

ANALYSIS 2

Wolfgang Soergel

31. Januar 2012

Inhaltsverzeichnis

1	Komplexe Exponentialfunktion	4
1.1	Definition und erste Eigenschaften	4
1.2	Fundamentalsatz der Algebra	10
1.3	Integration von vektorwertigen Funktionen	12
1.4	Integration rationaler Funktionen	16
1.5	Komplexe Differenzierbarkeit	19
2	Lösung einiger Schwingungsgleichungen	25
2.1	Gedämpfte Schwingungen	25
2.2	Lineare Differentialgleichungen höherer Ordnung	28
2.3	Gekoppelte Schwingungen	30
2.4	Angeregte Schwingungen	32
3	Grundlegendes zu Fourierreihen	35
3.1	Eindeutigkeit der Fourierreihe	35
3.2	Der Satz von Stone-Weierstraß	36
3.3	Konvergenz der Fourierreihe	42
4	Ableitungen in mehreren Veränderlichen	46
4.1	Partielle Ableitungen und Gradient	46
4.2	Das Differential	48
4.3	Die Kettenregel	57
4.4	Weitere Ableitungsregeln	58
4.5	Differenzierbarkeit über partielle Ableitungen	62
5	Mehrfache Integrale und Ableitungen	66
5.1	Integration über kompakte Quader	66
5.2	Taylorentwicklung in mehreren Veränderlichen	71
5.3	Rechnen mit Approximationen	74
5.4	Maxima und Minima in mehreren Veränderlichen	77
6	Wegintegrale	82
6.1	Vektorfelder und Kovektorfelder	82
6.2	Gradienten in krummlinigen Koordinaten*	93
6.3	Wegintegrale	99
6.4	Wegzusammenhängende Räume	108
6.5	Homotopie von Wegen	111
6.6	Rotation und Potential	114

7 Umkehrsatz und Anwendungen	129
7.1 Der Satz über die Umkehrabbildung	129
7.2 Der Satz über implizite Funktionen	137
7.3 Untermannigfaltigkeiten reeller Räume	146
7.4 Die Transformationsformel	156
7.5 Integration über Untermannigfaltigkeiten des \mathbb{R}^n	167
8 Gewöhnliche Differentialgleichungen	175
8.1 Grundlegende Definitionen und Eigenschaften	175
8.2 Existenz und Eindeutigkeit von Lösungen	186
8.3 Lineare Differentialgleichungen	192
8.4 Höhere Ableitungen ohne Koordinaten	195
8.5 Lösungen als Funktionen ihres Anfangswerts	197
9 Maß und Integral	201
9.1 Maßräume und Maße	201
9.2 Konstruktion des Lebesguemaßes auf \mathbb{R}	207
9.3 Meßbare Abbildungen	218
9.4 Das Integral von nichtnegativen Funktionen	223
9.5 Integrierbare Funktionen und ihr Integral	229
9.6 Integration auf Produkträumen	234
9.7 Regularität von Borelmaßen	244
9.8 Rechnen mit dem Lebesgue-Integral	246
9.9 Flächenmaß	251
10 Der Satz von Stokes	255
10.1 Multilineare Algebra und Dachprodukt	255
10.2 Differentialformen höheren Grades	260
10.3 Das Maß einer Differentialform	263
10.4 Orientierung von Mannigfaltigkeiten	268
10.5 Integration von Differentialformen	271
10.6 Äußere Ableitung von Differentialformen	278
10.7 Berandete Untermannigfaltigkeiten	285
10.8 Der Satz von Stokes	292
10.9 Divergenz und Laplace in krummen Koordinaten*	303
Literaturverzeichnis	311
Index	311

1 Komplexe Exponentialfunktion

1.1 Definition und erste Eigenschaften

1.1.1. Mir scheint es wünschenswert, daß dieser Abschnitt bereits zu Ende des ersten oder zu Beginn des zweiten Semesters behandelt wird, damit die komplexe Exponentialfunktion bei der Diskussion unitärer und orthogonaler Abbildungen in der linearen Algebra zur Verfügung steht. Ich setze voraus, daß der Leser mit den Grundlagen des algebraischen Rechnens mit komplexen Zahlen vertraut ist, wie sie etwa in ?? entwickelt werden.

Definition 1.1.2. Für jede komplexe Zahl $z \in \mathbb{C}$ definieren wir eine komplexe Zahl $\exp(z) \in \mathbb{C}$ durch die Vorschrift

$$\exp z = \sum_{k=0}^{\infty} \frac{1}{k!} z^k$$

Aufgrund der Abschätzungen $\operatorname{Re}(z^k), \operatorname{Im}(z^k) \leq |z^k| = |z|^k$ konvergiert diese Reihe für alle $z \in \mathbb{C}$ nach dem Majorantenkriterium mit der reellen Exponentialreihe als Majorante absolut in Real- und Imaginärteil und wir erhalten folglich eine Abbildung $\exp : \mathbb{C} \rightarrow \mathbb{C}$, die **komplexe Exponentialfunktion**.

1.1.3. Man prüft genau wie in 10.5.24 im Reellen auch im Komplexen die Funktionalgleichung

$$\exp(z + w) = (\exp z)(\exp w)$$

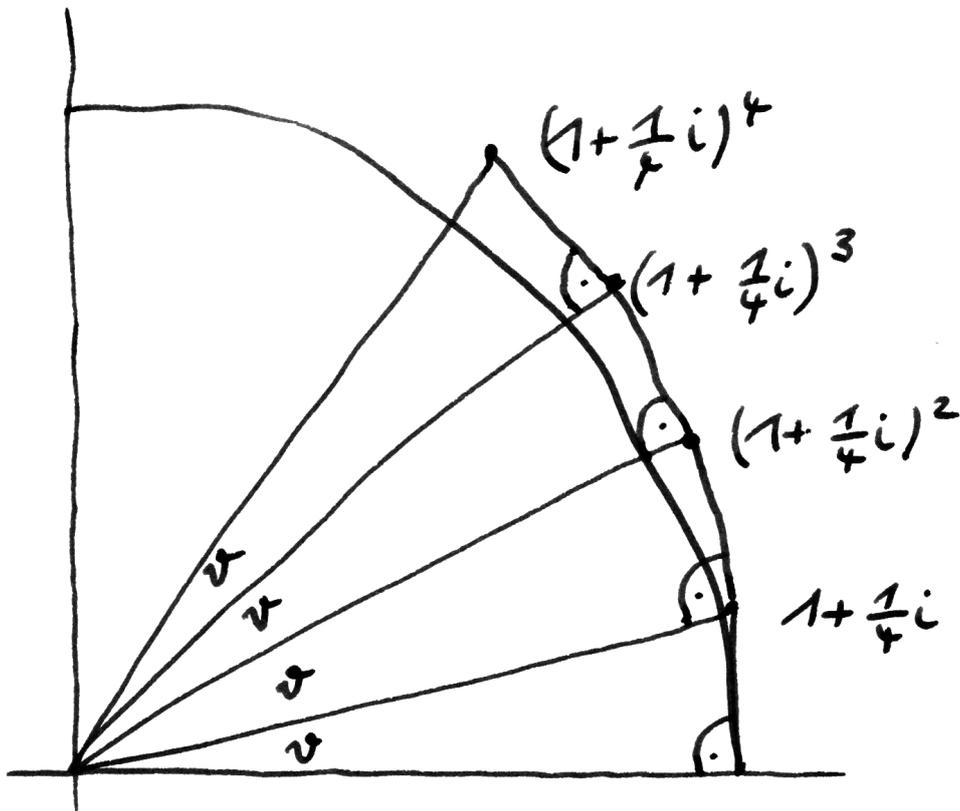
In der Sprache der Algebra ausgedrückt ist die Exponentialabbildung also ein Gruppenhomomorphismus von der additiven Gruppe der komplexen Zahlen in die multiplikative Gruppe der von Null verschiedenen komplexen Zahlen und man erhält insbesondere $\exp(-z) = (\exp z)^{-1}$. Wie im Reellen 6.1.13 zeigt man auch die Stetigkeit von $\exp : \mathbb{C} \rightarrow \mathbb{C}$ zunächst im Nullpunkt über die Reihenentwicklung und dann an jeder Stelle $z \in \mathbb{C}$ mithilfe der Funktionalgleichung. Zusätzlich folgern wir aus der Vertauschbarkeit der komplexen Konjugation mit Summe, Produkt und Grenzwertbildung auch noch

$$\exp(\bar{z}) = \overline{\exp z} \quad \forall z \in \mathbb{C}$$

Für den Betrag von $\exp z$ erhalten wir dann

$$\begin{aligned} |\exp z|^2 &= \exp z \overline{\exp z} \\ &= \exp z \exp \bar{z} \\ &= \exp(z + \bar{z}) \\ &= \exp(2 \operatorname{Re} z) \end{aligned}$$

Folglich gilt $|\exp z| = \exp(\operatorname{Re} z)$ und speziell $|\exp(it)| = 1$ für alle $t \in \mathbb{R}$.



Auch für komplexes z zeigt man wie in 5.6.3 die Formel

$$\exp(z) = \lim_{n \rightarrow \infty} \left(1 + \frac{z}{n}\right)^n$$

Das obige Bild stellt unter anderem den vierten Term dieser Folge im Fall $z = i$ dar. Ich finde, man kann recht gut erkennen, wie diese Folge bei wachsendem n gegen den Punkt auf der Kreislinie konvergiert, für den das Segment der Kreislinie, das von ihm zur reellen Achse herunterläuft, die Länge Eins hat.

1.1.4. Für eine reelle Zahl $a > 0$ und $z \in \mathbb{C}$ definieren wir wieder

$$a^z = \exp(z \log a)$$

und schreiben insbesondere auch $\exp z = e^z$ für $z \in \mathbb{C}$. Aus der Beschreibung von \sin , \cos und \exp durch Potenzreihen 10.6.2 oder auch aus 10.6.3 folgt die **Euler'sche Formel**

$$e^{it} = \cos t + i \sin t$$

Anschaulich beschreibt die Abbildung $\mathbb{R} \rightarrow \mathbb{C}$, $t \mapsto \exp(it)$ also das Aufwickeln der reellen Zahlengerade auf den Einheitskreis in der komplexen Zahlenebene $\{z \in \mathbb{C} \mid |z| = 1\}$. Insbesondere erfüllen unsere Hauptdarsteller die bemerkenswerte Identität

$$e^{i\pi} = -1$$

Aus $\exp(-it) = \overline{\exp(it)}$ folgern wir umgekehrt für alle $t \in \mathbb{R}$ die Formeln

$$\cos t = \frac{e^{it} + e^{-it}}{2} \quad \text{und} \quad \sin t = \frac{e^{it} - e^{-it}}{2i}$$

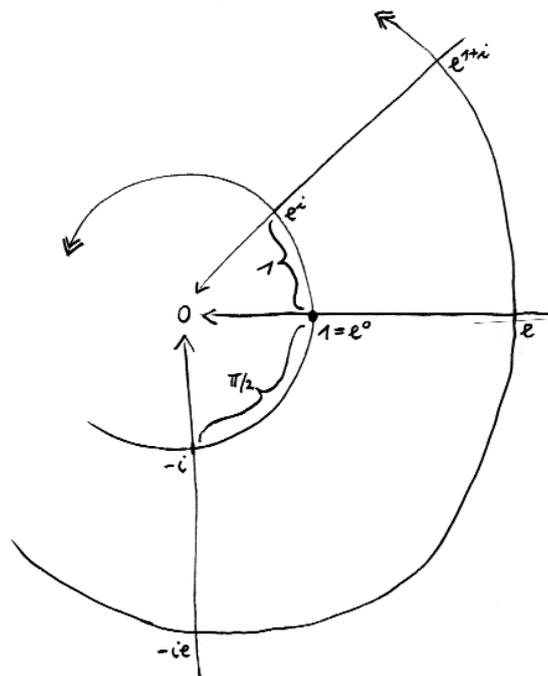
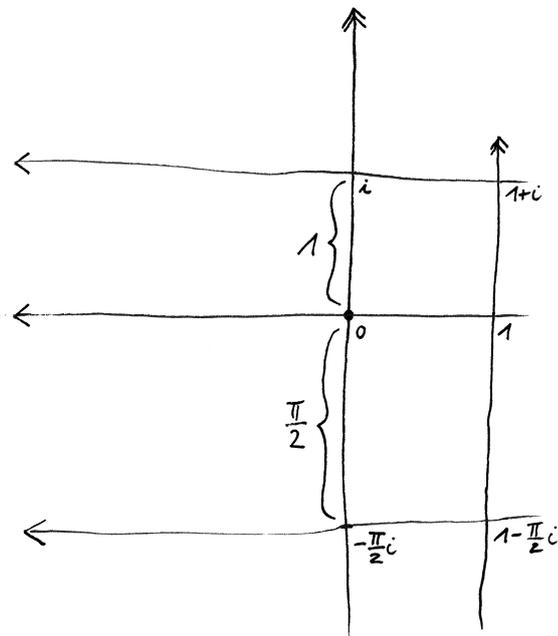
Benutzen wir diese Formeln, um den Sinus und Cosinus zu Funktionen $\mathbb{C} \rightarrow \mathbb{C}$ auszudehnen, und dehnen wir ihre hyperbolischen Analoga in derselben Weise zu Funktionen $\mathbb{C} \rightarrow \mathbb{C}$ aus, so können wir die formale Analogie zwischen diesen Funktionen präzisieren zu den Formeln $\cos z = \cosh iz$ und $\sin z = -i \sinh iz$ für alle $z \in \mathbb{C}$. In der angelsächsischen Literatur wird manchmal auch die Abkürzung $\exp(iz) = \text{cis}(z)$ verwendet, die wohl auf die Euler'sche Formel $\exp(iz) = \cos z + i \sin z$ zurückzuführen ist.

Ergänzende Übung 1.1.5. Man zeige mit der Euler'schen Formel aus 1.1.4 die Identität $\sin^3 \vartheta = \frac{3}{4} \sin \vartheta - \frac{1}{4} \sin(3\vartheta)$.

1.1.6. Nach 1.1.4 induziert die Exponentialfunktion eine Surjektion der imaginären Geraden $i\mathbb{R}$ auf den Einheitskreis $\{z \in \mathbb{C} \mid |z| = 1\}$. Daß die Exponentialfunktion eine Bijektion $\mathbb{R} \xrightarrow{\sim} \mathbb{R}_{>0}$ induziert, wissen wir bereits aus 6.2.11. Da nun jede von Null verschiedene komplexe Zahl w sich schreiben läßt als Produkt $w = (w/|w|)|w|$ mit $w/|w|$ auf dem Einheitskreis und $|w|$ positiv, ist die Exponentialfunktion nach der Funktionalgleichung sogar ein surjektiver Gruppenhomomorphismus $\exp : \mathbb{C} \rightarrow \mathbb{C}^\times$. Der Kern dieses Gruppenhomomorphismus, als da heißt das Urbild des neutralen Elements $1 \in \mathbb{C}^\times$ besteht aufgrund unserer Gleichungen $|\exp z| = \exp(\text{Re } z)$ und $e^{it} = \cos t + i \sin t$ genau aus allen ganzzahligen Vielfachen von $2\pi i$, in Formeln

$$\ker(\exp) = 2\pi i \mathbb{Z}$$

Im übrigen können wir mit Hilfe der komplexen Exponentialfunktion auch leicht zeigen, daß es für jede komplexe Zahl $z \in \mathbb{C}$ ein $w \in \mathbb{C}$ gibt mit $w^2 = z$: Wir



Die Bilder ausgewählter Teile der komplexen Zahlenebene unter der komplexen Exponentialfunktion. Die Vertikalen werden zu Kreislinien aufgewickelt, die Horizontalen in vom Nullpunkt ausgehende Strahlen transformiert.

dürfen ja ohne Beschränkung der Allgemeinheit $z \neq 0$ annehmen, dann finden wir nach dem vorhergehenden ein $a \in \mathbb{C}$ mit $z = \exp(a)$, und tut es $w = \exp(a/2)$.

1.1.7. Wir bestimmen mit dieser Erkenntnis die **n -ten Einheitswurzeln**, als da heißt die komplexen Lösungen der Gleichung $z^n = 1$. Nach 1.1.6 hat jede Lösung die Gestalt $z = e^b$ für geeignetes $b \in \mathbb{C}$ und so ein z löst unsere Gleichung genau dann, wenn gilt $z^n = e^{nb} = 1$ alias $nb \in 2\pi i \mathbb{Z}$. Wir erhalten so die Lösungen $\exp(2\pi i \nu/n)$ für $\nu = 0, 1, \dots, n-1$ und erkennen auch, daß sie paarweise verschieden sind und es keine anderen Lösungen geben kann. In der komplexen Zahlenebene kann man sich die n -ten Einheitswurzeln veranschaulichen als die Ecken desjenigen in den Einheitskreis einbeschriebenen regelmäßigen n -Ecks, das als eine Ecke die 1 hat.

Ergänzung 1.1.8. Der Satz von **Hermite-Lindemann** sagt, daß für eine von Null verschiedene im Sinne von 5.4.2 algebraische komplexe Zahl α der Wert der Exponentialfunktion $\exp(\alpha)$ stets transzendent ist. Daraus folgt sowohl, daß die Euler'sche Zahl $e = \exp(1)$ transzendent ist, als auch, daß $2\pi i$ und damit natürlich auch π transzendent sind, da nämlich $\exp(2\pi i) = 1$ nicht transzendent ist. In etwas allgemeinerer Form sagt der Satz, daß gegeben komplexe algebraische Zahlen $\alpha_1, \dots, \alpha_n$, die linear unabhängig sind über \mathbb{Q} , die Werte der Exponentialfunktion $\exp(\alpha_1), \dots, \exp(\alpha_n)$ algebraisch unabhängig sind über \mathbb{Q} im Sinne von ?? . Mehr dazu findet man etwa in [?]. Schanuels Vermutung, wie das zu verallgemeinern sein sollte, findet man in ?? .

Übung 1.1.9. Sei $n \in \mathbb{N}$. Für jede komplexe Zahl $a \neq 0$ besitzt die Gleichung $z^n = a$ genau n Lösungen $z \in \mathbb{C}$.

Ergänzende Übung 1.1.10. Man zeige, daß es nicht möglich ist, in stetiger Weise zu jeder komplexen Zahl eine Wurzel zu wählen, daß es also keine stetige Abbildung $w : \mathbb{C} \rightarrow \mathbb{C}$ gibt mit $w(z)^2 = z \quad \forall z \in \mathbb{C}$. Hinweis: Man prüfe, daß die Funktion $w(\exp(z)) \exp(-z/2)$ einerseits konstant sein müßte, aber andererseits nicht denselben Wert bei 0 und $2\pi i$ annehmen würde. Die anschauliche Bedeutung der Aussage mag aus der graphischen Darstellung der Abbildung $z \mapsto z^2$ in ?? klar werden.

Ergänzende Übung 1.1.11. In einem regelmäßigen Fünfeck stehen die Längen der Diagonalen zu den Längen der Seiten im Verhältnis des goldenen Schnitts. Man prüfe diese elementargeometrisch leicht einzusehende Behauptung durch algebraische Rechnung. Hinweis: Der goldene Schnitt ist die positive Lösung der Gleichung $a/1 = (1+a)/a$ alias $a^2 - a - 1 = 0$, seine geometrische Bedeutung wurde in 1.2.1 erklärt. Es gilt zu zeigen, daß für $\zeta = \exp(2\pi i/5)$ der Ausdruck $a = |1 - \zeta^2|/|1 - \zeta| = |1 + \zeta|$ die fragliche Gleichung löst. Man verwende $\zeta^4 = \bar{\zeta}$.

Übung 1.1.12. Für alle $\lambda \in \mathbb{C}$ hat die Abbildung $\mathbb{R} \rightarrow \mathbb{C}$, $t \mapsto e^{\lambda t}$ die Ableitung $t \mapsto \lambda e^{\lambda t}$. Später kann das auch mit 1.5.12 und 1.5.18 sehr schnell erledigt werden.



Die Überlagerung zweier Sinuswellen mit nahe beieinanderliegender Periode.
Rechnerisch finden wir etwa die Identität

$$e^{i\omega_1 t} + e^{i\omega_2 t} = e^{i(\omega_1 - \omega_2)t/2} (e^{i(\omega_1 + \omega_2)t/2} + e^{-i(\omega_1 + \omega_2)t/2})$$

und durch Betrachtung der Imaginärteile beider Seiten

$$\sin(\omega_1 t) + \sin(\omega_2 t) = 2 \sin((\omega_1 - \omega_2)t/2) \cos((\omega_1 + \omega_2)t/2)$$

Liegen hier ω_1 und ω_2 nah beieinander, so ergibt sich für diesen Ausdruck als Funktion von t das obige Bild, in dem sich anschaulich gesprochen immer abwechselnd beide Sinuswellen einmal gegenseitig auslöschen und dann wieder addieren. Man kann das auch mit eigenen Ohren erfahren, wenn man sich von zwei Komilitonen zwei nahe beieinanderliegende Töne vorsingen läßt: Es ist dann so eine Art Wummern zu hören, das eben mit der Frequenz $\omega_1 - \omega_2$ geschieht.

Ergänzende Übung 1.1.13. Die Nullstellen des komplexen Sinus liegen alle auf der reellen Achse.

Ergänzende Übung 1.1.14. Man leite einige Formeln der nachstehenden Tabelle her.

	sin	cos
π	0	-1
$\pi/2$	1	0
$\pi/3$	$\sqrt{3}/2$	1/2
$\pi/4$	$1/\sqrt{2}$	$1/\sqrt{2}$
$\pi/5$	$\sqrt{5 - \sqrt{5}}/2\sqrt{2}$	$(\sqrt{5} + 1)/4$
$\pi/6$	1/2	$\sqrt{3}/2$
$\pi/7$?	?
$\pi/8$	$\sqrt{\frac{1}{2} - \sqrt{2}}$	$\sqrt{\frac{1}{2} + \sqrt{2}}$
$\pi/9$?	?
$\pi/10$	$(\sqrt{5} - 1)/4$	$\sqrt{5 + \sqrt{5}}/2\sqrt{2}$
$\pi/11$?	?

Man bemerkt, daß sich für $\cos(\pi/5)$ gerade die Hälfte unseres “goldenen Schnitts” aus 1.1.11 ergibt. Bei der Bestimmung der Werte für $\pi/5$ und $\pi/10$ mag man von nebenstehendem Bild ausgehen, das insbesondere bei der Bestimmung von $\sin(\pi/10)$ helfen sollte. Wir zeigen in ??, warum es unmöglich ist, Formeln derselben Bauart auch für $\sin(\pi/7)$ oder $\sin(\pi/9)$ oder $\sin(\pi/11)$ anzugeben.

1.2 Fundamentalsatz der Algebra

Satz 1.2.1 (Fundamentalsatz der Algebra). *Jedes nicht konstante komplexe Polynom besitzt mindestens eine komplexe Nullstelle.*

1.2.2. Der im folgenden wiedergegebene Beweis von Jean-Robert Argand hat dahingegen den Vorteil, mit besonders wenigen technischen Hilfsmitteln auszukommen. Andere Beweise diskutieren wir in ??.

Beweis. Sei P unser Polynom. Wir zeigen zunächst, daß es eine Stelle $p \in \mathbb{C}$ gibt, an der die Funktion $\mathbb{C} \rightarrow \mathbb{R}, z \mapsto |P(z)|$ ihr Minimum annimmt, in Formeln $|P(z)| \geq |P(p)| \forall z \in \mathbb{C}$. In der Tat, nehmen wir irgendein $u \in \mathbb{C}$ her, so gibt es offensichtlich $R \in \mathbb{R}$ derart, daß aus $|z| \geq R$ folgt $|P(z)| \geq |P(u)|$. Als stetige

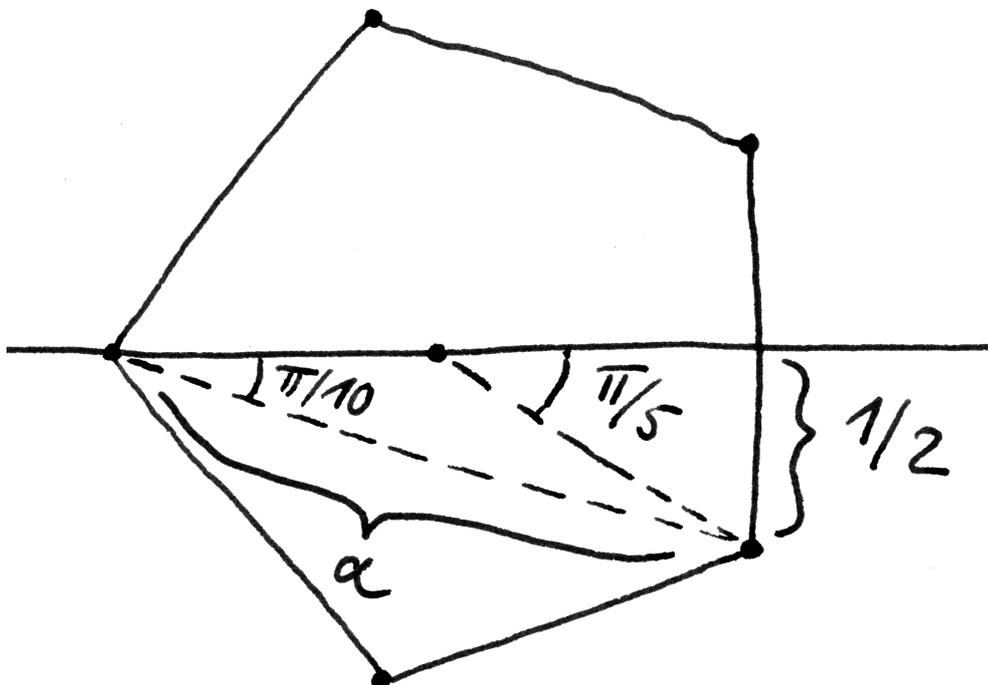


Illustration zur Berechnung von $\sin(\pi/10)$

Funktion nimmt aber die Funktion $z \mapsto |P(z)|$ auf der kompakten Kreisscheibe $\{z \mid |z| \leq R\}$ ein Minimum an, sagen wir an der Stelle p , und das muß dann auch das Minimum von $|P(z)|$ auf ganz \mathbb{C} sein. Wir zeigen nun $P(p) = 0$ durch Widerspruch und müssen dazu nachweisen: Ist $p \in \mathbb{C}$ gegeben mit $P(p) \neq 0$, so nimmt die Funktion $z \mapsto |P(z)|$ bei p nicht ihr Minimum an. Sei dazu erst einmal $p \in \mathbb{C}$ beliebig. Entwickeln wir $P(p+w)$ nach Potenzen von w , so erhalten wir

$$P(p+w) = P(p) + bw^m + w^{m+1}Q(w)$$

mit $b \neq 0$, $m \geq 1$ (da P nicht konstant ist) und einem geeigneten Polynom Q . Nach 1.1.9 finden wir $q \in \mathbb{C}$ mit $P(p) + bq^m = 0$ und sind fertig, sobald wir zeigen können, daß unter der Annahme $P(p) \neq 0$ für hinreichend kleine $t > 0$ gilt

$$|P(p+ tq)| < |P(p)|$$

Das ist klar im Fall $Q = 0$ und wir müssen nur noch erklären, warum die Terme der Ordnung $> m$ diese Ungleichung für kleines t nicht zerstören können. Wir betrachten dazu die Abbildung $\mathbb{R} \rightarrow \mathbb{C}$, $t \mapsto P(p+ tq)$ und erhalten

$$P(p+ tq) = P(p) - t^m P(p) + t^{m+1} \tilde{Q}(t)$$

für ein geeignetes Polynom \tilde{Q} , also

$$|P(p+ tq)| \leq (1 - t^m)|P(p)| + t^m |t\tilde{Q}(t)| \quad \text{für } t \in [0, 1]$$

Gilt nun $P(p) \neq 0$ und wählen wir $t \in (0, 1)$ hinreichend klein für die Ungleichung $|t\tilde{Q}(t)| < |P(p)|$, so folgt $|P(p+ tq)| < |P(p)|$ und wir sind fertig. \square

1.3 Integration von vektorwertigen Funktionen

1.3.1. Die Aufgabe, Funktionen wie zum Beispiel $(x + x^5)/(x^7 - 1)$ zu integrieren, wird uns im anschließenden Abschnitt in natürlicher Weise zur Integration komplexwertiger Funktionen führen. Anstatt einen solchen Integralbegriff ad hoc einzuführen, zimmern wir in diesem Abschnitt gleich einen größeren begrifflichen Rahmen, der nicht nur das Integrieren komplexwertiger Funktionen als Spezialfall umfaßt, sondern auch in natürlicher Weise unsere Überlegungen zum Differenzieren vektorwertiger Funktionen ergänzt und uns in Zukunft noch in mancherlei Weise die Arbeit erleichtern wird.

Definition 1.3.2. Sei $[a, b] \subset \mathbb{R}$ ein nichtleeres kompaktes Intervall, V ein reeller Vektorraum und $f : [a, b] \rightarrow V$ eine Abbildung. Wir betrachten für $r \geq 1$ die äquidistante Unterteilung $a = t_0 \leq t_1 \leq \dots \leq t_r = b$ und definieren die r -te Riemannsumme $S^r(f) \in V$ durch

$$S^r(f) = \sum_{i=1}^r (t_i - t_{i-1}) f(t_i) = \left(\frac{b-a}{r} \right) \sum_{i=1}^r f(t_i)$$

Satz 1.3.3 (Integration vektorwertiger Funktionen). Ist $f : [a, b] \rightarrow V$ eine stetige Abbildung von einem nichtleeren kompakten Intervall in einen Banachraum V , so existiert der Grenzwert der Riemannsummen, wir können das **Integral** $(\int f) \in V$ von f definieren als diesen Grenzwert

$$\int f = \int_a^b f = \int_a^b f(t) dt = \lim_{r \rightarrow \infty} S^r(f)$$

und das so erklärte Integral hat die folgenden Eigenschaften:

1. Für alle $c \in [a, b]$ gilt $\int_a^b f = \int_a^c f + \int_c^b f$.
2. Ist $f = v$ konstant mit $v \in V$, so gilt $\int_a^b f(t) dt = \int_a^b v dt = (b - a)v$.
3. Ist W ein weiterer Banachraum und $\Lambda : V \rightarrow W$ eine stetige lineare Abbildung, so gilt

$$\int (\Lambda \circ f) = \Lambda \left(\int f \right)$$

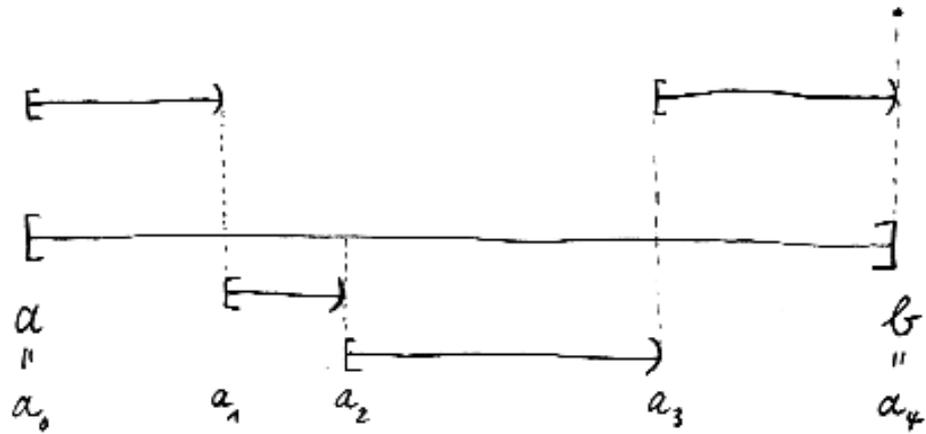
4. Für die Norm des Integrals gilt die Abschätzung $\|\int f\| \leq \int \|f\|$.

1.3.4. Sie mögen in diesem Satz die Regeln $\int \lambda f = \lambda \int f$ sowie $\int (f + g) = \int f + \int g$ für stetige vektorwertige Funktionen f, g und $\lambda \in \mathbb{R}$ vermißt haben. Sie folgen jedoch formal aus Teil 3. In der Tat dürfen wir dort $\Lambda = (\lambda \cdot) : V \rightarrow V$ nehmen und auch $\Lambda : V \times V \rightarrow V$ die Addition sowie die beiden Projektionen. So ergibt sich für die $V \times V$ -wertige Funktion (f, g) zunächst $\text{pr}_1 \int (f, g) = \int f$ und $\text{pr}_2 \int (f, g) = \int g$ und damit $\int (f, g) = (\int f, \int g)$ und durch Anwenden der Addition dann $\int (f + g) = \int f + \int g$.

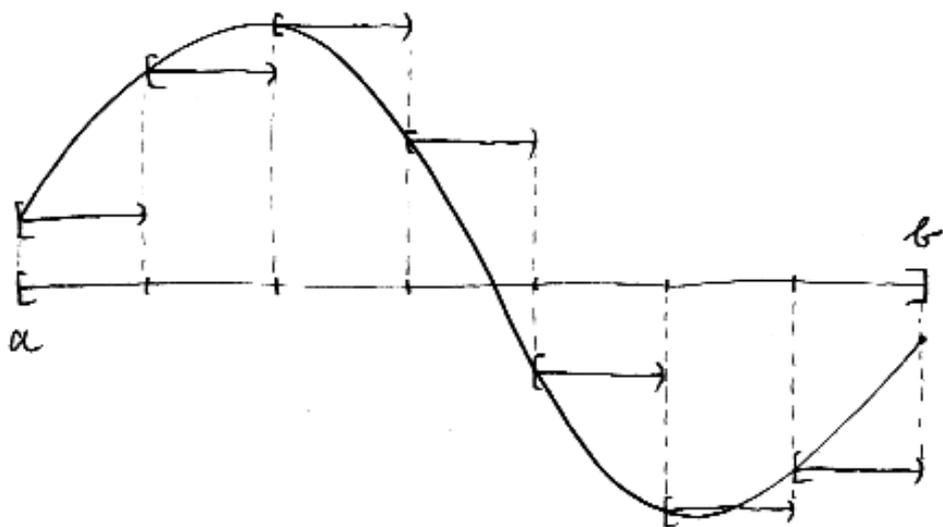
Beweis. Im Fall $a = b$ sind alle Riemannsummen Null und ihr Grenzwert existiert und ist auch Null. Wir konzentrieren uns im folgenden auf den Fall $a < b$. Nach 10.5.10 ist mit V auch der Vektorraum $\text{Ens}^b([a, b], V)$ aller beschränkten Abbildungen $[a, b] \rightarrow V$ mit seiner Supremumsnorm vollständig. Darin betrachten wir nun den Teilraum $T \subset \text{Ens}^b([a, b], V)$ aller Abbildungen $s : [a, b] \rightarrow V$ mit der Eigenschaft, daß es eine nicht notwendig äquidistante Unterteilung $a = a_0 < a_1 < \dots < a_r = b$ unseres Intervalls gibt derart, daß s auf jedem der Teilintervalle $[a_{i-1}, a_i)$ konstant ist. Die Elemente von T heißen **Treppenfunktionen** auf $[a, b]$. Offensichtlich existiert eine lineare Abbildung $I : T \rightarrow \mathbb{R}$ mit der Eigenschaft

$$I(s) = \sum s(a_{i-1})(a_i - a_{i-1})$$

wann immer für eine Unterteilung $a = a_0 < a_1 < \dots < a_r = b$ unseres Intervalls unsere Funktion s konstant ist auf allen Teilintervallen $[a_{i-1}, a_i)$. Offensichtlich



Der Graph einer reellwertigen Treppenfunktion.



Der Graph einer reellwertigen Funktion f und der zugehörigen Treppenfunktion f_7 aus dem nebenstehenden Beweis, mit $I(f_7)$ der siebten Riemannsumme von f .

hat diese lineare Abbildung auch die Eigenschaft $\|I(s)\| \leq (b-a)\|s\|_\infty$. Insbesondere ist $I : T \rightarrow V$ gleichmäßig stetig. Damit zeigt hinwiederum 10.5.5, daß I auf genau eine Weise zu einer stetigen Abbildung auf den Abschluß \bar{T} von T in $\text{Ens}^b([a, b], V)$ fortgesetzt werden kann. In diesem Abschluß liegen nun aber, etwa nach gleichmäßiger Stetigkeit 9.7.14, alle stetigen Abbildungen, in Formeln $\mathcal{C}([a, b], V) \subset \bar{T}$, so daß wir durch stetige Fortsetzung vom Raum der Treppenfunktionen insbesondere eine Abbildung

$$I : \mathcal{C}([a, b], V) \rightarrow V$$

erhalten. Wieder nach gleichmäßiger Stetigkeit 9.7.14 ist jede stetige Abbildung $f : [a, b] \rightarrow V$ auch der Grenzwert in der Supremumsnorm der Treppenfunktionen f_r , die wir erhalten, wenn wir von der äquidistanten Unterteilung $a = t_0 \leq t_1 \leq \dots \leq t_r = b$ ausgehen und f_r auf $[t_i, t_{i+1})$ konstant den Wert $f(t_i)$ annehmen lassen und auf t_r den Wert $f(t_r)$. Für diese f_r gilt also $\lim_{r \rightarrow \infty} f_r = f$ und wegen $S_r(f) = I(f_r)$ folgt

$$\lim_{r \rightarrow \infty} S_r(f) = \lim_{r \rightarrow \infty} I(f_r) = I(f)$$

Der Grenzwert unserer Riemannsummen existiert also in der Tat und stimmt mit $I(f)$ überein. Die erste Eigenschaft zeigt man nun, indem man die Notation I zu I_a^b verfeinert, dann die Identität

$$I_a^b(f) = I_a^c(f) + I_c^b(f)$$

zunächst für Treppenfunktionen $f \in T$ prüft, und sie dann für alle Funktionen aus \bar{T} folgert. Die drei anderen Eigenschaften erhält man, indem man die analogen Eigenschaften für die Riemannsummen hinschreibt und zum Grenzwert übergeht. \square

1.3.5. Wie im Fall reellwertiger Funktionen verwenden wir auch im Fall vektorwertiger Funktionen die Konvention $\int_b^a f = -\int_a^b f$ und ist $f : I \rightarrow V$ eine stetige Abbildung von einem reellen Intervall in einen Banachraum, so gilt für beliebige $a, b, c \in I$ die Formel $\int_a^b f = \int_a^c f + \int_c^b f$.

Satz 1.3.6 (Vektorwertige Variante des Hauptsatzes). *Gegeben ein halboffenes Intervall $I \subset \mathbb{R}$, ein Banachraum V , eine stetige Funktion $f : I \rightarrow V$ und ein Punkt $a \in I$ ist die Funktion*

$$F : I \rightarrow V \\ x \mapsto \int_a^x f(t) \, dt$$

die einzige differenzierbare Funktion $F : I \rightarrow V$ mit $F' = f$ und $F(a) = 0$.

Beweis. Sehr ähnlich zum Beweis für reellwertige Funktionen und dem Leser zur Übung überlassen. \square

Korollar 1.3.7 (Integrieren mit Stammfunktionen). Sei V ein Banachraum und $f : [a, b] \rightarrow V$ stetig. Ist $G : [a, b] \rightarrow V$ eine **Stammfunktion** von f , d.h. eine differenzierbare Funktion mit Ableitung $G'(t) = f(t)$, so gilt

$$\int_a^b f(t) dt = G(b) - G(a)$$

Beweis. Das folgt sofort aus dem vorhergehenden Satz 1.3.6. \square

Ergänzende Übung 1.3.8. Man berechne $\int_0^1 e^{it} dt$. Hinweis: 1.1.12. Man finde eine Stammfunktion von $\cos^4 x$. Hinweis: 1.1.4.

Ergänzende Übung 1.3.9. Man formuliere und beweise eine Variante für vektorwertige Funktionen des Satzes 9.11.1 über Integrale mit Parametern.

Übung 1.3.10. Gegeben ein halboffenes kompaktes Intervall $I \subset \mathbb{R}$ und ein Banachraum Y ist auch der Raum $\mathcal{C}^1(I, Y)$ aller stetig differenzierbaren Abbildungen von I nach Y vollständig für die Norm $\|\varphi\|_1 = \|\varphi\| + \|\varphi'\|$ der gleichmäßigen Konvergenz der Funktionen und ihrer ersten Ableitungen. Hinweis: Man verwende 10.5.30 und verallgemeinere 8.1.14.

1.4 Integration rationaler Funktionen

1.4.1. Zur Integration rationaler Funktionen erinnern wir zunächst die Partialbruchzerlegung aus ???. Wenn wir vergessen, daß wir uns dabei in die komplexe Zahlenebene vorgewagt haben, so wird das Integrieren rationaler Funktionen sehr einfach: Wir haben ja in der Partialbruchzerlegung unsere rationale Funktion geschrieben als eine Linearkombination von Funktionen der Gestalt $x \mapsto (x - \mu)^m$ mit $m \in \mathbb{Z}$, und Stammfunktionen für diese Funktionen sind natürlich

$$\frac{1}{m+1}(x - \mu)^{m+1} \text{ für } m \neq -1 \text{ bzw. } \log(x - \mu) \text{ für } m = -1 \text{ und } x > \mu.$$

Und nun ist es eben so, daß diese Formeln für $\mu \in \mathbb{C}$ ganz genauso gelten, wenn wir sie nur richtig interpretieren. Im Fall $m \neq -1$ ist das der Inhalt der Übung 10.2.19 und folgt auch sofort aus den Resultaten des anschließenden Abschnitts 1.5. Im Fall $m = -1$ wird es im Folgenden ausgeführt.

1.4.2. Um zu $\mathbb{R} \setminus \mu \rightarrow \mathbb{C}$, $x \mapsto (x - \mu)^{-1}$ für beliebiges $\mu \in \mathbb{C}$ eine Stammfunktion anzugeben, überlegt man sich zunächst, daß die komplexe Exponentialfunktion eine Bijektion

$$\exp : \mathbb{R} + (-\pi, \pi]i \xrightarrow{\sim} \mathbb{C}^\times$$

definiert. Die Umkehrfunktion

$$\log : \mathbb{C}^\times \xrightarrow{\sim} \mathbb{R} + (-\pi, \pi]i$$

wird auf der positiven reellen Achse gegeben durch unseren üblichen Logarithmus $u \mapsto \log u$, auf der oberen bzw. unteren komplexen Halbebene durch die Vorschrift

$$\log(u + iv) = \log \sqrt{u^2 + v^2} \pm i \frac{\pi}{2} - i \arctan \frac{u}{v} \quad \text{für } \pm v > 0$$

und auf der negativen reellen Achse durch $u \mapsto \log(-u) + i\pi/2$. Man beachte jedoch, daß dieser **Hauptzweig des Logarithmus** nicht stetig ist längs der negativen reellen Achse, obwohl seine Einschränkung auf die negative reelle Achse durchaus stetig ist: Wir haben für $u < 0$ genauer

$$\lim_{v \searrow 0} \log(u + iv) = \log(u) = \lim_{v \nearrow 0} \log(u + iv) + 2\pi i$$

Unabhängig davon ist es mit unseren expliziten Formeln eine elementare Übung, für alle $\mu \in \mathbb{C}$ zu prüfen, daß die Funktion

$$\begin{aligned} \mathbb{R} \setminus \mu &\rightarrow \mathbb{C} \\ x &\mapsto \log(x - \mu) \end{aligned}$$

die Ableitung $1/(x - \mu)$ hat. Weniger elementar aber dafür konzeptioneller folgt es auch aus der komplexen Kettenregel 1.5.12 mitsamt der Regel für die komplexe Ableitung von Zweigen des komplexen Logarithmus 1.5.19 aus dem anschließenden Abschnitt. In jedem Fall sehen wir, wie sich im Prinzip die Stammfunktion einer beliebigen rationalen Funktion wieder als eine rationale Funktion mitsamt einigen Ausdrücken im Arcustangens und im reellen Logarithmus schreiben läßt.

1.4.3. Setzt man auch im Komplexen $a^b = \exp(b \log a)$ mit dem eben definierten Hauptzweig des komplexen Logarithmus, so ergibt sich $i^i = \exp(-\pi/2)$. Insbesondere ist in diesem Sinne also i^i reell.

Beispiel 1.4.4. Wir bestimmen eine Stammfunktion zu $1/(1+x^2)$ mithilfe unserer Partialbruchzerlegung ???. Die Nullstellen des Nenners sind $\pm i$ und der Grad des Zählers ist echt kleiner als der Grad des Nenners, wir dürfen also ansetzen

$$\frac{1}{1+x^2} = \frac{a}{x+i} + \frac{b}{x-i}$$

und finden sofort $(a+b)x - ia + ib = 1$, also $a+b = 0$ und $a-b = i$ und folglich $a = i/2$ und $b = -i/2$. Eine Stammfunktion ist mithin

$$\frac{i}{2} \log(x+i) - \frac{i}{2} \log(x-i)$$

Um die weitere Rechnung zu vereinfachen beachten wir, daß unsere Stammfunktion bis auf eine Konstante eine reellwertige Funktion sein muß, wir dürfen also bereits vor dem Addieren die Realteile nehmen und erhalten als Stammfunktion sofort

$$-\frac{\pi}{4} + \frac{1}{2} \arctan(x) - \frac{\pi}{4} - \frac{1}{2} \arctan(-x) = \arctan(x) - \frac{\pi}{2}$$

in Übereinstimmung mit unseren Ergebnissen aus [10.6.16](#).

Beispiel 1.4.5. Wir bestimmen zu $(x^4 + 2x^2)/(x^2 + 2x + 1)$ eine Stammfunktion. Die Partialbruchzerlegung haben wir bereits in ?? durchgeführt und erhielten

$$\frac{x^4 + 2x^2}{x^2 + 2x + 1} = x^2 - 2x + 5 - \frac{8}{x + 1} + \frac{3}{(x + 1)^2}$$

Als Stammfunktion finden wir damit sofort

$$\frac{x^3}{3} - x^2 + 5x - 8 \log|x + 1| - \frac{3}{(x + 1)}$$

1.4.6. Gegeben eine rationale Funktion $R = P/Q$ betrachten wir die Funktion $D \rightarrow \mathbb{R}, t \mapsto R(e^t)$ mit dem Definitionsbereich $D = \{t \in \mathbb{R} \mid Q(e^t) \neq 0\}$. Das Integral eines solchen rationalen Ausdrucks in e^t kann man auf das Integral einer rationalen Funktion zurückführen durch die Substitution $x = e^t, dx = e^t dt = x dt$. Zum Beispiel berechnen wir

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

Das Integral eines rationalen Ausdrucks in $\sqrt[n]{t}$ für eine natürliche Zahl $n \geq 1$ kann man ähnlich durch die Substitution $\sqrt[n]{t} = x, dt = nx^{n-1} dx$ auf das Integral einer rationalen Funktion in x zurückführen.

1.4.7. Das Integral eines rationalen Ausdrucks im Funktionenpaar (\sin, \cos) wie zum Beispiel

$$\frac{\sin^3(\tau) + \cos(\tau)}{\cos(\tau) + \cos^2(\tau)}$$

kann man auffassen als Kurvenintegral im Sinne von [10.3.6](#) einer rationalen Funktion in zwei Veränderlichen, in unserem Beispiel der Funktion

$$R(x, y) = \frac{y^3 + x}{x + x^2}$$

über ein Stück der Kreislinie und dann mittels der Umparametrisierung aus [10.6.17](#), d.h. mithilfe der Substitution $t = \tan(\tau/2)$ und folglich $\sin(\tau) = 2t/(1 + t^2)$,

$\cos(\tau) = (1 - t^2)/(1 + t^2)$, $d\tau = 2/(1 + t^2) dt$ umwandeln in ein Integral einer rationalen Funktion einer Veränderlichen. Integrale über rationale Ausdrücke im Funktionenpaar $(\sqrt{1 - x^2}, x)$ kann man in ähnlicher Weise als Kurvenintegral auffassen und lösen, im Gegensatz zu eben hat nur $x \mapsto (\sqrt{1 - x^2}, x)$ nicht konstante absolute Geschwindigkeit 1, sondern vielmehr die absolute Geschwindigkeit $1/\sqrt{1 - x^2}$. Formal mag man auch $x = \sin t$, $dx = \cos t dt$ substituieren und sich so auf den bereits behandelten Fall eines rationalen Ausdrucks im Funktionenpaar (\sin, \cos) zurückziehen.

1.4.8. Integrale rationaler Ausdrücke in den Funktionenpaaren $(\sqrt{x^2 + 1}, x)$ bzw. $(\sqrt{x^2 - 1}, x)$ kann man auf die bereits in 1.4.6 behandelten Integrale rationaler Funktionen in e^t zurückführen durch die Substitution $x = \sinh t$, $dx = \cosh t dt$ bzw. $x = \cosh t$, $dx = \sinh t dt$. Die geometrische Bedeutung dieses Tricks wird in 6.3.13 erklärt.

Ergänzende Übung 1.4.9. Man finde eine Stammfunktion zu $1/(1 + x^4)$.

1.5 Komplexe Differenzierbarkeit

Definition 1.5.1. Sei $U \subset \mathbb{C}$ eine Teilmenge und $p \in U$ ein Punkt. Eine Funktion $f : U \rightarrow \mathbb{C}$ heißt **komplex differenzierbar bei p mit Ableitung $b \in \mathbb{C}$** genau dann, wenn p ein Häufungspunkt von U ist und es gilt

$$\lim_{z \rightarrow p} \frac{f(z) - f(p)}{z - p} = b$$

Wir kürzen diese Aussage ab durch $f'(p) = b$.

1.5.2. Der Grenzwert ist hier im Sinne von 9.6.8 zu verstehen. Diese Definition ist fast identisch zu unserer alten Definition 7.1.3 bis auf das Detail, daß wir überall statt reeller Zahlen komplexe Zahlen betrachten und, wie im Komplexen üblich, die Variable mit z bezeichnen. Den Definitionsbereich unserer Funktion haben wir statt mit I hier mit U bezeichnet, weil der meistgebrauchte Fall nicht mehr der eines halboffenen Intervalls, sondern vielmehr der einer offenen Teilmenge der komplexen Zahlenebene ist. Der Fall eines Intervalls $U \subset \mathbb{R}$ wird jedoch auch oft vorkommen. In diesem Fall stimmt die hier definierte Ableitung überein mit der Ableitung im Sinne von 10.2.1. Der Rest dieses Abschnitts besteht darin, unsere Resultate zur reellen Differenzierbarkeit mitsamt ihren Beweisen im Komplexen zu wiederholen.

1.5.3. Ich gebe noch einige alternative Formulierungen an. Ist $U \subset \mathbb{C}$ eine Teilmenge und p ein Häufungspunkt von U , so ist nach 9.6.11 eine Funktion $f : U \rightarrow \mathbb{C}$ komplex differenzierbar bei p mit Ableitung $b \in \mathbb{C}$ genau dann, wenn es eine

Funktion $\varphi : U \rightarrow \mathbb{C}$ gibt, die stetig ist bei p mit Funktionswert $\varphi(p) = b$ derart, daß für alle $z \in U$ gilt

$$f(z) = f(p) + (z - p)\varphi(z)$$

In anderen nochmals anderen Formeln ist unsere Funktion $f : U \rightarrow \mathbb{C}$ komplex differenzierbar bei p mit Ableitung b genau dann, wenn gilt

$$f(p + h) = f(p) + bh + \varepsilon(h)h$$

für eine Funktion ε , die stetig ist bei Null und die dort den Wert Null annimmt. Hier ist zu verstehen, daß die Funktion ε definiert sein soll auf der Menge aller h mit $h + p \in U$. Diese Formulierung hat den Vorteil, daß besonders gut zum Ausdruck kommt, inwiefern für festes p und kleines h der Ausdruck $f(p) + f'(p)h$ eine gute Approximation von $f(p+h)$ ist. Anschaulich wirkt f lokal um einen gegebenen Punkt p in erster Approximation wie eine Drehstreckung mit Zentrum in besagtem Punkt, deren Winkel und Streckfaktor durch $f'(p)$ beschrieben werden, gefolgt von einer Verschiebung um $f(p)$.

Beispiele 1.5.4. Eine konstante Funktion auf einer Menge von komplexen Zahlen ist bei jedem Häufungspunkt besagter Menge komplex differenzierbar mit Ableitung Null. Die Funktion $\text{id} : \mathbb{C} \rightarrow \mathbb{C}$, $z \mapsto z$ hat bei jedem Punkt p die Ableitung $\text{id}'(p) = 1$.

Lemma 1.5.5. Die Funktion $z \mapsto \frac{1}{z}$ ist komplex differenzierbar bei jedem Punkt von \mathbb{C}^\times und ihre Ableitung bei einer Stelle $p \in \mathbb{C}^\times$ ist $-\frac{1}{p^2}$.

Beweis. Wir rechnen $\lim_{z \rightarrow p} \frac{\frac{1}{z} - \frac{1}{p}}{z - p} = \lim_{z \rightarrow p} \frac{-1}{zp} = -\frac{1}{p^2}$. □

Lemma 1.5.6. Sei $U \subset \mathbb{C}$ eine Teilmenge. Ist eine Funktion $f : U \rightarrow \mathbb{C}$ komplex differenzierbar bei $p \in U$, so ist f stetig bei p .

Beweis. Das folgt sofort aus 1.5.3. □

Proposition 1.5.7. Sei $U \subset \mathbb{C}$ eine Teilmenge und seien $f, g : U \rightarrow \mathbb{C}$ komplex differenzierbar bei einem Punkt $p \in U$. So sind auch die Funktionen $f + g$ und fg komplex differenzierbar bei p und es gilt

$$(f + g)'(p) = f'(p) + g'(p) \quad \text{und} \quad (fg)'(p) = f'(p)g(p) + f(p)g'(p)$$

Beweis. Identisch zum Beweis im Reellen nach 7.2.1. □

Definition 1.5.8. Ist eine Funktion $f : U \rightarrow \mathbb{C}$ definiert auf einer Teilmenge $U \subset \mathbb{C}$ und differenzierbar bei jedem Punkt von U , so nennen wir f **komplex differenzierbar auf U** und nennen die Funktion $f' : U \rightarrow \mathbb{C}$, $p \mapsto f'(p)$ ihre **Ableitung**. Aus unseren Definitionen folgt natürlich insbesondere, daß dann die Menge U keine isolierten Punkte haben darf.

1.5.9. Für die Ableitungen komplex differenzierbarer Funktionen mit gemeinsamem Definitionsbereich gelten mithin die **Summenregel** und die **Produktregel** oder **Leibniz-Regel**

$$(f + g)' = f' + g' \quad \text{und} \quad (fg)' = f'g + fg'$$

Korollar 1.5.10 (Ableiten ganzzahliger Potenzen). Für alle $n \in \mathbb{Z}$ und unter der Voraussetzung $z \neq 0$ im Fall $n \leq 0$ ist die Ableitung der Funktion $z \mapsto z^n$ die Funktion $z \mapsto nz^{n-1}$.

Beweis. Man zeigt das durch vollständige Induktion über n separat für $n \geq 0$ und $n \leq -1$. □

Übung 1.5.11. Ein komplexes Polynom hat bei $\lambda \in \mathbb{C}$ eine mehrfache Nullstelle genau dann, wenn auch seine Ableitung bei λ verschwindet.

Satz 1.5.12 (Kettenregel). Seien $U, V \subset \mathbb{C}$ Teilmengen und $f : U \rightarrow \mathbb{C}$ und $g : V \rightarrow \mathbb{C}$ Funktionen und es gelte $f(U) \subset V$. Sei f komplex differenzierbar bei p und g komplex differenzierbar bei $f(p)$. So ist $g \circ f : U \rightarrow \mathbb{C}$ komplex differenzierbar bei p mit Ableitung

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

Beweis. Identisch zum Beweis im Reellen nach 7.2.5. Man beachte, daß nun rechts ein Produkt komplexer Zahlen steht. □

Beispiel 1.5.13. Wir berechnen für $\lambda, \mu \in \mathbb{C}$ und $m \geq 1$ eine natürliche Zahl die Ableitung der Funktion $\mathbb{R} \rightarrow \mathbb{C}$ gegeben durch $f : t \mapsto (t^2 + \lambda t + \mu)^m$ und erhalten mit der Kettenregel $f'(t) = (2t + \lambda)m(t^2 + \lambda t + \mu)^{m-1}$. Schalten wir noch eine differenzierbare Abbildung $t : \mathbb{R} \rightarrow \mathbb{R}$, $\tau \mapsto t(\tau)$ davor, so ergibt sich die Ableitung der zusammengesetzten Funktion wieder mit der Kettenregel zu

$$\frac{df}{d\tau} = \frac{df}{dt} \frac{dt}{d\tau} = (2t(\tau) + \lambda)m(t(\tau)^2 + \lambda t(\tau) + \mu)^{m-1} \frac{dt}{d\tau}$$

Proposition 1.5.14 (Quotientenregel). Sei $U \subset \mathbb{C}$ eine Teilmenge, $f : U \rightarrow \mathbb{C}$ eine Funktion ohne Nullstelle und $p \in U$ ein Punkt.

1. Ist f komplex differenzierbar bei p , so ist auch $z \mapsto 1/f(z)$ komplex differenzierbar bei p und hat dort die Ableitung $-f'(p)/f(p)^2$.
2. Ist zusätzlich $g : U \rightarrow \mathbb{C}$ komplex differenzierbar bei p , so ist auch g/f komplex differenzierbar bei p mit Ableitung

$$\left(\frac{g}{f}\right)'(p) = \frac{g'(p)f(p) - g(p)f'(p)}{f(p)^2}$$

Beweis. Teil 1 folgt sofort aus 1.5.5 mit der Kettenregel 1.5.12. Teil 2 folgt aus Teil 1 mit der Produktregel 1.5.7. \square

Satz 1.5.15 (Ableitung von Umkehrfunktionen). Sei $U \subseteq \mathbb{C}$ offen und sei $f : U \rightarrow \mathbb{C}$ eine stetige Injektion mit offenem Bild und stetiger Umkehrung $f^{-1} : f(U) \rightarrow U$. Ist dann f komplex differenzierbar beim Punkt $p \in U$ mit Ableitung $f'(p) \neq 0$, so ist auch die Umkehrfunktion $f^{-1} : f(U) \rightarrow \mathbb{C}$ komplex differenzierbar bei $q = f(p)$ mit Ableitung

$$(f^{-1})'(q) = 1/f'(f^{-1}(q))$$

1.5.16. In ?? werden wir zeigen, daß eine injektive komplex differenzierbare Funktion mit offenem Definitionsbereich stets offenes Bild und eine stetige Umkehrung hat. Ein Teil der Bedingungen an unsere Funktion sind also eigentlich überflüssig und dienen nur dazu, den Beweis zu vereinfachen.

Beweis. Nach unseren Annahmen gibt es eine stetige Funktion ohne Nullstelle $\varphi : U \rightarrow \mathbb{C}$ mit $f(z) - f(p) = (z - p)\varphi(z)$ und $\varphi(p) = f'(p)$. Setzen wir hier $z = f^{-1}(w)$, so ist $\psi = 1/(\varphi \circ f^{-1}) : f(U) \rightarrow \mathbb{C}$ eine stetige Funktion mit $(w - q)\psi(w) = f^{-1}(w) - f^{-1}(q)$ und $\psi(q) = 1/f'(p)$. \square

Beispiel 1.5.17. Das Quadrieren liefert eine Bijektion zwischen der Halbebene aller komplexen Zahlen mit positivem Realteil und der “geschlitzten Zahlenebene” $\mathbb{C} \setminus \mathbb{R}_{\leq 0}$. Die Umkehrfunktion zu dieser Bijektion ist also eine komplex differenzierbare Funktion auf der geschlitzten Zahlenebene, die wir \sqrt{z} notieren und die nach 1.5.15 differenzierbar ist mit Ableitung $1/(2\sqrt{z})$.

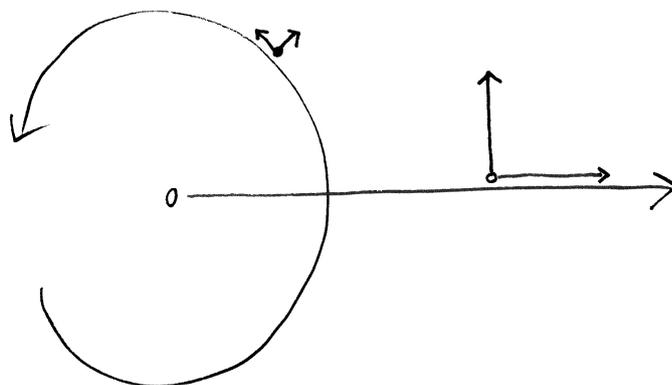
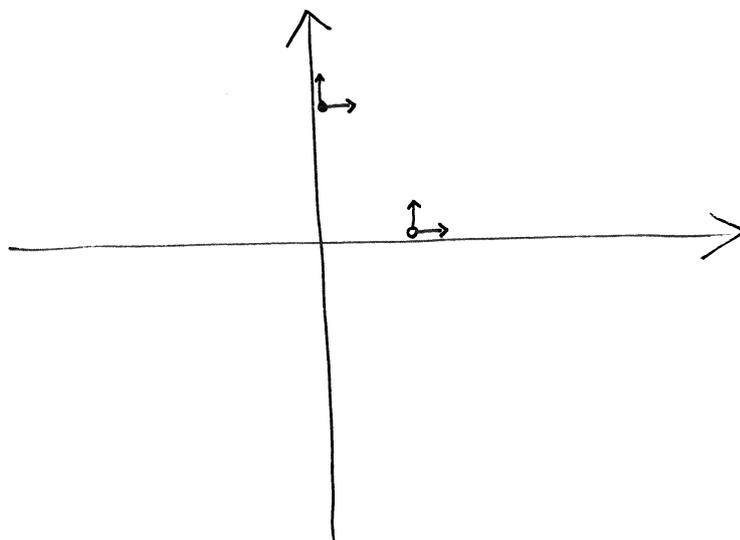
Lemma 1.5.18. Die komplexe Exponentialfunktion ist komplex differenzierbar und stimmt auf der ganzen komplexen Zahlenebene mit ihrer eigenen Ableitung überein.

Beweis. Der Beweis des reellen Analogons 7.2.8 kann wortwörtlich übernommen werden. \square

Beispiel 1.5.19. Ist $U \subseteq \mathbb{C}$ eine offene Teilmenge derart, daß die Exponentialfunktion eine Injektion mit offenem Bild und stetiger Umkehrfunktion $\log : \exp(U) \rightarrow \mathbb{C}$ liefert, so nennt man \log einen **Zweig des Logarithmus**. Nach 1.5.15 ist jeder solche Zweig des Logarithmus komplex differenzierbar mit Ableitung

$$\log'(q) = \frac{1}{\exp(\log q)} = \frac{1}{q}$$

Im Spezialfall $U = \mathbb{R} + (-\pi, \pi)i$ spricht man auch vom **Hauptzweig des Logarithmus**, den wir bereits in 1.4.2 eingeführt und sogar noch auf die negative reelle Achse fortgesetzt hatten, allerdings in nur noch partiell stetiger Weise.



Anschauliche Bedeutung der Ableitung der komplexen Exponentialfunktion

Übung 1.5.20. Man zeige, daß für alle $z \in \mathbb{C}$ mit $|z| < 1$ der Hauptzweig des Logarithmus von $1 + z$ auch dargestellt werden kann durch die Potenzreihe

$$\log(1 + z) = z - \frac{z^2}{2} + \frac{z^3}{3} - \dots$$

Hinweis: Es reicht zu zeigen, daß für alle $u \in \mathbb{C}$ mit $|u| = 1$ das Einsetzen von $z = ut$ auf beiden Seiten dieselbe Funktion in $t \in (-1, 1)$ liefert. Beide Seiten nehmen aber bei $z = 0$ den Wert Null an, so daß es reicht, die Gleichheit ihrer Ableitungen zu zeigen. Im Rahmen der Funktionentheorie dürfen Sie diese Übung in ?? mit mehr Theorie und weniger Rechnen ein weiteres Mal lösen.

1.5.21. Eine komplex differenzierbare komplexwertige Funktion, die auf einer offenen Teilmenge der komplexen Zahlenebene definiert ist, heißt eine **holomorphe Funktion**. Die Theorie der holomorphen Funktionen, die sogenannte **Funktionentheorie**, ist grundlegend verschieden von der Theorie der differenzierbaren reellwertigen Funktionen auf einer offenen Teilmenge der reellen Zahlengerade, die wir in dieser Vorlesung ausführlich studiert haben. Zum Beispiel ist jede holomorphe Funktion, d.h. jede auf einer offenen Teilmenge der komplexen Zahlenebene definierte einmal komplex differenzierbare Funktion, bereits beliebig oft komplex differenzierbar. Mehr dazu findet man etwa in ??.

2 Lösung einiger Schwingungsgleichungen

2.1 Gedämpfte Schwingungen

2.1.1. Wir interessieren uns für die Bewegung eines Massepunktes, der an einer Feder aufgehängt ist und dessen Bewegung durch eine zur Geschwindigkeit proportionale Reibung gedämpft wird. Mißt die Funktion $x : \mathbb{R} \rightarrow \mathbb{R}, t \mapsto x(t)$ seine Auslenkung von der Gleichgewichtslage zum Zeitpunkt t , so muß unsere Funktion aus physikalischen Gründen eine Differentialgleichung zweiten Grades der Gestalt

$$\ddot{x} = -a\dot{x} - bx$$

erfüllen, wobei die Konstanten a und b die Stärke der Feder und der Dämpfung ausdrücken und in physikalisch relevanten Fällen nichtnegativ sind. Wir lösen diese Differentialgleichung hier erst einmal ad hoc und erheben danach in 2.2.2 diesen Zugang zur Methode.

Proposition 2.1.2 (zur Lösung der Schwingungsgleichung). *Seien reelle Zahlen $a, b \in \mathbb{R}$ gegeben.*

1. Die Menge aller zweimal differenzierbaren Funktionen $x : \mathbb{R} \rightarrow \mathbb{R}$ mit $\ddot{x} + a\dot{x} + bx = 0$ bildet einen Untervektorraum des Raums $\text{Ens}(\mathbb{R}, \mathbb{R})$ aller Abbildungen $\mathbb{R} \rightarrow \mathbb{R}$, den **Lösungsraum** L unserer Differentialgleichung.
2. Die Abbildung $x \mapsto (x(0), \dot{x}(0))$ definiert einen Vektorraumisomorphismus $L \xrightarrow{\sim} \mathbb{R}^2$ dieses Lösungsraums mit dem \mathbb{R}^2 , den **Anfangswertisomorphismus**.
3. Hat das Polynom $X^2 + aX + b$ zwei verschiedene reelle Nullstellen λ und μ , so bilden die beiden Funktionen $x_1(t) = e^{\lambda t}$ und $x_2(t) = e^{\mu t}$ eine Basis des Lösungsraums. Hat es dahingegen eine doppelte reelle Nullstelle λ , so bilden die beiden Funktionen $x_1(t) = e^{\lambda t}$ und $x_2(t) = t e^{\lambda t}$ eine Basis des Lösungsraums.

Bemerkung 2.1.3. Um den Fall, daß unser Polynom gar keine reelle Nullstelle hat, werden wir uns gleich noch gesondert kümmern.

Beweis. Teil 1 scheint mir offensichtlich. Um Teil 2 zu zeigen beachten wir, daß die Vorschrift $x \mapsto (x, \dot{x})$ offensichtlich einen Isomorphismus zwischen unserem Lösungsraum L und dem Lösungsraum des Systems $\dot{\gamma}_1 = \gamma_2, \dot{\gamma}_2 = -b\gamma_1 - a\gamma_2$ induziert, das in Matrixschreibweise die Gestalt $\dot{\gamma} = A\gamma$ annimmt mit der Matrix

$$A = \begin{pmatrix} 0 & 1 \\ -b & -a \end{pmatrix}$$

Teil 2 folgt damit aus 10.4.13. Für Teil 3 müssen wir folglich nur prüfen, daß die beiden angegebenen Funktionen in der Tat linear unabhängige Lösungen sind. Das kann dem Leser überlassen bleiben. \square

2.1.4. Statt in Teil 3 mögliche Lösungen einfach zu erraten, hätten wir uns auch daran erinnern können, daß ja nach 10.4.9 jede Lösung von der Form

$$x(t) = \gamma_1(t) = \text{pr}_1(\exp(tA)c)$$

sein muß für $c = (x(0), \dot{x}(0))$. Das charakteristische Polynom unserer Matrix A ist aber nun gerade $X^2 + aX + b$. Hat es zwei verschiedene reelle Nullstellen λ, μ und bilden wir eine Matrix P mit Eigenvektoren zu λ und μ als Spalten, so gilt $A = P \text{diag}(\lambda, \mu)P^{-1}$ und $\exp(tA) = P \text{diag}(e^{\lambda t}, e^{\mu t})P^{-1}$ und wir erkennen auf Anhieb, daß jede Lösung eine Linearkombination der Gestalt $x(t) = \alpha e^{\lambda t} + \beta e^{\mu t}$ sein muß. Im Fall einer doppelten reellen Nullstelle finden wir ähnlich ein P mit

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

und 10.5.24 liefert

$$\exp \begin{pmatrix} u & t \\ 0 & u \end{pmatrix} = \exp \begin{pmatrix} u & 0 \\ 0 & u \end{pmatrix} \exp \begin{pmatrix} 0 & t \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} e^u & 0 \\ 0 & e^u \end{pmatrix} \begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} e^u & t e^u \\ 0 & e^u \end{pmatrix}$$

womit sich die allgemeine Lösung ergibt als eine Linearkombination der Gestalt $\alpha e^{\lambda t} + \beta t e^{\lambda t}$.

2.1.5. Im Fall der gedämpften Schwingung hat unser Polynom $X^2 + aX + b$ die beiden Nullstellen

$$-\frac{a}{2} \pm \sqrt{\frac{a^2}{4} - b}$$

Bei hinreichend großer Dämpfung $a^2/4 \geq b$ erhalten wir reelle nichtpositive Lösungen und unser Massepunkt kehrt mit höchstens einmaligem Überschwingen zum Ruhezustand zurück. Im Fall kleiner Dämpfung $a^2/4 < b$ hat unser Polynom dahingegen keine reellen Nullstellen mehr und stattdessen die beiden komplexen Nullstellen $\pm i\omega - a/2$ mit $\omega = \sqrt{b - a^2/4} > 0$. Um hier weiterzukommen verallgemeinern wir zunächst einmal alles bisher Gesagte ins Komplexe.

Proposition 2.1.6 (Lösung der Schwingungsgleichung). *Seien komplexe Zahlen $a, b \in \mathbb{C}$ gegeben.*

1. *Die Menge aller zweimal differenzierbaren Funktionen $x : \mathbb{R} \rightarrow \mathbb{C}$ mit $\ddot{x} + a\dot{x} + bx = 0$ bildet einen komplexen Untervektorraum des Raums $\text{Ens}(\mathbb{R}, \mathbb{C})$ aller Abbildungen $\mathbb{R} \rightarrow \mathbb{C}$, den **Lösungsraum** L unserer Differentialgleichung.*

2. Die Abbildung $x \mapsto (x(0), \dot{x}(0))$ definiert einen Vektorraumisomorphismus $L \xrightarrow{\sim} \mathbb{C}^2$, den **Anfangswertisomorphismus**.
3. Hat das Polynom $X^2 + aX + b$ zwei verschiedene Nullstellen λ und μ , so bilden die beiden Funktionen $x_1(t) = e^{\lambda t}$ und $x_2(t) = e^{\mu t}$ eine Basis des Lösungsraums L . Hat es eine doppelte Nullstelle λ , so bilden die beiden Funktionen $x_1(t) = e^{\lambda t}$ und $x_2(t) = t e^{\lambda t}$ eine Basis des Lösungsraums.

