent,
so dass $\sum_{i,j} a_i b_j$ in beliebiger Reihenfolge (als Reihe) berechnet werden kann.
Es gilt $(\sum a_i) (\sum b_k) = \sum a_i b_k = \sum_{d=0}^{\infty} (\sum_{j=0}^d a_j b_{d-j})$.

□ Das Resultat ist sogar rechnerisch noch "schöner" als im endlichen Fall. Wieso?

Beweis: Mit Satz (3.1.26). Sei $\sum |a_i| = A$ und $\sum |b_j| = B$. Dann gilt

$$\sum_i |a_i| \left(\sum_j |b_j| \right) = \sum_i |a_i| B = \left(\sum_i |a_i| \right) B = AB < \infty.$$

(3.2.22) Beispiel zum Umgang mit dem Summenzeichen: Wir wissen, dass für $|x| < 1$ gilt $\frac{1}{1-x} = \sum x^i$. Die Reihe ist absolut konvergent. Wir haben $\left(\frac{1}{1-x}\right)^2 = (\sum x^i) (\sum x^k)$. Wir dürfen die rechte Seite nach dem Cauchyprodukt ausrechnen und finden

$$(\sum x^i) (\sum x^k) = \sum_{d=0}^{\infty} (\sum_{j=0}^d x^j x^{d-j}) = \sum_{d=0}^{\infty} (d+1) x^d$$

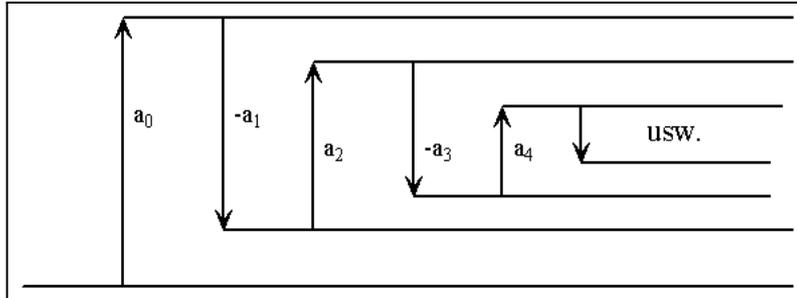
Vor der Behandlung der konvergenten aber nicht absolut konvergenten Reihen ist es sinnvoll, ein Konvergenzkriterien einzuschalten. *Konvergenzkriterien* sind Resultate, die es einem erlauben mit Hilfe geeigneter Eigenschaften der Folge a auf die Konvergenz der Reihe $\sum a$ zu schließen. Ein erstes grundlegendes Beispiel war das Cauchyriterium.

2.3.3 Konvergenzkriterien(I)

(3.3.1) Das folgende Kriterium liefert Konvergenz, macht aber keine Aussage darüber, ob absolute Konvergenz vorliegt. Tatsächlich erhält man mit seiner Hilfe leicht Beispiele von konvergenten, aber nicht absolut konvergenten Reihen.

Das **Leibnizkriterium** Sei a_n reell mit $a_n \geq 0$ monotone Nullfolge
Dann ist $\sum (-1)^n a_n$ konvergent mit einem Grenzwert $S = \theta a_0$, wobei $0 \leq \theta \leq 1$ gilt.

Die Figur illustriert das zugehörige Verhalten der Summen, die immer weiter eingeschachtelt werden. ($a_0 - a_1 + a_2 - a_3 \dots$ und $a_2 < a_1 < a_0$):



(3.3.2) Beweis: Wir betrachten die beiden Teilfolgen $k \mapsto s_{2k}$ und $k \mapsto s_{2k+1}$. Man hat:

$$\begin{aligned} s_{2k+1} &= (a_0 - a_1) + (a_2 - a_3) + \dots + (a_{2k} - a_{2k+1}) \\ s_{2k} &= a_0 - (a_1 - a_2) - \dots - (a_{2k-1} - a_{2k}) \end{aligned}$$

Die erste Folge ist monoton wachsend, da jeder Summand $(a_{2k} - a_{2k+1})$ positiv ist. Und entsprechend ist die zweite monoton fallend. Da weiter $a_0 \geq s_{2k} \geq s_{2k+1} \geq 0$ gilt, sind beide Folgen beschränkt. (Vgl. die Figur.) Nach dem Monotoniekriterium (2.6.7) sind beide Teilfolgen konvergent. Die Grenzwerte bezeichnen wir mit S_+ und S_- . Nach dem Grenzwertsatz (G1) aus () folgt $\lim_{k \rightarrow \infty} (s_{2k+1} - s_{2k}) = S_+ - S_-$. Andererseits hat man

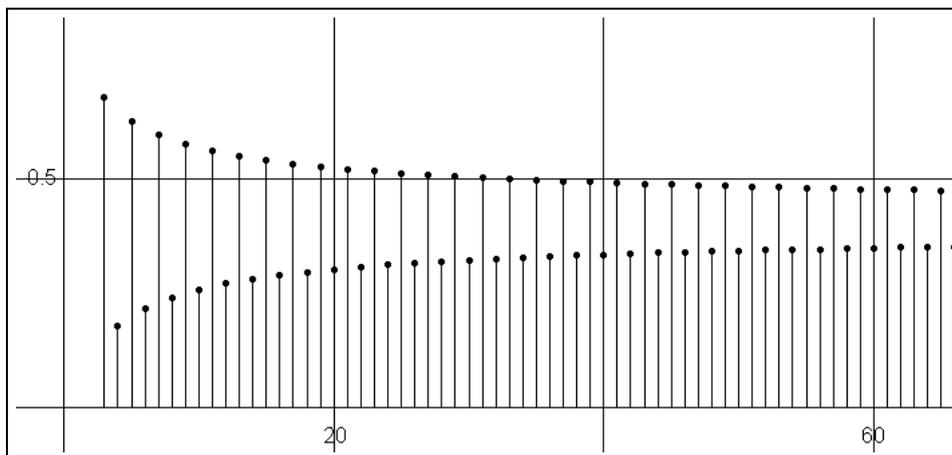
$$\lim_{k \rightarrow \infty} (s_{2k+1} - s_{2k}) = \lim_{k \rightarrow \infty} (-1)^{2k+1} a_{2k+1} = 0$$

Folglich ist $S_+ = S_-$. **Die Reihe ist konvergent** und wegen der Ungleichung $0 \leq s_{2k+1} \leq s_{2k} \leq a_0$ gilt $0 \leq S \leq a_0$ oder $S = \theta a_0$ mit $0 \leq \theta \leq 1$.

(3.3.3) Beispiel: Hiernach sind für $\alpha > 0$ alle Reihen $\sum \frac{(-1)^n}{n^\alpha}$ konvergent. Das sind dann aber Reihen, die z.T. ausgesprochen langsam konvergieren. D.h. man benötigt Strategiefunktionen, die mit $\frac{1}{\epsilon}$ sehr rasch anwachsen.

Das Bild zeigt die ersten Partialsummen der Reihe $\sum_{k=1}^{\infty} \frac{(-1)^{k+1}}{\sqrt{k}}$ ab $n=3$. Die Konvergenz ist in diesem Fall ausgesprochen langsam, was noch viel deutlicher wird, wenn man mehr Terme mitnimmt. Deutlich sichtbar ist das zum Leibnizkriterium gehörige Verhalten. Und die Tabelle verdeutlicht die Langsamkeit der

Konvergenz. Wir haben $S = \theta \cdot 1$ mit $\theta = 0.60\dots?$. (Start hier bei $a_1 = 1$)



n	100	1000	10000	100000	1000000
s_n	0.555	0.589	0.599	0.6033	0.6044
s_{n+1}	0.654	0.620	0.609	0.6065	0.6054

- Zeigen Sie, dass für $0 < \alpha \leq 1$ die Reihen $\sum \frac{(-1)^k}{k^\alpha}$ nicht absolut konvergent sind.
 (3.3.4) Das folgende Kriterium ist einfach, aber ausgesprochen nützlich, ja grundlegend. Es wird später benutzt, um Wurzel- und Quotientenkriterium herzuleiten. Beachten Sie: Es ist ein Kriterium für absolute Konvergenz.

Vergleichskriterium: Ist $\sum b_n$ absolut konvergent und gilt $|a_n| \leq |b_n|$ für fast alle n ,
dann ist $\sum a_n$ absolut konvergent.

Beweis: Wähle für die Strategiefunktion im Cauchy Kriterium $N_a = N_b$:

$$\sum_{k=m}^n |a_k| \leq \sum_{k=m}^n |b_k| < \varepsilon \quad \text{für } n, m \geq N_b(\varepsilon).$$

- Formulieren Sie ein analoges Kriterium für die **bestimmte Divergenz** von $\sum a_k$
 Beispiel: Die **Dezimalbruch**darstellung einer reellen Zahl mit Abschätzung durch die größte Ziffer 9 gibt:

$$\begin{aligned} 3.1745\dots &= 3 + 1 \cdot 10^{-1} + 7 \cdot 10^{-2} + 4 \cdot 10^{-3} + 5 \cdot 10^{-4} + \dots \\ &\leq 9 \cdot (10^{-0} + 10^{-1} + 10^{-2} + 10^{-3} + 10^{-4} + \dots) = \frac{9}{\frac{9}{10}} = 10 \\ &= 9.999999\dots \end{aligned}$$

Also ist die Darstellungsreihe konvergent.

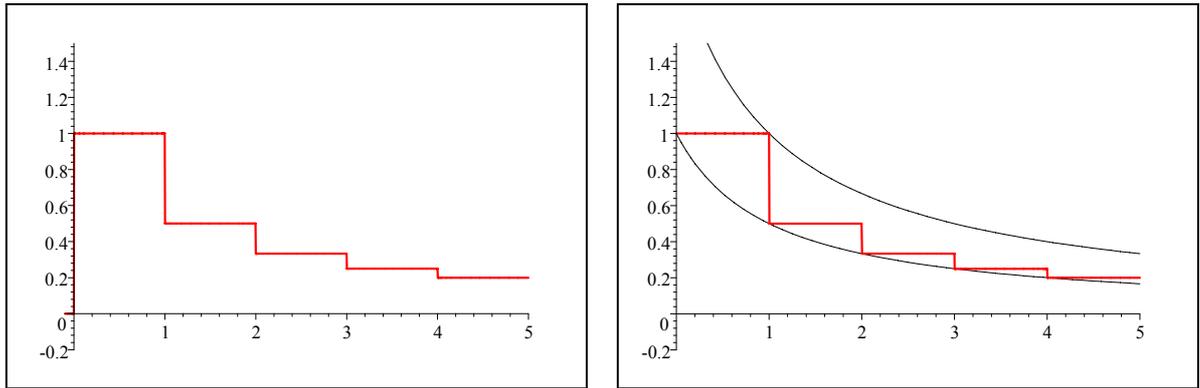
Beispiel. Wir haben bereits gezeigt, dass die harmonische Reihe $\sum \frac{1}{k}$ bestimmt divergent ist. Die Reihen $\sum \frac{(-1)^k}{k^\alpha}$ für $0 < \alpha \leq 1$ sind nach dem Leibnizkriterium konvergent. Da aber $\left| \frac{(-1)^k}{k^\alpha} \right| \geq \frac{1}{k}$ für diese α gilt, sind all diese Reihen nicht absolut konvergent.

Jetzt verwenden wir das Vergleichskriterium, um ein für viele praktische Zwecke sehr nützlich Kriterium herzuleiten, das *Integralvergleichskriterium*. (Dabei benutzen wir einige noch nicht besprochene elementare Eigenschaften reeller Funktionen)

Sei dazu a eine Reihe mit nicht negativen Gliedern ($a_n \geq 0$). Wir definieren eine reelle Funktion $f = f_a : [0, \infty[\rightarrow \mathbb{R}$ mit Sprungstellen durch

$$f(x) = a_n \quad \text{für } n \leq x < n+1 \text{ und } n \in \mathbb{N}$$

Nehmen wir z.B. $a_n = \frac{1}{n+1}$. Dann sieht der Beginn des Graphen von f_a wie im linken Bild aus. Die Sprungstellen sind zu Verdeutlichung verbunden. Wie groß ist der Flächeninhalt unter der roten Kurve bis $x=n \in \mathbb{N}$? Er besteht aus lauter Rechtecken der Breite 1 und der Höhe a_n . D.h. $F(n) = \int_0^n dt f(t) = \sum_{k=0}^{n-1} a_{k+1} = \sum_{k=0}^{n-1} \frac{1}{k+1} = s_{n-1}$. **Die (n-1)-te Partialsumme ist gleich diesem Flächeninhalt.**



Im rechten Bild haben wir zwei weitere Funktionen g und h mit eingezeichnet, die $h(x) \leq f(x) \leq g(x)$ erfüllen. Mit der Monotonie der Integration folgt:

$$H(n) = \int_0^n dx h(x) \leq s_{n-1} = \int_0^n dx f(x) \leq \int_0^n dx g(x) = G(n).$$

Angenommen man kann das Integral $G(n)$ berechnen und die Folge $n \mapsto G(n)$ erweist sich als nach oben beschränkt, dann gilt dasselbe für $n \mapsto s_n$. Wegen $a_n \geq 0$ ist s_n überdies monoton wachsend. Nach dem Monotoniekriterium konvergiert die Reihe $\sum a_n$. Und es gilt $\sum a_n \leq G_0$, wenn $G(n) \leq G_0$ für alle n . Divergiert umgekehrt die zweite Folge $n \mapsto H(n)$, dann muss auch $\sum a_n$ divergieren. Im Beispiel haben wir $h(x) = \frac{1}{x+1}$ gewählt. Also $H(n) = \int_0^n \frac{dx}{x+1} = \ln(n+1)$. Das divergiert und wir sehen erneut (und ganz leicht), dass die harmonische Reihe divergiert.

Fassen wir zusammen:

Integalkriterium Es sei $a_n \geq 0$. Weiter seien g, h reelle Funktionen, für die gilt:

$$0 \leq h(x) \leq a_n \leq g(x) \text{ für alle } x \text{ mit } n \leq x \leq n+1 \text{ und alle } n \in \mathbb{N}.$$

Dann ist $\sum a_k$ absolut konvergent, falls die Folge $n \mapsto G(n) = \int_0^n dx g(x)$ beschränkt ist.

Dann ist $\sum a_k$ bestimmt divergent, falls die Folge $n \mapsto H(n) = \int_0^n dx h(x)$ unbeschränkt ist.

Anschaulich: Der Graph von g hat einen größeren Flächeninhalt. Ist dieser für $n \rightarrow \infty$ beschränkt, dann gilt dasselbe für den Flächeninhalt des Graphen von f und das ist die betrachtete Reihe! Gilt die Ungleichung für die ersten n nicht, dann kann man das Kriterium abgewandelt bei N_0 beginnen lassen. Für die Konvergenz macht das nichts.

Zwei anspruchsvolle Beispiele: $\sum_{k=2}^{\infty} \frac{1}{k \ln(k)}$. Da $h(x) = \frac{1}{x \ln x}$ monoton fällt, haben wir

$$H(n) = \int_2^{n+1} \frac{dx}{x \ln x} \leq \sum_{k=2}^n \frac{1}{k \ln(k)}$$

Das Integral gibt mit Umkehrung Kettenregel $H(n) = \ln|\ln x| \Big|_2^{n+1} = \ln(\ln(n+1)) - \ln(\ln(2))$. Das divergiert (sehr langsam!) für $n \rightarrow \infty$. Folglich divergiert die Reihe $\sum_{k=2}^{\infty} \frac{1}{k \ln(k)}$ auch.

Bei $\sum_{k=2}^{\infty} \frac{1}{k \ln^2(k)}$ gehen die Summanden nur wenig stärker als im ersten Beispiel nach Null. Aber diese Reihe konvergiert. Man hat nämlich

$$\sum_{k=2}^n \frac{1}{k \ln^2(k)} \leq \int_2^{n+1} \frac{dx}{(x+1) \ln^2(x+1)} = \left[-\frac{1}{\ln(x+1)} \right]_2^{n+1}$$

Und das geht (wieder sehr langsam!) nach Null! Also konvergiert die Reihe $\sum_{k=2}^{\infty} \frac{1}{k \ln^2(k)}$

2.3.4 Bedingte Konvergenz. Der Umordnungssatz

(3.4.1) Wir besprechen jetzt das Verhalten von konvergenten, aber nicht absolut konvergenten Reihen. Solche Reihen werden auch **bedingt konvergent** genannt. Für bedingt konvergente Reihen darf man den üblichen Summenzeichenformalismus **nicht** verwenden. Es kommt vielmehr auf die Reihenfolge der Summanden an. Insbesondere gilt auch (3.2.21) nicht, wie das folgende Beispiel verdeutlicht.

(3.4.2) **Gegenbeispiel:** Zunächst erinnern wir daran, dass die Glieder a_n einer konvergenten Reihe notwendig ein Nullfolge bilden. Nach dem Cauchy Kriterium muss ja $|s_n - s_{n-1}| = |a_n| < \varepsilon$ gelten. Das zeigt, dass $n \rightarrow a_n$ notwendig Nullfolge ist

Jetzt betrachten wir die Reihe $\sum \frac{(-1)^k}{\sqrt{k}}$, die ja nach der Leibnizregel konvergent, aber nach Vergleich mit der harmonischen Reihe nicht absolut konvergent ist. Den Grenzwert G kennen wir nicht. Wir bilden einerseits

$$\left(\sum_{n=1}^{\infty} \frac{(-1)^n}{\sqrt{n}} \right) \left(\sum_{k=1}^{\infty} \frac{(-1)^k}{\sqrt{k}} \right) \stackrel{!!}{=} GG \quad \text{nach (G3)}$$

und berechnen andererseits das Cauchyprodukt zu

$$\sum_d (-1)^d \left(\sum_{a=1}^{d-1} \frac{1}{\sqrt{a}\sqrt{d-a}} \right) \stackrel{Def}{=} \sum_d (-1)^d c_d \quad \text{Was ist das?}$$

Die Zahlen $c_d > 0$ können wir wie folgt nach unten abschätzen:

$$c_d = \sum_{a=1}^{d-1} \frac{1}{\sqrt{a(d-a)}} \geq \sum_{a=1}^{d-1} \frac{1}{\sqrt{(d-1)^2}} = \frac{d-1}{d-1} = 1$$

da für die beiden Nennerfaktoren gilt $a \leq d-1$ und $(d-a) \leq d-1$. Vergrößerung des positiven Nenners verkleinert den Ausdruck. Und man hat $(d-1)$ solche Summanden. Diese Reihe ist daher divergent. Oder: Wir haben durch naiv distributives Ausmultiplizieren von zwei konvergenten Reihen eine divergente gemacht.

(3.4.3) Der folgende Satz zeigt, dass derartige Erscheinungen für alle bedingt konvergenten Reihen gelten:

(3.4.4) Großer **Umordnungssatz von Riemann**. Dabei begnügen wir uns mit der einelementigen Partition)

Es sei $a = (\mathbb{N}, n \mapsto a_n, \mathbb{R})$ eine Folge mit unendlicher Reihe $\sum a_k$. Weiter sei $\pi : \mathbb{N} \rightarrow \mathbb{N}$ eine bijektive Abbildung, also eine Permutation der Indexmenge. $b = (\mathbb{N}, k \mapsto a_{\pi(k)}, \mathbb{R})$ sei die Folge mit permutierten Indizes mit unendlicher Reihe $\sum b_k$. Schließlich sei $\sum a_k$ **bedingt konvergent**.

Dann kann man zu jedem $c \in \mathbb{R}$ eine Permutation π von \mathbb{N} finden, so dass $\sum b_k$ gegen c konvergiert. Ebenso gibt es Permutationen, für die $\sum b_k$ bestimmt gegen $\pm\infty$ divergiert. Sind schließlich c_+ und c_- zwei verschiedene Zahlen, die auch $\pm\infty$ sein können, dann gibt es weitere Permutationen, für die c_+ und c_- Häufungspunkte der Folge $n \mapsto \sum a_{\pi(k)}$ sind.

Kurz: durch eine geeignete Umordnung der Summanden kann man alle überhaupt üblicherweise auftretenden Resultate produzieren. $\sum_{k=1}^{1000} \frac{(-1)^{k+1}}{k} = 430\,559\,82$

(3.4.5) Statt des Beweises führen wir die Beweisidee an einem Beispiel konkret vor! Zunächst erinnern wir daran, dass diese Reihen nicht absolut konvergent sind. Und das bedeutet, dass $\sum |a_k|$ über alle Grenzen wächst. Das besagt aber Folgendes:

Zu jeder Zahl $C > 0$ und jedem $N \in \mathbb{N}$ gibt es ein $K \in \mathbb{N}$, für das gilt $\sum_{k=N}^{N+K} |a_k| = |a_N| + |a_{N+1}| + \dots + |a_{N+K}| \geq C$. (*)

□ Zeigen Sie: bei einer bedingt konvergenten Reihe $\mathbb{N} \rightarrow \mathbb{R}$ müssen beide Vorzeichen unendlich oft vorkommen.

□ Zeigen Sie: Die Teilfolge aller positiven a_k ist divergent. Dasselbe gilt, entsprechend abgewandelt, für die Teilfolge aller negativen a_k

Daher gilt (*) auch für diese beiden Teilfolgen

Andernfalls wäre die Folge der zugehörigen Partialsummen doch beschränkt!

(3.4.6) **Beispiel:** Die alternierende harmonische Reihe $1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \dots$ ist bedingt konvergent. In der so gegebenen Reihenfolge konvergiert sie langsam gegen $\ln(2) = 0.693147\dots$. Beispielsweise erhält man $s_{1000} = 0.692647\dots$. Wir wollen die Reihe so umordnen, dass sie gegen einen bestimmten Wert, sagen $S=1.5$ konvergiert. Wie geht man vor? Man führt eine Folge von Schritten aus, die wie gleich beschrieben werden. Dabei ist zu bedenken, dass nach (*) aus (3.4.5) die Teilsummen der positiven a_k beliebig groß und die der negativen beliebig klein ($-\infty$) gemacht werden können, egal von welchem N_0 man startet

Jetzt die Schrittbeschreibung:

- Wir summieren zunächst soviel positive Terme in der Reihenfolge ihres Auftretens, bis deren Summe 1.5 übersteigt. Numerisch gibt das $1 + \frac{1}{3} + \frac{1}{5} = 1.53$.

- Jetzt addiert man negative Terme bis die Summe unter 1.5 kommt. Numerisch genügt in diesem Fall ein Term $1 + \frac{1}{3} + \frac{1}{5} - \frac{1}{2} = 1.03$.

- Jetzt addiert man die nächsten noch nicht verbrauchten positiven Terme, bis die Summe wieder 1.5 übersteigt:

$$1 + \frac{1}{3} + \frac{1}{5} - \frac{1}{2} + \left[\frac{1}{7} + \frac{1}{9} + \frac{1}{11} + \frac{1}{13} + \frac{1}{15} \right] = 1.521$$

- Und wieder unverbrauchte negative Terme

$$1 + \frac{1}{3} + \frac{1}{5} - \frac{1}{2} + \frac{1}{7} + \frac{1}{9} + \frac{1}{11} + \frac{1}{13} + \frac{1}{15} - \frac{1}{4} = 1.27$$

•

$$1 + \frac{1}{3} + \frac{1}{5} - \frac{1}{2} + \frac{1}{7} + \frac{1}{9} + \frac{1}{11} + \frac{1}{13} + \frac{1}{15} - \frac{1}{4} + \left[\frac{1}{17} + \frac{1}{19} + \frac{1}{21} + \frac{1}{23} + \frac{1}{25} \right] = 1.514$$

•

$$1 + \frac{1}{3} + \frac{1}{5} - \frac{1}{2} + \frac{1}{7} + \frac{1}{9} + \frac{1}{11} + \frac{1}{13} + \frac{1}{15} - \frac{1}{4} + \frac{1}{17} + \frac{1}{19} + \frac{1}{21} + \frac{1}{23} + \frac{1}{25} - \frac{1}{6} = 1.347$$

•

$$1 + \frac{1}{3} + \frac{1}{5} - \frac{1}{2} + \frac{1}{7} + \frac{1}{9} + \frac{1}{11} + \frac{1}{13} + \frac{1}{15} - \frac{1}{4} + \frac{1}{17} + \frac{1}{19} + \frac{1}{21} + \frac{1}{23} + \frac{1}{25} - \frac{1}{6} + \left[\frac{1}{27} + \frac{1}{29} + \frac{1}{31} + \frac{1}{33} + \frac{1}{35} \right] = 1.510$$

- Usw.

