

LK Mathematik Abitur 2005

B. Waldmüller,
Jens B., Magnus D., Marc F., Christian F., Martin F.,
Sonja G., Anton G., Matthias G., Paul H., Helge J.,
Sebastian K., Valeria K., Jan-Willem K.,
Dennis K., Sebastian L. Jochen V., Peter W., Boris W.
Söderblom-Gymnasium Espelkamp

20. Mai 2005

Inhaltsverzeichnis

I	Analysis	7
1	Änderungen und Änderungsrate	9
1.1	Die Fundamentalskizze	9
1.2	Erste Anwendungen des Apparats	10
1.3	Inhalt der Fläche unter einem Graphen	11
1.4	Von allen Geraden approximiert die Tangente am besten	12
1.5	Verbesserung der Approximation	13
1.6	Das Taylor-Polynom	14
1.7	Die e -Funktion	16
1.8	Ableitungsregeln	17
2	Kurven in Parameterdarstellung (1)	19
2.1	Einstiegsbeispiel: Der laufende Käfer	19
2.2	Bahngeschwindigkeit und Tangentensteigung	20
3	Das Integral	23
3.1	Käferweglängen und Flächen unter der Kurve $y = v(t)$	23
3.2	Riemannsche Summen	23
3.3	Obersummen und Untersummen	24
4	Kurven in Parameterdarstellung (2)	27
4.1	Flächenstücke zwische Kurve und Achse	27
4.2	Die Leibnizsche Sektorformel	28
5	Volumenbestimmung mit Hilfe von Integralen	31
5.1	Funktionen in zwei Variablen	31
5.2	Das Volumen eines Kartoffelkörpers	32
5.3	Das Prinzip von Cavalieri	33
6	Kurven in Parameterdarstellung (3)	35
6.1	Ebene und Raum	35
6.2	Vektoren	35
6.3	Geschwindigkeitsvektoren von Kurven	37
6.4	Winkel	37
6.5	Übungen	39
7	Klausuren	41
7.1	Erste Klausur 12.1	41
7.2	Zweite Klausur 12.1	43
7.3	Nachschreibausgabe der zweiten Klausur	44
8	Fragen	47

II	Stochastik	49
9	Einige Beispiele zur Einführung	51
9.1	Ein konkretes Beispiel	51
9.2	Aufgaben	52
9.3	Die Verteilungsfunktion einer Zufallsvariablen	52
9.4	Verteilungsfunktion durch Simulation	53
9.5	Die Dichtefunktion einer Zufallsvariablen	53
9.6	Eine Formel für den Erwartungswert einer Zufallsvariablen	54
10	Wirklichkeit und Mathematik	57
11	Konstruktion einer Simulation zu X mit gegebenem F	59
11.1	Die Methode	59
11.2	Eine neue Formel für den Erwartungswert	60
11.3	Beispiel: Radioaktiver Zerfall	60
12	Die Varianz einer Zufallsvariablen	63
12.1	Funktionen von Zufallsvariablen	63
12.2	Definition der Varianz	63
12.3	Die Ungleichung von Tschebyschew	64
13	Übungsaufgaben zu stetig verteilten Zufallsvariablen	65
14	Diskrete Zufallsvariable	67
14.1	Ein typisches Beispiel	67
14.2	Rechenregeln für Erwartungswerte: $E(X + Y), E(aX + b)$	69
14.3	Bedingte Wahrscheinlichkeit und Unabhängigkeit	70
14.4	Erwartungswert eines Produkts: $E(XY)$	70
14.5	Regeln für die Varianz	71
14.6	Ketten von Zufallsversuchen	72
14.7	Bernoullis schwaches Gesetz der großen Zahl	73
14.8	Hypergeometrisch-verteilte Zufallsvariable	74
14.9	Eigenschaften binomial-verteilter Zufallsvariablen	74
14.10	Warten auf Erfolg: geometrisch verteilte Zufallsvariable	76
14.11	Poisson-verteilte Zufallsvariable	77
15	Beurteilende Statistik	79
15.1	Wieviel Erfolge gibt es bei einer Bernoulli-Kette normalerweise?	79
15.2	Das Konfidenzintervall	79
15.3	Testen	80
15.4	Schnelltest für den Alltag	80
16	Klausuren	83
16.1	Erste Klausur 12.2	83
16.2	Zweite Klausur 12.2	84
III	Lineare Algebra und Geometrie	87
17	Lineare Algebra	89
17.1	Lineare Gleichungssysteme	89
17.2	Der Spaltenraum \mathbb{R}^n	90
17.3	Lösungsmengen linearer Gleichungssysteme	90
17.4	Für welche rechte Seite \vec{b} ist $A\vec{x} = \vec{b}$ lösbar?	91

17.5 „Größe“ eines Erzeugnisses und lineare Unabhängigkeit	91
17.6 Ein durchgerechnetes Beispiel	92
17.7 Matrixabbildungen und lineare Gleichungssysteme	93
17.8 Lineare Abbildungen	94
17.9 Einige Maple-Befehle aus dem Paket linalg	95
18 Analytische Geometrie des Raumes	97
18.1 Die Beziehung zwischen dem \mathbb{R}^3 und dem Anschauungsraum	97
18.2 Gleichung einer Ebene	98
18.3 Ebenengleichungen sind sehr nützlich	98
18.4 Abstandsaufgaben	99
18.5 Kugeln	99
18.6 Aufgaben	100
19 Geometrie im \mathbb{R}^4	101
19.1 Theorie	101
19.2 Aufgaben	102
20 Der \mathbb{R}^n als Vektorraum	105
20.1 Teilräume des \mathbb{R}^n	105
20.2 Basen und Dimension eines Teilraums	106
21 Der Vektorraum der Funktionen $I \rightarrow \mathbb{R}$	107
21.1 Das Konzept des Vektorraums	107
21.2 Lineare Unabhängigkeit im Funktionenraum	107
21.3 Skalarprodukt im Funktionenraum	108
21.4 Fourier-Reihen	109
22 Kurven und Flächen im Raum	111
22.1 Parameterdarstellungen von Kurven und Flächen	111
22.2 Tangentialebene einer Fläche	112
22.3 Regelflächen	113
23 Klausuren	115
23.1 Erste Klausur 13.1	115
23.2 Zweite Klausur 13.1	116
23.3 Klausur 13.2	118
23.4 Abiturklausur Abi02	119

Teil I
Analysis

Kapitel 1

Änderungen und Änderungsrate

1.1 Die Fundamentalskizze

Eine Größe y hängt von einer anderen Größe x ab. Gewöhnlich schreibst du diese Abhängigkeit in Form einer Gleichung $y = f(x)$, und das werden wir auch beibehalten. Wo es passt, lasse ich aber das Funktionssymbol f weg und benutze nur die Variablen selbst, und außer x und y benutze ich noch u, v, w, z, \dots , was gerade gebraucht wird.

Zu einem bestimmten Wert x gehört ein bestimmter Wert y und ein Punkt $(x|y)$ einer Kurve. Hauptaufgabe der Analysis ist es, Angaben darüber zu machen, wie sich der Wert y ändert, wenn man dem Wert x eine Änderung Δx aufprägt. Bisher hast du für diese Änderung des x -Wertes den Buchstaben h verwandt, ich nehme für eine Änderung einer Größe x, y, z, \dots aber nach alter Tradition $\Delta x, \Delta y, \Delta z, \dots$.

Die Änderung Δy , die der y -Wert erfährt, wenn wir x um Δx ändern, zerlegen wir in den linearen Anteil

$$dx := f'(x)\Delta x \quad (1.1)$$

und den Fehler $r(\Delta x)$, den man am besten durch die Gleichung

$$\Delta y = f'(x)\Delta x + r(\Delta x) \quad (1.2)$$

beschreibt. Die Situation ist in Abbildung 1.1 dargestellt.

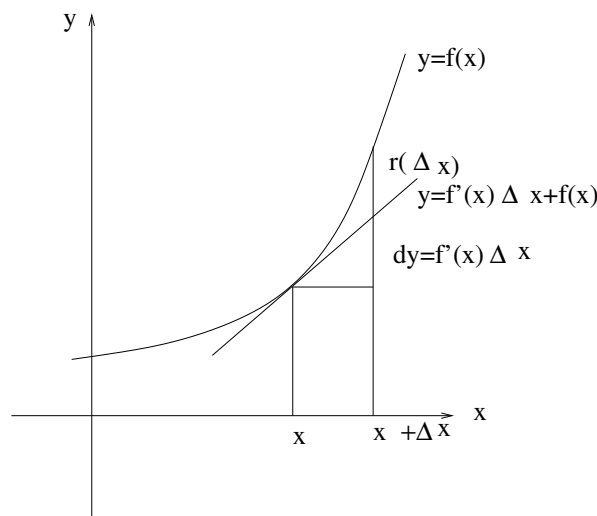


Abbildung 1.1: Unsere Fundamentalskizze

Den in Abbildung 1.1 dargestellten Zusammenhang nutzen wir in vielfältiger Weise aus:

1. Für „kleine“ Δx ist $f'(x) \approx \frac{\Delta y}{\Delta x}$.
2. Für „kleine“ Δx ist $\Delta y \approx f'(x)\Delta x$.
3. Für „kleine“ Δx ist $f(x + \Delta x) \approx f(x) + f'(x)\Delta x$.
4. „In der Nähe“ des Punktes $(x|y)$ fallen Kurve und Tangente fast zusammen, man kann also dort die Tangente als Approximation für die Kurve verwenden.

Klar, hier wird in etwas anderen Worten viermal das Gleiche gesagt, nur der Blickwinkel ist immer ein klein wenig anders.¹

1.2 Erste Anwendungen des Apparats

Mit unseren Hilfsmitteln gewinnen wir Formeln für das Volumen eines Kreiskegels und das Volumen einer Kugel. Die Stationen des Wegs halte ich hier für das Beispiel des Kegels fest. Es sei also das Volumen eines Kegels mit der Höhe H und dem Radius R der Grundfläche gesucht.

1. Verallgemeinere das Problem: Wir nennen das Volumen des Kegelteils von der Spitze bis zur Höhe x , von der Spitze aus gemessen, $y = V(x)$. Jetzt bestimmen wir $V(x)$; wenn wir das haben, ist das ursprünglich gesuchte Volumen einfach $V(H)$.²
2. Schau dir die Änderung Δy an, um die sich $V(x)$ ändert, wenn du x um Δx änderst. Es handelt sich um das Volumen eines Kegelstumpfes der Höhe Δx .³
3. Zum Glück brauchen wir Δy nur so genau, dass wir den Grenzwert von $\frac{\Delta y}{\Delta x}$ für Δx gegen 0 bestimmen können (Martin!). Im Extremfall nehmen wir Δx so klein, dass wir den Kegelstumpf als eine hauchdünne Scheibe der Dicke Δx ansehen können. Ihr Radius ist der des Kegels in der Höhe x , von der Spitze aus gemessen, also $x\frac{R}{H}$. Ihr Volumen ist dann mit guter Genauigkeit

$$\pi \left(x \frac{R}{H} \right)^2 \Delta x \quad .$$

4. Nun haben wir also in wunderbarer Weise

$$\left(x \frac{R}{H} \right)^2 \pi$$

als Ableitung der Funktion V gewonnen, von der wir keinen Term kennen! Wir schließen daraus, dass

$$V(x) = \frac{1}{3} \left(\frac{R}{H} \right)^2 x^3 \pi + c$$

sein sollte für eine feste Zahl $c \in \mathbb{R}$. Mit Christians Idee bestimmen wir c , indem wir $V(0)$ bilden - das muss ja = 0 sein. Es folgt $c = 0$.

5. Das gesuchte Kegelvolumen bekommen wir nun, indem wir H für x einsetzen:

$$V(H) = \frac{1}{3} \pi R^2 H \quad .$$

¹Wenn man exakte Mathematik machen will, muss man unbedingt präzisieren, wann ein Δx „klein“ heißen soll, aber für heute begnügen wir uns mit dieser anschaulichen Sprechweise.

²Eigentlich ist das Wahnsinn, jetzt haben wir es mit unendlich vielen Kegeln zu tun, die Aufgabe ist also viel schwieriger geworden!

³Doppelter Wahnsinn, das Volumen eines Kegelstumpfes ist schwieriger zu berechnen als das Volumen eines Kegels.

Wie ging das nun? Entscheidend war, dass wir die Situation überschaubar machten, indem wir nur kleine Δx betrachteten haben. So gelangen den Erfindern der Analysis wichtige Entdeckungen. Physiker arbeiten übrigens häufig so.

1.3 Inhalt der Fläche unter einem Graphen

Gegeben sei eine Funktion $y = f(x)$, deren Graph oberhalb der x -Achse liegt. Wir markieren zwei Stellen a und b auf der x -Achse und fragen nach dem Inhalt des Flächenstücks, das zwischen $x = a$ und $x = b$ liegt, nach oben durch den Graphen und nach unten durch die x -Achse begrenzt wird. Man nennt das die Fläche unter dem Graphen von f zwischen a und b .

Das Flächenstück ist im oberen Teil von Abbildung 1.2 dargestellt (die rechte Grenze b ist aus Platzgründen nicht bezeichnet). Wir gehen genau so vor, wie in Abschnitt 1.2 beschrieben ist, und definieren die Funktion $A(x) :=$ Inhalt des Teils der Fläche zwischen a und x . Der Graph von A ist im unteren Teil von Abbildung 1.2 skizziert, so ungefähr sollte er aussehen⁴. Wir fragen nun nach

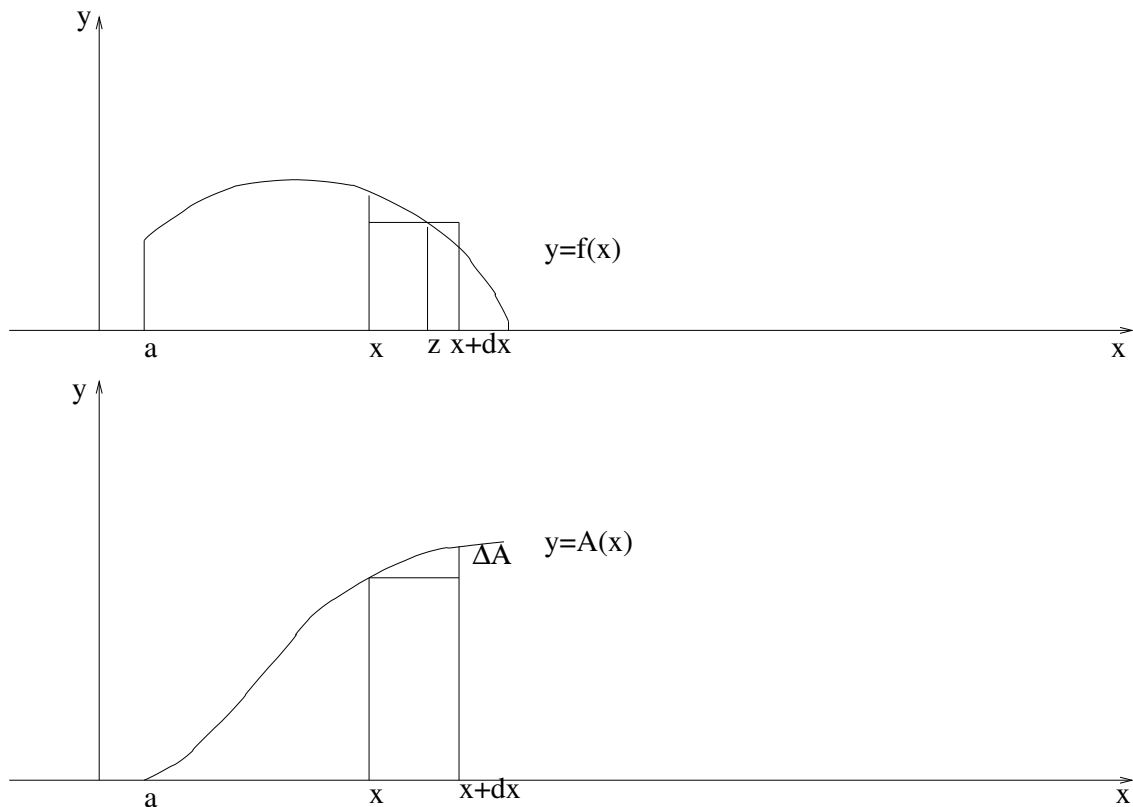


Abbildung 1.2: Graphen einer Funktion f und der zugehörigen Flächeninhaltsfunktion A

der Ableitung von A . Dazu schauen wir auf eine Stelle x , vergrößern x um Δx (in der Abbildung mit dx bezeichnet) und sehen uns die Änderung ΔA an, die $A(x)$ dabei erfährt. Das ist der Inhalt des Streifens zwischen x und $x + dx$. Es gibt nun einen Wert z zwischen x und $x + dx$, so dass das Rechteck mit der Breite dx und der Höhe $f(z)$ den gleichen Flächeninhalt hat wie der Streifen. Müssten wir z konkret angeben, gerieten wir in arge Verlegenheit, aber zum Glück reicht es uns, zu wissen, dass es solch ein z gibt. Wenn wir dx ändern, ändert sich auch das z , und für $dx \rightarrow 0$ strebt offensichtlich $f(z)$ gegen $f(x)$. Das reicht! Denn wegen

$$\frac{\Delta A}{\Delta x} = \frac{f(z)\Delta x}{\Delta x} = f(z) \xrightarrow{\Delta x \rightarrow 0} f(x)$$

⁴Warum? Warum ist $A(a) = 0$? Wo sollte der Graph einen Wendepunkt haben? Sollte er immer steigen?

folgt daraus, dass $A'(x) = f(x)$ ist.

Nun müssen wir uns eine Stammfunktion F von f besorgen, also eine Funktion, deren Ableitung unser f ist. Die ist übrigens gar nicht immer leicht zu finden. Aber wenn wir eine haben, sollte $A(x) = F(x) - F(a)$ sein, denn wir brauchen ja $A(a) = 0$. Der gesuchte Inhalt des gesamten Flächenstücks ist nun $A(b) = F(b) - F(a)$ - wie von Christians Bruder vorhergesagt. Seine Methode kannst du getrost auch benutzen: Um den Inhalt des Flächenstücks unter dem Graphen von f (oberhalb der x -Achse!) zwischen a und b zu bestimmen, nimmst du eine Stammfunktion F von f und bildest $F(b) - F(a)$, fertig. Natürlich solltest du wissen, warum das funktioniert!

1.4 Von allen Geraden approximiert die Tangente am besten

Wir schauen uns eine Funktion $y = f(x)$ in der Nähe einer Stelle x_0 an. Was passiert eigentlich, wenn wir die Funktion in der Nähe von x_0 nicht durch die Tangente $t : y = f(x_0) + f'(x_0)\Delta x$ annähern, sondern durch eine beliebige Gerade g durch $(x_0 | f(x_0))$? Eine solche Gerade hat eine Gleichung der Form $g : y = f(x_0) + m\Delta x$, der Abschneidefehler ist dann

$$R(\Delta x) := f(x_0 + \Delta x) - f(x_0) - m\Delta x .$$

Vergleichen wir ihn mit dem Abschneidefehler $r(\Delta x) = f(x_0 + \Delta x) - f(x_0) - f'(x_0)\Delta x$ der Tangente!

1. Wenn wir Δx gegen 0 laufen lassen, strebt $r(\Delta x)$ auch gegen 0, aber das gilt auch für $R(\Delta x)$.⁵
2. Das Besondere bei der Tangente ist, dass wir ihr $r(\Delta x)$ noch durch Δx dividieren dürfen, der Quotient strebt mit Δx immer noch gegen 0. Das sieht man so:

$$\frac{r(\Delta x)}{\Delta x} = \frac{f(x_0 + \Delta x) - f(x_0) - f'(x_0)\Delta x}{\Delta x} = \frac{\Delta y}{\Delta x} - f'(x_0) \xrightarrow{\Delta x \rightarrow \infty} f'(x_0) - f'(x_0) = 0$$

Für $\frac{R(\Delta x)}{\Delta x}$ dagegen erhalten wir

$$\frac{R(\Delta x)}{\Delta x} = \frac{f(x_0 + \Delta x) - f(x_0) - m\Delta x}{\Delta x} = \frac{\Delta y}{\Delta x} - m \xrightarrow{\Delta x \rightarrow 0} f'(x_0) - m ,$$

und das ist nicht 0, wenn g nicht die Tangente ist.

Dieses Resultat notieren wir in einem Lemma, so nennt man einen kleinen mathematischen Lehrsatz.

1 Lemma *Es sei $P(x_0 | f(x_0))$ ein Punkt des Graphen einer Funktion f und $g : y = f(x_0) + m\Delta x$ die Gerade durch P mit der Steigung m . Der Fehler $R(\Delta x) := f(x_0 + \Delta x) - f(x_0) - m\Delta x$ erfüllt die Bedingung*

$$\frac{R(\Delta x)}{\Delta x} \xrightarrow{\Delta x \rightarrow 0} 0$$

dann und nur dann, wenn $m = f'(x_0)$ ist, und dann ist g die Tangente an den Graphen von f in P .

Du solltest dir unbedingt einige Skizzen machen!

⁵Aufgabe: Rechne mal $r(\Delta x)$ und $R(\Delta x)$ für $f(x) = x^2$, $x_0 = 1$ und $m = 2, 1$ konkret aus!

1.5 Verbesserung der Approximation

Wie berechnet der Taschenrechner seine Sinuswerte? Dieser Frage wollen wir nachspüren. Einige Überlegungen hatten wir schon angestellt: Am gleichseitigen Dreieck können wir herleiten, dass $\sin(\pi/6) = 1/2$ und dass $\cos(\pi/6) = 1/2\sqrt{3}$ ist. Wenn wir die Sinuskurve näherungsweise durch ihre Tangente ersetzen, können wir dann auch halbwegs brauchbare Werte für $\sin(x)$ angeben, wenn x in der Nähe von $\pi/6$ liegt. Abbildung 1.3 zeigt dieses Stück der Sinuskurve. Die waagerechte Achse trägt eine Δx -Skala, zu $\Delta x = 0$ gehört der x -Wert $x_0 = \pi/6$. Außer der Sinuskurve ist noch die Tangente im Punkt $P(x_0 | \sin(x_0))$ eingezeichnet. Wenn man Abbildung 1.3 anschaut,

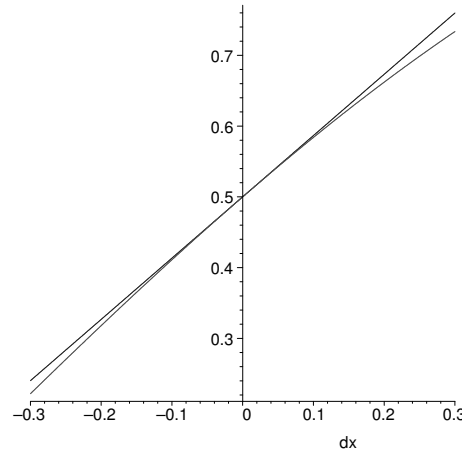


Abbildung 1.3: Die Sinuskurve in der Nähe von $x_0 = \pi/6$ und ihre Tangente an dieser Stelle

wird man vielleicht ein Δx bis $\pm 0,1$ akzeptieren.

Wie kann man die Näherung verbessern - sowohl ihre Güte als auch ihre Reichweite? Lemma 1 zeigt, dass wir mit Geraden nicht weiterkommen. Dann müssten wir es mit einer Parabel als approximierender Kurve versuchen. Der einfachste Weg wäre, zum Term

$$t(\Delta x) = f(x_0) + f'(x_0)\Delta x = \sin(\pi/6) + \cos(\pi/6)\Delta x \quad (1.3)$$

der Tangente einen geeigneten quadratischen Term $a(\Delta x)^2$ hinzuzufügen. Dass man Kurven so besser approximieren kann, als nur mit der Tangente, mag Abbildung 1.4 zeigen - im Unterricht kann ich dir eine Animation dazu zeigen.

Versuchen wir also, unsere Sinuskurve in der Nähe von $x_0 = \pi/6$ durch eine quadratische Funktion

$$q(\Delta x) = \sin(\pi/6) + \cos(\pi/6)\Delta x + a(\Delta x)^2 \quad (1.4)$$

möglichst gut zu approximieren. Das heißt, wir müssen das a in Gleichung 1.4 so wählen, dass der Graph von q möglichst gut passt. Aber wie soll das gehen? Sehen wir uns die Summanden in $q(\Delta x)$ der Reihe nach an:

1. Der konstante Term $\sin(\pi/6) = f(x_0)$ sorgt eigentlich nur dafür, dass der Graph von q an der Stelle x_0 die richtige Höhe hat; mehr kann man von ihm auch nicht verlangen.
2. Der Summand $\cos(\pi/6)\Delta x = f'(x_0)\Delta x$ sorgt dafür, dass der Graph von q an der Stelle x_0 die richtige Steigung hat. Zusammen mit dem ersten Summanden bildet er den Term der Tangente. Ganz nahe bei x_0 passt die sehr gut. Dass sich die Sinuskurve von der Tangente entfernt, wenn man weiter von x_0 weggeht, liegt daran, dass die Tangente immer mit der gleichen Steigung weiterläuft, während sich die Steigung der Sinuskurve mit Δx ändert. Dieser Änderung der Steigung der Sinuskurve kann die Tangente nicht folgen.

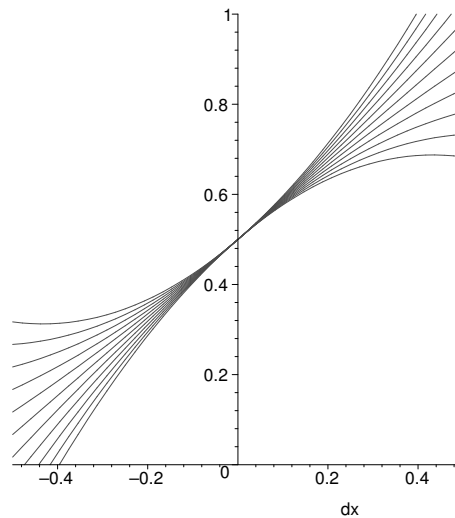


Abbildung 1.4: $f(x_0) + f'(x_0)\Delta x + k/10(\Delta x)^2$ für $k = -5, \dots, 5$

Der quadratische Term sollte also die Änderungstendenz⁶ der Ableitung der Sinusfunktion an der Stelle x_0 erfassen. Maß für die Änderungstendenz der Ableitung ist die Ableitung der Ableitung, also die zweite Ableitung! Nun haben wir einen Erfolg versprechenden Ansatz: **Wähle a so, dass $f''(x_0) = p''(\Delta x)$ ist!** Daraus ergibt sich

$$a = \frac{1}{2}f''(x_0) = -\frac{1}{2}\sin\left(\frac{\pi}{6}\right) .$$

Mit diesem a erhalten wir als quadratische Näherung für die Sinusfunktion in der Nähe von $x_0 = \pi/6$

$$q(\Delta x) = f(x_0) + f'(x_0)\Delta x + \frac{1}{2}f''(x_0)(\Delta x)^2 = \sin\left(\frac{\pi}{6}\right) + \cos\left(\frac{\pi}{6}\right)\Delta x - \frac{1}{2}\sin\left(\frac{\pi}{6}\right)(\Delta x)^2 . \quad (1.5)$$

Um einen Eindruck zu bekommen, wie gut die neue Näherung im Vergleich zur alten ist, lassen wir uns die Sinuskurve, ihre Tangente bei x_0 und den Graphen von q in ein Schaubild zeichnen (siehe Abbildung 1.5).

Aufgabe Berechne Näherungswerte für $\sin(35^\circ)$, $\sin(40^\circ)$, $\sin(25^\circ)$ und $\sin(20^\circ)$ mit $q(\Delta x)$ und vergleiche sie mit den Taschenrechnerwerten.

1.6 Das Taylor-Polynom

Wir haben in Abschnitt 1.5 gesehen, dass sich die Approximation verbessert, wenn wir zum Term $f(x_0) + f'(x_0)\Delta x$ der Tangente an die Kurve $y = f(x)$ einen quadratischen Term $a(\Delta x)^2$ hinzuaddieren, so dass die neue Funktion mit f an der betrachteten Stelle im Funktionswert und in den ersten beiden Ableitungen übereinstimmt. Wenn wir noch weitere Summanden $a_3(\Delta x)^3$, $a_4(\Delta x)^4$, \dots , $a_n(\Delta x)^n$ hinzufügen, können wir sogar Übereinstimmung der Werte der 3., 4., \dots , n -ten Ableitung der beiden Funktionen an der betrachteten Stelle erreichen. Formal gesprochen suchen wir ein Polynom

$$p(\Delta x) = a_0 + a_1\Delta x + a_2(\Delta x)^2 + \dots + a_n(\Delta x)^n =: \sum_{k=0}^n a_k(\Delta x)^k ,$$

so dass

$$p(0) = f(x_0), \quad p'(0) = f'(x_0), \quad p''(0) = f''(x_0), \quad \dots, \quad p^{(n)}(0) = f^{(n)}(x_0)$$

⁶Die Ableitung wird ja auch als lokale Änderungsrate bezeichnet! Kennst du diesen Ausdruck?

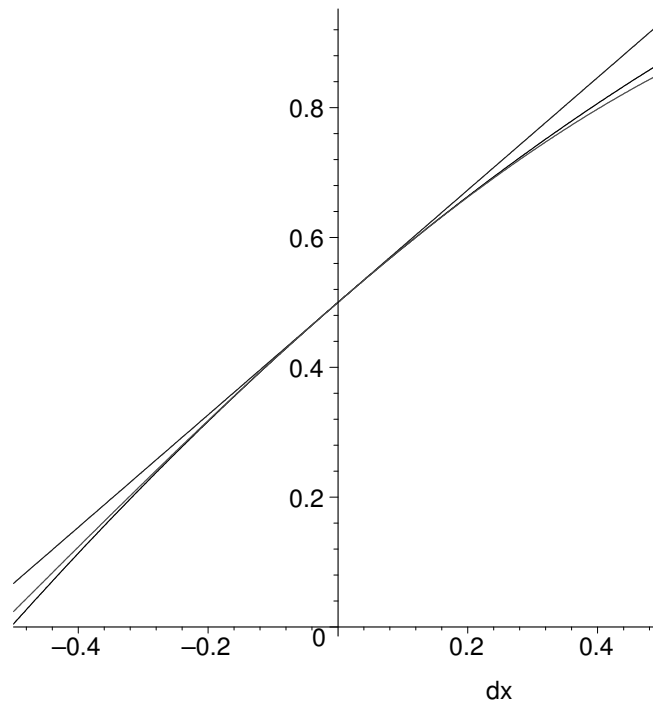


Abbildung 1.5: Sinuskurve (rot), Tangente (blau) und Graph von q aus Gleichung 1.5 (schwarz) an der Stelle $x_0 = \pi/6$

ist. Berechnen wir daraus die $a_k!$

$$p(0) = a_0 = f(x_0) \Rightarrow a_0 = f(x_0)$$

$$p'(0) = a_1 = f'(x_0) \Rightarrow a_1 = f'(x_0)$$

$$p''(0) = 2a_2 = f''(x_0) \Rightarrow a_2 = \frac{1}{2}f''(x_0)$$

$$p^{(3)}(0) = 3 \cdot 2a_3 = f^{(3)}(x_0) \Rightarrow a_3 = \frac{1}{2 \cdot 3}f^{(3)}(x_0)$$

...

$$p^{(k)}(0) = k(k-1)(k-2) \cdots 2a_k = f^{(k)}(x_0) \Rightarrow a_k = \frac{1}{k(k-1)(k-2) \cdots 2}f^{(k)}(x_0)$$

...

$$p^{(n)}(0) = n(n-1)(n-2) \cdots 2a_n = f^{(n)}(x_0) \Rightarrow a_n = \frac{1}{n(n-1)(n-2) \cdots 2}f^{(n)}(x_0)$$

Du wirst diese Überlegung schon verstehen, wenn du dir klar machst, dass in der k -ten Ableitung $p^{(k)}(\Delta x)$ die ersten k Summanden durch das Ableiten weggefallen sind. Als erster Summand steht also der mit a_k da. Von dem $(\Delta x)^k$, das dazugehörte, ist kein Faktor Δx mehr übrig. Das erste Ableiten brachte dem a_k den Faktor k , das zweite Ableiten den Faktor $k-1$, und so weiter, bis zum Faktor 1 beim k -ten Ableiten. Das ganze Produkt ist als $p^{(k)}(0)$ angegeben. Die folgenden Summanden $a_{k+1}(\Delta x)^{(k+1)}$ und so weiter von $p(\Delta x)$ enthielten mehr als k Faktoren Δx ; von denen sind noch welche übrig, wenn man k -mal abgeleitet hat. Zu $p^{(k)}(0)$ liefern die keinen Beitrag, weil man ja eben 0 für Δx einsetzt.

Wir brauchen nun einige Bezeichnungen, damit wir unser Ergebnis schön hinschreiben können. Für die k -te Ableitung von f haben wir das Symbol $f^{(k)}$, dem bist du ja schon begegnet. Wir vereinbaren, dass $f^{(0)} = f$ ist, die nullte Ableitung ist die Funktion selbst. Ferner brauchen wir

$$k! := 1 \cdot 2 \cdot 3 \cdots k \quad \text{lies: } k\text{-Fakultät}$$

mit dem Sonderfall $0! := 1$ und das Summenzeichen Σ . Damit definieren wir:

2 Definition Das n -te Taylor-Polynom zu f an der Stelle x_0 ist das Polynom

$$\begin{aligned} p(\Delta x) &= f(x_0) + f'(x_0)\Delta x + \frac{1}{2}f''(x_0)(\Delta x)^2 + \cdots + \frac{1}{k!}f^{(k)}(x_0)(\Delta x)^k + \cdots + \frac{1}{n!}f^{(n)}(x_0)(\Delta x)^n \\ &= \sum_{k=0}^n \frac{1}{k!}f^{(k)}(x_0)(\Delta x)^k . \end{aligned}$$

Wie wir gesehen haben, nähern Taylor-Polynome die Sinusfunktion ausgezeichnet an, das zehnte schon auf einer vollen Schwingung.

Aufgabe Ein Grundkurschüler hat einen Taschenrechner, der nur die Grundrechenarten beherrscht. Schreibe ihm eine Formel hin, mit der er Sinuswerte mit vernünftiger Genauigkeit berechnen kann. Gib auch eine Gebrauchsanweisung für deine Formel - eine Gebrauchsanweisung, keine Herleitung!

1.7 Die e -Funktion

Du kennst Exponentialfunktionen $x \mapsto a^x$ für eine positive Basis $a \neq 1$. Abbildung 1.6 zeigt einige Beispiele mit $a > 1$. Die Kurven gehen alle durch $(0|1)$, sie verlaufen alle oberhalb der x -Achse.

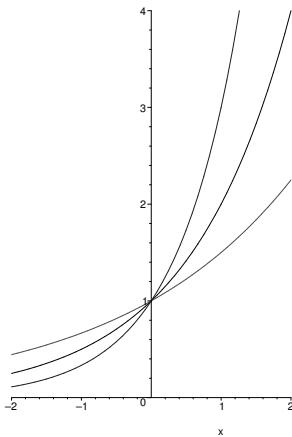


Abbildung 1.6: Graphen von $x \mapsto a^x$ für $a \in \{1.5, 2, 3\}$

Für $x \rightarrow -\infty$ strebt a^x gegen 0, für $x \rightarrow \infty$ strebt a^x gegen ∞ (für $0 < a < 1$ ist das umgekehrt, für $a < 0$ ist a^x nicht definiert).

Du kannst dir eine Skizze eines solchen Graphen zeichnen und dir überlegen, wie die Ableitungskurve dazu wohl aussieht.

Versuchen wir, $(a^x)'$ zu berechnen. Versuche es nicht mit bekannten Regeln, das geht garantiert schief. Wir müssen ganz vorn anfangen:

$$\frac{\Delta y}{\Delta x} = \frac{a^{x+\Delta x} - a^x}{\Delta x} = \frac{a^x(a^{\Delta x} - 1)}{\Delta x} = a^x \frac{a^{\Delta x} - 1}{\Delta x}$$

Daran sehen wir, dass die Ableitung von a^x wieder a^x ist, multipliziert mit einem konstanten Faktor, nämlich mit dem Grenzwert von $(a^{\Delta x} - 1)/\Delta x$ für $\Delta x \rightarrow 0$. Nun, dieser Bruch, dessen Grenzwert da zu bilden ist, ist die Steigung der Sekante durch die Kurvenpunkte $(0|1)$ und $(\Delta x|a^{\Delta x})$, und er strebt für $\Delta x \rightarrow 0$ gegen die Steigung der Tangente an die Kurve im Punkt $(0|1)$. Wenn wir die Steigung der Kurve in diesem einen Punkt kennen, haben wir schon die ganze Ableitung zur Hand.

Leider kannst du nur Näherungswerte dieser Steigung berechnen, indem du einfach mit dem Taschenrechner die Sekantensteigung für kleine Δx berechnest. Aber es gibt einen genialen Ausweg: Schau dir die Beispielkurven in Abbildung 1.6 einmal an: je größer a ist, desto steiler ist die Tangente in $(0|1)$. Du kannst die Tangentensteigung fast auf 0 bringen, aber auch beliebig groß machen⁷.

Wenn wir uns vorstellen, dass wir a stufenlos wachsen lassen und dabei die Tangentensteigung im Punkt $(0|1)$ stufenlos mitwächst, sollte wir doch annehmen, dass es **einen** a -Wert gibt, für den die Tangentensteigung gerade den Wert 1 hat, einverstanden? Dieses a nennen wir e , es ist die berühmte eulersche Zahl. Die Ableitung von e^x ist dann wieder e^x !

3 Definition Wir nennen die Basis der Exponentialfunktion $x \mapsto a^x$, die an der Stelle 0 die Ableitung 1 hat, e .

Wir werden gleich beliebig genaue Näherungswerte für e erzeugen, und zwar mit Hilfe des Taylor-Polynoms. Wenn nämlich $(e^x)' = e^x$ ist, ist doch $(e^x)^{(k)} = e^x$ für jedes $k \in \mathbb{N}$. Du kannst also sofort eine Formel für das n -te Taylor-Polynom $p(\Delta x)$ von e^x an der Stelle 0 hinschreiben. Näherungswerte für e bekommst du dann, wenn du einfach $p(1)$ bildest. Führe das selbst durch, finde Näherungswerte für e und vergleiche sie mit dem Tabellen- oder Taschenrechnerwert. -

Natürlich hast du den Auftrag treu erledigt. Weil die Taylor-Polynome der Exponentialfunktion von großer Bedeutung sind, schreibe ich das Ergebnis hier noch einmal auf: Die n -te Ableitung von e^x ist e^x , ihr Wert an der Stelle 0 also stets 1. Damit erhalten wir für das n -te Taylor-Polynom der Exponentialfunktion an der Stelle 0

$$p_n(\Delta x) = \sum_{k=0}^n \frac{1}{k!} (\Delta x)^k . \quad (1.6)$$

Wenn du einen Näherungswert für e haben willst, nimmst du einfach $p_n(1)$. Für $p_7(1)$ bekommst du schon den Taschenrechnerwert für e :

$$e \approx p_7(1) = \sum_{k=0}^7 \frac{1}{k!} = \frac{685}{252} \approx 2.718253968 . \quad (1.7)$$

Ich hoffe doch, dass du diese erstaunlichen Erfolge unserer Methoden zu würdigen weißt!⁸

Wir bekommen nun auch die Ableitung der allgemeinen Exponentialfunktion $x \mapsto a^x$ für $a > 0, a \neq 1$. Wir können nämlich a als e^b schreiben, wie du dir am Graphen der e -Funktion klarmachen kannst. Diese Zahl b , mit der du e potenzieren musst, um a zu erhalten, ist der natürliche Logarithmus von a , geschrieben $\ln(a)$. Nach der Kettenregel, die du inzwischen kennst, ist dann

$$(a^x)' = \left((e^{\ln(a)x}) \right)' = \left(e^{x \ln(a)} \right)' = e^{x \ln(a)} \ln(a) = a^x \ln(a) .$$

Die Tangentensteigung von $x \mapsto a^x$ an der Stelle 0 ist also $\ln(a)$!