Beweis. Der Beweis ist identisch zum Beweis der reellen Version 2.1.6, sobald man die dabei benötigten Hilfsmittel ins Komplexe verallgemeinert hat, was wir im weiteren Verlauf dieses Abschnitts und insbesondere in 2.1.10 tun werden. \square

2.1.7. Sind in der Situation aus 2.1.6 die Koeffizienten a, b beide reell, so bilden die reellwertigen Lösungen unseres Systems nach 2.1.2 einen zweidimensionalen reellen Untervektorraum $L_{\mathbb{R}} \subset \text{Ens}(\mathbb{R}, \mathbb{R})$ und seine komplexwertigen Lösungen bilden nach 2.1.6 einen zweidimensionalen komplexen Untervektorraum $L_{\mathbb{C}} \subset \text{Ens}(\mathbb{R}, \mathbb{C})$, der stabil ist unter dem Übergang zum komplex Konjugierten, in Formeln $f \in L_{\mathbb{C}} \Rightarrow \bar{f} \in L_{\mathbb{C}}$. Per definitionem gilt weiter $L_{\mathbb{R}} = L_{\mathbb{C}} \cap \text{Ens}(\mathbb{R}, \mathbb{R})$. Haben wir nun Erzeuger f_1, \dots, f_r für den \mathbb{C} -Vektorraum der komplexwertigen Lösungen gefunden, so erzeugen deren Realteile zusammen mit ihren Imaginärteilen den \mathbb{R} -Vektorraum der reellwertigen Lösungen: In der Tat schreibt sich ja jede reellwertige Lösung f als $f = c_1 f_1 + \dots + c_r f_r$ mit $c_\nu \in \mathbb{C}$, und bilden wir hier auf beiden Seiten den Realteil, so ergibt sich für f die Darstellung

$$f = \text{Re}(c_1) \text{Re}(f_1) - \text{Im}(c_1) \text{Im}(f_1) + \dots + \text{Re}(c_r) \text{Re}(f_r) - \text{Im}(c_r) \text{Im}(f_r)$$

2.1.8. Im Fall der gedämpften Schwingungen 2.1.1 mit kleiner Dämpfung und folglich komplexen Nullstellen $\pm i\omega - a/2$ erhalten wir die komplexen Lösungen $x_{\pm}(t) = e^{-at/2} e^{\pm i\omega t}$ und die Euler-Formel liefert, daß die Funktionen

$$x_1(t) = e^{-at/2} \cos \omega t \quad \text{und} \quad x_2(t) = e^{-at/2} \sin \omega t$$

den Raum der reellwertigen Lösungen aufspannen. Die Größe ω wird in diesem Zusammenhang auch als **Winkelgeschwindigkeit** bezeichnet. Die Additionstheoreme zeigen, daß sich jede reelle Linearkombination $\alpha \sin(\omega t) + \beta \cos(\omega t)$ der Funktionen $\cos(\omega t)$ und $\sin(\omega t)$ als Sinuswelle mit **Amplitude** k und **Phase** ϕ in der Form

$$\alpha \sin(\omega t) + \beta \cos(\omega t) = k \sin(\omega t + \phi)$$

schreiben läßt, für $k = \sqrt{\alpha^2 + \beta^2}$ und ϕ einer Lösung des Gleichungssystems $k \cos \phi = \alpha$ und $k \sin \phi = \beta$. Im Fall kleiner Dämpfung kann die allgemeine Lösung also geschrieben werden als $x(t) = k e^{-at/2} \sin(\omega t + \phi)$ und beschreibt eine Schwingung, deren Amplitude bei positiver Dämpfung $a > 0$ exponentiell abfällt.

2.1.9. Man kann ohne Schwierigkeiten die Exponentialabbildung auf quadratischen Matrizen ins Komplexe erweitern zu

$$\begin{aligned} \exp : M(n \times n; \mathbb{C}) &\rightarrow M(n \times n; \mathbb{C}) \\ A &\mapsto \sum_{k=0}^{\infty} \frac{1}{k!} A^k \end{aligned}$$

Satz 2.1.10 (Lineare Differentialgleichungen). *Ist $A \in M(n \times n; \mathbb{C})$ eine quadratische Matrix und $c \in \mathbb{C}^n$ ein Spaltenvektor, so gibt es genau eine differenzierbare Abbildung $\gamma : \mathbb{R} \rightarrow \mathbb{C}^n$ mit Anfangswert $\gamma(0) = c$ derart, daß gilt $\dot{\gamma}(t) = A\gamma(t)$ für alle $t \in \mathbb{R}$, und diese Abbildung wird gegeben durch die Vorschrift*

$$\gamma(t) = \exp(tA)c$$

Beweis. Mutatis mutandis, als da heißt nach Verändern des zu Verändernden identisch zum Beweis von 10.4.9. □

Korollar 2.1.11 (Anfangswertisomorphismus). *Ist $A \in M(n \times n; \mathbb{C})$ eine quadratische Matrix, so bilden die differenzierbaren Abbildungen $\gamma : \mathbb{R} \rightarrow \mathbb{C}^n$ mit $\dot{\gamma}(t) = A\gamma(t) \quad \forall t \in \mathbb{R}$ einen komplexen Untervektorraum $L \subset \text{Ens}(\mathbb{R}, \mathbb{C}^n)$ und an jeder Stelle liefert das Auswerten einen Isomorphismus $L \xrightarrow{\sim} \mathbb{C}^n$.*

Beweis. Dem Leser überlassen. Im Reellen war das Übung 10.4.13. □

2.1.12. Die Regel $\exp(PAP^{-1}) = P(\exp A)P^{-1}$ aus 10.5.28 gilt genauso für komplexe Matrizen. Die Berechnung des Exponentials einer beliebigen quadratischen Matrix wird Ihnen auf dieser Grundlage leicht gelingen, sobald sie in der linearen Algebra die Theorie der “Jordan’schen Normalform” ?? kennengelernt haben.

Übung 2.1.13. Ist $A \in M(n \times n; \mathbb{R})$ eine reelle Matrix und $\gamma : \mathbb{R} \rightarrow \mathbb{C}^n$ eine komplexe Lösung der Differentialgleichung $\dot{\gamma}(t) = A\gamma(t)$, so sind ihr koordinatenweise gebildeter Real- und Imaginärteil $\text{Re } \gamma$ und $\text{Im } \gamma$ reelle Lösungen. Erzeugt eine Menge \mathbb{C}^n -wertiger Funktionen den \mathbb{C} -Vektorraum der \mathbb{C}^n -wertigen Lösungen unserer Differentialgleichung, so erzeugen ihre Real- und Imaginärteile zusammen den \mathbb{R} -Vektorraum der \mathbb{R}^n -wertigen Lösungen.

2.2 Lineare Differentialgleichungen höherer Ordnung

2.2.1. Die Erfahrungen, die wir bei der Behandlung gedämpfter Schwingungen gemacht haben, fassen wir nun noch etwas allgemeiner.

Satz 2.2.2. *Seien komplexe Zahlen $a_0, \dots, a_{n-1} \in \mathbb{C}$ gegeben.*

1. Die komplexwertigen n -mal differenzierbaren Funktionen $f : \mathbb{R} \rightarrow \mathbb{C}$ mit $f^{(n)} + a_{n-1}f^{(n-1)} + \dots + a_0f = 0$ bilden einen Untervektorraum im Raum aller Funktionen $\mathbb{R} \rightarrow \mathbb{C}$, den **Lösungsraum** unserer Differentialgleichung.
2. Die Abbildung $f \mapsto (f(0), f'(0), \dots, f^{(n-1)}(0))$ ist ein Isomorphismus dieses Lösungsraums mit dem \mathbb{C}^n , der **Anfangswertisomorphismus**.
3. Ist $\lambda \in \mathbb{C}$ eine Nullstelle des Polynoms $X^n + a_{n-1}X^{n-1} + \dots + a_0$ der Vielfachheit r , so sind die Funktionen $e^{\lambda t}, t e^{\lambda t}, \dots, t^{r-1} e^{\lambda t}$ Lösungen unserer Differentialgleichung, und durchläuft λ alle Nullstellen unseres Polynoms, so bilden diese Lösungen eine Basis des Lösungsraums.

Ergänzung 2.2.3. Der Satz bleibt gültig, wenn wir darin überall \mathbb{R} durch ein beliebiges halboffenes Intervall $I \subset \mathbb{R}$ ersetzen.

Beweis. 1 ist offensichtlich. Um die in 2 behauptete Existenz und Eindeutigkeit zu zeigen beachten wir zunächst, daß unsere Überlegungen aus 10.4.9 ohne Änderungen auch im Komplexen gültig sind. Für eine quadratische Matrix $A \in M(n \times n; \mathbb{C})$ mit komplexen Einträgen haben also die differenzierbaren Funktionen $g : \mathbb{R} \rightarrow \mathbb{C}^n$, die die Differentialgleichung $g' = Ag$ lösen, die Form $g(t) = (\exp tA)g(0)$ wo wir den Anfangswert $g(0) \in \mathbb{C}^n$ frei wählen dürfen. Insbesondere definiert die Abbildung $g \mapsto g(0)$ einen Isomorphismus vom Lösungsraum der Differentialgleichung $g' = Ag$ mit dem \mathbb{C}^n . Jetzt beachten wir, daß die Vorschrift $f \mapsto g = (f, f', f'', \dots, f^{(n-1)})^\top$ eine Bijektion induziert zwischen der Menge aller n -mal differenzierbaren Funktionen $f : \mathbb{R} \rightarrow \mathbb{C}$, die die Differentialgleichung aus dem Satz erfüllen, und der Menge aller differenzierbaren Funktionen $g : \mathbb{R} \rightarrow \mathbb{C}^n$, die das System von Differentialgleichungen

$$\begin{aligned} g_0' &= g_1 \\ g_1' &= g_2 \\ &\vdots \\ g_{n-1}' &= a_{n-1}g_{n-1} + \dots + a_1g_1 + a_0g_0 \end{aligned}$$

lösen, wo wir etwas ungewöhnlich $g = (g_0, \dots, g_{n-1})$ indiziert haben der besseren Übersichtlichkeit halber. Damit ist auch Teil 2 bewiesen.

3. Motiviert durch unsere Erkenntnisse bei der Lösung von 2.1.6 beginnen wir mit dem Ansatz $f(t) = e^{\lambda t}$ für $\lambda \in \mathbb{C}$. Mögliche λ sind dann genau die Nullstellen des Polynoms $X^n + a_{n-1}X^{n-1} + \dots + a_1X + a_0$. Ist λ eine Nullstelle der Vielfachheit r , so sind sogar, wieder in Verallgemeinerung unserer Erkenntnisse bei der Lösung von 2.1.6, auch $t e^{\lambda t}, \dots, t^{r-1} e^{\lambda t}$ noch Lösungen unserer Gleichung. Um das einzusehen, betrachten wir den Vektorraum $\mathcal{C}^\infty(\mathbb{R})$ aller beliebig oft differenzierbaren Funktionen $\mathbb{R} \rightarrow \mathbb{C}$ und fassen das Ableiten auf als eine lineare

Abbildung

$$D : \mathcal{C}^\infty(\mathbb{R}) \rightarrow \mathcal{C}^\infty(\mathbb{R})$$

Zerfällt unser Polynom in Linearfaktoren

$$X^n + a_{n-1}X^{n-1} + \dots + a_0 = (X - \lambda_1)^{n_1} \dots (X - \lambda_r)^{n_r}$$

so können wir den Operator $D^n + a_{n-1}D^{n-1} + \dots + a_0 : \mathcal{C}^\infty(\mathbb{R}) \rightarrow \mathcal{C}^\infty(\mathbb{R})$ auch schreiben als Verknüpfung der Operatoren $(D - \lambda_i)^{n_i}$, und es reicht folglich $(D - \lambda)^r t^{r-1} e^{\lambda t} = 0$ nachzuweisen. Nun gilt aber offensichtlich

$$(D - \lambda)t^m e^{\lambda t} = mt^{m-1} e^{\lambda t}$$

und die Behauptung folgt per Induktion. Um zu zeigen, daß die $t^j e^{t\lambda_i}$ für $0 \leq j < n_i$ eine Basis des Lösungsraums bilden, reicht es die lineare Unabhängigkeit nachzuweisen. Beherrscht man die zugehörige lineare Algebra, so erkennt man leicht, daß die $t^m e^{\lambda t}$ jeweils zum Hauptraum $\text{Hau}(D; \lambda)$ gehören und muß wegen ?? nur noch die lineare Unabhängigkeit der $t^m e^{\lambda t}$ für festes λ und variables m zeigen, die wiederum sofort aus der linearen Unabhängigkeit der Funktionen t^m folgt. Man vergleiche auch ??.

Beherrscht man die zugehörige lineare Algebra noch nicht, so muß man mehr arbeiten. Man setzt dann etwa eine Linearkombination $\sum c_{ji} t^j e^{\lambda_i t} = 0$ an und muß zeigen, daß alle c_{ji} verschwinden. Sonst könnten wir aber nach eventueller Umnummerierung der Nullstellen ein k finden mit $c_{k1} \neq 0$ aber $c_{j1} = 0$ für $j > k$. Wenden wir dann auf unsere Summe den Differentialoperator $(D - \lambda_1)^k (D - \lambda_2)^N \dots (D - \lambda_r)^N$ an für hinreichend grosses N , so ergibt sich $c_{k1} e^{t\lambda_1} = 0$ im Widerspruch zu unserer Annahme $c_{k1} \neq 0$. \square

Ergänzende Übung 2.2.4. Man bestimme eine Basis des komplexen sowie des reellen Lösungsraums der Differentialgleichung $f''' = f$.

2.3 Gekoppelte Schwingungen

Beispiel 2.3.1. An gegenüberliegenden Wänden eines Zimmers ist jeweils ein Wägelchen mit einer Feder befestigt und die beiden Wägelchen sind auch untereinander durch eine Feder verbunden. Bezeichnen $x(t)$ bzw. $y(t)$ die Position des ersten bzw. zweiten Wägelchens auf einer Skala, auf der $x = y = 0$ den Gleichgewichtszustand bedeuten und größere x bzw. y einen größeren Abstand eines Wägelchens von "seiner" Wand, so genügt unser System einer Differentialgleichung

$$\begin{aligned}\ddot{x} &= -ax - b(x + y) \\ \ddot{y} &= -cy - d(y + x)\end{aligned}$$

für Konstanten $a, b, c, d > 0$, in die die Stärke der Federn und die Massen der Wägelchen eingehen. Erklären wir $v : \mathbb{R} \rightarrow \mathbb{R}^2, t \mapsto v(t) = (x(t), y(t))$ und betrachten die Matrix

$$A = \begin{pmatrix} -(a+b) & -b \\ -d & -(c+d) \end{pmatrix}$$

so können wir unser System schreiben als

$$\ddot{v}(t) = Av(t)$$

Der Leser mag als Übung zeigen, daß der Lösungsraum vierdimensional sein muß. Unsere Matrix A hat, wie man dem charakteristischen Polynom ansieht, negative reelle Eigenwerte λ_1, λ_2 . Also hat bereits der \mathbb{R}^2 eine Basis v_1, v_2 aus Eigenvektoren von A . Dann sind die vier Funktionen

$$t \mapsto \exp((\pm\sqrt{\lambda_i})t)v_i \quad \text{mit } i = 1, 2$$

offensichtlich Lösungen, und ähnliche Argumente wie im vorhergehenden Beispiel zeigen, daß sie sogar eine Basis Lösungsraums bilden. Setzen wir $\omega_i = \sqrt{-\lambda_i}$, so erhalten wir eine alternative Basis des Lösungsraums durch die vier Funktionen

$$\cos(t\omega_i)v_i \quad \text{und} \quad \sin(t\omega_i)v_i \quad \text{mit } i = 1, 2.$$

Ist noch spezieller unsere Situation symmetrisch unter der Vertauschung der beiden Wägelchen, haben sie also dieselbe Masse und sind durch dieselben Federn mit den Wänden verbunden, so folgt $b = d$ und $a = c$ und wir erhalten $v_1 = (1, 1)$ mit $\lambda_1 = -a - 2b$ sowie $v_2 = (1, -1)$ mit $\lambda_2 = -a$. Diese Eigenvektoren entsprechen den zwei **Eigenschwingungen** des Systems, bei denen beide Wägelchen zu allen Zeiten in derselben bzw. in entgegengesetzten Richtungen fahren. Die Bewegung der einzelnen Wägelchen $x(t) = x_+(t)$ und $y(t) = x_-(t)$ wird dann beschrieben durch

$$\operatorname{Re} (c_1 e^{i\omega_1 t} \pm c_2 e^{i\omega_2 t}) = \operatorname{Re} (e^{i(\omega_1 - \omega_2)t/2} (c_1 e^{i(\omega_1 + \omega_2)t/2} \pm c_2 e^{-i(\omega_1 + \omega_2)t/2}))$$

mit komplexen c_i . Nimmt man hier zum Beispiel $c_1 = c_2 = 1$, so ergibt sich die Lösung

$$\begin{aligned} x(t) &= 2 \cos((\omega_1 - \omega_2)t/2) \cos((\omega_1 + \omega_2)t/2) \\ y(t) &= 2 \sin((\omega_1 - \omega_2)t/2) \sin((\omega_1 + \omega_2)t/2) \end{aligned}$$

Ist die verbindende Feder schwach im Verhältnis zu den Federn gegen die Wände, in Formeln $a \gg b$, so liegen die beiden Eigenwerte λ_1, λ_2 und damit auch die Winkelgeschwindigkeiten ω_1, ω_2 verhältnismäßig nah beieinander. Im Versuch kann man in diesem Fall schön sehen, wie die beiden Wägelchen mit der Winkelgeschwindigkeit $(\omega_1 - \omega_2)/2$ ihre Energie untereinander austauschen.

2.3.2. Im Übrigen ist es auch a priori klar, daß in der symmetrischen Situation die zweielementige Symmetriegruppe unserer Gleichung, die der Vertauschung der beiden Wägelchen entspricht, auf dem Lösungsraum operieren muß, daß wir also uns schon von Anfang an hätten darauf beschränken dürfen, nur die symmetrischen und die antisymmetrischen Lösungen zu bestimmen und die allgemeine Lösung als Linearkombination solcher speziellen Lösungen zu erhalten.

2.4 Angeregte Schwingungen

2.4.1. Ist wieder A eine komplexe $(n \times n)$ -Matrix und ist zusätzlich eine stetige Funktion $f : \mathbb{R} \rightarrow \mathbb{C}^n$ vorgegeben und man sucht alle differenzierbaren $\gamma : \mathbb{R} \rightarrow \mathbb{C}^n$, die das "inhomogene" System von Differentialgleichungen

$$\dot{\gamma}(t) = A\gamma(t) + f(t)$$

lösen, so rät einem die Methode der **Variation der Konstanten** zum Ansatz

$$\gamma(t) = \exp(tA)g(t)$$

Man erkennt leicht, daß dieser Ansatz eine Lösung liefert, wenn $g : \mathbb{R} \rightarrow \mathbb{C}^n$ differenzierbar ist und die Gleichung $f(t) = \exp(tA)\dot{g}(t)$ erfüllt, als da heißt für

$$g(t) = \int^t \exp(-\tau A)f(\tau) d\tau$$

wobei $g(t)$ als unbestimmtes Integral natürlich nur bis auf eine additive Konstante aus dem \mathbb{C}^n wohldefiniert ist. Daß wir mit diesem Verfahren tatsächlich auch alle Lösungen $\gamma(t)$ unseres inhomogenen Systems von Differentialgleichungen erhalten ergibt sich daraus, daß ja ganz offensichtlich die Differenz von je zwei Lösungen unserer inhomogenen Gleichung eine Lösung der homogenen Gleichung $\dot{\gamma} = A\gamma(t)$ sein muß. In der Sprache der linearen Algebra bilden die Lösungen der inhomogenen Gleichung also einen affinen Teilraum des Raums aller Funktionen, dessen Raum von Richtungsvektoren der Lösungsraum der homogenen Gleichung ist.

Beispiel 2.4.2 (Angeregte Schwingungen). Eine Lampe ist mit einer Feder an einer vibrierenden Decke aufgehängt. Sei $h(t)$ die Auslenkung der Decke zur Zeit t und $x(t)$ die Höhe der Lampe zur Zeit t , beide gemessen auf einer gegen den Boden festen Skala, auf der $h = x = 0$ einen Zustand beschreibt, in dem sich die Federkraft, die die Lampe zur Decke zieht, und die Schwerkraft der Lampe die Waage halten. So genügt $x(t)$ einer Differentialgleichung der Gestalt

$$\ddot{x}(t) = -a(x(t) - h(t))$$

wobei a positiv ist und von der Masse der Lampe und der Federkonstante abhängt. Wie im Beweis von 2.2.2 schreiben wir das um zu einem System erster Ordnung

$$\begin{aligned}\dot{\gamma}_0 &= \gamma_1 \\ \dot{\gamma}_1 &= -a\gamma_0 + ah\end{aligned}$$

oder in Matrixschreibweise

$$\dot{\gamma} = \begin{pmatrix} 0 & 1 \\ -a & 0 \end{pmatrix} \gamma + \begin{pmatrix} 0 \\ ah \end{pmatrix}$$

Das charakteristische Polynom unserer Matrix A ist $X^2 + a$, die Eigenwerte ergeben sich zu $\pm i\eta$ für $\eta = \sqrt{a}$ und als zugehörige Eigenvektoren finden wir $(1, \pm i\eta)^\top$. Nehmen wir diese Eigenvektoren als Spalten einer Matrix

$$P = \begin{pmatrix} 1 & 1 \\ -i\eta & i\eta \end{pmatrix}$$

so haben wir offensichtlich $AP = PB$ mit $B = \text{diag}(-i\eta, i\eta)$ einer Diagonalmatrix und $\varphi = P^{-1}\gamma$ erfüllt die Differentialgleichung $\dot{\varphi}(t) = B\varphi(t) + f(t)$ mit

$$f(t) = P^{-1} \begin{pmatrix} 0 \\ ah(t) \end{pmatrix} = \frac{1}{2i\eta} \begin{pmatrix} i\eta & -1 \\ i\eta & 1 \end{pmatrix} \begin{pmatrix} 0 \\ ah(t) \end{pmatrix} = \frac{ah(t)}{2i\eta} \begin{pmatrix} -1 \\ 1 \end{pmatrix}$$

Ich betrachte von nun an diese Differentialgleichung, da es mit übersichtlicher scheint, mit $\exp tB$ anstelle von $\exp tA = P(\exp tB)P^{-1}$ zu hantieren. Nach unseren Überlegungen 2.4.1 lautet die allgemeine Lösung dieser Differentialgleichung

$$\varphi(t) = \exp(tB)g(t) \quad \text{mit} \quad g(t) = \int^t \exp(-\tau B)f(\tau) d\tau$$

Nehmen wir zum Beispiel an, unsere Decke vibriere mit $h(t) = k \sin(\omega t)$ für $\omega > 0$, so ergibt sich die erste Komponente $g_1(t)$ von $g(t)$ zu

$$\begin{aligned}g_1(t) &= -\frac{ka}{2i\eta} \int^t e^{i\tau\eta} \left(\frac{e^{i\tau\omega} - e^{-i\tau\omega}}{2i} \right) d\tau \\ &= \frac{ka}{4\eta} \int^t e^{i\tau(\eta+\omega)} d\tau - \frac{ka}{4\eta} \int^t e^{i\tau(\eta-\omega)} d\tau \\ &= \text{konst} + \frac{ka}{4i\eta(\eta+\omega)} e^{it(\eta+\omega)} - \begin{cases} \frac{ka}{4i\eta(\eta-\omega)} e^{it(\eta-\omega)} & \text{falls } \eta \neq \omega; \\ \frac{ka}{4\eta} t & \text{falls } \eta = \omega. \end{cases}\end{aligned}$$

Ähnlich berechnen wir $g_2(t)$ und erkennen, daß in dem Fall, daß die Eigenfrequenz der Lampe nahe an der Frequenz der Decke ist, d.h. für $|\eta - \omega|$ klein, die

Schwingung sehr groß werden kann und im Fall $\eta = \omega$ die Auslenkung eventuell sogar gegen Unendlich strebt. In der Physik spricht man in diesen Fällen von **Resonanz** bzw. von einer **Resonanzkatastrophe**. Seien die Eigenschwingung unseres Systems $t \mapsto e^{i t \eta}$ und die Anregung $t \mapsto h(t)$ oder gleichbedeutend $t \mapsto f(t)$ periodisch mit derselben Periode p im Sinne der gleich folgenden Definition 3.1.1. Nehmen wir der Einfachkeit halber $p = 2\pi$ an, so finden wir $\eta \in \mathbb{Z}$, und entwickeln wir f in eine Fourierreihe im Sinne von 3.1.3, so zeigt unsere obige Formel $g(t) = \int^t \exp(-\tau B) f(\tau) d\tau$, daß nur die Summanden $c_{\pm\eta} e^{\pm i t \eta}$ für die Resonanz verantwortlich sind in dem Sinne, daß alle anderen Summanden der Fourierreihe nur periodische Beiträge zu $g(t)$ liefern. Später in ?? folgende werden Sie lernen, daß $g_1(t)$ im allgemeinen bis auf eine Konstante auch interpretiert werden kann als der “Wert bei $-\eta$ der Fouriertransformierten des Produkts von f_1 mit der charakteristischen Funktion des Intervalls $[0, t]$ ”.

3 Grundlegendes zu Fourierreihen

3.1 Eindeutigkeit der Fourierreihe

Definition 3.1.1. Sei M eine Menge und $p > 0$ eine positive reelle Zahl. Wir sagen, eine Abbildung $f : \mathbb{R} \rightarrow M$ habe die **Periode** p genau dann, wenn gilt $f(x + p) = f(x) \quad \forall x \in \mathbb{R}$.

Satz 3.1.2 (Entwicklung in eine Fourier-Reihe, reelle Form). Sei $f : \mathbb{R} \rightarrow \mathbb{R}$ eine stetig differenzierbare Funktion mit der Periode 2π . So gibt es eindeutig bestimmte $a_\nu, b_\nu, c \in \mathbb{R}$ derart, daß gilt

$$f(x) = c + \sum_{\nu=1}^{\infty} a_\nu \sin(\nu x) + b_\nu \cos(\nu x)$$

in dem Sinne, daß die Folge der Partialsummen gleichmäßig gegen unsere Funktion f konvergiert.

Satz 3.1.3 (Entwicklung in eine Fourier-Reihe, komplexe Form). Sei $f : \mathbb{R} \rightarrow \mathbb{C}$ eine stetig differenzierbare Funktion mit der Periode 2π . So gibt es eindeutig bestimmte $c_\nu \in \mathbb{C}$ derart, daß im Sinne der gleichmäßigen Konvergenz gilt

$$f(x) = \lim_{n \rightarrow \infty} \sum_{\nu=-n}^{\nu=n} c_\nu e^{i\nu x}$$

3.1.4. Natürlich können wir in der ersten Formulierung 3.1.2 unseres Satzes auch komplexwertige Funktionen erlauben, wenn wir $a_\nu, b_\nu, c \in \mathbb{C}$ zulassen. Die beiden Sätze sind dann äquivalent, da ja nach der Euler'schen Formel gilt

$$\begin{aligned} e^{i\nu x} &= \cos \nu x + i \sin \nu x \\ e^{-i\nu x} &= \cos \nu x - i \sin \nu x \end{aligned}$$

Gegeben eine Darstellung wie in Satz 3.1.3 erhalten wir also eine Darstellung wie in Satz 3.1.2 mit $c = c_0$, $b_\nu = c_\nu + c_{-\nu}$, $a_\nu = i c_\nu - i c_{-\nu}$, und diese Gleichungen sind erfüllt genau dann, wenn gilt

$$c_0 = c, \quad c_\nu = \frac{1}{2}(b_\nu - i a_\nu), \quad \text{und} \quad c_{-\nu} = \frac{1}{2}(b_\nu + i a_\nu).$$

Beweis. Wir zeigen vorerst nur die Eindeutigkeit, der Beweis der Existenz wird in 3.3.7 nachgeholt. Aus 1.5.18 oder auch aus der Euler'schen Formel folgt, daß die Ableitung von $f(x) = e^{i\nu x} = \cos \nu x + i \sin \nu x$ gegeben wird durch $f'(x) = -\nu \sin \nu x + i \nu \cos \nu x = i \nu e^{i\nu x}$. Mit 1.1.12 und der vektorwertigen Variante des

Hauptsatzes der Differential- und Integralrechnung 1.3.7 erhalten wir $\int_0^{2\pi} e^{i\nu x} dx = \frac{1}{i\nu} e^{i\nu x} \Big|_0^{2\pi} = 0$ für $\nu \in \mathbb{Z} \setminus \{0\}$, und für $\nu = 0$ ergibt sich $\int_0^{2\pi} 1 dx = 2\pi$. Wir folgern

$$\int_0^{2\pi} e^{i\mu x} e^{-i\nu x} dx = \begin{cases} 2\pi & \nu = \mu; \\ 0 & \text{sonst.} \end{cases}$$

Indem wir die gleichmäßige Konvergenz mit dem Integral vertauschen und uns überlegen, daß das auch für komplexwertige Funktionen erlaubt ist, erhalten wir

$$\int_0^{2\pi} f(x) e^{-i\nu x} dx = 2\pi c_\nu$$

Das zeigt die Eindeutigkeit der c_ν . Der Beweis der Existenz wird in 3.3.7 nachgeholt. □

Satz 3.1.5 (Variante zur Fourier-Reihe). *Ist $f : \mathbb{R} \rightarrow \mathbb{C}$ eine stetig differenzierbare Funktion mit der Periode 2π , so gibt es eindeutig bestimmte $c_\nu \in \mathbb{C}$ derart, daß bezüglich der Norm der gleichmäßigen Konvergenz auf $\text{Ens}^b(\mathbb{R}, \mathbb{C})$ im Sinne von 10.5.11 gilt*

$$\sum_{\nu \in \mathbb{Z}} c_\nu e^{i\nu x} = f(x)$$

3.1.6. Natürlich müssen hier die c_ν dieselben sein wie in der schwächeren aber einfacher zu formulierenden Version 3.1.3. Der Beweis ihrer Existenz wird in 3.3.7 gegeben. Ich habe etwas gezögert, auf der rechten Seite $f(x)$ zu schreiben, wo doch schlicht die Funktion f gemeint ist, aber auf der linken Seite steht ja auch $e^{i\nu x}$ für die Funktion $x \mapsto e^{i\nu x}$.

Ergänzung 3.1.7. Diese Aussage über die Fourierreihe ist nur der Beginn einer ganzen Theorie: Insbesondere ergibt sich offensichtlich die Frage, welche Abbildungen $\mathbb{Z} \rightarrow \mathbb{C}$ denn nun verschiedenen Klassen von Funktionen entsprechen. Diese Frage erweist sich als recht delikate. In ?? führen wir den Raum der "quadratintegrierbaren Funktionen" ein und zeigen, daß diese genau den Abbildungen $\mathbb{Z} \rightarrow \mathbb{C}$ entsprechen, bei denen die Summe der Betragsquadrate endlich ist.

3.2 Der Satz von Stone-Weierstraß

3.2.1. Unter einem kompakten Raum darf man in diesem Abschnitt je nach Wissensstand einen kompakten metrischen oder allgemeiner einen kompakten topologischen Raum verstehen.

3.2.2. Im folgenden verwenden wir eine Begrifflichkeit, wie sie in ?? ausführlicher eingeführt wird. Ich bespreche hier nur das Nötigste. Gegeben ein Körper k bezeichnet man einen k -Vektorraum A mit einer bilinearen Verknüpfung $A \times A \rightarrow A$ ganz allgemein als eine k -**Algebra**. Ist die Verknüpfung assoziativ, so spricht man folgerichtig von einer **assoziativen Algebra**. Gibt es für unsere Verknüpfung ein neutrales Element, so spricht man von einer **unitären Algebra** und nennt das fragliche Element das **Eins-Element**. Eine Algebra ist also genau dann assoziativ und unitär, wenn die zugrundeliegende Menge mit der Vektorraum-Addition als Addition und der bilinearen Verknüpfung als Multiplikation im Sinne von ?? ein Ring ist. Ich schlage deshalb vor, derartige Algebren **Ringalgebren** zu nennen. Gegeben eine beliebige Menge X ist etwa der k -Vektorraum $\text{Ens}(X, k)$ aller k -wertigen Funktionen auf X mit der punktweisen Multiplikation als Verknüpfung eine Ringalgebra mit der konstanten Funktion Eins als Eins-Element. Wir sagen, eine Teilmenge $A \subset \text{Ens}(X, k)$ **trenne die Punkte von X** genau dann, wenn es für alle $x, y \in X$ mit $x \neq y$ ein $a \in A$ gibt mit $a(x) \neq a(y)$.

3.2.3. Gegeben ein Körper k und eine k -Algebra A versteht man unter einer **Unteralgebra** $B \subset A$ einen unter der Verknüpfung unserer Algebra stabilen Untervektorraum. Gegeben ein Körper k und eine k -Ringalgebra A verstehen wir unter einer **Unterringalgebra** $B \subset A$ einen unter der Verknüpfung unserer Ringalgebra stabilen Untervektorraum, der darüber hinaus das Einselement der Ringalgebra A enthält. In anderen Worten ist eine Unterringalgebra also eine Unteralgebra, die gleichzeitig im Sinne von ?? ein Teilring ist.

3.2.4. Man beachte, daß wir bereits nach 9.5.33 wissen, daß die stetigen reellen Funktionen auf einem topologischen Raum X in der \mathbb{R} -Ringalgebra aller reellen Funktionen eine \mathbb{R} -Unterringalgebra

$$\mathcal{C}(X, \mathbb{R}) \subset \text{Ens}(X, \mathbb{R})$$

bilden. Des weiteren wissen wir nach 9.6.4, daß diese Unterringalgebra stabil ist unter dem Bilden gleichmäßiger Grenzwerte. Gegeben ein kompakter Raum X bezeichne im folgenden $\mathcal{C}(X, \mathbb{R})$ den reellen Vektorraum aller stetigen reellwertigen Funktionen auf X mit seiner Supremumsnorm.

Definition 3.2.5. Eine Teilmenge eines metrischen oder auch eines topologischen Raums heißt **dicht** genau dann, wenn ihr Abschluß der ganze Raum ist.

Satz 3.2.6 (Stone-Weierstraß). *In der Ringalgebra aller stetigen reellwertigen Funktionen auf einem kompakten Raum liegt jede Unterringalgebra, die die Punkte unseres Raums trennt, bereits dicht in Bezug auf die Metrik der gleichmäßigen Konvergenz.*

Ergänzung 3.2.7. Ich erwähne noch eine Variante dieses Satzes, die man oft in der Literatur findet, die jedoch im weiteren Verlauf dieser Vorlesung nicht von Belang ist. Statt von unserer Unteralgebra A zu fordern, daß sie die konstante Funktion Eins enthält, bräuchten wir nur vorauszusetzen, daß es für jedes $x \in X$ ein $a \in A$ gibt mit $a(x) \neq 0$. Man sagt dann, A habe "keine simultane Nullstelle". Unser Beweis funktioniert im wesentlichen auch unter diesen Voraussetzungen, man muß nur Lemma 3.2.10 verfeinern zur Aussage, daß p_ε sogar ohne konstanten Term gefunden werden kann, und muß in Schritt 3 etwas feiner argumentieren.

Korollar 3.2.8 (Approximationssatz von Weierstraß). *Ist $X \subset \mathbb{R}^n$ eine kompakte Teilmenge und $f : X \rightarrow \mathbb{R}$ eine stetige Funktion, so gibt es für jedes $\varepsilon > 0$ eine Polynomfunktion $p \in \mathbb{R}[x_1, \dots, x_n]$ mit*

$$|p(x) - f(x)| < \varepsilon \quad \forall x \in X$$

Beweis. Das folgt sofort aus dem Satz von Stone-Weierstraß 3.2.6. □

Beweis des Satzes von Stone-Weierstraß. Sei X unser kompakter Raum und $A \subset \mathcal{C}(X, \mathbb{R})$ unsere Unterringalgebra. Wir ziehen uns zunächst auf den Fall zurück, daß A in $\mathcal{C}(X, \mathbb{R})$ abgeschlossen ist, und zeigen dazu:

Lemma 3.2.9. *Ist X kompakt und $A \subset \mathcal{C}(X, \mathbb{R})$ eine Unteralgebra, so ist auch der Abschluß \bar{A} von A eine Unteralgebra.*

Beweis. Nach 9.4.11 ist \bar{A} genau die Menge aller stetigen Funktionen $a : X \rightarrow \mathbb{R}$ derart, daß es eine Folge a_n aus A gibt, die gleichmäßig gegen a konvergiert. Sei b ein anderes Element von \bar{A} und b_n eine Folge aus A , die gleichmäßig gegen b konvergiert. Wir behaupten, daß dann auch $a_n + b_n$ gleichmäßig gegen $a + b$ konvergiert und $a_n b_n$ gleichmäßig gegen ab . Die erste Aussage überlassen wir dem Leser. Für die zweite benutze man die Abschätzung

$$\begin{aligned} \|ab - a_n b_n\| &\leq \|a - a_n\| \cdot \|b_n\| + \|a\| \cdot \|b - b_n\| \\ &\leq \varepsilon(\|b\| + 1 + \|a\|) \end{aligned}$$

falls gilt $\|a - a_n\| < \varepsilon$, $\|b - b_n\| < \varepsilon$ und $\varepsilon < 1$. □

Erfüllt also eine Unterringalgebra $A \subset \mathcal{C}(X, \mathbb{R})$ die Bedingungen im Satz, so ist auch ihr Abschluß \bar{A} eine Unterringalgebra und trennt a fortiori die Punkte von X . Um den Satz von Stone-Weierstraß zu beweisen reicht es demnach aus, wenn wir unter der zusätzlichen Annahme A abgeschlossen zeigen $A = \mathcal{C}(X, \mathbb{R})$. Um weiterzukommen, zeigen wir zunächst einen Spezialfall des Approximationssatzes von Weierstraß durch ein direktes Argument.

Lemma 3.2.10. *Für beliebiges $\varepsilon > 0$ gibt es stets ein Polynom $p = p_\varepsilon$ mit $|\sqrt{x} - p(x)| < \varepsilon \quad \forall x \in [0, 1]$.*

Erster Beweis. Sei $\varepsilon > 0$ gegeben. Da \sqrt{x} gleichmäßig stetig ist auf $[0, 2]$, finden wir $\eta \in (0, 1)$ mit

$$|\sqrt{x} - \sqrt{x+\eta}| < \varepsilon/2 \quad \forall x \in [0, 1]$$

Da die Taylorreihe von \sqrt{z} um den Entwicklungspunkt 1 nach 8.1.19 auf dem Intervall $[\eta, 1+\eta]$ gleichmäßig gegen \sqrt{z} konvergiert, finden wir weiter ein Polynom p mit

$$|\sqrt{x+\eta} - p(x)| < \varepsilon/2 \quad \forall x \in [0, 1] \quad \square$$

Zweiter Beweis. Bei der folgenden Alternative muß man etwas mehr denken, aber nichts wissen über die Konvergenz von Taylorreihen. Wir konstruieren induktiv eine Folge von Polynomen durch $p_0(x) = 0$, $p_{n+1}(x) = p_n(x) + (1/2)(x - p_n(x)^2)$ und behaupten, daß diese Folge auf $[0, 1]$ gleichmäßig gegen \sqrt{x} konvergiert. In der Tat gilt ja

$$p_{n+1} = p_n + (\sqrt{x} - p_n)(\sqrt{x} + p_n)/2$$

und dieser Gleichung sehen wir an, daß für $x \in [0, 1]$ gilt

$$p_0 \leq p_1 \leq \dots \leq \sqrt{x}$$

denn es folgt induktiv $(\sqrt{x} - p_n) \geq 0$ und $(\sqrt{x} + p_n)/2 \leq 1$. Andererseits folgt aus unserer Gleichung auch

$$\begin{aligned} (\sqrt{x} - p_{n+1}) &= (\sqrt{x} - p_n)(2 - \sqrt{x} - p_n)/2 \\ &\leq (\sqrt{x} - p_n)(2 - \sqrt{x})/2 \end{aligned}$$

und somit konvergiert unsere Folge p_n auf jedem Intervall $[a, 1]$ mit $0 < a < 1$ gleichmäßig gegen \sqrt{x} . Dann muß mit etwas Nachdenken unsere Folge aber auf ganz $[0, 1]$ gleichmäßig gegen \sqrt{x} konvergieren. \square

Nun zeigen wir den Satz von Stone-Weierstraß 3.2.6 für eine abgeschlossene Unterringalgebra A in fünf Schritten.

1. $a \in A \Rightarrow |a| \in A$. Um das zu zeigen, schreiben wir $a = \lambda b$ mit $\lambda \in (0, \infty)$ und $\|b\| \leq 1$ und erhalten $|a| = \lambda\sqrt{b^2}$. Nach Lemma 3.2.10 gibt es eine Folge p_n von Polynomen, die auf $[0, 1]$ gleichmäßig gegen \sqrt{x} strebt, und dann strebt $\lambda p_n(b^2)$ auf X gleichmäßig gegen $|a|$. Da A eine Unter algebra ist, liegen alle $\lambda p_n(b^2)$ auch in A , und da A abgeschlossen ist unter gleichmäßiger Konvergenz, folgt $|a| \in A$.

2. $a, b \in A \Rightarrow \sup(a, b) \in A, \inf(a, b) \in A$. In der Tat gilt

$$\begin{aligned} \sup(a, b) &= 1/2(a + b + |a - b|) \\ \inf(a, b) &= 1/2(a + b - |a - b|) \end{aligned}$$

3. Für $x \neq y$ zwei verschiedene Punkte aus X und $\alpha, \beta \in \mathbb{R}$ gibt es $a \in A$ mit $a(x) = \alpha$, $a(y) = \beta$. In der Tat betrachte man die \mathbb{R} -lineare Abbildung

$$\begin{aligned} A &\rightarrow \mathbb{R}^2 \\ a &\mapsto (a(x), a(y)) \end{aligned}$$

Da A Punkte trennt, gibt es $a \in A$ mit $a(x) \neq a(y)$. Da die Konstanten zu A gehören, liegt jedoch auch $(1, 1)$ im Bild unserer linearen Abbildung. Damit enthält das Bild unserer linearen Abbildung zwei linear unabhängige Vektoren und ist folglich ganz \mathbb{R}^2 .

4. Für beliebige $f \in \mathcal{C}(X, \mathbb{R})$, $x \in X$ und $\varepsilon > 0$ gibt es $a_x \in A$ mit $a_x(x) = f(x)$ und

$$a_x(y) < f(y) + \varepsilon \quad \forall y \in X$$

In der Tat, für alle $y \in X$ finden wir $a_{x,y} \in A$ mit $a_{x,y}(x) = f(x)$ und $a_{x,y}(y) = f(y)$. Auf einer geeigneten offenen Umgebung U_y von y gilt dann $a_{x,y}(z) < f(z) + \varepsilon \quad \forall z \in U_y$. Da X kompakt ist, gibt es nun $E \subset X$ endlich mit $X = \bigcup_{y \in E} U_y$. Dann nehmen wir $a_x = \inf_{y \in E} a_{x,y}$ und haben unser a_x gefunden.

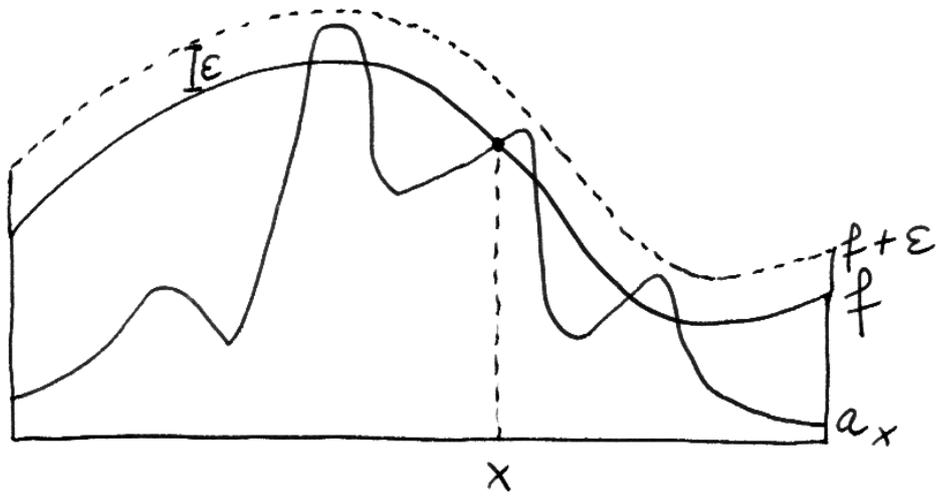
5. Für beliebiges $f \in \mathcal{C}(X, \mathbb{R})$ und $\varepsilon > 0$ gibt es $a \in A$ mit $\|a - f\| < \varepsilon$. Sei in der Tat für jedes $x \in X$ ein a_x wie eben gewählt. Dann hat jeder Punkt $x \in X$ eine offene Umgebung V_x mit $f(z) - \varepsilon < a_x(z) < f(z) + \varepsilon \quad \forall z \in V_x$ wobei die zweite Ungleichung sogar gilt für alle $z \in X$. Da X kompakt ist, gibt es wieder $F \subset X$ endlich mit $X = \bigcup_{x \in F} V_x$. Ist X nicht leer, so nehmen wir nun $a = \sup_{x \in F} a_x$ und haben unser a gefunden. Der Fall $X = \emptyset$ ist eh unproblematisch. \square

Definition 3.2.11. Gegeben ein Kompaktum X betrachten wir nun die \mathbb{C} -Ringalgebra $\mathcal{C}(X)$ aller stetigen komplexwertigen Funktionen auf X mit der Supremumsnorm.

Korollar 3.2.12 (Komplexer Stone-Weierstraß). Sei X kompakt und $B \subset \mathcal{C}(X)$ eine komplexe Unterringalgebra, die die Punkte von X trennt. Ist B zusätzlich stabil unter Konjugation, in Formeln $b \in B \Rightarrow \bar{b} \in B$, so liegt B dicht in $\mathcal{C}(X)$.

Beweis. Man wendet den Satz von Stone-Weierstraß 3.2.6 an auf $A = B \cap \mathcal{C}(X, \mathbb{R})$. Aus $b \in B$ folgt $\operatorname{Re} b$ und $\operatorname{Im} b \in A$, denn es gilt $\operatorname{Re} b = (b + \bar{b})/2$ und $\operatorname{Im} b = (b - \bar{b})/2i$. Also trennt auch unser A die Punkte von X . Für $f \in \mathcal{C}(X)$ finden wir $u, v \in A$ mit $|\operatorname{Re} f(x) - u(x)| < \varepsilon/2$ und $|\operatorname{Im} f(x) - v(x)| < \varepsilon/2$ für alle $x \in X$, setzen $b = u + iv$ und folgern $\|f - b\| < \varepsilon$. \square

Definition 3.2.13. Eine Funktion $f : \mathbb{R} \rightarrow \mathbb{C}$ der Gestalt $t \mapsto \sum_{\nu=-n}^{\nu=n} d_\nu e^{i\nu t}$ mit $d_\nu \in \mathbb{C}$ heißt ein **trigonometrisches Polynom**.



Zum Beweis von 3.2.6, Schritt 4

Satz 3.2.14 (Dichtheit trigonometrischer Polynome). *Gegeben eine stetige Funktion $f : [0, 2\pi] \rightarrow \mathbb{C}$ mit $f(0) = f(2\pi)$ gibt es für beliebiges $\varepsilon > 0$ ein trigonometrisches Polynom $g = g_\varepsilon$ mit*

$$|f(x) - g(x)| < \varepsilon \quad \forall x \in [0, 2\pi]$$

Beweis. Sei $S^1 = \{z \in \mathbb{C} \mid |z| = 1\}$ der Einheitskreis in der komplexen Ebene. Wir betrachten die Abbildung $E : [0, 2\pi] \rightarrow S^1, t \mapsto e^{it}$, die anschaulich gesprochen “unser Intervall zu einer Kreislinie zusammenbiegt”. Das Vorschalten von E liefert eine Bijektion

$$(\circ E) : \mathcal{C}(S^1) \xrightarrow{\sim} \{f \in \mathcal{C}([0, 2\pi]) \mid f(0) = f(2\pi)\}$$

wie man unschwer direkt einsehen und auch formal aus dem anschließenden Lemma 3.2.15 folgern kann. Unter unserer Bijektion entsprechen nun aber die trigonometrischen Polynome auf $[0, 2\pi]$ genau den Funktionen der Form $\sum_{\nu=-n}^n d_\nu z^\nu$ auf der Kreislinie S^1 . Da gilt $\bar{z} = z^{-1}$ für alle $z \in S^1$, dürfen wir den Satz von Stone-Weierstraß für komplexwertige Funktionen 3.2.12 anwenden und folgern, daß das \mathbb{C} -Erzeugnis der z^ν dicht liegt in $\mathcal{C}(S^1)$. Der Satz folgt. \square

Lemma 3.2.15. *Ist $f : X \rightarrow Y$ eine stetige Surjektion von kompakten metrischen Räumen, so ist eine Abbildung $g : Y \rightarrow Z$ in einen weiteren metrischen Raum Z stetig genau dann, wenn $g \circ f$ stetig ist.*

3.2.16. Das gilt auch allgemeiner und mit fast demselben Beweis, wenn wir statt metrischen Räumen topologische Räume betrachten und zusätzlich Y Hausdorff annehmen, vergleiche etwa ??.

Beweis. Das Problem ist hier, die Stetigkeit von g aus der Stetigkeit von $g \circ f$ zu folgern. Da f surjektiv ist, gilt für jede Teilmenge $A \subset Z$ offensichtlich

$$g^{-1}(A) = f((g \circ f)^{-1}(A))$$

Ist A abgeschlossen in Z , so ist $(g \circ f)^{-1}(A)$ abgeschlossen in X wegen der Stetigkeit von $g \circ f$, also kompakt nach 9.7.3. Dann ist $f((g \circ f)^{-1}(A))$ kompakt nach 9.7.10 als Bild einer kompakten Menge unter einer stetigen Abbildung, mithin abgeschlossen nach 9.7.3. Zusammenfassend haben wir also gezeigt, daß das Urbild $g^{-1}(A)$ einer abgeschlossenen Teilmenge $A \subset Z$ abgeschlossen ist in Y , und daraus folgt mit 9.5.30 die Stetigkeit von g . \square

3.3 Konvergenz der Fourierreihe

3.3.1. Wir verwenden im folgenden die Begrifflichkeit der Skalarprodukte, wie sie etwa in ?? eingeführt wird.

Definition 3.3.2. Wir versehen den komplexen Vektorraum $V = \mathcal{C}([0, 2\pi])$ aller stetigen Funktionen $f : [0, 2\pi] \rightarrow \mathbb{C}$ mit dem Skalarprodukt

$$\langle f, g \rangle = \frac{1}{2\pi} \int_0^{2\pi} \bar{f}g$$

Die zugehörige Norm notiert man in diesem Fall mit $\sqrt{\langle f, f \rangle} = \|f\|_2$.

3.3.3. Unsere Formeln aus dem Beweis von 3.1.3 besagen genau, daß die $e^{i\nu x}$ mit $\nu \in \mathbb{Z}$ in diesem Raum ein Orthonormalsystem im Sinne von ?? bilden,

$$\langle e^{i\nu x}, e^{i\mu x} \rangle = \begin{cases} 1 & \nu = \mu; \\ 0 & \text{sonst.} \end{cases}$$

Die Fourier-Koeffizienten schreiben sich nun kürzer $c_\nu = \langle e^{i\nu x}, f \rangle$ und wir erhalten eine Abbildung

$$\begin{array}{ccc} \mathcal{C}([0, 2\pi]) & \rightarrow & \text{Ens}(\mathbb{Z}, \mathbb{C}) \\ f & \mapsto & f^\wedge \end{array}$$

indem wir allgemeiner jeder stetigen Funktion $f : [0, 2\pi] \rightarrow \mathbb{C}$ die Familie ihrer “Fourierkoeffizienten” zuordnen, in Formeln $f^\wedge(\nu) = \langle e^{i\nu x}, f \rangle$. In ?? werden wir diese Abbildung erweitern zu einer Bijektion zwischen geeigneten Räumen quadratintegrierbarer Funktionen.

Satz 3.3.4 (über die quadratische Konvergenz der Fourierreihe). Ist $f : [0, 2\pi] \rightarrow \mathbb{C}$ eine stetige Funktion und sind $c_\nu = \langle e^{i\nu x}, f \rangle$ ihre Fourierkoeffizienten, so gilt

$$\lim_{n \rightarrow \infty} \left\| f - \sum_{\nu=-n}^n c_\nu e^{i\nu x} \right\|_2 = 0$$

3.3.5. Gegeben eine Folge f_n stetiger Funktionen von einem kompakten Intervall $[a, b]$ nach \mathbb{C} und eine weitere stetige Funktion f sagt man, die Folge der f_n **konvergiere im quadratischen Mittel gegen f** genau dann, wenn gilt

$$\lim_{n \rightarrow \infty} \int_a^b |f - f_n|^2 = 0$$

Unser Satz sagt also in dieser Terminologie, daß die Fourierreihe einer stetigen Funktion im quadratischen Mittel gegen besagte Funktion konvergiert. Wir werden den vorhergehenden Satz in ?? verallgemeinern von stetigen auf alle “quadratintegrierbaren” Funktionen. Mit der Terminologie aus 10.5.11 können wir im normierten Vektorraum $\mathcal{C}([0, 2\pi])$ aus 3.3.2 die Aussage des Satzes auch schreiben in der Gestalt

$$f = \sum_{\nu \in \mathbb{Z}} c_\nu e^{i\nu x}$$

Beweis. Für alle n können wir f nach ?? zerlegen in seine Projektion auf den von allen $e^{i\nu x}$ mit $-n \leq \nu \leq n$ aufgespannten Teilraum von $\mathcal{C}([0, 2\pi])$ und einen auf diesem Teilraum senkrechten Anteil,

$$f = \sum_{\nu=-n}^n c_\nu e^{i\nu x} + \left(f - \sum_{\nu=-n}^n c_\nu e^{i\nu x} \right)$$

Wir nehmen nun zunächst zusätzlich $f(0) = f(2\pi)$ an. Für alle $\varepsilon > 0$ finden wir dann nach dem Korollar 3.2.14 des Satzes von Stone-Weierstraß ein trigonometrisches Polynom $g = \sum_{\nu=-n}^n d_\nu e^{i\nu x}$ mit

$$|f(x) - g(x)| < \varepsilon \quad \forall x$$

Es folgt sofort $\|f - g\|_2 < \varepsilon$. Da in einem euklidischen Vektorraum nach ?? die orthogonale Projektion eines Vektors auf einen endlichdimensionalen Teilraum stets die bestmögliche Approximation durch Vektoren dieses Teilraums ist, folgt für alle $m \geq n$ erst recht

$$\left\| f - \sum_{\nu=-m}^m c_\nu e^{i\nu x} \right\|_2 < \varepsilon$$

Das zeigt die Behauptung im Fall $f(0) = f(2\pi)$. Im Fall $f(0) \neq f(2\pi)$ müssen wir noch eine extra Verrenkung machen und zunächst eine stetige Funktion \tilde{f} finden mit $\tilde{f}(0) = \tilde{f}(2\pi)$ sowie $\|\tilde{f} - f\|_2 < \varepsilon$. Dann gibt es wieder ein trigonometrisches Polynom g mit $\|\tilde{f} - g\|_2 < \varepsilon$, also $\|f - g\|_2 < 2\varepsilon$, und der Beweis kann wie zuvor zu Ende geführt werden. \square

Korollar 3.3.6. Sei $f : [0, 2\pi] \rightarrow \mathbb{C}$ stetig und seien $c_\nu = \langle e^{i\nu x}, f \rangle$ die Fourierkoeffizienten. So gilt $\|f\|_2^2 = \sum_{\nu=-\infty}^{\infty} |c_\nu|^2$.

Beweis. Die Differenz $\|f\|_2^2 - \sum_{\nu=-n}^n |c_\nu|^2$ ist das Quadrat des Ausdrucks, von dem wir gerade gezeigt haben, daß er gegen Null strebt. \square

Satz 3.3.7 (über die gleichmäßige Konvergenz der Fourierreihe). Ist $f : \mathbb{R} \rightarrow \mathbb{C}$ stetig differenzierbar mit der Periode 2π , so konvergiert die Fourierreihe von f gleichmäßig gegen f .

3.3.8. Der Satz gilt mit fast demselben Beweis auch noch, wenn unsere Funktion nur stückweise stetig differenzierbar ist in dem Sinne, daß es Punkte $0 = a_0 < a_1 < \dots < a_k = 2\pi$ gibt derart, daß die Einschränkung von f auf jedes der Intervalle $[a_i, a_{i+1}]$ stetig differenzierbar ist. Die Details mag der Leser zur Übung ausarbeiten.

Beweis. Die Fourier-Koeffizienten $c_\nu = \langle e^{i\nu x}, f \rangle$ von f ergeben sich für $\nu \neq 0$ aus den Fourier-Koeffizienten $c'_\nu = \langle e^{i\nu x}, f' \rangle$ von f' durch partielles Integrieren zu

$$c_\nu = \frac{1}{2\pi} \int_0^{2\pi} f(x) e^{-i\nu x} dx = -\frac{1}{2\pi} \int_0^{2\pi} \frac{f'(x) e^{-i\nu x}}{-i\nu} dx = \frac{-i c'_\nu}{\nu}$$

Jetzt gilt jedoch $2|\alpha\beta| \leq (|\alpha|^2 + |\beta|^2)$ für beliebige $\alpha, \beta \in \mathbb{C}$ und es folgt

$$\sum_\nu |c_\nu| \leq |c_0| + \sum_{\nu \neq 0} \left(\frac{1}{\nu^2} + |c'_\nu|^2 \right) < \infty$$

Also konvergiert die Funktionenfolge $\sum_{\nu=-n}^n c_\nu e^{i\nu x}$ gleichmäßig gegen eine stetige Funktion g . Natürlich konvergiert unsere Reihe erst recht auf jedem kompakten Intervall im quadratischen Mittel gegen diese Funktion g , aus 3.3.4 folgt also $g = f$ und wir sind fertig. \square

3.3.9. Fassen wir eine stetig differenzierbare 2π -periodische Funktion f auf als eine Funktion auf dem Einheitskreis und nehmen sie reellwertig an, so gilt für ihre Fourier-Koeffizienten offensichtlich $c_{-\nu} = \bar{c}_\nu$. Sie können zum Beispiel in ?? lernen, warum die Formel

$$P(z) = c_0 + \sum_{\nu=1}^{\infty} c_\nu z^\nu + c_{-\nu} \bar{z}^\nu$$

dann die eindeutig bestimmte “stabile Wärmeverteilung mit Randverteilung f auf der Einheitskreisscheibe” beschreibt. In diesem Zusammenhang hat Fourier, von dem erzählt wird, daß er häufig fröstelte, ursprünglich die heute nach ihm benannten Reihenentwicklungen gefunden und in seinem Werk “Théorie analytique de la chaleur” veröffentlicht.

4 Ableitungen in mehreren Veränderlichen

4.1 Partielle Ableitungen und Gradient

Definition 4.1.1. Sei $A \subseteq \mathbb{R}^n$ eine offene Teilmenge, $f : A \rightarrow \mathbb{R}$ eine Funktion und $p = (p_1, \dots, p_n) \in A$ ein Punkt. Wir nennen f **partiell differenzierbar bei p nach der i -ten Variablen** genau dann, wenn die Funktion $x \mapsto f(p_1, \dots, p_{i-1}, x, p_{i+1}, \dots, p_n)$ differenzierbar ist bei $x = p_i$. Die Ableitung dieser Funktion heißt dann die **i -te partielle Ableitung** von f und wird notiert

$$(D_i f)(p) := \frac{\partial f}{\partial x_i}(p) := \lim_{h \rightarrow 0} \frac{f(p_1, \dots, p_i + h, \dots, p_n) - f(p_1, \dots, p_i, \dots, p_n)}{h}$$

4.1.2. Diese partiellen Ableitungen sind, soweit sie existieren, wieder reellwertige Funktionen auf A . Um $\frac{\partial f}{\partial x_i}$ zu berechnen muß man sich nur vorstellen, alle x_j mit $j \neq i$ seien Konstanten. Zum Beispiel berechnen wir die partiellen Ableitungen von $f(x, y) = x \sin(xy)$ und erhalten

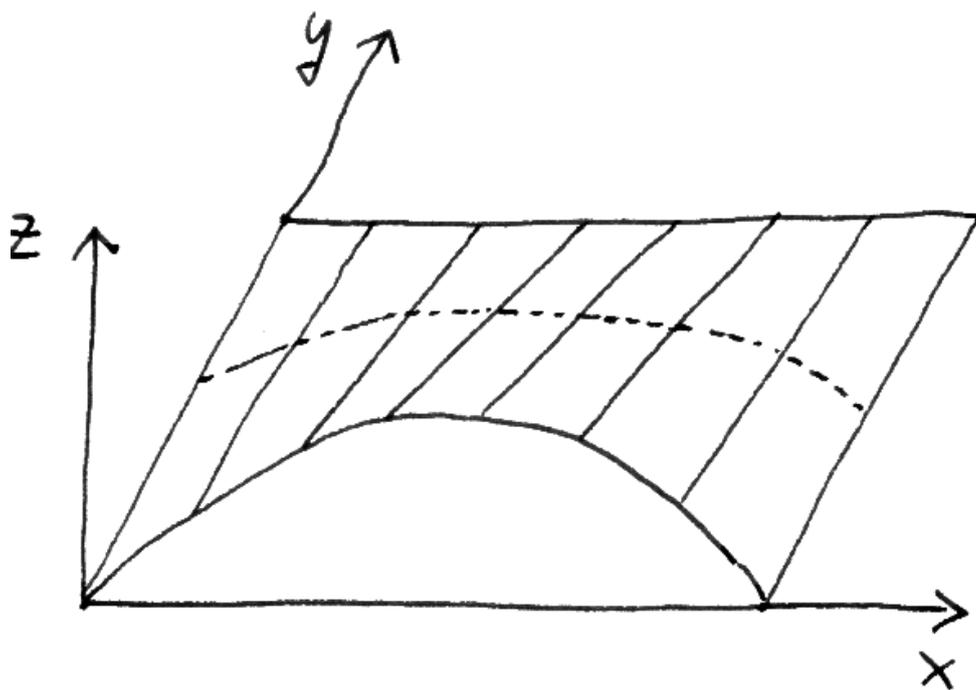
$$\begin{aligned} \frac{\partial f}{\partial x} &= \sin(xy) + xy \cos(xy) \\ \frac{\partial f}{\partial y} &= x^2 \cos(xy) \end{aligned}$$

Dieses Beispiel zeigt auch die Vorteile der Notation $\frac{\partial}{\partial x}$ gegenüber der etwas exakteren Notation D_i , bei der man stets eine Reihenfolge der Variablen festlegen muß und schneller in Indizes ertrinkt. Im Fall, daß weder die Variablen noch die Funktion selbst bereits Indizes tragen, benutzt man auch die sehr konzise Schreibweise

$$\frac{\partial f}{\partial x} = f_x$$

Ergänzung 4.1.3. Allgemeiner definiert man ebenso auch partielle Ableitungen für Abbildungen f von einer offenen Teilmenge $A \subseteq \mathbb{R}^n$ in einen beliebigen normierten Vektorraum. Diese partiellen Ableitungen sind dann, soweit sie existieren, Abbildungen $\frac{\partial f}{\partial x_i}$ von A in denselben normierten Vektorraum.

4.1.4. Ich will an einem Beispiel erläutern, aus welchem Grund man im Fall mehrerer Veränderlichen unsere bisherige Notation $\frac{d}{dx}$ zu $\frac{\partial}{\partial x}$ abändert. Denken wir uns einen Wanderer bei einer Wanderung durch die Alpen, bei der schlechtes Wetter aufkommt. Der Luftdruck $D = D(t, h)$ hängt dann sowohl von der Zeit als auch von der Höhe ab. Macht unser Wanderer zum Zeitpunkt $t = t_0$ in der Höhe $h = h_0$ eine Pause, so ändert sich der Luftdruck, den sein Barometer mißt, mit der Rate $\frac{\partial D}{\partial t}(t_0, h_0)$. Geht er jedoch zum Zeitpunkt $t = t_0$ bergab oder bergauf und gibt die Funktion $h(t)$ seine Höhe zum Zeitpunkt t an, so ändert sich der Luftdruck, den sein Barometer mißt, mit der Rate $\left. \frac{d}{dt} \right|_{t=t_0} (D(t, h(t)))$. Wir werden zeigen,



Veranschaulichen wir uns eine reellwertige Funktion von zwei reellen Veränderlichen durch ihren Graphen, eine "hügelige Landschaft", im Bild etwa $f(x, y) = \frac{1}{2} \sin((1 - y)x)$, so mag man sich die partielle Ableitung $\frac{\partial f}{\partial x}$ an einer Stelle denken als die Steigung an besagter Stelle einer Straße, die besagte hügeliger Landschaft in Richtung der x -Achse durchquert. Zum Beispiel wäre diese partielle Ableitung an der Stelle $(0, 1/2)$ in unserem Fall die Steigung der gestrichelt eingezeichneten Straße an ihrem Beginn auf der y -Achse.

daß sich diese Rate auch ausdrücken läßt in der Gestalt $\left. \frac{d}{dt} \right|_{t=t_0} (D(t, h(t))) = \frac{\partial D}{\partial t}(t_0, h(t_0)) + h'(t_0) \frac{\partial D}{\partial h}(t_0, h(t_0))$. In Kurzschreibweise gilt also

$$\frac{dD}{dt} = \frac{\partial D}{\partial t} + \frac{dh}{dt} \frac{\partial D}{\partial h}$$

Der Zweck der Variation unserer Notation liegt nun eben darin, daß mit ihr solche Verkürzungen noch verständlich bleiben. Um die behauptete Formel zu beweisen, führen wir den Begriff des Differential ein, studieren seinen Zusammenhang mit den partiellen Ableitungen und erhalten unsere Formel als Korollar 4.5.4 der Kettenregel für Differentiale.

Definition 4.1.5. Ist $A \subseteq \mathbb{R}^n$ eine offene Teilmenge und $f : A \rightarrow \mathbb{R}$ eine auf ganz A nach jeder der n Variablen partiell differenzierbare Funktion, so definieren wir den **Gradienten** von f als die Abbildung

$$\begin{aligned} \text{grad } f : A &\rightarrow \mathbb{R}^n \\ x &\mapsto \left(\frac{\partial f}{\partial x_1}(x), \dots, \frac{\partial f}{\partial x_n}(x) \right)^\top \end{aligned}$$

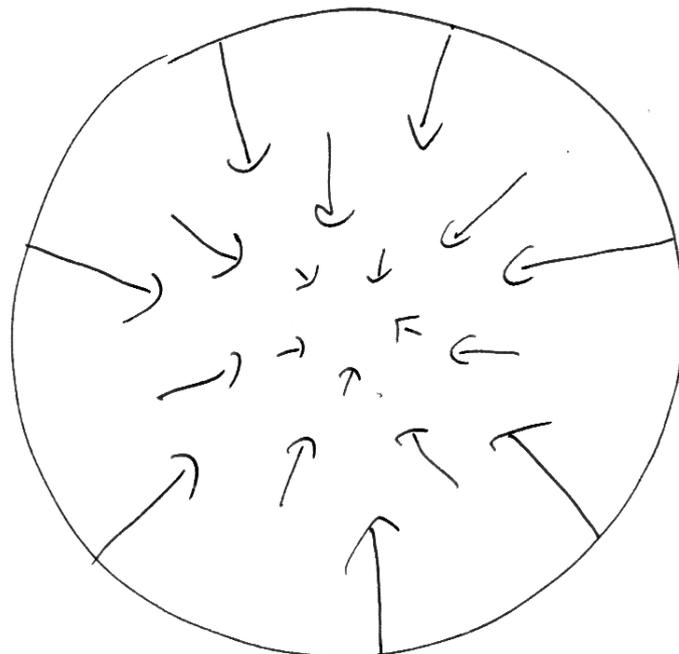
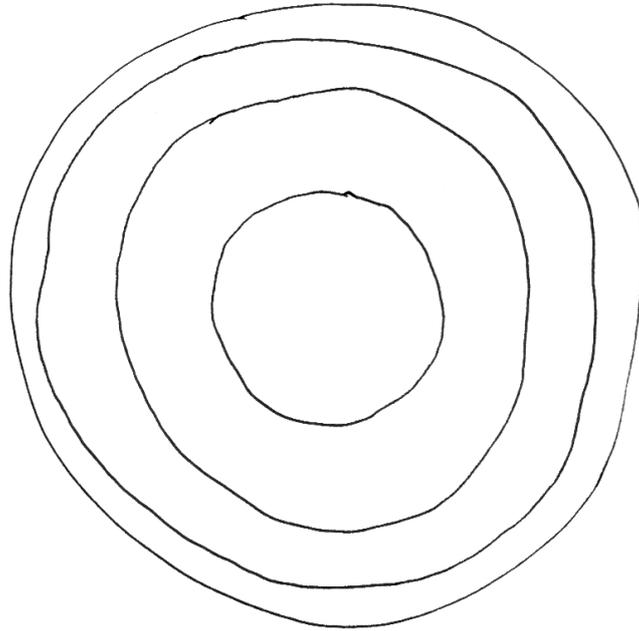
4.1.6. Man beachte, daß in dieser Definition das Symbol x für ein Element des \mathbb{R}^n steht und nicht wie zuvor für eine reelle Zahl. Ich stelle mir $\text{grad } f$ meist vor als ein Vektorfeld, das also jedem Punkt aus A einen Vektor aus dem \mathbb{R}^n zuordnet. Das ist auch der Grund dafür, daß ich in obiger Definition den Zeilenvektor in einen Spaltenvektor transponiert habe. Denken wir uns im Fall $n = 2$ den Graphen von f als eine Hügellandschaft, so zeigt $\text{grad } f$ stets in die Richtung, in der es am steilsten den Berg hinaufgeht, und ist desto länger, je steiler es hinaufgeht. Diese Anschauung wird durch Bemerkung 4.5.6 formalisiert.

4.1.7. Der Begriff des Gradienten ist nur für reellwertige Funktionen auf dem \mathbb{R}^n sinnvoll. Bereits reellwertigen Funktionen auf abstrakten endlichdimensionalen reellen Vektorräumen kann nicht mehr sinnvoll ein Gradient in Gestalt eines Vektorfeldes werden. Ich vermeide deshalb im folgenden nach Möglichkeit den Gradientenbegriff und arbeite stattdessen mit den sogenannten ‘‘Differentialen’’, die in sehr viel größerer Allgemeinheit sinnvoll sind, und deren Beziehung zum Gradienten in 4.2.6 und 6.2.6 besprochen werden wird.

Übung 4.1.8. Sei $R(x, y) = \sum_{i,j} c_{ij} x^i y^j$ ein Polynom in zwei Variablen mit reellen Koeffizienten $c_{ij} \in \mathbb{R}$. Gibt es eine nichtleere offene Teilmenge $A \subseteq \mathbb{R}^2$ derart, daß gilt $R(p) = 0 \forall p \in A$, so ist R das Nullpolynom, in Formeln $c_{ij} = 0 \forall i, j$.

