

# Mathematik II

## Hinweise zum Skript und Klausuraufgaben

### Sommersemester 2024

#### Inhalt

Affine Abbildungen, affine Räume und Unterräume.....	3
Lineare Gleichungssysteme (LGS).....	4
Gleichungssysteme und lineare Abbildungen .....	7
Lösbarkeit und Lösungen von Gleichungssystemen.....	9
Zusammenfassung Lösbarkeit (mit kleinen Ergänzungen): .....	11
Fallunterscheidung beim Lösen eines LGS.....	11
Skalarprodukte und Abstände .....	12
Bilinearformen.....	12
Normen und Metriken.....	16
Determinanten und das Eigenwertproblem.....	20
Eigenwerte, Eigenvektoren, Eigenraum.....	23
Bestimmung der Eigenwerte, Eigenvektoren und Eigenraum.....	24
Spektrum und weitere Begriffe .....	28
Nützliche Eigenschaften im Zusammenhang mit Eigenwerten .....	28
Diagonalisierung von Matrizen.....	29
Typisches Vorgehen zum Diagonalisieren einer $n \times n$ -Matrix $A$ .....	30
Mathematische Grundlagen / Theorie aus dem Skript zum Diagonalisieren.....	32
Stetige Funktionen.....	33
Funktionsgrenzwerte .....	35
Stetigkeit.....	37
Differenzialrechnung .....	41
Ableitungen .....	41
Die Mittelwertsätze der Differenzialrechnung .....	43
Der Satz von Taylor .....	45
Partielle Ableitungen .....	49
Integralrechnung .....	50
Stammfunktionen.....	50
Rationale Funktionen .....	54
Vorgehen bei der Partialbruchzerlegung.....	63
Das Riemannsche Integral .....	70
Klassen integrierbarer Funktionen .....	71

Eigenschaften des Riemann-Integrals .....	72
Uneigentliche Integrale .....	73
Fouriertransformation .....	75
Elementare Differentialgleichungen.....	77
Motivation .....	77
Lineare Differentialgleichungen erster Ordnung.....	79
Lineare Differentialgleichungen erster Ordnung - Exkurs .....	83
Getrennte Variablen .....	85
Separation der Variablen- Exkurs .....	86
$e\lambda$ -Ansatzes - Exkurs .....	88
Klausuraufgaben .....	90

## Affine Abbildungen, affine Räume und Unterräume

Es geht schon mal gleich unschön in Skript los: Plötzlich ist von affinen Unterräumen die Rede, die erst später definiert werden. Der Begriff „affin“ ist mir bekannt, auch wenn ich jetzt nicht spontan eine Definition angeben könnte, aber es hat etwas mit „Verschiebung“ zu tun. Diese Grundbedeutung solltest du dir auch erstmal merken. Die Ausführungen hier sind vermutlich mathematisch nicht ganz sauber, aber ich hoffe anschaulicher als das Skript.

Etwas deutlicher ist mir der Begriff „affine“ im Zusammenhang mit Abbildungen geblieben. Du hast im letzten Semester lineare Abbildungen kennen gelernt, die man auch mit Hilfe von Abbildungsmatrizen (Darstellungsmatrizen) beschreiben und analysieren kann. Erinnere dich dabei mal an die Spiegel- oder Drehmatrizen für Abbildungen im  $\mathbb{R}^2$ . Es ist so möglich, Punkte in der Ebene entsprechend abzubilden, d. h. zu einem Urbildpunkt einen entsprechenden Bildpunkt zu berechnen. Was lineare Abbildungen (ohne Tricks) nicht können, sind Verschiebungen:

Eine lineare Abbildung  $F$  kann durch  $F(x) = A \cdot x$  beschrieben werden. Eine affine Abbildung  $G$  hingegen kann beschrieben werden:  $G(x) = B \cdot x + t$ , wobei  $x$  und  $t$  für Vektoren aus dem passenden Koordinatenraum sind.  $t$  nennt man hier Translationsvektor. Er sorgt für eine Verschiebung. Es ist auch möglich, die Verschiebung in der Abbildungsmatrix einzubauen, aber dazu muss man auch den Vektor manipulieren. Das ist zwar jetzt nicht so wichtig, aber der Trick geht so: Die Matrix  $B$  wird wie folgt erweitert:

$$\tilde{B} = \begin{pmatrix} B & t \\ 0^T & 1 \end{pmatrix}$$

$0^T$  steht für den transponierten Nullvektor. Parktisch ergänzt man die Matrix  $B$  um eine Nullzeile mit einer 1 am Ende und fügt  $t$  als weitere Spalte hinzu. Beim Abbilden müssen aber auch die Vektoren  $x$  um eine weitere Komponente (Koordinate) ergänzt werden:

$$\tilde{x} = \begin{pmatrix} x \\ 1 \end{pmatrix}$$

Die affine Abbildung berechnet sich dann mit einer einzigen Matrixmultiplikation  $\tilde{B} \cdot \tilde{x} = \tilde{x}'$ , wobei beim Bild  $\tilde{x}'$  (verglichen mit  $B \cdot x + t = x'$ ) die letzte Komponente für die ursprüngliche Abbildung ignoriert werden kann. Ein Beispiel ist hier auf Seite 2 zu finden: [https://www.physastromath.ch/uploads/myPdfs/LinAlg/LinAlg\\_07.pdf](https://www.physastromath.ch/uploads/myPdfs/LinAlg/LinAlg_07.pdf)

Es gibt noch eine ganze Menge zu affinen Abbildungen sagen, z. B. über die erweiterte Abbildungsmatrix, über Fixpunkte und Fixgeraden (also Punkte oder Mengen von Punkten), die beim Abbilden auf sich selbst abgebildet werden. Hier kann man dazu einiges lesen: [https://de.wikipedia.org/wiki/Affine\\_Abbildung](https://de.wikipedia.org/wiki/Affine_Abbildung)

Nennenswert in dem Zusammenhang ist noch die Scherung (Abbildung). Mehr dazu hier: [https://de.wikipedia.org/wiki/Scherung\\_\(Geometrie\)](https://de.wikipedia.org/wiki/Scherung_(Geometrie))

Das Skript beschränkt sich (zunächst) auf affine Unterräume und schreibt fast beiläufig auf Seite 223:

**Bemerkung 7.4.7** (Definition). Ist  $V$  ein Vektorraum,  $U$  ein Unterraum von  $V$  und  $v \in V$ , so heißt

$$U' = U + v := \{ u + v \mid u \in U \}$$

ein **affiner Unterraum**. Wir setzen  $\dim U' := \dim U$ . Es gilt also: Die Lösungsmenge eines lösbar inhomogenen Gleichungssystems  $Ax = b$  ist ein affiner Unterraum der Dimension  $n - \text{rg } A$  (im Beispiel: 2).

(In dem Kapitel geht es um lineare Gleichungssystem.) Bleiben wir zunächst bei den Begrifflichkeiten. Wenn wir einen Vektorraum  $V$  und einen Unterraum  $U$  von  $V$  haben, dann erhält man mit einem festen  $v \in V$  den **affinen Unterraum**  $U' = U + v = \{u + v \mid u \in U\}$ . Es ist praktisch so, dass alle Elemente von  $U$  durch die Addition mit dem Vektor  $v$  verschoben werden. Das Ergebnis ist in der Regel kein Unterraum (da das Nullelement oft verloren geht), aber anschaulich klar und nützlich.

**Beispiel:** Betrachte den Koordinatenraum  $\mathbb{R}^3$  wie in der Schule. Der zugehörige Vektorraum hat die Dimension 3. Die  $x_1$ - $x_2$ -Ebene (eigentlich die Ortsvektoren zu den Punkten in der Ebene) stellt einen Unterraum dar. Du kannst das gerne mit der Definition aus Mathematik I überprüfen. Wenn du zu al-

len Vektoren dieses Unterraums den Vektor  $\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$  addierst, erhältst du einen affinen Unterraum. Die zugehörige Ebene (bzw. die Ortsvektoren ...), die parallel zur  $x_1$ - $x_2$ -Ebene verläuft, stellen keinen Unterraum dar, da das Nullelement nicht mehr enthalten ist. Natürlich kann man dennoch mit dem affinen Unterraum allerhand anfangen und ihn anschaulich darstellen.

Es wird so selbstverständlich über diese Dinge gesprochen, dass die Genauigkeit bei der Ausdrucksweise schnell leidet (kann auch mir passieren!). **Beachte:** Wenn wir den Koordinatenraum zur Illustration eines Vektorraums verwenden, dann meinen wir dennoch mehr als „Punkte im Raum“, sondern eben einen Vektorraum über einen Körper (vgl. Def. 5.1.1). Eine Ebene oder eine Gerade ist nur eine Menge von Punkten – mehr nicht. Wenn man hier von (Unter)Vektorräumen sprechen will, dann braucht man wieder alles, was einen solchen Raum ausmacht (Definition beachten), also auch die beiden Verknüpfungen (meist Addition und Skalarmultiplikation).

Mag sein, dass nicht mehr als die Bemerkung 7.4.7. (Definition) benötigt wird, da im Skript wohl immer ausgehend von einem Vektorraum ein affiner Unterraum betrachtet wird. Ein affiner Raum kann auch direkt eingeführt bzw. definiert werden:

[https://de.wikipedia.org/wiki/Affiner\\_Raum](https://de.wikipedia.org/wiki/Affiner_Raum)

## Lineare Gleichungssysteme (LGS)

Lineare Gleichungssysteme spielen in der Mathematik eine wichtige Rolle, da sie z. B. bei vielen Problemstellungen eine Rolle spielen.

Kleine LGS löst man in der Schule mit dem Einsetzungs-, Gleichsetzungs- oder Additionsverfahren. Ziel es dabei das LGS schrittweise auf eine Gleichung mit einer Unbekannten zu reduzieren mit dieser Teillösung wiederum die übrigen Unbekannten schrittweise zu berechnen. Für größere LGS (und besonders wenn man diese händisch lösen muss) bietet sich das Additionsverfahren an, dass näher formalisiert als Gaußsche Eliminationsverfahren beschrieben wird. Im Skript werden dafür viele, viele Seiten verwendet: Seiten 202-215. Hier wird alles gaaaanz ausführlich beschrieben.

Ich verwende immer direkt die erweiterte Koeffizientenmatrix<sup>1</sup> und folgendes Schema:

Das **Gaußsche Eliminationsverfahren**, benannt nach dem deutschen Mathematiker Carl Friedrich Gauß, ist eine Methode zur Lösung von linearen Gleichungssystemen. Diese Methode wird häufig in der linearen Algebra angewendet und ermöglicht es, ein System von linearen Gleichungen in eine äquivalente Form zu transformieren, die leicht gelöst werden kann.

---

<sup>1</sup> Leider finde ich das Skript nicht so toll, weil z. B. die erweiterte Koeffizientenmatrix erst später, auf Seite 219 in Definition 7.4.1 eingeführt.

Hier ist eine Beschreibung des Gaußschen Eliminationsverfahrens:

1. **Schritt 1: Aufstellen der erweiterten Koeffizientenmatrix:** Das lineare Gleichungssystem wird zunächst in Form einer erweiterten Koeffizientenmatrix dargestellt. Diese Matrix enthält die Koeffizienten der Unbekannten sowie die Ergebnisse der Gleichungen auf der rechten Seite. (Natürlich muss man sich merken, in welcher Spalte die Koeffizienten zu welcher Unbekannten stehen.)
2. **Schritt 2: Umformung der erweiterten Matrix in eine obere Dreiecksmatrix:** Durch elementare Zeilenoperationen (wie Addition, Subtraktion und Multiplikation von Zeilen, Zeilentausch) wird die erweiterte Koeffizientenmatrix schrittweise in eine obere Dreiecksmatrix umgeformt. Dabei wird darauf geachtet, dass die Hauptdiagonale (von links oben nach rechts unten) nur aus Nichtnull-Einträgen besteht und dass unterhalb der Hauptdiagonale nur Nullen stehen.
3. **Schritt 3: Rückwärtssubstitution:** Nachdem die erweiterte Koeffizientenmatrix in obere Dreiecksform gebracht wurde, können die Lösungen der Unbekannten durch Rückwärtssubstitution berechnet werden. Dabei wird von der letzten Gleichung ausgehend die Lösung für die letzte Unbekannte bestimmt und dann schrittweise zurückgearbeitet, um die Lösungen für die vorherigen Unbekannten zu finden. Im Beispiel oben kann z. B. aus der letzten Zeile abgelesen werden, dass das dreifache der dritten Variable 9 ergibt, also  $3x_3 = 9$ . Also ist  $x_3 = 3$ . Dieses kann in der vorigen Zeile eingesetzt werden und man erhält  $-1 \cdot x_2 + 3 \cdot 3 = 11$ . Auch diese Gleichung lässt sich lösen. Es gilt:  $x_2 = -2$ . Und so weiter.

Beispiel einer (nicht normierten) oberen Dreiecksmatrix:

$$\begin{pmatrix} 2 & -2 & 6 & 30 \\ 0 & -1 & 3 & 11 \\ 0 & 0 & 3 & 9 \end{pmatrix}$$

Andere Bezeichnungen: Stufenform, Zeilenstufenform, Treppenform, ...

Alternativ zu Nr. 3 kann man durch vergleichbare Umformungen, wie bei Nr. 2, auch die Matrix so umformen, dass nur noch Einsen auf der Hauptdiagonalen stehen, der Rest Nullen sind und in der letzten Spalte die Lösungen der Unbekannten stehen. Ggf. entsteht am Ende auch ein oder mehrere Nullzeilen. Man nennt die fertige Matrix auch reduzierte (oder normierte) Zeilenstufenform und sie sieht z. B. so aus:

$$\begin{pmatrix} 1 & 0 & 0 & 4 \\ 0 & 1 & 0 & -2 \\ 0 & 0 & 1 & 3 \end{pmatrix}$$

Hier kann man die Lösungen direkt ablesen, also hier im Beispiel  $x_1 = 4$ ,  $x_2 = -2$  und  $x_3 = 3$ .

Wählt man diesen alternativen Weg, so spricht man vom Gauß-Jordan-Verfahren. Natürlich kann man bei solchen LGS auch ganz systematisch vorgehen, d. h. nach einem festen Algorithmus die obere Dreiecksgestalt bzw. die normierte Zeilenstufenform berechnen.

Wichtig ist noch die Deutung der Matrix nach dem Umformen:

1. Woran erkennt man, dass ein LGS unlösbar ist?
  - Wenn mind. eine Nullzeile auftritt, bei der in der letzten Spalte eine Zahl ungleich Null steht. Das sieht also so aus:

$$\begin{pmatrix} 1 & 0 & 0 & 4 \\ 0 & 1 & 0 & -2 \\ 0 & 0 & 0 & 4 \end{pmatrix}$$

Die letzte Zeile würde bedeuten  $0 = 4$ , was falsch ist. Daher ist das LGS nicht lösbar.

2. Wann gibt es unendlich viele Lösungen?

- Wenn mind. eine reine Nullzeile auftritt. Diese bedeutet nämlich, dass der Wert der letzten Unbekannten beliebig ist. Das sieht also so aus:

$$\begin{pmatrix} 1 & 0 & 0 & 4 \\ 0 & 1 & 0 & -2 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

Im Skript wird nebenbei gezeigt, wie man die Inverse zu einer Matrix bestimmen kann. Darauf gehe ich an dieser Stelle nicht näher ein und verweise auf S. 46 in meinem Miniskript vom ersten Semester.

## Gleichungssysteme und lineare Abbildungen

Zu Erinnerung:

### Definition 6.3.1: Bild und Kern

Es seien  $V$  und  $W$   $\mathbb{K}$ -Vektorräume und  $F : V \rightarrow W$  eine lineare Abbildung, dann heißt der Unterraum

$$\text{Bild } F := F(V) = \{ F(v) \mid v \in V \}$$

von  $W$  das *Bild* von  $F$  (vergleiche 1.1.8). Die Menge

$$\text{Ker } F := \{ v \in V \mid F(v) = 0 \}$$

heißt der *Kern* von  $F$ .

An den Unterräumen  $\text{Bild } F$  und  $\text{Ker } F$  lässt sich ablesen, ob die Abbildung  $F$  surjektiv beziehungsweise injektiv ist:

### Lemma 6.3.3

Sind  $V$  und  $W$   $\mathbb{K}$ -Vektorräume und ist  $F : V \rightarrow W$  linear, dann ist

- (i)  $F$  genau dann surjektiv, wenn  $\text{Bild } F = W$ .
- (ii)  $F$  genau dann injektiv, wenn  $\text{Ker } F = \{ 0 \}$ .

[https://y-](https://youtu.be/589YkgeEoiA?si=bj2RHTSHvp4-fLgF)

[outu.be/589YkgeEoiA?si=bj2RHTSHvp4-fLgF](https://youtu.be/589YkgeEoiA?si=bj2RHTSHvp4-fLgF)

### Definition 6.3.4: Rang und Defekt

Die Dimensionen von  $\text{Bild } F$  und  $\text{Ker } F$  erhalten besondere Namen:

- (i)  $\text{rg } F := \dim \text{Bild } F$ , der *Rang* von  $F$  und
- (ii)  $\text{def } F := \dim \text{Ker } F$ , der *Defekt* von  $F$ .

### Satz 6.3.6: Dimensionsformel für lineare Abbildungen

Ist  $V$  ein endlich-dimensional,  $W$  ein beliebiger  $\mathbb{K}$ -Vektorraum und  $F : V \rightarrow W$  linear, so gilt

$$\dim V = \dim \text{Ker } F + \dim \text{Bild } F = \text{def } F + \text{rg } F.$$

## Rang einer Matrix:

Der **Spaltenrang** einer Matrix sagt dir, wie viele linear unabhängige Spaltenvektoren du in der Matrix maximal finden kannst. Die maximale Anzahl linear unabhängiger Zeilenvektoren ist der **Zeilenrang**. **In jeder Matrix sind Zeilenrang und Spaltenrang gleich.** Deshalb sprichst du oft nur vom **Rang einer Matrix**. (Daher ändert das Transponieren einer Matrix, also das Vertauschen von Spalten und Zeilen, nichts am Rang der Matrix. Anschaulich entsteht die transponierte Matrix durch Spiegelung der Ausgangsmatrix an ihrer Hauptdiagonale. Viele Kenngrößen von Matrizen, wie Spur, Rang, Determinante und Eigenwerte, bleiben unter Transponierung erhalten. Vertiefung hier: [https://de.wikipedia.org/wiki/Transponierte\\_Matrix](https://de.wikipedia.org/wiki/Transponierte_Matrix))

Der Rang einer beliebigen  $m \times n$  Matrix  $B$  ist immer kleiner als oder gleich groß wie das Minimum aus Zeilenanzahl und Spaltenanzahl, also  $\text{Rang}(B) \leq \min(m, n)$ .

Das stammt von <https://studyflix.deg/mathematik/rang-einer-matrix-1785> wo dies auch mit Video und Beispielen beschrieben wird.

Auch das Nachfolgevideo ist gut: <https://studyflix.de/mathematik/kern-einer-matrix-1788>

Warum sprechen wir jetzt so viel vom Rang der Matrix, wenn es doch um den Rang der Abbildung geht?

**Eine lineare Abbildung und die zugehörige Abbildungsmatrix besitzen denselben Rang, d. h. die Dimension des Bildes entspricht den linear unabhängigen Spalten bzw. Zeilen.**

Vgl. Definition 7.3.2. im Skript:

### Satz 7.3.1: Isomorphismen sind rangerhaltend

Es seien  $V$  und  $W$   $\mathbb{K}$ -Vektorräume,  $F : V \rightarrow W$  linear und  $G : W \rightarrow W$  ein Isomorphismus, dann gilt

$$\text{rg } G \circ F = \text{rg } F.$$

Das ist verständlich: Ein Isomorphismus ist ja eine bijektive, also umkehrbare Abbildung. Hier gehen keine Informationen verloren. Es wird nur etwas verzerrt, umgeordnet, etc.

Das habe wir eben schon festgehalten.

Diesen Isomorphismus nutzen wir ständig. Die Koordinatenschreibweise ist eben sehr praktisch. Die Dimensionen (Rang) müssen dann auch zusammenpassen.

### Lemma 7.3.3

Der Bildraum von  $\tilde{A} = F_{\tilde{A}}$  ist isomorph zum  $\mathbb{K}^r$ , in Zeichen

$$\text{Bild } \tilde{A} = \text{Bild } F_{\tilde{A}} \cong \mathbb{K}^r,$$

weshalb  $\text{rg } A = \text{rg } \tilde{A} = r$ .

In den Videos wird zusätzlich noch der Zusammenhang zwischen den Spaltenvektoren der Matrix und dem Bild betont!

### Definition 7.3.4: Zeilen- und Spaltenrang

So eben wurde gezeigt, dass der Rang einer Matrix  $A$  gleich der Dimension des von den Spaltenvektoren erzeugten Unterraums von  $\mathbb{K}^m$  ist; wir bezeichnen ihn daher auch als **Spaltenrang** von  $A$ . Analog erklären wir den **Zeilenrang** von  $A$  als die Dimension des von den Zeilen von  $A$  im  $\mathbb{K}^n$  aufgespannten Unterraumes.

### Satz 7.3.5: Zeilenrang=Spaltenrang

Zeilenrang und Spaltenrang einer Matrix sind gleich.

Besonders interessant für quadratische Matrizen und der Frage, ob sie invertierbar ist.

## Lösbarkeit und Lösungen von Gleichungssystemen

Im Skript folgt nun erstmal die Definition von:

### Definition 7.4.1: Bezeichnungen in Gleichungssystemen

Ein lineares Gleichungssystem, in dem alle Koeffizienten  $b_i$  der rechten Seite gleich 0 sind heißt *homogen*. Ist wenigstens ein  $b_i \neq 0$ , so heißt das Gleichungssystem *inhomogen*. Es sei (I) das lineare Gleichungssystem

$$\begin{aligned} a_{11}x_1 + \cdots + a_{1n}x_n &= b_1, \\ &\vdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n &= b_m. \end{aligned} \tag{I}$$

Ferner sei (H) das zu (I) gehörige homogene Gleichungssystem

$$\begin{aligned} a_{11}x_1 + \cdots + a_{1n}x_n &= 0, \\ &\vdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n &= 0. \end{aligned} \tag{H}$$

Mit  $A$  bezeichnen wir die *Koeffizientenmatrix*

$$\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix},$$

mit  $(A, b)$  die *erweiterte Koeffizientenmatrix*

$$\begin{pmatrix} a_{11} & \cdots & a_{1n} & b_1 \\ \vdots & & \vdots & \vdots \\ a_{m1} & \cdots & a_{mn} & b_m \end{pmatrix}.$$

Dazu gehört der Satz 7.4.2 – beachte die Bezüge auf die Gleichungssystem (I) und (H):

### Satz 7.4.2: Lösbarkeit und Darstellung der Lösungen

(i) *Rangkriterium*: Die Gleichung (I) ist genau dann lösbar, wenn

$$\operatorname{rg} A = \operatorname{rg}(A, b).$$

Diese vielen Sätze sind ohne Beispiele etwas schwer zu verstehen.

Zu (i): Der Rang von  $A$  soll dem Rang von  $(A, b)$  entsprechen. Wann ist das nicht der Fall? Wenn bei der Umformung in eine Zeilenstufenform die Matrix  $A$  eine Nullzeile aufweist, also nicht den vollen Rang hat,

- (ii) Ist  $x = (x_1, \dots, x_n)$  eine Lösung von (I), so durchläuft  $x + z$  alle Lösungen von (I), falls  $z = (z_1, \dots, z_n)$  alle Lösungen von (H) durchläuft.  
(iii) Die Lösungen von (H) bilden einen Unterraum des  $\mathbb{K}^n$  der Dimension  $n - \operatorname{rg} A$ .

aber die erweiterte Matrix in der letzten Spalte, in der entsprechenden Zeile nicht Null ist, also der Rang mindestens um eins größer hat als die Matrix  $A$ . Siehe Bsp. oben auf Seite 4. Es kann sein, dass die erweiterte Matrix eine durchgehende Nullzeile hat. Dann ist ihr Rang wiederum wie der von Matrix  $A$  und das zugehörige LGS ist lösbar, wenn auch nicht eindeutig lösbar.

Bei den Teilen (ii) und (iii) fehlt mir gerade der praktische Bezug, aber ich denke es geht im Kern um die Frage wann eine Lösung eindeutig ist bzw. was über eine mehrdeutige Lösung zu sagen ist. Zu (ii): Im Skript wird dieser Teil durch folgendes Lemma bewiesen:

**Lemma 7.4.3: Lösungsmenge der inhomogenen linearen Gleichung**

Sind  $V$  und  $W$   $\mathbb{K}$ -Vektorräume,  $F : V \rightarrow W$  eine lineare Abbildung von  $V$  nach  $W$ ,  $w \in W$  und  $v_s \in V$  eine Lösung von

$$F(v_s) = w, \quad (\text{I}^*)$$

dann durchläuft  $u + v_s$  alle Lösungen von  $(\text{I}^*)$ , falls  $u \in V$  alle Lösungen von

$$F(u) = 0 \quad (\text{H}^*)$$

durchläuft.

Meine Deutung: Wenn  $v_s$  eine Lösung vom  $(\text{I}^*)$  ist und wenn  $u$  eine Lösung von  $(\text{H}^*)$  ist, dann ist natürlich  $u + v_s$  auch eine Lösung von  $(\text{I}^*)$ , denn es gilt:

$$\begin{aligned} F(u + v_s) &= F(u) + F(v_s) \\ &= 0 + F(v_s) = F(v_s) = w \end{aligned}$$

Sind  $v_s$  und  $v$  beides Lösungen von  $(\text{I}^*)$ , dann ist  $v - v_s$  auch

eine Lösung von  $(\text{H}^*)$ , denn es gilt  $F(v - v_s) = F(v) - F(v_s) = w - w = 0$  und man kann  $u := v - v_s$  definieren  $v$  auch schreiben als  $v = u + v_s$ . Warum sind diese Aussagen bedeutsam?

Hmm... Anwendung? Ich sehe gerade nichts. Vielleicht kommt dazu noch etwas auf einem Übungsblatt.

Wir betrachten weiter eine Abbildung  $F : V \rightarrow W$  mit und betrachten  $V = \mathbb{K}^n$  und  $W = \mathbb{K}^m$ .

**Satz 7.4.4: Universelle und eindeutige Lösbarkeit**

- (i) Das inhomogene Gleichungssystem  $(\text{I})$  ist genau dann *universell lösbar*, das heißt es ist für alle  $b \in \mathbb{K}^m$  lösbar, wenn  $\text{rg } A = m$  gilt.
- (ii) Das homogene Gleichungssystem  $(\text{H})$  besitzt genau dann nur die Nulllösung, wenn  $\text{rg } A = n$  ist. Genau in diesem Fall besitzt das inhomogene Gleichungssystem  $(\text{I})$  höchstens eine Lösung, weshalb man etwas ungenau sagt, dass in diesem Fall  $(\text{I})$  *eindeutig lösbar* ist.

Universell lösbar heißt also, dass das LGS für jeden beliebigen Spaltenvektor  $b \in \mathbb{K}^m$  lösbar ist.

Wenn das homogene LGS nur die Nulllösung hat, hat das inhomogene LGS höchstens eine Lösung, da man mit  $u = 0$  keine weiteren Lösungen angeben

oder konstruieren kann (vgl. Lemma 7.4.3 oben).

## Zusammenfassung Lösbarkeit (mit kleinen Ergänzungen):

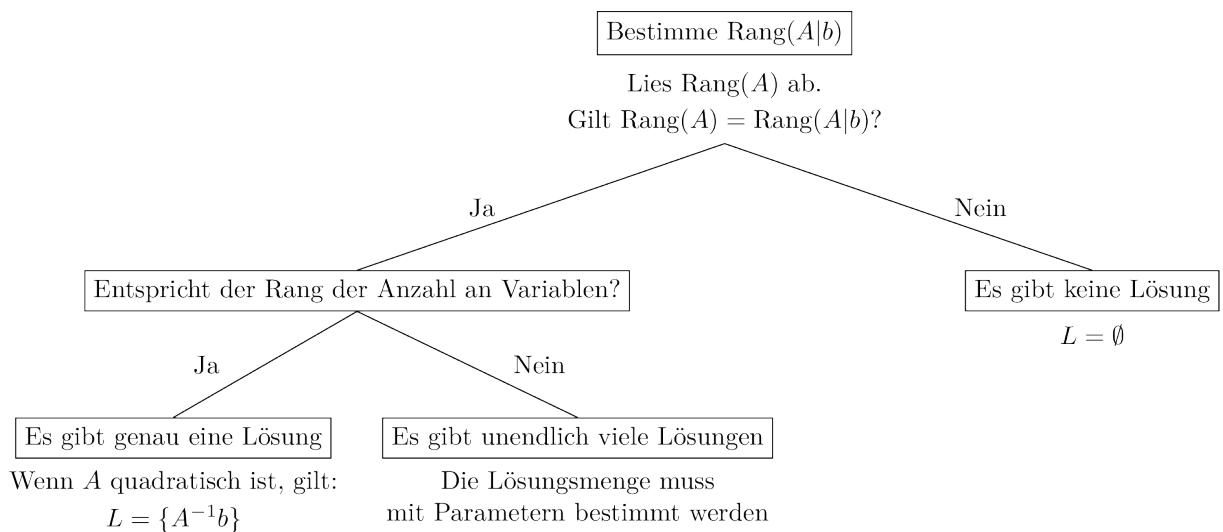
Sei  $Ax = b$  ein lineares Gleichungssystem mit einer  $m \times n$ -Matrix. Ferner sei  $F: \mathbb{K}^n \rightarrow \mathbb{K}^m$ ,  $x \mapsto Ax$  die zu  $A$  gehörige Abbildung.

Dann gilt:

- **Lösbarkeit:** Folgende Aussagen sind äquivalent:
  - Das lineare Gleichungssystem  $Ax = b$  ist lösbar.
  - $\text{rang } A = \text{rang } (A | b)$
  - $b \in \text{Bild } F$
- **Universelle Lösbarkeit** (Das Gleichungssystem ist für jedes  $b \in \mathbb{K}^m$  lösbar). Folgende Aussagen sind äquivalent:
  - Das lineare Gleichungssystem  $Ax = b$  ist universell lösbar.
  - $\text{Rang } A = m$
  - $F$  ist surjektiv
- **Eindeutige Lösbarkeit** (Das Gleichungssystem  $Ax = b$  besitzt genau eine Lösung.) Unter Voraussetzung der Lösbarkeit von  $Ax = b$ , also  $\text{rang } A = \text{rang } (A | b)$ , sind folgende Aussagen äquivalent:
  - Das lineare Gleichungssystem  $Ax = b$  ist eindeutig lösbar.
  - $\text{Rang } A = n$
  - $F$  ist injektiv
  - $\text{Ker } F = \{0\}$

## Fallunterscheidung beim Lösen eines LGS

1. **Bestimmung des Rangs einer Matrix:** Die maximale Anzahl linear unabhängiger Spalten- bzw. Zeilenvektoren heißt Rang der Matrix. Sofern nicht erkennbar, die Matrix in Zeilenstufenform umwandeln. Dann sieht man es gleich: Anzahl der Nicht-Nullzeilen = Rang der Matrix. Sonderfall: Der Rang einer regulären (also invertierbaren) Matrix entspricht der Zeilen- bzw. Spaltenzahl der Matrix.
2. Die **Anzahl der Variablen** entspricht der Anzahl der Spalten von A.



Zum Bestimmen der Lösung ( $A|b$ ) in Stufen- oder reduzierter (normierter) Stufenform bringen und Lösungen durch Rückwärtseinsetzen bestimmen bzw. ablesen.

# Skalarprodukte und Abstände

## Bilinearformen

### Definition 8.1.1: Linearform

Eine **Linearform** auf einem  $\mathbb{K}$ -Vektorraum  $V$  ist eine lineare Abbildung  $\varphi : V \rightarrow \mathbb{K}$ . Die Menge aller Linearformen auf  $V$  bilden einen Vektorraum<sup>a</sup>  $V^* := \text{Hom}(V, \mathbb{K})$ , welcher der zu  $V$  *duale Vektorraum* heißt.

<sup>a</sup>Satz 6.2.7.

Also halten wir fest: Eine **Linearform** ist eine lineare Abbildung, die einen Vektor auf ein Körperelement, d. h. in der Regel auf eine Zahl abbildet. Aus einem Vektor wird eine Zahl? So etwas gibt es auch in der Schule, z. B. wenn man die Länge (Norm) eines Vektors berechnet. Dort würde man fol-

gendes schreiben: Im  $\mathbb{R}^3$  ist die Länge eines Vektors  $\vec{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$  definiert als:  $|\vec{x}| := \sqrt{x_1^2 + x_2^2 + x_3^2}$ .

Einem Vektor wird dadurch eine Zahl zugeordnet, z. B. dem Vektor  $\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$  die Zahl  $\sqrt{2}$ . Allerdings handelt es hier nicht um eine lineare Abbildung, da die Additivität nicht erfüllt ist. Die Norm eines Vektors ist also keine Linearform.

In Skript wird für eine Linearform ein Beispiel angegeben, welches jedoch ohne tieferen Sinn daherkommt:

**Beispiele 8.1.3.** (i) Die Abbildung  $\varphi : \mathbb{R}^3 \rightarrow \mathbb{R}$ ,  $\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \mapsto 2x_1 - x_2 + 3x_3$ , ist eine Linearform, die Matrix von  $\varphi$  bezüglich der kanonischen Basen im  $\mathbb{R}^3$  und  $(\mathbb{R}^3)^*$  ist  $(2, -1, 3)$ . Dies rechnet man leicht selbst nach.

Es wird keine Motivation angegeben, warum die Komponenten (Koordinaten) gerade durch  $2x_1 - x_2 + 3x_3 = x'$  auf eine reelle Zahl abgebildet wird. Aber ist hoffentlich klar, dass es wirklich eine lineare Abbildung ist, also Additivität und Homogenität erfüllt sind. Die Matrix  $A$  zu dieser Abbildung mit der Abbildung  $Ax = x'$  ist eine Matrix mit nur einer Zeile. Hier ist:

$$Ax = (2 \quad -1 \quad 3) \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = 2x_1 - x_2 + 3x_3 = x'$$

Erinnerung zur Matrixmultiplikation „Zeile mal Spalte, Zeile mal Spalte, ...“

Zum **dualen Vektorraum**: Betrachtet man die Menge aller Linearformen eines Vektorraums  $V$  über einen Körper  $\mathbb{K}$ , so bilden diese mit der Addition und der Skalarmultiplikation selbst einen Vektorraum, d. h. insbesondere gelten die ganzen Verträglichkeitsbedingungen und Distributivgesetze. Das Thema geht noch viel weiter und auch in Gebiete der Physik hinein ( $\rightarrow \text{https://de.wikipedia.org/wiki/Dualraum}$ ). Ich vermute aber, dass das hier mehr der Vollständigkeit halber angesprochen wurde.

#### Definition 8.1.4: Bilinearform

Es seien  $V$  und  $W$   $\mathbb{K}$ -Vektorräume, dann ist eine *Bilinearform* auf  $V \times W$  eine Abbildung

$$B : V \times W \rightarrow \mathbb{K}, (v, w) \mapsto B(v, w),$$

die für jedes feste  $v \in V$  beziehungsweise jedes feste  $w \in W$  eine Linearform auf  $W$  beziehungsweise  $V$  ist. Das heißt, für alle  $v, v_1, v_2 \in V$  und alle  $w, w_1, w_2 \in W$  und alle  $\lambda \in \mathbb{K}$  gilt

- (B1)  $B(v_1 + v_2, w) = B(v_1, w) + B(v_2, w),$
- (B2)  $B(\lambda v, w) = \lambda B(v, w),$
- (B3)  $B(v, w_1 + w_2) = B(v, w_1) + B(v, w_2),$
- (B4)  $B(v, \lambda w) = \lambda B(v, w).$

Bei einer Bilinearform werde ZWEI Vektoren (daher „Bi-“) zusammen auf ein Körperelement, also in der Regel auf eine Zahl abgebildet.

Die Additivität und Homogenität sind für den ersten und für den zweiten Vektor für jeweils feste Vektoren gegeben.

Dieses ist die allgemeine Definition. Betrachten wir schnell einen bekannten Sonderfall. Das Skalarprodukt, wie es auch in der Schule schon bekannt ist. Zwei Vektoren aus dem  $\mathbb{R}^3$  werden auf eine Zahl abgebildet. In der Schulschreibweise würde man schreiben: Das Skalarprodukt  $\vec{x} * \vec{y}$  zweier Vektoren  $\vec{x}$  und  $\vec{y}$  ist:

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} * \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = x_1 y_1 + x_2 y_2 + x_3 y_3$$

Beispiel das Skalarprodukt von  $\begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix}$  und  $\begin{pmatrix} 3 \\ -2 \\ 1 \end{pmatrix}$  ist 0. Das hat eine besondere Bedeutung, wie du dich vielleicht dich erinnerst. **Ist das Skalarprodukt zweier Vektoren Null, so sind die Vektoren orthogonal zueinander.** Das Skalarprodukt hat noch mehr Aussagekraft bzw. Nutzen (Projektion!) und sicherlich wird dies im Skript auch noch aufgegriffen.

(Bilinearformen gibt es aber auch in anderen Kontexten. Siehe Skript S. 226. Ziemlich wildes Zeug, z. B.  $V^* \times V \rightarrow \mathbb{K}$ . Hier wird eine Linearform (aus dem Dualraum) mit einem Vektor zusammen auf eine Zahl aus  $\mathbb{K}$  abgebildet. Am ehesten macht das vielleicht so Sinn:  $\varphi$  sei eine beliebige Linearform mit  $\varphi \in V^*$  und  $v \in V$  ein Vektor, dann ist  $B(\varphi, v) = \varphi(v)$  eine Bilinearform für die die Eigenschaften (B1) bis (B4) gelten. Das ist hier alles sehr allgemein. Ich vermute, es wird schnell auf besondere Bilinearformen, wie das Skalarprodukt, hinauslaufen.

#### Definition 8.1.7: Orthogonalität

Es seien  $V$  und  $W$   $\mathbb{K}$ -Vektorräume und  $B : V \times W \rightarrow \mathbb{K}$  eine Bilinearform, dann heißen  $v \in V$  und  $w \in W$  *orthogonal* bezüglich der Bilinearform  $B$ , falls  $B(v, w) = 0$  gilt. Für  $M \subseteq V$  ist

$$M^\perp := \{ w \in W \mid \forall v \in M (B(v, w) = 0) \}$$

ein Unterraum von  $W$ . Er heißt der zu  $M$  *orthogonale Unterraum* bezüglich  $B$ . Analog erklärt man für  $N \subseteq W$  den zu  $N$  *orthogonalen Unterraum*  $N^\perp \subseteq V$  bezüglich  $B$ .

Die Orthogonalität gilt nicht nur beim Skalarprodukt wie in der Schule, sondern lässt sich für beliebige Bilinearformen definieren.

Der orthogonale Unterraum wird auch *orthogonales Komplement* genannt. Ich bin mir nicht sicher, wie wichtig das Thema ist, aber das im Skript finde ich nicht sehr

verständlich. Besser wird es hier besprochen: <https://youtu.be/G8TZKSYFjPc?si=Ua7yskzDaK5MuPwn>

Das Video ist zumindest für das bessere Vorstellen hilfreich, denn im Skript werden noch eine ganze Reihe von Begrifflichkeiten runtergebetet, wobei die Anschauung ziemlich auf der Strecke bleibt.