1.8 Ableitungsregeln

Die stehen nun wirklich an jeder Ecke, interessant wäre höchstens unsere Beweismethode mit den Δ s. Ich zähle die Regeln nur auf: Summenregel, Faktorregel, Produktregel, Quotientenregel und die berühmte Kettenregel für zusammengesetzte Funktionen:

$$(u \pm v)' = u' \pm v' \quad (cu)' = cu' \text{ für } c \in \mathbb{R} \quad (uv)' = u'v + uv' \quad (1.8)$$

$$\left(\frac{u}{v} \right)' = \frac{u'v - uv'}{v^2} \text{ für } v \neq 0 \quad (f(g(x)))' = f'(g(x))g'(x) \quad (1.9)$$

⁷Wie musst du a dazu jeweils wählen?

⁸Übrigens gilt exakt $e^x = \sum_{k=0}^{\infty} \frac{1}{k!} x^k$, dies ist eine unendliche Reihe, die Taylorreihe der Exponentialfunktion.

Die Ableitung des natürlichen Logarithmus bekommst du aus Symmetrieüberlegungen an den Graphen von $y = e^x$ und $y = \ln(x)$ oder über implizite Differentiation. Letztere geht so: Es ist $e^{\ln(x)} = x$ für alle positiven x . Du leitest beide Seiten der Gleichung ab (links Kettenregel beachten):

$$\begin{array}{r|l} e^{\ln(x)} = x & ' \\ e^{\ln(x)}(\ln(x))' = 1 & | \quad e^{\ln(x)} = x \\ x(\ln(x))' = 1 & | \quad : x \\ (\ln(x))' = \frac{1}{x} & \end{array}$$

Ein erstaunliches Ergebnis, merke es dir.

Aufgabe Du weißt, dass $(e^x)' = e^x$ ist. Es sei nun f eine weitere Funktion mit der Eigenschaft $f'(x) = f(x)$. Berechne die Ableitung des Quotienten $f(x)/e^x$. Was sagt uns das Ergebnis?

Kapitel 2

Kurven in Parameterdarstellung (1)

2.1 Einstiegsbeispiel: Der laufende Käfer

Keine Sorge, du brauchst weder Biologie noch Physik, wir denken uns alles einfach aus. Also: Eine große ebene Scheibe dreht sich mit konstanter Geschwindigkeit um ihren Mittelpunkt. Ein Käfer läuft auf der Scheibe mit konstanter Geschwindigkeit geradlinig vom Mittelpunkt weg, und du schaust dir den Vorgang von außen an. Ich habe eine Reihe von Fragen aufgeschrieben, die wir bedenken wollen, und ich freue mich, wenn du weitere Fragen hinzufügst. Zum Teil müssen die Fragen noch präzisiert werden, bevor sie ordentlich beantwortet werden können, aber das ist eben ein Teil der Arbeit.

1. Beschreibe die Bahn des Käfers, die du von außen siehst.
2. Wie schnell ist der Käfer zur Zeit t ?
3. In welcher Richtung bewegt sich der Käfer zur Zeit t ?
4. Welchen Weg legt er in der Zeit vom Start bis t zurück?
5. Welchen Weg auf der Scheibe und mit welcher Geschwindigkeit müsste der Käfer laufen, damit du ihn geradlinig laufen siehst?
6. Der Käfer befinde sich im Punkt $P(t)$. Es sei α der Winkel, um den man die Tangente an die Bahn im Punkt P gegen den Uhrzeigersinn drehen muss, bis sie senkrecht zum Fahrstrahl steht (mit Fahrstrahl bezeichnet man die Strecke vom Nullpunkt zu P). Bestimme $\tan(\alpha)$.
7. Die Bahnkurve des Käfers heißt Archimedische Spirale, du findest eine in Abbildung 2.1 auf Seite 20. Wenn man sie zur Verfügung hat, ist es ein Leichtes, einen gegebenen Winkel in drei gleich große Teile zu teilen (das ist eines der berühmten Probleme der Antike, die mit Zirkel und Lineal nicht lösbar sind). Wie stellt man das an?
8. Wir lassen den Vorgang im Dunkeln ablaufen und benutzen nur eine Diskolampe, die regelmäßig einen Lichtblitz abgibt und so den Ort des Käfers zu genau diesem Zeitpunkt sichtbar macht. An welchen Stellen siehst du den Käfer, wenn die Dunkelzeit der Lampe gleich der Umlaufzeit der Scheibe, gleich der halben oder gleich der doppelten Umlaufzeit ist?
9. Wie groß ist die Fläche, die der Fahrstrahl in der Zeit vom Start bis t überstreicht?
10. Zwei Käfer starten gleichzeitig. Der eine ist doppelt so schnell wie der andere. Vergleiche ihre Bahnkurven.

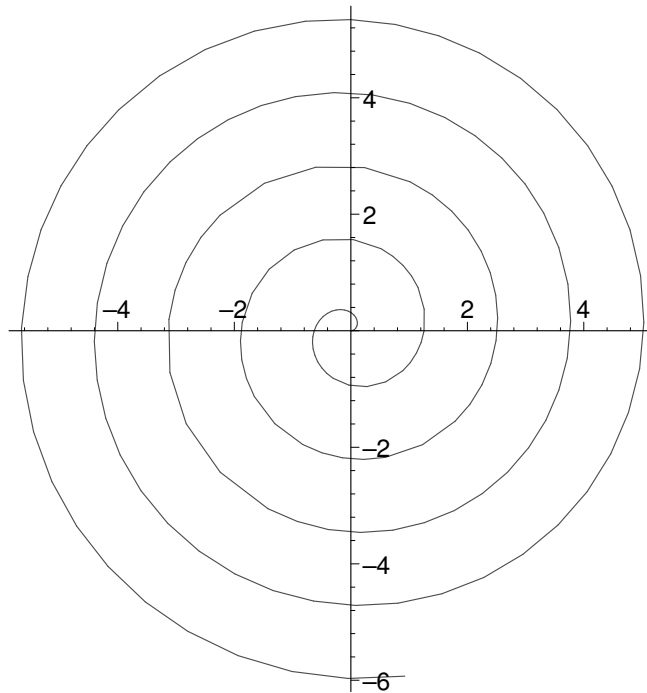


Abbildung 2.1: Maple-plot einer Archimedischen Spirale. Offensichtlich ist sie kein Funktionsgraph, denn zu jedem x -Wert gibt es mehrere Punkte der Kurve, die diesen Wert als x -Koordinate haben. Am besten stellst du dir die Kurve als Bahn eines beweglichen Punktes $P(t)$ vor, der sich zum Zeitpunkt t an der Stelle $(x(t)|y(t))$ befindet. Hier ist $x(t) = \frac{1}{5}t \cos(t)$ und $y(t) = \frac{1}{5}t \sin(t)$. Man sagt, man habe eine Parameterdarstellung der Kurve. Der plot-Befehl für Maple sieht dann so aus: `plot([t/5 * cos(t), t/5 * sin(t), t = 0..30], scaling=constrained);`

Achte darauf, dass du den Laufbereich des Parameters t mit in die eckige Klammer schreibst.

2.2 Bahngeschwindigkeit und Tangentensteigung

Das Beispiel und die Fragen in Abschnitt 2.1 dienen uns als Einstieg und Leitfaden. Formal gesprochen handelt es sich um Folgendes. Du stellst dir einen Punkt vor, der sich bewegt. Der Ort $P(t)$, an dem er sich zur Zeit t befindet, muss irgendwie durch eine Vorschrift gegeben sein. Das Symbol $P(t)$ erinnert nicht von ungefähr an eine Funktion, es handelt sich tatsächlich um eine Funktion

$$P : t \mapsto P(t) \quad ,$$

die jedem t aus einem Intervall $[a; b]$ ein Zahlenpaar $(x|y)$ zuordnet. Dabei sind $x = x(t)$ und $y = y(t)$ gewöhnliche Funktionen, wie du sie gut kennst. Nehmen wir ein Beispiel: Es sei

$$P(t) = (\cos(t)|\sin(t)) \quad \text{für } 0 \leq t \leq 4\pi.$$

Der Startpunkt der Reise ist $P(0) = (1|0)$. Die x -Koordinate $\cos(t)$ wird kleiner, die y -Koordinate $\sin(t)$ wird größer, also bewegt sich der Punkt vom Startpunkt aus nach links oben. Aber was rede ich, du hast die Kurve längst erkannt: es ist der Einheitskreis, und er wird gegen den Uhrzeigersinn zweimal durchlaufen.

Ich notiere noch ein paar Beispiele:

1. $P(t) = (t|t^2)$ für $-1 \leq t \leq 1$. Das ist das Stück der Normalparabel zwischen den Punkten $(-1|1)$ und $(1|1)$.
2. $P(t) = (\sin(t)|\sin^2(t))$ für $0 \leq t \leq 12\pi$. Hm, das ist wieder das gleiche Stück der Normalparabel. Der Startpunkt ist der Nullpunkt, dann geht es zum Punkt $(1|1) = P(\frac{1}{2}\pi)$, dann

wieder zurück zum Punkt $(-1|1) = P(\frac{3}{2}\pi)$, und so weiter, noch ein paarmal hin und her, bis die Reise wieder im Nullpunkt endet.

3. $P(t) = (t^2|t)$ für $-1 \leq t \leq 1$. Das ist nun eine nach rechts offene Normalparabel mit Scheitelpunkt $(0|0)$.
4. $P(t) = (t|f(t))$ für $a \leq t \leq b$. Na, das ist einfach der Graph der Funktion f im Intervall $[a; b]$.

An diesen Beispielen siehst du schon, dass Kurven viel flexibler sind als Funktionsgraphen.

Fragen wir nun nach der **Bahngeschwindigkeit** $v(t)$ des Punktes $P(t)$ zur Zeit t . Bei einer gescheiten Kurve ist das Stückchen von $P(t)$ bis $P(t + \Delta t)$ praktisch gerade, seine Länge ist also in sehr guter Näherung

$$\Delta s = \sqrt{(\Delta x)^2 + (\Delta y)^2} .$$

Um $v(t)$ zu erhalten, müssen wir Δs durch Δt teilen und den Grenzwert des Quotienten für Δt gegen 0 bilden:

$$\frac{\Delta s}{\Delta t} = \frac{\sqrt{(\Delta x)^2 + (\Delta y)^2}}{\Delta t} \quad (2.1)$$

$$= \sqrt{\left(\frac{\Delta x}{\Delta t}\right)^2 + \left(\frac{\Delta y}{\Delta t}\right)^2} \quad (\text{für } \Delta t > 0) \quad (2.2)$$

$$\xrightarrow{\Delta t \rightarrow 0} \sqrt{(x'(t))^2 + (y'(t))^2} \quad (2.3)$$

Wir erhalten also als Ergebnis für die Bahngeschwindigkeit des Punktes

$$v(t) = \sqrt{(x'(t))^2 + (y'(t))^2} , \quad (2.4)$$

und wir behalten die Gleichung

$$v(t) = s'(t) \quad (2.5)$$

dabei im Hinterkopf. - Um ein einfaches Beispiel zu haben, berechnen wir die Bahngeschwindigkeit des Punktes $P(t) = (\cos(t)|\sin(t))$:

$$v(t) = \sqrt{(-\sin(t))^2 + (\cos(t))^2} = \sqrt{1} = 1$$

Der Punkt bewegt sich also mit konstanter Bahngeschwindigkeit auf dem Einheitskreis.

Über die Steigung der Tangente an eine solche Kurve im Punkt $P(t)$ brauchten wir dank Magnus' Ansatz und dank Antons Bruchrechnenkünsten nicht lange nachzudenken. Wir müssen im Kern $\frac{\Delta y}{\Delta x}$ bilden und Δt gegen 0 laufen lassen:

$$\frac{\Delta y}{\Delta x} = \frac{\frac{\Delta y}{\Delta t}}{\frac{\Delta x}{\Delta t}} \xrightarrow{\Delta t \rightarrow 0} \frac{y'(t)}{x'(t)}$$

Die Steigung $m(t)$ der Tangente an die Kurve im Punkt $P(t)$ ist also gegeben durch

$$m(t) = \frac{y'(t)}{x'(t)} . \quad (2.6)$$

Für $P(t) = (\cos(t)|\sin(t))$ erhalten wir

$$m(t) = \frac{\cos(t)}{-\sin(t)} ,$$

und das ist nichts Anderes als $-x(t)/y(t)$, wie es sein muss, denn $y(t)/x(t)$ ist ja die Steigung des Berührradius der Tangente.

Um die Lage des Punktes zu beschreiben, benutzen wir auch Polarkoordinaten, wenn das vorteilhaft ist; das will ich hier aber nur erwähnen.

Kapitel 3

Das Integral

3.1 Käferweglängen und Flächen unter der Kurve $y = v(t)$

Als wir die Bahngeschwindigkeit unseres Käfers in 2.1 bestimmt haben, haben wir den in einer sehr kleinen Zeitspanne Δt zurückgelegten Weg Δs berechnet und den Grenzwert für $\Delta t \rightarrow 0$ gebildet. Damit ist klar, dass die Bahngeschwindigkeit v gleich der Ableitung des zurückgelegten Weges s nach der Zeit t ist; das weißt du natürlich schon längst aus der Physik. Ich halte diese Tatsache in einer Gleichung fest:

$$v(t) = s'(t) \quad (3.1)$$

Wenn wir nun den zurückgelegten Weg $s(t)$ rekonstruieren wollen, brauchen wir uns ja im Prinzip bloß eine Stammfunktion V von v zu suchen und die so einzustellen, dass sie an der Stelle 0 auch den Wert 0 annimmt, dann haben wir $s(t)$. Dies können wir so aussprechen:

$$\text{es ist } s(t) = V(t) - V(0) \text{ für eine beliebige Stammfunktion } V \text{ von } v. \quad (3.2)$$

Wenn es dir nicht schon selbst aufgefallen ist, weise ich dich nun ausdrücklich darauf hin: Diese Lösungsmethode hatten wir an ganz anderer Stelle benutzt, nämlich um Raum- und Flächeninhalte zu berechnen. Schauen wir genau hin! Hätten wir uns die Aufgabe gestellt, den Inhalt des Flächenstücks zwischen der t -Achse und der Kurve $y = v(t)$ zwischen 0 und t zu berechnen¹, hätten wir doch als Ergebnis genau die gleiche Antwort gegeben:

$$\begin{aligned} \text{Inhalt des Flächenstücks zwischen } y = v(t) \text{ und } t\text{-Achse von } 0 \text{ bis } t = \\ V(t) - V(0) \text{ für eine beliebige Stammfunktion } V \text{ von } v. \end{aligned} \quad (3.3)$$

Sei dir der Bedeutung dieses Phänomens bewusst: Die Mathematik erzeugt Verwandtschaften zwischen Problemen, die äußerlich nichts miteinander zu tun haben, eine zentrale Leistung der Mathematik. Aber nun will ich die Katze aus dem Sack lassen und die Integralschreibweise einführen. Wir setzen

$$\int_a^b f(x) dx := F(x) \Big|_a^b := F(b) - F(a) \text{ für eine beliebige Stammfunktion } F \text{ von } f. \quad (3.4)$$

Wir lesen das als Integral von a bis b über f . Die Bedeutung des Integralbegriffs kannst du kaum überschätzen, du wirst noch viel damit zu tun haben, und dir wird gleich klar werden, warum Leibniz für das Integral gerade dieses Symbol geschaffen hat.

3.2 Riemannsche Summen

Für die Bahngeschwindigkeit unseres Käfers hatten wir so etwas wie

$$v(t) = \sqrt{1 + 4t^2}$$

¹Vorsicht, hier ist t in zwei Bedeutungen gebraucht, als Laufvariable und fester Zeitpunkt!

berechnet. Der in der Zeit von 0 bis 3 zurückgelegte Weg ist nach unserer Definition in Gleichung 3.4 auf Seite 23 durch

$$\int_0^3 \sqrt{1+4t^2} dt$$

gegeben. Nun ist eine Stammfunktion von $\sqrt{1+4t^2}$ alles andere als leicht zu finden, unsere Möglichkeiten übersteigt diese Aufgabe bei weitem. Dieser Fall tritt übrigens in der Praxis recht häufig auf. Man versucht dann, einen genügend guten Näherungswert auszurechnen, und dazu bildet man eine sogenannte **Riemannsche Summe**: Man teilt das Intervall $[0; 3]$ in n Teilintervalle auf. Die Breite des i -ten Teilintervalls bezeichnen wir mit $(\Delta t)_i$. In jedem der Teilintervalle wählt man nun einen t -Wert, der t -Wert im i -ten Teilintervall heie t_i . Die Riemannsche Summe ist dann

$$\sum_{i=1}^n v(t_i)(\Delta t)_i \quad . \quad (3.5)$$

Schaue dir den Ausdruck genau an! Erstens ist er der Schlüssel zum Integralsymbol: Das Summenzeichen wurde zum Integralzeichen designed, das dt erinnert an das $(\Delta t)_i$. Auf den zweiten Blick siehst du, dass der Ausdruck einen Näherungswert für das Integral liefern sollte. Der Summand $v(t_i)(\Delta t)_i$ ist das Produkt einer Geschwindigkeit und einer Zeitspanne, also die Länge eines Wegstücks. Man tut so, als sei die Geschwindigkeit unseres Käfers in der Zeitspanne $(\Delta t)_i$ konstant $v(t_i)$. Während einer kleinen Zeitspanne ändert sich die Geschwindigkeit des Käfers nur wenig, so dass dieser Ansatz nachvollziehbar ist. Insgesamt ersetzt man quasi die Kurve $y = v(t)$ durch eine Treppe.

3.3 Obersummen und Untersummen

Du sollst nun spezielle Riemannsche Summen kennen lernen, nämlich Ober- und Untersummen. Wir bleiben bei unserem Käferbeispiel, aber wir stellen uns die Aufgabe diesmal in der Flächeninhaltsversion. Dann gilt es, den Inhalt der „Fläche unter dem Graphen“ von v zwischen 0 und 3 näherungsweise zu berechnen. Dazu teilen wir das Flächenstück in Streifen ein und betrachten die sogenannte Untersumme und die sogenannte Obersumme (siehe Abbildung 3.1 auf Seite 25 und Abbildung 3.2 auf Seite 25).

Wir berechnen die Werte der Unter- und der Obersumme. Für die Untersumme erhalten wir mit der Abkürzung Δt für die Kästchenbreite $3/4$

$$U = v(0)\Delta t + v(\Delta t)\Delta t + v(2\Delta t)\Delta t + v(3\Delta t)\Delta t = \sum_{k=0}^3 v(k\Delta t)\Delta t$$

und für die Obersumme entsprechend

$$O = v(\Delta t)\Delta t + v(2\Delta t)\Delta t + v(3\Delta t)\Delta t + v(4\Delta t)\Delta t = \sum_{k=1}^4 v(k\Delta t)\Delta t$$

Diese Ausdrücke kannst du leicht auf den Fall n gleich breiter Streifen verallgemeinern. Übrigens kannst du dir die Werte bequem von Maple ausrechnen lassen. Wenn du das student-Paket geladen und die Funktion v definiert hast, wie es im Text zu Abbildung 3.1 auf Seite 25 beschrieben ist, lauten die entsprechenden Befehle für den vorliegenden Fall `leftsum(v(t),t=0..3)`; bzw. `rightsum(v(t),t=0..3)`; Die Werte bekommst du, indem du nach dem `leftsum`-Befehl das übliche `evalf(%)`; eingibst. Übrigens kannst du auch die Streifenzahl n vorgeben. Dann schreibst du mit einer konkreten Anzahl n die Befehle `leftsum(v(t),t=0..3,n)`; `evalf(%)`;

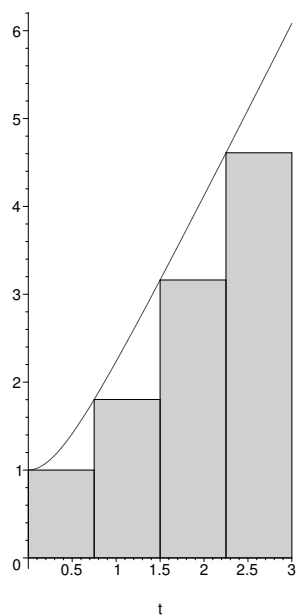


Abbildung 3.1: Untersumme mit vier Streifen zu $v(t) = \sqrt{1 + 4t^2}$. Das Bild erzeugst du leicht mit Maple. Definiere die Funktion: $v := t \rightarrow \text{sqrt}(1 + 4 * t * t)$; Lade das student-Paket: `with(student)`; Der plot-Befehl ist dann `leftbox(v(t), t = 0..3)`;

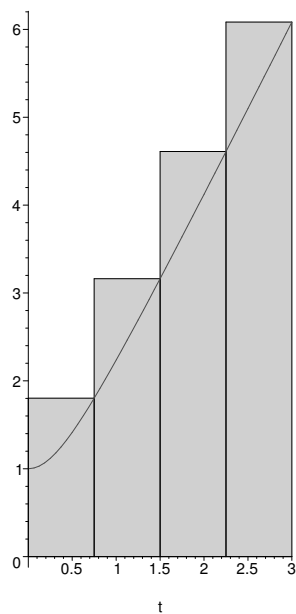


Abbildung 3.2: Obersumme mit vier Streifen zu $v(t) = \sqrt{1 + 4t^2}$. Die Maple-Befehle sind entsprechend denen im Text zu Abbildung 3.1, statt `leftbox` musst du `rightbox` verwenden.

Kapitel 4

Kurven in Parameterdarstellung (2)

4.1 Flächenstücke zwischen Kurve und Achse

Es sei durch $P(t) = (x(t), y(t))$ für $a \leq t \leq b$ eine Kurve in Parameterdarstellung gegeben. Wir suchen eine Möglichkeit, den Inhalt des Flächenstücks zwischen der Kurve und der x -Achse zu berechnen. Dazu teilen wir das Zeitintervall $[a; b]$ durch Teilpunkte $a = t_0 < t_1 < t_2 < \dots < t_n = b$ in n Teile. Die Kurvenpunkte $P(t_1), P(t_2) \dots$ teilen die Kurve in Abschnitte ein. Die Lote von den $P(t_i)$ auf die x -Achse zerlegen das Flächenstück in Streifen (siehe Abbildung 4.1 auf Seite 27), und der Inhalt des i -ten Streifens ist näherungsweise gegeben durch

$$y(t_i)(x(t_i) - x(t_{i-1})) \approx y(t_i)x'(t_i)(\Delta t)_i \quad .$$

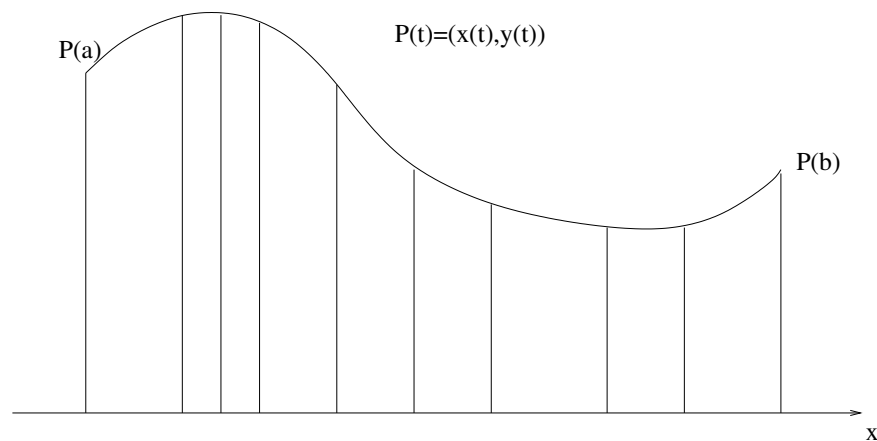


Abbildung 4.1: Zur Berechnung des Inhaltes der Fläche zwischen Kurve und x -Achse

Der Inhalt des betrachteten Flächenstücks ist demnach näherungsweise gegeben durch

$$\sum_{i=1}^n y(t_i)x'(t_i)(\Delta t)_i \quad ,$$

und das ist eine Riemannsche Summe für das Integral

$$\int_a^b y(t)x'(t) dt \quad . \tag{4.1}$$

Es ist somit vernünftig, den Wert des Integrals als Inhalt des Flächenstücks anzusehen. Dabei ist allerdings sowohl auf das Vorzeichen von $\Delta x \approx x'(t)\Delta t$ als auch auf das Vorzeichen von $y(t)$ zu achten. Das Integral liefert zum Beispiel einen positiven Wert, wenn der Punkt $P(t)$ sich oberhalb der x -Achse von links nach rechts oder wenn er sich unterhalb der x -Achse von rechts nach links bewegt.

Wenn durch $P(t) = (x(t), y(t))$, $a \leq t \leq b$, eine geschlossene Kurve gegeben ist, so dass der Punkt $P(t)$ ein einfaches Flächenstück einmal umläuft, ist der Inhalt dieses Flächenstücks (bis auf das Vorzeichen) durch das Integral in Gleichung 4.1 gegeben. Dies kann man anhand von Abbildung 4.2 auf Seite 28 einsehen. Der Punkt umlaufe das Flächenstück von A über B , C und D wieder zu A . Wir können die Parameterdarstellung so konstruiert annehmen, dass $a < t_B < t_C < t_D < b$ und $P(a) = A$, $P(t_B) = B$, $P(t_C) = C$, $P(t_D) = D$ und $P(b) = A$ ist. Dann ist

$$\int_a^b y(t)x'(t) dt = \int_a^c y(t)x'(t) dt + \int_c^b y(t)x'(t) dt .$$

Das erste Integral ist der (positive) Inhalt des Flächenstücks $APQCBA$. Das zweite Integral ist negativ, sein Betrag ist der Inhalt des Flächenstücks $CDAPQC$. Insgesamt erhalten wir eine negative Zahl, ihr Betrag ist der Inhalt des eingeschlossenen Flächenstücks. Damit du nicht zu unkritisch

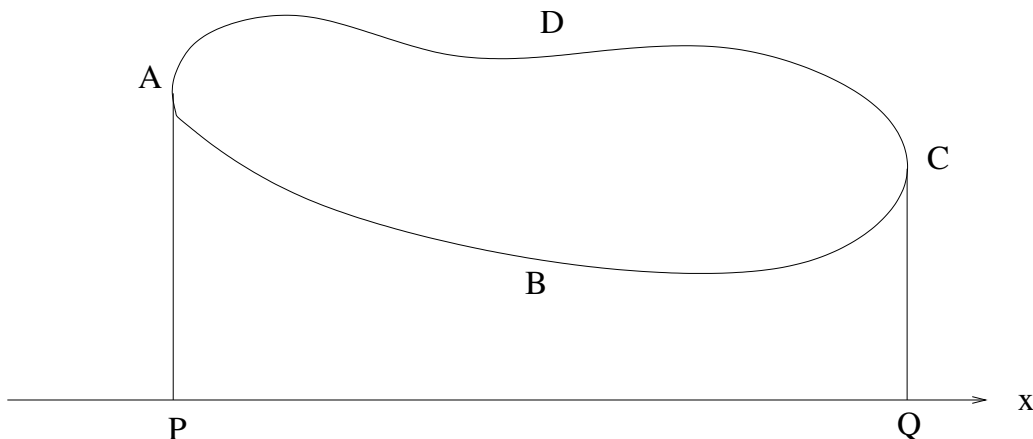


Abbildung 4.2: Zum Inhalt einer von einer Kurve eingeschlossenen Fläche

an diese Dinge herangehst, solltest du dich mit den beiden folgenden Aufgaben beschäftigen.

Aufgabe Untersuche die durch $P(t) = (\sin(2t), \sin(t))$, $0 \leq t \leq 2\pi$ gegebene geschlossene Kurve. Wie gross ist der Inhalt der eingeschlossenen Fläche? Was liefert das Integral in Gleichung 4.1 auf Seite 27?

Aufgabe Welchen Wert nimmt das Integral in Gleichung 4.1 an, wenn der Punkt das Flächenstück zweimal umläuft?

Natürlich kann man auch nach dem Inhalt der Fläche zwischen Kurve und y -Achse fragen, oder für eine geschlossene Kurve auf das Integral in Gleichung 4.1 partielle Integration anwenden:

$$\int_a^b y(t)x'(t) dt = y(t)x(t) \Big|_{t=a}^{t=b} - \int_a^b y'(t)x(t) dt$$

Für eine geschlossene Kurve ist $x(a) = x(b)$ und $y(a) = y(b)$, so dass der erste Summand nach dem Gleichheitszeichen wegfällt.

4.2 Die Leibnizsche Sektorformel

Eine Halbgerade beginne im Nullpunkt der xy -Ebene, und sie drehe sich gleichmäßig um den Nullpunkt. Ein Punkt, der sich zur Zeit t auf der Halbgeraden in der Entfernung $r(t)$ vom Nullpunkt

befindet, beschreibt dann eine ebene Kurve mit der Parameterdarstellung

$$P(t) = (r(t) \cos(t), r(t) \sin(t)) \quad . \quad (4.2)$$

Du kannst ruhig an einen Käfer denken, der sich auf einer sich drehenden Stange bewegt. Unsere Spiralkurven waren Beispiele solcher Kurven, für $r(t) = t$ etwa ist die Kurve eine Archimedische Spirale.

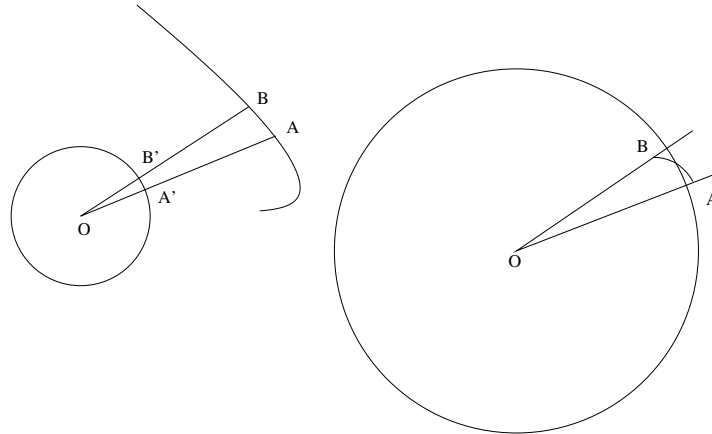


Abbildung 4.3: Zur Leibnizschen Sektorformel

Es sei nun $A = P(t)$ und $B = P(t + \Delta t)$ (siehe linke Figur in Abbildung 4.3 auf Seite 29). Die Punkte $A' = (\cos(t), \sin(t))$ und $B' = (\cos(t + \Delta t), \sin(t + \Delta t))$ sind die zugehörigen Punkte auf dem Einheitskreis um O . Wir fragen nun nach dem Inhalt ΔA der Fläche, die der Fahrstrahl von P in der Zeitspanne von t bis $t + \Delta t$ überstreicht. Stelle dir einen Kreis um O mit wachsendem Radius vor. Der Kreissektor, den die Halbgeraden OA und OB aus dem Kreis ausschneiden, hat anfangs einen kleineren Inhalt als ΔA , für große Radien aber einen größeren. Es gibt nun einen Kreis, für den der Inhalt des Sektors genau ΔA ist, und der Radius ist $r(z)$ für ein $z \in [t, t + \Delta t]$ (rechte Figur in Abbildung 4.3).¹

Der Anteil der Sektorfläche an der Kreisfläche ist gleich dem Anteil von Δt an 2π - der Bogen von A' bis B' auf dem Einheitskreis hat gerade die Länge Δt . Somit ist

$$\Delta A = \frac{1}{2} r(z)^2 \Delta t \quad \text{für ein } z \in [t, t + \Delta t] \quad ,$$

und daraus folgt

$$\frac{\Delta A}{\Delta t} = \frac{1}{2} r(z)^2 \xrightarrow{\Delta t \rightarrow 0} \frac{1}{2} r(t)^2 \quad .$$

Damit ist für den Inhalt $A(t)$ der vom Fahrstrahl überstrichenen Fläche gezeigt, dass $A'(t) = \frac{1}{2} r(t)^2$ ist. Demnach ist die vom Fahrstrahl in der Zeit von a bis b überstrichene Fläche gegeben durch das Integral

$$\int_a^b \frac{1}{2} r(t)^2 dt \quad . \quad (4.3)$$

Dies ist die Leibnizsche Sektorformel, auf die Leibniz sehr stolz gewesen sein soll. Wenn du sie benutzt, musst du im Auge behalten, dass sie mehrfach überstrichene Flächen auch mehrfach zählt.

Für die Archimedische Spirale mit $P(t) = (t \cos(t), t \sin(t))$ gibt die Sektorformel für die in der Zeit von 0 bis b vom Fahrstrahl überstrichene Fläche den Wert

$$\int_0^b \frac{1}{2} t^2 dt = \frac{1}{6} b^3$$

¹Natürlich darf der Käfer nicht die Fähigkeit besitzen, zu verschwinden und an einem anderen Ort wieder zu materialisieren!

an, und das sollte mit unserem früheren Ergebnis übereinstimmen.

Kapitel 5

Volumenbestimmung mit Hilfe von Integralen

5.1 Funktionen in zwei Variablen

Eine Funktion der Art $f : (x, y) \mapsto x^2 + \frac{1}{2}y^2$ ordnet jedem Paar (x, y) reeller Zahlen eine reelle Zahl $z = f(x, y)$ zu. Wenn wir das Konzept des Graphen einer Funktion auf eine solche Funktion verallgemeinern wollen, können wir als Definitionsbereich ein Stück der gewöhnlichen xy -Ebene nehmen und den Funktionswert z nach oben in Richtung einer dritten Achse antragen, die auf der xy -Ebene senkrecht steht. Der Graph der Funktion ist dann eine Fläche im Raum (siehe Abbildung 5.1). Wir fragen gleich nach dem Volumen des Körpers zwischen dem Graphen von f und der xy -

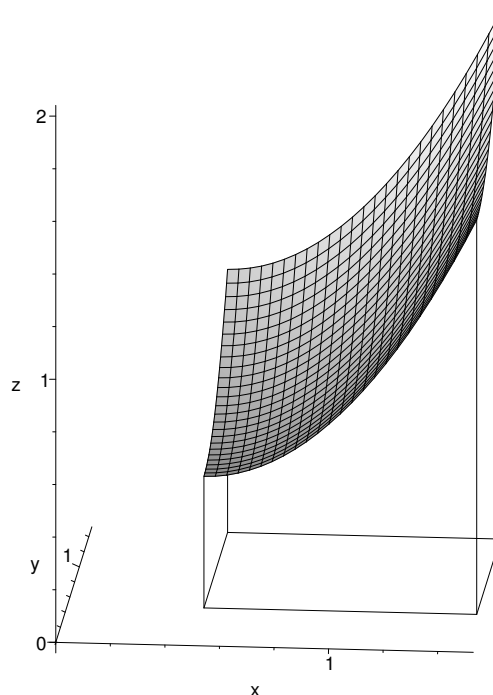


Abbildung 5.1: Ein Teil des Graphen der Funktion $f : (x, y) \mapsto (x - .5)^2 + (\frac{1}{2}(y - .5)^2 + .5$

Ebene über dem Rechteck der (x, y) mit $a \leq x \leq b$, $c \leq y \leq d$. Einen Näherungswert erhält man folgendermaßen: Man unterteilt das Rechteck durch Parallelen zur x - und zur y -Achse. Der i -te senkrechte und der j -te waagerechte Streifen überschneiden sich in einem kleinen Rechteck

der Länge $(\Delta x)_i$ und der Breite $(\Delta y)_j$. In diesem Rechteck wählen wir einen Punkt (x_i, y_j) . Ein Näherungswert für das Volumen ist dann gegeben durch

$$\sum_i \sum_j f(x_i, y_j) (\Delta y)_j (\Delta x)_i \quad . \quad (5.1)$$

Der Ausdruck erinnert nicht von ungefähr an eine Riemannsche Summe.

Wir greifen nun Jochens Vorschlag auf und unterteilen den Körper durch Ebenen senkrecht zur y -Achse in dünne Scheiben. Die j -te dieser Ebenen schneide die y -Achse beim Wert y_j für $j = 1, \dots, n$. Der Inhalt der j -ten Scheibe ist dann etwa der Inhalt der Schnittfläche der Ebene mit dem Körper multipliziert mit der Dicke $(\Delta y)_j$ der Scheibe. Aufsummiert über alle Scheiben ergibt das

$$\sum_{j=1}^n \left(\int_a^b f(x, y_j) dx \right) (\Delta y)_j \quad ,$$

und das ist eine Riemannsche Summe für

$$\int_c^d \left(\int_a^b f(x, y) dx \right) dy \quad . \quad (5.2)$$

Es ist vernünftig, das Doppelintegral in 5.2 als das Volumen des Körpers anzusehen (jedenfalls, wenn das Flächenstück über dem Rechteck ganz über der xy -Ebene liegt). Der Ausdruck in 5.1 ist gewissermaßen eine Riemannsche Summe für das Doppelintegral.

5.2 Das Volumen eines Kartoffelkörpers

Wir verallgemeinern dies auf den Fall eines Kartoffelkörpers im Raum. Wir zeichnen eine beliebige Gerade mit x -Skala als Achse aus. Stelle dir nun vor, eine Ebene, die auf der Geraden senkrecht steht, werde längs der Geraden verschoben. Wenn die Ebene an der Stelle x der Geraden ist, sei $F(x)$ der Inhalt der Schnittfläche der Ebene mit dem Kartoffelkörper (siehe Abbildung 5.2). Der

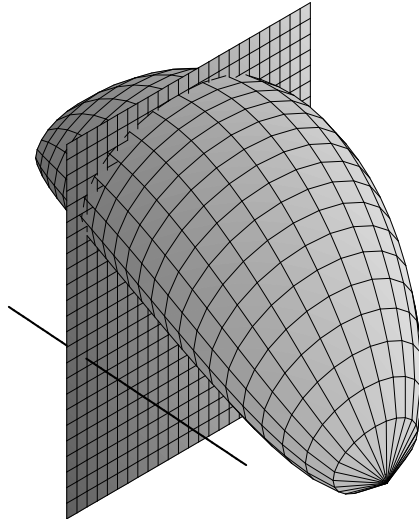


Abbildung 5.2: Die Ebene senkrecht zur Achse, die die Achse an der Stelle x schneidet, schneidet den Körper in einem Flächenstück des Inhalts $F(x)$

Körper liege ganz zwischen den Ebenen bei $x = a$ und bei $x = b$ mit $a < b$. Das Volumen des Körpers ist dann gegeben durch

$$V = \int_a^b F(x) dx \quad . \quad (5.3)$$

Du siehst leicht ein, dass das, was wir vorher machten, ein Spezialfall des jetzigen Ansatzes ist.