4.2 Das Differential

4.2.1. Wir werden es vorerst nur mit endlichdimensionalen normierten Vektorräumen zu tun haben. Ich arbeite dennoch hier und im folgenden mit beliebigen



Einige Niveaulinien und das Gradientenfeld eines Hügels, hier möglicherweise
 der Funktion $\frac{1}{2}\sqrt{1-x^2-y^2}$ auf der Kreisscheibe $x^2+y^2 < \frac{1}{2}$.

normierten Vektorräumen, weil das zum Ersten in keiner Weise schwieriger ist, weil es zum Zweiten einen größeren Abstand zum unübersichtlichen Gestrüpp der Koordinaten schafft, und weil es zum Dritten bei unserer Behandlung von Differentialgleichungen 8.5.2 auch für unendlichdimensionale Räume gebraucht wird. Noch natürlicher wäre es, mit Abbildungen zwischen affinen Räumen X, Y im Sinne von 9.8.2 zu arbeiten, deren Richtungsräume \vec{X}, \vec{Y} jeweils mit einer Norm versehen sind, und das Differential als eine lineare Abbildung zwischen diesen Richtungsräumen zu erklären. Ich werde diesen Gesichtspunkt im folgenden nur andeuten, indem ich schlicht von “normierten Räumen” rede und Symbole ohne Pfeil schreibe, wenn man ebensogut affine Räume mit normiertem Richtungsraum betrachten könnte, und Symbole mit Pfeil schreibe, wenn der zugehörige Richtungsraum gemeint ist.

Definition 4.2.2. Seien X, Y normierte reelle Räume, $A \subseteq X$ eine offene Teilmenge, $f : A \rightarrow Y$ eine Abbildung und $p \in A$ ein Punkt. Genau dann heißt die Abbildung f **differenzierbar** oder genauer **Fréchet-differenzierbar bei p** , wenn es eine stetige lineare Abbildung $L : \vec{X} \rightarrow \vec{Y}$ gibt derart, daß gilt

$$f(p+h) = f(p) + Lh + \|h\|\varepsilon(h)$$

für eine \vec{Y} -wertige Abbildung ε mit $\lim_{h \rightarrow 0} \varepsilon(h) = 0$.

4.2.3. Hier ist implizit mit zu verstehen, daß die Abbildung ε definiert sein soll auf der Menge aller $h \in \vec{X}$ mit $p+h \in A$. Weil wir in der Definition fordern, daß L stetig sein soll, ist jede bei p differenzierbare Abbildung bei p auch stetig. Die Stetigkeit von L ist im übrigen auch notwendig für die Gültigkeit der Kettenregel 4.3.1. Die lineare Abbildung L ist eindeutig bestimmt wenn sie existiert, da man für beliebiges $v \in \vec{X}$ durch Einsetzen von $h = tv$ zur Formel

$$Lv = \lim_{t \rightarrow 0} \frac{f(p+tv) - f(p)}{t}$$

gelangt. Diesen Grenzwert in \vec{Y} hinwiederum nennt man, wann immer er existiert, die **Richtungsableitung von f bei p in Richtung v** und kürzt ihn ab mit $(D_v f)(p)$. Das D steht hier für englisch **directional derivative**. Anschaulich mißt diese Richtungsableitung im Fall $Y = \mathbb{R}$, wie schnell unsere Funktion wächst bzw. abnimmt, wenn wir von p aus in der Richtung v gehen. Es gilt allerdings zu beachten, daß unsere Richtungsableitung keineswegs nur von der Richtung des Vektors v abhängt, sondern durchaus auch von seiner Länge. Die lineare Abbildung L selbst heißt das **Differential von f bei p** und wird bezeichnet mit

$$L = d_p f$$

Es mag dem Verständnis helfen, statt h das Symbol δp zu verwenden. Dann liest sich unsere Definition des Differentials

$$f(p + \delta p) = f(p) + (d_p f)(\delta p) + \|\delta p\|_\varepsilon(\delta p)$$

Beispiel 4.2.4. Für $f : \mathbb{R}^n \rightarrow \mathbb{R}$ oder allgemeiner $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ differenzierbar existieren insbesondere unsere partiellen Ableitungen und sind gerade die Richtungsableitungen in Richtung der Einheitsvektoren e_i , in Formeln

$$\frac{\partial f}{\partial x_i}(p) = (D_{e_i} f)(p)$$

Dasselbe gilt auch, wenn f nur auf einer offenen Teilmenge von \mathbb{R}^n definiert ist. Umgekehrt werden wir in 4.5.1 zeigen, wie man aus der Existenz *und Stetigkeit* der partiellen Ableitungen die Existenz des Differentials folgern kann.

4.2.5. Wenn das Differential $d_p f$ existiert, so existieren insbesondere auch alle Richtungsableitungen und es gilt

$$(D_v f)(p) = (d_p f)(v)$$

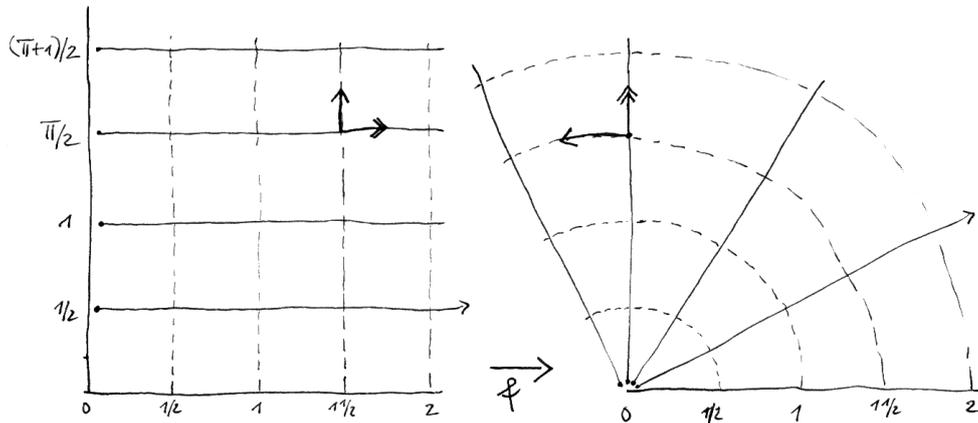
für alle $v \in \vec{X}$. Nennen wir eine Abbildung einfach nur **differenzierbar**, so ist die Differenzierbarkeit an jeder Stelle gemeint. Die Ableitung oder genauer das Differential einer differenzierbaren Abbildung $f : U \rightarrow \mathbb{R}^m$ für $U \subseteq \mathbb{R}^n$ ist aber nun eine Abbildung $U \rightarrow \text{Hom}(\mathbb{R}^n, \mathbb{R}^m)$, $p \mapsto d_p f$. Für jede fest vorgegebene Stelle $p \in U$ ist weiter die darstellende Matrix des Differentials $d_p f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ nach ?? die Matrix mit den Spaltenvektoren $(d_p f)(\vec{e}_i) = \frac{\partial f}{\partial x_i}(p)$. Hat unsere Abbildung also die Gestalt $f = (f_1, \dots, f_m)$ mit Funktionen $f_j : U \rightarrow \mathbb{R}$, so hat die darstellende Matrix der linearen Abbildung $d_p f$ die Gestalt

$$[d_p f] = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(p) & \cdots & \frac{\partial f_1}{\partial x_n}(p) \\ \vdots & & \vdots \\ \frac{\partial f_m}{\partial x_1}(p) & \cdots & \frac{\partial f_m}{\partial x_n}(p) \end{pmatrix}$$

Diese Matrix heißt die **Jacobi-Matrix** unserer Abbildung. Wir denken uns in diesem Zusammenhang Vektoren stets als Spaltenvektoren und hätten etwas pedantisch wohl auch besser $f = (f_1, \dots, f_m)^\top$ schreiben sollen, um das nocheinmal zu betonen. Für die Jacobi-Matrix findet man häufig auch die Notation

$$\frac{\partial(f_1, \dots, f_m)}{\partial(x_1, \dots, x_n)}$$

In der Literatur werden unsere differenzierbaren Abbildungen vielfach **total differenzierbar** genannt, um sie abzugrenzen von den **partiell differenzierbaren**



Dies Bild soll die Bedeutung des Differential in der Anschauung einer Abbildung “als Abbildung” verdeutlichen. Wir betrachten die Polarkoordinatenabbildung

$$f : \mathbb{R}_{>0} \times (0, 2\pi) \rightarrow \mathbb{R}^2, \quad (r, \vartheta) \mapsto (r \cos \vartheta, r \sin \vartheta)$$

wobei verwirrender Weise die Klammern $(,)$ einmal ein offenes Intervall und dann wieder Elemente kartesischer Produkte andeuten. Ihr Differential wird beschrieben durch die Jacobi-Matrix

$$[df] = \begin{pmatrix} \cos \vartheta & -r \sin \vartheta \\ \sin \vartheta & r \cos \vartheta \end{pmatrix}$$

Insbesondere wird das Differential an der Stelle $a = (1\frac{1}{2}, \frac{\pi}{2})$ beschrieben durch die Matrix

$$[d_p f] = \begin{pmatrix} 0 & -1\frac{1}{2} \\ 1 & 0 \end{pmatrix}$$

Die Pfeile im Bild sollen zeigen, daß das in der Tat diejenige lineare Abbildung L ist, für die für kleines h die Abbildung $p + h \mapsto f(p) + Lh$ unsere Abbildung $p + h \mapsto f(p + h)$ besonders gut approximiert.

Abbildungen, bei denen nur die Existenz aller partiellen Ableitungen gefordert wird. Unser Differential heißt in manchen Quellen auch das **totale Differential**. Wenn Ihnen die Identifikation von Matrizen mit linearen Abbildungen $\mathbb{R}^m \rightarrow \mathbb{R}^n$ aus ?? und ?? einmal richtig in Fleisch und Blut übergegangen ist, werden Sie sich auch nicht daran stören, wenn wir einmal mit $d_p f$ sowohl das Differential als auch die Jacobi-Matrix bezeichnen sollten, was leicht einmal vorkommen kann.

4.2.6. Im Spezialfall einer reellwertigen Abbildung $f : U \rightarrow \mathbb{R}$, die auf einer offenen Teilmenge $U \subseteq \mathbb{R}^n$ definiert ist, ist die Jacobi-Matrix also eine Zeilenmatrix

$$[d_p f] = \left(\frac{\partial f}{\partial x_1}(p), \dots, \frac{\partial f}{\partial x_n}(p) \right)$$

und kann auch als die “Transponierte des Gradienten” verstanden werden, der ja als Vektor in unseren Konventionen a priori als eine Spaltenmatrix aufzufassen ist. Im Fall des \mathbb{R}^n kann eben der Dualraum in natürlicher Weise mit dem Raum selber identifiziert werden, so daß wir die Linearform $d_p f$ auch als einen Vektor auffassen können und die Zuordnung $p \mapsto d_p f$ auch als das Vektorfeld, das wir bereits in 4.1.6 als Gradientenfeld kennengelernt haben. Im Fall eines abstrakten endlichdimensionalen reellen Vektorraums V und einer reellwertigen Abbildung $f : U \rightarrow \mathbb{R}$, die auf einer offenen Teilmenge $U \subseteq V$ definiert ist, geht das jedoch nicht mehr. Die Zuordnung $p \mapsto d_p f$ ist dann zwar, unter den entsprechenden Differenzierbarkeitsannahmen, immer noch eine sinnvoll definierte Abbildung $U \rightarrow V^*$ von U in den Dualraum von V , eben das Differential unserer Abbildung f . Diese Abbildung kann aber nicht mehr in natürlicher Weise mit einem Vektorfeld alias einer Abbildung $U \rightarrow V$ identifiziert werden. Abbildungen $U \rightarrow V^*$ für $U \subseteq V$ heißen im übrigen “Kovektorfelder auf U ” und werden in 6.1.4 folgende noch ausführlich besprochen werden.

Ergänzung 4.2.7. Unser Differenzierbarkeitsbegriff ändert sich nicht, wenn wir die Normen auf \vec{X} und \vec{Y} durch äquivalente Normen ersetzen. Das kann man sich ohne große Mühe direkt überlegen, es wird aber auch formal aus der Kettenregel 4.3.1 folgen. Sind insbesondere X und Y endlichdimensional, so ist unser Differenzierbarkeitsbegriff unabhängig von der Wahl der Normen.

4.2.8. Die der Definition zugrundeliegende Eindeutigkeit des Differentials von f bei p gilt nach 9.9.12 auch dann noch, wenn wir vom Definitionsbereich A von f nur fordern, daß es eine nichtleere offene Teilmenge $C \subseteq \vec{X}$ gibt mit $p + [0, 1]C \subset A$, wenn es also anschaulich gesprochen ein kleines Kegelchen mit Spitze in p gibt, das ganz in A liegt. Gibt es solch ein Kegelchen für alle $p \in A$, so nennen wir A eine **halboffene** Teilmenge von X . Dieser Begriff kommt in der Literatur sonst nicht vor, er scheint mir jedoch nützlich, da er hilft, Verkrampfungen bei der Definition der Differenzierbarkeit auf abgeschlossenen Halbräumen und dergleichen zu vermeiden. Diese Allgemeinheit hinwiederum benötigen wir bei der

Verallgemeinerung des Hauptsatzes der Differential- und Integralrechnung auf höhere Dimensionen in 10.8.1.

Ergänzung 4.2.9. Ist $A \subset \mathbb{R}^n$ eine halboffene Teilmenge und $f : A \rightarrow \mathbb{R}$ differenzierbar, so setzen wir

$$\frac{\partial f}{\partial x_i}(p) = (d_p f)(e_i)$$

auch dann, wenn gewisse partielle Ableitungen bei p recht eigentlich gar nicht gebildet werden können, etwa weil die entsprechenden Geraden durch p die Menge A nur in p treffen.

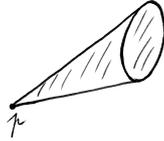
Beispiel 4.2.10. Das Differential einer stetigen affinen Abbildung ist an jeder Stelle ihr linearer Anteil. Ist also in Formeln $f : X \rightarrow Y$ gegeben durch $f(v + h) = Lh + w$ mit einer stetigen linearen Abbildung $L : \vec{X} \rightarrow \vec{Y}$ und $v \in X, w \in Y$ fest, so haben wir $d_p f = L$ für alle $p \in X$. Insbesondere verschwindet das Differential einer konstanten Abbildung an jedem Punkt und das Differential einer stetigen linearen Abbildung ist an jedem Punkt die lineare Abbildung selbst.

Beispiel 4.2.11. Im allgemeinen kann man sich $d_p f$ vorstellen als “den linearen Anteil der affinen Abbildung, die unsere Funktion in der Nähe der vorgegebenen Stelle bestmöglich approximiert” oder in anderen Worten als “diejenige lineare Abbildung L , für die $x \mapsto f(p) + L(x - p)$ unsere Funktion $x \mapsto f(x)$ in der Nähe von p am besten approximiert”. Veranschaulichen wir uns zum Beispiel eine Funktion $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ durch ihren Graphen, eine hügelige Landschaft, so ist die “Tangentialebene” an unsere hügelige Landschaft im Punkt $(p, f(p))$ im verschobenen Koordinatensystem mit Ursprung $(p, f(p))$ gerade der Graph des Differentials $d_p f : \mathbb{R}^2 \rightarrow \mathbb{R}$.

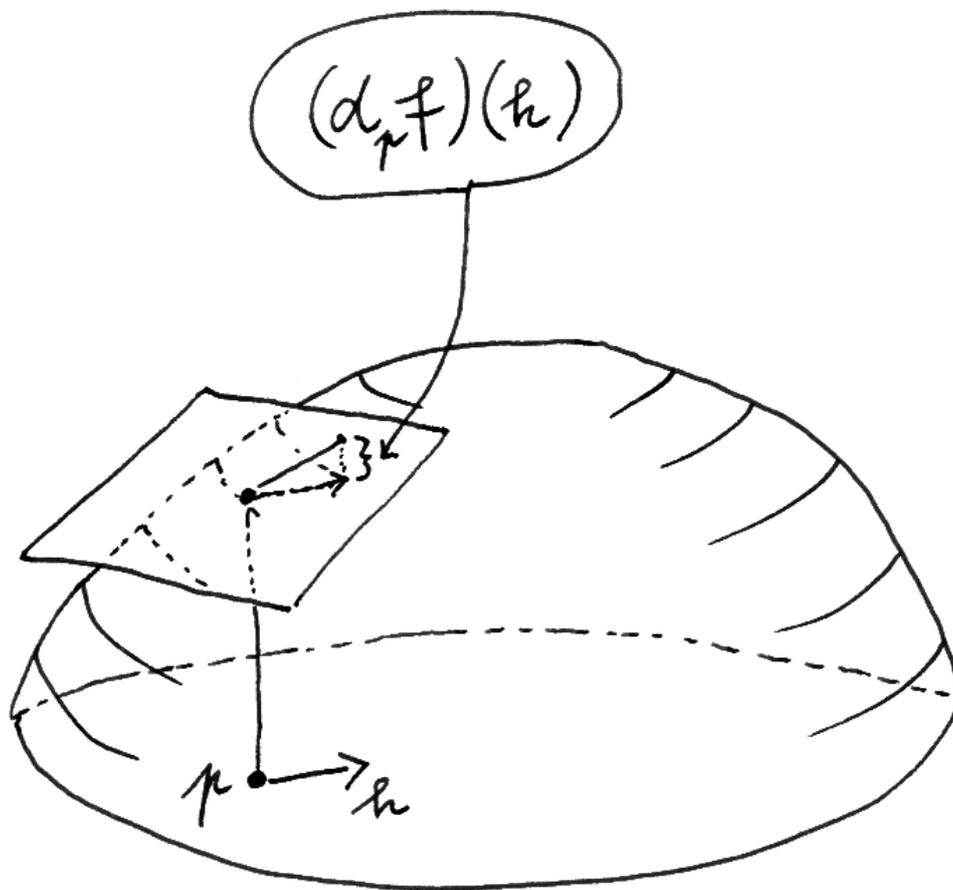
Beispiel 4.2.12. Für $\gamma : \mathbb{R} \rightarrow Y$ eine Abbildung von \mathbb{R} in einen normierten Raum Y wird unser $d_p \gamma : \mathbb{R} \rightarrow \vec{Y}$ gegeben durch Multiplikation mit dem Vektor $\gamma'(p)$ aus 10.2.1, wir haben also in Formeln

$$\gamma'(p) = (d_p \gamma)(1)$$

Später werden wir derlei Feinheiten aber auch oft ignorieren, die durch das Auswerten bei 1 gegebene kanonische Identifikation $\text{Hom}(\mathbb{R}, \vec{Y}) \xrightarrow{\sim} \vec{Y}$ nicht mehr explizit machen und schlicht $\gamma'(p) = d_p \gamma$ schreiben. Bereits bei reellwertigen Funktionen f einer reellen Veränderlichen hätten wir die Differenzierbarkeit bei p mit Ableitung b auch dadurch charakterisieren können, daß gilt $f(p + h) = f(p) + bh + |h|\varepsilon(h)$ für eine Funktion ε , die stetig ist bei Null und die dort den Wert Null annimmt. Dort konnten wir die Betragstriche um h noch ohne Schaden weglassen. Ist jedoch h ein Vektor wie in unserer allgemeinen Situation, so sind die Betragstriche als da heißt das Bilden der Norm unumgänglich, schon allein, da wir ja im allgemeinen gar kein Produkt von Vektoren zur Verfügung haben.



Ein Kegelchen der Gestalt $p + [0, 1]C$ in der Papierebene, für C die Menge aller Richtungsvektoren, die von p ins Innere der ellipsenförmigen Menge zeigen.



Anschauliche Bedeutung des Differential einer reellwertigen Funktion von zwei reellen Variablen in der Veranschaulichung der Funktion durch ihren Graphen nach 4.2.11. Der Graph des Differential ist bis auf eine Verschiebung gerade die Tangentialebene. Der Wert des Differential auf dem Vektor h ist etwa der im Bild durch eine geschweifte Klammer angedeutete Abstand oder noch genauer die zugehörige positive reelle Zahl.

Beispiel 4.2.13. Ist X ein eindimensionaler reeller Raum, $x : X \xrightarrow{\sim} \mathbb{R}$ ein Isomorphismus affiner Räume, $f : X \rightarrow Y$ eine differenzierbare Abbildung in einen normierten Raum Y und $p \in X$ ein Punkt, so erklären wir den **Differentialquotient bei p** durch die Vorschrift

$$\left. \frac{df}{dx} \right|_{x=p} = d_p f \circ (d_p x)^{-1}$$

Hier meint $(d_p x)^{-1} : \mathbb{R} \xrightarrow{\sim} \vec{X}$ die Umkehrabbildung zu $d_p x : \vec{X} \xrightarrow{\sim} \mathbb{R}$ und unser Differentialquotient ist mithin eine lineare Abbildung $\mathbb{R} \rightarrow \vec{Y}$, die wir meist mittels der durch das Auswerten bei Eins gegebenen Identifikation schlicht als einen Vektor aus \vec{Y} auffassen. Ist x die Identität auf $X = \mathbb{R}$, so ist unser Differentialquotient nur eine andere Schreibweise für die Ableitung zum Zeitpunkt p . Die allgemeinere Definition zeigt jedoch, wie gut unsere neue Notation df für das Differential verträglich ist mit unserer alten Notation $\frac{df}{dx}$ für die Ableitung.

Ergänzung 4.2.14. Ist insbesondere \mathbb{T} unsere Zeitachse aus ?? und \mathbb{E} unser Anschauungsraum aus ?? und $\gamma : \mathbb{T} \rightarrow \mathbb{E}$ die mathematische Beschreibung der Flugbahn einer unsterblichen Fliege, eine differenzierbare Abbildung zwischen besagten reellen affinen Räumen, so ist das Differential dieser Abbildung zu einem festen Zeitpunkt $t \in \mathbb{T}$ eine lineare Abbildung $d_t \gamma : \vec{\mathbb{T}} \rightarrow \vec{\mathbb{E}}$, die man als Element von $\text{Hom}(\vec{\mathbb{T}}, \vec{\mathbb{E}})$ auffassen kann oder mit ?? auch als Element von $\vec{\mathbb{E}} \otimes \vec{\mathbb{T}}^*$, das man dann die **vektorielle Geschwindigkeit** nennt. An dieser Stelle möchte ich Sie am liebsten wieder einmal davon überzeugen, daß das Abstrakte das eigentlich Konkrete ist.

4.2.15. Für das Differential, das wir hier mit $d_p f$ bezeichnen, findet man in der Literatur auch die Notationen $(Df)(p)$ und $f'(p)$. Das vorstehenden Beispiel 4.2.13 erklärt, warum ich die Notation $d_p f$ vorziehe. Einen zusätzlichen Grund findet man in 6.3.2: Dort wird erklärt, in welchem Sinne das Symbol dx , das wir bisher beim Integrieren nur benutzt haben, um die Integrationsvariable auszuzeichnen und die Substitutionsregel leichter zu erinnern, eigentlich das Differential der Funktion $\mathbb{R} \rightarrow \mathbb{R}$, $x \mapsto x$ bedeutet.

Ergänzende Übung 4.2.16. Das Differential bei der Nullmatrix der Exponentialabbildung $\exp : M(n \times n; \mathbb{C}) \rightarrow M(n \times n; \mathbb{C})$ ist die Identität, in Formeln gilt also $d_0 \exp = \text{id} : M(n \times n; \mathbb{C}) \rightarrow M(n \times n; \mathbb{C})$. Man zeige das und zeige es allgemeiner auch für die Exponentialabbildung auf dem Raum der stetigen Endomorphismen eines beliebigen Banachraums, vergleiche 10.5.23.

Ergänzende Übung 4.2.17. Eine Abbildung $f : \mathbb{C} \rightarrow \mathbb{C}$ ist komplex differenzierbar bei $p \in \mathbb{C}$ im Sinne von 1.5.1 mit Ableitung $f'(p) \in \mathbb{C}$ genau dann, wenn so f bei p differenzierbar ist im Sinne von 4.2.2 und sein Differential $d_p f : \mathbb{C} \rightarrow \mathbb{C}$ eine komplexlineare Abbildung. In diesem Fall ist das Differential von f gerade

die Multiplikation mit seiner komplexen Ableitung $f'(p)$ aus 1.5.1, in Formeln $d_p f = (f'(p) \cdot) : \mathbb{C} \rightarrow \mathbb{C}$. Analoges gilt, wenn f nur auf einer halboffenen Teilmenge von \mathbb{C} definiert ist.

4.3 Die Kettenregel

Satz 4.3.1 (Kettenregel). *Seien A, B, C halboffene Teilmengen normierter reeller Räume X, Y, Z . Seien $f : A \rightarrow B$ und $g : B \rightarrow C$ Abbildungen und $p \in A$ ein Punkt derart, daß f differenzierbar ist bei p und g differenzierbar bei $f(p)$. So ist auch $g \circ f$ differenzierbar bei p und es gilt*

$$d_p(g \circ f) = (d_{f(p)}g) \circ (d_p f)$$

4.3.2. Es ist anschaulich klar, daß die bestmögliche affine Approximation an die Verknüpfung $g \circ f$ zweier Abbildungen f und g bei einer vorgegebenen Stelle p gerade die Verknüpfung der bestmöglichen affinen Approximation an f bei p mit der bestmöglichen affinen Approximation an g bei $f(p)$ sein muß. Die Kettenregel formalisiert diese Anschauung für die linearen Anteile unserer bestmöglichen affinen Approximationen.

4.3.3. Sind unsere drei normierten Räume $\mathbb{R}^n, \mathbb{R}^m, \mathbb{R}^l$, so bedeutet die Kettenregel die Identität der Jacobi-Matrizen

$$[d_p(g \circ f)] = [d_{f(p)}g] \circ [d_p f]$$

oder ausgeschrieben die Identität

$$\begin{aligned} & \begin{pmatrix} \frac{\partial(g \circ f)_1}{\partial x_1}(p) & \cdots & \frac{\partial(g \circ f)_1}{\partial x_n}(p) \\ \vdots & & \vdots \\ \frac{\partial(g \circ f)_l}{\partial x_1}(p) & \cdots & \frac{\partial(g \circ f)_l}{\partial x_n}(p) \end{pmatrix} = \\ & \begin{pmatrix} \frac{\partial g_1}{\partial y_1}(f(p)) & \cdots & \frac{\partial g_1}{\partial y_m}(f(p)) \\ \vdots & & \vdots \\ \frac{\partial g_l}{\partial y_1}(f(p)) & \cdots & \frac{\partial g_l}{\partial y_m}(f(p)) \end{pmatrix} \circ \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(p) & \cdots & \frac{\partial f_1}{\partial x_n}(p) \\ \vdots & & \vdots \\ \frac{\partial f_m}{\partial x_1}(p) & \cdots & \frac{\partial f_m}{\partial x_n}(p) \end{pmatrix} \end{aligned}$$

Beweis. Zur Vereinfachung setzen wir $q = f(p)$, $L = d_p f$ und $M = d_q g$ und haben

$$f(p+h) = f(p) + Lh + \|h\|\varepsilon(h)$$

$$g(q+j) = g(q) + Mj + \|j\|\eta(j)$$

für Abbildungen ε und η , die stetig sind bei Null und die dort verschwinden. Wir schreiben

$$f(p+h) = q + j(h)$$

mit $j(h) = Lh + \|h\|\varepsilon(h)$ und erhalten durch Einsetzen

$$\begin{aligned}(g \circ f)(p+h) &= g(q+j(h)) \\ &= g(q) + Mj(h) + \|j(h)\|\eta(j(h)) \\ &= (g \circ f)(p) + MLh + M\|h\|\varepsilon(h) + \|j(h)\|\eta(j(h))\end{aligned}$$

Wir sind fertig, sobald wir zeigen

$$\lim_{h \rightarrow 0} M\varepsilon(h) = 0 \quad \text{und} \quad \lim_{h \rightarrow 0} \frac{\|j(h)\|}{\|h\|} \eta(j(h)) = 0$$

Der erste Grenzwert ergibt sich mühelos, $h \mapsto M\varepsilon(h)$ ist eben auch stetig bei $h = 0$ und nimmt dort den Wert Null an. Um den zweiten Grenzwert zu berechnen, schätzen wir erst ab $\|j(h)\| \leq \|h\|(\|L\| + \|\varepsilon(h)\|)$ und dann

$$\frac{\|j(h)\|}{\|h\|} \|\eta(j(h))\| \leq (\|L\| + \|\varepsilon(h)\|) \|\eta(j(h))\|$$

Die rechte Seite ist wieder stetig bei $h = 0$ und nimmt dort den Wert Null an, gleichbedeutend strebt sie also für $h \rightarrow 0$ gegen Null, und nach dem Quetschlemma 9.6.13 strebt dann die linke Seite erst recht gegen Null. \square

4.3.4. Speziell liefert die Kettenregel für stetiges lineares $M = g$ die Formel $d_p(M \circ f) = M \circ (d_p f)$, noch spezieller folgt $d_p(\lambda f) = \lambda(d_p f)$ für $\lambda \in \mathbb{R}$.

Übung 4.3.5. Seien X, Y normierte reelle Räume. Sei $A \subset X$ halboffen und $f : A \rightarrow Y$ differenzierbar. Liegt für zwei Punkte $p, q \in A$ das ganze verbindende Geradensegment $[p, q]$ in A und ist die Operatornorm des Differentials von f auf $[p, q]$ beschränkt durch eine Konstante K , in Formeln $\|d_x f\| \leq K \forall x \in A$, so gilt $\|f(p) - f(q)\| \leq K\|p - q\|$. Hinweis: 10.2.11.

4.4 Weitere Ableitungsregeln

Proposition 4.4.1 (Komponentenregel). Seien X, Y_1, Y_2 normierte Räume, $A \subset X$ eine halboffene Teilmenge und $f = (f_1, f_2) : A \rightarrow Y_1 \times Y_2$ eine Abbildung. Genau dann ist f differenzierbar bei $p \in A$, wenn f_1 und f_2 es sind, und dann gilt für die Differentiale die Formel

$$d_p f = (d_p f_1, d_p f_2) : \vec{X} \rightarrow \vec{Y}_1 \times \vec{Y}_2$$

4.4.2. Man beachte, daß $(d_p f_1, d_p f_2)$ in Matrixschreibweise unter unseren Konventionen ??, anders als die Schreibweise suggerieren mag, als Spaltenmatrix von Homomorphismen aufzufassen wäre.

Beweis. Die Projektionen $\text{pr}_i : X \rightarrow Y_i$ sind stetig und linear. Ist f differenzierbar bei p , so sind damit nach der Kettenregel auch die $f_i = \text{pr}_i \circ f$ differenzierbar bei p und die Kettenregel liefert zusätzlich $d_p f_i = d_{f(p)} \text{pr}_i \circ d_p f = \text{pr}_i \circ d_p f$, also $d_p f = (d_p f_1, d_p f_2)$. Sind umgekehrt f_1 und f_2 differenzierbar bei p mit Differentialen L_1 und L_2 , so können wir nach Definition schreiben

$$f_i(p+h) = f_i(p) + L_i h + \|h\| \varepsilon_i(h)$$

für geeignete Abbildungen ε_i , die stetig sind bei Null und die dort den Wert $\varepsilon_i(0) = 0$ annehmen. Setzen wir $L = (L_1, L_2)$ und $\varepsilon = (\varepsilon_1, \varepsilon_2)$, so ist L stetig linear und ε stetig bei 0 mit Funktionswert $\varepsilon(0) = 0$ und es gilt

$$f(p+h) = f(p) + Lh + \|h\| \varepsilon(h)$$

Das bedeutet aber genau, daß f differenzierbar ist bei p mit Differential $d_p f = L$. □

4.4.3. Mit Induktion folgt die analoge Aussage für eine Abbildung $f : A \rightarrow Y_1 \times \dots \times Y_m$ in ein längeres kartesisches Produkt normierter Räume. Insbesondere ist eine Abbildung $f = (f_1, \dots, f_m) : A \rightarrow \mathbb{R}^m$ differenzierbar bei $p \in A$ genau dann, wenn alle f_j es sind, und in diesem Fall gilt für die Differentiale die Formel

$$d_p f = (d_p f_1, \dots, d_p f_m) : \vec{X} \rightarrow \mathbb{R}^m$$

Wieder ist hier $(d_p f_1, \dots, d_p f_m)$ gemäß unseren Konventionen, anders als die Schreibweise suggerieren mag, als Spaltenmatrix von Homomorphismen aufzufassen.

Korollar 4.4.4 (Summenregel). *Seien X ein normierter Raum, \vec{Y} ein normierter Vektorraum und $A \subseteq X$ eine halboffene Teilmenge. Sind $f, g : A \rightarrow \vec{Y}$ differenzierbar bei $p \in A$, so ist auch $f + g : A \rightarrow \vec{Y}$ differenzierbar bei p und es gilt*

$$d_p(f+g) = d_p f + d_p g$$

Beweis. Die Addition $+: \vec{Y} \times \vec{Y} \rightarrow \vec{Y}$, $(w, w') \mapsto w + w'$ ist linear und stetig, und wir können $f + g$ schreiben als die Verknüpfung $f + g = + \circ (f, g)$. Das Differential von $f + g$ an der Stelle p ergibt sich also mit der Kettenregel zu $d_p(f+g) = + \circ (d_p f, d_p g) = d_p f + d_p g$. □

Proposition 4.4.5 (Differential bilinearer Abbildungen). *Seien $\vec{X}, \vec{Y}, \vec{Z}$ normierte Vektorräume und sei $\varphi : \vec{X} \times \vec{Y} \rightarrow \vec{Z}$, $(v, w) \mapsto \varphi(v, w)$ eine stetige bilineare Abbildung. So ist φ differenzierbar und das Differential von φ im Punkt (p, q) ist die lineare Abbildung*

$$\begin{aligned} d_{(p,q)} \varphi : \vec{X} \times \vec{Y} &\rightarrow \vec{Z} \\ (h, k) &\mapsto \varphi(h, q) + \varphi(p, k) \end{aligned}$$

Beweis. Wir rechnen

$$\varphi(p+h, q+k) = \varphi(p, q) + \varphi(h, q) + \varphi(p, k) + \varphi(h, k)$$

und müssen nur noch $\lim_{(h,k) \rightarrow 0} \varphi(h, k) / \|(h, k)\| = 0$ zeigen. Das folgt aber mit [9.9.25](#) aus der Stetigkeit von φ . \square

Ergänzende Übung 4.4.6. Man zeige, daß auch im Kontext normierter Vektorräume stetige multilineare Abbildungen stets differenzierbar sind, und gebe ein zur Produktregel [4.4.5](#) analoge Formel für deren Differential. Hinweis: [9.9.18](#).

Beispiel 4.4.7. Die Leibnizregel [7.2.1](#) können wir aus der Kettenregel für Differentiale herleiten wie folgt: Gegeben $f, g : \mathbb{R} \rightarrow \mathbb{R}$ schreiben wir das Produkt fg als die Verknüpfung $fg = \text{mult} \circ (f, g)$ der Funktion $(f, g) : \mathbb{R} \rightarrow \mathbb{R}^2$ mit der Multiplikation $\text{mult} : \mathbb{R}^2 \rightarrow \mathbb{R}$. Sind f und g differenzierbar bei $t \in \mathbb{R}$, so nach der Komponentenregel auch ihre Zusammenfassung (f, g) , und deren Jacobi-Matrix ist die Spaltenmatrix $[d_t(f, g)] = (f'(t), g'(t))^\top$. Andererseits ist die Multiplikation differenzierbar als stetige bilineare Abbildung oder auch nach [4.5.1](#) wegen der Existenz und Stetigkeit der partiellen Ableitungen und ihr Differential bei (x, y) hat als Jacobi-Matrix die Zeilenmatrix $[d_{(x,y)} \text{mult}] = (y, x)$. Mit der Kettenregel in mehreren Veränderlichen folgt dann

$$\begin{aligned} (fg)'(t) &= [d_t(f \circ g)] \\ &= [d_{(f(t), g(t))} \text{mult}] \circ [d_t(f, g)] \\ &= (g(t), f(t)) \circ (f'(t), g'(t))^\top \\ &= g(t)f'(t) + f(t)g'(t) \end{aligned}$$

Korollar 4.4.8. Seien $A : \mathbb{R} \rightarrow M(n \times m; \mathbb{R})$ und $B : \mathbb{R} \rightarrow M(m \times k; \mathbb{R})$ differenzierbare matrixwertige Funktionen. So ist auch das Produkt $AB : t \mapsto A(t)B(t)$ differenzierbar und die Geschwindigkeit $(AB)'$ der Produktfunktion $AB : \mathbb{R} \rightarrow M(n \times k; \mathbb{R})$ wird gegeben durch die Formel

$$(AB)' = A'B + AB'$$

4.4.9. Das sollten Sie zur Übung schon in Koordinaten nachgerechnet haben. Der hier gegebene Beweis ist komplizierter und dient in erster Linie nicht der Herleitung des Resultats, sondern vielmehr der Illustration unserer allgemeinen Regeln durch ein übersichtliches Beispiel. Man beachte jedoch auch, wie unübersichtlich dieses Beispiel wird, sobald wir versuchen, statt mit abstrakten Differentialen mit Jacobi-Matrizen zu arbeiten.

Beweis. Die Matrixmultiplikation ist eine stetige bilineare Abbildung

$$\text{Mult} : M(n \times m; \mathbb{R}) \times M(m \times k; \mathbb{R}) \rightarrow M(n \times k; \mathbb{R})$$

und wir können AB schreiben als die Verknüpfung $AB = \text{Mult} \circ (A, B)$. Mit der Kettenregel und der Komponentenregel ergibt sich

$$d_t(AB) = (d_{(A(t), B(t))} \text{Mult}) \circ (d_t A, d_t B)$$

Wenden wir diese lineare Abbildung $\mathbb{R} \rightarrow M(n \times k; \mathbb{R})$ an auf $1 \in \mathbb{R}$, so erhalten wir mit 4.4.5 wie gewünscht

$$\begin{aligned} (AB)'(t) &= d_t(AB)(1) \\ &= (d_{(A(t), B(t))} \text{Mult})(A'(t), B'(t)) \\ &= A'(t)B(t) + A(t)B'(t) \quad \square \end{aligned}$$

Übung 4.4.10. Sei $\text{inv} : GL(n; \mathbb{R}) \rightarrow M(n \times n; \mathbb{R})$ das Invertieren von Matrizen, $\text{inv}(X) = X^{-1}$. Man zeige für das Differential des Invertierens bei der Einheitsmatrix I die Formel $d_I \text{inv} : H \mapsto -H$. Man zeige allgemeiner, daß das Differential dieser Abbildung am Punkt P in Verallgemeinerung der Ableitungsregel für $x \mapsto 1/x$ gegeben wird durch

$$\begin{array}{ccc} d_P \text{inv} : M(n \times n; \mathbb{R}) & \rightarrow & M(n \times n; \mathbb{R}) \\ H & \mapsto & -P^{-1}HP^{-1} \end{array}$$

Hinweis: Man zeige erst, daß inv differenzierbar ist. Dann nehme man in der Gleichung $\text{inv}(X)X = I$ auf beiden Seiten das Differential an der Stelle P .

Übung 4.4.11. Gegeben ein Banachraum V bilden die invertierbaren Elemente von $\mathcal{B}(V)$ eine offene Teilmenge und das Invertieren ist darauf differenzierbar mit Differential $d_P \text{inv} : H \mapsto -P^{-1}HP^{-1}$. Hinweis: Man beachte, daß für alle Endomorphismen von V der Norm < 1 gilt

$$(I - H)(I + H + H^2 + H^3 \dots) = I$$

wobei 10.5.9 und 10.5.11 die Konvergenz sicherstellen. Diese Übung verallgemeinert die vorhergehende Übung 4.4.10. Wir werden dies Resultat noch in 8.4.8 im Zusammenhang mit Lösungen gewöhnlicher Differentialgleichungen brauchen.

Ergänzende Übung 4.4.12. Sei $B \in M(n \times n; \mathbb{R})$ fest. Das Differential der Abbildung $\psi : GL(n; \mathbb{R}) \rightarrow M(n \times n; \mathbb{R})$ gegeben durch $A \mapsto ABA^{-1}$ bei der Einheitsmatrix ist die lineare Abbildung $H \mapsto HB - BH$.

Ergänzende Übung 4.4.13. Gegeben ein endlichdimensionaler reeller Vektorraum V und $W \subset \text{End } V$ ein Untervektorraum seines Endomorphismenraums, der aus paarweise kommutierenden Abbildungen besteht, zeige man für das Differential von $\exp : W \rightarrow \text{End } V$ bei $A \in W$ die Formel $d_A \exp = (\cdot \exp A) : W \rightarrow \text{End } V$. Idem für V ein Banachraum und $\mathcal{B}(V)$ statt $\text{End } V$. Eine allgemeine Formel für das Differential von $\exp : \text{End } V \rightarrow \text{End } V$ wird in 5.1.15 diskutiert.

Ergänzende Übung 4.4.14. Sei V ein Banachraum und $A \in \mathcal{B}(V)$ ein stetiger Endomorphismus von V der Norm $\|A\| < 1$. Man zeige, daß das formale Einsetzen von A in die Taylorreihe von $\log(1+x)$, als da heißt, die Reihe

$$A - \frac{A^2}{2} + \frac{A^3}{3} - \dots$$

gegen einen Endomorphismus $B \in \mathcal{B}(V)$ konvergiert mit der Eigenschaft $\exp(B) = I + A$. Hinweis: Man berechne unter Verwendung von 4.4.13 die Ableitung der Abbildung $f : t \mapsto \exp(tA - \frac{t^2 A^2}{2} + \frac{t^3 A^3}{3} - \dots)$ und die Ableitung der Abbildung $t \mapsto f(t)(I + tA)^{-1}$ und zeige, daß letztere Funktion konstant ist. Ein besser verallgemeinerbares Argument findet man in ??.

4.5 Differenzierbarkeit über partielle Ableitungen

Proposition 4.5.1. Sei $U \subseteq \mathbb{R}^n$ eine offene Teilmenge und $f : U \rightarrow Y$ eine Abbildung von U in einen normierten Raum Y . Existieren alle partiellen Ableitungen und sind stetig als Abbildungen $\frac{\partial f}{\partial x_i} : U \rightarrow \vec{Y}$, so ist die Abbildung f differenzierbar.

Ergänzung 4.5.2. Seien X, Y normierte reelle Räume, $A \subset X$ eine halboffene Teilmenge und $f : A \rightarrow Y$ eine Abbildung. Genau dann heißt f **stetig differenzierbar**, wenn f differenzierbar ist und wenn zusätzlich die Abbildung $p \mapsto d_p f$ von A in den Raum der stetigen linearen Abbildungen $\mathcal{B}(\vec{X}, \vec{Y})$ mit seiner Operatornorm aus 9.9.27 stetig ist. Aus den Voraussetzungen der Proposition folgt mit demselben Beweis unmittelbar, daß f sogar stetig differenzierbar ist.

Ergänzende Übung 4.5.3. Eine stetige Funktion $f : \mathbb{R}^2 \rightarrow \mathbb{R}$, deren sämtliche Richtungsableitungen überall existieren, die jedoch im Ursprung nicht total differenzierbar ist, kann man wie folgt erhalten: Man wählt eine 2π -periodische stetig differenzierbare Funktion $g : \mathbb{R} \rightarrow \mathbb{R}$ mit der Eigenschaft $g(x + \pi) = -g(x)$, die nicht identisch Null ist, und setzt $f(r \cos \theta, r \sin \theta) = rg(\theta)$ für $r > 0$ und $f(0, 0) = 0$. Dann hängen insbesondere die Richtungsableitungen am Ursprung gar nicht linear vom Richtungsvektor ab.

Beweis. Es gilt, an jeder Stelle $p \in U$ die totale Differenzierbarkeit zu zeigen. Indem wir zu f eine geeignete Konstante sowie eine geeignete lineare Abbildung addieren, dürfen wir ohne Beschränkung der Allgemeinheit annehmen, daß gilt $f(p) = 0$ und $\frac{\partial f}{\partial x_i}(p) = 0 \quad \forall i$. Unter diesen zusätzlichen Annahmen müssen wir nun zeigen, daß f total differenzierbar ist bei p mit Differential Null. Indem wir vor f eine geeignete Verschiebung davorschalten, dürfen wir zusätzlich auch ohne Beschränkung der Allgemeinheit $p = 0$ annehmen. Gegeben eine offene konvexe Umgebung $C \subseteq \vec{Y}$ des Nullvektors von \vec{Y} finden wir nun sicher $\delta > 0$ derart, daß

alle partiellen Ableitungen $\frac{\partial f}{\partial x_i}$ auf dem Ball $B(0; \delta)$ nur Werte in C annehmen und daß dieser Ball ganz in U enthalten ist. Aus dem Schrankensatz 10.2.11 folgt für $|h| < \delta$ schon $f(h_1, \dots, h_{i-1}, h_i, 0, \dots, 0) - f(h_1, \dots, h_{i-1}, 0, 0, \dots, 0) \in h_i C$ und insgesamt

$$f(h) = f(h) - f(0) \in (h_1 + \dots + h_n)C$$

und für $h \neq 0$ also $f(h)/|h| \in nC$. Damit ergibt sich $\lim_{h \rightarrow 0} f(h)/|h| = 0$ wie gewünscht. \square

Korollar 4.5.4. Seien $x_1, \dots, x_n : \mathbb{R} \rightarrow \mathbb{R}$ differenzierbare Abbildungen und sei $F : \mathbb{R}^n \rightarrow \mathbb{R}$ stetig partiell differenzierbar. So ist die Abbildung $t \mapsto F(x_1(t), \dots, x_n(t))$ differenzierbar und ihre Ableitung an der Stelle $t = a$ wird unter Verwendung der Abkürzung $x(t) = (x_1(t), \dots, x_n(t))$ gegeben durch die Formel

$$\frac{d}{dt} \Big|_{t=a} F(x(t)) = \frac{\partial F}{\partial x_1}(x(a)) \frac{dx_1}{dt}(a) + \dots + \frac{\partial F}{\partial x_n}(x(a)) \frac{dx_n}{dt}(a)$$

4.5.5. Natürlich gilt die Aussage auch dann noch, wenn unsere Funktionen x_i auf einem echten Intervall $I \subset \mathbb{R}$ definiert sind und F auf einer offenen Teilmenge von \mathbb{R}^n , solange nur $x(t)$ stets im Definitionsbereich von F liegt. Man schreibt diese Formel meist etwas salopp in der Form

$$\frac{dF}{dt} = \frac{\partial F}{\partial x_1} \frac{dx_1}{dt} + \dots + \frac{\partial F}{\partial x_n} \frac{dx_n}{dt}$$

Beweis. Wir betrachten x als eine Abbildung $x : \mathbb{R} \rightarrow \mathbb{R}^n$. Nach Definition ist $\frac{d}{dt} \Big|_{t=a} F(x_1(t), \dots, x_n(t)) = (d_a(F \circ x))(1)$ der einzige Eintrag in der Matrix der linearen Abbildung $d_a(F \circ x) : \mathbb{R} \rightarrow \mathbb{R}$. Mit der Kettenregel finden wir nun

$$\begin{aligned} d_a(F \circ x) &= d_{x(a)}F \circ d_a x \\ &= \left(\frac{\partial F}{\partial x_1}(x(a)), \dots, \frac{\partial F}{\partial x_n}(x(a)) \right) \left(\frac{dx_1}{dt}(a), \dots, \frac{dx_n}{dt}(a) \right)^\top \\ &= \frac{\partial F}{\partial x_1}(x(a)) \frac{dx_1}{dt}(a) + \dots + \frac{\partial F}{\partial x_n}(x(a)) \frac{dx_n}{dt}(a) \end{aligned}$$

wobei in der vorletzten Zeile das Produkt einer Zeilenmatrix mit einer Spaltenmatrix zu verstehen ist, wie der obere Index \top andeutet. \square

4.5.6. Für die Richtungsableitung $(D_v f)(p)$ einer differenzierbaren reellwertigen Funktion $f : \mathbb{R}^n \rightarrow \mathbb{R}$ am Punkt p in Richtung v erhalten wir speziell

$$(D_v f)(p) = (d_p f)(v) = \langle (\text{grad } f)(p), v \rangle$$

Insbesondere wird die Richtungsableitung bei p in Richtung eines Vektors v der Länge Eins maximal genau dann, wenn der Gradient von f ein nichtnegatives Vielfaches von v ist.

Ergänzende Übung 4.5.7. Man zeige, daß die komplexe Exponentialfunktion $\exp : \mathbb{C} \rightarrow \mathbb{C}$ differenzierbar ist mit Differential

$$d_z \exp = (\exp z) \cdot : \mathbb{C} \rightarrow \mathbb{C}$$

Ergänzende Übung 4.5.8. Man zeige, daß $\det : M(n \times n; \mathbb{R}) \rightarrow \mathbb{R}$ differenzierbar ist, und daß das Differential der Abbildung \det an der Einheitsmatrix I die Spur tr ist, in Formeln

$$d_I \det = \text{tr} : M(n \times n; \mathbb{R}) \rightarrow \mathbb{R}$$

Für das Differential von \det an einer beliebigen Stelle P zeige man die Formel $(d_P \det)(H) = \text{tr}((\det P)P^{-1}H)$. Hier meint $(\det P)P^{-1}$ den Wert bei P der stetigen Fortsetzung der Abbildung $P \mapsto (\det P)P^{-1}$ vom Raum der invertierbaren Matrizen auf den Raum aller Matrizen alias die “adjungierte Matrix” P^\sharp aus ?? . Hinweis: Man mag mit 4.4.6 arbeiten, oder auch mit partiellen Ableitungen. Erinnerung: Die Spur einer Matrix ist die Summe der Einträge auf der Diagonalen.

Ergänzende Übung 4.5.9. Jede stetig differenzierbare reellwertige Funktion auf einer offenen Teilmenge der Hyperebene $0 \times \mathbb{R}^n$ oder einer offenen Teilmenge des Halbraums $\mathbb{R}_{\leq 0} \times \mathbb{R}^n$ läßt sich zu einer stetig differenzierbaren Funktion auf einer offenen Teilmenge des \mathbb{R}^{n+1} fortsetzen.

Ergänzende Übung 4.5.10 (Differential über partielle Ableitungen). Seien X, Y, Z normierte Räume, $U \subseteq X$ sowie $V \subseteq Y$ offene Teilmengen und $f : U \times V \rightarrow Z$ eine Abbildung. Wir betrachten für alle $x \in U$ die Einbettung $j_x : V \rightarrow U \times V$, $y \mapsto (x, y)$ und für alle $y \in V$ die Einbettung $i_y : U \rightarrow U \times V$, $x \mapsto (x, y)$. Existieren für alle $p = (x, y) \in U \times V$ die Differentiale $d_x(f \circ i_y) : \vec{X} \rightarrow \vec{Z}$ und $d_y(f \circ j_x) : \vec{Y} \rightarrow \vec{Z}$ und sind stetig als Funktionen $U \times V \rightarrow \mathcal{B}(\vec{X}, \vec{Z})$ bzw. $U \times V \rightarrow \mathcal{B}(\vec{Y}, \vec{Z})$, so ist die Abbildung f differenzierbar mit Differential

$$d_p f : (v, w) \mapsto (d_x(f \circ i_y))(v) + (d_y(f \circ j_x))(w)$$

Hinweis: Man kopiere mutatis mutandis den Beweis von 4.5.1. Mutige mögen umgekehrt 4.5.1 aus dem Ergebnis dieser Übung ableiten durch vollständige Induktion über n .

Ergänzung 4.5.11. Oft werden auch partielle Ableitungen in größerer Allgemeinheit verwendet als in unserer Definition 4.1.1. Sind genauer X_1, \dots, X_n eindimensionale reelle Räume und ist E ein normierter reeller Raum und $U \subseteq X_1 \times \dots \times X_n$ eine offene Teilmenge und $f : U \rightarrow E$ eine Abbildung, so bezeichnet

$$\frac{\partial f}{\partial x_i}$$

auch das “Differential der Restriktion auf X_i bei festen anderen Variablen”, eine Abbildung $\frac{\partial f}{\partial x_i} : U \rightarrow \text{Hom}(\vec{X}_i, \vec{E})$. Unter der Identifikation des Richtungsraums

unseres Produkts $X_1 \times \dots \times X_n$ mit dem Produkt der Richtungsräume und des Raums Homomorphismen von dort nach \vec{E} mit dem Produkt der Homomorphismenräume haben wir dann

$$df = \left(\frac{\partial f}{\partial x_1} \mid \dots \mid \frac{\partial f}{\partial x_n} \right)$$

Im Fall $E = \mathbb{R}^m$ erhalten wir speziell wieder unsere Jacobi-Matrix als eine Zeilenmatrix von Spaltenvektoren.

5 Mehrfache Integrale und Ableitungen

Da es mir leichter fällt, die Vertauschbarkeit der Integrationsreihenfolge anschaulich zu begründen als die Vertauschbarkeit von partiellen Ableitungen, bespreche ich nun einige Grundtatsachen zur Integration von Funktionen mehrerer Veränderlicher.

5.1 Integration über kompakte Quader

Definition 5.1.1. Ein Produkt von n Intervallen in \mathbb{R}^n nennen wir einen **Quader**. Ist $Q = [a_1, b_1] \times \dots \times [a_n, b_n] \subset \mathbb{R}^n$ ein nichtleerer kompakter Quader und $f : Q \rightarrow \mathbb{R}$ stetig, so definieren wir das **Integral** von f über Q , eine reelle Zahl $\int_Q f \in \mathbb{R}$, durch die Formel

$$\int f = \int_Q f = \int_{a_n}^{b_n} \left(\dots \left(\int_{a_1}^{b_1} f(x_1, \dots, x_n) dx_1 \right) \dots \right) dx_n$$

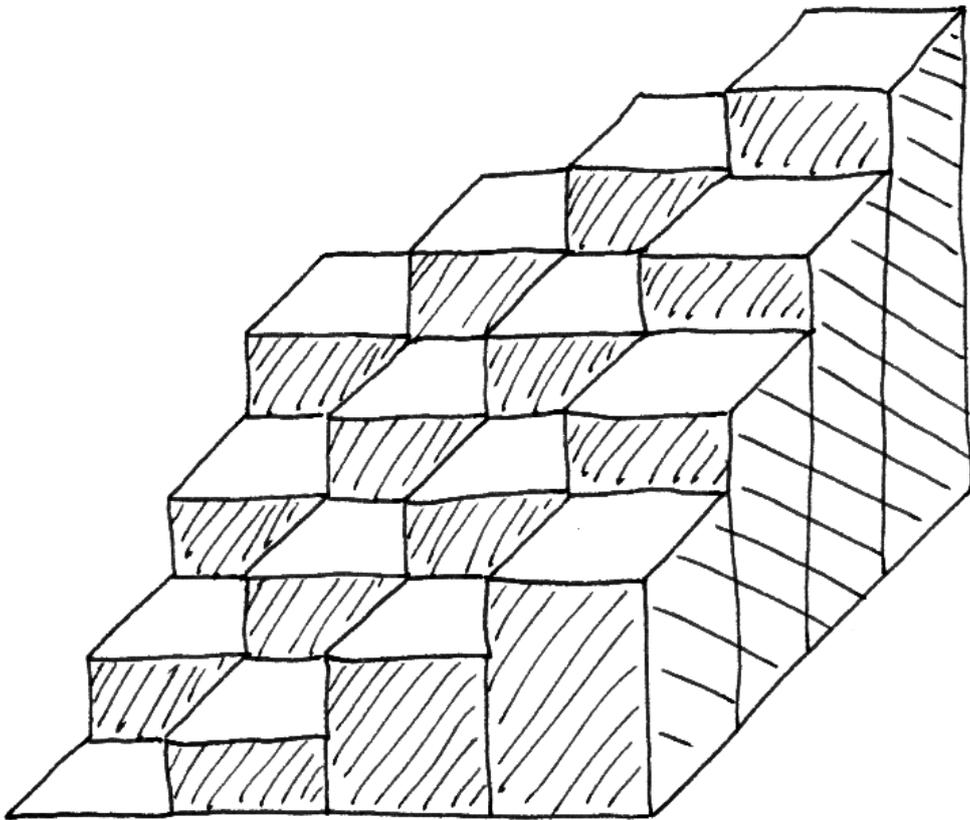
5.1.2. Proposition 9.11.1 zeigt, daß in dieser Definition alle Integranden stetig vom Integrationsparameter abhängen, so daß alle unsere Integrale definiert sind. Aus den Eigenschaften des Integrals von Funktionen einer reellen Veränderlichen folgt sofort $\int(f+g) = \int f + \int g$, $\int(\lambda f) = \lambda \int f$ für $\lambda \in \mathbb{R}$ und $f \leq g \Rightarrow \int f \leq \int g$, insbesondere auch $|\int f| \leq \int |f|$. Bezeichnet $\text{vol } Q = (b_1 - a_1) \dots (b_n - a_n)$ das Volumen des Quaders Q , so erhalten wir für eine konstante Funktion c das Integral $\int_Q c = c \text{ vol } Q$.

5.1.3. Für $n = 2$ bedeutet $\int f$ anschaulich den Rauminhalt zwischen dem Graphen der Funktion f und der xy -Ebene, wobei Rauminhalte unterhalb der xy -Ebene negativ zu zählen sind. Diese Anschauung wird im folgenden formal gerechtfertigt.

Ergänzung 5.1.4. In derselben Weise erklären wir von 1.3.3 und 1.3.9 ausgehend auch das Integral einer stetigen Abbildung von einem kompakten Quader in einen Banachraum. Es ist dann ein Vektor aus besagtem Banachraum, und die im Rest dieses Abschnitts erklärten Regeln gelten in diesem Fall entsprechend. Der Beweis der Vertauschbarkeit der Integrationsreihenfolge muß allerdings umgeschrieben werden, das mag eine gute Übung abgeben.

Definition 5.1.5. Sei $Q = [a, b] \times [c, d] \subset \mathbb{R}^2$ ein nichtleerer kompakter zweidimensionaler Quader alias ein **Rechteck** und $f : Q \rightarrow \mathbb{R}$ eine stetige Funktion. Bezeichne $\text{vol } Q = (b - a)(d - c)$ die Fläche von Q . Für $r \geq 1$ definieren wir die r -te **Riemannsumme** $S^r(f)$ von f wie folgt: Wir betrachten die äquidistanten Unterteilungen

$$a = a_0 \leq a_1 \leq \dots \leq a_r = b$$



Die vierte Riemannsumme der Funktion $f(x, y) = (x + y)/2$ auf dem Einheitsquadrat mag man sich als das Volumen des hier gezeichneten räumlichen Gebildes denken.

$$c = c_0 \leq c_1 \leq \dots \leq c_r = d$$

der Kanten unseres Rechtecks, erhalten eine Unterteilung unseres Rechtecks in r^2 kleine Rechtecke $Q_{i,j} = [a_i, a_{i+1}] \times [c_j, c_{j+1}]$ mit Flächeninhalt $(\text{vol } Q)/r^2$, und setzen

$$S^r(f) = \sum_{i,j=0}^{r-1} f(a_i, c_j) \frac{\text{vol } Q}{r^2} = \text{vol } Q \sum_{i,j=0}^{r-1} f(a_i, c_j)$$

Proposition 5.1.6. *It $Q \subset \mathbb{R}^2$ ein Rechteck und $f : Q \rightarrow \mathbb{R}$ eine stetige Funktion, so ist das Integral von f über Q der Grenzwert unserer Riemannsummen, in Formeln*

$$\int_Q f = \lim_{r \rightarrow \infty} S^r(f)$$

Beweis. Wir definieren Unter- und Obersummen durch

$$\underline{S}^r(f) = \sum_{i,j=0}^{r-1} \inf f(Q_{i,j}) \frac{\text{vol } Q}{r^2} \quad \text{und} \quad \bar{S}^r(f) = \sum_{i,j=0}^{r-1} \sup f(Q_{i,j}) \frac{\text{vol } Q}{r^2}$$

Bei den Untersummen lassen wir etwa auf unseren kleinen Quadern $Q_{i,j}$ Türmchen hochwachsen, bis sie am Graphen unserer Funktion anstoßen, und bilden die Summe der Volumina aller dieser Türmchen, und bei der Obersumme nehmen wir entsprechend die kleinstmöglichen Türmchen, aus denen unsere Funktion nicht mehr oben herausguckt. Nun behaupten wir die Ungleichungen

$$\underline{S}^r(f) \leq S^r(f) \leq \bar{S}^r(f)$$

$$\underline{S}^r(f) \leq \int_Q f \leq \bar{S}^r(f)$$

Die Ungleichungen der ersten Zeile sind offensichtlich. Um die Ungleichungen der zweiten Zeile einzusehen, benutzen wir zunächst die Regeln für Integrale einer Veränderlichen und erkennen

$$\inf f(Q_{i,j}) \frac{\text{vol } Q}{r^2} \leq \int_{Q_{i,j}} f \leq \sup f(Q_{i,j}) \frac{\text{vol } Q}{r^2}$$

Aus unseren Regeln für Integrale einer Veränderlichen folgt zusätzlich auch noch $\int_Q f = \sum_{i,j} \int_{Q_{i,j}} f$. Summieren wir dann alle unsere Ungleichungen für $0 \leq i, j \leq r-1$, so ergibt sich die zweite Zeile oben. Für alle $\varepsilon > 0$ gibt es nun wegen der gleichmäßigen Stetigkeit unserer Funktion auf unserem kompakten Rechteck ein $\delta = \delta_\varepsilon > 0$ mit

$$|(x_1, y_1) - (x, y)| < \delta \quad \Rightarrow \quad |f(x_1, y_1) - f(x, y)| < \varepsilon$$

Ist $R = R_\varepsilon$ so groß, daß alle Kantenlängen unserer kleinen Rechtecke $Q_{i,j}$ bei äquidistanter Unterteilung in R Stücke unter δ sinken, so folgt aus $r \geq R$ also $|S^r(f) - \underline{S}^r(f)| < (\text{vol } Q)\varepsilon$ und mit unseren beiden Zeilen von Ungleichungen ergibt sich $|\int_Q f - S^r(f)| < (\text{vol } Q)\varepsilon$. Das zeigt $\int_Q f = \lim_{r \rightarrow \infty} S^r(f)$ wie im Satz behauptet. \square

Korollar 5.1.7 (Vertauschen von partiellen Integrationen). *Ist $Q = [a, b] \times [c, d] \subset \mathbb{R}^2$ ein Rechteck und $f : Q \rightarrow \mathbb{R}$ stetig, so gilt*

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

Beweis. Beide Seiten sind der Grenzwert $\lim_{r \rightarrow \infty} S^r(f)$ derselben Folge von Riemannsummen. \square

5.1.8. Den gemeinsamen Wert dieses Integrals notieren wir dann kürzer auch $\int_Q f(x, y) d(x, y)$ und benutzen analoge Notationen im Fall von noch mehr Veränderlichen. Steht dahingegen x für eine Veränderliche des \mathbb{R}^k , so benutzen wir die Notation $\int f(x) d^k x$.

5.1.9. Da das Differenzieren so in etwa der inverse Prozess zum Integrieren ist, müssen mit den partiellen Integralen auch die partiellen Ableitungen sowie partielle Ableitung und partielles Integral vertauschen. Diese Idee wird im Folgenden ausgeführt.

Korollar 5.1.10 (Vertauschen von partiellen Ableitungen). *Sei $A \subset \mathbb{R}^2$ eine offene Teilmenge und $f : A \rightarrow \mathbb{R}$ eine Funktion. Existiert die gemischte partielle Ableitung $\frac{\partial}{\partial y} \left(\frac{\partial f}{\partial x} \right)$ auf A und ist dort stetig und existiert darüber hinaus die partielle Ableitung $\frac{\partial f}{\partial y}$ auf A , so existiert auch die umgekehrte gemischte partielle Ableitung $\frac{\partial}{\partial x} \left(\frac{\partial f}{\partial y} \right)$ auf A und es gilt*

$$\frac{\partial}{\partial x} \left(\frac{\partial f}{\partial y} \right) = \frac{\partial}{\partial y} \left(\frac{\partial f}{\partial x} \right)$$

5.1.11. Eine mögliche anschauliche Interpretation dieses Korollars wird der Satz über die Taylorentwicklung 5.2.1 geben: Geeignet differenzierbare reelle Funktionen von zwei Variablen besitzen eben lokal an jeder Stelle eine "beste" Approximation durch ein Polynom vom Grad höchstens zwei, also durch eine Funktion der Form $a + bx + cy + dx^2 + ey^2 + hxy$, und die gemischte partielle Ableitung unserer Funktion an besagter Stelle ist dann genau der Koeffizient h des "gemischten Terms".

Beispiel 5.1.12. Die Funktion $f(x, y) = xy(x^2 - y^2)/(x^2 + y^2)$ kann durch $f(0, 0) = 0$ stetig auf ganz \mathbb{R}^2 fortgesetzt werden und ist überall zweimal partiell differenzierbar, aber ihre beiden gemischten partiellen Ableitungen stimmen im Ursprung nicht überein. Das zeigt, daß unsere Forderung der Stetigkeit an eine gemischte partielle Ableitung im vorhergehenden Korollar 5.1.10 auch notwendig ist.

Beweis. Ohne Beschränkung der Allgemeinheit sei A ein offenes Rechteck. Wir verwenden für die partiellen Ableitungen nach der ersten bzw. zweiten Variablen die Abkürzungen f_x und f_y und schreiben $f_{xy} = (f_x)_y$ für die gemischte partielle Ableitung “erst nach x , dann nach y ”. Gegeben $(a, c) \in A$ beliebig aber fest finden wir

$$\begin{aligned} \int_a^x \int_c^y f_{xy}(s, t) dt ds &= \int_a^x f_x(s, y) - f_x(s, c) ds \\ &= f(x, y) - f(x, c) - f(a, y) + f(a, c) \end{aligned}$$

Ich bin nicht glücklich, daß die Symbole x und y hier sowohl als Integrationsgrenzen als auch als Anzeiger für zu bildende partielle Ableitungen vorkommen, aber ich fürchte, eine alternative Notation wie etwa f_1, f_2, f_{12}, f_{21} statt f_x, f_y, f_{xy}, f_{yx} wäre wieder in anderer Weise verwirrend. Jetzt vertauschen wir vorne die Integrationsreihenfolge, bringen hinten die drei letzten Summanden auf die andere Seite und erhalten

$$\left(\int_c^y \int_a^x f_{xy}(s, t) ds dt \right) + f(x, c) + f(a, y) - f(a, c) = f(x, y)$$

Die linke Seite ist hier ganz offensichtlich partiell differenzierbar erst nach y und dann nach x und ihre gemischte partielle Ableitung ergibt sich zu f_{xy} wie gewünscht. \square

Korollar 5.1.13 (Differenzieren unter dem Integral). Sei $[a, b] \subset \mathbb{R}$ ein kompaktes Intervall, $I \subset \mathbb{R}$ halboffen und $f : [a, b] \times I \rightarrow \mathbb{R}$, $(x, y) \mapsto f(x, y)$ stetig mit stetiger partieller Ableitung nach der zweiten Variablen. So ist die Funktion $y \mapsto \int_a^b f(x, y) dx$ differenzierbar und man darf die Integration über die erste Variable mit der partiellen Ableitung nach der zweiten Variablen vertauschen, in Formeln

$$\frac{d}{dy} \left(\int_a^b f(x, y) dx \right) = \int_a^b \frac{\partial f}{\partial y}(x, y) dx$$

5.1.14. Einen allgemeineren Satz zum Differenzieren unter dem Integral werden sie im Rahmen der Lebesgue’schen Integrationstheorie in Übung 9.5.14 herleiten.

Beweis. Ohne Beschränkung der Allgemeinheit sei I ein echtes Intervall. Wir wählen $c \in I$ beliebig und finden

$$\int_a^b \int_c^y f_y(x, t) dt dx = \int_a^b f(x, y) - f(x, c) dx$$

Vertauschen wir vorne die Integrationsreihenfolge und bringen den letzten Summanden auf die andere Seite, so ergibt sich

$$\int_c^y \int_a^b f_y(x, t) dx dt - \int_a^b f(x, c) dx = \int_a^b f(x, y) dx$$

Die linke Seite ist aber offensichtlich partiell differenzierbar nach y mit Ableitung $\int_a^b f_y(x, t) dx$. \square

Ergänzende Übung 5.1.15. Gegeben ein Vektorraum V und $A \in \text{End } V$ erklärt man die lineare Abbildung $\text{ad } A : \text{End } V \rightarrow \text{End } V$ durch die Vorschrift $\text{ad } A : B \mapsto (AB - BA)$. Man zeige, daß für V endlichdimensional und reell das Differential von $\exp : \text{End } V \rightarrow \text{End } V$ bei $A \in \text{End } V$ gegeben wird durch die Formel

$$d_A \exp = (\cdot \exp A) \circ \left(\frac{\exp(\text{ad } A) - 1}{\text{ad } A} \right)$$

Beim letzten Faktor ist gemeint, daß $\text{ad } A$ in die Potenzreihe $\sum_{\nu \geq 0} z^\nu / (\nu + 1)!$ der Funktion $(\exp(z) - 1)/z$ eingesetzt werden soll. Hinweis: Man wende $\partial^2 / \partial s \partial t = \partial^2 / \partial t \partial s$ an auf $\exp(s(A + tB)) \exp(-sA)$, setze $t = 0$ und integriere über s .

5.2 Taylorentwicklung in mehreren Veränderlichen

Satz 5.2.1 (Taylorentwicklung in zwei Veränderlichen). Sei $A \subseteq \mathbb{R}^2$ eine offene Teilmenge, die den Nullpunkt enthält, und sei $f : A \rightarrow \mathbb{R}$ eine d -mal stetig partiell differenzierbare Funktion. So gibt es genau ein Polynom in zwei Veränderlichen $P(x, y) = \sum_{i+j \leq d} c_{i,j} x^i y^j$ vom Grad $\leq d$ derart, daß gilt

$$\lim_{(x,y) \rightarrow (0,0)} \frac{f(x, y) - P(x, y)}{|(x, y)|^d} = 0$$

und die Koeffizienten $c_{i,j}$ dieses Polynoms P werden gegeben durch die Formel

$$c_{i,j} = \frac{1}{i!j!} \frac{\partial^{i+j} f}{\partial x^i \partial y^j}(0, 0)$$

5.2.2. Ist $f(x, y) = \sum_{i,j} a_{i,j} x^i y^j$ selbst eine Polynomfunktion, so erkennt man leicht, daß gilt

$$a_{i,j} = \frac{1}{i!j!} \frac{\partial^{i+j} f}{\partial x^i \partial y^j}(0, 0)$$

In diesem Fall liefert unsere Formel also $P(x, y) = \sum_{i+j \leq d} a_{i,j} x^i y^j$ und man sieht sofort, daß dieses P die geforderte Eigenschaft hat.

5.2.3. Um unseren Satz auch in mehr als zwei Veränderlichen übersichtlich formulieren zu können, führen wir neue Notationen ein. Gegeben ein **Multiindex** $\alpha = (\alpha_1, \dots, \alpha_n) \in \mathbb{N}^n$ definieren wir

$$\begin{aligned} |\alpha| &:= \alpha_1 + \dots + \alpha_n \\ \alpha! &:= \alpha_1! \dots \alpha_n! \\ x^\alpha &:= x_1^{\alpha_1} \dots x_n^{\alpha_n} \\ \partial^\alpha f &:= \frac{\partial^{|\alpha|} f}{\partial x_1^{\alpha_1} \dots \partial x_n^{\alpha_n}} \end{aligned}$$

wobei wir für die letzte Notation annehmen, daß f eine $|\alpha|$ -mal stetig partiell differenzierbare Funktion von $A \subseteq \mathbb{R}^n$ nach \mathbb{R} ist, so daß es insbesondere beim partiellen Ableiten nicht auf die Reihenfolge ankommt.