Offenbar entsteht eine Umordnung der ursprünglichen Reihe mit der gewünschten Konvergenzeigenschaft. Alle Terme der ursprünglichen Reihe werden genau einmal benutzt.

2.3.5 Konvergenzkriterien (II)

(3.5.1) Wir kommen zu weiteren, häufig nützlichen Konvergenzkriterien. Dabei handelt es sich jeweils um Konsequenzen des Vergleichskriteriums. Verglichen wird dabei mit der geometrischen Reihe.

(3.5.2)

Quotientenkriterium:	Für fast alle n gelte $a_n \neq 0$ (1)
	Weiter gebe es eine Zahl q mit $0 < q < 1$, für die für fast alle n gilt $\left \frac{a_{n+1}}{a_n} \right \leq q$ (2)
Dann	ist $\sum a_n$ absolut konvergent .

Beweis: Die Voraussetzungen seien ab N erfüllt. Dann ist $|a_{N+1}| \leq q|a_N|$ wegen (1) und weiter:
 $|a_{N+2}| \leq q|a_{N+1}| \leq q^2|a_N|$. Induktiv folgt $|a_{N+k}| \leq q^k|a_N|$.

Folglich hat man $\sum a_k = \sum_{k=0}^N a_k + \sum_{k=N}^{\infty} a_k$. Zu zeigen ist die absolute Konvergenz der zweiten Summe:

$$\left| \sum_{k=N}^{\infty} a_k \right| \leq \sum_{k=N}^{\infty} |a_k| \leq |a_N| \sum_{k=0}^{\infty} q^k = \frac{|a_N|}{1-q} \quad \text{da } q < 1 \text{ nach (2).}$$

Das - Vergleich mit einer geometrischen Reihe - zeigt die Konvergenz und gibt sogar eine Abschätzung.

(3.5.3) Entsprechend beweist man :

Gilt für fast alle n $\left| \frac{a_{n+1}}{a_n} \right| > q$ mit $q > 1$, dann ist die Reihe divergent.

(3.5.4) !!! **Beachten Sie : Für Konvergenz genügt nicht zu zeigen, dass $\left| \frac{a_{n+1}}{a_n} \right| < 1$ gilt für fast alle n . Das Beispiel der harmonischen Reihe zeigt, dass unbedingt das von n unabhängige q mit $q < 1$ dazwischengeschaltet werden muss!** Im Fall der divergenten harmonischen Reihe gilt $\frac{1}{\frac{n+1}{n}} = \frac{n}{n+1} < 1$ für alle n . Aber es gibt kein $q < 1$, so dass $\frac{n}{n+1} < q$ für alle n . Der Bruch nähert sich der 1 beliebig an!

(3.5.5) Wie steht es mit $\sum_n \frac{x^n}{n!}$? Hier erhält man $\frac{|a_{n+1}|}{|a_n|} = \frac{|x|^{n+1}n!}{|x|^n(n+1)!} = \frac{|x|}{n+1}$. Für $n > 2|x|$ wird das $\leq \frac{1}{2}$. Man hat absolute Konvergenz für alle Werte von x .

(3.5.6) Zur Anwendung des Quotientenkriteriums muss man den Bruch $\frac{a_{n+1}}{a_n}$ von einer Stelle ab bilden können! Das schränkt die Anwendbarkeit manchmal ein. Ein Trick zum Umgehen dieses Problems besteht darin, zu einer neuen Reihe überzugehen, in der alle a_n mit $a_n = 0$ ausgelassen sind. Der Summenwert ändert sich hierdurch nicht.

□ Beweisen Sie: Hat die Folge $n \mapsto \left| \frac{a_{n+1}}{a_n} \right|$ einen Grenzwert q mit $q < 1$, dann gilt für fast alle n auch

$$\left| \frac{a_{n+1}}{a_n} \right| \leq 1 - \frac{\varepsilon}{2}, \quad \text{wenn } \varepsilon = 1 - q \text{ ist.}$$

Und das heißt, dass das Quotientenkriterium anwendbar ist.

(3.5.7) Auch das nächste Kriterium basiert auf dem Vergleich mit der geometrischen Reihe:

Das Wurzelkriterium: Es gebe eine Zahl q mit $0 < q < 1$, für die gilt:
Für fast alle n ist $\sqrt[n]{|a_n|} \leq q$.
Dann ist $\sum a_n$ absolut konvergent

□ Was ist, wenn $\lim_n \sqrt[n]{|a_n|} = w < 1$ existiert?

Beweis: Folgt über die Diskussion der folgenden Ungleichung:

$$\sum_{n=N}^{\infty} |a_n| \leq \sum_{n=N}^{\infty} q^n = \frac{q^{N+1}}{1-q}$$

□ Formulieren Sie das entsprechende Kriterium für bestimmte Divergenz.

Beispiel: $\sum \frac{x^n}{n^n}$ ist danach für alle $x \in \mathbb{C}$ absolut konvergent. Wir haben nach (..) $\lim \sqrt[n]{n} = 1$ und ebenso $\lim \frac{1}{\sqrt[n]{n+1}} = 1$. Mit den Grenzwertsätzen folgt $\sqrt[n]{n|x|^n} \rightarrow |x|$. und ebenso $\sqrt[n]{\frac{|x|^n}{n+1}} \rightarrow \frac{|x|}{1}$. Beide Reihen konvergieren daher ebenso wie $\sum x^n$ für $|x| < 1$ absolut. Oder auch: die beiden zusätzlichen Faktoren n und $\frac{1}{n}$ ändern an der Konvergenz der Reihe nichts

Beispiel: Wie steht es mit der Konvergenz der folgenden beiden Reihen? $\sum n x^n$ und $\sum \frac{1}{n+1} x^n$? Wir wenden das Wurzelkriterium an

(3.5.8) Die folgende Reihe erweist sich als sehr wichtig: $\sum \frac{[\alpha]_n}{n!} x^n$. Wir verwenden das Quotientenkriterium und finden $\left| \frac{[\alpha]_{n+1} x^{n+1} n!}{(n+1)! [\alpha]_n x^n} \right| = \frac{|(a-n)x|}{n+1} = \frac{1-\frac{a}{n}}{1+\frac{1}{n}} x$. Nach dem Grenzwertsatz g3 aus (2.5.26) hat das den Grenzwert x . Folglich ist die Reihe für $|x| < 1$ absolut konvergent.

2.3.6 Die Exponentialfunktion

(3.6.1) Im Vorkurs sind wir der Exponentialfunktion mehrfach begegnet, ohne sie exakt definieren zu können. Wir die Funktionsschar $x \mapsto E_a(x) = a^x$ für $a > 0$ eingeführt und darin a so gewählt, dass die Ursprungsteigung den Wert 1 hatte. Diesen speziellen a -Wert haben wir mit e bezeichnet. Daraus folgte dann, dass $E'_e(x) = E_e(x)$ für alle x galt. Zusätzlich konnte man über $E_a(x) = e^{x \ln(a)}$ alle diese Funktionen auf \exp zurückführen. Aber wir hatten keine genaue Definition für $\exp(x) = e^x$ selbst

Dann war uns die Zahl e im Zusammenhang mit der Eulerschen Formel begegnet. $e^{ix} = \cos x + i \sin x$. Dort hatten wir die Exponentialschreibweise allerdings nur als Gedächtnishilfe für die dortigen Rechenregeln benutzt.

Schließlich konnte man den natürlichen Logarithmus über das Integral $\int_1^x \frac{dt}{t}$ definieren und die zugehörige Umkehrfunktion einführen. Das sollte erneut $x \mapsto e^x$ sein.

Wir wollen jetzt die Exponentialfunktion exakt mit Hilfe von Grenzprozessen einführen und dabei alle die soeben zusammengestellten "losen Enden" verbinden und beweisen.

(3.6.2) Für $n \in \mathbb{N}$ und $x \in \mathbb{C}$ definieren wir $a_n(x) = \left(1 + \frac{x}{n}\right)^n$. Wir wollen die Folge $n \mapsto a_n(x)$ genauer untersuchen. Dabei ist x als äußerer Parameter anzusehen. Es wird sich zeigen, dass $e^x = \lim_n a_n(x)$ ist. Speziell also $e = \lim_n \left(1 + \frac{1}{n}\right)^n$. Die Konvergenz ist recht langsam, wie die nachfolgende Tabelle zeigt, aber die Werte haben die erwartete Größe: ($e = 2.718281\dots$)

n	1	10	100	1000	10000
$\left(1 + \frac{1}{n}\right)^n$	2	2.593	2.704813	2.716923	2.718145

(3.6.3) **Wie steht es mit der Konvergenz dieser Folge?** Hierzu verwenden wir das Monotoniekriterium (2.6.7). Dabei benötigen wir noch eine zweite Folge, nämlich $b_n(x) = \left(1 - \frac{x}{n}\right)^{-n} = \frac{1}{a_n(-x)}$.

Genauer gesagt, zeigen wir

- (1) $a_n(x) \leq a_{n+1}(x)$ für alle x und $n > |x|$
- (2) $b_n(x) \geq b_{n+1}(x)$ für alle x und $n > |x|$.
- (3) $a_n(x) \leq b_n(x)$

Insgesamt können wir argumentieren: Nach (1) wächst die Folge a monoton für große n . Und wegen (2) und (3) ist sie nach oben beschränkt. **Also ist $n \mapsto a_n(x)$ konvergent für alle x .**

▼ Beweis von (1). Wir verwenden AGM: (Beachte $\frac{(n+1)1+n\frac{x}{n}}{n+1} = 1 + \frac{x}{n+1}$; n so groß, dass $1 + \frac{x}{n} > 0$, also $n > -x$ oder sicher $n > |x|$.)

$$a_n(x) = \left(1 + \frac{x}{n}\right)^n = 1 \cdot \left(1 + \frac{x}{n}\right) \cdot \dots \cdot \left(1 + \frac{x_n}{n}\right) \stackrel{AGM}{<} \left(1 + \frac{x}{n+1}\right)^{n+1} = a_{n+1}(x)$$

oder auch $\frac{a_{n+1}(x)}{a_n(x)} \geq 1$.

▼ Beweis von (2): Definitionsgemäß und wegen (1) ist $\frac{b_n(x)}{b_{n+1}(x)} = \frac{a_{n+1}(-x)}{a_n(-x)} \geq 1$ für $n > x$, sicher für $n > |x|$.

▼ Beweis von (3): $\frac{a_n(x)}{b_n(x)} = \left(1 + \frac{x}{n}\right)^n \left(1 - \frac{x}{n}\right)^n = \left(1 - \frac{x^2}{n^2}\right)^n \leq 1$.

Damit ist die Konvergenz von $n \mapsto a_n(x)$ gesichert.

(3.6.4) Erste Eigenschaften der Exponentialfunktion folgen mit Hilfe der Grenzwertsätze:

■ (3.6.4a) Sei $r \in \mathbb{N}$

$$a_n(rx) = \left(1 + \frac{rx}{n}\right)^n = \left(\left(1 + \frac{x}{\frac{n}{r}}\right)^{\frac{n}{r}}\right)^r = \left(a_{\frac{n}{r}}(x)\right)^r$$

Mit der Grenzwertregel G3 folgt:

$$\boxed{e^{rx} = (e^x)^r}. \text{ Wie zu erwarten.}$$

■ (3.6.4b) Ableitung. Wir bilden die Ableitung $a'_n(x)$ nach x und gehen zum Grenzwert über:

$$\begin{aligned} a'_n(x) &= \frac{n}{n} \left(1 + \frac{x}{n}\right)^{n-1} = \frac{a_n(x)}{\left(1 + \frac{x}{n}\right)} \\ \lim_{n \rightarrow \infty} \frac{a_n(x)}{\left(1 + \frac{x}{n}\right)} &= \frac{\lim_{n \rightarrow \infty} a_n(x)}{\lim_{n \rightarrow \infty} \left(1 + \frac{x}{n}\right)} = e^x \end{aligned}$$

Mit der (hier noch zu vermutenden, keineswegs immer gültigen Regel) "Ableitung des Grenzwertes = Grenzwert der Ableitung" ergibt das wie erwartet $\boxed{(e^x)' = e^x}$.

■ (3.6.4c) Ist das **Additionstheorem** der Exponentialfunktion nachweisbar? Einsetzen der Definition gibt

$$a_n(x)a_n(y) = \left(1 + \frac{x+y}{n} + \frac{xy}{n^2}\right)^n \quad \text{und} \quad a_n(x+y) = \left(1 + \frac{x+y}{n}\right)^n$$

Stimmen die Grenzwerte überein?

(3.6.5) Das folgt sofort mit dem folgenden Satz, der auch die Reihendarstellung der Exponentialfunktion liefert:

<p>Satz: Sei $(\mathbb{N}, n \mapsto z_n, \mathbb{C})$ eine konvergente Folge mit $\lim_{n \rightarrow \infty} z_n = Z$.</p> <p>Dann gilt $\boxed{\lim_{n \rightarrow \infty} \left(1 + \frac{z_n}{n}\right)^n = e^Z = \sum_{k=0}^{\infty} \frac{Z^k}{k!}}$</p>

(3.6.6) Kommentare:

- Das gibt eine zweite (besser konvergente) Darstellung von e^Z in Form einer unendlichen Reihe
- Da $\boxed{n \mapsto x + y + \frac{xy}{n}}$ gegen $x+y$ konvergiert, folgt das Additionstheorem für \exp in der Produktdarstellung. (3.6.4c)
- Die Reihe ist absolut konvergent. Die üblichen Rechenregeln sind zulässig. Insbesondere folgt über Multiplikation der Reihen das Additionstheorem.
- Über Euler folgen die Reihen für \sin und \cos .
- Der Nachweis, dass $e^{i\varphi}$ der komplexe Einheitsvektor ist, folgt entweder über die Diskussion der Bahnkurve $t \mapsto e^{it}$ und deren Ableitung in der komplexen Ebene. Oder direkt durch Anwenden des Satzes auf

$$a_n(i\alpha)a_n(-i\alpha) = \lim_{n \rightarrow \infty} \left(1 + \frac{i\alpha}{n}\right)^n \left(1 - \frac{i\alpha}{n}\right)^n = \lim_{n \rightarrow \infty} \left(1 + \frac{\alpha^2}{n^2}\right)^n = e^0 = 1.$$

Dabei ist $z_n = \frac{0}{n} + \frac{\alpha^2}{n^2}$.

(3.6.7) Jetzt der Beweis des Satzes! Wir wollen zeigen, dass stets gilt

$$\lim_{n \rightarrow \infty} \left(1 + \frac{z_n}{n}\right)^n = \sum_{k=0}^{\infty} \frac{Z^k}{k!} \quad \text{wobei} \quad \lim_{n \rightarrow \infty} z_n = Z.$$

Und das heißt, dass $\left| \left(1 + \frac{z_n}{n}\right)^n - \sum_{k=0}^{\infty} \frac{Z^k}{k!} \right|$ in n abzuschätzen ist.

▼1) Gebe $\varepsilon > 0$ vor. Wähle dazu K so groß, dass für $n \geq K$ folgende Bedingungen erfüllt sind. (Die zweite ist erfüllbar wegen der absoluten Konvergenz der Exponentialreihe):

$$(a) |z_n| < |Z| + 1 \quad (b) \sum_{k=K+1}^{\infty} \frac{(|Z| + 1)^k}{k!} < \frac{\varepsilon}{3}$$

▼2) Jetzt schätzen wir für $n > K$ ab, wobei wir den Binomialsatz sowie die Dreiecksungleichung verwenden:

$$\begin{aligned} \left| \left(1 + \frac{z_n}{n}\right)^n - \sum_{k=0}^{\infty} \frac{Z^k}{k!} \right| &\leq \sum_{k=0}^K \left| \binom{n}{k} \frac{|z_n|^k}{n^k} - \frac{Z^k}{k!} \right| + \sum_{k=K+1}^n \binom{n}{k} \frac{|z_n|^k}{n^k} + \sum_{k=K+1}^{\infty} \frac{|Z|^k}{k!} \\ &= S_1 + S_2 + S_3. \quad \text{Klar: } \underline{S_3 < \frac{\varepsilon}{3}} \end{aligned}$$

▼3) Für S_2 gilt:

$$S_2 = \sum_{k=K+1}^n \binom{n}{k} \frac{|z_n|^k}{n^k} = \sum_{k=K+1}^n \frac{n!}{(n-k)!n^k} \frac{|z_n|^k}{k!} \stackrel{(*)}{\leq} \sum_{k=K+1}^n 1 \frac{|z_n|^k}{k!} \leq \sum_{k=K+1}^n \frac{(|Z| + 1)^k}{k!} < \frac{\varepsilon}{3}$$

Nun die Rechnung, die (*) rechtfertigt. Wir benötigen das Resultat auch zur Behandlung von S_1 :

$$\frac{n!}{(n-k)!n^k} = \frac{[n]_k}{n^k} = 1 \left(1 - \frac{1}{n}\right) \left(1 - \frac{2}{n}\right) \dots \left(1 - \frac{k-1}{n}\right) \leq 1 \quad \text{da } k \leq n!$$

Das ist (*)

$$\text{Oder auch } \boxed{\binom{n}{k} \frac{1}{n^k} = \varepsilon_n + \frac{1}{k!} \rightarrow \frac{1}{k!} \quad \text{für } n \rightarrow \infty, \varepsilon_n = \varepsilon_n(k) \text{ Nullfolge}}$$

▼4) Die erste Summe S_1 besteht aus endlich vielen Summanden, wobei jeder einzeln mit n nach Null geht!

$$\binom{n}{k} \frac{|z_n|^k}{n^k} - \frac{Z^k}{k!} = \varepsilon_n |z_n|^k + \frac{1}{k!} [|z_n|^k - |Z|^k]$$

Im letzten Beitrag geht ein, das z_n gegen Z geht. Daher kann (nach den Grenzwertsätzen) auch S_1 kleiner als $\frac{\varepsilon}{3}$ gemacht werden, wenn n ausreichend groß wird. Oben war ja K fixiert und $n > K$ zugelassen.

□ Wogegen konvergiert $n \rightarrow b_n(x)$? Und wie?

▼5) **Damit ist der Satz bewiesen!**

(3.6.8) Damit haben wir insbesondere die Reihendarstellung der Exponentialfunktion gewonnen:

$$e^x = \sum_{k=0}^{\infty} \frac{x^k}{k!} = \frac{1}{0!} + \frac{x}{1!} + \frac{x^2}{2!} + \dots = 1 + x + \frac{x^2}{2} + \frac{x^3}{6} + \frac{x^4}{24} + \dots$$

Das ist eine der wichtigsten Reihen überhaupt. Sie ist für alle $x \in \mathbb{C}$ konvergent. Als Beispiel geben wir einige Partialsummen für $x=2$ an. Zum Vergleich noch die entsprechenden Werte von $a_n(2)$ und $\frac{1}{b_n(-2)}$.

n	1	2	5	10	20	50	100	500
$\sum_{k=0}^n \frac{x^k}{k!}$	3	5	7.266	7.38899	7.3890560989306050	7.389	7.389	7.389...
$a_n(2)$	2	4	5.378	6.192	6.727	7.107	7.245	7.374...
$b_n(2)$	-	-	12	9.313	8.225	7.699	7.540	7.403...

Vergleich mit dem exakten Wert $e^2 = 7.3890560989306502272\dots$ zeigt, dass die Reihe für $n=20$ bereits 12 Nachkommastellen korrekt gibt, wogegen die beiden Folgen sich noch nicht einmal auf die erste Nachkommastelle einigen können.

□ Vergleichen Sie die ersten Terme der Reihenentwicklung von e^3 und e^{-3} . Welchen Nutzen hat das Leibnizkriterium hier?

□ Bestimmen Sie die Reihenentwicklungen für $\cos x$ und $\sin x$ mit Hilfe der Eulerschen Formel. Wieso darf man das?

(3.6.9) Das Wachstum der Fakultät. Die Folge $n \mapsto n!$ erweist sich als wichtig. Sie divergiert für $n \rightarrow \infty$ bestimmt nach ∞ . Aber man hat keinen zunächst keinen Vergleichsmaßstab dafür wie rasch diese Divergenz ist. Genauer: Man hat keine der üblichen elementar konstruierbaren Funktionen, von der $n!$ die Restriktion auf die natürlichen Zahlen wäre. So wie $n \mapsto n^3$ die Restriktion von $x \mapsto x^3$ ist. Wir leiten jetzt eine erste Abschätzung für das Wachstumsverhalten von $n!$ her. Es zeigt sich, dass dieses Wachstum sehr stark ist, stärker als exponentiell. Grob wächst $n!$ wie $n^n = e^{n \ln n}$. D.h. schneller als e^n , aber langsamer als e^{n^2} .

Wir haben oben gezeigt: $\boxed{a_k(1) \leq e \leq b_{k+1}(1)}$ für $k=1,2,3,\dots$. Das gibt ausgeschrieben

$$\left(\frac{k+1}{k}\right)^k \leq e \leq \left(\frac{k+1}{k}\right)^{k+1} \quad k=1,2,3,\dots$$

Oder

$$\begin{aligned} \left(\frac{2}{1}\right)^1 &\leq e \leq \left(\frac{2}{1}\right)^2 \\ \left(\frac{3}{2}\right)^2 &\leq e \leq \left(\frac{3}{2}\right)^3 \\ \left(\frac{4}{3}\right)^3 &\leq e \leq \left(\frac{4}{3}\right)^4 \\ \left(\frac{5}{4}\right)^4 &\leq e \leq \left(\frac{5}{4}\right)^5 \\ &\dots \\ \left(\frac{n}{n-1}\right)^{n-1} &\leq e \leq \left(\frac{n}{n-1}\right)^n \\ \left(\frac{n+1}{n}\right)^n &\leq e \leq \left(\frac{n+1}{n}\right)^{n+1} \end{aligned}$$

Multiplikation dieser Ungleichungskette für $k=1,\dots,n-1$ (nicht n) gibt:

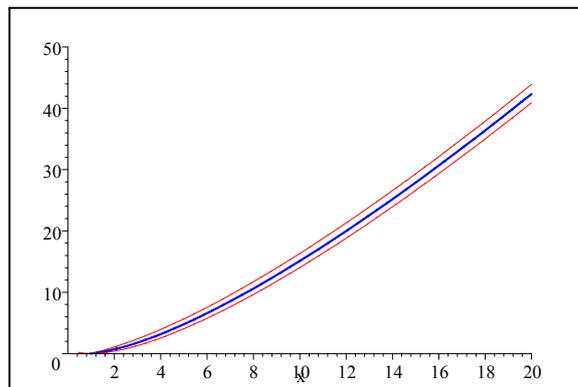
$$\frac{n^{n-1}}{(n-1)!} \leq e^{n-1} \leq \frac{n^n}{(n-1)!}$$

Die beiden Ungleichungen lassen sich umstellen zu einer Ungleichung für $n!$, nämlich

$$\boxed{n \left(\frac{n}{e}\right)^{n-1} = e \left(\frac{n}{e}\right)^n \leq n! \leq e n \left(\frac{n}{e}\right)^n}$$

$$\boxed{n \ln\left(\frac{n}{e}\right) + 1 \leq \ln n! \leq n \ln\left(\frac{n}{e}\right) + 1 + \ln n}$$

Damit erhält man eine ungefähre Vorstellung (mit Hilfe der elementaren Funktionen) davon, wie rasch $n!$ mit n wächst. Das nachfolgende Bild zeigt die zweite logarithmische Ungleichungskette. Blau in der Mitte ist dabei der Logarithmus der Gammafunktion dargestellt, die eine exakte Interpolation der Fakultäten liefert.



2.3.7 Die Binomialreihe

(3.7.1) Die Definition der Potenzfunktion $\boxed{(-1, \infty[, x \mapsto (1+x)^\alpha, \mathbb{R})}$ ist für die meisten werte von α bisher noch problematisch. Ausnahmen sind rationale Exponenten wie $\alpha = -1, -\frac{1}{2}, \frac{1}{2}, 2, \dots, -\frac{1}{3}, -\frac{2}{3}, \dots$

(3.7.2) Erneut ist eine Reihendefinition günstig. Wir führen eine Reihe $B(z; \alpha)$ mit zwei äußeren Parametern ein. (Beachten Sie an Stellen wie dieser den Nutzen der Konvention, dem Symbol Σa_n je nach Situation die Rolle der Folge der Partialsummen, des zugehörigen Grenzwertes oder aber eines Rechenausdruckes mit zugehörigen Termumformungsregeln zu geben!).