Man verwendet unser Verfahren zum Beispiel, um das Volumen von Rotationskörpern zu bestimmen. Als Achse wählt man natürlich die Rotationachse, die Schnittflächen sind dann immer Kreisscheiben.

5.3 Das Prinzip von Cavalieri

Wir betrachten zwei Kartoffelkörper, und wir bilden die Schnittflächen bezüglich einer gemeinsamen Achse. An der Stelle x der Achse sei der Inhalt der Schnittfläche der Messebene mit dem ersten Körper $F_1(x)$, mit dem zweiten Körper $F_2(x)$. Angenommen, es sei $F_1(x) = F_2(x)$ für alle x . Dann müssen die beiden Körper nach Gleichung 5.3 das gleiche Volumen haben.

Diese Tatsache ist als das Prinzip von Cavalieri bekannt. Man kann es benutzen, um das Volumen einer Halbkugel vom Radius R zu bestimmen. Die Achse stehe im Mittelpunkt der Halbkugel senkrecht auf der Grundfläche. Neben die Halbkugel stellt man den Restkörper, den man bekommt, wenn man einen Zylinder mit dem Radius R und der Höhe R von oben so ausbohrt, dass ein Kreiskegel vom Radius R und der Höhe R entfernt wird. Die beiden Körper haben in jeder Höhe gleich große Schnittflächen mit der entsprechenden Messebene. So hat meines Wissens Archimedes das Kugelvolumen bestimmt¹, und du magst die Methode in der zehnten Klasse gesehen haben. In unserem Buch ist sie auf Seite 240 beschrieben.

Zur Illustration gebe ich noch ein Beispiel. Es sei $r(t) := t^4 - \frac{2}{3}t^2 - \frac{1}{3}$. Dann ist durch

$$P(s, t) := (r(t) \cos(s) \mid r(t) \sin(s) \mid t) \quad \text{für } -1 \leq t \leq 1, 0 \leq s < 2\pi$$

eine Fläche im Raum gegeben, die einen Körper einschließt. Der Körper erstreckt sich von $z = -1$ bis $z = 1$. Für $-1 \leq z \leq 1$ ist die Schnittfläche des Körpers mit der Ebene parallel zur xy -Ebene in der Höhe z eine Kreisscheibe mit dem Radius $r(z)$. Wir verschieben nun den Körper um 2 in x -Richtung und verformen ihn zusätzlich, indem wir in der x -Koordinate noch einen Term hinzufügen:

$$Q(s, t) := (2 - t^2 + r(t) \cos(s) \mid r(t) \sin(s) \mid t) \quad \text{für } -1 \leq t \leq 1, 0 \leq s < 2\pi$$

Beide Körper sind in Abbildung 5.3 auf Seite 33 gezeichnet. Sie haben das gleiche Volumen! Ich

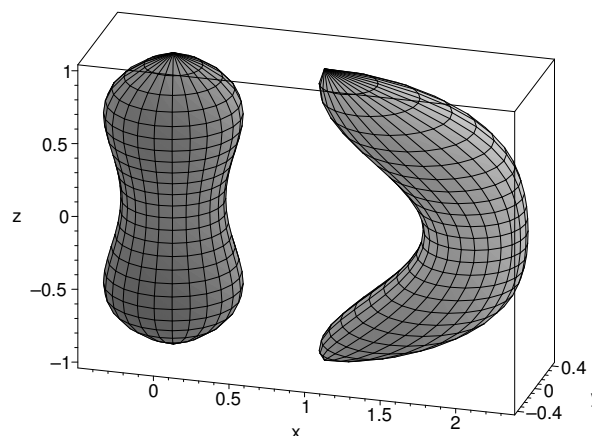


Abbildung 5.3: Zum Prinzip von Cavalieri

gebe noch die Maple-Eingabe an:

¹Eine bemerkenswerte Leistung!

```
r:=t->t^4-2*t^2/3-1/3;
plot3d( {[r(t)*cos(s),r(t)*sin(s),t], [2+t^2*cos(s),r(t)*sin(s),t]}, t=-1..1,
s=0..2*Pi, scaling=constrained, labels=[x,y,z]);
```

Kapitel 6

Kurven in Parameterdarstellung (3)

6.1 Ebene und Raum

Du hast nun schon praktische Erfahrungen mit geometrischen Objekten im Raum, und du bist mit dem kartesischen xyz -Koordinatensystem schon recht vertraut. Um die Lage eines Punktes A im Raum zu beschreiben, brauchst du halt drei statt zwei Koordinaten, und das meiste von dem, was wir in der Ebene angestellt haben, lässt sich leicht auf den Raum übertragen. Zum Beispiel ist die Länge der Strecke \overline{AB} mit den Endpunkten $A(a_1|a_2|a_3)$ und $B(b_1|b_2|b_3)$ durch

$$\sqrt{(b_1 - a_1)^2 + (b_2 - a_2)^2 + (b_3 - a_3)^2}$$

gegeben; dahinter steckt der Satz des Pythagoras, und du hast früher auf diese Weise die Länge der Raumdiagonalen eines Quaders mit den Kantenlängen $|b_1 - a_1|, |b_2 - a_2|, |b_3 - a_3|$ berechnet. Unsere alte xy -Ebene ist in den Raum auf natürliche Weise eingebettet; sie besteht aus allen Raumpunkten, deren z -Koordinate $= 0$ ist. Und für $a_3 = b_3 = 0$ wird aus unserer Entfernungsformel oben gleich wieder die Formel für die Länge einer Strecke in der Ebene.

Auch Parameterdarstellungen von Kurven im Raum können wir leicht bilden, indem wir eine zusätzliche dritte Koordinatenfunktion $z(t)$ einbringen. Aus einer ebenen Kurve

$$P(t) = (x(t)|y(t)) \quad \text{für } a \leq t \leq b$$

wird dann eine Raumkurve

$$Q(t) = (x(t)|y(t)|z(t)) \quad \text{für } a \leq t \leq b .$$

Die alte Kurve zeigt praktisch den Weg des Raumpunktes $Q(t)$ auf einer Landkarte, oder, anders gesagt, wenn du den Raum von oben mit Licht bescheinst, dessen Strahlen alle parallel zur z -Achse verlaufen, ist $P(t)$ der Schatten des Punktes $Q(t)$ in der xy -Ebene. Technisch gesprochen wird der Raum in z -Richtung auf die xy -Ebene projiziert. Du solltest dich nur nicht daran stoßen, dass dabei der Schatten eines Punktes auch zwischen Lichtquelle und Punkt liegen kann - nichts auf Erden ist vollkommen.

Es erhebt sich nun folgendes Problem. Wir haben gefragt, in welcher Richtung sich der Punkt $P(t)$ der Ebene zur Zeit t bewegt, und damals haben wir diese Richtung mit Hilfe der Steigung der Tangente beschrieben - siehe Gleichung 2.6 auf Seite 21. Dieses Konzept lässt sich nun nicht so gut auf den Raum übertragen, da brauchen wir etwas Neues, und das will ich jetzt angehen.

6.2 Vektoren

Der Begriff des Vektors gehört zum Kern der Linearen Algebra, die du in 13.1 lernen sollst. Ich werde die Bezeichnungen so bilden, dass es später nicht knirscht; sie wirken dadurch hier vielleicht

etwas künstlich.

Wir nehmen zwei Raumpunkte A und B und fragen, in welcher Richtung von A aus gesehen B liegt. Dazu brauchen wir doch nur anzugeben, wie weit man von A aus in x -, in y - und in z -Richtung laufen muss, um zu B zu gelangen, wir bilden also jeweils die Differenzen der Koordinaten (mit Vorzeichen!). Wenn wir die Koordinaten von A mit a_i und die von B mit b_i bezeichnen für $i = 1, 2, 3$, geben wir die gesuchte Richtung an durch

$$\overrightarrow{AB} := \begin{pmatrix} b_1 - a_1 \\ b_2 - a_2 \\ b_3 - a_3 \end{pmatrix} .$$

Denke getrost an einen Pfeil, der in A beginnt und in B endet! Nehmen wir einmal als Beispiel $A = (1|2|-1)$ und $B = (3|3|4)$. Dann ist

$$\overrightarrow{AB} = \begin{pmatrix} 2 \\ 1 \\ 5 \end{pmatrix} .$$

Eine kleine Zeichnung wird vielleicht hilfreich sein, siehe Abbildung 6.1.

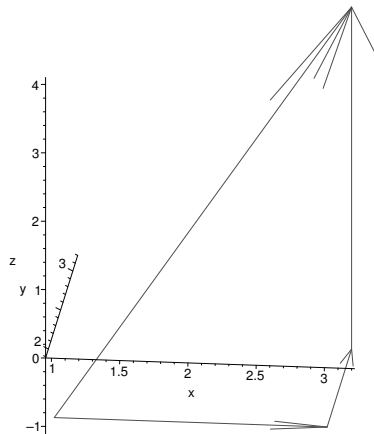


Abbildung 6.1: Beschreibung der Richtung vom Raumpunkt A zum Raumpunkt B durch einen Pfeil

Und was ist nun der Vektor? Nein, nicht der Pfeil. Der Vektor im Sinne der Linearen Algebra ist die Dreierspalte, und als Symbol verwendet man kleine Buchstaben mit Vektorpfeilen darüber, also hier meinetwegen

$$\vec{c} = \begin{pmatrix} 2 \\ 1 \\ 5 \end{pmatrix} .$$

Wie der Name Algebra schon verrät, rechnet man mit den Spalten. Man addiert sie komponentenweise, und man multipliziert sie mit einer Zahl, indem man jeden Eintrag mit dieser Zahl multipliziert. Ich zeige dir, wie es geht. Den Pfeil \overrightarrow{OA} vom Nullpunkt O zum Punkt A schreiben wir als \vec{a} und sprechen vom Ortsvektor von A . Dann ist

$$\vec{a} = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} , \quad \vec{b} = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} , \quad \vec{b} - \vec{a} = \begin{pmatrix} b_1 - a_1 \\ b_2 - a_2 \\ b_3 - a_3 \end{pmatrix} , \quad r\vec{a} = \begin{pmatrix} ra_1 \\ ra_2 \\ ra_3 \end{pmatrix} .$$

Vielleicht denkst du jetzt, in Wirklichkeit sei der Vektor eben doch der Pfeil, aber das solltest du nicht tun. Du kannst leicht andere Raumpunkte C und D finden, deren Pfeil \overrightarrow{CD} die gleiche

Spalte \vec{c} liefert wie die Punkte A und B oben. Die Pfeile \overrightarrow{CD} und \overrightarrow{AB} sind dann verschieden, sie sind allerdings gleich lang und parallel, so dass man sie in einen Topf wirft und sie dem gleichen Vektor \vec{c} zuordnet. Dass ich einfach $\vec{c} = \overrightarrow{AB}$ schreibe, ist praktisch, aber nicht ganz sauber. Eine korrektere Schreibweise wäre natürlich möglich, aber recht sperrig.

6.3 Geschwindigkeitsvektoren von Kurven

Nehmen wir uns nun eine Kurve $P(t) = (x(t)|y(t)|z(t))$ für $a \leq t \leq b$ her und betrachten wir die Punkte $P := P(t)$ und $Q := P(t + \Delta t)$. Für genügend kleine Δt ist die Änderung Δx der x -Koordinate praktisch $x'(t)\Delta t$. Für die übrigen Koordinaten gilt das entsprechend, so dass ich praktisch schreiben kann

$$\overrightarrow{PQ} = \begin{pmatrix} x'(t)\Delta t \\ y'(t)\Delta t \\ z'(t)\Delta t \end{pmatrix} = \Delta t \begin{pmatrix} x'(t) \\ y'(t) \\ z'(t) \end{pmatrix} .$$

Der Vektor

$$\vec{v}(t) := \begin{pmatrix} x'(t) \\ y'(t) \\ z'(t) \end{pmatrix} \quad (6.1)$$

ist der Geschwindigkeitsvektor der Kurve im Punkt $P(t)$. Heftest du in $P(t)$ den Pfeil an, der zu v gehört, liegt er auf der Tangente an die Kurve in $P(t)$. Die Länge des Pfeils ist die Bahngeschwindigkeit des Punktes zur Zeit t ! In Abbildung 6.2 findest du eine Archimedische Spirale mit einigen angehefteten Geschwindigkeitsvektoren.

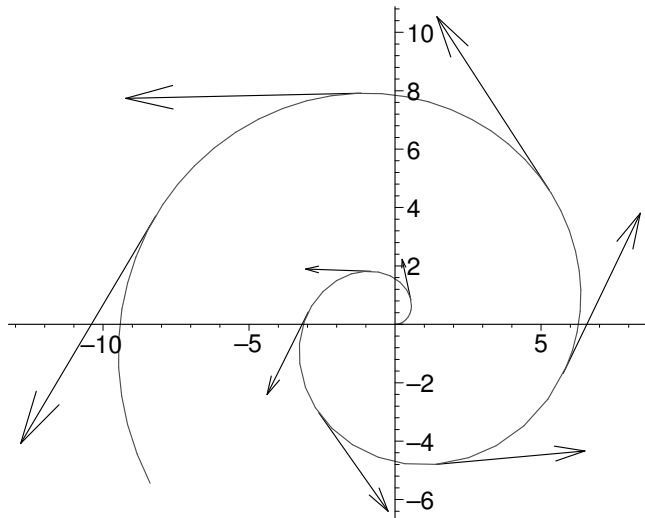


Abbildung 6.2: Eine Archimedische Spirale mit einigen Geschwindigkeitsvektoren

6.4 Winkel

Eigentlich hatte ich schreiben wollen „Winkel zwischen Vektoren“, aber das hätte dich vielleicht abgeschreckt, denn Vektoren sind ja Spalten mit Einträgen aus \mathbb{R} . Dennoch kann man von dem Winkel sprechen, den zwei Vektoren \vec{a} und \vec{b} einschließen, und das ist schlicht der Winkel zwischen den Ortsvektoren, also zwischen den Pfeilen, die im Nullpunkt beginnen und bei A bzw. B enden. Zugriff auf diesen Winkel haben wir durch den Kosinussatz, denn die Seitenlängen im Dreieck OAB können wir ja mühelos berechnen.

Als wir in den Kosinussatz

$$c^2 = a^2 + b^2 - 2ab \cos(\gamma)$$

die entsprechenden Terme $a = |\vec{a}|$, $b = |\vec{b}|$ und $c = |\vec{b} - \vec{a}|$ eingesetzt und anschließend vereinfacht hatten, erhielten wir

$$a_1 b_1 + a_2 b_2 + a_3 b_3 = |\vec{a}| \cdot |\vec{b}| \cos(\gamma) .$$

Dem Ausdruck auf der linken Seite sieht man seine Nützlichkeit sofort an, deshalb bekommt er ein eigenes Symbol. Wir setzen für $\vec{a}, \vec{b} \in \mathbb{R}^n$

$$\vec{a} * \vec{b} := \sum_{i=1}^n a_i b_i \quad (6.2)$$

und nennen dies Produkt das Skalarprodukt oder das innere Produkt der Vektoren \vec{a} und \vec{b} . Das konnte ich getrost gleich für den Raum \mathbb{R}^n der n -Spalten reeller Zahlen erledigen, das n ist dabei völlig egal. Freilich muss ich die Vektoren aus dem gleichen Raum nehmen, eine Zweierspalte kann ich nicht mit einer Dreierspalte multiplizieren. Übrigens ist mit dem neuen Symbol

$$|\vec{a}| = \sqrt{\vec{a} * \vec{a}} . \quad (6.3)$$

Mit Hilfe des inneren Produkts haben wir Zugriff auf Längen und Winkel im \mathbb{R}^n , wir werden also später (in bescheidenem Umfang) Geometrie im n -dimensionalen Raum treiben können.

Hier und heute benutzen wir unsere neuen Begriffe aber nur, um den Schnittwinkel zweier Raumkurven bestimmen zu können. Dazu bilden wir einfach die Geschwindigkeitsvektoren im Schnittpunkt und dann den Winkel zwischen denen, fertig.

Dazu ein Beispiel. Wir nehmen die Hutfläche, die entsteht, wenn wir den Graphen von $z = e^{-x^2}$ um die z -Achse rotieren lassen (siehe Abbildung 6.3). Dies ist, wie wir uns überlegt haben, der Graph der Funktion

$$f := (x, y) \rightarrow e^{-(x^2+y^2)} .$$

Wenn wir uns die Fläche zeichnen lassen, benutzt Maple zwei Scharen von Kurven. Wir wählen

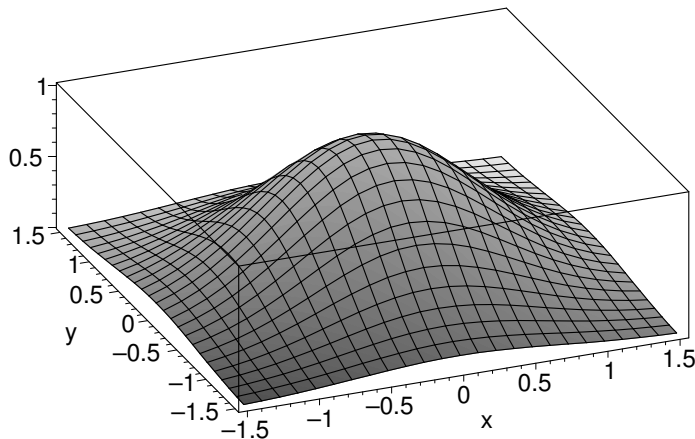


Abbildung 6.3: Der Graph von $f := (x, y) \rightarrow \exp(-(x^2 + y^2))$

zwei davon, nämlich

$$\vec{p}(t) = \begin{pmatrix} t \\ -1 \\ f(t, -1) \end{pmatrix} \quad \text{und} \quad \vec{q}(s) = \begin{pmatrix} 1/2 \\ s \\ f(1/2, s) \end{pmatrix} .$$

Sie schneiden sich im Punkt $S(1/2|-1|f(1/2, -1))$. Die Geschwindigkeitsvektoren in diesem Punkt sind

$$\vec{p}'(1/2) = \begin{pmatrix} 1 \\ 0 \\ -e^{-5/4} \end{pmatrix} \quad \text{und} \quad \vec{q}'(-1) = \begin{pmatrix} 0 \\ 1 \\ 2e^{-5/4} \end{pmatrix} .$$

Für den Winkel bekommen wir dann

$$\arccos\left(\frac{\vec{p}'(1/2) * \vec{q}'(-1)}{|\vec{p}'(1/2)| \cdot |\vec{q}'(-1)|}\right) = \frac{-2e^{-5/2}}{\sqrt{1+e^{-5/2}}\sqrt{1+4e^{-5/2}}} \approx 1.70816 \approx 97.87^\circ .$$

Vielleicht hastest du erwartet, dass die Kurven im Schnittpunkt orthogonal sind, aber das stimmt nicht.

6.5 Übungen

1. Es sei $f(x, y) = x^2 + y^2 + 1$. Den Graphen der Funktion kennst du recht gut. Wir nehmen die Kurven

$$P(t) = (t|-1|f(t, -1)) \quad \text{für } -1 \leq t \leq 3 \quad \text{und} \quad Q(s) = (2|s|f(2, s)) \quad \text{für } -2 \leq s \leq 0 .$$

Offensichtlich verlaufen beide Kurven im Graphen von f . Berechne die Geschwindigkeitsvektoren $\vec{p}'(t)$ und $\vec{q}'(s)$. Welchen Winkel bilden sie im Schnittpunkt der Kurven? Zeichne auch die von P beschriebene Kurve mit den Geschwindigkeitsvektoren in den Punkten mit ganzzahligem t .

2. Geschwindigkeitsvektoren kannst du auch von ebenen Kurven bilden, da haben sie eben nur zwei Einträge. Untersuche die logarithmische Spirale

$$P(t) = (e^t \cos(6t)|e^t \sin(6t)) \quad \text{für } 0 \leq t \leq 3 .$$

Skizziere die Kurve, berechne den Geschwindigkeitsvektor $\vec{p}'(t)$ und den Winkel, den er mit dem Ortsvektor $\vec{p}(t)$ seines Punktes bildet.

Kapitel 7

Klausuren

7.1 Erste Klausur 12.1

1. Aufgabe

Es sei $f(x) = x^3 + 3x^2$. Berechne die Nullstellen und die Koordinaten der Extrempunkte des Graphen, zeichne eine flüchtige Skizze und berechne den Inhalt des Flächenstücks, das Graph und x -Achse einschließen.

2. Aufgabe

Jutta hat

$$\int_0^1 (5x^4 - 8x^3 + 1) dx = 0$$

ausgerechnet. Ich meine, es stimmt. Rechne du das noch einmal nach und erkläre ihr, wie es sein kann, dass der Wert des Integrals 0 ist.

3. Aufgabe

Berechne die Ableitungen mit Hilfe unserer Regeln. Vereinfache die Ergebnisse **nicht**.

$$\frac{\sin(x)}{x} \quad xe^{-x^2} \quad x \ln(x^2)$$

4. Aufgabe

Es sei

$$p(x) = 1 + x + \frac{1}{2}x^2 + \frac{1}{6}x^3 + \frac{1}{24}x^4 + \frac{1}{120}x^5 .$$

Berechne $p'(x)$ und eine Stammfunktion $P(x)$. Schön wäre, es fele dir etwas auf. Notiere deinen Einfall.

5. Aufgabe

Begründe mit Hilfe eines gleichschenkligen rechtwinkligen Dreiecks, dass

$$\sin\left(\frac{\pi}{4}\right) = \cos\left(\frac{\pi}{4}\right) = \frac{1}{\sqrt{2}}$$

ist, und bilde mit Hilfe dieser Werte ein Taylor-Polynom vom Grad 2. Wozu ist das gut?

6. Aufgabe

Wie sieht die durch die Parameterdarstellung

$$Q(t) = \left(\frac{1}{t} \cos(t) \mid \frac{1}{t} \sin(t) \right) \quad \text{für } 2\pi \leq t < \infty$$

gegebene Kurve aus? Beschreibe!

7. Aufgabe

Abbildung 7.1 auf Seite 42 zeigt einen Käferweg. Beantworte dazu die folgenden Fragen.

- Wo ist Startpunkt, wo ist Zielpunkt, und in welcher Richtung wird die Kurve durchlaufen?
- Unter welchem Winkel überquert der Käfer beim ersten Mal die x -Achse? (Mache mit Hilfe einer Skizze deutlich, welchen Winkel du berechnest!)
- Schreibe einen Term für die Bahngeschwindigkeit hin.

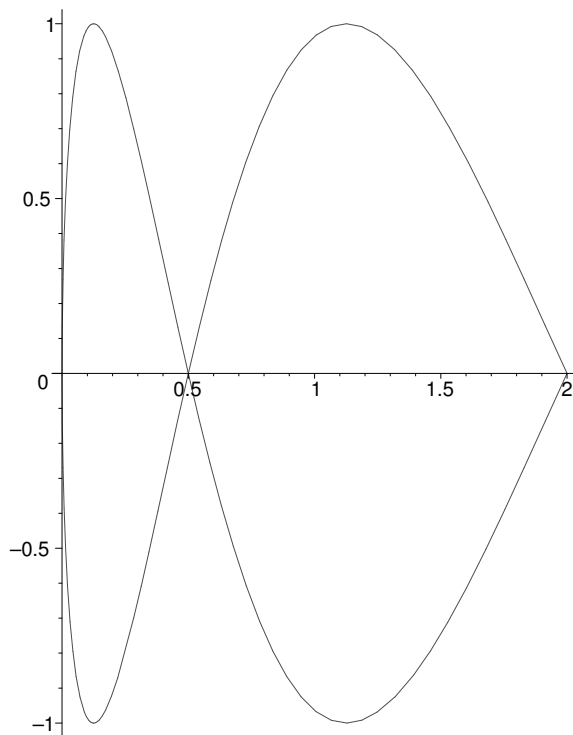


Abbildung 7.1: $P(t) = (2t^2 |\sin(2\pi t)|)$, $-1 \leq t \leq 1$

8. Offener Forschungsauftrag

Kann man zwei verschiedene Punkte P und Q im Koordinatensystem so durch eine Kurve verbinden, dass keine Tangente an die Kurve parallel zu PQ ist? Untersuche das mal.

7.2 Zweite Klausur 12.1

Hinweis: Oft frage ich nach einem Rechenausdruck für eine Größe. Schreibe dann einen Term in annehmbarer Form hin, und werte den nicht aus!

1. Aufgabe

- (a) Berechne die folgenden Integrale.

$$\int_0^{\pi} (x + \cos(x)) dx \quad \int_{-1}^0 (t+1)e^{-t} dt \quad \int_0^1 (10x-2)^7 dx \quad \int_0^1 \int_0^x xy dy dx$$

- (b) Berechne das Volumen des Körpers zwischen der xy -Ebene und der durch

$$z = f(x, y) = xy + y^2 \text{ für } 1 \leq x \leq 2 \text{ und } 0 \leq y \leq 1$$

gegebenen Fläche.

- (c) Das Integral $\int_0^1 \cos(x^2) dx$ versuchst du besser nicht auszurechnen. Gib einen brauchbaren Näherungswert dafür an (Streifenmethode, vier Streifen).

2. Es sei M das Flächenstück, das von den Kurven mit den Gleichungen $y = x$ und $y = x^2$ eingeschlossen wird.

- (a) Zeichne ein Bild von M im Koordinatensystem und berechne den Inhalt von M .
 (b) Wir definieren durch $z = f(x, y) = (x - y)(y - x^2)$ eine Funktion. Begründe, dass diese Funktion auf dem Rand von M den Wert 0 und im Innern von M positive Werte annimmt.
 (c) Schreibe einen Rechenausdruck für das Volumen des Körpers hin, der M als Grundfläche hat und der nach oben durch den Graphen von f begrenzt wird. **Nicht ausrechnen!**

3. Wir lassen das durch $z = (x - 2)^2$ für $0 \leq x \leq 2$ gegebene Kurvenstück um die z -Achse rotieren. Die entstandene Fläche schließt mit der xy -Ebene einen Körper ein. Zeichne eine Skizze des Körpers und gib einen Rechenausdruck für sein Volumen an. **Nicht ausrechnen!**

4. Raumkurven

- (a) Beschreibe die durch

$$P(t) = (t \cos(t), t \sin(t), t) \text{ für } 0 \leq t \leq 2\pi \text{ und } Q(t) = (t, t^2, t(1-t)) \text{ für } 0 \leq t \leq 1$$

gegebenen Raumkurven.

- (b) Gib eine Parameterdarstellung einer Schraubenlinie an, die sich im Abstand 2 um die x -Achse windet.

5. Ebene Kurven

- (a) Bestimme a und b so, dass der Punkt $P(t) = (t^2, t - t^3)$ gerade die in Abbildung 7.2 auf Seite 44 gezeichnete Kurve durchläuft, und gib einen Rechenausdruck für den Inhalt des umflogenen Flächenstücks an. **Nicht ausrechnen!**
 (b) Es sei $Q(t) = (\sqrt{t} \cos(t), \sqrt{t} \sin(t))$ für $0 \leq t < \infty$. Beschreibe die Kurve und gib einen Rechenausdruck für den Inhalt der Fläche an, die der Fahrstrahl in der Zeit von 0 bis 1 überstreicht.

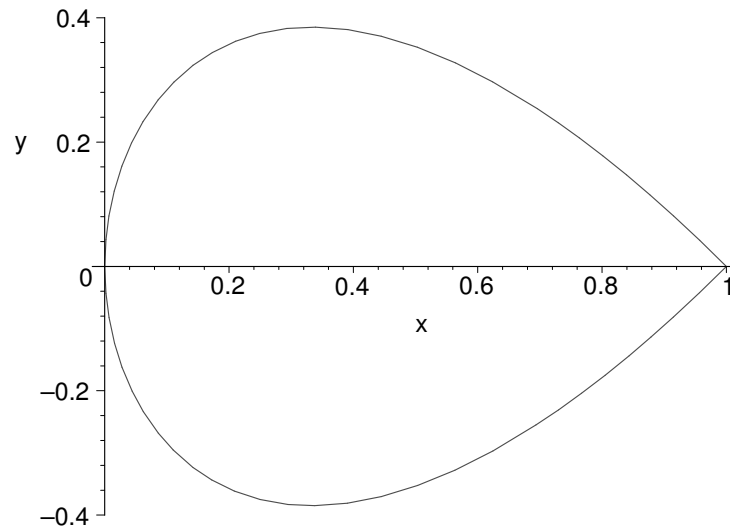


Abbildung 7.2: Zur Aufgabe zu ebenen Kurven der zweiten Klausur

7.3 Nachschreibausgabe der zweiten Klausur

1. Berechne.

$$\int_1^2 (5x-1)^7 dx \quad \int_0^\pi x \sin(3x) dx \quad \int_0^3 te^{t^2} dt \quad \int_0^1 \int_{x^2}^{\sqrt{x}} y dy dx$$

2. Das Kartoffelkörperprinzip

- (a) Schreibe eine Anleitung auf, wie man das Volumen eines Kartoffelkörpers bestimmt, und berechne damit das Volumen des Körpers, der von der durch

$$P(s, t) = (t^2 + t \cos(s), t + t \sin(s), t^3) \quad \text{für } 0 \leq t \leq 1, 0 \leq s < 2\pi$$

gegebenen Fläche und der zur xy -Ebene parallelen Ebene in der Höhe $z = 1$ eingeschlossen ist. Beschreibe diesen Körper auch! (Hinweis: Wenn du mit dem Körper nicht zurechtkommst, baust du dir eben selbst einen geeigneten Körper.)

- (b) Wenn man das Kartoffelkörperprinzip eine Dimension tiefer formuliert, hat man eine Methode, den Inhalt der Fläche zwischen zwei Graphen zu berechnen. Klar? Führe das am Beispiel der Fläche zwischen den Graphen mit den Gleichungen $y = x^2 - 4$ und $y = 2x - 2$ vor.
- (c) Hinz, ein großer Freund des Doppelintegrals, will seinem Freund Kunz das Kartoffelkörperprinzip madig machen. Er meint, es könne nichts Einfacheres geben als

$$\int_a^b \int_c^d f(x, y) dx dy .$$

„Quark“, knurrt Kunz, „das ist doch nur ein harmloser Spezialfall des Kartoffelkörperprinzips.“ Erkläre, was Hinzens Ausdruck berechnet, und entscheide den Streit der beiden.

3. In der xy -Ebene liegt die Grundfläche einer quadratischen Säule. Ihre Eckpunkte sind die Punkte $A(-1|-1)$, $B(1|-1)$, $C(1|1)$ und $D(-1|1)$. Nach oben ist sie begrenzt durch die Fläche mit der Gleichung $z = 3 + x^2 + y^2$. Zeichne ein Bild des Körpers und berechne sein Volumen.

4. Zeichne die durch

$$P(t) = (t^2 \cos(2\pi t) | t^2 \sin(2\pi t)), \quad 0 \leq t \leq 1$$

gegebene Kurve, berechne den Inhalt der vom Fahrstrahl überstrichenen Fläche und gib einen Term für die Bahngeschwindigkeit an.

Kapitel 8

Fragen

1. Wir gehen bei der Bestimmung von Flächen- und Rauminhalten immer von einer Funktion A aus, leiten die ab und schließen aus der Ableitung A' auf A zurück. Ist denn sicher, dass es diese Funktion wirklich gibt? Oder reden wir möglicherweise über ein Phantom?
2. Wenn $F'(x) = f(x)$ ist, dann ist auch $(F(x) + c)' = f(x)$ für jedes $c \in \mathbb{R}$. Gibt es außer diesen $F(x) + c$ noch weitere Funktionen, deren Ableitung $= f(x)$ ist?
3. Kann man etwas über die Größe des Fehlers der quadratischen usw. Näherung sagen?
4. Geht unsere Argumentation, dass $\Delta y = f(z)\Delta x$ ist für ein z zwischen x und $x + \Delta x$, auch für komplizierte Funktionen?
5. Hat die Spirale aus der Klausur endliche Länge?
6. Was bringen eigentlich Ober- und Untersummen?

Teil II

Stochastik

Kapitel 9

Einige Beispiele zur Einführung

Das Thema des Halbjahrs heißt Stochastik, es handelt sich um die Mathematik des Zufalls. Das klingt merkwürdig: Was hat Zufall in der Mathematik zu suchen? In der Tat ist der Zufall ein Phänomen unseres Alltags, und (idealisierte) Beispiele aus dem Alltag wollen wir zum Ausgangspunkt unserer Untersuchungen machen.

Der erste Begriff, den ich nenne, ist der Begriff der **Zufallsvariablen** oder Zufallsgröße. Ich mache an einigen Beispielen deutlich, um was es geht. Eine Zufallsvariable ist zum Beispiel

- die Lebensdauer einer Glühbirne aus einer bestimmten Produktion
- die Anzahl der Zerfälle in einer Probe einer radioaktiven Substanz in der nächsten Sekunde
- der Zeitpunkt des Todes eines bestimmten Menschen
- die Anzahl der Autos, die in der nächsten halben Stunde aus der Isenstedter Straße in die Kantstrasse einbiegen
- die Zeit, bis Frau Schulzes Telefon im Sekretariat das nächste Mal klingelt, und
- die Dauer des anschließenden Telefongesprächs
- der Wert, den ein Glücksrad liefert
- die nächste mit einem Würfel gewürfelte Zahl
- die Summe der Lottozahlen der nächsten Ziehung

Bei all diesen Beispielen läuft ein Vorgang ab und das Ergebnis ist eine gewisse Zahl aus einem Wertevorrat. Nun bekommst du gleich ein paar Begriffe: Der Wertevorrat ist eine Menge Ω ; man sagt, man habe es jeweils mit einer Zufallsvariablen X zu tun, die Werte aus Ω annimmt. Ergebnisse konkreter Durchführungen werden mit kleinen Buchstaben x, x_1, x_2, \dots bezeichnet, und man sagt, so ein x_1 sei ausgelost worden.

9.1 Ein konkretes Beispiel

Das folgende Beispiel ist recht künstlich, aber wir kontrollieren es vollständig. Mit Hilfe eines geeigneten Glücksrads bestimmen wir eine Zahl z , $0 < z \leq 3$. Unsere Zufallsvariable X ist das Quadrat dieser Zahl. Dann ist also $\Omega =]0; 9]$, aber das tut nicht viel zur Sache.

Mit Hilfe des Befehls `rand()` können wir das Drehen des Glücksrads und das Ablesen bequem vom Rechner erledigen lassen, er stellt uns sogar die Ergebnisse grafisch dar (siehe Abbildung 9.1). Im ersten Bild siehst du die Ergebnisse von zehn Versuchen. Anschließend haben wir für $n = 10, n = 100$ und für $n = 1000$ zehnmal den Versuch jeweils n -mal durchgeführt und den

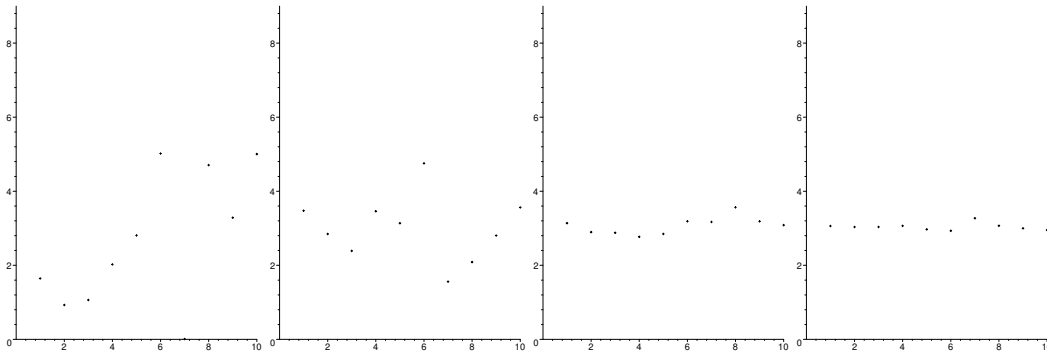


Abbildung 9.1: Zehn ausgeloste Werte (links), zehn Mittelwerte von jeweils zehn ausgelosten Werten (2. von links), zehn Mittelwerte von jeweils hundert ausgelosten Werten (3. von links) und zehn Mittelwerte von jeweils tausend ausgelosten Werten (rechts)

Mittelwert der n Werte gebildet. Die Zehnermittelwerte siehst du im zweiten Bild, die Hunderter- und die Tausendermittelwerte im dritten und vierten Bild.

Wenn du die Grafiken ansiehst, wirst du feststellen, dass die Mittelwerte mit wachsendem n immer weniger schwanken, und damit sind wir schon beim Kern der Sache angelangt: Mögen die Einzelwerte auch schwanken, die Mittelwerte genügend großer Anzahlen von Einzelwerten sind in der Regel recht stabil und damit gut vorhersagbar. Dieser Befund ist im **empirischen Gesetz der großen Zahl** notiert: Lost man Werte einer Zufallsvariablen aus und bildet den Mittelwert der Werte, nähert der sich bei steigender Versuchszahl in der Regel einer festen Zahl.

Bei unserem Beispiel können wir diesen geheimnisvollen Idealwert sogar angeben. Wie groß ist x^2 für $0 < x \leq 3$ im Mittel? Wie hoch liegen die Parabelpunkte im Durchschnitt? Auf Jochens Vorschlag hin haben wir aus dem Flächenstück unter der Parabel zwischen 0 und 3 ein Rechteck mit der Breite 3 gemacht, seine Höhe ist

$$\frac{1}{3} \int_0^3 x^2 dx = 3 \quad .$$

9.2 Aufgaben

1. Mit welcher Wahrscheinlichkeit nimmt die Zufallsvariable X des vorigen Abschnitts einen Wert zwischen 1 und 2 an? Und was bedeutet diese Aussage überhaupt?
2. Mit $P(X < x)$ bezeichnen wir die Wahrscheinlichkeit, dass die Zufallsvariable X einen kleineren Wert als x annimmt. Offensichtlich ist dies eine Funktion von x , man nennt sie die **Verteilungsfunktion** von X . Bestimme sie für das X des vorigen Abschnitts und zeichne ihren Graphen.
3. Wenn du nun mit dem Glücksrad einen Wert z , $0 < z \leq 1$ bestimmst und dessen Kehrwert bildest, hast du wieder eine Zufallsvariable. Untersuche sie!