**Noch ein Versuch, um den orthogonalen Unterraum zu veranschaulichen:** Zuerst die Teilmenge  $M$  eines Vektorraums: Betrachten wir z. B. alle Vektoren aus dem  $V = W = \mathbb{R}^3$ , die als

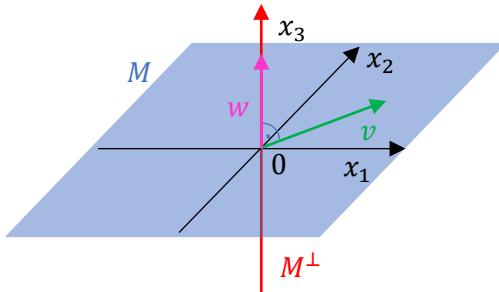
Linearkombination von  $\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$  und  $\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$  dargestellt werden können, also die lineare Hülle  $M = \langle \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \rangle \subset \mathbb{R}^3$ .  $M$  ist natürlich ein Unterraum des  $\mathbb{R}^3$  („übertragen“ in der Anschauung wieder die  $x_1 - x_2$ -Ebene). Der orthogonale Unterraum bzgl. einer Bilinearform  $B$  ist

$$M^\perp = \{w \in \mathbb{R}^3 \mid \forall v \in M \ (B(v, w) = 0)\}.$$

Betrachten wir für  $B$  das Standardskalarprodukt. Welche Vektoren  $w \in \mathbb{R}^3$  sind zu allen Vektoren aus  $M$  orthogonal? Alle, die in den ersten beiden Komponenten eine Null haben, also:

$$M^\perp = \left\{ \begin{pmatrix} 0 \\ 0 \\ \lambda \end{pmatrix} \mid \lambda \in \mathbb{R} \right\}. \text{ Denn es gilt für } w = \begin{pmatrix} 0 \\ 0 \\ \lambda \end{pmatrix} \in M^\perp \text{ und } v = \alpha \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + \beta \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} \alpha \\ \beta \\ 0 \end{pmatrix} \in M \text{ stets:} \\ B\left(\begin{pmatrix} \alpha \\ \beta \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ \lambda \end{pmatrix}\right) = \alpha \cdot 0 + \beta \cdot 0 + 0 \cdot \lambda = 0$$

$M^\perp$  besteht aus allen Vektoren in Richtung der  $x_3$ -Achse. Wie zuvor schon gesagt: Vektoren sind keine Punkte, aber damit am es sich etwas vorstellen kann ein Bild dazu:



Die Definition im Skript bzgl. der **Ausartung** (Definition 8.1.9) finde ich sehr unglücklich. Ich denke mal, dass diese Eigenschaft auch eher der Vollständigkeit halber angegeben wurde. Ich will es dennoch versuchen zu verdeutlichen: Es sei  $B$  eine Bilinearform mit  $B: V \times W \rightarrow \mathbb{K}$ .

Wir betrachten anstatt  $M \subseteq V$  ganz  $V$ , also  $M = V$  und damit nicht  $M^\perp$ , sondern

$V^\perp = \{w \in W \mid \forall v \in V \ (B(v, w) = 0)\} \subseteq W$ . Es sind also alle Elemente von  $W$  gemeint, die orthogonal zum gesamten Raum  $V$  sind. Man könnte auch umgekehrt alle Elemente von  $V$  betrachten, die orthogonal zum gesamten Raum  $W$  sind:  $\overset{\perp}{V} = \{v \in V \mid \forall w \in W \ (B(v, w) = 0)\} \subseteq V$ . Man nennt diese Untervektorräume auch Rechtskern (Rechtsradikal)  $V^\perp$  bzw. Linkskern (Linksradikal)  $\overset{\perp}{V}$ .

Zu Erinnerung: Alle Elemente, die bei einer linearen Abbildung auf Null abgebildet werden. Daher hier die Kern-Analogie bei den Bilinearformen und den Elementen, bei denen  $B(v, w) = 0$  gilt.

Sind **beide Kerne trivial**, d. h. enthalten sie **beide nur den Nullvektor**, so heißt die Bilinearform **nicht ausgeartet** oder **nicht entartet**. Andernfalls heißt die Bilinearform ausgeartet oder entartet.

Man kann auch formulieren:

Die Bilinearform ist somit genau dann nicht ausgeartet, wenn Folgendes gilt:

- Zu jedem Vektor  $v \in V \setminus \{0\}$  existiert ein Vektor  $w \in W$  mit  $B(v, w) \neq 0$  und
- zu jedem Vektor  $w \in W \setminus \{0\}$  existiert ein Vektor  $v \in V$  mit  $B(v, w) \neq 0$ .

### Definition 8.1.11: Symmetrie, Definitheit, Skalarprodukt

Es sei  $V$  ein  $\mathbb{K}$ -Vektorraum, dann heißt eine Bilinearform  $B$  auf  $V \times V$

- (i) *symmetrisch*, falls für alle  $v, w \in V$   $B(v, w) = B(w, v)$  gilt.
- (ii) *positiv definit*, falls  $\mathbb{K} = \mathbb{R}$  und  $\forall v \in V \setminus \{0\}$  ( $B(v, v) > 0$ ).
- (iii) *negativ definit*, falls  $\mathbb{K} = \mathbb{R}$  und  $\forall v \in V \setminus \{0\}$  ( $B(v, v) < 0$ ).
- (iv) *positiv beziehungsweise negativ semidefinit*, falls  $\mathbb{K} = \mathbb{R}$  und  $\forall v \in V \setminus \{0\}$  ( $B(v, v) \geq 0$ ) beziehungsweise  $\leq 0$ .
- (v) *indefinit*, wenn  $\mathbb{K} = \mathbb{R}$  und  $B$  weder positiv noch negativ semidefinit ist.
- (vi) ein *Skalarprodukt*, falls  $B$  positiv definit und symmetrisch ist.

Wieder ein Haufen von Begriffen, wobei besonders wichtig (vi), also das **Skalarprodukt** ist.

Das **Skalarprodukt** ist eine wichtige algebraische Operation in Vektorräumen, die zwei Vektoren miteinander verknüpft und eine reelle Zahl liefert. Es hat mehrere Bedeutungen und Anwendungen in der linearen Algebra und anderen mathematischen Disziplinen. In der Schule wird es nicht als Bili-

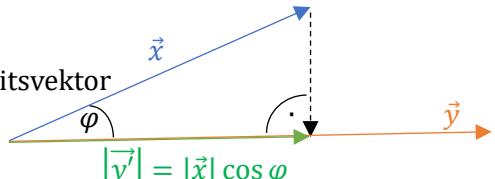
nearform eingeführt, sondern direkt als Verknüpfung zweier Vektoren:  $\vec{x} * \vec{y} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} * \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} =$   
 $x_1 y_1 + x_2 y_2 + x_3 y_3$ . An der Uni wird eher diese Schreibweise verwendet  $\langle x, y \rangle$ .

Hier sind einige wichtige Bedeutungen bzw. Verwendungszwecke des Skalarprodukts für Vektorräume:

- **Länge eines Vektors:** Das Skalarprodukt eines Vektors mit sich selbst liefert das Quadrat der Länge des Vektors. Die Länge eines Vektors ist also:  $|\vec{x}| = \sqrt{\vec{x} * \vec{x}}$ .
- **Winkel zwischen Vektoren:** Durch die Verwendung des Skalarprodukts können wir den Winkel zwischen zwei Vektoren berechnen, was in vielen Anwendungen wie der Geometrie, Physik und Signalverarbeitung nützlich ist. Es gilt:  $\cos \varphi = \frac{\vec{x} * \vec{y}}{|\vec{x}| |\vec{y}|}$ .
- **Orthogonalität von Vektoren:** Zwei Vektoren sind genau dann orthogonal zueinander, wenn ihr Skalarprodukt null ist.
- **Projektion eines Vektors auf einen anderen:** Das Skalarprodukt ermöglicht es, die Projektion eines Vektors auf einen anderen zu berechnen. Diese Projektion ist wichtig in Anwendungen wie der linearen Regression, der Berechnung von Abständen zwischen Punkten und der Bestimmung von Komponenten eines Vektors entlang einer bestimmten Richtung. Stellt man die Gleichung  $\cos \varphi = \frac{\vec{x} * \vec{y}}{|\vec{x}| |\vec{y}|}$  nach  $\vec{x} * \vec{y}$  so erhält man  $\vec{x} * \vec{y} = \cos \varphi |\vec{x}| |\vec{y}|$ . Entsprechend kann man das Skalarprodukt auch so definieren:  $\vec{x} * \vec{y} = \begin{cases} \cos \varphi |\vec{x}| |\vec{y}| & \text{falls } \vec{x}, \vec{y} \neq 0 \\ 0 & \text{sonst} \end{cases}$ . Was hat das mit einer Projektion zu tun? Wegen des folgenden Zusammenhangs:

Wir sehen hier die Projektion von  $\vec{x}$  auf  $\vec{y}$  mit  
der projizierten Länge von  $|\vec{x}| \cos \varphi \cdot \frac{\vec{y}}{|\vec{y}|}$  ist der Einheitsvektor  
in Richtung von  $\vec{y}$ .

$$\vec{y}' = |\vec{x}| \cos \varphi \cdot \frac{\vec{y}}{|\vec{y}|} = |\vec{x}| |\vec{y}| \cos \varphi \cdot \frac{\vec{y}}{|\vec{y}|^2} = \vec{x} * \vec{y} \cdot \frac{\vec{y}}{|\vec{y}|^2}$$



Im Skript werden noch einige Besonderheiten / Beispiele angesprochen:

**Bemerkung 8.1.12.** Offenbar ist ein Skalarprodukt stets nicht ausgeartet.

**Bemerkung 8.1.13** (Skalarprodukt auf  $\mathbb{C}$ -Vektorräumen). Ist  $V$  ein Vektorraum über  $\mathbb{C}$ , dann heißt eine Bilinearform  $B$  auf  $V \times V$  *hermitesch*<sup>2</sup>, falls für alle  $v, w \in V$

- $B(v, w) = \overline{B(w, v)}$  gilt.

In diesem Fall ist  $B(v, v) = \overline{B(v, v)}$  und daher  $B(v, v) \in \mathbb{R}$ , das heißt, die positive Definitheit lässt sich auch in diesem Fall sinnvoll erklären. Ein *Skalarprodukt* auf einem  $\mathbb{C}$ -Vektorraum ist eine positiv definite, hermitesche Sesquilinearform (siehe 8.1.5).

Im Komplexen gibt es als noch eine zusätzliche Bezeichnung – **hermitesch**, wenn obige Eigenschaft erfüllt ist. Ferner sei noch auf den Zusammenhang von Bilinearformen und Matrizen hingewiesen:

**Bemerkung 8.1.15.** Bilinearformen auf endlich-dimensionalen Vektorräumen lassen sich wieder mit Matrizen darstellen, so ist etwa das kanonische Skalarprodukt auf dem  $\mathbb{R}^n$  gegeben durch  $\langle x, y \rangle = B(x, y) = x^T I y$  und die Bilinearform  $B(x, y) = x_1(y_2 - 2y_1) + x_2(y_3 + y_1)$  auf  $\mathbb{R}^2 \times \mathbb{R}^3$  lässt sich darstellen als

$$(x_1 \ x_2) \begin{pmatrix} -2 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}.$$

Entsprechend lassen sich die Definitionen von Indefinitheit sowie positiver und negativer (Semi-)Definitheit auf reelle quadratische Matrizen übertragen. Weiter kann man sagen, dass eine Matrix  $A \in M(n \times n, \mathbb{C})$  *hermitesch* ist, wenn  $A^T = \overline{A}$  gilt. Je nach Anwendungsfall lassen sich mit diesen Eigenschaften weitere Aussagen beziehungsweise Ansätze zur Lösung des Problems machen.

## Normen und Metriken

### Definition 8.2.1: Kanonische Norm und Abstand im $\mathbb{R}^n$

Es sei  $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{R}^n$ ,  $n \in \mathbb{N}$ , dann definieren wir die *Norm* (Länge, Betrag, Euklidische Norm) von  $x$  als

$$\|x\| := \|x\|_2 := \left( \sum_{i=1}^n x_i^2 \right)^{\frac{1}{2}} = \sqrt{x_1^2 + \dots + x_n^2}.$$

Der *Abstand* von zwei Vektoren  $x, y \in \mathbb{R}^n$  ist dann nach dem Satz des Pythagoras

$$d(x, y) := \|x - y\| = \sqrt{\sum_{i=1}^n (x_i - y_i)^2}.$$

Anstatt die Koordinaten zu quadrieren können sie auch anders potenziert werden. Dies für auf die Verallgemeinerung:

### Definition 8.2.2: $p$ -Norm

Es seien  $x \in \mathbb{R}^n$  oder  $x \in \mathbb{C}^n$  und  $1 \leq p \leq \infty$ , dann heißt

$$\|x\|_p := \begin{cases} \left(\sum_{i=1}^n |x_i|^p\right)^{\frac{1}{p}}, & \text{für } 1 \leq p < \infty, \\ \max_{1 \leq i \leq n} |x_i|, & \text{für } p = \infty \end{cases} \quad (8.2)$$

die  $p$ -Norm von  $x$ .

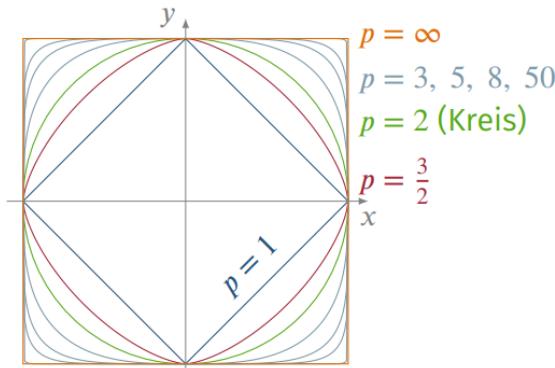
Ist  $p$  nicht explizit angegeben, gehen wir immer von der kanonischen Norm mit  $p = 2$  aus.

Durch  $d_p(x, y) := \|x - y\|_p$  ist dann wieder ein Abstandsbegriff gegeben.

Beispiel für verschiedene Normen:

**Beispiel 8.2.3.** Es sei  $x = (1, -3, 2)^T \in \mathbb{R}^3$ , dann gilt

$$\begin{aligned} \|x\|_1 &= |1| + |-3| + |2| = 6, \\ \|x\|_2 &= \|x\|_2 = \sqrt{1^2 + (-3)^2 + (2)^2} = \sqrt{14}, \\ \|x\|_\infty &= \max\{|1|, |-3|, |2|\} = 3. \end{aligned}$$



Für die  $p$ -Norm gelten (N1), (N2) und (N3):

**Bemerkung 8.2.4.** Es seien  $x, y \in \mathbb{R}^n$  oder  $\mathbb{C}^n$ ,  $\lambda \in \mathbb{R}$  und  $\|\cdot\| = \|\cdot\|_p$  für ein  $1 \leq p \leq \infty$ , dann gilt

(N1)  $\|x\| \geq 0$  und  $\|x\| = 0 \Leftrightarrow x = 0$  (Definitheit).

(N2)  $\|\lambda x\| = |\lambda| \|x\|$  (Homogenität).

(N3)  $\|x + y\| \leq \|x\| + \|y\|$  (Dreiecksungleichung).

Allgemeiner ist

### Definition 8.2.5: Normierter Raum

Es sei  $V$  ein Vektorraum und  $N : V \rightarrow \mathbb{R}$  mit den Eigenschaften (N1)-(N3), dann heißt  $N$  Norm auf  $V$  und  $(V, N)$  heißt normierter Raum.

### Satz 8.2.7: Induzierte Norm

Auf jedem  $\mathbb{K}$ -Vektorraum  $V$ ,  $\mathbb{K} = \mathbb{R}$  oder  $\mathbb{K} = \mathbb{C}$ , mit einem Skalarprodukt  $\langle \cdot, \cdot \rangle$ , ist für  $x \in V$  durch  $\sqrt{\langle x, x \rangle}$  eine Norm, die sogenannte *induzierte Norm*, erklärt. Im Fall  $V = \mathbb{R}^n$  ist dies gerade die 2-Norm. Vektorräume, die bezüglich der induzierten Norm vollständig im Sinne von Bemerkung 3.3.21(ii) sind, heißen *Hilberträume*<sup>a</sup>.

<sup>a</sup>David Hilbert, 1862-1943, dt. Mathematiker

### Satz 8.2.8: Cauchy-Schwarzsche Ungleichung

Es sei  $V$  ein  $\mathbb{R}$ - oder  $\mathbb{C}$ -Vektorraum mit einem Skalarprodukt  $\langle \cdot, \cdot \rangle$  und der induzierten Norm  $\|\cdot\|$ , dann gilt für alle  $x, y \in V$  die Ungleichung

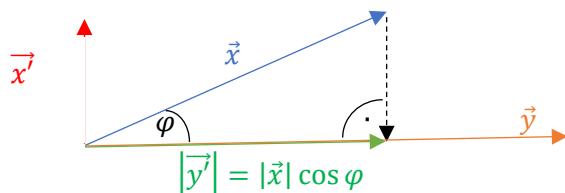
$$|\langle x, y \rangle| \leq \|x\| \|y\|,$$

wobei Gleichheit genau dann gilt, wenn  $x = \lambda y$  oder  $y = \lambda x$  mit  $\lambda \in \mathbb{R}$  beziehungsweise  $\lambda \in \mathbb{C}$ .

Als Beispiel(!) folgt im Skript ein ziemlich wichtiges Verfahren, dass das Skalarprodukt ausnutzt, das das **Gram-Schmidtsche Orthogonalisierungsverfahren**. Mit diesem Verfahren kann man zu einer Menge von linear unabhängigen Vektoren eine Menge von paarweise orthogonalen Vektoren der Länge 1, ein sogenanntes **Orthonormalsystem** berechnen. „Ortho“, da paarweise orthogonale Vektoren und „normal“, da die Vektoren normiert sind, d. h. alle die Länge 1 haben. Handelt es sich bei der Menge der Vektoren um eine Basis, dann kann so eine **Orthonormalbasis** bestimmt werden. Das Verfahren gibt es in leicht unterschiedlichen Formulierungen, die sich z. B. darin unterscheiden, in welchem Schritt normiert (also durch die Länge des Vektors dividiert) wird. Außerdem gibt es neben der klassischen Variante noch eine modifizierte, die dafür sorgen soll, dass bei der technischen Umsetzung mit Fließkommazahlen weniger Rundungsfehler auftreten. Ich habe das Verfahren als Grundlage für ein anderes Verfahren im Rahmen einer Semesterarbeit behandelt (Arnoldi.pdf).

Das Verfahren wurde schon vor den Herren Gram und Schmidt verwendet und ist anschaulich gut zu verstehen:

$$\vec{x}' = \vec{x} - \vec{x} * \vec{y} \cdot \frac{\vec{y}}{|\vec{y}|^2}$$



Normiert man vorab den ersten Vektor  $\vec{y}$ , so vereinfacht sich die Berechnung von

$\vec{x}' = \vec{x} - \vec{x} * \vec{y} \cdot \vec{y}$ , da die Länge von  $\vec{y}$  nun 1 beträgt. In der Schreibweise des Skripts gilt für den senkrechten Vektor  $\vec{x}' = \vec{x} - \langle \vec{x}, \vec{y} \rangle \cdot \vec{y}$ . Ein anschließendes Normieren führt zu  $\frac{\vec{x}'}{|\vec{x}'|}$ , einem Vektor, der senkrecht zu  $\vec{y}$  steht und die Länge 1 hat.

In der (händischen) Praxis geht man so vor. Ein Basisvektor wird ausgewählt und normiert. Im Skript wird er  $w_1$  genannt. Dann berechnet man die nachfolgenden orthogonalen und normierten Vektoren so:

$$\tilde{w}_k = v_k - \sum_{j=1}^{k-1} \frac{\langle w_j, v_k \rangle}{\|w_j\|^2} w_j$$

Und normiert anschließend durch:  $w_k = \frac{\tilde{w}_k}{\|\tilde{w}_k\|}$ . Beachte: Die Summe läuft bis  $k-1$ . Wenn man bereits die vorigen Vektoren normiert hat, dann ist in der Summe  $\|w_j\|^2$  immer 1 und die Formel ist kürzer:

$$\tilde{w}_k = v_k - \sum_{j=1}^{k-1} \langle w_j, v_k \rangle w_j$$

Am Ende erhält man eine Menge von paarweise orthogonalen Vektoren der Länge 1.

Als weitere Beispiele(!) werden im Skript **orthogonale Matrizen** (reguläre Matrizen mit der Eigenschaft  $A^T = A^{-1}$ ) und **unitäre Matrizen** ( $A^T = A^{-1}$ ) angesprochen, wozu es sicherlich noch mehr zu sagen gäbe.

**Normen** gibt es nicht nur für Vektoren, sondern auch für Matrizen, Funktionen und für andere mathematische Objekte.

**Beispiel 8.2.10** (Matrixnormen). Es sei  $\|\cdot\|_a$  eine Norm auf dem  $\mathbb{R}^n$ , dann ist durch

$$\|A\|_a := \sup_{x \neq 0} \frac{\|Ax\|_a}{\|x\|_a}$$

**Matrixnorm**. Die Frobenius-Norm ist sicherlich die bekannteste.

eine Norm auf  $M(n \times n, \mathbb{R})$  definiert, wie man leicht zeigt (Aufgabe A.8.10).

Weiter erhält man, da  $M(m \times n, \mathbb{R})$  und der  $\mathbb{R}^{mn}$  isomorph sind (Bemerkung 6.2.10 (ii)), durch jede Norm auf dem  $\mathbb{R}^{mn}$  auch eine auf  $M(m \times n, \mathbb{R})$ . So führt etwa die kanonische Norm auf dem  $\mathbb{R}^{mn}$ ,  $\|x\|_2 = \sqrt{\sum_{i=1}^{mn} |x_i|^2}$  für  $A \in M(m \times n, \mathbb{R})$ ,  $A = (a_{ij})$  mit  $x_1 = a_{11}, x_2 = a_{12}, \dots, x_n = a_{1n}, x_{n+1} = a_{21}, \dots, x_{mn} = a_{mn}$  zur **Frobenius<sup>4</sup>-Norm**

$$\|A\|_F := \sqrt{\sum_{i=1}^m \sum_{j=1}^n |a_{ij}|^2}.$$

### Verallgemeinerung des Abstands begriffes

**Bemerkung 8.2.11.** Es seien  $x, y, z \in \mathbb{R}^n$  beliebig. Der Abstand  $d$  aus Definition 8.2.1 hat folgende Eigenschaften

(Me1)  $d(x, y) \geq 0$  und  $d(x, y) = 0 \Leftrightarrow x = y$ .

(Me2)  $d(x, y) = d(y, x)$  (Symmetrie).

(Me3)  $d(x, z) \leq d(x, y) + d(y, z)$  (Dreiecksungleichung).

### Definition 8.2.12: Metrik und metrischer Raum

Es sei  $M \neq \emptyset$  und  $d : M \times M \rightarrow \mathbb{R}$  habe die Eigenschaften (Me1)-(Me3), dann heißt  $d$  **Abstand** oder **Metrik** auf  $M$  und das Paar  $(M, d)$  heißt **metrischer Raum**.

**Beispiele 8.2.13.** (i) Es sei  $M \neq \emptyset$ , dann ist durch

$$d(x, y) = \begin{cases} 0, & x = y, \\ 1, & x \neq y \end{cases}$$

ein Abstand gegeben (*diskrete Metrik*).

(ii) Betrachten wir die Menge der Codes der Länge  $n$ , das heißt die Menge  $M = \{0, 1\}^n$ , dann ist für  $x \in M$  durch

$$\|x\| = \sum_{k=1}^n x_k$$

eine Norm und für  $x, y \in M$  durch

$$d(x, y) = \|x - y\| = \sum_{k=1}^n |x_k - y_k|$$

ein Abstand (Metrik) definiert. Dieser sogenannte *Hamming-Abstand* liefert gerade die Anzahl der Stellen, an denen sich die beiden Codes (Vektoren) unterscheiden und spielt zum Beispiel bei der Erkennung und Korrektur von Fehlern bei der Datenübertragung eine Rolle.

## Determinanten und das Eigenwertproblem

Die Motivation im Skript ist hier ganz gut. Ausgangspunkt ist die allgemeine Lösung eines 2x2 LGS. Hier tritt eine gewisse Verknüpfung von Koeffizienten bzw. von den Komponenten der zugehörigen Matrix auf. Hat die Matrix z. B. die Einträge  $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ , so nennt man den Ausdruck  $ad - bc$  die Determinante von  $A$ . Die Determinante bestimmt maßgeblich die Lösbarkeit des zugehörigen LGS, denn ist sie Null, so ist das LGS nicht lösbar. Daher der Name „Determinante“. Schreibweise:

$$\det A = \begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc$$

### Definition 9.1.2: Determinante einer Matrix

Es sei  $(a_{ki})$  eine  $n$ -reihige Matrix, dann heißt

$$\det(a_{ki}) = \begin{vmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{nn} \end{vmatrix} := \sum_{\pi \in S_n} \operatorname{sgn} \pi a_{1\pi(1)} \cdots a_{n\pi(n)} \quad (9.1)$$

die *Determinante* der Matrix  $(a_{ki})$ .

**Beispiele 9.1.3.** Wir wollen die Determinante einer  $n$ -reihigen Matrix für  $n = 1, 2, 3$  berechnen. In Bemerkung 2.3.4 (iii) hatten wir die entsprechenden Permutationen bereits aufgeführt.

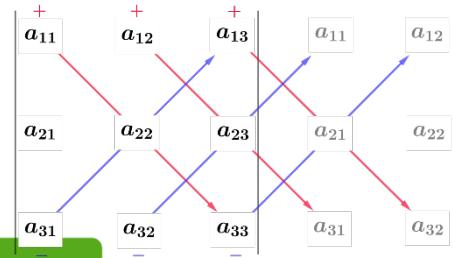
$$\begin{aligned} n = 1 \quad & \det(a_{11}) = a_{11}. \\ n = 2 \quad & \det \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = a_{11}a_{22} - a_{12}a_{21} \\ n = 3 \quad & \det \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} = a_{11}a_{22}a_{33} - a_{11}a_{23}a_{32} - a_{13}a_{22}a_{31} \\ & \quad - a_{12}a_{21}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32}. \end{aligned}$$

Die Vorschrift zur Berechnung der Determinante für  $n = 3$  ist als *Sarrusche<sup>2</sup> Regel* bekannt. Für  $n \geq 4$  werden die  $n!$  Summanden schnell unübersichtlich, wir betrachten später Methoden, um die Berechnung schrittweise zu vereinfachen.

$$\det \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} = a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} \\ - a_{31}a_{22}a_{13} - a_{32}a_{23}a_{11} - a_{33}a_{21}a_{12}$$

#### Satz 9.1.4: Determinante der transponierten Matrix

Es sei  $A \in M(n \times n, \mathbb{K})$ , dann gilt  $\det A = \det A^T$ .



#### Satz 9.1.5

Es sei  $A = (a_{ik}) \in M(n \times n, \mathbb{K})$ ,  $\lambda \in \mathbb{K}$ , dann gilt

- (i)  $\det A$  multipliziert sich mit  $\lambda$ , wenn man eine Zeile<sup>a</sup> mit  $\lambda$  multipliziert (siehe auch Aufgabe A.9.1).
- (ii)  $\det A$  bleibt unverändert, wenn man ein Vielfaches einer Zeile von  $A$  zu einer anderen addiert.
- (iii)  $\det I = 1$ .

<sup>a</sup>die Aussagen gelten auch für Spalten, siehe 9.1.6.

#### Korollar 9.1.6

Es sei  $A \in M(n \times n, \mathbb{K})$ , dann gilt

- (i)  $\det A$  ändert bei Zeilenumtauschungen das Vorzeichen.
- (ii)  $\det A = 0$ , falls die Zeilenvektoren von  $A$  linear abhängig sind, insbesondere falls zwei gleiche Zeilen oder eine nur aus Nullen bestehende Zeile auftritt.
- (iii) Die in den Sätzen 9.1.4, 9.1.5 und dem Korollar 9.1.6 genannten Eigenschaften bezüglich der Zeilen gelten auch für die Spalten.
- (iv) Für  $\lambda \in \mathbb{K}$  gilt  $\det \lambda A = \lambda^n \det A$ .

Bei der Regel von Sarrus notiert man rechts neben den Einträgen der Matrix erneut die erste und zweite Spalte. Die Determinante berechnet sich dann nach dem folgenden Kreuzschema (von links oben nach rechts unten addieren und von links unten nach rechts oben subtrahieren):

### Lemma 9.1.7: Determinantenentwicklung

Ist  $A \in M(n \times n, \mathbb{K})$  von der speziellen Gestalt

$$A = \begin{pmatrix} a_1 & * \\ 0 & A' \end{pmatrix}$$

### Korollar 9.1.8: Determinante einer Dreiecksmatrix

Ist  $A$  eine *obere Dreiecksmatrix*, das heißt

$$A = \begin{pmatrix} a_1 & a_2 & * & & \\ & \ddots & & & \\ \mathbf{0} & & a_n & & \end{pmatrix},$$

so folgt durch sukzessive Anwendung von Lemma 9.1.7, dass

$$\det A = a_1 \cdot a_2 \cdots a_n.$$

Die Determinante einer oberen Dreiecksmatrix ist also das Produkt der Diagonalelemente. Insbesondere folgt mit Korollar 9.1.6 (ii), dass  $\det A = 0 \Leftrightarrow \text{rg } A < n$ .

**Bemerkungen 9.1.10.** (i) Allgemeiner gilt für *Blockmatrizen* bezeichnete

$A \in M(n \times n, \mathbb{K})$  der Form  $A = \begin{pmatrix} A_1 & * \\ 0 & A_2 \end{pmatrix}$  mit  $A_1 \in M(k \times k, \mathbb{K})$  und  $A_2 \in M((n-k) \times (n-k), \mathbb{K})$ , dass  $\det A = \det A_1 \cdot \det A_2$ .

Beispiele dazu im Skript!

### Satz 9.1.12: Laplacescher<sup>a</sup> Entwicklungssatz

<sup>a</sup>Pierre-Simon Laplace, 1749 - 1827, frz. Mathematiker

(i) Entwicklung nach der  $i$ -ten Spalte:

$$\det(a_{ki}) = \sum_{k=1}^n a_{ki} A_{ki}.$$

(ii) Entwicklung nach der  $k$ -ten Zeile:

$$\det(a_{ki}) = \sum_{i=1}^n a_{ki} A_{ki}.$$

Dabei folgt die zweite Formel durch Betrachtung von  $(a_{ki})^T$  und  $A_{ki}$  heit der *Cofaktor* zu  $a_{ki}$  in der Matrix  $(a_{ki})$ .

### Satz 9.1.13: Determinantenmultiplikationssatz

Es seien  $A, B \in M(n \times n, \mathbb{K})$ , dann gilt

$$\det(AB) = \det A \cdot \det B.$$

**Bemerkung 9.1.14.** Determinante einer linearen Abbildungdet:linAbb Für eine lineare Abbildung  $F : V \rightarrow V$  eines endlich-dimensionalen Vektorraums können wir nun

$$\det F = \det A$$

setzen, wobei  $A$  die Darstellungsmatrix von  $F$  bezüglich einer beliebigen Basis von  $V$  ist. Ist  $\tilde{A}$  die Darstellungsmatrix von  $F$  bezüglich einer anderen Basis, dann existiert eine invertierbare Matrix  $B$  mit  $\tilde{A} = BAB^{-1}$  und folglich gilt

$$\det \tilde{A} = \det(BAB^{-1}) = \det A.$$

Die Determinante von  $F$  ist daher eindeutig bestimmt.

(Die Determinante ist **unabhängig** von der Wahl der Basis!)

## Eigenwerte, Eigenvektoren, Eigenraum

Die Motivation im Skript zündet bei mir nicht so recht. Bei diesem Thema betrachtet man linearen Abbildungen von  $V$  in sich, also Endomorphismen.  $V$  ist ein Vektorraum über einen Körper  $\mathbb{K}$  (in der Regel der Körper  $\mathbb{R}$  oder  $\mathbb{C}$ ). Nun untersucht man die lineare Abbildung  $f: V \rightarrow V$  näher: Gibt es besondere Vektoren, die unter der Abbildung unverändert bleiben? D. h. gibt es Vektoren für die gilt:

$$f(v) = v$$

Solche Vektoren würden auf sich selbst abgebildet werden. In der Mathematik bezeichnet man Werte / Punkte / Vektoren, die durch eine Abbildung auf sich selbst abgebildet werden, kurz als „**Fixpunkte**“, auch wenn bei Vektoren in dem Fall auch „**Fixvektoren**“ genannt werden. In der Geometrie lassen sich Fixpunkte leicht vorstellen. Betrachtet man etwa eine Punktspiegelung, so ist das Spiegelzentrum ein Fixpunkt. Das Bild des Spiegelzentrums liegt am gleichen Ort wie sein Urbild. Das Spiegelzentrum wird auf sich selbst abgebildet. Bei der Spiegelung sind alle Punkte der Spiegelgerade Fixpunkte und werden auf sich selbst abgebildet.

Nun ist die Forderung  $f(v) = v$  sehr streng. Nicht jede lineare Abbildung hat (abgesehen vom Nullelement) einen Fixvektor bzw. Fixpunkt. Die Drehung um den Ursprung ist dafür ein Beispiel. Bei affinen Abbildungen mit Verschiebungen gibt es auch keine Fixpunkte. Vielleicht könnte man etwas weniger fordern, z. B. gibt es Vektoren, die unter der Abbildung ihre Richtung nicht ändern (also nur ihre Länge)? D. h. wir suchen Vektoren für die gilt:

$$f(v) = \lambda \cdot v$$

Ist der Vektorraum endlichdimensional, so kann jeder Endomorphismus  $f$  durch eine quadratische Matrix  $A$  beschrieben werden. Die obige Gleichung lässt sich dann als Matrizengleichung schreiben:

$$A \cdot v = \lambda v$$

wobei  $v$  hier einen Spaltenvektor bezeichnet. Man nennt eine Lösung  $v \neq 0$  und  $\lambda$  in diesem Fall Eigenvektor bzw. Eigenwert der Matrix  $A$ .

Diese Gleichung kann man auch in der Form

$$A \cdot v = \lambda E \cdot v$$

schreiben, wobei  $E$  die Einheitsmatrix bezeichnet (das Skript verwendet dafür  $I$ ), und äquivalent zu

$$(A - \lambda E) \cdot v = 0 \text{ bzw. } (\lambda E - A) \cdot v = 0$$

umformen.

Bestimmung der Eigenwerte, Eigenvektoren und Eigenraum

Die Eigenwerte definierende Gleichung

$$(A - \lambda I) \cdot v = 0$$

stellt ein **homogenes**(!) lineares Gleichungssystem dar. Da  $v \neq 0$  vorausgesetzt wird, ist dieses genau dann lösbar, wenn gilt:

$$\det(A - \lambda I) = 0$$

Berechnet man die Determinante auf der linken Seite, so erhält man bei einer  $n \times n$ -Matrix immer ein Polynom  $n$ -ten Grades in  $\lambda$ . Dieses wird als **charakteristisches Polynom  $P_A(\lambda)$**  bezeichnet und **dessen Nullstellen sind die Eigenwerte**, also die Lösungen der Gleichung:

$$\alpha_n \cdot \lambda^n + \alpha_{n-1} \cdot \lambda^{n-1} + \cdots + \alpha_1 \cdot \lambda + \alpha_0 = 0$$

Da ein Polynom vom Grad  $n$  höchstens  $n$  Nullstellen besitzt, gibt es höchstens  $n$  Eigenwerte. Zerfällt das Polynom vollständig, z. B. wie jedes Polynom über  $\mathbb{C}$ , so gibt es genau  $n$  Nullstellen, wobei mehrfache Nullstellen mit ihrer Vielfachheit gezählt werden.

Um **Eigenvektoren** und ihre **Eigenräume** zu bestimmen, geht man wie folgt vor:

1. **Bestimmung der Eigenwerte** (Bilden der Matrix  $A - \lambda I$ . Aufstellen des charakteristischen Polynoms mittels  $\det(A - \lambda I)$  und lösen der Gleichung  $\det(A - \lambda I) = 0$ .)
2. Für jeden Eigenwert  $\lambda_i$  wird die Gleichung

$$(A - \lambda_i I) \cdot v_i = 0$$

betrachtet. Je nach Dimension des Vektorraums hat der Vektor  $v_i$  entsprechend viele Komponenten (Koordinaten). Diese Gleichung stellt ein LGS da, das gelöst werden kann. Beachte, dass  $(A - \lambda_i E) = \tilde{A}$  eine Matrix ist und das Produkt mit einem gesuchten Vektor  $v_i$  dann ein LGS mit  $n$  Unbekannten und  $n$  Gleichungen liefert. **Diese LGS wird z. B. mit dem Gaußverfahren gelöst.**

3. **Der Eigenraum** zu jedem  $\lambda_i$  ist die Menge aller Vektoren  $v$ , für die  $f(v) = \lambda_i \cdot v$  erfüllt ist.

Beispiele online: <https://www.mathebibel.de/eigenvektoren-berechnen>

**Eigenes Beispiel:** Bestimme die Eigenwerte, Eigenvektoren und Eigenräume zur linearen Abbildung  $f$  mit der Darstellungsmatrix  $A = \begin{pmatrix} 1 & 1 \\ 2 & -1 \end{pmatrix}$ .

**Ansatz Eigenwerte:** Löse die Gleichung  $\det(A - \lambda I) = 0$

$$\begin{aligned} \det(A - \lambda E) &= \det\left(\begin{pmatrix} 1 & 1 \\ 2 & -1 \end{pmatrix} - \lambda \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\right) = \det\left(\begin{pmatrix} 1-\lambda & 1 \\ 2 & -1-\lambda \end{pmatrix}\right) = (1-\lambda)(-1-\lambda) - 1 \cdot 2 \\ &= (-1 - \lambda + \lambda + \lambda^2) - 2 = \lambda^2 - 3 = 0 \end{aligned}$$

$$\Rightarrow \lambda_1 = \sqrt{3} \text{ und } \lambda_2 = -\sqrt{3}$$

**Ansatz Eigenvektoren:** Für jeden Eigenwert  $\lambda_i, i = 1, 2$ , wird die Gleichung  

$$(A - \lambda_i I) \cdot v_i = 0$$

nach  $v_i$  aufgelöst.