5.2.4. Unter einem **Polynom in mehreren Veränderlichen** x_1, x_2, \dots, x_n mit reellen Koeffizienten versteht man eine “endliche formale Summe” der Gestalt

$$\sum_{\alpha} c_{\alpha} x^{\alpha} = \sum_{\alpha} c_{\alpha} x_1^{\alpha_1} x_2^{\alpha_2} \dots x_n^{\alpha_n}$$

wobei die Summe über alle Multiindizes $\alpha \in \mathbb{N}^n$ laufen soll und alle Koeffizienten c_{α} reelle Zahlen sind, die darüber hinaus fast alle verschwinden müssen, da wir ja salopp gesprochen nur endliche formale Summen zulassen wollen. Mit dem **Grad** oder genauer dem **Totalgrad** eines Polynoms in mehreren Veränderlichen meint man $\sup\{|\alpha| \mid c_{\alpha} \neq 0\}$. Das Nullpolynom hat also den Grad $-\infty$, konstante Polynome haben den Grad Null und

$$x^4 y^3 - z^5 y + 3z^2 x^2 y^2$$

ist ein Polynom in den drei Veränderlichen x, y, z vom Grad 7. Wir werden in 5.2.7 zeigen, daß verschiedene polynomiale Ausdrücke auch verschiedene Funktionen liefern, so daß wir im Fall reeller Koeffizienten nicht so genau zu hinterfragen brauchen, was wir unter solch einem “formalen Ausdruck” eigentlich genau verstehen wollen. Den Fall beliebiger Koeffizienten diskutieren wir in ?? im Fall einer Variablen und in ?? im allgemeinen.

Satz 5.2.5 (Taylorentwicklung). Sei $A \subseteq \mathbb{R}^n$ offen, $f : A \rightarrow \mathbb{R}$ eine d -mal stetig partiell differenzierbare Funktion, und $p \in A$ ein Punkt. So gibt es genau ein Polynom P vom Grad $\leq d$ mit

$$f(p+h) = P(h) + |h|^d \varepsilon(h)$$

für eine Funktion ε mit $\lim_{h \rightarrow 0} \varepsilon(h) = 0$, und dieses Polynom wird gegeben durch die Formel

$$P(h) = \sum_{|\alpha| \leq d} \frac{(\partial^\alpha f)(p)}{\alpha!} h^\alpha$$

5.2.6. Dieser Satz ist deutlich schwächer als unsere verschiedenen Versionen im Fall einer Variablen in 8.2.2 folgende. Ich denke jedoch, daß an dieser Stelle größere Allgemeinheit den Aufwand nicht wert ist, den ich auch in einer Variablen nur getrieben habe, um den Aspekt der “Verallgemeinerung der Ableitung durch die Taylorentwicklung” herauszuarbeiten. In 8.4.12 deuten wir an, wie der vorhergehende Satz koordinatenfrei formuliert werden könnte.

Beweis. Wir beginnen mit dem Nachweis der Eindeutigkeit und zeigen dazu das folgende

Lemma 5.2.7. Sei R ein Polynom in n Veränderlichen mit reellen Koeffizienten vom Grad $\leq d$. Gilt $\lim_{h \rightarrow 0} R(h)/|h|^d = 0$, so folgt $R = 0$.

Beweis. Wir zeigen das durch Widerspruch. Wäre $R \neq 0$, so gäbe es $v \neq 0$ mit $R(v) \neq 0$, und $t \mapsto R(tv)$ wäre ein von Null verschiedenes Polynom in einer Veränderlichen $t \in \mathbb{R}$ vom Grad $\leq d$ mit $\lim_{t \rightarrow 0} R(tv)/|t|^d = 0$. Wir wissen aber schon nach 6.3.23, daß es solch ein Polynom in einer Variablen nicht gibt. \square

Aus unserem Lemma folgt sofort die Eindeutigkeit von P , denn ist \hat{P} ein anderes mögliches Approximationspolynom, so können wir das Lemma auf $R = P - \hat{P}$ anwenden. Um die Existenz der Taylorentwicklung nachzuweisen, nehmen wir ein $h \in \mathbb{R}^m$, das so klein ist, daß sogar das ganze Geradensegment $[p, p+h] = \{p+th \mid t \in [0, 1]\}$ in A enthalten ist, und betrachten die Taylorentwicklung der Funktion $g = g_h : [0, 1] \rightarrow \mathbb{R}$, $t \mapsto f(p+th)$. Wir behaupten zunächst, daß die höheren Ableitungen von g gegeben werden durch

$$g^{(\nu)}(t) = \sum_{|\alpha|=\nu} \frac{\nu!}{\alpha!} (\partial^\alpha f)(p+th) h^\alpha$$

In der Tat gilt nach der Kettenregel in mehreren Veränderlichen 4.5.4 schon mal

$$g'(t) = \frac{\partial f}{\partial x_1}(p+th) \cdot h_1 + \dots + \frac{\partial f}{\partial x_n}(p+th) \cdot h_n$$

und wir folgern induktiv

$$g^{(\nu)}(t) = \sum_{i_1, \dots, i_\nu} \frac{\partial^\nu f}{\partial x_{i_1} \dots \partial x_{i_\nu}}(p + th) \cdot h_{i_1} \dots h_{i_\nu}$$

wobei die Summe über alle möglichen ν -Tupel aus $\{1, \dots, n\}$ laufen soll. Nach dem anschließenden Lemma 5.2.8 gibt es aber genau $\nu!/\alpha!$ Möglichkeiten, ein ν -Tupel $(i_1, \dots, i_\nu) \in \{1, \dots, n\}^\nu$ so zu wählen, daß unter den i_1, \dots, i_ν jedes j genau α_j -mal vorkommt. Fassen wir also gleiche Summanden zusammen, so ergibt sich die behauptete Formel für die ν -te Ableitung $g^{(\nu)}$ von g . Jetzt schreiben wir zur Funktion $g(t)$ die Taylorreihe mit der Lagrange'schen Form des Restglieds 8.2.5 um den Entwicklungspunkt $t = 0$ hin und erhalten an der Stelle $t = 1$ mit einer kleinen Umformulierung die Gleichung

$$f(p + h) = \sum_{|\alpha| \leq d} \frac{(\partial^\alpha f)(p)}{\alpha!} h^\alpha + \sum_{|\alpha|=d} \frac{(\partial^\alpha f)(p + \xi_h h) - (\partial^\alpha f)(p)}{\alpha!} h^\alpha$$

für geeignetes $\xi_h \in (0, 1)$. Es reicht also, wenn wir für $|\alpha| = d$ zeigen, daß gilt $\lim_{h \rightarrow 0} (\partial^\alpha f)(p + \xi_h h) - (\partial^\alpha f)(p) = 0$, und das folgt sofort aus der Stetigkeit der partiellen Ableitungen. \square

Lemma 5.2.8. *Seien $\alpha_1, \dots, \alpha_n \in \mathbb{N}$ gegeben und sei $\nu = \alpha_1 + \dots + \alpha_n$ ihre Summe. So gibt es genau $\nu!/\alpha_1! \dots \alpha_n!$ Abbildungen von einer Menge X mit ν Elementen nach $\{1, \dots, n\}$ derart, daß der Wert j jeweils genau α_j -mal angenommen wird.*

Beispiel 5.2.9. Wollen wir 10 nummerierte Bälle so anmalen, daß 5 Bälle blau, 3 Bälle rot und 2 Bälle gelb werden, so gibt es dafür also $10!/(5!3!2!) = 2520$ Möglichkeiten.

Beweis. Es gibt genau $\nu!$ Möglichkeiten, unsere Menge X anzuordnen. Jede dieser Möglichkeiten liefert eine Abbildung i wie folgt: Wir bilden die ersten α_1 Zahlen auf 1 ab, die nächsten α_2 Zahlen auf 2, und so weiter, bis wir zum Schluß die letzten α_n Zahlen auf n abbilden. So erhalten wir nur Abbildungen der gewünschten Form, genauer erhalten wir so jede der gewünschten Abbildungen genau $(\alpha_1! \dots \alpha_n!)$ -mal. Das Lemma ist bewiesen. \square

5.3 Rechnen mit Approximationen

Definition 5.3.1. Eine Abbildung $P : \mathbb{R}^n \rightarrow \mathbb{R}^m$ heißt **polynomial** oder auch **regulär** genau dann, wenn sie die Gestalt $P = (P_1, \dots, P_m)$ hat, für geeignete Polynome P_1, \dots, P_m in n Veränderlichen. Haben alle unsere P_j Grad $\leq d$, so sagen wir auch, die polynomiale Abbildung P habe **Grad** $\leq d$.

Definition 5.3.2. Seien $f, g : D \rightarrow \mathbb{R}^m$ zwei auf einer Teilmenge $D \subset \mathbb{R}^n$ definierte Abbildungen. Sei $p \in D$ ein Punkt und $d \in \mathbb{N}$ eine natürliche Zahl. Wir sagen, f und g **stimmen bei p überein bis zur Ordnung d** und schreiben

$$f \sim_p^d g$$

genau dann, wenn gilt $f(p+h) - g(p+h) = |h|^d \varepsilon(h)$ für eine Funktion ε , die stetig ist bei $h = 0$ mit Funktionswert $\varepsilon(0) = 0$.

5.3.3. Ist $p \in D$ ein Häufungspunkt von D , so können wir das umschreiben zur Forderung, daß gilt $f(p) = g(p)$ und

$$\lim_{x \rightarrow p} \frac{f(x) - g(x)}{|x - p|^d} = 0$$

5.3.4. Natürlich stimmen zwei \mathbb{R}^m -wertige Funktionen bis zu einer gewissen Ordnung überein genau dann, wenn alle ihre Komponenten bis zu der entsprechenden Ordnung übereinstimmen. Schreiben wir also $f = (f_1, \dots, f_m)$ und $g = (g_1, \dots, g_m)$, so gilt

$$f \sim_p^d g \Leftrightarrow (f_j \sim_p^d g_j \quad \forall j)$$

Offensichtlich folgt auch aus $f \sim_p^d g$ und $g \sim_p^d h$ schon $f \sim_p^d h$. Sind weiter $P, Q : \mathbb{R}^n \rightarrow \mathbb{R}^m$ polynomiale Abbildungen vom Grad $\leq d$ und ist $D \subset \mathbb{R}^n$ offen, so folgt aus $P \sim_p^d Q$ schon $P = Q$.

5.3.5. Der Satz über die Taylorentwicklung 5.2.5 liefert uns für d -mal stetig partiell differenzierbares f die eindeutig bestimmte polynomiale Abbildung P vom Grad $\leq d$ mit $P \sim_p^d f$. Genauer besagt unser Satz, daß diese polynomiale Abbildung $P = (P_1, \dots, P_m)$ charakterisiert wird dadurch, daß die partiellen Ableitungen der Polynome P_j bis zur Ordnung d bei p denselben Wert annehmen wie die entsprechenden partiellen Ableitungen der Funktionen f_j .

Satz 5.3.6 (Rechnen mit Approximationen). Seien $D \subset \mathbb{R}^n, E \subset \mathbb{R}^m$ Teilmengen und $f : D \rightarrow \mathbb{R}^m, g : E \rightarrow \mathbb{R}^l$ Abbildungen mit $f(D) \subset E$. Gegeben $p \in D$ und polynomiale Abbildungen P, Q mit $f \sim_p^d P$ und $g \sim_{f(p)}^d Q$ folgt

$$g \circ f \sim_p^d Q \circ P$$

5.3.7. Im Fall $d = 0$ bedeutet das schlicht die Stetigkeit der Verknüpfung. Im Fall $d = 1$ ist die Aussage des Satzes äquivalent zur Kettenregel in mehreren Veränderlichen.

Beweis. Wir beginnen mit einem Lemma.

Lemma 5.3.8. Seien $D \subset \mathbb{R}^n$ eine Teilmenge und $f_1, f_2 : D \rightarrow \mathbb{R}$ Funktionen. Gegeben $p \in D$ und Polynome P_1, P_2 mit $f_i \sim_p^d P_i$ folgt

$$f_1 + f_2 \sim_p^d P_1 + P_2 \quad \text{und} \quad f_1 f_2 \sim_p^d P_1 P_2$$

5.3.9. Dies Lemma besteht in der Tat aus zwei Spezialfällen des Satzes, man kann nämlich die Addition $(+) : \mathbb{R}^2 \rightarrow \mathbb{R}$ betrachten und rechnen $f_1 + f_2 = (+) \circ (f_1, f_2) \sim_p^d (+) \circ (P_1, P_2) = P_1 + P_2$ und ähnlich für die Multiplikation. Wir brauchen jedoch einen unabhängigen Beweis, damit wir das Lemma beim Beweis des Satzes verwenden dürfen.

Beweis. Dem Leser überlassen. Statt P_i polynomial reicht es auch, P_i stetig bei p anzunehmen. \square

Wir zeigen nun zunächst $g \circ f \sim_p^d Q \circ f$ und dann $Q \circ f \sim_p^d Q \circ P$. Für die erste Aussage schreiben wir $g(y) = Q(y) + |y - f(p)|^d \varepsilon(y - f(p))$ und erhalten durch Einsetzen von $y = f(x)$ und Erweitern des rechten Terms

$$(g \circ f)(x) = (Q \circ f)(x) + |x - p|^d \left[\left(\frac{|f(x) - f(p)|}{|x - p|} \right)^d \varepsilon(f(x) - f(p)) \right]$$

für alle $x \neq p$. Falls $d \geq 1$ stimmt f bei p bis mindestens zur Ordnung 1 überein mit der polynomialen Abbildung P , folglich ist f differenzierbar bei p , die vordere Klammer in den eckigen Klammern bleibt beschränkt für $x \rightarrow p$ und der Ausdruck in eckigen Klammern strebt für $x \rightarrow p$ gegen Null. Falls $d = 0$ stimmt f bei p bis zur Ordnung 0 überein mit dem Polynom P , also ist f zumindest stetig bei p und der Ausdruck in eckigen Klammern strebt für $x \rightarrow p$ wieder gegen Null. Wir müssen also nur noch für jede polynomiale Abbildung Q zeigen

$$Q \circ f \sim_p^d Q \circ P$$

Es reicht sicher, das im Fall $l = 1$ zu zeigen, also für Q ein Polynom, und in diesem Fall folgt die Aussage sofort aus dem vorhergehenden Lemma 5.3.8. \square

Beispiel 5.3.10. Wollen wir für die Funktion $f(x, y) = \sin(x e^y)$ die partielle Ableitung $\frac{\partial^3 f}{\partial x (\partial y)^2}$ im Nullpunkt bestimmen, so benutzen wir unseren Satz 5.3.6 und rechnen

$$\begin{aligned} \sin t &= t - \frac{t^3}{3!} + \dots \\ x e^y &= x + xy + \frac{xy^2}{2} + \dots \\ \sin(x e^y) &= x + xy + \frac{xy^2}{2} - \frac{x^3}{3!} + \dots \end{aligned}$$

und die gesuchte partielle Ableitung bei $x = y = 0$ ergibt sich mit der Taylorreihe zu 1.

Übung 5.3.11. Eine Potenzreihe in mehreren Veränderlichen, die an allen Stellen einer offenen Menge punktweise absolut konvergiert, stellt auf dieser offenen Menge eine beliebig oft partiell differenzierbare Funktion dar.

5.4 Maxima und Minima in mehreren Veränderlichen

Definition 5.4.1. Sei A ein metrischer Raum, $f : A \rightarrow \mathbb{R}$ eine Funktion und $p \in A$ ein Punkt. Wir sagen, f hat bei p ein **lokales Minimum** (bzw. **Maximum**) genau dann, wenn gilt $f(x) \geq f(p)$ (bzw. $f(x) \leq f(p)$) für alle x in einer hinreichend kleinen Umgebung von p . Wir sagen, f hat bei p ein **isoliertes lokales Minimum** (bzw. **Maximum**) genau dann, wenn gilt $f(x) > f(p)$ (bzw. $f(x) < f(p)$) für alle von p verschiedenen x in einer hinreichend kleinen Umgebung von p .

Proposition 5.4.2. Sei $A \subseteq \mathbb{R}^n$ offen, $f : A \rightarrow \mathbb{R}$ differenzierbar und $p \in A$ ein Punkt. Besitzt f bei p ein lokales Minimum oder Maximum, so gilt $\frac{\partial f}{\partial x_i}(p) = 0$ für alle i .

5.4.3. Die Bedingung, A sei offen, ist in diesem Zusammenhang wesentlich, wie wir bereits im Fall einer Veränderlichen in 7.3.3 diskutiert haben.

Beweis. Für beliebiges i und hinreichend kleines $\varepsilon > 0$ betrachten wir das parametrisierte Geradensegment $g : (-\varepsilon, \varepsilon) \rightarrow A$, $t \mapsto p + t e_i$. Natürlich muß auch $f \circ g : (-\varepsilon, \varepsilon) \rightarrow \mathbb{R}$ ein lokales Minimum oder Maximum bei $t = 0$ haben, also gilt $(f \circ g)'(0) = \frac{\partial f}{\partial x_i}(p) = 0$. \square

Definition 5.4.4. Sei $A \subseteq \mathbb{R}^n$ eine offene Teilmenge und $f : A \rightarrow \mathbb{R}$ eine reellwertige Funktion. Verschwinden alle ersten partiellen Ableitungen unserer Funktion an einer Stelle $p \in A$, so sagt man, die Funktion habe bei p eine **kritische Stelle**. Ist allgemeiner A eine halboffene Teilmenge eines normierten reellen Vektorraums und $f : A \rightarrow \mathbb{R}$ eine reellwertige Funktion und verschwindet ihr Differential an einem Punkt $p \in A$, in Formeln $d_p f = 0$, so sagt man auch, die Funktion habe bei p eine **kritische Stelle**.

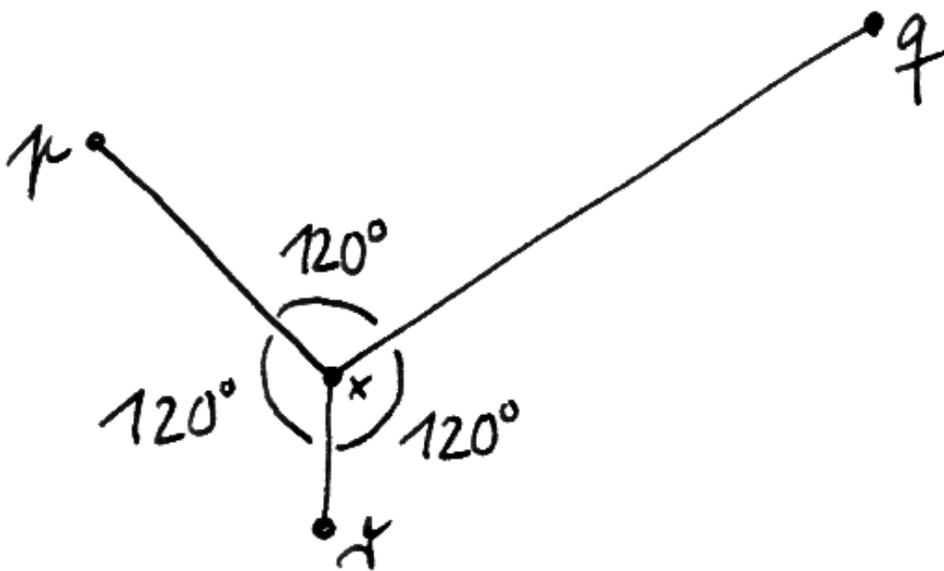
Übung 5.4.5. Sei V ein normierter Raum, $A \subseteq V$ offen, $f : A \rightarrow \mathbb{R}$ differenzierbar und $p \in A$ ein Punkt. Besitzt f bei p ein lokales Minimum oder Maximum, so folgt $d_p f = 0$.

Beispiel 5.4.6. Gegeben drei Punkte $p, q, r \in \mathbb{R}^2$ suchen wir die Punkte $x \in \mathbb{R}^2$, für die die Summe der Abstände

$$S(x) = \|x - p\| + \|x - q\| + \|x - r\|$$

kleinstmöglich wird. Sicher gilt $\lim_{\|x\| \rightarrow \infty} S(x) = \infty$, folglich existiert ein Kompaktum $K \subset \mathbb{R}^2$ mit

$$\inf_{x \in \mathbb{R}^2} S(x) = \inf_{x \in K} S(x)$$



Die Summe der Abstände zu drei vorgegebenen Punkten, die nicht auf einer Gerade liegen und ein Dreieck bilden, in dem kein Winkel größergleich 120° ist, wird minimal an der Stelle, an der die Halbgeraden zu den Ecken jeweils den Winkel 120° einschließen. Ist dahingegen ein Winkel größergleich 120° , so liegt das Minimum bei der fraglichen Ecke selbst.

und damit nimmt unsere Funktion nach 9.7.11 ihr Infimum auch wirklich als Funktionswert an. Unsere Funktion ist auf $\mathbb{R}^2 \setminus \{p, q, r\}$ stetig differenzierbar und ihr Gradient bei x ergibt sich nach kurzer Rechnung zu

$$(\text{grad } S)(x) = \frac{x - p}{\|x - p\|} + \frac{x - q}{\|x - q\|} + \frac{x - r}{\|x - r\|}$$

Für das Minimum kommen nach unseren Erkenntnissen nur unsere drei Punkte p, q, r sowie die Nullstellen des Gradienten in Frage. Die weiteren Überlegungen führen wir nicht mehr in formaler Strenge durch, da das von unseren formalen Kenntnissen ausgehend einen unangemessenen Aufwand bedeuten würde. Anschaulich scheint es mir klar, daß unser Gradient nur dann verschwinden kann, wenn nicht alle drei Punkte p, q, r auf einer Geraden liegen und x im Inneren der zugehörigen Dreiecksfläche alias ihrer konvexen Hülle liegt und wenn die drei Vektoren $x - p, x - q$ und $x - r$ jeweils einen Winkel von 120° alias $2\pi/3$ einschließen. Das ist für einen Punkt im Innern der Dreiecksfläche jedoch nur dann möglich, wenn jeder der Winkel unseres Ausgangsdreiecks kleiner ist als 120° . Nur unter den Voraussetzungen, daß unsere drei Punkte p, q, r nicht auf einer Geraden liegen und jeder der Winkel des Dreiecks mit den Ecken p, q, r kleiner ist als 120° , kann also das Minimum außerhalb der drei Punkte $\{p, q, r\}$ angenommen werden. Sind sie erfüllt, so kann das Minimum hinwiederum nicht an einem dieser Punkte angenommen werden, da der Wert von $S(x)$ dann abnimmt, wenn wir auf einer Winkelhalbierenden ins Dreieck hineinlaufen, wie man etwa an unserer Beschreibung des Gradienten sehen kann. Folglich muß dann das Minimum bei der kritischen Stelle angenommen werden, die eben dadurch charakterisiert ist, daß die Vektoren $x - p, x - q$ und $x - r$ jeweils einen Winkel von $120^\circ = 2\pi/3$ einschließen.

5.4.7. Um eine hinreichende Bedingung für ein lokales Minimum oder Maximum zu erhalten, müssen wir wie im Fall einer Veränderlichen die zweiten Ableitungen untersuchen. Am Beispiel der Funktionen $(x, y) \mapsto x^2 + y^2$ bzw. x^2 bzw. $x^2 - y^2$ kann man sehen, was lokal um $(0, 0) \in \mathbb{R}^2$ so alles passieren kann. Wir betrachten nun allgemeiner eine beliebige **quadratische Form** $q(x_1, \dots, x_n) = \sum a_{ij}x_ix_j$ mit $a_{ij} \in \mathbb{R}$ wie in ??.

Definition 5.4.8. Eine quadratische Form $q : \mathbb{R}^n \rightarrow \mathbb{R}$ heißt

positiv definit	genau dann, wenn gilt	$q(x) > 0 \quad \forall x \neq 0;$
positiv semidefinit	genau dann, wenn gilt	$q(x) \geq 0 \quad \forall x;$
negativ definit	genau dann, wenn gilt	$q(x) < 0 \quad \forall x \neq 0;$
positiv semidefinit	genau dann, wenn gilt	$q(x) \leq 0 \quad \forall x;$
indefinit	genau dann, wenn gilt	es gibt $x, y \in \mathbb{R}^n$ mit $q(x) > 0, q(y) < 0.$

5.4.9. Wir werden nachher erklären, wie man für eine gegebene quadratische Form entscheiden kann, welche Definitheitseigenschaften sie hat. Zunächst diskutieren wir jedoch, inwieweit diese Eigenschaften für den quadratischen Approximationsterm das lokale Verhalten einer Funktion an einer kritischen Stelle bestimmen.

Satz 5.4.10 (Maxima und Minima in mehreren Veränderlichen). Sei $A \subseteq \mathbb{R}^n$ offen, $f : A \rightarrow \mathbb{R}$ zweimal stetig partiell differenzierbar und $p \in A$ eine kritische Stelle. Wir bilden zu unserer Funktion die quadratische Form

$$q(h) = q_p(h) = \frac{1}{2} \sum \frac{\partial^2 f}{\partial x_i \partial x_j}(p) h_i h_j$$

die also gerade aus den quadratischen Termen der Taylorreihe besteht.

1. Ist unsere quadratische Form q positiv definit, so hat f bei p ein isoliertes lokales Minimum.
2. Ist unsere quadratische Form q negativ definit, so hat f bei p ein isoliertes lokales Maximum.
3. Ist unsere quadratische Form q indefinit, so hat f bei p weder ein lokales Minimum noch ein lokales Maximum.

5.4.11. Der Faktor $1/2$ stand in der Taylorformel nur vor den reinen und nicht vor den gemischten doppelten partiellen Ableitungen. In der im Satz gegebenen Darstellung erstreckt sich die Summe dafür über alle i, j und nicht nur über $i \leq j$. Man beachte, daß der Satz keine Aussage für die semidefiniten Fälle macht, in Verallgemeinerung der Tatsache, daß man auch für Funktionen einer Veränderlichen bei Verschwinden der ersten und zweiten Ableitung an einer vorgegebenen Stelle ohne weitere Informationen noch nichts über Maxima oder Minima aussagen kann.

Beweis. Aus unseren Annahmen folgt mit der Taylor-Formel

$$f(p+h) = f(p) + q(h) + \varepsilon(h)|h|^2$$

für eine Funktion ε mit $\lim_{h \rightarrow 0} \varepsilon(h) = 0$. Wir behandeln nun als erstes den Fall q positiv definit. Sei a das Minimum nach 9.7.11 von q auf der Oberfläche des Einheitswürfels, $a = \inf\{q(h) \mid |h| = 1\}$. Aus unserer Annahme folgt $a > 0$. Offensichtlich gilt $q(h) \geq a|h|^2$ für alle $h \in \mathbb{R}^n$. Wegen $\lim_{h \rightarrow 0} \varepsilon(h) = 0$ finden wir $\delta > 0$ derart, daß aus $|h| < \delta$ folgt $|\varepsilon(h)| \leq a/2$. Damit ergibt sich für $|h| < \delta$ aber

$$f(p+h) \geq f(p) + (a/2)|h|^2$$

und f hat in der Tat ein isoliertes lokales Minimum. Ist q negativ definit, so argumentieren wir entsprechend. Ist q indefinit, so finden wir zwei Geraden durch Null derart, daß die Einschränkung von q auf diese Geraden außerhalb des Nullpunkts positiv bzw. negativ ist. Dann muß aber die Restriktion von f auf die erste Gerade ein isoliertes lokales Minimum haben bei p , und auf der zweiten Geraden ein isoliertes lokales Maximum. Folglich hat f bei p weder ein lokales Maximum noch ein lokales Minimum. \square

Übung 5.4.12. Man zeige in der Situation des Satzes: Ist q positiv semidefinit und verschieden von Null, so kann f bei p kein lokales Maximum haben.

5.4.13. Jeder quadratischen Form $q(x_1, \dots, x_n) = \sum_{i,j=1}^n a_{ij}x_ix_j$ ordnen wir die symmetrische Matrix $M(q)$ zu mit Einträgen $(a_{ij} + a_{ji})/2$. Nach den Definitionen hat die quadratische Form q eine gewisse Definitheit im Sinne unserer Definition 5.4.8 genau dann, wenn ihre Matrix die entsprechende Definitheit hat im Sinne der linearen Algebra.

Definition 5.4.14. Die Matrix der zweiten partiellen Ableitungen, die zum Doppelpen der quadratischen Form q aus unserem Satz gehört, heißt die **Hesse-Matrix** $H(f)$ von f , in Formeln

$$H(f) = \left(\frac{\partial^2 f}{\partial x_i \partial x_j} \right)_{i,j=1}^n$$

Korollar 5.4.15 (Maxima, Minima und Hesse-Matrix). Sei $A \subseteq \mathbb{R}^n$ offen, $f : A \rightarrow \mathbb{R}$ zweimal stetig partiell differenzierbar und $p \in A$ ein kritischer Punkt.

1. Ist die Hesse-Matrix von f bei p positiv definit, so hat f bei p ein isoliertes lokales Minimum.
2. Ist die Hesse-Matrix von f bei p negativ definit, ein isoliertes lokales Maximum.
3. Ist die Hesse-Matrix von f bei p indefinit, so hat f bei p weder ein lokales Minimum noch ein lokales Maximum.

Beweis. Das ist nur eine Umformulierung von Satz 5.4.10. \square

5.4.16. Um die Definitheitseigenschaften einer symmetrischen quadratischen Matrix zu bestimmen, bringt man sie am einfachsten durch Basiswechsel in Diagonalgestalt, wie im Beweis von ?? erklärt. Bei kleineren Matrizen kann auch das Hurwitz-Kriterium ?? ein guter Trick sein: Danach ist eine symmetrische $(n \times n)$ -Matrix positiv definit genau dann, wenn für alle $k < n$ die quadratische Untermatrix, die man erhält durch Wegstreichen der letzten k Spalten und der untersten k Zeilen, eine positive Determinante hat.

6 Wegintegrale

6.1 Vektorfelder und Kovektorfelder

6.1.1. Unter einem reellen Raum verstehen wir wie in 9.9.1 einen affinen Raum über dem Körper der reellen Zahlen. Zu jedem reellen Vektorraum V bilden wir, wie in der linearen Algebra in ?? folgende ausführlich diskutiert und erläutert, seinen **Dualraum** $V^* = \text{Hom}_{\mathbb{R}}(V, \mathbb{R})$.

Definition 6.1.2. Sei X ein endlichdimensionaler reeller Raum und $U \subset X$ eine halboffene Teilmenge. Ein **Vektorfeld auf U** ist eine Abbildung

$$\begin{aligned} A : U &\rightarrow \vec{X} \\ p &\mapsto A_p \end{aligned}$$

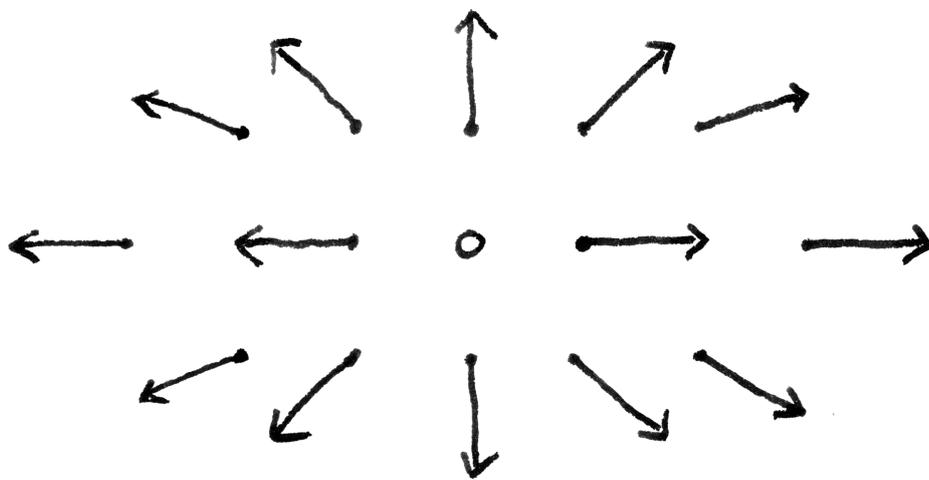
von U in den Richtungsraum \vec{X} von X . Wir schreiben im Zusammenhang mit Differentialgleichungen statt A_p auch $A(p)$. Die Notation A_p dahingegen ist praktisch, wenn wir unsere Vektorfelder wie in 6.1.13 auf Funktionen anwenden wollen. In der physikalischen Terminologie heißen Vektorfelder **kontravariant** aus Gründen, die in 6.1.15 noch diskutiert werden.

Ergänzung 6.1.3. Sicherlich könnte man diese Definition auch allgemeiner für beliebige Teilmengen U vereinbaren. Das führt jedoch auf die Schwierigkeit, daß wir etwa ein Vektorfeld auf einer Kreislinie später werden verstehen wollen als eine Abbildung, die jedem Punkt der Kreislinie in der Ebene einen Tangentialvektor an besagte Kreislinie an besagtem Punkt zuordnet und nicht einfach irgendeinen Richtungsvektor der Ebene. Daß wir uns an dieser Stelle auf halboffene Teilmengen U beschränken, dient einzig und allein dem Zweck, derartige begriffliche Inkonsistenzen zu vermeiden.

Definition 6.1.4. Sei X ein endlichdimensionaler reeller Raum und $U \subset X$ eine halboffene Teilmenge. Ein **Kovektorfeld auf U** ist eine Abbildung

$$\begin{aligned} \omega : U &\rightarrow \vec{X}^* \\ p &\mapsto \omega_p \end{aligned}$$

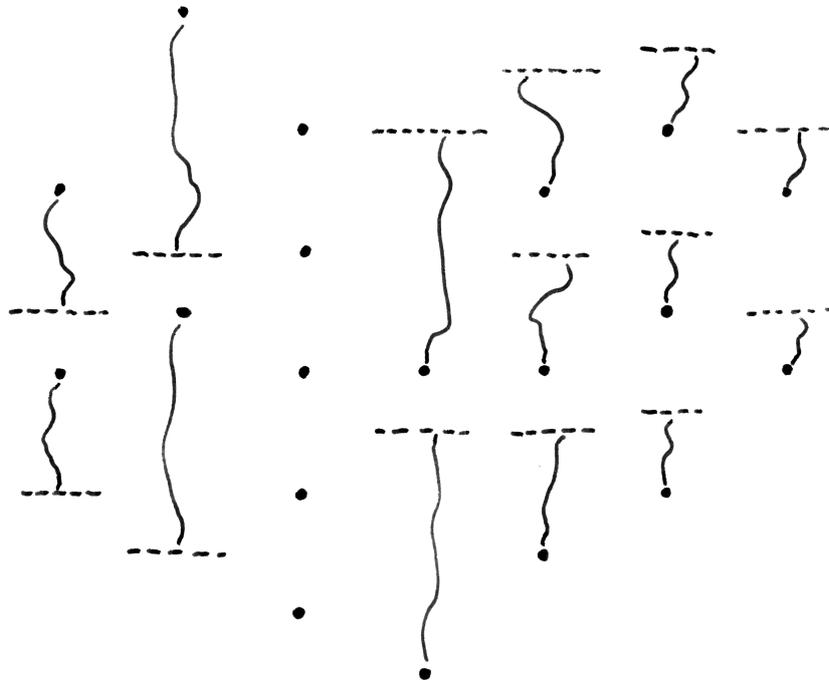
von U in den Dualraum \vec{X}^* des Richtungsraums von X . Wir schreiben hier ω_p statt $\omega(p)$, damit $\omega_p(v) \in \mathbb{R}$ den Wert der Linearform ω_p auf einem Vektor $v \in \vec{X}$ bezeichnen kann. Ein Kovektorfeld nennt man auch eine **Pfaff'sche Form** oder eine **Differentialform erster Ordnung** oder eine **1-Form**. In der physikalischen Terminologie heißen Kovektorfelder **kovariant** aus Gründen, die in 6.1.15 noch diskutiert werden.



Graphische Darstellung eines Vektorfelds auf der Papierebene, das in geeigneten Koordinaten in der Notation von 6.1.14 durch die Formel

$$\left(\frac{x}{\sqrt{x^2 + y^2}} \right) \partial_x + \left(\frac{y}{\sqrt{x^2 + y^2}} \right) \partial_y$$

gegeben würde. Hier haben wir zu ausgewählten Punkten den ihnen zugeordneten Richtungsvektor als Pfeil von besagtem Punkt zu dem um diesen Richtungsvektor verschobenen Punkt dargestellt.



Versuch der graphischen Darstellung eines Kovektorfelds auf der Papierebene, das in geeigneten Koordinaten in der Notation 6.1.10 durch die Formel

$$x \, dy$$

gegeben würde. Hier haben wir zu ausgewählten fett eingezeichneten Punkten den ihnen zugeordneten Kovektor dargestellt durch eine gestrichelte Linie, die jeweils einen Teil der Geraden zeigt, deren Punkte vom jeweiligen fetten Punkt durch einen Richtungsvektor erreicht werden können, auf dem der dem jeweilige

Kovektor den Wert 1 annimmt. Die eingezeichneten Fäden deuten an, welche gestrichelte Linie jeweils zu welchem fetten Punkt gehört. Je weiter die gestrichelte Linie von ihrem fetten Punkt entfernt ist, desto kleiner ist also unser Kovektor, zum Beispiel bedeutet der doppelte Abstand den halben Kovektor.

Fette Punkte ganz ohne gestrichelte Linie stehen für den Wert Null unseres Kovektorfelds an besagter Stelle. Daß eine gestrichelte Linie durch "ihren" fetten Punkt geht, ist nicht zulässig. Man mag versuchen, in diesem Bild auch noch das Vektorfeld $(x, y) \mapsto (1, x)$ oder in der Notation aus 6.1.14 geschriebenen $\partial_x + x\partial_y$ einzuzichnen und anschaulich zu verstehen, daß as Einsetzen im Sinne von 6.1.7 dieses Vektorfelds in unser Kovektorfeld auch tatsächlich die Funktion

$$(x, y) \mapsto x^2 \text{ liefert.}$$

Beispiel 6.1.5. Ein Kovektorfeld auf der Zeitachse \mathbb{T} aus ?? können wir uns in der in ?? noch genauer erklärten Weise denken als eine Vorschrift, die jedem Zeitpunkt eine Frequenz oder, vielleicht noch besser, eine Drehgeschwindigkeit zuordnet.

6.1.6. Wir addieren Vektorfelder und auch Kovektorfelder punktweise, die Summe $\omega + \eta$ zweier Kovektorfelder ist also erklärt durch $(\omega + \eta)_p = \omega_p + \eta_p$, wobei letzteres Summenzeichen die Addition in \vec{X}^* meint. Wir multiplizieren Vektorfelder und auch Kovektorfelder mit Funktionen $f : U \rightarrow \mathbb{R}$, indem wir setzen $(fA)_p = f(p)A_p$ bzw. $(f\omega)_p = f(p)\omega_p$.

6.1.7. Sei X ein endlichdimensionaler reeller Raum und $U \subset X$ eine halboffene Teilmenge. Ist $A : U \rightarrow \vec{X}$ ein Vektorfeld und $\omega : U \rightarrow \vec{X}^*$ ein Kovektorfeld, so können wir auch das Vektorfeld A in das Kovektorfeld ω einsetzen oder, vielleicht besser gesagt, das Kovektorfeld ω auf dem Vektorfeld A auswerten oder, ganz ausgewogen und immer noch gleichbedeutend, das **Kovektorfeld ω mit dem Vektorfeld A paaren**. Wir erhalten dann eine Funktion

$$\begin{aligned} \omega(A) = \langle \omega, A \rangle : U &\rightarrow \mathbb{R} \\ p &\mapsto \omega_p(A_p) \end{aligned}$$

6.1.8. Sei X ein endlichdimensionaler reeller Raum und $U \subset X$ eine halboffene Teilmenge. Im Sinne von 2.2.27 können wir ein Kovektorfeld $\omega : U \rightarrow \vec{X}^*$ auch auffassen als eine Abbildung $U \times \vec{X} \rightarrow \mathbb{R}$ oder sogar als eine Abbildung $\vec{X} \rightarrow \text{Ens}(U, \mathbb{R})$. Es gehört etwas Übung dazu, alle diese verschiedenen Aspekte gleichzeitig präsent zu haben. Wir können also ein Kovektorfeld einerseits an einem Punkt $p \in U$ auswerten und so eine Linearform $\omega_p : \vec{X} \rightarrow \mathbb{R}$ auf dem Richtungsraum erhalten, wir können es aber andererseits auch auf einem Richtungsvektor $v \in \vec{X}$ auswerten und so eine reellwertige Funktion $U \rightarrow \mathbb{R}$, $p \mapsto \omega_p(v)$ erhalten. Wir können es sogar etwas allgemeiner, wie in 6.1.7 besprochen, auf einem Vektorfeld $p \mapsto v_p$ auswerten und auch so eine reellwertige Funktion $U \rightarrow \mathbb{R}$, $p \mapsto \omega_p(v_p)$ erhalten. Man beachte, daß beim Auswerten von Kovektorfeldern auf Vektorfeldern keinerlei Differentiation stattfindet, sondern ausschließlich lineare Algebra, nur eben “in Abhängigkeit vom Punkt p ”.

Definition 6.1.9. Sei X ein reeller endlichdimensionaler Raum und $U \subset X$ eine halboffene Teilmenge. Ist $f : U \rightarrow \mathbb{R}$ differenzierbar, so ist das Differential von f bei p eine lineare Abbildung $d_p f : \vec{X} \rightarrow \mathbb{R}$. Unter dem **Differential df von f** verstehen wir das Kovektorfeld auf U , das gegeben wird durch die Vorschrift

$$\begin{aligned} df : U &\rightarrow \vec{X}^* \\ p &\mapsto d_p f \end{aligned}$$

6.1.10. Für das Differential von einem Produkt gilt nach 4.4.5 die Produktregel $d(fg) = f dg + g df$ und für das Differential einer Summe haben wir $d(f + g) = df + dg$. Ist speziell $X = \mathbb{R}^n$ und $U \subset \mathbb{R}^n$ halboffen und bezeichnet $x_i : U \rightarrow \mathbb{R}$ die Restriktion der i -ten Koordinate auf U , so ist $dx_i : U \rightarrow (\mathbb{R}^n)^*$ konstant die die i -te Koordinate selber. Die Koordinaten bilden nun eine Basis des Dualraums von \mathbb{R}^n , folglich läßt sich jedes Kovektorfeld auf U schreiben in der Gestalt $\sum a_i dx_i$ mit eindeutig bestimmten $a_i : U \rightarrow \mathbb{R}$. Für eine differenzierbare Funktion $f : U \rightarrow \mathbb{R}$ auf einer halboffenen Menge $U \subset \mathbb{R}^n$ haben wir dann

$$df = \sum_{i=1}^n \frac{\partial f}{\partial x_i} dx_i$$

Man prüft das leicht durch Auswerten beider Seiten an einer Stelle $p \in U$ und Anwenden der so entstehenden Linearformen auf alle Vektoren der Standard-Basis des \mathbb{R}^n . Speziell haben wir für $f : \mathbb{R} \rightarrow \mathbb{R}$ also $df = f'(x) dx$. Ist U nicht offen, sondern nur halboffen, so sind die partiellen Ableitungen oben im Sinne unserer Notation 4.2.9 zu verstehen.

6.1.11. Anschaulich gesprochen beschreibt diese Gleichung, wie sich der Funktionswert der Funktion f in erster Näherung ändert, wenn wir an den Koordinaten x_i wackeln: Genauer gilt bei festen x_1, \dots, x_n für $\delta x_1, \dots, \delta x_n \in \mathbb{R}$ so nah bei Null, daß alles definiert ist, eben

$$f(x_1 + \delta x_1, \dots, x_n + \delta x_n) - f(x_1, \dots, x_n) = \sum_{i=1}^n \frac{\partial f}{\partial x_i} \delta x_i + R(\delta x_1, \dots, \delta x_n)$$

mit einem Rest R , der auch nach dem Teilen durch das Maximum der Beträge aller δx_i noch gegen Null strebt, wenn alle δx_i gegen Null streben. Hierbei ist zu verstehen, daß die fraglichen partiellen Ableitungen an unserer festen Stelle (x_1, \dots, x_n) ausgewertet werden sollen, und um die partiellen Ableitungen zu bilden, müssen die x_i natürlich noch als variabel gedacht werden. Vielleicht wäre es hier konsistenter gewesen, die partiellen Ableitungen $\partial_i f$ zu notieren oder sogar $(\partial_i f)(x_1, \dots, x_n)$ um anzudeuten, daß sie ja an der festen Stelle (x_1, \dots, x_n) auszuwerten sind, aber es kommt bei komplizierteren Formeln auch nicht selten vor, daß größere Präzision insbesondere für fortgeschrittene Leser nicht zu besserer Verständlichkeit führt. Die Notation δx_i könnten wir zu δ_i abkürzen, aber dann wirkt die Formel weniger suggestiv. Kürzen wir auch noch die linke Seite zu δf ab, so können wir unsere Identität mit der in 5.3.2 eingeführten Notation auch schreiben als die Übereinstimmung erster Ordnung von Funktionen der "Verrückungen" δx_i der Gestalt

$$\delta f \sim_0^1 \sum_{i=1}^n \frac{\partial f}{\partial x_i} \delta x_i$$

Beispiel 6.1.12. Die Funktion $f : \mathbb{R}^3 \setminus 0 \rightarrow \mathbb{R}$, $v \mapsto 1/\|v\|$ hat mit der Konvention $v = (x, y, z)$ das Differential $df = -(x dx + y dy + z dz)/\|v\|^3$.

Definition 6.1.13. Gegeben ein endlichdimensionaler reeller Raum X , eine halboffene Teilmenge $U \subset X$, ein Vektorfeld $A : U \rightarrow \vec{X}$ und eine differenzierbare Funktion $f : U \rightarrow \mathbb{R}$ erklären wir eine Funktion $(Af) : U \rightarrow \mathbb{R}$ durch die Vorschrift

$$(Af)(p) := (d_p f)(A_p)$$

Ist U eine Umgebung von p , so ist nach 4.2.5 also $(Af)(p)$ die Richtungsableitung von f bei p in der Richtung A_p . Wir sagen deshalb auch, die Funktion Af entstehe aus f durch **Ableiten in Richtung des Vektorfelds A** . In anderen Worten entsteht diese Funktion durch das Paaren des Vektorfelds A mit der durch das Differential der Funktion f gegebene Kovektorfeld df . Mit unserer Notation 6.1.7 kann diese Funktion auch $\langle df, A \rangle$ geschrieben werden.

6.1.14. Meist werden Vektorfelder identifiziert mit den zugehörigen Differentialoperatoren. So bezeichnet man etwa das konstante Vektorfeld mit Wert e_i auf \mathbb{R}^n oft als “das Vektorfeld $\frac{\partial}{\partial x_i}$ ” oder “das Vektorfeld ∂_i ” und im Fall nicht nummerierter Koordinaten wie etwa x, y, z auf \mathbb{R}^3 schreiben wir für die fraglichen Vektorfelder auch $\partial_x, \partial_y, \partial_z$ oder dergleichen. Sicher läßt sich jedes Vektorfeld auf $U \subset \mathbb{R}^n$ halboffen schreiben in der Gestalt

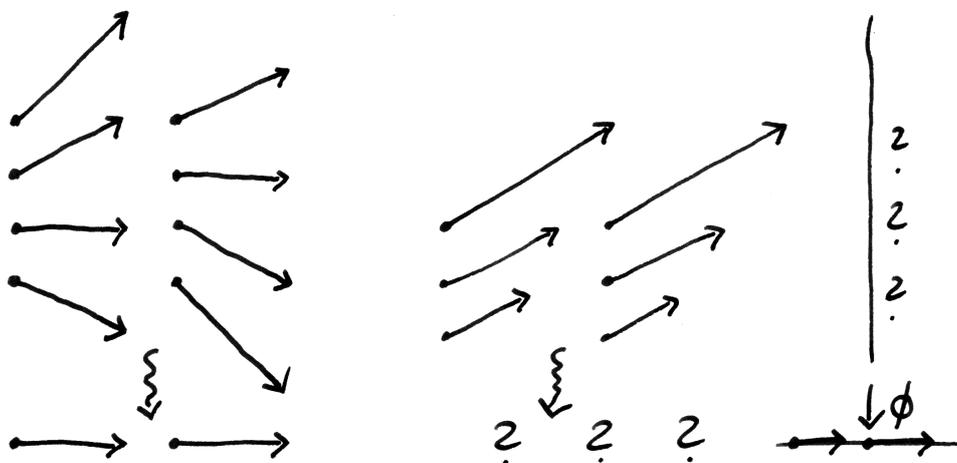
$$\sum c_i \partial_i$$

mit eindeutig bestimmten $c_i : U \rightarrow \mathbb{R}$. Paaren wir etwa das Vektorfeld $\sum c_i \partial_i$ auf \mathbb{R}^n mit dem Kovektorfeld $\sum a_i dx_i$, so ergibt sich die Funktion $\sum a_i c_i$. In unserer Notation 6.1.7 und mit dem Kroneckerdelta haben wir nämlich in der Tat

$$\langle dx_i, \partial_j \rangle = \delta_{ij}$$

6.1.15. Zumindest unter linearen Koordinatentransformationen verhalten sich Kovektorfelder “so wie Koordinaten”. Ist etwa x_1, \dots, x_n ein System linearer Koordinaten auf X im Sinne eines Systems von Abbildungen $x_i : X \rightarrow \mathbb{R}$, die zusammen einen Isomorphismus von affinen Räumen $X \rightarrow \mathbb{R}^n$ liefern, und ist y_1, \dots, y_n ein anderes System linearer Koordinaten, und haben wir etwa $y_i = \sum_j a_{ij} x_j$ für eine Matrix von reellen Zahlen a_{ij} , so gilt die Identität von Kovektorfeldern $dy_i = \sum_j a_{ij} dx_j$. Für die durch unsere Koordinatensysteme bestimmten Vektorfelder haben wir dahingegen umgekehrt

$$\frac{\partial}{\partial x_j} = \sum_i a_{ij} \frac{\partial}{\partial y_i}$$



Ganz links ist zuerst ein Vektorfeld auf der Ebene abgebildet, das unter der orthogonalen Projektion auf die x -Achse verwandt ist zu einem ebenfalls eingezeichneten konstanten Vektorfeld auf der x -Achse. In der Mitte dann ein Vektorfeld auf der Ebene, das unter dieser Projektion zu keinem Vektorfeld auf der x -Achse verwandt ist. Schließlich ganz rechts die konstante Abbildung der y -Achse auf einen Punkt der x -Achse und ein Vektorfeld auf der x -Achse, das darunter zu keinem Vektorfeld auf der y -Achse verwandt ist.

und benötigen die inverse Matrix, um $\frac{\partial}{\partial y_i}$ durch die $\frac{\partial}{\partial x_j}$ auszudrücken. In diesem Sinne “transformieren sich Kovektorfelder wie Koordinaten” und heißen deshalb auch “kovariant”, wohingegen Vektorfelder sich “vermittels der inversen transponierten Matrix transformieren” und deshalb “kontravariant” heißen.

Definition 6.1.16. Sei $\phi : U \rightarrow V$ eine differenzierbare Abbildung zwischen halboffenen Teilmengen endlichdimensionaler reeller Räume X und Y .

1. Zwei Vektorfelder $A : U \rightarrow \vec{X}$ und $B : V \rightarrow \vec{Y}$ heißen ϕ -**verwandt** und wir schreiben $\phi : A \rightsquigarrow B$ genau dann, wenn für alle $x \in U$ gilt

$$(\mathrm{d}_x \phi)(A_x) = B_{\phi(x)}$$

2. Zwei Kovektorfelder $\eta : U \rightarrow \vec{X}^*$ und $\omega : V \rightarrow \vec{Y}^*$ heißen ϕ -**verwandt** und wir schreiben $\phi : \eta \rightsquigarrow \omega$ genau dann, wenn für alle $x \in U$ gilt $\eta_x = \omega_{\phi(x)} \circ \mathrm{d}_x \phi$ oder gleichbedeutend mit der transponierten Abbildung aus ?? notiert

$$\eta_x = (\mathrm{d}_x \phi)^\top(\omega_{\phi(x)})$$

3. Zwei reelle Funktionen $g : U \rightarrow \mathbb{R}$ und $f : V \rightarrow \mathbb{R}$ heißen ϕ -**verwandt** und wir schreiben $\phi : g \rightsquigarrow f$ genau dann, wenn gilt $g = f \circ \phi$, als da heißt, wenn für alle $x \in U$ gilt

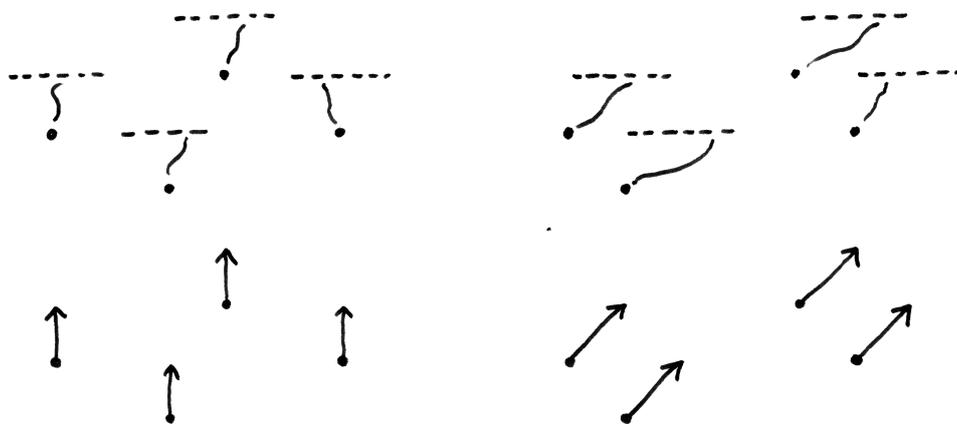
$$g(x) = f(\phi(x))$$

6.1.17. Unter einer differenzierbaren Bijektion mit differenzierbarer Umkehrabbildung haben alle Vektorfelder, Kovektorfelder und Funktionen jeweils genau einen Verwandten, und unter der Identität sind sie jeweils selbst dieser einzige Verwandte. Ist allgemeiner $\phi : U \rightarrow V$ differenzierbar aber sonst beliebig, so hat jedes Kovektorfeld ω auf V immer noch genau einen “Rückwärtsverwandten” auf U , der eben gegeben wird durch die Formel $\eta_x = (\mathrm{d}_x \phi)^\top(\omega_{\phi(x)})$ und der notiert wird als

$$\eta = \phi^*(\omega)$$

Er heißt das **mit ϕ zurückgezogene** oder **zurückgeholte Kovektorfeld**. Ebenso hat jede Funktion f auf V genau einen “Rückwärtsverwandten”, eben die Funktion $f \circ \phi$, die man auch die **mit ϕ zurückgezogene Funktion** nennt und manchmal $\phi^*(f)$ notiert. Bei Vektorfeldern liegen die Verhältnisse nicht so einfach, aber ist ϕ surjektiv, so hat jedes Vektorfeld auf U zumindest nicht mehr als einen “Vorwärtsverwandten” auf V .

6.1.18. Verwandtschaft ist transitiv. Sei genauer zusätzlich zu den obigen Daten Z ein endlichdimensionaler reeller Raum und $W \subset Z$ eine halboffene Teilmenge und $\psi : V \rightarrow W$ eine stetig differenzierbare Abbildung. Ist ein Vektorfeld C auf



Dieses Bild soll den Effekt der Scherung $\phi : \mathbb{R}^2 \xrightarrow{\sim} \mathbb{R}^2, (x, y) \mapsto (x + y, y)$ auf dem Kovektorfeld dy und dem Vektorfeld ∂_y darstellen. Bei der bildlichen Darstellung unseres Kovektorfelds folgen wir den auf Seite 84 im Anschluß an 6.1.2 eingeführten Konventionen. Man erkennt, daß dy unter dieser Scherung verwandt ist zu sich selber, in Formeln $\phi : dy \rightsquigarrow dy$, wohingegen ∂_y verwandt ist zu $\partial_x + \partial_y$, in Formeln $\phi : \partial_y \rightsquigarrow \partial_x + \partial_y$. Alternativ und im Wesentlichen gleichbedeutend mag man sich auch auf den Standpunkt stellen, daß wir auf dem Wertebereich von ϕ ein “verschertes Koordinatensystem” (u, v) eingeführt haben mit u und v den Komponenten der zu ϕ inversen Abbildung, also $u(x, y) = x - y$ und $v(x, y) = y$. Dann erhalten wir statt der obigen Verwandtschaften die Formeln $dv = dy$ sowie $\partial_v = \partial_x + \partial_y$.

W unter ψ verwandt zu B , so ist auch A unter $\psi \circ \phi$ verwandt zu C , in Formeln implizieren $\phi : A \rightsquigarrow B$ und $\psi : B \rightsquigarrow C$ also $\psi \circ \phi : A \rightsquigarrow C$. Analoges gilt für Funktionen und Kovektorfelder und läßt sich in den beiden letzteren Fällen auch schreiben als $(\psi \circ \phi)^* = \phi^* \circ \psi^*$, so daß etwa für jedes Kovektorfeld κ auf C gilt

$$(\psi \circ \phi)^*(\kappa) = \phi^*(\psi^*(\kappa))$$

Aus Gründen der formalen Vollständigkeit ergänze ich noch, daß unter der Identität natürlich jedes Vektorfeld und jedes Kovektorfeld und jede Funktion verwandt ist zu sich selber und nur zu sich selber.

6.1.19. Verwandte Funktionen haben verwandte Differentiale, in Formeln impliziert $\phi : g \rightsquigarrow f$ also $\phi : dg \rightsquigarrow df$, und gleichbedeutend haben wir für alle f die Identität $\phi^*(df) = d(\phi^*(f)) = d(f \circ \phi)$. In der Tat gilt für jeden Punkt y nach der Definition der Verwandtschaft und der Kettenregel

$$(\phi^*(df))_y = (d_{\phi(y)}f) \circ d_y\phi = d_y(f \circ \phi)$$

6.1.20. Wenden wir verwandte Vektorfelder auf verwandte differenzierbare Funktionen an, so erhalten wir wieder verwandte Funktionen, in Formeln folgt aus $\phi : A \rightsquigarrow B$ und $\phi : g \rightsquigarrow f$ also $\phi : Ag \rightsquigarrow Bf$ oder anders geschrieben $(Bf) \circ \phi = A(f \circ \phi)$. Das folgt direkt aus der Kettenregel in mehreren Veränderlichen. Letzteres ist sogar eine hinreichende Bedingung: Gilt $(Af) \circ \phi = B(f \circ \phi)$ für alle differenzierbaren Funktionen f , so folgt $\phi : A \rightsquigarrow B$.

Beispiel 6.1.21. Für $X = \mathbb{R}^n$ mit Koordinaten x_1, \dots, x_n und $Y = \mathbb{R}^m$ mit Koordinaten y_1, \dots, y_m und $\phi = (\phi_1, \dots, \phi_n)$ eine differenzierbare Abbildung von einer halboffenen Teilmenge von \mathbb{R}^m in eine halboffene Teilmenge von \mathbb{R}^n ergibt sich $\phi^*(dx_i) = d(\phi^*x_i) = d\phi_i = \sum_j \frac{\partial \phi_i}{\partial y_j} dy_j$ unter Verwendung von 6.1.19 und 6.1.10. Folglich kann das Zurückholen von Kovektorfeldern in Koordinaten beschrieben werden durch die Formel

$$\phi^* \left(\sum_i a_i dx_i \right) = \sum_{i,j} (a_i \circ \phi) \frac{\partial \phi_i}{\partial y_j} dy_j$$

Übung 6.1.22. Verwandtschaft ist verträglich mit dem Bilden von Produkten von Funktionen. Verwandtschaft ist verträglich mit dem Bilden des Produkts von Funktionen und Vektorfeldern, in Formeln folgt aus Verwandtschaften $\phi : g \rightsquigarrow f$ und $\phi : A \rightsquigarrow B$ also $\phi : gA \rightsquigarrow fB$ oder anders geschrieben folgt aus $\phi : A \rightsquigarrow B$ bereits $\phi : (f \circ \phi)A \rightsquigarrow fB$. Verwandtschaft ist verträglich mit dem Bilden des Produkts von Funktionen und Kovektorfeldern, in Formeln folgt aus $\phi : g \rightsquigarrow f$ und $\phi : \eta \rightsquigarrow \omega$ also $\phi : g\eta \rightsquigarrow f\omega$ oder anders geschrieben gilt $\phi^*(f\omega) = (f \circ \phi)\phi^*\omega$. Schließlich ist Verwandtschaft auch verträglich mit dem Auswerten von Kovektorfeldern auf Vektorfeldern, in Formeln folgt aus $\phi : \eta \rightsquigarrow \omega$ und $\phi : A \rightsquigarrow B$

also $\phi : \langle \eta, A \rangle \rightsquigarrow \langle \omega, B \rangle$ alias aus $\phi : A \rightsquigarrow B$ folgt $\langle \phi^* \omega, A \rangle = \langle \omega, B \rangle \circ \phi$. Das ist sogar eine hinreichende Bedingung: Gilt $\langle \phi^* \omega, A \rangle = \langle \omega, B \rangle \circ \phi$ für alle Kovektorfelder ω , so folgt $\phi : A \rightsquigarrow B$.

Beispiel 6.1.23. Wir betrachten die Polarkoordinatenabbildung

$$P : \mathbb{R}^2 \rightarrow \mathbb{R}^2 \\ (r, \vartheta) \mapsto (r \cos \vartheta, r \sin \vartheta)$$

und benutzen die üblichen Koordinaten x, y auf dem Wertebereich. Unter dieser Abbildung ist etwa das Kovektorfeld dx rechts verwandt zum Kovektorfeld $d(r \cos \vartheta) = (\cos \vartheta) dr - (r \sin \vartheta) d\vartheta$ links. Ebenso ist das Kovektorfeld dy rechts verwandt zum Kovektorfeld $d(r \sin \vartheta) = (\sin \vartheta) dr + (r \cos \vartheta) d\vartheta$ links. Um einen Verwandten für ∂_ϑ zu suchen, wenn dieses Vektorfeld denn einen Verwandten haben sollte, machen wir den Ansatz $\partial_\vartheta \rightsquigarrow a\partial_x + b\partial_y$ mit unbestimmten Funktionen a, b und finden durch Paaren mit dx leicht $-(r \sin \vartheta) \rightsquigarrow a$ und durch Paaren mit dy ebenso $(r \cos \vartheta) \rightsquigarrow b$, womit wir für das Vektorfeld ∂_ϑ links als einzigen Verwandten das Vektorfeld $-y\partial_x + x\partial_y$ rechts finden. Das Vektorfeld ∂_r links hat keinen Verwandten rechts, denn derselbe Ansatz $\partial_r \rightsquigarrow a\partial_x + b\partial_y$ führt zu $P : \sin \vartheta \rightsquigarrow a$ und $P : \cos \vartheta \rightsquigarrow b$, und derartige Funktionen a, b gibt es nicht. Schränken wir jedoch unsere Polarkoordinatenabbildung ein zu einer Abbildung $P : \{(r, \vartheta) \mid r > 0\} \rightarrow \mathbb{R}^2 \setminus 0$, so gibt es derartige Funktionen doch und unser Vektorfeld ∂_r hat unter dieser Einschränkung den einzigen Verwandten

$$\left(\frac{x}{\sqrt{x^2 + y^2}} \right) \partial_x + \left(\frac{y}{\sqrt{x^2 + y^2}} \right) \partial_y$$

Meist wird man mit diesen Begriffen etwas großzügiger umgehen, zwischen verwandte Objekte schlicht ein Gleichheitszeichen schreiben und es auch mit den Definitionsbereichen nicht so genau nehmen, so daß wir etwa schreiben würden

$$\begin{aligned} \partial_r &= (\cos \vartheta) \partial_x + (\sin \vartheta) \partial_y = \left(\frac{x}{\sqrt{x^2 + y^2}} \right) \partial_x + \left(\frac{y}{\sqrt{x^2 + y^2}} \right) \partial_y \\ \partial_\vartheta &= -(r \sin \vartheta) \partial_x + (r \cos \vartheta) \partial_y = -y \partial_x + x \partial_y \\ \partial_x &= (\cos \vartheta) \partial_r - (r^{-1} \sin \vartheta) \partial_\vartheta \\ \partial_y &= (\sin \vartheta) \partial_r + (r^{-1} \cos \vartheta) \partial_\vartheta \\ dx &= (\cos \vartheta) dr - (r \sin \vartheta) d\vartheta \\ dy &= (\sin \vartheta) dr + (r \cos \vartheta) d\vartheta \\ d\vartheta &= \left(-y/(x^2 + y^2) \right) dx + \left(x/(x^2 + y^2) \right) dy \\ dr &= \left(x/\sqrt{x^2 + y^2} \right) dx + \left(y/\sqrt{x^2 + y^2} \right) dy \end{aligned}$$

Man kann die unteren Formeln auch so verstehen, daß eben dr das Differential der Funktion $r : (\mathbb{R}^2 \setminus 0) \rightarrow \mathbb{R}, (x, y) \mapsto r(x, y)$ meint. Bei $d\vartheta$ wird es schon kritischer, da ja eigentlich ϑ nur auf geschlitzten Ebenen definiert werden kann. Allerdings unterscheiden sich die auf verschiedenen geschlitzten Ebenen definierten ϑ dann wieder nur um additive Konstanten, so daß sie alle dasselbe Differential haben und wir doch ein wohldefiniertes Kovektorfeld $d\vartheta$ auf $\mathbb{R}^2 \setminus 0$ erhalten. Das ist auch der tiefere Grund dafür, daß alle unsere Standardvektorfelder in diesem Fall wohldefinierte Verwandte haben und wir mit unseren Gleichheitszeichen nicht in Teufels Küche kommen. Bei komplizierteren Vektorfeldern sähe das anders aus: So hat etwa das Vektorfeld $\vartheta\partial_\vartheta$ gar keinen Verwandten, es sei denn, wir schränken unsere Polarkoordinatenabbildung noch weiter ein.

6.1.24. Gegeben ein endlichdimensionaler affiner Raum X und eine offene Teilmenge $U \subseteq X$ und ein Diffeomorphismus alias ein System lokaler Koordinaten $(x_1, \dots, x_n) : U \xrightarrow{\sim} V \subseteq \mathbb{R}^n$ bezeichnet man gerne mit $\frac{\partial}{\partial x_i}$ oder auch mit ∂_i diejenigen Vektorfelder auf U , die unter diesem Diffeomorphismus zu den eben eingeführten Vektorfeldern auf \mathbb{R}^n verwandt sind. Man beachte jedoch, daß für eine einzelne Funktion $x : U \rightarrow \mathbb{R}$ nicht sinnvoll ein Vektorfeld $\frac{\partial}{\partial x}$ auf U erklärt werden kann: Selbst wenn sich unsere Funktion zu einem Koordinatensystem ergänzen lassen sollte, wird doch das durch diese Ergänzung erklärte Vektorfeld $\frac{\partial}{\partial x}$ wesentlich von der Wahl der anderen Koordinaten abhängen. All das steht im Gegensatz zum Differential dx einer Funktion x , das durchaus auch für eine einzelne Funktion sinnvoll definiert ist.

Übung 6.1.25. Unter der Inversion am Einheitskreis $\mathbb{R}^2 \setminus 0 \xrightarrow{\sim} \mathbb{R}^2 \setminus 0, (x, y) \mapsto (u, v) = (x^2 + y^2)^{-1}(x, y)$ zeige man die Verwandtschaft von Vektorfeldern

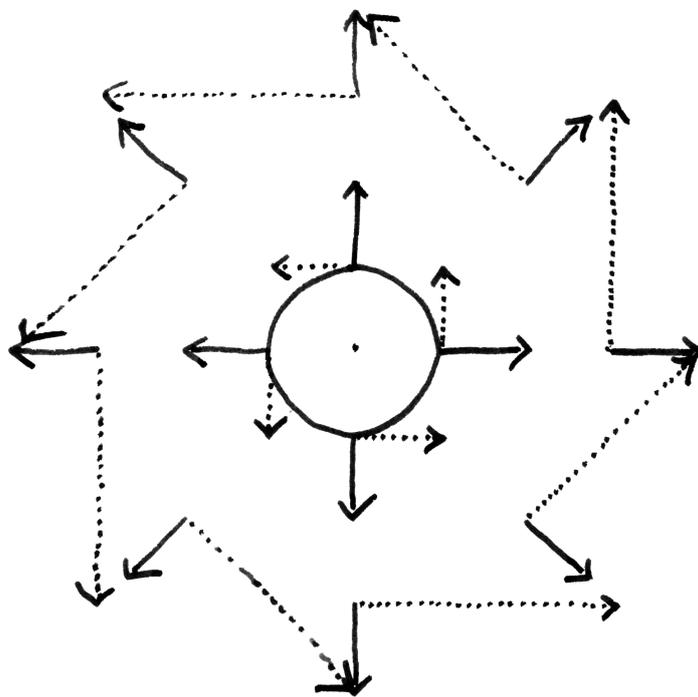
$$\begin{aligned}\partial_x &\rightsquigarrow (v^2 - u^2)\partial_u - 2uv\partial_v \\ \partial_y &\rightsquigarrow (u^2 - v^2)\partial_v - 2uv\partial_u\end{aligned}$$

6.2 Gradienten in krummlinigen Koordinaten*

6.2.1. Gegeben eine offene Teilmenge $U \subseteq \mathbb{R}^n$ und eine hinreichend differenzierbare Funktion $f : U \rightarrow \mathbb{R}$ definiert man wie in 4.1.5 ihren **Gradienten** als das Vektorfeld

$$\text{grad } f = \frac{\partial f}{\partial x_1}\partial_1 + \dots + \frac{\partial f}{\partial x_n}\partial_n$$

auf U . Ich will im Folgenden diskutieren, welche Form dieses Konstrukt in krummlinigen Koordinaten annimmt. Formal ist damit folgendes gemeint: Man betrachte zusätzlich $V \subseteq \mathbb{R}^n$ und einen Diffeomorphismus $\phi : V \xrightarrow{\sim} U$ und berechne aus $(f \circ \phi)$ das unter ϕ zu $(\text{grad } f)$ verwandte Vektorfeld auf V . In der Notation wird vielfach ϕ einfach weggelassen und nur die Bezeichnungen der Koordinaten deuten das Gemeinte an. Ist etwa $\phi = P : \mathbb{R}_{>0} \times \mathbb{R} \rightarrow \mathbb{R}^2 \setminus 0$ wie in 6.1.23



Einige Werte des Vektorfelds ∂_r als durchgezogene Pfeile und des Vektorfeld ∂_θ als gepunktete Pfeile, gezeichnet in der xy -Ebene.

die Polarkoordinatenabbildung, so erhalten wir mit den Formeln aus 6.1.23 sofort $f_x = \cos \vartheta f_r - r^{-1} \sin \vartheta f_\vartheta$ und $f_y = \sin \vartheta f_r + r^{-1} \cos \vartheta f_\vartheta$ und nach kurzer Rechnung die Verwandtschaft von Vektorfeldern

$$\text{grad } f = f_x \partial_x + f_y \partial_y = f_r \partial_r + \frac{1}{r^2} f_\vartheta \partial_\vartheta$$

Man nennt sie die **Darstellung des Gradienten in Polarkoordinaten**. Hier haben wir die Notation $f_x = \frac{\partial f}{\partial x}$ für die entsprechende partielle Ableitung aus 4.1.1 und die Abkürzung $\partial_x = \frac{\partial}{\partial x}$ aus 6.1.14 für den besagten Differentialoperator alias besagtes Vektorfeld verwendet. Bereits bei der Transformation des Gradienten in Kugelkoordinaten wird die Rechnung jedoch recht aufwendig. Ich will im folgenden erklären, mit welchen Kunstgriffen man sie strukturieren und übersichtlicher gestalten kann.