9.3 Die Verteilungsfunktion einer Zufallsvariablen

Bei der Besprechung der Aufgaben und weiterer Beispiele haben wir ganz unbefangen von der Wahrscheinlichkeit $P(a \leq X < b)$ gesprochen, dass unsere Zufallsvariable X einen Wert zwischen a und b annimmt. Wir greifen diese Sprechweise nun auf und nennen die Funktion

$$F := x \mapsto P(X < x) \quad \text{für } x \in \mathbb{R} \quad (9.1)$$

die **Verteilungsfunktion** der Zufallsvariablen X . Offensichtlich ist F monoton steigend, die Werte liegen zwischen 0 und 1, und es gilt $F(x) \rightarrow 1$ für $x \rightarrow \infty$ und $F(x) \rightarrow 0$ für $x \rightarrow -\infty$.

Zunächst rechnen wir $F(x)$ für das Einführungsbeispiel in Abschnitt 9.1 aus: die dort definierte Zufallsvariable X nimmt Werte zwischen 0 und 9 an. Für ein x zwischen 0 und 9 ist $P(X < x)$ gleich der Wahrscheinlichkeit, dass das Glücksrad ein z zwischen 0 und \sqrt{x} liefert, und das ist $\sqrt{x}/3$. Also ist

$$F(x) = \begin{cases} 0 & \text{für } x < 0 \\ \frac{1}{3}\sqrt{x} & \text{für } 0 \leq x < 9 \\ 1 & \text{für } 9 \leq x \end{cases} .$$

Wer dieses F hat, kontrolliert die Zufallsvariable vollständig. Zum Beispiel ist

$$P(1 \leq X < 2) = F(2) - F(1) = \frac{1}{3}\sqrt{2} - \frac{1}{3} .$$

Und das bedeutet folgendes: Wenn man sehr viele Werte von X auslost, sagen wir N Stück, dann ist der Anteil der Werte, die zwischen 1 und 2 liegen, ziemlich genau $1/3(\sqrt{2} - 1)$. Du kannst das nachprüfen, indem du eine geeignete Simulation durchführst; das solltest du schon allein können.

9.4 Verteilungsfunktion durch Simulation

Ich will nun zeigen, wie man sich dem F empirisch nähert. Das ist natürlich wichtig für Zufallsvariable von der Art „Dauer von Frau Schulzes nächstem Telefongespräch“, die man theoretisch kaum in den Griff bekommt. Dazu lose ich eine große Zahl von Werten von X aus, sagen wir wieder N , und ordne sie der Größe nach. Als Schätzwert für $P(X < x)$ nehme ich dann den Anteil der Werte der Liste, die kleiner als x sind. Wenn $x = x_i$ der i -te Wert der geordneten Liste ist, ist dieser Anteil einfach i/N (eigentlich $(i - 1)/N$, aber das spielt keine Rolle). Abbildung 9.2 zeigt Ergebnisse für $N = 10$, $N = 100$ und $N = 1000$. Natürlich brächte eine erneute Durchführung der Simulation nicht die gleichen Ergebnisse - das ist Stochastik! Meine Prognose: Für $N = 10$ gäbe es deutliche, für $N = 100$ weniger deutliche Abweichungen, aber für $N = 1000$ erhielten wir im wesentlichen die gleiche Kurve. Probiere es aus!

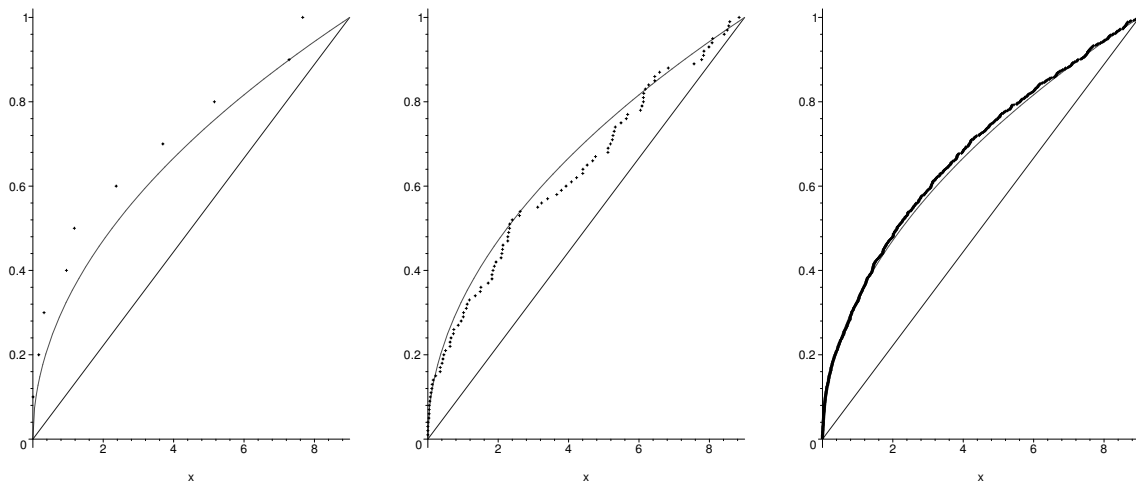


Abbildung 9.2: Empirisch ermittelte Punktvolken für den Graphen der Verteilungsfunktion von X mit $N = 10$ (links), $N = 100$ (Mitte) und $N = 1000$ (rechts). Außer der Punktvolke zeigen die Bilder die Strecke, die den Nullpunkt mit $(9|1)$ verbindet, und die theoretische Kurve.

9.5 Die Dichtefunktion einer Zufallsvariablen

Ob unser Glücksrad, das eine Zahl zwischen 0 und 3 auslost, einmal genau die Zahl 2 ausgibt? Oder genau die Zahl $\sqrt{5}$? Möglich, aber die Wahrscheinlichkeit dafür ist 0. Man sagt, diese Ereignisse

seien „fast unmöglich“. Genauso ist die Wahrscheinlichkeit, dass die 2 nicht ausgelost wird, 1, aber es könnte trotzdem anders kommen. Deshalb sagt man hier, es sei „fast sicher“, dass die 2 nicht kommt. Auch für unsere Beispielzufallsvariable X aus 9.1 ist $P(X = 1) = P(X = 8) = 0$. Dennoch scheinen Werte in der Nähe von 1 irgendwie wahrscheinlicher zu sein als Werte in der Nähe von 8, und das hat mit der Steigung des Graphen von F zu tun.

Wir wollen dies nun präziser formulieren. Wenn die Verteilungsfunktion F auf einem Intervall differenzierbar ist, so wie unser F auf dem Intervall $]0; 9[$, dann haben wir folgende Möglichkeiten. Zum einen können wir Wahrscheinlichkeiten als Integrale über die Ableitungsfunktion $f := F'$ der Verteilungsfunktion schreiben:

$$P(a \leq X < b) = F(b) - F(a) = \int_a^b f(x) dx \quad (9.2)$$

Zum zweiten ist für kleine Δx die Wahrscheinlichkeit, dass X einen Wert in einem Stückchen der Länge Δx um einen Wert x annimmt, ungefähr gleich $f(x)\Delta x$:

$$P(x \leq X < x + \Delta x) \approx f(x)\Delta x \quad (9.3)$$

Du solltest dir diese Zusammenhänge gut anhand von Graphen verdeutlichen! Als Beispiel zeige ich dir in Abbildung 9.3 die Dichtefunktion

$$f = x \mapsto \left(\frac{1}{3}\sqrt{x}\right)' = \frac{1}{6\sqrt{x}}$$

der Zufallsvariablen X unseres Einführungsbeispiels auf Seite 51.

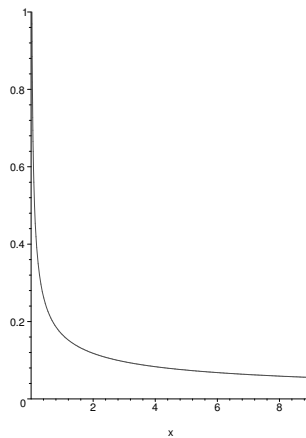


Abbildung 9.3: Dichtefunktion der Zufallsgröße X des Einführungsbeispiels in Abschnitt 9.1. Die Wahrscheinlichkeit, dass X einen Wert zwischen 2 und 3 annimmt, ist gleich dem Inhalt der Fläche unter dem Graphen zwischen 2 und 3

9.6 Eine Formel für den Erwartungswert einer Zufallsvariablen

Wir lösen sehr viele, sagen wir N , Werte unserer Zufallsvariablen X aus 9.1 aus und fragen nach dem Durchschnittswert

$$\bar{x} := \frac{1}{N} \sum_{i=1}^N x_i$$

dieser Werte. Wir teilen den Bereich von 0 bis 9, in dem die ausgelosten Werte liegen, in n Abschnitte der Länge Δx ein. Im i -ten Abschnitt liegen dann n_i Werte, die sich nur wenig unterscheiden. Deshalb werfen wir die alle in einen Topf und ersetzen die Summe dieser n_i Werte durch $n_i \bar{x}_i$; dabei soll \bar{x}_i den Mittelpunkt des i -ten Abschnitts bedeuten. Damit ergibt sich dann für den Durchschnittswert

$$\bar{x} \approx \frac{1}{N} \sum_{i=1}^n n_i \bar{x}_i = \sum_{i=1}^n \frac{n_i}{N} \bar{x}_i .$$

Der Quotient n_i/N ist gerade der Anteil der x_i , die im i -ten Abschnitt liegen, also etwa die Wahrscheinlichkeit dafür, dass X einen Wert im i -ten Abschnitt annimmt, folglich etwa $f(\bar{x}_i)\Delta x$. Wir erhalten

$$\bar{x} \approx \sum_{i=1}^n \bar{x}_i f(\bar{x}_i) \Delta x ,$$

und für $\Delta \rightarrow 0$ strebt das gegen das Integral über $xf(x)$. Demnach sollten wir den gesuchten Mittelwert so berechnen können:

$$\bar{x} = \int_0^9 xf(x) dx . \tag{9.4}$$

Aufgabe Berechne die Werte, die die Formel für die Beispiele liefert, die wir bisher behandelt haben. Passen die Werte zu den Ergebnissen unserer Simulationen?

Aufgabe Was sagst du eigentlich zu der Überlegung, die ich oben vorgestellt habe. Was ist das eigentlich, ein Beweis oder was?

Kapitel 10

Wirklichkeit und Mathematik

Begriffe wie Zufallsvariable und Wahrscheinlichkeit sind dir nun schon geläufig. Vielleicht hast du kurz gestutzt, als ich dir im letzten Abschnitt von Kapitel 1 den Begriff des Erwartungswerts untergemogelt habe, aber wenn ich dir sage, das sei der Wert, den die betrachtete Zufallsvariable auf lange Sicht im Mittel pro Durchführung annimmt, bist du vermutlich beruhigt. Dass wir vom Erwartungswert einer Zufallsvariablen und von Wahrscheinlichkeiten reden können, sagt uns das empirische Gesetz der großen Zahl, und unsere Simulationen haben uns darin auch bestätigt.

Jetzt kann ich dir einen kleinen Schock nicht ersparen. Ein empirisches Gesetz ist auf Erfahrung gegründet und somit meinetwegen Teil der Physik. Ein mathematischer Satz jedoch ist eine Aussage über mathematische Begriffe, deren Gültigkeit bewiesen ist; niemals reicht ein Hinweis auf Erfahrungen, um einen mathematischen Satz zu beweisen.

War also alles für die Katz, was wir bisher unter dem Namen Stochastik gemacht haben? Nein, aber es war noch keine Mathematik. Es diente dazu, die Prägung mathematischer Begriffe vorzubereiten. Bevor du weitere Erklärungen bekommst, musst du erst ins kalte Wasser.

4 Definition *Es sei $F : \mathbb{R} \rightarrow [0; 1]$ eine monoton steigende Funktion mit*

$$F(x) \xrightarrow{x \rightarrow -\infty} 0 \quad \text{und} \quad F(x) \xrightarrow{x \rightarrow \infty} 1 \quad .$$

*Dann wollen wir sagen, durch F sei eine **Zufallsvariable** X mit **Verteilungsfunktion** F gegeben und für $x \in \mathbb{R}$ sei $F(x)$ die **Wahrscheinlichkeit** $P(X < x)$, dass X einen Wert kleiner als x annimmt. Wenn F keine Sprungstellen hat und überall bis auf höchstens endlich viele Stellen differenzierbar ist, nennen wir die Ableitung $f := F'$ von F die **Dichtefunktion** der Zufallsvariablen X und das **Integral***

$$E(X) := \int_{-\infty}^{\infty} x f(x) dx$$

*den **Erwartungswert** der Zufallsvariablen X .*

Durch diese Definitionen werden Wahrscheinlichkeit und Zufallsvariable zu reinen Sprechweisen. Grundlegender mathematischer Begriff ist die Verteilungsfunktion, sie hat klar formulierte Eigenschaften. Auch Dichtefunktion und Erwartungswert sind sauber definierte Begriffe, mit denen man Mathematik treiben kann. Wir werden damit **beweisen**, dass der so definierte Erwartungswert in gewissem Sinne die Eigenschaften hat, die im empirischen Gesetz der großen Zahl beschrieben sind.

Und dann? Die Mathematik bietet dem Anwender eine gut entwickelte Theorie, die Wahrscheinlichkeitstheorie. Er kann sie benutzen, um darin Modelle für seine realen Zufallsphänomene zu bilden. Die Mathematik macht verlässliche Aussagen über die Modelle, und der Anwender kann diese Aussagen in die Welt seiner realen Phänomene übersetzen - freilich auf eigenes Risiko! Die Übersetzungen sind nur in dem Maße sicher, wie das Modell passte. Aussagen der Mathematik sind innerhalb der Mathematik verlässlich, aber die Wirklichkeit ist nicht Teil der Mathematik.

Wer Mathematik auf die Wirklichkeit anwenden will, sollte diese Problematik kennen. Leider sind gerade Anwender der Wahrscheinlichkeitstheorie oft in doppelter oder gar dreifacher Hinsicht ahnungslos. Sie verstehen die Mathematik nicht, die sie anwenden, sie wissen nicht, dass sie nur Auskünfte über ein Modell bekommen, und sie denken, weil Zahlen präzise sind, sei auch präzise, was in Zahlen formuliert ist. Diese Barbaren füttern ein Statistik-Programm mit ihren Daten und nehmen das, was das Programm über das mathematische Modell sagt, für Aussagen in barer Münze über ihr reales Phänomen. Solcher Unverstand sei fern von dir!

Kapitel 11

Konstruktion einer Simulation zu X mit gegebenem F

11.1 Die Methode

Es sei X eine Zufallsvariable mit Verteilungsfunktion F , und F habe keine Sprungstellen. Wir suchen eine Möglichkeit, Werte von X mit dem Rechner auszulösen.

Der Graph im linken Bild von Abbildung 11.1 sei einmal der Graph unserer Verteilungsfunktion mit der Gleichung $z = F(x)$. Dann ist $b = F(a) = P(X < a)$, das heißt, mit der Wahrscheinlichkeit b wird ein Wert von X ausgelost, der kleiner als a ist.

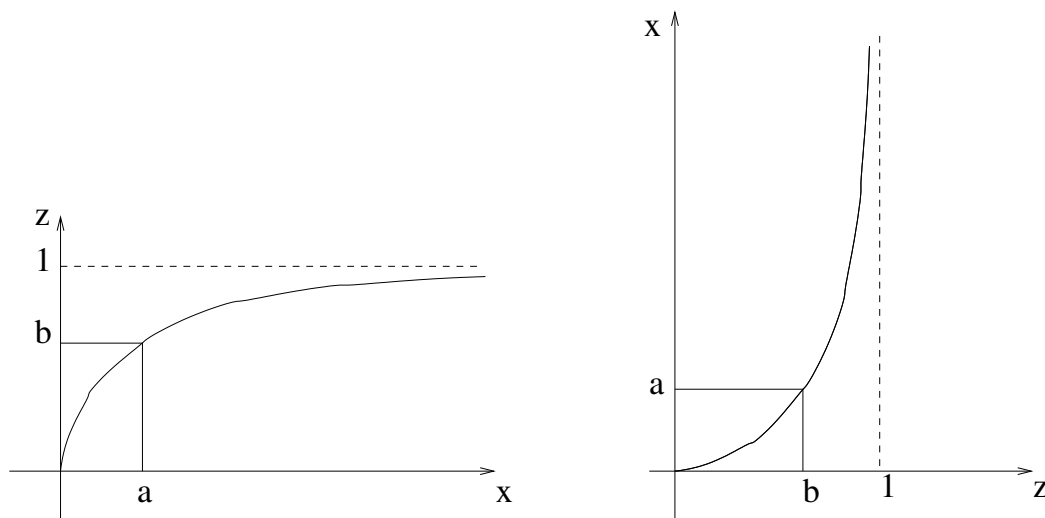


Abbildung 11.1: Graph der Verteilungsfunktion F der Zufallsvariablen X (links) und der Umkehrfunktion F^{-1} (rechts)

Im rechten Bild ist der Graph der Umkehrfunktion F^{-1} von F dargestellt, die Gleichung der Kurve ist also $x = F^{-1}(z)$. Wir lassen nun den Rechner eine Zahl z zwischen 0 und 1 auslösen und bilden die Zufallsgröße $Y = F^{-1}(Z)$. Nun nimmt Y einen Wert kleiner als a an, wenn Z einen Wert annimmt, der kleiner als b ist, und die Wahrscheinlichkeit, dass das passiert, ist gerade b :

$$P(Y < a) = P(Z < b) = b = P(X < a)$$

Das heißt aber nichts anderes, als dass X und Y die gleiche Verteilungsfunktion haben. Mathematisch sind sie dann identisch! Und damit haben wir eine Möglichkeit gefunden, Werte von X

mit dem Rechner auszulösen.

11.2 Eine neue Formel für den Erwartungswert

Du hast in Abschnitt 11.1 gesehen, wie wir die Auslosung von Werten einer Zufallsvariablen X mit Verteilungsfunktion F mit dem Rechner durchführen können: Wir lassen den Rechner eine Zufallszahl z zwischen 0 und 1 auslosen und berechnen daraus $x = F^{-1}(z)$. Für die Zufallsvariablen selbst können wir dies in der Form $X = F^{-1}(Z)$ hinschreiben.

Aus dieser Sicht der Dinge gewinnen wir eine neue Formel für den Erwartungswert von X . Betrachte ein Teilstück der Länge Δz des Intervalls $[0; 1]$. Die Wahrscheinlichkeit, dass das nächste ausgeloste z in diesem Teilstück Δz liegt, ist einfach $\Delta z/1 = \Delta z$. Wenn wir sehr viele, sagen wir N , Werte von Z auslosen, sollte der Anteil n/N derer, die im Teilstück Δz liegen, etwa Δz sein, also können wir mit $n = N\Delta z$ Werten in Δz rechnen. Für genügend kleines Δz können wir die alle mit dem Mittelpunkt des Teilstücks gleichsetzen. Teilen wir das ganze Intervall $[0; 1]$ in kleine Teilstückchen $(\Delta z)_i$ für $i = 1, \dots, n$ ein und nennen wir den Mittelpunkt des i -ten Teilstückchens \bar{z}_i , dann erhalten wir für den Mittelwert von den N Werten von X etwa

$$\frac{1}{N} \sum_{j=1}^N z_j \approx \frac{1}{N} \sum_{i=1}^n (N\Delta z) F^{-1}(\bar{z}_i) = \sum_{i=1}^n F^{-1}(\bar{z}_i) \Delta z \quad ,$$

und das ist eine Riemannsche Summe für das Integral

$$\int_0^1 F^{-1}(z) dz \quad . \quad (11.1)$$

Demnach sollte dieses Integral auch den Erwartungswert $E(X)$ liefern!

Das ist nun wieder eine Plausibilitätsbetrachtung, und sie ist sinnvoll und nützlich, damit du siehst, wie Stochastik funktioniert. Der Beweis sieht anders aus: Substituieren wir in dem Integral in Gleichung 11.1 das z durch $F(x)$, erhalten wir nach der Substitutionsregel

$$\int_0^1 F^{-1}(z) dz = \int_{-\infty}^{\infty} F^{-1}(F(x)) F'(x) dx = \int_{-\infty}^{\infty} x f(x) dx \quad ,$$

denn $F^{-1}(F(x)) = x$ und $F'(x) = f(x)$. Und das ist nach unserer grundlegenden Definition auf Seite 57 genau der Erwartungswert $E(X)$.

11.3 Beispiel: Radioaktiver Zerfall

Wenn wir jetzt N_0 Teilchen einer radioaktiven Substanz haben, werden es zum Zeitpunkt t noch

$$N(t) = N_0 e^{-at}$$

sein, dabei hängt a vom Material ab. Radioaktiver Zerfall ist ein echter Zufallsvorgang. Niemand kann im voraus wissen, wann ein bestimmtes Teilchen zerfällt! Dennoch kann man $N(t)$ präzise hinschreiben, und es stimmt genau so gut wie das Fallgesetz, wenn es nur genügend viele Teilchen sind.

Wir markieren nun eines der Teilchen und messen den Zerfallszeitpunkt t dieses Teilchens. Dadurch ist die Zufallsvariable T gegeben. Ihre Verteilungsfunktion F können wir leicht bestimmen. Die Wahrscheinlichkeit $P(T < t)$ ist der Anteil der vor t zerfallenen Teilchen an der Gesamtzahl, also

$$F(t) = P(T < t) = \frac{N_0 - N(t)}{N_0} = 1 - e^{-at} \quad \text{für } t \geq 0 \quad .$$

Als Dichtefunktion erhalten wir

$$f(t) = F'(t) = a e^{-at} \quad \text{für } t > 0 \quad ,$$

und für den Erwartungswert

$$E(T) = \int_0^{\infty} tf(t) dt = \int_0^{\infty} tae^{-at} dt = t(-e^{-at}) \Big|_0^{\infty} - \int_0^{\infty} -e^{-at} dt = 0 - \frac{1}{a}e^{-at} \Big|_0^{\infty} = \frac{1}{a} .$$

Bei dieser Rechnung benutzen wir das folgende Lemma.

5 Lemma Für alle natürlichen Zahlen n gilt

$$t^n e^{-t} \xrightarrow{t \rightarrow \infty} 0 .$$

Beweis. Wir benutzen die Taylor-Reihe der e -Funktion - Martin erzählt gelegentlich noch etwas darüber - und schätzen diese ab: Für positive t gilt

$$e^t = \sum_{k=0}^{\infty} \frac{t^k}{k!} \geq \frac{t^{n+1}}{(n+1)!} ,$$

denn wir lassen ja nur positive Summanden der Summe weg. Mit dieser Abschätzung erhalten wir für positive t

$$t^n e^{-t} = \frac{t^n}{e^t} \leq \frac{t^n(n+1)!}{t^{n+1}} = \frac{(n+1)!}{t} \xrightarrow{t \rightarrow \infty} 0 .$$

□

Wir wollen nun Werte von T auslosen. Es ist überraschend, dass das überhaupt möglich ist, denn der Bereich der Werte von T ist ja nicht nach oben beschränkt. Nun, wir lösen die Gleichung

$$z = F(t) = 1 - e^{-at}$$

nach t auf. Es ergibt sich

$$t = -\frac{1}{a} \ln(1 - z) .$$

Nun müssen wir nur $z \in [0; 1]$ zufällig wählen und dazu $F^{-1}(z) = -\frac{1}{a} \ln(1 - z)$ berechnen.

Im Unterricht haben wir die Simulation mit Maple durchgeführt. Aber weil ihr euch mit der Sache so schwertut, will hier ein paar Worte anfügen. Ich setze $a = 1$, gehe also aus von $F(t) = 1 - e^{-t}$ und $t = F^{-1}(z) = \ln(1 - z)$, und lasse auf die oben beschriebene Art 400 t -Werte auslosen. Die ordne ich der Größe nach und lasse mir die empirische Verteilungsfunktion dazu zusammen mit dem Graphen der Verteilungsfunktion zeichnen (siehe Abbildung 11.2 links).

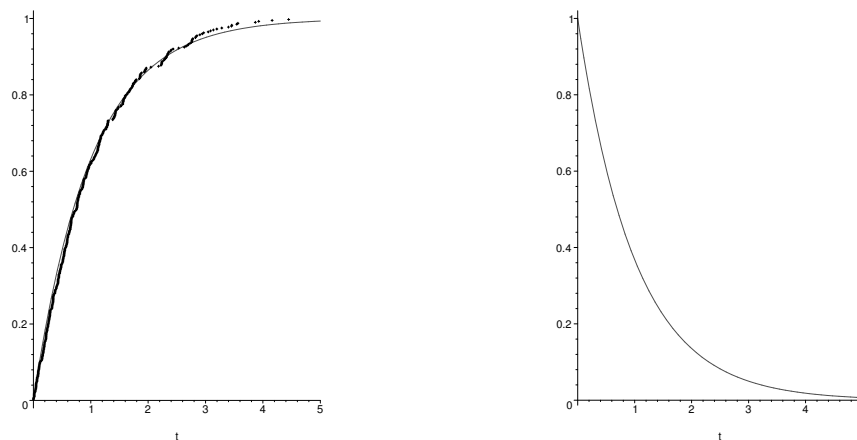


Abbildung 11.2: Verteilungsfunktion $F(t) = 1 - e^{-t}$ und empirische Verteilungsfunktion (400 Punkte) dazu (linkes Bild) sowie die zugehörige Dichte $f(t) = e^{-t}$ (rechtes Bild)

Du siehst dort, dass sich die Punktwolke der Verteilungsfunktion recht gut anpasst. Ferner siehst du, dass die Punkte anfangs sehr dicht liegen; in dem Bereich ab $t = 3$ wird es dann deutlich dünner. Das muss auch so sein. Es ist $F(1) = 1 - e^{-1} \approx 0.632$, also sollten etwa 63% der Punkte auf das Stück $0 \leq t \leq 1$ entfallen. Dagegen ist $P(T \geq 3) = 1 - P(T < 3) = 1 - F(3) = e^{-3} \approx .05$, es sollten also nur etwa 5% der Punkte jenseits von $t = 3$ liegen. Mache dir wieder klar, wie die empirische Verteilungsfunktion hergestellt wird. Der k -te der der Größe nach geordneten T -Werte, t_k , liefert den Punkt $(t_k | k/400)$. Links von ihm liegen $k - 1$ Werte, also sollte $F(t_k) \approx k/400$ sein (eigentlich $(k - 1)/400$). Deshalb gibt die Punktwolke den Verlauf des Graphen der Verteilungsfunktion wieder.

Kapitel 12

Die Varianz einer Zufallsvariablen

12.1 Funktionen von Zufallsvariablen

Es sei X eine Zufallsvariable mit Dichtefunktion f . Wir bilden eine neue Zufallsvariable Y , indem wir auf die Werte von X eine (vernünftige) Funktion g anwenden: $Y = g(X)$. Wie sieht es mit dem Erwartungswert von Y aus? Man kann natürlich die Verteilungsfunktion oder die Dichtefunktion von Y aufstellen und damit $E(Y)$ berechnen. Aber wir haben uns überlegt, dass wir $E(Y)$ auch direkt mit f berechnen können:

$$E(Y) = \int_{-\infty}^{\infty} g(x)f(x) dx \quad (12.1)$$

Übung

1. Es sei X Zufallsvariable mit Dichte f . Berechne die Erwartungswerte der Zufallsgrößen $X+3$, $5X$, $mX + b$ für Zahlen m, b .
2. Damit du nicht denkst, es sei immer $E(g(X)) = g(E(X))$, berechne einmal den Erwartungswert von X^2 für die Zufallsgröße X mit der Dichte $f(x) = 1 - x/2$ für $0 \leq x \leq 2$ und vergleiche ihn mit $E(X)^2$.

12.2 Definition der Varianz

Die Varianz einer Zufallsvariablen ist eine Zahl, die nach einer besonderen Vorschrift berechnet wird. Sie soll erfassen, wie stark die Werte der Zufallsvariablen um den Erwartungswert streuen. Ich schreibe die Definition hin:

6 Definition *Es sei X eine Zufallsvariable mit Dichte f und Erwartungswert $\mu = E(X)$. Dann ist die **Varianz** $V(X)$ von X gegeben durch*

$$V(X) := \int_{-\infty}^{\infty} (x - \mu)^2 f(x) dx \quad .$$

Die anschauliche Bedeutung des Integrals ist klar. Da wird einfach der Erwartungswert des Quadrats der Abweichung des Wertes von X vom Erwartungswert von X berechnet. Einen ersten Hinweis auf die Bedeutung der Varianz gibt die Ungleichung von Tschebyschew. Aussage und Beweis sind nicht so schlimm, wie sie auf den ersten Blick aussehen.

12.3 Die Ungleichung von Tschebyschew

Es sei X eine Zufallsvariable mit Erwartungswert $\mu = E(X)$ und Varianz $V(X)$. Wir schreiben den Rechenausdruck für die Varianz hin und betrachten den Beitrag der x -Werte, die μ um mindestens $a > 0$ verfehlen. Dazu unterteilen wir den Integrationsbereich und schätzen ab. Auf geht's!

$$\begin{aligned} V(X) &= \int_{-\infty}^{\infty} (x - \mu)^2 f(x) dx \\ &= \int_{-\infty}^{\mu-a} (x - \mu)^2 f(x) dx + \int_{\mu-a}^{\mu+a} (x - \mu)^2 f(x) dx + \int_{\mu+a}^{\infty} (x - \mu)^2 f(x) dx \end{aligned}$$

Nun, das mittlere Integral ist ≥ 0 , denn der Integrand ist ≥ 0 und die untere Grenze ist kleiner als die obere. Wir lassen das mittlere Integral einfach weg. Dann kann das, was übrig bleibt, höchstens so groß wie die Varianz sein:

$$V(X) \geq \int_{-\infty}^{\mu-a} (x - \mu)^2 f(x) dx + \int_{\mu+a}^{\infty} (x - \mu)^2 f(x) dx$$

Der Faktor $(x - \mu)^2$ in den beiden Integralen ist in den betrachteten Integrationsbereichen mindestens so groß wie a^2 . Wenn wir $(x - \mu)^2$ durch a^2 ersetzen, werden die Integrale höchstens kleiner, denn $f(x)$ ist ≥ 0 . Was übrig bleibt, ist also erst recht höchstens so groß wie die Varianz:

$$\begin{aligned} V(X) &\geq \int_{-\infty}^{\mu-a} a^2 f(x) dx + \int_{\mu+a}^{\infty} a^2 f(x) dx \\ &= a^2 \left(\int_{-\infty}^{\mu-a} f(x) dx + \int_{\mu+a}^{\infty} f(x) dx \right) \\ &= a^2 P(|X - \mu| \geq a) \end{aligned}$$

Nun liefert Division durch a^2 die **Tschebyschew-Ungleichung**:

$$P(|X - \mu| \geq a) \leq \frac{V(X)}{a^2} \tag{12.2}$$

Kapitel 13

Übungsaufgaben zu stetig verteilten Zufallsvariablen

Diese Aufgaben stammen aus dem Buch “A First Course in Probability” von Sheldon Ross, erschienen 1976 in New York.

Example 1a, p. 122 Suppose that X is a continuous random variable whose probability density function is given by

$$f(x) = \begin{cases} C(4x - 2x^2) & 0 < x < 2 \\ 0 & \text{otherwise} \end{cases}$$

What is the value of C ? Find $P(X > 1)$.

Example 1b, p. 122 The amount of time, in hours, that a computer functions before breaking down is a continuous random variable with probability density function given by

$$f(x) = \begin{cases} \lambda e^{-x/100} & x \geq 0 \\ 0 & x < 0 \end{cases}$$

What is the probability that a computer will function between 50 and 150 hours before breaking down? What is the probability that it will function less than 100 hours?

Example 1c, p. 123 The lifetime in hours of a certain kind of radio tube is a random variable having a probability density function given by

$$f(x) = \begin{cases} 0 & x \leq 100 \\ \frac{100}{x^2} & x > 100 \end{cases}$$

What is the probability that exactly 2 of 5 such tubes in a radio set will have to be replaced within the first 150 hours of operation? ¹

Example 2a, p. 125 If X is uniformly distributed over $(0; 10)$ ², calculate the probability that (1) $X < 3$, (2) $X > 6$, and (3) $3 < X < 8$.

Example 2b, p. 126 Buses arrive at a specified stop at 15-minute intervals starting at 7 a.m. That is, they arrive at 7, 7:15, 7:30, 7:45, and so on. If a passenger arrives at the stop at a time

¹Diese Aufgabe ist nicht, wie die vorigen Aufgaben, nur eine technische Übung. Da müssen wir uns eine ganze Reihe von Gedanken machen und einige Voraussetzungen formulieren.

²gleich-verteilt, das heißt, dass die Dichtefunktion zwischen 0 und 10 konstant ist und sonst 0 ist.

that is uniformly distributed between 7 and 7:30, find the probability that he waits (1) less than 5 minutes for a bus? (2) More than 10 minutes for a bus?

Example 2c, p. 126 Consider a “random chord” of a circle. What is the probability that the length of the chord will be greater than the side of the equilateral triangle inscribed that circle?³

Example 3c, p. 133 An expert witness in a paternity suit testifies that the length (in days) of pregnancy (that is, the time from impregnation to the delivery of the child) is approximately normally distributed with parameters $\mu = 270$ and $\sigma^2 = 100$. The defendant in the suit is able to prove that he was out of the country during a period that began 290 days before the birth of the child and ended 240 days before the birth. If the defendant was, in fact, the father of the child, what is the probability that the mother could have had the very long or very short pregnancy indicated by the testimony?

Text: Exponential Random Variables, p. 136 A continuous random variable whose probability density function is given, for some $\lambda > 0$, by

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & x \geq 0 \\ 0 & x < 0 \end{cases}$$

[...] The exponential distribution often arises, in practice, as being the distribution of the amount of time until some specific event occurs. For instance, the amount of time (starting from now) until an earthquake occurs, or until a new war breaks out, or until a telephone call you receive turns out to be a wrong number are all random variables that tend in practice to have exponential distributions.

Example 4a, p. 136 Suppose that the length of a phone call in minutes is an exponential random variable with parameter $\lambda = \frac{1}{10}$. If someone arrives immediately ahead of you at a public telephone booth, find the probability that you will have to wait (1) more than 10 minutes, and (2) between 10 and 20 minutes.

Example 4b, p. 137 Consider a post office that is manned by two clerks. Suppose that when Mr. Smith enters the system he discovers that Mr. Jones is being served by one of the clerks and Mr. Brown by the other. Suppose also that Mr. Smith is told that his service will begin as soon as either Jones or Brown leaves. If the amount of time that a clerk spends with a customer is exponentially distributed with parameter λ , what is the probability that, of the three customers, Mr. Smith is the last to leave the office?⁴

Zerfall Ein Teilchen einer radioaktiven Substanz, das wir jetzt vor uns sehen, zerfällt zu einem zufälligen Zeitpunkt. Die auf diese Weise gegebene Zufallsvariable ist exponentialverteilt mit Parameter $\lambda = 8$.

Wir sperren zwei Teilchen dieser Substanz in ein Gefäß, sie heißen Hans und Fritz. Berechne die Wahrscheinlichkeiten der folgenden Ereignisse.

- Hans lebt mindestens doppelt so lange wie Fritz.
- Nach 20 (Tagen) gibt es beide Teilchen nicht mehr.
- Nach vier (Tagen) ist mindestens eins der Teilchen zerfallen.
- Hans zerfällt im Zeitraum $[4, 6]$.
- Nimm an dass Fritz nach sechs (Tagen) noch da ist und dass Hans nicht mehr zu sehen ist. Wie groß ist die W, dass Hans im Zeitraum $[4, 6]$ zerfallen ist, wenn er mit Sicherheit nicht entkommen konnte (er vielleicht entkommen ist)?

³Diese Aufgabe ist schwierig!

⁴Diese Aufgabe ist schwierig!

Kapitel 14

Diskrete Zufallsvariable

Wenn du in unser Buch schaust, wirst du sehen, dass es dort in der Hauptsache um ganz andere Zufallsvariable geht, als dir bisher begegnet sind. Hier will ich dir zunächst an einem ausgearbeiteten Beispiel verdeutlichen, wie unsere Begriffe bei diskreten Zufallsvariablen aussehen. Dabei bedeutet das Wort diskret, dass die Werte der Zufallsvariablen sozusagen einzeln liegen. Wenn es nur endlich viele Werte gibt, ist die Zufallsvariable immer diskret.

Mein Fernziel ist, dir zu zeigen, dass sich unsere theoretischen Zufallsvariablen so verhalten, wie es das empirische Gesetz der großen Zahl für empirische Zufallsvorgänge sagt.

14.1 Ein typisches Beispiel

Unser Beispiel ist denkbar einfach: Wir werfen einen Würfel zweimal und notieren die Ergebnisse. Im Buch ist dann immer die Rede von einem L-Würfel, das ist ein Laplace-Würfel, ein idealer Würfel, bei dem alle Flächen wirklich gleichberechtigt sind.