Fall  $\lambda_1 = \sqrt{3}$ : Betrachte das zugehörige LGS mit  $v_1 = \begin{pmatrix} x \\ y \end{pmatrix}$  zu:

$$\left( \begin{pmatrix} 1 & 1 \\ 2 & -1 \end{pmatrix} - \sqrt{3} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \right) \cdot \begin{pmatrix} x \\ y \end{pmatrix} = 0$$

$$\begin{pmatrix} 1 - \sqrt{3} & 1 \\ 2 & -1 - \sqrt{3} \end{pmatrix} \cdot \begin{pmatrix} x \\ y \end{pmatrix} = 0$$

Also:

$$\left| \begin{array}{lll} (1 - \sqrt{3})x + y & = 0 \\ 2x + (-1 - \sqrt{3})y & = 0 \end{array} \right|$$

Lösen mit Gaußverfahren. Erweiterte Koeffizientenmatrix:

$$\left( \begin{array}{ccc} 1 - \sqrt{3} & 1 & 0 \\ 2 & -1 - \sqrt{3} & 0 \end{array} \right) \boxed{\quad} \mid : (1 - \sqrt{3}) \cdot (-2) \downarrow +$$

$$\left( \begin{array}{ccc} 1 - \sqrt{3} & 1 & 0 \\ 0 & \underbrace{(-1 - \sqrt{3}) - \frac{2}{1 - \sqrt{3}}}_{=\alpha} & 0 \end{array} \right) \boxed{\quad} \boxed{\quad}$$

$$\text{Beachte: } \alpha = (-1 - \sqrt{3}) - \frac{2}{1 - \sqrt{3}} = -\frac{(1 + \sqrt{3})(1 - \sqrt{3})}{1 - \sqrt{3}} - \frac{2}{1 - \sqrt{3}} = -\frac{1 - 3}{1 - \sqrt{3}} - \frac{2}{1 - \sqrt{3}} = -\frac{-2}{1 - \sqrt{3}} - \frac{2}{1 - \sqrt{3}} = 0$$

$$\left( \begin{array}{ccc} 1 - \sqrt{3} & 1 & 0 \\ 0 & 0 & 0 \end{array} \right)$$

Damit folgt, dass das LGS wegen der Nullzeile unendlich viele Lösungen für  $x$  und  $y$  hat, nämlich alle, die die Gleichung  $(1 - \sqrt{3})x + y = 0$  erfüllen. Um einen Eigenvektor besser angeben zu können, kann man nun eine Variable wählen, z. B.  $y = 1$ .

$$\Rightarrow (1 - \sqrt{3})x + 1 = 0$$

$$(1 - \sqrt{3})x = -1$$

$$x = -\frac{1}{1 - \sqrt{3}} = -\frac{1}{1 - \sqrt{3}} \cdot \frac{1 + \sqrt{3}}{1 + \sqrt{3}} = -\frac{1 + \sqrt{3}}{1 - 3} = \frac{1}{2}(1 + \sqrt{3})$$

Damit lässt sich ein Eigenvektor so angeben:  $v_1 = \begin{pmatrix} \frac{1}{2}(1 + \sqrt{3}) \\ 1 \end{pmatrix}$ .

Ein analoges Vorgehen für den Fall  $\lambda_2 = -\sqrt{3}$  führt auf:  $v_2 = \begin{pmatrix} \frac{1}{2}(1 - \sqrt{3}) \\ 1 \end{pmatrix}$

Hier eine Proberechnung (nicht nötig):

$$\begin{aligned}
A \cdot v_1 &= \begin{pmatrix} 1 & 1 \\ 2 & -1 \end{pmatrix} \begin{pmatrix} \frac{1}{2}(1 + \sqrt{3}) \\ 1 \end{pmatrix} = \begin{pmatrix} \frac{1}{2}(1 + \sqrt{3}) + 1 \\ (1 + \sqrt{3}) - 1 \end{pmatrix} = \begin{pmatrix} \frac{1}{2}(1 + \sqrt{3}) + 1 \\ \sqrt{3} \end{pmatrix} = \begin{pmatrix} \frac{1}{2}(3 + \sqrt{3}) \\ \sqrt{3} \end{pmatrix} \\
&= \sqrt{3} \begin{pmatrix} \frac{1}{2}(1 + \sqrt{3}) \\ 1 \end{pmatrix} = \lambda_1 \cdot v_1
\end{aligned}$$

### Ansatz Eigenraum:

Da der Eigenraum die Menge aller Vektoren umfasst, für die für ein  $\lambda$  die Gleichung  $f(v) = \lambda \cdot v$  erfüllen, kann ausgehend von der Berechnung der Eigenvektoren der Eigenraum bestimmt werden. Ergibt sich bei der Berechnung eine eindeutige Lösung  $\neq 0$ , so bestünde der Eigenraum nur aus dem Nullelement und der eindeutigen Lösung des LGS. Oft, wie im Beispiel hier, gibt es unendlich viele Lösungen. Alle Koordinaten  $x$  und  $y$ , die  $(1 - \sqrt{3})x + y = 0$  erfüllen, liefern verschiedene Eigenvektoren. Wir hatten zuvor  $y = 1$  gewählt. Wie lassen sich die Vektoren beschreiben, wenn wir  $y = k \in \mathbb{R}$  wählen? Es folgt:

$$(1 - \sqrt{3})x + k = 0$$

$$(1 - \sqrt{3})x = -k$$

$$x = -\frac{k}{1 - \sqrt{3}} = -\frac{k}{1 - \sqrt{3}} \cdot \frac{1 + \sqrt{3}}{1 + \sqrt{3}} = \frac{1}{2}(1 + \sqrt{3})k$$

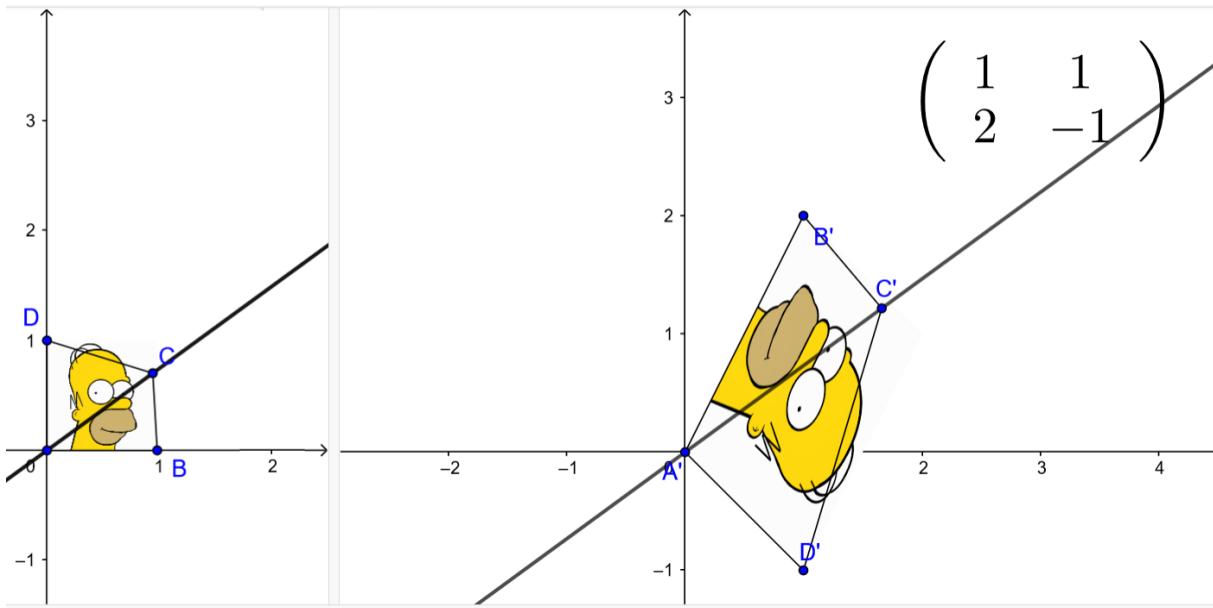
Damit ist  $v_{1,k} = \begin{pmatrix} \frac{1}{2}(1 + \sqrt{3})k \\ k \end{pmatrix} = k \begin{pmatrix} \frac{1}{2}(1 + \sqrt{3}) \\ 1 \end{pmatrix}$ . Dies ist kein Zufall. **Der Ansatz mit  $y = k$  ist nicht notwendig.** Es ist auch so klar, dass die Vielfachen von dem zuvor konkret (durch  $y = 1$ ) angefundenen Eigenvektor auch Eigenvektoren sind. Das liegt an der Homogenität der linearen Abbildung: Ist  $v_1$  zu  $\lambda_1$  ein Eigenvektor, so gilt  $f(v_1) = \lambda_1 \cdot v_1$ . Ist  $v_{1,k} = k \cdot v_1$ , so ist wegen der Homogenität:  $f(v_{1,k}) = f(k \cdot v_1) = k \cdot f(v_1) = k \cdot \lambda_1 \cdot v_1 = \lambda_1 \cdot k \cdot v_1 = \lambda_1 \cdot v_{1,k}$  auch ein Eigenvektor zu  $\lambda_1$ . **Wir können also direkt den Eigenraum  $\text{Eig}(f, \lambda_1) := \{v \in V \mid f(v) = \lambda \cdot v\}$  durch vielfache der bereits bestimmten Eigenvektoren angeben:**

$$\text{Eig}(f, \lambda_1) := \{v \in V \mid v = k \begin{pmatrix} \frac{1}{2}(1 + \sqrt{3}) \\ 1 \end{pmatrix}, k \in \mathbb{R}\}$$

$$\text{Eig}(f, \lambda_2) := \{v \in V \mid v = k \begin{pmatrix} \frac{1}{2}(1 - \sqrt{3}) \\ 1 \end{pmatrix}, k \in \mathbb{R}\}$$

### Veranschaulichung des Beispiels:

Unter der Abbildung  $f$  werden Vektoren verzerrt und gespiegelt. Fassen wir die Vektoren als Ortsvektoren zu Punkten auf, dann kann man auch die Punkte zu den abgebildeten Ortsvektoren zeichnen. Das sieht dann so aus:



Auf der linken Seite sind die Urbilder. Rechts sieht man das Bild der Abbildung. Die schräg nach oben laufende Gerade verläuft in Richtung des ersten Eigenvektors. Entsprechend werden alle Punkte, die auf dieser Geraden liegen, wieder auf die Gerade abgebildet (nur um  $\lambda_1 = \sqrt{3}$  gestreckt.) Beachte in der Zeichnung die Lage von Punkt  $C$  und Punkt  $C'$

#### Beispiel Klausuraufgabe:

- (c) Ist  $\lambda = 2$  ein Eigenwert der Matrix  $A = \begin{pmatrix} 3 & 2 & 1 \\ 2 & 2 & 2 \\ -1 & 2 & 0 \end{pmatrix}$ ? Wenn ja, geben Sie einen Eigenvektor zum Eigenwert  $\lambda$  an. (5)

Diese Klausuraufgabe erfordert nicht die Berechnung aller Eigenwerte, sondern die Überprüfung, ob  $\lambda = 2$  ein Eigenwert ist und wenn ja, wie ein Eigenvektor dazu aussieht. Betrachten wir hierzu  $(A - \lambda I)$  und bestimmen die Determinante, denn es muss  $\det(A - \lambda I) = 0$  gelten.

$$(A - \lambda E) = (A - 2I) = \begin{pmatrix} 1 & 2 & 1 \\ 2 & 0 & 2 \\ -1 & 2 & -2 \end{pmatrix}$$

Determinante:

$$\begin{pmatrix} 1 & 2 & 1 \\ 2 & 0 & 2 \\ -1 & 2 & -2 \end{pmatrix} \xrightarrow{\text{Row operations}} \begin{pmatrix} 1 & 2 & 1 \\ 0 & -4 & 0 \\ -1 & 2 & -2 \end{pmatrix} \xrightarrow{\text{Row operations}} \begin{pmatrix} 1 & 2 & 1 \\ 0 & 1 & 0 \\ -1 & 2 & -2 \end{pmatrix}$$

$$\det \begin{pmatrix} 1 & 2 & 1 \\ 2 & 0 & 2 \\ -1 & 2 & -2 \end{pmatrix} = 0 + 2 \cdot 2 \cdot (-1) + 1 \cdot 2 \cdot 2 - 0 - 2 \cdot 2 \cdot 1 - (-2) \cdot 2 \cdot 2 = 4 \neq 0$$

$\det(A - 2I) = 0$  ist nicht erfüllt. Damit ist das nachgewiesen, dass  $\lambda = 2$  kein Eigenwert ist und die Aufgabe fertig. Würde man sich die Mühe machen, das charakterliche Polynom zu  $A = \begin{pmatrix} 3 & 2 & 1 \\ 2 & 2 & 2 \\ -1 & 2 & 0 \end{pmatrix}$  zu bestimmen, so wäre das Ergebnis:  $P_A(\lambda) = \det(A - \lambda E) = -\lambda^3 + 5\lambda^2 + \lambda - 10$ . Setzt man  $\lambda = 2$  ein, sieht man ebenfalls dass hier keine Nullstelle vorliegt.

#### Eigenraum zum Eigenwert

Ist  $\lambda$  ein Eigenwert der linearen Abbildung  $f: V \rightarrow V$ , dann nennt man die Menge aller Eigenvektoren zu diesem Eigenwert den Eigenraum zum Eigenwert  $\lambda$ . Der Eigenraum ist definiert durch:

$$\text{Eig}(f, \lambda) := \{v \in V \mid f(v) = \lambda \cdot v\}$$

Für jeden Eigenwert gibt es einen Eigenraum.

Spektrum und weitere Begriffe

Mehrache Vorkommen eines bestimmten Eigenwertes fasst man zusammen und erhält so nach Umbenennung die Aufzählung  $\lambda_1, \dots, \lambda_k$  der **verschiedenen** Eigenwerte mit ihren Vielfachheiten  $\mu_1, \dots, \mu_k$ . Dabei ist  $1 \leq k \leq n$  und  $\sum_{i=1}^k \mu_i = n$ . Die eben dargestellte Vielfachheit  $\mu_i$  eines Eigenwertes als Nullstelle des charakteristischen Polynoms bezeichnet man als **algebraische Vielfachheit**.

Die Menge der Eigenwerte wird **Spektrum** genannt und  $\sigma(A)$  geschrieben. Es gilt also:

$$\sigma(A) = \{\lambda \in \mathbb{K} \mid \exists v \neq 0: Av = \lambda v\}$$

Als **Spektralradius** bezeichnet man den Betrag des betragsmäßig größten Eigenwerts.

Im Skript wird noch eingeführt:

#### Definition 9.3.3: Spur

Die Summe  $\sum_{i=1}^n a_{ii}$  der Diagonalelemente der Matrix  $(a_{ij})$  heißt die **Spur** von  $(a_{ij})$ , in Zeichen

$$\text{Spur } A = \sum_{i=1}^n a_{ii}$$

oder auch – da  $\det(A - \lambda I)$  unabhängig von der Basis und die Monome  $\lambda^k$  linear unabhängig sind, ist der Koeffizient von  $\lambda^{n-1}$  unabhängig von der Basis – die **Spur der linearen Abbildung** Spur  $F$ .

Warum die Spur eingeführt wird, sehe ich noch nicht. Viele Eigenschaften der Spur werden hier angegeben: <https://studyflix.de/mathematik/spur-einer-matrix-2340> oder noch ausführlicher hier:

[https://de.wikipedia.org/wiki/Spur\\_\(Mathematik\)](https://de.wikipedia.org/wiki/Spur_(Mathematik)) – Kann sein, dass die Spur später im Skript verwendet wird.

Nützliche Eigenschaften im Zusammenhang mit Eigenwerten

Will man Eigenwerte berechnen, so ist es häufig nützlich, wenn man ein paar Eigenschaften darüber kennt. Daher sollen im Folgenden ein paar derer aufgezählt werden.

- Sei  $\lambda$  ein Eigenwert der invertierbaren Matrix  $A$  mit dem Eigenvektor  $v$ . Dann ist  $\frac{1}{\lambda}$  ein Eigenwert der inversen Matrix  $A^{-1}$  zum Eigenvektor  $v$ .
- Seien  $\lambda_i$  die Eigenwerte der Matrix  $A \in \mathbb{C}^{n \times n}$ . Dann gilt:
  - $\sum_{i=1}^n \lambda_i = \text{Spur}(A)$
  - $\prod_{i=1}^n \lambda_i = \det A$ 
    - Insbesondere gilt: Hat das charakteristische Polynom eine Nullstelle bei 0, dann ist die Determinante 0 und die Matrix ist nicht invertierbar.
- Ist  $\lambda$  ein Eigenwert der Matrix  $A$ , so ist er auch Eigenwert der transponierten Matrix  $A^T$ . Daher stimmen die Spektren von  $A$  und  $A^T$  überein.
- Jeder Eigenwert einer reellen symmetrischen Matrix ist reell. Im Allgemeinen können aber auch komplexe Eigenwerte durchaus auftreten.

Weitere nützliche Links:

- <https://mathepedia.de/Eigenwerte.html>
- <https://studyflix.de/mathematik/eigenwert-1635>

## Diagonalisierung von Matrizen

Im Skript wird ausgiebig das Thema der Diagonalisierbarkeit von Matrizen untersucht. Dieses ist schon ein wichtiges Thema, denn kann man eine Matrix diagonalisieren, so lassen sich viele Fragen und Probleme schnell beantworten:

- Die Determinante ist das Produkt der Einträge auf der Hauptdiagonalen.
- Der Rang der Matrix ist gleich der Nichtrnull-Zeilen.
- Die Eigenwerte sind die Einträge der Hauptdiagonalen mit den kanonischen Einheitsvektoren als Eigenvektoren.

Natürlich fallen viele Rechnungen leichter aus, sollte die Matrix eine Diagonalgestalt haben, z. B. Matrixmultiplikationen (Matrizenmultiplikation), Skalarmultiplikation, Matrixaddition und Matrixsubtraktion, inverse Matrix oder transponierte Matrix berechnen. Dieses kann auch für computergestützte Berechnungen große Rechenzeitvorteil haben.

Es stellt sich die Frage, wann eine Matrix einer linearen Abbildung durch einen geeigneten Basiswechsel auf Diagonalgestalt gebracht werden kann.

Um es kurz zu machen, bevor ich mir das Skript näher anschau, diese Zusammenfassung:

Eine  $n \times n$ -Matrix ist genau dann diagonalisierbar, wenn

- es  $n$  paarweise linear unabhängige Eigenvektoren gibt (diese bilden dann eine Basis, so dass durch den Basiswechsel die Darstellungsmatrix Diagonalgestalt bekommt)
- das charakteristische Polynom vollständig in Linearfaktoren zerfällt und für jeden Eigenwert entspricht die geometrische Vielfachheit der algebraischen Vielfachheit (siehe unten). (Sind es  $n$  paarweise verschiedene Eigenwerte, so ist dies automatisch erfüllt.)

(Es gibt noch weitere Bedingungen, so dass die Matrix diagonalisierbar ist.)

Hinweis:

- **geometrische Vielfachheit:** Anzahl linear unabhängiger Eigenvektoren
- **algebraische Vielfachheit:** Die Anzahl der Eigenwerte, wobei die Vielfachheit der Nullstellen berücksichtigt werden muss.

Typisches Vorgehen zum Diagonalisieren einer  $n \times n$ -Matrix  $A$

- (1) Bestimme alle Eigenwerte von  $A$ , d. h. bestimme die Nullstellen des charakteristischen Polynoms  $\det(A - \lambda I)$ .  
Hat  $A$   $n$  paarweise verschiedene Eigenwerte, dann ist  $A$  diagonalisierbar.  
→ Wenn nicht, dann erkennt man erst in Schritt (2), ob  $A$  diagonalisierbar ist oder nicht.
- (2) Bestimme zu jedem Eigenwert eine Basis des Eigenraums (bzw. eine maximale Anzahl linear unabhängiger Eigenvektoren). Hat  $A$  insgesamt  $n$  linear unabhängige Eigenvektoren  $u_1, \dots, u_n$ , dann ist  $A$  diagonalisierbar.
- (3) Schreibe die Eigenvektoren  $u_1, \dots, u_n$  als Spalten in eine Matrix  $P^{-1}$ .
- (4) Das Matrixprodukt  $D = PAP^{-1}$  ist eine Diagonalmatrix mit den Diagonalelementen  $\lambda_1, \dots, \lambda_n$ , wobei  $\lambda_i$  der Eigenwert zum Eigenvektor  $u_i$  ist. Sofern nur  $D$  gesucht ist, kann man die Diagonalmatrix  $D$  mit den Eigenwerten direkt hinschreiben. **Hinweis:** Das Skript verwendet die gleichwertige Schreibweise:  $D = S^{-1}AS$ , wobei  $S$  dann aus den Eigenvektoren besteht.

Beispiel:

$$A = \begin{pmatrix} 0 & 0 & -2 \\ 1 & 2 & 1 \\ 1 & 0 & 3 \end{pmatrix}$$

- (1) Bestimmung der Eigenwerte durch Lösen von  $\det(A - \lambda E) = 0$ :

$$\det(A - \lambda E) = \det \begin{pmatrix} -\lambda & 0 & -2 \\ 1 & 2 - \lambda & 1 \\ 1 & 0 & 3 - \lambda \end{pmatrix}$$

$$= -\lambda(2 - \lambda)(3 - \lambda) + 0 + 0 - 1(2 - \lambda)(-2) - 0 - 0 = -\lambda^3 + 5\lambda^2 - 8\lambda + 4 = 0$$

Mit Taschenrechner / CAS lösen liefert die beiden Lösungen  $\lambda = 1$  und  $\lambda = 2$ . Dabei hat der Eigenwert 2 die Vielfachheit 2, d. h. das charakteristische Polynom könnte man so faktorisieren:  $-\lambda^3 + 5\lambda^2 - 8\lambda + 4 = -(\lambda - 2)^2(\lambda - 1)$

(Ansonsten: Wenn ganzzahlige Lösungen vorhanden sind, können es nur Teiler von 4 sein! → Raten, Polynomdivision usw.)

Da wir nur zwei und nicht drei verschiedene Eigenwerte haben, müssen wir den Schritt (2) ausführen, um die geometrische Vielfachheit zu erkennen.

- (2) Bestimmung der Eigenvektoren  $u_i$  durch Lösen von

$(A - \lambda_i I) \cdot u_i = 0$   
für jeden Eigenwert.

$\lambda = 1$ :

$$(A - I) \cdot u_1 = 0$$

$$\begin{pmatrix} -1 & 0 & -2 \\ 1 & 1 & 1 \\ 1 & 0 & 2 \end{pmatrix} \cdot \begin{pmatrix} u_{11} \\ u_{21} \\ u_{31} \end{pmatrix} = 0$$

$$\begin{vmatrix} -u_{11} + 0 - 2u_{31} & = & 0 \\ u_{11} + u_{21} + u_{31} & = & 0 \\ u_{11} + 0 + 2u_{31} & = & 0 \end{vmatrix}$$

bzw. als erweiterte Koeffizientenmatrix für das Gaußverfahren:

$$\begin{pmatrix} -1 & 0 & -2 & 0 \\ 1 & 1 & 1 & 0 \\ 1 & 0 & 2 & 0 \end{pmatrix}$$

$$\begin{pmatrix} -1 & 0 & -2 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

$\Rightarrow$  Eine Variable frei wählbar. Es gilt  $u_{21} - u_{31} = 0 \Rightarrow u_{21} = u_{31}$  und  $-u_{11} - 2u_{31} = 0$   
 $\Rightarrow u_{11} = -2u_{31}$

Damit erhalten wir mit der Wahl von  $u_{31} = 1$  einen Eigenvektor  $u_1 = \begin{pmatrix} -2 \\ 1 \\ 1 \end{pmatrix}$ .

$\lambda = 2$ :

$$(A - I) \cdot u_2 = 0$$

$$\begin{pmatrix} -2 & 0 & -2 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} u_{12} \\ u_{22} \\ u_{32} \end{pmatrix} = 0$$

$$\begin{vmatrix} -2u_{12} + 0 - 2u_{32} & = & 0 \\ u_{12} + 0 + u_{32} & = & 0 \\ u_{12} + 0 + u_{32} & = & 0 \end{vmatrix}$$

$\Rightarrow$  Das LGS ist nicht eindeutig lösbar, da alle drei Gleichungen letztendlich auf eine reduziert werden können:  $u_{12} + 0 + u_{32} = 0$ . Damit folgt:  $u_{12} = -u_{32}$ . Wählt man  $u_{32} = 1$ , dann ist  $u_{12} = -1$  und  $u_{22}$  ist beliebig, z. B. 0. Ein Eigenvektor ist also  $u_2 = \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}$ . Da  $u_{22}$  frei wählbar ist, gibt es noch einen weiteren Eigenvektor  $u_3 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$  eine Lösung der Gleichung ist.

Damit wurden drei linear unabhängige Eigenvektoren gefunden. Die Matrix  $A$  ist diagonalisierbar und es gilt:

$$P^{-1} = \begin{pmatrix} -2 & -1 & 0 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}$$

$$\Rightarrow P = \begin{pmatrix} -1 & 0 & -1 \\ 1 & 0 & 2 \\ 1 & 1 & 1 \end{pmatrix}$$

$$PAP^{-1} = D = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}$$

**Beispiele von Matrizen, die nicht diagonalisierbar sind – einfach mal durchrechnen:**

$$M = \begin{pmatrix} 2 & 4 & 0 \\ -4 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

→ Ergebnis: Zu wenig Eigenwerte!

$$M = \begin{pmatrix} 3 & 2 & 3 \\ 4 & 7 & 9 \\ -3 & -4 & -5 \end{pmatrix}$$

→ Ergebnis: Genügend Eigenwerte, aber die geometrischen Vielfachheiten passen nicht!

Genauer nachzulesen hier: <https://studyflix.de/mathematik/matrix-diagonalisieren-3359>

Mathematische Grundlagen / Theorie aus dem Skript zum Diagonalisieren  
Mit entsprechenden Beweisen auf den Seiten 255-262.

#### Lemma 9.4.1: Notwendige Bedingung

Der  $\mathbb{C}$ - oder  $\mathbb{R}$ -Vektorraum  $V$  besitze eine Basis aus Eigenvektoren, dann ist das charakteristische Polynom von  $A$  ein Produkt von  $n$  Linearfaktoren.

#### Hinreichende Bedingung 9.4.3

Das charakteristische Polynom ist Produkt von  $n$  verschiedenen Linearfaktoren. Dazu äquivalent ist: Das charakteristische Polynom hat  $n$

verschiedene Nullstellen.

#### Satz 9.4.4

Die Eigenvektoren  $x_1, \dots, x_r$  von  $A$  zu paarweise verschiedenen Eigenwerten  $\lambda_1, \dots, \lambda_r$  sind linear unabhängig.

#### Definition 9.4.5: Ähnlichkeit

Zwei  $n$ -reihige Matrizen  $A$  und  $A'$  heißen *ähnlich*, falls es eine invertierbare  $n$ -reihige Matrix  $B$  gibt mit

$$B^{-1}AB = A'.$$

Ist  $A$  zu einer Diagonalmatrix ähnlich, so heißt  $A$  *diagonalisierbar*.

#### Lemma 9.4.6: Gestalt der Transformationsmatrix

Für eine invertierbare Matrix  $B$  gilt

$$B^{-1}AB = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$$

genau dann, wenn die Spalten  $b_k$  von  $B$  Eigenvektoren von  $A$  zum Eigenwert  $\lambda_k$  sind.

#### Definition 9.4.8: Vielfachheiten und Eigenraum

Es sei  $A \in M(n \times n, \mathbb{K})$  mit  $\mathbb{K} = \mathbb{R}$  oder  $\mathbb{K} = \mathbb{C}$  und  $\lambda_0 \in \mathbb{K}$  eine  $m$ -fache Nullstelle von  $P_A(\lambda) = \det(A - \lambda I)$ , dann heißt

- (i)  $m$  die *algebraische Vielfachheit* von  $\lambda_0$ ,
- (ii)  $\dim \text{Ker}(A - \lambda_0 I) =: \dim N_{\lambda_0}$  *geometrische Vielfachheit* von  $\lambda_0$  und
- (iii)  $\text{Ker}(A - \lambda_0 I) = \{ v \mid Av = \lambda_0 v \} = N_{\lambda_0}$  der *Eigenraum* von  $A$  zu  $\lambda_0$ .

#### Lemma 9.4.9

Es seien  $\mathbb{K} = \mathbb{R}$  oder  $\mathbb{K} = \mathbb{C}$ ,  $A \in M(n \times n, \mathbb{K})$  und  $\lambda_0$  ein Eigenwert von  $A$ , dann gilt: Die geometrische Vielfachheit von  $\lambda_0$  ist kleiner oder gleich der algebraischen Vielfachheit.

## Stetige Funktionen

Wie auch bei anderen Stellen, versuche ich die Inhalte eher kompakt und anschaulich zu beschreiben. Dieses Mini-Skript bleibt daher unvollständig und hat nicht den Anspruch eine lückenlose, auf Axiomen gegründete Herleitung der Inhalte zu präsentieren. Diese wird in echten Skripten und Lehrbüchern geboten.

Stetige Funktionen sind uns aus der Schule gut vertraut, zumal dort überwiegend stetige Funktionen betrachtet werden. Seltene Ausnahmen bilden vielleicht gebrochenrationale Funktionen mit Termen wie  $f(x) = \frac{1}{x}$ , die mal eine Polstelle oder eine sogenannte hebbare Definitionslücke haben, wie z. B.  $g(x) = \frac{x+1}{x^2-1}$ . Hier ist die Definitionslücke bei  $x = -1$  hebbbar, die bei  $x = 1$  aber nicht. Das Thema gebrochenrationale Funktionen, Polstellen, Asymptoten usw. hat in den letzten 15 Jahren jedoch in

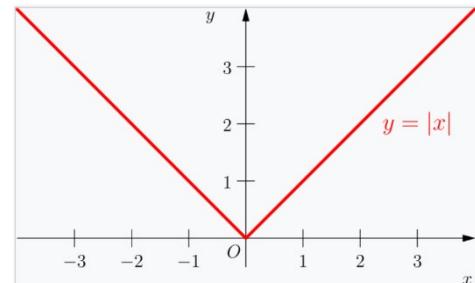
der Schulmathematik ebenfalls an Bedeutung verloren. Was bleibt? Überwiegend harmlose Funktionen, mit denen man fast bedenkenlos alles Schulische machen kann.

### **Was heißt nun stetig? Und was differenzierbar?**

Stetigkeit anschaulich: Jeder Graph, den ich ohne den Stift abzusetzen in einem Abschnitt zeichnen kann, gehört zu einer in dem entsprechenden Intervall stetigen Funktion. Es sind keine Lücken und keine Polstellen vorhanden. Die Funktionswerte sind endlich. Das Zeichnen mit dem Stift ohne ihn abzusetzen, wird später mathematisch mittels einer  $\varepsilon$ - $\delta$ -Definition konkretisiert. Im Kern geht es darum, dass bei stetigen Funktionen der Abstand der Funktionswerte an einer Stelle  $x_0$  zu einer Stelle  $x$  immer kleiner wird – verschwindet klein, je näher sich  $x$  an  $x_0$  annähert. Da das für alle Stellen in dem Intervall gelten, wo die Funktion stetig ist, hängen die zugehörigen Punkte  $(x_0 | f(x_0))$  beliebig dicht im Intervall zusammen.

Differenzierbarkeit: In der Schule wird über die Steigung des Graphen an einer bestimmten Stelle gesprochen. Um diese zu bestimmen, wird erst eine Sekante betrachtet, deren einer Schnittpunkt dem Punkt  $(x_0 | f(x_0))$  entspricht, an dem man die Steigung berechnen möchte. Der zweite Schnittpunkt wird in der Regel ein kleines Stückchen weiter rechts gewählt, typischerweise bei  $(x_0 + h | f(x_0 + h))$ . Nun wird mittels Steigungsdiagramm die Sekantensteigung ( $\rightarrow$  Differenzenquotient) bestimmt und diese für den Grenzwertprozess  $h \rightarrow 0$  untersucht. Existiert der Grenzwert, so nennt man diesen Wert die Steigung in dem Punkt. Das allgemeine Betrachten dieser Grenzwertprozesse führt auf die Ableitungsfunktionen. So wird in der Schule die Differentialrechnung eingeführt. Die Frage, ob eine Funktion überhaupt differenzierbar ist, wird oft nur darauf beschränkt, ob die Funktion an der Stelle definiert ist, also dort stetig ist. Vielleicht wird auch kurz darauf hingewiesen, dass stetige Funktionen nicht immer differenzierbar sind und dass die Stetigkeit Voraussetzung dafür ist, dass man überhaupt von einer Steigung sprechen kann.

Differenzierbarkeit anschaulich: Natürlich kann man nicht von einer Steigung sprechen, wenn die Funktion an einer Stelle nicht definiert ist oder einen Sprung macht oder der Funktionswert nicht endlich ist. Aber wir müssen genauer sein: Eine Funktion ist differenzierbar, wenn sie stetig und knickfrei ist. Man kann einen Graphen ohne Absetzen zeichnen, der jedoch einen Knick hat. Standardbeispiel ist die Betragsfunktion mit  $f(x) = |x|$ . Diese Funktion ist stetig, aber der Graph hat an der Stelle  $x = 0$  einen Knick.



Hier passiert etwas, was bei der Differenzierbarkeit nicht passieren darf. Wenn man die Steigung wie oben beschrieben mittels Grenzwertbetrachtung des Differenzenquotienten untersucht, dann existiert zwar bei der Betragsfunktion an jeder Stelle ein Grenzwert, aber wenn man sich der Stelle  $x = 0$  von rechts nähert, also  $x > 0$  vorgibt, dann ist der Grenzwert (die Steigung) stets 1, während wenn man sich der Null von links nähert, der Grenzwert stets -1 ist. Das zeichnet einen Knick im Graphen aus. Die Steigung an der Knickstelle ist nicht eindeutig. Was sollte man wählen? -1? +1? Oder den Mittelwert 0? Letzteres würde bedeuten, dass die Tangente an der Spitze der Betragsfunktion waagerecht verlaufen solle, was aber fragwürdig erscheint. Daher nennt man die Funktion an dieser Stelle nicht differenzierbar, obgleich der Grenzwert „rechtsseitig“ und „linksseitig“ existiert, jedoch die Werte verschieden sind. So wird später auch Differenzierbarkeit mathematisch formal definiert: Wenn der rechts- und linksseitige Grenzwert des Differenzenquotienten existieren und die Grenzwerte übereinstimmen, dann ist die Funktion an der Stelle differenzierbar.

Soweit der erste Überblick. Vieles, was nun folgt, greift auf das Thema Konvergenz und Folgen zurück und ich hoffe, dass daher die Übertragung auf Funktionen mit ihren Funktionswerten gut verständlich ist.

Ein paar Begriffe:

#### Definition 10.1.1: $\varepsilon$ -Umgebung

Für  $x_0 \in \mathbb{R}$  und  $\varepsilon > 0$  heißt

$$U_\varepsilon(x_0) = \{x \in \mathbb{R} \mid |x - x_0| < \varepsilon\} = (x_0 - \varepsilon, x_0 + \varepsilon)$$

die  $\varepsilon$ -Umgebung von  $x_0$  und

$$\begin{aligned}\dot{U}_\varepsilon(x_0) &= \{x \in \mathbb{R} \mid 0 < |x - x_0| < \varepsilon\} \\ &= U_\varepsilon(x_0) \setminus \{x_0\} = (x_0 - \varepsilon, x_0) \cup (x_0, x_0 + \varepsilon)\end{aligned}$$

die punktierte  $\varepsilon$ -Umgebung von  $x_0$ .

Der Unterschied zwischen einer Umgebung und einer punktierten Umgebung von  $x_0$  ist, dass  $x_0$  in der Umgebung enthalten ist und in der punktierten Umgebung fehlt. Beides beschreibt einen Zahlenbereich um eine Zahl  $x_0$ , jedoch fehlt die Zahl in der punktierten Umgebung.

## Funktionsgrenzwerte

#### Definition 10.1.3: Funktionsgrenzwert

Es sei  $f : I \rightarrow \mathbb{R}$  eine Funktion.

- (i) Es sei  $x_0 \in I$ , dann heißt  $a \in \mathbb{R}$  der Grenzwert oder Limes von  $f$  an der Stelle  $x_0$  beziehungsweise wir sagen  $f$  konvergiert gegen  $a$  für  $x \rightarrow x_0$ , wenn

$$\forall \varepsilon > 0 \exists \delta > 0 \forall x \in \mathbb{R} (x \in \dot{U}_\delta(x_0) \Rightarrow |f(x) - a| < \varepsilon).$$

beziehungsweise analog mit  $x \in (x_0, x_0 + \delta) \cap I$  für den rechtsseitigen Grenzwert.

Wir schreiben  $\lim_{x \rightarrow x_0^-} f(x) = a$ ,  $\lim_{x \uparrow x_0} f(x) = a$  oder  $f(x) \rightarrow a$  für  $x \rightarrow x_0^-$  beziehungsweise  $\lim_{x \rightarrow x_0^+} f(x) = a$ ,  $\lim_{x \downarrow x_0} f(x) = a$  oder  $f(x) \rightarrow a$  für  $x \rightarrow x_0^+$ .

- (iii) Es sei  $I = (a, +\infty)$  beziehungsweise  $I = (-\infty, b)$  und  $c \in \mathbb{R}$ , dann konvergiert  $f$  gegen  $c$  für  $x \rightarrow +\infty$  beziehungsweise für  $x \rightarrow -\infty$ , wenn

$$\forall \varepsilon > 0 \exists x_1 \in \mathbb{R} \forall x \in \mathbb{R} (x > x_1 \Rightarrow |f(x) - c| < \varepsilon).$$

Wir schreiben  $\lim_{x \rightarrow +\infty} f(x) = c$  oder  $f(x) \rightarrow c$  für  $x \rightarrow +\infty$  beziehungsweise  $\lim_{x \rightarrow -\infty} f(x) = c$  oder  $f(x) \rightarrow c$  für  $x \rightarrow -\infty$ .

Dabei kann jeweils  $\delta$  beziehungsweise  $x_1$  von  $\varepsilon$  und  $x_0$  abhängen, das heißt  $\delta = \delta(\varepsilon, x_0)$  und  $x_1 = x_1(\varepsilon)$ .

Es werden nun Grenzwerte bei Funktionen in Analogie zu Grenzwerte bei Folgen eingeführt. Der Begriff „Konvergenz“ wird hier ebenfalls analog verwendet. Es gibt jedoch die Besonderheit, dass wir keine Folgenglieder und Indexnummern haben, sondern, dass wir Werte aus einer Umgebung und die zugehörigen Funktionswerte haben. Es gibt keine Indexnummer, die gegen Unendlich läuft, sondern es gibt  $x$ -Werte, die sich einem bestimmten Wert  $x_0$  annähern. Dieses kann mit  $x$ -Werten erfolgen, die größer oder kleiner als  $x_0$  sind, d. h. man kann sich der interessanten Stelle von links oder von rechts annähern.

Spricht man nur von der Konvergenz und nicht von der links- oder rechtseitigen Konvergenz, so meint man damit, dass der links- und rechtsseitige Grenzwert übereinstimmen. Eine weitere Besonderheit ist noch das  $\delta$ . Bei Folgen hieß es bei der Konvergenz: Zu jedem  $\varepsilon > 0$  gibt es ein  $N_0$ , so dass für alle  $n \geq N_0$  der Abstand zwischen Folgeglied und Grenzwert kleiner als  $\varepsilon$  ist. Das  $\delta$  entspricht der  $N_0$ : Zu jedem  $\varepsilon > 0$  gibt es ein  $\delta > 0$ , dass eine Umgebung um  $x_0$  festlegt, aus dem die  $x$ -Werte betrachtet werden, dass der Abstand zwischen Funktionswert und Grenzwert kleiner als  $\varepsilon$  ist. So wie in der Regel  $N_0$  von  $\varepsilon$  abhängig ist, ist  $\delta$  in der Regel auch von  $\varepsilon$  abhängig.

#### Lemma 10.1.4: Eindeutigkeit des Grenzwerts

Der Grenzwert einer Funktion ist eindeutig bestimmt.

Dies verwundert nicht, da ja bei Funktionen die Funktionswerte alle eindeutig bestimmt sind.

#### Satz 10.1.6: Folgenkriterium

Es seien  $f : I \rightarrow \mathbb{R}$ ,  $x_0 \in I$  und  $a \in \mathbb{R}$ , dann gilt  $\lim_{x \rightarrow x_0} f(x) = a$  genau dann, wenn  $\lim_{k \rightarrow \infty} f(x_k) = a$  für jede Folge  $(x_k)_{k \in \mathbb{N}} \subseteq I \setminus \{x_0\}$  mit  $x_k \rightarrow x_0$  für  $k \rightarrow \infty$ .

**Ein wichtiger Satz!** Beachte, dass es ein **genau-dann-wenn-Satz** ist, also er ist in beide Richtung gültig und eine Aussage über alle Folgen mit der ange-

geben Eigenschaft macht. Auch alle Teilstufen, die so konvergieren!