(3.7.3) Die Definition

$$\boxed{B(z; \alpha) := \sum_{k=0}^{\infty} \binom{\alpha}{k} z^k \quad \text{absolut konvergent für } |z| < 1. \quad \binom{\alpha}{k} := \frac{[\alpha]_k}{k!} = \frac{\alpha(\alpha-1)\dots(\alpha-k+1)}{k!}}$$

Der Konvergenzbereich folgt sofort mit Hilfe des Quotientenkriteriums:

$$\frac{\frac{[\alpha]_{k+1}}{(k+1)!} |z|^{k+1}}{\frac{[\alpha]_k}{k!} |z|^k} = \frac{|\alpha - k|}{k+1} |z| = \frac{|\frac{\alpha}{k} - 1|}{1 + \frac{1}{k}} |z| \rightarrow 1|z|$$

(3.7.4) Einige α -Werte geben bekannte Spezialfälle. $\alpha = n \in \mathbb{N}$ gibt die übliche Binomialformel für ein Polynom vom Grade n in z . In diesem Fall geht die Reihe über in eine endliche Summe, die natürlich für alle z konvergent ist.

$$B(z, n) = (1+z)^n \quad n \in \mathbb{N} \quad \text{Endliche Summe} = \text{Polynom Grad } n.$$

□ Auch $n=-1$ sollte ein vertrautes Resultat geben, nämlich die geometrische Reihe. Testen Sie das. Und was ist mit $n=-2$?

(3.7.5) Für ganzes n und m gilt $(1+z)^n (1+z)^m = (1+z)^{n+m}$. Wenn $B(z; \alpha)$ für alle α die Potenz $(1+z)^\alpha$ definiert, dann muss diese Formel allgemein gelten. Sie tut es:

Satz: Sei	$\alpha, \beta \in \mathbb{R}$ und $z \in \mathbb{C}$ mit $ z < 1$.
Dann gilt	$B(z; \alpha) B(z; \beta) = B(z; \alpha + \beta)$

Beweis: Die Reihen sind absolut konvergent. Wir können das Cauchyprodukt über den üblichen Summenzeichenformalismus bestimmen.

$$\begin{aligned} B(z; \alpha) B(z; \beta) &= \sum_{m,n=0}^{\infty} \binom{\alpha}{m} \binom{\beta}{n} z^{m+n} \\ &\stackrel{\text{Cauchypr.}}{=} \sum_{d=0}^{\infty} \left(\sum_{s=0}^d \binom{\alpha}{d-s} \binom{\beta}{s} \right) z^d \\ &\stackrel{(\text{unten}^*)}{=} \sum_{r=0}^{\infty} \binom{\alpha + \beta}{r} z^r = B(z; \alpha + \beta) \end{aligned}$$

(3.7.6) Hieraus folgt sofort für alle $n \in \mathbb{N}$

$$B(z; n\alpha) = (B(z; \alpha))^n \quad \text{und} \quad B(z; \alpha) = \left(B(z; \frac{\alpha}{n}) \right)^n$$

Also gilt für jede rationale Zahl $p = \frac{n}{m}$

$$\boxed{B(z; \frac{n}{m}) = (1+z)^{\frac{n}{m}}}$$

Die Ausdehnung auf irrationale Werte von α diskutieren wir hier noch nicht.

(3.7.8) Die letzte Umformung (*) aus dem Beweis ist noch zu rechtfertigen:

Hilfssatz: Für $\alpha, \beta \in \mathbb{R}$ gilt: $\boxed{\sum_{s=0}^r \binom{\alpha}{r-s} \binom{\beta}{s} = \binom{\alpha + \beta}{r}}$

Beweis: Wir erinnern an folgende Sachverhalte. (1) Hat ein Polynom vom Grade höchstens n mehr als n Nullstellen, dann ist es das Nullpolynom. Man sagt: *identisch Null*. (2) $\alpha \mapsto [\alpha + k]_n$ ist ein Polynom vom Grade n . (3) Für $\alpha, \beta \in \mathbb{N}$ ist die Identität richtig, da der Binomialsatz gilt.

Jetzt der eigentliche Beweis: Die Gleichung gilt für $\alpha, \beta \in \mathbb{N}$. Interpretiere die Differenz beider Seiten als **Polynom** (vom Grad höchstens r) in α . Das schreiben wir wie folgt:

$$\alpha \mapsto \sum_{s=0}^r \frac{1}{(r-s)!s!} [\alpha]_{r-s} [\beta]_s - \frac{1}{r!} [\alpha + \beta]_r.$$

$$\alpha \mapsto P(\alpha) = \sum_{s=0}^r \binom{r}{s} [\alpha]_{r-s} [\beta]_s - [\alpha + \beta]_r$$

$$\boxed{[\alpha + \beta]_n = \sum_{s=0}^n \binom{n}{s} [\alpha]_{n-s} [\beta]_s} \quad \text{Zu beweisende Polynomidentität.}$$

Wähle jetzt $\beta \in \mathbb{N}$. Dann hat das Polynom P für jedes $\alpha \in \mathbb{N}$ eine Nullstelle, also sicher mehr als r und ist mithin identisch Null (für alle $\alpha \in \mathbb{R}$). Oder: Die Formel gilt für alle $\alpha \in \mathbb{R}$ und $\beta \in \mathbb{N}$. Jetzt interpretiere man den Ausdruck umgekehrt als Polynom in β und argumentiere entsprechend. Das gibt das gewünschte Resultat.

(3.7.9) Es ist sinnvoll, sich einmal konkret anzuschauen, was da bewiesen ist. Wir zeigen das für $\alpha, \beta \in \mathbb{R}$ und $r=2$ bzw. $r=3$. Das ergibt schon beträchtliche Polynomumformungen.

$$r = 2 : \quad \boxed{[\alpha]_2 + 2[\alpha]_1 [\beta]_1 + [\beta]_2 = [\alpha + \beta]_2}$$

$$\alpha(\alpha - 1) + 2\alpha\beta + \beta(\beta - 1) = (\alpha + \beta)(\alpha + \beta - 1)$$

$$\alpha^2 - \alpha + 2\alpha\beta + \beta^2 - \beta = \alpha^2 - \alpha + 2\alpha\beta + \beta^2 - \beta$$

$$r = 3 : \quad \binom{3}{0} [\alpha]_3 + \binom{3}{1} [\alpha]_2 [\beta]_1 + \binom{3}{2} [\alpha]_1 [\beta]_2 + \binom{3}{3} [\beta]_3 = [\alpha + \beta]_3$$

$$\boxed{[\alpha]_3 + 3[\alpha]_2 [\beta]_1 + 3[\alpha]_1 [\beta]_2 + [\beta]_3 = [\alpha + \beta]_3}$$

$$\alpha(\alpha - 1)(\alpha - 2) + 3\alpha(\alpha - 1)\beta + 3\alpha\beta(\beta - 1) + \beta(\beta - 1)(\beta - 2)$$

$$= \alpha^3 - 3\alpha^2 + 2\alpha + 3\alpha^2\beta - 6\alpha\beta + 3\alpha\beta^2 + \beta^3 - 3\beta^2 + 2\beta$$

$$[\alpha + \beta]_3 = (\alpha + \beta)(\alpha + \beta - 1)(\alpha + \beta - 2)$$

$$= \alpha^3 - 3\alpha^2 + 2\alpha + 3\alpha^2\beta - 6\alpha\beta + 3\alpha\beta^2 + \beta^3 - 3\beta^2 + 2\beta$$

□ Wir haben $\frac{d}{dx}(1+x)^\alpha = \alpha(1+x)^{\alpha-1}$. Berechnen Sie die rechte Seite über die Binomialreihe und zeigen Sie, dass sich $\sum \binom{\alpha}{k} \frac{d}{dx} x^k$ ergibt. Welche weitere Vermutung über -eigenschaften absolut konvergenter Reihen liegt nahe?

(3.7.10) Wir diskutieren jetzt noch einige Resultate zum Logarithmus, der Umkehrfunktion der Exponentialfunktion auf heuristischem Niveau. D.h. wir übertragen für den Fall absoluter Konvergenz einige Umformungen vom endlichen auf den abzählbar unendlichen Fall, die wir erst später rechtfertigen werden. Im Augenblick geht es darum, mit den Reihen und den damit verbundenen Rechnungen vertraut zu werden.

Wie sieht die Reihendarstellung für den Logarithmus aus? Wir erhalten den Logarithmus durch

auffösen von $e^x = 1 + y$ nach x . Also $x = \ln(1+y)$. Mit der Näherung durch a_n folgt

$$\begin{aligned} \left(1 + \frac{x}{n}\right)^n &= 1 + y & \text{Also } \boxed{a_n(x) \approx 1+y} & \text{ Das gibt} \\ 1 + \frac{x}{n} &= (1 + y)^{\frac{1}{n}} & \text{Hoffnung: } \lim_{n \rightarrow \infty} x_n(y) &= \ln(1 + y) \\ x_n(y) &= n \left[(1 + y)^{\frac{1}{n}} - 1 \right] & \text{Binomialreihe!} \\ &= n \sum_{k=1}^{\infty} \frac{\binom{1}{n}^k}{k!} y^k & \text{Ausgeschrieben} \\ x_n(y) &= n \left[\frac{1}{1!} y - \frac{1}{n} \frac{(1 - \frac{1}{n})}{2!} y^2 + \frac{1}{n} \frac{(1 - \frac{1}{n})(2 - \frac{1}{n})}{3!} y^3 - \dots \right] \\ \lim_{n \rightarrow \infty} x_n(y) &\stackrel{(\text{????})}{=} \frac{1}{1} y + \frac{-1}{2} y^2 + \frac{1}{3} y^3 - \frac{1}{4} y^4 + \dots = \sum_{k=1}^{\infty} \frac{(-1)^{k+1}}{k} y^k \\ &\boxed{\text{Wir erwarten für } |y| < 1 \quad \ln(1+y) = \sum_{k=1}^{\infty} \frac{(-1)^{k+1}}{k} y^k} \end{aligned}$$

Diese Reihe ist nach dem Wurzelkriterium oder dem Quotientenkriterium für $|y| < 1$ absolut konvergent!!! und für $y = +1$ nach dem Leibnizkriterium bedingt konvergent. Etwa:

$$\sqrt[n]{\frac{|y|^n}{n}} \rightarrow |y|$$

Die kritische Stelle: Dürfen wir den Grenzübergang $n \rightarrow \infty$ in den einzelnen Summanden ausführen. Oder auch: Dürfen wir Grenzübergang und Summation vertauschen? Die bisherigen Grenzwertsätze rechtfertigen das nur für endlich viele Summanden!

- Wieso ist keine Reinenentwicklung von $\ln y$ in y zu erwarten? Wieso von $\ln(1+y)$?

Jetzt bestimmen wir dieselbe Reihe für reelles y mit Hilfe der Integraldefinition des Logarithmus wie folgt

$$\begin{aligned} \ln(1 + y) &= \int_1^{1+y} \frac{dx}{x} = \int_0^y \frac{du}{1 + u} \\ &= \int_0^y du \left(\sum_{k=0}^{\infty} (-1)^k u^k \right) \\ &\stackrel{(\text{????})}{=} \sum_{k=0}^{\infty} (-1)^k \int_0^y du u^k \quad \# \# \# \quad \text{zulässig??} \\ &= \sum_{k=0}^{\infty} (-1)^k \frac{1}{k+1} y^{k+1} \\ &= \sum_{k=1}^{\infty} (-1)^{k+1} \frac{1}{k} y^k \quad \text{Dieselbe Reihe! **Konvergent für } |y| < 1. \end{aligned}**$$

Einige Rechenbeispiele:

$\ln 1.5 = 0.405\,465\,108\dots$	$\sum_{k=1}^5 \frac{(-1)^{k-1}}{k} 0.5^k = 0.407\,291$	$\sum_{k=1}^{10} \frac{(-1)^{k-1}}{k} 0.5^k = 0.405\,434$
----------------------------------	--	---

- $1+y=(1+x)^\alpha$ gibt $x=(1+y)^{\frac{1}{\alpha}} - 1$. Was sollte das für die Reihen bedeuten?

Kap. 2.4: Stetige und differenzierbare Funktionen

2.4.0 Umgebungen

Im Zusammenhang mit den reellen Funktionen werden wir häufig den Begriff der ε -Umgebung verwenden, den wir später noch stark verallgemeinern. Er gehört zum Gebiet der Topologie, die die in gewissem Sinne die allgemeinsten geometrischen Strukturen erfasst. Dazu gehört insbesondere der geometrische Aspekt der Grenzwertprozesse. Der Begriff der ε -Umgebung und allgemeiner der der Umgebung (eines Punktes) erweist sich als ausgesprochen wichtig.

Definition: Es sei $x_0 \in \mathbb{R}$ und $\varepsilon > 0$.
Dann bezeichnen wir die Menge $U_\varepsilon(x_0) =]x_0 - \varepsilon, x_0 + \varepsilon[$ als (*offene*) ε -Umgebung von x_0 (in \mathbb{R}).
Ist $z_0 \in \mathbb{C}$, definieren wir entsprechend
 $U_\varepsilon(z_0) = \{z \in \mathbb{C}, |z - z_0| < \varepsilon\}$.

Die ε -Umgebungen in \mathbb{C} sind offene (=ohne Rand) Kreisscheiben vom Radius ε und Mittelpunkt z_0 . (Verallgemeinerung in Kap.6.2.)

Die Bedingung dafür, dass die Folge $n \mapsto a_n$ gegen A konvergiert, können wir mit Hilfe des Umgebungsbegriffes wie folgt formulieren:

Zu jedem $\varepsilon > 0$ gibt es ein N_0 , derart dass $a_n \in U_\varepsilon(A)$ für $n \geq N_0$.

Also: Zu jedem $\varepsilon > 0$ kann man ein N_0 finden, so dass alle a_n mit $n \geq N_0$ in der ε -Umgebung $U_\varepsilon(A)$ liegen.

□ Sei $z = (\mathbb{N}, n \mapsto \frac{e^{in}}{n+1}, \mathbb{C})$ und $k \in \mathbb{N}, k > 0$. Bestimmen Sie K_0 , derart, dass $z_n \in U_{\frac{1}{k}}(0)$ für $n \geq K_0$.

2.4.1 Glatte Funktionen

(4.1.1) Wir betrachten zunächst nur Funktionen $f: D \rightarrow \mathbb{R}$ mit $D \subset \mathbb{R}$, wobei D meist noch zusätzliche Bedingungen erfüllen soll. Der Graph ist dann eine Teilmenge von $D \times \mathbb{R}$, liegt also in der Ebene \mathbb{R}_{xy}^2 .

(4.1.2) Ein Hauptziel wird sein, den intuitiven Begriff der *Glattheit des Graphen* zu präzisieren. werden. Oder auch zu klären: Wann ist ein Graph *zeichnenbar* ?

(4.1.3) Dabei sind zwei unabhängige Eigenschaften ("Freiheitsgrade" oder "Dimensionen") zu berücksichtigen:

- ★ Die geometrische Ausdehnung der Glattheitseigenschaft in D .
- ★ Die Intensität oder Stärke der Glattheit. ("glatt, noch glatter...")

(4.1.4) Hinsichtlich der ersten Dimension, dem **Ausdehnungsbereich der Glattheitseigenschaft** unterscheidet man üblicherweise folgende Fälle:

- Die Eigenschaft gilt
 - im Punkte $x_0 \in D$ (Über die anderen Punkte wird nichts gesagt)
 - in einer Umgebung $U_\varepsilon(x_0)$ des Punktes. Über die Größe von ε wird nichts gesagt. Nur $\varepsilon > 0$ wird verlangt.
 - überall **bis auf endlich viele Punkte** von D , die "Ausnahmepunkte". Man sagt, die Eigenschaft gelte "stückweise".
 - in **jedem** Punkt von D , man sagt "global".

(4.1.5) Was die zweite Eigenschaft, die "Glattheitsstärke" anbelangt, so hat man folgende Skala zunehmender Glattheit, die (oben) mit ganz beliebigen Funktionen beginnt. Was die stückweise glatten Funktionen anbelangt, so haben sich gewisse Kombinationen als besonders wichtig erwiesen, die wir in der rechten Spalte mit angeben!

<i>Punkteigenschaft</i>	<i>Global</i>	<i>Stückweise</i>
Beliebig	Beliebig	Beliebig
Stetig in x_0	Stetig (C^0)	stw. C^0
Differenzierbar in x_0	Stetig diff. (C^1)	$C^0 + \text{stw.}C^1$
2-fach diff.	2-fach stetig diff (C^2)	$C^0 + \text{stw.}C^2$
:	:	:
n-fach diff	n-fach stetig diff (C^n)	$C^0 + \text{stw.}C^n$
:	:	:
Beliebig oft diff. (C^∞)	C^∞	$C^0 + \text{stw.}C^\infty$
Analytisch (C^ω)	C^ω	$C^0 + \text{stw.}C^\omega$

Die Begriffe *stetig* und *differenzierbar* werden im Laufe dieses Kapitels präzisiert.

Dabei steht C^1 nicht nur für "einmal differenzierbar", sondern die Ableitung soll auch noch stetig sein. "f ist C^1 " heißt: f ist im Definitionsbereich (mindestens) einmal differenzierbar und die Ableitung ist stetig. Jede C^2 -Funktion ist automatisch C^1 und erst recht C^0 . Natürlich sollte man die Aussagen "f ist stetig" und "die Ableitung von f ist stetig" sorgfältig unterscheiden. Es gibt Funktionen, die differenzierbar sind, aber trotzdem nicht C^1 . Beispiel in (4.6.12).

$C^0 + \text{stw.}C^1$ soll besagen: In allen Punkten stetig und bis auf **endlich viele** Ausnahmepunkte sogar differenzierbar. "stw" steht für *stückweise* d.h. "bis auf endlich viele Ausnahmepunkte". *Analytisch* schließlich ist eine noch stärkere Glattheitseigenschaft als beliebig oft differenzierbar. Wir werden darauf zurückkommen. Das Kurzsymbol C^n für die Eigenschaft "n-fach-differenzierbar" hat sich inzwischen ziemlich eingebürgert.

(4.1.6) Beachten Sie, dass jede dieser Stufen die Zahl der erfüllenden Funktionen stark einschränkt. D.h. es gibt etwa überall stetige, aber nirgends differenzierbare Funktionen.

(4.1.7) Es gibt noch einige weitere Eigenschaften, die man in diese Glattheithierarchie einfügen sollte. Drei wollen wir unmittelbar nennen: *Monotone Funktionen* und *konvexe Funktionen* und *Lipschitz-stetige Funktionen*.

□ Versuchen Sie durch Nachschlagen in einem oder mehreren Analysisbüchern herauszubekommen, wo in unserer Glattheitsleiter diese Funktionen einzuordnen sind. Und ob das überhaupt geht. .

2.2.2 Stetigkeit

(4.2.1) Wir beginnen mit dem fundamentalen Begriff der Stetigkeit, der im wesentlichen die glatten Funktionen aus der Menge der beliebigen aussondert.

(4.2.2) Wir benötigen eine zusätzliche **Definition**, mit der wir den Umgebungsbegriff ergänzen.

<p>Sei $D \subset \mathbb{R}$, $x_0 \in D$ und $U_\varepsilon(x_0) =]x_0 - \varepsilon, x_0 + \varepsilon[$ eine ε-Umgebung von x_0 in \mathbb{R}.</p> <p>Dann bezeichnen wir $U_\varepsilon^D(x_0) = U_\varepsilon(x_0) \cap D$ als eine Relativumgebung von x_0 in D.</p>
--

Beispiele: $D =]0, 1[$ und $U_\varepsilon(0) =]-\varepsilon, \varepsilon[$ mit $0 < \varepsilon < 1$. Dann ist $U_\varepsilon^D(0) =]0, \varepsilon[$. Oder $D = \{0\} \cup]1, \infty[$ und $U_\varepsilon(0)$ wie oben. Dann ist $U_\varepsilon^D(0) = \{0\}$.

Ist D Definitionsbereich einer Funktion f , dann besteht $U_\varepsilon^D(a)$ aus all denjenigen Punkten der ε -Umgebung von a , für die man auch den Funktionswert bilden kann.

(4.2.3) Die folgende **intuitive Bedeutung von "f ist stetig in $x_0 \in D$ "** macht den Stetigkeitsbegriff für die Anwendungen bedeutsam:

Enthält $U_\varepsilon^D(x_0)$ für ein $\varepsilon > 0$ ausreichend viele Elemente, dann wird $f(x_0)$ durch die Werte von $f(x)$ in $U_\varepsilon^D(x_0) - \{x_0\}$ festgelegt, ist also aus den Funktionswerten der benachbarten Punkte vorhersagbar! Insbesondere ist das der Fall, wenn ein ganzes $U_\varepsilon(x_0)$ in D enthalten ist, so dass hierfür $U_\varepsilon^D(x_0) = U_\varepsilon(x_0)$ gilt.

(4.2.4) Oder auch: Wir betrachten die Funktion vom Feldstandpunkt. $D - \{x_0\}$ sei unser Konfigurationsraum. Wir stellen uns vor, dass wir den Feldwert $f(x)$ für beliebige $x \in D - \{x_0\}$ bestimmen können. **Sind wir mit diesen Daten in der Lage den Feldwert $f(x_0)$ vorherzusagen, obwohl wir ja in x_0 selbst nicht messen können?** Wir werden sehen: Die Eigenschaft "stetig" erlaubt diese Vorhersage (unter der gemachten Einschränkung)!

(4.2.5) **Man hat mehrere mathematisch gleichwertige Möglichkeiten, "stetig" zu definieren bzw. diese Eigenschaft zu charakterisieren!** Die wichtigste davon paßt das für Folgen entwickelte Strategiefunktionskonzept an die neue Situation an. Bei Folgen kam man zur Grenze, wenn man n nach ∞ gehen ließ. Oder $\frac{1}{n}$ nach Null. **Jetzt muss man x gegen den Aufpunkt x_0 gehen lassen.** Das liefert folgende Formulierung:

Definition:	Sei $f: D \rightarrow \mathbb{R}$ und $x_0 \in D$.
Dann	heißt f in x_0 stetig, wenn gilt:
★	Zu jedem $\varepsilon > 0$ gibt es ein $\delta = \delta(\varepsilon) = \delta(\varepsilon, x_0)$ derart, dass gilt: $ f(x) - f(x_0) < \varepsilon$ für alle $x \in D$ mit $ x - x_0 < \delta = \delta(\varepsilon)$.
★	f heißt stetig, wenn f in jedem $x_0 \in D$ stetig ist.
★	f heißt gleichmäßig stetig, wenn man eine gemeinsame Strategiefunktion für alle $x_0 \in D$ finden kann.

Eine alternative Schreibweise für "f ist stetig in x_0 " ist: $\lim_{x \rightarrow x_0} f(x) = f(x_0)$.

(3.2.6) Wir können das Szenenbild auch etwas anders formulieren, indem wir eine neue unabhängige Variable einführen und schreiben

$$\boxed{f(x_0 + \Delta x) = f(x_0) + F(x_0, \Delta x)} \quad \text{mit} \quad \begin{array}{l} \Delta x = x - x_0 \\ \Delta y = F(x_0, \Delta x) = f(x_0 + \Delta x) - f(x_0) \end{array}$$

Diese Darstellung entspricht folgendem Vorgehen:

x_0 wird als neuer Nullpunkt gewählt. Dann ist Δx neue unabhängige Variable und $\Delta y = F(x_0, \Delta x)$ ist die Wertänderung, die man erhält, wenn man von x_0 nach $x_0 + \Delta x$ geht. Oder auch den Fehler, den man macht, wenn man $f(x_0 + \Delta x)$ durch $f(x_0)$ ersetzt. **Ist die Funktion stetig, dann geht der Fehler nach Null, wenn Δx nach Null geht** $\lim_{x \rightarrow x_0} F(x_0, \Delta x) = 0$.