6.2.2. Gegeben ein Vektorraum V über einem Körper k notieren wir

$$\text{Bil}(V) = \text{Bil}_k(V)$$

den Vektorraum aller bilinearen Abbildungen $V \times V \rightarrow k$. Gegeben Linearformen $\lambda, \eta : V \rightarrow k$ notieren wir $(\lambda \otimes \eta) \in \text{Bil}(V)$ die bilineare Abbildung $(v, w) \mapsto \lambda(v)\eta(w)$. Sicher ist $(\lambda, \eta) \mapsto \lambda \otimes \eta$ selbst eine bilineare Abbildung $V^* \times V^* \rightarrow \text{Bil}(V)$. Statt $\eta \otimes \eta$ schreibt man meist kürzer $\eta^{\otimes 2}$. Das Symbol \otimes wird in ?? noch mit mehr Bedeutung aufgeladen. Hier darf und soll es ausschließlich als bequeme Notation verstanden werden.

Definition 6.2.3. Sei X ein endlichdimensionaler reeller Raum und $U \subseteq X$ eine offene Teilmenge. Ein **2-Tensor auf U** ist eine Abbildung

$$g : U \rightarrow \text{Bil}(\vec{X})$$

von U in den Raum $\text{Bil}(\vec{X})$ aller bilinearen Abbildungen $\vec{X} \times \vec{X} \rightarrow \mathbb{R}$. Eine **Riemann'sche Metrik auf U** ist ein 2-Tensor g , der jedem Punkt $p \in U$ ein Skalarprodukt g_p auf \vec{X} zuordnet.

Beispiel 6.2.4. Das Standardskalarprodukt auf \mathbb{R}^n liefert eine Riemann'sche Metrik auf \mathbb{R}^n und auf jeder offenen Teilmenge $U \subseteq \mathbb{R}^n$. Gegeben Kovektorfelder ω und η auf einer offenen Teilmenge U eines endlichdimensionalen reellen Raums X können wir den 2-Tensor

$$\begin{aligned} \omega \otimes \eta : U &\rightarrow \text{Bil}(\vec{X}) \\ p &\mapsto \omega_p \otimes \eta_p \end{aligned}$$

betrachten. Weiter können wir 2-Tensoren punktweise addieren und mit Funktionen multiplizieren. Die übliche Riemann'sche Metrik auf \mathbb{R}^n kann in diesen

Konventionen geschrieben werden als $s = dx_1^{\otimes 2} + \dots + dx_n^{\otimes 2}$. Eine beliebige Riemann'sche Metrik g auf einer offenen Teilmenge $U \subseteq \mathbb{R}^n$ hat in diesen Notationen die Gestalt

$$\sum_{i,j=1}^n g_{ij} dx_i \otimes dx_j$$

für Funktionen $g_{ij} : U \rightarrow \mathbb{R}$, die an jedem Punkt $p \in U$ eine positiv definite symmetrische Matrix bilden.

6.2.5. Gegeben ein Vektorraum V über einem Körper k liefert jede Bilinearform $g \in \text{Bil}(V)$ eine Abbildung

$$\begin{aligned} \text{can}_g : V &\rightarrow V^* \\ v &\mapsto (w \mapsto g(v, w)) \end{aligned}$$

von unserem Vektorraum in seinen Dualraum, die jedem Vektor v die Linearform "Paare mit v " zuordnet. Zum Beispiel hätten wir $\text{can}_{\lambda \otimes \eta}(v) = \lambda(v)\eta$. Gleichberechtigt könnten wir auch die Abbildung $\text{can}_g^2 : v \mapsto (w \mapsto g(w, v))$ betrachten und müßten dann, um dieser Gleichberechtigung Ausdruck zu verleihen, eigentlich genauer $\text{can}_g = \text{can}_g^1$ schreiben. Das würde jedoch die Notation schwerfälliger machen, und ich denke, diese zusätzliche Schwere wiegt den Gewinn an Klarheit nicht auf. An dieser Stelle möchte ich auch allen Leserinnen versichern, daß sie ganz genauso gemeint sind, wenn einmal von "dem Leser" die Rede ist. Ich denke auch in diesem Zusammenhang, daß die zusätzliche Schwere der geschlechtsneutralen Formulierungen den Gewinn an Klarheit nicht aufwiegt. Ist speziell g nichtausgeartet und V endlichdimensional, so ist can_g ein Isomorphismus $\text{can}_g : V \xrightarrow{\sim} V^*$ und wir können auch sein Inverses $\text{can}_g^{-1} : V^* \xrightarrow{\sim} V$ betrachten.

6.2.6. Gegeben ein Vektorfeld A und ein 2-Tensor g können wir ebenso ein Kovektorfeld $\text{can}_g(A)$ bilden durch das Einsetzen von A in die erste Stelle von g . Ist unser 2-Tensor g an keiner Stelle ausgeartet, insbesondere also im Fall einer Riemann'schen Metrik, so ist diese Abbildung eine Bijektion

$$\text{can}_g : \{\text{Vektorfelder}\} \xrightarrow{\sim} \{\text{Kovektorfelder}\}$$

Bezeichnet speziell s das Standardskalarprodukt auf dem \mathbb{R}^n , so haben wir $\text{can}_s(a\partial_i) = a dx_i$ für jede Funktion a und für unseren Gradienten aus 4.1.5 gilt folglich

$$\text{grad } f = \text{can}_s^{-1}(df)$$

Im allgemeinen verwendet man die Notation $\text{grad}_g f := \text{can}_g^{-1}(df)$ und nennt dieses Vektorfeld den **Gradienten von f** in Bezug auf die Riemann'sche Metrik g oder allgemeiner in Bezug auf den nichtausgearteten 2-Tensor g .

Definition 6.2.7. Seien $U \subseteq X, V \subseteq Y$ offene Teilmengen endlichdimensionaler reeller Räume und $\phi : U \rightarrow V$ stetig differenzierbar. Vorgegebene 2-Tensoren s auf U und g auf V heißen **ϕ -verwandt** und wir schreiben $\phi : s \rightsquigarrow g$ genau dann, wenn für alle $x \in U$ und $v, w \in \vec{X}$ gilt

$$s_x(v, w) = g_{\phi(x)}((d_x\phi)(v), (d_x\phi)(w))$$

6.2.8. Wieder ist Verwandtschaft verträglich mit allen natürlichen Operationen, etwa mit dem Einsetzen von Vektorfeldern, dem Multiplizieren mit Funktionen, unserer Konstruktion \otimes etc. Insbesondere haben verwandte Funktionen unter verwandten Riemann'schen Metriken verwandte Gradienten, in Formeln impliziert $\phi : s \rightsquigarrow g$ also

$$\phi : \text{grad}_s(f \circ \phi) \rightsquigarrow \text{grad}_g f$$

Offensichtlich hat jeder 2-Tensor g auf V genau einen Verwandten auf U , den wir mit ϕ^*g bezeichnen und den **zurückgeholtten 2-Tensor** nennen. Gegeben eine parametrisierte Fläche im Raum oder allgemeiner eine differenzierbare Abbildung $\phi : U \rightarrow \mathbb{R}^3$ mit $U \subseteq \mathbb{R}^2$ bezeichnet man den symmetrischen 2-Tensor auf \mathbb{R}^2 , der durch das Zurückholen der Standardmetrik entsteht, auch als die **erste Fundamentalform** unserer parametrisierten Fläche.

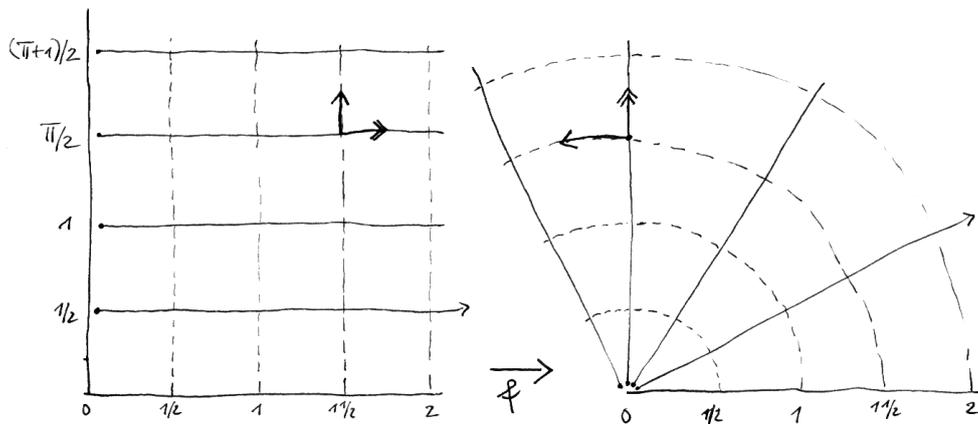
Beispiel 6.2.9. Unter der Polarkoordinatenabbildung P aus 6.1.23 ist die Standardmetrik $s = dx^{\otimes 2} + dy^{\otimes 2}$ auf der xy -Ebene verwandt zum 2-Tensor

$$\begin{aligned} g &= (\cos \vartheta dr - r \sin \vartheta d\vartheta) \otimes (\cos \vartheta dr - r \sin \vartheta d\vartheta) \\ &\quad + (\sin \vartheta dr + r \cos \vartheta d\vartheta) \otimes (\sin \vartheta dr + r \cos \vartheta d\vartheta) \\ &= dr^{\otimes 2} + r^2 d\vartheta^{\otimes 2} \end{aligned}$$

auf der $r\vartheta$ -Ebene, der auf dem Komplement der Nullstellenmenge von r auch wieder eine Riemann'sche Metrik ist. Daß hier keine gemischten Tensoren $dr \otimes d\vartheta$ auftreten, hat den Grund, daß die Vektorfelder ∂_r und ∂_ϑ auch in der xy -Ebene an jedem Punkt aufeinander senkrecht stehen. Die Koeffizienten 1 und r^2 bedeuten gerade die quadrierten Längen $s(\partial_r, \partial_r)$ und $s(\partial_\vartheta, \partial_\vartheta)$ der Vektoren dieser Vektorfelder. Für eine Funktion $f = f(x, y)$ muß schließlich df unter P verwandt sein zu $d(f \circ P)$, und dann muß auch $\text{grad} f = \text{grad}_s f = \text{can}_s^{-1}(df)$ verwandt sein zu

$$\text{grad}_g(f \circ P) = \text{can}_g^{-1} d(f \circ P) = \text{can}_g^{-1}(f_r dr + f_\vartheta d\vartheta) = f_r \partial_r + \frac{1}{r^2} f_\vartheta \partial_\vartheta$$

Damit haben wir die Darstellung des Gradienten in Polarkoordinaten ein weiteres Mal hergeleitet.



Dies Bild soll die Verwandtschaft von Riemann'schen Metriken $f : dr^{\otimes 2} + r^2 d\vartheta^{\otimes 2} \rightsquigarrow dx^{\otimes 2} + dy^{\otimes 2}$ verdeutlichen, mit $f = P$ der Polarkoordinatenabbildung. Das Differential an der Stelle $(r, \vartheta) = (1\frac{1}{2}, \frac{\pi}{2})$ ist dargestellt durch seinen Effekt auf der Standardbasis, die wir auch $(\partial_r, \partial_\vartheta)$ notieren könnten. Die Standardbasis geht an jeder Stelle über in eine Orthogonalbasis und das Bild des ersten Basisvektors hat auch wieder die Länge Eins, das Bild des zweiten Basisvektors jedoch im allgemeinen die Länge r und in unserem Fall die Länge $1\frac{1}{2}$. Die Standardmetrik auf der xy -Ebene entspricht folglich einer Metrik auf der $r\vartheta$ -Ebene, bei der ∂_r und ∂_ϑ aufeinander senkrecht stehen und ∂_r die Länge Eins hat, wohingegen ∂_ϑ die Länge r hat. Diese Eigenschaften aber charakterisieren genau unsere Metrik $dr^{\otimes 2} + r^2 d\vartheta^{\otimes 2}$.

Ergänzung 6.2.10. Ingenieure arbeiten gerne mit einer anderen Darstellung von Vektorfeldern und betrachten etwa auf dem \mathbb{R}^2 die auf euklidische Länge Eins normierten Vektorfelder $\mathbf{e}_r = \partial_r$ und $\mathbf{e}_\vartheta = r^{-1}\partial_\vartheta$. Natürlich kann jedes Vektorfeld v auf dem Komplement des Ursprungs auch als $v = a\mathbf{e}_r + b\mathbf{e}_\vartheta$ geschrieben werden mit geeigneten reellwertigen Funktionen a, b . In Formelsammlungen findet man häufig Formeln für Gradienten und dergleichen in dieser Darstellung, zum Beispiel hätten wir $\text{grad } f = (\partial_r f)\mathbf{e}_r + r^{-1}(\partial_\vartheta f)\mathbf{e}_\vartheta$. Meist heißen die Koeffizienten eines Vektorfelds $v = a\mathbf{e}_r + b\mathbf{e}_\vartheta$ dann auch noch $a = v_r, b = v_\vartheta$. Das verbietet sich für uns jedoch, da wir die Indexnotation bereits als Kürzel für partielle Ableitungen verwenden.

Beispiel 6.2.11. Die **Kugelkoordinaten** im Raum werden beschrieben durch eine geeignete Einschränkung der Abbildung

$$K : \quad \mathbb{R}^3 \quad \rightarrow \quad \mathbb{R}^3 \\ (r, \vartheta, \varphi) \mapsto (r \cos \varphi \sin \vartheta, r \sin \varphi \sin \vartheta, r \cos \vartheta)$$

deren anschauliche Bedeutung in nebenstehendem Bild erläutert wird.

Übung 6.2.12. Man zeige, daß die Standardmetrik im xyz -Raum unter Kugelkoordinaten, wie sie 6.2.11 eingeführt werden, verwandt ist zur Metrik

$$g = dr^{\otimes 2} + r^2 d\vartheta^{\otimes 2} + (r \sin \vartheta)^2 d\varphi^{\otimes 2}$$

Übung 6.2.13. Man zeige, daß der Gradient in Kugelkoordinaten, wie sie 6.2.11 eingeführt werden, ausgedrückt wird durch die Formel

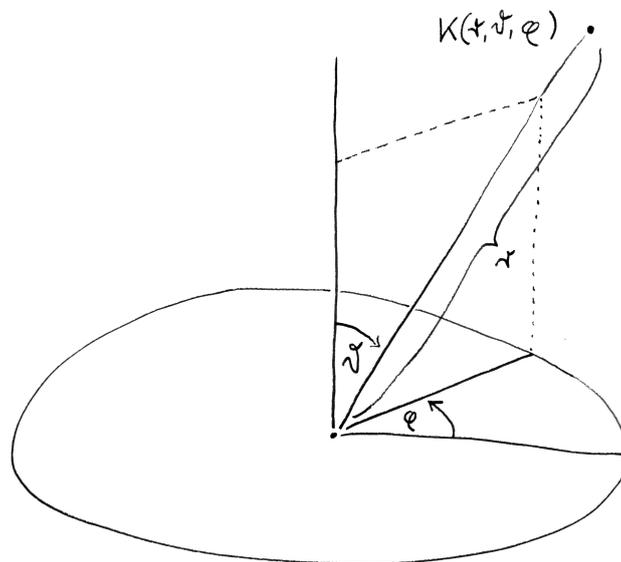
$$\text{grad } f = f_r \partial_r + r^{-2} f_\vartheta \partial_\vartheta + (r \sin \vartheta)^{-2} f_\varphi \partial_\varphi$$

Ergänzung 6.2.14. Ein 2-Tensor heißt **symmetrisch** bzw. **antisymmetrisch** genau dann, wenn er an jedem Punkt als Wert eine symmetrische bzw. antisymmetrische Bilinearform auf \vec{X} annimmt. Antisymmetrische 2-Tensoren werden wir später als sogenannte 2-Formen wiedertreffen. Eine Riemann'sche Metrik ist per definitionem ein symmetrischer 2-Tensor mit der zusätzlichen Eigenschaft, positiv definit zu sein.

6.3 Wegintegrale

Definition 6.3.1. Gegeben $\gamma : [a, b] \rightarrow A$ ein stetig differenzierbarer Weg in einer halboffenen Teilmenge A eines endlichdimensionalen reellen Raums X und ein stetiges Kovektorfeld $\omega : A \rightarrow \vec{X}^*$ auf A definieren wir eine reelle Zahl $\int_\gamma \omega$, das **Integral des Kovektorfelds ω längs des Weges γ** , durch die Vorschrift

$$\int_\gamma \omega := \int_a^b \omega_{\gamma(t)}(\gamma'(t)) dt$$



Die Kugelkoordinatenabbildung hat die folgende anschauliche Bedeutung:
 Stellen wir uns ein Teleskop vor, das im Ursprung eines kartesischen Koordinatensystems auf einem waagerechten, d.h. in der xy -Ebene liegenden Drehteller steht und senkrecht nach oben zeigt. Um einen Stern zu betrachten, schwenken wir zunächst das Teleskop nach unten in Richtung der positiven x -Achse um einen Winkel $\vartheta \in [0, \pi]$ und drehen dann den Drehteller um einen geeigneten Winkel, sagen wir um den Winkel $\varphi \in [0, 2\pi)$ gegen den Uhrzeigersinn. Ist schließlich r die Entfernung unseres Sterns, so gibt $K(r, \vartheta, \varphi)$ seine kartesischen Koordinaten an. Natürlich ist im Fall eines senkrecht über oder unter dem Teleskop befindlichen Sterns φ nicht eindeutig, und befindet sich das Teleskop bereits im Stern, so sind beide Winkel nicht eindeutig. Die Einschränkung unserer Abbildung auf $r > 0$, $\vartheta \in (0, \pi)$ und $\varphi \in [0, 2\pi)$ hinwiederum ist zwar injektiv, aber nicht surjektiv. Oft findet man auch eine Variante, bei der wir uns das Teleskop zu Beginn horizontal in Richtung der positiven x -Achse ausgerichtet denken und wo die zweite Koordinate $\theta \in [-\pi/2, \pi/2]$ den Winkel bezeichnet, um den das Teleskop nach oben bzw. bei negativem Winkel nach unten geschwenkt werden muß. Die Formeln lauten dann abweichend $(r, \theta, \varphi) \mapsto (r \cos \varphi \cos \theta, r \sin \varphi \cos \theta, r \sin \theta)$.

In der physikalisch motivierten Terminologie nach 10.2.2 gilt es also, zu jedem Zeitpunkt t den Kovektor $\omega_{\gamma(t)}$ auf dem Geschwindigkeitsvektor $\gamma'(t)$ auszuwerten und die so entstehende reellwertige Funktion zu integrieren.

6.3.2 (Wegintegrale über Geradensegmente). Gegeben ein stetiges Kovektorfeld $\omega : A \rightarrow \tilde{X}^*$ auf einer halboffenen Teilmenge A eines endlichdimensionalen reellen Raums X und Punkte $p, q \in A$, bei denen das ganze verbindende Geradensegment $[p, q]$ in A liegt, vereinbaren wir für das Integral unseres Kovektorfelds über den geraden Weg $\gamma : [0, 1] \rightarrow A, t \mapsto p + t(q - p)$ von p nach q die Notation

$$\int_p^q \omega := \int_{\gamma} \omega$$

Für ein stetiges Kovektorfeld $\omega = f(t) dt$ auf einem reellen Intervall $A \subset \mathbb{R}$ und $p, q \in A$ erhalten wir dann insbesondere

$$\int_p^q \omega = \int_0^1 f(p + t(q - p)) (q - p) dt = \int_p^q f(t) dt$$

so daß unsere neue Notation links sich als eine Verallgemeinerung unserer schon vor langem eingeführten Notation rechts erweist. Weiter erkennen wir, daß in diesem Kalkül das Differential dt der Identität t auf \mathbb{R} in einem ideellen Sinn dasselbe bedeutet wie das dt , das wir bisher bei Integralen oft hinzugefügt hatten, um klar zu machen, über welche Variable integriert werden soll. Ist $A \subset \mathbb{R}$ ein halboffenes Intervall und $f : A \rightarrow \mathbb{R}$ eine stetig differenzierbare Funktion, so erhalten wir mit unserer Formel $df = f'(t) dt$ und dem Hauptsatz der Differential- und Integralrechnung für beliebige $a, b \in A$ insbesondere

$$\int_a^b df = \int_a^b f'(t) dt = f(b) - f(a)$$

Beispiel 6.3.3. Für einen beliebigen stetig differenzierbaren Weg γ und ein beliebiges Kovektorfeld ω hat nun das mit dem Weg zurückgeholte Kovektorfeld gerade die Gestalt $\gamma^* \omega = \omega_{\gamma(t)} (\gamma'(t)) dt$, so daß wir mit unserer Notation aus 6.3.2 die Definition des Wegintegrals umschreiben können zur Identität

$$\int_{\gamma} \omega = \int_a^b \gamma^* \omega$$

Diese Darstellung ist für explizite Rechnungen besonders praktisch. Integrieren wir etwa das Kovektorfeld $\omega = x dx + x^4 dy$ auf dem \mathbb{R}^2 über den Weg $\gamma :$

$[1, 2] \rightarrow \mathbb{R}^2$ gegeben durch $\gamma(t) = (\sqrt{t}, \log t)$, so erhalten wir

$$\begin{aligned} \int_{\gamma} \omega &= \int_{\gamma} x \, dx + x^4 \, dy \\ &= \int_1^2 \sqrt{t} \, d(\sqrt{t}) + (\sqrt{t})^4 \, d(\log t) \\ &= \int_1^2 \sqrt{t} \frac{1}{2\sqrt{t}} \, dt + t^2 t^{-1} \, dt \\ &= \int_1^2 \left(\frac{1}{2} + t\right) \, dt = 2 \end{aligned}$$

Lemma 6.3.4 (Anschauung für das Wegintegral). Sei $\gamma : [a, b] \rightarrow A$ ein stetig differenzierbarer Weg in einer halboffenen Teilmenge A eines endlichdimensionalen reellen Raums X und sei $\omega : A \rightarrow \vec{X}^*$ ein stetiges relatives Kovektorfeld auf A . Betrachtet man in der Situation der Definition des Wegintegrals 6.3.1 für alle $r \geq 1$ die äquidistanten Unterteilungen $a = a_0 \leq a_1 \leq \dots \leq a_r = b$ und bildet die Riemannsummen

$$S_{\gamma}^r(\omega) = \sum_{i=1}^r \omega_{\gamma(a_i)} (\gamma(a_i) - \gamma(a_{i-1}))$$

so ist unser Wegintegral der Grenzwert der Folge von Riemannsummen

$$\int_{\gamma} \omega = \lim_{r \rightarrow \infty} S_{\gamma}^r(\omega)$$

Beweis. Sei $\|\cdot\|$ eine Norm auf dem Richtungsraum \vec{X} und bezeichne $\|\cdot\|$ auch die zugehörige Operatornorm auf \vec{X}^* . Nach 6.5.5 ist unser Integral der Grenzwert der Folge von Riemannsummen

$$S^r = \sum_{i=1}^r \omega_{\gamma(t_i)} (\gamma'(t_i)) \cdot (t_i - t_{i-1})$$

Gegeben $\varepsilon > 0$ finden wir wegen der gleichmäßigen Stetigkeit von γ' ein $\delta > 0$ derart, daß gilt $|s - t| < \delta \Rightarrow \|\gamma'(s) - \gamma'(t)\| < \varepsilon$. Ist r so groß, daß die Länge der Intervalle $t_i - t_{i-1}$ unter δ sinkt, so folgt mit dem Mittelwertsatz 10.2.11 in mehreren Veränderlichen $\gamma(t_i) - \gamma(t_{i-1}) \in (t_i - t_{i-1})B(\gamma'(t_i); \varepsilon)$, was wir umschreiben können zu

$$\|\gamma(t_i) - \gamma(t_{i-1}) - (t_i - t_{i-1})\gamma'(t_i)\| < (t_i - t_{i-1})\varepsilon$$

Das hinwiederum liefert für r oberhalb einer geeigneten Schranke die Abschätzung

$$|S_{\gamma}^r(\omega) - S^r| \leq \sum_{i=1}^r \|\omega_{\gamma(t_i)}\| (t_i - t_{i-1})\varepsilon \leq \left(\sup_{t \in [a, b]} \|\omega_{\gamma(t)}\|\right) (b - a)\varepsilon$$

Diese Differenz strebt also gegen Null für $r \rightarrow \infty$, folglich strebt die Folge $S_{\gamma}^r(\omega)$ gegen denselben Grenzwert wie die Folge S^r . \square

Beispiel 6.3.5. Ein Kovektorfeld auf der Zeitachse kann, wie in 6.1.5 erklärt, nach der Wahl eines ausgezeichneten Drehsinns als eine Vorschrift aufgefaßt werden, die jedem Zeitpunkt eine Drehgeschwindigkeit zuordnet. Das Integral eines derartigen Kovektorfelds über einen Weg in der Zeitachse liefert dann anschaulich gesprochen die Zahl der Umdrehungen in Richtung des ausgezeichneten Drehsinns zwischen Anfangszeitpunkt und Endzeitpunkt. Liegt der Endzeitpunkt hier vor dem Anfangszeitpunkt, so ist entsprechend das Negative zu nehmen.

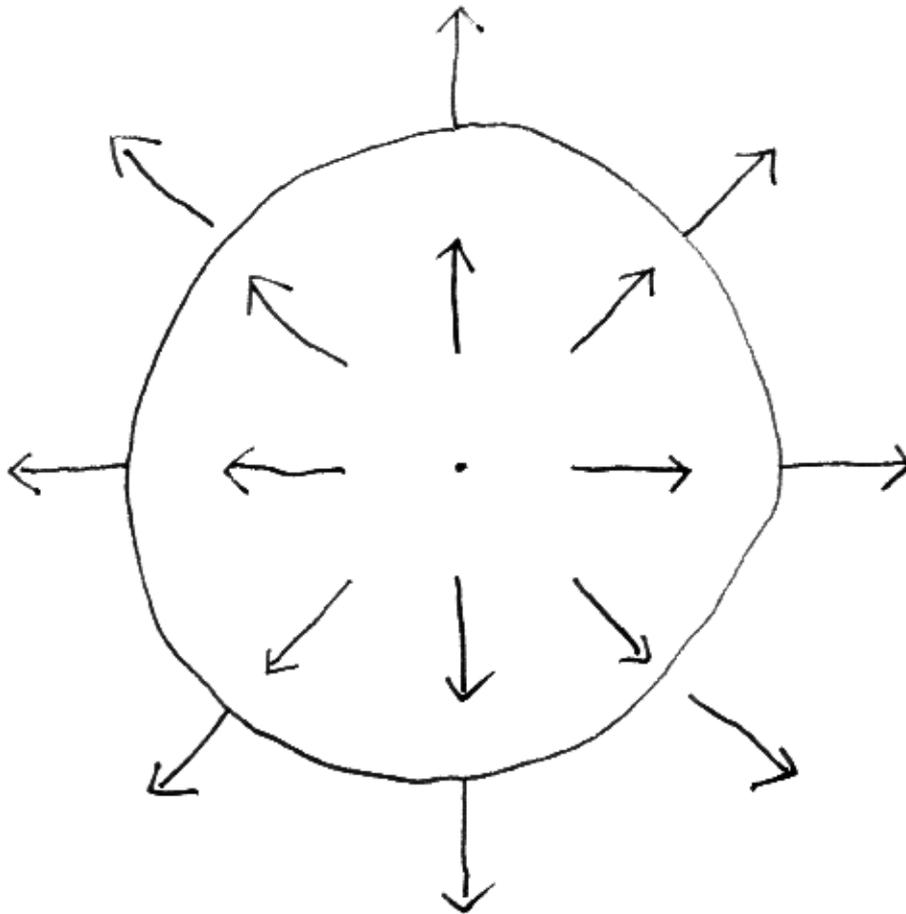
Ergänzung 6.3.6 (Wegintegrale von Vektorfeldern). Redet man für $X = \mathbb{R}^n$ vom Integral eines Vektorfelds $v : A \rightarrow \mathbb{R}^n$ längs eines Weges oder von der **Zirkulation eines Vektorfelds in einem Weg**, so ist das Integral des Kovektorfelds $\omega = \langle v, \cdot \rangle$ gemeint, das in Formeln gegeben wird durch $\omega = v_1 dx_1 + \dots + v_n dx_n$. In der Physik wird das Standardskalarprodukt auf dem \mathbb{R}^n meist $v \cdot w$ notiert und unser Wegintegral über einen Weg $\gamma : [a, b] \rightarrow A$ würde geschrieben als

$$\int_{\gamma} v \cdot dx = \int_a^b v \cdot d\gamma = \int_a^b v(\gamma(t)) \cdot \dot{\gamma}(t) dt$$

Die Bedeutung der Terme des rechtesten Integrals sollte hier klar sein. In der Mitte ist zu verstehen $d\gamma = d_t\gamma = \dot{\gamma}(t) dt$. Ganz links meint dx ein “kleines vektoriell Kurvenelement” und das x ist fett gedruckt um anzudeuten, daß ein Vektor gemeint ist. Ich mag diese Notation nicht besonders, die fette Schreibweise ist auch an der Tafel schlecht umzusetzen. Allgemeiner kann man Wegintegrale von Vektorfeldern v bilden, wann immer ein Skalarprodukt oder allgemeiner ein ausgezeichneter 2-Tensor g zur Verfügung steht, indem wir eben zu unserem Vektorfeld das Kovektorfeld $\omega = \text{can}_g^1(v)$ oder auch $\omega = \text{can}_g^2(v)$ bilden und diese Kovektorfelder dann integrieren wie in 6.3.1 erklärt. Ohne einen ausgezeichneten 2-Tensor gelingt es eben nicht, zwei Vektoren in natürlicher Weise zu paaren: Das gelingt ohne zusätzliche Wahlen in natürlicher Weise nur für einen Vektor und einen Kovektor.

Ergänzung 6.3.7. Redet man für $X = \mathbb{R}^2$ vom **Fluß eines Vektorfelds** $v = (v_1, v_2) : A \rightarrow \mathbb{R}^2$ **durch einen Weg**, so ist das Integral über das Kovektorfeld $\omega = v_1 dy - v_2 dx$ gemeint. Dies Kovektorfeld kann alternativ auch beschrieben werden durch die Formel $\omega_p(u) = \det(v(p)|u)$, in der unsere Vektoren $v(p)$ und u als Spaltenvektoren aufzufassen sind.

6.3.8. In der Literatur scheint mir eine gewisse Verwirrung zu herrschen was die Begriffe “Wegintegral” und “Kurvenintegral” angeht. Die hier gewählte Terminologie soll zum Ausdruck bringen, daß für einen injektiven stetig differenzierbaren Weg $\gamma : [a, b] \rightarrow \mathbb{R}^n$ unser Kurvenintegral nur von der Bildmenge $\gamma([a, b]) \subset \mathbb{R}^n$ abhängt, die wir im Sinne unserer Definition 10.7.2 eine “Kurve” werden nennen dürfen. Unser Wegintegral dahingegen hängt auch von der “durch den Weg



Der Fluß des Vektorfeldes $\partial_r = (x/r)\partial_x + (y/r)\partial_y$ durch den Weg $\gamma : [0, 2\pi] \rightarrow \mathbb{R}^2, t \mapsto (3 \cos t, 3 \sin t)$ ergibt sich nach kurzer Rechnung zu 6π . Die Zirkulation desselben Vektorfeldes in demselben Weg ist dahingegen Null.

γ gegebenen Richtung auf unserer Kurve” ab und ändert sein Vorzeichen, wenn wir die Kurve “in der umgekehrten Richtung durchlaufen”. Andererseits bleibt das Wegintegral unverändert selbst bei nicht notwendig monotoner “Neuparametrisierung”, wenn diese nur den Anfang bzw. das Ende des neuen Parameterintervalls auf den Anfang bzw. das Ende des Alten wirft, siehe 6.3.11. Das Kurvenintegral dahingegen ändert sich bei derartigen Neuparametrisierungen im allgemeinen sehr wohl.

Satz 6.3.9 (Transformation von Wegintegralen). *Seien X und Y endlichdimensionale reelle Räume, A bzw. B darin halboffene Teilmengen und $\phi : A \rightarrow B$ stetig differenzierbar. Ist $\gamma : [a, b] \rightarrow A$ ein stetig differenzierbarer Weg in A und ω ein stetiges Kovektorfeld auf B , so gilt*

$$\int_{\gamma} \phi^* \omega = \int_{\phi \circ \gamma} \omega$$

6.3.10. Zwei Wege $\gamma : I \rightarrow A$ und $\kappa : J \rightarrow B$ heißen ϕ -**verwandt** und wir schreiben $\phi : \gamma \rightsquigarrow \kappa$ genau dann, wenn gilt $\kappa = \phi \circ \gamma$, als da heißt, wenn sie denselben Definitionsbereich $I = J$ haben und für alle $t \in I$ gilt

$$\kappa(t) = \phi(\gamma(t))$$

Sicher hat jeder Weg γ in A genau einen Verwandten, nämlich den Weg $\phi \circ \gamma$. Der Inhalt des Satzes läßt sich mit diesen Begriffsbildungen dahingehend zusammenfassen, daß das Wegintegral Verwandtschaft respektiert. Genauer folgt aus der Verwandtschaft von Wegen $\phi : \gamma \rightsquigarrow \kappa$ und der Verwandtschaft von Kovektorfeldern $\phi : \eta \rightsquigarrow \omega$ die Gleichheit der Wegintegrale $\int_{\gamma} \eta = \int_{\kappa} \omega$.

Beweis. Wir rechnen $\int_{\gamma} \phi^* \omega = \int_a^b \gamma^* \phi^* \omega = \int_a^b (\phi \circ \gamma)^* \omega = \int_{\phi \circ \gamma} \omega$, wo wir 6.3.3, 6.1.18 und nochmals 6.3.3 angewandt haben. \square

Proposition 6.3.11 (Unabhängigkeit von der Parametrisierung). *Sei $\gamma : [a, b] \rightarrow X$ ein stetig differenzierbarer Weg in einem endlichdimensionalen reellen Raum X und ω ein stetiges relatives Kovektorfeld auf einer halboffenen Teilmenge, die sein Bild umfaßt. Ist $u : [c, d] \rightarrow [a, b]$ stetig differenzierbar mit $u(c) = a$ und $u(d) = b$, so gilt*

$$\int_{\gamma \circ u} \omega = \int_{\gamma} \omega$$

6.3.12. Der folgende Beweis zeigt auch, daß bei einer richtungsumkehrenden Umparametrisierung, also für u mit $u(c) = b$ und $u(d) = a$ das Wegintegral über den umparametrisierten Weg das Negative des Wegintegrals über den ursprünglichen Weg ist. In 10.5.3 werden wir allgemeiner das Integral von k -Formen

über k -dimensionale orientierte Mannigfaltigkeiten einführen und speziell im Fall $k = 1$ ein Integral von Kovektorfeldern über orientierte Kurven erhalten, das nach 10.5.12 im wesentlichen die in der Proposition enthaltene Unabhängigkeit des Wegintegrals von der Parametrisierung zu einer Definition umgießt.

Beweis. Wir schreiben unsere Behauptung um zu $\int_c^d u^*(\gamma^*\omega) = \int_a^b \gamma^*\omega$. Für $\eta = f(t) dt$ ein beliebiges stetiges Kovektorfeld auf $[a, b]$, insbesondere auch für $\eta = \gamma^*\omega$ gilt jedoch $u^*\eta = f(u(\tau))u'(\tau) d\tau$ und damit $\int_c^d u^*\eta = \int_a^b \eta$ nach der Substitutionsregel 7.6.1. \square

6.3.13. Wir können nun auch den in 1.4.8 erklärten Trick zur Berechnung der Integrale von rationalen Ausdrücken in $(x, \sqrt{x^2 + 1})$ geometrisch verstehen. Gegeben solch ein rationaler Ausdruck $R(x, y)$ betrachten wir dazu auf einer geeigneten Teilmenge des \mathbb{R}^2 die Differentialform $R(x, y) dx$ und den Weg $\gamma : [a, b] \rightarrow \mathbb{R}^2$ mit $\gamma(t) = (t, \sqrt{t^2 + 1})$ und fassen unser Integral auf als Wegintegral

$$\int_a^b R(t, \sqrt{t^2 + 1}) dt = \int_\gamma R(x, y) dx$$

Solch ein Wegintegral ist nach 6.3.11 unabhängig von der Parametrisierung. Unser Weg durchläuft ein Stück der Hyperbel $y^2 - x^2 = 1$, genauer ein Stück des Hyperbelastes mit $y > 0$. Diesen Ast können wir nach 7.7.6 auch parametrisieren durch $\varphi : (-1, 1) \rightarrow \mathbb{R}^2$ mit

$$\varphi(\tau) = \left(\frac{2\tau}{\tau^2 - 1}, \frac{1 + \tau^2}{1 - \tau^2} \right)$$

und bei dieser Parametrisierung führt uns unser Wegintegral ganz offensichtlich auf das Integral einer rationalen Funktion in τ , das wir nach 1.4 im Prinzip durch bekannte Funktionen ausdrücken können.

Satz 6.3.14 (Wegintegral über ein Differential). *Ist X ein endlichdimensionaler reeller Raum, $A \subset X$ eine halboffene Teilmenge, $f : A \rightarrow \mathbb{R}$ eine stetig differenzierbare Funktion und $\gamma : [a, b] \rightarrow A$ ein stetig differenzierbarer Weg in A , so gilt*

$$\int_\gamma df = f(\gamma(b)) - f(\gamma(a))$$

Beweis. Es gilt $\int_\gamma df = \int_a^b \gamma^*(df) = \int_a^b d(f \circ \gamma) = f(\gamma(b)) - f(\gamma(a))$ nach 6.3.3, 6.1.19 und 6.3.2. \square

6.3.15. Wir interessieren uns nun für die Frage, unter welchen Bedingungen ein stetiges Kovektorfeld das Differential einer Funktion ist, und inwieweit diese Funktion eindeutig bestimmt ist. Diese Fragen werden nach einigen Vorbereitungen durch 6.6.1 und 6.4.6 beantwortet.

Ergänzung 6.3.16. Seien X ein endlichdimensionaler reeller Raum, W ein endlichdimensionaler reeller Vektorraum und $A \subset X$ eine halboffene Teilmenge. Ein **W -wertiges Kovektorfeld auf A** ist eine Abbildung

$$\omega : A \rightarrow \text{Hom}_{\mathbb{R}}(\vec{X}, W)$$

Sie ordnet also jedem Punkt $p \in A$ eine lineare Abbildung des Richtungsraums in den Raum W zu. Ist etwa Y ein weiterer endlichdimensionaler reeller Raum und A halboffen und $f : A \rightarrow Y$ differenzierbar, so ist df oder genauer $p \mapsto d_p f$ ein \vec{Y} -wertiges Kovektorfeld auf A . Ist nun $\varphi : [a, b] \rightarrow A$ ein stetig differenzierbarer Weg in einer halboffenen Teilmenge A eines endlichdimensionalen reellen Raums X und $\omega : A \rightarrow \text{Hom}_{\mathbb{R}}(\vec{X}, W)$ ein stetiges Kovektorfeld auf A mit Werten in einem endlichdimensionalen reellen Vektorraum W , so definieren wir in Verallgemeinerung des Falls reellwertiger Kovektorfelder aus 6.3 einen Vektor $(\int_{\varphi} \omega) \in W$, das **Integral des W -wertigen Kovektorfelds ω längs des Weges φ** , durch die Vorschrift

$$\int_{\varphi} \omega = \int_a^b \omega_{\varphi(t)}(\varphi'(t)) dt$$

Auf der rechten Seite ist also für jeden Zeitpunkt t der Homomorphismus $\omega_{\varphi(t)} : \vec{X} \rightarrow W$ auszuwerten auf dem Geschwindigkeitsvektor $\varphi'(t) \in \vec{X}$, und die so entstehende stetige Abbildung $[a, b] \rightarrow W$ ist als vektorwertige Funktion zu integrieren im Sinne von 1.3.3.

Ergänzendes Beispiel 6.3.17. In der Physik begegnen einem insbesondere oft Kovektorfelder mit Werten in eindimensionalen reellen Vektorräumen. Zum Beispiel wird man sich ein Kraftfeld auf dem Anschauungsraum \mathbb{E} aus ?? a priori wie in ?? erklärt als ein “Vektorfeld mit Einheiten” denken, genauer als Abbildung

$$F : \mathbb{E} \rightarrow \vec{\mathbb{E}} \otimes \langle\langle g/s^2 \rangle\rangle$$

Da es sich jedoch mit Kovektorfeldern bei Koordinatenwechseln sehr viel besser rechnen läßt als mit Vektorfeldern, ist es oft günstiger, die durch das kanonische Skalarprodukt $s : \vec{\mathbb{E}} \times \vec{\mathbb{E}} \rightarrow \mathbb{L}^{\otimes 2}$ aus ?? gegebene Identifikation $\text{can}_s : \vec{\mathbb{E}} \xrightarrow{\sim} \text{Hom}_{\mathbb{R}}(\vec{\mathbb{E}}, \mathbb{L}^{\otimes 2})$ nachzuschalten und unser Kraftfeld stattdessen als eine Abbildung

$$\tilde{F} : \mathbb{E} \rightarrow \text{Hom}_{\mathbb{R}}(\vec{\mathbb{E}}, \langle\langle g \text{ m}^2/\text{s}^2 \rangle\rangle)$$

aufzufassen. Die Elemente des eindimensionalen Vektorraums

$$\langle\langle g \text{ m}^2/\text{s}^2 \rangle\rangle = \mathbb{M} \otimes \mathbb{L}^{\otimes 2} \otimes (\vec{\mathbb{T}}^*)^{\otimes 2}$$

heißen in der Physik auch **Energien**. In diesem Sinne können wir ein Kraftfeld dann also auch als ein Energie-wertiges Kovektorfeld auffassen. Das Wegintegral über dieses Kovektorfeld heißt die bei Durchlaufen des Weges in besagtem

Kraftfeld freiwerdende Energie, und ihr Negatives die zu verrichtende **Arbeit**. Anschaulich und etwas vage gesprochen ordnet das Negative dieses Kovektorfelds nämlich gerade “jeder kleinen Verrückung die Arbeit zu, die bei dieser kleinen Verrückung gegen das Kraftfeld zu leisten wäre”. Eine energiewertige Abbildung $V : \mathbb{E} \rightarrow \langle\langle \text{g m}^2/\text{s}^2 \rangle\rangle$ mit $dV = -\tilde{F}$ heißt in der Physik ein **Potential** unseres Kraftfelds.

Ergänzendes Beispiel 6.3.18. Zentral in der sogenannten “Funktionentheorie” ?? sind die Wegintegrale komplexwertiger Kovektorfelder, die auf offenen Teilmengen der komplexen Zahlenebene definiert sind. Üblicherweise bezeichnet in diesem Kontext $z : \mathbb{C} \rightarrow \mathbb{C}$ die Identität und dz ihr Differential, ein komplexwertiges Kovektorfeld auf \mathbb{C} . Mit $f(z) dz$ bezeichnet man dann das Produkt dieses Kovektorfelds mit einer komplexwertigen Funktion $z \mapsto f(z)$. Das Integral derartiger Kovektorfelder über Integrationswege liefert in diesem Fall entsprechend komplexe Zahlen.

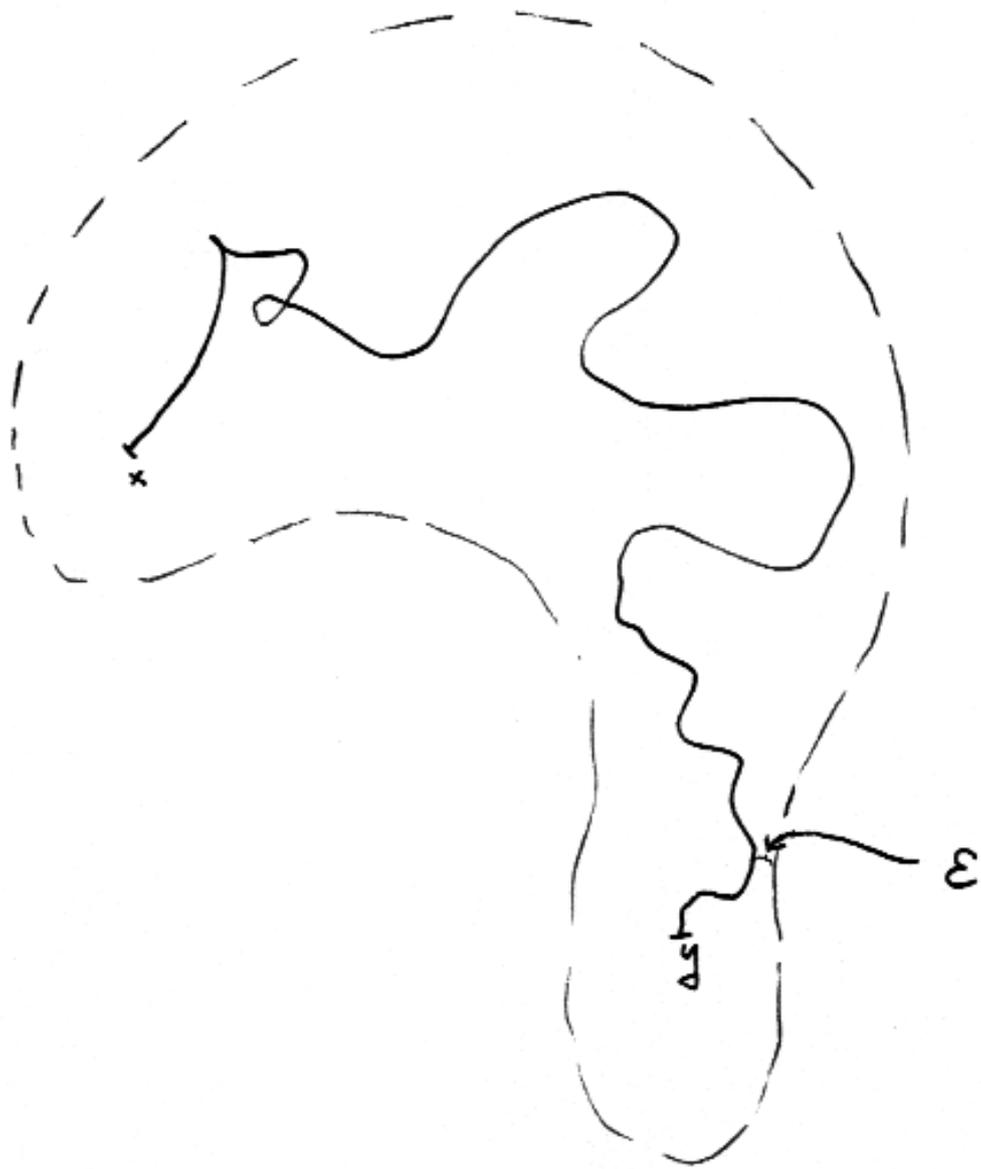
6.4 Wegzusammenhängende Räume

Definition 6.4.1. Ist X ein topologischer Raum und sind $x, y \in X$ Punkte, so nennen wir eine stetige Abbildung $\gamma : [a, b] \rightarrow X$ mit $\gamma(a) = x$ und $\gamma(b) = y$ einen **Weg von x nach y** . Ein topologischer Raum X heißt **wegweise zusammenhängend** oder auch kurz **wegzusammenhängend** genau dann, wenn er nicht leer ist und es für je zwei Punkte unseres Raums einen Weg vom einen zum anderen gibt.

Übung 6.4.2. Auf jedem topologischen Raum X definiert man die Relation der “Wegverbindbarkeit” durch die Vorschrift xWy genau dann, wenn es in X einen Weg von x nach y gibt. Man zeige, daß das eine Äquivalenzrelation ist. Hinweis: Die Transitivität ergibt sich durch das “Aneinanderhängen von Wegen” und die Stetigkeit der so entstehenden Wege folgt mit 9.5.32. Die Äquivalenzklassen für diese Relation heißen die **Wegzusammenhangskomponenten** unseres Raums. Man zeige, daß die Wegzusammenhangskomponenten eines topologischen Raums offen sind genau dann, wenn jeder Punkt eine wegzusammenhängende Umgebung besitzt.

Definition 6.4.3. Unter einem **stückweise linearen Weg** oder auch **Polygonzug** in einem normierten reellen Raum verstehen wir einen Weg, der aus endlich vielen Geradensegmenten zusammengesetzt ist.

6.4.4. Im Sinne unserer Definitionen müßten wir eigentlich von einem “stückweise affinen Weg” reden, aber das tut kein Mensch.



Lemma 6.4.5. *In einer wegzusammenhängenden offenen Teilmenge eines normierten reellen Raums lassen sich je zwei Punkte auch durch einen stückweise linearen Weg verbinden.*

Beweis. Sei $A \subseteq V$ unsere Teilmenge und seien $x, y \in A$ gegeben. Nach Annahme gibt es einen Weg $\gamma : [a, b] \rightarrow A$ von x nach y . Ohne Beschränkung der Allgemeinheit dürfen wir $A \neq V$ annehmen. Dann ist der Abstand zum Komplement von A nach 9.2.22 eine stetige Funktion $d_{V \setminus A} : V \rightarrow \mathbb{R}$ ohne Nullstelle auf A und $d_{V \setminus A} \circ \gamma$ hat nach 6.4.3 auf $[a, b]$ ein Minimum $\varepsilon > 0$, als da heißt, es gibt $\varepsilon > 0$ derart, daß alle Punkte aus $\gamma([a, b])$ mindestens den Abstand ε zum Komplement von A haben. Andererseits ist γ gleichmäßig stetig, wir finden also eine Unterteilung $a = a_0 \leq a_1 \leq \dots \leq a_n = b$ unseres Intervalls mit $\|\gamma(a_i) - \gamma(a_{i-1})\| < \varepsilon$ für $1 \leq i \leq n$. Ein zwischen den Eckpunkten $x = \gamma(a_0), \gamma(a_1), \dots, \gamma(a_n) = y$ jeweils linear verlaufender Weg bleibt also ganz in A . Damit ist gezeigt, daß sich je zwei Punkte aus A auch durch einen stückweise linearen Weg in A verbinden lassen. \square

Lemma 6.4.6. *Auf einer offenen wegzusammenhängenden Teilmenge eines endlichdimensionalen reellen Raums ist jede differenzierbare reellwertige Funktion mit verschwindendem Differential konstant.*

Beweis. Eine differenzierbare Funktion mit verschwindendem Differential muß nach 6.3.14 am Anfang und Ende jedes stetig differenzierbaren Weges und dann auch am Anfang und Ende jedes stückweise linearen Weges denselben Wert annehmen. Das Lemma folgt damit aus 6.4.5. \square

Definition 6.4.7. Eine Teilmenge eines topologischen Raums nennen wir **diskret** genau dann, wenn jeder ihrer Punkte eine Umgebung besitzt, in der kein anderer Punkt besagter Teilmenge liegt. Zum Beispiel ist die Menge aller Brüche $\{1, 1/2, 1/3, \dots\}$ mit einer Eins im Zähler eine diskrete Teilmenge der reellen Zahlengeraden.

6.4.8. In anderen Worten nennen wir also eine Teilmenge eines topologischen Raums diskret genau dann, wenn sie mit der Spurtopologie im Sinne von 9.5.9 ein diskreter topologischer Raum im Sinne von 9.5.20 wird. Andere Autoren verstehen unter einer diskreten Teilmenge eines topologischen Raums abweichend eine Teilmenge derart, daß jeder Punkt des gesamten Raums eine Umgebung besitzt, in der höchstens ein Punkt besagter Teilmenge liegt. In unserer Terminologie sind das genau die diskreten abgeschlossenen Teilmengen.

Übung 6.4.9. Das Komplement einer abgeschlossenen diskreten Teilmenge in einer wegzusammenhängenden offenen Teilmenge eines \mathbb{R}^n ist für $n > 1$ wegzusammenhängend. Dasselbe gilt im Übrigen auch ohne die Bedingung “abgeschlossen”, ist dann aber schwerer zu zeigen.

Übung 6.4.10. Ist $U \subseteq \mathbb{R}^n$ offen und wegzusammenhängend und $A \subset \mathbb{R}^n$ ein affiner Teilraum einer Dimension $\dim A \leq n-2$ alias einer Kodimension mindestens Zwei, so ist auch $U \setminus A$ wegzusammenhängend. Für Teilräume A der Kodimension Eins alias affine Hyperebenen A gilt das natürlich nicht!

Definition 6.4.11. Eine Funktion auf einer Teilmenge des \mathbb{R}^n , die um jeden Punkt ihres Definitionsbereichs in einer Umgebung durch ihre Taylorreihe dargestellt werden kann, heißt **analytisch**.

6.4.12. Wir werden in ?? zeigen, daß Potenzreihen in einer Veränderlichen analytische Funktionen liefern. Analog kann man es auch für Potenzreihen in mehreren Veränderlichen zeigen.

Übung 6.4.13. Stimmen zwei auf derselben wegzusammenhängenden offenen Teilmenge des \mathbb{R}^n definierte analytische Funktionen auf einer Umgebung eines Punktes überein, so sind sie gleich. Hinweis: Man ziehe sich mithilfe stückweise linearer Wege auf den Fall $n = 1$ zurück.

Ergänzende Übung 6.4.14. Man zeige: Die Gruppe $SO(n)$ aller orthogonalen $(n \times n)$ -Matrizen mit Determinante Eins ist wegzusammenhängend. Hinweis: ?. Die Gruppe $GL(n; \mathbb{R})^+$ aller invertierbaren reellen $(n \times n)$ -Matrizen mit positiver Determinante ist wegzusammenhängend. Hinweis: ?. Die vorgeschlagenen Lösungsansätze laufen auf eine Flickschusterei hinaus. Einen konzeptionelleren Beweis werden wir in ?? kennenlernen.

Ergänzende Übung 6.4.15. Das Bild eines wegzusammenhängenden Raums unter einer stetigen Abbildung ist stets wieder wegzusammenhängend. Die wegzusammenhängenden Teilmengen von \mathbb{R} sind gerade die nichtleeren Intervalle.

Übung 6.4.16. Gegeben ein wegzusammenhängender topologischer Raum ist jede Teilmenge, die sowohl offen als auch abgeschlossen ist, entweder leer oder bereits der ganze Raum. Hinweis: Man wähle sonst einen Weg von einem Punkt unserer Teilmenge in ihr Komplement und konstruiere einen Widerspruch.

Ergänzung 6.4.17. In ?? wird ganz allgemein ein “zusammenhängender topologischer Raum” definiert durch die Eigenschaft, daß er nicht leer ist und daß jede Teilmenge, die sowohl offen als auch abgeschlossen ist, entweder leer oder bereits der ganze Raum sein muß. Dort bespreche ich auch noch ausführlicher die Beziehungen dieses Begriffs zum hier eingeführten Wegzusammenhang. Die vorhergehende Übung 6.4.16 besagt in dieser Terminologie insbesondere, daß jeder wegzusammenhängende Raum auch zusammenhängend ist.

6.5 Homotopie von Wegen

6.5.1. Einen durch das Einheitsintervall parametrisierten Weg $\gamma : [0, 1] \rightarrow X$ in einem topologischen Raum X nennen wir im Folgenden einen **normierten**

Weg. Zu jedem Weg $\gamma : [a, b] \rightarrow X$ bilden wir in offensichtlicher Weise den zugehörigen normierten Weg $\hat{\gamma} : t \mapsto \gamma((1-t)a + tb)$.

Definition 6.5.2. Seien x, y Punkte eines topologischen Raums X . Zwei normierte Wege α, β von x nach y heißen **homotop** oder präziser **homotop in X** oder auch **homotop mit festen Endpunkten** und wir schreiben $\alpha \simeq \beta$ genau dann, wenn es eine stetige Abbildung

$$h : [0, 1]^2 \rightarrow X$$

des Einheitsquadrats in unseren Raum gibt, die auf der Unter- bzw. Oberkante unseres Quadrats mit α bzw. β übereinstimmt und die auf der Vorder- und der Hinterkante konstant ist. In Formeln ausgedrückt fordern wir also $h(t, 0) = \alpha(t)$ und $h(t, 1) = \beta(t)$ für alle $t \in [0, 1]$ sowie $h(0, \tau) = x$ und $h(1, \tau) = y$ für alle $\tau \in [0, 1]$. Wir sagen dann auch, h sei eine **Homotopie** zwischen α und β und schreiben $h : \alpha \simeq \beta$. Zwei beliebige Wege von x nach y nennen wir **homotop** genau dann, wenn die zugehörigen normierten Wege homotop sind.

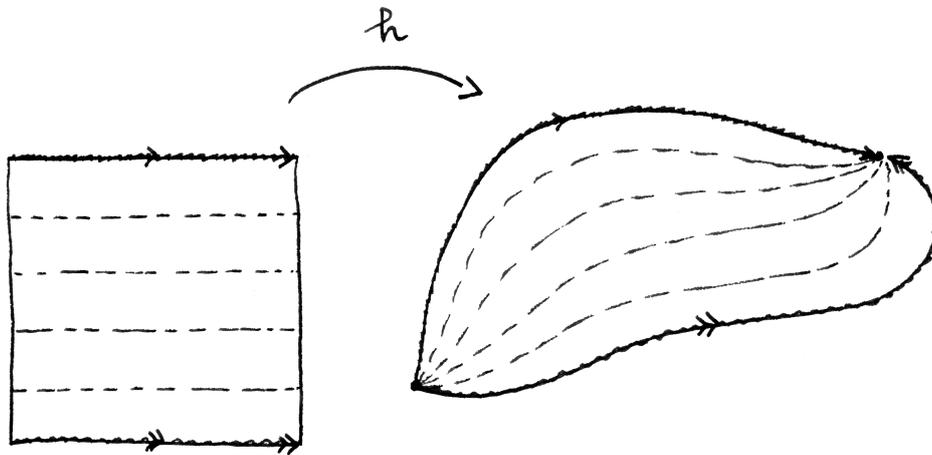
6.5.3. Vielleicht anschaulicher kann man Homotopie auch dahingehend interpretieren, daß es eine durch $\tau \in [0, 1]$ parametrisierte Familie von normierten Wegen h_τ von x nach y geben soll derart, daß gilt $h_0 = \alpha$, $h_1 = \beta$ und daß unsere Familie stetig von τ abhängt in dem Sinne, daß die Abbildung $[0, 1]^2 \rightarrow X$, $(t, \tau) \mapsto h_\tau(t)$ stetig ist.

Beispiel 6.5.4. Für eine konvexe Teilmenge X eines endlichdimensionalen reellen Raums und zwei beliebige Punkte $x, y \in X$ sind je zwei Wege α, β von x nach y homotop in X . Sind unsere Wege normiert, so kann man eine Homotopie explizit angeben vermittels $h(t, \tau) = (1 - \tau)\alpha(t) + \tau\beta(t)$.

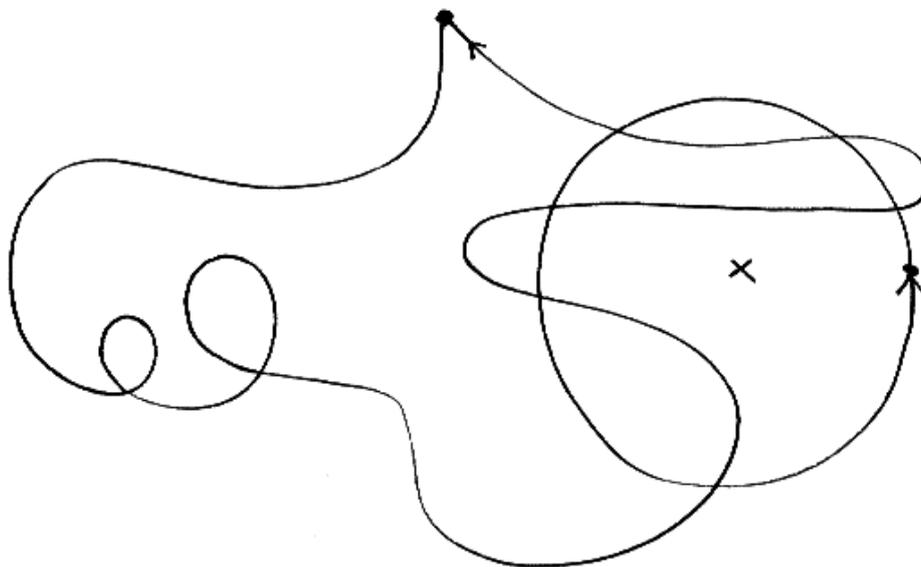
6.5.5. Bilder homotoper Wege sind homotop. Ist genauer eine Abbildung $f : X \rightarrow Y$ stetig, so folgt aus $h : \alpha \simeq \beta$ schon $f \circ h : f \circ \alpha \simeq f \circ \beta$. Speziell ist ein Weg homotop zu allen seinen Umparametrisierungen, denn nach 6.5.4 sind je zwei Wege in $[0, 1]$ von 0 nach 1 homotop und damit gilt dasselbe für ihre Bilder unter einem beliebigen stetigen $\gamma : [0, 1] \rightarrow Y$.

Übung 6.5.6. Homotopie ist eine Äquivalenzrelation auf der Menge aller Wege zwischen zwei fest vorgegebenen Punkten. Hinweis: 9.5.32.

Definition 6.5.7. Ein Weg in einem topologischen Raum heißt ein **geschlossener Weg** genau dann, wenn sein Anfangs- und Endpunkt zusammenfallen. Ein Weg heißt **zusammenziehbar** genau dann, wenn er homotop ist zu einem konstanten Weg. Per definitionem ist also jeder zusammenziehbare Weg geschlossen. Ein topologischer Raum heißt **wegweise einfach zusammenhängend** genau dann, wenn er wegzusammenhängend ist und wenn darüber hinaus jeder geschlossene Weg in unserem Raum zusammenziehbar ist.



Eine Homotopie zwischen zwei Wegen, in diesem Fall zwischen den beiden Randwegen unserer Banane.



Ein zusammenziehbarer und ein nicht zusammenziehbarer geschlossener Weg in Komplement des durch ein Kreuzchen markierten Punktes in der Papierebene

6.5.8. In ?? werden wir auch “einfach zusammenhängende” topologische Räume kennenlernen, die in etwas anderer Weise definiert werden. Für halboffene Teilmengen normierter Räume wird sich dieser Begriff jedoch als zu “wegweise einfach zusammenhängend” gleichbedeutend herausstellen, weshalb wir in diesen Fällen auch jetzt schon das “wegweise” meist weglassen.

Übung 6.5.9. Ein Raum ist wegweise einfach zusammenhängend genau dann, wenn er wegzusammenhängend ist und je zwei Wege mit demselben Anfangs- und demselben Endpunkt darin homotop sind.

Ergänzung 6.5.10. Ist $U \subseteq \mathbb{R}^n$ offen und einfach zusammenhängend und ist $A \subset \mathbb{R}^n$ ein affiner Teilraum einer Kodimension ≥ 3 , so ist auch $U \setminus A$ einfach zusammenhängend. Für einen Beweis dieses Analogons zu 6.4.10 verweise ich auf die Topologie, siehe etwa ??.

Ergänzende Übung 6.5.11. Jeder Weg in einer offenen Teilmenge eines endlichdimensionalen reellen Vektorraums ist dort homotop zu einem Polygonzug. Hinweis: 6.4.5.

6.6 Rotation und Potential

Proposition 6.6.1 (Wegintegrale und Potential). Sei X ein endlichdimensionaler reeller Raum, $A \subseteq X$ eine offene Teilmenge und ω ein stetiges Kovektorfeld auf A . So sind gleichbedeutend:

1. Unser Kovektorfeld ist das Differential einer differenzierbaren Funktion.
2. Das Integral unseres Kovektorfelds über beliebige stetig differenzierbare Wege in A hängt nur vom Anfangs- und Endpunkt ab.
3. Das Integral unseres Kovektorfelds über jeden geschlossenen stetig differenzierbaren Weg in A verschwindet.

6.6.2. In physikalischer Terminologie 6.3.17 hat also ein Kraftfeld oder genauer das zugehörige energiewertige Kovektorfeld ein Potential genau dann, wenn die längs beliebiger Wege geleistete Arbeit nur vom Anfangs- und Endpunkt abhängt.

Beweis. Es ist im folgenden bequem, für etwas allgemeinere Wege den Begriff des Wegintegrals zur Verfügung zu haben.

Definition 6.6.3. Sei X ein endlichdimensionaler reeller Raum. Ein Weg $\gamma : [a, b] \rightarrow X$ heißt **stückweise stetig differenzierbar** genau dann, wenn es eine Zerlegung $a = a_0 < a_1 < \dots < a_r = b$ unseres Intervalls gibt derart, daß die Restriktionen $\gamma|_{[a_{i-1}, a_i]}$ stetig differenzierbar sind für alle i . Wir bezeichnen stückweise stetig differenzierbare Wege abkürzend als **Integrationswege**.

6.6.4. Ist $\gamma : [a, b] \rightarrow X$ ein Integrationsweg in einem endlichdimensionalen reellen Raum X und ω ein auf dem Bild von γ definiertes stetiges Kovektorfeld, so setzen wir

$$\int_{\gamma} \omega = \int_{\gamma|_{[a, a_1]}} \omega + \int_{\gamma|_{[a_1, a_2]}} \omega + \dots + \int_{\gamma|_{[a_{r-1}, b]}} \omega$$

für $a_1 < \dots < a_{r-1}$ die Stellen in (a, b) , an denen γ nicht differenzierbar ist. Sicher gilt dann $\int_{\gamma} \omega = \int_{\gamma|_{[a, t]}} \omega + \int_{\gamma|_{[t, b]}} \omega$ für alle $t \in (a, b)$.

Wir behaupten nun zunächst, daß die Aussagen 2 bzw. 3 der Proposition jeweils gleichbedeutend sind zu

- 2'. Das Integral von ω über beliebige Integrationswege in A hängt nur vom Anfangs- und Endpunkt ab.
- 3'. Das Integral von ω über jeden geschlossenen Integrationsweg in A verschwindet.

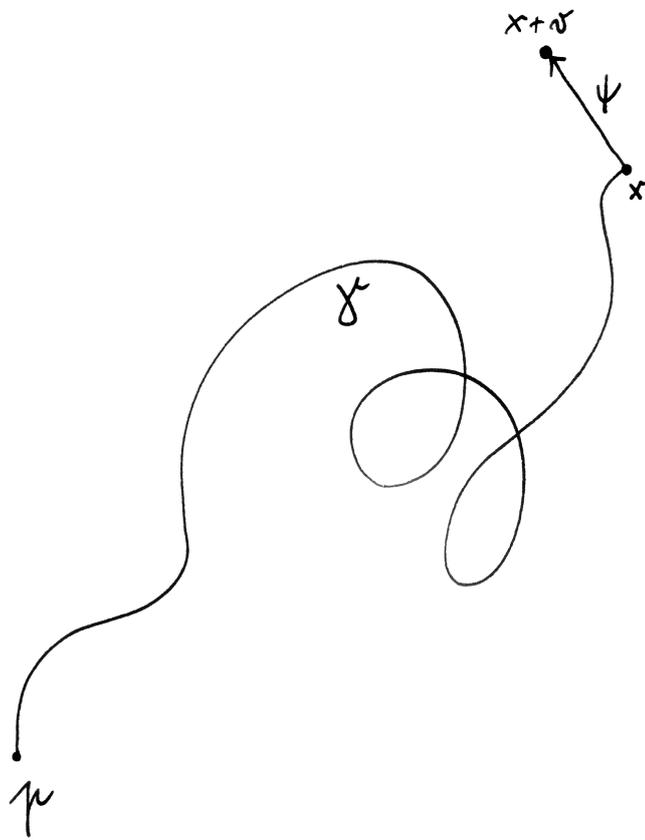
Hier ist $2' \Rightarrow 2$ offensichtlich. Andererseits können wir aber jeden Integrationsweg so umparametrisieren, daß er stetig differenzierbar wird. Das Integral ändert sich dabei nicht, und so folgt auch die andere Richtung $2 \Rightarrow 2'$. Ebenso zeigt man $3 \Leftrightarrow 3'$. Nach diesen Vorarbeiten beginnen wir nun mit dem eigentlichen Beweis der Proposition. Die Folgerungen $1 \Rightarrow 2 \Rightarrow 3$ sind offensichtlich. Wir zeigen als nächstes $3' \Rightarrow 2$ durch Widerspruch: Gäbe es zwei Integrationswege mit demselben Anfangs- und Endpunkt aber verschiedenen Integralen, so könnten wir den einen dieser Wege umdrehen und an den anderen anhängen und so einen geschlossenen Integrationsweg erhalten, über den das Integral von ω nicht Null wäre. Damit ist $3' \Rightarrow 2$ gezeigt. Zeigen wir nun noch $2' \Rightarrow 1$, so haben wir schon einmal die Äquivalenzen $1 \Leftrightarrow 2 \Leftrightarrow 3$ nachgewiesen. Nach 6.4.2 und 6.4.5 dürfen wir annehmen, daß A nicht leer ist und sich je zwei Punkte aus A durch einen Integrationsweg verbinden lassen. Dann wählen wir $p \in A$ fest und definieren eine Funktion $f : A \rightarrow \mathbb{R}$ durch die Vorschrift

$$f(x) = \int_{\gamma} \omega$$

für einen und nach 2' dann auch jeden Integrationsweg γ von p nach x . Ist nun ψ ein stetig differenzierbarer Weg in A von x nach y , so behaupten wir

$$\int_{\psi} \omega = f(y) - f(x)$$

In der Tat können wir ja ψ am γ anhängen und so einen Integrationsweg von p nach y erhalten, so daß also gilt $\int_{\gamma} \omega + \int_{\psi} \omega = f(y)$. Mit dieser Erkenntnis läßt



Ein beliebiger Weg mit angehängtem geraden Stück aus dem Beweis von $3 \Rightarrow 1$.

sich das Differential von f nun sehr leicht berechnen. Gegeben $x \in A$ sei $B \subseteq X$ ein offener Ball um Null mit $x + B \subset A$. Gegeben $v \in B$ erhalten wir mit unserer Notation 6.3.2 sofort

$$f(x+v) - f(x) = \int_x^{x+v} \omega = \int_0^1 \omega_{x+tv}(v) dt = \omega_x(v) + \int_0^1 (\omega_{x+tv} - \omega_x)(v) dt$$

Das letzte Integral läßt sich aber schreiben als $\|v\|$ mal eine Funktion, die beschränkt ist für $v \in B$ durch $\sup\{\|\omega_{x+w} - \omega_x\| \mid \|w\| \leq \|v\|\}$ und die folglich mit v gegen Null strebt. Das zeigt $d_x f = \omega_x$ wie gewünscht. \square

Definition 6.6.5. Sei X ein endlichdimensionaler reeller Raum und $A \subseteq X$ eine offene Teilmenge. Ein stetig differenzierbares Kovektorfeld $\omega : A \rightarrow \vec{X}^*$ auf A heißt **geschlossen** genau dann, wenn an jeder Stelle $p \in A$ sein Differential $d_p \omega : \vec{X} \rightarrow \vec{X}^*$ eine symmetrische Bilinearform auf \vec{X} liefert im Sinne einer Gleichheit von reellen Zahlen

$$(d_p \omega)(\vec{v})(\vec{w}) = (d_p \omega)(\vec{w})(\vec{v}) \quad \forall \vec{v}, \vec{w} \in \vec{X}$$

Diese Terminologie geht vermutlich auf den gleich folgenden Satz 6.6.10 zurück, nach dem ein stetig differenzierbares Kovektorfeld geschlossen ist genau dann, wenn seine Wegintegrale über alle geschlossenen und im Definitionsbereich zusammenziehbaren Wege verschwinden.