Eine weitere Analogie zu den Grenzwerten von Folgen:

#### Lemma 10.1.8: Grenzwertsätze für Funktionsgrenzwerte

Es seien  $f, g : I \rightarrow \mathbb{R}$ ,  $x_0 \in \bar{I}$  und  $a, b \in \mathbb{R}$  mit  $\lim_{x \rightarrow x_0} f(x) = a$ ,  $\lim_{x \rightarrow x_0} g(x) = b$ . Dann gilt

- (i)  $\forall \alpha \in \mathbb{R} \left( \lim_{x \rightarrow x_0} (\alpha f(x)) = \alpha a \right)$ .
- (ii)  $\lim_{x \rightarrow x_0} (f(x) + g(x)) = a + b$ .
- (iii)  $\lim_{x \rightarrow x_0} (f(x) \cdot g(x)) = a \cdot b$ .
- (iv) Wenn  $b \neq 0$ , dann gilt  $\lim_{x \rightarrow x_0} \frac{f(x)}{g(x)} = \frac{a}{b}$ .

Und noch eine – wie bei Folgen – wenn man den Grenzwert nicht kennt:

#### Lemma 10.1.10: Cauchy-Kriterium für Funktionsgrenzwerte

Es seien  $f : I \rightarrow \mathbb{R}$  und  $x_0 \in I$ , dann existiert  $\lim_{x \rightarrow x_0} f(x)$  genau dann, wenn

$$\forall \varepsilon > 0 \exists \delta > 0 \forall x_1, x_2 \in I (x_1, x_2 \in U_\delta(x_0) \Rightarrow |f(x_1) - f(x_2)| < \varepsilon).$$

Noch ein paar Vokabeln:

#### Definition 10.1.12: Bestimmte Divergenz

Es seien  $f : I \rightarrow \mathbb{R}$  und  $x_0 \in \bar{I}$ . Dann divergiert  $f$  in  $x_0$  bestimmt gegen  $+\infty$ , in Zeichen  $\lim_{x \rightarrow x_0} f(x) = +\infty$ , wenn

$$\forall c > 0 \exists \delta > 0 \forall x \in \mathbb{R} (x \in U_\delta(x_0) \Rightarrow f(x) \geq c).$$

Außerdem ist

$$\lim_{x \rightarrow +\infty} f(x) = +\infty \Leftrightarrow \forall c > 0 \exists x_1 \in \mathbb{R} \forall x \in \mathbb{R} (x \geq x_1 \Rightarrow f(x) \geq c).$$

Analog sind die einseitigen Grenzwerte  $\lim_{x \rightarrow x_0^+} f(x) = +\infty$ ,  $\lim_{x \rightarrow x_0^-} f(x) = +\infty$  sowie die bestimmte Divergenz gegen  $-\infty$  definiert. Statt von bestimmter Divergenz spricht man auch von *uneigentlicher Konvergenz*.

**In Kurzform:** Geht der Grenzwert gegen  $+\infty$  oder  $-\infty$ , dann liegt eine bestimmte Divergenz vor. Kann man nicht sagen, welcher Grenzwert angenommen wird, z. B. durch Oszillation, ist es eine unbestimmte Divergenz.

Ich bin mir nicht sicher, was ist, wenn der links- und rechtsseitige Grenzwert gegen  $\infty$  gehen, aber mit verschiedenen Vorzeichen.

#### Definition 10.1.14: Beschränktheit von Funktionen

Es seien  $I$  ein beliebiges Intervall und  $f : I \rightarrow \mathbb{R}$ , dann heißt  $f$  auf  $I$  **beschränkt** : $\Leftrightarrow \exists c > 0 (x \in I \Rightarrow |f(x)| \leq c)$ .

#### Lemma 10.1.15

Es seien  $a, b \in \mathbb{R}$ ,  $a < b$  und  $f : (a, b) \rightarrow \mathbb{R}$  monoton und beschränkt, dann existiert  $\lim_{x \rightarrow b^-} f(x)$  und  $\lim_{x \rightarrow a^+} f(x)$ .

### Stetigkeit

Durch die Definitionen und Sätze im vorigen Abschnitt kann man nun die Stetigkeit mathematisch formulieren. Was heißt es, ohne den Stift absetzen zu können, den Graphen zu zeichnen? Der Graph besteht aus Punkten  $(x | f(x))$ . Diese Punkte liegen „unendlich dicht zusammen“, d. h. es gibt keine Lücken und die Werte bleiben endlich groß. Dieses wird ausgedrückt durch:

#### Definition 10.2.1: Stetigkeit

Es sei  $D \subseteq \mathbb{R}$ . Eine Funktion  $f : D \rightarrow \mathbb{R}$  heißt **stetig in  $x_0 \in D$**

$$:\Leftrightarrow \forall \varepsilon > 0 \exists \delta > 0 (x \in U_\delta(x_0) \cap D \Rightarrow |f(x) - f(x_0)| < \varepsilon).$$

$f$  heißt **stetig auf  $D$**  : $\Leftrightarrow \forall x_0 \in D (f \text{ stetig in } x_0)$ .

Die Menge aller auf  $D$  stetigen Funktionen wird mit  $C(D)$  bezeichnet.

Wenn ich den Abstand zwischen den Funktionswerten  $f(x)$  und den Funktionswert  $f(x_0)$  betrachte, dann kann ich immer eine  $\delta$ -Umgebung um  $x_0$  finden, so dass alle Funktionswert  $f(x)$  mit  $x$  aus der Umgebung, noch näher an  $f(x_0)$  liegen, als jedes zuvor ausgewählte  $\varepsilon > 0$ .

Stetige Funktionen sind etwas Feines. Es gilt:

#### Lemma 10.2.3: Operationen mit stetigen Funktionen

- (i) Es seien  $D \subseteq \mathbb{R}$ ,  $f, g : D \rightarrow \mathbb{R}$ ,  $x_0 \in D$  und  $c \in \mathbb{R}$ . Wenn  $f$  und  $g$  in  $x_0$  stetig sind, dann sind auch die Funktionen  $c \cdot f$ ,  $f + g$  und  $f \cdot g$  in  $x_0$  stetig. Gilt zusätzlich  $g(x_0) \neq 0$ , dann ist auch  $\frac{f}{g}$  in  $x_0$  stetig.
- (ii) Es seien  $D_1, D_2 \subseteq \mathbb{R}$ ,  $f : D_1 \rightarrow D_2$ ,  $g : D_2 \rightarrow \mathbb{R}$  und  $x_0 \in D_1$ . Wenn  $f$  in  $x_0$  stetig ist und  $g$  in  $f(x_0)$  stetig ist, dann ist auch  $g \circ f$  in  $x_0$  stetig.

#### Lemma 10.2.5

Es sei  $I$  ein Intervall und  $f : I \rightarrow \mathbb{R}$  stetig in  $x_0 \in I$ , dann gilt

$$f(x_0) > 0 \Rightarrow \exists \delta > 0 \forall x \in U_\delta(x_0) \cap I (f(x) > 0).$$

#### Zwischenwertsatz 10.2.7

Es seien  $I = [a, b]$  ein kompaktes Intervall und  $f : I \rightarrow \mathbb{R}$  stetig mit  $f(a) < f(b)$ . Dann existiert für jeden Zwischenwert  $y \in (f(a), f(b))$  ein  $\xi \in (a, b)$  mit  $f(\xi) = y$ .

Wegen der Stetigkeit (Lückenlosigkeit), muss es für jeden Zwischenwert eine Stelle geben, an dem die Funktion diesen

Zwischenwert annimmt. Geht ja nicht anders. Wir zeichnen, ohne den Stift abzusetzen!

Hinweis: Ein Intervall heißt **kompakt**, wenn es abgeschlossene Grenzen hat, also so geschrieben wird:  $[a, b]$ . Die Intervalle  $[a, b)$  oder  $(-\infty, b]$  sind nicht kompakt, sondern haben mind. eine offene Grenze)

#### Definition 10.2.9: Maximum, Minimum, Supremum, Infimum einer Funktion

Es seien  $D \subseteq \mathbb{R}$  und  $f : D \rightarrow \mathbb{R}$ . Falls existent, heißt

- (i)  $\max_{x \in D} f(x) = \max_D f(x) = \max\{f(x) \mid x \in D\}$  das *Maximum* von  $f$  in  $D$ ,
- (ii)  $\min_{x \in D} f(x) = \min_D f(x) = \min\{f(x) \mid x \in D\}$  das *Minimum* von  $f$  in  $D$ ,
- (iii)  $\sup_{x \in D} f(x) = \sup_D f(x) = \sup\{f(x) \mid x \in D\}$  das *Supremum* von  $f$  in  $D$ ,
- (iv)  $\inf_{x \in D} f(x) = \inf_D f(x) = \inf\{f(x) \mid x \in D\}$  das *Infimum* von  $f$  in  $D$ .

#### Lemma 10.2.11

Es seien  $I = [a, b]$  ein kompaktes Intervall und  $f : I \rightarrow \mathbb{R}$  stetig, dann ist  $f$  beschränkt.

Stetigkeit impliziert endlich große Werte und das impliziert die Beschränktheit.

#### Satz vom Minimum und Maximum (Weierstraß) 10.2.14

Es seien  $I = [a, b]$  ein kompaktes Intervall und  $f : I \rightarrow \mathbb{R}$  stetig, dann existieren zwei Punkte  $x^+, x^- \in I$  mit

$$\forall x \in I (f(x^-) \leq f(x) \leq f(x^+)).$$

Mit anderen Worten, es gilt

$$f(x^-) = \inf_I f(x) = \min_I f(x), \quad f(x^+) = \sup_I f(x) = \max_I f(x).$$

Es gibt bei stetigen Funktionen immer einen kleinsten oder größten Wert auf einem Intervall, wobei Gleichheit zugelassen wird.

#### Korollar 10.2.15: Bild eines kompakten Intervalls unter einer stetigen Funktion II

Es seien  $a, b \in \mathbb{R}$ ,  $a < b$ . Für eine stetige Funktion  $f : [a, b] \rightarrow \mathbb{R}$  mit  $y_1 = \min_{[a,b]} f(x)$  und  $y_2 = \max_{[a,b]} f(x)$  gilt  $f([a, b]) = [y_1, y_2]$ .

Soll heißen, dass wenn man ein Intervall  $[a, b]$  mit einer stetigen Funktion abbildet, dann erhält man ein kompaktes Intervall  $[y_1, y_2]$  der Funktionswerte – ohne Lücken oder so.

Bin mir nicht sicher, wie wichtig der folgende Abschnitt ist:

#### Definition 10.2.17: Gleichmäßige Stetigkeit

Es sei  $D \subseteq \mathbb{R}$ . Eine Funktion  $f : D \rightarrow \mathbb{R}$  heißt *gleichmäßig stetig* auf  $D$  : $\Leftrightarrow$

$$\forall \varepsilon > 0 \exists \delta > 0 \forall x_1, x_2 \in D (|x_1 - x_2| < \delta \Rightarrow |f(x_1) - f(x_2)| < \varepsilon).$$

Eine Funktion  $f : \mathbb{R} \rightarrow \mathbb{R}$  heißt **Lipschitz-stetig**, falls es ein  $L \in \mathbb{R}$  gibt, sodass

$$|f(x_1) - f(x_2)| \leq L|x_1 - x_2| \quad \forall x_1, x_2 \in \mathbb{R}.$$

**Bemerkungen 10.2.18.** (i)  $\delta = \delta(\varepsilon)$  kann bei einer gleichmäßig stetigen Funktion unabhängig von  $x_0 \in D$  gewählt werden. Gewöhnliche Stetigkeit auf  $D$  bedeutet

$$\forall x_0 \in D \forall \varepsilon > 0 \exists \delta > 0 \forall x \in D (|x_0 - x| < \delta \Rightarrow |f(x) - f(x_0)| < \varepsilon).$$

(ii) Jede gleichmäßig stetige Funktion ist stetig, die Umkehrung ist im Allgemeinen falsch.

**Beispiele 10.2.19.** (i)  $f : (0, 1) \rightarrow \mathbb{R}$ ,  $f(x) = x^2$  ist gleichmäßig stetig: Es sei  $\varepsilon > 0$ . Wir setzen  $\delta = \frac{\varepsilon}{2}$ , dann gilt

$$\begin{aligned} |f(x_1) - f(x_2)| &= |x_1^2 - x_2^2| = |(x_1 - x_2)(x_1 + x_2)| \\ &= (x_1 + x_2)|x_1 - x_2| < 2|x_1 - x_2| < \varepsilon, \end{aligned}$$

falls  $|x_1 - x_2| < \frac{\varepsilon}{2}$ , woraus sich die Wahl von  $\delta$  erklärt.

(ii)  $f : (0, +\infty) \rightarrow \mathbb{R}$ ,  $f(x) = \frac{1}{x}$  ist nicht gleichmäßig stetig, denn sei zum Beispiel  $\varepsilon_0 = 1$ , dann ist für  $\delta \in (0, 1)$  beliebig und  $x_1 = \delta$ ,  $x_2 = \frac{\delta}{2}$

$$|x_1 - x_2| < \delta \wedge |f(x_1) - f(x_2)| = \frac{1}{\delta} > 1 = \varepsilon_0.$$

Dass hier  $\delta$  eingeschränkt wurde, ist unerheblich: Für  $\delta \geq 1$  wählt man einfach  $x_1$  und  $x_2$  wie oben für ein  $\delta' \in (0, 1)$ , da aus  $|x_1 - x_2| < \delta'$  dann auch  $|x_1 - x_2| < \delta$  folgt.

Bei kompakten Intervallen gilt:

#### Satz 10.2.20

Jede auf  $I = [a, b]$  stetige Funktion ist gleichmäßig stetig.

Mag nützlich für Aufgaben sein, bei denen die Stetigkeit gezeigt werden soll.

#### Definition 10.2.21: Konvergenz von Funktionenfolgen

Es seien  $D \subseteq \mathbb{R}$  und  $f_k : D \rightarrow \mathbb{R}$  für  $k \in \mathbb{N}$ .

- (i) Die Funktionenfolge  $(f_k)_{k \in \mathbb{N}}$  konvergiert punktweise gegen eine Funktion  $f : D \rightarrow \mathbb{R}$ , falls für alle  $x \in D$ :  $f_k(x) \rightarrow f(x)$  für  $k \rightarrow \infty$ .
- (ii) Die Funktionenfolge  $(f_k)_{k \in \mathbb{N}}$  heißt gleichmäßig konvergent:

$$f_k \xrightarrow{\text{glm.}} f \text{ für } k \rightarrow \infty$$

$$\Leftrightarrow \forall \varepsilon > 0 \exists N \in \mathbb{N} \forall x \in D \forall k \in \mathbb{N} (k \geq N \Rightarrow |f_k(x) - f(x)| < \varepsilon).$$

Der folgende Abschnitt läuft auf Konvergenzbe trachtungen hinaus. Erinnere dich an die Konvergenzbetrachtungen von Reihen mit verschiedenen Kriterien, wie z. B. das Leibniz-Kriterium.

So etwas kann auch für Folgen von Funktionen bzw. Reihen Funktionsfolgen von betrachtet werden.

#### Satz 10.2.23: Eigenschaften der Grenzfunktion

Es seien  $I$  ein Intervall und  $(f_k)_{k \in \mathbb{N}}$  eine auf  $I$  gleichmäßig konvergente Folge (gleichmäßig) stetiger Funktionen, dann ist die Grenzfunktion ebenfalls (gleichmäßig) stetig.

### Weierstraßsches Majorantenkriterium 10.2.25

Für  $k \in \mathbb{N}_0$  seien  $f_k : D \rightarrow \mathbb{R}$  mit  $|f_k(x)| \leq M_k$  für alle  $x \in D$  und  $\sum_{k=0}^{\infty} M_k$  sei konvergent, dann ist  $\sum_{k=0}^{\infty} f_k(x)$  gleichmäßig konvergent auf  $D$ .

„gleichmäßig konvergent“  
– ähm. Ja?  
Andere schreiben sogar  
auch „absolut konver-  
gent“!

Ich glaube, dass der Prof diesen Abschnitt vollständigkeitshalber angegeben hat und eher in einem Mathematikstudium seinen Platz hat. Eine Beispielaufgabe, falls dich das weiter interessiert:

Untersuche die Funktionenreihe

$$\sum_{n=0}^{\infty} \frac{n^2}{\sqrt{n!}} \left( x^n + \frac{1}{x^n} \right), \quad x \in \left[ \frac{1}{2}, 2 \right]$$

auf punktweise und gleichmäßige Konvergenz. Falls die Reihe konvergiert: Ist die dadurch gegebene Funktion stetig? Ist sie differenzierbar? Gebe gegebenenfalls die Ableitung an.

**Lösung:** Der Schlüssel zur Lösung der Aufgabe ist der Satz auf Seite 142 im Skript, das Weierstraß'sche Majorantenkriterium für Funktionenreihen, auch *Weierstraß M-Test* genannt. Wir schätzen für alle  $x \in [\frac{1}{2}, 2]$  ab:

$$\left| \frac{n^2}{\sqrt{n!}} \left( x^n + \frac{1}{x^n} \right) \right| \leq \frac{n^2}{\sqrt{n!}} \left( 2^n + \frac{1}{(1/2)^n} \right) = \frac{n^2}{\sqrt{n!}} (2^n + 2^n) = \frac{n^2}{\sqrt{n!}} 2^{n+1}.$$

Zur Abkürzung schreiben wir nun  $c_n := \frac{n^2}{\sqrt{n!}} 2^{n+1}$  und müssen nur noch zeigen, dass  $\sum_{n=0}^{\infty} c_n$  konvergiert, dann folgt mit dem oben genannten Satz die gleichmäßige Konvergenz der Funktionenreihe. Wir benutzen hierfür das Quotientenkriterium und sehen dass

$$\left| \frac{c_{n+1}}{c_n} \right| = \left( \frac{n+1}{n} \right)^2 \frac{2^{n+2}}{2^{n+1}} \sqrt{\frac{n!}{(n+1)!}} = \left( \frac{n+1}{n} \right)^2 \cdot 2 \cdot \frac{1}{\sqrt{n+1}} \rightarrow 0 < 1,$$

für  $n \rightarrow \infty$ . Daraus folgt, dass  $\sum_{n=0}^{\infty} c_n$  konvergiert und wir haben bewiesen, dass die Reihe gleichmäßig konvergiert. Also konvergiert sie auch punktweise. Des weiteren ist die durch die Reihe definierte Funktion stetig als Grenzfunktion einer gleichmäßig konvergenten Reihe stetiger Funktionen. (Satz auf Seite 138 im Skript)

Um festzustellen, ob die Grenzfunktion differenzierbar ist, müssen wir überlegen, ob die Folge der Ableitungen gleichmäßig konvergiert (um dann den Satz auf Seite 143 anzuwenden). Es gilt

$$\left( \sum_{n=0}^N \frac{n^2}{\sqrt{n!}} \left( x^n + \frac{1}{x^n} \right) \right)' = \sum_{n=0}^N \frac{n^2}{\sqrt{n!}} (x^n + \frac{1}{x^n})' = \sum_{n=1}^N \frac{n^3}{\sqrt{n!}} (x^{n-1} - x^{-n-1}),$$

und die Frage ist also, ob diese Reihe konvergiert. Um dies zu beweisen, kommt wieder der Weierstraß M-Test zum Zuge. Wir schätzen wieder ab:

$$\begin{aligned} \left| \frac{n^3}{\sqrt{n!}} (x^{n-1} - x^{-n-1}) \right| &\leq \frac{n^3}{\sqrt{n!}} (|x|^{n-1} + |x|^{-n-1}) \\ &\leq \frac{n^3}{\sqrt{n!}} (2^{n-1} + 2^{n+1}) \\ &\leq \frac{n^3}{\sqrt{n!}} (2^{n+1} + 2^{n+1}) \\ &\leq \frac{n^3}{\sqrt{n!}} 2^{n+2} =: c'_n \end{aligned}$$

Die Reihe  $\sum_{n=1}^{\infty} c'_n$  konvergiert wie oben mit dem Quotientenkriterium. Also konvergiert die Folge der Ableitungen nach dem M-Test gleichmäßig und zwar gegen  $\sum_{n=1}^{\infty} \frac{n^3}{\sqrt{n!}} (x^{n-1} - x^{-n-1})$ . Nach dem Satz auf Seite 143 ist also die durch die in der Aufgabe gegebene Reihe definierte Funktion differenzierbar und hat die Ableitung  $\sum_{n=1}^{\infty} \frac{n^3}{\sqrt{n!}} (x^{n-1} - x^{-n-1})$ .

## Differentialrechnung

### Ableitungen

#### Definition 11.1.1: Differenzierbarkeit, Ableitung

Es seien  $D \subseteq \mathbb{R}$  ein offenes Intervall und  $x_0 \in D$ .  $f : D \rightarrow \mathbb{R}$  heißt **differenzierbar in  $x_0$**  : $\Leftrightarrow$

$$\exists c \in \mathbb{R} \left( c = \lim_{x \rightarrow x_0} \frac{f(x) - f(x_0)}{x - x_0} \right).$$

In diesem Fall schreiben wir

$$\frac{d}{dx} f(x_0) = f'(x_0) := \lim_{x \rightarrow x_0} \frac{f(x) - f(x_0)}{x - x_0}$$

und nennen  $f'(x_0)$  die **Ableitung von  $f$  in  $x_0$** .  $f$  heißt **differenzierbar auf  $D$** , wenn  $f$  in jedem Punkt  $x \in D$  differenzierbar ist. In diesem Fall heißt die Funktion  $f' : D \rightarrow \mathbb{R}$ ,  $x \mapsto f'(x)$ , die **Ableitung von  $f$** . Ist  $f' : D \rightarrow \mathbb{R}$  stetig, so heißt  $f$  auf  $D$  **stetig differenzierbar** und wir schreiben  $f \in C^1(D)$ .

Existieren für  $x_0 \in \overline{D}$  die Grenzwerte

$$\begin{aligned} f'(x_0^+) &:= \lim_{x \rightarrow x_0^+} \frac{f(x) - f(x_0)}{x - x_0} \text{ beziehungsweise} \\ f'(x_0^-) &:= \lim_{x \rightarrow x_0^-} \frac{f(x) - f(x_0)}{x - x_0}, \end{aligned}$$

so heißen sie **rechtsseitige** beziehungsweise **linksseitige Ableitungen** von  $f$  in  $x_0$  (vergleiche Definition 10.1.3).

Hier wird die Ableitung und die Differenzierbarkeit über die Existenz eines Grenzwerts definiert. Die übliche geometrische Anschauung, die Ableitung an einer Stelle  $x_0$  als Steigung des Graphen an dieser Stelle zu betrachten, wird hier nicht benannt, ist aber stets hilfreich, sich etwas unter den Differenzenquotienten  $\frac{f(x)-f(x_0)}{x-x_0}$  vorzustellen. Dieser Quotient ist nämlich die Sekantensteigung der Sekante, die den Graphen an der Stelle  $x$  und  $x_0$  schneidet. Beim Übergang  $x \rightarrow x_0$  wird sie zur Tangentensteigung, die identisch mit der Steigung des Graphen an der Stelle  $x_0$  ist.

Die eben verwendete Technik  $x = x_0 + h$  mit  $h \in \mathbb{R}$  zu schreiben, kann nützlich sein. Der Grenzwert des Differenzenquotienten wird wie oben dann zu

$$\lim_{x \rightarrow x_0} \frac{f(x) - f(x_0)}{x - x_0} = \lim_{h \rightarrow 0} \frac{f(x_0 + h) - f(x_0)}{h}.$$

Für viele Berechnungen der Ableitung über diese Definition ist die Schreibweise mit dem  $h$  einfacher. Ich mag diese  $h$ -Methode sehr.

### Darstellungssatz 11.1.3

Es seien  $I \subseteq \mathbb{R}$  ein Intervall,  $f : I \rightarrow \mathbb{R}$  und  $x_0 \in I$ , dann ist  $f$  genau dann in  $x_0$  differenzierbar, wenn ein  $c \in \mathbb{R}$  und eine in  $x_0$  stetige Funktion  $\varphi = \varphi_{x_0} : I \rightarrow \mathbb{R}$  mit  $\varphi(x_0) = 0$  existieren, so dass für alle  $x \in I$  die

#### Darstellung

$$f(x) = f(x_0) + c \cdot (x - x_0) + \varphi_{x_0}(x) \cdot (x - x_0)$$

gilt. In diesem Fall ist  $c = f'(x_0)$ .

Ein interessanter Satz, den ich auch kaum verwende.

Wenn ich das richtig lese, dann kann man hiermit nicht nur etwas über die Differenzierbarkeit aussagen, sondern eine Funktion durch eine andere ausdrücken – egal, wie kompliziert  $f$  sein mag,

sondern  $f$  differenzierbar ist. Es gilt dann  $f(x) = f(x_0) + f'(x_0)(x - x_0) + \varphi_{x_0}(x)(x - x_0)$ . Natürlich ist hier nicht gesagt, was diese  $\varphi_{x_0}$ -Funktion genau ist. Riecht schon irgendwie nach Taylor für mich. Wenn ich also eine komplizierte Funktion habe, dann kann man sie offenbar durch ein paar Werte (blau) und einer anderen Funktion (rot) ausdrücken. Es treten dann nur noch ein paar lineare Faktoren auf. Hmm. Schon interessant. Lass noch mal kurz umstellen:

$$f(x) = f(x_0) + f'(x_0)(x - x_0) + \varphi_{x_0}(x)(x - x_0)$$

$$f(x) - f(x_0) = f'(x_0)(x - x_0) + \varphi_{x_0}(x)(x - x_0)$$

$$\frac{f(x) - f(x_0)}{(x - x_0)} - f'(x_0) = \varphi_{x_0}(x)$$

Tja, keine Ahnung für was man das nutzen kann.

### Satz 11.1.4

Es sei  $f : I \rightarrow \mathbb{R}$  in  $x_0 \in I$  differenzierbar, dann ist  $f$  in  $x_0$  stetig.

Aus Differenzierbarkeit folgt Stetigkeit. Irgendwie logisch.

### Satz 11.1.6: Differenzierbarkeit von Potenzreihen

Es sei  $f(x) = \sum_{k=0}^{\infty} a_k(x - x_0)^k$  eine Potenzreihe mit Konvergenzradius  $R > 0$ , dann ist  $f$  differenzierbar auf  $(x_0 - R, x_0 + R)$  und falls  $R = +\infty$  auf ganz  $\mathbb{R}$  differenzierbar und es gilt

$$\forall x \in \mathbb{R} \left( |x - x_0| < R \Rightarrow f'(x) = \sum_{k=1}^{\infty} k a_k (x - x_0)^{k-1} \right).$$

Man darf quasi das Ableiten auf alle Glieder der Summe übertragen oder anders gesagt, wenn man das Ableiten als Anwenden des Ableitungsoperator  $\frac{d}{dx}$  betrachtet, den Operator in

die Summe reinziehen.

### Ableitungsregeln 11.1.8

Es seien  $f, g : D \rightarrow \mathbb{R}$  in  $x_0 \in D$  differenzierbare Funktionen, dann sind auch  $f + g$ ,  $f \cdot g$  und falls  $g(x_0) \neq 0$  auch  $\frac{f}{g}$  in  $x_0$  differenzierbar und es gilt

(i)  $(f + g)'(x_0) = f'(x_0) + g'(x_0)$  (Linearität der Ableitung),

(ii)  $(f \cdot g)'(x_0) = f'(x_0)g(x_0) + f(x_0)g'(x_0)$  (Produktregel),

(iii)  $\left(\frac{f}{g}\right)'(x_0) = \frac{f'(x_0)g(x_0) - f(x_0)g'(x_0)}{g^2(x_0)}$  (Quotientenregel).

Naja, diese Regeln sind (überwiegend) aus der Schule bekannt und ersparen einen das mühsame Ableiten über die Definition. In Kombi mit bekannten Ableitungen und Ableitungsregeln, wie die Potenzregel usw. ist das nicht schwer.

### Satz 11.1.10: Kettenregel

Es seien  $I, J \subseteq \mathbb{R}$  und  $f : I \rightarrow J, g : J \rightarrow \mathbb{R}$  Funktionen. Wenn  $f$  in  $x_0 \in I$  und  $g$  in  $f(x_0) \in J$  differenzierbar ist, dann ist  $g \circ f : I \rightarrow \mathbb{R}$  in  $x_0$  differenzierbar und es gilt

$$(g \circ f)'(x_0) = g'(f(x_0)) \cdot f'(x_0).$$

Siehe oben – sollte bekannt sein.

### Satz 11.1.12: Ableitung der Umkehrfunktion

Es seien  $I, J \subseteq \mathbb{R}$  Intervalle und  $f : I \rightarrow J$  bijektiv. Wenn  $f$  in  $x_0 \in I$  differenzierbar mit  $f'(x_0) \neq 0$  ist, dann ist auch  $f^{-1}$  in  $y_0 = f(x_0)$  differenzierbar und es gilt

$$\frac{d}{dy} f^{-1}(y_0) = \frac{1}{f'(f^{-1}(y_0))}.$$

Umkehrfunktionen werde erst jetzt wieder im Schulstoff stärker behandelt. Diesen Zusammenhang sollte man mal zur Kenntnisnehmen.

### Korollar 11.1.14

Für alle  $x > 0$  und alle  $c \in \mathbb{R}$  gilt

$$\frac{d}{dx} x^c = \frac{d}{dx} e^{c \ln x} = e^{c \ln x} \frac{c}{x} = c \cdot x^{c-1}.$$

Manchmal sind „Umwege“ über Umkehrfunktionen oder andere Funktionen nützlich, z. B. hier ein Beweis der Potenzregel 😊

## Die Mittelwertsätze der Differentialrechnung

### Definition 11.2.1: Lokale Extrema

Es seien  $D \subseteq \mathbb{R}, f : D \rightarrow \mathbb{R}$  und  $x_0 \in D$ . Die Funktion  $f$  besitzt an der Stelle  $x_0$  ein *lokales Maximum (lokales Minimum)* : $\Leftrightarrow$

$$\exists \delta > 0 \forall x \in U_\delta(x_0) \cap D (f(x) \leq f(x_0))$$

beziehungsweise

$$\exists \delta > 0 \forall x \in U_\delta(x_0) \cap D (f(x) \geq f(x_0)).$$

- Wir sagen kurz, dass  $f$  in  $x_0$  ein *lokales Extremum* besitzt, wenn  $f$  in  $x_0$  ein lokales Maximum oder ein lokales Minimum besitzt.
- Ein lokales Extremum heißt *isoliert*, falls die strikte Ungleichung in der punktierten Umgebung gilt.

Kommen wir zu Elementen der Funktionsuntersuchung. Hier zunächst die Definition von *lokalen Maximum* und *lokalen Minimum* – nicht zu verwechseln mit dem absoluten Maximum (größter Funktionswert) und absoluten Minimum (kleinstes Funktionswert).

### Satz von Fermat<sup>a</sup> 11.2.3

<sup>a</sup>Pierre de Fermat, 1601 - 1665, frz. Mathematiker

Es seien  $I \subseteq \mathbb{R}$  ein offenes Intervall und  $f : I \rightarrow \mathbb{R}$  differenzierbar in  $x_0 \in I$ . Wenn  $f$  in  $x_0$  ein lokales Extremum besitzt, dann gilt  $f'(x_0) = 0$ .

Das ist der Satz zur notwendigen Bedingung für Extremstellen. In jeder Extremstelle (Hoch- / Tiefpunkt) muss die erste Ableitung Null sein.

### Satz von Rolle<sup>a</sup> 11.2.5

<sup>a</sup>Michel Rolle, 1652-1719, frz. Mathematiker

Es seien  $a, b \in \mathbb{R}, a < b$ . Die Funktion  $f : [a, b] \rightarrow \mathbb{R}$  sei stetig und auf  $(a, b)$  differenzierbar. Weiter gelte  $f(a) = f(b)$ . Dann existiert ein  $\xi \in (a, b)$  mit  $f'(\xi) = 0$ .

Wenn an zwei Stellen die Funktion den gleichen Funktionswert annimmt, dann muss mind. eine Stelle in dem Intervall

existieren, an der die Ableitung Null ist. Das geht bei stetigen Funktionen gar nicht anders. Klar, oder?

### Zweiter Mittelwertsatz 11.2.6

Es seien  $a, b \in \mathbb{R}$ ,  $a < b$ . Die Funktionen  $f, g : [a, b] \rightarrow \mathbb{R}$  seien stetig und auf  $(a, b)$  differenzierbar. Dann existiert ein  $\xi \in (a, b)$  mit

$$f'(\xi)(g(b) - g(a)) = g'(\xi)(f(b) - f(a)).$$

wird auf eine Hilfsfunktion angewendet. Die Herleitung ist aber nicht so wichtig. Interessant ist noch die alternative Schreibweise:

$$\frac{f'(\xi)}{g'(\xi)} = \frac{f(b) - f(a)}{g(b) - g(a)}, \text{ wobei } g'(\xi) \neq 0 \text{ und } g(b) \neq g(a) \text{ gilt.}$$

### Korollar 11.2.8: Erster Mittelwertsatz

Es seien  $a, b \in \mathbb{R}$ ,  $a < b$  und  $f : [a, b] \rightarrow \mathbb{R}$  stetig und differenzierbar auf  $(a, b)$ , dann gilt

$$\exists \xi \in (a, b) \left( f'(\xi) = \frac{f(b) - f(a)}{b - a} \right).$$

Der Mittelwertsatz ist ein Sonderfall vom zweiten mit  $g(x) = x$  und nach  $f'$  aufgelöst.

### Identitätssatz für differenzierbare Funktionen 11.2.9

Es seien  $I \subseteq \mathbb{R}$  ein Intervall und  $f, g : I \rightarrow \mathbb{R}$  auf  $I$  differenzierbare Funktionen, dann gilt

(i)  $\forall x \in I (f'(x) = 0) \Rightarrow \exists c \in \mathbb{R} \forall x \in I (f(x) = c)$ , das heißt wenn die

Ableitung von  $f$  für alle  $x \in I$  verschwindet, dann ist  $f$  konstant.

(ii)  $\forall x \in I (f'(x) = g'(x)) \wedge \exists x_0 \in I (f(x_0) = g(x_0)) \Rightarrow \forall x \in I (f(x) = g(x))$ .

Tja. Sagt mir nichts, aber dass die Ableitung einer konstanten Funktion verschwindet (Steigung = 0), ist ja auch so klar und die Umkehrung intuitiv auch.

### Lemma 11.2.11: Monotoniekriterien für differenzierbare Funktionen

Es seien  $I \subseteq \mathbb{R}$  ein Intervall und  $f : I \rightarrow \mathbb{R}$  eine differenzierbare Funktion. Dann gilt

(i)  $\forall x \in I (f'(x) \geq 0 \Rightarrow f \nearrow)$  auf  $I$ .

(ii)  $\forall x \in I (f'(x) \leq 0 \Rightarrow f \searrow)$  auf  $I$ .

(iii) Analoge Aussagen gelten mit den strikten Ungleichungen ( $>$ ,  $<$ ) und strenger Monotonie ( $\uparrow$ ,  $\downarrow$ ).

Ein wichtiger Zusammenhang zwischen Monotonie und dem Vorzeichen der Ableitung!



## De L'Hospitalsche Regel 11.2.12

<sup>a</sup>Guillaume François Antoine, Marquis de L'Hospital, 1661-1704, frz. Mathematiker

Die Funktionen  $f$  und  $g$  seien auf dem offenen Intervall  $(a, b)$ ,  $-\infty \leq a < b \leq +\infty$ , differenzierbar und weiter sei  $g'(x) \neq 0$  für  $x \in (a, b)$ . Es sei

$$\lim_{x \rightarrow a} f(x) = \lim_{x \rightarrow a} g(x) = 0 \text{ (der Fall } \frac{0}{0}\text{)}$$

oder

$$\lim_{x \rightarrow a} g(x) = \pm\infty.$$

Außerdem existiere der Grenzwert  $\lim_{x \rightarrow a} \frac{f'(x)}{g'(x)}$ , wobei bestimmte Divergenz zugelassen ist. Dann gibt es ein  $b' \in \mathbb{R}$ ,  $a < b' \leq b$  mit  $g(x) \neq 0$  für  $x \in (a, b')$ ; im Fall  $\frac{0}{0}$  gilt dies für alle  $x \in (a, b)$ . Unter diesen Voraussetzungen existiert der Grenzwert  $\lim_{x \rightarrow a} \frac{f(x)}{g(x)}$  und es gilt die de L'Hospitalsche Regel

$$\lim_{x \rightarrow a} \frac{f(x)}{g(x)} = \lim_{x \rightarrow a} \frac{f'(x)}{g'(x)}.$$

Diese Regel darf in keiner Mathematikvorlesung fehlen 😊

Beispiele hier:

<https://de.serlo.org/math/1689/regel-von-lhospital>

Und mit noch mehr Fällen und konkreten Beispielen hier:

<https://studyflix.de/mathematik/lhospital-1869>

## Der Satz von Taylor

### Definition 11.3.1: Ableitungen höherer Ordnung

- (i) Es seien  $f : I \rightarrow \mathbb{R}$  differenzierbar auf  $I$ ,  $x_0 \in I$  und  $(f')'(x_0)$  existiere. Dann heißt

$$f''(x_0) = \frac{d^2 f}{dx^2}(x_0) := (f')'(x_0) = \left( \frac{d}{dx} \left( \frac{df}{dx} \right) \right)(x_0)$$

die zweite Ableitung von  $f$  an der Stelle  $x_0$ .

- (ii) Falls  $f''(x)$  für alle  $x \in I$  existiert, dann heißt  $f$  zweimal differenzierbar auf  $I$ . Die zweite Ableitung  $f'' : I \rightarrow \mathbb{R}$  ist dann auf ganz  $I$  erklärt.

Bevor es an der Satz von Taylor geht, ein paar Worte und Schreibweise zu höheren Ableitungen. Hier wird der Ableitungsoperator  $\frac{d}{dx}$  verwendet.

Ich kann keine Stelle im Skript finden, wo er eingeführt wird! Im Beispiel 11.1.7 scheint er einfach aufzutauchen. Er eignet

sich gut, um das Ableiten auszudrücken und umgeht die unmathematische Schreibweise wie z. B.  $(x^2)' = 2x$ , die zwar jeder Mathematiker als Ableiten verstehen würde, aber das Strichen wird allgemein nicht als Operator eingeführt. Anders der Operator  $\frac{d}{dx}$ , der ein *unitärer Operator* ist, also nicht wie z. B.  $+$ ,  $-$ ,  $\cdot$  oder  $:$  zwei Elemente verbindet, sondern einseitig wirkt (und zwar nach rechts). Entsprechend wäre die formal bessere Schreibweise von dem Beispiel oben:  $\frac{d}{dx} x^2 = 2x$ . Die Ableitung der Ableitung ist die zweite Ableitung und wird so geschrieben:  $\frac{d}{dx} \left( \frac{d}{dx} \right) = \frac{d^2}{dx^2}$ . Es ist also  $\frac{d^2}{dx^2} x^2 = \frac{d}{dx} 2x = x \cdot \frac{d}{dx}$  ist mehr als eine Schreibweise für „leite ab“, sondern ein Baustein für die formale Formulierung eines umfassenden Kalküls. Physiker können dir ein Lied davon singen.

- (iii) Ist zusätzlich  $f''$  stetig auf  $I$ , so heißt  $f$  zweimal stetig differenzierbar auf  $I$  und wir schreiben kurz  $f \in C^2(I)$ .
- (iv) Für  $n \in \mathbb{N}$ ,  $n \geq 2$ , definieren wir rekursiv die  $n$ -te Ableitung  $f^{(n)}(x_0) = \frac{d^n f}{dx^n}(x_0)$  und die Klasse  $C^n(I)$  der auf  $I$   $n$ -mal stetig differenzierbaren Funktionen.
- (v) Existiert  $f^{(n)}(x_0)$  für alle  $n \in \mathbb{N}$ , so heißt  $f$  unendlich oft differenzierbar im Punkt  $x_0$ . Ist  $f^{(n)}(x)$  für alle  $n \in \mathbb{N}$  und alle  $x \in I$  erklärt, so heißt  $f$  unendlich oft differenzierbar auf  $I$ ,  $f \in C^\infty(I)$ . In diesem Fall ist  $f$  automatisch auch unendlich oft stetig differenzierbar.