Oder auch: Der Herausforderer legt über die Vorgabe von ε eine Stellenzahl (im Funktionswert) vor. Und der Verteidiger stellt sicher, dass für $|\Delta x| < \delta$ der Fehler keinen Einfluss auf die vorgegebenen Stellen hat. In den vorgegebenen Stellen unterscheiden sich die Werte $f(x_0)$ und $f(x_0 + \Delta x)$ nicht. Man sagt: *Der Fehler kann beliebig klein gemacht werden.*

(4.2.7) Die Größe Δx entspricht der Größe $\frac{1}{n}$ bei den Folgen. Beide gehen in der Grenze nach Null.

(4.2.8) Häufig hat man es mit folgender Situation zu tun: $\delta(\varepsilon)$ sei Strategiefunktion für f . Man möchte, dass $A \cdot |f(x_0 + \Delta x) - f(x_0)| < \varepsilon$ wird für ein gegebenes $A > 0$. Dann genügt es, $|\Delta x| < \delta(\frac{\varepsilon}{A})$ zu wählen.

(4.2.9) **Beispiel:** $f(x) = x^2$ und $x_0 = 1$. Dann ist $F(1, \Delta x) = (1 + \Delta x)^2 - 1 = 2\Delta x + \Delta x^2$ mit der erwarteten Nullstelle bei $\Delta x = 0$!

□ Formen Sie die Fehlerfunktion $F(x_0, \Delta x)$ für die Sinusfunktion in naheliegender Weise weiter um, so dass man die Nullstelleneigenschaft besser sieht:

$$F(x_0, \Delta x) = \sin(x_0 + \Delta x) - \sin(x_0) = \dots$$

□ Was bedeutet $\lim_{\substack{\Delta x \rightarrow 0 \\ \Delta x \neq 0}} F(x_0, \Delta x) = 0$ nach (3.2.5) explizit?

(4.2.10) Ein gutes Verständnis der ε - δ -Definition der Stetigkeit ist wichtig. Wir entwickeln hierfür zunächst eine **geometrische Interpretation:**

- – Wir definieren zu gegebenem f und x_0 in der $\varepsilon - \Delta x$ -Halbebene

$$H = \{(\Delta x, \varepsilon) | (\Delta x + x_0) \in D, \varepsilon > 0\}$$

die folgenden beiden Mengen

$$\begin{aligned} Z &= \{(\Delta x, \varepsilon) \mid \varepsilon > 0, \Delta x = x - x_0 \in \mathbb{R}, x \in D, |f(x_0 + \Delta x) - f(x_0)| < \varepsilon\} \\ V &= \{(\Delta x, \varepsilon) \mid \varepsilon > 0, \Delta x = x - x_0 \in \mathbb{R}, x \in D, |f(x_0 + \Delta x) - f(x_0)| \geq \varepsilon\}. \end{aligned}$$

V ist die Menge der *verbotenen Punkte*, in denen die Ungleichung verletzt ist, und Z die der *zulässigen Punkte*, in denen die Ungleichung erfüllt ist, in denen sich die beiden Werte also um weniger als ε unterscheiden. In den nachfolgenden Bildern ist der verbotene Bereich V dunkel gezeichnet. Der helle Bereich ist Z. Die ε -Halbachse gehört auch zu Z.

– Z hat offensichtlich die Eigenschaft: $\varepsilon_1 > \varepsilon$ und $(\Delta x, \varepsilon) \in Z \Rightarrow (\Delta x, \varepsilon_1) \in Z$. Die Grenze zwischen den beiden Bereichen wird gegeben durch

$$\Delta x \mapsto \varepsilon = \gamma(\Delta x) = |f(x_0 + \Delta x) - f(x_0)|$$

Der Graph dieser *Grenzfunktion* trennt die beiden Bereiche Z und V.

□ Bestimmen Sie $Z=Z(x_0)$ und $V=V(x_0)$ für $f(x)=x^2$ durch Angabe der Grenzfunktion.

- Die Stetigkeitsbedingung verlangt, dass zu jedem $\varepsilon > 0$ ein kleines Δx -Intervall um $\Delta x = 0$ existiert, das ganz in Z liegt. Oder auch: Für alle Δx mit $-\delta < \Delta x < \delta$ muss $(\Delta x, \varepsilon) \in Z$ gelten. Der Graph einer eventuellen Strategiefunktion $\varepsilon \mapsto \delta(\varepsilon, x_0)$ muss demnach für jedes $\varepsilon > 0$ die folgende Bedingung erfüllen

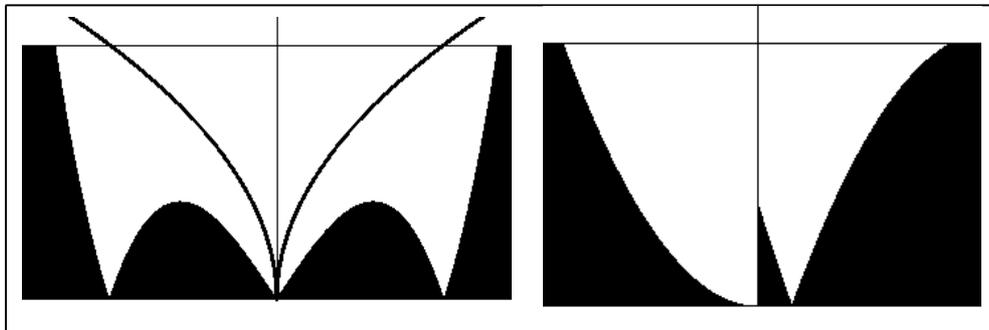
$$\text{Für alle } \Delta x \text{ mit } -\delta(\varepsilon, x_0) < \Delta x < \delta(\varepsilon, x_0) \text{ ist } (\Delta x, \varepsilon) \in Z.$$

★ (4.2.11) Wir beschreiben das Konzept eines Computerprogramms zur Veranschaulichung des Sachverhaltes in der $\Delta x - \varepsilon$ -Ebene.

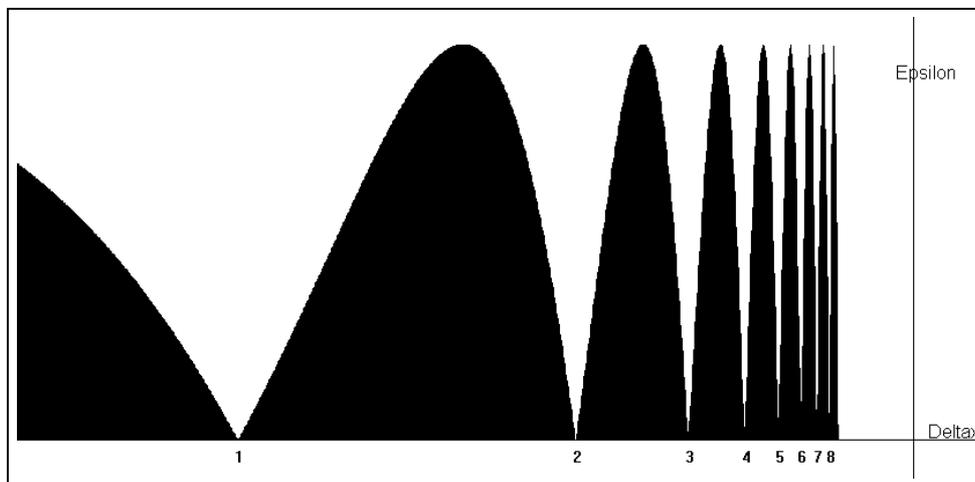
- Gebe einen Rechenausdruck für f vor und ebenso den Wert x_0 .
- Der Computer soll alle Punktepaare $(\Delta x, \varepsilon)$ mit $0 < \varepsilon < 1$ zeichnen, die die gewünschte Ungleichung $|f(x_0 + \Delta x) - f(x_0)| < \varepsilon$ verletzen. Oder: Er schwärzt die Punkte der Menge V.
- Dazu geht das Programm alle Punkte $(\Delta x, \varepsilon)$ des Bildschirmbereiches durch, berechnet jeweils $\varepsilon - \gamma(\Delta x)$ und schwärzt den Punkt, sofern dieser Wert positiv ist.

Die geometrische Form der komplementären Figur Z der zulässigen Punkte entscheidet, ob Stetigkeit in x_0 vorliegt oder nicht: Im linken Bild der Figur ist das der Fall. Dabei handelt es sich um die Funktion $y=x(1-x^2)$ und $x_0 = 0$. Also $f(0)=0$. Für $x = \pm 1$ ist $|f(x)-0|=0 < \varepsilon$, d.h. Z reicht bis zur Δx -Achse herunter. Zusätzlich ist hier noch der Graph einer Strategiefunktion eingezeichnet. Zwischen deren beiden Ästen darf nie ein Punkt aus V liegen und die Graphenpunkte selbst dürfen nie auf der ε -Achse liegen.

Im rechten Bild reicht der verbotene Bereich bis an die ε -Achse heran. Man kann keine zulässige Funktion $\delta = \delta(\varepsilon)$ finden. Es handelt sich hier um die Funktion $f(x) = \begin{cases} x^2 & \text{für } x \leq 0 \\ x^2 + 0.2(1 - 6x) & \text{für } x > 0 \end{cases}$ an der Stelle $x_0 = 0$, die in diesem Punkt unstetig ist. Die Nullstelle in der Nähe von $x = \frac{1}{6}$ ist sichtbar, nutzt aber nichts.



Jetzt noch das Beispiel $f(x) = \sin \frac{1}{x}$ für $x \neq 0$ und $f(0) = 0$. Diese Funktion hat Nullstellen bei $x_n = \frac{1}{n\pi}$ für alle $n \in \mathbb{Z}$ und $n \neq 0$. Zwischen je zwei dieser Nullstellen liegt ein Maximum bzw. Minimum mit Funktionswert 1. In der Figur sind die Nullstellen für $n = -1, -2, \dots, -8$ angedeutet, für die der Z -Bereich jeweils bis zur Δx -Achse reicht. In dem nicht gezeichneten Bereich bis zur ε -Achse gibt es dann unendlich viele weitere immer schmäler werdende V -Bereiche, die sich der ε -Achse beliebig nähern. Sie sind nicht gezeichnet, da sie das Auflösungsvermögen des Bildschirms überfordern. Das zeigt sich bereits daran, dass etwa ab $n=3$ die Menge Z in der Zeichnung nicht mehr bis zur Δx -Achse herunter zu reichen scheint. Es ist offensichtlich, dass man für diese Funktion keine Strategiefunktion findet. Wie klein man $\delta(\varepsilon)$ für $0 < \varepsilon < 1$ auch wählt, es gibt immer noch V -Bereiche, die zwischen (δ, ε) und der ε -Achse liegen.



(4.2.12) Der übliche explizite **Nachweis der Stetigkeit** einer Funktion erfolgt durch Angabe einer Strategiefunktion. Der allereinfachsten Fall liegt vor, wenn die Grenzfunktion $\varepsilon = \gamma(\Delta x)$ für $\Delta x > 0$ und für $\Delta x < 0$ jeweils für $0 < \varepsilon < \varepsilon_0$ invertierbar ist. Ist $\varepsilon \mapsto \Delta(\varepsilon)$ das Minimum dieser inversen Funktionen, dann kann $\varepsilon \mapsto \delta(\varepsilon) = \alpha \Delta(\varepsilon)$ mit $0 < \alpha < 1$ als Strategiefunktion genommen werden.

Ist γ nicht invertierbar, dann geht man in der Regel nach folgendem Schema vor:

$$|f(x_0 + \Delta x) - f(x_0)| \stackrel{(v)}{\leq} g(x_0, \Delta x) \stackrel{(m)}{\leq} h(x_0, \delta) \stackrel{(r)}{<} \varepsilon.$$

- Zunächst wird $|f(x_0 + \Delta x) - f(x_0)|$ vorsichtig vergrößert, D.h. man geht in der Menge Z vertikal nach oben. Das ist Schritt (v).
- Dann sichert man - etwa über Monotonie des durch Vergrößerung entstandenen Ausdrucks, dass die Ungleichung für alle Δx mit $|\Delta x| \leq \delta$ gilt. D.h. man testet, ob sich in horizontaler Richtung zur ε -Achse hin keine verbotenen Punkte befinden. Schritt (m).
- Und schließlich versucht man die verbleibende Bestimmungsungleichung in δ zu lösen. Das ist Schritt (r). Jede Lösung (der Ungleichung) ist eine Strategiefunktion.
- Zusätzlich genügt es, die Abschätzungen auf $\varepsilon \leq \varepsilon_0$ und $\delta \leq \delta_0$ für zwei fallspezifisch geeignete Konstanten $\varepsilon_0, \delta_0 > 0$ zu beschränken. Es genügt, die Lösbarkeit in diesem Bereich nachzuweisen. Denn man kann $\delta(\varepsilon) = \delta(\varepsilon_0)$ für $\varepsilon > \varepsilon_0$ wählen.

Das Vorgehen ist weitgehend analog zum Vorgehen zur Untersuchung der Konvergenz von Folgen.

(4.2.13) Es folgen **Beispiele. Jeweils mit einer Abschätzung der beschriebenen Art mit anschließender kurzer Diskussion der Schritte.**

▼(4.2.14) Beispiel: $f(x) = x^2$ in $x_0 = 0$:

$$|f(x) - f(0)| \stackrel{(1)}{=} |f(x)| \stackrel{(2)}{=} x^2 \stackrel{(3)}{\leq} \delta^2 \stackrel{(4)}{=} \delta^2(\varepsilon) \stackrel{(5)}{<} \varepsilon$$

(1) $f(0) = 0$; (2) $f(x) = x^2$ einsetzen; (3) = (m) radiale Monotonie von $x \mapsto x^2$; (5) Die Bedingung wird erfüllt durch $\delta(\varepsilon) = \sqrt{\frac{\varepsilon}{2}}$.

▼(4.2.15) Ein Beispiel einer **Unstetigkeitsstelle** (in $x_0 = 0$)

$$f(x) = \begin{cases} x^2 + c & \text{für } x \geq 0 \\ 0 & \text{für } x < 0 \end{cases} \quad \begin{array}{l} c=0 \text{ stetig in } x_0 = 0, \text{ sogar differenzierbar (} f(0)=0 \text{)} \\ c \neq 0 : \text{ unstetig in } x_0 = 0. \quad (f(0)=\boxed{?..}) \end{array}$$

$$\boxed{c \neq 0:} \quad |f(x) - f(0)| \stackrel{!}{=} |f(x) - 0| = \begin{cases} x^2 & x \geq 0 \\ |c| & x < 0 \end{cases}$$

Wähle $\varepsilon = \frac{|c|}{2}$, Dann gilt für **jedes** $\delta > 0$ und etwa $x_1 = -\frac{\delta}{2}$

$$|f(x_1) - f(0)| \stackrel{(1)}{=} ||c| - 0| > \varepsilon = \frac{|c|}{2}$$

Also kann es keine Strategiefunktion geben.

□ Die zugehörige **Verneinungsübung**: Vervollständigen Sie die Stetigkeitsaussage und darunter ihre Verneinung. Diese Verneinung wurde im Beispiel benutzt.

Für jedes $\varepsilon > 0$ gibt es ein $\delta > 0$ derart dass ...für alle zulässigen x $< \varepsilon$

Es gibt ein $\varepsilon > 0$, so dass für alle $\delta > 0$ es ein zulässiges x gibt mit $\geq \varepsilon$

▼(4.2.16) Beispiel: $\boxed{f(x) = \text{sqr}(|x|)}$ für $x \geq 0$ und o.B.d.A. $x > y \geq 0$

$$|\sqrt{x} - \sqrt{y}| \stackrel{(1)}{=} \frac{x - y}{\sqrt{x} + \sqrt{y}} \stackrel{(2)}{\leq} \frac{x}{\sqrt{x}} \stackrel{(3)}{=} \sqrt{x} \stackrel{(4)}{\leq} \sqrt{\delta} < \varepsilon$$

$$\boxed{\delta(\varepsilon) = \frac{1}{2}\varepsilon^2}$$

(1) Umformen. (2)=(v) Vergrößern. (3) Umformen. (4)=(m) Monotonie der Wurzelfunktion.

Wir haben gleichmäßige Stetigkeit, da das angegebene δ nicht von x_0 abhängt.

▼(4.2.17) Beispiel: $\boxed{f =]0, \infty[, x \mapsto \frac{1}{x}, \mathbb{R})}$

$$\left| \frac{1}{x} - \frac{1}{x_0} \right| \stackrel{(1)}{=} \frac{|x - x_0|}{xx_0} \stackrel{(2)}{\leq} \frac{2|x - x_0|}{x_0^2} \stackrel{(3)}{\leq} \frac{2\delta}{x_0^2} \stackrel{(4)}{<} \varepsilon$$

Für (2) $\delta \leq \frac{1}{2}x_0$ verlangen. Dann ist $x \geq x_0 - \delta \geq \frac{1}{2}x_0$

$$\delta(\varepsilon, x_0) = \min\left(\frac{1}{2}x_0, \frac{1}{4}x_0^2\varepsilon\right) \quad \text{erfüllt (2) und (4)}$$

Wann ist die Bedingung verletzt?

$$\begin{aligned} \frac{|x - x_0|}{xx_0} &\geq \varepsilon \\ |x - x_0| &\geq \varepsilon xx_0 \\ x - x_0 &\geq \varepsilon(x - x_0 + x_0)x_0 && \text{für } x > x_0 \\ (1 - \varepsilon)(x - x_0) &\geq \varepsilon x_0^2 \\ (x - x_0) &\geq \frac{\varepsilon}{1 - \varepsilon} x_0^2 && \text{(gleichw.)} \end{aligned}$$

Wählt man jetzt irgendein $\delta = \delta(\varepsilon)$ unabhängig von x_0 , so kann man zu festem $\varepsilon > 0$ immer $x_0 > 0$ so klein wählen, dass $\delta(\varepsilon) > 2\frac{\varepsilon}{1-\varepsilon}x_0^2$ wird. Dann ist die gewünschte Ungleichung für $0 < x - x_0 < \delta(\varepsilon)$ sicher verletzt. Oder: **Für zulässiges δ muss gelten**

$$\frac{\varepsilon}{1 - \varepsilon} x_0^2 > \delta(\varepsilon, x_0) > 0$$

und das geht mit x_0 nach Null. Die Ungleichung kann nicht unabhängig von x_0 erfüllt werden.

Die Funktion ist (überall) stetig, aber nicht gleichmäßig stetig.

▼(4.2.18) Beispiel $f(x)=x^2$ im Punkte x_0

$$|x^2 - x_0^2| \stackrel{(1)}{=} |x + x_0||x - x_0| \stackrel{(2)}{\leq} C(x_0)|x - x_0| \stackrel{(3)}{\leq} C(x_0)\delta \stackrel{4}{<} \varepsilon$$

$$C(x_0) \stackrel{0}{=} \max(1, 3x_0) \quad \text{und} \quad \delta(\varepsilon) = \frac{\varepsilon}{2C(x_0)}$$

Verletzungsbereich: (Beachte $|a+b| \geq ||a| - |b||$)

$$|x^2 - x_0^2| \stackrel{0}{=} |x + x_0||x - x_0| = |x - x_0 + 2x_0||x - x_0|$$

$$= \Delta x^2 + 2x_0\Delta x > \varepsilon \quad \text{für } x > x_0 > 0$$

$$\Delta x_{12} = -x_0 \pm \sqrt{x_0^2 - \varepsilon} = x_0 \left(-1 \pm \sqrt{1 - \frac{\varepsilon}{x_0^2}} \right)$$

$$\Delta x_1 \approx -\frac{\varepsilon}{2x_0} \quad \Delta x_2 \approx -2x_0 + \frac{\varepsilon}{2x_0}$$

Die durch Δx_1 gegebene Grenze rückt mit wachsendem x_0 immer näher an Null heran. Eine von x_0 unabhängige Wahl von δ ist erneut nicht möglich.

▼(4.2.19) Beispiel für Stetigkeit in einem Randpunkt: $f = ([0,1], x \mapsto \sqrt{x(1-x)}, \mathbb{R})$ und $x_0 = 0$

$$|f(\Delta x) - f(0)| = \sqrt{\Delta x(1-\Delta x)} \leq \sqrt{\Delta x} \leq \sqrt{\delta} < \varepsilon$$

$$\text{Aber nur für } 0 \leq \Delta x \leq 1! \quad \delta(\varepsilon) = \frac{1}{2}\varepsilon.$$

- Wir definieren eine Funktion D wie folgt: $D(x) = \begin{cases} 1 & \text{für } x \text{ irrational} \\ 0 & \text{für } x \text{ rational, } x \neq 0 \\ \frac{1}{2} & \text{für } x=0 \end{cases}$
- Zeigen Sie, dass diese Funktion in jedem Punkt unstetig ist, indem Sie die zugehörige Menge V genauer beschreiben. Wie sieht V insbesondere für $x_0 = 0$ aus?
- Bisher haben wir Funktionen $\mathbb{R} \rightarrow \mathbb{R}$ betrachtet. Wie steht es mit Funktionen $\mathbb{R} \rightarrow \mathbb{C}$ oder $\mathbb{C} \rightarrow \mathbb{C}$. Lässt sich der Stetigkeitsbegriff darauf in naheliegender Weise verallgemeinern?

2.4.2a Alternative Charakterisierungen von stetig

(4.2.20) Wir geben zwei **alternative Charakterisierungen** der Stetigkeitseigenschaft. Genauer: Ist f stetig in x_0 , dann sind die folgenden beiden Bedingungen erfüllt. Ist umgekehrt eine dieser beiden Bedingungen erfüllt, dann ist f stetig in x_0 .

(4.2.21) **Umgebungsformulierung der Stetigkeit:**

Sei $f: D \rightarrow \mathbb{R}$, $x_0 \mapsto y_0 = f(x_0)$, $U_\varepsilon(y_0)$ irgendeine ε -Umgebung von y_0 .
Dann enthält $f^{-1}(U_\varepsilon(y_0))$ eine Umgebung von x_0 in D, also ein $U_\delta^D(x_0)$ mit $\delta > 0$.

Diese Definition eignet sich gut zur Verallgemeinerung auf Fälle, in denen man die Punktumgebungen nicht durch Zahlen charakterisieren kann.

- Beweisen Sie die Gleichwertigkeit zur Stetigkeitsdefinition.

(4.2.22) **Folgenstetigkeit** oder Vorhersagbarkeit des Funktionswertes mit Hilfe von Folgen:

Sei $u = (\mathbb{N}, n \mapsto u_n, D)$ mit $\lim_{n \rightarrow \infty} u_n = x_0$.
Dann folgt $\lim_{n \rightarrow \infty} f(u_n) = f(x_0)$.

Folgenstetigkeit liefert in vielen Fällen eine günstige Explikation der Stetigkeitsaussage.

Beweis: Stetigkeit ist gleichwertig zur Folgenstetigkeit

i) Sei f stetig in $x_0 \in D$ und $\delta = \delta(\varepsilon)$ zugehörige Strategiefunktion. Weiter $a_n \in D$ mit $a_n \rightarrow x_0$. Oder $\lim_{n \rightarrow \infty} a_n = x_0$. Weiter sei $N(\varepsilon)$ eine zur Folge gehörige Strategiefunktion. Wähle $\varepsilon > 0$ und bestimme $\delta(\varepsilon)$. Möglich, da f in x_0 stetig. Bilde damit $N(\delta(\varepsilon))$ und $n > N(\delta(\varepsilon))$ also $|a_n - x_0| < \delta(\varepsilon)$. Es folgt $|f(a_n) - f(x_0)| < \varepsilon$. Also $\lim_{n \rightarrow \infty} f(a_n) = f(x_0)$.

ii) Sei f in x_0 folgenstetig. Indirekter Beweis. Angenommen f ist in x_0 nicht stetig. D.h.: Es gibt ein $\varepsilon > 0$, so dass es für **jedes** $\delta > 0$ ein $x \in D$ gibt mit $|f(x) - f(x_0)| \geq \varepsilon$ (*), obwohl $|x - x_0| < \delta$. Wähle nun für $\delta_n = \frac{1}{n}$ ein $x = a_n$ mit dieser Eigenschaft. Da $|a_n - x_0| < \frac{1}{n}$, ist das eine konvergente Folge mit Grenzwert x_0 . Wegen der Folgenstetigkeit folgt $\lim_{n \rightarrow \infty} f(a_n) = f(x_0)$. Das widerspricht (*), das für ausreichend hohes n erfüllt sein müßte.