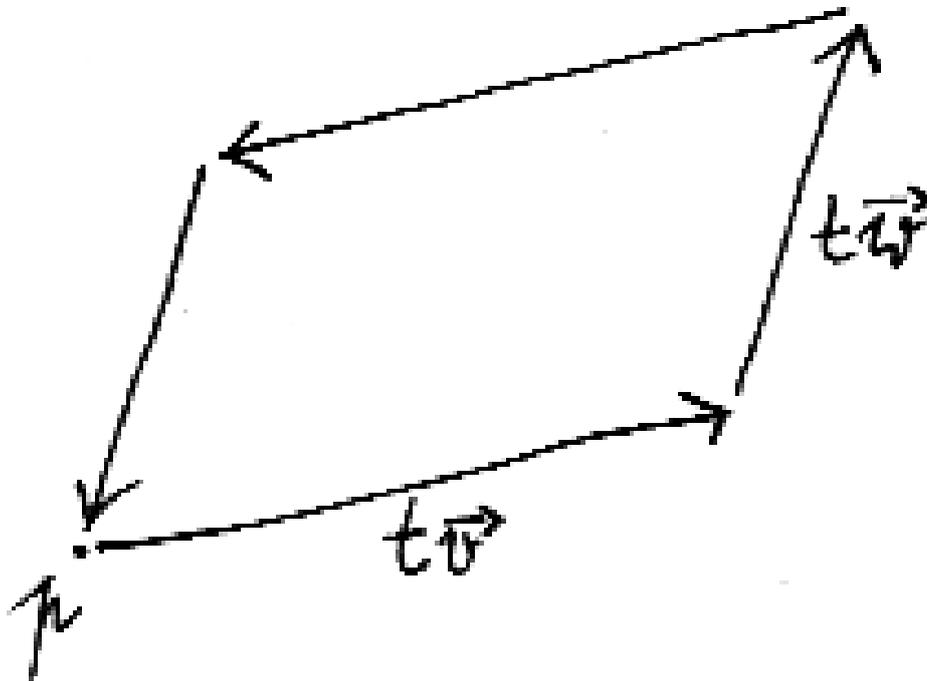
Beispiel 6.6.6. Ein stetig differenzierbares Kovektorfeld $\omega = \sum u_i dx_i$ auf einer offenen Teilmenge eines \mathbb{R}^n ist geschlossen genau dann, wenn gilt

$$\frac{\partial u_i}{\partial x_j} = \frac{\partial u_j}{\partial x_i} \quad \forall i, j$$

Ergänzung 6.6.7. Später werden wir in 10.6.3 ganz allgemein die “äußere Ableitung von Differentialformen” einführen. In dieser Terminologie sind dann unsere geschlossenen Kovektorfelder aus der vorhergehenden Definition 6.6.5 genau diejenigen stetig differenzierbaren Kovektorfelder, deren äußere Ableitung verschwindet.

Übung 6.6.8. Sei X ein endlichdimensionaler reeller Raum, $A \subseteq X$ eine offene Teilmenge, $p \in A$ ein Punkt und $\omega : A \rightarrow \vec{X}^*$ ein stetig differenzierbares Kovektorfeld. So gilt in den Notationen der vorhergehenden Definition 6.6.5 für alle $\vec{v}, \vec{w} \in \vec{X}$ die Identität

$$(d_p \omega)(\vec{v})(\vec{w}) - (d_p \omega)(\vec{w})(\vec{v}) = \lim_{t \rightarrow 0} \frac{1}{t^2} \int_{\gamma(p, t\vec{v}, t\vec{w})} \omega$$



Der Weg $\gamma(p, t\vec{v}, t\vec{w})$ aus Übung 6.6.8. Mit $t \rightarrow 0$ wird er natürlich immer kleiner.

mit der Notation $\gamma(p, t\vec{v}, t\vec{w})$ für den Weg, der einmal das Parallelogramm mit einer Ecke p und Kantenvektoren $t\vec{v}$ und $t\vec{w}$ umläuft, oder genauer, der stückweise linear läuft erst von p nach $p+t\vec{v}$, dann weiter nach $p+t\vec{v}+t\vec{w}$, von da nach $p+t\vec{w}$, und dann wieder zurück nach p . Hinweis: Es mag die Rechnung vereinfachen, wenn man das fragliche Integral zu einer Funktion von zwei Veränderlichen s, t erweitert.

Ergänzende Übung 6.6.9. Man zeige: Gegeben ein geschlossenes stetig differenzierbares Kovektorfeld ist auch sein Rückwärtsverwandter unter jeder stetig differenzierbaren Abbildung geschlossen. Insbesondere ist diese Eigenschaft “unabhängig von der Wahl der Koordinaten”. Eine gewisse Anschauung für diese Eigenschaft mag 6.6.8 liefern.

Satz 6.6.10 (Rotation und Potential). Sei X ein endlichdimensionaler reeller Raum und $A \subseteq X$ eine offene Teilmenge. Für ein stetig differenzierbares Kovektorfeld ω auf A sind gleichbedeutend:

1. Unser Kovektorfeld ist geschlossen;
2. Jeder Punkt von A besitzt eine offene Umgebung, auf der unser Kovektorfeld das Differential einer Funktion ist;
3. Auf jeder einfach zusammenhängenden offenen Teilmenge von A ist unser Kovektorfeld das Differential einer Funktion;
4. Sind γ und ψ zwei in A zueinander homotope Integrationswege, so gilt

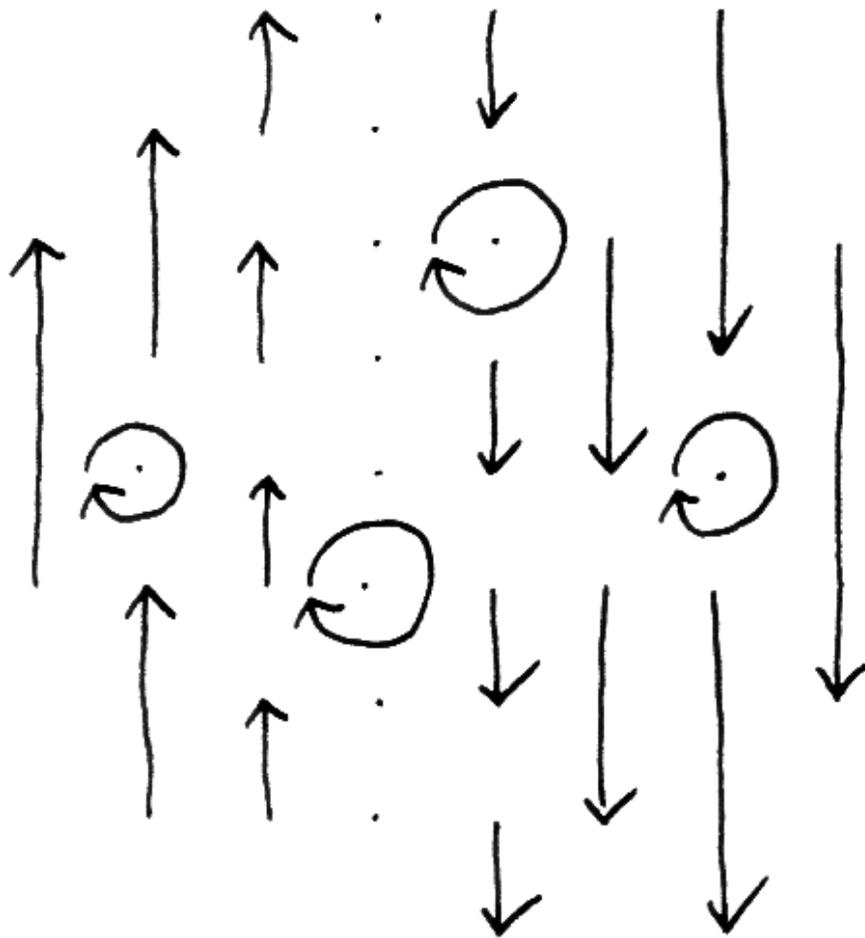
$$\int_{\gamma} \omega = \int_{\psi} \omega$$

6.6.11. Insbesondere verschwindet also das Wegintegral eines stetig differenzierbaren geschlossenen Kovektorfelds über jeden in A zusammenziehbaren Integrationsweg. Einen alternativen und in gewisser Weise glatteren Beweis unter stärkeren Voraussetzungen geben wir in 10.8.25.

6.6.12. Gegeben ein differenzierbares Vektorfeld $v = (v_1, v_2) : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ bezeichnet man die Differenz $\text{rot } v = \frac{\partial v_2}{\partial x_1} - \frac{\partial v_1}{\partial x_2}$ als seine **Wirbeldichte** oder **skalare Rotation**, und gegeben ein differenzierbares Vektorfeld $v : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ definiert man seine **Rotation**, ein weiteres Vektorfeld, durch die Vorschrift

$$\text{rot } v = \left(\frac{\partial v_3}{\partial x_2} - \frac{\partial v_2}{\partial x_3}, \frac{\partial v_1}{\partial x_3} - \frac{\partial v_3}{\partial x_1}, \frac{\partial v_2}{\partial x_1} - \frac{\partial v_1}{\partial x_2} \right)$$

Um für diese Konzepte eine Anschauung zu entwickeln, mag man sich unser Vektorfeld F als ein Kraftfeld vorstellen. Läßt man im ebenen Fall dieses Kraftfeld



Das ebene Vektorfeld $(x, y) \mapsto (0, -x)$ hat konstant die Rotation -1 .

auf den Rand einer kleinen Kreisscheibe wirken, die an einer Stelle unserer Ebene drehbar befestigt ist, so beginnt sie sich zu drehen. Drehsinn sowie die Stärke der drehenden Kraft entsprechen Vorzeichen und Betrag der skalaren Rotation. Läßt man im räumlichen Fall dieses Kraftfeld auf die Oberfläche eines kleinen Balls wirken, den man an einer Stelle p hineinhält, so beginnt er sich auch zu drehen. Die Drehachse ist dann die von der Rotation unseres Vektorfeldes bei p erzeugte Gerade, und der Drehsinn sowie die Stärke der drehenden Kraft entsprechen Richtung und Länge der Rotation. In dieser Terminologie formuliert besagt der vorstehende Satz in \mathbb{R}^2 und \mathbb{R}^3 insbesondere, daß jedes rotationsfreie Vektorfeld auf einer einfach zusammenhängenden offenen Menge der Gradient einer Funktion ist alias in physikalischer Terminologie ausgedrückt ein Potential besitzt.

Beweis. $4 \Rightarrow 3$ ergibt sich unmittelbar aus Satz 6.6.1, nach dem jedes stetige Kovektorfeld das Differential einer Funktion ist, bei dem alle Wegintegrale über geschlossene Integrationswege verschwinden. $3 \Rightarrow 2$ ist offensichtlich. Um $2 \Rightarrow 1$ zu zeigen, dürfen wir ohne Beschränkung der Allgemeinheit $X = \mathbb{R}^n$ annehmen. Die Implikation ist dann auch offensichtlich, für $\omega = \sum u_i dx_i = df$ gilt in der Tat

$$\frac{\partial u_i}{\partial x_j} = \frac{\partial^2 f}{\partial x_j \partial x_i} = \frac{\partial^2 f}{\partial x_i \partial x_j} = \frac{\partial u_j}{\partial x_i}$$

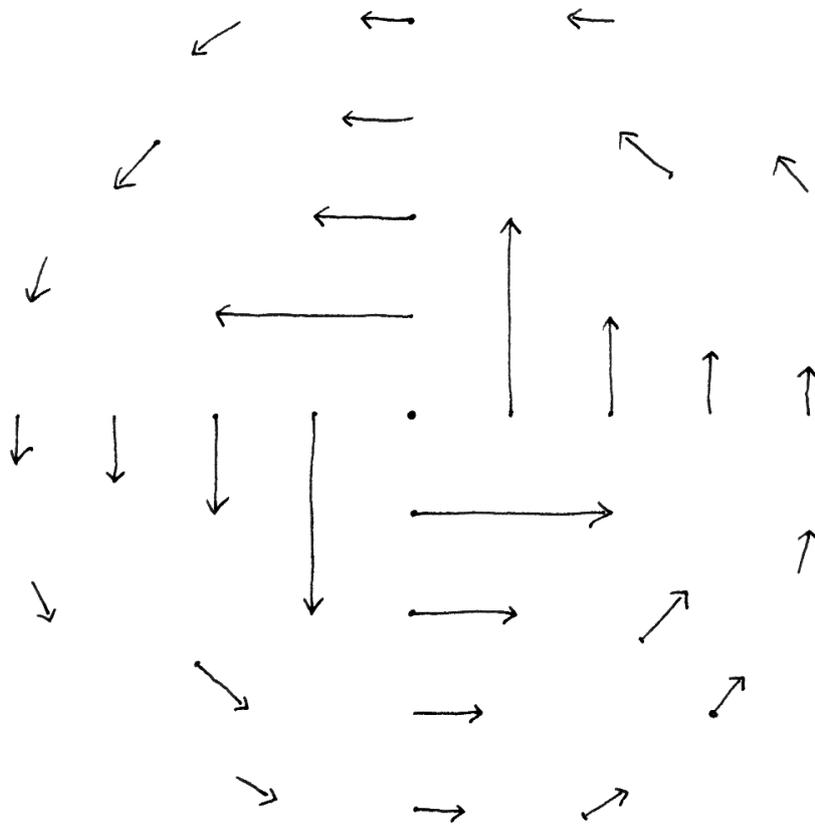
Im übrigen sieht man auch leicht ein, daß die $d_p(df)$ entsprechende symmetrische Bilinearform gerade das Doppelte des “quadratischen Anteils der Taylorentwicklung der Funktion f um p ” ist. Wie dem auch sei, bleibt nur $1 \Rightarrow 4$ zu zeigen. Wir dürfen ohne Beschränkung der Allgemeinheit weiter $X = \mathbb{R}^n$ annehmen. Wir beginnen unseren Beweis von $1 \Rightarrow 4$, indem wir einen Spezialfall von $1 \Rightarrow 3$ als Proposition formulieren und unabhängig zeigen.

Lemma 6.6.13. *Ist $A \subseteq \mathbb{R}^n$ eine offene Kugel und ω darauf ein stetig differenzierbares geschlossenes Kovektorfeld, so ist ω das Differential einer Funktion $f : A \rightarrow \mathbb{R}$.*

6.6.14. Ich gebe dafür zwei Beweise: Erst einen sehr kurzen mehr rechnerischen Beweis, und im Anschluß einen etwas längeren mehr konzeptionellen.

Rechnerischer Beweis. Wir verwenden unsere Notation 6.3.2. Wir dürfen ohne Beschränkung der Allgemeinheit $0 \in A$ annehmen, bezeichnen nun wieder mit x einen Vektor und betrachten die Funktion $f : A \rightarrow \mathbb{R}$ gegeben durch

$$f(x) = \int_0^x \omega = \int_0^1 \omega_{tx}(x) dt = \int_0^1 \sum_{j=1}^n u_j(tx) \cdot x_j dt$$



Daß ein rotationsfreies Vektorfeld auf einer nicht wegweise einfach zusammenhängenden offenen Teilmenge eines \mathbb{R}^n nicht notwendig ein Potential besitzt, zeigt das Vektorfeld $\text{grad } \theta$ auf $\mathbb{R}^2 \setminus 0$, wo $\theta(x, y)$ der eben nur bis auf eine additive Konstante wohlbestimmte Winkel ist, den der Strahl vom Nullpunkt nach (x, y) mit der horizontalen Koordinatenachse einschließt. Der Gradient $\text{grad } \theta$ ist dann ein wohldefiniertes rotationsfreies Vektorfeld auf dem Komplement des Ursprungs, hat aber kein global definiertes Potential. Dies Vektorfeld ist nicht ganz leicht zu zeichnen, da die Längen seiner Vektoren gegen den Ursprung hin ins Unendliche wachsen. Auf den ersten Blick mag es absurd wirken, dieses Feld wirbelfrei zu nennen. Eine außerhalb des Ursprungs zum Testen hereingelegte kleine Kreisscheibe würde aber in der Tat nicht gedreht, die stärkeren Vektoren zerran zwar an der dem Ursprung zugewandten Seite, aber von diesen Vektoren greifen andererseits auch weniger an. In gewisser Weise konzentriert sich hier das gesamte Wirbeln im Ursprung, und der gehört nun eben gerade nicht zu unserem Definitionsbereich. In mathematischer Sprechweise ist $d\theta$ ein geschlossenes Kovektorfeld auf der punktierten Ebene, das jedoch nicht das Differential einer global definierten Funktion ist.

Ihre partielle Ableitung nach x_i ergibt sich zu

$$\begin{aligned}
 \frac{\partial f}{\partial x_i}(p) &= \int_0^1 \frac{\partial}{\partial x_i} \Big|_{x=p} \left(\sum_{j=1}^n (u_j \circ (t \cdot)) \cdot x_j \right) dt \\
 &= \int_0^1 \sum_{j=1}^n t \cdot \frac{\partial u_j}{\partial x_i}(tp) \cdot p_j + u_i(tp) dt \\
 &= \int_0^1 \sum_{j=1}^n t \cdot \frac{\partial u_i}{\partial x_j}(tp) \cdot p_j + u_i(tp) dt \\
 &= \int_0^1 t \cdot \frac{d}{dt}(u_i(tp)) + u_i(tp) dt \\
 &= \int_0^1 \frac{d}{dt}(t \cdot (u_i(tp))) dt \\
 &= t \cdot u_i(tp) \Big|_0^1 \\
 &= u_i(p)
 \end{aligned}$$

und wir sehen, daß in der Tat gilt $df = \omega$. □

Konzeptioneller Beweis. Wir behandeln zunächst den Fall $n = 2$ als eigenständiges Lemma.

Lemma 6.6.15. *Ist $A \subseteq \mathbb{R}^2$ eine ebene Kreisscheibe und ω darauf ein stetig differenzierbares geschlossenes Kovektorfeld, so ist ω das Differential einer Funktion $f : A \rightarrow \mathbb{R}$.*

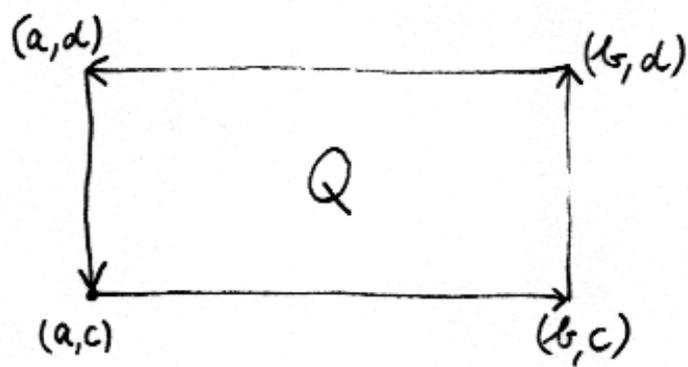
Beweis. Um Indizes zu vermeiden schreiben wir bei der Behandlung dieses Spezialfalls (x, y) statt (x_1, x_2) in der Hoffnung, daß dies Einsparen von Indizes mehr Klarheit schafft, als die Verwendung der Buchstaben x, y mit verschiedenen Bedeutungen an Verwirrung erzeugt. Betrachten wir ein Rechteck $Q = [a, b] \times [c, d] \subset A$ und integrieren unser Kovektorfeld einmal im Gegenuhrzeigersinn auf dem Rand entlang, den wir parametrisieren als Weg ρ , so erhalten wir

$$\begin{aligned}
 \int_{\rho} \omega &= \int_a^b u_1(x, c) dx + \int_c^d u_2(b, y) dy - \int_a^b u_1(x, d) dx - \int_c^d u_2(a, y) dy \\
 &= \int_c^d \int_a^b \left(\frac{\partial u_2}{\partial x} - \frac{\partial u_1}{\partial y} \right) dx dy
 \end{aligned}$$

Für ein stetig differenzierbares geschlossenes Kovektorfeld verschwindet also das Wegintegral einmal um den Rand unseres Rechtecks und der "obere" bzw. der "untere" Weg auf den Kanten des Rechtecks von einem Punkt zum diagonal gegenüberliegenden Punkt liefern dasselbe Wegintegral. Halten wir nun einen Punkt $(p, q) \in A$ fest, so definiert dieses gemeinsame Wegintegral eine Funktion

$$\begin{aligned}
 f(x, y) &= \int_p^x u_1(t, q) dt + \int_q^y u_2(x, s) ds \\
 &= \int_q^y u_2(p, s) ds + \int_p^x u_1(t, y) dt
 \end{aligned}$$

für die wegen der ersten Darstellung offensichtlich gilt $f_y = u_2$ und wegen der zweiten Darstellung $f_x = u_1$. Damit gilt $\omega = df$ wie behauptet. □



Das Rechteck aus dem Beweis von $4 \Rightarrow 1$.

Jetzt führen wir unseren konzeptionellen Beweis des Lemmas im Fall allgemeiner Dimension zu Ende. Wir betrachten dazu alle Wege, die längs der Kanten eines achsenparallelen Quaders vom Ursprung nach p laufen. Genauer betrachten wir für jede Permutation $\sigma \in \mathcal{S}_n$ den Weg $[\sigma] = [\sigma; p]$ vom Ursprung nach p , der gerade verläuft zwischen den Eckpunkten

$$0, p_{\sigma(1)}\vec{e}_{\sigma(1)}, p_{\sigma(1)}\vec{e}_{\sigma(1)} + p_{\sigma(2)}\vec{e}_{\sigma(2)}, \dots, p$$

Ist $\tau = (i, i + 1)$ eine Transposition benachbarter Zahlen, so unterscheiden sich $[\sigma]$ und $[\sigma \circ \tau]$ nur dadurch, daß sie beim i -ten und $(i + 1)$ -ten Geradenstück auf verschiedenen Kantenwegen diagonal gegenüberliegende Punkte eines ebenen Rechtecks verbinden. Ziehen wir unser Kovektorfeld auf eine geeignete Ebene zurück, so landen wir im bereits behandelten Fall und folgern

$$\int_{[\sigma]} \omega = \int_{[\sigma \circ \tau]} \omega$$

für jede Transposition τ der Gestalt $\tau = (i, i + 1)$. Wissen wir nun bereits nach ??, daß derartige Transpositionen die symmetrische Gruppe erzeugen, so können wir sofort folgern, daß $\int_{[\sigma]} \omega$ gar nicht von $\sigma \in \mathcal{S}_n$ abhängt. Die durch

$$f(p) = \int_{[\sigma; p]} \omega$$

für ein und alle σ definierte Funktion f hat dann Differential $df = \omega$, da ihre partielle Ableitung nach x_i auch aus jeder Darstellung durch ein σ mit $\sigma(n) = i$ berechnet werden kann, für die $\frac{\partial f}{\partial x_i} = u_i$ offensichtlich ist. \square

Jetzt können wir schließlich in unserem Satz 6.6.10 auch noch die Implikation $1 \Rightarrow 4$ zeigen. Sei $h : [0, 1]^2 \rightarrow A$ eine Homotopie zwischen unseren beiden Integrationswegen, die wir ohne Beschränkung der Allgemeinheit normiert annehmen dürfen. Analog wie beim Beweis von 6.4.5 zeigen wir mithilfe von 9.2.22 und 9.7.11, daß es für den Abstand von Punkten aus dem Bild unseres Einheitsquadrats und Punkten außerhalb von A eine positive untere Schranke gibt. Da h nach 9.7.14 gleichmäßig stetig ist, finden wir weiter ein $r \in \mathbb{N}$, $r \geq 1$ derart, daß bei Unterteilung des Einheitsquadrats in r^2 kleine Schachfelder der Kantenlänge $1/r$ die einzelnen Felder unter h jeweils ganz in einen offenen Ball in A abgebildet werden. Jetzt betrachten wir die Integrale längs der Geradensegmente zwischen den Bildern in A von benachbarten Ecken unserer Schachfelder

$$c_{i,j} = \int_{h(\frac{i}{r}, \frac{j}{r})}^{h(\frac{i+1}{r}, \frac{j}{r})} \omega \quad \text{und} \quad d_{i,j} = \int_{h(\frac{i}{r}, \frac{j}{r})}^{h(\frac{i}{r}, \frac{j+1}{r})} \omega$$

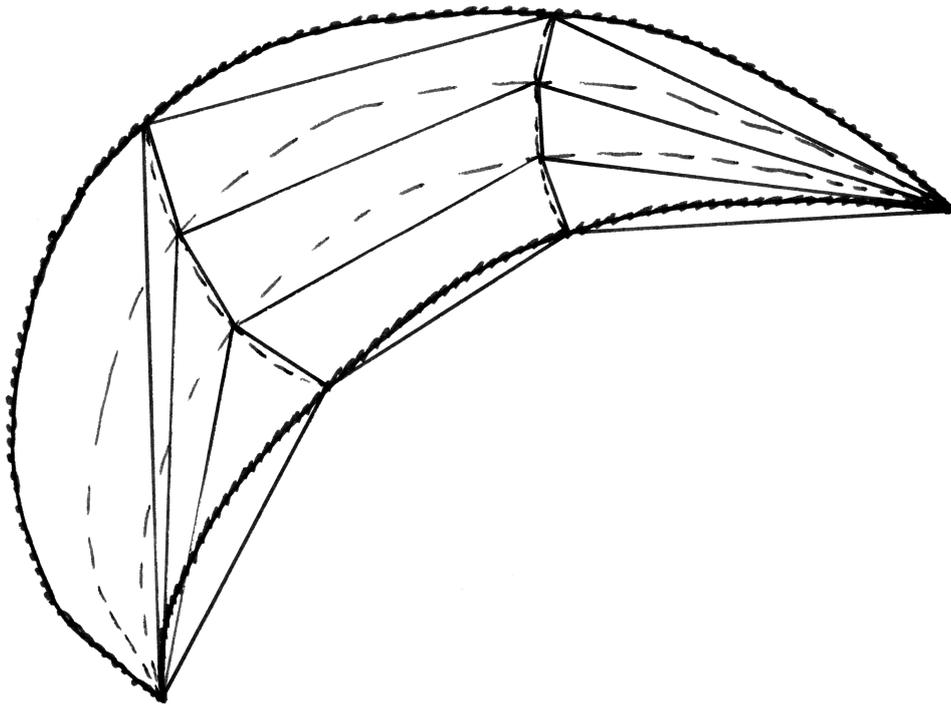


Illustration zum Beweis von Satz [6.6.10](#) über die Homotopieinvarianz von Wegintegralen bei gewissen Kovektorfeldern. Die beiden Wege werden durch dicke gezackte Linien dargestellt, die Homotopie zwischen ihnen durch feine gestrichelte Linien. Es gilt, diese Unterteilung so fein zu wählen, daß jeder dieser “Ziegel” ganz in einem im Definitionsbereich unserer geschlossenen Differentialform enthaltenen Ball liegt.

Indem wir die für konvexe Definitionsbereiche bereits bewiesene Implikation auf unsere offenen Bälle in A anwenden, finden wir $c_{i,j} + d_{i+1,j} - d_{i,j} - c_{i,j+1} = 0$ und durch Aufsummieren

$$\sum_{0 \leq i < r} c_{i,0} + \sum_{0 \leq j < r} d_{r,j} - \sum_{0 \leq j < r} d_{0,j} - \sum_{0 \leq i < r} c_{i,r} = 0$$

Indem wir nochmals den bereits bewiesenen Fall auf unsere offenen Bälle anwenden sehen wir dann weiter, daß diese vier Summen jeweils den Wegintegralen von ω über die durch die vier Kanten unseres Quadrats gegebenen Wege gleichen. Zwei von diesen Wegen sind eh konstant und die übrigen sind eben gerade γ und ψ . \square

Übung 6.6.16. Man zeige, daß gegeben $n \in \mathbb{Z}$ der geschlossene Weg $\gamma_n : [0, 2\pi] \rightarrow \mathbb{R}^2 \setminus \{0\}$ mit $\gamma(t) = (\cos nt, \sin nt)$ in $\mathbb{R}^2 \setminus \{0\}$ nur für $n = 0$ zusammenziehbar ist. Hinweis: Man berechne das Integral des Winkelfeldes über diesen Weg und beachte 6.6.10. Ich empfinde es allerdings als Umweg, diese Aussage mithilfe von Wegintegralen nachzuweisen, und ziehe den topologischen Beweis über Liftungseigenschaften in ?? folgende vor.

Ergänzende Übung 6.6.17. Man zeige, daß es in 6.6.1 sogar ausreicht, die Differenzierbarkeit unseres stetigen Kovektorfelds an fast allen Punkten von A vorauszusetzen.

Ergänzende Übung 6.6.18. Gegeben auf einem Rechteck $Q = [a, b] \times [c, d] \subset \mathbb{R}^2$ ein stetig differenzierbares Vektorfeld $v : Q \rightarrow \mathbb{R}^2$ stimmt das Integral seiner Rotation $\text{rot } v$ im Sinne von 6.6.12 über das Rechteck Q überein mit seinem Wegintegral im Sinne von 6.3.6 einmal im Gegenuhrzeigersinn um den Rand des Rechtecks. In 10.8.15 werden wir diese Aussage als Spezialfall des allgemeinen Stokes'schen Satzes zu verstehen lernen.

Satz 6.6.19 (Fundamentalsatz der Algebra). *Jedes nicht konstante komplexe Polynom besitzt mindestens eine komplexe Nullstelle.*

6.6.20. Andere Beweise werden in ?? diskutiert.

Beweis. Sei $P(z) = z^n + a_{n-1}z^{n-1} + \dots + a_0$ unser Polynom. Wir argumentieren durch Widerspruch und betrachten für jeden Radius $r > 0$ den geschlossenen Weg $\gamma : [0, 2\pi] \rightarrow \mathbb{C}$, $\gamma_r(t) = re^{it} = r \cos t + ir \sin t$, der einmal auf dem Kreis mit Radius r umläuft. Nach 6.5.4 ist er in \mathbb{C} zusammenziehbar. Hätte unser Polynom keine Nullstelle, so lieferte es eine stetige Abbildung $P : \mathbb{C} \rightarrow \mathbb{C}^\times$, und nach 6.5.5 wären alle $P \circ \gamma_r$ zusammenziehbar in \mathbb{C}^\times . Für hinreichend großes r gilt nun jedoch $r^n > |a_{n-1}|r^{n-1} + \dots + |a_1|r + |a_0|$, und für solche r ist der Weg $P \circ \gamma_r$ in \mathbb{C}^\times homotop zum Weg $t \mapsto \gamma_r(t)^n$, da nämlich für kein t die Strecke von

$P(\gamma_r(t))$ nach $\gamma_r(t)^n$ den Nullpunkt trifft. Hätte also P keine Nullstelle, so wäre der Weg $[0, 2\pi] \rightarrow \mathbb{C}^\times, t \mapsto \gamma_r(t)^n$ zusammenziehbar in \mathbb{C}^\times . Das steht jedoch im Widerspruch zu [6.6.16](#). \square

7 Umkehrsatz und Anwendungen

7.1 Der Satz über die Umkehrabbildung

Definition 7.1.1. Eine stetig differenzierbare Abbildung zwischen halboffenen Teilmengen normierter reeller Räume nennt man abkürzend auch eine \mathcal{C}^1 -**Abbildung**. Der Buchstabe \mathcal{C} steht hier für englisch “continuous” oder französisch “continu”, zu deutsch stetig, und der obere Index 1 deutet an, daß nur die Existenz und Stetigkeit der *ersten* Ableitung gefordert wird. Eine bijektive \mathcal{C}^1 -Abbildung, deren Umkehrung auch eine \mathcal{C}^1 -Abbildung ist, heißt ein \mathcal{C}^1 -**Diffeomorphismus**.

Satz 7.1.2 (über die Umkehrabbildung). *Seien X, Y endlichdimensionale reelle Räume, $V \subseteq X$ offen und $f : V \rightarrow Y$ stetig differenzierbar. Ist an einer Stelle $p \in V$ das Differential ein Isomorphismus $d_p f : \vec{X} \xrightarrow{\sim} \vec{Y}$, so induziert f einen \mathcal{C}^1 -Diffeomorphismus zwischen einer offenen Umgebung von p in V und einer offenen Umgebung von $f(p)$ in Y .*

7.1.3. Wir zeigen den Satz nach einigen Vorbereitungen zu Ende dieses Abschnitts. Die Kettenregel liefert im Übrigen für das Differential des Inversen eines \mathcal{C}^1 -Diffeomorphismus sofort die Formel

$$d_{f(p)}(f^{-1}) = (d_p f)^{-1}$$

7.1.4. Unser Satz gilt auch für unendlichdimensionale normierte Räume X, Y , wenn wir zusätzlich annehmen, daß ihre Richtungsräume vollständig sind und daß das Inverse des Vektorraumisomorphismus $d_p f$ stetig ist. Der Beweis bleibt derselbe. Diese Allgemeinheit ist durchaus auch von Interesse, eröffnet sie doch einen direkten Zugang zum Studium von Lösungen gewöhnlicher Differentialgleichungen, wie in 8.5.2 ausgeführt wird. Als erste Prüfung für unseren Satz über die Umkehrabbildung überlege man sich die Gültigkeit der Behauptung im Spezialfall $X = Y = \mathbb{R}$.

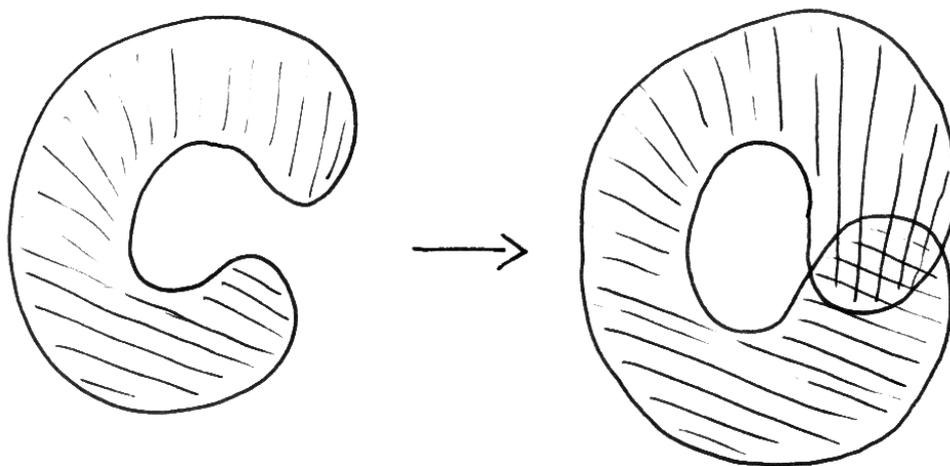
Definition 7.1.5. Ein Punkt, der unter einer Abbildung auf sich selbst abgebildet wird, heißt ein **Fixpunkt** besagter Abbildung.

Definition 7.1.6. Eine Abbildung f zwischen zwei metrischen Räumen heißt **Lipschitz-stetig** genau dann, wenn es eine Konstante $\lambda > 0$ gibt mit

$$d(f(x), f(y)) \leq \lambda d(x, y)$$

für alle x, y im Ausgangsraum. Wir sagen dann auch, f sei lipschitzstetig **zur Lipschitz-Konstante** λ . Eine Abbildung f zwischen metrischen Räumen heißt **kontrahierend** genau dann, wenn sie lipschitzstetig ist zu einer Lipschitzkonstante $\lambda < 1$, wenn es also $\lambda < 1$ gibt mit

$$d(f(x), f(y)) \leq \lambda d(x, y) \quad \forall x, y$$



Eine stetig differenzierbare Abbildung von einer offenen zusammenhängenden Teilmenge $V \subseteq \mathbb{R}$ nach \mathbb{R} mit nirgends verschwindender Ableitung ist notwendig injektiv. Eine stetig differenzierbare Abbildung von einer offenen zusammenhängenden Teilmenge $V \subseteq \mathbb{R}^2$ in den \mathbb{R}^2 mit überall injektivem Differential ist dahingegen im allgemeinen nur noch “lokal injektiv” in einer Weise, die der Satz über die Umkehrabbildung spezifiziert.

Beispiel 7.1.7. Natürlich ist jede Lipschitzstetige Abbildung stetig. Eine Abbildung $\mathbb{R} \rightarrow \mathbb{R}$ ist Lipschitzstetig zu einer Lipschitzkonstante λ genau dann, wenn alle ihre Sekantensteigungen betragsmäßig beschränkt sind durch λ . Es ist hoffentlich auch anschaulich klar, daß im Fall $\lambda < 1$ der Graph einer derart “flachen” Funktion an genau einer Stelle die Hauptdiagonale alias den Graphen der Identität kreuzen muß. Der Cosinus ist als Abbildung $\mathbb{R} \rightarrow \mathbb{R}$ keineswegs kontrahierend, das Supremum seiner Sekantensteigungen ist ja Eins. Die Einschränkung des Cosinus zu einer Abbildung $[0, 1] \rightarrow [0, 1]$ ist jedoch kontrahierend, und den im Banach’schen Fixpunktsatz versteckten Algorithmus zur Bestimmung des Fixpunktes illustriert nebenstehende Abbildung.

Lemma 7.1.8 (Banach’scher Fixpunktsatz). *Jede kontrahierende Selbstabbildung eines nichtleeren vollständigen metrischen Raums besitzt genau einen Fixpunkt.*

Beweis. Sei $f : M \rightarrow M$ unsere kontrahierende Selbstabbildung und $\lambda < 1$ eine Lipschitzkonstante. Wir wählen $x_0 \in M$ und betrachten die rekursiv definierte Folge $x_{n+1} = f(x_n)$. Mit Induktion folgt $d(x_n, x_{n+1}) \leq \lambda^n d(x_0, x_1)$, und mit der Dreiecksungleichung folgt für $n \leq m$ bereits

$$d(x_n, x_{m+1}) \leq (\lambda^n + \lambda^{n+1} + \dots + \lambda^m) d(x_0, x_1) \leq \frac{\lambda^n}{1 - \lambda} d(x_0, x_1)$$

Also ist unsere Folge x_n eine Cauchy-Folge und konvergiert gegen einen Punkt $\lim_{n \rightarrow \infty} x_n = p \in M$. Da nun eine kontrahierende Abbildung notwendig stetig ist, folgt aus der Vertauschbarkeit nach 9.3.9 von Grenzwerten mit dem Anwenden stetiger Funktionen

$$f(p) = \lim_{n \rightarrow \infty} f(x_n) = \lim_{n \rightarrow \infty} x_{n+1} = p$$

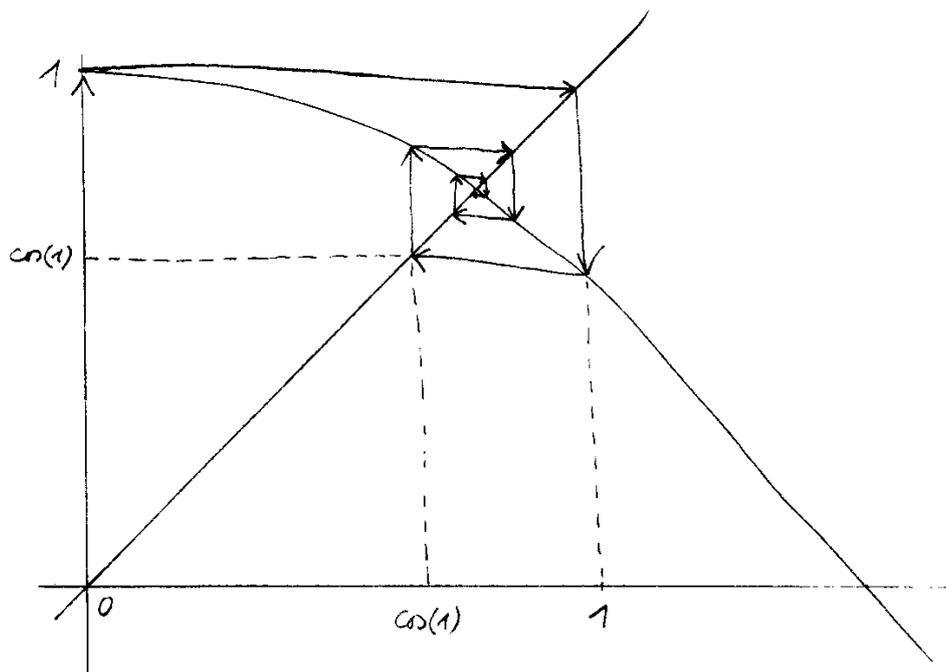
und wir haben schon mal einen Fixpunkt gefunden. Ist q ein zweiter Fixpunkt, so folgt $d(p, q) = d(f(p), f(q)) \leq \lambda d(p, q)$ für $\lambda < 1$ und damit $d(p, q) = 0$, also $p = q$. \square

Ergänzung 7.1.9. Lassen wir in der Ungleichungskette aus obigem Beweis m nach Unendlich streben, so erhalten wir für den Abstand der n -ten Approximation x_n zum Fixpunkt p zusätzlich die Abschätzung

$$d(x_n, p) \leq \frac{\lambda^n}{1 - \lambda} d(x_0, x_1)$$

Satz 7.1.10 (Umkehrsatz für stetige Abbildungen). *Seien X ein vollständiger normierter reeller affiner Raum, $U \subseteq X$ eine offene Teilmenge und $f : U \rightarrow X$ eine Abbildung, deren Differenz zur Identität*

$$(f - \text{id}) : U \rightarrow \vec{X}$$



Schaltet man den Taschenrechner ein, stellt auf Bogenmaß und drückt wiederholt die Taste “cos”, so wird man feststellen, daß sich die Zahl in der Anzeige nach einer gewissen Zeit überhaupt nicht mehr ändert. Um das zu verstehen beachte man, daß der Cosinus eine kontrahierende Selbstabbildung des Intervalls $[0, 1]$ liefert, da nämlich seine Ableitung dort betragsmäßig durch $\sin(1) \in (0, 1)$ beschränkt ist. Der Beweis des Fixpunktsatzes [7.1.8](#) zeigt, daß unter diesen Umständen das wiederholte Anwenden stets eine Folge liefert, die gegen den Fixpunkt konvergiert.

kontrahierend ist, d.h. Lipschitzstetig zu einer Lipschitzkonstante $\lambda < 1$. So ist unsere Abbildung injektiv mit offenem Bild $f(U) \subseteq X$ und ihre Umkehrabbildung $f^{-1} : f(U) \rightarrow U$ ist Lipschitzstetig zur Lipschitzkonstante $1/(1 - \lambda)$.

Übung 7.1.11. Der Leser mag zur Übung aus den Argumenten des anschließenden Beweises folgern, daß unter den Annahmen des Satzes genauer gilt

$$B(p; R) \subset U \Rightarrow B(f(p); (1 - \lambda)R) \subset f(U)$$

Beweis. Ohne Beschränkung der Allgemeinheit dürfen wir annehmen, X sei ein vollständiger normierter Vektorraum. Die Injektivität von f ergibt sich, da aus unserer Annahme $\|(f - \text{id})(x) - (f - \text{id})(y)\| \leq \lambda\|x - y\|$ sofort folgt $\|y - x\| - \|f(x) - f(y)\| \leq \lambda\|x - y\|$ alias

$$\|f(x) - f(y)\| \geq (1 - \lambda)\|x - y\|$$

Durch Einsetzen von $x = f^{-1}(p)$ und $y = f^{-1}(q)$ folgt ohne Schwierigkeiten $\|p - q\| \geq (1 - \lambda)\|f^{-1}(p) - f^{-1}(q)\|$ und damit sogar die Lipschitzstetigkeit der Umkehrfunktion $f^{-1} : f(U) \rightarrow U$. Bis hierher brauchen wir weder U offen noch X vollständig anzunehmen, und unsere Aussagen sind wenig mehr als ein Spezialfall der allgemeinen Resultate für “nicht abstandsverkleinernde” Abbildungen metrischer Räume aus Übung 9.2.26. Der wesentliche Punkt besteht darin, zu zeigen, daß f offenes Bild hat. Dazu betrachten wir für alle $y \in X$ die Abbildung

$$\begin{aligned} k_y : U &\rightarrow X \\ x &\mapsto x - f(x) + y \end{aligned}$$

Ihre Fixpunkte sind die Urbilder von y unter f und k_y ist unter unseren Annahmen auch kontrahierend. Genauer gilt offensichtlich

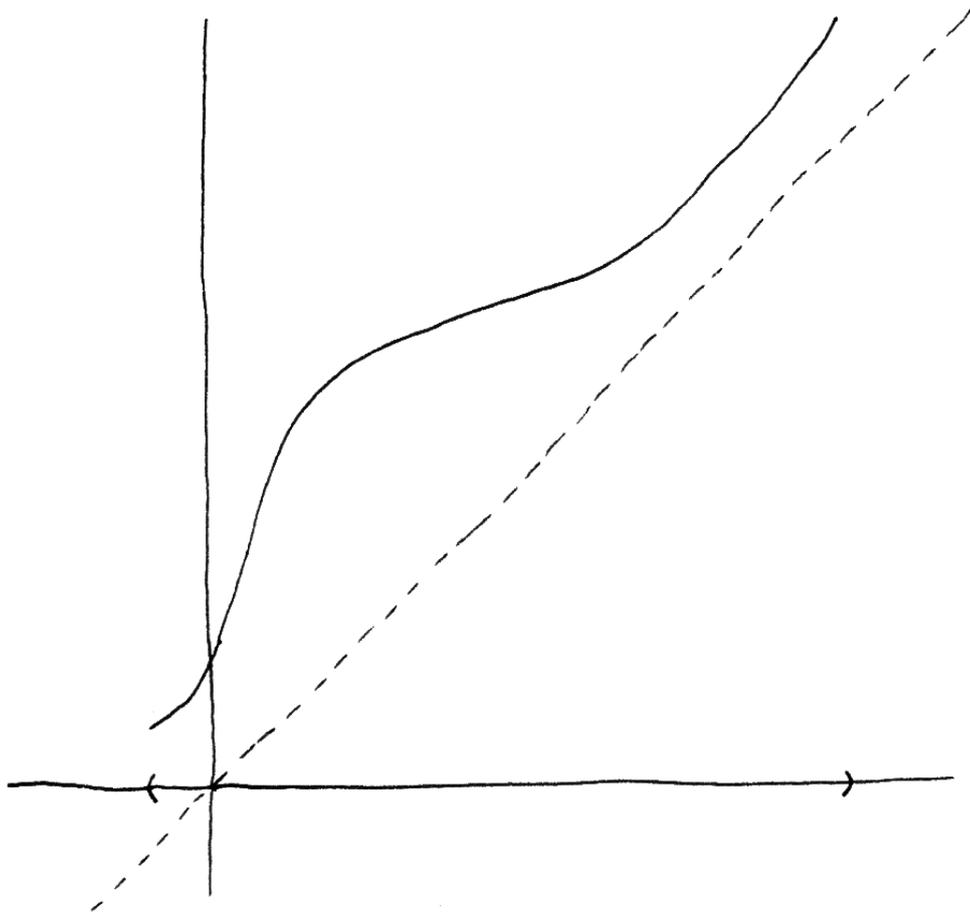
$$\|k_y(x) - k_y(z)\| \leq \lambda\|x - z\| \quad \forall x, z \in U$$

Gegeben $p \in U$ finden wir nun einen Radius $R > 0$ derart, daß der “abgeschlossene Ball”

$$A(p; R) := \{x \in X \mid \|p - x\| \leq R\}$$

ganz in U enthalten ist. Für $y \in B(f(p); (1 - \lambda)R)$ bildet dann k_y unseren abgeschlossenen Ball $A(p; R)$ in sich ab, denn es gilt

$$\begin{aligned} \|p - x\| \leq R &\Rightarrow \|k_y(p) - k_y(x)\| \leq \lambda R \\ &\Rightarrow \|(p - k_y(x)) + (y - f(p))\| \leq \lambda R \\ &\Rightarrow \|p - k_y(x)\| - \|y - f(p)\| \leq \lambda R \\ &\Rightarrow \|p - k_y(x)\| \leq R \end{aligned}$$



Für eine Abbildung f von einer offenen Teilmenge $U \subseteq \mathbb{R}$ nach \mathbb{R} ist $(f - \text{id})$ kontrahierend mit Kontraktionsfaktor λ genau dann, wenn alle ihre Sekantensteigungen im Intervall $[1 - \lambda, 1 + \lambda]$ liegen. In diesem Fall sollte es anschaulich klar sein, daß f injektiv und offen ist und daß die Sekantensteigungen der Umkehrfunktion betragsmäßig beschränkt sind durch $(1 - \lambda)^{-1}$: In diesem Fall liegen alle Sekantensteigungen der Umkehrfunktion sogar im Intervall $[(1 + \lambda)^{-1}, (1 - \lambda)^{-1}]$.

Wenden wir den Banach'schen Fixpunktsatz 7.1.8 auf die Selbstabbildung k_y von $A(p; R)$ an, das nach 10.5.2 als abgeschlossene Teilmenge eines vollständigen metrischen Raums auch vollständig ist, so finden wir, daß k_y einen Fixpunkt in $A(p; R)$ haben muß. Es folgt $f(A(p; R)) \supset B(f(p); (1 - \lambda)R)$, und das zeigt, daß das Bild von f offen sein muß, in Formeln $f(U) \Subset X$. \square

Beweis des Umkehrsatzes 7.1.2. Ohne Beschränkung der Allgemeinheit seien X, Y Vektorräume. Sicher reicht es, wenn wir das Lemma für $(d_p f)^{-1} \circ f$ statt für f zeigen. Wir dürfen also ohne Beschränkung der Allgemeinheit $X = Y$ und $d_p f = \text{id}$ annehmen. Es folgt $d_p(f - \text{id}) = 0$. Wählen wir eine Norm auf X und beachten, daß f stetig differenzierbar angenommen war, so folgt leicht die Existenz eines offenen Balls B um p mit $\|d_x(f - \text{id})\| \leq 1/2 \quad \forall x \in B$. Nach 4.3.5 ist dann jedoch $(f - \text{id}) : B \rightarrow X$ kontrahierend mit Lipschitzkonstante $\leq (1/2)$. Mit unserem Umkehrsatz für stetige Abbildungen 7.1.10 folgt, daß f eine Injektion mit offenem Bild $f : B \hookrightarrow X$ liefert, deren Umkehrung lipschitzstetig ist zur Lipschitzkonstante Zwei. Um die Differenzierbarkeit von $f^{-1} : f(B) \rightarrow B$ an der Stelle $f(p)$ zu zeigen, dürfen wir ohne Beschränkung der Allgemeinheit zusätzlich annehmen, daß gilt $p = 0$ und $f(p) = 0$. Da wir f differenzierbar bei p mit Differential id angenommen hatten, können wir dann schreiben

$$f(x) = x + \|x\|\varepsilon(x)$$

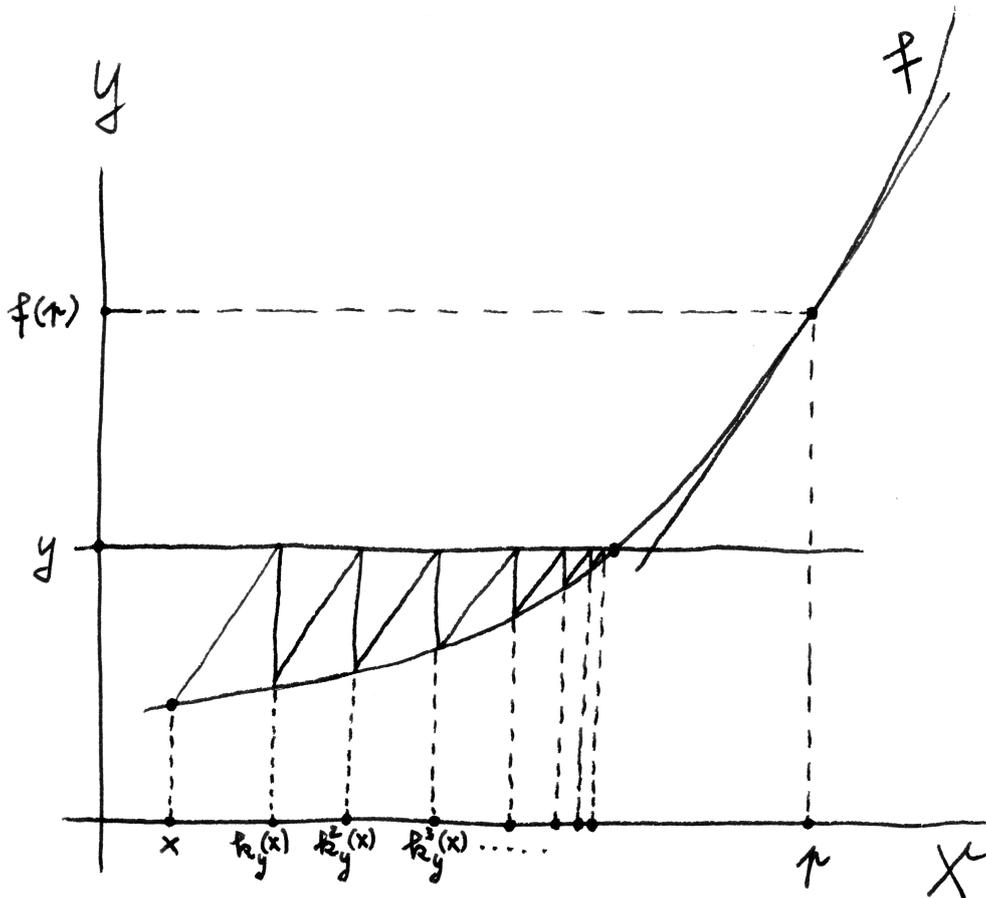
für eine geeignete Abbildung $\varepsilon : B \rightarrow X$, die stetig ist bei 0 und die dort den Wert Null annimmt. Setzen wir hier $x = f^{-1}(y)$ ein mit $y \in f(B)$, so ergibt sich

$$y = f^{-1}(y) + \|f^{-1}(y)\| \varepsilon(f^{-1}(y))$$

Nun liefert die Lipschitzstetigkeit der Umkehrfunktion für unseren eben gewählten offenen Ball B nach 7.1.10 auch für alle $y \in f(B)$ die Abschätzung $\|f^{-1}(y)\| \leq 2\|y\|$. Damit ergibt sich dann leicht

$$\lim_{y \rightarrow 0} \frac{f^{-1}(y) - y}{\|y\|} = 0$$

und das besagt gerade, daß die Umkehrabbildung f^{-1} bei $y = 0$ differenzierbar ist mit Differential id . Verkleinern wir unsere offene Umgebung von p noch so weit, daß df dort überall invertierbar ist, so folgt die Differenzierbarkeit der Umkehrabbildung an jeder Stelle des Bildes unserer verkleinerten offenen Umgebung. Die Stetigkeit des Differentials der Umkehrabbildung folgt dann leicht aus der Stetigkeit des Differentials der ursprünglichen Abbildung und der Stetigkeit des Invertierens linearer Abbildungen, vergleiche im Fall von Banachräumen etwa 4.4.11. \square



Das Verfahren aus dem Beweis von 7.1.2 ist auch durchaus von praktischer Bedeutung: Ausgeschrieben besagt es, daß wir eine Lösung x der Gleichung $f(x) = y$ unter geeigneten Annahmen finden können als den Fixpunkt der kontrahierenden Abbildung

$$k_y : U \rightarrow X$$

$$x \mapsto x + (d_p f)^{-1}(y - f(x))$$

für p mit $f(p)$ hinreichend nah bei y . Es ist dem Newtonverfahren aus 5.3.3 eng verwandt, stimmt jedoch nicht damit überein: Beim Newtonverfahren etwa im Fall einer Veränderlichen “gehen wir ja immer auf der Tangente bei $(x, f(x))$ wieder herunter zur x -Achse”, wohingegen wir bei unserer Korrektur k_y aus besagtem Beweis stattdessen auf der Parallelen durch $(x, f(x))$ zur Tangente bei $(p, f(p))$ heruntergehen, wie im Bild dargestellt.

Übung 7.1.12. Eine stetig differenzierbare Abbildung von einer offenen Teilmenge eines endlichdimensionalen reellen Raums in einen weiteren endlichdimensionalen reellen Raum hat offenes Bild, wenn ihr Differential in jedem Punkt surjektiv ist. Ist unsere stetig differenzierbare Abbildung zusätzlich injektiv, so liefert sie einen Diffeomorphismus unserer offenen Teilmenge mit ihrem Bild.

7.2 Der Satz über implizite Funktionen

7.2.1. Gegeben eine differenzierbare Funktion $f : \mathbb{R}^3 \rightarrow \mathbb{R}$ von drei Veränderlichen scheint es mir anschaulich klar, daß ihre Niveauläche N durch einen vorgegebenen Punkt $p \in \mathbb{R}^3$ “lokal um p als Graph einer reellwertigen Funktion der ersten beiden Veränderlichen dargestellt werden kann”, wenn die Tangentialebene an unsere Niveauläche bei p nicht auf der xy -Ebene alias der Nullstellenmenge der dritten Veränderlichen senkrecht steht, wenn also die partielle Ableitung unserer Funktion f nach der dritten Veränderlichen bei p nicht verschwindet. Etwas formaler sollte es also unter der Bedingung $f_z(p) \neq 0$ offene Teilmengen $C_1 \subseteq \mathbb{R}^2$ und $A_1 \subseteq \mathbb{R}$ geben derart, daß gilt $p \in C_1 \times A_1$ und daß $N \cap (C_1 \times A_1)$ der Graph einer Funktion $g : C_1 \rightarrow A_1$ ist. Derartige “implizit definierte Funktionen” werden wir im folgenden in voller Allgemeinheit studieren, allerdings mit einer anderen Konvention der Bezeichnung der Variablen: Statt (x, y, z) schreiben wir im folgenden (z_1, z_2, x) , mit y bezeichnen wir Koordinaten des Raums, in dem unsere Abbildung f landet, da kommt nämlich keineswegs nur die reelle Zahlengerade in Betracht, und dabei fassen wir f bei vorgegebenem y auf als eine Gleichung in x , die von zwei Parametern z_1, z_2 abhängt.

7.2.2. Gegeben allgemeiner nicht nur eine einzige Gleichung in mehreren Veränderlichen, sondern ein ganzes Gleichungssystem mit weniger Gleichungen als Unbekannten, sagen wir mit nur m Gleichungen für $m + n$ Unbekannte, wird man im allgemeinen erwarten, daß sich die Lösungsmenge in der folgenden Weise beschreiben läßt: Wir dürfen n Unbekannte frei wählen und können die übrigen m Unbekannten dann aus unseren m Gleichungen als Funktionen der bereits gewählten n Unbekannten bestimmen. Mir hilft es beim Denken und Reden, n der Unbekannten als “Parameter” zu bezeichnen und unser System als ein System von m Gleichungen in m Unbekannten anzusehen, bei dem unsere m Gleichungen noch von insgesamt n Parametern abhängen. Besteht unser Gleichungssystem etwa aus m nicht notwendig homogenen linearen Gleichungen in m Unbekannten, wobei diese m Gleichungen noch stetig von n Parametern abhängen, und ist die $(m \times m)$ -Matrix der Koeffizienten für eine Wahl der Parameter invertierbar, so wird sie auch für benachbarte Parameter invertierbar bleiben, und wir können die eindeutig bestimmten Lösungen für benachbarte Parameter als Funktionen der Parameter schreiben. Sind unsere Gleichungen nicht linear, so gelten entsprechende

Aussagen. Eine präzise Formulierung gibt der “Satz über implizite Funktionen”. Mit den “impliziten” Funktionen sind dabei diejenigen Funktionen gemeint, die die m Unbekannten als Funktionen der n Parameter ausdrücken: Diese Funktionen sind nämlich a priori nicht explizit etwa als Polynome oder allgemeiner als algebraische Ausdrücke in bekannten Funktionen gegeben, sondern nur implizit als Lösungen eines Gleichungssystems.

7.2.3. Ich will den Fall, daß unsere Gleichungen linear sind und sogar linear von den Parametern abhängen, auch noch koordinatenfrei und etwas präziser formulieren. Seien dazu ganz allgemein X, Y, Z Vektorräume. Die Elemente von Z spielen im folgenden die Rolle der “Parameter”. Ist $f : Z \times X \rightarrow Y$ eine lineare Abbildung, deren Restriktion auf X bijektiv ist, so gibt es genau eine Abbildung $g : Z \rightarrow X$ mit $f(z, g(z)) = 0$ für alle $z \in Z$. In der Tat ist unsere Gleichung wegen der Linearität von f gleichbedeutend zu $f(0, g(z)) + f(z, 0) = 0$, und nach Annahme gibt es eben genau ein $x \in X$ mit $f(0, x) = -f(z, 0)$. Bezeichnet in Formeln in_X die lineare Einbettung $x \mapsto (0, x)$ von X nach $Z \times X$, so haben wir $g(z) = -(f \circ \text{in}_X)^{-1}(f(z, 0))$, und notieren wir auch noch in_Z die lineare Einbettung $z \mapsto (z, 0)$ von Z nach $Z \times X$, so erhalten wir

$$g = -(f \circ \text{in}_X)^{-1} \circ (f \circ \text{in}_Z)$$

Der Satz über implizite Funktionen besagt, daß ähnliche Aussagen “lokal” auch allgemeiner für beliebige stetig differenzierbare Abbildungen gelten.

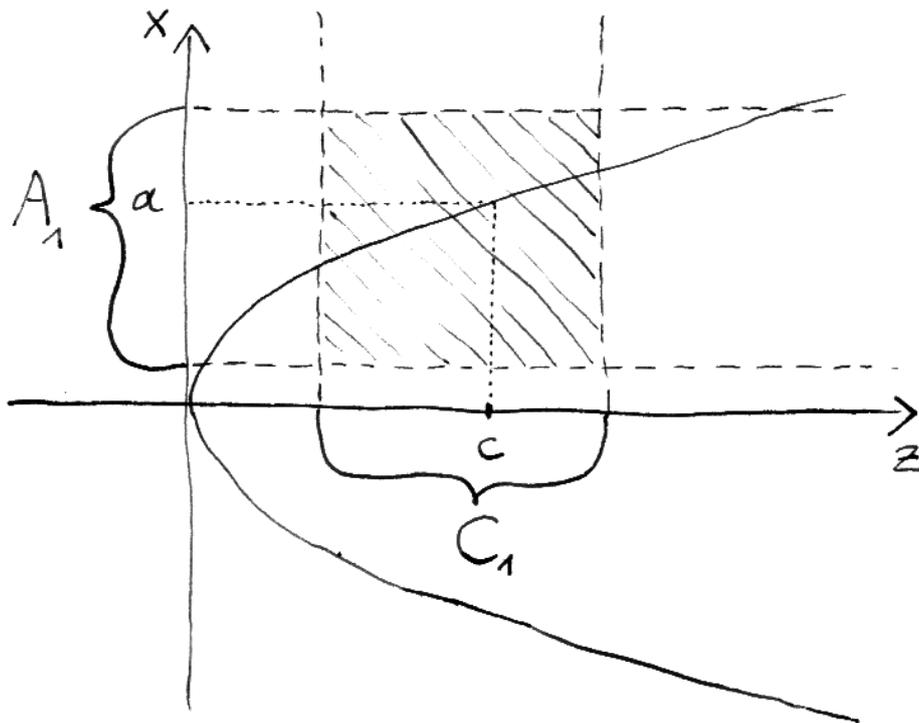
Satz 7.2.4 (über implizite Funktionen). *Seien Z, X, Y endlichdimensionale reelle Räume, $C \subseteq Z$ und $A \subseteq X$ offen und $f : C \times A \rightarrow Y$ stetig differenzierbar. Induziert an einer Stelle $(c, a) \in C \times A$ die Restriktion des Differentials von f eine Bijektion*

$$(d_{(c,a)}f) \circ \text{in}_{\vec{X}} : \vec{X} \xrightarrow{\sim} \vec{Y}$$

so existieren Paare (C_1, A_1) bestehend aus einer offenen Umgebung $C_1 \subseteq C$ von c und einer offenen Umgebung $A_1 \subseteq A$ von a derart, daß die Niveaumenge $\{(z, x) \in C_1 \times A_1 \mid f(z, x) = f(c, a)\}$ der Graph einer stetig differenzierbaren Funktion $g : C_1 \rightarrow A_1$ ist. Diese **implizite Funktion** g hat dann bei $z \in C_1$ das Differential

$$d_z g = -(d_{(z,g(z))}f \circ \text{in}_{\vec{X}})^{-1} \circ (d_{(z,g(z))}f \circ \text{in}_{\vec{Z}})$$

7.2.5. Wir sagen unter diesen Umständen, die Funktion g sei **implizit definiert durch die Bedingung** $f(z, g(z)) = b$ für $b = f(c, a)$. Es kann durchaus verschiedene Paare (C_1, A_1) und (C'_1, A'_1) geben, für die die Bedingungen des Satzes erfüllt sind. Im Fall von nicht zusammenhängendem C_1 kann die implizite Funktion auch durchaus von der Wahl von A_1 abhängen: Zwei Paare (C_1, A_1) und (C_1, A'_1) mit demselben C_1 können also durchaus verschiedene implizite Funktionen $C_1 \rightarrow A$ liefern, wie nebenstehendes Bild illustriert.



Der Schnitt der Niveaufläche $\{(z, x) \mid f(z, x) = x^2 - z = 0\}$ mit einem geeigneten Rechteck $C_1 \times A_1$ ist der Graph einer Funktion, eben der entsprechenden impliziten Funktion, hier der Funktion $g(z) = \sqrt{z}$. Dieses Beispiel zeigt auch, daß im Fall von nicht zusammenhängendem C_1 die implizite Funktion durchaus von der Wahl von A_1 abhängen kann: Zwei Paare (C_1, A_1) und (C_1, A'_1) mit demselben C_1 können verschiedene implizite Funktionen $C_1 \rightarrow A$ liefern! Die Tangente an den Graphen unserer impliziten Funktion g schließlich steht in (c, a) senkrecht auf dem Gradienten $(f_z(c, a), f_x(c, a))$, so daß auch die Formel für die Tangentensteigung $g'(c) = -f_z(c, a)/f_x(c, a)$ anschaulich unmittelbar einleuchtet.

7.2.6. Die Kettenregel zeigt, daß die lineare Abbildung $(d_{(c,a)}f) \circ \text{in}_{\vec{X}}$ auch als das Differential bei a der Abbildung $x \mapsto f(c, x)$ verstanden werden kann. Wenn also dies Differential invertierbar ist für einen Parameter c , so kann “jedes kleine Verwackeln von c zu z eindeutig durch ein kleines Verwackeln von a zu x in dem Sinne ausgeglichen werden, daß unter diesem simultanen Wackeln der Wert von f konstant bleibt”.

Ergänzung 7.2.7. Wir werden zusätzlich zeigen, daß wir im Satz eine offene Umgebung $C_1 \subseteq C$ von c sogar so wählen können, daß es (1) für jede darin enthaltene zusammenhängende offene Umgebung U von c genau eine stetige Funktion $g : U \rightarrow A$ gibt mit $g(c) = a$ und $f(z, g(z)) = f(c, a)$ für alle $z \in U$ und daß (2) diese Funktionen g stets stetig differenzierbar sind.

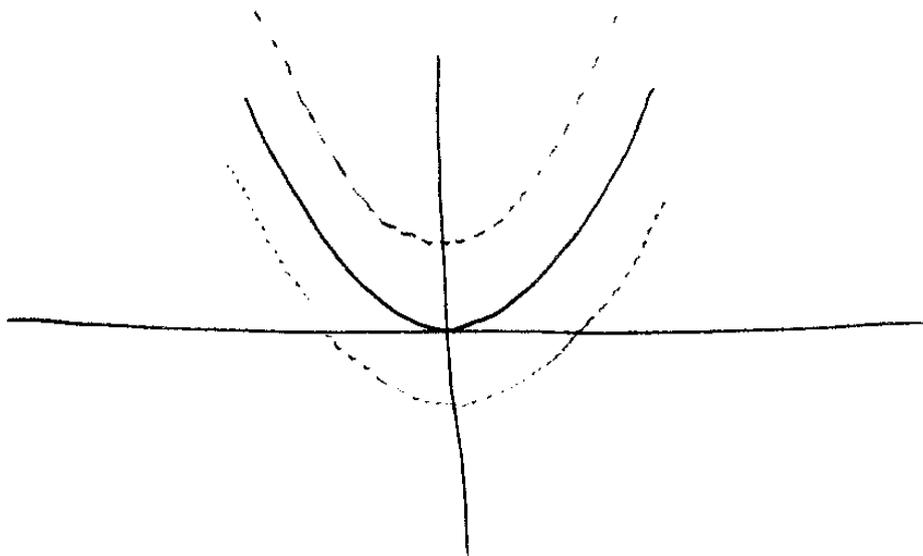
7.2.8. Wieder gilt der Satz mit demselben Beweis auch für vollständige normierte Räume, wenn man zusätzlich zur Bijektivität auch noch die Stetigkeit der inversen Abbildung von $(d_{(c,a)}f) \circ \text{in}_{\vec{X}} : \vec{X} \xrightarrow{\sim} \vec{Y}$ fordert und “stetig differenzierbar” im Sinne von 4.5.2 interpretiert. Diese Allgemeinheit wird sich bei der Lösung gewöhnlicher Differentialgleichungen 8.5.2 als nützlich erweisen.

Beispiel 7.2.9. Wir betrachten $f(z, x) = x^3 + zx^2 - 3$ als eine Schar von Polynomen in x mit Parameter z . Für $z = 2$ ist $x = 1$ eine Nullstelle, $f(2, 1) = 0$. Wir wollen nun untersuchen, wie diese Nullstelle “mit dem Parameter z wackelt”, und wenden dazu den Satz über implizite Funktionen an. Die partielle Ableitung nach x ist $f_x = 3x^2 + 2zx$, insbesondere haben wir $f_x(2, 1) = 7 \neq 0$ und nach dem Satz über implizite Funktionen gibt es folglich $\varepsilon > 0$ und $\delta > 0$ derart, daß für alle $z \in (2 - \varepsilon, 2 + \varepsilon)$ das Polynom $f(z, \cdot)$ genau eine Nullstelle $x = g(z) \in (1 - \delta, 1 + \delta)$ hat. Diese Funktion g ist zwar schwer explizit anzugeben, aber der Satz sagt uns, daß sie in einer Umgebung von $z = 2$ differenzierbar ist, und ihre Ableitung bei $z = 2$ kennen wir auch: Wir haben nämlich $f_z = x^2$, $f_z(2, 1) = 1$ und folglich

$$g'(2) = -f_x(2, 1)^{-1} f_z(2, 1) = -\frac{1}{7}$$

Betrachten wir stattdessen $h(z, x) = x^2 - z$, so ist für $z = 0$ natürlich $x = 0$ eine Nullstelle, aber Entsprechendes gilt keineswegs: Schieben wir z etwas ins Negative, so hat $h(z, \cdot)$ überhaupt keine reelle Nullstelle mehr, schieben wir z dahingegen etwas ins Positive, so werden aus unserer Nullstelle plötzlich zwei Nullstellen, wie das nebenstehende Bild illustriert. Das zeigt, daß die Bedingung an die Ableitung auch wirklich notwendig ist.