### Satz 11.3.2

Es sei  $f(x) = \sum_{k=0}^{\infty} a_k(x - x_0)^k$  eine Potenzreihe mit Konvergenzradius  $R > 0$ . Es gilt  $f$  ist innerhalb des Konvergenzradius unendlich oft differenzierbar, die  $n$ -te Ableitung berechnet sich für  $x \in \mathbb{R}$ ,  $|x - x_0| < R$  zu

$$f^{(n)}(x) = \sum_{k=n}^{\infty} a_k k(k-1)\dots(k-n+1)(x - x_0)^{k-n}$$

und es gilt  $f^{(n)}(x_0) = a_n n!$ , also  $a_n = \frac{f^{(n)}(x_0)}{n!}$  und daher

$$f(x) = \sum_{k=0}^{\infty} \frac{f^{(k)}(x_0)}{k!} (x - x_0)^k \text{ für alle } x \in \mathbb{R}, |x - x_0| < R.$$

Jetzt wird es langsam interessant. Das wiederholte Ableiten führt dazu, dass nach der Potenzregel, die Exponenten ein Produkt bilden mit einer absteigenden Folge von Faktoren, die mittels dem Fakultätsymbol abgekürzt werden. Für den Fall  $x = x_0$  fallen alle Glieder mit  $(x - x_0)^j$  weg, sofern

$j > 0$  ist. Für  $n = k$  bleibt ein Glied übrig, nämlich  $a_n n!$ . Daher kann die  $n$ -te Ableitung an der Stelle  $x_0$  durch  $f^{(n)}(x_0) = a_n n!$  ausgedrückt werden und diese Gleichung wiederrum nach  $a_n$  umgestellt werden, wodurch man eine alternative Schreibweise für die Funktion mit der Potenzreihe findet (letzte Zeile).

### Identitätssatz für Potenzreihen 11.3.3

Es seien  $f(x) = \sum_{k=0}^{\infty} a_k(x - x_0)^k$  und  $g(x) = \sum_{k=0}^{\infty} b_k(x - x_0)^k$  zwei Potenzreihen und ein  $R > 0$ , so dass  $f$  und  $g$  für  $x \in \mathbb{R}$ ,  $|x - x_0| < R$ , konvergieren. Ist  $(x_k)_{k \in \mathbb{N}} \subseteq \{x \in \mathbb{R} \mid |x - x_0| < R\} = I$  eine Folge mit  $x_k \rightarrow x_0$ ,  $x_k \neq x_0$  und gilt  $f(x_k) = g(x_k)$  für alle  $k \in \mathbb{N}$ , so gilt

$$\forall x \in I (f(x) = g(x)) \wedge \forall k \in \mathbb{N}_0 (a_k = b_k).$$

K. A. – braucht man vielleicht später.

### Definition 11.3.4: Taylorpolynom, Taylorreihe

Es sei  $f : I \rightarrow \mathbb{R}$  an der Stelle  $x_0 \in I$   $n$ -mal differenzierbar. Dann ist

$$T^{(n)} f(x_0, x) := \sum_{k=0}^n \frac{f^{(k)}(x_0)}{k!} (x - x_0)^k$$

das  $n$ -te Taylorpolynom von  $f$  an der Stelle  $x$  mit dem Entwicklungspunkt  $x_0$ , wobei wir  $f^{(0)} = f$  setzen. Ist  $f$  an der Stelle  $x_0$  unendlich oft differenzierbar, so heißt die formale Reihe

$$Tf(x_0, x) := \sum_{k=0}^{\infty} \frac{f^{(k)}(x_0)}{k!} (x - x_0)^k$$

die Taylorreihe von  $f$  an der Stelle  $x$  mit dem Entwicklungspunkt  $x_0$ .

Taylorpolynome sind recht nützlich und leicht zu handhaben – besonders wenn es um Näherung geht. Oft wird anstatt der Schreibweise im Kasten auch  $T_{nf}(x, a)$  verwendet ( $n$  als Index, andere Reihenfolge der Parameter).  $T_1 f(x, a)$

liefert das erste Taylorpolynom an der Stelle  $a$ , wobei das Taylorpolynom ein Polynom in  $x$  ist.

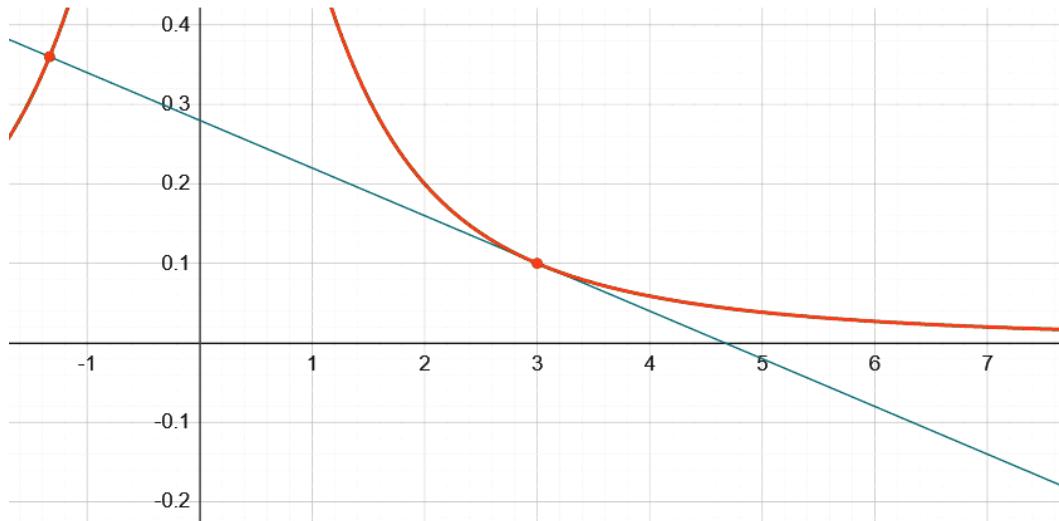
Es ist also  $T_1 f(x, a) = f(a) + f'(a)(x - a)$ .

Dieses Polynom ist eine lineare Funktion, die eine Tangente an den Graphen von  $f$  an der Stelle  $a$  beschreibt.

**Beispiel:**

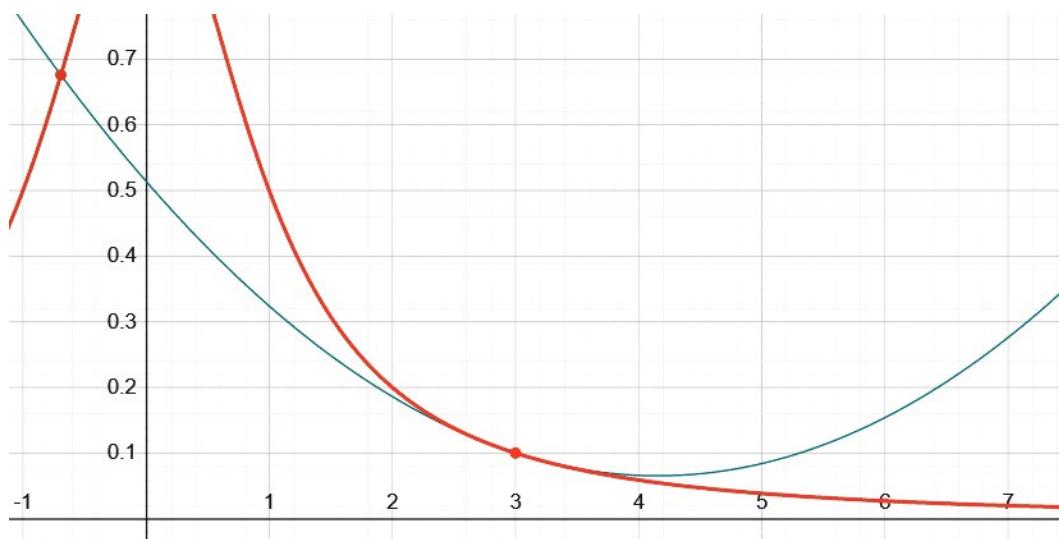
$$f(x) = \frac{1}{x^2 + 1}$$

$$T_1 f(x, 3) = f(3) + f'(3)(x - 3) = \frac{1}{10} - \frac{3(x - 3)}{50}$$



Würde man einen Grad höher gehen, erhält man eine Schmiegeparabel, die lokal an der Stelle 3 noch näher am Graphen von  $f$  verläuft.

$$T_2 f(x, 3) = \frac{1}{10} - \frac{3(x - 3)}{50} + \frac{13(x - 3)^2}{500}$$



Je höher der Grad des Taylorpolynoms, desto besser liegt der Graph des Polynoms am Graphen der Funktion  $f$ .

Es gäbe hier noch etwas mehr zu besprechen und zu untersuchen, z. B. der Fehler (die Abweichung) zwischen dem Taylorpolynom und der Ausgangsfunktion, das sogenannte Restglied; was wenn der Grad  $n \rightarrow \infty$ . Wir gehen aber im Skript weiter:

### Satz von Taylor 11.3.5

Es seien  $a, b \in \mathbb{R}$ ,  $a < b$ ,  $I = [a, b]$ ,  $x_0 \in I$  und  $f : I \rightarrow \mathbb{R}$  eine Funktion. Weiter sei  $f$  auf  $(a, b)$   $n$ -mal differenzierbar und  $(n-1)$ -mal stetig differenzierbar in  $I$ . Dann gilt die *Taylorsche Formel*

$$\begin{aligned} f(x) &= \sum_{k=0}^{n-1} \frac{f^{(k)}(x_0)}{k!} (x - x_0)^k + \frac{f^{(n)}(\xi)}{n!} (x - x_0)^n \\ &= T^{(n-1)} f(x_0, x) + R_n(x_0, x) \end{aligned}$$

für alle  $x \in I$  mit einem  $\xi = x_0 + t(x - x_0)$  für ein  $t \in (0, 1)$  und

$$R_n(x_0, x) := \frac{f^{(n)}(\xi)}{n!} (x - x_0)^n$$

Hier wird das Restglied nach Lagrange eingeführt.

ist das *Lagrangesche Restglied*<sup>a</sup>.

<sup>a</sup>Joseph-Louis de Lagrange, 1736-1813, ital. Mathematiker

### Lemma 11.3.6

Es seien  $a, b \in \mathbb{R}$ ,  $a < b$  und  $f : I = [a, b] \rightarrow \mathbb{R}$  zweimal stetig differenzierbar auf  $(a, b)$ . Dann gilt für  $x, x_0 \in I$  die Taylorsche Formel

$$\begin{aligned} f(x) &= f(x_0) + f'(x_0)(x - x_0) + \frac{f''(\xi)}{2} (x - x_0)^2 \\ &= f(x_0) + f'(x_0)(x - x_0) + \frac{f''(x_0)}{2} (x - x_0)^2 + R(x_0, x) \end{aligned}$$

mit einem  $\xi$  zwischen  $x_0$  und  $x$  und

$$R(x_0, x) = \frac{f''(\xi) - f''(x_0)}{2} (x - x_0)^2.$$

Hier nochmal der explizite Fall für Grad 2, bei dem die Funktion  $f$  durch eine Taylorpolynom vom Grad 2 und dem Restglied beschrieben wird.

Die meisten Anwendungen verwenden die „Linearisierung“, d. h. eine

nichtlineare Funktion wird lokal linearisiert betrachtet. Lineare Probleme lassen sich nämlich viel leichter und schneller lösen und sofern man nahe an der Entwicklungsstelle bleibt, ist die Abweichung minimal. Rechengeschwindigkeit bzw. leichte Berechenbarkeit ist für hochkomplexe Aufgaben sehr wichtig!

### Satz 11.3.8: Notwendiges zweite-Ableitungskriterium

Die Funktion  $f : I \rightarrow \mathbb{R}$  sei zweimal stetig differenzierbar auf  $I$  und besitze in einem inneren Punkt  $x_0 \in I$  ein Maximum beziehungsweise ein Minimum, dann gilt

$$f''(x_0) \leq 0 \text{ beziehungsweise } f''(x_0) \geq 0.$$

Okay, das hier kennst du aus der Schule!

### Satz 11.3.9: Hinreichendes zweite-Ableitungskriterium

Es sei  $f : I \rightarrow \mathbb{R}$  zweimal stetig differenzierbar auf  $I$ . In einem inneren Punkt  $x_0 \in I$  sei  $f'(x_0) = 0$  und  $f''(x_0) < 0$  beziehungsweise  $f''(x_0) > 0$ . Dann besitzt  $f$  an der Stelle  $x_0$  ein isoliertes relatives Maximum beziehungsweise Minimum.

## Partielle Ableitungen

### Definition 11.4.1: Partielle Ableitung, partielle Differenzierbarkeit

Es seien  $D \subseteq \mathbb{R}^n$ ,  $a \in D$ ,  $f : D \rightarrow \mathbb{R}$  wie oben und  $e_j$  der  $j$ -te kanonische Einheitsvektor.

- (i)  $f$  heißt in  $a$  partiell nach  $x_j$  differenzierbar ( $1 \leq j \leq n$ ), wenn der Grenzwert

$$\lim_{h \rightarrow 0} \frac{f(a + h \cdot e_j) - f(a)}{h} = \lim_{h \rightarrow 0} \frac{f(a_1, \dots, a_{j-1}, a_j + h, a_{j+1}, \dots, a_n) - f(a_1, \dots, a_n)}{h}$$

existiert. Dieser Grenzwert heißt partielle Ableitung von  $f$  nach  $x_j$  in  $a$  und wird mit  $\frac{\partial f}{\partial x_j}(a)$  bezeichnet.

- (ii) Existiert die partielle Ableitung von  $f$  nach  $x_j$  in allen  $a \in D$ , so heißt die Funktion

$$\frac{\partial f}{\partial x_j} : D \rightarrow \mathbb{R}, x \mapsto \frac{\partial f}{\partial x_j}(x)$$

die partielle Ableitung von  $f$  nach  $x_j$ .

- (iii) Existieren für  $x_1, \dots, x_n$  alle partiellen Ableitungen von  $f$  im Punkt  $a$ , dann heißt  $\nabla f(a) = \left( \frac{\partial f}{\partial x_1}(a), \dots, \frac{\partial f}{\partial x_n}(a) \right)^T$  der Gradient von  $f$ .

Partielle Ableitungen werden bei mehrdimensionalen Funktionen wichtig. In der Handhabung gar nicht so schwer, aber dahinter steckt noch viel mehr. Müsste man eigentlich noch vertiefen, z. B. wie berechnet man ein Minimum oder Maximum bei mehrdimensionalen Funktionen. Im Fall von zwei Dimensionen haben wir es dann graphisch mit einer 3D-Landschaft zu tun, bei denen die Funktionswerte der Höhe des Landschaftspunktes beschreibt. Und natürlich kann man sich da Täler (Senken) und Hügel spitzen vorstellen, die man berechnen kann. Für

noch höhere Dimensionen wird die Vorstellung wiederum schwieriger, aber es wird dort immer noch lokal größte und kleinste Werte geben. Im Skript wird noch einiges als Bemerkungen reingebuttert, aber nicht alles.

Hier nochmal, wie auf Blatt 10 beschrieben, der Umgang mit diesen Ableitungen:

**Partielle Ableitungen:** Alle Variablen werden als Konstanten betrachtet (die als Faktor erhalten bleiben oder als Glied wegfallen), außer die Variable nach der partiell abgeleitet wird, also z. B. bei  $f(x_1, x_2) = x_1^3 x_2 + x_2^2$ . Hier ist:

$$\frac{\partial}{\partial x_1} f(x_1, x_2) = 3x_1^2 x_2$$

$$\frac{\partial}{\partial x_2} f(x_1, x_2) = x_1^3 + 2x_2$$

Das ist schon alles okay, aber man darf nicht den Überblick verlieren.

# Integralrechnung

## Stammfunktionen

### Definition 12.1.1: Stammfunktion, unbestimmtes Integral

Es seien  $I \subseteq \mathbb{R}$  ein Intervall und  $f, F : I \rightarrow \mathbb{R}$ . Wenn  $F$  auf  $I$  die Ableitung  $f$  besitzt, das heißt wenn  $\forall x \in I (F'(x) = f(x))$  gilt, so heißt  $F$  eine **Stammfunktion** oder ein **unbestimmtes Integral** von  $f$ , in Zeichen

$$F(x) = \int f(x) dx \text{ für } x \in I.$$

Ich will jetzt hier nicht zu viele Worte verlieren. Vieles von dem hier ist aus der Schule bekannt, wird nur etwas genau beschrieben. Die Begriffe und Symbole sollten klar sein – auch später das Riemann-Integral.

### Satz 12.1.2

Jede auf einem kompakten Intervall  $I$  stetige Funktion  $f : I \rightarrow \mathbb{R}$  besitzt eine Stammfunktion.

– anschaulich – einen berechenbaren (orientierten) Flächeninhalt unter dem Graphen.

### Lemma 12.1.4

Es seien  $I \subseteq \mathbb{R}$  ein Intervall und  $F_1, F_2 : I \rightarrow \mathbb{R}$  zwei Stammfunktionen einer Funktion  $f : I \rightarrow \mathbb{R}$ , dann gilt

$$\exists c \in \mathbb{R} \forall x \in I (F_1(x) - F_2(x) = c).$$

### Stammfunktionen elementarer Funktionen 12.1.5

Es gilt, sofern nicht anders angegeben für alle  $x \in \mathbb{R}$ ,

(i)  $\int (x-a)^n dx = \frac{(x-a)^{n+1}}{n+1}$  für  $x \neq a$ ,  $n \in \mathbb{Z}$ ,  $n \neq -1$  oder  
 $x \in \mathbb{R}$ ,  $n \in \mathbb{N}_0$  und alle  $a \in \mathbb{R}$ .

(ii)  $\int \frac{dx}{x-a} = \ln|x-a|$  für  $x \neq a$ .

(iii)  $\int x^\mu dx = \frac{x^{\mu+1}}{\mu+1}$  für  $x > 0$ ,  $\mu \in \mathbb{R}$ ,  $\mu \neq -1$ .

(iv)  $\int e^{ax} dx = \frac{e^{ax}}{a}$  für  $a \in \mathbb{R}$ ,  $a \neq 0$ .

(v)  $\int a^x dx = \int e^{\ln a x} dx = \frac{a^x}{\ln a}$  für  $a > 0$ ,  $a \neq 1$ .

(vi)  $\int \sin x dx = -\cos x$ ,  $\int \cos x dx = \sin x$ .

(vii)  $\int \frac{dx}{\cos^2 x} = \tan x$  für  $x \neq \left(k + \frac{1}{2}\right)\pi$ ,  $k \in \mathbb{Z}$ ,

$$\int \frac{dx}{\sin^2 x} = -\cot x \text{ für } x \neq k\pi, k \in \mathbb{Z}.$$

(viii)  $\int \cot x dx = \ln|\sin x|$  für  $\sin x > 0$ ,

$$\int \tan x dx = -\ln|\cos x| \text{ für } \cos x > 0.$$

(ix)  $\int \frac{dx}{\sqrt{1-x^2}} = \arcsin x$  für  $|x| < 1$ .

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

Mit anderen Worten:  
Jede stetige Funktion ist in dem entsprechenden Intervall integrierbar, hat

Zwei Stammfunktionen unterscheiden sich in einem Intervall nur um einen konstanten Wert – an jeder Stelle im Intervall.  
Aus der Schule solltest du

wissen: Es gibt nicht DIE Stammfunktion, sondern wenn es eine gibt, dann gibt es unendlich viele, weil man die Funktion stets um  $+C$  (Konstante) verschieben kann, aber beim Ableiten oder beim Berechnen eines bestimmten Integrals diese Konstante wegfällt bzw. sich gegenseitig aufheben. In der Schule wird das durch den Hauptsatz der Integralrechnung motiviert. Siehe unten.

Es gibt einen ganzen Katalog an Funktionen und ihren zugehörigen Stammfunktionen.

### Satz 12.1.6: Stammfunktionen von Potenzreihen

Wenn die Potenzreihe  $f(x) = \sum_{k=0}^{\infty} a_k(x - x_0)^k$  für  $x \in \mathbb{R}$  mit  $|x - x_0| < R$  konvergiert, dann gilt dort

$$\int f(x) dx = \sum_{k=0}^{\infty} \frac{a_k}{k+1} (x - x_0)^{k+1}.$$

Durch das Anwenden der „Potenzregel rückwärts“ wie in der Schule, kann man eine Stammfunktion einer Potenzreihen angeben, in dem man Gliedweise „aufleitet“.

### Satz 12.1.7: Linearität der Stammfunktion

Es seien  $f, g : I \rightarrow \mathbb{R}$  Funktionen und  $c \in \mathbb{R}$ . Existieren  $\int f(x) dx$  und  $\int g(x) dx$  dann existieren auch  $\int (f + g)(x) dx$  und  $\int (c \cdot f)(x) dx$  und es gilt

$$\int (f + g)(x) dx = \int f(x) dx + \int g(x) dx$$

und

$$\int (c \cdot f)(x) dx = c \cdot \int f(x) dx.$$

Das geht natürlich nur, weil die Linearität der Stammfunktion gilt, d. h. man eine Summe zerlegen kann und die Summanden integrieren kann und weil man einen Faktor „rausholen“ darf.



Wir kommen jetzt zu **höheren Integrationstechniken**, die im Skript für meinen Geschmack schlecht teilweise beschrieben werden. Diese Techniken erfordern ein höheres Maß an Konzentration und es ist auf Übersichtlichkeit zu achten. Sonst geht es schnell schief!

### Satz 12.1.8: Partielle Integration

Es seien  $I \subseteq \mathbb{R}$  ein Intervall,  $f, g : I \rightarrow \mathbb{R}$  differenzierbar in  $I$  und die Funktion  $f \cdot g'$  besitze eine Stammfunktion in  $I$ . Dann hat auch  $f' \cdot g$

Schaut man sich die Produktregel für Ableitungen genauer an, ergibt sich daraus ein Weg, um Stammfunktionen zu bestimmen. Siehe dazu meine Ausführungen zu Blatt 10 und beachte das



eine Stammfunktion und es gilt

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

Skript auf den Seiten 315-316. Es gibt hier trickreiche Schritte. Neben dem vereinfachen des Integranden wie in Beispiel 12.1.9 (i)

$$\int \frac{x}{=g(x)=f'(x)} \frac{e^x}{f(x)g(x)} dx = \frac{xe^x}{f(x)g(x)} - \int \frac{1}{=g'(x)} \cdot \frac{e^x}{=f(x)} dx = xe^x - e^x.$$

leichter

gibt es die „intelligente Eins“ („fruchtbare Eins“, ähnlich wie „fruchtbare Null“):

$$\int \ln x dx = \int \frac{1}{=f'(x)} \cdot \frac{\ln x}{=g(x)} dx = \frac{x \ln x}{f(x)g(x)} - \int \frac{x}{=f(x)} \frac{1}{=g'(x)} dx = x \ln x - \int 1 dx = x \ln x - x$$

Oder wegen, die auf ein ähnliches Integral wie zu beginn führen, was man ausnutzen kann:

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

Durch Addition auf die linke Seite bringen und dann durch 2 teilen

Damit folgt

$$\int \cos^2 x \, dx = \frac{1}{2} (\cos x \sin x + x)$$



Die Substitutionsregel beschreibt andere schöner und überhaupt lernt man die am besten, wenn man sie 10 mal durchgeführt hat. Siehe unten.

#### Satz 12.1.10: Substitutionsregel

Es seien  $I, J \subseteq \mathbb{R}$  zwei Intervalle,  $f : I \rightarrow J$  differenzierbar und  $g : J \rightarrow \mathbb{R}$  besitze eine Stammfunktion im Intervall  $J$ , dann besitzt die Funktion  $g \circ f \cdot f' : I \rightarrow \mathbb{R}$  eine Stammfunktion und für  $t \in I$  gilt die **Substitutionsregel**

$$\int g(x) \, dx \Big|_{x=f(t)} = \int g(f(t)) f'(t) \, dt.$$

Diese Regel steht im Zusammenhang mit der Kettenregel. Auf eine genauere Herleitung verzichte ich und konzentriere mich auf die Anwendung.

#### Korollar 12.1.11

Ist  $f$  auf  $I$  umkehrbar, so erhält man durch Einsetzen die häufig nützlichere Form der **Substitutionsregel**:

$$\int g(x) \, dx = \int g(f(t)) \cdot f'(t) \, dt \Big|_{t=f^{-1}(x)}.$$

Ich bin mir auch nicht sicher, ob es im Skript steht, aber bei einem bestimmten Integral, also einem Integral mit Grenzen, kann man sich die Rücksubstitution

sparen, wenn man die Grenzen mit transformiert. Evtl. gebe ich dafür ein Beispiel weiter unten an.

Bei der **Substitution** versucht man die Variable  $x$  oder einen ganzen Ausdruck wie  $\ln(x)$  durch einen anderen Ausdruck zu ersetzen, so dass der Integrand einfacher zu integrieren wird. Aber Vorsicht! Beim Substituieren ändert sich mit der Variablen  $x$  auch das Differential, also das  $dx$ -Symbol. Die ganze Regel röhrt vereinfacht gesagt von der Umkehrung der Kettenregel her. Hier das praktische Vorgehen:

Allgemein kann die Substitutionsregel und das Vorgehen so formuliert werden:

1. Überlegen welcher Term substituiert werden soll. Bei Blatt 10, Aufgabe 5, Teil b) stört uns die  $\ln$ -Funktion und wir wählen  $t = \ln(x)$  (Substitution)
2. Diese Gleichung wird nach  $x$  aufgelöst:  $e^t = x$ . Der Term mit der neuen Variablen wird als neue Funktion  $g(t)$  aufgefasst, also hier  $g(t) = e^t$ .
3. Wir bestimmen die Ableitung  $g'$ :  $g'(t) = e^t$

4. Ersetzen das Differential  $dx$  mit  $g'(t)dt$  und führe die Substitution wie in Schritt 1 bzw. ausgewählt durch:

$$\int \frac{\ln(x)}{x\sqrt{1+\ln^2(x)}} dx = \int \frac{t}{e^t\sqrt{1+t^2}} e^t dt = \int \frac{t}{\sqrt{1+t^2}} dt$$

Wenn alles gut gegangen ist, sollte der neue **Integrand leichter zu integrieren sein**. In dem Beispiel hier sind wir noch nicht ganz so weit, dass man leicht die Stammfunktion hinschreiben kann, es sei denn man hat einen echt scharfen Blick auf die Kettenregel. Es gibt nämlich einen Sonderfall. Es gilt:

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

Bei Beispiel bzw. bei der Aufgabe von Blatt 10 ist:

$$\int \frac{t}{\sqrt{1+t^2}} dt = \int \frac{2t}{2\sqrt{1+t^2}} dt = \int \underbrace{2t}_{=g'(x)} \cdot \underbrace{\frac{1}{2\sqrt{1+t^2}}}_{=f'(g(x))} dt = \sqrt{1+t^2}$$

Ganz ehrlich: Wenn man nicht ganz viel solcher Aufgaben gemacht hat und immer überprüft, ob der Integrand vielleicht schon das Ergebnis einer Ableitung gemäß der Kettenregel ist, sieht man das nicht.

**Noch zwei häufige Sonderfälle, wo die Regel simpel ist:**

**Lineare Substitution** (wird in der Schule auch gemacht):

Besteht der Integrand aus einer verketteten Funktion, wobei die äußere Funktion  $f$  die Stammfunktion  $F$  besitzt und die innere Funktion linear von der Form  $ax + b$  ist, so lautet die Lösung des Integrals folgendermaßen:

$$\int f(ax+b) dx = \frac{1}{a} F(ax+b)$$

Beispiel:

$$\int \cos(5x+1) dx = \frac{1}{5} \sin(5x+1)$$

### Logarithmische Integration

Ist der Integrand ein Bruch mit einer Funktion  $f$  im Nenner und deren Ableitung  $f'$  im Zähler, so ist der natürliche Logarithmus der Funktion  $f$  die gesuchte Stammfunktion.

$$\int \frac{f'(x)}{f(x)} dx = \ln |f(x)|$$

Beispiel: - Hier wird erst noch geschickt erweitert, damit im Zähler die Ableitung vom Nenner steht!

$$\int \frac{\sin(5x)}{\cos(5x)} dx = -\frac{1}{5} \int \frac{-5 \cdot \sin(5x)}{\cos(5x)} dx = -\frac{1}{5} \ln |\cos(5x)|$$

Hier findest du auch Erklärungen, die die Grenzen berücksichtigen:

<https://studyflix.de/mathematik/integration-durch-substitution-1783>

## Rationale Funktionen

Hier geht es um die Integration gebrochen rationaler Funktionen. Im Zentrum steht eine Umformungstechnik, die man Partialbruchzerlegung nennt.

Zur Erinnerung:

**Definition:** Der Funktionsterm  $f(x) = a_n x^n + a_{n-1} x^{n-1} + \dots + a_1 x + a_0$  mit  $n \in \mathbb{N}$ ,  $a_n, \dots, a_0 \in \mathbb{R}$ ,  $a_n \neq 0$ , heißt **ganzrationales Polynom** in  $x$ . Die Zahlen  $a_n, \dots, a_0$  heißen Koeffizienten und  $n$  ist der **Grad** des Polynoms. Entsprechende Funktionen heißen **ganzrationale Funktionen**.

**Definition:** Eine Funktion  $f$ , deren Term sich in der Form  $f(x) = \frac{Z(x)}{N(x)}$  darstellen lässt, wobei  $Z(x)$  und  $N(x)$  Polynome sind und Grad von  $N(x) > 0$  ist, heißt **gebrochenrationale Funktion**.

Gilt Grad von  $Z(x) \geq$  Grad von  $N(x)$ , dann heißt  $f$  **unecht gebrochenrational**. Ansonsten handelt es sich um eine **echt gebrochenrationale Funktion**.

Zur Sprechweise eine Tabelle mit der Bruch-Analogie

Zahl als Bruch	Funktion
Ganze Zahl (Nenner 1), z. B. $\frac{3}{1}$	Ganzrational (Nenner konstant), z. B. $\frac{x^2 + 1}{2}$
Echter Bruch, z. B. $\frac{2}{3}$	echt gebrochenrational $\frac{x^2 + 1}{x^6 + x^4 + 1}$
Unechter Bruch, z. B. $\frac{5}{2}$	unecht gebrochenrational $\frac{x^2 + 1}{4x}$
Gemischte Schreibweise: $\frac{5}{2} = 2\frac{1}{2}$	Gemischte Schreibweise: $\frac{x^2 + x + 1}{x^2 + 1} = \frac{x^2 + 1}{x^2 + 1} + \frac{x}{x^2 + 1} = 1 + \frac{x}{x^2 + 1}$

### Partialbruchzerlegung 12.2.5

Es sei  $R = \frac{P}{Q}$  eine echt gebrochene rationale Funktion und  $Q$  durch die Produktdarstellung

$$Q(z) = \prod_{k=1}^r (z - z_k)^{p_k}$$

gegeben. Dann besitzt  $R$  für  $z \notin \{z_1, \dots, z_r\}$  eine **Partialbruchdarstellung** der Form

$$\begin{aligned} R(z) &= \sum_{k=1}^r \sum_{\ell=1}^{p_k} \frac{C_k^{(\ell)}}{(z - z_k)^\ell} \\ &= \frac{C_1^{(1)}}{z - z_1} + \dots + \frac{C_1^{(p_1)}}{(z - z_1)^{p_1}} + \dots + \frac{C_r^{(p_r)}}{(z - z_r)^{p_r}} \end{aligned} \quad (12.3)$$

mit Konstanten  $C_k^{(\ell)} \in \mathbb{C}$ ,  $k = 1, \dots, r$ ,  $\ell = 1, \dots, p_k$ .

Dieses ist erstmal ein Existenzsatz. Zu einer echt gebrochenrationalen Funktion gibt es eine Partialbruchdarstellung, d. h. man kann diese Funktion auch als Summe von gewissen Teilbrüchen darstellen.

### Reelle Partialbruchzerlegung 12.2.6

Es sei  $R = \frac{P}{Q}$  eine echt gebrochene rationale Funktion und  $Q$  möge durch die reelle Produktdarstellung

$$Q(x) = (x - x_1)^{p_1} \cdots (x - x_k)^{p_k} \cdot (x^2 + b_1x + c_1)^{t_1} \cdots (x^2 + b_\ell x + c_\ell)^{t_\ell}$$

wie oben gegeben sein, dann besitzt  $R$  eine *Partialbruchdarstellung* der Form

$$\begin{aligned} R(x) &= \sum_{i=1}^k \sum_{j=1}^{p_i} \frac{A_i^{(j)}}{(x - x_i)^j} + \sum_{i=1}^\ell \sum_{j=1}^{t_i} \frac{B_i^{(j)}x + C_i^{(j)}}{(x^2 + b_i x + c_i)^j} \\ &= \sum_{i=1}^k \left( \frac{A_i^{(1)}}{x - x_i} + \cdots + \frac{A_i^{(p_i)}}{(x - x_i)^{p_i}} \right) \\ &\quad + \sum_{i=1}^\ell \left( \frac{B_i^{(1)}x + C_i^{(1)}}{x^2 + b_i x + c_i} + \cdots + \frac{B_i^{(t_i)}x + C_i^{(t_i)}}{(x^2 + b_i x + c_i)^{t_i}} \right) \end{aligned}$$

mit  $A_i^{(j)} \in \mathbb{R}$  für  $i = 1, \dots, k, j = 1, \dots, p_i$  und  $B_i^{(j)}, C_i^{(j)} \in \mathbb{R}$  für  $i = 1, \dots, \ell, j = 1, \dots, t_i$ .

Je nachdem, was für eine echt gebrochenrationale Funktion vorliegt, gibt es u. a. besondere Darstellungsmöglichkeiten, wobei die Summanden nicht beliebige Brüche darstellen. Dieses ist der Schlüssel zur Integration dieser Funktionen:

**Integration rationaler Funktionen 12.2.9.** Aufgrund der Linearität des Integrals und der reellen Partialbruchzerlegung genügt es nun offenbar, Ausdrücke der Form

$$\frac{1}{(x - c)^n}, \frac{1}{(x^2 + bx + c)^n}, \frac{x}{(x^2 + bx + c)^n} \quad (12.4)$$

mit Konstanten  $b, c \in \mathbb{R}, n \in \mathbb{N}$  und  $\forall x \in \mathbb{R} (x^2 + bx + c \neq 0)$  zu integrieren.  
Es gilt

$$(i) \int \frac{dx}{x - c} = \ln|x - c|.$$

$$(ii) \int \frac{dx}{(x - c)^n} = \frac{1}{(1 - n)(x - c)^{n-1}} \text{ für } n > 1.$$

$$\begin{aligned} (iii) \int \frac{dx}{x^2 + bx + c} &= \int \frac{dx}{\left(x + \frac{b}{2}\right)^2 + \frac{4c-b^2}{4}} \\ &= \frac{4}{4c-b^2} \int \frac{dx}{\left(\frac{2x+b}{\sqrt{4c-b^2}}\right)^2 + 1} \\ &= \frac{2}{\sqrt{4c-b^2}} \arctan \frac{2x+b}{\sqrt{4c-b^2}}. \end{aligned}$$

$$\begin{aligned} (iv) \int \frac{x}{x^2 + bx + c} dx &= \frac{1}{2} \int \frac{2x+b}{x^2 + bx + c} dx - \frac{b}{2} \int \frac{dx}{x^2 + bx + c} \\ &= \frac{1}{2} \ln|x^2 + bx + c| - \frac{b}{\sqrt{4c-b^2}} \arctan \frac{2x+b}{\sqrt{4c-b^2}}. \end{aligned}$$

(v) Für  $n > 1$  gibt es für die verbleibenden Terme in (12.4) Rekursionsformeln, siehe etwa [20, 7.4].

In einem Exkurs geben ich an, wie man praktisch vorgeht, um die gebrochenrationale Funktion in Partialbrüche zu zerlegen. Zunächst noch:

### Satz über die Integration rationaler Funktionen 12.2.10

Es sei  $R(x) = \frac{P(x)}{Q(x)}$  eine nicht notwendigerweise echt gebrochen rationale Funktion.  $Q(x)$  sei durch die Produktdarstellung

$$Q(x) = (x - x_1)^{p_1} \cdot \dots \cdot (x - x_k)^{p_k} \cdot (x^2 + b_1x + c_1)^{q_1} \cdot \dots \cdot (x^2 + b_\ell x + c_\ell)^{q_\ell}$$

gegeben, dabei sind  $x_1, \dots, x_k \in \mathbb{R}$  die paarweise verschiedenen Nullstellen von  $Q(x)$  mit den Vielfachheiten  $p_1, \dots, p_k \in \mathbb{N}$ . Weiter sind  $x^2 + b_1x + c_1, \dots, x^2 + b_\ell x + c_\ell$  paarweise verschiedene quadratische Polynome, welche keine reelle Nullstellen besitzen, das heißt es gilt  $4c_1 - b_1^2, \dots, 4c_\ell - b_\ell^2 > 0$ , und  $q_1, \dots, q_\ell \in \mathbb{N}$ . Dann gilt eine Darstellung der Form

$$\int R(x) dx = S(x) + \sum_{i=1}^k A_i \ln|x - x_i| + \sum_{j=1}^\ell \left( B_j \ln|x^2 + b_j x + c_j| + C_j \arctan \frac{2x + b_j}{\sqrt{4c_j - b_j^2}} \right).$$

Dabei ist  $S(x)$  eine rationale Funktion, welche echt gebrochen ist, falls  $R(x)$  es ist, und  $A_1, \dots, A_k, B_1, \dots, B_\ell, C_1, \dots, C_\ell \in \mathbb{R}$ .

Auch diese hier ist mehr ein Existenzsatz. Man kann unter den gegebenen Umständen eine solche Stammfunktion finden.

Das konkrete und vermutlich für die Übungen / Klausuren relevante Vorgehen lernst du im Exkurs und vermutlich auf den Aufgabenblättern!

Zum besseren Verständnis: Ich habe bislang die Partialbruchzerlegung nur im Zusammenhang von Integrationsaufgaben.

**Mein Vorgehen:** Ist die Funktion echt gebrochenrational? Wenn nicht durch Polynomdivision in einen ganzrationalen Teil und in einen echt gebrochenrationalen Teil aufspalten. Der ganzrationale Teil lässt sich leicht im Kopf integrieren (Umkehrung der Potenzregel). Für den Rest, sofern nicht bereits als bekannter Bruch (siehe gelber Balken oben) bekannt, kommt dann die Partialbruchzerlegung zum Einsatz, wobei auch dabei bestimmte Fälle auftreten.

#### Beispiel 1:

Es soll  $R(x) = \frac{x^2+x+1}{x^2+1}$  integriert werden. Zunächst erhalten wir durch Polynomdivision:

$$\frac{x^2 + x + 1}{x^2 + 1} = 1 + \frac{x}{x^2 + 1}$$

Damit sind wir bzgl. des Integrierens schon am Ziel und benötigen keine Partialbruchzerlegung, denn es gilt:

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

Zum zweiten Integral: Ich habe hier die logarithmische Integration (Sonderfall bei der Substitution bzw. Kettenregel) genutzt. Vermutlich hätte auch (iv) mit  $b = 0$  oben beim gelben Balken verwendet werden können.