Die Menge Ω aller Ergebnisse des Versuchs enthält 36 Elemente. Ob wir einen Sieg davontragen oder eine Niederlage einstecken, hängt in hohem Maße von der Schreibweise ab, die wir wählen. Ich schreibe hier $[i, j]$, wenn der erste Wurf die Zahl i und der zweite Wurf ein j ergab. Die ganze Menge Ω schreibe ich in Form einer Matrix auf.

$$A = \begin{pmatrix} [1, 1] & [1, 2] & [1, 3] & [1, 4] & [1, 5] & [1, 6] \\ [2, 1] & [2, 2] & [2, 3] & [2, 4] & [2, 5] & [2, 6] \\ [3, 1] & [3, 2] & [3, 3] & [3, 4] & [3, 5] & [3, 6] \\ [4, 1] & [4, 2] & [4, 3] & [4, 4] & [4, 5] & [4, 6] \\ [5, 1] & [5, 2] & [5, 3] & [5, 4] & [5, 5] & [5, 6] \\ [6, 1] & [6, 2] & [6, 3] & [6, 4] & [6, 5] & [6, 6] \end{pmatrix}$$

Eine Zufallsvariable ist nun eine Abbildung X , die jedem $\omega \in \Omega$ eine Zahl $X(\omega)$ zuordnet. Du kannst dir das so vorstellen, dass von jedem ω ein Pfeil zu einer Zahl zeigt, oder dass an jedem ω ein Zettel angeheftet wird, auf dem der Wert $X(\omega)$ steht. Hier mache ich es einfach so, dass ich zu X eine neue Matrix hinschreibe, und an der Stelle, wo eigentlich $[i, j]$ steht, schreibe ich $X([i, j])$ hin. Für die Zufallsvariable X_{min} : kleinste gewürfelte Zahl geht das so aus:

$$A_{min} = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 2 & 2 & 2 & 2 & 2 \\ 1 & 2 & 3 & 3 & 3 & 3 \\ 1 & 2 & 3 & 4 & 4 & 4 \\ 1 & 2 & 3 & 4 & 5 & 5 \\ 1 & 2 & 3 & 4 & 5 & 6 \end{pmatrix}$$

Entsprechend bilde ich für die Zufallsvariablen

1. X_{summe} : Summe der Augenzahlen
2. X_{seltsam} : Summe der Augenzahlen + kleinste gewürfelte Zahl
3. X_1 : erste gewürfelte Zahl
4. X_{doppelt} : das Doppelte der ersten gewürfelten Zahl
5. X_{prod} : Produkt der gewürfelten Zahlen
6. X_{quad} : Quadrat der ersten gewürfelten Zahl

die entsprechenden Matrizen.

$$\begin{aligned}
 A_{\text{summe}} &= \begin{pmatrix} 2 & 3 & 4 & 5 & 6 & 7 \\ 3 & 4 & 5 & 6 & 7 & 8 \\ 4 & 5 & 6 & 7 & 8 & 9 \\ 5 & 6 & 7 & 8 & 9 & 10 \\ 6 & 7 & 8 & 9 & 10 & 11 \\ 7 & 8 & 9 & 10 & 11 & 12 \end{pmatrix} & A_{\text{seltsam}} &= \begin{pmatrix} 3 & 4 & 5 & 6 & 7 & 8 \\ 4 & 6 & 7 & 8 & 9 & 10 \\ 5 & 7 & 9 & 10 & 11 & 12 \\ 6 & 8 & 10 & 12 & 13 & 14 \\ 7 & 9 & 11 & 13 & 15 & 16 \\ 8 & 10 & 12 & 14 & 16 & 18 \end{pmatrix} \\
 A_1 &= \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 \\ 2 & 2 & 2 & 2 & 2 & 2 \\ 3 & 3 & 3 & 3 & 3 & 3 \\ 4 & 4 & 4 & 4 & 4 & 4 \\ 5 & 5 & 5 & 5 & 5 & 5 \\ 6 & 6 & 6 & 6 & 6 & 6 \end{pmatrix} & A_{\text{doppelt}} &= \begin{pmatrix} 2 & 2 & 2 & 2 & 2 & 2 \\ 4 & 4 & 4 & 4 & 4 & 4 \\ 6 & 6 & 6 & 6 & 6 & 6 \\ 8 & 8 & 8 & 8 & 8 & 8 \\ 10 & 10 & 10 & 10 & 10 & 10 \\ 12 & 12 & 12 & 12 & 12 & 12 \end{pmatrix} \\
 A_{\text{prod}} &= \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 2 & 4 & 6 & 8 & 10 & 12 \\ 3 & 6 & 9 & 12 & 15 & 18 \\ 4 & 8 & 12 & 16 & 20 & 24 \\ 5 & 10 & 15 & 20 & 25 & 30 \\ 6 & 12 & 18 & 24 & 30 & 36 \end{pmatrix} & A_{\text{quad}} &= \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 \\ 4 & 4 & 4 & 4 & 4 & 4 \\ 9 & 9 & 9 & 9 & 9 & 9 \\ 16 & 16 & 16 & 16 & 16 & 16 \\ 25 & 25 & 25 & 25 & 25 & 25 \\ 36 & 36 & 36 & 36 & 36 & 36 \end{pmatrix}
 \end{aligned}$$

Jede dieser Zufallsvariablen stiftet ihre Einteilung von Ω in Teilmengen, und zwar gehören Elemente mit gleichem X -Wert immer zur gleichen Teilmenge. Die Stochastiker benutzen hier ein seltsames Symbol: Für die Teilmenge der ω mit $X_{\min}(\omega) = 3$ zum Beispiel schreiben sie $\{X_{\min} = 3\}$:

$$\{X = k\} := \{\omega \in \Omega \mid X(\omega) = k\}$$

Für die Wahrscheinlichkeit, dass X den Wert k annimmt, müssten sie dann $P(\{X = k\})$ schreiben, aber da lassen sie die Mengenklammern dann wieder einfach weg.

Ein Beispiel: Die Wahrscheinlichkeit, dass ein ω mit $X_{\min}(\omega) = 3$ ausgelost wird, ist

$$P(X_{\min} = 3) = \frac{|\{X_{\min} = 3\}|}{|\Omega|} = \frac{|\{[3, 3], [3, 4], [3, 5], [3, 6], [4, 3], [5, 3], [6, 3]\}|}{|\Omega|} = \frac{7}{36} .$$

Dabei bedeutet $|M|$ für eine Menge M einfach die Anzahl der Elemente von M .

Wir wollen nun die Erwartungswerte der Zufallsvariablen ausrechnen. Das kann man auf zwei Weisen tun. Man kann für jedes $\omega \in \Omega$ das Produkt $X(\omega)P(\{\omega\})$ bilden und diese Werte aufsummieren:

$$E(X) = \sum_{\omega \in X(\Omega)} X(\omega)P(\{\omega\}) . \quad (14.1)$$

Oder man kann sich die von X gestiftete Einteilung von Ω zunutze machen und immer alle ω mit gleichem X -Wert zusammenfassen. Das gibt dann

$$E(X) = \sum_{k \in X(\Omega)} kP(X = k) . \quad (14.2)$$

Hier nehme ich die erste Formel. Da muss ich den Rechner immer nur die Einträge einer Matrix addieren und das Ergebnis durch 36 teilen lassen. Und das sind die Ergebnisse:

$$E(X_{\min}) = \frac{91}{36} \quad E(X_{\text{summe}}) = 7 \quad E(X_{\text{seltsam}}) = \frac{343}{36} \quad E(X_{\text{doppelt}}) = 7$$

$$E(X_1) = \frac{7}{2} \quad E(X_{\text{prod}}) = \frac{49}{4} \quad E(X_{\text{quad}}) = \frac{91}{6}$$

Anhand dieser Ergebnisse wollen wir zu Vermutungen kommen, wie man mit Erwartungswerten rechnet.

14.2 Rechenregeln für Erwartungswerte: $E(X + Y)$, $E(aX + b)$

Zwischen den Zufallsvariablen des Abschnitts 14.1 gab es Zusammenhänge. Zum Beispiel können wir schreiben

$$X_{\text{seltsam}} = X_{\text{summe}} + X_{\min} \quad \text{und} \quad X_{\text{summe}} = X_1 + X_2 \quad ,$$

wenn wir unter X_2 die zweite gewürfelte Zahl verstehen. Besonders die, die sich mit formalen Dingen schwertun, müssen sich die Bedeutung genau klarmachen. Die Gleichung $X_{\text{summe}} = X_1 + X_2$ zwischen den Zufallsvariablen X_{summe} , X_1 und X_2 sagt, dass

$$X_{\text{summe}}(\omega) = X_1(\omega) + X_2(\omega) \quad \text{gilt für jedes } \omega \in \Omega \quad .$$

Wenn wir uns die entsprechenden Erwartungswerte anschauen, stellen wir fest, dass auch dafür gilt

$$E(X_{\text{seltsam}}) = E(X_{\text{summe}}) + E(X_{\min}) \quad \text{und} \quad E(X_{\text{summe}}) = E(X_1) + E(X_2) \quad .$$

In der Tat gilt allgemein das folgende Lemma.¹

7 Lemma *Es seien $X_1 : \Omega \rightarrow \mathbb{R}$ und $X_2 : \Omega \rightarrow \mathbb{R}$ zwei Zufallsvariable zum gleichen Zufallsexperiment. Wenn die Erwartungswerte existieren, gilt dann*

$$E(X_1 + X_2) = E(X_1) + E(X_2) \quad .$$

Beweis. Wir schreiben die Terme für die Erwartungswerte hin und vergleichen.

$$\begin{aligned} E(X_1 + X_2) &= \sum_{\omega \in \Omega} (X_1(\omega) + X_2(\omega))P(\{\omega\}) \\ &= \sum_{\omega \in \Omega} (X_1(\omega)P(\{\omega\}) + X_2(\omega)P(\{\omega\})) \\ &= \sum_{\omega \in \Omega} X_1(\omega)P(\{\omega\}) + \sum_{\omega \in \Omega} X_2(\omega)P(\{\omega\}) \\ &= E(X_1) + E(X_2) \end{aligned}$$

Fertig! □

Es folgt gleich das nächste Lemma:

8 Lemma *Es sei $X : \Omega \rightarrow \mathbb{R}$ eine Zufallsvariable mit Erwartungswert $E(X)$. Dann gilt für beliebige Zahlen $a, b \in \mathbb{R}$*

$$E(aX + b) = aE(X) + b \quad .$$

¹Das Lemma gilt auch für unsere stetig verteilten Zufallsvariablen! Du kannst den Beweis wörtlich übertragen.

Beweis. Wenn du dir erst klar gemacht hast, dass sich hinter $aX + b$ die Zufallsvariable verbirgt, die jedem $\omega \in \Omega$ den Wert $aX(\omega) + b$ zuordnet, brauchst du bloß noch die Definition des Erwartungswerts hinzuschreiben:

$$E(aX + b) = \sum_{\omega \in \Omega} (aX(\omega) + b)P(\{\omega\}) = a \sum_{\omega \in \Omega} X(\omega)P(\{\omega\}) + b \sum_{\omega \in \Omega} P(\{\omega\}) = aE(X) + b \cdot 1$$

□

14.3 Bedingte Wahrscheinlichkeit und Unabhängigkeit

Beim Produkt zweier Zufallsvariablen liegen die Dinge nicht so einfach. Die Gleichung $E(XY) = E(X)E(Y)$ ist manchmal erfüllt, manchmal nicht. Sie gilt, wenn die Zufallsvariablen von einander **unabhängig** sind. Da sich der Würfel nicht merkt, welche Zahl er gewürfelt hat, wird das Ergebnis des zweiten Wurfs nicht von dem des ersten Wurfs beeinflusst. Wenn ich weiß, dass der erste Wurf 1 war, sagt mir das nichts über den zweiten Wurf. Die Zufallsvariablen „Ergebnis des ersten Wurfs“ und „Ergebnis des zweiten Wurfs“ sind unabhängig.

Dagegen sind die Zufallsvariablen X_1 und X_{\min} unseres Beispiels nicht unabhängig: Wenn ich weiß, dass X_{\min} den Wert 5 annahm, weiß ich auch, dass X_1 nicht 1 sein kann.

Für gewöhnliche Ereignisse, also für Teilmengen A und B von Ω , sieht die Sache so aus. Wir setzen voraus, dass $P(A) \neq 0$ ist. Man vergleicht nun die Wahrscheinlichkeit $P(B)$, dass der Versuch ein $\omega \in B$ ausgibt, und die **bedingte Wahrscheinlichkeit**

$$P_A(B) := \frac{P(A \cap B)}{P(A)} ,$$

dass ein ω aus B kommt, wenn nur aus den Elementen von A gewählt wird. Bei Unabhängigkeit sind diese beiden Wahrscheinlichkeiten gleich, und dann erhält man durch Multiplikation mit $P(A)$ die schöne Gleichung

$$P(A \cap B) = P(A)P(B) \quad \text{bei Unabhängigkeit von } A \text{ und } B .$$

In den Anwendungen kommt oft eine Konstellation wie diese vor: Es ist Ω die deutsche Bevölkerung, A die Teilmenge der Raucher, B die Teilmenge der Menschen mit einer Krebskrankheit. Dann vergleicht man den Anteil $|B|/|\Omega|$ der Krebskranken in der Gesamtbevölkerung mit dem entsprechenden Anteil $|A \cap B|/|A|$ der Krebskranken unter den Rauchern. Sind die Werte verschieden, besteht eine stochastische Abhängigkeit zwischen Rauchen und Krebskrankheit.² Die Stochastik kommt hier so ins Spiel, dass man Leute zufällig auswählt und feststellt, ob sie rauchen und ob sie krebskrank sind: Man zieht eine Stichprobe. Alle wird man niemals untersuchen können.

Zurück zu unseren Zufallsvariablen: Es seien x ein Wert von X und y ein Wert von Y . Wenn X und Y unabhängig sind, ist stets

$$P(X = x, Y = y) = P(X = x)P(Y = y) .$$

14.4 Erwartungswert eines Produkts: $E(XY)$

9 Lemma Es seien $X, Y : \Omega \rightarrow \mathbb{R}$ von einander unabhängige Zufallsvariable mit Erwartungswerten $E(X)$ und $E(Y)$. Dann ist

$$E(XY) = E(X)E(Y) .$$

²Damit ist keineswegs gesagt, dass auch eine kausale Abhängigkeit besteht.

Beweis. Es ist

$$E(XY) = \sum_{\omega \in \Omega} X(\omega)Y(\omega)P(\{\omega\}) .$$

In dieser Summe fassen wir immer alle ω zusammen, die den gleichen X -Wert **und** den gleichen Y -Wert haben. Dann ist

$$\begin{aligned} E(XY) &= \sum_{x \in X(\Omega), y \in Y(\Omega)} xyP(X = x, Y = y) \\ &= \sum_{x \in X(\Omega), y \in Y(\Omega)} xyP(X = x)P(Y = y) \\ &= \sum_{x \in X(\Omega), y \in Y(\Omega)} xP(X = x)yP(Y = y) . \end{aligned}$$

Dabei haben wir die Unabhängigkeit ausgenutzt und etwas umgeordnet. Nun fassen wir für jedes $x \in X(\Omega)$ alle Summanden zusammen, die dieses x als Faktor erhalten:

$$\begin{aligned} E(XY) &= \sum_{x \in X(\Omega)} \left(\sum_{y \in Y(\Omega)} xP(X = x)yP(Y = y) \right) \\ &= \sum_{x \in X(\Omega)} \left(xP(X = x) \sum_{y \in Y(\Omega)} yP(Y = y) \right) \\ &= \sum_{x \in X(\Omega)} xP(X = x)E(Y) \\ &= E(Y) \sum_{x \in X(\Omega)} xP(X = x) \\ &= E(Y)E(X) \end{aligned}$$

□

14.5 Regeln für die Varianz

10 Lemma *Es sei X eine Zufallsvariable mit Erwartungswert $\mu = E(X)$ und Varianz $V(X)$. Dann gilt für alle Zahlen $a, b \in \mathbb{R}$*

$$V(aX) = a^2V(X) \quad \text{und} \quad V(X + b) = V(X) .$$

Beweis. Die Gleichungen ergeben sich aus der Definition der Varianz, ohne dass große Kunststücke nötig wären. Wir wissen schon, dass $E(aX) = aE(X) = a\mu$ ist, und damit erhalten wir

$$\begin{aligned} V(aX) &= \sum_{\omega \in \Omega} (aX(\omega) - a\mu)^2 P(\{\omega\}) = \sum_{\omega \in \Omega} a^2(X(\omega) - \mu)^2 P(\{\omega\}) \\ &= a^2 \sum_{\omega \in \Omega} (X(\omega) - \mu)^2 P(\{\omega\}) = a^2V(X) . \end{aligned}$$

Die zweite Gleichung beweist man auch so. □

Wenn du an das Histogramm denkst, bist du nicht überrascht, dass sich Verschieben des ganzen Histogramms um b nicht auf das Streuungsmaß $V(X)$ auswirkt.

11 Lemma *Es seien X und Y unabhängige Zufallsgrößen auf dem gleichen Ω . Wenn diese Varianzen existieren, gilt*

$$V(X + Y) = V(X) + V(Y) .$$

Beweis. Es ist

$$V(X + Y) = \sum_{x \in X(\Omega)} \sum_{y \in Y(\Omega)} (x + y - (E(X) + E(Y)))^2 P(X = x, Y = y) .$$

Wir formen den allgemeinen Summanden um zu

$$((x - E(X))^2 + (y - E(Y))^2 + xy - xE(Y) - E(X)y + E(X)E(Y)) P(X = x)P(Y = y) .$$

Wir lösen die große Klammer auf und schauen uns jeden Summanden einzeln an. Für den ersten ergibt sich

$$\begin{aligned} \sum_{x \in X(\Omega)} \sum_{y \in Y(\Omega)} (x - E(X))^2 P(X = x)P(Y = y) &= \sum_{x \in X(\Omega)} (x - E(X))^2 P(X = x) \sum_{y \in Y(\Omega)} P(Y = y) \\ &= V(X) \cdot 1 , \end{aligned}$$

denn alle von y unabhängigen Faktoren darf man aus der Summe über die y ausklammern, und die Summe über alle $P(Y = y)$ ist 1. Der Summand xy liefert $E(X)E(Y)$ (Unabhängigkeit!), die übrigen Summanden sind harmlos. Fasst man alles zusammen, erhält man genau die Behauptung. \square

14.6 Ketten von Zufallsversuchen

Häufig wiederholt man einen Zufallsversuch mehrfach. Man kann zum Beispiel den Zufallsversuch „würfeln“ n -mal wiederholen, eine Münze n -mal werfen oder n Leute zu einer Sache befragen. Oft interessiert man sich dafür, wie viele Sechsen geworfen wurden, wie oft Wappen kam oder wie viele Leute mit „ja“ geantwortet haben.

Ich führe ein Beispiel vor. Die Oppositionsführerin im Bundestag behauptet, 75% der deutschen Bevölkerung sei für eine deutsche Beteiligung am Irakkrieg. Nehmen wir einmal an, sie sagte die Wahrheit. Wir können einen Deutschen zufällig auswählen und ihn nach seiner Meinung fragen. Dann sollte er mit der Wahrscheinlichkeit $p = \frac{3}{4}$ mit „ja“ antworten. Wir führen diesen Zufallsversuch n -mal aus und bezeichnen die Anzahl der „ja“-s im i -ten Versuch mit X_i . Natürlich kann X_i nur die Werte 0 und 1 annehmen, und dies mit den Wahrscheinlichkeiten $p = \frac{3}{4}$ und $q := 1 - p = \frac{1}{4}$. Daraus ergibt sich sofort, dass

$$E(X_i) = 1 \cdot p + 0 \cdot q = p \quad \text{und} \quad V(X_i) = (1 - p)^2 p + p^2 q = q^2 p + p^2 q = pq(p + q) = pq \quad (14.3)$$

ist. Für die Gesamtzahl

$$S := \sum_{i=1}^n X_i$$

der „ja“-s unter n Befragten erhalten wir nach unseren Regeln sofort

$$E(S) = np \quad \text{und} \quad V(S) = npq . \quad (14.4)$$

Dabei gehen wir davon aus, dass die X_i von einander unabhängig sind.³

Was bringt nun die Kenntnis dieser Parameter? Wenn wir unter $n = 1000$ Leuten nur 400 finden, die für den Einsatz der Bundeswehr im Irak sind, könnte die Oppositionsführerin ja von Zufall sprechen. Aber das nähmen wir ihr nicht ab. Denn für $n = 1000$ ist

$$P(|S - 1000p| \geq 100) \leq \frac{V(S)}{100^2} = \frac{1000 \cdot \frac{3}{4} \cdot \frac{1}{4}}{100^2} = 0,01875 ,$$

³Du solltest diese Ergebnisse übrigens nicht gering schätzen. Wir haben uns überlegt, dass

$$P(S = k) = \frac{n!}{k!(n-k)!} p^k q^{n-k}$$

ist. Das lässt sich zwar leicht hinschreiben, die Fakultäten werden aber sehr schnell sehr groß. Der Rechner konnte den Erwartungswert für $n = 1000$ längst nicht mehr über die Definition berechnen.

und die Tschebyschew-Ungleichung überschätzt die Wahrscheinlichkeit noch erheblich. Wenn sie Recht hat, darf die Anzahl der „ja“-s in der Stichprobe von 750 keinesfalls um 100 oder mehr abweichen.

Dieses Beispiel sollte dir einen kleinen Einblick in die Beurteilende Statistik geben. Du wirst mehr davon zu sehen bekommen. Für heute halte ich nur noch einen wichtigen Begriff fest.

12 Definition Eine Zufallsvariable X , die die Werte $k = 0, 1, \dots, n$ mit der Wahrscheinlichkeit

$$P(X = k) = \frac{n!}{k!(n-k)!} p^k q^{n-k}$$

annimmt für eine feste Zahl p , $0 < p < 1$, $q = 1 - p$, heißt binomialverteilt mit den Parametern n und p , kurz $B(n, p)$ -verteilt. Dabei nennt man p die Erfolgswahrscheinlichkeit.

Einen Versuch, der nur zwei Ergebnisse hat, nennt man Bernoulli-Versuch. Gewöhnlich sieht man einen Ausgang als Erfolg an, seine Wahrscheinlichkeit wird mit p bezeichnet, und den anderen als Misserfolg. Die Wahrscheinlichkeit $1 - p$ des Misserfolgs bezeichnet man mit q . Wiederholt man den Versuch n -mal unabhängig, nennt man dies eine Bernoulli-Kette der Länge n . Die Zufallsvariable „Anzahl der Erfolge“ ist dann $B(n, p)$ -verteilt. Bernoulli-Kette treten in den Anwendungen sehr häufig auf.

14.7 Bernoullis schwaches Gesetz der großen Zahl

Du hast im vorigen Abschnitt gesehen, wie eine Zufallsvariable aus unserem theoretischen Vorrat benutzt wurde, eine Umfrage zu modellieren. Man geht davon aus, dass sich eine reale Umfrage im Prinzip so verhält wie das Modell aus der Theorie; eine Annahme, die sich praktisch hundertfach bewährt hat, die aber außerhalb der Mathematik liegt und auch keinesfalls „beweisbar“ ist. Das Risiko bei diesen Anwendungen trägt immer der Anwender. Die Mathematik ist in Ordnung, die Anwendung auf reale Phänomene bleibt ein Wagnis.

Du kennst das empirische Gesetz der großen Zahl, und du sollst jetzt sehen, dass sich eine Zufallsvariable unserer Theorie so verhält, wie es das empirische Gesetz der großen Zahl für reale Zufallsvorgänge sagt. Wenn dies nicht so wäre, hätte es wenig Sinn, theoretische Zufallsvariable als Modelle für reale Zufallsvorgänge zu benutzen.

Es sei also X eine (diskrete) Zufallsvariable mit Erwartungswert μ und Varianz $V(X)$. Du kannst getrost für X die Augenzahl beim Werfen eines L -Würfels nehmen. Wir nehmen n unabhängige Kopien X_1, \dots, X_n von X und bilden den Mittelwert

$$\bar{X} = \frac{1}{n} \sum_{k=1}^n X_k .$$

Für das Beispiel hieße das, dass wir n -mal würfeln und den Mittelwert der gewürfelten Zahlen bilden. Erwartungswert und Varianz von \bar{X} kennen wir, es ist

$$E(\bar{X}) = \mu \quad \text{und} \quad V(\bar{X}) = \frac{1}{n} V(X) .$$

Wir schätzen nun die Wahrscheinlichkeit, dass der Wert von \bar{X} den Erwartungswert μ von \bar{X} um $\epsilon > 0$ oder noch mehr verfehlt, mit der Tschebyschew-Ungleichung ab:

$$P(|\bar{X} - \mu| \geq \epsilon) \leq \frac{V(\bar{X})}{\epsilon^2} = \frac{1}{n} \frac{V(X)}{\epsilon^2} \xrightarrow{n \rightarrow \infty} 0 \quad (14.5)$$

Das ist schon **Bernoullis schwaches Gesetz der großen Zahl**. Mache dir die Aussage gründlich klar.

Wir haben an Histogrammen verfolgt, wie sich \bar{X} bei wachsendem n verhält. Die Histogramme von \bar{X} wurden immer höher und spitzer, die Wahrscheinlichkeit konzentrierte sich immer mehr in einem engen Bereich um den Erwartungswert μ . „Mittelwerte schwanken weniger als die Einzelwerte.“

14.8 Hypergeometrisch-verteilte Zufallsvariable

Man hat eine Menge Ω mit N Elementen und eine Teilmenge $A \subseteq \Omega$ mit n Elementen. Nun wählt man eine m -elementige Teilmenge $B \subseteq \Omega$ zufällig aus und schaut, wieviele Elemente aus A in B sind. Diese Anzahl ist der Wert einer Zufallsvariablen X , sie heißt hypergeometrisch-verteilt mit den Parametern N, n, m .

Nimm zum Beispiel als Ω die Menge der Schüler unseres Kurses und als A die Teilmenge derer, die in der Gemeinde Stemwede wohnen. Wähle $m = 5$ Schüler des Kurses zufällig und stelle fest, wieviele davon in Stemwede wohnen; diese Anzahl ist dann der Wert von X . Vornehm gesprochen: X ist die Anzahl der Stemweder unter fünf zufällig gewählten Schülern unseres Kurses.

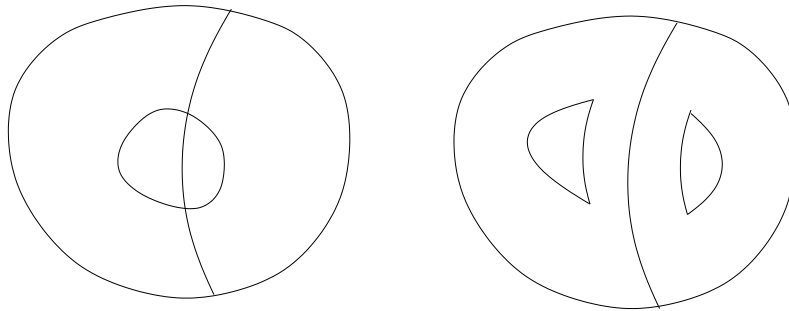


Abbildung 14.1: Zur hypergeometrischen Verteilung

Wir bestimmen $P(X = k)$. Schau dir dazu die linke Figur in Abbildung 14.1 an. Der ganze Kringel steht für Ω . Die vertikale Linie teilt Ω in die Teilmenge A rechts der Linie und das Komplement \bar{A} links der Linie. Der kleine Kringel steht für die Teilmenge B . Der rechte Teil des kleinen Kringels ist $B \cap A$, der linke Teil ist $B \cap \bar{A}$. Wieviele solcher Mengen B gibt es, die m Elemente haben, von denen genau k Elemente zu A gehören? Schau auf die rechte Figur in 14.1. Die Mengen Ω und A sind genau so veranschaulicht wie in der linken Figur. Wir wählen eine k -elementige Teilmenge von A (rechter kleiner Kringel) und eine $(m - k)$ -elementige Teilmenge von \bar{A} (linker kleiner Kringel) und fügen die beiden zu einem B zusammen. Auf diese Weise erhält man alle möglichen $B \subseteq \Omega$ mit $|B| = m$ und $|B \cap A| = k$. Ihre Anzahl ist das Produkt der Anzahlen der möglichen k -elementigen Teilmengen von A und der $(m - k)$ -elementigen Teilmengen von \bar{A} , also gerade

$$\binom{n}{k} \binom{N - n}{m - k} .$$

Um $P(X = k)$ zu erhalten, muss man diese Anzahl nur durch die Anzahl aller m -elementigen Teilmengen von Ω teilen:

$$P(X = k) = \frac{\binom{n}{k} \binom{N - n}{m - k}}{\binom{N}{m}} \quad (14.6)$$

Man kann $P(X = k)$ auch mit Hilfe eines Baums berechnen, das schreibe ich aber nicht auf.

14.9 Eigenschaften binomial verteilter Zufallsvariablen

In diesem Abschnitt sei X eine $B(n, p)$ -verteilte Zufallsvariable, also typischerweise die Anzahl der Erfolge bei einer n -stufigen Bernoulli-Kette mit der Erfolgswahrscheinlichkeit p .⁴ Zunächst halte ich ein Ergebnis fest, das wir im Unterricht bewiesen haben.

⁴Siehe Seite 73

13 Lemma Für eine $B(n, p)$ -verteilte Zufallsvariable X gilt

$$P(X = k) < P(X = k + 1) \iff k < np - q .$$

Die Wahrscheinlichkeit $P(X = k)$ wächst also mit k bis etwa zum Erwartungswert und fällt dann wieder, und das Histogramm ist eingipflig.

Im Buch findest du auf den Seiten 242 und 243 Histogramme einiger binomial-verteilter Zufallsvariablen. Wir lesen daran einige Beobachtungen ab.

1. Das Histogramm einer $B(n, q)$ -verteilten Zufallsvariablen bekommt man, indem man das einer $B(n, p)$ -verteilten an der Geraden $x = n/2$ spiegelt.
2. Für kleines n ist das Histogramm sehr asymmetrisch, wenn p nahe bei 0 oder bei 1 liegt.
3. Bei festem p werden die Histogramme mit wachsendem n immer breiter, flacher und symmetrischer. Sie gehen aber nicht im gleichen Maße in die Breite, in dem n wächst. Der Bereich der k , in dem die Wahrscheinlichkeit zum allergrößten Teil liegt, wird relativ immer kleiner.

Wenn man sich dies so anschaut, mag man auf den Gedanken kommen, dass hinter den Histogrammen eine Grund- oder Urform steckt. Dies ist in der Tat so, aber es bedarf einiger Mühe, diese Urform ans Licht zu bringen.

1. Die **Verschiedenheit der Lage** ist leicht aufzufangen. Man verschiebt das Histogramm einfach um μ nach links. Der Erwartungswert der neuen Zufallsvariablen $X - \mu$ ist dann 0.
2. Jetzt muss man versuchen, das **Zerfließen in die Breite** auffangen. Das geht am einfachsten durch Stauchen der Merkmalsachse mit einem geeigneten Faktor r . Wir erhalten die neue Zufallsvariable

$$Z := \frac{X - \mu}{r} .$$

Wenn es tatsächlich die Urform gibt, sollten verschiedene X praktisch gleiche Z liefern. Das bedeutet jedenfalls, dass die Varianz von Z möglichst überhaupt nicht von X abhängen sollte. Berechnen wir also die Varianz von Z !

$$V(Z) = V\left(\frac{X - \mu}{r}\right) = \frac{1}{r^2}V(X - \mu) = \frac{1}{r^2}V(X)$$

Damit ist die Sache klar: Wir stauchen mit $r = \sqrt{V(X)} = \sqrt{npq}$, dann wird $V(Z) = 1$, und das fängt das Zerfließen in die Breite in der Tat genau auf. Die gemeinsame Urform ist durch die berühmte Gauß-Kurve gegeben, die den Zehnmarkschein schmückte. Die Größe $\sigma := \sqrt{V(X)}$ nennt man übrigens die **Standardabweichung** von X . Ergebnis: **Wenn die Varianz $V(X)$ nicht zu klein ist, wird das Histogramm der standardisierten Zufallsvariablen**

$$Z = \frac{X - \mu}{\sigma}$$

recht gut durch die Gauß-Kurve approximiert.

Zur Illustration gebe ich hier ein Beispiel. In Abbildung 14.2 siehst du das Histogramm der standardisierten Zufallsvariablen $Z = (X - \mu)/\sigma$ für eine $B(30, 0.2)$ -verteilte Zufallsvariable X und den Graphen der Gauß-Funktion φ in einem Schaubild.

Ich notiere das Ergebnis etwas härter in einem Lemma.

14 Lemma (Integrale Näherungsformel von de Moivre und Laplace) Es sei X eine $B(n, p)$ -verteilte Zufallsvariable, $Z := (X - \mu)/\sigma$ die zu X gehörige standardisierte Zufallsvariable. Dann gilt

$$P(X \leq k) = P\left(Z \leq \frac{k - \mu}{\sigma}\right) \approx \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\frac{k - \mu}{\sigma} + \frac{1}{2\sigma}} e^{-\frac{1}{2}x^2} dx .$$

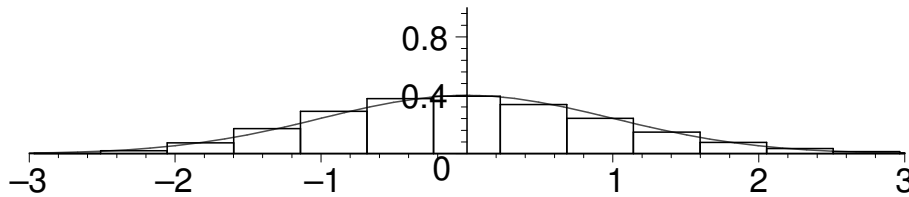


Abbildung 14.2: Histogramm von $(X - \mu)/\sigma$ für $B(30, 0.2)$ -verteiltes X und Gaußkurve $y = \varphi(x)$ in einem Schaubild

Man beweist das Lemma mit knochenharter Analysis, wir schenken uns das.

Mit Hilfe des Lemmas berechnet man Wahrscheinlichkeiten bei binomial-verteilten Zufallsvariablen. Du kannst davon ausgehen, dass die Approximation gut genug ist, wenn du die Wahrscheinlichkeiten nicht mehr zu Fuß ausrechnen kannst. Das Integral über die berühmte Gauß-Funktion

$$\varphi(x) := \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2} \quad (14.7)$$

kann man dummerweise auch nicht geschlossen hinschreiben, aber es gibt Tabellen für

$$\Phi(x) := \int_{-\infty}^x \varphi(t) dt \quad , \quad (14.8)$$

und damit kann man die gesuchten Wahrscheinlichkeiten ausrechnen.

14.10 Warten auf Erfolg: geometrisch verteilte Zufallsvariable

Wie oft muss man einen L-Würfel werfen, bis man eine vier hat? Es sei X die Anzahl der benötigten Würfe. Dann ist mit den üblichen Bezeichnungen

$$P(X = k) = pq^{k-1} \quad \text{für } k = 1, 2, 3, \dots \quad (14.9)$$

Eine Zufallsvariable, die die Werte $k \in \mathbb{N}$ mit der Wahrscheinlichkeit $P(X = k)$ aus Gleichung 14.9 annimmt, heißt **geometrisch verteilt**.

Im Mittel braucht man nun

$$E(X) = \sum_{k=1}^{\infty} kpq^{k-1} = p \sum_{k=1}^{\infty} kq^{k-1}$$

Würfe.

Der Wert dieser Reihe will erst einmal bestimmt sein! Das mag man so versuchen: Die Reihe $\sum_{k=0}^n q^k$ ist die bekannte geometrische Reihe, ihr Wert ist

$$S_n = \sum_{k=0}^n q^k = \frac{1 - q^{n+1}}{1 - q} \quad .$$

Das beweist man zum Beispiel mit dem Trick, dass man $qS_n - S_n$ ausrechnet, das ist einfach $q^{n+1} - 1$, und daraus S_n berechnet. Unser q ist positiv und kleiner als 1, da kann man n gegen Unendlich laufen lassen und erhält $1/(1 - q)$ als Grenzwert. Für $0 < q < 1$ ist also⁵

$$f(q) := \sum_{k=0}^{\infty} q^k = \frac{1}{1 - q} \quad .$$

⁵Hier hast du ein Beispiel einer Taylor-Reihe! Frage bei Martin nach!

Man kann auch mit Mitteln der Stochastik argumentieren: Wenn man n -mal würfelt, ist unter den Ergebnissen mit der Wahrscheinlichkeit $\sum_{k=1}^n pq^{k-1}$ eine vier. Andernfalls hat man noch keine vier geworfen, das passiert mit der Wahrscheinlichkeit q^n ; weitere Möglichkeiten gibt es nicht. Also ist

$$q^n + \sum_{k=1}^n pq^{k-1} = 1 \quad .$$

Daraus folgt, dass

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

ist, und wir erhalten wieder unsere Summenformel für die geometrische Reihe.

Unsere Reihe oben, die wir für $E(X)$ erhalten haben, ist die Ableitung der geometrischen Reihe nach q , also sollte

$$E(X) = pf'(q) = p \left(\frac{1}{1-q} \right)' = p \frac{1}{(1-q)^2} = \frac{1}{p}$$

sein. Der Schluss ist allerdings etwas gewagt.

Wir beschäftigen uns hier mit Stochastik, deshalb füge ich wieder ein zweites Argument hinzu. Spiele das Spiel sehr oft, sagen wir, bis du etwa $6N$ -mal gewürfelt hat. Dann wirst du etwa N viieren haben, das heißt, du hast das Spiel etwa N -mal gespielt. Im Mittel hast du dann je Spiel $6N/N = 6$ Würfe gebraucht.

Ich halte das Resultat hier fest: Wenn du einen Bernoulli-Versuch mit der Erfolgswahrscheinlichkeit p so oft durchführst, bis sich ein Erfolg einstellt, ist die Anzahl X der benötigten Würfe geometrisch verteilt, das heißt, $P(X = k)$ ist wie in Gleichung 14.9 gegeben. Der Erwartungswert von X ist

$$E(X) = \frac{1}{p} \quad .$$

14.11 Poisson-verteilte Zufallsvariable

Beobachtungen mögen ergeben haben, dass Frau Schulzes Telefon im Mittel sechsmal in der Stunde klingelt. Es sei X die Anzahl der Anrufe, die in der nächsten Stunde eingehen. Wie sieht es mit $P(X = k)$ aus?

Wenn wir so fragen, gehen wir davon aus, dass die Anrufe zufällig eingehen, und wir sehen von der Dauer der Gespräche ab. Beides ist keineswegs unproblematisch, aber wer nicht vereinfachen will, kann nicht viel sagen.

Man löst die Sache so: Man teilt die Stunde in n kleine Intervalle ein und bezeichnet mit X_n die Anzahl der Anrufe, die in einem Intervall eingehen. Im Mittel müssen das natürlich $6/n$ sein, das ist klar, aber was bringt das sonst? Nun, der Grund ist, dass man für genügend großes n davon absehen kann, dass zwei oder mehr Anrufe eingehen. Dann ist bei X_n die gesamte Wahrscheinlichkeit auf die Werte 0 und 1 verteilt, und wir können $p_n := P(X_n = 1)$ ausrechnen:

$$6/n = E(X_n) = 1 \cdot P(X_n = 1) + 0 \cdot P(X_n = 0) = P(X_n = 1)$$

Wenn wir wissen wollen, wieviele Anrufe in der ganzen Stunde eingehen, schauen wir jedes Teilintervall an und zählen die Teilintervalle, in denen ein Anruf eingeht. Das heißt, unser X ist, wenn n groß genug ist, $B(n, 6/n)$ -verteilt:

$$P(X = k) \approx \binom{n}{k} \left(\frac{6}{n} \right)^k \left(1 - \frac{6}{n} \right)^{n-k}$$

Ich habe euch im Unterricht plausibel gemacht, dass der Ausdruck auf der rechten Seite für $n \rightarrow \infty$ gegen

$$\frac{1}{k!} \left(\frac{6}{n} \right)^k e^{-\frac{6}{n}}$$

strebt. Dabei kam die bemerkenswerte Tatsache

$$\left(1 + \frac{x}{n}\right) \xrightarrow{n \rightarrow \infty} e^x \quad (14.10)$$

zur Sprache. Auf Einzelheiten gehe ich hier nicht ein.

15 Definition Es sei $\lambda > 0$. Eine Zufallsvariable X mit

$$P(X = k) = \frac{\lambda^k}{k!} e^{-\lambda} \quad \text{für } k = 0, 1, 2, \dots$$

heißt *Poisson-verteilt mit Parameter λ* .

Die Zufallsvariable X , von der wir ausgegangen sind, ist also Poisson-verteilt mit Parameter $\lambda = 6$.

Mit Martins Hilfe haben wir uns wieder an die Taylorreihe der e -Funktion erinnert. Damit konnten wir zeigen, dass $E(X) = \lambda$ ist für Poisson-verteiltes X , was ja auch nicht so verwunderlich ist.