7.2.10. Ich will den Satz über implizite Funktionen auch noch in Koordinaten angeben. Seien dazu $C \subseteq \mathbb{R}^n$ und $A \subseteq \mathbb{R}^m$ offen, $f : C \times A \rightarrow \mathbb{R}^m$ stetig differenzierbar und $z_1, \dots, z_n, x_1, \dots, x_m$ unsere “Parameter” und “Unbekannten”. So gilt:



Die Graphen von $h(0, \cdot)$, $h(-1/2, \cdot)$ und $h(1/2, \cdot)$ für die Funktion $h(z, x) = x^2 - z$ vom Ende des Beispiels 7.2.9. Hier haben wir anders als sonst x in horizontaler Richtung aufgetragen und denken an Nullstellen der veränderten Gleichung statt an Schnitte der Niveaulinie $h(z, x) = x^2 - z = 0$ mit verschiedenen horizontalen Geraden. Dies Bild interpretiert also diesselbe Formeln wie das vorhergehende Bild auf Seite 139, nur eben in der alternativen Anschauung der “Abhängigkeit der Nullstellen von den Parametern” und in einem entsprechend an der Hauptdiagonale gespiegeltem Koordinatensystem.

1. Ist die Matrix

$$\left(\frac{\partial f_i}{\partial x_j} \right)_{i,j=1}^m$$

nicht ausgeartet an einer Stelle $(c, a) \in C \times A$, so existieren Tripel (C_1, A_1, g) bestehend aus einer offenen Umgebung $C_1 \subseteq C$ von c , einer offenen Umgebung $A_1 \subseteq A$ von a und einer stetig differenzierbaren Funktion $g : C_1 \rightarrow A_1$ derart, daß an jeder Stelle $z \in C_1$ der Wert $g(z)$ das einzige $x \in A_1$ ist mit $f(z, x) = f(c, a)$.

2. Die partiellen Ableitungen der Komponenten von g werden dann gegeben durch die Matrix-Gleichung

$$\left(\frac{\partial g_j}{\partial z_k} \Big|_z \right) = - \left(\frac{\partial f_i}{\partial x_j} \Big|_{(z, g(z))} \right)^{-1} \left(\frac{\partial f_i}{\partial z_k} \Big|_{(z, g(z))} \right)$$

In dieser Sprache ausgedrückt kann also ein System von m Gleichungen in $m+n$ Unbekannten im allgemeinen "lokal" in der Weise aufgelöst werden, daß wir n der Unbekannten frei wählen und die restlichen m Unbekannten dadurch dann im Wesentlichen eindeutig festgelegt werden.

Beispiel 7.2.11. Wir betrachten $f : \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$, $f(z, x) = z^2 + x^2$ und $(c, a) = (0, 1)$. Da $\frac{\partial f}{\partial x}$ nicht verschwindet bei $(0, 1)$, sind die Voraussetzungen des Satzes erfüllt. Ein mögliches Tripel bestünde aus $C_1 = (-1, 1)$, $A_1 = (0, \infty)$ und $g(z) = \sqrt{1 - z^2}$. Unsere implizite Funktion sucht sich in diesem Fall für jedes $z \in (-1, 1)$ dasjenige positive $x = g(z)$ aus, für das der Punkt (z, x) auf dem Einheitskreis liegt. Die Ableitung dieser impliziten Funktion ergibt sich mit unserer Regel richtig zu

$$\frac{\partial g}{\partial z} = - \left(\frac{\partial f}{\partial x} \right)^{-1} \left(\frac{\partial f}{\partial z} \right) = - \frac{2z}{2x} = - \frac{z}{\sqrt{1 - z^2}}$$

Beispiel 7.2.12. Wir betrachten $f : \mathbb{R}^2 \times \mathbb{R} \rightarrow \mathbb{R}$, $f(z, w, x) = z^2 + w^2 + x^2$ und $(c, a) = (0, 0, 1)$. Um Indizes zu vermeiden, haben wir hier die Parameter mit (z, w) statt mit (z_1, z_2) bezeichnet. Da $\frac{\partial f}{\partial x}$ nicht verschwindet bei $(0, 0, 1)$, sind die Voraussetzungen des Satzes erfüllt. Ein mögliches Tripel bestünde aus C_1 der offenen Einheitskreisscheibe in der zw -Ebene, $A_1 = (0, \infty)$ und $g(z, w) = \sqrt{1 - z^2 - w^2}$. Anschaulich wählt unsere implizite Funktion für jedes Paar (z, w) aus der offenen Einheitskreisscheibe das positive x aus, für das der Punkt (z, w, x) auf der Einheitssphäre liegt.

Beispiel 7.2.13. Wir betrachten das Gleichungssystem

$$\begin{aligned} \zeta \xi + \zeta v - \xi u^2 &= 1 \\ \zeta u^3 + 3\zeta u - uv &= 3 \end{aligned}$$

Es ist etwa erfüllt für $(\zeta, \xi, u, v) = (1, 1, 1, 1)$. Wenn man nun ζ und ξ ein bißchen ändert, kann man dann stets u und v so anpassen, daß unser Gleichungssystem erfüllt bleibt? Der Satz über implizite Funktionen liefert genau diese Aussage, man muß dazu nur prüfen, daß die (2×2) -Matrix der partiellen Ableitungen der beiden Gleichungen unseres Systems nach u und v bei $(1, 1, 1, 1)$ invertierbar ist. Genauer erhält man hier die Matrix

$$\begin{pmatrix} -2 & 1 \\ 5 & -1 \end{pmatrix}$$

und damit implizite Funktionen $u(\zeta, \xi)$ und $v(\zeta, \xi)$, die in einer Umgebung von $(1, 1)$ definiert sind und die dort jeweils den Wert 1 annehmen. Man beachte jedoch die Verschiebung der Notation: Unser (ζ, ξ) hier entspricht im Satz über implizite Funktionen dem z und unser (u, v) hier hieß dort x . Wie finden wir nun die partiellen Ableitungen von u und v bei $(\zeta, \xi) = (1, 1)$? Sicher können wir unsere abstrakte Formel anwenden, aber dabei verheddert man sich leicht. Ich ziehe es vor, den Beweis im Spezialfall zu wiederholen und die definierenden Gleichungen partiell abzuleiten. Wir erhalten so für die partiellen Ableitungen nach ζ etwa

$$\begin{aligned} \xi + v + \zeta v_\zeta - 2\xi u u_\zeta &= 0 \\ u^3 + 3\zeta u^2 u_\zeta + 3u + 3\zeta u_\zeta - u_\zeta v - u v_\zeta &= 0 \end{aligned}$$

und nach Auswerten bei $(\zeta, \xi) = (1, 1)$ ergibt sich für die Werte unserer partiellen Ableitungen dort das lineare Gleichungssystem

$$\begin{aligned} 2 + v_\zeta - 2u_\zeta &= 0 \\ 4 + 5u_\zeta - v_\zeta &= 0 \end{aligned}$$

dessen Lösung keine Probleme mehr aufwerfen sollte.

Beweis des Satzes 7.2.4 über implizite Funktionen. Wir setzen $f(c, a) = b$ und betrachten die Abbildung

$$\begin{aligned} F : C \times A &\rightarrow Z \times Y \\ (z, x) &\mapsto (z, f(z, x)) \end{aligned}$$

Ihr Differential bei (c, a) hat im Sinne von ?? Block-Gestalt

$$\begin{pmatrix} \text{id}_{\bar{Z}} & 0 \\ d_{(c,a)}f \circ \text{in}_{\bar{Z}} & d_{(c,a)}f \circ \text{in}_{\bar{X}} \end{pmatrix}$$

und ist insbesondere invertierbar. Nach dem Umkehrsatz gibt es also offene Umgebungen $C_1 \subset C$ von c und $A_1 \subset A$ von a und $P_1 \subseteq Z \times Y$ von $(c, b) = F(c, a)$ derart, daß F einen C^1 -Diffeomorphismus

$$F : C_1 \times A_1 \xrightarrow{\sim} P_1$$

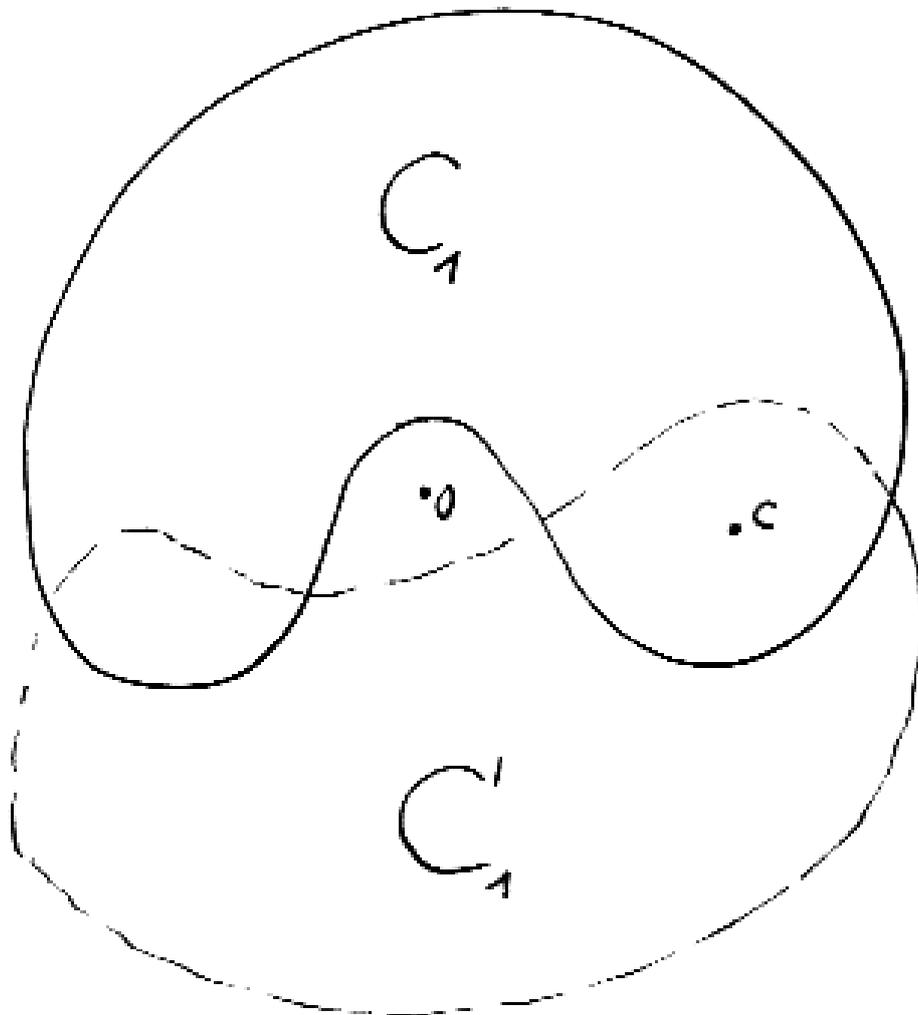
liefert. Die Umkehrabbildung $G = F^{-1} : P_1 \xrightarrow{\sim} C_1 \times A_1$ hat dann offensichtlich die Gestalt $(z, y) \mapsto (z, \tilde{g}(z, y))$ für geeignetes $\tilde{g} : P_1 \rightarrow X$. Nun ist $f(z, x) = b$ gleichbedeutend zu $F(z, x) = (z, b)$, und unter den Zusatzannahmen $(z, x) \in C_1 \times A_1$ und $(z, b) \in P_1$ ist das weiter gleichbedeutend zu $(z, x) = G(z, b)$ alias $x = \tilde{g}(z, b)$. Verkleinern wir falls nötig C_1 noch weiter derart, daß zusätzlich gilt $C_1 \times \{b\} \subset P_1$, dann gibt es zu jedem $z \in C_1$ genau ein $x = g(z) \in A_1$ mit $f(z, x) = b$, nämlich $g(z) = \tilde{g}(z, b)$. Die so definierte Funktion g ist stetig differenzierbar nach dem Umkehrsatz. Ihre Ableitung bei $z \in C_1$ ergibt sich leicht, wenn man die Definitionsgleichung $f(z, g(z)) = b$ als Abbildung $C_1 \rightarrow Y$ auffaßt und auf beiden Seiten das Differential an der Stelle z nimmt: Mit der Kettenregel folgt nämlich

$$d_{(z, g(z))}f \circ \begin{pmatrix} \text{id} \\ d_z g \end{pmatrix} = 0$$

und zerlegen wir darin $d_{(z, g(z))}f = (d_{(z, g(z))}f \circ \text{in}_{\bar{Z}}, d_{(z, g(z))}f \circ \text{in}_{\bar{X}})$ als Zeilen-Blockmatrix im Sinne von ??, so ergibt sich sofort die behauptete Formel für das Differential. \square

Ergänzung 7.2.14. Wir zeigen noch unsere Behauptung 7.2.5 und zeigen genauer, daß für C_1 wie im vorhergehenden Beweis und $U \subseteq C_1$ eine wegzusammenhängende Umgebung von c die Einschränkung der bisher betrachteten Funktion $g : C_1 \rightarrow A_1$ auch die einzige stetige Funktion $\hat{g} : U \rightarrow A$ ist mit $\hat{g}(c) = a$ und $f(z, \hat{g}(z)) = b$ für alle $z \in U$. Sei in der Tat solch ein \hat{g} gegeben. Die Menge der Punkte $z \in U$ mit $\hat{g}(z) = g(z)$ ist nicht leer, da sie c enthält. Sie ist abgeschlossen in U wegen der Stetigkeit beider Funktionen. Wenn wir auch noch ihre Offenheit zeigen, sind wir fertig mit der Beschreibung ??, ?? des Zusammenhangs. Wegen der Stetigkeit nimmt \hat{g} in einer Umgebung von c nur Werte aus A_1 an, in dieser Umgebung von c muß also \hat{g} mit g übereinstimmen. Dieselbe Argumentation greift nun aber für jeden Punkt $z \in U$ mit $\hat{g}(z) = g(z)$, denn alles bereits Gesagte gilt genauso für $(z, g(z))$ wie für (c, a) .

Übung 7.2.15. Gegeben eine einfache Nullstelle eines reellen oder komplexen Polynoms wird bei hinreichend kleinem Wackeln an den Koeffizienten des Polynoms sich auch unsere Nullstelle nur ein bißchen bewegen und differenzierbar, im komplexen Fall sogar komplex differenzierbar von besagten Koeffizienten abhängen. Man formuliere diese Aussage präzise und beweise sie.



Die Funktion $f : \mathbb{R}^3 \setminus \langle e_3 \rangle \rightarrow \mathbb{R}$ auf dem Komplement der z -Achse gegeben durch $(r \cos \vartheta, r \sin \vartheta, z) \mapsto \sin(z - \vartheta)$ für $r > 0$ und ϑ beliebig hat als Niveaumenge zum Niveau Null eine Wendeltreppe, die sich um die z -Achse $\langle e_3 \rangle$ windet. Eingezeichnet sind zusammenhängende offene Umgebungen C_1 und C'_1 eines Punkte $c \in \mathbb{R}^2$, die sich durchaus zu "erlaubten" Paaren (C_1, A_1) und (C'_1, A'_1) ergänzen lassen, bei denen die zugehörigen impliziten Funktionen jedoch auf $C_1 \cap C'_1$ nicht übereinstimmen.

7.3 Untermannigfaltigkeiten reeller Räume

7.3.1. Im weiteren Verlauf der Vorlesung werden wir Formeln herleiten für die Oberfläche einer Kugel oder eines Ellipsoids oder eines Torus alias Schwimmrings im \mathbb{R}^3 und allgemeiner Maxima und Minima von Funktionen auf derartigen Gebilden untersuchen. Später werden wir darüber hinaus Funktionen über derartige Gebilde integrieren. Um alle diese Formeln in angemessener Allgemeinheit diskutieren zu können, führen wir hier den Begriff der Untermannigfaltigkeit einer endlichdimensionalen reellen Raums ein und erklären, wie man mit diesem Begriff umgeht. Anschließend diskutieren wir dann Extremwertaufgaben.

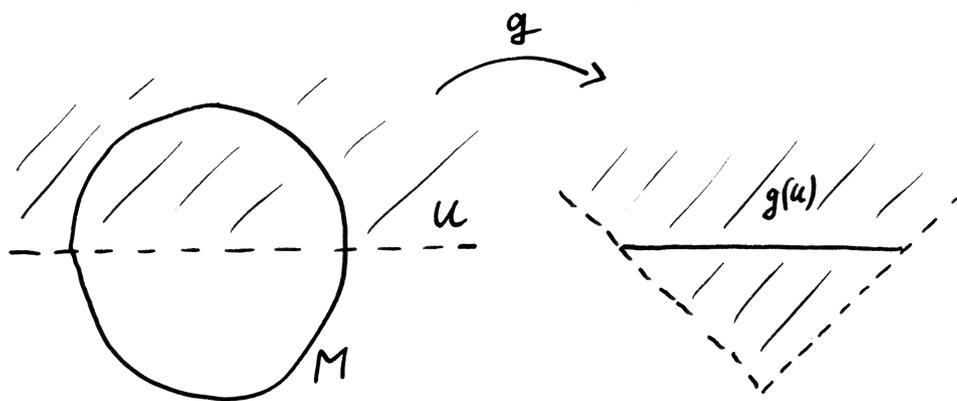
Definition 7.3.2. Sei X ein reeller Raum endlicher Dimension und $k \geq 0$ eine natürliche Zahl. Eine Teilmenge $M \subset X$ heißt eine **k -dimensionale C^1 -Untermannigfaltigkeit** oder kurz **Untermannigfaltigkeit** von X genau dann, wenn es für jeden Punkt $p \in M$ ein Paar (U, g) gibt bestehend aus einer offenen Umgebung $U \subseteq X$ von p und einem C^1 -Diffeomorphismus $g : U \xrightarrow{\sim} g(U)$ von U auf eine offene Teilmenge $g(U) \subseteq \mathbb{R}^n$ derart, daß gilt:

$$g(U \cap M) = \{z \in g(U) \mid z_{k+1} = \dots = z_n = 0\}$$

Ein derartiges Paar (U, g) heißt eine **Plättung** von M um p . Natürlich gilt hier notwendig $n = \dim_{\mathbb{R}} X$. Auf der rechten Seite unserer Gleichung hätte ich auch kürzer $g(U) \cap (\mathbb{R}^k \times 0)$ schreiben können, wobei die 0 für die einelementige Menge $\{0\} \subset \mathbb{R}^{n-k}$ steht. In vergleichbaren Situationen werde ich von nun an diese abgekürzte Darstellung bevorzugen. Im Grenzfall $k = n = \dim_{\mathbb{R}} X$ ist insbesondere eine Untermannigfaltigkeit dasselbe wie eine offene Teilmenge.

7.3.3. Manche Autoren fordern von Mannigfaltigkeiten zusätzlich, daß sie zusammenhängend sein sollen. Andere Autoren, zum Beispiel Frank Warner [?], verwenden den Begriff einer Untermannigfaltigkeit in einer anderen Bedeutung: Unsere Untermannigfaltigkeiten im Sinne der vorhergehenden Definition würden diese Autoren als "eingebettete Untermannigfaltigkeiten" bezeichnen.

7.3.4. Sei X ein reeller Raum endlicher Dimension. Einen C^1 -Diffeomorphismus g von einer offenen Umgebung U eines Punktes $p \in X$ auf eine offene Teilmenge $g(U)$ eines \mathbb{R}^n nennt man auch ein **lokales Koordinatensystem** um den Punkt p und die Komponenten g_1, \dots, g_n von g heißen die **Koordinaten** $g_i : U \rightarrow \mathbb{R}$ unseres Koordinatensystems. Ein typisches Beispiel sind etwa die Polarkoordinaten auf offenen Teilmengen von $X = \mathbb{R}^2$, bei denen man statt (g_1, g_2) meist (r, ϑ) schreibt. In dieser Terminologie kann etwa eine Untermannigfaltigkeit von X der Dimension Eins beschrieben werden als eine Teilmenge $M \subset X$ derart, daß es um jeden Punkt $p \in M$ ein lokales Koordinatensystem von X gibt, unter dem M einer Koordinatenachse entspricht.



Der Einheitskreis ist eine eindimensionale Untermannigfaltigkeit der Ebene $M \subset \mathbb{R}^2$. Eine Plättung um beliebige Punkte aus der oberen Halbebene U wird etwa gegeben durch $g(x, y) = (x, \sqrt{x^2 + y^2} - 1)$. In diesem Fall besteht $g(U)$ aus allen Punkten oberhalb des Graphen der Funktion $x \mapsto |x| - 1$.

Beispiele 7.3.5. Jeder affine Teilraum $Y \subset X$ ist eine Untermannigfaltigkeit der Dimension $\dim Y$. Der Graph jeder \mathcal{C}^1 -Funktion $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ ist eine zweidimensionale Untermannigfaltigkeit des \mathbb{R}^3 , als Plättung mag man $g : \mathbb{R}^3 \xrightarrow{\sim} \mathbb{R}^3$ mit $g(x, y, z) = (x, y, z - f(x, y))$ nehmen. Der Einheitskreis ist eine eindimensionale Untermannigfaltigkeit des \mathbb{R}^2 . Eine Untermannigfaltigkeit der Dimension Null ist dasselbe wie eine diskrete Teilmenge im Sinne von 6.4.7.

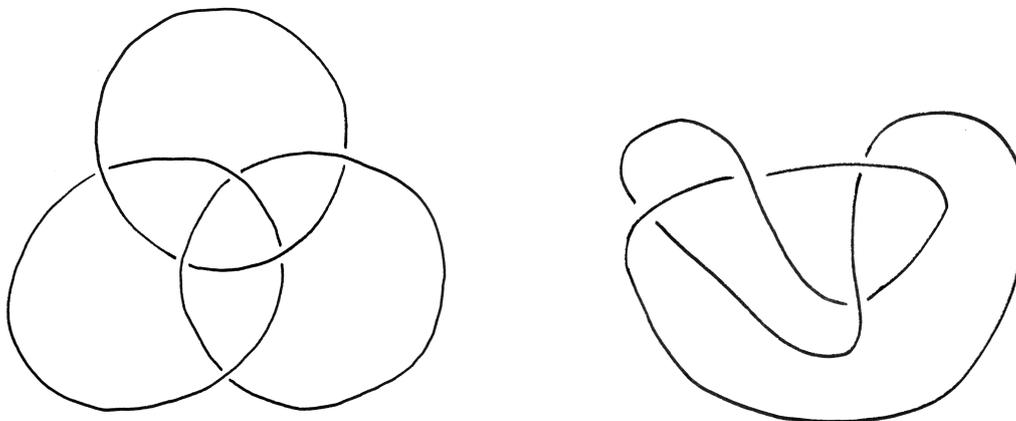
Übung 7.3.6 (Jede Untermannigfaltigkeit ist lokal ein Graph). Gegeben eine k -dimensionale Untermannigfaltigkeit $M \subset \mathbb{R}^n$ gibt es für jeden Punkt $p \in M$ eine offene Umgebung $U \subseteq \mathbb{R}^n$ und eine Permutation $\sigma \in \mathcal{S}_n$ derart, daß $M \cap U$ unter der entsprechenden Permutation der Koordinaten dem Graph einer \mathcal{C}^1 -Abbildung $V \rightarrow \mathbb{R}^{n-k}$ entspricht, die auf einer offenen Teilmenge $V \subseteq \mathbb{R}^k$ definiert ist. Zum Beispiel ist jede eindimensionale Untermannigfaltigkeit der Ebene \mathbb{R}^2 lokal entweder Graph einer reellwertigen \mathcal{C}^1 -Funktion der x -Koordinate oder der an der Hauptdiagonalen gespiegelte Graph einer reellwertigen \mathcal{C}^1 -Funktion der y -Koordinate.

Ergänzung 7.3.7. Eine kompakte Untermannigfaltigkeit der Dimension Eins in \mathbb{R}^3 heißt eine **Verschlingung** und, wenn sie wegzusammenhängend ist, ein **Knoten**. Zwei Verschlingungen heißen **isotop** genau dann, wenn sie durch einen \mathcal{C}^1 -Diffeomorphismus $\mathbb{R}^3 \xrightarrow{\sim} \mathbb{R}^3$ mit überall positiver Funktionaldeterminante ineinander überführt werden können. Die “Knotentheorie” versucht, Kriterien dafür zu entwickeln, wann zwei gegebene Verschlingungen isotop sind.

7.3.8. Sei X ein endlichdimensionaler reeller Raum. Spätestens mit 7.3.21 wird klar, daß für $k, l \in \mathbb{N}$ eine Teilmenge $M \subset X$ nur dann sowohl eine k -dimensionale als auch eine l -dimensionale Untermannigfaltigkeit sein kann, wenn entweder gilt $k = l$ oder aber $M = \emptyset$. Jede nichtleere Untermannigfaltigkeit hat also eine wohlbestimmte Dimension.

7.3.9. Sei X ein endlichdimensionaler reeller Raum. Eine Untermannigfaltigkeit der Dimension Eins von X heißt eine **Kurve** in X und eine Untermannigfaltigkeit der Dimension Zwei eine **Fläche**. Gegeben $k \leq \dim X$ versteht man unter einer Untermannigfaltigkeit der **Kodimension** k eine Untermannigfaltigkeit der Dimension $(\dim X) - k$. Eine Untermannigfaltigkeit der Kodimension Eins heißt eine **Hyperfläche**. Jede offene Teilmenge einer Untermannigfaltigkeit ist selbst eine Untermannigfaltigkeit derselben Dimension.

Proposition 7.3.10 (Untermannigfaltigkeiten als Urbilder). *Seien X und Y endlichdimensionale reelle Räume, $U \subseteq X$ eine offene Teilmenge und $f : U \rightarrow Y$ eine stetig differenzierbare Abbildung mit überall surjektivem Differential. So ist für alle $c \in Y$ das Urbild $M = f^{-1}(c)$ eine Untermannigfaltigkeit von X der Dimension $\dim X - \dim Y$.*



Die links abgebildete Verschlingung ist besonders bemerkenswert dadurch, daß je zwei der Ringe getrennt werden könnten, wenn eben der Dritte nicht wäre. Sie heißt die Verschlingung der **Borromäischen Ringe** nach einer italienischen Familie, die diese Ringe in ihrem Wappen führte. Rechts ein Beispiel für einen Knoten.

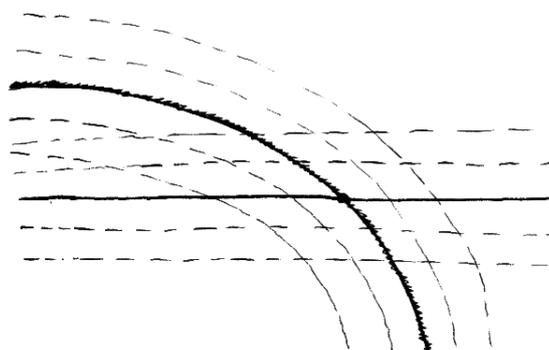


Illustration zum Beweis des Satzes [7.3.10](#) über Mannigfaltigkeiten als Urbilder. Dick gezackelt eine Niveaulinie einer Funktion $f : U \rightarrow \mathbb{R}$ mit $U \subseteq \mathbb{R}^2$. Liegt p auf dieser Niveaulinie und ist $d_p f$ surjektiv alias nicht null, so läßt sich f durch eine lineare Abbildung, im Bild die Abbildung y , zu einem lokalen Koordinatensystem ergänzen, und der Umkehrsatz [7.1.2](#) liefert dann die gesuchte lokale Plättung unserer Niveaulinie.

Beweis. Ohne Beschränkung der Allgemeinheit dürfen wir $X = \mathbb{R}^n$ und $Y = \mathbb{R}^m$ annehmen und haben dann $f = (f_1, \dots, f_m)$. Für jedes $p \in M$ finden wir sicher lineare Abbildungen $l_{m+1}, \dots, l_n : X \rightarrow \mathbb{R}$ derart, daß das Differential von $g = (f_1, \dots, f_m, l_{m+1}, \dots, l_n)$ in p bijektiv ist. Nach dem Umkehrsatz 7.1.2 gibt es dann $U' \subseteq U$ mit $p \in U'$ derart, daß g einen \mathcal{C}^1 -Diffeomorphismus $U' \xrightarrow{\sim} g(U') \subseteq \mathbb{R}^n$ induziert, und nach Konstruktion gilt $g(U' \cap M) = g(U') \cap (\{c\} \times \mathbb{R}^{n-m})$. Damit erhält man dann leicht die gesuchte Plättung von M um p . \square

Beispiel 7.3.11. Die **Sphären** $S^n = \{x \in \mathbb{R}^{n+1} \mid x_1^2 + \dots + x_{n+1}^2 = 1\}$ sind n -dimensionale Untermannigfaltigkeiten in \mathbb{R}^{n+1} , denn das Differential von $x \mapsto x_1^2 + \dots + x_{n+1}^2$ verschwindet nirgends auf der Sphäre S^n .

Übung 7.3.12. In der Situation von 7.3.10 ist allgemeiner auch für jede Untermannigfaltigkeit $C \subset Y$ ihr Urbild $M = f^{-1}(C)$ eine Untermannigfaltigkeit von X der Dimension $\dim X - \dim Y + \dim C$.

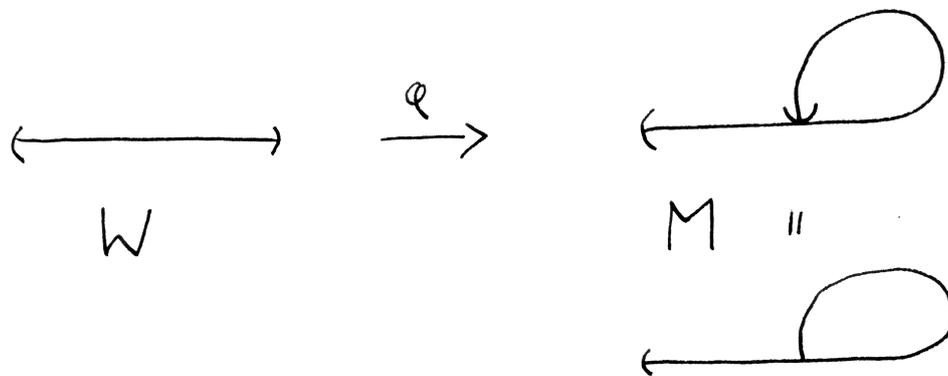
Proposition 7.3.13 (Untermannigfaltigkeiten als Bilder). *Seien ein endlichdimensionaler reeller Raum X der Dimension $\dim_{\mathbb{R}} X = n$ und $k \geq 0$ gegeben. Eine Teilmenge $M \subset X$ ist eine k -dimensionale Untermannigfaltigkeit genau dann, wenn es für jeden Punkt $p \in M$ eine stetig differenzierbare Abbildung $\varphi : W \rightarrow X$ von einer offenen Teilmenge $W \subseteq \mathbb{R}^k$ nach X gibt derart, daß gilt:*

1. $\varphi(W)$ ist enthalten in M und offen in M und enthält p ;
2. $d_x \varphi$ ist injektiv für alle $x \in W$;
3. φ ist injektiv und $\varphi^{-1} : \varphi(W) \rightarrow W$ ist stetig.

7.3.14. Ein Paar (W, φ) wie in der Proposition nenne ich eine **Karte** der Untermannigfaltigkeit M . Eine Karte der Stadt Freiburg kann als eine Variante dieses Begriffs verstanden werden, bei der W ein Rechteck aus Papier ist und das Bild einiger Punkte des Papiers unter der Abbildung φ in das wirkliche Freiburg durch bildliche Symbole angezeigt wird.

7.3.15. Im Fall $k = 0$ stellt bereits die erste Bedingung sicher, daß jeder Punkt von M offen ist in M , so daß M in der Tat eine nulldimensionale Untermannigfaltigkeit von X ist.

Beweis. Ist $M \subset X$ eine Untermannigfaltigkeit der Dimension k und (U, g) eine Plättung von M , so liefert die Einschränkung von g^{-1} auf $W = g(U) \cap (\mathbb{R}^k \times 0)$ eine Karte von M . Folglich hat eine Untermannigfaltigkeit um jeden Punkt mindestens eine Karte. Ist andererseits $\varphi : W \rightarrow M \subset X$ eine Karte von M um



Ein Beispiel einer Teilmenge M der Papierebene, die keine Untermannigfaltigkeit ist und für die die Bedingung aus 7.3.13 erfüllt wäre, wenn wir von unseren Karten nicht auch noch fordern würden, daß ihre Umkehrabbildungen stetig sind.

p mit $W \subseteq \mathbb{R}^k$, so können wir Vektoren $v_1, \dots, v_{n-k} \in \vec{X}$ finden derart, daß das Differential von

$$\begin{aligned} \tilde{\varphi}: \quad W \times \mathbb{R}^{n-k} &\rightarrow X \\ (w, t_1, \dots, t_{n-k}) &\mapsto \varphi(w) + t_1 v_1 + \dots + t_{n-k} v_{n-k} \end{aligned}$$

im Punkt $(\varphi^{-1}(p), 0)$ bijektiv ist. Nach dem Umkehrsatz 7.1.2 induziert $\tilde{\varphi}$ dann einen \mathcal{C}^1 -Diffeomorphismus von einer offenen Umgebung $G \subseteq W \times \mathbb{R}^{n-k}$ von $(\varphi^{-1}(p), 0)$ auf eine offene Umgebung $\tilde{U} \subseteq X$ von p . Kürzen wir in etwas gewagter Notation

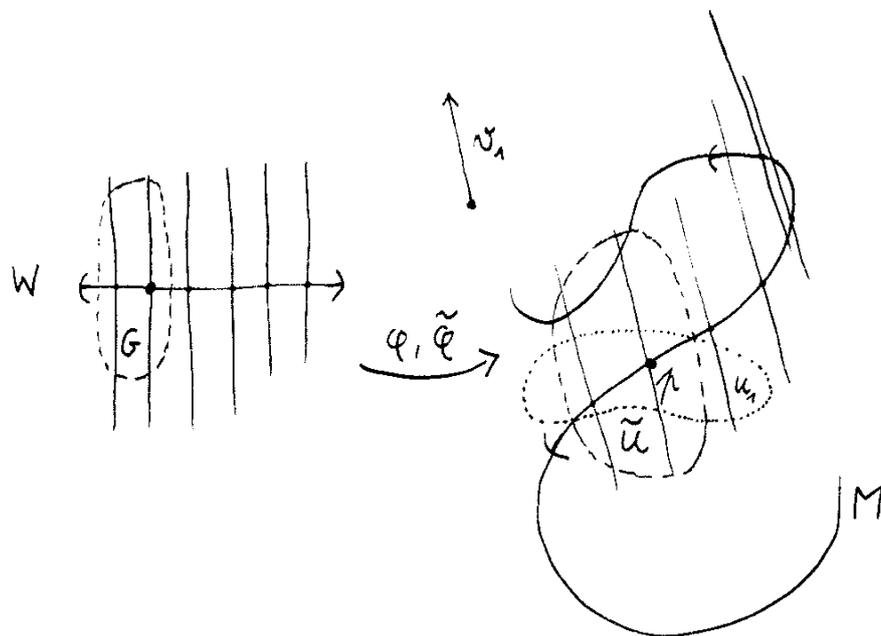
$$\{w \in W \mid (w, 0) \in G\} = W \cap G$$

ab, so ist $W \cap G$ offen in W und damit $\varphi(W \cap G)$ offen in $\varphi(W)$ und damit in M . Also gibt es $U_1 \subseteq X$ mit $\varphi(W \cap G) = M \cap U_1$. Dann setzen wir $U = \tilde{U} \cap U_1$ und $g = \tilde{\varphi}^{-1} : U \rightarrow \mathbb{R}^n$ ist die gesuchte Plättung von M um p . \square

Übung 7.3.16. Man zeige: Gegeben eine Karte (W, φ) einer Untermannigfaltigkeit $M \subset X$ und ein Punkt $p \in W$ gibt es stets ein Paar (U, g) bestehend aus einer offenen Umgebung $U \subseteq X$ von $\varphi(p)$ und einer \mathcal{C}^1 -Abbildung $g : U \rightarrow W$ derart, daß gilt $g(\varphi(y)) = y$ für alle $y \in \varphi^{-1}(U)$.

Übung 7.3.17. Der Doppelkegel $\{(x, y, z) \mid x^2 + y^2 = z^2\}$ ist keine Untermannigfaltigkeit des \mathbb{R}^3 . Auch die Teilmenge aller seiner Punkte mit nichtnegativer z -Koordinate ist keine Untermannigfaltigkeit des \mathbb{R}^3 . Entfernen wir aus dem Doppelkegel jedoch den Ursprung, so erhalten wir eine Untermannigfaltigkeit des \mathbb{R}^3 .

7.3.18. Sei X ein endlichdimensionaler reeller Raum und $M \subset X$ eine Untermannigfaltigkeit. Die Umkehrabbildung $\varphi^{-1} : \varphi(W) \rightarrow \mathbb{R}^k$ zu einer Karte $\varphi : W \rightarrow M$ von M nennen wir in Verallgemeinerung von 7.3.4 ein **lokales Koordinatensystem von M** und seine Komponenten $\text{pr}_i \circ \varphi^{-1} : \varphi(W) \rightarrow \mathbb{R}$ für $1 \leq i \leq k$ nennen wir **lokale Koordinaten auf M** . Viele Autoren verwenden allerdings auch eine andere Terminologie und verstehen unter einer Karte das, was wir ein Koordinatensystem genannt haben. Lokale Koordinaten um einen Punkt der Erdoberfläche, der nicht gerade auf dem sogenannten ‘Nullmeridian’ liegt, sind etwa die Längen- und Breitengrade. Bilden Funktionen $x_1, \dots, x_k : U \rightarrow \mathbb{R}$ ein System von lokalen Koordinaten auf einer offenen Teilmenge $U \subseteq M$ einer Mannigfaltigkeit, und ist $f : U \rightarrow \mathbb{R}$ eine Funktion, so bezeichnen wir mit $\frac{\partial f}{\partial x_i}$ auch die Funktion $U \rightarrow \mathbb{R}$, die unter der zugehörigen Karte $\varphi : W \xrightarrow{\sim} U$ verwandt ist zu $\frac{\partial(f \circ \varphi)}{\partial x_i}$, wenn denn $f \circ \varphi$ partiell differenzierbar ist nach der i -ten Variablen. Auch hier gilt es zu beachten, daß $\frac{\partial f}{\partial x_i}$ von der Wahl aller Koordinaten abhängt, und keineswegs nur von der i -ten Koordinate.



Ein Bild zum Beweis von [7.3.13](#) im Fall $n = k = 1$

Übung 7.3.19. Seien $X \subset Y$ ein endlichdimensionaler reeller Raum mit einem affinen Teilraum. So ist eine Teilmenge $M \subset X$ als Teilmenge von Y eine Untermannigfaltigkeit genau dann, wenn M als Teilmenge von X eine Untermannigfaltigkeit ist.

Definition 7.3.20. Sind $(W_\alpha, \varphi_\alpha)$ und (W_β, φ_β) zwei Karten einer Untermannigfaltigkeit M , so setzen wir $W_{\alpha\beta} = \varphi_\alpha^{-1}(\varphi_\beta(W_\beta))$ und nennen die Abbildung

$$\varphi_{\beta\alpha} := \varphi_\beta^{-1} \circ \varphi_\alpha : W_{\alpha\beta} \rightarrow W_{\beta\alpha}$$

den **Kartenwechsel** zwischen unseren beiden Karten.

Proposition 7.3.21. *Kartenwechsel sind stets C^1 -Diffeomorphismen.*

Beweis. Nach dem Beweis von 7.3.13 kann man für (W, φ) eine Karte und $p \in \varphi(W)$ einen Punkt aus ihrem Bild stets eine offene Umgebung U von p in X finden derart, daß $\varphi^{-1} : U \cap \varphi(W) \rightarrow W$ die Restriktion einer Plättung $g : U \rightarrow \mathbb{R}^n$ unserer Mannigfaltigkeit ist. \square

Satz 7.3.22 (Extrema unter Nebenbedingungen). *Sei X ein endlichdimensionaler reeller Raum, $U \subseteq X$ offen, $f : U \rightarrow \mathbb{R}^m$ stetig differenzierbar, und $p \in U$ ein Punkt mit $d_p f$ surjektiv, so daß der Schnitt der Menge*

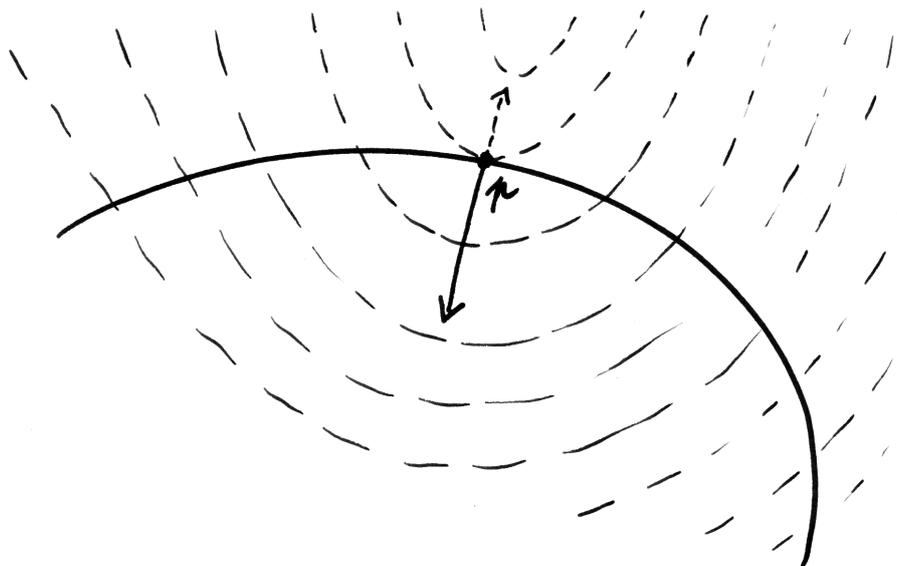
$$M = \{q \in U \mid f(q) = f(p)\}$$

mit einer hinreichend kleinen offenen Umgebung von p eine Mannigfaltigkeit ist. Besitzt dann für eine differenzierbare Funktion $h : U \rightarrow \mathbb{R}$ ihre Einschränkung $h|_M$ ein lokales Extremum bei p , so gibt es $\lambda_1, \dots, \lambda_m \in \mathbb{R}$ mit

$$d_p h = \lambda_1 d_p f_1 + \dots + \lambda_m d_p f_m$$

7.3.23. Mit der Notation $f(p) = c$ mag man in dieser Situation ein lokales Extremum der Restriktion $h|_M$ auch ein **lokales Extremum von h unter den Nebenbedingungen** $f_1(q) = c_1, \dots, f_m(q) = c_m$. Die λ_i heißen die **Lagrange'schen Multiplikatoren**. Im Fall $X = \mathbb{R}^n$ kann man unsere Bedingung dahingehend interpretieren, daß "der Gradient der Funktion h in p auf M senkrecht stehen muß" oder auch, daß "der Gradient der Funktion h in p eine Linearkombination der Gradienten der Nebenbedingungen sein muß". Die Bedingung " $d_p f$ surjektiv" hinwiederum kann man dahingehend interpretieren, daß die Gradienten der f_i bei p linear unabhängig sein sollen.

Beweis. Indem wir U verkleinern, dürfen wir ohne Beschränkung der Allgemeinheit annehmen, daß $d_q f$ surjektiv ist für alle $q \in U$. Nach 7.3.10 ist dann $M \subset X$ eine Untermannigfaltigkeit. Ist $\varphi : W \rightarrow M$ eine Karte von M um p mit



Dies Bild soll den Satz über Extrema mit Nebenbedingungen veranschaulichen für den Fall einer Funktion h auf der Papierebene, hier angedeutet durch einige gestrichelt eingezeichnete Niveaulinien, die maximiert werden soll unter einer Nebenbedingung f , hier angedeutet durch die fett eingezeichnete Kurve M der Punkte, bei denen sie erfüllt ist. Es scheint mir anschaulich recht offensichtlich, daß Extrema von h auf M nur an Stellen $p \in M$ zu erwarten sind, an denen der Gradient von h senkrecht steht auf M , also ein Vielfaches des Gradienten von f ist. Im Bild hätten wir etwa grob geschätzt $(\text{grad } h)(p) = -\frac{1}{2}(\text{grad } f)(p)$.

Allerdings ist es für reellwertige Funktionen auf der Papierebene streng genommen erst nach der Wahl eines Maßstabs sinnvoll, von Gradienten zu reden, und in allgemeineren Fällen erst nach Wahl eines Skalarprodukts auf dem Richtungsraum, weshalb ich im Satz die koordinatenfreie Formulierung mit Differentialen vorgezogen habe.

$p = \varphi(w)$, so hat $h|_{M \cap U}$ ein lokales Extremum bei p genau dann, wenn $h \circ \varphi$ ein lokales Extremum bei w hat. Nach 5.4.5 ist eine notwendige Bedingung dafür $d_w(h \circ \varphi) = 0$, als da heißt $d_p h \circ d_w \varphi = 0$. Wir haben aber eh $f \circ \varphi = 0$, also $d_p f \circ d_w \varphi = 0$, und aus Dimensionsgründen bilden die $d_p f_i$ für $1 \leq i \leq m$ sogar eine Basis für den Untervektorraum von \vec{X}^* aller Linearformen, die auf dem Bild von $d_w \varphi$ verschwinden. Verschwindet auch $d_p h$ auf diesem Teilraum, so muß es folglich als Linearkombination der $d_p f_i$ geschrieben werden können. \square

Beispiel 7.3.24. Wir suchen lokale Extrema der Funktion $h : (x, y) \mapsto x + y$ auf dem Einheitskreis $M = S^1$, d.h. unter der Nebenbedingung $x^2 + y^2 = 1$. Diese Nebenbedingung bedeutet, daß die Funktion $f : (x, y) \mapsto x^2 + y^2$ den Wert 1 annehmen muß. Lokale Extrema können nach unserem Satz nur an Stellen $p \in M$ mit $d_p h = \lambda d_p f$ angenommen werden, also an Stellen $p = (x, y) \in S^1$ mit $(1, 1) = \lambda(2x, 2y)$. Damit kommen nur die beiden Stellen $(-1/\sqrt{2}, -1/\sqrt{2})$ und $(1/\sqrt{2}, 1/\sqrt{2})$ in Frage. Hier ist offensichtlich die erste ein lokales Minimum und die zweite ein lokales Maximum.

Beispiel 7.3.25. Wir suchen lokale Extrema der Funktion $h : (x, y, z) \mapsto ax + by + cz$ auf dem Einheitskreis $M = S^1 \times 0$ in der xy -Ebene, d.h. unter den beiden Nebenbedingungen $f_1(x, y, z) = x^2 + y^2 = 1$ und $f_2(x, y, z) = z = 0$. Lokale Extrema können nach unserem Satz nur an Stellen $p \in M$ angenommen werden, für die es $\lambda_1, \lambda_2 \in \mathbb{R}$ gibt mit $d_p h = \lambda_1 d_p f_1 + \lambda_2 d_p f_2$, also an Stellen $(x, y, z) \in M$ mit

$$(a, b, c) = \lambda_1(2x, 2y, 0) + \lambda_2(0, 0, 1)$$

Daraus folgt jedoch schnell $\lambda_2 = c$, und unter der zusätzlichen Voraussetzung $(a, b) \neq (0, 0)$ es kommen nur die beiden Stellen $\pm(a^2 + b^2)^{-1}(a, b, 0)$ in Frage. Wieder ist hier offensichtlich die eine ein lokales Minimum und die andere ein lokales Maximum.

7.4 Die Transformationsformel

Definition 7.4.1. Sei X ein topologischer Raum und $f : X \rightarrow V$ eine Abbildung in einen Vektorraum oder allgemeiner eine abelsche Gruppe. Der **Träger** $\text{supp } f$ von f (für englisch und französisch **support**) ist die kleinste abgeschlossene Teilmenge von X , außerhalb derer die Funktion verschwindet, in Formeln

$$\text{supp } f = \overline{f^{-1}(V \setminus \{0\})}$$

Den Vektorraum aller stetigen Funktionen $f : X \rightarrow \mathbb{R}$ mit kompaktem Träger bezeichnen wir mit $\mathcal{C}_c(X, \mathbb{R})$.

7.4.2. Diese Definition muß kleinlich genau genommen werden. Ist zum Beispiel $X = (0, 1)$ ein offenes Intervall und f die konstante Funktion 1, so hat f *nicht* kompakten Träger, denn $\text{supp } f = X$ ist nicht kompakt.

Definition 7.4.3. Ist $U \subseteq \mathbb{R}^n$ eine offene Teilmenge und $f \in C_c(U, \mathbb{R})$ eine stetige Funktion auf U mit kompaktem Träger, so definieren wir eine reelle Zahl

$$\int_U f = \int_U f(x_1, \dots, x_n) dx_1 \dots dx_n = \int_U f(x) d^n x,$$

das **Integral von f über U** , indem wir f durch Null zu einer stetigen Funktion mit kompaktem Träger auf \mathbb{R}^n ausdehnen und diese Ausdehnung im Sinne von 5.1.1 integrieren über irgendeinen kompakten Quader, der den Träger umfaßt. Im Spezialfall $n = 0$ vereinbaren wir, daß das Integral einfach der einzige Funktionswert der zu integrierenden Funktion sein möge.

7.4.4. Man überlege sich als Übung, daß es hierbei nicht auf die Wahl des kompakten Quaders ankommt. Das so definierte Integral ist offensichtlich linear, $\int f + g = \int f + \int g$ und $\int \lambda f = \lambda \int f$ für $\lambda \in \mathbb{R}$, und monoton, als da heißt $f \leq g \Rightarrow \int f \leq \int g$. Insbesondere folgt wie im Fall einer Veränderlichen $|\int f| \leq \int |f|$.

Ergänzung 7.4.5. In derselben Weise erklärt man auch den Träger für vektorwertige Funktionen oder, noch allgemeiner, für Funktionen mit Werten in einer beliebigen Gruppe, und in Fortführung von 5.1.4 erklärt man für jede stetige Abbildung mit kompaktem Träger von \mathbb{R}^n in einen endlichdimensionalen reellen Vektorraum ihr Integral, einen Vektor aus besagtem Vektorraum. Auch die folgende Transformationsformel überträgt sich unmittelbar auf den Fall vektorwertiger Funktionen.

Satz 7.4.6 (Transformationsformel). Seien $U, V \subseteq \mathbb{R}^n$ offene Teilmengen und $\phi : U \xrightarrow{\sim} V$ ein C^1 -Diffeomorphismus. Bezeichne $|\det d\phi|$ die Abbildung $U \rightarrow \mathbb{R}$, $p \mapsto |\det d_p \phi|$. So gilt für jede stetige Funktion $f : V \rightarrow \mathbb{R}$ mit kompaktem Träger

$$\int_V f = \int_U (f \circ \phi) |\det d\phi|$$

Beispiele 7.4.7. Ist $\phi : \mathbb{R}^n \rightarrow \mathbb{R}^n$ eine abstandserhaltende Abbildung, also nach ?? die Verknüpfung einer orthogonalen Abbildung mit einer Translation, so liefert die Transformationsformel für jede stetige Funktion mit kompaktem Träger $f : \mathbb{R}^n \rightarrow \mathbb{R}$ die Identität $\int f = \int f \circ \phi$. Ist $\phi : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ die Streckung um den Faktor 2, so liefert die Transformationsformel die Identität $\int f = 4 \int f \circ \phi$. Beide Aussagen sollten auch anschaulich unmittelbar einleuchten.

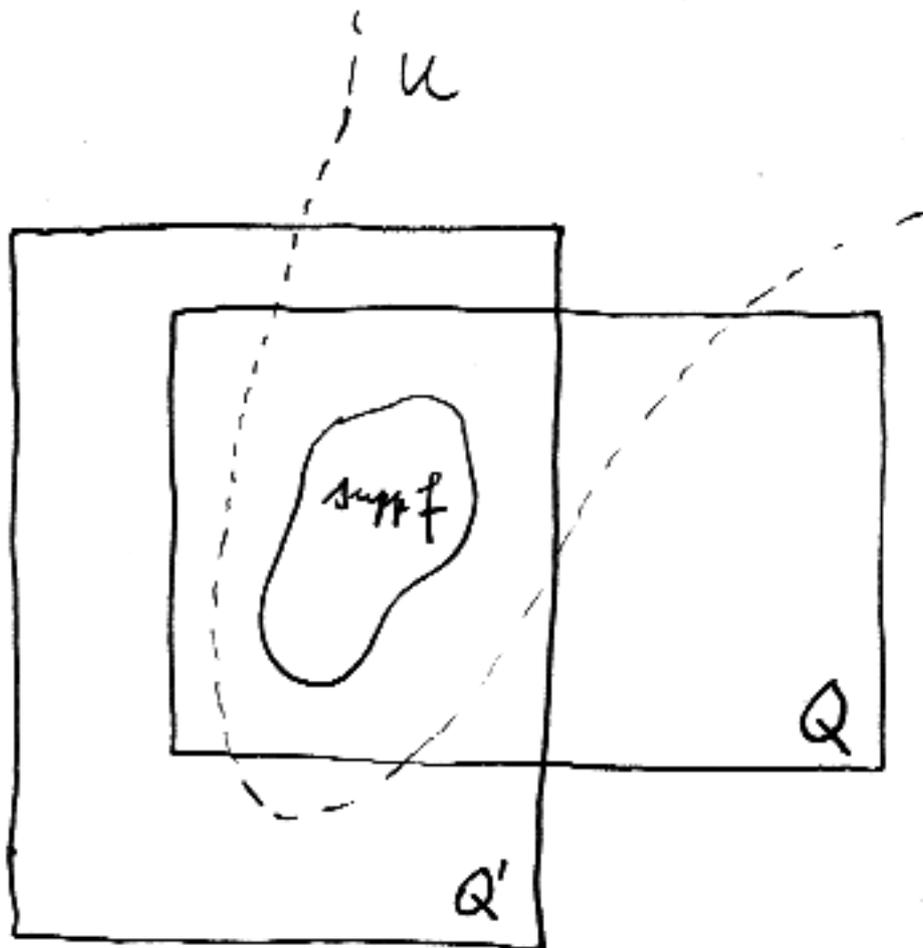


Illustration zur Definition des Integrals stetiger Funktionen mit kompaktem Träger 7.4.3. Man prüft ohne Schwierigkeiten, daß die Wahl des kompakten Quaders hier keine Rolle spielt, solange er nur den Träger unserer Funktion umfaßt. Im Bild kommen unter vielen anderen etwa die beiden Quader Q und Q' in Frage.

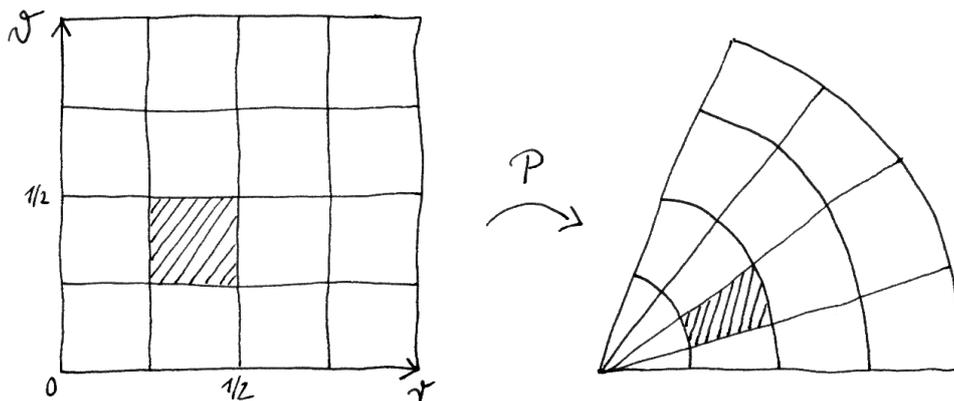


Illustration zur Transformationsformel, insbesondere zu Beispiel 7.4.8. Das Integral einer Funktion f über das linke Kuchenstück kann angenähert werden, indem wir die angedeutete Unterteilung des Integrationsbereichs betrachten, in jedem der unterteilenden Stücke den Funktionswert an einer Stelle mit der Fläche des entsprechenden Stücks multiplizieren, und diese Produkte aufsummieren. Unter der Polarkoordinatenabbildung P entspricht nun die Unterteilung unseres Kuchenstücks einer Unterteilung unseres Quadrats, und die Fläche des Bildes eines Unterquadrats ist in etwa der Betrag der Funktionaldeterminante $|\det P|$ an einer Stelle unseres Unterquadrats multipliziert mit der Fläche besagten Unterquadrats. So wäre etwa die Fläche des schraffierten Teils im Kuchenstück rechts etwas weniger als halb so groß wie die Fläche des schraffierten Unterquadrats links, und $|\det P| = r$ nimmt auf unserem Unterquadrat Werte zwischen $1/4$ und $1/2$ an. Es wird also in etwa dasselbe herauskommen, wenn wir von der Funktion $(f \circ P)|\det P|$ auf unserem Quadrat in jedem der Unterquadrate den Funktionswert an einer Stelle mit der Fläche des entsprechenden Unterquadrats multiplizieren, und diese Produkte aufsummieren. Im Grenzübergang für immer feinere Unterteilungen kommt dann sogar nicht nur in etwa, sondern ganz genau dasselbe heraus, und das ist die anschauliche Bedeutung der Transformationsformel.

Beispiel 7.4.8. Wir betrachten die Polarkoordinatenabbildung

$$P : \mathbb{R}_{>0} \times (-\pi, \pi) \xrightarrow{\sim} \mathbb{R}^2 \setminus (\text{abgeschlossene negative } x\text{-Achse})$$

$$(r, \vartheta)^\top \mapsto (r \cos \vartheta, r \sin \vartheta)^\top$$

wobei verwirrender Weise die Klammern $(,)$ einmal ein offenes Intervall und dann wieder Elemente kartesischer Produkte andeuten, die wir anschließend noch zu Spaltenvektoren transponieren. Das Differential der Polarkoordinatenabbildung wird gegeben durch die Jacobi-Matrix

$$dP = \begin{pmatrix} \cos \vartheta & -r \sin \vartheta \\ \sin \vartheta & r \cos \vartheta \end{pmatrix}$$

mit der Determinante $\det dP = r$. Beim Bilden des Betrags ändert sich Nichts, und wir erhalten

$$\int_V f(x, y) \, dx \, dy = \int_{P^{-1}(V)} f(P(r, \vartheta)) \, r \, dr \, d\vartheta$$

für stetige Funktionen auf \mathbb{R}^2 mit kompaktem Träger, der darüber hinaus nicht die abgeschlossene negative x -Achse treffen darf. Oft schreibt man kurz $f(r, \vartheta)$ statt $f(P(r, \vartheta))$ in der Erwartung, daß schon aus der bloßen Bezeichnung der Variablen klar wird, was gemeint ist. So ergibt sich dann eine Formel für die Transformation eines Integrals auf Polarkoordinaten, die man als

$$dx \, dy = r \, dr \, d\vartheta$$

abkürzen mag. Leider erhalten wir besagte Formel vorerst nur für sehr spezielle Funktionen. In der Praxis ist deshalb unser Satz kaum anwendbar und wir bauen im nächsten Kapitel erst einmal die Integrationstheorie aus, um die auch in der Praxis anwendbare Verallgemeinerung [9.8.1](#) formulieren und beweisen zu können.

7.4.9. Man nennt $\det d\phi$ die **Funktionaldeterminante von ϕ** . Wir verallgemeinern die Transformationsformel in [9.8.1](#) auf beliebige “integrierbare” Funktionen f , sie läßt sich auch erst in dieser Allgemeinheit gut anwenden. Bevor wir die hier gegebene Version beweisen, wollen wir versuchen, sie mit Anschauung zu füllen. Wir beschränken uns dazu auf den Fall $n = 2$. Zunächst ist hoffentlich anschaulich klar, daß es für jede lineare Abbildung $L : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ eine reelle Konstante $c(L) \geq 0$ gibt derart, daß “das Bild unter L eines Flächenstücks U der Fläche $\text{vol}(U)$ die Fläche $\text{vol}(LU) = c(L) \text{vol}(U)$ hat”. Unsere Transformationsformel enthält nun, wenn man sie ohne Rücksicht auf die Bedingungen des Satzes mutig auf die konstante Funktion $f = 1$ auf U anwendet und $\phi = L$ linear annimmt, die Erkenntnis

$$c(L) = |\det L|$$

Das sieht man auch anschaulich leicht ein: Zunächst sollte anschaulich klar sein, daß “eine Scherung die Fläche nicht ändert” und “die Streckung einer Achse die Fläche genau durch Multiplikation mit dem Betrag des Streckfaktors ändert”, so daß also unsere Erkenntnis anschaulich klar ist für lineare Abbildungen L mit Matrizen der Gestalt

$$\begin{pmatrix} 1 & a \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} a & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & a \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ a & 1 \end{pmatrix}$$

Anschaulich klar ist weiter $c(L \circ M) = c(L)c(M)$ und nach der Multiplikationsformel für Determinanten haben wir auch $|\det L \circ M| = |\det L| |\det M|$. So rechtfertigen wir dann unsere Erkenntnis $c(L) = |\det L|$ im allgemeinen. Mehr dazu mag man in ?? nachlesen. Das Integral von f erhalten wir nun im Grenzwert, wenn wir V in lauter kleine Flächenstücke V_i zerlegen und die Produkte der Flächen dieser Flächenstücke mit einem Funktionswert an einem Punkt $y_i \in V_i$ des jeweiligen Flächenstück aufsummieren, in Formeln

$$\int_V f \simeq \sum f(y_i) \text{vol}(V_i)$$

Wir betrachten nun die Urbilder $x_i = \phi^{-1}(y_i)$ unserer Punkte y_i und die Zerlegung von U durch die Urbilder $U_i = \phi^{-1}(V_i)$ unserer kleinen Flächenstücke V_i . Bei x_i wird ϕ bis auf Verschiebung gut approximiert durch $d_{x_i}\phi$, deshalb haben die Bilder $\phi(U_i) = V_i$ dieser Flächenstücke U_i in etwa die Fläche $\text{vol}(V_i) \simeq |\det d_{x_i}\phi| \text{vol}(U_i)$ und wir folgern

$$\int_V f \simeq \sum f(y_i) \text{vol}(V_i) \simeq \sum (f \circ \phi)(x_i) |\det d_{x_i}\phi| \text{vol}(U_i) \simeq \int_U (f \circ \phi) |\det d\phi|$$

Das beendet unsere anschauliche aber doch recht vage Argumentation und wir kommen nach einem Beispiel zum eigentlichen Beweis.

Beweis. Wir zeigen den Satz durch vollständige Induktion über n . Der Fall $n = 0$ ist unproblematisch und wir behandeln gleich den Fall $n = 1$. In diesem Fall kann U nach 5.1.3 als disjunkte Vereinigung von offenen Intervallen U_i geschrieben werden, und deren Bilder in V sind wieder Intervalle nach dem Zwischenwertsatz und offen etwa nach 7.1.12. Unsere Funktion f verschwindet außerhalb von endlich vielen der $\phi(U_i)$ nach 9.10.3 und wir können folglich ohne Beschränkung der Allgemeinheit annehmen, daß U und V bereits offene Intervalle sind. Wir finden dann sicher ein echtes kompaktes Intervall $[c, d] \subset V$, das den Träger von f umfaßt. Die Substitutionsregel 7.6.1 liefert nun

$$\int_c^d f(y) dy = \int_a^b f(\phi(x))\phi'(x) dx$$

für die $a, b \in U$ mit $\phi(a) = c$ bzw. $\phi(b) = d$. Da ϕ ein \mathcal{C}^1 -Diffeomorphismus sein soll, ist ϕ' stetig mit $\phi' \neq 0$ auf U und wir folgern, daß auf U entweder gilt $\phi' > 0$ oder aber $\phi' < 0$. Im ersten Fall haben wir $a < b$ und unsere Transformationsformel steht bereits da. Im zweiten Fall haben wir $a > b$ und $|\phi'| = -\phi'$ und

$$\int_U (f \circ \phi) |\det d\phi| = \int_b^a f(\phi(x)) |\phi'(x)| dx = \int_a^b f(\phi(x)) \phi'(x) dx$$

und sind wieder fertig. Damit ist der Fall $n = 1$ erledigt. Nehmen wir nun also an, wir hätten $n \geq 2$ und der Satz sei für Integration im \mathbb{R}^{n-1} schon bewiesen. Wir gehen dann in mehreren Schritten vor.

1. Läßt ϕ die erste Koordinate unverändert, in Formeln $\phi_1(x_1, \dots, x_n) = x_1$, so folgt unsere Transformationsformel aus der Induktionsvoraussetzung. In der Tat, betrachten wir die Schnitte $U_c = U \cap \{x_1 = c\}$ und $V_c = V \cap \{x_1 = c\}$ unserer offenen Mengen mit der Hyperebene $x_1 = c$, so induziert unsere Abbildung ϕ für jedes feste c Diffeomorphismen $\phi_c : U_c \xrightarrow{\sim} V_c$ und es gilt

$$|\det d_{(x_2, \dots, x_n)} \phi_c| = |\det d_{(c, x_2, \dots, x_n)} \phi|$$

Für $f_c : V_c \rightarrow \mathbb{R}$ gegeben durch $f_c(x_2, \dots, x_n) = f(c, x_2, \dots, x_n)$ erhalten wir also nach der Induktionsvoraussetzung

$$\begin{aligned} \int f_c &= \int (f_c \circ \phi_c) |\det d\phi_c| \\ &= \int (f \circ \phi)(c, x_2, \dots, x_n) |\det d_{(c, x_2, \dots, x_n)} \phi| \end{aligned}$$

Integrieren wir diese Gleichung über alle c , so ergibt sich die Transformationsformel für die Koordinatentransformation ϕ .

2. Sind $W \xrightarrow{\psi} U \xrightarrow{\phi} V$ zwei \mathcal{C}^1 -Diffeomorphismen zwischen offenen Teilmengen des \mathbb{R}^n , und gilt unsere Transformationsformel für ϕ und ψ , so gilt sie auch für $\phi \circ \psi$. In der Tat erhalten wir

$$\begin{aligned} \int f &= \int (f \circ \phi) |\det d\phi| \\ &= \int (f \circ \phi \circ \psi) (|\det d\phi| \circ \psi) |\det d\psi| \\ &= \int (f \circ \phi \circ \psi) |\det d(\phi \circ \psi)| \end{aligned}$$

Hier gilt die erste Zeile nach der Transformationsformel für ϕ angewandt auf die Funktion f , die zweite nach der Transformationsformel für ψ angewandt auf die Funktion $(f \circ \phi) |\det d\phi|$, und die dritte nach der Kettenregel

$$d_p(\phi \circ \psi) = d_{\psi(p)} \phi \circ d_p \psi$$

für $p \in W$ und der Multiplikationsformel für Determinanten $\det(AB) = (\det A)(\det B)$.

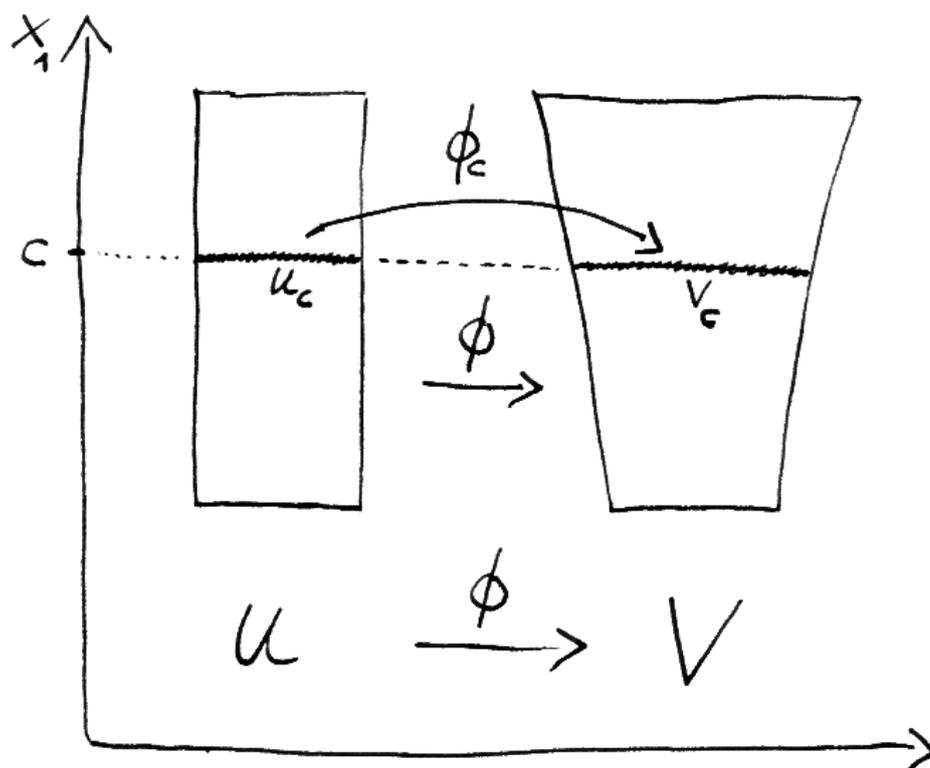


Illustration zum Beweis der Transformationsformel.

3. Für ϕ eine Vertauschung der Koordinaten gilt unsere Formel. In der Tat ist so ein ϕ ja linear mit $|\det d\phi| = 1$, und wir wissen schon nach 5.1.7, daß es bei Mehrfachintegralen nicht auf die Reihenfolge ankommt.

4. Ist eine Komponente von ϕ eine Koordinate auf U , haben wir also in Formeln $\phi_i(x_1, \dots, x_n) = x_j$ für geeignete i und j , so gilt unsere Formel. In der Tat finden wir eine Darstellung $\phi = \psi \circ \tilde{\phi} \circ \tilde{\psi}$ derart, daß $\tilde{\phi}$ die erste Koordinate unverändert läßt und $\psi, \tilde{\psi}$ Koordinatenvertauschungen sind. Für $\tilde{\phi}$ gilt dann unser Satz nach Schritt 1, für ψ und $\tilde{\psi}$ nach Schritt 3, und dann gilt er auch für ϕ nach Schritt 2.

5. Jeder Punkt $p \in U$ besitzt eine offene Umgebung U_p derart, daß unsere Transformationsformel gilt für die Restriktion von ϕ auf U_p . In der Tat finden wir zunächst ein i derart, daß gilt $\frac{\partial \phi_i}{\partial x_1}(p) \neq 0$, und dann gibt es nach dem Umkehrsatz eine offene Umgebung U_p von p derart, daß die Abbildung

$$\begin{aligned} \psi : \quad U &\rightarrow \mathbb{R}^n \\ (x_1, \dots, x_n) &\mapsto (\phi_i(x_1, \dots, x_n), x_2, \dots, x_n) \end{aligned}$$

einen \mathcal{C}^1 -Diffeomorphismus von U_p auf eine offene Teilmenge $\psi(U_p) = W_p \subset \mathbb{R}^n$ induziert. Wir bezeichnen das Bild von U_p unter ϕ mit $\phi(U_p) = V_p$ und erhalten ein kommutatives Diagramm von \mathcal{C}^1 -Diffeomorphismen

$$\begin{array}{ccc} U_p & \xrightarrow{\psi} & W_p \\ & \searrow \phi & \swarrow \phi\psi^{-1} \\ & & V_p \end{array}$$

wobei die i -te Komponente der Abbildung $\phi\psi^{-1}$ gerade die erste Koordinate ist, in Formeln

$$(\phi\psi^{-1})_i(y_1, \dots, y_n) = y_1$$

Für beide Abbildungen ψ und $(\phi\psi^{-1})$ gilt also nach Schritt 4 unsere Transformationsformel, mithin gilt sie nach Schritt 2 auch für ihre Verknüpfung, als da heißt für die Restriktion $\phi : U_p \xrightarrow{\sim} V_p$ von ϕ auf U_p . Hier ist im übrigen die Stelle im Beweis, die uns daran hindert, unsere Induktion mit dem Trivialfall $n = 0$ zu starten: Im Fall $n = 1$ können wir nämlich Schritt 4 auf ψ nicht anwenden, da in diesem Fall keine Komponente von ψ eine Koordinate wäre.