2.4.2b Elementare Konsequenzen der Stetigkeit

(4.2.23) Die elementaren Grenzwertsätze sind von der Form: "Seien f, g, \dots in \dots stetig. Dann ist auch \dots stetig in \dots ". Wie üblich kann man mit ihrer Hilfe aus stetigen Funktionen rekursiv weitere stetige Funktionen konstruieren. Und die Sätze sind tendenziell derart, dass die Mengen stetiger Funktionen bestimmte algebraische Strukturen erhalten. Etwa: Die Menge aller in $x_0 \in D$ stetigen Funktionen $D \rightarrow \mathbb{R}$ bildet einen Vektorraum.

(4.2.24) Wir geben ein **Beispiel** eines solchen Satzes (**mit Beweis**):

Seien f und g stetig in $x_0 \in D$. Dann ist $\alpha f + \beta g$ in x_0 stetig.

Beweis über Angabe einer Strategiefunktion: δ_f und δ_g seien Strategiefunktionen für f und g in x_0 . Solche gibt es laut Voraussetzung. Damit folgt:

$$\begin{aligned} & |(\alpha f + \beta g)(x_0 + \Delta x) - (\alpha f + \beta g)(x_0)| \\ &= |\alpha(f(x_0 + \Delta x) - f(x_0)) + \beta(g(x_0 + \Delta x) - g(x_0))| \\ &\leq |\alpha| |f(x_0 + \Delta x) - f(x_0)| + |\beta| |g(x_0 + \Delta x) - g(x_0)| \\ &< |\alpha| \left(\frac{1}{2|\alpha|} \varepsilon \right) + |\beta| \left(\frac{1}{2|\beta|} \varepsilon \right) = \varepsilon \end{aligned}$$

Für die Strategiefunktion $\delta = \delta_{\alpha f + \beta g}(\varepsilon) = \min \left(\delta_f \left(\frac{1}{2|\alpha|} \varepsilon \right), \delta_g \left(\frac{1}{2|\beta|} \varepsilon \right) \right)$.

(4.2.25) Die Aussage des Satzes kann man wie folgt zusammenfassen: **Die in $x_0 \in D$ stetigen Funktionen $D \rightarrow \mathbb{R}$ bilden einen Vektorraum.** Inspektion des Beweises zeigt weiter: Sind f und g in D stetig bzw. gleichmäßig stetig, dann hat $\alpha f + \beta g$ dieselbe Eigenschaft.

- Was ist mit $f \cdot g$? Bilden die in $x_0 \in D$ stetigen Funktionen einen Ring?
- Wie steht es mit den analogen Aussagen für (global) stetig und für gleichmäßig stetig?
- Wie steht es mit $\frac{f}{g}$ und $g(x_0) \neq 0$?
- $g \circ f$ und f^{-1} ?

(4.2.26) Alle diese Sätze sind elementar und man sollte sich nicht ausführlich, sondern nur stichwortartig auf sie beziehen. Etwa: "Das Produkt stetiger Funktionen ist erneut stetig".

Betrachten wir dazu ein Beispiel. $h_0 : x \mapsto 1$ und $h_1 : x \mapsto x$ sind stetig mit $\delta(\varepsilon) = 1$ und $\delta(\varepsilon) = \frac{1}{2}\varepsilon$. Beide sind daher auch gleichmäßig stetig. Dann folgt über Produktbildung die Stetigkeit von $h_2 = h_1 \cdot h_1$ und allgemein h_n . Durch Linearkombinationsbildung folgt anschließend die Stetigkeit jedes Polynoms in jedem Punkt. Aber bereits h_2 ist nicht mehr auf ganz \mathbb{R} gleichmäßig stetig! Haben Sie das bei der Ausführung von obigem Beweis bemerkt? Arbeiten Sie den Grund heraus.

□ : $x \mapsto x$ und $x \mapsto |x| = \sqrt{x^2}$ sind beide stetig. Aber wie steht es mit $x \mapsto \frac{x}{|x|}$? Das Beispiel zeigt, dass manchmal auch im Bereich elementarer Sätze Vorsicht geboten ist.

2.4.2c Das Computerprogramm zur Epsilontik.

```

PROCEDURE ein_epsdelt
  a = -2
  E = 2
  O = 1.5
  u = -0.5
  d = E - a
  h = O - u
  dx = 3 / _X
  dy = 3 / _Y
RETURN
PROCEDURE vor_epsdelt
  drawp(a,0,E,0)
  drawp(0,0,0,O)
  drawp(a,1,E,1)
  DEFFN f(x) = x ^2 - (x > 0) * P
  DEFFN g(y) = Q * y ^2
RETURN
PROCEDURE kreis
  ein_epsdelt
  vor_epsdelt
  FOR x = a TO E STEP dx
  FOR y = 0 TO 1 STEP dy
    GOSUB skalar(x,y)
    IF s < 0.5
      drawpm(x,y)
    ENDIF
  NEXT y
NEXT x
COLOR 0
RETURN
PROCEDURE skalar(x,y)
  s = (-FN f(x) + FN f(0) + y > 0) * (y + FN f(x) - FN f(0) > 0 )
RETURN
PROCEDURE lauf2
  DEFLINE 0,3
  FOR yy = 0 TO O STEP dy / 2
    xx = FN g(yy)
    drawpm(xx,yy)
    drawpm(-xx,yy)
  NEXT yy
COLOR 0
DEFLINE 0,1
RETURN

```

2.4.3 Einige topologische Grundbegriffe

(4.3.1) Wir haben den Begriff der ε -Umgebung eingeführt. Mit dessen Hilfe lassen sich in \mathbb{R} einige Typen von Teilmengen aussondern, die sich für den Ausbau der Analysis als ausgesprochen wichtig erweisen. Genauer benötigt man sie immer zur Formulierung geometrischer Aspekte der Analysis. Wir behandeln davon die offenen, die abgeschlossenen und die kompakten Mengen.

(4.3.2) Man kann diese Mengen teilweise auch über ihre Punkte charakterisieren. Diese Punkte stehen dann in einer ganz bestimmten Beziehung zur Ausgangsteilmenge, einer Beziehung, die meist mit Hilfe des Umgebungsbegriffes ausgedrückt wird. Anschließend bildet man dann die Menge aller Punkte eines bestimmten Typs. Ist $T \subset \mathbb{R}$ Teilmenge, dann gibt es z.B. *innere Punkte von T*, *Randpunkte von T*, *Berührungspunkte* und *Häufungspunkte von T*.

- Sei $T = \{-1\} \cup]0,1[$. Was werden die zugehörigen *Randpunkte von T*, was die *inneren Punkte von T* sein. Was für einen Unterschied wird es zwischen *Berührungspunkt* und *Häufungspunkt* geben. Vergleichen Sie die drei Punkte -1, 0 und 1 von T . Versuchen Sie selbst eine allgemeine und verallgemeinerungsfähige Definition dieser Punkte zu entwickeln. Versuchen Sie auch, sich über Beispiele zu verdeutlichen, dass dies alles unterschiedliche Typen von Punkten sind. Testen Sie, ob Ihre Definitionen auch auf Funktionen $\mathbb{C} \rightarrow \mathbb{C}$ anwendbar sind. Die genauen Definitionen geben wir später. (Achtung: Das ist eine wichtige Frage, mit der Sie sich beschäftigen sollten, ohne gleich die späteren Definitionen nachzuschlagen. Einmal, weil das definitorische Festlegen von Begriffen eine wichtige und nicht leichte mathematische Arbeit ist, die man auch üben sollte und zum anderen, weil sich vorab gemachte Überlegungen beim späteren Verstehen und merken dieser "abstrakten" Definitionen auszahlen dürften.)

(4.3.3) Die Gültigkeit vieler allgemeiner Sätze ist auf einen bestimmten Typ derartiger Teilmengen beschränkt.

(4.3.4) **Definition:** Sei $T \subset \mathbb{R}$ eine Teilmenge. T heißt *offene Teilmenge von \mathbb{R}* , wenn es zu jedem $x \in T$ eine ε -Umgebung $U_\varepsilon(x)$ gibt, die ganz in T liegt. Also $U_\varepsilon(x) \subset T$ oder auch $U_\varepsilon^T(x) = U_\varepsilon(x)$.

Insbesondere ist ganz \mathbb{R} offen und auch die leere Menge ist offen. Da sie keinen Punkt enthält, ist jede Bedingung an ihre Punkte automatisch erfüllt.

(4.3.5) **Definition:** Sei $T \subset \mathbb{R}$. T heißt *abgeschlossene Teilmenge*, wenn ihr Komplement $\mathbb{R} - T$ offen ist.

Da leere Menge und ganz \mathbb{R} Komplemente von einander sind, sind beide Mengen abgeschlossen. Diese beiden Mengen sind daher sowohl offen als auch abgeschlossen. Gültig ist hier die exakte mathematische Definition, nicht aber die vage umgangssprachliche Bedeutung, der dieser Sachverhalt natürlich widerstreben könnte.

- Beweisen Sie kurz: Ein offenes Intervall ist offen ein abgeschlossenes abgeschlossen.

(4.3.6) Natürlich sind die meisten Teilmengen von \mathbb{R} weder offen noch abgeschlossen. Ein ausgeprägtes Beispiel hierfür gibt die Teilmenge der rationalen Zahlen.

Definition: Eine Teilmenge $T \subset \mathbb{R}$ heißt *beschränkt*, wenn sie sowohl eine obere wie eine untere Schranke besitzt.

(4.3.7) (Vorläufige) **Definition:** Eine Teilmenge $T \subset \mathbb{R}$ heißt *kompakte Teilmenge* von \mathbb{R} , wenn sie abgeschlossen und beschränkt ist.

Später werden wir eine andere für \mathbb{R} gleichwertige Definition geben, die sich besser verallgemeinern läßt.

(4.3.8) Sei $T \subset \mathbb{R}$ eine Teilmenge. Ein Punkt $a \in T$ heißt *innerer Punkt von T*, wenn es eine ε -Umgebung von a (aus \mathbb{R}) gibt, die ganz in T liegt. eine offene Menge besteht dann aus lauter inneren Punkten. Die beiden Randpunkte 0 und 1 des abgeschlossenen Intervalles $[0,1]$ sind keine inneren Punkte dieses Intervalles! Ist $U_\varepsilon(a)$ eine derartige ε -Umgebung eines inneren Punktes $a \in D$ eines Definitionsbereiches, dann kann man für alle Punkte dieser Umgebung den Funktionswert bilden.

(4.3.9) Wodurch sind die Punkte *abgeschlossener* Mengen ausgezeichnet? Sei $A \subset \mathbb{R}$ abgeschlossen, d.h. $\mathbb{R} - A$ ist offen. Weiter sei $(\mathbb{N}, n \mapsto x_n, A)$ eine konvergente Folge aus A mit Grenzwert $\lim_n x_n = X$. Liegt dann X erneut in A ? Allgemein ist das sicher nicht der Fall. Etwa $(\mathbb{N}, n \mapsto \frac{1}{n+1},]0,1[)$ mit Grenzwert 0. Aber 0 liegt nicht im Intervall $]0,1[$ und idese Intervall ist auch nicht abgeschlossen.

Also nochmals etwas sorgfältiger: **Eine (in \mathbb{R}) abgeschlossene Menge enthält auch alle ihre Grenzwerte in \mathbb{R} .** Man gelangt auch vermittels einer Grenzwertbildung nicht aus ihr heraus. Daher das Attribut *abgeschlossen*.

Zum Beweis dieser Behauptung. Sei A abgeschlossen, $(\mathbb{N}, n \mapsto x_n, A)$ konvergente Folge mit $\lim_n x_n = X$. Wir führen einen indirekten Beweis, nehmen also an, dass $X \notin A$. Dann liegt X im offenen Komplement von A . Und das heißt, dass es eine Umgebung $U_\varepsilon(X)$ gibt, die auch ganz im Komplement liegt. Andererseits soll x_n gegen X konvergieren. Also muss es n geben, mit $|x_n - X| < \varepsilon$. Widerspruch: Alle x_n liegen in A , also auch dieses. Andererseits liegt es in $U_\varepsilon(X)$, also nicht in A .

□ Übertragen Sie diese Überlegungen von \mathbb{R} nach \mathbb{C} .

□ Sie möchten die Überlegungen auf \mathbb{Q} übertragen. Welche Schwierigkeiten entstehen? Betrachten Sie insbesondere "A ist abgeschlossene Teilmenge von \mathbb{Q} ".

(4.3.10) Sei jetzt K eine kompakte Teilmenge von \mathbb{R} und $a: \mathbb{N} \rightarrow K$ eine Folge mit Werten in K . Da K beschränkt ist, ist a beschränkt und enthält nach dem Satz (2.6.8) von Bolzano und Weierstraß eine konvergente Teilfolge. Da K auch abgeschlossen ist, liegt auch der Grenzwert G in K . Also in K hat jede Folge eine konvergente Teilfolge, deren Grenzwert erneut in K liegt. Ist also K Definitionsbereich einer Abbildung, dann ist $f(G) = f(\lim_n a_n)$ sicher bildbar.

(4.3.11) **Zusammenfassung:** Wir haben offene, abgeschlossene und kompakte Teilmengen von \mathbb{R} eingeführt. Die offenen Mengen bestehen aus lauter inneren Punkten. Die abgeschlossenen enthalten auch alle ihre Grenzwerte in \mathbb{R} .

2.4.4 Allgemeine Konsequenzen der Stetigkeit

(4.4.1) Stetigkeit hat eine Reihe von Konsequenzen. Das sind einmal die bereits behandelten Grenzwertsätze, die aber alle mehr formaler Natur sind und die für sich alleine die Bedeutung des Stetigkeitsbegriffes nicht begründen können. Dazu gibt es jedoch eine Reihe sehr allgemeiner Sätze, die den eigentlichen Wert der Stetigkeit ausmachen. Folgerungen dieser Sätze werden für den praktischen Umgang mit Funktionen immer wieder benötigt. Und vieles, was man im Umgang mit Funktionen zu tun gewohnt ist, wird erst durch diese Sätze gerechtfertigt. Schließlich ist auch die leichte Verallgemeinerbarkeit des Stetigkeitsbegriffes und vieler der allgemeinen Sätze wichtig.

Sachlich geht es bei diesen Sätzen um die Vorhersagbarkeit, die Annahme von Zwischenwerten und Extremwerten und Grenzwerten und um die Vollständigkeit.

(4.4.2) Wir geben jetzt die wichtigsten dieser Sätze. Und es sei betont: Keines dieser Resultate ist selbstverständlich! Fehlt eine zugehörige Voraussetzung wie *stetig*, dann wird es Gegenbeispiele geben.

■ (4.4.3) **Umgebungssatz:**

Es sei f in x_0 stetig mit $f(x_0) > 0$.
Dann gibt es eine kleine Umgebung $U_\delta(x_0)$ in der $f(x) > 0$ gilt.

Beweis: Sei $\delta = \delta(\varepsilon)$ eine Strategiefunktion. Setze $f(x_0) = 2\varepsilon$. Dann gilt

$$f(x) = f(x_0) + (f(x) - f(x_0))$$

Wählt man $|x - x_0| < \delta(\varepsilon)$, dann heißt das, $f(x) = 2\varepsilon + \eta$ mit $|\eta| < \varepsilon$, also $f(x) > \varepsilon$.

■ (4.4.4) **Zwischenwertsatz:**

Sei $f: [a, b] \rightarrow \mathbb{R}$ stetig mit $a < b$ und $A = f(a)$ und $B = f(b)$.
Dann wird jeder Zwischenwert C zwischen A und B angenommen.
D.h. es gibt x_0 mit $a < x_0 < b$, so dass $f(x_0) = C$ gilt.

Beweis: OBdA sei $B \geq A$. (Sonst $-f$ betrachten!). Also $a < b$ und $A \leq B$. Ist $A = B$, so ist $C = A = B$. Der Zwischenwert ist eventuell ein Randpunkt. Sei jetzt $A < B$ und $A < C < B$. Betrachte die Menge

$$M = \{x | a \leq x \leq b, f(x) \leq C\}$$

Diese Menge ist nach oben durch b beschränkt und besitzt folglich ein Supremum c . Für $x > c$ ist notwendig $f(x) > C$.

Da c Supremum, gibt es eine Folge $(\mathbb{N}, n \mapsto x_n, M)$, die gegen c konvergiert. Wegen der Stetigkeit von f folgt $f(x_n) = f(c)$. Wegen $x_n \in M$ folgt $f(x_n) \leq C$ und damit wie üblich über einen Grenzwertsatz $f(c) \leq C < B$. Daher gilt auch $c < b$. Wähle nun eine Folge $n \mapsto y_n$ mit $c < y_n < b$, die auch gegen c konvergiert. Wir haben $f(y_n) > C$, da $y_n \notin M$. Wegen der Stetigkeit folgt $\lim f(y_n) = f(c)$. Und nach den Grenzwertsätzen ist $f(c) \geq C$.

Mit dem ersten Resultat gibt das $f(c)=C$ wie gewünscht. c ist dann **der größte zwischen a und b liegende x -Wert**, in dem f den Wert C annimmt.

Der Satz wird gerne benutzt, um die Existenz von Nullstellen zu sichern. Das bekannteste Beispiel hierfür ist die Aussage, dass **jedes Polynom mit einer ungeraden Ordnung mindestens eine reelle Nullstelle besitzt**.

□ Vervollständigen Sie die Begründung der letzten Aussage.

Der Satz sagt aus, dass es mindestens ein derartiges x_0 gibt. Aber natürlich kann es auch mehr als einen solchen Punkt geben. Wieviel mehr es sind, darüber macht der Satz keine Aussage. So kann ein Polynom der Ordnung $2k+1$ bis zu $2k+1$ verschiedene reelle Nullstellen haben.

■ (4.4.5) **Wertvorhersage**. Hierzu benötigen wir zunächst eine **Definition**:

Def: *Häufungspunkt einer Menge.*

Sei $D \subset \mathbb{R}$. und $x_0 \in \mathbb{R}$. x_0 heißt Häufungspunkt der Menge D ,

wenn **jede** Umgebung $U_\varepsilon^D(x_0) = U_\varepsilon(x_0) \cap D$ unendlich viele Elemente enthält.

Beachten Sie: Es wird $x_0 \in \mathbb{R}$, nicht aber $x_0 \in D$ gefordert. Eventuelle Häufungspunkte von D können, müssen aber nicht in D selbst liegen. Ist $D = \{-1\} \cup]0, 1]$, dann ist $-1 \in D$ kein Häufungspunkt, $0 \notin D$ dagegen ist einer. Das entspricht der semantischen Bedeutung von *häufen*. Damit gilt nun zum Stichwort *Wertvorhersage*:

Ist	$f: D \rightarrow \mathbb{R}$ in x_0 stetig und x_0 Häufungspunkt von D ,
dann	kann man den Wert $f(x_0)$ über $\lim_{x \rightarrow x_0, x \neq x_0} f(x) = f(x_0)$ vorhersagen.
Oder auch:	Es gibt genau eine stetige Ergänzung von $(D - \{x_0\}, x \mapsto f(x), \mathbb{R})$, Gilt $x_0 \in D$, dann die Ergänzung durch den Wert $f(x_0)$. Andernfalls durch den Grenzwert.

Und ebenso gilt

Ist	f in $D - \{x_0\}$ definiert und ist x_0 Häufungspunkt von D ,
dann	gibt es höchstens eine stetige Ergänzung von f in x_0 .

Beachten Sie: $D = \{-1\} \cup]0, 1]$. dann läßt sich $f(-1)$ nicht vorhersagen, obwohl f dort stetig ist.

Die **Häufungswerte einer Folge a sind die Häufungspunkte von $\text{Bild}(a)$** . Denken Sie an das Beispiel $n \mapsto \sin(n)$. Was ist da die Menge aller Häufungspunkte?

- Bestimmen Sie die Menge $H(a)$ aller Häufungspunkte der Folge $a: n \mapsto \tan(n)$. Was bedeutet das für die Mächtigkeiten von $\text{Bild}(a)$ und $H(a)$?
- Beweisen Sie: Eine abgeschlossene Menge enthält alle ihre Häufungspunkte.

Weiß man, dass f in a stetig ist, dann genügt die Kenntnis beispielsweise der Werte $f(x_n)$ für $x_n = a + \frac{1}{n}$ zur Vorhersage von $f(a)$.

■ (4.4.6) **Extremwerte werden auf kompakten Mengen angenommen!** Dieser Satz verlangt als Wertemenge \mathbb{R} . Dagegen ist \mathbb{C} nicht zugelassen. Zunächst ein erläuterndes Gegenbeispiel: $f =]0, 1[, x \mapsto \frac{1}{x}, \mathbb{R}$. Hier werden die Extremwerte nicht angenommen. D.h. z.B.: es gibt kein $x_M \in D =]0, 1[$, so dass $f(x_M) \geq f(x)$ für alle $x \in D$. Ebensovienig gibt es $x_m \in D$ mit $f(x_m) \leq f(x)$ für alle $x \in D$. Der Rand- und Häufungspunkt 1 liegt ja nicht mehr in D ! Jetzt der Satz:

Sei	$f: K \rightarrow \mathbb{R}$ und $K \subset \mathbb{R}$ kompakt.
Dann	gibt es $x_m, x_M \in K$ so dass $f(x_m)$ minimal und $f(x_M)$ maximal in $\text{Bild} f$ wird. Die Extremwerte sind endlich und werden angenommen.

Beweis: Wir betrachten das Maximum. Dazu setzen wir $M = \sup \text{Bild} f$, sofern das Bild beschränkt ist. Und $M = \infty$, wenn das Bild unbeschränkt ist. Dieser Fall wird bald ausgeschlossen werden. Jetzt betrachten wir im Falle $M < \infty$ in der Wertemenge die Umgebungen $U_{\frac{1}{n}}(M)$ und wählen zu jedem n ein $x_n \in K$ mit $f(x_n) \in U_{\frac{1}{n}}(M)$. Ist $M = \infty$ wählen wir x_n so, dass $f(x_n) > n$. Nach Supremumsdefinition muss es solche x_n geben. Damit haben wir eine Folge $(n, n \mapsto x_n, K)$. Da K kompakt ist, existiert eine konvergente Teilfolge. $k \mapsto x_{n_k}$ mit Grenzwert $\lim_{k \rightarrow \infty} x_{n_k} = x_M$. Da f stetig ist, folgt $\lim_{k \rightarrow \infty} f(x_{n_k}) = f(x_M)$. Das schließt $M = \infty$ aus. Nach der Definition der x_n gilt außerdem $\lim_{k \rightarrow \infty} f(x_{n_k}) = M$. Also $f(x_M) = M$ wie gewünscht. Da $M = \sup \text{Bild} f$ ist, nimmt f in x_M sein absolutes Maximum M an.

■ (4.4.7) **Gleichmäßige Stetigkeit auf einem Kompaktum** (Vgl. hierzu die Beispiele. (3.2.16) und (3.2.17).

Sei K kompakt und $f: K \rightarrow \mathbb{R}$ stetig.

Dann ist f gleichmäßig stetig.

D.h.: Man kann eine für alle $x_0 \in K$ gültige Strategiefunktion angeben.

Oder auch:

Zu $\varepsilon > 0$ gibt es $\delta = \delta(\varepsilon)$, so dass $|f(x_1) - f(x_2)| < \varepsilon$ für alle $x_1, x_2 \in K$ mit $|x_1 - x_2| < \delta$.

Beweis: Indirekt. Angenommen f wäre nicht gleichmäßig stetig. D.h. es gibt ein $\varepsilon_0 > 0$ ohne zugehöriges für alle x gültiges δ . Wähle $\delta_n = \frac{1}{n}$. Dazu gibt es daher x_n und Δx_n mit $|\Delta x_n| < \delta_n = \frac{1}{n}$, so dass $|f(x_n + \Delta x_n) - f(x_n)| > \varepsilon_0$ ist (*). Wähle aus der Folge $n \mapsto x_n$ eine konvergente Teilfolge mit Grenzwert $X \in K$. Das geht, da K kompakt! (x_n und $x + \Delta x_n$ aus K). Da f stetig ist, folgt $\lim_{n \rightarrow \infty} f(x_n) = f(X)$. Und da $n \mapsto \Delta x_n$ Nullfolge ist, folgt über die Grenzwertsätze auch $\lim_{n \rightarrow \infty} (x_n + \Delta x_n) = X$. Also $\lim_{n \rightarrow \infty} f(x_n + \Delta x_n) = f(X)$. Und das heißt $n \mapsto (f(x_n + \Delta x_n) - f(x_n))$ ist Nullfolge im Widerspruch zur Eingangsannahme (*).