Kapitel 15

Beurteilende Statistik

Die Beurteilende Statistik macht Aussagen über reale Zufallsvorgänge.

15.1 Wieviel Erfolge gibt es bei einer Bernoulli-Kette normalerweise?

Eine Werbefirma verschickt 1000 Werbesendungen, die zu einer Einkaufsfahrt einladen. Bekanntlich reagieren im Mittel 5% der Leute. Mit welcher Anzahl von Teilnehmern bei der Einkaufsfahrt sollte man rechnen? Wir fassen die Sache als Bernoulli-Kette der Länge $N = 1000$ mit Erfolgswahrscheinlichkeit $p = 0,05$ auf, das heißt, wir stellen uns vor, dass jeder Empfänger der Post sich unabhängig von den anderen mit der Wahrscheinlichkeit p dafür entscheidet, mitzufahren. Im Prinzip könnten alle 1000 Leute kommen, dann brauchte man zwanzig Busse. Sollte man so viele Busse chartern? Natürlich nicht, die Wahrscheinlichkeit ist ja nur $0,05^{1000}$, und das ist nicht viel. Der Geschäftsführer will wissen, wieviele „normalerweise“ kommen. Wenn er eine Antwort von uns haben will, muss er sagen, was für ihn „normalerweise“ heißt. Er könnte sagen, mit 99%-iger Sicherheit. Dann bestimmen wir eine Umgebung des Erwartungswerts 50, in der der ausgeloste Wert von X mit der Wahrscheinlichkeit 0,99 liegt. Man nennt diese Umgebung die 99%-Umgebung des Erwartungswerts. Um sie zu bekommen, bilden wir

$$z = \Phi^{-1}(0,995) \approx 2,8$$

und haben damit den Radius

$$2,8\sigma \approx 2,8\sqrt{1000 \cdot 0,05 \cdot 0,95} \approx 19,3$$

der Umgebung gefunden. Der Geschäftsführer sollte davon ausgehen, dass zwischen 30 und 70 Einkaufswillige erscheinen werden.

15.2 Das Konfidenzintervall

Bleiben wir bei dem Beispiel aus Abschnitt 15.1, und nehmen wir an, es seien 110 Leute gekommen. Der Geschäftsführer wird schnell noch einen Bus besorgen müssen und dann, wenn das aktuelle Problem gelöst ist, von seiner Statistikabteilung verlangen, dass der Anteil p derer, die gewöhnlich auf die Werbesendung reagieren, neu bestimmt wird.

Man könnte von heute ab mit dem neuen Wert $p = 0,11$ rechnen, aber das wäre wenig professionell. Hätte man nur 100 Werbesendungen verschickt und wären 11 Leute gekommen, hätte man auch $p = 0,11$ gesagt, aber diese 0,11 wären sicher weniger verlässlich als unsere 0,11. Man sagt also nicht einfach „Neues p ist 0,11“, sondern man bestimmt das **Konfidenzintervall** zum beobachteten Wert $\hat{p} = 0,11$ zu einer vorzuziehenden Irrtumswahrscheinlichkeit, meinetwegen wieder 1%. Es enthält genau die p , in deren 99%-Umgebung das \hat{p} liegt.

Die Rechnung geht so: Der Radius der 99%-Umgebung von $\mu = Np$ ist

$$\Phi^{-1}(0,995)\sigma = \Phi^{-1}(0,995) \cdot \sqrt{Np(1-p)} \quad .$$

Dies setzt man gleich $|Np - N\hat{p}|$, das ergibt eine quadratische Gleichung. Die beiden Lösungen sind das größte und das kleinste p , in deren 99%-Umgebung von Np der beobachtete Wert $N\hat{p}$ gerade noch liegt. Das Intervall zwischen diesen beiden Werten ist das gesuchte Konfidenzintervall. Ersatzweise kannst du auch einfach $\Phi^{-1}(0,995)\sigma$ mit dem σ von \hat{p} als Radius des Konfidenzintervalls nehmen, der Mittelpunkt ist in dem Fall \hat{p} . Dann sollte aber \hat{p} nicht weit von 0,5 weg liegen.

Warnung Bitte sage nicht, das gesuchte neue p liege mit der Wahrscheinlichkeit 99% im gefundenen Intervall. Das hört man zwar oft, aber davon wird es nicht richtiger. Was bedeutet denn „Wahrscheinlichkeit 99%“? Wenn man die Sache oft macht, sollte es in etwa 99% der Fälle gut gehen. Aber wenn man die Sache (bestimme \hat{p} , berechne Konfidenzintervall) mehrfach durchführt, bekommt man jedesmal ein neues \hat{p} und damit ein neues Konfidenzintervall, das ist das Problem. Ausgang eines Zufallsversuchs ist das Intervall, nicht etwa ein wahres p . Etwa 99% der gefundenen Intervalle sollten das wahre p überdecken. Die Wahrscheinlichkeit 99% sagt etwas über die Güte des Verfahrens, nicht über den gefundenen Wert. Bleibe hier intellektuell redlich und schludere nicht!

15.3 Testen

Sagen wir, ein Würfel soll getestet werden, ob es auch ein ordentlicher L-Würfel ist. Dann wirft man ihn meinetwegen 1200-mal und zählt, wie viele Sechsen er gewürfelt hat. Wenn diese Anzahl genügend nahe beim Erwartungswert 200 liegt, akzeptiert man den Würfel, sonst lehnt man ihn ab. Dieses „genügend nahe“ präzisiert man zu einer so genannten **Entscheidungsregel**: Akzeptiere den Würfel, wenn die beobachtete Anzahl der Sechsen im Bereich von 180 bis 220 liegt, und lehne ihn sonst ab.

Im Idealfall besteht ein L-Würfel den Test immer, und ein gezinkter Würfel wird immer abgelehnt. Ein realer Test leistet das aber nicht. Ein L-Würfel wird mit einer gewissen Wahrscheinlichkeit abgelehnt, und ein schlechter Würfel, der eine sechs meinetwegen nur mit der Wahrscheinlichkeit 0,16 liefert, kann den Test trotzdem bestehen.

Wir berechnen die Wahrscheinlichkeiten, mit der diese Fehler auftreten. Dazu bezeichnen wir mit X_p eine $B(1200, p)$ -verteilte Zufallsvariable. Die Wahrscheinlichkeit, dass ein Würfel, der mit der Wahrscheinlichkeit p eine sechs würfelt, den Test besteht, ist dann

$$f(p) := P(180 \leq X_p \leq 220) \approx \Phi\left(\frac{220,5 - 1200p}{\sqrt{1200p(1-p)}}\right) - \Phi\left(\frac{179,5 - 1200p}{\sqrt{1200p(1-p)}}\right) \quad .$$

Diese Funktion f heißt **Gütefunktion** des Tests. In Abbildung 15.1 siehst du ihren Graphen. Die Wahrscheinlichkeit, dass ein L-Würfel abgelehnt wird, ist

$$1 - f\left(\frac{1}{6}\right) \approx 0,1123 \quad ,$$

das mag akzeptabel sein. Ein gezinkter Würfel mit der Wahrscheinlichkeit $p = 0,16$ für eine sechs besteht den Test aber auch mit der Wahrscheinlichkeit $f(0,16) \approx 0,8251$. Will man diese Wahrscheinlichkeit kleiner machen, muss man die Anzahl der Würfe erhöhen und/oder den Annahmebereich verkleinern. Letzteres wird aber dazu führen, dass auch mehr L-Würfel abgelehnt werden.

15.4 Schnelltest für den Alltag

Es sei X eine $B(n, p)$ -verteilte Zufallsvariable, und es sei h ein beobachteter Wert von X . Ist h ein auffälliger Wert? Am einfachsten misst man, wie weit (in σ gemessen) h vom Erwartungswert np

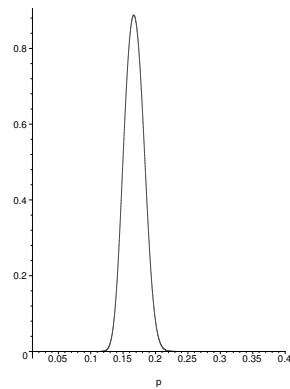


Abbildung 15.1: Gütefunktion eines Tests

entfernt ist, das heißt, man bildet die Kenngröße

$$\frac{h - np}{\sigma} .$$

Ihr Betrag sollte eigentlich nicht größer als 3 sein.

Kapitel 16

Klausuren

16.1 Erste Klausur 12.2

1. Es sei X eine Zufallsvariable mit Dichte $f(x) = \frac{2}{9}x$ für $0 \leq x \leq 3$.
 - a) Berechne den Erwartungswert von X und gib $F(x)$ an.
 - b) Teile das Intervall $[0; 3]$ der x -Achse in drei gleich wahrscheinliche Teile.
2. Eine radioaktive Substanz zerfällt so, dass die Anzahl der Teilchen zum Zeitpunkt t durch

$$N(t) = N_0 e^{-\frac{1}{3}t}$$

gegeben ist. Wir wählen ein bestimmtes Teilchen aus. Die Zufallsvariable T sei der Zerfallszeitpunkt dieses Teilchens, die Verteilungsfunktion von T ist dann durch

$$F(t) = 1 - e^{-\frac{1}{3}t}$$

gegeben.

- a) Berechne den Erwartungswert von T und die Wahrscheinlichkeit, dass das Teilchen innerhalb der nächsten zwei Zeiteinheiten zerfällt.
 - b) Nimm an, wir schauen nach zwei Zeiteinheiten nach und das Teilchen ist noch da. Wie wahrscheinlich ist es dann, dass es innerhalb der nächsten zwei Zeiteinheiten zerfällt?
 - c) Wie lange dauert es (vom Beginn an), bis das Teilchen mit der Wahrscheinlichkeit 0.99 zerfallen ist?
3. Wie hoch liegt ein zufällig gewählter Punkt der oberen Hälfte des Einheitskreises im Mittel über der x -Achse? „Das ist doch klar“, meint Kunz, „nimm $\pi/2$ und teile das durch 2.“ Aber da hat Kunz mal wieder nicht richtig hingeschaut. Berechne du den richtigen Wert, gib auch die Verteilungsfunktion der Zufallsvariablen an und finde heraus, was Kunz sich gedacht hat.
 4. Erkläre den Begriff der Verteilungsfunktion einer Zufallsvariablen X : Was bedeutet $F(x)$? Wie sieht der Graph von F typischerweise aus? Und stelle dar, wie man den Graphen von F empirisch gewinnen kann, wenn man eine Maschine hat, die einem auf Knopfdruck beliebig viele Werte von X auslost.
 5. Hinz baut ein Glücksrad für die charity-week. Er malt neun von zehn Sektoren schwarz; wenn der Zeiger da stehen bleibt, gibt es nichts. Wenn der Zeiger im zehnten Sektor stehen bleibt, zahlt Hinz einen Euro aus. Nun überlegt er, wieviel er auf lange Sicht pro Spiel auszahlen muss, und er bildet den Ausdruck

$$\frac{9}{10} \cdot 0 + \frac{1}{10} \cdot 1 \quad .$$

Was sagst du zu seiner Rechnung?

6. In Joseph Bertrands Nachlass fand sich folgendes Maple-Programm. Was tut es, und wozu wurde es geschrieben?

```
> wert:=proc()
>   local z;
>   z:=rand(0..10000)()/10000;
>   RETURN( 2*sqrt(1-z^2) );
> end:
> N:=100; s:=0:
> for i from 1 to N do s:=s+wert(); od:
> evalf( s/N );
```

16.2 Zweite Klausur 12.2

In der Stochastik ist es besonders wichtig, dass du den Lösungsweg gut dokumentierst. Du brauchst nicht viele Worte zu machen, aber schreibe hin, was nötig ist. Viel Glück!

1. Du kennst doch sicher diese Überraschungseier aus Schokolade, die in ihrem Innern eine Überraschung aus Plastik enthalten. Jan und Jens sind ganz verrückt auf eine bestimmte Sorte von Figuren, die in durchschnittlich jedem fünften Ei sein sollen.
 - a) Jan kauft drei dieser Eier. Dann könnte er ja keine, eine, zwei oder drei dieser Figuren erbeutet haben. Berechne die Wahrscheinlichkeiten und zeichne ein Histogramm. Gib auch den Erwartungswert an und erkläre, was er bedeutet.
 - b) Jens ist da schon aus anderem Holz geschnitzt, er kauft nämlich gleich achtzig dieser Dinger. Er öffnet sie an Ort und Stelle, und siehe da, er hat 42 Figuren erbeutet. Natürlich freut er sich, aber der Geschäftsführer des Ladens kommt und wirft Jens vor, er habe verbotenerweise die Eier mit der Hand gewogen und vielversprechende Kandidaten ausgewählt. Hören wir uns mal den Dialog der beiden an:

J: „Das war Zufall“.

G: „Nein, dass man zufällig 42 Figuren bekommt, ist extrem unwahrscheinlich.“

J: „Was wollen Sie, dass man 16 Figuren bekommt, ist auch nicht so wahrscheinlich.“

G: „Schon, aber es ist eben unwahrscheinlich, dass man 42 oder sogar noch mehr Figuren in 80 Überraschungseiern findet.“

Na, das reicht jetzt. Berechne du die Wahrscheinlichkeiten, mit denen die beiden argumentieren, erkläre, wieso Jens von 16 Figuren redet, und gib ein Urteil ab.
2. Ein Mathematikurs und ein Englischkurs fahren gemeinsam auf Kursfahrt. Im Mathematikurs sind sechzehn Jungen und zwei Mädchen, im Englischkurs sind sieben Jungen und elf Mädchen.
 - a) Der Wirt der Unterkunft trifft ein Mitglied der Reisegruppe auf dem Flur. Mit welcher Wahrscheinlichkeit gehört es zum Englischkurs? Mit welcher Wahrscheinlichkeit gehört es zum Mathematikurs, wenn es ein Mädchen ist?
 - b) Nun kommen ein Junge und ein Mädchen aus der Reisegruppe gemeinsam um die Ecke. Mit welcher Wahrscheinlichkeit gehören die beiden zu verschiedenen Kursen?
 - c) Vier Jungen spielen Tischfußball. Mit welcher Wahrscheinlichkeit sind sie alle aus dem Mathematikurs?
 - d) Morgens kommen die Leute aus dem Mathematikurs irgendwie der Reihe nach in den Frühstücksraum. Wie groß ist die Wahrscheinlichkeit, dass als dritte Person ein Mädchen eintrifft?
 - e) In der Unterkunft gibt es ein besonders schönes Dreibettzimmer, und eine Dreiergruppe von Jungen des Mathematikurses soll es bekommen. Damit es gerecht zugeht, soll das Los

- entscheiden. Der Wirt schlägt vor, dass man jede mögliche Dreierkombination auf einen Zettel schreibt, die Zettel in eine Urne tut und einen Zettel zieht. Was sagst du zu dieser Idee? Mache auch einen Alternativvorschlag.
- f) Wie wahrscheinlich ist es, dass die beiden Sebastians in das Dreierzimmer kommen?
- g) Wie wahrscheinlich ist es, dass keiner der beiden Sebastians in das Dreierzimmer kommt?
- h) Der Wirt grummelt, „Junge sein“ und „im Mathematikkurs sein“ sei wohl nicht unabhängig voneinander. Was genau verstehen wir unter stochastischer Unabhängigkeit, und ist sie bei dieser Reisegruppe gegeben?
3. Schau dich mal in diesem Raum um, der dir schon fast ein zweites Zuhause ist. Wir teilen in Gedanken ein Viertel des Raums ab, den Teil dahinten, wo Sebastian, Magnus und die anderen sitzen. Die Luftmoleküle hier im Raum sind sehr zahlreich, und sie sind ständig in heftiger Bewegung. Denke dir, eines trüge eine rote Mütze. Nun machst du ein Foto und schaust nach, ob sich das Molekül mit der roten Mütze in dem hinteren Viertel des Raums befindet. Das wird mit der Wahrscheinlichkeit $p = 1/4$ der Fall sein, dieser Überlegung wirst du wohl zustimmen. Was dem Molekül mit der roten Mütze recht ist, ist den anderen billig. Wir gehen nun davon aus, dass jedes der – sagen wir – $N = 6 \cdot 10^{25}$ Luftteilchen seinen Aufenthaltsort im Raum gewählt hat, ohne sich um die anderen zu scheren.
- a) Wieviele Teilchen sollten sich im hinteren Raumteil etwa aufhalten?
- b) Nun könnte es ja sein, dass sich zufällig besonders viele oder besonders wenige Teilchen für den hinteren Raumteil entschieden haben. Dann hätten deine Mitschüler Druck auf den Ohren, oder sie schnappten nach Luft. Wie groß ist denn die Wahrscheinlichkeit, dass der Anteil der Teilchen im hinteren Raumteil um mindestens 0.001 vom Sollwert abweicht? Gib zunächst eine Antwort mit Hilfe der Tschebyschew-Ungleichung. Erinnerung: Vergiss nicht, die benutzte Zufallsvariable zu definieren!
4. Beweise, dass $E(X + Y) = E(X) + E(Y)$ ist für Zufallsvariable $X, Y : \Omega \rightarrow \mathbb{R}$ und dass $E(X) = np$ ist für eine $B(n, p)$ -verteilte Zufallsvariable X .

Teil III

Lineare Algebra und Geometrie

Kapitel 17

Lineare Algebra

Die Lineare Algebra ist eine vergleichsweise junge mathematische Theorie. Wir finden einen guten Zugang zu ihren zentralen Begriffen, indem wir der Frage nachgehen, wie die Lösungsmengen linearer Gleichungssysteme aussehen und wie sie man sie geometrisch deuten kann.

Zunächst geht es darum, angemessene Schreibweisen zu entwickeln. Ein Problem kann leicht zu lösen sein, wenn es in einer passenden Schreibweise formuliert ist, und nahezu unlösbar, wenn die Schreibweise verkorkst ist. Lass dich also nicht abschrecken, wenn ich jetzt ganz komprimiert notiere, was wir in einer Reihe von Stunden kennengelernt haben; es handelt sich um ganz vorzügliche, unglaublich leistungsfähige Werkzeuge.

17.1 Lineare Gleichungssysteme

Eine Lösung des linearen Gleichungssystems

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \dots + a_{1m}x_m &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2m}x_m &= b_2 \\ \dots & \\ a_{n1}x_1 + a_{n2}x_2 + \dots + a_{nm}x_m &= b_n \end{aligned} \tag{17.1}$$

von n Gleichungen in m Variablen muss ein Satz von m Werten sein. Man schreibt diesen Wertesatz als Spalte

$$\vec{x} := \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{pmatrix} \in \mathbb{R}^m, \tag{17.2}$$

fasst entsprechend die rechten Seiten zu einer Spalte $\vec{b} \in \mathbb{R}^n$ zusammen und ist dann schon bei der Vektor-Matrixschreibweise

$$A\vec{x} = \vec{b}. \tag{17.3}$$

für das lineare Gleichungssystem 17.1. Dabei ist die Matrix A einfach das Koeffizientenschema der linken Seite von 17.1, also

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1j} & \dots & a_{1m} \\ a_{21} & a_{22} & \dots & a_{2j} & \dots & a_{2m} \\ \vdots & \vdots & & \vdots & & \vdots \\ a_{i1} & a_{i2} & \dots & a_{ij} & \vdots & a_{im} \\ \vdots & \vdots & & \vdots & & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nj} & \dots & a_{nm} \end{pmatrix}, \tag{17.4}$$

und man spricht von einer $n \times m$ -Matrix, weil sie n Zeilen und m Spalten hat. Das Element a_{ij} steht in der i -ten Zeile und der j -ten Spalte.

17.2 Der Spaltenraum \mathbb{R}^n

Der \mathbb{R}^n ist natürlich die Menge aller Spalten der Länge n , deren Einträge reelle Zahlen sind. Für die Elemente des \mathbb{R}^n benutzen wir die Symbole $\vec{x}, \vec{y}, \vec{a}, \dots$, und die n reellen Zahlen von \vec{x} bezeichnen wir in der Regel mit x_1, x_2, \dots, x_n .

Damit ist der \mathbb{R}^n erst einmal nur ein Haufen toten Materials; er taugt gerade mal als Vorrat, aus dem wir Lösungen unserer LGS nehmen. Er bekommt eine Struktur, wenn man Rechenoperationen für die Spalten definiert. Mit diesen Rechenoperationen ist der \mathbb{R}^n ein **euklidischer Vektorraum**.

16 Definition Für $\vec{x}, \vec{y} \in \mathbb{R}^n$, $r \in \mathbb{R}$ definieren wir

$$\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} + \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} := \begin{pmatrix} x_1 + y_1 \\ x_2 + y_2 \\ \vdots \\ x_n + y_n \end{pmatrix}, \quad r \cdot \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} := \begin{pmatrix} rx_1 \\ rx_2 \\ \vdots \\ rx_n \end{pmatrix},$$

und

$$\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} * \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} := x_1 y_1 + x_2 y_2 + \dots + x_n y_n.$$

Für die so definierten Rechenoperationen $+$ und \cdot gelten im Wesentlichen die Regeln, die für die entsprechenden Operationen mit reellen Zahlen gelten; deshalb wählt man auch dieselben Bezeichnungen. Das innere Produkt $*$ ist uns schon im letzten Jahr begegnet.

17.3 Lösungsmengen linearer Gleichungssysteme

Um die Teilmengen des \mathbb{R}^n , die wir als Lösungsmengen linearer Gleichungssysteme berechnet haben, bequem schreiben zu können, definieren wir einige Begriffe.

17 Definition Es seien $\vec{a}_1, \vec{a}_2, \dots, \vec{a}_m \in \mathbb{R}^n$. Eine Summe

$$r_1 \vec{a}_1 + r_2 \vec{a}_2 + \dots + r_m \vec{a}_m$$

von Vielfachen der \vec{a}_i heißt **Linearkombination** der \vec{a}_i , und die Menge aller Linearkombinationen heißt das **Erzeugnis**

$$\langle \vec{a}_1, \vec{a}_2, \dots, \vec{a}_m \rangle := \{r_1 \vec{a}_1 + r_2 \vec{a}_2 + \dots + r_m \vec{a}_m \mid r_i \in \mathbb{R}\}$$

der \vec{a}_i . Ferner definieren wir für Teilmengen $M \subseteq \mathbb{R}^n$ und $\vec{a} \in \mathbb{R}^n$ die Vektormenge

$$\vec{a} + M := \{\vec{a} + \vec{m} \mid \vec{m} \in M\}.$$

Wir lösen lineare Gleichungssysteme mit dem Gaußschen Algorithmus. Dabei bringen wir die Koeffizientenmatrix des Systems auf Stufenform. Wenn wir dieses Verfahren mit wachen Sinnen durchführen, erkennen wir, dass die Lösungsmenge leer ist oder genau einen Vektor enthält oder von der Form

$$\vec{v} + \langle \vec{w}_1, \vec{w}_2, \dots, \vec{w}_d \rangle$$

ist. Deuten wir für $n = 3$ die Vektoren als Koordinaten von Punkten, bedeutet die Lösungsmenge im dritten Fall geometrisch auf den ersten Blick für $d = 1$ eine Gerade durch den Punkt V , deren Richtung durch \vec{w}_1 gegeben ist, oder für $d = 2$ im allgemeinen eine Ebene durch V . Eine genauere Untersuchung kommt noch.

17.4 Für welche rechte Seite \vec{b} ist $A\vec{x} = \vec{b}$ lösbar?

Es sei A eine $n \times m$ -Matrix. Wir betrachten das lineare Gleichungssystem $A\vec{x} = \vec{b}$. Dann muss $\vec{b} \in \mathbb{R}^n$ und $\vec{x} \in \mathbb{R}^m$ sein, sonst passen die Größen nicht.

Wenn wir $A\vec{x}$ ausschreiben und genau hinschauen, erkennen wir, dass $A\vec{x}$ eine Linearkombination der m Spaltenvektoren $\vec{a}_1, \vec{a}_2, \dots, \vec{a}_m$ der Matrix A ist. Daraus ergibt sich das folgende Lemma.

18 Lemma Das lineare Gleichungssystem $A\vec{x} = \vec{b}$ ist genau dann lösbar, wenn \vec{b} im Erzeugnis der Spaltenvektoren von A liegt.

17.5 „Größe“ eines Erzeugnisses und lineare Unabhängigkeit

Wenn wir uns das Erzeugnis einer Anzahl von Vektoren des \mathbb{R}^3 geometrisch veranschaulichen wollen, sieht das Erzeugnis eines Vektors aus wie eine Gerade durch den Nullpunkt. Kommt ein zweiter Vektor hinzu, ergibt sich eine Ebene durch den Nullpunkt, und das Erzeugnis dreier Vektoren sollte bereits der ganze Raum sein. Hier ist freilich Vorsicht geboten. Natürlich darf der Nullvektor nicht unter den erzeugenden Vektoren sein, das ist klar. Aber der zweite Vektor darf nicht zum Erzeugnis des ersten gehören, und der dritte nicht zum Erzeugnis der ersten beiden. Und das kann schiefgehen – sonst wäre auch jedes 3×3 -System lösbar!

Damit sind wir beim Begriff der linearen Unabhängigkeit eines Systems von Vektoren.

19 Definition Die Vektoren $\vec{a}_1, \vec{a}_2, \dots, \vec{a}_m$ des \mathbb{R}^n nennen wir **linear unabhängig**, wenn keiner von ihnen als Linearkombination der übrigen darstellbar ist. Andernfalls nennen wir sie **linear abhängig**.

Es gibt ein recht bequemes Kriterium für lineare Unabhängigkeit.

20 Lemma Die Vektoren $\vec{a}_1, \vec{a}_2, \dots, \vec{a}_m$ des \mathbb{R}^n sind genau dann linear unabhängig, wenn

$$r_1\vec{a}_1 + r_2\vec{a}_2 + \dots + r_m\vec{a}_m = \vec{0}$$

nur für $r_1 = r_2 = \dots = r_m = 0$ gilt.

Beweis. Nehmen wir an, die Vektoren seien linear abhängig. Dann ist einer, meinetwegen der erste, Linearkombination der übrigen:

$$\vec{a}_1 = r_2\vec{a}_2 + \dots + r_m\vec{a}_m \quad .$$

Bringen wir \vec{a}_1 auf die andere Seite, erhalten wir eine Darstellung des Nullvektors als Linearkombination

$$r_1\vec{a}_1 + r_2\vec{a}_2 + \dots + r_m\vec{a}_m = \vec{0} \quad ,$$

und es ist zumindest $r_1 = -1 \neq 0$.

Es gebe nun umgekehrt eine Darstellung des Nullvektors als Linearkombination der \vec{a}_i , bei der nicht alle Vorfaktoren $= 0$ sind. Der Bequemlichkeit halber nehmen wir an, es sei $r_1 \neq 0$. Dann können wir die Gleichung nach \vec{a}_1 auflösen:

$$\vec{a}_1 = -\frac{r_2}{r_1}\vec{a}_2 - \dots - \frac{r_m}{r_1}\vec{a}_m \quad .$$

Folglich ist dann der Vektor \vec{a}_1 als Linearkombination der übrigen darstellbar.

Damit ist das Lemma bewiesen. □

Wenn wir mit einem Erzeugnis

$$W = \langle \vec{w}_1, \dots, \vec{w}_m \rangle$$

zu tun haben, werden wir uns die erzeugenden Vektoren anschauen. Sind sie linear abhängig, kann ich welche von ihnen einfach fortlassen, ohne dass das Erzeugnis kleiner wird. Das geht so lange, bis die restlichen Vektoren linear unabhängig sind. Die Mindestzahl von Vektoren, die ich brauche, um W zu erzeugen, heißt die Dimension von W ; das will ich hier ruhig schon verraten, damit du siehst, wohin die Reise geht.

17.6 Ein durchgerechnetes Beispiel

Es sei eine 4×4 -Matrix A gegeben. Wir fragen zunächst, für welche rechten Seiten \vec{b} das lineare Gleichungssystem $A\vec{x} = \vec{b}$ eine Lösung hat. Die Menge aller Lösungen ist, wie wir wissen, das Erzeugnis der vier Spaltenvektoren $\vec{a}_1, \dots, \vec{a}_4$ der Matrix A . Wir prüfen, ob die Vektoren linear unabhängig sind oder ob darunter überflüssige Vektoren sind, die wir weglassen können, ohne dass das Erzeugnis kleiner wird. Dazu verwenden wir den Ansatz

$$r_1\vec{a}_1 + \dots + r_4\vec{a}_4 = \vec{0}$$

aus Lemma 20. Der führt auf ein lineares Gleichungssystem mit der rechten Seite $\vec{0}$, ein sogenanntes **homogenes** Gleichungssystem. Wir bearbeiten es wie üblich mit dem Gauß-Algorithmus. Die rechte Seite lassen wir einfach weg, weil die ja $\vec{0}$ bleibt.

Nehmen wir an, der Gauß-Algorithmus führt schließlich auf die Matrix

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

Wir lesen die Lösungsmenge ab:

$$\mathbb{L} = \left\{ \begin{pmatrix} r_2 - 2r_3 \\ r_2 \\ r_3 \\ 0 \end{pmatrix} \mid r_2, r_3 \in \mathbb{R} \right\} = \left\langle \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} -2 \\ 0 \\ 1 \\ 0 \end{pmatrix} \right\rangle$$

Die vier Vektoren $\vec{a}_1, \dots, \vec{a}_4$ sind nicht linear unabhängig, weil es ja nicht nur die Lösung mit $r_1 = \dots = r_4 = 0$ gibt. Wir sehen aber noch mehr, wenn wir die beiden erzeugenden Lösungen in unseren Ansatz einsetzen. Dann steht da nämlich

$$\vec{a}_1 + \vec{a}_2 = \vec{0} \quad \text{und} \quad -2\vec{a}_1 + \vec{a}_3 = \vec{0} .$$

Das heißt, dass wir die Vektoren \vec{a}_2 und \vec{a}_3 durch \vec{a}_1 ausdrücken können. In Wirklichkeit wird \mathbb{L} schon durch \vec{a}_1 und \vec{a}_4 erzeugt. Die beiden sind allerdings linear unabhängig, denn setzen wir in unserem Ansatz $r_2 = r_3 = 0$, muss auch $r_1 = 0$ sein. Dass $r_4 = 0$ ist, war ja ohnehin klar.

Und wie sieht nun die Lösungsmenge des Systems $A\vec{x} = \vec{b}$ aus? Wenn \vec{b} nicht im Erzeugnis $\langle \vec{a}_1, \vec{a}_4 \rangle$ liegt, ist die Lösungsmenge leer. Sonst führt der Gauß-Algorithmus auf ein System der Form

$$\left(\begin{array}{cccc|c} 1 & -1 & 2 & -3 & b'_1 \\ 0 & 0 & 0 & 1 & b'_2 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array} \right) ,$$

und die Lösungsmenge ist

$$\begin{pmatrix} b'_1 \\ b'_2 \\ 0 \\ 0 \end{pmatrix} + \left\langle \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} -2 \\ 0 \\ 1 \\ 0 \end{pmatrix} \right\rangle .$$

Damit ist das Problem algebraisch gelöst. Wenn wir nach einer geometrischen Deutung der Angelegenheit fragen, wird die Sache ganz durchsichtig. Dazu führen wir zunächst Matrixabbildungen ein.

17.7 Matrixabbildungen und lineare Gleichungssysteme

Es sei A eine $n \times m$ -Matrix. Für das \vec{x} auf der linken Seite des linearen Gleichungssystems $A\vec{x} = \vec{b}$ kann man Vektoren aus dem \mathbb{R}^m einsetzen und ausrechnen, was herauskommt. Dann hat man es mit einer sogenannten Matrixabbildung

$$F : \vec{x} \mapsto A\vec{x} \quad \text{für } \vec{x} \in \mathbb{R}^m \quad (17.5)$$

zu tun, die jedem Vektor \vec{x} des \mathbb{R}^m einen Vektor $F(\vec{x})$ des \mathbb{R}^n zuordnet. Durch F sind zwei besondere Vektormengen ausgezeichnet, der Kern und das Bild von F . Ich definiere sie einmal ordentlich und notiere dann wichtige Eigenschaften von F , die rechtfertigen, dass man sich mit Matrixabbildungen beschäftigt.

21 Definition Es sei $F : \mathbb{R}^m \rightarrow \mathbb{R}^n$ eine Abbildung. Dann definieren wir den Kern und das Bild von F durch

$$\text{Kern}(F) := \{\vec{x} \in \mathbb{R}^m \mid F(\vec{x}) = \vec{0}\} \quad \text{und} \quad \text{Bild}(F) := \{F(\vec{x}) \mid \vec{x} \in \mathbb{R}^m\} .$$

22 Lemma Es sei A eine $n \times m$ -Matrix und $F : \vec{x} \mapsto A\vec{x}$ die zugehörige Matrixabbildung. Dann gelten die folgenden Aussagen.

1. $\text{Kern}(F)$ ist die Lösungsmenge des homogenen Gleichungssystems $A\vec{x} = \vec{0}$.
2. $\text{Bild}(F)$ ist das Erzeugnis der Spaltenvektoren von A .
3. Es gilt $F(r\vec{x}) = rF(\vec{x}) \quad \forall r \in \mathbb{R}, \forall \vec{x} \in \mathbb{R}^m$.
4. Es gilt $F(\vec{x} + \vec{y}) = F(\vec{x}) + F(\vec{y}) \quad \forall \vec{x}, \vec{y} \in \mathbb{R}^m$.
5. Für $\vec{a} \in \mathbb{R}^n$ und $\vec{w} \in \text{Kern}(F)$ gilt $F(\vec{a} + \vec{w}) = F(\vec{a})$.
6. Aus $F(\vec{x}) = F(\vec{y})$ folgt $\vec{x} - \vec{y} \in \text{Kern}(F)$.

Das lässt sich alles leicht nachrechnen, von mir aus rechne ich hier nichts davon vor.

Wir können jetzt die zentrale Aussage über die Lösungsmenge eines linearen Gleichungssystems $A\vec{x} = \vec{b}$ hinschreiben:

23 Lemma Es sei \vec{a} eine Lösung des linearen Gleichungssystems $A\vec{x} = \vec{b}$. Dann ist die Lösungsmenge \mathbb{L} des Gleichungssystems gegeben durch

$$\mathbb{L} = \vec{a} + \text{Kern}(F) .$$

Dabei ist $F : \vec{x} \mapsto A\vec{x}$ die zugehörige Matrixabbildung.

Beweis. Es sei $F(\vec{a}) = \vec{b}$ und $\vec{w} \in \text{Kern}(F)$. Ein solches \vec{w} gibt es immer, denn $\text{Kern}(F)$ enthält zumindest den Nullvektor. Dann ist

$$F(\vec{a} + \vec{w}) = F(\vec{a}) + F(\vec{w}) = \vec{b} + \vec{0} = \vec{b} ,$$

also gehört jeder Vektor aus $\vec{a} + \text{Kern}(F)$ zur Lösungsmenge des Systems. Es sei nun \vec{c} eine weitere Lösung des Systems, also $F(\vec{c}) = \vec{b} = F(\vec{a})$. Dann ist $\vec{c} - \vec{a} \in \text{Kern}(F)$, also $\vec{c} \in \vec{a} + \text{Kern}(F)$. \square

Damit haben wir die Sache im Griff! Zur $n \times m$ -Matrix A betrachten wir die zugehörige Matrixabbildung F . Das lineare Gleichungssystem $A\vec{x} = \vec{b}$ ist nur lösbar für $\vec{b} \in \text{Bild}(F)$, und das ist keine wilde Teilmenge des \mathbb{R}^n , sondern ein Erzeugnis. Für jedes dieser $\vec{b} \in \text{Bild}(F)$ ist die Lösungsmenge von $A\vec{x} = \vec{b}$ von der Form $\vec{a} + \text{Kern}(F)$, und $\text{Kern}(F)$ ist auch keine wilde Menge, sondern ebenfalls ein Erzeugnis. In letzter Konsequenz wird der Ursprungsraum \mathbb{R}^m durch F eingeteilt in

gleich große Schichten oder Fasern, und eine Schicht enthält alle Vektoren mit dem gleichen Bild unter F .

Jetzt wird dir ein **Beispiel** die ganze Herrlichkeit näher bringen und dir die Furcht vor den abstrakten Ausführungen nehmen. Es sei

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

Wir bilden die zugehörige Matrixabbildung $F : \vec{x} \mapsto A\vec{x}$. Dann ist

$$\text{Bild}(F) = \left\langle \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 2 \\ 1 \end{pmatrix} \right\rangle = \left\langle \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\rangle = \mathbb{R}^2$$

und

$$\text{Kern}(F) = \left\{ \begin{pmatrix} -2t \\ -t \\ t \end{pmatrix} \mid t \in \mathbb{R} \right\} = \left\langle \begin{pmatrix} -2 \\ -1 \\ 1 \end{pmatrix} \right\rangle .$$

Vektoren, deren z -Komponente = 0 ist, bleiben in gewissem Sinn fest. Wir stellen uns den \mathbb{R}^2 als die im \mathbb{R}^3 enthaltene xy -Ebene vor:

$$\begin{pmatrix} x \\ y \\ 0 \end{pmatrix} \xrightarrow{F} \begin{pmatrix} 1 & 0 & 2 \\ 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ 0 \end{pmatrix} = \begin{pmatrix} x \\ y \end{pmatrix}$$

Die Lösungsmenge von $A\vec{x} = \vec{b}$ ist also gegeben durch

$$\begin{pmatrix} b_1 \\ b_2 \\ 0 \end{pmatrix} + \text{Kern}(F) .$$

Jeder Raumpunkt liegt auf einer Parallelen zu der Geraden durch den Nullpunkt, die durch $\text{Kern}(F)$ gegeben ist, und (genau) alle Punkte jeder dieser Geraden gehen auf den gleichen Bildpunkt. Geometrisch handelt es sich bei dieser Abbildung um eine Scherung parallel zur xy -Ebene und anschließender Projektion parallel zur z -Achse in die xy -Ebene.

17.8 Lineare Abbildungen

Eine Abbildung $F : \mathbb{R}^m \rightarrow \mathbb{R}^n$, die die Aussagen 3 und 4 von Lemma 22 erfüllt, heißt **lineare Abbildung**. Diese beiden Eigenschaften garantieren dafür, dass F in gewissem Sinne die Struktur des \mathbb{R}^m erhält. Die Punkte zu $\vec{0}$, $\vec{x} \neq \vec{0}$, und $r\vec{x}$ liegen im \mathbb{R}^m auf einer Geraden durch den Nullpunkt, und die Bildpunkte $\vec{0}$, $F(\vec{x})$ und $F(r\vec{x}) = rF(\vec{x})$ liegen wieder auf einer Geraden des \mathbb{R}^n . Überhaupt gilt

$$F(\vec{a} + r\vec{b}) = F(\vec{a}) + rF(\vec{b})$$

für beliebige $\vec{a}, \vec{b} \in \mathbb{R}^m$ und $r \in \mathbb{R}$. Deshalb ist das Bild einer Geraden des \mathbb{R}^m wieder eine Gerade - oder ein einziger Punkt.