Bevor es mit einem Beispiel mit Partialbruchzerlegung weitergeht, ein paar Hinweise zur Polynomdivision:

## Zur Polynomdivision:

Zum Lernen: Skript S. 319-320 und siehe <https://studyflix.de/mathematik/polynomdivision-2342>

Zum Üben: z. B. <https://www.studimup.de/abitur/lineare-algebra/polynomdivision/>

Zum Überprüfen: Dein TI Nspire Rechner oder z. B. <https://matheguru.com/rechner/polynomdivision>  
 (Falls du deinen Rechner verwenden möchtest, nutze den expand-Befehl („entwickle“) und gib die Polynome (am besten in Klammern) ein: (Z(x)) / (N(x)), z. B. expand ((x<sup>2</sup>+x+1) / (x+1)) liefert: x +  $\frac{1}{x+1}$ ).

Bei Geogebra: Geht mit dem Befehl **Division**(Zähler-Polynom, Nenner-Polynom). Die Ausgabe muss interpretiert werden: {ganzzrationaler Anteil, Rest}, d. h.

Zähler-Polynom : Nenner-Polynom = ganzzrationaler Anteil + Rest/Nenner-Polynom

**Division**(x<sup>2</sup>+x+1, x+1) liefert {x, 1} und das bedeutet der ganzzrationale Anteil ist x und der Rest ist 1 also:  
 $(x^2 + x + 1) : (x + 1) = x + \frac{1}{x+1}$

**Hinweis: Ich verwende in den Beispielen ausschließlich die Einsetz-Methode um bestimmte Koeffizienten zu bestimmen. Am Ende gehe ich auf andere Methoden ein.**

### Beispiel 2: (Nur einfache Nullstellen)

Es soll  $R(x) = \frac{1}{x^3 - x}$  integriert werden. Es liegt bereits eine echt gebrochenrationale Funktion vor. Im nächsten Schritt wird der Nenner faktorisiert. (Beachte: Wenn das nicht geht, dann müsste man ihn komplexe faktorisieren, also mit komplexen Zahlen arbeiten.)

Es gilt:  $x^3 - x = x(x^2 - 1) = x(x + 1)(x - 1)$

In diesem Fall haben wir es mit 3 verschiedenen einfachen Nullstellen zu tun. Der Ansatz für die Partialbruchzerlegung ist nun:

$$\frac{1}{x(x+1)(x-1)} = \frac{A}{x} + \frac{B}{x+1} + \frac{C}{x-1}$$

Nun müssen A, B, und C richtig bestimmt werden.

Zuerst multiplizieren wir mit dem Nenner auf der linken Seite:

$$\begin{aligned} \frac{1}{x(x+1)(x-1)} &= \frac{A}{x} + \frac{B}{x+1} + \frac{C}{x-1} \mid \cdot x(x+1)(x-1) \\ \frac{x(x+1)(x-1)}{x(x+1)(x-1)} &= \frac{A \cdot x(x+1)(x-1)}{x} + \frac{B \cdot x(x+1)(x-1)}{x+1} + \frac{C \cdot x(x+1)(x-1)}{x-1} \end{aligned}$$

Nach Kürzen:

$$1 = A \cdot (x+1)(x-1) + B \cdot x(x-1) + C \cdot x(x+1)$$

Nun setzen wir für x die drei Nullstellen ein:

$$x = 0 \Rightarrow 1 = A \cdot (0+1)(0-1) + 0 + 0, \text{ also } 1 = -A \text{ bzw. } A = -1$$

$$x = -1 \Rightarrow 1 = 0 + B \cdot (-1)(-1-1) + 0, \text{ also } 1 = 2B \text{ bzw. } B = \frac{1}{2}$$

$$x = 1 \Rightarrow 1 = 0 + 0 + C \cdot 1(1+1), \text{ also } 1 = 2C \text{ bzw. } C = \frac{1}{2}$$

Damit ist die Zerlegung gefunden:  $\frac{1}{x(x^2-1)} = -\frac{1}{x} + \frac{1}{2(x+1)} + \frac{1}{2(x-1)}$

Das war einfach. Es gibt komplizierte Fälle, bei denen mehrfache Nullstellen auftreten oder komplexe Nullstellen. Bei mehrfachen Nullstellen werden beim Ansatz alle Potenzen der Nullstelle von 1 bis zur Vielfachheit einzeln berücksichtigt. Dadurch steigt der Aufwand! Bei komplexen Nullstellen werden die Zähler beim Ansatz komplizierter.

### Beispiel 3: (mehrfache Nullstellen)

Es soll  $R(x) = \frac{x}{(x-1)^3(x+1)^2}$  integriert werden. Es liegt bereits eine echt gebrochenrationale Funktion vor. Der Nenner ist schon faktorisiert, aber Vorsicht! Es sind Nullstellen mit Vielfachheiten!

Ansatz:

$$\frac{x}{(x-1)^3(x+1)^2} = \frac{A}{x-1} + \frac{B}{(x-1)^2} + \frac{C}{(x-1)^3} + \frac{D}{x+1} + \frac{E}{(x+1)^2}$$

Multiplizieren mit dem Nenner auf der linken Seite:

$$x = \frac{A \cdot (x-1)^3(x+1)^2}{x-1} + \frac{B \cdot (x-1)^3(x+1)^2}{(x-1)^2} + \frac{C \cdot (x-1)^3(x+1)^2}{(x-1)^3} + \frac{D \cdot (x-1)^3(x+1)^2}{x+1} \\ + \frac{E \cdot (x-1)^3(x+1)^2}{(x+1)^2}$$

Kürzen:

$$x = A \cdot (x-1)^2(x+1)^2 + B \cdot (x-1)(x+1)^2 + C \cdot (x+1)^2 + D \cdot (x-1)^3(x+1) + E \\ \cdot (x-1)^3$$

Einsetzen einfacher Werte für  $x$ , am besten erst mit den Nullstellen anfangen:

$$x = 1 \quad \Rightarrow \quad 1 = 0 + 0 + C \cdot (1+1)^2 + 0 + 0, \text{ also } 1 = 4C \text{ bzw. } C = \frac{1}{4}$$

$$x = -1 \quad \Rightarrow \quad -1 = 0 + 0 + 0 + 0 + E \cdot (-1-1)^3, \text{ also } -1 = E(-8) \text{ bzw. } E = \frac{1}{8}$$

Jetzt wählen wir andere einfache Werte:

$$x = 0 \quad \Rightarrow \quad 0 = A \cdot (0-1)^2(0+1)^2 + B \cdot (0-1)(0+1)^2 + C \cdot (0+1)^2 + D \cdot (0-1)^3(0+1) + E \\ \cdot (0-1)^3$$

Also:

$$0 = A - B + C - D - E = A - B + \frac{1}{4} - D - \frac{1}{8}$$

$$x = 2 \quad \Rightarrow$$

$$2 = A \cdot (2-1)^2(2+1)^2 + B \cdot (2-1)(2+1)^2 + C \cdot (2+1)^2 + D \cdot (2-1)^3(2+1) + E \\ \cdot (2-1)^3$$

Also:

$$2 = A \cdot 9 + B \cdot 9 + C \cdot 9 + D \cdot 3 + E = 9A + 9B + \frac{9}{4} + 3D + \frac{1}{8}$$

$$x = -2 \quad \Rightarrow$$

$$-2 = A \cdot (-2-1)^2(-2+1)^2 + B \cdot (-2-1)(-2+1)^2 + C \cdot (-2+1)^2 + D \cdot (-2-1)^3(-2+1) \\ + E \cdot (-2-1)^3$$

Also:

$$\begin{aligned}-2 &= A \cdot (-3)^2 + B \cdot (-3) + C + D \cdot (-3)^3(-1) + E \cdot (-3)^3 = 9A - 3B + C + 27D - 27E \\ &= 9A - 3B + \frac{1}{4} + 27D - \frac{27}{8}\end{aligned}$$

Nun haben wir **drei Gleichungen** mit den drei verbleibenden Unbekannten. Dieses LGS müssen wir lösen:

$$\left| \begin{array}{l} A - B + \frac{1}{4} - D - \frac{1}{8} = 0 \\ 9A + 9B + \frac{9}{4} + 3D + \frac{1}{8} = 2 \\ 9A - 3B + \frac{1}{4} + 27D - \frac{27}{8} = -2 \end{array} \right|$$

$$\left| \begin{array}{l} A - B - D = -\frac{1}{8} \\ 9A + 9B + 3D = -\frac{3}{8} \\ 9A - 3B + 27D = \frac{9}{8} \end{array} \right|$$

$$\left( \begin{array}{cccc} 1 & -1 & -1 & -\frac{1}{8} \\ 9 & 9 & 3 & -\frac{3}{8} \\ 9 & -3 & 27 & \frac{9}{8} \end{array} \right) \rightarrow \left( \begin{array}{cccc} 1 & -1 & -1 & -\frac{1}{8} \\ 0 & 18 & 12 & \frac{6}{8} \\ 0 & 6 & 36 & \frac{18}{8} \end{array} \right) \rightarrow \left( \begin{array}{cccc} 1 & -1 & -1 & -\frac{1}{8} \\ 0 & 18 & 12 & \frac{6}{8} \\ 0 & 0 & 32 & 2 \end{array} \right)$$

$$\Rightarrow 32D = 2 \Rightarrow D = \frac{1}{16} \Rightarrow 18B + 12 \cdot \frac{1}{16} = \frac{6}{8} \Rightarrow B = 0 \Rightarrow A - \frac{1}{16} = -\frac{1}{8} \Rightarrow A = \frac{1}{16}$$

Also gilt:

$$\frac{x}{(x-1)^3(x+1)^2} = -\frac{1}{16(x-1)} + \frac{1}{4(x-1)^3} + \frac{1}{16(x+1)} + \frac{1}{8(x+1)^2}$$

Und damit:

$$\begin{aligned}\int \frac{x}{(x-1)^3(x+1)^2} dx &= \int -\frac{1}{16(x-1)} dx + \int \frac{1}{4(x-1)^3} dx + \int \frac{1}{16(x+1)} dx + \int \frac{1}{8(x+1)^2} dx \\ &= -\frac{1}{16} \int \frac{1}{x-1} dx + \frac{1}{4} \int \frac{1}{(x-1)^3} dx + \frac{1}{16} \int \frac{1}{x+1} dx + \frac{1}{8} \int \frac{1}{(x+1)^2} dx \\ &= -\frac{1}{16} \ln|x-1| - \frac{1}{4} \cdot \frac{1}{2(x-1)^2} + \frac{1}{16} \ln|x+1| - \frac{1}{8} \frac{1}{x+1}\end{aligned}$$

(Die Stammfunktionen findet man schnell, da die Partialbrüche den Integranden oben beim gelben Balken bzw. den Grundfunktionen auf S. 325-326 entsprechen.)

#### Beispiel 4: (komplexe Nullstellen)

Es soll  $R(x) = \frac{1}{(x-1)(x^2+1)^2}$  integriert werden. Es liegt bereits eine echt gebrochenrationale Funktion vor. Der Nenner ist faktorisiert, aber Vorsicht! Es liegt eine komplexe Nullstelle mit der Vielfachheit 2 vor. Hier kommen andere Glieder (Typ Ax+B) beim Ansatz vor und wegen der Vielfachheit auch gleich zwei davon.

Ansatz:

$$\frac{1}{(x-1)(x^2+1)^2} = \underbrace{\frac{A}{x-1}}_{\substack{\text{wegen der einfachen} \\ \text{reellen Nullstelle}}} + \underbrace{\frac{Bx+C}{x^2+1} + \frac{Dx+E}{(x^2+1)^2}}_{\substack{\text{wegen der komplexen} \\ \text{Nullstelle mit Vielfachheit 2}}}$$

Multiplizieren mit dem Nenner auf der linken Seite:

$$1 = \frac{A(x-1)(x^2+1)^2}{x-1} + \frac{(Bx+C)(x-1)(x^2+1)^2}{x^2+1} + \frac{(Dx+E)(x-1)(x^2+1)^2}{(x^2+1)^2}$$

Kürzen:

$$1 = A(x^2+1)^2 + (Bx+C)(x-1)(x^2+1) + (Dx+E)(x-1)$$

Einsetzen von Nullstellen und einfachen Werten:

$$x = 1 \Rightarrow 1 = A(1+1)^2 + 0 + 0, \text{ also } 1 = 4A \text{ bzw. } A = \frac{1}{4}$$

$$x = 0 \Rightarrow 1 = \frac{1}{4}(0+1)^2 + (0+C)(0-1)(0+1) + (E)(0-1)$$

$$\Rightarrow 1 = \frac{1}{4} - C - E$$

$$x = -1 \Rightarrow 1 = \frac{1}{4}(1+1)^2 + (-B+C)(-1-1)(1+1) + (-D+E)(-1-1)$$

$$\Rightarrow 1 = 1 + 2B - 2C + 2D - 2E$$

$$\Rightarrow 0 = B - C + D - E$$

$$x = 2 \Rightarrow 1 = \frac{1}{4}(2^2+1)^2 + (2B+C)(2-1)(2^2+1) + (2D+E)(2-1)$$

$$\Rightarrow 1 = \frac{25}{4} + 10B + 5C + 2D + E$$

$$x = i \Rightarrow 1 = 0 + 0 + (Di+E)(i-1)$$

$\Rightarrow 1 = -D + iE - Di - E$  (Riskant, ob das LGS sich gut lösen lässt, daher lieber einen anderen Wert für  $x$  einsetzen. Könnte aber auch klappen. Wer weiß?)

$$x = -2 \Rightarrow 1 = \frac{1}{4}(5)^2 + (-2B+C)(-2-1)(5) + (-2D+E)(-2-1)$$

$$\Rightarrow 1 = \frac{25}{4} - 15(-2B+C) - 3(-2D+E)$$

$$\Rightarrow 1 = \frac{25}{4} + 30B - 15C + 6D - 3E$$

$$\left| \begin{array}{l} 1 = \frac{1}{4} - C - E \\ 0 = B - C + D - E \\ 1 = \frac{25}{4} + 10B + 5C + 2D + E \\ 1 = \frac{25}{4} + 30B - 15C + 6D - 3E \end{array} \right.$$

Das LGS hat, wenn man richtig rechnet, die Lösung:  $A = \frac{1}{4}, B = -\frac{1}{4}, C = -\frac{1}{4}, D = -\frac{1}{2}, E = -\frac{1}{2}$

$$\begin{aligned}
\frac{1}{(x-1)(x^2+1)^2} &= \frac{\frac{1}{4}}{x-1} + \frac{-\frac{1}{4}x - \frac{1}{4}}{x^2+1} + \frac{-\frac{1}{2}x - \frac{1}{2}}{(x^2+1)^2} = \frac{1}{4(x-1)} - \frac{x+1}{4(x^2+1)} - \frac{x+1}{2(x^2+1)^2} \\
\int \frac{1}{(x-1)(x^2+1)^2} dx &= \int \frac{1}{4(x-1)} dx - \int \frac{x+1}{4(x^2+1)} dx - \int \frac{x+1}{2(x^2+1)^2} dx \\
&= \frac{1}{4} \int \frac{1}{x-1} dx - \frac{1}{4} \int \frac{x+1}{x^2+1} dx - \frac{1}{2} \int \frac{x+1}{(x^2+1)^2} dx \\
&= \frac{1}{4} \ln|x-1| - \frac{1}{4} \int \frac{x}{x^2+1} dx - \frac{1}{4} \int \frac{1}{x^2+1} dx - \frac{1}{2} \int \frac{x}{(x^2+1)^2} dx - \frac{1}{2} \int \frac{1}{(x^2+1)^2} dx \\
&= \frac{1}{4} \ln|x-1| - \frac{1}{8} \ln|x^2+1| - \frac{1}{4} \arctan x - \frac{1}{4} \int \frac{2x}{(x^2+1)^2} dx - \frac{1}{2} \int \frac{1}{(x^2+1)^2} dx \\
&= \frac{1}{4} \ln|x-1| - \frac{1}{8} \ln|x^2+1| - \frac{1}{4} \arctan x - \frac{1}{4} \left(-\frac{1}{x^2+1}\right) - \frac{1}{2} \int \frac{1}{(x^2+1)^2} dx \\
&= \frac{1}{4} \ln|x-1| - \frac{1}{8} \ln|x^2+1| - \frac{1}{4} \arctan x + \frac{1}{4(x^2+1)} - \frac{1}{2} \int \frac{1}{(x^2+1)^2} dx
\end{aligned}$$

Beim letzten Integral kommen wir zu einem wunden Punkt. Arbeiten wir mal einfach weiter:

Versuche für  $\frac{1}{(x^2+1)^2}$  eine Zerlegung zu finden:

$$\begin{aligned}
\frac{1}{(x^2+1)^2} &= \frac{Ax+B}{x^2+1} + \frac{Cx+D}{(x^2+1)^2} \\
1 &= (Ax+B)(x^2+1)^2 + Cx+D
\end{aligned}$$

$$\begin{aligned}
x = 0 \quad &\Rightarrow \quad 1 = B + D \Rightarrow D = 1 - B \\
&\Rightarrow \quad 1 = (Ax+B)(x^2+1)^2 + Cx + 1 - B \\
&\Rightarrow \quad 0 = (Ax+B)(x^2+1)^2 + Cx - B \\
x = 1 \quad &\Rightarrow \quad 0 = (A+B)4 + C - B = 4A + 3B + C \\
x = -1 \quad &\Rightarrow \quad 0 = (-A+B)4 - C - B = -4A - 5B - C
\end{aligned}$$

Aus den letzten beiden Gleichungen

$$0 = 4A + 3B + C$$

$$0 = -4A - 5B - C$$

Folgt durch Addition  $0 = -2B$ , also  $B = 0$  und damit  $D = 1$ . Bleibt der Zusammenhang zwischen  $A$  und  $C$  zu klären.

$$x = i \quad \Rightarrow \quad 0 = (Ai+B) \cdot 0 + Ci - B \quad \Rightarrow \quad 0 = Ci - 0 \quad \Rightarrow 0 = C \Rightarrow A = 0$$

Setzt man ein, so erhält man

$$\frac{1}{(x^2+1)^2} = \frac{0 \cdot x + 0}{x^2+1} + \frac{0 \cdot x + 1}{(x^2+1)^2} = \frac{1}{(x^2+1)^2}$$

OH-HA! **Wir haben keine neue Zerlegung gefunden!** Dieses kann offenbar bei komplexen Nullstellen auftreten. Ich konnte keine Information darüber finden, wann das der Fall ist. Wir hatten hier auf



jeden Fall ausschließlich eine einzige komplexe Nullstelle. Bei doppelter Vielfachheit kann man keine Zerlegung vornehmen. Wenn der Nenner hoch 3 ist, gibt es wieder eine komplexe Zerlegung. Vielleicht kannst du mal nachfragen unter welchen Umständen keine Zerlegung existiert. Gibt ruhig diesen Bruch als Beispiel an. Klar ist, dass  $\frac{1}{x^2}$ , usw. nicht weiter zerlegbar sind. Aber bei anderen Brüchen sehe ich das nicht so leicht, wie z. B. bei  $\frac{1}{x^2+1} = \frac{1}{(x+i)(x-i)}$ . Warum sollte es hier keine Zerlegung geben? Bei  $\frac{1}{(x^2+1)^2} = \frac{1}{(x+i)^2(x-i)^2}$  auch nicht, aber bei  $\frac{1}{(x^2+1)^3} = \frac{1}{(x+i)^3(x-i)^3}$  schon?

Wir müssen das Integral also anders knacken. Auf der einen Seite ist  $\frac{1}{(x^2+1)^2}$  ein typischer Ausdruck, zu dem man in einer Formelsammlung die Stammfunktion nachschlagen kann. Auf der anderen Seite erwähnt das Skript am Rande, dass es eine Rekursionsformel gibt, gibt diese aber nicht an. Sie lautet:

$$\begin{aligned} \int \frac{1}{(x^2 + bx + c)^{n+1}} dx \\ = \frac{1}{(4c - b^2) \cdot n} \cdot \frac{2x + b}{(x^2 + bx + c)^n} + \frac{2}{4c - b^2} \cdot \frac{2n - 1}{n} \int \frac{1}{(x^2 + bx + c)^n} dx + const \end{aligned}$$

Bei  $\frac{1}{(x^2+1)^2}$  ist  $b = 0$  und  $c = 1$  und  $n = 1$  zu verwenden

$$\int \frac{1}{(x^2 + 1)^2} dx = \frac{1}{4} \cdot \frac{2x}{x^2 + 1} + \frac{1}{2} \cdot \int \frac{1}{x^2 + 1} dx = \frac{1}{2} \cdot \frac{x}{x^2 + 1} + \frac{1}{2} \cdot \arctan x$$

Insgesamt erhalten wir so

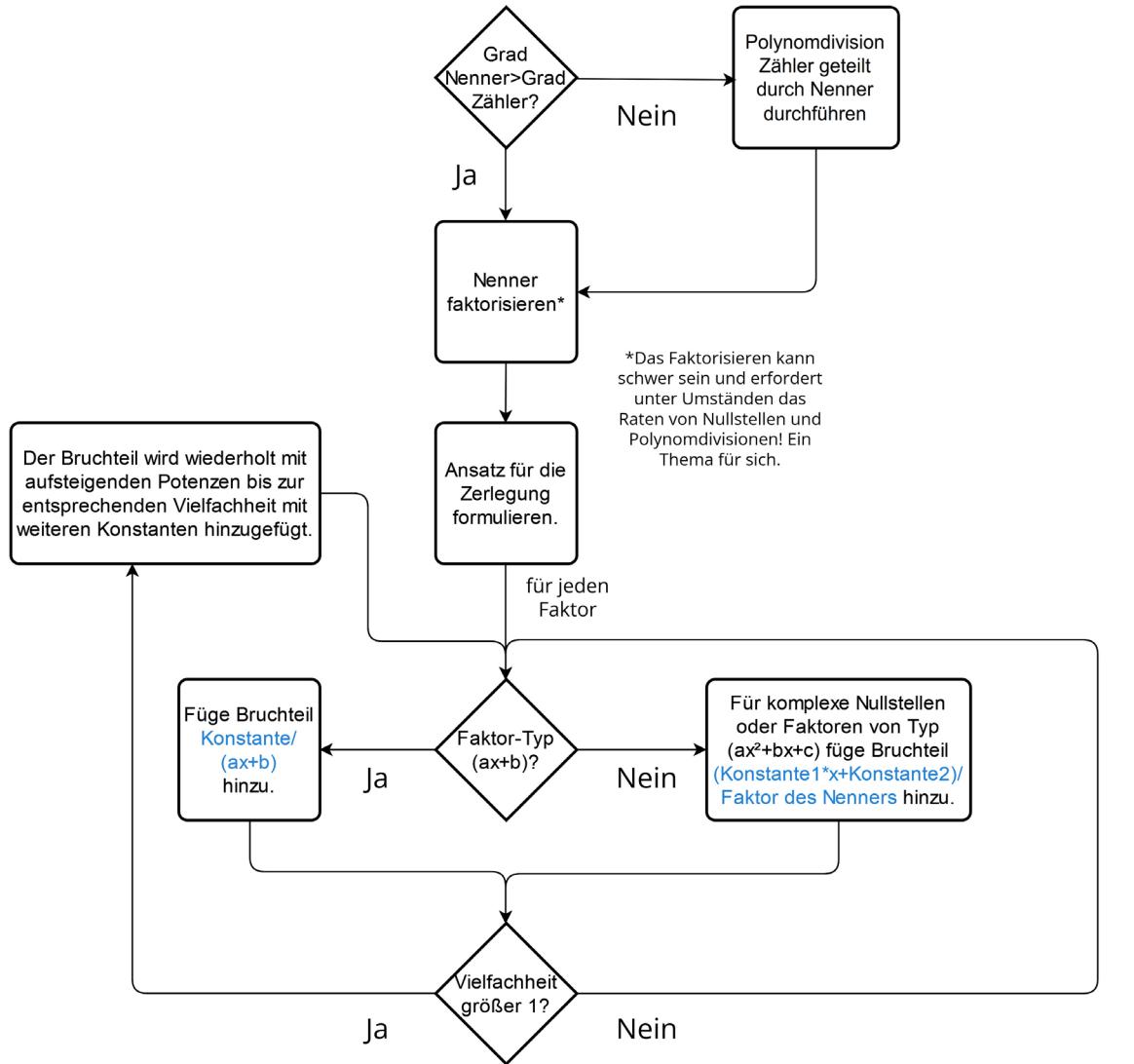
$$\begin{aligned} \int \frac{1}{(x - 1)(x^2 + 1)^2} dx \\ = \frac{1}{4} \ln|x - 1| - \frac{1}{8} \ln|x^2 + 1| - \frac{1}{4} \arctan x + \frac{1}{4(x^2 + 1)} \\ - \frac{1}{2} \left( \frac{1}{2} \cdot \frac{x}{x^2 + 1} + \frac{1}{2} \cdot \arctan x \right) \\ = \frac{1}{4} \ln|x - 1| - \frac{1}{8} \ln|x^2 + 1| - \frac{1}{4} \arctan x + \frac{1}{4(x^2 + 1)} - \frac{1}{4} \cdot \frac{x}{x^2 + 1} - \frac{1}{4} \cdot \arctan x \\ = \frac{1}{4} \ln|x - 1| - \frac{1}{8} \ln|x^2 + 1| - \frac{1}{2} \arctan x + \frac{1}{4(x^2 + 1)} - \frac{x}{4(x^2 + 1)} \end{aligned}$$

Du siehst, dass solche Aufgaben knifflig werden können. Je komplexer die rationale Funktion, desto größer wird das LGS zum Lösen.

Die gute Nachricht ist: Alle rationalen Funktionen sind integrierbar. Es erfordert nur ggf. einen hohen Aufwand.

Zum Ende möchte ich auf andere Methoden zur Bestimmung der Partialbruchzerlegung eingehen und nochmal das Vorgehen zusammenfassen:

## Vorgehen bei der Partialbruchzerlegung



## **Ansatz:**

1. Prüfe, ob eine Polynomdivision notwendig ist.  
Dies ist immer der Fall, wenn der Zählergrad  $\geq$  Nennergrad ist.
  2. Faktorisiere das Nennerpolynom. (Allein das kann schon hart sein, z. B. weil man Nullstellen raten muss und dann mit dieser eine Polynomdivision durchgeführt werden muss, aber ich denke die meisten Aufgaben werden hier einfach sein. Bei sehr komplexen Polynomen ist ohne Computerunterstützung schnell Schluss, auch wenn es entsprechende Verfahren wie die Faktorisierung mit Grobner-Basen gibt.)
  3. Jetzt kommt es darauf an, wie die Faktoren im Nenner aussehen. Es gilt:

Faktor im Nenner	Terme in der Partialbruchzerlegung
	$\frac{A}{(a \cdot x + b)}$
$(a \cdot x + b)^k$	$\frac{A_1}{(a \cdot x + b)} + \frac{A_2}{(a \cdot x + b)^2} + \dots + \frac{A_k}{(a \cdot x + b)^k}$
$(a \cdot x^2 + b \cdot x + c)$ gilt auch für komplexe Nullstellen	$\frac{A \cdot x + B}{(a \cdot x^2 + b \cdot x + c)}$
$(a \cdot x^2 + b \cdot x + c)^k$ gilt auch für komplexe Nullstellen	$\frac{A_1 \cdot x + B_1}{(a \cdot x^2 + b \cdot x + c)} + \dots + \frac{A_k \cdot x + B_k}{(a \cdot x^2 + b \cdot x + c)^k}$

Nach dem der Ansatz gefunden wurde, müssen die Koeffizienten bestimmt werden. Typisches Vorgehen:

Die **Einsetzungsmethode (Wertevergleich)** wurden in den Beispielen oben verwendet. Hier wird zunächst mit dem „Hauptnenner“ multipliziert. Dadurch erhält man eine Gleichung ohne  $x$  im Nenner. Anschließend werden die Nullstellen und andere einfache Werte für  $x$  eingesetzt. Man erhält so mehrere Gleichungen, die ein LGS milden, das gelöst werden kann.

Der **Koeffizientenvergleich** liefert oft auch schnell die relevanten Koeffizienten. Dazu geht man zunächst wie bei der Einsetzungsmethode vor, multipliziert aber die linke Seite vollständig aus. Beispiel:

$$\begin{aligned}\frac{1}{x(x+1)(x-1)} &= \frac{A}{x} + \frac{B}{x+1} + \frac{C}{x-1} \\ \frac{x(x+1)(x-1)}{x(x+1)(x-1)} &= \frac{Ax(x+1)(x-1)}{x} + \frac{Bx(x+1)(x-1)}{x+1} + \frac{Cx(x+1)(x-1)}{x-1} \\ 1 &= A(x+1)(x-1) + Bx(x-1) + Cx(x+1) \\ 1 &= A(x^2 - 1) + B(x^2 - x) + C(x^2 + x) \\ 1 &= Ax^2 - A + Bx^2 - Bx + Cx^2 + Cx \\ 1 &= Ax^2 + Bx^2 + Cx^2 - Bx + Cx - A \\ 1 &= (A + B + C)x^2 + (-B + C)x - A\end{aligned}$$

Damit mit der richtigen Wahl von A, B und C auf der rechten Seite das gleiche herauskommt, wie auf der linken Seite, muss offenbar  $A + B + C = 0$ ,  $B = C$  und  $A = -1$  gelten. Diese kleine LGS kann man schon im Kopf lösen.  $\Rightarrow -1 + B + B = 0 \Rightarrow B = \frac{1}{2} = C$  - Fertig.

Eine andere auch im Skript angesprochene Methode ist die **Grenzwertmethode (Abdeck-Methode, Zuhalte-Methode)**, die evtl. nicht immer funktioniert, aber für die meisten einfachen Übungsaufgaben schon:

Wir schauen uns die Methode an der Beispielaufgabe an:

$$\frac{1}{x(x+1)(x-1)} = \frac{A}{x} + \frac{B}{x+1} + \frac{C}{x-1}$$

Nun decken wir für jede Nullstelle auf der linken Seite den entsprechenden Faktor ab und berechnen den Grenzwert, wenn man sich mit  $x$  dieser Stelle nähert. Dieser Grenzwert ist der korrespondierende Koeffizient.

A: Für den Faktor  $x$  müssen wir  $x \rightarrow 0$  betrachten:

$$\lim_{x \rightarrow 0} \frac{1}{\cancel{x}(x+1)(x-1)} = 1 = A$$

B: Für den Faktor  $x+1$  müssen wir  $x \rightarrow -1$  betrachten:

$$\lim_{x \rightarrow -1} \frac{1}{x \cancel{(x+1)}(x-1)} = \frac{1}{-1(-1-1)} = \frac{1}{2} = B$$

C: Für den Faktor  $x-1$  müssen wir  $x \rightarrow 1$  betrachten:

$$\lim_{x \rightarrow 1} \frac{1}{x(x+1)\cancel{(x-1)}} = \frac{1}{1(1+1)} = \frac{1}{2} = C$$

**Warum funktioniert diese Methode?** Multipliziert man mit einem Faktor aus dem Nenner (das passt beim „Abdecken“, so fällt er auf der linken Seite weg, z. B.

$$\frac{1}{x(x+1)(x-1)} = \frac{A}{x} + \frac{B}{x+1} + \frac{C}{x-1} \quad | \cdot (x+1)$$

$$\frac{1}{x(x-1)} = \frac{A}{x}(x+1) + B + \frac{C}{x-1}(x+1)$$

Bei der Grenzwertbetrachtung  $x \rightarrow -1$  fallen die Terme mit  $(x+1)$  weg:

$$\lim_{x \rightarrow -1} \frac{1}{x(x-1)} = 0 + B + 0 = B$$

Diese Methode entspricht praktisch der Einsetzungsmethode aus den Beispielen, wenn man die Nullstellen einsetzt. Man spart sich nur die Multiplikation mit dem Hauptnenner.

**Einschränkungen:** Für einfache Partialbrüche geht das immer ganz gut, aber heikel wird es, wenn Vielfachheiten und komplexere Partialbrüche auftreten, z. B. hier:

$$\frac{x}{(x-1)^3(x+1)^2} = \frac{A}{x-1} + \frac{B}{(x-1)^2} + \frac{C}{(x-1)^3} + \frac{D}{x+1} + \frac{E}{(x+1)^2}$$

Hier können wir nur mit der größten Potenz einer Nullstelle multiplizieren (abdecken), weil ansonsten der Grenzwert nicht existiert (es sind ja Polstellen!).

So bekommen wir in dem Beispiel aber immerhin  $C$  und  $E$  sofort heraus.

$$\lim_{x \rightarrow 1} \frac{x}{\cancel{(x-1)^3}(x+1)^2} = \frac{1}{(1+1)^2} = \frac{1}{4} = C$$

$$\lim_{x \rightarrow -1} \frac{x}{(x-1)^3(x+1)} = \frac{-1}{(-1-1)^3} = \frac{1}{8} = E$$

$$\frac{x}{(x-1)^3(x+1)^2} = \frac{A}{x-1} + \frac{B}{(x-1)^2} + \frac{1}{4(x-1)^3} + \frac{D}{x+1} + \frac{1}{8(x+1)^2}$$

Theoretisch könnte man jetzt so weiter machen: Die bestimmten Glieder können subtrahiert werden und die linke Seite zusammengefasst werden:

$$\frac{x}{(x-1)^3(x+1)^2} - \frac{1}{4(x-1)^3} - \frac{1}{8(x+1)^2} = \frac{A}{x-1} + \frac{B}{(x-1)^2} + \frac{D}{x+1}$$

$$\frac{8x}{8(x-1)^3(x+1)^2} - \frac{2(x+1)^2}{8(x-1)^3(x+1)^2} - \frac{(x-1)^3}{8(x-1)^3(x+1)^2} = \frac{A}{x-1} + \frac{B}{(x-1)^2} + \frac{D}{x+1}$$

$$\frac{8x - 2(x+1)^2 - (x-1)^3}{8(x-1)^3(x+1)^2} = \frac{A}{x-1} + \frac{B}{(x-1)^2} + \frac{D}{x+1}$$

Auf der linken Seite würde man als nächstes den Zähler ausmultiplizieren und zusammenfassen und durch  $x+1$  oder durch  $x-1$  teilen, um wieder einen Ausdruck zu erhalten, mit dem man mit der Grenzwert-Methode fortfahren kann. Sprich: Polynomdivision oder Horner-Schema. Bäh! Bei einer Division durch  $x+1$  (entspricht dem Kürzen mit  $x+1$  auf der linken Seite) würde man erhalten:

$$\frac{-x^2 + 2x + 1}{8(x-1)^3(x+1)} = \frac{A}{x-1} + \frac{B}{(x-1)^2} + \frac{D}{x+1}$$

Nun könnte man wieder die Methode für den Faktor  $x+1$  anwenden und  $D$  bestimmen. Das Ganze ist aber wegen den notwendigen Termumformungen und Polynomdivisionen aufwendig. Stattdessen kann die Methode erweitert werden, in dem der Residuensatz der Funktionentheorie verwendet wird. Es ist die Heaviside-Methode mit Residuen. Allerdings treten hier Ableitungen auf, die bei hohen Vielfachheiten auch einen hohen Aufwand zur Folge haben.

Zunächst noch mal die Grenzwert-Methode bei komplexen Termen:

$$\frac{1}{(x-1)(x^2+1)^2} = \underbrace{\frac{A}{x-1}}_{\substack{\text{wegen der einfachen} \\ \text{reellen Nullstelle}}} + \underbrace{\frac{Bx+C}{x^2+1}}_{\substack{\text{wegen der komplexen} \\ \text{Nullstelle mit Vielfachheit 2}}} + \frac{Dx+E}{(x^2+1)^2}$$

Wir können hier wegen der Vielfachheit nur  $(x^2+1)^2$  abdecken und lasse  $x \rightarrow i$  laufen.

$$\lim_{x \rightarrow i} \frac{1}{(x-1)(x^2+1)^2} = \frac{1}{i-1} = Di + E$$

$$(Di + E)(i-1) = 1$$

$$(-D + Ei) - Di - E = 1$$

$$-D - E + (E - D)i = 1$$

Koeffizientenvergleich:  $E = D$ , damit  $i$  wegfällt und  $-D - E = -D - D = -2D = 1$ , also  $D = E = \frac{1}{2}$

$A$  können wir auch leicht bestimmen, aber an  $B$  und  $C$  kommen wir nicht direkt rann.

Wie man sieht, kann diese Methode sehr nützlich sein, aber sie hat (vom Aufwand her) auch ihre Grenzen. Es ist aber auch möglich Methoden kombiniert zu verwenden.

### Finale Vertiefung (Hilfen zur Grenzwertmethode und die Heaviside-Methode mit Residuen):

Wir haben gesehen, dass die Grenzwertmethode bei Vielfachheiten in der Praxis an ihre Grenze stößt. Wenn die Vielfachheiten nicht zu hoch sind, kann man sie durch folgende Tricks immer noch händisch gut verwenden. Außerdem wird klar, dass man entsprechende Algorithmen beschreiben und programmieren kann, so dass es technisch immer möglich ist, die Partialbruchzerlegung zu berechnen, auch wenn das für die Klausur nicht relevant ist.

#### Fall „Last man standing“ bei nicht zu komplexen Brüchen:

Für den Fall, dass nur eine doppelte Nullstelle auftritt, kann man alle Koeffizienten bis auf eine mit der Grenzwertmethode sofort bestimmen. Den letzten Koeffizienten kann man dann durch eine andere Grenzwertbetrachtung finden:

$$\text{Beispiel: } \frac{1}{(x+1)^2(x-4)} = \frac{A}{x+1} + \frac{B}{(x+1)^2} + \frac{C}{x-4}$$

$$B \text{ und } C \text{ lassen sich sofort bestimmen: } B = \lim_{x \rightarrow -1} \frac{1}{x-4} = -\frac{1}{5}, C = \lim_{x \rightarrow 4} \frac{1}{(x+1)^2} = \frac{1}{25}$$

Damit gilt  $\frac{1}{(x+1)^2(x-4)} = \frac{A}{x+1} - \frac{1}{5(x+1)^2} + \frac{1}{25(x-4)}$ . A bestimmen wir nun, indem wir mit  $x+1$  multiplizieren und dann den Grenzwert  $x \rightarrow \infty$  auf beiden Seiten betrachten.

$$\lim_{x \rightarrow \infty} \frac{1}{(x+1)(x-4)} = \lim_{x \rightarrow \infty} \left( A - \frac{x+1}{5(x+1)^2} + \frac{x+1}{25(x-4)} \right) \Rightarrow 0 = A - 0 + \frac{1}{25}$$

$$\text{Also ist } A = -\frac{1}{25}. \quad \frac{1}{(x+1)^2(x-4)} = -\frac{1}{25(x+1)} - \frac{1}{5(x+1)^2} + \frac{1}{25(x-4)}$$

#### Fall „geringe Vielfachheit“ (Heaviside-Methode mit Residuen):

Ohne auf die Theorie einzugehen, gilt für die Bestimmung der Koeffizienten bei der Partialbruchzerlegung von  $R(x) = \frac{P(x)}{Q(x)} = \frac{P(x)}{(x-a)^n \tilde{Q}(x)}$ , also bei einer  $n$ -fachen Nullstelle im Nenner an der Stelle  $a$ :

$$\frac{P(x)}{(x-a)^n \tilde{Q}(x)} = \frac{A_1}{x-a} + \frac{A_2}{(x-a)^2} + \frac{A_3}{(x-a)^3} + \dots + \frac{A_n}{(x-a)^n} + \text{weitere Terme wegen } \tilde{Q}$$

Die Koeffizienten können mit folgender Formel bestimmt werden. Vielfachheiten haben aber ihren Preis. In diesem Fall Ableitungen:

$$A_k = \lim_{x \rightarrow a} \frac{1}{(n-k)!} \frac{d^{(n-k)}}{dx^{(n-k)}} \left( (x-a)^n \cdot \frac{P(x)}{Q(x)} \right) = \lim_{x \rightarrow a} \frac{1}{(n-k)!} \frac{d^{(n-k)}}{dx^{(n-k)}} \left( \frac{P(x)}{\tilde{Q}(x)} \right)$$

Für  $n=1$  und für  $k=n$  erhält man die normale Grenzwert-Methode. Für  $n > 1$  und  $k \neq n$  werden Ableitungen benötigt. Betrachten wir ein Beispiel mit  $n=2$ :

$$\begin{aligned} \frac{1}{(x+1)^2(x-4)} &= \frac{A_1}{x+1} + \frac{A_2}{(x+1)^2} + \frac{B}{x-4} \\ A_1 &= \lim_{x \rightarrow -1} \frac{1}{(2-1)!} \frac{d^{(2-1)}}{dx^{(2-1)}} \left( (x+1)^2 \cdot \frac{1}{(x+1)^2(x-4)} \right) = \lim_{x \rightarrow -1} \frac{d}{dx} \left( \frac{1}{(x-4)} \right) \\ &= \lim_{x \rightarrow -1} \left( \frac{-1}{(x-4)^2} \right) = -\frac{1}{25} \end{aligned}$$

Die übrigen Koeffizienten können ohne Ableitung bestimmt werden. Diese Methode erfordert die sichere Anwendung der Ableitungsregeln, insbesondere der Quotientenregel.