2.4.5 Funktionsfolgen und die Darstellung stetiger Funktionen

(4.5.1) Sei $D \subset \mathbb{R}$ ein fester Definitionsbereich und $\mathcal{F}(D, \mathbb{R})$ die Menge aller Abbildungen $D \rightarrow \mathbb{R}$. Eine zugehörige *Funktionsfolge* ist dann eine Abbildung der folgenden Art.

$$\boxed{(\mathbb{N}, n \mapsto f_n, \mathcal{F}(D, \mathbb{R})) \quad \text{mit} \quad f_n = (D, x \mapsto f_n(x), \mathbb{R})}$$

Sagen wir $D = [-1, 1]$ und $f_n = h_n$. Oder $D = \mathbb{R}$ und $f_n = c_n$ mit $c_n(x) = \frac{1}{n} \sin(nx)$.

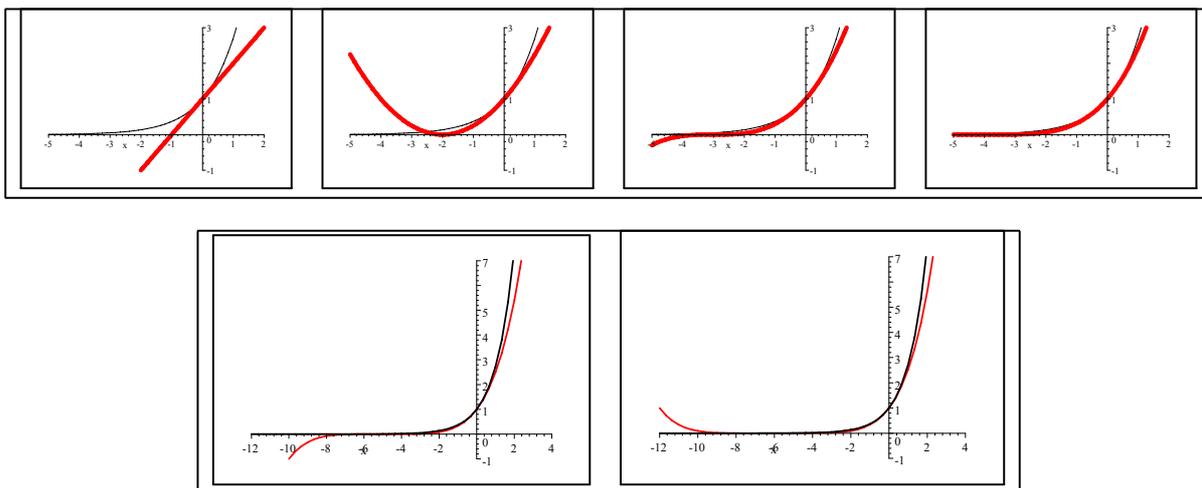
(4.5.2) Es liegt nahe, einen Rollenwechsel der folgenden Art vorzunehmen: Man macht $x \in D$ zu einem äußeren Parameter. Das ergibt eine *Schar von Zahlfolgen*, für jeden zulässigen Wert von x eine zugehörige Zahlfolge:

$$\boxed{(\mathbb{N}, n \mapsto f_n(x), \mathbb{R}) \quad \text{gewöhnliche Zahlfolge mit äußerem Parameter } x \in D}$$

Auf jede dieser Zahlfolgen kann man dann den gesamten für Zahlfolgen entwickelten Apparat loslassen.

(4.5.3) Wegen der großen Bedeutung dieses Rollenwechsels schieben wir eine Reihe von Beispielen ein.

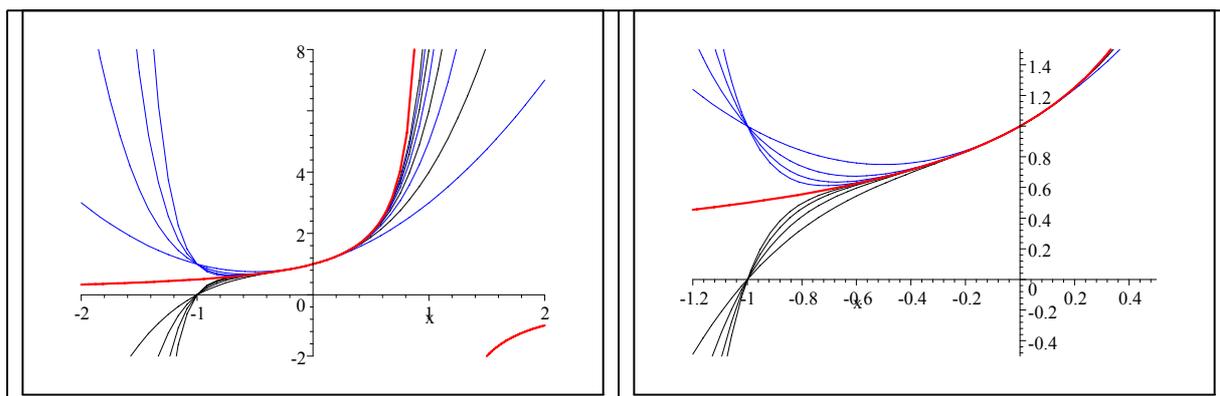
(4.5.4) **Beispiel Exponentialfunktion.** Wir hatten in Kap. 2.3.6 die Folge $n \mapsto a_n(x) = (1 + \frac{x}{n})^n$ betrachtet, bisher mit x als äußerem Parameter. Jetzt betrachten wir die Folge von Funktionen $x \mapsto a_n(x) = (1 + \frac{x}{n})^n$. Das ist ein Polynom vom Grade n mit $a_n(0) = 1 = \exp(0)$ und $a'_n(0) = 1 = \exp(0)$ und $a''_n(0) = \frac{n-1}{n}$ (statt 1). Weiter hat dieses Polynom eine Nullstelle n -ten Grades bei $x = -n$. Das ähnelt schon einigen wesentlichen Eigenschaften der Exponentialfunktion. Die ersten 4 Bilder zeigen die Graphen von $a_1(x)$ bis $a_4(x)$ zusammen mit dem der Exponentialfunktion. In der zweiten Reihe noch $a_5(x)$ und $a_6(x)$ in anderem Maßstab. Für $x > 0$ nähern sich die Polynome von unten der Exponentialfunktion. Für negative x erfolgt das polynomiale Wachstum erst jenseits der Stelle $x = -n$, wandert also immer weiter nach links. Beachten Sie: Die früher bewiesene Monotonieeigenschaft $a_n(x) \leq a_{n+1}(x)$ in (3.6.3) galt erst für $n > -x$, d.h. für $x > -n$ oder rechts von der Nullstelle.



Die Figuren zeigen die ersten Glieder der Funktionsfolge. Wählt man einen festen x -Wert, etwa $x=-2$, dann erhält man daraus die zugehörige Zahlfolge. Vgl. (3.6.2) für $x=1$.

- Lassen Sie von einem Computeralgebraprogramm den Graphen von $x \mapsto \left|1 - \frac{5}{x}\right|^x$ für $x \geq 0$ zeichnen. Welche Bedeutung hat das für unser Problem? Warum der Absolutbetrag? Der "erste Blick" auf den Graphen weckt eine falsche Erwartung. Nämlich?

(4.5.5) Als **zweites Beispiel betrachten wir die geometrische Reihe**. Wir wissen, dass sie nur für $|x| < 1$ gegen $\frac{1}{1-x}$ konvergiert. Die Zeichnung gibt die Grenzfunktion (rot) sowie die ersten Partialsummen wieder. Blau s_2, s_4, s_6 und s_8 und schwarz s_3, s_5, s_7 und s_9 . Die erste Gruppe geht durch $(-1,1)$ und die zweite durch $(-1,0)$. Für $x=1$ hat man keine Konvergenz mehr. Für $-1 < x < 0$ sieht man, wie sich die Funktionen immer mehr, aber nicht gleichmäßig der Grenzfunktion nähern. Für $x > 1$ ist von Approximation keine Rede mehr. Die Grenzfunktion ist negativ, während die Partialsummen nach $+\infty$ streben.



(4.5.6) **Beispiel** $f_n = ([0, 1], x \mapsto x^n, \mathbb{R})$ Das ergibt per Rollenwechsel die Folge $n \mapsto x^n$ mit dem äußeren Parameter x . Sie konvergiert gegen $f(x) = \begin{cases} 0 & \text{für } 0 \leq x < 1 \\ 1 & \text{für } x = 1 \end{cases}$.

Wechseln wir jetzt wieder (von der Folge mit äußerem Parameter x) zur Funktion (mit unabhängiger Variabler x) über, dann sieht man: Alle Funktionen f_n sind **stetig**, sogar C^∞ , aber die Grenzfunktion f ist in $x=1$ unstetig! Die Eigenschaft **stetig** bleibt beim Grenzübergang nicht notwendig erhalten! Schauen wir uns einmal die zugehörige Strategieabschätzung an:

$$x^n - 0 \leq x^N < \varepsilon \quad N(\varepsilon) \ln x > \ln(\varepsilon) \quad N(\varepsilon, x) = 2 \frac{\ln \varepsilon}{\ln x} \quad 0 < x, \varepsilon < 1.$$

Die Abschätzung ist **nicht gleichmäßig in x !**

(4.5.7) Zur Unterscheidung führen wir die folgende wichtige Sprechweise ein:

Gegeben	eine Folge $f: n \mapsto f_n(x)$ mit äußerem Parameter $x \in D$, die sich auch als Funktionsfolge $n \mapsto f_n$ interpretieren läßt.
★	Konvergiert jede Folge $f_n(x)$ gegen einen Grenzwert $f(x)$, dann sprechen wir von <i>punktweiser Konvergenz (der Funktionsfolge $n \mapsto f_n$)</i> .
★	Ist es (zusätzlich) möglich, für alle $x \in D$ dieselbe Strategiefunktion zu wählen, dann sprechen wir von gleichmäßiger Konvergenz (der Funktionsfolge) .

Aus gleichmäßiger Konvergenz folgt natürlich punktweise Konvergenz. Der umgekehrte Schluss ist falsch. Unser Beispiel hat gezeigt: Punktweise Konvergenz stetiger Funktionen reicht offensichtlich nicht, um Stetigkeit der Grenzfunktion zu sichern!

(4.5.8) Der folgende Satz zeigt die Bedeutung der gleichmäßigen Konvergenz:

Satz: Sei	$n \mapsto f_n$ Funktionsfolge mit $f_n \in \mathcal{F}(D, \mathbb{R})$. Alle f_n seien in $x_0 \in D$ stetig. Weiter konvergiere f_n gleichmäßig gegen f .
Dann	ist f in x_0 stetig! Sind die f_n stetig (in ganz D), dann ist auch f stetig (in D).

Man kann auch sagen: Der Grenzwert einer gleichmäßig konvergenten Folge stetiger Funktionen ist erneut stetig.

Beweis: Dies ist ein typischer Beweis aus dem Bereich der *Epsilontik*. Denken Sie immer an das Ziel, nämlich eine gültige Abschätzung für $|f(x_0 + \Delta x) - f(x_0)|$ zu erhalten. Dahin arbeitet man sich in mehreren Schritten vor.

1) Sei $\varepsilon > 0$ vorgeben. Bestimmen $N(\varepsilon)$, so dass gilt

$$|f_n(x) - f(x)| < \frac{\varepsilon}{3} \quad \text{für alle } x \in D \quad \text{und alle } n \geq N(\varepsilon)$$

Das ist infolge der gleichmäßigen Konvergenz möglich. $N(\varepsilon)$ ist unabhängig von x_0 . Wähle dann $n \geq N(\varepsilon)$.

2) Da f_n in x_0 stetig ist, wähle Δx so klein, dass $|f_n(x_0 + \Delta x) - f_n(x_0)| < \frac{\varepsilon}{3}$.

3) Jetzt rechne man mit Hilfe der Dreiecksungleichung wie folgt:

$$\begin{aligned} & |f(x_0 + \Delta x) - f(x_0)| \\ &= |(f(x_0 + \Delta x) - f_n(x_0 + \Delta x)) + (f_n(x_0 + \Delta x) - f_n(x_0)) + (f_n(x_0) - f(x_0))| \\ &\leq |(f(x_0 + \Delta x) - f_n(x_0 + \Delta x))| + |(f_n(x_0 + \Delta x) - f_n(x_0))| + |(f_n(x_0) - f(x_0))| \\ &< \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon. \end{aligned}$$

Das ist die geforderte Abschätzung und f ist wie behauptet stetig!

(4.5.9) Wann aber hat man gleichmäßige Konvergenz? Hier ein nützliches Kriterium, das in vielen Fällen das Gewünschte liefert:

Weierstraßsches Majorantenkriterium für gleichmäßige Konvergenz

Es sei	(1) $n \mapsto f_n = (D, x \mapsto f_n(x), \mathbb{R})$ eine Funktionsfolge.
	(2) Es gelte $ f_n(x) \leq a_n$ für alle n .
	(3) Die Reihe $\sum_n a_n$ sei konvergent.
Dann	ist die Reihe $\sum_n f_n(x)$ absolut und gleichmäßig konvergent auf D .

(4.5.10) **Folgerung:** Falls die f_n aus (3.5.7) stetig sind, ist auch f nach Satz (3.5.6) stetig.

(4.5.11) Beweis von (4.5.7): Durch Abschätzen des Betrages des Reihenrestes! Es sei N Strategiefunktion der konvergenten Reihe $\sum a_k$. Damit folgt

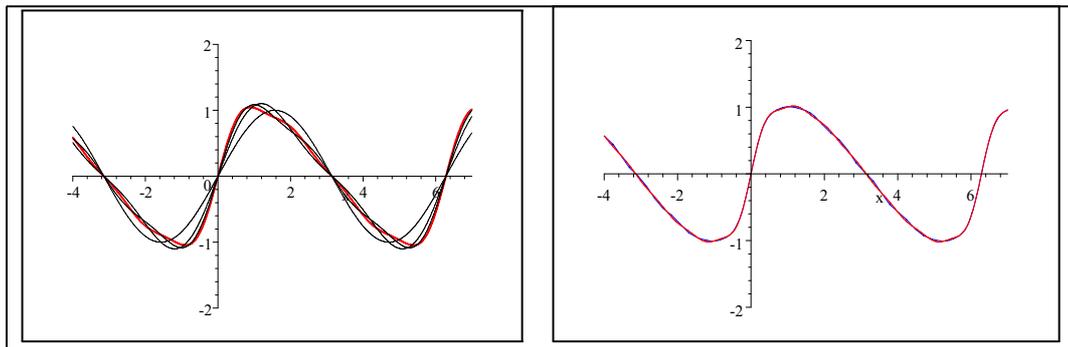
$$|r_n(x)| = |\sum_{k=n}^{\infty} f_k(x)| \leq \sum_{k=n}^{\infty} |f_k(x)| \leq \sum_{k=n}^{\infty} a_k < \varepsilon \quad n \geq N(\varepsilon)$$

Das ist eine von x unabhängige Abschätzung für die Reihen $\sum f_n(x)$.

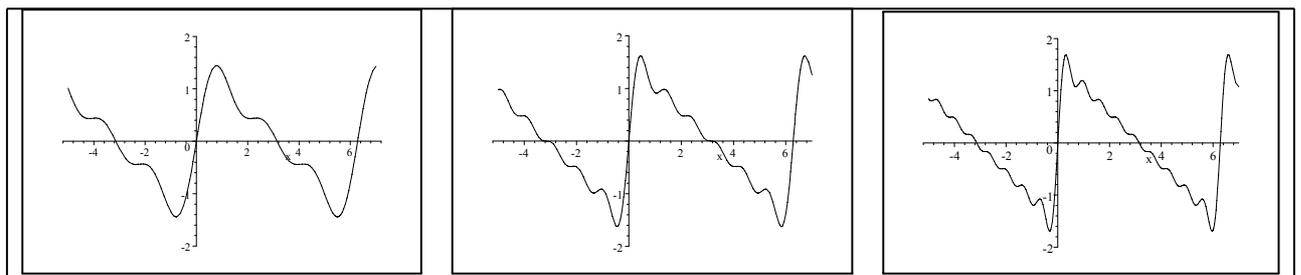
(4.5.12) **Anwendungsbeispiel:** Wir setzen $t_n(x) = a_n \sin(nx)$, wobei $n \mapsto a_n$ eine Zahlfolge sein soll, etwa $a_n = \frac{1}{n}$. Dann betrachten wir die Reihe $\sum t_n(x)$. Zunächst wählen wir x als äußeren Parameter. Etwa $x=0$. Dann sind alle Glieder Null, wir erhalten eine konstante Nullfolge, die Reihe gibt Null. Dasselbe gilt auch für $x=\pm\pi$. Für $x=\frac{\pi}{2}$ erhalten wir $\sum a_k \sin(k\frac{\pi}{2}) = \sum (-1)^k a_k$ usw. Nach Rollenwechsel sieht die Sache dagegen wie folgt aus: Jeder Summand ist periodisch mit Periode 2π und überdies ungerade. Es genügt, das Intervall $[-\pi, \pi]$ zu betrachten, und darauf sollte eine eventuelle Grenzfunktion erneut ungerade sein. $\sum t_n(x)$ ist ein Beispiel einer *trigonometrischen Reihe*.

$t_1(x) = \sin x$	$t_2(x) = \sin x + \frac{1}{2} \sin(2x)$	$t_3(x) = \sin x + \frac{1}{2} \sin(2x) + \frac{1}{3} \sin(3x)$	usw.
-------------------	--	---	------

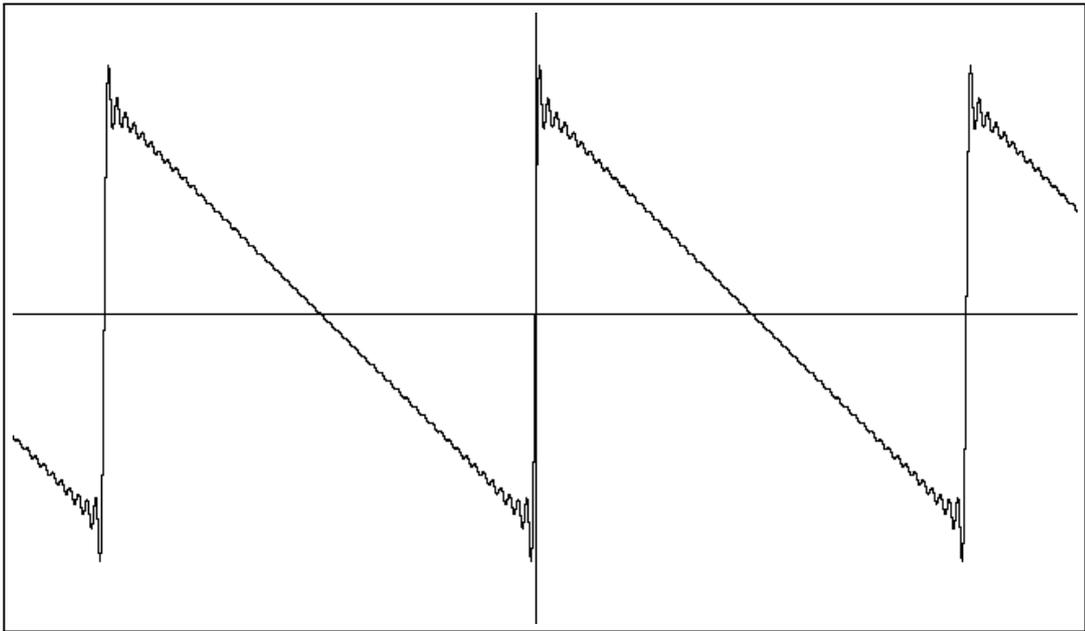
(4.5.13) Wann kann man das Majorantenkriterium für unser Beispiel anwenden? Wir haben $|t_n(x)| = |a_n| |\sin(x)| \leq |a_n|$. Das Kriterium ist anwendbar, wenn $\sum |a_n| < \infty$. Also beispielsweise für $a_n = \frac{1}{n^2}$. Es ist nicht anwendbar für $a_n = \frac{1}{n}$. Allerdings dürfen wir dabei nicht vorschnell auf Divergenz schließen: Die zusätzlichen Sinusfaktoren liefern noch unterschiedliche Vorzeichen, so dass Kompensationen auftreten können. Wir erproben das Verhalten der beiden Funktionsreihen einmal. Zunächst $\sum \frac{1}{n^2} \sin(nx)$. Nach dem Kriterium sollte das gleichmäßig gegen eine stetige Grenzfunktion gehen. Wir zeichnen das Intervall $[-4, 7]$, um auch die Periodizität zu verdeutlichen. Die Konvergenz ist sehr schnell. Im rechten Bild sind unterschiedlich gefärbt die 7. und die 8. Partialsumme gezeichnet. Es ist kein Unterschied mehr zu erkennen.



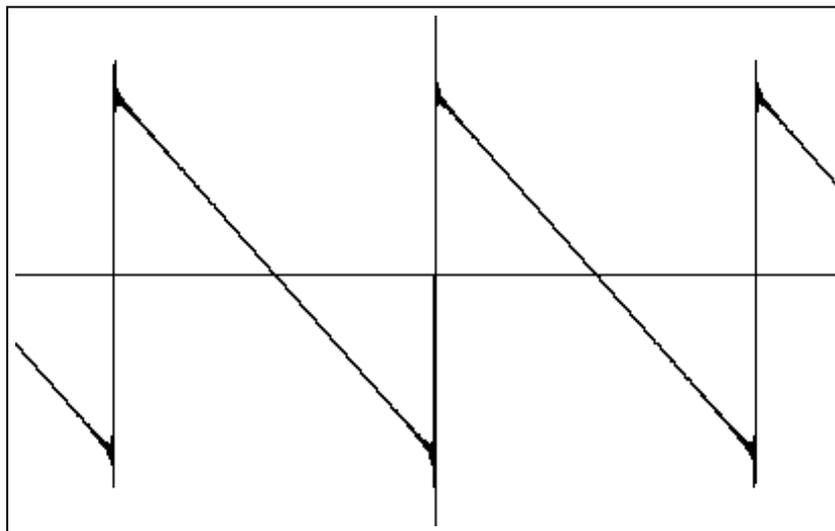
(4.5.14) Ganz anders sieht es im Falle $a_n = \frac{1}{n}$ aus. Hier muss man viele Terme mitnehmen, um ein stabiles Grenzverhalten zu erkennen. Dabei deutet sich folgender Sachverhalt an: Die Grenzfunktion ist unstetig und die Konvergenz ist zumindest in der Nähe der Unstetigkeitsstelle nicht gleichmäßig. Die ersten drei Bilder zeigen die Partialsummen s_3, s_6 und s_9 . Beispielsweise $s_3(x) = \sin x + \frac{1}{2} \sin(2x) + \frac{1}{3} \sin(3x)$. In der Nähe von $x=0$ erhalten wir mit Tangentenapproximation $s_3(x) \approx x + \frac{1}{2}2x + \frac{1}{3}3x = 3x$. Entsprechend $s_9(x) \approx 9x$. Die Steigung wächst unbeschränkt. Bei dagegen $x_k=(2k+1)\pi$ entstehen Nullstellen vom Typ $z - \frac{1}{2}2z + \frac{1}{3}3z - \frac{1}{4}4z + \dots$ mit $z=(x-x_k)$.