Schauen wir nun auf die zweite Bedingung. Die Punkte zu $\vec{0}$, \vec{x} , $\vec{x} + \vec{y}$ und \vec{y} bilden gewöhnlich so etwas wie ein Parallelogramm im \mathbb{R}^m , und die Bildpunkte verhalten sich wieder so. Allerdings können alle vier Bildpunkte auf einer Geraden liegen oder sogar der Nullvektor sein.

Eine besonders wichtige Eigenschaft linearer Abbildungen will ich in einem Lemma notieren:

24 Lemma *Es sei $F : \mathbb{R}^m \rightarrow \mathbb{R}^n$ eine lineare Abbildung und $\vec{x}_1, \dots, \vec{x}_k \in \mathbb{R}^m$. Dann ist*

$$F \left(\sum_{i=1}^k r_i \vec{x}_i \right) = \sum_{i=1}^k r_i F(\vec{x}_i) ,$$

eine Linearkombination der \vec{x}_i wird also auf die entsprechende Linearkombination der $F(\vec{x}_i)$ abgebildet.

17.9 Einige Maple-Befehle aus dem Paket linalg

```

Paket linalg laden:
> with(linalg):
Vektoren definieren
(Vorsicht, maple schreibt Zeilen, es sind aber Spaltenvektoren!):
> a:=vector(3,[1,2,3]); b:=vector(3,[4,5,6]); c:=vector(3,[7,8,9]);
Vektoren zu einer Matrix M zusammenfassen:
> M:=augment(a,b,c);
Gauss-Algorithmus anwenden
(Ergebnis: die Vektoren sind nicht l.u.. Siehst du, warum nicht?).
> gausselim(M);
Zur Bestaetigung LGS ra+sb+tc=0 loesen:
es gibt nicht nur die Loesung mit r=s=t=0.
> o:=vector(3,[0,0,0]); linsolve(M,b);
Vektor c als Linearkombination von a und b schreiben: es ist c=-a+2*b.
> linsolve( augment(a,b), c);
Ebenso b durch a und c sowie a durch b und c ausdruecken (das geht nicht immer!):
> linsolve( augment(a,c), b); linsolve( augment(b,c), a);
LGS in Vektor-Matrix-Schreibweise ueberfuehren:
> lgs:={3*x+4*y+5*z=4, x-y-z=-2, -x+y+3*z=11};
> A:=genmatrix(lgs,[x,y,z],'e'); print(e);
Und umgekehrt: aus der Vektor-Matrix-Form das LGS:
> B:=matrix(2,3,[[1,2,3],[4,5,6]]); f:=vector(2,[7,8]);
> geneqns(B,[x,y,z],f);
Produkt Matrix*Vektor berechnen
(praktisch etwas fuer x,y,z in die linke Seite des LS einsetzen und die
zugehoerige rechte Seite bestimmen):
> evalm(A&*[2,-3,1]);
zugehoeriges LGS loesen:
> linsolve(A,%);
>

```


Kapitel 18

Analytische Geometrie des Raumes

18.1 Die Beziehung zwischen dem \mathbb{R}^3 und dem Anschauungsraum

Führt man im Raum ein Koordinatensystem ein, gehört zu jedem Raumpunkt ein Koordinaten-tripel (x, y, z) . Schreibt man dieses Koordinatentripel als Spalte, und das ist ja keine große Sache, hat man einen Spaltenvektor des \mathbb{R}^3 vor sich. Unseren $\vec{x} \in \mathbb{R}^3$ entsprechen also in natürlicher Weise Punkte des Raumes.¹

Die hier skizzierte Beziehung zwischen Vektoren und Punkten erweist sich als außerordentlich fruchtbar, denn einer Vektormenge der Form $\vec{a} + \langle \vec{b} \rangle$ entspricht für $\vec{b} \neq \vec{0}$ eine Gerade im Raum. Ihr sogenannter Aufpunkt ist der Punkt, der zum Vektor \vec{a} gehört, und ihre Richtung ist durch den Pfeil beschrieben, der zum Vektor \vec{b} gehört. Analog gehört zu $\vec{a} + \langle \vec{b}, \vec{c} \rangle$ für linear unabhängige \vec{b}, \vec{c} eine Ebene im Raum. Das ist schon erstaunlich genug, aber es ist längst nicht alles. Wenn man nach der Länge des Pfeils fragt, der zu einem Vektor \vec{a} gehört, und nach dem Winkel, den die Pfeile zu \vec{a} und \vec{b} einschließen, wendet man den Satz des Pythagoras und den Kosinussatz an, und nur jemand, der überhaupt kein Gespür für Mathematik hat, entdeckt dann nicht das Skalarprodukt. So kommt es dann, dass man im \mathbb{R}^n scheinbar unmotiviert das Skalarprodukt definiert und den Betrag $|\vec{a}|$ des Vektors \vec{a} sowie den von $\vec{a} \neq \vec{0}$ und $\vec{b} \neq \vec{0}$ eingeschlossenen Winkel γ wie folgt festlegt:

$$|\vec{a}| := \sqrt{\vec{a} * \vec{a}} \quad \text{und} \quad \cos(\gamma) := \frac{\vec{a} * \vec{b}}{|\vec{a}| \cdot |\vec{b}|} \quad (18.1)$$

Dadurch hat man im \mathbb{R}^n eine Geometrie eingeführt, und für $n = 2, 3$ passen die Begriffe dieser Geometrie genau zu unseren geometrischen Vorstellungen.

Um es ganz deutlich zu sagen: Bei dem, was wir hier tun, handelt es sich um harte Tatsachen, solange wir über die Vektoren des \mathbb{R}^3 oder meinetwegen sogar des \mathbb{R}^n reden. Sobald geometrische Begriffe ins Spiel kommen, betreten wir schwankenden Boden. Was „der Raum“ der Geometer ist, ist eine sehr schwierige Frage; damit können wir uns leider nicht befassen. Wir werden in unserem euklidischen Vektorraum \mathbb{R}^3 arbeiten und bei dem, was wir tun, unbekümmert geometrische Sprache verwenden und uns bei unserer Arbeit von unserer geometrischen Vorstellung leiten und inspirieren lassen. Und damit werden wir sehr gut fahren und auch noch anschaulich gegebene geometrische Probleme lösen können.

Ein Hinweis: Die Bilder, die ich sonst an die Tafel male, kann ich dir hier nicht bieten. Die Zeit habe ich einfach nicht. Zeichne dir stets selbst Skizzen!

¹Diese Entsprechung wird vermittelt durch das gewählte Koordinatensystem. Das sage ich der Vollständigkeit halber dazu.

18.2 Gleichung einer Ebene

Wenn wir eine Gleichung der Form

$$ax + by + cz = d \quad (18.2)$$

lösen, bei der nicht alle Vorfaktoren $a, b, c = 0$ sind, haben wir es mit einem 1×3 -System zu tun. Zwei Parameter sind frei wählbar, und es ist für dich eine Kleinigkeit, im konkreten Fall die Lösungsmenge der Gleichung hinzuschreiben. Sie hat die Form

$$E = \vec{a} + W \quad \text{mit} \quad W = \langle \vec{b}, \vec{c} \rangle, \quad \vec{b}, \vec{c} \text{ linear unabhängig.} \quad (18.3)$$

Im Raum gehört zu E eine Ebene mit Aufpunkt A (zu \vec{a}), und zu W gehört eine zu dieser Ebene parallele Ebene durch den Nullpunkt des Koordinatensystems.

Wir bemerken, dass wir die Gleichung 18.2 auch in der Form

$$\vec{n} * \vec{x} = d \quad (18.4)$$

schreiben können, dabei ist \vec{n} der Vektor, dessen Komponenten die Vorfaktoren der Variablen in Gleichung 18.2 sind. Was hat es mit diesem Vektor \vec{n} auf sich? Nach der Gleichung 18.4 hat $\vec{n} * \vec{x}$ für jedes $\vec{x} \in E$ denselben Wert, nämlich d . Das ist schon erstaunlich genug. Insbesondere ist $\vec{n} * \vec{a} = d$, denn $\vec{a} \in E$. Damit ist für jedes $\vec{w} \in W$

$$d = \vec{n} * (\vec{a} + \vec{w}) = \vec{n} * \vec{a} + \vec{n} * \vec{w} = d + \vec{n} * \vec{w} \quad ,$$

und daraus folgt sofort

$$\vec{n} * \vec{w} = 0 \quad \forall \vec{w} \in W \quad . \quad (18.5)$$

Das heißt nichts anderes als dass die Pfeile zu den Vektoren aus W alle zu dem Pfeil zu \vec{n} senkrecht sind. Die von \vec{n} erzeugte Gerade durchstößt die (parallelen!) Ebenen W und E senkrecht!² Deshalb heißt der Vektor \vec{n} aus Gleichung 18.4 ein **Normalenvektor** der Ebenen E und W .

Hier ist nun eine Warnung am Platze. Die Pfeile, die zu \vec{n} und zu einem Punkt $\vec{x} \in E$ gehören, bilden keineswegs rechte Winkel, das trifft nur für $\vec{a} \in W$ zu, und dann ist $d = 0$.

18.3 Ebenengleichungen sind sehr nützlich

So seltsam es klingt, die Gleichung 18.2, deren Lösungsmenge die Ebene E aus 18.3 ist, ist häufig viel praktischer als die Vektormenge selbst. Nehmen wir an, wir wollen den Schnittpunkt von E mit der Geraden $g : \vec{x}(t) = \vec{e} + t\vec{f}$ bestimmen. Wenn wir die explizite Darstellung von E in Gleichung 18.3 verwenden, bekommen wir eine Vektorgleichung und müssen ein 3×3 -System lösen. Setzen wir aber $\vec{x}(t)$ einfach für \vec{x} in Gleichung 18.4 ein, müssen wir nur noch t ausrechnen; es handelt sich um ein 1×1 -System.

Noch deutlicher fällt der Vorteil aus, wenn man den Schnitt zweier Ebenen berechnen will. Verwendet man die explizite Darstellung, muss man ein 3×4 -System lösen. Das ist nicht weiter tief Sinnig, aber ziemlich lästig. Da stellt man besser Gleichungen der beiden Ebenen zu einem 2×3 -System zusammen.

Freilich muss man sich eine Gleichung der Ebene unter Umständen erst beschaffen. Ist die Ebene explizit in der Form 18.4 gegeben, bestimmt man \vec{n} , indem man das Gleichungssystem

$$\vec{b} * \vec{n} = 0, \quad \vec{c} * \vec{n} = 0$$

für \vec{n} löst. Hat man \vec{n} , braucht man bloß noch $d = \vec{n} * \vec{a}$ zu bilden für einen beliebigen Vektor der Ebene - es kommt ja für alles dasselbe heraus.

²Eigentlich müsste ich hier sagen: durchstößt die zu E und W gehörenden Ebenen senkrecht. Wenn man das immer korrekt sagen will, wird der Text recht sperrig, und das ist die Sache nicht wert. Wir wissen ja, wie es gemeint ist!

Kennt man drei Punkte der Ebene, meinetwegen die zu $\vec{p}, \vec{q}, \vec{u}$, löst man das System

$$(\vec{p} - \vec{u}) * \vec{n} = 0, \quad (\vec{q} - \vec{u}) * \vec{n} = 0$$

für \vec{n} . Eine Gleichung für die Ebene ist dann $\vec{n} * \vec{x} = \vec{n} * \vec{u}$.

18.4 Abstandsaufgaben

Stellen wir uns die **Aufgabe, den Abstand des Nullpunktes von einer Geraden g zu bestimmen**. Der allgemeine Punkt der Geraden sei gegeben durch $\vec{x}(t) = \vec{a} + t\vec{b}$. Dem Nullpunkt am nächsten liegt der Punkt $\vec{x}(t)$, der zu \vec{b} orthogonal ist³. Daraus ergibt sich die Gleichung

$$(\vec{a} + t\vec{b}) * \vec{b} = 0 \quad ,$$

und die löst man bequem im Kopf: Es ist $t = -(\vec{a} * \vec{b}) / (\vec{b} * \vec{b})$. Damit haben wir den Fußpunkt des Lotes vom Nullpunkt auf g gefunden, die Länge des Lotes ist der gesuchte Abstand.

Beinahe noch einfacher ist es, den **Abstand des Nullpunktes von einer Ebene E zu bestimmen**. Es sei E durch die Gleichung $\vec{n} * \vec{x} = d$ gegeben. Der Fußpunkt des Lotes vom Nullpunkt auf die Ebene liegt auf der von \vec{n} erzeugten Geraden, er ist also von der Form $t\vec{n}$ für ein $t \in \mathbb{R}$. Da der Fußpunkt ein Punkt der Ebene ist, muss $\vec{n} * (t\vec{n}) = d$ sein. Es ergibt sich $t = d / (\vec{n} * \vec{n})$, und man ist praktisch fertig. Der gesuchte Abstand ist

$$\left| \frac{d}{\vec{n} * \vec{n}} \vec{n} \right| = \frac{|d|}{|\vec{n}|} \quad .$$

Am Vorzeichen von d erkennt man, ob \vec{n} vom Nullpunkt aus zur Ebene hin oder von ihr weg zeigt.

Wir wollen nun einen Punkt \vec{p} an der Ebene E mit der Gleichung $\vec{n} * \vec{x} = 0$ **spiegeln**. Das geht recht einfach, wenn man den folgenden Kunstgriff benutzt: Man zerlegt \vec{p} in eine Summe

$$\vec{p} = \vec{p}_{\parallel} + \vec{p}_{\perp} \quad \text{mit } p_{\parallel} \in E \text{ und } \vec{p}_{\perp} \in \langle \vec{n} \rangle \quad .$$

Dann ist $\vec{p}_{\perp} = t\vec{n}$, und das t kann man leicht ausrechnen. Man multipliziert die obige Gleichung skalar mit \vec{n} . Der Summand aus E fällt dann weg, und es ergibt sich $t = (\vec{p} * \vec{n}) / (\vec{n} * \vec{n})$. Nun kann man den Bildpunkt \vec{p}' hinschreiben:

$$\vec{p}' = \vec{p} - 2 \frac{\vec{p} * \vec{n}}{\vec{n} * \vec{n}} \vec{n} \quad .$$

18.5 Kugeln

Wenn jemand von der Kugel um den Nullpunkt mit dem Radius r spricht, meint er gewöhnlich die Kugelschale, das heißt die Lösungsmenge der Gleichung

$$\vec{x} * \vec{x} = r^2 \quad . \tag{18.6}$$

Manchmal ist aber auch die Vollkugel gemeint, das ist dann die Lösungsmenge der Ungleichung

$$\vec{x} * \vec{x} \leq r^2 \quad . \tag{18.7}$$

Geometrisch gesprochen ist die Kugel um den Nullpunkt mit Radius r die Menge aller Punkte, deren Entfernung vom Nullpunkt r ist (Kugelschale) bzw. höchstens r ist (Vollkugel), und diese Bedingungen werden genau durch die Gleichungen abgebildet. Die Gleichung für die Kugel um \vec{m} mit Radius r ist entsprechend

$$(\vec{x} - \vec{m}) * (\vec{x} - \vec{m}) = r^2, \quad \text{ausgeschrieben } (x - m_1)^2 + (y - m_2)^2 + (z - m_3)^2 = r^2 \quad . \tag{18.8}$$

³Denke an die Skizze, die du zeichnen sollst, und vergiss den Nullpunkt des Systems dabei nicht.

Das ist nicht weiter geheimnisvoll.

Interessant für uns sind die Tangentialebenen einer Kugel. Die Gleichung der Ebene, die die Kugel um den Nullpunkt mit dem Radius r im Kugelpunkt \vec{a} berührt, enthält natürlich \vec{a} , darüber hinaus ist \vec{a} als Normalenvektor der Ebene geeignet. Also ist

$$\vec{a} * \vec{x} = \vec{a} * \vec{a} \quad ,$$

und die rechte Seite ist $= r^2$, da \vec{a} ja Kugelpunkt ist. Somit erhalten wir als Gleichung der Ebene, die die Kugel um den Nullpunkt mit dem Radius r im Kugelpunkt \vec{a} berührt, die erstaunlich einfache Gleichung

$$\vec{a} * \vec{x} = r^2 \quad . \quad (18.9)$$

18.6 Aufgaben

1. Nimm unseren Würfel, dessen Zentrum der Nullpunkt ist, der die Kantenlänge 2 hat und dessen Kanten parallel zu den Koordinatenachsen sind. Setze auf das obere Quadrat ein ordentliches Walmdach, dessen First parallel zur x -Achse ist. Der First und die vier schrägen Kanten, die die Enden des Firstes mit den Ecken des oberen Quadrates verbinden, sollen gleich lang sein. Vorsicht, es gibt eine ganze Schar dieser Walmdächer! Setze nun dieses Walmdach auch auf die Vorderseite des Würfels, die in der Ebene $y = -1$ liegt, und zwar so, dass die neue Firstkante parallel zur z -Achse verläuft. Stelle die Maße des Walmdachs nun so ein, dass das Fünfeck, das von den Enden des Dachfirstes, den Ecken $(1| - 1|1)^t$ und $(-1| - 1|1)^t$ des Würfels und dem oberen Endpunkt des Firstes auf der Vorderseite gebildet wird, in einer Ebene liegt. Rechnerisch schaffst du das am einfachsten, wenn du dafür sorgst, dass die obere Firstecke des neuen Walmdachs, der Mittelpunkt der Würfelkante von $(-1| - 1|1)^t$ bis $(1| - 1|1)$ und der Mittelpunkt des Firstes des alten Walmdachs auf einer Geraden liegen.

Hinweis: Hier geht es um die Konstruktion des Dodekaeders, eines der fünf platonischen Körper. Siehe Abbildung 18.1 auf Seite 100.

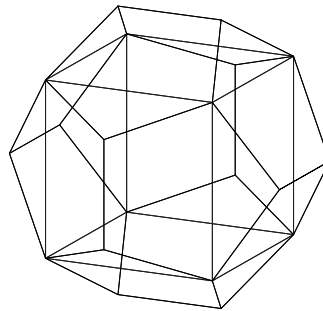


Abbildung 18.1: Dodekaeder als Würfel mit aufgesetzten Walmdächern

2. Nimm wieder den Würfel der vorigen Aufgabe. Du startest in der Würfecke $(1| - 1|1)^t$ und du läufst auf Kanten des Würfels einmal um den Würfel, bis du wieder in der Startecke bist. Dabei darfst du außer dem Startpunkt keinen Punkt mehrfach durchlaufen, und du darfst weder den Punkt $(1|1|1)^t$ noch den Punkt $(-1| - 1| - 1)^t$ durchlaufen. Zeige, dass die Mittelpunkte der durchlaufenen Kanten in einer Ebene liegen und zeichne die Schnittfigur des Würfels mit dieser Ebene.

Kapitel 19

Geometrie im \mathbb{R}^4

19.1 Theorie

In Kapitel 18 hast du gesehen, dass wir gewisse Mengen von Elementen des \mathbb{R}^3 als Geraden, Ebenen, Kugeln des Raumes deuten können und dass wir dann geometrische Aussagen über Objekte im Raum als algebraische Aussagen über diese Mengen formulieren können. Bei den algebraischen Aussagen spielt in der Regel überhaupt keine Rolle, dass es sich um 3-tupel handelt; deshalb kann man diese Konstruktionen auch im \mathbb{R}^n durchführen und bekommt so eine Geometrie im n -dimensionalen Raum. Für $n = 4$ wollen wir uns das einmal ansehen.

Es sei also $V = \mathbb{R}^4$. Die Elemente von V stehen für die Punkte des Raumes, manchmal spreche ich auch einfach vom Punkt $\vec{x} \in V$. Das Erzeugnis $\langle \vec{b} \rangle$ eines Vektors $\vec{b} \neq \vec{0}$ „ist“ dann eine Gerade durch den Ursprung des Koordinatensystems, $\vec{a} + \langle \vec{b} \rangle$ ist die dazu parallele Gerade durch \vec{a} . Für linear unabhängige $\vec{b}, \vec{c} \in V$ ist $\vec{a} + \langle \vec{b}, \vec{c} \rangle$ eine Ebene durch \vec{a} und so weiter. Längen und Winkel messen wir mit Hilfe des Skalarprodukts mit den üblichen Formeln, damit haben wir auch Orthogonalität zur Hand - rechnerisch ist das alles kein Problem.

Mit der geometrischen Vorstellung ist das natürlich nicht so einfach. Wir müssen unser xyz -System um eine vierte Achse, die w -Achse, erweitern, und die steht auf allen anderen drei Achsen senkrecht. Das ist etwas gewöhnungsbedürftig. Und wenn wir für $\vec{n} \neq \vec{0}$ die Lösungsmenge der Gleichung $\vec{n} * \vec{x} = d$ bilden, erhalten wir eine Menge H mit drei wählbaren Parametern, also ein dreidimensionales¹ Gebilde, das wir in der Form

$$H = \vec{a} + \langle \vec{b}_1, \vec{b}_2, \vec{b}_3 \rangle$$

schreiben können mit linear unabhängigen Vektoren $\vec{b}_1, \vec{b}_2, \vec{b}_3$. Ein solches Gebilde nennen wir eine **Hyperebene** des \mathbb{R}^4 .

Nützlich ist die Vorstellung, unser \mathbb{R}^3 sei in $V = \mathbb{R}^4$ eingebettet, unseren Raumpunkt $(x|y|z)^t$ finden wir dann als Punkt $(x|y|z|0)^t$ von V wieder. Wir leben dann quasi in der Hyperebene mit der Gleichung $w = 0$, und für einen echten Bewohner des \mathbb{R}^4 wäre unser Raum platt wie eine Fläche unseres Raumes für uns platt ist.

Es gibt viele merkwürdige Dinge in diesem \mathbb{R}^4 . Zum Beispiel haben die Ebenen $\langle \vec{e}_1, \vec{e}_2 \rangle$ und $\langle \vec{e}_3, \vec{e}_4 \rangle$ nur den Nullpunkt gemeinsam, und die Gerade $\vec{e}_3 + \langle \vec{e}_4 \rangle$ hat mit der ersten Ebene keinen Punkt gemeinsam, obwohl sie keineswegs parallel zu dieser Ebene ist.

Jeder, der so etwas sieht, fragt sich wohl, ob es einen vierdimensionalen Raum denn „gibt“. Nun, als Gegenstand der Mathematik „gibt“ es ihn; er ist nicht weniger real als ein gleichseitiges Dreieck mit der Kantenlänge 1 oder eine Kugel mit Radius 2. Und wenn man einen Zufallsversuch

¹Einen sauberen Dimensionsbegriff für lineare Gebilde lernst du im Kapitel über Vektorräume. Der allgemeine Dimensionsbegriff, der praktisch auf die Anzahl der frei wählbaren Parameter hinausläuft, stammt von Riemann; er ist intuitiv einleuchtend, liegt aber weit außerhalb unserer Reichweite. Ich denke, die Sprechweise kann ich ruhig benutzen; du wirst dir schon etwas Richtiges darunter vorstellen.

durchführt, bei dem als Ergebnis ein Satz von vier Zahlen herauskommt, ist es nützlich, einen vierdimensionalen Raum zu haben, damit man das Ω des Versuchs überhaupt hinschreiben kann. Um hier Wahrscheinlichkeiten zu bestimmen, muss man in der Regel das Volumen vierdimensionaler Körper berechnen!

Das **Volumen** eines 4-Körpers ist im Prinzip kein Problem; man verwendet unser Kartoffelkörperverfahren. Wenn sich der Körper meinetwegen in w -Richtung zwischen $w = a$ und $w = b$ erstreckt, bestimmt man für jedes c zwischen a und b den Schnitt des Körpers mit der Hyperebene $w = c$. Diese Schnittmenge ist ein 3-Körper, sein Volumen sei durch $F(c)$ gegeben. Das Volumen des 4-Körpers ist dann

$$\int_a^b F(c) dc .$$

Damit ist freilich nicht gesagt, dass die konkrete Berechnung einfach durchzuführen ist. Du wirst aber einige Beispiele sehen. Ein erstes Beispiel nenne ich schon: Das Volumen eines 4-Quaders ist das Produkt seiner Kantenlängen.

19.2 Aufgaben

1. Es sei K die 4-Kugel um den Nullpunkt mit Radius 1. Dann ist \vec{e}_4 ein Punkt von K . Die Hyperebene $H : \vec{e}_4 * \vec{x} = 1$ hat mit K nur den Punkt \vec{e}_4 gemeinsam, und jede Gerade in H , die durch \vec{e}_4 verläuft, ist orthogonal zu der Strecke, die den Nullpunkt mit \vec{e}_4 verbindet. Weise nach, dass dies zutrifft!
Anmerkung: H ist die Tangentialhyperebene an K im Punkt \vec{e}_4 .
2. Nimm den Würfel mit der Kantenlänge 2 und achsenparallelen Kanten und Zentrum im Nullpunkt unseres gewöhnlichen xyz -Systems. Seine acht Ecken haben die Koordinaten

$$(\pm 1 | \pm 1 | \pm 1) ,$$

wie du dir leicht klarmachst. Diesen Würfel wollen wir als Boden einer 4-Pyramide benutzen. Um die Sache rechnerisch so einfach wie nur möglich zu machen, legen wir die Spitze S der 4-Pyramide in den Nullpunkt und den Boden in die Hyperebene $H : w = 3$. Die acht Ecken des Bodens der 4-Pyramide sind dann

$$(\pm 1 | \pm 1 | \pm 1 | 3) .$$

Für $0 < t < 3$ schneidet die Hyperebene $H_t : w = t$ die Verbindungsstrecken von S und den acht Ecken des Bodens in den Punkten

$$\left(\pm \frac{t}{3} | \pm \frac{t}{3} | \pm \frac{t}{3} | t \right) ,$$

und diese acht Punkte sind die Eckpunkte eines Würfels mit der Kantenlänge $\frac{2t}{3}$, also eines verkleinerten Abbilds des Ausgangswürfels, und der Würfel liegt in der Hyperebene H_t . Er ist gerade der Schnitt der Hyperebene H_t mit der 4-Pyramide. Nutze dies aus, um das Volumen der Pyramide zu berechnen.

3. Es seien zwei Punkte \vec{p}, \vec{q} und eine Hyperebene $H : \vec{x} * \vec{n} = d$ des \mathbb{R}^4 gegeben. Wie kann man entscheiden, ob diese Punkte auf verschiedenen Seiten der Hyperebene liegen? Und wenn sie auf verschiedenen Seiten der Hyperebene liegen, muss dann jede Kurve k , die \vec{p} und \vec{q} verbindet, die Hyperebene schneiden?
4. Wir waren ja dem Phänomen begegnet, dass man einen Punkt aus dem Inneren eines 3-Körpers mit einem Punkt außerhalb durch eine Kurve verbinden kann, die die Oberfläche des Körpers nicht schneidet. Nimm als Beispiel die Punkte mit den Koordinaten $(1|0|0)$ und $(-1|0|0)$ des \mathbb{R}^3 . Man kann sie im \mathbb{R}^3 nicht verbinden, ohne dabei durch die yz -Ebene zu

gehen. Wir betten den \mathbb{R}^3 in den \mathbb{R}^4 ein und verbinden die Punkte $(1|0|0|0)$ und $(-1|0|0|0)$ durch eine Kurve k , die die Ebene $w = 0, x = 0$ nicht schneidet. Weise nach, dass die Kurve

$$k : \vec{x}(t) = \begin{pmatrix} \cos(t) \\ 0 \\ 0 \\ \sin(t) \end{pmatrix}, \quad 0 \leq t \leq \pi$$

das Gewünschte leistet.

5. Wir stellen uns wieder vor, wir lebten in der Hyperebene $H : w = 0$ des \mathbb{R}^4 , das heißt, unser Raum sei in den \mathbb{R}^4 eingebettet. Wir heben die 3-Kugel um den Nullpunkt mit Radius 1 um 5 in w -Richtung an, dann haben wir eine 3-Kugel K , die in der Hyperebene $H_5 : w = 5$ des \mathbb{R}^4 liegt. Wir wollen einen 4-Körper untersuchen, den man als 4-Kegel bezeichnen könnte. Dazu verwenden wir K als Boden und den Nullpunkt als Spitze. Zum Körper gehören alle Strecken, die den Nullpunkt mit einem Punkt der Kugel verbinden. Liegt der Kugelpunkt im Inneren von K , gehören die Punkte der Strecke außer Anfangs- und Endpunkt zum Inneren des Kegels. Die Strecken, die einen Punkt der Kugelschale mit dem Nullpunkt verbinden, gehören zum Rand des Kegels.
- Mache dir die Situation gehörig klar.
 - Schreibe K und den Rand von K als Punktmenge des \mathbb{R}^4 hin.
 - Welche Punkte hat die Gerade, die den Nullpunkt mit dem Punkt $(0|0|0|5)^t$ verbindet, mit K gemeinsam?
 - Berechne das Volumen des Kegels. Welchen Anteil des umgebenden 4-Zylinders füllt der Kegel?
 - Ist die Spitze des Kegels von jedem Punkt des Bodens gleich weit entfernt?
 - Sind alle „Mantellinien“ des Kegels gleich lang?
 - Finde eine möglichst kleine 4-Kugel, die den Kegel enthält.
 - Finde eine möglichst große 4-Kugel, die in den Kegel hineinpasst.

Kapitel 20

Der \mathbb{R}^n als Vektorraum

20.1 Teilräume des \mathbb{R}^n

Ich zähle einige Teilmengen des \mathbb{R}^3 auf, die dir schon begegnet sind. Dabei sind \vec{a} , \vec{b} und \vec{n} feste Vektoren ungleich $\vec{0}$.

$$M_1 = \langle \vec{a} \rangle, \quad M_2 := \vec{a} + \langle \vec{b} \rangle, \quad M_3 = \{\vec{x} \mid \vec{x} * \vec{x} = 1\}, \quad M_4 = \{\vec{x} \mid \vec{n} * \vec{x} = 2\},$$

$$M_5 = \{\vec{x} \mid \vec{x} * \vec{n} = \frac{1}{2} \cdot |\vec{x}| \cdot |\vec{n}|\}, \quad M_6 = \{\vec{x} \mid x_1, x_2, x_3 \in \mathbb{Z}\}, \quad M_7 = \{\vec{0}\}$$

Für Vektoren ist entscheidend, dass man sie addieren und mit reellen Zahlen multiplizieren kann. Einige der Mengen oben haben die Eigenschaft, dass man ihre Elemente addieren und mit reellen Zahlen multiplizieren kann und dass dabei wieder Elemente der gleichen Menge herauskommen. Man sagt, solche Mengen seien abgeschlossen bezüglich der Addition und der Multiplikation mit reellen Zahlen, und man bezeichnet sie als Teilräume, denn sie sind selbst Vektorräume¹.

25 Definition Eine Teilmenge $W \neq \emptyset$ des \mathbb{R}^n heißt Teilraum des \mathbb{R}^n , in Zeichen $W \leq \mathbb{R}^n$, wenn gilt

1. Für $\vec{x}, \vec{y} \in W$ ist stets auch $\vec{x} + \vec{y} \in W$.
2. Für $\vec{x} \in W$ und $r \in \mathbb{R}$ ist stets auch $r\vec{x} \in W$.

Die Beispielmengen oben hast du hoffentlich alle erkannt. Nur M_1 und M_7 sind Teilräume, M_2 nur, wenn die Gerade durch den Nullpunkt geht, also \vec{a} Vielfaches von \vec{b} ist. Die Menge M_3 ist geometrisch eine Kugel vom Radius 1 um den Nullpunkt, beliebige Vielfache und Summen darf man keinesfalls bilden (manche schon!). Bei M_4 handelt es sich um eine Ebene, die nicht durch den Nullpunkt geht, da ist überhaupt nichts zu machen. Interessant ist M_5 , das ist Sebastians Kegel. Man darf Vielfache bilden, in der Regel aber keine Summen. Und M_6 ist ein Gitter; man darf Summen bilden, auch mit ganzen Zahlen multiplizieren, aber eben nicht mit $\sqrt{2}$.

Die wichtigsten Beispiele für Teilräume sind Erzeugnisse - in Wahrheit ist es sogar so, dass sich jeder Teilraum des \mathbb{R}^n als Erzeugnis eines geeigneten Systems von Vektoren darstellen lässt. Weil du mit formalen Schlüssen nicht sehr vertraut bist, zelebriere ich jetzt einmal einen richtigen Beweis.

26 Lemma Für $\vec{a}_1, \vec{a}_2, \dots, \vec{a}_m \in \mathbb{R}^n$ ist stets $\langle \vec{a}_1, \vec{a}_2, \dots, \vec{a}_m \rangle \leq \mathbb{R}^n$.

¹Ich verwende hier schon einmal die Sprechweise, damit du dich an den Begriff gewöhnst. Das Wesentliche ist schon gesagt, in einem Vektorraum kann man vernünftig addieren und Vielfache bilden. Standardbeispiel ist der \mathbb{R}^n .

Beweis. Wir bezeichnen das Erzeugnis mit W . Es sei $\vec{x}, \vec{y} \in W$. Dann ist $\vec{x} = \sum_{i=1}^m r_i \vec{a}_i$ und $\vec{y} = \sum_{i=1}^m s_i \vec{a}_i$ für geeignete reelle Zahlen $r_i, s_i, i = 1, 2, \dots, m$, und es gilt

$$\vec{x} + \vec{y} = \sum_{i=1}^m r_i \vec{a}_i + \sum_{i=1}^m s_i \vec{a}_i = \sum_{i=1}^m (r_i \vec{a}_i + s_i \vec{a}_i) = \sum_{i=1}^m (r_i + s_i) \vec{a}_i ,$$

und dies liegt wieder in W , denn die Summe reeller Zahlen ist wieder eine reelle Zahl. Analog ist für $t \in \mathbb{R}$

$$t\vec{x} = t \sum_{i=1}^m r_i \vec{a}_i = \sum_{i=1}^m t(r_i \vec{a}_i) = \sum_{i=1}^m (tr_i) \vec{a}_i \in W ,$$

denn es ist $tr_i \in \mathbb{R}$ für $i = 1, 2, \dots, m$. Damit ist gezeigt, dass für $\vec{x}, \vec{y} \in W$ und $t \in \mathbb{R}$ stets auch $\vec{x} + \vec{y} \in W$ und $t\vec{x} \in W$ ist. \square

Übungsaufgabe Es sei A eine $m \times n$ -Matrix und $F : \mathbb{R}^n \rightarrow \mathbb{R}^m$ die zugehörige Matrixabbildung. Beweise, dass $\text{Kern}(F) \leq \mathbb{R}^n$ und $\text{Bild}(F) \leq \mathbb{R}^m$ ist. Formuliere diese Aussagen auch um in Aussagen über lineare Gleichungssysteme.

20.2 Basen und Dimension eines Teilraums

Es sei W das Erzeugnis von Vektoren $\vec{a}_1, \vec{a}_2, \dots, \vec{a}_m$ des \mathbb{R}^n . Du könntest weitere Vektoren aus W als erzeugende Vektoren zu den \vec{a}_i hinzufügen, ohne dass das Erzeugnis größer würde, aber das wirst du nicht tun, wozu auch? Vielleicht könntest du aber welche von den \vec{a}_i weglassen, ohne dass das Erzeugnis kleiner würde, und das ist schon interessanter. Das geht nur, wenn $\vec{a}_1, \vec{a}_2, \dots, \vec{a}_m$ linear abhängig ist, wie du weißt. Sind die Vektoren linear unabhängig, ist kein erzeugender Vektor überflüssig, und wir haben es mit einem minimalen Erzeugendensystem von W zu tun. Ein solches minimales Erzeugendensystem heißt eine **Basis** von W . Für Basen von W gilt der folgende Satz.

27 Lemma *Alle Basen eines Teilraums W des \mathbb{R}^n haben die gleiche Anzahl von Elementen.*

Die Aussage des Satzes wird dich nicht überraschen, deshalb beweisen wir ihn nicht; das kostete uns viel Zeit. Man nennt die Anzahl der Elemente einer Basis von $W \leq \mathbb{R}^n$ die **Dimension** von W , in Zeichen $\dim(W)$.

Beispiele Der \mathbb{R}^n ist Teilraum von sich selbst, und $\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n$ ist eine Basis. Also ist $\dim(\mathbb{R}^n) = n$. Die Dimension einer Ursprungsgeraden im Raum ist 1, die Dimension einer Ebene durch den Nullpunkt ist 2.

Eine wichtige Eigenschaft von Basen solltest du schon kennen:

28 Lemma *Es sei $W \leq \mathbb{R}^n$, und es sei $\vec{a}_1, \dots, \vec{a}_m$ eine Basis von W . Dann ist jeder Vektor $\vec{x} \in W$ eindeutig als Linearkombination der \vec{a}_j darstellbar.*

Beweis. Jedes $\vec{x} \in W$ ist als Linearkombination der \vec{a}_j darstellbar, denn eine Basis von W ist ja insbesondere ein Erzeugendensystem von W . Wenn nun ein $\vec{x} \in W$ auf zwei Arten als Linearkombination der \vec{a}_j geschrieben werden kann, haben wir

$$\sum_{j=1}^m r_j \vec{a}_j = \vec{x} = \sum_{j=1}^m s_j \vec{a}_j ,$$

und daraus folgt

$$\sum_{j=1}^m (r_j - s_j) \vec{a}_j = \vec{0} .$$

Da $\vec{a}_1, \dots, \vec{a}_m$ als Basis ein linear unabhängiges Vektorsystem ist, muss $r_j - s_j = 0$ sein für $j = 1, \dots, m$. Damit ist die Behauptung bewiesen. \square

Kapitel 21

Der Vektorraum der Funktionen

$$I \rightarrow \mathbb{R}$$

21.1 Das Konzept des Vektorraums

Es sei I ein Intervall, meinetwegen $I = [0; 1]$ oder $I = [0, \pi]$. und es seien f und g zwei Funktionen, die jedem $x \in I$ je eine reelle Zahl $f(x)$ bzw. $g(x)$ zuordnen. Dann kann man zu f und g ihre Summenfunktion $f+g$ bilden; das ist die Funktion, die jedem $x \in I$ den Wert $f(x)+g(x)$ zuordnet. Genauso kann man zu f und $r \in \mathbb{R}$ die Funktion bilden, die jedem $x \in I$ den Wert $rf(x)$ zuordnet, und diese Funktion bezeichnen wir als rf .

Wie du siehst, kann man Funktionen addieren und Vielfache mit Zahlen bilden, sie verhalten sich in dieser Hinsicht genau so wie unsere Spalten des \mathbb{R}^n .

In der Mathematik wurde, recht spät übrigens, nämlich erst in der zweiten Hälfte des achtzehnten Jahrhunderts, der Begriff des Vektorraums geprägt. Von einem Vektorraum spricht der Mathematiker, wenn er eine Menge vor sich hat, deren Elemente er addieren und mit reellen Zahlen so multiplizieren kann, dass wieder Elemente der Menge dabei herauskommen, und wenn die Addition und die Multiplikation mit reellen Zahlen vernünftige Eigenschaften haben. Diese Eigenschaften sind in unserem Buch auf Seite 196f aufgelistet; im Kern besagen sie nur, dass man mit Vektoren so rechnen kann wie mit unseren Spalten. Insofern bilden die Funktionen, die I in \mathbb{R} abbilden, einen Vektorraum, wenn man mit ihnen so rechnet, wie ich es oben angegeben habe.