### Fall „Harakiri“:

Sobald höhere Vielfachheiten auftreten, „explodieren“ die Ableitungen. Auch wenn der gekürzte Bruch  $\frac{P(x)}{\tilde{Q}(x)}$  komplex ist, werden die Ableitungen schnell aufwendig. Das heißt auch Heaviside-Methode mit Residuen ist nicht perfekt. Im Prinzip bleibt dann die Wahl zwischen Pest und Cholera: Komplizierte Ableitungen oder doch Polynomdivisionen (letztere lassen sich übrigens leichter programmieren als komplexe Ableitungen!).

**Das Vorgehen ist also für den Weg mit der Polynomdivision:**

$$\frac{P(x)}{(x-a)^n \tilde{Q}(x)} = \frac{A_1}{x-a} + \frac{A_2}{(x-a)^2} + \frac{A_3}{(x-a)^3} + \dots + \frac{A_n}{(x-a)^n} + \text{weitere Terme wegen } \tilde{Q}$$

- Höchster Koeffizient  $A_n$  bestimmen (Grenzwert-Methode).
- Term auf beiden Seiten subtrahieren:

$$\frac{P(x)}{(x-a)^n \tilde{Q}(x)} - \frac{A_n}{(x-a)^n} = \frac{A_1}{x-a} + \frac{A_2}{(x-a)^2} + \frac{A_3}{(x-a)^3} + \dots + \text{weitere Terme wegen } \tilde{Q}$$

- Term erweitern und Brüche zusammenfassen:

$$\frac{P(x) - A_n \tilde{Q}(x)}{(x-a)^n \tilde{Q}(x)} = \frac{A_1}{x-a} + \frac{A_2}{(x-a)^2} + \frac{A_3}{(x-a)^3} + \dots + \text{weitere Terme wegen } \tilde{Q}$$

- Zähler zusammenfassen zu  $P_n(x) = P(x) - A_n \tilde{Q}(x)$  (erfordert Ausmultiplizieren und Zusammenfassen):

$$\frac{P_n(x)}{(x-a)^n \tilde{Q}(x)} = \frac{A_1}{x-a} + \frac{A_2}{(x-a)^2} + \frac{A_3}{(x-a)^3} + \dots + \text{weitere Terme wegen } \tilde{Q}$$

- Der Zähler kann durch  $(x-a)$  geteilt werden (Polynomdivision) und man erhält so einen veränderten Zähler und einen Nenner mit einer geringeren Vielfachheit der Nullstelle:

$$\frac{\hat{P}_n(x)}{(x-a)^{n-1} \tilde{Q}(x)} = \frac{A_1}{x-a} + \frac{A_2}{(x-a)^2} + \frac{A_3}{(x-a)^3} + \dots + \text{weitere Terme wegen } \tilde{Q}$$

- Nun kann das Verfahren für  $A_{n-1}$  wiederholt werden.

### Fall „Teile und herrschel“:

Zuletzt eine Alternative, wenn man weder höhere Ableitungen noch die Polynomdivision verwenden möchte. Sie ist im Kern rekursiv, führt auf einfache Fälle, aber dafür steigt der Aufwand an einzelnen Teilschritten. Dazu wird durch Klammersetzung dafür gesorgt, dass in der innersten Klammer jede Nullstelle einfach auftritt. Wieder das Beispiel von oben:

$$\frac{1}{(x+1)^2(x-4)} = \frac{1}{(x+1)} \left( \frac{1}{(x+1)(x-4)} \right)$$

Wir finden für diese Teil mit den einfachen eine Partialbruchzerlegung. Das geht fast im Kopf mit der Grenzwert-Methode:

$$\frac{1}{(x+1)(x-4)} = \frac{A}{x+1} + \frac{B}{x-4}$$

$A = -\frac{1}{5}$ ,  $B = \frac{1}{5}$  und erhalten so:

$$\begin{aligned}\frac{1}{(x+1)^2(x-4)} &= \frac{1}{(x+1)} \left( -\frac{1}{5(x+1)} + \frac{1}{5(x-4)} \right) = -\frac{1}{5(x+1)(x+1)} + \frac{1}{5(x-4)(x+1)} \\ &= -\frac{1}{5(x+1)^2} + \frac{1}{5(x-4)(x+1)}\end{aligned}$$

Der grüne Teil ist fertig und lässt sich nicht weiter zerlegen. Es bleibt der blaue Teil, der aber nur einfache Nullstellen im Nenner aufweist:

$$\frac{1}{5(x-4)(x+1)} = \frac{A}{5(x-4)} + \frac{B}{(x+1)}$$

Grenzwert-Methode:  $A = \frac{1}{5}$ ,  $B = -\frac{1}{25}$

Fertige Zerlegung:

$$\frac{1}{(x+1)^2(x-4)} = -\frac{1}{5(x+1)^2} + \frac{1}{25(x-4)} - \frac{1}{25(x+1)}$$

**Zusammenfassend:** Es gibt verschiedene Wege, die Partialbruchzerlegung zu finden. Wichtig ist es einen Weg sicher zu beherrschen. Die stumpfe Einsetz-Methode und der Koeffizientenvergleich führen auf ein LGS, was zu lösen ist. Die anderen Methoden sind in der Durchführung etwas „dynamischer“, führen aber teilweise schnell zu einer Lösung.

Angesichts der Tatsache, dass die größte Vielfachheit in den Klausuraufgaben 2 war, reichen die beschriebenen Methoden sicher aus.

Zur weiteren Vertiefung empfehle ich: <https://palaiologos.rocks/posts/heaviside-pdf/>

## Das Riemannsche Integral

### Definition 12.3.2: Partition, Zerlegung eines Intervalls

Es seien  $I = [a, b]$  ein kompaktes Intervall und Punkte  $x_0, x_1, \dots, x_n \in I$  mit  $a = x_0 < x_1 < \dots < x_n = b$ , so dass  $I$  in  $n$  kompakte Teilintervalle  $I_k = [x_{k-1}, x_k]$  für  $k = 1, 2, \dots, n$  aufgeteilt wird. Dann heißt das Tupel  $\pi := (x_0, x_1, \dots, x_n)$  eine *Partition* oder *Zerlegung* von  $I$ . Wir schreiben

$$\pi : a = x_0 < x_1 < \dots < x_n = b.$$

Dieses ganze Kapitel wird prinzipiell von der Schulmathematik abgedeckt, nur wird dort auf eine exakte Ausdrucksweise verzichtet und Dinge, wie der Hauptsatz der Differential und Integralrechnung vermutlich nicht bewiesen.

### Definition 12.3.5: Ober-, Unter- und Zwischensummen

Es seien  $\pi : a = x_0 < x_1 < \dots < x_n = b$  eine Partition von  $I = [a, b]$  und  $f : I \rightarrow \mathbb{R}$  eine beschränkte Funktion. Für  $k \in \{1, \dots, n\}$  setzen wir

$$M_k := \sup_{I_k} f(x), \quad m_k := \inf_{I_k} f(x).$$

Wir nennen

$$S(\pi, f) = \sum_{k=1}^n M_k |I_k|, \text{ beziehungsweise } s(\pi, f) = \sum_{k=1}^n m_k |I_k|$$

die *Riemann-Darboux'sche Ober-* beziehungsweise *Untersumme*<sup>a</sup> von  $f$  bezüglich  $\pi$ . Sind für  $k \in \{1, \dots, n\}$   $\xi_k \in I_k$ , dann heißt

$$\sigma(\pi, f) = \sigma(\pi, f, \xi_k) = \sum_{k=1}^n f(\xi_k) |I_k|$$

*Riemannsche Zwischensumme* von  $f$  bezüglich  $\pi$  und  $\xi_k, k = 1, \dots, n$ .

<sup>a</sup>Jean Gaston Darboux, 1842-1917, frz. Mathematiker

Im Wesentlichen wird zunächst das Riemannintegral eingeführt. Die ganzen Ausführungen mit Intervallzerlegung, Ober- und Untersumme münden in der Grenzwertbetrachtung, bei dem die Summen für unendlich feine Zerlegungen berechnet werden. In der Schule betrachten wir dazu unendlich viele unendlich dünne Rechteckstreifen, die summiert werden und so die Fläche unter dem Graphen der Funktion darstellen.

### Definition 12.3.8: Riemann-Integral

Es seien  $I = [a, b]$  ein kompaktes Intervall und  $f : I \rightarrow \mathbb{R}$  eine beschränkte Funktion. Dann heißt  $f$  über (auf)  $I$  *Riemann-integrierbar*, wenn  $s(f) = S(f)$  gilt. In diesem Fall heißt der gemeinsame Wert das

*Riemann-Integral* von  $f$  über  $I$ , in Zeichen

$$\int_a^b f(x) dx = \int_I f(x) dx = \int_{[a,b]} f(x) dx := s(f) = S(f).$$

Ansonsten werden hier im nächsten Unterkapitel noch ein paar Eigenschaften angegeben, die aber auch logisch sind.

### Definition 12.3.9: Feinheit einer Partition

Ist  $\pi : a = x_0 < x_1 < \dots < x_n = b$  eine Partition von  $I = [a, b]$ , so heißt

$$|\pi| := \max \{ |x_k - x_{k-1}| \mid k = 1, \dots, n \}$$

*Feinheit* der Partition  $\pi$ .

### Lemma 12.3.10: Technisches Hilfslemma

Es seien  $I = [a, b]$  ein kompaktes Intervall,  $f : I \rightarrow \mathbb{R}$  eine beschränkte Funktion und  $\pi, \pi'$  zwei Partitionen von  $I$ . Es seien  $\pi : a = x_0 < x_1 < \dots < x_n = b$ ,  $\pi' : a = x'_0 < x'_1 < \dots < x'_{n'} = b$  und  $M = \sup_I f(x)$ ,  $m = \inf_I f(x)$ , dann gilt

$$S(\pi', f) \leq S(\pi, f) + (n - 1)(M - m) |\pi'|.$$

### Satz 12.3.11: Berechnung des Integrals über einen Grenzwert

Es seien  $I$  ein kompaktes Intervall,  $f : I \rightarrow \mathbb{R}$  eine beschränkte Funktion und  $(\pi_k)_{k \in \mathbb{N}}$  eine *ausgezeichnete Partitionenfolge*, das heißt es gilt  $\lim_{k \rightarrow \infty} |\pi_k| = 0$ . Dann gilt

$$S(f) = \lim_{k \rightarrow \infty} S(\pi_k, f), \quad s(f) = \lim_{k \rightarrow \infty} s(\pi_k, f).$$

### Satz 12.3.13: Riemannsche Definition des Integrals

Es seien  $I$  ein kompaktes Intervall und  $f : I \rightarrow \mathbb{R}$  eine beschränkte Funktion. Dann ist  $f$  genau dann über  $I$  Riemann-integrierbar, wenn für jede ausgezeichnete Partitionsfolge  $(\pi_k)_{k \in \mathbb{N}}$  und jede Wahl der Zwischenpunkte  $\xi_1, \dots, \xi_n$  mit  $\xi_\ell \in I_\ell$  für  $\ell = 1, \dots, n_k$  die Zwischensumme

$$\sigma(\pi_k, f) = \sigma(\pi_k, f, \xi_\ell) = \sum_{\ell=1}^{n_k} f(\xi_\ell) |I_\ell|$$

konvergiert. In diesem Fall haben alle Summenfolgen den selben Grenzwert und es gilt

$$\lim_{k \rightarrow \infty} \sigma(\pi_k, f) = \int_I f(x) dx.$$

## Klassen integrierbarer Funktionen

### Satz 12.4.1

Jede monotone Funktion  $f : I \rightarrow \mathbb{R}$  ist Riemann-integrierbar.

### Satz 12.4.2

Jede stetige Funktion  $f : I \rightarrow \mathbb{R}$  ist Riemann-integrierbar.

### Satz 12.4.3

Es sei  $f : I \rightarrow \mathbb{R}$  eine beschränkte Funktion, die stückweise monoton<sup>a</sup> oder bis auf endlich viele Punkte stetig ist, dann ist  $f$  Riemann-integrierbar.

<sup>a</sup>das heißt, es existiert eine Partition von  $I$  in endlich viele Teilintervalle  $I_1, \dots, I_n$  so dass  $f$  auf  $I_j$  monoton ist für  $j = 1, \dots, n$ .

## Eigenschaften des Riemann-Integrals

### Satz 12.5.1: Linearität des Riemann-Integrals

Es seien  $f, g : I \rightarrow \mathbb{R}$  Riemann-integrierbar, dann sind auch die Funktionen  $f + g$  und  $f \cdot g$  Riemann-integrierbar. Gilt zusätzlich  $\exists c > 0 \forall x \in I (|f(x)| \geq c)$ , dann ist auch  $\frac{1}{f}$  Riemann-integrierbar. Weiter gilt für alle  $\alpha, \beta \in \mathbb{R}$  die *Linearitätsrelation*

$$\int_I (\alpha f(x) + \beta g(x)) dx = \alpha \int_I f(x) dx + \beta \int_I g(x) dx.$$

Das meiste hier sollte die gut vertraut sein.

### Satz 12.5.2: Additivität des Integrationsbereiches

Es sei  $f : I = [a, b] \rightarrow \mathbb{R}$  auf  $I$  Riemann-integrierbar und  $a < c < b$ ,

dann ist  $f$  auf  $[a, c]$  und auf  $[c, b]$  Riemann-integrierbar und es gilt

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

### Satz 12.5.3: Monotonie des Riemann-Integrals

Es seien  $f, g : I \rightarrow \mathbb{R}$  Riemann-integrierbar und  $\forall x \in I (f(x) \leq g(x))$ , dann gilt

$$\int_I f(x) dx \leq \int_I g(x) dx.$$

### Satz 12.5.4: Dreiecksungleichung für das Riemann-Integral

Die Funktion  $f : I \rightarrow \mathbb{R}$  sei Riemann-integrierbar, dann ist auch die Funktion  $|f| : I \rightarrow [0, +\infty)$  Riemann-integrierbar auf  $I$  und es gilt

$$\left| \int_I f(x) dx \right| \leq \int_I |f(x)| dx.$$

### Mittelwertsatz der Integralrechnung 12.5.5

Es sei  $f : I \rightarrow \mathbb{R}$  Riemann-integrierbar und seien  $m, M \in \mathbb{R}$  mit  $\forall x \in I (m \leq f(x) \leq M)$ . Dann folgen durch Integration die Ungleichungen

$$m |I| \leq \int_I f(x) dx \leq M |I|.$$

Das *Integralmittel*  $\mu := \frac{1}{|I|} \int_I f(x) dx$  genügt also den Ungleichungen  $m \leq \mu \leq M$ . Ist  $f$  stetig, so folgt aus dem Zwischenwertsatz 10.2.7 die Existenz eines  $\xi \in I$  mit  $\mu = f(\xi)$ . Also gilt

$$\int_I f(x) dx = f(\xi) |I|.$$

### Erster Hauptsatz der Differenzial- und Integralrechnung 12.5.6

Es sei  $f : I = [a, b] \rightarrow \mathbb{R}$  Riemann-integrierbar und in  $x_0 \in I$  stetig, dann ist die Funktion

$$F(x) := \int_c^x f(t) dt$$

für  $c \in [a, b]$  in  $x_0$  differenzierbar und es gilt  $F'(x_0) = f(x_0)$ . Ist also  $f$  auf  $I$  stetig, dann ist  $F$  auf  $I$  differenzierbar und es gilt  $\forall x \in I (F'(x) = f(x))$ , das heißt  $F$  ist eine Stammfunktion von  $f$ .

### Zweiter Hauptsatz der Differenzial- und Integralrechnung 12.5.7

Es sei  $F : I \rightarrow \mathbb{R}$  eine differenzierbare Funktion auf einem kompakten Intervall  $I = [a, b]$  und die Ableitung  $f := F' : I \rightarrow \mathbb{R}$  Riemann-integrierbar über  $I$ , dann gilt

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

## Uneigentliche Integrale

### Definition 12.6.1: Uneigentliches Integral, Konvergenz

- (i) Es sei  $I = [a, b)$  mit  $a \in \mathbb{R}$ ,  $b \in \mathbb{R} \cup \{+\infty\}$ ,  $a < b$ . Weiter sei  $f : I \rightarrow \mathbb{R}$  eine Funktion, die für jedes  $c \in \mathbb{R}$  mit  $a < c < b$  auf  $[a, c]$  beschränkt und Riemann-integrierbar ist. Der Grenzwert

$$\lim_{c \rightarrow b^-} \int_a^c f(x) dx =: \int_a^b f(x) dx$$

Eigentlich auch Stoff der Oberstufe im LK. Uneigentliche Integral, als Integrale, wo eine oder beide Grenzen gegen einen „kritischen“ Wert streben.

heißt das *uneigentliche Integral* von  $f$  über  $I$ , falls er existiert. In dem Fall heißt das Integral *konvergent* und es heißt *absolut konvergent*, wenn das Integral von  $|f|$  über  $I$  konvergiert.

- (ii) Analog ist das uneigentliche Integral über das halboffene Intervall  $I = (a, b]$ ,  $a \in \mathbb{R} \cup \{-\infty\}$ ,  $b \in \mathbb{R}$ ,  $a < b$ , definiert durch

$$\int_a^b f(x) dx = \lim_{c \rightarrow a^+} \int_c^b f(x) dx.$$

- (iii) Es sei  $I = (a, b)$ ,  $a \in \mathbb{R} \cup \{-\infty\}$ ,  $b \in \mathbb{R} \cup \{+\infty\}$ ,  $a < b$ , dann definieren wir das uneigentliche Integral von  $f$  über  $I$  durch

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

falls für ein, und damit für alle  $c \in (a, b)$  die beiden uneigentlichen Integrale auf der rechten Seite konvergieren.

### Folgenkriterium 12.6.3

Es seien  $I = [a, b]$ ,  $a \in \mathbb{R}$ ,  $b \in \mathbb{R} \cup \{+\infty\}$ ,  $a < b$ , und  $f : I \rightarrow \mathbb{R}$  eine Funktion, die auf jedem kompaktem Teilintervall  $J = [a, c]$ ,  $a < c < b$ , Riemann-integrierbar ist, dann konvergiert das uneigentliche Integral  $\int_a^b f(x) dx$  genau dann, wenn der Grenzwert

$$\lim_{k \rightarrow \infty} \int_a^{c_k} f(x) dx$$

für jede Folge  $(c_k)_{k \in \mathbb{N}} \subseteq I$  mit  $c_k \rightarrow b$  für  $k \rightarrow \infty$  existiert.

### Satz 12.6.6

Es sei  $I$  ein halboffenes oder offenes Intervall und  $f : I \rightarrow \mathbb{R}$  eine Funktion, die auf jedem kompakten Teilintervall  $J \subseteq I$  beschränkt und Riemann-integrierbar ist. Weiter existiere ein  $M > 0$ , so dass für jedes kompakte Intervall  $J \subseteq I$  die Ungleichung

$$\int_J |f(x)| dx \leq M < +\infty$$

erfüllt ist - dabei ist  $M$  unabhängig von  $J$ . Dann konvergiert das uneigentliche Integral  $\int_I f(x) dx$ .

### Integralkriterium 12.6.8

Es sei  $f : [0, +\infty) \rightarrow [0, +\infty)$  eine nicht-negative, monoton fallende Funktion und für  $k \in \mathbb{N}_0$  sei  $a_k := f(k)$ . Dann gelten für alle  $n \in \mathbb{N}$  die

#### Ungleichungen

$$0 \leq \sum_{k=1}^n a_k \leq \int_0^n f(x) dx \leq \sum_{k=0}^{n-1} a_k.$$

Insbesondere konvergiert die Reihe  $\sum_{k=0}^{\infty} a_k$  genau dann, wenn das uneigentliche Integral  $\int_0^{+\infty} f(x) dx$  konvergiert.

Und so mündet alles in ein Kriterium, dass für die Konvergenz von Reihen verwendet werden kann. Damals durftest du das nicht, weil ihr noch keine Integrale hattet. Nun würde es gehen.

## Fouriertransformation

### Definition 12.7.1: Fouriertransformation

Es sei  $f : \mathbb{R} \rightarrow \mathbb{C}$  stückweise stetig und  $\int_{-\infty}^{+\infty} |f(t)| dt < +\infty$ , dann heißt  $f$  Fourier-transformierbar und die Fouriertransformation (FT) von  $f$  ist definiert durch

$$\mathcal{F} : \mathbb{R} \rightarrow \mathbb{C}, \omega \mapsto \mathcal{F}f(\omega) := \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} f(t)e^{-i\omega t} dt.$$

Die Funktion  $\hat{f}(\omega) := (\mathcal{F}f)(\omega) : \mathbb{R} \rightarrow \mathbb{C}$  heißt die Fouriertransformierte (FT) von  $f$ .

Da nun Integrale eingeführt wurden, kann die Fouriertransformation weiter beschrieben werden. Ich finde das ganz Kapitel ziemlich reingeschüttet und behaupte, dass man eigentlich eine eigene Vorlesung für die Fouriertransformation benötigt, um sie und ihre Anwendungen wirklich zu verstehen. Ich empfehle das Video <https://www.youtube.com/watch?v=zXd743X610w> damit man fußfassen kann. Ansonsten tendiere ich dazu hier zu kapitulieren und freue mich eher auf das letzte Kapitel: DGL

### Satz 12.7.4: Verhalten der Fouriertransformierten bei $\pm\infty$

Es sei  $f : \mathbb{R} \rightarrow \mathbb{C}$  stückweise stetig und Fourier-transformierbar, dann ist  $\hat{f}$  beschränkt und stetig und es gilt

$$\lim_{\omega \rightarrow \pm\infty} \hat{f}(\omega) = 0.$$

### Satz 12.7.5: Rechenregeln für die Fouriertransformation

Es seien  $f, g : \mathbb{R} \rightarrow \mathbb{C}$  Fourier-transformierbar,  $a, b \in \mathbb{C}$ ,  $h \in \mathbb{R}$  und  $\mathbb{R} \ni c > 0$ , dann gilt für  $\omega \in \mathbb{R}$

(i) Linearität:

$$(af(t) + bg(t))^\wedge(\omega) = a\hat{f}(\omega) + b\hat{g}(\omega).$$

(ii) Translation (Verschiebung):

$$(a) (f(t+h))^\wedge(\omega) = e^{ih\omega}\hat{f}(\omega),$$

$$(b) (e^{-iht}f(t))^\wedge(\omega) = \hat{f}(\omega + h).$$

(iii) Streckung:

$$(f(ct))^\wedge(\omega) = \frac{1}{c}\hat{f}\left(\frac{\omega}{c}\right).$$

(iv) Ableitung im Zeitbereich: Falls  $\operatorname{Re} f$  und  $\operatorname{Im} f$  stetig differenzierbar sind und  $f'$  Fourier-transformierbar ist, gilt

$$(f')^\wedge(\omega) = i\omega\hat{f}(\omega).$$

(v) Ableitung im Frequenzbereich: Falls  $t \mapsto tf(t)$  Fourier-transformierbar ist, dann ist  $\hat{f}'$  differenzierbar und es gilt

$$\frac{d}{d\omega}\hat{f}(\omega) = \hat{f}'(\omega) = (-itf(t))^\wedge(\omega).$$

### Definition 12.7.8: Inverse Fouriertransformation

Existieren die auftretenden Integrale und ist  $f$  stetig und Fourier-transformierbar, so ist die *inverse Fouriertransformation* oder *Fourierumkehrtransformation*  $\mathcal{F}^{-1}$  von  $\hat{f}$  definiert als

$$(\mathcal{F}^{-1}\hat{f})(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \hat{f}(\omega) e^{i\omega t} d\omega.$$

### Definition 12.7.10: Faltung

Es seien  $f, g : \mathbb{R} \rightarrow \mathbb{C}$ , dann ist die *Faltung*  $f \star g$  von  $f$  und  $g$  (falls existent) definiert als

$$(f \star g)(t) := \int_{-\infty}^{+\infty} f(t-u)g(u) du.$$

### Satz 12.7.11: Faltung und Fouriertransformation

Es seien  $f, g : \mathbb{R} \rightarrow \mathbb{C}$  Fourier-transformierbar, dann gilt

- (i) Die Faltung ist kommutativ, also  $(f \star g)(t) = (g \star f)(t)$ .
- (ii) Glättungseigenschaften:
  - (a) Ist  $f$  (oder  $g$ ) stetig, so ist  $f \star g$  stetig.
  - (b) Ist  $f$  stetig differenzierbar, so ist  $f \star g$  stetig differenzierbar mit  $(f \star g)'(t) = (f' \star g)(t)$ .
- (iii) Faltungssatz im Zeitbereich: Die Faltung  $f \star g$  ist Fourier-transformierbar mit

$$(\widehat{f \star g})(\omega) = \sqrt{2\pi} \hat{f}(\omega) \cdot \hat{g}(\omega).$$

- (iv) Faltungssatz im Frequenzbereich: Ist auch die inverse Fouriertransformation auf  $f$  und  $g$  anwendbar, so gilt

$$(\widehat{f \cdot g})(\omega) = \frac{1}{\sqrt{2\pi}} (\hat{f} \star \hat{g})(\omega).$$

### Abtast- oder Sampling-Theorem 12.7.13

Es sei  $f$  stetig, fourier-transformierbar und  $\forall \omega \in \mathbb{R} \setminus [-c, c] (\hat{f}(\omega) = 0)$ , dann gilt

$$f(x) = \sum_{k=-\infty}^{\infty} f\left(\frac{k}{2c}\right) \text{sinc}(2cx - k). \quad (12.8)$$

# Elementare Differentialgleichungen

## Motivation

Differentialgleichungen spielen u. a. in der Physik eine große Rolle. Entsprechend ist es nicht verwunderlich, dass die meisten Motivationen der Physik entspringen. In einer Differentialgleichung tritt eine (in der Regel zeitabhängige) Größe und mindestens eine (in der Regel zeitliche) Ableitung der Größe auf. Das klassische Beispiel ist beim radioaktiven Zerfall, wo angenommen wird, dass die Änderungsrate der radioaktiven Kerne  $\dot{N}$  (zeitliche Änderung von  $N$ ) und die Anzahl der Kerne  $N$  proportional zueinander sind. Das bedeutet, dass man bei einer doppelt so großen Anzahl der Kerne auch doppelt so viel Zerfälle pro Zeiteinheit erwartet. Man schreibt:

$$\dot{N} = \lambda \cdot N$$

bzw. in der Schreibweise der Mathematik (nach Leibniz):

$$N' = \lambda \cdot N$$

bzw. unter Verwendung des Differentialoperators:

$$\frac{d}{dt} N = \lambda \cdot N$$

Es wäre eigentlich besser zu betonen, dass die Größe  $N$  zeitabhängig ist, d. h.  $N$  steht eigentlich für eine Funktion mit dem Term  $N(t)$ , die zum Zeitpunkt  $t$  angibt, wie viele radioaktive Kerne noch vorhanden sind. Eine Differentialgleichung lässt sich schnell aufschreiben, aber die Frage, welche analytische Funktion die Größe beschreibt, ist eine schwere und mitunter nicht analytisch zu beantwortende Frage. Man nennt die Funktion, die die Differentialgleichung genügt, die Lösung der Differentialgleichung. Im Beispiel oben, suchen wir eine Funktion, die abgeleitet der Funktion selbst entspricht mit einem zusätzlichen Faktor. Unwillkürlich würde man hier an eine  $e$ -Funktion mit dem Term  $e^{\lambda \cdot t}$  denken. Tatsächlich löst  $N(t) = e^{\lambda \cdot t}$  die Differentialgleichung. Möchte man auch noch zum Zeitpunkt  $t = 0$  einen Startwert (anfängliche Anzahl der Kerne) berücksichtigen so lässt sich das durch einen zusätzlichen Faktor  $N_0$  realisieren, womit wir schon das vollständige Zerfallsgesetz aus dem Physikunterricht gefunden haben:  $N(t) = N_0 e^{\lambda \cdot t}$

Die Schwierigkeit besteht bei dem Thema, eine solche Funktion zu finden (und sie nicht einfach zu erraten, was bei zunehmender Komplexität eh schwierig ist). Dafür gibt es bestimmte Ansätze – Standardverfahren, aber wie schon angedeutet: Nicht jede Differentialgleichung (DGL) ist lösbar. Manche Werte können nur numerisch berechnet werden und für DGLs gibt es entsprechende Verfahren.

Ebenso kann es Systeme von DGLs geben.

Man unterscheidet explizite und implizite Differentialgleichungen, wobei bei expliziten Differentialgleichungen sich dadurch auszeichnen, dass die höchste Ableitung der Größe isoliert auf einer Seite steht. Diese Differentialgleichungen heißen auch gewöhnliche Differentialgleichungen. Wenn die Gleichung nicht zu kompliziert ist, kann man oft bei expliziten Differentialgleichungen so umstellen, dass die höchste Ableitung isoliert auf einer Seite steht.

Die gängigsten Verfahren zum Lösen einer DGL sind:

- **Trennung der Variablen**
- **Integration bei linearen Differentialgleichungen erster Ordnung**
- **Substitution und Trennung der Variablen** bei homogene Differentialgleichungen
- **Rückführungen** auf lineare Differentialgleichungen z. B. **bei Bernoulli-Differentialgleichungen**

- Bei linearen Differentialgleichungen höherer Ordnung: **Lösen der homogenen DGL, Finden einer partikulären Lösung der inhomogenen Gleichung durch Variation der Konstanten** oder die Methode der unbestimmten Koeffizienten. Angabe der Gesamtlösung als Summe der homogenen Lösung und der partikulären Lösung.
- Laplace-Transformation** (Nicht im Skript! Diese Methoden decken eine Vielzahl von gewöhnlichen Differentialgleichungen ab, aber ich kenne sie nicht.)
- $e^\lambda$ -Ansatz (nicht im Skript!)**, denkt aber einige Differentialgleichungen mit  $e$ -Funktionen, Linearkombinationen von  $e$ -Funktionen und sin- uns cos-Funktionen ab.

### Definition 13.1.3: Differenzialgleichung

- (i)  $I_1, I_2, \dots, I_n \subseteq \mathbb{R}$  seien nichtleere Intervalle. Die Menge  $M = I_1 \times I_2 \times \dots \times I_n \subseteq \mathbb{R}^n$  heißt ( $n$ -dimensionales) Rechteck.
- (ii) Es seien  $M \subseteq \mathbb{R}^2$  ein Rechteck und  $\varphi : M \rightarrow \mathbb{R}$  eine stetige<sup>a</sup> Funktion. Weiter seien  $I \subseteq \mathbb{R}$  ein Intervall und  $y : I \rightarrow \mathbb{R}$  eine Funktion.  $y$  nennen wir eine *Lösung der expliziten (gewöhnlichen) Differenzialgleichung (DGL) erster Ordnung*

$$y' = \varphi(x, y),$$

wenn  $y$  stetig differenzierbar ist sowie jeweils für alle  $x \in I$  gilt, dass  $(x, y(x)) \in M$  und  $y'(x) = \varphi(x, y(x))$ .

- (iii) Es seien  $M \subseteq \mathbb{R}^2$  ein Rechteck,  $\varphi : M \rightarrow \mathbb{R}$  eine stetige Funktion und  $(\xi, \eta) \in M$ . Ist  $I \subseteq \mathbb{R}$  ein Intervall, dann heißt  $y : I \rightarrow \mathbb{R}$  eine *Lösung des Anfangswertproblems (AWP)*

$$y' = \varphi(x, y), \quad y(\xi) = \eta,$$

wenn  $y$  eine Lösung der Differenzialgleichung  $y' = \varphi(x, y)$  ist und zusätzlich  $\xi \in I$  mit  $y(\xi) = \eta$  gilt.

- (iv)  $y : I \rightarrow \mathbb{R}$ , mit einem Intervall  $I \subseteq \mathbb{R}$ , heißt eine *Lösung der expliziten Differenzialgleichung  $n$ -ter Ordnung*

$$y^{(n)} = \varphi(x, y, y', \dots, y^{(n-1)}),$$

wenn gilt

- $y$  ist  $n$ -mal stetig differenzierbar,
- $(x, y(x), y'(x), \dots, y^{(n-1)}(x)) \in M$  für alle  $x \in I$  und

Für  $(\xi, \eta_0, \eta_1, \dots, \eta_{n-1})$  heißt  $y : I \rightarrow \mathbb{R}$  eine Lösung des Anfangswertproblems

$$y^{(n)} = \varphi(x, y, y', \dots, y^{(n-1)}), \quad y^{(k)}(\xi) = \eta_k \text{ für } k = 0, \dots, n-1,$$

wenn  $y$  die Differenzialgleichung  $y^{(n)} = \varphi(x, y, y', \dots, y^{(n-1)})$  löst und  $y^{(k)}(\xi) = \eta_k$  für  $k = 0, 1, \dots, n-1$  gilt.

<sup>a</sup>Ersetzt man in der Definition von Stetigkeit die Beträge durch Normen, erhält man eine Definition für Funktionen mehrerer Variablen.

Im Skript werden DLGs stets als explizite DLGs beschrieben. Auf der linken Seite steht die (höchste) Ableitung von  $y$ , die rechte Seite wird als eine mehrdimensionale Funktion  $\varphi$  aufgefasst, in dessen Term  $x$  und  $y$  bzw. niedere Ableitungen von  $y$  auftauchen können – aber nicht müssen. Wir müssen uns hier stets denken, dass  $y$  und die Ableitungen von  $y$  von  $x$  abhängig sind.

Manchmal wird zusätzlich ein bestimmter Wert vorgegeben (AWP), z. B. ein Startwert für  $x = 0$ .

Der Formalismus kann für DGL mit höheren Ableitungen erweitert werden.

**Bemerkungen 13.1.5.** (i) Jede explizite Differentialgleichung  $n$ -ter Ordnung,  $y^{(n)} = \varphi(x, y^{(0)}, y^{(1)}, \dots, y^{(n-1)})$ , lässt sich auf ein System von gewöhnlichen Differentialgleichungen erster Ordnung zurückführen, etwa

$$\begin{aligned}(y^{(0)})' &= y^{(1)}, \\ (y^{(1)})' &= y^{(2)}, \\ &\vdots \\ (y^{(n-2)})' &= y^{(n-1)}, \\ (y^{(n-1)})' &= \varphi(x, y^{(0)}, y^{(1)}, \dots, y^{(n-1)}).\end{aligned}$$

Nicht wundern, wenn plötzlich Vektoren auftauchen. Hintergrund ist diese Bemerkung.

Schreibt man  $y = y^{(0)} = y_1$ ,  $y' = y^{(1)} = y_2$  und so weiter bis  $y_n = y^{(n-1)}$ , erhält man mit der Vektorschreibweise  $y = (y_1, \dots, y_n)^T$  das System

$$y' = \begin{pmatrix} y_2 \\ \vdots \\ y_n \\ \varphi(x, y_1, \dots, y_n). \end{pmatrix}$$

Komponentenweise handelt es sich dabei um Differentialgleichungen erster Ordnung. Für Systeme von Differentialgleichungen erster Ordnung existieren Lösungsverfahren, die mit dieser Umformung auf Differentialgleichungen  $n$ -ter Ordnung anwendbar sind.

## Lineare Differentialgleichungen erster Ordnung

Es wird zunächst eine DGL der Gestalt  $y' = f \cdot y$  betrachtet, wobei  $f$  eine stetige Funktion sein soll. Anschließend wird die DGL erweitert um eine Funktion  $g$ , d. h. wir betrachten  $y' = f \cdot y + g$

### Satz 13.2.1: Lineare Differentialgleichungen erster Ordnung

Es seien  $I \subseteq \mathbb{R}$  ein Intervall und  $\xi \in I$  kein Randpunkt von  $I$ . Es seien weiter  $f, g : I \rightarrow \mathbb{R}$  stetige Funktionen und  $\eta \in \mathbb{R}$ . Für

$$\begin{aligned}y_0 : I &\rightarrow \mathbb{R}, \quad y_0(x) := \exp\left(\int_{\xi}^x f(t) dt\right), \\ y : I &\rightarrow \mathbb{R}, \quad y(x) := \left(\eta + \int_{\xi}^x \frac{g(t)}{y_0(t)} dt\right) \cdot y_0(x)\end{aligned}$$

gilt dann

- (i)  $y_0$  löst das Anfangswertproblem  $y' = f(x)y$ ,  $y(\xi) = 1$  und
- (ii)  $y$  löst das Anfangswertproblem  $y' = f(x)y + g(x)$ ,  $y(\xi) = \eta$ .

$\exp$  steht hier für die e-Funktion, d. h. die Lösung  $y_0$  ist  $y_0(x) = e^{\int_{\xi}^x f(t) dt}$ .

Ob dieser Weg erfolgreich durchgeführt werden kann, hängt auch davon ab, wie gut wir die Funktion  $f$  integrieren können.

Die DGL (i) wird auch homogene DGL genannt und die (ii) inhomogene DGL.

Neben dem Beispiel 13.2.2 auf Seite 364 betrachten wir die (homogene) DGL:  $y' = -x^2 \cdot y$  mit  $y(1) = 5$ .

Es ist also  $f(x) = -x^2$ . Wir betrachten zunächst den Fall  $y(1) = 1$ :

$$y_0(x) = e^{\int_1^x -t^2 dt} = e^{-\frac{1}{3}t^3|_1^x} = e^{-\frac{1}{3}x^3 - (-\frac{1}{3})} = e^{\frac{1}{3}(-x^3 + 1)}$$

Probe:  $\frac{d}{dx} y_0(x) = \frac{d}{dx} e^{\frac{1}{3}(-x^3 + 1)} = -x^2 e^{\frac{1}{3}(-x^3 + 1)} = -x^2 y_0(x)$ .  $y_0$  ist Lösung der DGL.

Es gilt  $y_0(1) = 1$ . Gefordert war jedoch  $y(1) = 5$ . Wie im Skript kann durch das Hinzufügen eines Faktors diese Randbedingung eingehalten werden. Die gesuchte Lösung ist also  $y(x) = 5 \cdot e^{\frac{1}{3}(-x^3+1)}$ , da der Faktor beim Ableiten unverändert bleibt und  $y(1) = 5$  erfüllt wird.