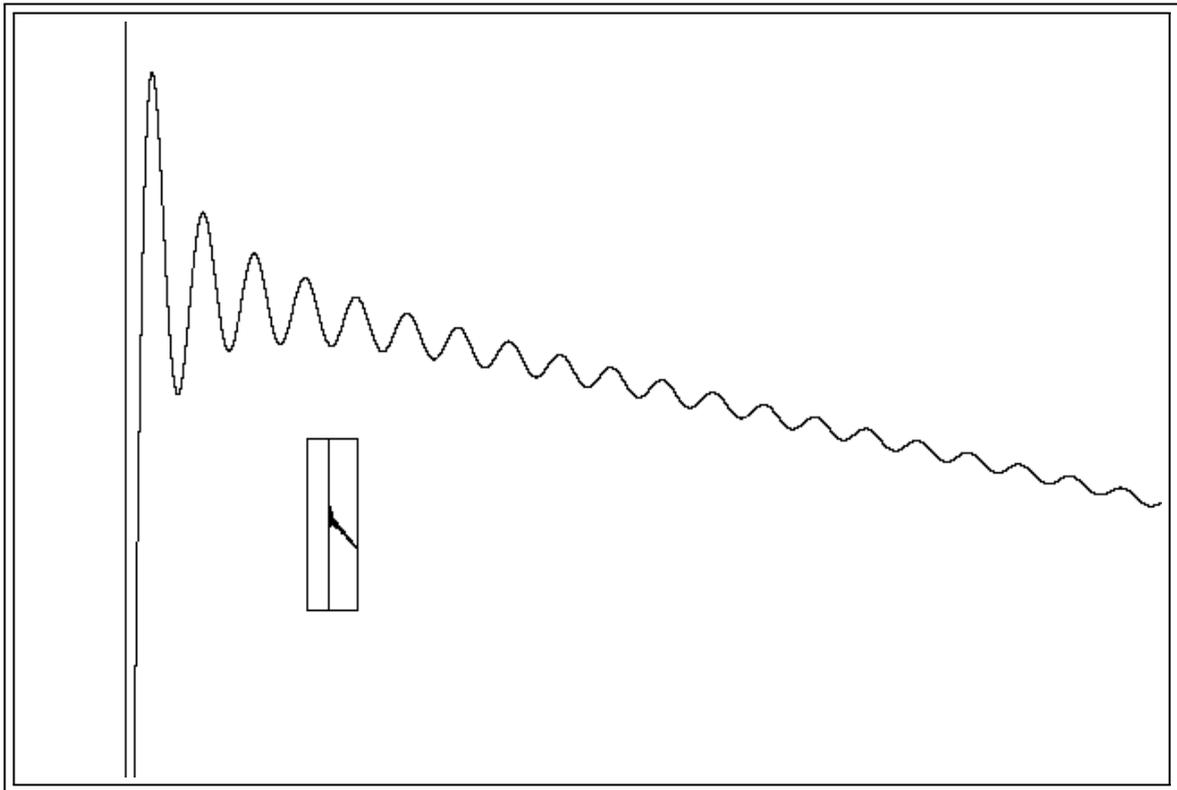


Das nächste Bild zeigt s_{50} . Man sieht, dass sich bei $x=2\pi k$ **Sprungstellen** entwickeln, Dazwischen konvergiert die Folge punktweise gegen eine Gerade.



Das nächste Bild zeigt s_{200} . In der Nähe der Sprungstellen sieht es nicht nach gleichmäßiger Konvergenz aus. Eine dieser Stellen ist im nachfolgenden Bild noch einmal vergrößert dargestellt.





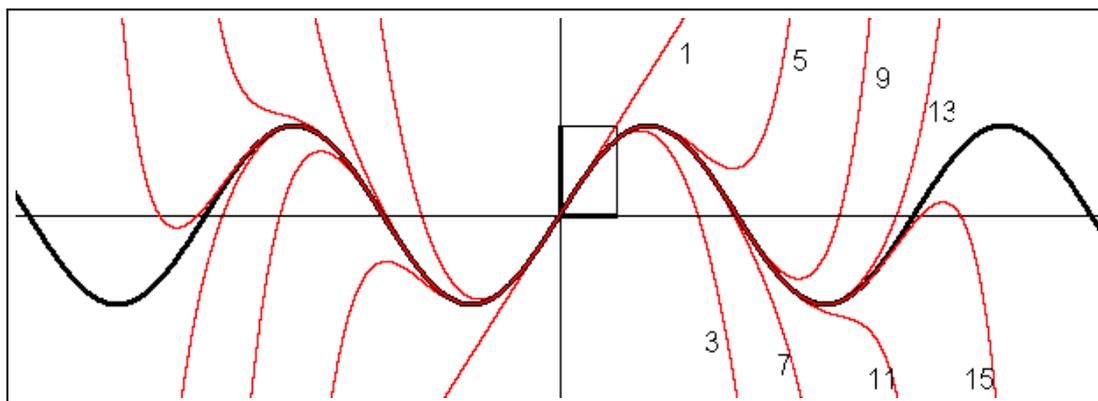
□ Wann ist eine Fourierreihe $\sum a_n \sin(nx)$ sicher stetig? (Bedingung für die a_n .)

(4.5.15) Im Falle der Binomialreihe hatten wir Konvergenz nur für $|x| < 1$. Wie verhalten sich Reihen, die für alle x konvergieren? Als Beispiel wählen wir die trigonometrische Funktion \sin .

Die Eulersche Formel $e^{ix} = \cos x + i \sin x$ liefert uns die beiden zugehörigen Reihen. Die Rechnung ist zulässig, alle beteiligten die Reihen absolut konvergieren.

$$e^{ix} = \sum_{k=0}^{\infty} \frac{(ix)^k}{k!} = \sum_{r=0}^{\infty} \frac{(-1)^r}{(2r)!} x^{2r} + i \sum_{r=0}^{\infty} \frac{(-1)^r}{(2r+1)!} x^{2r+1} = \cos x + i \sin x.$$

Einige Beobachtungen: \sin ist eine gerade Funktion, die Reihe enthält nur die ungeraden Potenzen. Die \cos -Reihe dagegen die geraden. Beide Reihen sind alternierend als Folge der Werte von i^k . Die beiden resultierenden Reihen sind erneut für alle Werte von x absolut konvergent. Die Figur zeigt etwa zwischen -3π und 3π . Dazu rot die ersten Approximationen. Angeschrieben ist jeweils $2r+1$, die höchste noch mitgenommene Potenz.



Je weiter sich x vom Ursprung entfernt, umso mehr Terme muss man mitnehmen, um eine akzeptable Approximation zu erhalten. Die Konvergenz ist nur in kompakten Intervallen gleichmäßig und wird dort von den Punkten bestimmt, die am weitesten vom Ursprung entfernt sind.

2.4.6 Differenzierbarkeit

(4.6.1) Die Definitionsbereiche seien in 2.3.6 immer als endliche Vereinigung von nicht trivialen Intervallen darstellbar. An eventuellen Randpunkten sind einseitige Limiten zu bilden und zu betrachten. Es soll keine isolierten Punkte geben. Dann ist jeder Punkt von D ein Häufungspunkt von D .

(4.6.2) Szenenbild: $f : D \rightarrow \mathbb{R}$, $x_0 \in D$ und $\Delta x = x - x_0 \neq 0$ für $x \in D$ und $m \in \mathbb{R}$. Damit können wir folgende Termumformung vornehmen

$$f(x_0 + \Delta x) = f(x_0) + m\Delta x + \Delta x R_f^m(x_0, \Delta x).$$

Diese Gleichung bestimmt den Term R_f^m für $\Delta x \neq 0$ eindeutig, da man nach ihm umstellen kann. m ist bisher beliebig. Als Zusatzbedingung wählen wir

$$\lim_{\substack{\Delta x \rightarrow 0 \\ \Delta x \neq 0}} R_f^m(x_0, \Delta x) = 0 \quad \text{Resttermbedingung}$$

Man kann auch sagen: $\Delta x \mapsto R_f^m(x_0, \Delta x)$ ist für $\Delta x = 0$ stetig ergänzbar mit Wert 0. Die übrigen Größen x_0 und m sind äußere Parameter. Vgl. auch (3.2.6).

(4.6.3) Je nach Wahl von m kann das erfüllt sein oder nicht. Nun gilt:

Satz:	Es gibt höchstens einen Wert von m , für den die Restermbedingung erfüllt ist.
--------------	--

Oder auch: R ist für höchstens einen Wert in $\Delta x = 0$ stetig ergänzbar mit Wert 0.

Beweissizze: Angenommen man hat zwei Zerlegungen der angegebenen Art. D.h. zwei gültige Gleichungen mit erfüllter Restermbedingung

$$\begin{aligned} f(x+\Delta x) &= f(x) + m_1\Delta x + \Delta x R_1^{m_1}(x, \Delta x) \\ f(x+\Delta x) &= f(x) + m_2\Delta x + \Delta x R_2^{m_2}(x, \Delta x) \end{aligned}$$

Subtrahiere die beiden Gleichungen, dividiere durch Δx und gehe zum Grenzwert $\Delta x \rightarrow 0$ ($\Delta x \neq 0$) über. Das gibt $m_1 = m_2$. Beachten Sie: Hier liegt ein Beispiel der Wertvorhersage (3.4.5) vor. Setzt man $m_1 = m_2$ in die durch Subtraktion erhaltene gültige Gleichung ein, folgt $R_f^{m_1} = R_f^{m_2}$. D.h. die Zerlegungen sind gleich. Natürlich kann die Beweisausgangslage unerfüllbar sein in dem Sinne, dass man überhaupt keine Zerlegung finden kann.

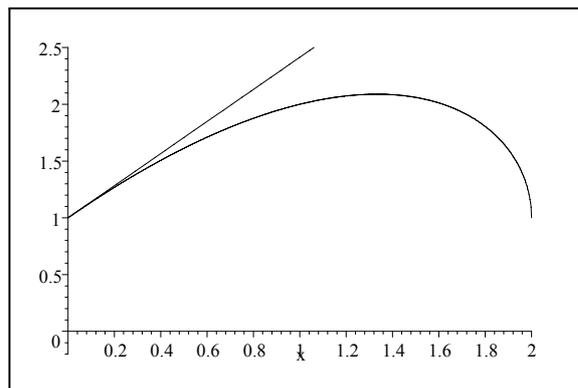
(4.6.4) Ist die Restermbedingung für genau einen Wert erfüllt, dann verwendet man folgende Beszeichnungen und Sprechweisen

Statt m :	$f'(x_0) = \frac{df}{dx}(x_0) = (Df)(x_0) = \dot{f}(x_0) = \dots$	"Ableitung von f in x_0 "
	$f(x_0 + \Delta x) = f(x_0) + f'(x_0)\Delta x + \Delta x R_f(x_0, \Delta x)$	Die Tangentenerlegung von f um x_0
	$f(x_0 + \Delta x) \approx f(x_0) + f'(x_0)\Delta x$	Tangentenapproximation von f um x_0
	$F(x_0, \Delta x) = \Delta x R_f(x_0, \Delta x)$	Absoluter Fehler (von)
	$R_f(x_0, \Delta x)$	Restterm oder lokaler Fehler.

f ist in x_0 differenzierbar mit Ableitung $f'(x_0)$
--

(4.6.5) Zu eventuellen Randpunkten von D gehören **einseitige Ableitungen**. D.h. man hat mit Relativumgebungen U_ε^D zu arbeiten. (Vgl. (2.3.3).) Sei beispielsweise $f = ([0, 2], x \mapsto 1 + \sqrt{x^2(2-x)}, \mathbb{R})$ und $x_0 = 0$. Dann ist $\Delta x \rightarrow 0$, $\Delta x > 0$ zu betrachten. Wir werden sehen, dass $m = \sqrt{2}$ die zugehörige einseitige

Ableitung ist. Die Figur zeigt den Graphen samt zugehöriger Tangente (Bzw. Halbtangente) $x \mapsto 1 + x\sqrt{2}$. Bei $x=2$ gibt es dagegen kein zulässiges m. (Unendliche Steigung!)



(4.6.6) Ist f in jedem Punkt $x_0 \in G$ differenzierbar, dann heißt f *differenzierbar (in D)*. Unter diesen Umständen kann man die *Ableitungsfunktion* definieren

$$f' = Df = (G, x \mapsto f'(x), \mathbb{R}).$$

(4.6.7) Ist f' in G stetig, so sagt man f sei **stetig differenzierbar** und verwendet zur symbolischen Charakterisierung das Zeichen C^1 .

2.4.6a Die zwei Denkfiguren

(4.6.8) Die Tangentzerlegung ist eine gültige Gleichung, bei der der interessierende Term $f(x_0 + \Delta x)$ in einer für die Anwendungen günstigen Weise codiert oder aufgespalten ist. Man benutzt sie vornehmlich im Rahmen von zwei häufig auftretenden Denkfiguren, die wir jetzt beschreiben und in denen das Vorgehen unterschiedlich ist.

(4.6.9) **Die erste Denkfigur.**

- ★ Das zugehörige Szenenbild: $f, x_0, f(x_0 + \Delta x)$ sind gegeben.
- ▼ Durch geeignete Rechenmanipulationen für $f(x_0 + \Delta x)$ verschafft man sich eine Zerlegung der Form (3.6.2)
- ▼ Für das darin enthaltene R weist man die Restermeigenschaft nach.
- ▼ Dann kann man nach (3.6.3) folgern, dass f in x_0 differenzierbar ist und über den linearen Term liest man die zugehörige Ableitung ab. **Die Denkfigur liefert Differenzierbarkeit und eine Formel für die Ableitung!**

(4.6.10) **Die zweite Denkfigur**

- ★ Das zugehörige Szenenbild: $f, x_0, f(x_0 + \Delta x)$ und $f'(x_0)$ sind gegeben, f ist in x_0 differenzierbar
- ▼ Dann schreibt man die Tangentzerlegung als gültige Gleichung (mit der man weiter arbeiten kann hin). Für den Restterm R reduziert man das Wissen über die exakte Funktion auf die Restermeigenschaft. **Damit sucht man dann die Lösung anstehender Probleme.**

1. (4.6.11) Wie im Vorkurs beschrieben, gibt es zum Start der ganzen Theorie ein **Routinevorgehen:**

- ▼ Man bestimmt für die üblichen Grundfunktionen die Ableitung mit Hilfe der ersten Denkfigur
- ▼ Man beweist mit Hilfe beider Denkfiguren die üblichen Ableitungsregeln
- ▼ Man berechnet in Problemen die Ableitung einer Funktion mit Hilfe der Ableitungsregeln und verwendet dann die zweite Denkfigur, um das zu behandeln.

(4.6.12) Ein **Beispiel** einer differenzierbaren Funktion, die nicht C^1 ist, bei der die Ableitung also unstetig ist. Sei $f(x) = x^2 \sin \frac{1}{x}$ für $x \neq 0$ und $f(0) = 0$. Man zeigt leicht, dass f stetig ist. Über die erste Denkfür folgt Differenzierbarkeit in $x_0 = 0$ mit $f'(0) = 0$. Für $x \neq 0$ folgt die Ableitung über die Ableitungsregeln. Dieser sieht man per Inspektion an, dass sie keinen Grenzwert für x nach Null hat.

□ Führen Sie die Einzelheiten aus.

(4.6.13) Die Ableitungsregeln und Beispiele dazu im Vorkurs.

2.4.7 Allgemeine Konsequenzen der Differenzierbarkeit

(4.7.1) Wie im Fall der Stetigkeit gibt es auch im Falle der Differenzierbarkeit eine bestimmte sachgemäße Reihenfolge des Vorgehens. Diese beginnt mit den Definitionen, die bei der Formulierung der Resttermbedingung auf den Stetigkeitsbegriff zurückgreifen. Dann werden die **elementaren** Konsequenzen in Form der Ableitungen einfacher Funktionen und der Ableitungsregeln erarbeitet. Im nächsten Schritt geht es um die **allgemeinen** Konsequenzen der Differenzierbarkeit. Die sind in der Regel relativ abstrakt, d.h. sie werden meist nicht unmittelbar zur Lösung konkreter Probleme benötigt, sie werden eher benötigt, um dabei benutzte Verfahren zu entwickeln und zu rechtfertigen.

Die elementaren Konsequenzen setzen wir als bekannt voraus.

Wir besprechen davon jetzt die wichtigsten.

(4.7.2) **Das Kriterium für einen lokalen glatten Extremwert.**

Satz: $f:]a, b[\rightarrow \mathbb{R}$ habe in $c \in]a, b[$ einen lokalen Extremwert.
Weiter sei f in c differenzierbar.
Dann ist notwendig $f'(c) = 0$.

Beweis. O.B.d.A. liege in c ein lokales Maximum vor. D.h. Es gibt ein $\varepsilon_0 > 0$, so dass $f(c) \geq f(x)$ für alle $x \in U_{\varepsilon_0}(c)$. Beachten Sie: c ist notwendig innerer Punkt von $]a, b[$. Da f in c differenzierbar sein soll, dürfen wir die Tangentenerlegung als gültige Gleichung anschreiben:

$$f(c + \Delta x) = f(c) + f'(c)\Delta x + \Delta x R_f(c, \Delta x).$$

Angenommen $f'(c) \neq 0$. Dann wähle $\varepsilon_1 = \frac{1}{2}|f'(c)| > 0$ und $\varepsilon = \min(\varepsilon_0, \varepsilon_1)$. Sazu δ so dass für $|\Delta x| < \delta$ gilt $|R_f(c, \Delta x)| < \varepsilon$. R_f ist ja stetig. Dann folgt

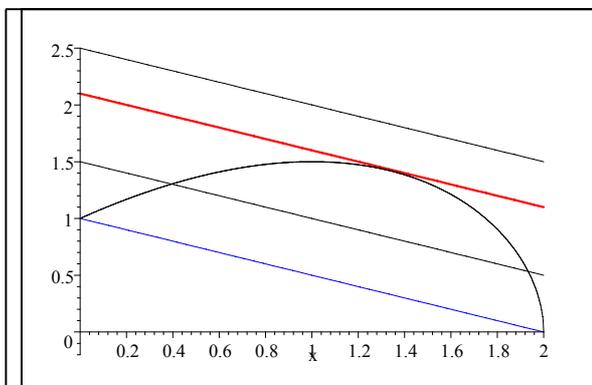
$$f(c + \Delta x) - f(c) = \Delta x(f'(c) + R_f(c, \Delta x)).$$

Nun kann man das Vorzeichen von Δx so wählen, dass $\Delta x f'(c) > 0$ ist. Da $|R_f(c, \Delta x)| < \frac{1}{2}|f'(c)|$ gilt, muß $f(c + \Delta x) - f(c) > 0$ gelten. Und das heißt, entgegen der Annahme kann f in c kein lokales Maximum haben.

□ Gilt der Satz auch für ein lokales Extremum in einem Randpunkt?

(4.7.3) Der zweite Satz ist der Mittelwertsatz, nicht zu verwechseln mit dem nur Stetigkeit fordernden Zwischenwertsatz.

Mittelwertsatz Sei $K = [a, b]$ kompakt und $f: K \rightarrow \mathbb{R}$ stetig und in $]a, b[$ differenzierbar.
Dann gibt es ein $\zeta \in]a, b[$ mit $f'(\zeta) = \frac{f(b) - f(a)}{b - a}$



Geometrische Interpretation
Verschiebt man die Sekante (blau) zwischen a und b parallel, so gibt es eine Stelle, an der diese zur Tangente (rot) wird.

Im Spezialfall $f(a)=f(b)$ ist die Sekante parallel zur x-Achse, so dass der Tangentialpunkt die Steigung Null hat.

□ $f(x)=x^3$ und $x_0 = 0$. Versuchen Sie die Konstruktion der Figur für diesen Fall. Was ist anders?

(4.7.4) Der Beweis erfolgt durch **Rückführung auf einen Spezialfall**. Dazu führt man die folgende Hilfsfunktion g ein

$$g(x) = f(x) - \frac{f(b) - f(a)}{b - a}(x - a) \quad \text{mit } g(a)=g(b).$$

Findet man hierfür ein ζ mit $g'(\zeta) = 0$, dann erfüllt $f'(\zeta)$ offensichtlich die Forderung des Mittelwertsatzes. Die Existenz eines ζ für g folgt aber aus dem jetzt folgenden Satz von Rolle.

Satz von Rolle: Sei $K=[a,b]$ kompakt und $g:K \rightarrow \mathbb{R}$ stetig und in $]a,b[$ differenzierbar.
Weiter sei $g(a)=g(b)$.
Dann gibt es $\zeta \in]a,b[$ mit $g'(\zeta) = 0$.

Beweis: Dieser ist trivial, wenn g konstant ist. Sei g nicht konstant. Da g stetig auf dem kompakten Intervall K ist, nimmt g seine Extremwerte an. Mindestens einer dieser Extremwerte ist ungleich $g(a)$. O.B.d.A. sei das das Maximum M der Funktionswerte (Andernfalls die Funktion $-g$ betrachten!). M werde in $c \neq a, b$ angenommen: $M=g(c)$. Nun soll g in c differenzierbar sein. Dann ist die Ableitung Null, sonst läge kein glatter Extremwert vor.

(4.7.5) Jetzt wollen wir eine weitere wichtige Folgerung herausarbeiten.

Angenommen wir wissen $m \leq f'(x) \leq M$ für alle $x \in]a,b[$. Dann folgt über den Zwischenwertsatz ($b > a$):

$$m \leq f'(\zeta) = \frac{f(b) - f(a)}{b - a} \leq M \quad \text{und damit}$$

$$m(b - a) \leq f(b) - f(a) \leq M(b - a)$$

Für b darf man natürlich auch jedes x mit $a < x \leq b$ wählen. Das gibt den

Schranksatz: Es sei f in $[a,b]$ stetig und in $]a,b[$ differenzierbar.
Für alle $x \in]a,b[$ gelte $m \leq f'(x) \leq M$.
Dann folgt für alle $x \in [a,b]$ die folgende Ungleichung für den Zuwachs
 $m(x - a) \leq f(x) - f(a) \leq M(x - a)$.

In etwas allgemeinerer Form haben wir das im Vorkurs als "Satz vom beschränkten Zuwachs" eingeführt. Dort haben wir beispielsweise die wichtige Monotonie der Integration damit begründet. In Kapitel 6 kommen wir erneut auf diesen Satz zurück und geben einen Beweis.

- Was folgt, wenn man eine Ungleichung $|f'(x)| \leq K$ hat?
- Was folgt für monotone differenzierbare Funktionen?

2.4.8 Konvexe Funktionen und Differenzierbarkeit

(4.8.1) Wir erinnern an die Zweipunkteformel einer Geraden:

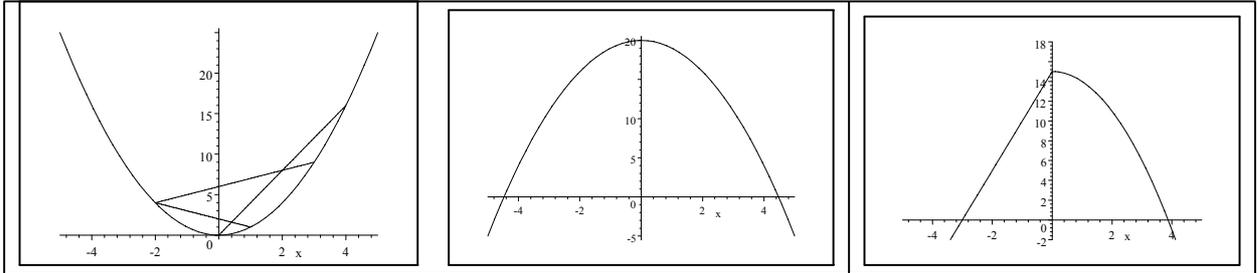
$$\vec{x}_g(\lambda) = \vec{a} + \lambda(\vec{b} - \vec{a}) = (1 - \lambda)\vec{a} + \lambda\vec{b} = \lambda_1\vec{a} + \lambda_2\vec{b} \quad \text{mit } \lambda_1 + \lambda_2 = 1$$

Die Verbindungsstrecke zwischen \vec{a} und \vec{b} wird erhalten, wenn man zusätzlich $\lambda_1, \lambda_2 \geq 0$ verlangt. D.h. $0 \leq \lambda \leq 1$. Jetzt betrachten wir zwei Punkte $(x, f(x))$ auf einem Funktionsgraphen. Je zwei von ihnen können wir verbinden. Liegt dann die Verbindungsstrecke (=Sekante) immer auf einer Seite des Graphen oder schneidet sie diesen? Liegt sie immer auf einer Seite, bezeichnet man die Funktion als konvex bzw. konkav (im betrachteten Bereich). Vgl. die Figur in (3.7.3).