Es wird dir nicht leichtfallen, die Funktionen, die dir ja halbwegs vertraute Gegenstände mit zahlreichen Eigenschaften sind, als Vektoren anzusehen. Aber es ist der Mühe wert. Wir können nämlich alles, was wir am \mathbb{R}^n gelernt haben, auf den Vektorraum der Funktionen übertragen, und das wird uns überraschende Einsichten über Funktionen liefern. Es ist eine große Stärke der Mathematik, dass sie häufig Dinge verknüpft, die auf den ersten Blick völlig verschieden sind. Du sollst am Beispiel des Funktionenraums sehen, wie das funktioniert.

21.2 Lineare Unabhängigkeit im Funktionenraum

Wir sehen uns die Funktionen $p_k = (x \mapsto x^k)$ für $k = 0, 1, 2, \dots, n$ an. Sie sind sicherlich auf I definiert, wir können für x ja jede Zahl einsetzen. Was ist denn nun das Erzeugnis $\langle p_0, p_1, p_2, \dots, p_k \rangle$? Es enthält alle Linearkombinationen der p_k , also alle

$$\sum_{k=0}^n r_k p_k = \sum_{k=0}^n r_k (x \mapsto x^k) = \sum_{k=0}^n (x \mapsto r_k x^k) = (x \mapsto \sum_{k=0}^n r_k x^k) .$$

Siehst du, was dabei herausgekommen ist? Klar, die Menge aller ganzrationalen Funktionen vom Grad $\leq n$. Wenn man es sich noch einmal überlegt, ist das keine Überraschung. Steckt noch mehr dahinter? Ja, $p_0, p_1, p_2, \dots, p_n$ ist linear unabhängig. Das siehst du, wenn du eine solche

Linearkombination gleich dem Nullvektor $\vec{0}$ setzt. Der Nullvektor ist natürlich die Funktion, die jedem x den Wert 0 zuweist. Da steht dann

$$\sum_{k=0}^n r_k p_k = (x \mapsto \sum_{k=0}^n r_k x^k) = \vec{0} = (x \mapsto 0) .$$

Das heißt aber nichts anderes als

$$\sum_{k=0}^n r_k x^k = 0 \quad \forall x \in I ,$$

und dann müssen alle $r_k = 0$ sein, denn sonst stünde auf der linken Seite ein Polynom vom Grad $\leq n$, und das hat höchstens n Nullstellen.

Ich schreibe in ein Lemma, was wir gerade eingesehen haben.

29 Lemma Die Funktionen $(x \mapsto 1), (x \mapsto x), (x \mapsto x^2), \dots, (x \mapsto x^n)$ sind für jedes n linear unabhängig. Das bedeutet, dass der Teilraum der ganzrationalen Funktionen schon unendlichdimensional ist.

21.3 Skalarprodukt im Funktionenraum

Ich schreibe es dir einfach hin. Man braucht dabei die Grenzen des Intervalls I , es sei also $I = [a, b]$. Dann setzt man für Funktionen $f, g : I \rightarrow \mathbb{R}$ fest

$$f * g := \int_a^b f(x)g(x) dx . \quad (21.1)$$

Ich muss natürlich sicherstellen, dass das Integral auch existiert, indem ich zum Beispiel nur differenzierbare Funktionen zulasse. Aber mit den Feinheiten wollen wir uns nicht zu sehr plagen. Wenn man gescheite Funktionen nimmt, geht die Sache gut.

Dieses Skalarprodukt verhält sich so, wie du es vom Skalarprodukt gewohnt bist. Bedenke nun, dass du von Längen, Winkeln und Orthogonalität sprechen kannst, wenn du ein Skalarprodukt zur Verfügung hast. Du hast so etwas wie den Abstand zweier Funktionen f und g , nämlich $|f - g| = \sqrt{(f - g) * (f - g)}$, und das ist sehr nützlich.

Beispiel Für $I = [0, 1]$ ist $(x \mapsto x^2) * (x \mapsto x^3) = \frac{1}{6}$ und $|(x \mapsto x^2)| = \sqrt{\frac{1}{3}}$.

Aufgaben

1. Wende auf $(x \mapsto 1), (x \mapsto x), (x \mapsto x^2)$ unser Orthogonalisierungsverfahren an. Verwende dabei $I = [0, 1]$.
2. Integriere

$$\int_0^\pi \sin(nx) \sin(mx) dx$$

zweimal partiell. Beachte dabei, dass du nicht einen Faktor erst auf- und dann wieder ableitest. Leite einen zweimal auf und den anderen zweimal ab.

21.4 Fourier-Reihen

Ich beginne mit einem Lemma.

30 Lemma Für natürliche Zahlen n, m gilt

$$\int_0^\pi \sin(nx) \sin(mx) dx = \begin{cases} 0 & \text{für } m \neq n, \\ \frac{\pi}{2} & \text{für } n = m. \end{cases} \quad (21.2)$$

Der Beweis geschieht durch partielle Integration und durch Ausnutzung der Identität $\sin^2(x) + \cos^2(x) = 1$. Das ist nicht ganz einfach durchzurechnen, aber wenig erhellend. Wichtig für uns ist, was aus diesem Lemma folgt.

31 Lemma Die Funktionen

$$(x \mapsto \sin(x)), (x \mapsto \sin(2x)), \dots, (x \mapsto \sin(nx))$$

sind bezüglich des durch $f * g := \int_0^\pi f(x)g(x) dx$ definierten Skalarprodukts paarweise orthogonal, insbesondere also linear unabhängig. Man kann auch sagen, sie bilden einen n -dimensionalen Teilraum des Raums der Funktionen $[0; \pi] \rightarrow \mathbb{R}$.

Die praktische Bedeutung dieses Sachverhalts kannst du kaum überschätzen. Nimm eine (nicht zu wilde) Funktion $f : [0; \pi] \rightarrow \mathbb{R}$. Sie erzeugt zusammen mit den n Sinusfunktionen aus Lemma 31 einen Vektorraum, den ich mit V bezeichne. In aller Regel wird das f nicht im Erzeugnis der Sinusfunktionen liegen; du kannst dir also f als Pfeil vorstellen, der aus dem von den Sinusfunktionen erzeugten Gebilde herauschaut. Zerlege f in eine Summe $f = f_\perp + f_\parallel$, wobei f_\parallel im von den Sinusfunktionen erzeugten Gebilde liegt und f_\perp orthogonal zu diesem Gebilde ist. Anschaulich gesprochen wird f orthogonal in den von den Sinusfunktionen erzeugten Teilraum projiziert, und das Bild ist f_\parallel . Dieses f_\parallel ist eine Linearkombination der Sinusfunktionen, und die Vorfaktoren lassen sich leicht ausrechnen: Multipliziere die Gleichung

$$f = f_\perp + \sum_{k=1}^n r_k (x \mapsto \sin(kx)) \quad (21.3)$$

skalar mit $(x \mapsto \sin(kx))$! Wegen Lemma 30 und der Konstruktion von f_\perp ergibt sich dann

$$f * (x \mapsto \sin(kx)) = r_k (x \mapsto \sin(kx)) * (x \mapsto \sin(kx)) = r_k \cdot \frac{\pi}{2},$$

und daraus folgt

$$r_k = \frac{2}{\pi} \int_0^\pi f(x) \sin(kx) dx. \quad (21.4)$$

Wenn man n nur groß genug wählt, kann man f beliebig genau als Summe der Sinusfunktionen darstellen. Die Summe der Sinusfunktionen nennt man eine endliche **Fourier-Reihe** von f . Fourier-Reihen finden vielfältige Anwendungen, wir können uns leider nicht mehr damit befassen. Ich zeige dir wenigstens einige Beispiele. Du kannst dir auch selbst welche erzeugen!

```
> f:=x->x;
> a:=k->2/Pi*int( f(x)*sin(k*x), x=0..Pi);
> F:=(n,x)->sum( a(k)*sin(k*x), k=1..n);
> plot( F(10,x), x=0..Pi);
> plot( F(20,x), x=-Pi..2*Pi);
>
```


Kapitel 22

Kurven und Flächen im Raum

22.1 Parameterdarstellungen von Kurven und Flächen

Über Kurven und Flächen haben wir in unserem Analysisjahr schon gesprochen. Jetzt, wo du dich auch in der Geometrie und in der Linearen Algebra auskennst, kommen wir darauf noch einmal zurück.

Eine Raumkurve ist gewöhnlich durch eine Parameterdarstellung

$$\vec{p} : t \mapsto \begin{pmatrix} x(t) \\ y(t) \\ z(t) \end{pmatrix}$$

gegeben. Wir benutzen nur vernünftige Funktionen als Koordinatenfunktionen $x(t)$, $y(t)$ und $z(t)$, die man differenzieren kann, so oft wie man will.

Wenn du dir $\vec{p}(t)$ als Ort eines beweglichen Punktes zur Zeit t vorstellst, ist $\vec{p}'(t)$ der Geschwindigkeitsvektor des Punktes zur Zeit t . Mit Hilfe dieses Geschwindigkeitsvektors kann man zum Beispiel die Länge eines Stücks der Kurve ausrechnen.

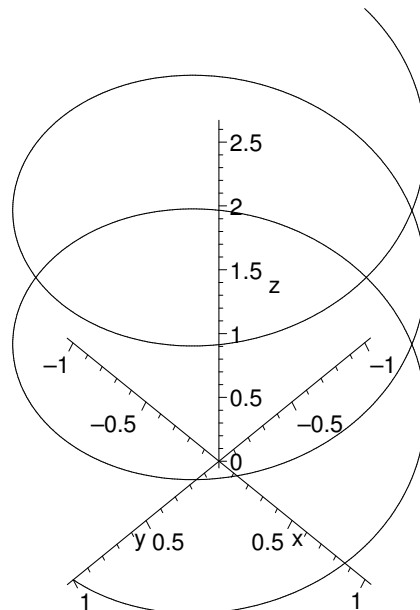


Abbildung 22.1: Schraubenlinie, $\vec{p}(t) = (\cos(t), \sin(t), \frac{1}{6}t)^t$

Um ein Flächenstück zu parametrisieren, braucht man zwei Parameter; die Anzahl der wirklich benötigten Parameter ist immer gleich der Dimension des Objekts. Nehmen wir als Beispiel einen Kreiskegel (siehe Abbildung 22.2). Die Punkte (s, t) selbst füllen ein Flächenstück der st -Ebene,

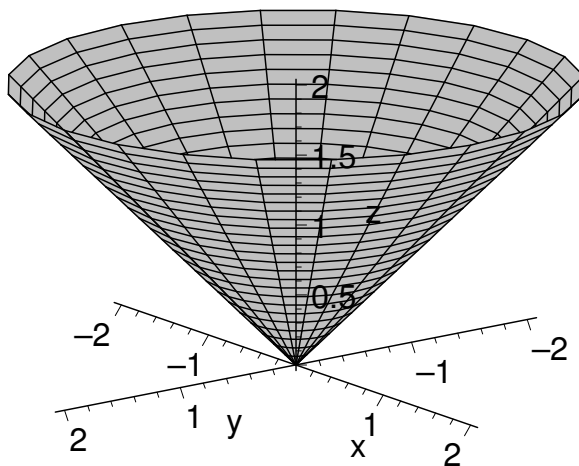


Abbildung 22.2: Kreiskegel, $\vec{p}(t, s) = (s \cos(t), s \sin(t), \frac{1}{6}s)^t$

man kann es als Karte des Kegels ansehen. Streng genommen darf man den Wert $s = 0$ nicht zulassen, zu jedem Flächenpunkt darf es nur einen Kartenpunkt geben.

Wir haben auch die Möglichkeit, den Kreiskegel durch die Gleichung

$$x^2 + y^2 = z^2$$

zu beschreiben.

22.2 Tangentialebene einer Fläche

Wir nehmen uns eine Fläche her, ein Paraboloid zum Beispiel, und betrachten einen festen Punkt der Fläche. Dann kann man beliebige Kurven nehmen, die in der Fläche verlaufen und durch diesen Punkt gehen, und ihre Tangentialvektoren in diesem Punkt berechnen. Das Erstaunliche ist, dass alle diese Tangentialvektoren in einer Ebene liegen, und das ist die Tangentialebene der Fläche in diesem Punkt.

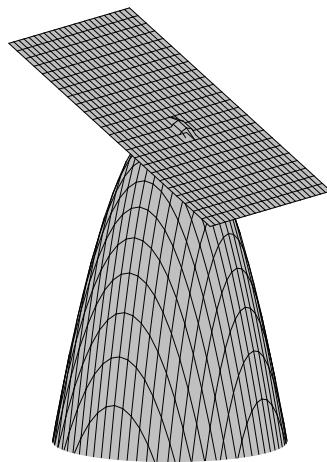


Abbildung 22.3: Paraboloid mit Tangentialebene

22.3 Regelflächen

Eine Ebene kann man sich so entstanden denken: Man beginnt mit einer Geraden und verschiebt eine andere Gerade längs der ersten Geraden. Die erste Gerade ist dann die Leitkurve der Ebene. Man kann nun als Leitkurve eine beliebige Kurve nehmen und eine Gerade entlang der Kurve verschieben. Dann bekommt man so etwas wie einen Zylindermantel. Du erinnerst dich an die Wellblechfläche, die wir im Unterricht konstruiert haben.

Wenn man erlaubt, dass die verschobene Gerade beim Verschieben ihre Richtung ändert, bekommt man schon eine sehr große Vielfalt an Flächen, zum Beispiel auch den Kreiskegel aus Abbildung 22.2: Als Leitkurve wählen wir den Einheitskreis um den Punkt $(0, 0, 1)^t$, als Gerade in einem Punkt der Leitkurve die Ursprungsgerade durch den Punkt. Wenn wir den Ortsvektor eines Punktes der Leitkurve als $\vec{a}(t)$ und den Richtungsvektor der Geraden in diesem Punkt mit $\vec{b}(t)$ bezeichnen, ist die Fläche durch

$$\vec{p}(s, t) = \vec{a}(t) + s\vec{b}(t) \quad (22.1)$$

gegeben. Eine Fläche, die eine Darstellung wie in Gleichung 22.1 gestattet, nennt man eine **Regelfläche**. Offensichtlich hat jede Tangentialebene einer Regelfläche gleich eine ganze Gerade mit der Fläche gemeinsam. - Und damit soll es jetzt auch genug sein. Ich wünsche dir alles Gute!

Kapitel 23

Klausuren

23.1 Erste Klausur 13.1

1. Bestimme alle Funktionen vom Grad ≤ 2 , deren Graph durch den Nullpunkt geht und dort die Steigung 2 hat.
2. Jonas sucht ein lineares Gleichungssystem, das genau drei Lösungen hat. Das ist eine schwierige Aufgabe, du kennst dich ja mit linearen Gleichungssystemen aus! Gib ihm sachkundige Auskunft, die ihn wissender zurücklässt als er vorher war.
3. Hier siehst du drei Matrizen. Damit hat es folgende Bewandnis. Die beiden Freunde Hinz und Kunz wollten das lineare Gleichungssystem $A\vec{x} = \vec{b}$ untersuchen, und A ist die erste der drei Matrizen. Das machte jeder auf seine Art. Hinzens bevorzugtes Werkzeug sind die Matrixabbildungen, Kunz jedoch bleibt bei den linearen Gleichungssystemen.

$$\begin{pmatrix} 1 & -1 & -1 & 2 \\ 2 & -1 & 4 & -4 \\ 1 & 0 & 5 & -6 \end{pmatrix}, \begin{pmatrix} 1 & -1 & -1 & 2 \\ 0 & 1 & 6 & -8 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 1 & -1 & -1 & 2 & b_1 \\ 0 & 1 & 6 & -8 & b_2 - 2b_1 \\ 0 & 0 & 0 & 0 & b_3 + b_1 - b_2 \end{pmatrix}$$

- (a) Hinz wendet auf die Matrix A das Gaußsche Eliminationsverfahren an, das ergibt die mittlere Matrix. Damit bestimmt er den Kern und das Bild der Matrixabbildung $F : \vec{x} \mapsto A\vec{x}$, und zwar schreibt er sie als Erzeugnisse mit möglichst wenigen erzeugenden Vektoren. Jetzt kann er die Matrixabbildung F schon ganz gut beschreiben. Führe diese Lösung durch, die Matrix brauchst du dabei nicht selbst zu berechnen. Verwende bei der Beschreibung der Wirkung von F geometrische Sprache, versuche aber nicht, Drehungen oder so etwas zu erkennen.
 - (b) Kunz fügt den Vektor \vec{b} mit den Einträgen b_1, b_2, b_3 als fünften Spaltenvektor an A an und wendet auf die neue Matrix das Gaußsche Eliminationsverfahren an. Dabei kommt die dritte Matrix heraus. In ihrer letzten Spalte steht also praktisch die rechte Seite des nach Gauß umgeformten linearen Gleichungssystems $A\vec{x} = \vec{b}$. Welche Lösungsmenge des Systems erhält Kunz? Was erhält Kunz als Menge der \vec{b} , für die das System überhaupt lösbar ist?
 - (c) Die Ergebnisse der beiden Freunde sollten sich doch vertragen wie zwei Seiten einer Medaille, es geht ja um das gleiche Ausgangsproblem. Erkläre, was wozu gehört.
 - (d) Wieso eigentlich stimmen die mittlere und die rechte Matrix zum großen Teil überein?
4. Die drei Vektoren

$$\vec{a} = \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix} \quad \vec{b} = \begin{pmatrix} 2 \\ 0 \\ 3 \end{pmatrix} \quad \text{und} \quad \vec{c} = \begin{pmatrix} 0 \\ 2 \\ 1 \end{pmatrix}$$

sind nicht linear unabhängig. Weise dies zunächst nach; erkläre dabei den Begriff der linearen Unabhängigkeit und gib Auskunft, welche Vektoren man weglassen könnte. Ersetze dann so viele Vektoren, wie nötig, so dass du drei unabhängige Vektoren hast.

5. Es sei

$$P = \begin{pmatrix} 1 & 0 & \frac{1}{2} \\ 0 & 1 & \frac{1}{4} \end{pmatrix}, \quad \vec{b} = \frac{1}{2} \begin{pmatrix} -1 \\ -3 \\ 5 \end{pmatrix}$$

und $F : \vec{x} \mapsto A\vec{x}$ die zugehörige Matrixabbildung.

- (a) Berechne den Kern der Abbildung.
 (b) Zeichne das Bild des von $\vec{e}_1, \vec{e}_2, \vec{e}_3$ gebildeten Einheitswürfels unter F in ein gewöhnliches xy -System, Einheit 4 cm. Zeichne auch das Bild der Geraden $g : \vec{x}(t) = 2\vec{e}_2 + t\vec{b}$ ein. Geht g durch den Punkt $(1|0|1)$?
 (c) Ist jeder Punkt der Ebene Bildpunkt eines Raumpunkts unter F ?
6. Es sei

$$M(\alpha) = \begin{pmatrix} \cos(\alpha) & -\sin(\alpha) \\ \sin(\alpha) & \cos(\alpha) \end{pmatrix} .$$

- (a) Begründe der Reihe nach, dass die durch M gegebene lineare Abbildung der Ebene den Punkt $(1|0)$ um den Winkel α um den Nullpunkt dreht, die x -Achse insgesamt um den Winkel α um den Nullpunkt dreht, die ganze xy -Ebene um den Winkel α um den Nullpunkt dreht. Wenn dir eine Begründung nicht gelingt, kannst du trotzdem die Aussage bei der nächsten Begründung benutzen.
 (b) Berechne $M(\alpha) \cdot M(\beta)$. Das Produkt muss ja $M(\alpha + \beta)$ sein (warum?). Wenn du genau hinschaust, kannst du Formeln für $\sin(\alpha + \beta)$ und für $\cos(\alpha + \beta)$ ablesen, die so genannten Additionstheoreme.
 (c) Was bewirkt die Abbildung zu der Matrix $M(\alpha) \cdot \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \cdot M(-\alpha)$?
7. Die beiden Abbildungen

$$F : \begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} x \\ y \\ 3x - 2y \end{pmatrix} \quad \text{und} \quad G : \begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} x \\ y \\ x^2 + y^2 \end{pmatrix}$$

bilden beide die xy -Ebene in den Raum ab. Wie sehen die Bildmengen aus?

23.2 Zweite Klausur 13.1

1. Es sei

$$\vec{a} = \begin{pmatrix} 1 \\ -1 \\ 3 \end{pmatrix}, \quad \vec{b} = \begin{pmatrix} 2 \\ -1 \\ -2 \end{pmatrix}, \quad \text{und} \quad \vec{c} = \begin{pmatrix} -2 \\ 0 \\ 0 \end{pmatrix} .$$

- (a) Gib eine Parameterdarstellung und eine Gleichung der Ebene E an, die durch \vec{a}, \vec{b} und \vec{c} geht.
 (b) Betrachte das Dreieck, das vom Nullpunkt, \vec{a} und \vec{c} gebildet wird. Wie groß ist der Innenwinkel beim Nullpunkt, wie groß der bei \vec{a} ?
 (c) Unter welchem Winkel schneidet die von \vec{a} erzeugte Gerade die Ebene E ?
 (d) Unter welchem Winkel schneiden sich E und die von \vec{a} und \vec{c} erzeugte Ebene?
 (e) Berechne den Fußpunkt des Lotes vom Nullpunkt auf E . Liegt dieser Punkt in dem Dreieck, das von \vec{a}, \vec{b} und \vec{c} gebildet wird?

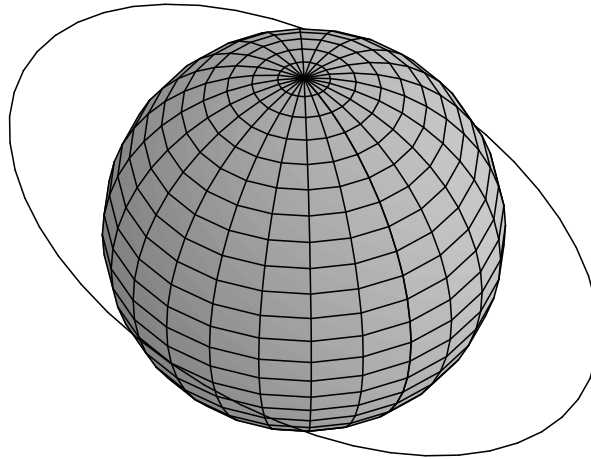


Abbildung 23.1: Zu Aufgabe 2

- (f) Wie groß ist der Abstand des Nullpunktes von der Geraden durch \vec{a} und \vec{b} ?
2. Es sei $\vec{n} = (1, 2, 3)^t$ und k die Kugel um den Nullpunkt mit Radius 5. Ferner sei E die Ebene mit der Gleichung $\vec{n} * \vec{x} = 0$.

- (a) Die Kugel k sollte doch zwei Tangentialebenen haben, die zu E parallel sind. Schreibe ihre Gleichungen hin.
- (b) Wir sehen die Kugel k als Himmelskörper an. Er wird von einem Satelliten auf einer kreisförmigen Bahn umrundet, die in der Höhe 3 über dem Boden verläuft und die in der Ebene E liegt. Gesucht ist ein System von Gleichungen, das genau diese Bahn beschreibt. Begründe, dass es sich keinesfalls um ein lineares Gleichungssystem handeln kann, und schreibe ein System hin. Löse es **nicht!**
- (c) Unser Freund Hinz bestimmt Vektoren $\vec{a}, \vec{b} \in E$, die zu einander orthogonal sind und die die Länge 8 haben. Er behauptet, dass der Punkt

$$\vec{x}(t) = \cos(t)\vec{a} + \sin(t)\vec{b}$$

genau die Bahn des Satelliten durchläuft. Stimmt das?

- (d) Bestimme Vektoren \vec{a} und \vec{b} , die Hinzens Anforderungen genügen.
- (e) Berechne Kern und Bild der Abbildung

$$F : \vec{x} \mapsto \frac{\vec{n} * \vec{x}}{14} \vec{n} .$$

Was macht F geometrisch?

3. Es seien die Kugel um den Nullpunkt mit Radius r und die Kugel um \vec{a} mit Radius R gegeben. Wir gehen davon aus, dass sich die Kugeln schneiden. Schreibe Gleichungen der Kugeln hin, bilde „die Differenz der Gleichungen“ und vereinfache. Welches geometrische Gebilde ist durch die neue Gleichung gegeben?
4. „Wastel“, spricht der Lehrer, „nenne eine Gleichung der Kugel um \vec{e}_3 mit dem Radius $\frac{1}{2}$.“ Und Wastel haut unerschrocken zu:

$$\vec{e}_3 * \vec{x} = |\vec{x}| \cdot \frac{1}{2} \tag{23.1}$$

Zur Bequemlichkeit des Kurses vereinfacht er das auch gleich zu

$$z^2 = \frac{1}{4}(x^2 + y^2 + z^2) . \quad (23.2)$$

Tja, das sieht ja doch etwas fremd aus. Ist das denn eine Kugel? Wir bezeichnen die Menge aller $\vec{x} \in \mathbb{R}^3$, die die Gleichung 23.2 erfüllen, mit k . Bearbeite die folgenden Aufträge.

- Zeige: Wenn $\vec{x} \in k$ ist, dann ist auch $r\vec{x} \in k$ für alle $r \in \mathbb{R}$.
- Bestimme die Menge aller $\vec{x} \in k$, deren dritte Komponente den Wert 1 hat, und beschreibe die Menge geometrisch.
- Was ist k geometrisch?
- Durch welche Umformung wandelte Wastel die Gleichung 23.1 in die Gleichung 23.2 um? Haben die beiden Gleichungen überhaupt die gleiche Lösungsmenge?

23.3 Klausur 13.2

- Die vier Punkte

$$\begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}, \quad \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \quad \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix} \quad \text{und} \quad \begin{pmatrix} 0 \\ -1 \\ -1 \end{pmatrix}$$

sind Eckpunkte eines Tetraeders. Es ist zwar nicht ganz regelmäßig, dafür sind die Koordinaten noch handlich. Zeichne das Tetraeder, bestimme seine Kantenlängen und sein Volumen. Nutze dabei möglichst die Symmetrie des Tetraeders aus, das spart Arbeit.

- Auch Hinz und Kunz haben sich mit dem Tetraeder aus Aufgabe 1 beschäftigt. Nun verwenden sie es als Boden eines Körpers im \mathbb{R}^4 . Dazu fügen sie jedem Punkt als vierte Koordinate den w -Wert 0 hinzu und verbinden ihn mit dem Punkt $\vec{s} = (0, 0, 0, 5)^t$.
 - Berechne das Volumen des Körpers. Wenn du das Ergebnis von Aufgabe 1 nicht hast, kannst du dir dafür einen Wert aussuchen.
 - Hinz behauptet: Wenn du im \mathbb{R}^{n+1} einen Körper hast, der als Boden einen n -dimensionalen Körper vom (n -dimensionalen) Volumen G hat und eine Spitze S in der Höhe h über dem Boden, dann ist das Volumen einfach $V = \frac{1}{n+1}Gh$. Was meinst du dazu?
 - Kunz fragt sich, wie lang das Stück der Geraden durch S und den Nullpunkt ist, das im Innern des Tetraeders verläuft. Was sagst du ihm?
- Es seien die Hyperebenen $H_1 : x - y + 2z - w = 1$ und $H_2 : x + y + z + w = 0$ des \mathbb{R}^4 gegeben. Berechne die Schnittmenge und gib Auskunft, ob der Punkt $\vec{s} = (0, 0, 0, 5)^t$ auf der gleichen Seite von H_1 liegt wie der Nullpunkt.

- Es sei

$$\vec{x} = \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \quad \vec{a} = \begin{pmatrix} 1 \\ 0 \\ 0 \\ -1 \end{pmatrix} \quad \text{und} \quad \vec{b} = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \end{pmatrix} .$$

Schreibe \vec{x} als Summe eines Vektors aus $E = \langle \vec{a}, \vec{b} \rangle$ und eines Vektors, der zu E orthogonal ist.

- Untersuche die Schnittkurve des Kegels K mit der Gleichung $z^2 = x^2 + y^2$ (mit $z \geq 0$) und der Ebene E mit der Gleichung $y + z = 1$. Um was handelt es sich deiner Meinung nach? Eine Parametrisierung ist nicht verlangt.

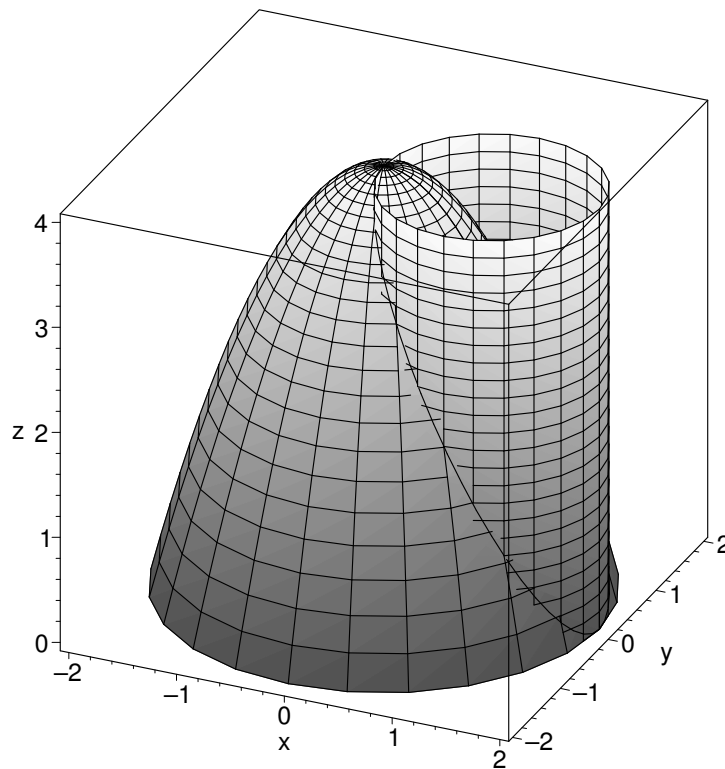


Abbildung 23.2: Zu Aufgabe 6

6. Du siehst auf Seite 119 den Graphen der Funktion $f : (x, y) \mapsto 4 - x^2 - y^2$. Er wird durchdrungen von einem Zylinder. Die Achse des Zylinders ist parallel zur z -Achse, und sie schneidet die x -Achse an der Stelle $x = 1$. Der Radius des Zylinders ist 1. Untersuche die Schnittkurve: Gib eine Parameterdarstellung an und gib begründete Auskunft, um was es sich handelt. Untersuche insbesondere, ob die Kurve in einer Ebene liegt.
7. Die Funktion $(x \mapsto x^3 + 50x - 2)$ hat mindestens eine Nullstelle! Heißt das, dass die Funktionen $(x \mapsto 1)$, $(x \mapsto x)$ und $(x \mapsto x^3)$ linear abhängig sind? Begründe die Behauptung und nimm Stellung zu der Frage.
8. Die Funktionen $f = (x \mapsto 4x)$ und $g = (x \mapsto x^2)$ erzeugen einen Teilraum des Funktionenraums. Wende das Gram-Schmidtsche Orthogonalisierungsverfahren an. Als Skalarprodukt nimm wie üblich $u * v = \int_0^1 u(x)v(x) dx$.
9. Es seien \vec{a}, \vec{b} und \vec{c} Vektoren irgendeines Vektorraums mit Skalarprodukt $*$. Die Vektoren seien alle $\neq \vec{0}$ und paarweise orthogonal. Beweise, dass sie dann linear unabhängig sind.
10. Berechne die Koeffizienten der $(x \mapsto \sin(kx))$ der Fourierreihe von $(x \mapsto \frac{1}{2}x)$ auf $[0; \pi]$ und skizziere, wie du dir den Graphen der Funktion, die durch die Fourierreihe gegeben ist, für $-2\pi \leq x \leq 4\pi$ vorstellst. Hier denke ich wirklich nur an eine grobe Skizze; nicht dass du eine Funktionsuntersuchung durchführst!

23.4 Abiturklausur Abi02

Aufgabe 1

- a) Es sei $y = f(x) = xe^{-x}$. Berechne die ersten drei Ableitungen von f und gib einen Term für die n -te Ableitung an. Ein strenger Beweis für dessen Richtigkeit ist nicht verlangt.

- b) Betrachten wir die durch $y = f_t(x) = (x - t)e^{-x}$ gegebene Kurvenschar. Skizziere, wie der Graph von f_t aussieht, so dass man die wesentlichen Eigenschaften erkennt.
- c) Auf welcher Kurve liegen die Hochpunkte der Scharkurven?
- d) Wie groß ist der Inhalt des Flächenstücks zwischen der x -Achse und dem Kurvenstück, das oberhalb der x -Achse liegt? (Zwischenergebnis: der Inhalt ist e^{-t})
- e) Begründe, dass man f_t nur für $t = 0$ für die Dichtefunktion einer stetig verteilten Zufallsvariable X verwenden könnte (das ist übrigens das f aus Teil a). Schreibe einen Rechenausdruck für den Erwartungswert von X hin, rechne ihn aber nicht aus.
- f) Schreibe die Taylorreihe für das f aus a) hin (Entwicklungspunkt ist 0, wie bei uns üblich).

Aufgabe 2

Unsere Freunde Hinz und Kunz unterhalten sich angeregt beim Bier. Hören wir eine Weile zu.

Kunz: Eine Basis für einen Vektorraum zu wählen, ist doch eine mühsame Sache. Ich frage mich, ob es riskant ist, einfach Vektoren in der richtigen Anzahl zufällig zu wählen. Bilden solche Vektoren nicht in der Regel eine Basis? Lass uns das doch einmal für den \mathbb{R}^3 überlegen.

Hinz: Hm, einen Vektor aus dem \mathbb{R}^3 zufällig wählen kann man nicht. Aber man könnte sich ja auf eine Kugel um den Nullpunkt vom Radius 4 meinetwegen beschränken und daraus drei Vektoren zufällig wählen. Die Wahrscheinlichkeit, dass die drei dann eine Basis bilden, kann man schon angeben. Aber sehr realistisch ist das Modell nicht, denke ich.

Kunz: Ja, wenn ich Vektoren wähle, nehme ich natürlich welche mit ganzzahligen Koordinaten. Wählen wir also zufällig drei Vektoren $\vec{a}, \vec{b}, \vec{c}$, deren Koordinaten ganzzahlig, mindestens 0 und höchstens 4 sind, und suchen wir die Wahrscheinlichkeit p , dass sie linear unabhängig sind.

Hinz: Ich denke, diese Wahrscheinlichkeit p ist mindestens

$$\frac{124}{125} \cdot \frac{120}{125} \cdot \frac{100}{125} .$$

Machen wir doch eine Simulation auf dem Rechner.

Gesagt, getan. Die beiden schreiben ein kleines Programm, das drei Vektoren, wie sie Kunz in seinem zweiten Beitrag beschrieben hat, zufällig wählt und gleich prüft, ob sie ein linear unabhängiges System bilden. Bei 3000 Versuchen erhalten sie für den Anteil der Basen den Wert 0,885.

- a) Mit welcher Wahrscheinlichkeit erhält man eine Basis des \mathbb{R}^3 , wenn man so vorgeht, wie Hinz in seinem ersten Beitrag vorschlägt?
- b) Verträgt sich das Ergebnis der Simulation mit dem Wert, den Hinz in seinem zweiten Beitrag für p nennt? Gib eine qualifizierte Antwort.
- c) Wie kommt Hinz auf den Ausdruck für p in seinem zweiten Beitrag? Und hast du eine Ahnung, wieso er vorsichtig sagt, p sei mindestens so groß wie der genannte Wert?
- d) Ich habe den Rechner einmal mit dem Programm der beiden drei konkrete Vektoren wählen lassen:

$$\vec{a} = \begin{pmatrix} 3 \\ 0 \\ 4 \end{pmatrix}, \quad \vec{b} = \begin{pmatrix} 2 \\ 4 \\ 0 \end{pmatrix}, \quad \vec{c} = \begin{pmatrix} 1 \\ 2 \\ 4 \end{pmatrix}$$

Bilden sie eine Basis des \mathbb{R}^3 ?

- e) Wende auf die Vektoren aus e) unser Orthogonalisierungsverfahren nach Gram-Schmidt an: Gib die ersten beiden Vektoren der neuen Basis konkret an und erläutere dabei den geometrischen Hintergrund des Verfahrens. Für den dritten Basisvektor brauchst du nur einen Rechenausdruck hinzuschreiben.

Aufgabe 3

Aus der xy -Ebene erhebt sich ein gläserner Berg. Seine Oberfläche ist gegeben durch die Vorschrift

$$z = f(x, y) = 9 - x^2 - 4y^2 \quad .$$

- a) Beschreibe den Berg: Wie sieht er aus?
 b) Wir betrachten das Teilstück des Bergkörpers, das auf dem Rechteck der Punkte der xy -Ebene mit $0 \leq x \leq 2$ und $0 \leq y \leq 1$ steht. Wie groß ist sein Volumen?
 c) Gib einen Rechenausdruck für das gesamte Volumen des Berges an, werte ihn aber nicht aus.
 d) Es sei nun P der Punkt der Bergoberfläche mit den Koordinaten $x = 2$, $y = 1$ und $z = f(2; 1)$. Wir denken uns drei Wege auf der Bergoberfläche, die durch P laufen, und zwar

$$\vec{p}_1(t) = \begin{pmatrix} t \\ 1 \\ f(t, 1) \end{pmatrix}, \quad \vec{p}_2(t) = \begin{pmatrix} 2 \\ t \\ f(2, t) \end{pmatrix} \quad \text{und} \quad \vec{p}_3(t) = \begin{pmatrix} 2t \\ t \\ f(2t, t) \end{pmatrix} \quad .$$

Die Vektoren

$$\vec{b}_1 = \begin{pmatrix} 1 \\ 0 \\ -4 \end{pmatrix}, \quad \vec{b}_2 = \begin{pmatrix} 0 \\ 1 \\ -8 \end{pmatrix} \quad \text{und} \quad \vec{b}_3 = \begin{pmatrix} 2 \\ 1 \\ -16 \end{pmatrix}$$

sind jeweils Tangentialvektoren der Kurven im Punkt P . Rechne dies für \vec{b}_3 nach.

- e) Bestätige rechnerisch, dass $\vec{b}_3 \in \langle \vec{b}_1, \vec{b}_2 \rangle$ ist, und kommentiere.
 f) Gib eine Punkt-Richtungsform und eine Gleichung für die Ebene E durch P mit den Richtungsvektoren \vec{b}_1 und \vec{b}_2 an.
 g) Ein Laserstrahl geht vom Punkt $Q(\frac{1}{2}; 12; 0)$ der xy -Ebene aus, trifft im Punkt P auf den Glasberg und wird dort reflektiert. Gib je einen Richtungsvektor für den Strahl und für den reflektierten Strahl an.