### Betrachten wir nun inhomogene DGLs:

Das Anfangswertproblem  $y' = x + y$ ,  $y(0) = 2$ , das heißt  $f(x) = 1$ ,  $g(x) = x$ ,  $\xi = 0$  und  $\eta = 2$  lösen wir in zwei Schritten:

$$\begin{aligned} y_0(x) &= \exp\left(\int_0^x f(t) dt\right) = \exp\left(\int_0^x 1 dt\right) = e^x, \\ y(x) &= \left(2 + \int_0^x te^{-t} dt\right)e^x = \left(2 + [-te^{-t}]_0^x + \int_0^x e^{-t} dt\right)e^x \\ &= \left(2 + [-e^{-t} - te^{-t}]_0^x\right)e^x = (3 - e^{-x} - xe^{-x})e^x \\ &= 3e^x - x - 1. \end{aligned}$$

Das funktioniert auch, wenn  $f$  und  $g$  nicht so langweilige Funktionen sind wie eben:

- (iii) Das Anfangswertproblem  $y' = 3x^2y + x^2$ ,  $y(0) = 1$ , wird mit  $f(x) = 3x^2$ ,  $g(x) = x^2$ ,  $\xi = 0$ ,  $\eta = 1$  gelöst mit

$$\begin{aligned} y_0(x) &= \exp\left(\int_0^x 3t^2 dt\right) = e^{x^3}, \\ y(x) &= \left(1 + \int_0^x t^2 e^{-t^3} dt\right)e^{x^3} = \left(1 - \frac{1}{3}[e^{-t^3}]_0^x\right)e^{x^3} \\ &= \left(1 - \frac{1}{3}e^{-x^3} + \frac{1}{3}e^0\right)e^{x^3} = \frac{4}{3}e^{x^3} - \frac{1}{3}. \end{aligned}$$

Test:

$$y'(x) = \frac{4}{3}3x^2e^{x^3} = 3x^2\left(\frac{4}{3}e^{x^3} - \frac{1}{3}\right) + x^2.$$

Außerdem ist  $y(0) = \frac{4}{3} - \frac{1}{3} = 1$ .

### Lemma 13.2.4: Darstellung der Lösung

Es seien  $I \subseteq \mathbb{R}$  ein Intervall und  $\xi \in I$  mit  $(\xi - \delta, \xi + \delta) \subseteq I$  für ein  $\delta > 0$  und  $y : I \rightarrow \mathbb{R}$  eine Lösung des Anfangswertproblems  $y' = f(x)y + g(x)$ ,  $y(\xi) = \eta$ . Dann ist

$$y(x) = \exp\left(\int_{\xi}^x f(t) dt\right)\left(\eta + \int_{\xi}^x \frac{g(t)}{\exp\left(\int_{\xi}^t f(u) du\right)} dt\right).$$

### Definition 13.2.5: Bernoulli-Differenzialgleichung

Es seien  $I \subseteq \mathbb{R}$  ein Intervall,  $f, g : I \rightarrow \mathbb{R}$  stetige Funktionen und  $\alpha \in \mathbb{R}$ , dann heißt

$$y' = f(x)y + g(x)y^{\alpha}$$

eine *Bernoulli-Differenzialgleichung*.

Bei inhomogenen DGLs bestimmt man dennoch zuerst die homogene Lösung  $y_0$ !

Dann wird im zweiten Schritt die Lösung wie im Satz 13.2.1 bestimmt.

Das Lemma ist mehr oder weniger die Wiederholung des vorigen Satzes.

Hier wird eine verallgemeinerte Form der vorigen DGL angesprochen, wobei die Fälle  $\alpha = 0$  und  $\alpha = 1$  gerade den behandelten Differentialgleichungstypen entspricht:

$\alpha = 0$ : inhomogene DGL der Form:  $y' = f(x) \cdot y + g(x)$

$\alpha = 1$ : homogene DGL der Form:  $y' = f(x) \cdot y + g(x) \cdot y = \left(\underbrace{f(x) + g(x)}_{=\tilde{f}(x)}\right) \cdot y = \tilde{f}(x) \cdot y$

Häufig wird die Bernoulli-DGL auch so aufgeschrieben:  $y' + f(x) \cdot y = g(x)y^\alpha$ , wobei sich hier  $f(x)$  und ein Vorzeichen von dem  $f(x)$  aus dem Skript unterscheidet.

**Lösungsansatz im Fall  $\alpha \notin \{0, 1\}$**  **13.2.7.** Wir versuchen, auch diesen Fall auf eine lineare Differenzialgleichung zurückzuführen, dabei gehen wir zunächst davon aus, dass alle Rechenschritte erlaubt sind und rechtfertigen das Vorgehen später. Wir haben

$$\begin{aligned} y' &= f(x)y + g(x)y^\alpha \Rightarrow \frac{y'}{y^\alpha} = f(x)\frac{y}{y^\alpha} + g(x) \\ &\Leftrightarrow y^{-\alpha}y' = f(x)y^{1-\alpha} + g(x). \end{aligned}$$

Setzen wir  $z = y^{1-\alpha}$ , so ist  $z' = (1 - \alpha)y^{-\alpha}y'$ , also

$$\frac{z'}{1 - \alpha} = f(x)z + g(x) \Leftrightarrow z' = (1 - \alpha)f(x)z + (1 - \alpha)g(x).$$

Also erhalten wir die gesuchte Lösung  $y$  der Bernoullischen Differenzialgleichung, wenn wir zunächst die inhomogene lineare Differenzialgleichung

$$z' = (1 - \alpha)f(x)z + (1 - \alpha)g(x)$$

lösen und danach  $y = z^{\frac{1}{1-\alpha}}$  setzen.

Gilt  $y > 0$ , so ist die Division durch  $y^\alpha$  erlaubt und die Herleitung damit grob gesprochen in Ordnung. Genauer gilt: Wenn  $y : I \rightarrow (0, +\infty)$  eine Lösung von  $y' = f(x)y + g(x)y^\alpha$  ist, dann gilt für

$$\begin{aligned} z : I &\rightarrow (0, +\infty), z(x) = y^{1-\alpha}(x), \\ z'(x) &= (1 - \alpha)y^{-\alpha}(x)y'(x) = (1 - \alpha)y^{-\alpha}(x)(f(x)y(x) + g(x)y^\alpha(x)) \\ &= (1 - \alpha)f(x)y^{1-\alpha}(x) + (1 - \alpha)g(x) \\ &= (1 - \alpha)f(x)z(x) + (1 - \alpha)g(x), \end{aligned}$$

das heißt  $z$  genügt einer linearen Differenzialgleichung.

Ich habe jetzt erfolglos zwei Beispielaufgaben mir ausgedacht, um sie hier als weiteres Beispiel auszuführen. Zum einen bin ich mir unsicher, wie man vorgeht, wenn keine Randbedingungen vorgegeben sind und wie man dann eine allgemeine Lösung angibt. Zum anderen bin ich mit den Bemühungen die DGL  $z' = (1 - \alpha)f(x)z + (1 - \alpha)g(x)$  zu lösen, unzufrieden und scheine mich dabei zu vertun. Ich muss mich mit dem Thema näher beschäftigen und weiß gerade nicht, ob es klausurrelevant ist. Ein Problem ist, dass wenn man sich irgendwelche DGLs ausdenkt, dass die Integrale recht anspruchsvoll sein können.

Diese DGLs müssten auch ohne Randbedingungen lösbar sein. Die Formeln im Skript sind immer mit Randbedingungen. Es gibt auch noch Techniken wie Variation der Konstanten usw. --- Irgendwie bin ich noch unzufrieden und erinnere mich schwach an „mehr“. Siehe Exkurs unten.

Hier wird beschrieben, wie man eine **Bernoulli-DGL** höherer Ordnung auf eine lineare DGL zurückführt.

In Kurzform:

Setze  $z = y^{1-\alpha}$

Löse:

$$z' = (1 - \alpha)f(x)z + (1 - \alpha)g(x)$$

Setze  $y = z^{\frac{1}{1-\alpha}}$

**Beispiele 13.2.8.** (i)  $y' = \sqrt{y}$ ,  $y(1) = 4$ , also ist  $y' = f(x)y + g(x)y^\alpha$  mit  $f(x) = 0$ ,  $g(x) = 1$ ,  $\alpha = \frac{1}{2}$ . Wir setzen  $z = \sqrt{y} = y^{1-\frac{1}{2}}$ , dann muss

$$z' = \frac{1}{2}y^{-\frac{1}{2}}y' = \frac{1}{2\sqrt{y}}y' = \frac{1}{2}, \quad z(1) = \sqrt{y(1)} = 2$$

gelten. Damit ist

$$z(x) = 2 + \int_1^x \frac{1}{2}dt = 2 + \frac{1}{2}(x-1) = \frac{x}{2} + \frac{3}{2} = \frac{x+3}{2}$$

und folglich  $y = \left(\frac{x+3}{2}\right)^2$ .

Schematisch war das Vorgehen also: Setze  $z = y^{1-\alpha}$ ,  $z(1) = y^{1-\alpha}(1)$ , löse  $z' = (1-\alpha)f(x)z + (1-\alpha)g(x) = \frac{1}{2}$  mit  $z(1) = \sqrt{y(1)}$  und setze anschließend  $y = z^{\frac{1}{1-\alpha}} = z^2$ .

(ii)  $y' = x^4y + x^4y^4$ ,  $y(0) = 1$ , das heißt  $f(x) = x^4 = g(x)$ ,  $\alpha = 4$ . Ansatz:  $z = y^{1-4} = \frac{1}{y^3}$ . Damit ist

$$z' = \frac{-3}{y^4}y' = -\frac{3}{y^4}(x^4y + x^4y^4) = -3x^4y^{-3} - 3x^4 = -3x^4z - 3x^4$$

und  $z(0) = 1$ . Die Lösung  $z$  ist bestimmt durch

$$\begin{aligned} z(x) &= \underbrace{\exp\left(\int_0^x (-3t^4) dt\right)}_{=\exp(-\frac{3}{5}x^5)} \left(1 - \int_0^x 3t^4 e^{\frac{3}{5}t^5} dt\right) \\ &= e^{-\frac{3}{5}x^5} \left(1 - \left[e^{\frac{3}{5}t^5}\right]_0^x\right) = e^{-\frac{3}{5}x^5} \left(1 - e^{\frac{3}{5}x^5} + e^0\right) \\ &= 2e^{-\frac{3}{5}x^5} - 1. \end{aligned}$$

$$\text{Damit haben wir } y = z^{\frac{1}{1-4}} = \left(2e^{-\frac{3}{5}x^5} - 1\right)^{-\frac{1}{3}}.$$

Hier zwei Beispiele aus dem Skript.

## Lineare Differentialgleichungen erster Ordnung - Exkurs

Wie oben angegeben bin ich etwas unzufrieden damit, dass im Skript nur DGL mit Randbedingungen betrachtet wurden, daher hier ein Ansatz, wie er mir vertrauter ist.

### Allgemeines

Eine lineare Differentialgleichung erster Ordnung lässt sich in der folgenden Form darstellen:

$$y' = f(t) y + g(t)$$

Darin ist  $y' = f(t) y$  die homogene DGL,  $g(t)$  die Inhomogenität.

Superpositionsprinzip: Die Lösung einer inhomogenen DGL setzt sich zusammen aus der allgemeinen Lösung  $y_H$  der homogenen DGL und einer speziellen (partikulären) Lösung  $y_P$  der inhomogenen DGL (siehe unten).

Typische Lösungsverfahren für diese Differentialgleichung sind:

- Separation der Variablen (siehe nächstes Kapitel),
- Exponentialansatz ( $y = Ce^{\lambda t}$ ),
- Potenzreihenansatz (siehe z. B. [https://www.zib.de/userpage/weber/Mathematik1\\_12.pdf](https://www.zib.de/userpage/weber/Mathematik1_12.pdf))

### Klassifikation und Lösungsverfahren

- Lineare homogene DGL erster Ordnung **mit konstanten Koeffizienten**:  
 $y' = f y, f = \text{konstant}$   
Die Lösung sollte leicht durch einen Ansatz mit einer Exponentialfunktion  $y = Ce^{\lambda t}$  finden lassen, denn es gilt  $y' = \lambda Ce^{\lambda t} = \lambda y$ , was der DGL mit  $f = \lambda$  entspricht. Auf die Konstante  $C$  kann bei dieser einfachen DLG auch verzichtet bzw. als  $C = 1$  betrachtet.
- Lineare homogene DGL erster Ordnung **mit konstanten Koeffizienten und konstanten Summanden**:  
 $y' = f y + g, f, g = \text{konstant}$   
Diese inhomogene DGL lässt sich noch durch Trennung der Variablen lösen (siehe nächstes Kapitel). Ansonsten kann man auch sie wie beim allgemeinen Fall lösen:
- Lineare homogene DGL erster Ordnung **mit variablen Koeffizienten**:

$$y' = f(t) y$$

Egal welchen Ansatz man hier verfolgt, die Lösung ist:

$$y_H = c \cdot \exp\left(\int f(t) dt\right)$$

- Lineare **inhomogene** DGL erster Ordnung **mit variablen Koeffizienten**:  
 $y' = f(t)y + g(t)$   
Die Gesamtlösung ist  $y = y_H + y_P$  (Superpositionsprinzip), wobei  $y_H$  für die Lösung der homogenen DLG steht und  $y_P$  eine partikuläre Lösung ist. Die Superposition gilt allgemein, aber bei dieser Art von DGL gibt es folgende Idee: Variation der Konstanten: Dies wird hier gut erklärt: <https://studyflix.de/mathematik/variation-der-konstanten-938>. Damit findet man als allgemeine Lösungsformel:

$$y = \left( \int \frac{g(t)}{y_H} dt + C \right) \cdot y_H = \left( \int g(t) \exp\left(-\int f(t) dt\right) dt + C \right) \cdot \exp\left(\int f(t) dt\right)$$

Das ist zwar doch sehr ähnlich zu dem aus dem Skript, aber ich finde es so übersichtlicher. Zur Übung stützen wir uns auf eine Bernoulli-DGL:

**Beispiel:**  $y' = f(x) \cdot y + g(x)y^\alpha$  (Bernoulli-DGL)

Wir suchen die Lösung von  $y' = \frac{1}{x}y - x^2y^2$  mit  $f(x) = \frac{1}{x}$  und  $g(x) = -x^2$ ,  $\alpha = 2$ .

Zunächst wandeln wir die DGL in eine homogene DGL um:

Setze  $z = y^{1-\alpha} = y^{-1}$ . Es ist  $1 - \alpha = -1$ .

Löse:

$$z' = (1 - \alpha)f(x)z + (1 - \alpha)g(x)$$

$$z' = -f(x)z - g(x) = -\frac{1}{x}z + x^2$$

Wir verwenden direkt die Lösungsformel, achten aber darauf, dass wir dabei den **homogenen Teil** und die **Inhomogenität** der z-DGL richtig erfassen und dass in dem Beispiel die Variable  $x$  und nicht  $t$  ist.

$$\begin{aligned} z &= \left( \int g(x) \exp\left(-\int f(x) dx\right) dx + C \right) \cdot \exp\left(\int f(x) dx\right) \\ &= \left( \int x^2 \exp\left(-\int -\frac{1}{x} dx\right) dx + C \right) \cdot \exp\left(\int -\frac{1}{x} dx\right) \\ &= \left( \int x^2 \exp(\ln x) dx + C \right) \cdot \exp(-\ln x) \\ &= \left( \int x^2 x dx + C \right) \cdot \frac{1}{x} = \left( \int x^3 dx + C \right) \cdot \frac{1}{x} = \left( \frac{1}{4}x^4 + C \right) \cdot \frac{1}{x} = \frac{1}{4}x^3 + \frac{C}{x} \end{aligned}$$

Rücksubstitution: Setze  $y = z^{\frac{1}{1-\alpha}} = z^{-1} = \left(\frac{1}{4}x^3 + \frac{C}{x}\right)^{-1}$

Probe: (Links ableiten, rechts einsetzen)

$$\begin{aligned} y' &= \frac{1}{x}y - x^2y^2 \\ -1\left(\frac{3}{4}x^2 - \frac{C}{x^2}\right)\left(\frac{1}{4}x^3 + \frac{C}{x}\right)^{-2} &= \frac{1}{x}\left(\frac{1}{4}x^3 + \frac{C}{x}\right)^{-1} - x^2\left(\frac{1}{4}x^3 + \frac{C}{x}\right)^{-2} \\ -\left(\frac{3}{4}x^2 - \frac{C}{x^2}\right) &= \frac{1}{x}\left(\frac{1}{4}x^3 + \frac{C}{x}\right)^1 - x^2 \\ -\frac{3}{4}x^2 + \frac{C}{x^2} &= \frac{1}{4}x^2 + \frac{C}{x^2} - x^2 \\ -\frac{3}{4}x^2 &= \frac{1}{4}x^2 - x^2 \\ 0 &= x^2 - x^2 \\ 0 &= 0 \end{aligned}$$

$y = \left(\frac{1}{4}x^3 + \frac{C}{x}\right)^{-1}$  erfüllt die DGL.

Weitere Links zum Vertiefen und erweitern:

- <https://www.mathe-seite.de/oberstufe/analysis-hoehere-mathematik/differentialgleichung-dgl/dgl-partikulaere-loesung/>
- <https://studyflix.de/mathematik/intro-gewohnliche-differentialgleichungen-losen-936>
- <https://studyflix.de/mathematik/bernoulli-dgl-940>
- Die Seite <https://de.fufaev.org/lineare-dgl> ist auch ganz gut.
- <https://statmath.wu.ac.at/~leybold/MOK/HTML/node175.html> Einfach DLG erster und zweiter Ordnung samt Lösungsverfahren. Kompakt und gut.

## Getrennte Variablen

### Definition 13.3.1: Differenzialgleichung mit getrennten Variablen

Es seien  $I, J \subseteq \mathbb{R}$  Intervalle und  $f : I \rightarrow \mathbb{R}$ ,  $g : J \rightarrow \mathbb{R}$  stetige Funktionen, dann heißt eine Differenzialgleichung der Form

$$y' = f(x)g(y)$$

*Differenzialgleichung mit getrennten Variablen* und

$$y' = f(x)g(y), \quad y(\xi) = \eta \text{ für } \xi \in I, \eta \in J$$

das zugehörige Anfangswertproblem.

Die linearen homogenen DGL sind ein Sonderfall dieser DGL mit  $g(y) = y$ .

Hier wird nun versucht solche DGL zu lösen:

### Lösungsansatz für das AWP mit getrennten Variablen 13.3.2.

- (i) Wenn  $g(\eta) = 0$ , dann ist  $y : I \rightarrow \mathbb{R}$ ,  $y(x) = \eta$  eine Lösung des Anfangswertproblems.
- (ii) Ist  $g(y) \neq 0$  für alle  $y \in J$  (dies kann im Fall  $g(\eta) \neq 0$  aufgrund der Stetigkeit von  $g$  durch eine geeignete Wahl von  $J$  stets erreicht werden) und  $y : I \rightarrow \mathbb{R}$  eine Lösung des Anfangswertproblems, also insbesondere  $y(x) \in J$  für alle  $x \in I$ , dann folgt für alle  $x \in I$

$$\frac{y'(x)}{g(y(x))} = f(x). \quad (13.3)$$

Es seien nun  $F : I \rightarrow \mathbb{R}$ ,  $F(x) = \int_{\xi}^x f(t) dt$  und  $G : J \rightarrow \mathbb{R}$ ,  $G(y) = \int_{\eta}^y \frac{1}{g(u)} du$ , dann folgt aus (13.3)

$$\begin{aligned} \int_{\xi}^x \frac{y'(t)}{g(y(t))} dt &= \int_{\xi}^x f(t) dt \\ \Leftrightarrow \int_{y(\xi)}^{y(x)} \frac{1}{g(u)} du &= \int_{\xi}^x f(t) dt \Leftrightarrow G(y(x)) = F(x), \end{aligned}$$

Zuerst ein kleiner Sonderfall für ein AWP.

Hier wird begründet, warum das Verfahren „Separation der Variablen“ funktioniert. Es wird für ein allgemeines AWP diskutiert.

wobei  $u = y(t)$ ,  $\frac{du}{dt} = y'(t)$  substituiert wurde. Aus dieser Gleichung erhält man dann die Lösung des Anfangswertproblems durch Auflösen nach  $y(x)$ .

Liegt nur eine Differenzialgleichung ohne Anfangswertproblem vor, so erhält man die Lösungen, in dem man die Integrale durch unbestimmte Integrale (Stammfunktionen) ersetzt und eine Konstante  $c \in \mathbb{R}$  addiert:

$$\int \frac{1}{g(u)} du = \int f(t) dt + c$$

und dann auflöst.

Unter der Voraussetzung  $g(y) \neq 0$  für alle  $y \in J$  ist die Funktion  $G : J \rightarrow \{g(y) \mid y \in J\}$  stets bijektiv nach Lemma 11.2.11 und daher folgt  $y(x) = G^{-1}(F(x))$ .

Hier steht die zentrale Ansatzzeile für das Verfahren. Siehe „Exkurs unten“. Ich finde die Darstellung im Skript nicht so übersichtlich, auch wenn das sicher alles richtig und sauberer ist, als mein Exkurs unten.

**Beispiele 13.3.3.** (i)  $y' = x(1+y)$ ,  $y(0) = 0$ , das heißt  $f(x) = x$ ,  $g(y) = 1+y$ . Es gilt

$$\begin{aligned} F(x) &= \int_0^x t dt = \frac{1}{2}x^2, \\ G(y) &= \int_0^y \frac{1}{1+t} dt = [\ln(1+t)]_0^y = \ln(1+y). \\ \Rightarrow \ln(1+y) &= \frac{1}{2}x^2 \\ \Rightarrow 1+y &= e^{\frac{1}{2}x^2} \Rightarrow y = e^{\frac{1}{2}x^2} - 1. \end{aligned}$$

### Separation der Variablen- Exkurs

Auch wenn das oben sicher alles mathematisch gesehen richtig und sauber ist, möchte ich diese Methode nochmal etwas handlungsorientierter beschreiben:

Ausgangspunkt ist eine DGL der Gestalt:

$$y' = f(x)g(y)$$

Dabei betrachten wir  $y$  als eine Größe, die von  $x$  abhängt, d. h. wir können für die Ableitung schreiben:

$$y' = \frac{dy}{dx}$$

Die DGL lässt sich also schreiben als:

$$\frac{dy}{dx} = f(x)g(y)$$

Nun kommt ein etwas haariger Schritt. Wir betrachten  $\frac{dy}{dx} = \frac{d}{dx} y$  nicht als Operator  $\frac{d}{dx}$ , der  $y$  ableitet, sondern als Quotient infinitesimaler Größen, so wie u. a. auch Euler gedacht hat. Mit diesen Größen kann man rechnen, z. B. die Gleichung mit  $dx$  multiplizieren:

$$dy = f(x)g(y)dx$$

$$\frac{1}{g(y)}dy = f(x)dx$$

Die linke und rechte Seite kann integriert werden (unendliche Summe unendlich kleiner Produkte):

$$\int \frac{1}{g(y)}dy = \int f(x)dx + c$$

Das ist die zentrale Ansatzformel für die Methode Separation der Variablen. Bleibt nur zu hoffen, das die Integrale lösbar sind.

**Beispiel 1: linearen homogenen DGL, d. h.  $g(y) = y$**

$$y' = f(x)y$$

$$\int \frac{1}{y}dy = \int f(x)dx + c$$

$$\ln y = \int f(x)dx + c = F(x) + c$$

$$y = \exp(F(x) + c)$$

Probe:  $y' = f(x) \cdot \exp(F(x) + c) = f(x)y - \text{passt.}$

Beispiel 2:  $y' = x^2 \sqrt{y} = f(x)g(y)$

Ansatz:

$$\int \frac{1}{\sqrt{y}}dy = \int x^2 dx + c$$

$$2\sqrt{y} = \frac{1}{3}x^3 + c$$

$$\sqrt{y} = \frac{1}{6}x^3 + \frac{c}{2}$$

$$y = \left(\frac{1}{6}x^3 + \frac{c}{2}\right)^2$$

Probe:

$$y' = 2\left(\frac{1}{2}x^2\right)\left(\frac{1}{6}x^3 + \frac{c}{2}\right)^1 = x^2\left(\frac{1}{6}x^3 + \frac{c}{2}\right) = x^2\sqrt{y}$$

Passt.

Wie passt man diese allgemeinen Lösungen auf AWP an? Durch die Konstante!

Angenommen das AWP lautet  $y' = x^2\sqrt{y} = f(x)g(y)$  und  $y(5) = 2$ .

Wir setzen  $x=5$  in die Lösung ein und fordern  $y(5)=2$ . Dann lösen wir nach  $c$  auf:

$$y(5) = \left(\frac{1}{6}5^3 + \frac{c}{2}\right)^2 = 2$$

$$\left(\frac{125}{6} + \frac{c}{2}\right)^2 = 2$$

$$\frac{125}{6} + \frac{c}{2} = \sqrt{2}$$

$$c = 2\sqrt{2} - \frac{125}{6}$$

### $e^\lambda$ -Ansatzes - Exkurs

Der  $e^\lambda$ -Ansatz ist eine Methode zur Lösung von linearen homogenen Differentialgleichungen mit konstanten Koeffizienten.

Betrachten wir eine lineare homogene Differentialgleichung n-ter Ordnung mit konstanten Koeffizienten:  $a_n y^{(n)} + a_{n-1} y^{(n-1)} + \dots + a_0 y = 0$

#### 1. Ansatz $y = e^{\lambda x}$

Wir nehmen an, dass die Lösung der Gleichung die Form  $y = e^{\lambda x}$  hat.

#### 2. Einsetzen des Ansatzes in die Differentialgleichung

Leiten wir den Ansatz mehrfach ab und setzen ihn in die Differentialgleichung ein:

- $y = e^{\lambda x}$
- $y' = \lambda e^{\lambda x}$
- $y'' = \lambda^2 e^{\lambda x}$
- ...
- $y^{(n)} = \lambda^n e^{\lambda x}$

Einsetzen dieser Ableitungen in die Differentialgleichung ergibt:

$$\lambda^n a_n e^{\lambda x} + \lambda^{n-1} a_{n-1} e^{\lambda x} + \dots + a_0 e^{\lambda x} = 0$$

#### 3. Charakteristische Gleichung

Die obige Gleichung kann man gefahrlos durch  $e^{\lambda x}$  teilen, da  $e^{\lambda x} \neq 0$  gilt. Dies führt uns zur sogenannten charakteristischen Gleichung:

$$\lambda^n a_n + \lambda^{n-1} a_{n-1} + \dots + a_0 = 0$$

#### 4. Lösung der charakteristischen Gleichung

Die Lösungen der charakteristischen Gleichung sind die Werte von  $\lambda$ , die die Gleichung erfüllen. Diese Werte,  $\lambda_1, \lambda_2, \dots, \lambda_n$  können reell oder komplex sein.

## 5. Allgemeine Lösung der Differentialgleichung

Je nach den Werten der Wurzeln der charakteristischen Gleichung können wir die allgemeine Lösung der Differentialgleichung formulieren. Dieses hängt jedoch stark von den Wurzeln der Gleichung ab. Im einfachsten Fall haben wir nur **einfache reelle Wurzeln**: Wenn die charakteristische Gleichung  $n$  verschiedene reelle Wurzeln  $\lambda_1, \lambda_2, \dots, \lambda_n$  hat, ist die allgemeine Lösung:

$$y(x) = C_1 e^{\lambda_1 x} + C_2 e^{\lambda_2 x} + \cdots + C_n e^{\lambda_n x}$$

Bei **mehrfachen reellen Wurzeln** treten zusätzlich Potenzen von  $x$  auf und bei **komplexe Wurzeln**, d. h. bei  $\lambda = \alpha \pm i\beta$  kann die Lösung in Form von Sinus- und Kosinusfunktionen ausgedrückt werden, z. B.  $y(x) = e^{\alpha x}(C_1 \cos(\beta x) + C_2 \sin(\beta x))$ . Im Studium wird häufig der Fall  $n = 2$  betrachtet.

Weitere ausführliche Betrachtungen dazu: <https://statmath.wu.ac.at/~ley-dold/MOK/HTML/node185.html>

## Klausuraufgaben

1. Es sei  $V$  ein  $\mathbb{C}$ -Vektorraum.

- (a) Was ist ein Skalarprodukt auf  $V$  (erläutern Sie die Eigenschaften explizit)? (6)
- (b) Es sei nun  $V = \mathbb{C}^n$  mit der kanonischen Basis und  $A \in M(n \times n, \mathbb{C})$ . Ist durch  $\langle x, y \rangle := \sum_{i,j=1}^n a_{ij} x_i \bar{y}_j = x^T A \bar{y}$  immer ein Skalarprodukt auf  $V$  gegeben (Begründung)? (2)
- (c) Es seien  $V$  und  $A$  wie in Teilaufgabe (b). Welche Eigenschaften von  $A$  sind notwendig bzw. hinreichend (Beweis/Herleitung), damit durch  $\langle x, y \rangle := \sum_{i,j=1}^n a_{ij} x_i \bar{y}_j = x^T A \bar{y}$  ein Skalarprodukt auf  $V$  gegeben ist? (9)
- (d) Zeigen Sie, dass der  $\mathbb{C}^n$  zusammen mit  $d : \mathbb{C}^n \times \mathbb{C}^n$ ,  $d(x, y) := \sum_{i=1}^n |x_i - y_i|$  ein metrischer Raum ist. (4)
- (e) Wird die Metrik aus Teilaufgabe (d) durch eine Norm induziert? Wenn ja, geben Sie die Norm an. (2)

1. Es sei  $V$  ein  $\mathbb{K}$ -Vektorraum.

- (a) Was ist eine Linearform auf  $V$ ? (1)
- (b) Es sei  $\langle \cdot, \cdot \rangle$  ein Skalarprodukt auf  $V$ . Geben Sie die durch das Skalarprodukt induzierte Norm  $\|\cdot\|$  an. (1)
- (c) Zeigen Sie, dass die induzierte Norm die Parallelogramm-Identität erfüllt, das heißt, für alle  $x, y \in V$  gilt (3)

$$\|x + y\|^2 + \|x - y\|^2 = 2(\|x\|^2 + \|y\|^2).$$

- (d) Es sei  $C(I)$  die Menge der stetigen Funktionen auf einem Intervall  $I = [a, b] \subset \mathbb{R}$ . Begründen Sie, warum für  $f, g \in C(I)$  durch  $\langle f, g \rangle := \int_a^b f(x)g(x) dx$  ein Skalarprodukt definiert ist. (7)  
*Hinweis:* Dass  $C(I)$  ein reeller Vektorraum ist, muss nicht gezeigt werden.

- (e) Was ist ein metrischer Raum? (4)

2. (a) Es sei  $A \in M(n \times n, \mathbb{R})$ . Angenommen, das charakteristische Polynom von  $A$  hat die Form

$$P_A(\lambda) = \lambda^4 + 12\lambda^3 - 4\lambda^2 + 2\lambda.$$

(1) Geben Sie  $n$  an. (1)

(2) Ist  $A$  invertierbar (mit Begründung)? (2)

(b) Ist  $A = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \\ 2 & 1 & 3 \end{pmatrix}$  invertierbar? Bestimmen Sie ggf.  $A^{-1}$ . (7)

(c) Es sei  $R(x) = \frac{x^8 - 8x^6 + 15x^4 + 7x^2 - 15}{(x^2 - 4)^2(x+1)(x^2 + 1)}$ .

Geben Sie  $\int R(x) dx$  an, wobei die Koeffizienten einer dafür eventuell benötigten Partialbruchzerlegung nicht bestimmt werden müssen. (11)

Hinweise:

- Gilt etwa  $R(x) = P(x) + \frac{A}{x+1} + \frac{B}{(x+1)^2}$  mit einem Polynom  $P$  und Konstanten  $A, B \in \mathbb{R}$ , dann ist die (explizite) Bestimmung von  $\int P(x) dx + A \int \frac{dx}{x+1} + B \int \frac{dx}{(x+1)^2}$  gefordert.
- $(x^2 - 4)^2(x+1)(x^2 + 1) = x^7 + x^6 - 7x^5 - 7x^4 + 8x^3 + 8x^2 + 16x + 16$ .
- $\int \frac{dx}{x^2 + bx + c} = \frac{2}{\sqrt{4c - b^2}} \arctan \frac{2x + b}{\sqrt{4c - b^2}}$  für  $4c - b^2 > 0$ .

2. (a) Gegeben sei die rationale Funktion  $R(x) = \frac{1}{(x^2 - a^2)^2}$  mit  $a \neq 0$ . Leiten Sie ein lineares Gleichungssystem zur Bestimmung der Koeffizienten der Partialbruchdarstellung von  $R$  her. (7)

(b) Es sei  $A \in M(m \times n, \mathbb{R})$ ,  $b \in \mathbb{R}^m$  und  $x \in \mathbb{R}^n$ . Geben Sie Kriterien für die Lösbarkeit und die eindeutige Lösbarkeit des Gleichungssystems  $Ax = b$  an. (2)

(c) Es seien  $A$ ,  $b$  und  $x$  wie in Teilaufgabe (b) und  $v \in \mathbb{R}^n$  eine Lösung von  $Ax = b$ . Zeigen Sie, dass für  $w \in \text{Ker } A$  auch  $v + w$  eine Lösung von  $Ax = b$  ist. (2)

(d) Aufgrund welchen Satzes ist die eindeutige Lösbarkeit des Gleichungssystems aus Teilaufgabe (a) gegeben? (1)

3. (a) Es sei  $A \in M(n \times n, \mathbb{R})$ .

(1) Wie hängen das charakteristische Polynom von  $A$  und die Eigenwerte von  $A$  zusammen? (1)

(2) Was ist ein Eigenvektor von  $A$ ? (2)

(3) Warum existiert zu jedem Eigenwert von  $A$  mindestens ein Eigenvektor? (3)

(b) Es sei  $V$  ein endlich-dimensionaler Vektorraum mit dem Skalarprodukt  $\langle \cdot, \cdot \rangle$ . Was ist eine Orthonormalbasis auf  $V$ ? (2)

**3.** (a) Es sei  $V$  ein  $\mathbb{K}$ -Vektorraum und  $F : V \rightarrow V$  linear. Was ist ein Eigenwert von  $F$ ? (2)

(b) Es sei nun  $A \in M(n \times n, \mathbb{R})$ .

(1) Erläutern Sie, was es heißt, dass  $A$  diagonalisierbar ist. (1)

(2) Was können Sie über die Transformationsmatrix aussagen, die  $A$  diagonalisiert? (1)

(c) Ist  $\lambda = 2$  ein Eigenwert der Matrix  $A = \begin{pmatrix} 3 & 2 & 1 \\ 2 & 2 & 2 \\ -1 & 2 & 0 \end{pmatrix}$ ? Wenn ja, geben Sie einen Eigenvektor zum Eigenwert  $\lambda$  an. (5)

(d) Zeigen Sie, dass für jedes  $\lambda \in \mathbb{R}$  die Funktion  $f(x) = e^{\lambda x}$  ein Eigenvektor der linearen Abbildung  $D : C^\infty(\mathbb{R}) \rightarrow C^\infty(\mathbb{R})$ ,  $f \mapsto f'$  ist. (2)

**4.** (a) Zeigen Sie, dass die Funktion

$$f(x) = 1 - \sqrt{|x|} + (x^2 + x) \left| x - \lfloor x \rfloor - \frac{1}{2} \right|$$

auf dem Intervall  $[0, 2]$  den Wert  $\frac{1}{2}$  annimmt. (5)

*Erinnerung:*  $\lfloor x \rfloor = \max\{z \in \mathbb{Z} \mid z \leq x\}$  (Abrunden).

(b) Wann heißt  $f : [a, b] \rightarrow \mathbb{R}$  streng monoton wachsend? (1)

(c) Geben Sie eine streng monoton wachsende Funktion  $f : [0, 1] \rightarrow \mathbb{R}$  mit  $f(0) < \frac{1}{2}$ ,  $f(1) > \frac{1}{2}$  an, die den Wert  $\frac{1}{2}$  nicht annimmt (mit Nachweis). (2)

(d) Es sei  $f : \mathbb{R} \rightarrow \mathbb{R}$  mit  $f(x) = \begin{cases} x^3 + ax^2 + bx + c, & x < 1, \\ x^2 - 2x + 2, & x \geq 1. \end{cases}$

Lassen sich  $a, b, c \in \mathbb{R}$  so wählen, dass  $f$  auf  $\mathbb{R}$  stetig differenzierbar ist? Bestimmen Sie ggf. die Menge aller Tripel  $(a, b, c)$  für die  $f$  stetig differenzierbar ist. (8)

(e) Welche (algebraische) Struktur hat die Lösungsmenge aus Teilaufgabe (d)? (1)

*Hinweis:* Schreiben Sie „ $\emptyset$ “, falls Sie in (d) das als Lösung hatten.

**4.** (a) Es sei  $f : D \subseteq \mathbb{R} \rightarrow \mathbb{R}$ . Wann heißt  $f$  auf  $D$  stetig? (2)

(b) Es seien  $a, b \in \mathbb{R}$  mit  $a < b$  und  $f : (a, b] \rightarrow \mathbb{R}$  stetig. Was bedeutet die Sprechweise „ $f$  ist stetig in  $a$  fortsetzbar“? (2)

(c) Geben Sie eine Funktion an, die in  $x_0 = 2$  definiert, aber nicht stetig ist (mit Beweis). (2)

(d) Untersuchen Sie die Funktion  $f : \mathbb{R} \rightarrow \mathbb{R}$ ,  $f(x) = \begin{cases} x - 1, & x > 1 \\ \sin(\pi x), & x \leq 1 \end{cases}$  auf Stetigkeit und Differenzierbarkeit im Punkt  $x_0 = 1$ . (6)

(e) Untersuchen Sie die Funktion  $f$  aus Teilaufgabe (d) auf Stetigkeit und Differenzierbarkeit für alle Punkte  $x_0 \neq 1$ . (3)

5. (a) Es seien  $[a, b] = I \subset \mathbb{R}$  ein Intervall und  $f : I \rightarrow \mathbb{R}$ . Wann heißt  $f$  auf  $I$  Riemann-integrierbar? (2)

(b) Bei welchen Integralen spricht man von Konvergenz (des Integrals) und was bedeutet das? (3)

*Hinweis:* Den zweiten Teil können Sie an einem abstrakten oder konkreten Beispiel erläutern. Sie müssen nicht alle möglichen Fälle angeben.

(c) Zeigen Sie, dass gilt (9)

$$\int \frac{2 \sin^2(x)}{\cos^3(x)} dx = \frac{\sin(x)}{\cos^2(x)} - \ln \tan\left(\frac{\pi}{4} + \frac{x}{2}\right).$$

*Erinnerung:* Additionstheoreme:

$$\sin(\alpha \pm \beta) = \sin \alpha \cos \beta \pm \sin \beta \cos \alpha \text{ und } \cos(\alpha \pm \beta) = \cos \alpha \cos \beta \mp \sin \alpha \sin \beta.$$

5. (a) Zeigen Sie, dass für  $a \neq 0$  gilt  $\int \frac{x}{\sin^2(ax)} dx = -\frac{x}{a} \cot(ax) + \frac{1}{a^2} \ln \sin(ax)$  und geben Sie eine weitere, andere Stammfunktion von  $\frac{x}{\sin^2(ax)}$  an. (4)

(b) Zeigen Sie mit Hilfe der  $\varepsilon$ - $\delta$ -Definition der Stetigkeit, dass für eine Riemann-integrierbare Funktion  $f : [a, b] \rightarrow \mathbb{R}$ , die Funktion  $F(x) = \int_a^x f(t) dt$  auf  $[a, b]$  stetig ist. (7)

*Hinweis:* Wenn Sie die Definition nicht kennen, erfragen Sie diese bei der Aufsicht. Sie erhalten dann aber keine Punkte auf Aufgabe 4(a).

(c) Geben Sie eine andere Begründung für die Stetigkeit von  $F$  aus Teilaufgabe (b) an. (2)