(4.8.2) Dieser Sachverhalt lässt sich wie folgt präzisieren:

Definition: $I \subset \mathbb{R}$ Intervall. $f: I \rightarrow \mathbb{R}$ heißt *konvex* in I ,
wenn für alle $x_1, x_2 \in I$ und λ_1, λ_2 mit $\lambda_1 + \lambda_2 = 1$, $\lambda_1, \lambda_2 \geq 0$ gilt:
(K1) $f(\lambda_1 x_1 + \lambda_2 x_2) \leq \lambda_1 f(x_1) + \lambda_2 f(x_2)$.

Links in der Figur der Graph einer konvexen glatten Funktion mit drei eingezeichneten Sekanten oberhalb des Graphen. In der Mitte und rechts Beispiele konkaver Funktionen. Auch die rechte Funktion ist konkav, aber bei $x=0$ nicht differenzierbar.



- Weitere Präzisierung: Definieren Sie *streng konvex* sowie *konkav* und *streng konkav*.
- Verallgemeinerung Sie die Ungleichung (K1) durch Induktion.
- Welche Beziehung ist außerhalb der Verbindungsstrecke (der Sekanten) zu erwarten?
- Welche der elementaren Funktionen sind konvex, welche konkav?
- In welchem Teilintervall ist $x \mapsto x(x^2 - 4)$ konvex?

(4.8.4) Sei f konvex. Weiter sei $D(x_1, x_2) = \frac{f(x_2) - f(x_1)}{x_2 - x_1}$ die Steigung der Sekanten zwischen den Graphenpunkten $(x_1, f(x_1))$ und $(x_2, f(x_2))$, wobei wir allerdings $x_1 < x_2$ fordern.

(4.8.5) Wir betrachten jetzt drei Punkte $x_1 < x < x_2$ im Konvexitätsbereich und behaupten:

Satz: Sei I Intervall und $f: I \rightarrow \mathbb{R}$.

Dann ist f genau dann konvex, wenn für alle $x, x_1, x_2 \in I$ mit $x_1 < x < x_2$ gilt

(K2) $D(x_1, x) \leq D(x, x_2)$ oder ausgeschrieben $\frac{f(x) - f(x_1)}{x - x_1} \leq \frac{f(x_2) - f(x)}{x_2 - x}$

Beachten Sie, dass es nicht um Beträge geht. So ist etwa $-100 \leq +1$ zulässig. Der Beweis folgt weiter unten.

- Geben Sie ein Beispiel einer nicht konvexen Funktion mit verletzter Ungleichung.
- Interpretieren Sie die Ungleichung geometrisch.
- Wieso ist die folgende Gleichung für $x_1 < x < x_2$ trivialerweise korrekt? (Denken Sie auch an unsere Konvention $y > x$ in $D(x, y)$ zu verlangen. Das vermeidet Betragsstriche!)

$$D(x_1, x)(x - x_1) + D(x, x_2)(x_2 - x) = D(x_1, x_2)(x_2 - x_1)$$

(4.8.6) Aus (K2) und der Gleichung der letzten Frage folgt eine weitere Beziehung zwischen den Steigungen:

$$\begin{aligned} D(x_1, x_2)(x_2 - x_1) &= D(x_1, x)(x - x_1) + D(x, x_2)(x_2 - x) \\ &\leq D(x, x_2)(x - x_1) + D(x, x_2)(x_2 - x) = D(x, x_2)(x_2 - x_1) \\ D(x_1, x_2) &\leq D(x, x_2) \quad x_1 < x < x_2 \end{aligned}$$

Und analog

$$\begin{aligned} D(x_1, x_2)(x_2 - x_1) &= D(x_1, x)(x - x_1) + D(x, x_2)(x_2 - x) \\ &\geq D(x_1, x)(x - x_1) + D(x_1, x)(x_2 - x) = D(x_1, x)(x_2 - x_1) \\ D(x_1, x) &\leq D(x_1, x_2) \end{aligned}$$

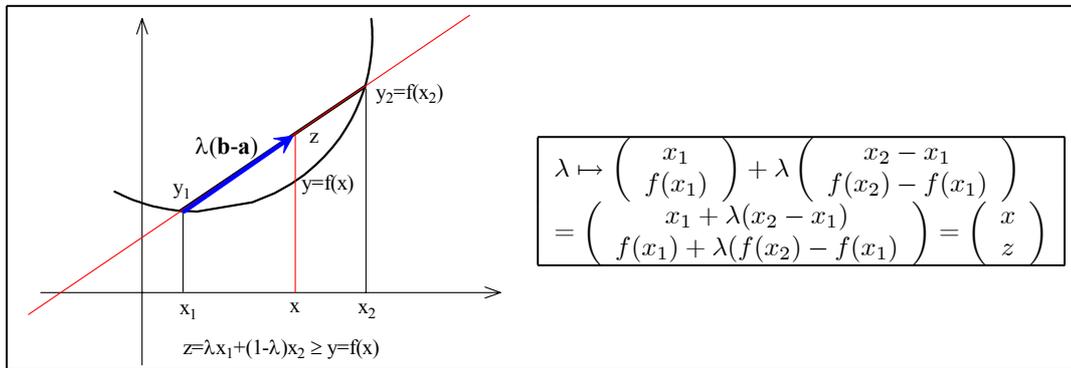
(4.8.7) Zusammen folgt der

Satz: Sei $f: I \rightarrow \mathbb{R}$ konvex und $x_1 < x < x_2$ mit $x_1, x_2 \in I$.
Dann gilt $D(x_1, x) \leq D(x_1, x_2) \leq D(x, x_2)$

- Fertigen Sie eine Skizze, die den Inhalt auch dieser Ungleichung verdeutlicht.
- Interpretieren Sie die beiden Ungleichungen $D(x_1, x) \leq D(x_1, x_2)$ und $D(x_1, x_2) \leq D(x, x_2)$ als Monotonieeigenschaft von D .

(4.8.8) Zum Beweis von (K2) selbst benötigen wir einige Hilfsüberlegungen.

(4.8.9) Bisher haben wir die Punkte der Verbindungsstrecke von $(x_1, f(x_1))$ und $(x_2, f(x_2))$ durch den Parameter λ erfaßt, der den Verbindungsvektor umskalierte, also $\vec{a} + \lambda(\vec{b} - \vec{a})$. Ausgeschrieben in Koordinaten:



Wir können alternativ die x-Koordinate des Punktes als Parameter nehmen. Man erhält die Umrechnung durch Umstellen nach λ :

$$x = x_1 + \lambda(x_2 - x_1) = (1 - \lambda)x_1 + \lambda x_2$$

$$\lambda = \frac{x - x_1}{x_2 - x_1} \quad 1 - \lambda = \frac{x_2 - x}{x_2 - x_1}$$

(4.8.10) Damit findet man eine andere gleichwertige Bedingung für konvex:

Satz: $f : I \rightarrow \mathbb{R}$ ist genau dann konvex, wenn für alle x_1, x, x_2 mit $x_1 \leq x \leq x_2$ gilt:

(K3) $f(x) \leq \frac{x_2 - x}{x_2 - x_1} f(x_1) + \frac{x - x_1}{x_2 - x_1} f(x_2)$

Der Unterschied liegt im zugehörigen Szenenbild: Bei (K1) wird man typischerweise mit x_1 und x_2 beginnen und dann ein x dazwischen wählen. Bei (K2) dagegen beginnt man eher mit x und wählt dann dazu die Punkte x_1 und x_2 .

(3.8.11) Mit (K3) können wir Satz (3.8.5) für die Sekantensteigung beweisen. Dazu rechnen wir wie folgt, wobei wir mit (K2) beginnen

$$\begin{aligned} (x_2 - x_1)f(x) &\leq (x_2 - x)f(x_1) + (x - x_1)f(x_2) \\ ((x_2 - x) + (x - x_1))f(x) &\leq (x_2 - x)f(x_1) + (x - x_1)f(x_2) \\ (x_2 - x)(f(x) - f(x_1)) &\leq (x - x_1)(f(x_2) - f(x)) \\ \frac{f(x) - f(x_1)}{x - x_1} &\leq \frac{f(x_2) - f(x)}{x_2 - x} \end{aligned}$$

Also $D(x_1, x) \leq D(x, x_2)$ wie behauptet. Jeder Schritt dieser Rechnung geht auch in umgekehrter Richtung, so dass die "genau"-Aussage des Satzes bewiesen ist.

(4.8.12) Kann man zum Grenzwert $x_2 - x_1 = 0$ übergehen? Im differenzierbaren Fall geht $D(x_1, x_2)$ dann in die Ableitung $f'(x_1)$ über. Erstaunlich. Dazu benutzen wir die Ungleichungskette (3.8.7) in geeigneter Form. Beachten Sie dabei, dass (3.8.7) einfach gewisse Monotonieeigenschaften von D beinhaltet.

(4.8.13) Für jeden inneren Punkt $x \in I$ gibt es $x_1, x_2 \in I$ mit $x_1 < x < x_2$. Nach (3.8.7) gilt für $\Delta x > 0$ aber $D(x_1, x) \leq D(x + \Delta x, x)$ D.h. $\Delta x \mapsto D(x + \Delta x, x)$ ist nach unten (durch $D(x_1, x)$) beschränkt und monoton wachsend in $\Delta x > 0$. Also existiert der einseitige Grenzwert

$$D(x + 0, x) = \lim_{\substack{\Delta x \rightarrow 0 \\ \Delta x > 0}} \frac{f(x + \Delta x) - f(x)}{\Delta x} = f'(x + 0).$$

Ebenso ist $D(x-\Delta x, x) \leq D(x, x_2)$ und $\Delta x \mapsto D(x-\Delta x, x)$ monoton wachsend. Damit existiert.

$$D(x-0) = \lim_{\substack{\Delta x \rightarrow 0 \\ \Delta x > 0}} \frac{f(x-\Delta x) - f(x)}{\Delta x} = f'(x-0).$$

Ergebnis: Für eine konvexe Funktion existieren zu jedem inneren Punkt die beiden einseitigen Ableitungen (die aber nicht gleich sein müssen!)

(4.8.14) Seien jetzt x und y mit $x < y$ innere Punkte von I und $0 < \Delta z < \frac{1}{2}(x-y)$. Da x, y innere Punkte von I sind, kann man Δz zusätzlich so klein wählen, dass $x-\Delta z$ und $y+\Delta z$ noch in I liegen. Dann folgt mit den durch (3.8.7) ausgedrückten Monotonieeigenschaften folgende Ungleichungskette:

$$D(x-\Delta z, x) \leq D(x, x+\Delta z) \leq D(x, y-\Delta z) \leq D(x, y) \leq D(y-\Delta z, y) \leq D(y, y+\Delta z)$$

(4.8.15) Jetzt gehen wir zur Grenze $\Delta z \rightarrow 0, \Delta z > 0$ über. Nach (3.8.13) existieren diese Grenzwerte und man erhält:

$$\boxed{f'(x-0) \leq f'(x+0) \leq D(x, y) \leq f'(y-0) \leq f'(y+0)}$$

Nochmals: Diese Ungleichungskette gilt für alle inneren Punkte $x < y$ der konvexen Funktion $f: I \rightarrow \mathbb{R}$.

□ Interpretieren Sie die Ungleichungen graphisch!

Sei x innerer Punkt von I . Dazu gibt es zwei weitere innere Punkte x_1, x_2 von I mit $x_1 < x < x_2$. Dazu wählen wir Δx mit $\Delta x < x_2 - x$ und $\Delta x < x - x_1$. Nach (3.8.15) folgt

$$f'(x_1+0) \leq D(x-\Delta x, x) \leq D(x, x+\Delta x) \leq f'(x_2)$$

Nun war $D(x-\Delta x, x) = \frac{f(x-\Delta x) - f(x)}{-\Delta x}$ und $D(x, x+\Delta x) = \frac{f(x+\Delta x) - f(x)}{\Delta x}$. D.h. die Sekantensteigung ist in Δx beschränkt durch $C = \max(|f'(x_1+0)|, |f'(x_2-0)|)$

$$| \text{D.h.} \frac{|f(x+\Delta x) - f(x)|}{|\Delta x|} \leq C \quad \text{Und daher.} \quad \boxed{|f(x+\Delta x) - f(x)| \leq C|\Delta x|}$$

Die letzte Ungleichung liefert aber die Stetigkeit von f in x ! (Eine mögliche Strategiefunktion ist $\delta(\varepsilon) = \frac{1}{2C}\varepsilon$.)

(4.8.16) Das ist ein erstaunliches Resultat. Denn der ursprünglichen Definition von konvex war die Stetigkeit keineswegs anzusehen!

Satz: Es sei $f: I \rightarrow \mathbb{R}$ konvex.
Dann ist f in jedem inneren Punkt stetig und einseitig differenzierbar.

Beispiel: $x \mapsto |x|$ ist konvex und daher stetig. In $x=0$ ist f aber nicht differenzierbar. Man hat dort die einseitigen Limiten $f'(0-0) = -1$ und $f'(0+0) = +1$. Dabei ist $f'(0-0) \leq f'(0+0)$ natürlich erfüllt.

□ Sei $f: [0, \infty[\rightarrow \mathbb{R}$ mit $f(x) = \begin{cases} 1 & \text{für } x=0 \\ x^2 & \text{für } x^2 > 0 \end{cases}$. Zeigen Sie, dass f konvex ist. Was folgt aus dem Beispiel? Ist $g: \mathbb{R} \rightarrow \mathbb{R}$ mit $g(x) = \begin{cases} 1 & \text{für } x=0 \\ x^2 & \text{für } x^2 \neq 0 \end{cases}$ konvex?

(4.8.17) **Angenommen f ist in $x_1 \in I$ sogar differenzierbar.** D.h. die beiden einseitigen Ableitungen $f'(x_1+0)$ und $f'_1(x-0)$ sind gleich. Dann folgt aus (3.8.15) für geeignetes $\Delta x > 0$

$$\begin{aligned} f'(x) &\leq D(x, x+\Delta x) && \text{also } 0 \leq f(x+\Delta x) - f(x) - f'(x)\Delta x && \text{und} \\ D(x-\Delta x, x) &\leq f'(x) && \text{also } 0 \leq f(x-\Delta x) - f(x) - f'(x)(-\Delta x) \end{aligned}$$

Vergleicht man das mit der Tangenzenzerlegung, so steht rechts in beiden Fällen der absolute Fehler, also die Größe $\Delta x R_f(x, \Delta x)$.

Satz: Sei $f: [a, b] \rightarrow \mathbb{R}$ konvex und in x_0 mit $a < x_0 < b$ differenzierbar.
Dann gilt für den absoluten Fehler in der Tangentenapproximation
 $\boxed{F(x_0, \Delta x) = \Delta x R_f(x_0, \Delta x) \geq 0}$.

Also: Bei dem Graphen einer konvexen Funktion liegt die Tangente immer unterhalb des Graphen.

(4.8.18) Jetzt nehmen wir an, dass f überall differenzierbar ist.

Satz: Sei f in $[a,b]$ stetig und in $]a,b[$ differenzierbar.
 Dann ist f genau dann konvex, wenn die Ableitung in $]a,b[$ monoton wächst.

Beweis: i) Angenommen f ist konvex und differenzierbar. Dann folgt die behauptete Monotonie aus (3.8.15).

ii) Angenommen die Ableitung wächst monoton. Wir beweisen (K2) aus (3.8.5). Das beinhaltet die Konvexität. Seien $x_1, x_2, x \in I$ mit $x_1 < x < x_2$. Da f stetig ist, ist der Mittelwertsatz (3.7.3) auf $[x_1, x]$ und auf $[x, x_2]$ anwendbar. Also gibt es zwei Werte $\zeta_1 \in]x_1, x[$ und $\zeta_2 \in]x, x_2[$ mit $f'(\zeta_1) = D(x_1, x)$ und $f'(\zeta_2) = D(x, x_2)$. Die Monotonie von f' liefert die Ungleichung (K2).

□ Wie lautet (3.8.18) für konkav. Zeigen Sie damit, dass $x \mapsto \ln(x)$ konkav ist! Was ergibt dann die induktiv verallgemeinerte Ungleichung aus (3.8.2) für die Logarithmusfunktion? Beweisen Sie damit die AGM-Ungleichung.

(4.8.19) Es gibt noch eine weitere **gleichwertige Charakterisierung** von konvex:

Definition $f: D \rightarrow \mathbb{R}$, $x_0 \in D$ und $\sigma = \sigma_{x_0}$ Gerade mit Gleichung $\sigma(x) = f(x_0) + m(x - x_0)$.
 Dann heißt $\sigma = \sigma_{x_0}$ **Stützgerade der Funktion f in x_0** , wenn gilt
 $f(x) \leq \sigma_{x_0}(x)$ für alle $x \in D$. (und $f(x_0) = \sigma_{x_0}(x_0)$).

D.h. man beginnt nicht mit Sekanten, sondern verschiebt diese so lange, bis sie den Graphen nur noch berühren.

(4.8.20) Damit gilt:

Satz: $f:]a,b[\rightarrow \mathbb{R}$ ist genau dann konvex, wenn jeder Punkt eine Stützgerade besitzt.

Wir erinnern an eine geometrische Konstruktion, mit der man in vielen Fällen Tangenten einführen kann: Man verschiebt eine Gerade parallel, bis sie in genau einem Punkt die betrachtete Figur berührt. Dann ist der Graph lokal um diesen Punkt herum auf einer Seite der Tangente, also von der Art der hier eingeführten Stützgeraden. Aber eine eventuelle Tangente ist - wie wir wissen - eindeutig. Die Stützgerade ist da allgemeiner: Betrachte $x \mapsto |x|$ und $x_0 = 0$. Eine Tangente existiert nicht. Aber alle Geraden $y = mx$ mit $-1 \leq m \leq 1$ sind Stützgeraden!

(4.8.21) Beachten Sie, dass im Satz ein offenes Definitionsintervall gefordert wird. D.h. für jedes x daraus gibt es die rechtsseitige Ableitung $f'(x+0)$.

□ Vergleichen Sie das mit der Tangenzenzerlegung. Was ist analog und welcher Unterschied besteht!

Beweis: 1) **Ist f konvex**, dann hat f in jedem Punkt eine Stützgerade. Nehme dazu

$$\sigma(x) = f(x_0) + f'(x_0 + 0)(x - x_0).$$

Denn es gilt ($x_2 > x_0$):

$$\begin{aligned} f'(x_0 + 0) &\leq D(x_2, x_0) \\ \sigma_{x_0}(x_2) &= f(x_0) + f'(x_0 + 0)(x_2 - x_0) \leq f(x_0) + D(x_2, x_0)(x_2 - x_0) \\ &= f(x_0) + \frac{f(x_2) - f(x_0)}{x_2 - x_0}(x_2 - x_0) = f(x_2). \quad \text{D.h. } \boxed{\sigma_{x_0}(x_2) \leq f(x_2)} \end{aligned}$$

Analog behandelt man den Fall $x_1 < x_0$. D.h.: Es gibt zu jedem Punkt x_0 eine Stützgerade.

2) **Jetzt habe f zu jedem x_0 eine Stützgerade.** Wir benötigen den folgenden Hilfssatz, den wir unten beweisen

Hilfssatz Sei $\sigma(x) = b + mx$ und $x = \lambda_1 x_1 + \lambda_2 x_2$ mit $\lambda_1 + \lambda_2 = 1$
 Dann ist $\sigma(x) = \sigma(\lambda_1 x_1 + \lambda_2 x_2) = \lambda_1 \sigma(x_1) + \lambda_2 \sigma(x_2)$.

Wir zeigen jetzt die Konvexität über die ursprüngliche Definition, zeigen also

$$f(\lambda_1 x_1 + \lambda_2 x_2) \geq \lambda_1 f(x_1) + \lambda_2 f(x_2).$$

Dazu rechnen wir wie folgt, wobei in der zweiten Zeile der Hilfssatz benutzt wird:

$$\begin{aligned}
f(x) &= f(\lambda_1 x_1 + \lambda_2 x_2) = \sigma_x(\lambda_1 x_1 + \lambda_2 x_2) \quad \text{da } f(x) = \sigma_x(x)! \\
&= \lambda_1 \sigma_x(x_1) + \lambda_2 \sigma_x(x_2) \quad \text{nach } (*) \\
&\leq \lambda_1 f(x_1) + \lambda_2 f(x_2) \quad \text{nach Definition Stützfunktion.}
\end{aligned}$$

Zum Beweis des Hilfssatzes einfach die rechte Seite ausrechnen!

$$\begin{aligned}
\lambda_1 \sigma(x_1) + \lambda_2 \sigma(x_2) &= \lambda_1(b + mx_1) + \lambda_2(b + mx_2) \\
&= (\lambda_1 + \lambda_2)b + m(\lambda_1 x_1 + \lambda_2 x_2) \\
&= b + mx = \sigma(x)
\end{aligned}$$

(4.8.22) Ist f in x_0 differenzierbar, dann hat f in x_0 höchstens eine Stützgerade und das ist die Tangente.

Beweis: Die Stützgerade habe die Steigung m . Dann formen wir die Tangenzenzerlegung wie folgt um:

$$f(x_0 + \Delta x) - (f(x_0) + m\Delta x) = (f'(x_0) - m + R_f(x_0, \Delta x))\Delta x, \text{ wobei } R_f \text{ die Resttermeigenschaft besitzt.}$$

Sei $|f'(x_0) - m| = 2\varepsilon$. Angenommen $\varepsilon > 0$. Dann wählen wir Δx so klein, dass $|R_f(x_0, \Delta x)| < \varepsilon$ gilt. Dann ist für diese Δx sicher $|f'(x_0) - m + R_f(x_0, \Delta x)| > \varepsilon > 0$. Und damit wechselt $f(x_0 + \Delta x) - (f(x_0) + m\Delta x)\Delta x$ für $\Delta x = 0$ das Vorzeichen, was bei einer Stützgeraden nicht zulässig ist. Die einzige eventuelle Ausnahme bildet $m=f'(x_0)$.

Ist f differenzierbar mit positivem absoluten Fehler, dann ist die Tangente in jedem Punkt Stützgerade (wie die Tangenzenzerlegung sofort zeigt. Und damit ist f konvex.

□ Wie können Sie jetzt den Satz aus (3.8.17) ergänzen?

2.4.8a Konvexe Mengen

Ihrer Natur nach sind *konvex* und *konkav* Begriffe, die sich auf die Form von Figuren, also Teilmengen eines Konfigurationsraumes beziehen. Denken Sie an *konvexe und konkave Linsen*. Tatsächlich liegt es nahe die Definition von *konvex* für Vektorräume über \mathbb{R} zu formulieren und dann die konvexen Funktionen gleichwertig dadurch zu definieren, dass eine bestimmte aus dem Graphen gebildete Figur konvex sein soll.

Die zu wählende Figur für $f: I \rightarrow \mathbb{R}$ ist $K_f = \{(x, y) | x \in I, y \geq f(x)\}$. D.h. sie besteht aus allen Punkten *oberhalb des Graphen*. Diese Funktion soll im Sinne der Geometrie konvex sein. Überlegen Sie sich die Form dieser Figur für $f = ([-1, 1], x \mapsto x^2, \mathbb{R})$. Dies f war konvex.

Und nun die rein vektorielle Definition von konvex:

Sei V Vektorraum über \mathbb{R} und $K \subset V$ Teilmenge. Dann heißt K konvex, wenn folgende Eigenschaft erfüllt ist:

Sei $\vec{a}, \vec{b} \in K$. Bilde dazu die Verbindungsstrecke $S_{\vec{a}\vec{b}} = \{\vec{a} + \lambda(\vec{b} - \vec{a}) | 0 \leq \lambda \leq 1\}$. Dann gilt auch

$$S_{\vec{a}\vec{b}} \subset K. \text{ Kurz: } \boxed{(\vec{a}, \vec{b} \in K) \implies (S_{\vec{a}\vec{b}} \subset K)}$$

Kurz: Mit zwei Punkten gehört auch die gesamte Verbindungsstrecke zu K . Diese Verbindungsstrecke ist in jedem Vektorraum über \mathbb{R} erklärt.

Und *konkave Funktion* bedeutet, dass die Menge $U_f = \{(x, y) | x \in I, y \leq f(x)\}$ konvex sein soll.

Der Beweis der Gleichwertigkeit verlangt natürlich noch die Ausführung einiger Einzelheiten.