6. Wir behandeln nun den allgemeinen Fall. Sei $f : V \rightarrow \mathbb{R}$ eine stetige Funktion mit kompaktem Träger $\text{supp } f \subset V$. Für $p \in U$ wählen U_p wie in Schritt 5 und setzen wieder $V_p = \phi(U_p)$. Da $\text{supp } f$ kompakt ist, finden wir nach 9.10.3 eine endliche Teilmenge $E \subset U$ mit $\text{supp } f \subset \bigcup_{p \in E} V_p$. Jetzt benutzen das im Anschluß formulierte und bewiesene technische Lemma 7.4.10, wählen wir für

unsere endliche Überdeckung von $\text{supp } f = K$ durch die V_p mit $p \in E$ eine angepaßte Teilung der Eins α_p und schreiben

$$f = \sum_{p \in E} \alpha_p f$$

Die Summanden $\alpha_p f$ sind dann stetig mit kompaktem in V_p enthaltenem Träger. Nach der Wahl der V_p haben wir nun $\int \alpha_p f = \int ((\alpha_p f) \circ \phi) |\det d\phi|$ für alle $p \in E$, und addieren wir diese Gleichungen, so ergibt sich wie gewünscht

$$\int f = \int (f \circ \phi) |\det d\phi| \quad \square$$

Lemma 7.4.10 (Teilung der Eins). *Ist $K \subset \mathbb{R}^n$ kompakt und sind $V_1, \dots, V_r \subseteq \mathbb{R}^n$ offen mit $K \subset \bigcup V_i$, so gibt es stetige Funktionen $\alpha_i : \mathbb{R}^n \rightarrow [0, 1]$ mit kompaktem, jeweils in V_i enthaltenen Träger derart, daß gilt*

$$\sum_{i=1}^r \alpha_i(x) = 1 \quad \forall x \in K$$

7.4.11. Eine derartige Familie von Funktionen α_i heißt eine an die gegebene Überdeckung von K angepaßte **Teilung der Eins**.

Beweis. Wir wählen für jedes $x \in K$ ein $j = j(x)$ mit $x \in V_j$ und eine stetige Funktion $\varphi_x : \mathbb{R}^n \rightarrow [0, \infty)$ mit kompaktem, in V_j enthaltenem Träger, die bei x nicht verschwindet, in Formeln $\varphi_x(x) > 0$. Die $N_x := \varphi_x^{-1}(\mathbb{R}_{>0})$ sind natürlich offen in \mathbb{R}^n und überdecken K und wir haben $\bar{N}_x \subset V_{j(x)}$. Da K kompakt ist, finden wir $E \subset K$ endlich mit $K \subset \bigcup_{x \in E} N_x$. Dann bilden wir

$$\psi = \sum_{x \in E} \varphi_x$$

Diese Funktion ist stetig auf ganz \mathbb{R}^n , nimmt auf $N = \bigcup_{x \in E} N_x$ positive Werte an, und verschwindet außerhalb von N . Nun betrachten wir für jedes $x \in E$ auf der offenen Menge N die stetige Funktion $\psi_x = \varphi_x / \psi$. Natürlich gilt $\sum_{x \in E} \psi_x(z) = 1$ nicht nur für alle $z \in K$, sondern sogar für alle $z \in N$, und ψ_x verschwindet außerhalb von N_x . Als nächstes konstruieren wir eine stetige Funktion $\beta : \mathbb{R}^n \rightarrow [0, 1]$, die auf K konstant Eins ist und deren Träger in N enthalten ist. Ist zum Beispiel m das Minimum von ψ auf K , so erhalten wir ein mögliches β , indem wir setzen $\beta = h \circ \psi$ für $h : \mathbb{R} \rightarrow \mathbb{R}$ eine stetige Funktion mit $h|_{[m, \infty)} = 1$ und $h|_{(-\infty, m/2]} = 0$. Dann bilden wir schließlich

$$\alpha_i = \sum_{j(x)=i} \beta \psi_x$$

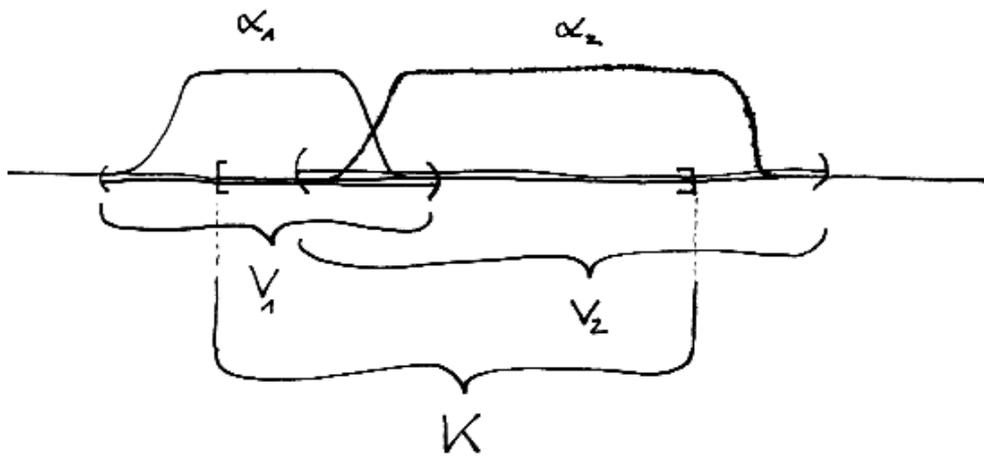


Illustration einer Teilung der Eins im Fall einer Überdeckung eines kompakten Intervalls $K \subset \mathbb{R}$ durch zwei offene Teilmengen.

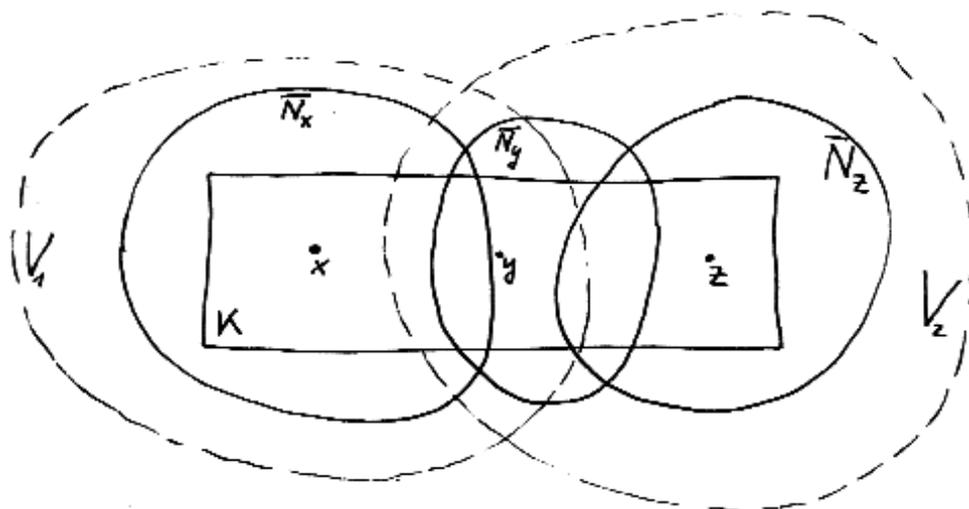


Illustration einiger Mengen, die bei unserer Konstruktion einer Teilung der Eins eine Rolle spielen, im Fall einer Überdeckung eines kompakten Quaders $K \subset \mathbb{R}^2$ durch zwei offene Teilmengen.

Diese Funktionen sind zwar a priori nur auf N definiert, aber da \mathbb{R}^n durch N und das Komplement des Trägers von β überdeckt wird, lassen sie sich stetig durch Null auf ganz \mathbb{R}^n fortsetzen, und diese Fortsetzungen haben dann offensichtlich die gewünschten Eigenschaften. \square

Ergänzung 7.4.12 (Glatte Teilung der Eins). Im vorherigen Lemma können die Funktionen α_i sogar glatt, als da heißt beliebig gemischt partiell differenzierbar gewählt werden. Um das zu sehen, sind nur wenige Zusatzüberlegungen von Nöten. Aus 7.2.11 kennen wir ja eine glatte Funktion $f : \mathbb{R} \rightarrow \mathbb{R}$, die auf der negativen Halbgeraden verschwindet und auf der echt positiven Halbgeraden positiv ist. Dann ist das Produkt $f(t)f(1-t)$ eine von Null verschiedene nichtnegative glatte Funktion mit kompaktem Träger auf \mathbb{R} . Man erhält von Null verschiedene nichtnegative glatte Funktionen mit kompaktem Träger auf \mathbb{R}^n , indem man von Null verschiedene nichtnegative glatte Funktionen mit kompaktem Träger in den einzelnen Koordinaten nimmt und das Produkt bildet. So sehen wir, daß die φ_x im vorhergehenden Beweis sogar glatt gewählt werden können. Damit sind dann auch ψ und die ψ_x glatt. Wählen wir zusätzlich die Funktion h glatt, bis auf Reskalierung könnte man für h etwa das Integral einer von Null verschiedenen nichtnegativen glatten Funktion mit kompaktem Träger nehmen, so liefert die Konstruktion aus dem vorhergehenden Beweis sogar eine glatte Teilung der Eins.

Übung 7.4.13. Man zeige, daß die Funktionaldeterminante der Kugelkoordinatenabbildung K aus 6.2.11 gegeben wird durch $\det dK = r^2 \sin \vartheta$. Salopp gesprochen transformieren sich also Volumenintegrale in Kugelkoordinaten vermittle der Regel

$$dx dy dz = r^2 \sin \vartheta \, dr \, d\varphi \, d\vartheta$$

7.5 Integration über Untermannigfaltigkeiten des \mathbb{R}^n

Satz 7.5.1 (Integration über Untermannigfaltigkeiten). Sei $M \subset \mathbb{R}^n$ eine k -dimensionale Untermannigfaltigkeit und $\mathcal{C}_c(M, \mathbb{R})$ der reelle Vektorraum aller stetigen Funktionen $f : M \rightarrow \mathbb{R}$ mit kompaktem Träger. So gibt es genau eine \mathbb{R} -lineare Abbildung

$$\int_M : \mathcal{C}_c(M, \mathbb{R}) \rightarrow \mathbb{R}$$

mit der Eigenschaft, daß für jede Karte $\varphi : W \rightarrow M$ und jede Funktion $f \in \mathcal{C}_c(M, \mathbb{R})$ mit Träger im Bild der Karte alias $\text{supp } f \subset \varphi(W)$ gilt

$$\int_M f = \int_W f(\varphi(x)) \sqrt{\det (d_x \varphi)^\top (d_x \varphi)} \, d^k x$$

7.5.2. Dieser Satz ist formal betrachtet noch nicht besonders nützlich, da für das Lösen praktischer Aufgaben meist Funktionen zu integrieren sind, die eben gerade nicht kompakten Träger haben. Dennoch sollte aus dem folgenden ersichtlich sein, wie in der Praxis gerechnet werden muß. Wir verschieben die formale Rechtfertigung dieser Rechnungen auf den Abschnitt über das Lebesgue-Integral, vergleiche 9.9.1 folgende, in dessen Rahmen sie besonders einfach formuliert und bewiesen werden können. Bevor ich den obigen Satz beweise, will ich erst einmal versuchen, ihn zu motivieren und den darin erklärten Integralbegriff mit Anschauung zu füllen.

Ergänzung 7.5.3. Gegeben ein endlichdimensionaler reeller affiner Raum X und ein Skalarprodukt auf seinem Richtungsraum mit Einheiten im orientierten eindimensionalen Vektorraum L im Sinne von ?? liefern die analogen Definitionen für jede k -dimensionale Untermannigfaltigkeit $M \subset X$ ein Integral

$$\int_M : \mathcal{C}_c(M, \mathbb{R}) \rightarrow L^{\otimes k}$$

Wie bereits erwähnt messen sich also auch in der Mathematik “Längen in Metern, Flächen in Quadratmetern und Volumen in Kubikmetern”. Betrachten wir noch allgemeiner Funktionen mit Werten in einem endlichdimensionalen reellen Vektorraum V , so wird unser Integral noch allgemeiner eine Abbildung $\int_M : \mathcal{C}_c(M, V) \rightarrow L^{\otimes k} \otimes V$.

Beispiel 7.5.4. Ist $\varphi : [a, b] \rightarrow \mathbb{R}^n$ eine Kurve derart, daß das Bild des offenen Intervalls $\varphi((a, b))$ eine 1-Mannigfaltigkeit M ist und $\varphi : (a, b) \rightarrow M$ eine Karte von M , so ist das Integral einer stetigen Funktion mit kompaktem Träger $f : M \rightarrow \mathbb{R}$ über M genau das Kurvenintegral der durch Null auf die Endpunkte fortgesetzten Funktion f längs der Kurve φ im Sinne von 10.3.6.

Beispiel 7.5.5 (Oberfläche der Einheitskugel). Lassen wir aus der Kugelschale $S^2 = \{(x, y, z) \mid x^2 + y^2 + z^2 = 1\}$ den Äquator weg, also alle Punkte (x, y, z) mit $z = 0$, und dazu noch einen halben Großkreis von Pol zu Pol, sagen wir alle Punkte (x, y, z) mit $y = 0$ und $x \leq 0$, so ist der Rest die disjunkte Vereinigung von zwei geschlitzten offenen Hemisphären $U_+ \cup U_-$ und U_{\pm} ist das Bild der Karte

$$\begin{aligned} \varphi_{\pm} : (0, 1) \times (-\pi, \pi) &\rightarrow S^2 \\ (r, \theta) &\mapsto (r \cos \theta, r \sin \theta, \pm \sqrt{1 - r^2}) \end{aligned}$$

Die Jacobi-Matrix ergibt sich zu

$$d\varphi_{\pm} = \begin{pmatrix} \cos \theta & -r \sin \theta \\ \sin \theta & r \cos \theta \\ \mp r / \sqrt{1 - r^2} & 0 \end{pmatrix}$$

und wir erhalten als Gram'sche Matrix

$$(d\varphi_{\pm})^{\top}(d\varphi_{\pm}) = \begin{pmatrix} 1/(1-r^2) & 0 \\ 0 & r^2 \end{pmatrix}$$

Die Wurzel aus der Determinante der Gram'schen Matrix ergibt sich damit zu $r/\sqrt{1-r^2}$ und wir folgern für jede stetige Funktion $f : S^2 \rightarrow \mathbb{R}$ mit Träger in $U_+ \cup U_-$ die Formel

$$\int_{S^2} f = \int_{-\pi}^{\pi} \int_0^1 f(\varphi_+(r, \theta)) \frac{r}{\sqrt{1-r^2}} dr d\theta + \int_{-\pi}^{\pi} \int_0^1 f(\varphi_-(r, \theta)) \frac{r}{\sqrt{1-r^2}} dr d\theta$$

Argumentieren wir nun heuristisch, gehen davon aus, daß dieses Weglassen eines Längen- und eines Breitengrades das Integral einer stetigen Funktion ja wohl kaum ändern kann, und wenden unsere Formel mutig auf die konstante Funktion Eins an, obwohl unser Satz das gar nicht erlaubt, so erhalten wir heuristisch für die Oberfläche der Einheitskugel das Ergebnis

$$\int_{S^2} 1 = 2 \int_{-\pi}^{\pi} \int_0^1 \frac{r}{\sqrt{1-r^2}} dr d\theta = 4\pi$$

Im Rahmen der Lebesgue'schen Integrationstheorie werden wir Sätze kennenlernen, mit deren Hilfe sich die obige heuristische Argumentation auch formal rechtfertigen läßt, vergleiche 9.9.6.

7.5.6. Hat M die Dimension k , so ist $d_x\varphi$ eine Matrix mit k Spalten und n Zeilen und das Produkt $(d_x\varphi)^{\top}(d_x\varphi)$ dieser Matrix mit ihrer Transponierten ist folglich eine $(k \times k)$ -Matrix. Diese sogenannte **Gram'sche Matrix** kann aufgefaßt werden als die Matrix aller Skalarprodukte zwischen Spaltenvektoren von $d_x\varphi$. Sie ist nach ?? insbesondere positiv semidefinit und hat damit eine nichtnegative Determinante. Gegeben eine nicht notwendig quadratische Matrix V mit Spaltenvektoren $v_1, \dots, v_k \in \mathbb{R}^n$ definieren wir ganz allgemein eine reelle Zahl

$$\text{vol } V = \text{vol}(v_1 | \dots | v_k) := \sqrt{\det(V^{\top} V)} = \sqrt{\det(\langle v_i, v_j \rangle)}$$

und nennen sie das **k -dimensionale Volumen** des von den Vektoren v_i aufgespannten Parallelepipeds. Die Wurzel aus der Determinante der Gram'schen Matrix können wir mit dieser Notation auch kürzer schreiben als

$$\sqrt{\det (d_x\varphi)^{\top}(d_x\varphi)} = \text{vol}(d_x\varphi)$$

Im Fall $k = 1$ ist das eindimensionale Volumen eines Vektors nach dieser Definition schlicht seine Länge. Im Fall $k = 2$ bedeutet das zweidimensionale Volumen eines Paares von Vektoren v, w die Fläche des von ihnen aufgespannten Parallelogramms mit den Ecken $0, v, w$ und $v + w$. Um die Bezeichnung "Volumen" für die Zahl $\text{vol}(v_1 | \dots | v_k)$ im Allgemeinen zu rechtfertigen, beachten wir:

1. Es gilt $\text{vol}(v_0|v_1|\dots|v_k) = \text{vol}(v_1|\dots|v_k)$ falls v_0 die Länge 1 hat und senkrecht steht auf allen anderen v_i .
2. Im Fall $k = n$ haben wir $\text{vol}(v_1|\dots|v_n) = |\det(v_1|\dots|v_n)|$. In der Tat, bezeichnet V die in diesem Fall quadratische Matrix mit Spalten v_i , so gilt nach dem Multiplikationssatz für Determinanten $\det(V^\top V) = (\det V)^2$.

Mit diesen Bemerkungen kann unsere anschauliche Interpretation der Zahl $\text{vol}(v_1|\dots|v_k)$ heuristisch auf unsere anschauliche Interpretation der Determinante in ?? und 7.4.9 zurückgeführt werden: Die Fläche eines Parallelogramms im Raum sollte eben das Volumen des Körpers sein, der entsteht, wenn ich mein Parallelogramm “zu einem Toast des Dicke Eins verdicke”.

Ergänzung 7.5.7. Gegeben ein endlichdimensionaler reeller Vektorraum und ein Skalarprodukt s auf seinem Richtungsraum mit Einheiten im orientierten eindimensionalen Vektorraum L im Sinne von ?? liefern die analogen Definitionen für jedes k -Tupel von Vektoren v_1, \dots, v_k ein Element

$$\text{vol}(v_1, \dots, v_k) = \sqrt{\det(s(v_i, v_j))} \in L^{\otimes k}$$

Salopp gesprochen messen sich also auch in der Mathematik Längen in Metern, Flächen in Quadratmetern und Volumen in Kubikmetern.

7.5.8. Wir wollen nun auch unsere Definition des Integrals anschaulich rechtfertigen. Sei dazu (U, φ) eine Karte einer der Einfachheit der Notation halber zweidimensionalen Untermannigfaltigkeit $M \subset \mathbb{R}^n$, die das Rechteck $Q = [0, a] \times [0, c]$ enthält, und sei $f : M \rightarrow \mathbb{R}$ stetig mit Träger in $\varphi(Q)$. Wir betrachten für $r \geq 1$ die äquidistanten Unterteilungen $0 = a_0 < a_1 < \dots < a_r = a$, $0 = c_0 < c_1 < \dots < c_r = c$ der Kanten von Q , bezeichnen mit $q_{i,j} = (a_i, c_j)$ die Gitterpunkte im so gegebenen Raster auf Q und mit $p_{i,j} = \varphi(q_{i,j})$ die Bilder dieser Gitterpunkte in unserer Mannigfaltigkeit M . Dann definieren wir die r -te **Riemannsumme** $S_M^r(f)$ durch die Formel

$$S_M^r(f) = \sum_{i,j=0}^{r-1} f(p_{i,j}) \text{vol}(p_{i+1,j} - p_{i,j} | p_{i,j+1} - p_{i,j})$$

Natürlich hängt diese Summe von der Karte (U, φ) ab, auch wenn das in der Notation nicht zum Ausdruck kommt. Die versprochene Anschauung für unseren Begriff des Integrals einer Funktion über eine Mannigfaltigkeit soll das nun folgende Lemma geben.

Lemma 7.5.9. Sei $M \subset \mathbb{R}^n$ eine zweidimensionale Untermannigfaltigkeit, (U, φ) eine Karte von M , $Q \subset U$ ein kompaktes Rechteck und $f : M \rightarrow \mathbb{R}$ eine stetige Funktion mit Träger im Bild des Inneren von Q . So gilt mit unseren eben definierten Riemannsummen

$$\int_M f = \lim_{r \rightarrow \infty} S_M^r(f)$$

Beweis. Um Indizes zu vermeiden bezeichnen wir die Koordinaten auf \mathbb{R}^2 mit x, y und schreiben φ_x, φ_y für die Spaltenvektoren der Jacobi-Matrix von φ . Die linke Seite ist per definitionem das Integral $\int_U (f \circ \varphi) \operatorname{vol}(\varphi_x | \varphi_y)$. Dies Integral können wir nach 1.3.3 schreiben als den Grenzwert für $r \rightarrow \infty$ gewisser Riemannsummen, die wir Übersichtlichkeit halber mit S_Q^r abkürzen wollen und die gegeben werden durch

$$S_Q^r = \sum_{i,j=0}^{r-1} f(\varphi(q_{i,j})) \operatorname{vol}(\varphi_x(q_{i,j}) | \varphi_y(q_{i,j})) \frac{\operatorname{vol} Q}{r^2}$$

für $\operatorname{vol} Q = ac$ die Fläche unseres Rechtecks Q und damit $(\operatorname{vol} Q)/r^2$ die Fläche der kleinen rechteckigen Felder $Q_{i,j} = [a_i, a_{i+1}] \times [c_j, c_{j+1}]$. Nun ist φ_x gleichmäßig stetig auf dem Kompaktum Q , für alle $\varepsilon > 0$ gibt es also ein $R > 0$ derart, daß gilt

$$\|\varphi_x(p) - \varphi_x(q)\| \leq \varepsilon$$

wann immer p und q im selben kleinen rechteckigen Feld für eine Unterteilung mit $r \geq R$ liegen. Mit dem Mittelwertsatz in mehreren Veränderlichen 10.2.11 folgt für die Vektoren $\varepsilon_{i,j}(r)$, die erklärt werden durch

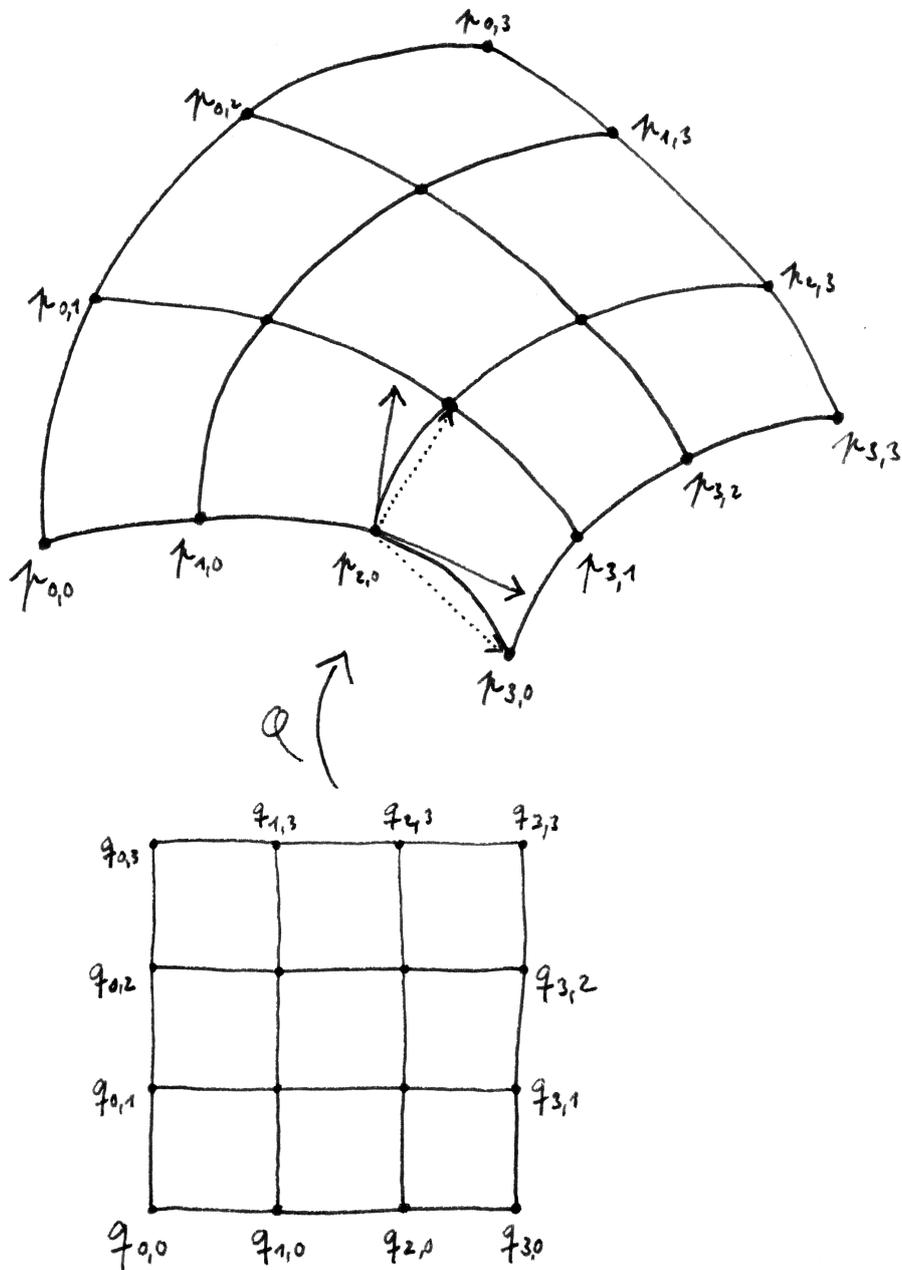
$$p_{i+1,j} - p_{i,j} = \frac{a}{r} (\varphi_x(q_{i,j}) + \varepsilon_{i,j}(r))$$

unter der Voraussetzung $r \geq R$ die Abschätzung $\|\varepsilon_{i,j}(r)\| \leq \varepsilon$. Eine analoge Abschätzung erhalten wir für $p_{i,j+1} - p_{i,j}$. Jetzt setzen wir diese Darstellungen ein in $S_M^r(f)$ und überlassen es dem Leser, hieraus zu folgern, daß gilt

$$\lim_{r \rightarrow \infty} (S_Q^r - S_M^r) = 0$$

Da aber die Folge S_Q^r gegen $\int_M f$ konvergiert, muß dasselbe auch für die Folge S_M^r gelten. \square

Beweis des Satzes zur Integration über Untermannigfaltigkeiten. Die Eindeutigkeit ist nicht schwer zu zeigen: Der Träger $K := \operatorname{supp} f$ unserer Funktion f wird nach 9.10.3 als Kompaktum von den Bildern endlich vieler Karten $(W_1, \varphi_1), \dots, (W_r, \varphi_r)$



Die gepunkteten Pfeile stellen die Vektoren $p_{3,0} - p_{2,0}$ und $p_{2,1} - p_{2,0}$ dar, die Fläche des durch sie bestimmten Parallelogramms geht in die Riemannsumme S_M^3 ein. Die durchgezogenen Pfeile stellen die Vektoren $\varphi_x(q_{2,0})$ und $\varphi_y(q_{2,0})$ dar, die Fläche des durch sie bestimmten Parallelogramms geht entsprechend in die Riemannsumme S_Q^3 ein. Beim Übergang zu immer feineren Rastern kommen wir zum selben Grenzwert, wie im Beweis von 7.5.9 ausgeführt wird.

überdeckt. Nach 9.5.12 gibt es $V_1, \dots, V_r \subseteq \mathbb{R}^n$ mit $\varphi_i(W_i) = V_i \cap M$. Nach 7.4.10 existiert eine an die Überdeckung $K \subset \bigcup V_i$ angepaßte Teilung der Eins α_i . Dann gilt $f = \sum_i \alpha_i f$, und da unsere Bedingung bereits die Integrale der Summanden festlegt, legt die ebenfalls geforderte Linearität dann auch das Integral von f fest. Es bleibt, die Existenz einer linearen Abbildung \int_M mit den geforderten Eigenschaften zu zeigen. Zunächst halten wir dazu unsere Funktion f fest und betrachten eine weitere Überdeckung ihres Trägers durch die Bilder endlich vieler Karten $(U_1, \psi_1), \dots, (U_s, \psi_s)$ und eine zugehörige Teilung der Eins β_j und behaupten die Gleichheit

$$\sum_i \int ((\alpha_i f) \circ \varphi_i)(x) \operatorname{vol}(d_x \varphi_i) d^k x = \sum_j \int ((\beta_j f) \circ \psi_j)(x) \operatorname{vol}(d_x \psi_j) d^k x$$

Sie ist aufgrund der Linearität aller Integrale äquivalent zur Gleichheit

$$\sum_{i,j} \int ((\beta_j \alpha_i f) \circ \varphi_i)(x) \operatorname{vol}(d_x \varphi_i) d^k x = \sum_{i,j} \int ((\beta_j \alpha_i f) \circ \psi_j)(x) \operatorname{vol}(d_x \psi_j) d^k x$$

und folgt, wenn wir die Gleichheit aller Summanden zeigen. Hier haben nun die Funktionen $\beta_j \alpha_i f$ Träger im Schnitt der Karten $\varphi_i(W_i) \cap \psi_j(U_j)$. Wir können also die Indizes weglassen und müssen nur für eine stetige Funktion mit kompaktem Träger $h : M \rightarrow \mathbb{R}$, deren Träger im Bild zweier Karten (W, φ) und (U, ψ) liegt, die Identität

$$\int h(\varphi(x)) \operatorname{vol}(d_x \varphi) d^k x = \int h(\psi(x)) \operatorname{vol}(d_x \psi) d^k x$$

zeigen. Indem wir unsere Karten notfalls noch etwas verkleinern, dürfen wir dabei sogar $\varphi(W) = \psi(U)$ annehmen, so daß der Kartenwechsel nach 7.3.20 ein \mathcal{C}^1 -Diffeomorphismus $g = \psi^{-1} \circ \varphi : W \xrightarrow{\sim} U$ ist mit $\psi \circ g = \varphi : U \rightarrow M$. Es folgt $h(\varphi(x)) = h(\psi(g(x)))$ und $d_x \varphi = d_{g(x)} \psi \circ d_x g$. Wir erhalten mit der Multiplikatitivität der Determinante also

$$\operatorname{vol}(d_x \varphi) = |\det d_x g| \operatorname{vol}(d_{g(x)} \psi)$$

und folgern die behauptete Gleichheit der Integrale aus der Transformationsformel 7.4.6, angewandt auf die charakteristische Funktion $h \circ \psi$. Damit haben wir gezeigt, daß jede Überdeckung des Trägers unserer Funktion f durch Bilder von Karten und jede zugehörige Teilung der Eins in der Formel oben dieselbe Summe liefert, die wir damit als unser $\int_M f$ erklären können. Daß die so erklärte Abbildung $f \mapsto \int_M f$ dann auch \mathbb{R} -linear ist und die geforderte Eigenschaft für Funktionen mit Träger im Bild einer Karte hat, folgt unmittelbar. \square

Übung 7.5.10. Gegeben eine kompakte k -dimensionale Untermannigfaltigkeit $M \subset \mathbb{R}^n$ und eine Isometrie $A : \mathbb{R}^n \xrightarrow{\sim} \mathbb{R}^n$ zeige man

$$\int_M 1 = \int_{A(M)} 1$$

Insbesondere und in Worten bleibt also beim Drehen und Verschieben von Flächen im Raum die Größe ihrer Oberfläche unverändert.

8 Gewöhnliche Differentialgleichungen

8.1 Grundlegende Definitionen und Eigenschaften

8.1.1. Ganz allgemein versteht man unter einer **gewöhnlichen Differentialgleichung** eine Gleichung, in der die Ableitungen einer zu bestimmenden Funktion *einer* Veränderlichen zueinander und mit der Veränderlichen selbst in Beziehung gesetzt werden, wie zum Beispiel in der Gleichung

$$(f''(t))^2 f'(t) + t f(t) = t^2$$

an eine zu bestimmende zweimal differenzierbare Funktion $f : \mathbb{R} \rightarrow \mathbb{R}$. Im Gegensatz dazu stehen zumindest terminologisch die **partiellen Differentialgleichungen**, bei denen die partiellen Ableitungen einer Funktion *mehrerer* Veränderlichen auftreten, wie zum Beispiel in der Gleichung

$$\frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} = 0$$

an eine Funktion $f : \mathbb{R}^2 \rightarrow \mathbb{R}$. Auf Englisch benutzt man die Abkürzungen **ODE** für **ordinary differential equation** und **PDE** für **partial differential equation**. Wir besprechen in diesem Abschnitt nur gewöhnliche Differentialgleichungen. Rein formal ist dieser Abschnitt unabhängig von der Behandlung linearer gewöhnlicher Differentialgleichungen mit konstanten Koeffizienten in 10.4.9 und 2 folgende. Ich denke jedoch, daß eine gewisse Vertrautheit mit diesen einfachsten und wichtigsten Spezialfällen es sehr erleichtern kann, auch die nun folgende allgemeine Theorie zu verstehen.

8.1.2. Die Ordnung der höchsten in einer gewöhnlichen Differentialgleichung auftretenden Ableitung heißt die **Ordnung** unserer Differentialgleichung. Von einer **expliziten** Gleichung spricht man, wenn in unserer Gleichung die Ableitung höchster Ordnung “explizit durch die tieferen Ableitungen ausgedrückt wird”. Andernfalls spricht man von einer **impliziten** Gleichung. Bei unserem obigen Beispiel handelt es sich also um eine implizite Gleichung. Wir werden uns im folgenden jedoch nur mit expliziten gewöhnlichen Differentialgleichungen beschäftigen. Eine derartige explizite Gleichung der Ordnung n hat, wenn wir von der Spezifikation allgemeinstmöglicher Definitionsbereiche einmal absehen, die Gestalt

$$f^{(n)}(t) = F(t, f(t), f'(t), \dots, f^{(n-1)}(t))$$

mit einer Abbildung $F : \mathbb{R}^{n+1} \rightarrow \mathbb{R}$. Gesucht sind dann alle Funktionen $f : \mathbb{R} \rightarrow \mathbb{R}$, die n -mal differenzierbar sind und eben zusammen mit ihren Ableitungen diese Gleichung erfüllen. Etwas allgemeiner betrachten wir zugleich auch **Systeme**

von gewöhnlichen Differentialgleichungen, bei denen vektorwertige Funktionen $f = (f_1, \dots, f_k) : \mathbb{R} \rightarrow \mathbb{R}^k$ gesucht werden derart, daß eine Gleichung der obigen Gestalt gilt, die nun aber eine vektorwertige Gleichung meint mit einer vorgegebenen Abbildung $F : \mathbb{R}^{kn+1} \rightarrow \mathbb{R}^k$.

8.1.3 (Reduktion auf Systeme erster Ordnung). Die Betrachtung von *Systemen* gewöhnlicher Differentialgleichungen erlaubt uns zumindest für Fragen des allgemeinen Lösungsverhaltens die Beschränkung auf den Fall erster Ordnung. Betrachten wir beispielhaft den Fall einer Gleichung dritter Ordnung

$$f'''(t) = F(t, f(t), f'(t), f''(t))$$

Jede Lösung f liefert sicher eine Abbildung $\gamma : \mathbb{R} \rightarrow \mathbb{R}^3$ mittels der Vorschrift $\gamma(t) = (f(t), f'(t), f''(t))$, und natürlich gilt dann

$$\begin{aligned}\gamma_1'(t) &= \gamma_2 \\ \gamma_2'(t) &= \gamma_3(t) \\ \gamma_3'(t) &= F(t, \gamma_1(t), \gamma_2(t), \gamma_3(t))\end{aligned}$$

Erklären wir nun also eine neue Abbildung $A : \mathbb{R}^4 \rightarrow \mathbb{R}^3$ durch die Vorschrift $A(t, (x, y, z)) = (y, z, F(t, x, y, z))$, so ist unser γ eine Lösung der Differentialgleichung

$$\gamma'(t) = A(t, \gamma(t))$$

Umgekehrt zeigt man leicht, daß für jede Lösung $\gamma : \mathbb{R} \rightarrow \mathbb{R}^3$ dieses Systems von Differentialgleichungen erster Ordnung die erste Komponente $\gamma_1(t) = f(t)$ eine Lösung unserer ursprünglichen Gleichung dritter Ordnung liefert. In derselben Weise kann auch im Allgemeinen die Frage nach der Existenz und Eindeutigkeit der Lösungen von Systemen gewöhnlicher Differentialgleichungen höherer Ordnung auf den Fall von Systemen erster Ordnung zurückgeführt werden. Anschaulich mag man sich dann A als ein zeitabhängiges Vektorfeld auf dem \mathbb{R}^n denken, das jedem Ort $x \in \mathbb{R}^n$ zu jedem Zeitpunkt $t \in \mathbb{R}$ einen Vektor $A(t, x) \in \mathbb{R}^n$ zuordnet. In dieser Anschauung beschreibt eine Lösung $\gamma : \mathbb{R} \rightarrow \mathbb{R}^n$ die Bewegung eines Teilchens, das zu jedem Zeitpunkt t die für seinen Ort zu diesem Zeitpunkt durch unser zeitabhängiges Vektorfeld A vorgegebene Geschwindigkeit hat.

8.1.4 (Reduktion auf den zeitunabhängigen Fall). Gegeben $A : \mathbb{R}^{n+1} \rightarrow \mathbb{R}^n$ löst eine differenzierbare Abbildung $\gamma : \mathbb{R} \rightarrow \mathbb{R}^n$ unsere Differentialgleichung

$$\gamma'(t) = A(t, \gamma(t))$$

genau dann, wenn die Abbildung $\psi : \mathbb{R} \rightarrow \mathbb{R}^{n+1}$, $\psi(t) = (t, \gamma(t))$ die Differentialgleichung

$$\psi'(t) = B(\psi(t))$$

löst für $B : \mathbb{R}^{n+1} \rightarrow \mathbb{R}^{n+1}$ gegeben durch $B(t, x) = (1, A(t, x))$. In diesem Sinne können wir uns also stets auf den Fall zeitunabhängiger Felder zurückziehen. Allerdings erhält man für zeitabhängige Felder bei einer eigenständigen Betrachtung etwas schärfere Existenz- und Eindeutigkeitsaussagen, weshalb dieser Fall insbesondere in einigen Ergänzungen weiter vorkommen wird. Zunächst konzentrieren wir uns nun jedoch auf den zeitunabhängigen Fall und besprechen seine geometrische Bedeutung in einer koordinatenfreien Sprache.

8.1.5. Unter einem **Vektorfeld** auf einer halboffenen Teilmenge $U \subset X$ eines normierten reellen Raums X verstehen wir wie in 6.1.2 eine Abbildung A von U in den Richtungsraum \vec{X} von X , in Formeln

$$\begin{aligned} A : U &\rightarrow \vec{X} \\ p &\mapsto A_p \end{aligned}$$

Definition 8.1.6. Sei X ein normierter reeller Raum, $U \subset X$ eine halboffene Teilmenge und $A : U \rightarrow \vec{X}$ ein Vektorfeld. Eine **Integralkurve** unseres Vektorfelds ist eine differenzierbare Abbildung $\gamma : I \rightarrow U$ von einem halboffenen reellen Intervall $I \subset \mathbb{R}$ nach U mit der Eigenschaft, daß “zu jedem Zeitpunkt $t \in I$ die Geschwindigkeit unserer Integralkurve zum Zeitpunkt t genau der durch das Vektorfeld vorgegebene Vektor an der Stelle $\gamma(t)$ ist”, in Formeln

$$\dot{\gamma}(t) = A(\gamma(t)) \quad \forall t \in I$$

Eine **maximale Integralkurve** ist eine Integralkurve, die nicht zu einer auf einem echt größeren reellen Intervall definierten Integralkurve erweitert werden kann. Ist $p \in U$ gegeben, so verstehen wir unter einer **Integralkurve mit Anfangswert p** oder kurz einer **Integralkurve zu p** eine Integralkurve (γ, I) mit $0 \in I$ und $\gamma(0) = p$.

8.1.7. Unsere Terminologie ist hier nicht vollständig konsistent, da mit dieser Definition eine Integralkurve keine Kurve im Sinne unserer Definition 7.3.9 ist, sondern vielmehr, zumindest im Fall eines offenen Definitionsintervalls und unter geeigneten Injektivitätsvoraussetzungen, eine Karte einer Kurve. Die Integralkurven eines Vektorfelds bilden im Übrigen eine “durch Einschränkung” partiell geordnete Menge, und unsere maximalen Integralkurven sind darin gerade die maximalen Elemente.

Beispiel 8.1.8. Ist unser Vektorfeld konstant, so laufen seine Integralkurven auf den Geraden mit diesem konstanten Vektor als Richtungsvektor und mit der durch diesen Vektor vorgegebenen konstanten Geschwindigkeit. Ist unser Vektorfeld auf einem endlichdimensionalen reellen Vektorraum $X = V$ definiert durch eine lineare Abbildung, sagen wir durch die lineare Abbildung $A \in \text{End } V$, so haben wir bereits in 10.4.9 gezeigt, daß seine maximalen Integralkurven genau

diejenigen Abbildungen $\mathbb{R} \rightarrow V$ sind, die gegeben werden durch die Formeln $\gamma(t) = \exp(tA)c$ mit $c \in V$.

8.1.9 (Zeitverschiebung). Gegeben ein Vektorfeld auf einer halboffenen Teilmenge eines normierten reellen Raums und eine Integralkurve (γ, I) ist für alle $c \in \mathbb{R}$ auch die Abbildung $t \mapsto \gamma(t + c)$ eine Integralkurve, die nun eben definiert ist auf dem verschobenen Intervall $I - c$. Das gilt im Fall zeitabhängiger Vektorfelder natürlich so nicht mehr.

Übung 8.1.10 (Verhalten unter differenzierbaren Abbildungen). Entsprechen sich unter einem Diffeomorphismus zwei Vektorfelder, so entsprechen sich auch deren Integralkurven. Allgemeiner haben offensichtlich im Sinne von 6.1.16 verwandte Vektorfelder auch verwandte Integralkurven. Ist genauer unter einer stetig differenzierbaren Abbildung ϕ ein Vektorfeld A verwandt zu einem Vektorfeld B , so ist für jede Integralkurve γ von A auch $\phi \circ \gamma$ eine Integralkurve von B . Ist insbesondere ein Vektorfeld A unter einer stetig differenzierbaren Abbildung ϕ verwandt zum Nullfeld, in Formeln $\phi : A \rightsquigarrow 0$, und ist γ eine seiner Integralkurven, so ist $\phi \circ \gamma$ eine Integralkurve des Nullfelds und mithin konstant, als da heißt, die Funktion ϕ ist konstant auf Integralkurven von A . Man nennt die Funktion ϕ dann auch ein **erstes Integral** unserer Differentialgleichung. In physikalischen Modellen liefern oft Energie, Impuls und Drehimpuls solche ersten Integrale.

Satz 8.1.11 (Picard-Lindelöf). *1. Gegeben ein stetig differenzierbares Vektorfeld auf einer offenen Teilmenge eines endlichdimensionalen reellen Raums gibt es zu jedem Anfangswert genau eine größte Integralkurve.*

2. Diese größte Integralkurve hat als Definitionsbereich ein offenes Intervall, und ist dieses Intervall nach oben beschränkt, so verläßt die fragliche Integralkurve für positive Zeiten jedes Kompaktum aus unserer offenen Teilmenge irgendwann einmal endgültig.

8.1.12. Wir zeigen diesen Satz als 1 sogar unter noch etwas schwächeren Voraussetzungen. Der letzte Teil des Satzes besagt salopp formuliert, daß der Grund dafür, daß sich eine Integralkurve nicht beliebig weit in Richtung positiver Zeiten fortsetzen läßt, nur darin liegen kann, daß sie bereits in endlicher Zeit “aus dem Definitionsbereich des Vektorfeldes hinausläuft”. Entsprechendes gilt in Richtung negativer Zeiten, was man auch formal durch Betrachtung des mit (-1) multiplizierten Vektorfeldes leicht folgern kann. Von unserer Integralkurve fordern wir natürlich nur, daß sie in der partiell geordneten Menge aller Integralkurven *zu unserem festen Anfangswert* das größte Element sein soll. Da größte Elemente in partiell geordneten Mengen eh eindeutig bestimmt sind, war die Präzisierung “genau eine” recht eigentlich überflüssig und nur dazu bestimmt, eventuellen Mißverständnissen vorzubeugen. In der Menge aller Integralkurven ist diese Integral-

kurve nur noch maximal und wird deshalb meist die **maximale Integralkurve** zu unserem Anfangswert genannt.

Beispiel 8.1.13 (Ein Fall mit nicht eindeutigen Integralkurven). Bei Vektorfeldern, die nicht stetig differenzierbar sind, kann es durchaus vorkommen, daß zu einem vorgegebenen Anfangswert keine größte Integralkurve existiert, weil etwa mehrere maximale Integralkurven mit ein und demselben Anfangswert existieren, die auf dem Schnitt ihrer Definitionsbereiche nicht übereinstimmen. Betrachten wir zum Beispiel auf \mathbb{R}^2 das Vektorfeld A , für das sämtliche verschobenen Kubiken $\gamma_c(t) = (t + c, t^3)$ Integralkurven sind. Wir haben $\dot{\gamma}_c(t) = (1, 3t^2)$ und damit $A(x, y) = (1, 3|y|^{2/3})$. Maximale Integralkurven sind in diesem Fall nicht nur die verschobenen Kubiken γ_c , sondern auch alle Kurven, die längs einer verschobenen Kubik auf die x -Achse hochsteigen und dann eine Weile auf der x -Achse entlanglaufen bevor sie auf einer anderen verschobenen Kubik weitersteigen. In diesem Fall existieren zwar maximale Integralkurven zu jedem Punkt, von Eindeutigkeit kann aber keine Rede sein.

Beispiel 8.1.14 (Der Fall eindimensionaler Felder ohne Nullstellen). Gegeben ein stetiges Vektorfeld ohne Nullstellen auf einer offenen Teilmenge eines eindimensionalen Raums gibt es zu jedem Anfangswert genau eine maximale Integralkurve. In diesem Fall brauchen wir also von unserem Vektorfeld nicht einmal stetige Differenzierbarkeit zu fordern. In der Tat sei ohne Beschränkung der Allgemeinheit $U \subseteq \mathbb{R}$ ein Intervall und unser stetiges Vektorfeld ohne Nullstellen zeige in Richtung der positiven x -Achse, als da heißt, es werde gegeben durch $a : U \rightarrow \mathbb{R}_{>0}$. Integralkurven sind auf halboffenen Intervallen $I \subseteq \mathbb{R}$ definierte differenzierbare Funktionen $\gamma : I \rightarrow U$ mit

$$\dot{\gamma}(t) = a(\gamma(t)) \quad \forall t \in I$$

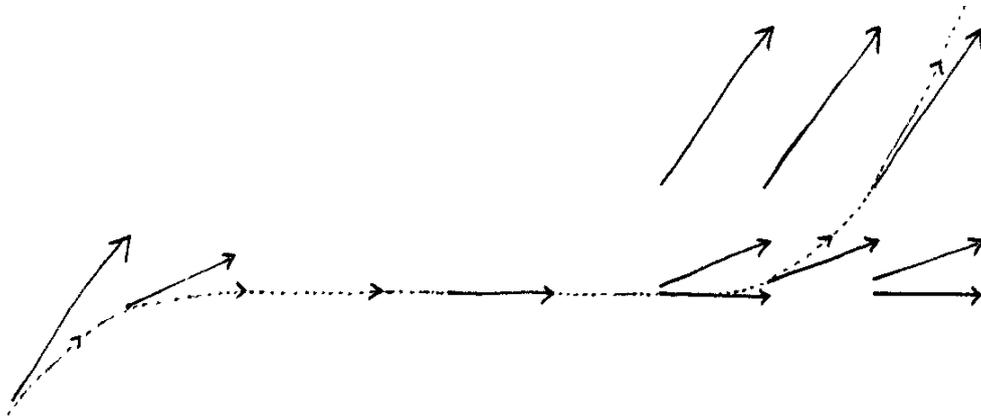
Aus dieser Gleichung folgt für alle $s, t \in I$ sofort

$$t - s = \int_s^t \frac{\dot{\gamma}(\tau) \, d\tau}{a(\gamma(\tau))} = \int_{\gamma(s)}^{\gamma(t)} \frac{dx}{a(x)} = G(\gamma(t)) - G(\gamma(s))$$

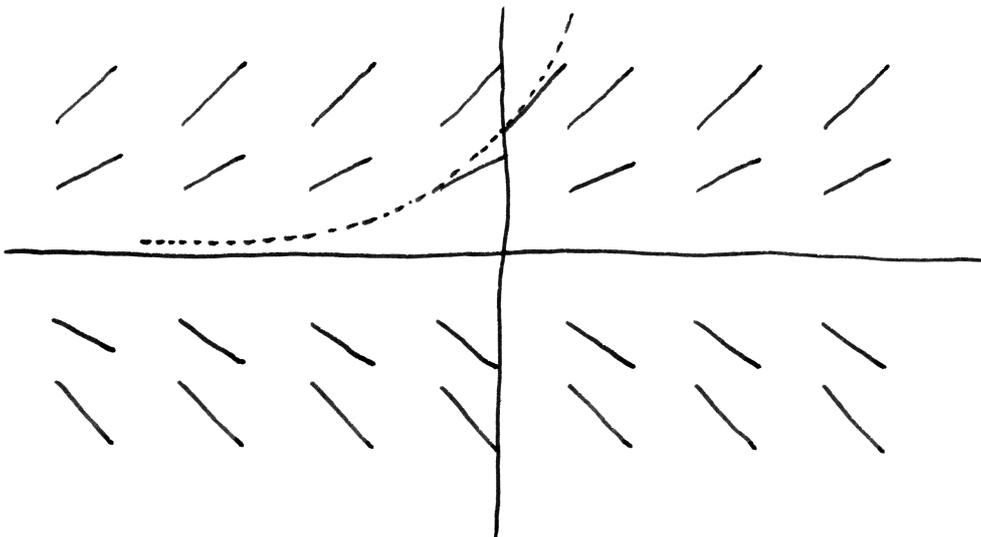
für $G : U \rightarrow \mathbb{R}$ eine Stammfunktion von $1/a$. Nun wächst G sicher streng monoton und hat folglich als Bild ein offenes Intervall $J \subseteq \mathbb{R}$ und für unsere Integralkurve folgt $\gamma(t) = G^{-1}(t + c)$ mit der Konstanten $c = G(\gamma(s)) - s$. In anderen Worten ist

$$G^{-1} : J \xrightarrow{\sim} U$$

bis auf "Zeitverschiebung" die einzige maximale Integralkurve. Zum Beispiel ist $a(x) = x$ ein stetiges Vektorfeld ohne Nullstellen auf $U = \mathbb{R}_{>0}$ und $G(x) = \log x$ ist eine Stammfunktion von $1/x$ und jede maximale Integralkurve ist von der Gestalt $\gamma : \mathbb{R} \rightarrow \mathbb{R}_{>0}$, $\gamma(t) = \exp(t + c)$ mit einer Konstanten $c \in \mathbb{R}$.



Das ebene stetige aber nicht stetig differenzierbare Vektorfeld aus Beispiel [8.1.13](#) mit einer seiner Integralkurven



Im Fall eindimensionaler Felder mag man sich die Lösung der entsprechenden Differentialgleichung durch ihren im Bild gestrichelt eingezeichneten Graphen veranschaulichen und das Vektorfeld als eine Vorschrift, die diesem Graphen in jeder Höhe x eine Steigung $a(x)$ vorschreibt.

Übung 8.1.15 (Größere Felder haben schnellere Integralkurven). Gegeben $U \subset \mathbb{R}$ halboffen und $a, b : U \rightarrow \mathbb{R}$ stetig ohne Nullstelle mit $a \leq b$ und $I \subset \mathbb{R}$ ein halboffenes Intervall und $\gamma, \kappa : I \rightarrow U$ differenzierbar mit $\dot{\gamma}(t) = a(\gamma(t))$ und $\dot{\kappa}(t) = b(\kappa(t))$ für alle $t \in I$ folgt aus $\gamma(t_0) \leq \kappa(t_0)$ für ein $t_0 \in I$ bereits dieselbe Aussage für alle $t \in I$ mit $t \geq t_0$.

Beispiel 8.1.16 (Spezielle eindimensionale Felder mit Nullstellen). Gegeben ein stetiges Vektorfeld mit Nullstellen auf einer offenen Teilmenge eines eindimensionalen Raums liegen die Verhältnisse komplizierter als in 8.1.14. Wir suchen etwa für $\alpha \in \mathbb{R}$ Integralkurven des Vektorfelds $a(x) = x^\alpha$ auf $\mathbb{R}_{>0}$ alias auf einem halboffenen Intervall $I \subset \mathbb{R}$ definierte Funktionen $\gamma : I \rightarrow \mathbb{R}_{>0}$ mit

$$\dot{\gamma}(t) = (\gamma(t))^\alpha \quad \forall t \in I$$

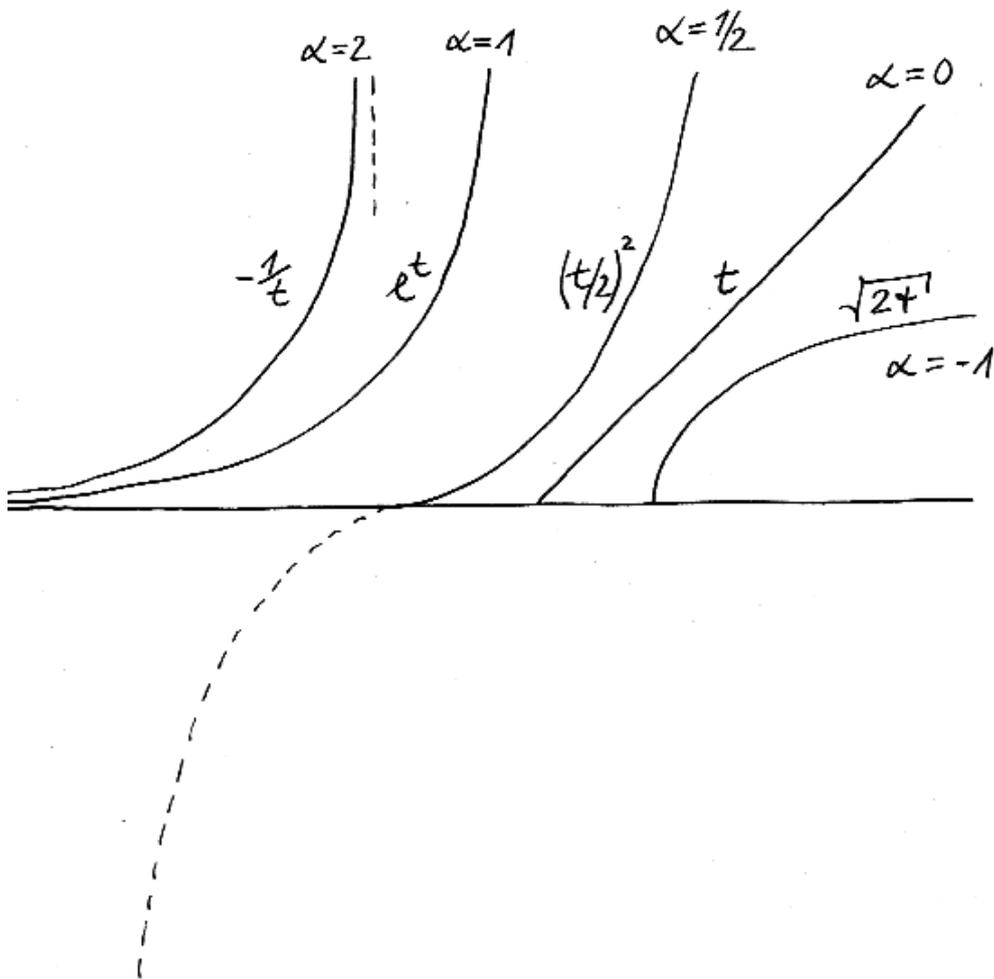
Unsere allgemeine Theorie aus 8.1.14 sagt uns, daß das gerade die Umkehrfunktion zu Stammfunktionen von $x^{-\alpha}$ sind. Den Fall $\alpha = 1$ kennen wir zur Genüge, im Fall $\alpha \neq 1$ erhalten wir als Stammfunktion $G(x) = x^{1-\alpha}/(1-\alpha)$. Im Fall $\alpha > 1$ induziert nun G eine Bijektion $G : \mathbb{R}_{>0} \xrightarrow{\sim} \mathbb{R}_{<0}$ und im Fall $\alpha < 1$ eine Bijektion $G : \mathbb{R}_{>0} \xrightarrow{\sim} \mathbb{R}_{>0}$, aber die Umkehrfunktion wird jedesmal durch dieselbe Formel gegeben und wir erhalten die Integralkurven

$$\gamma(t) = ((1-\alpha)t)^{(1-\alpha)^{-1}}$$

Im Fall $\alpha = 2$ etwa ergibt sich $\gamma(t) = -1/t$ und unsere Integralkurve “läuft in endlicher Zeit nach $+\infty$, braucht aber, wenn wir die Zeit rückwärts laufen lassen, unendlich lange bis zum Ursprung”. Dasselbe gilt in allen Fällen mit $\alpha > 1$. Im Fall $\alpha = 0$ dahingegen ergibt sich $\gamma(t) = t$ und unsere Integralkurve “läuft für alle positiven Zeiten, braucht aber, wenn wir die Zeit rückwärts laufen lassen, nur endlich viel Zeit bis zum Ursprung”. Dasselbe gilt in allen Fällen mit $\alpha < 1$. In den Fällen mit $0 \leq \alpha$ können wir unser Vektorfeld stetig auf \mathbb{R} fortsetzen durch die Vorschrift $a(x) = |x|^\alpha$ und für $0 < \alpha$ hat diese Fortsetzung eine Nullstelle bei Null. In den Fällen $0 < \alpha < 1$ gibt es nun auch Integralkurven, die in endlicher Zeit aus dem Negativen nach Null laufen und dort eine Weile stehenbleiben bevor sie ins Positive weiterlaufen. Erklären wir etwa auf \mathbb{R} ein stetiges Vektorfeld durch $a(x) = \sqrt[3]{x^2}$, so ist $\gamma(t) = t^3/27$ ein Integralkurve, aber auch die Abbildung $\psi : \mathbb{R} \rightarrow \mathbb{R}$, die gegeben wird durch die Vorschrift

$$\psi(t) = \begin{cases} t^3/27 & t \leq 0; \\ 0 & 0 \leq t \leq 1; \\ (t-1)^3/27 & 1 \leq t; \end{cases}$$

ist eine Integralkurve. Im Rückblick ist unser Beispiel 8.1.13 im wesentlichen dasselbe, nur in trivialer Weise um eine Dimension erweitert, damit es besser bildlich dargestellt werden kann.



Mögliche Lösungsfunktionen mit positiven Werten der Differentialgleichung $\dot{\gamma}(t) = (\gamma(t))^\alpha$ für verschiedene Werte von $\alpha \in \mathbb{R}$. Alle anderen Lösungsfunktionen mit positiven Werten erhält man durch horizontales Verschieben der entsprechenden Graphen. Im Fall $\alpha > 1$ “läuft unsere Lösung in endlicher Zeit nach Unendlich”, was die gestrichelt eingezeichnete vertikale Asymptote andeuten soll. Im Fall $1 > \alpha > 0$ kann man, wie gestrichelt angedeutet, die Lösung zu einer Lösung von $\dot{\gamma}(t) = |\gamma(t)|^\alpha$ ins Negative fortsetzen, aber eben auf vielerlei Weisen. Das war im Wesentlichen auch unser Gegenbeispiel [8.1.13](#).

Beispiel 8.1.17 (Verhalten unter Längenänderungen). Ändern wir bei einem Vektorfeld ohne Nullstellen nur die Längen seiner Vektoren, durchaus auch in Abhängigkeit vom Ort, so bleiben die Integralkurven offensichtlich bis auf Reparametrisierung dieselben. Ist also in Formeln X ein endlichdimensionaler reeller Raum, $U \subseteq X$ offen, $A : U \rightarrow \vec{X}$ ein Vektorfeld ohne Nullstellen und $c : U \rightarrow \mathbb{R}$ eine stetige Funktion ohne Nullstellen und ist $\gamma : I \rightarrow U$ eine Integralkurve von A , so finden wir mit dem Ansatz $\psi(t) = \gamma(r(t))$ eine Lösung der Differentialgleichung

$$\dot{\psi}(t) = c(\psi(t))A(\psi(t))$$

In der Tat liefert diese Gleichung nämlich für die Reparametrisierung r die Gleichung

$$\dot{r}(t)\dot{\gamma}(r(t)) = c(\gamma(r(t)))A(\gamma(r(t)))$$

und damit $\dot{r}(t) = (c \circ \gamma)(r(t))$, und diese Gleichung haben wir bereits in [8.1.14](#) lösen gelernt.

8.1.18 (Felder mit separierten Variablen). Gegeben endlichdimensionale reelle Räume X, Z und offene Teilmengen $U \subseteq X, V \subseteq Z$ und Vektorfelder $A : U \rightarrow \vec{X}$ sowie $B : V \rightarrow \vec{Z}$ sind die Integralkurven des Vektorfelds $(A \times B) : U \times V \rightarrow \vec{X} \times \vec{Z}$ genau die Abbildungen (γ, ψ) mit $\gamma : I \rightarrow U$ einer Integralkurve von A und $\psi : I \rightarrow V$ einer Integralkurve von B . In dieser Situation spricht man von einer Differentialgleichung mit getrennten Veränderlichen oder lateinisch **separierten Variablen**. Man beachte die enge Beziehung zu [8.1.10](#). Die in [8.1.21](#) erläuterte Methode der ‘‘Separation der Variablen’’ mag man auffassen als das Überführen einer Differentialgleichung in eine Gleichung mit separierten Variablen.

8.1.19 (Geometrische Interpretation im zeitabhängigen Fall). Allgemeiner können wir für einen endlichdimensionalen reellen Raum X eine Abbildung $A : \mathbb{R} \times X \rightarrow \vec{X}$ als ein **zeitabhängiges Vektorfeld** auf X auffassen und uns die Lösungen der Differentialgleichung

$$\dot{\gamma}(t) = A(t, \gamma(t))$$

in dieser Weise veranschaulichen. Dasselbe gilt, wenn A nur auf einer Teilmenge $U \subset \mathbb{R} \times X$ definiert ist. Der Fall eines zeitabhängigen Vektorfelds A kann leicht auf den Fall des zeitunabhängigen Vektorfelds $(1, A) : U \rightarrow \mathbb{R} \times \vec{X}$ zurückgeführt werden: In der Tat ist γ eine Integralkurve unseres zeitabhängigen Vektorfelds genau dann, wenn (id, γ) eine Integralkurve des zeitunabhängigen Vektorfelds $(1, A)$ ist, und jede Integralkurve von $(1, A)$ ist etwa nach [8.1.10](#) bis auf eine Zeitverschiebung von dieser Gestalt. Allerdings gelingt es im Fall zeitabhängiger Felder, die Existenz und Eindeutigkeit der Lösung bei einer direkten Betrachtung in ein bißchen größerer Allgemeinheit zu zeigen, vergleiche [8.2.9](#).

Beispiel 8.1.20 (Eindimensionale zeitabhängige Felder). Gegeben sind etwa $U \subseteq \mathbb{R}^2$ und $a : U \rightarrow \mathbb{R}$ und man interessiert sich Lösungen der Gleichung

$$\dot{x} = a(t, x)$$

Unter einer “Lösung” versteht man hierbei ein Paar (γ, I) mit $I \subset \mathbb{R}$ einem halboffenen Intervall und $\gamma : I \rightarrow \mathbb{R}$ einer differenzierbaren Funktion, deren Graph in U enthalten ist und für die gilt

$$\dot{\gamma}(t) = a(t, \gamma(t)) \quad \forall t \in I$$

Ich habe hier Lösungen als $\gamma(t)$ und nicht als $x(t)$ geschrieben, wie es die Gleichung suggeriert, in der Hoffnung, daß das zum besseren Verständnis beiträgt. Hängt $a(t, x)$ gar nicht von x ab, also $a(t, x) = a(t)$, so sind die Lösungen unserer Differentialgleichung natürlich genau die Stammfunktionen von a . Hängt $a(t, x)$ dahingegen nicht von t ab, also $a(t, x) = a(x)$, so sind die Lösungen unserer Differentialgleichung nichts anderes als die Integralkurven des auf einer geeigneten Teilmenge von \mathbb{R} definierten Vektorfelds $a(x)$, die wir bereits in 8.1.14 diskutiert hatten.

8.1.21. Seien $V, W \subseteq \mathbb{R}$ offen und $a : V \rightarrow \mathbb{R}$, $b : W \rightarrow \mathbb{R}$ stetig. Differentialgleichungen der Gestalt

$$\dot{x} = a(x)b(t)$$

lassen sich oft mit der Methode der **Separation der Variablen** oder deutsch **Variablentrennung** lösen. Ich führe zunächst dieses Verfahren vor und erkläre dann, inwiefern wir dabei implizit unsere Gleichung in eine Gleichung mit separierten Variablen im Sinne von 8.1.18 transformieren. Wir nehmen an, a habe keine Nullstelle und V sei ein Intervall. Gegeben eine Lösung $\gamma : I \rightarrow \mathbb{R}$ kann die Gleichung $\dot{\gamma}(t) = a(\gamma(t))b(t)$ dann auch geschrieben werden als

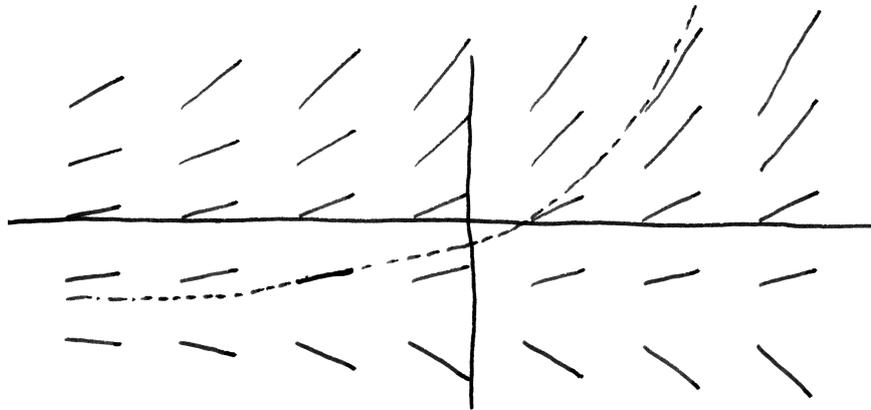
$$\frac{\dot{\gamma}(t)}{a(\gamma(t))} = b(t)$$

Ist nun G eine Stammfunktion von $1/a$ und B eine Stammfunktion von b , so folgt für alle $s, t \in I$ sofort

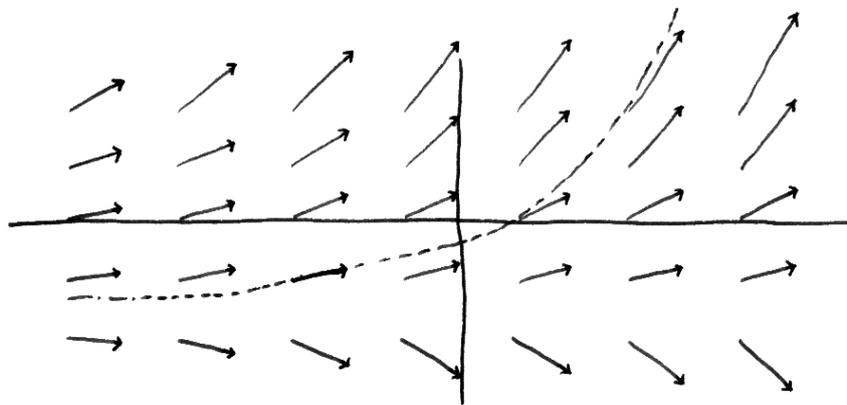
$$G(\gamma(t)) - G(\gamma(s)) = \int_{\gamma(s)}^{\gamma(t)} \frac{dx}{a(x)} = \int_s^t \frac{\dot{\gamma}(\tau)}{a(\gamma(\tau))} d\tau = \int_s^t b(\tau) d\tau = B(t) - B(s)$$

Hier ist G sicher streng monoton. Folglich hat es offenes Bild $G(V) \subseteq \mathbb{R}$, und bilden wir die Umkehrabbildung $G^{-1} : G(V) \rightarrow \mathbb{R}$, so folgt

$$\gamma(t) = G^{-1}(B(t) + c) \quad \forall t \in I$$



Im Fall zeitabhängiger eindimensionaler Felder mag man sich ein Vektorfeld als ein “Steigungsfeld” veranschaulichen, bei dem die jeweils vorgeschriebene Steigung $a(t, x)$ von beiden Koordinaten abhängen darf, und die Lösung der entsprechenden Differentialgleichung durch ihren im Bild gestrichelt eingezeichneten Graphen, der dann eben an jeder Stelle tangential an das dort vorgegebene Steigungsfeld sein soll.



In 8.1.19 habe ich ausgeführt, wie die Untersuchung der Integralkurven zeitabhängiger Felder auf einem Raum X auf die Untersuchung der Integralkurven zeitunabhängiger Felder auf dem Raum $\mathbb{R} \times X$ zurückgeführt werden kann. Dieses Bild zeigt im Spezialfall $X = \mathbb{R}$ das zeitunabhängige Vektorfeld auf \mathbb{R}^2 , dessen Integralkurven den Integralkurven des durch das Bild darüber dargestellten zeitabhängigen Feldes auf \mathbb{R} entsprechen.

mit der Konstante $c = G(\gamma(s)) - B(s)$. Umgekehrt prüft man auch ohne Schwierigkeiten, daß für $(s, v) \in W \times V$ die obige Formel für $c = G(v) - B(s)$ und $t \in I + B^{-1}(G(V) - c)$ die größte Lösung unserer Differentialgleichung mit $\gamma(s) = v$ liefert. Um den Zusammenhang mit der Situation separierter Variablen im Sinne von 8.1.18 herzustellen, interpretieren wir unsere Gleichung wie in 8.1.19 als die Suche nach Integalkurven des ebenen Vektorfelds $(z, x) \mapsto (1, a(x)b(z))$ und kommen unter der zusätzlichen Annahme, daß auch b keine Nullstelle habe, mit der in 8.1.17 erläuterten Längenänderung um $c(z, x) = b(z)^{-1}$ zum Vektorfeld $(z, x) \mapsto (b(z)^{-1}, a(x))$. In dieser Weise landen wir dann bei der Suche nach den Integalkurven eines Vektorfelds mit separierten Variablen.

8.1.22. Eine Fülle an weiteren Beispielen und Lösungsmethoden findet man etwa in [?].

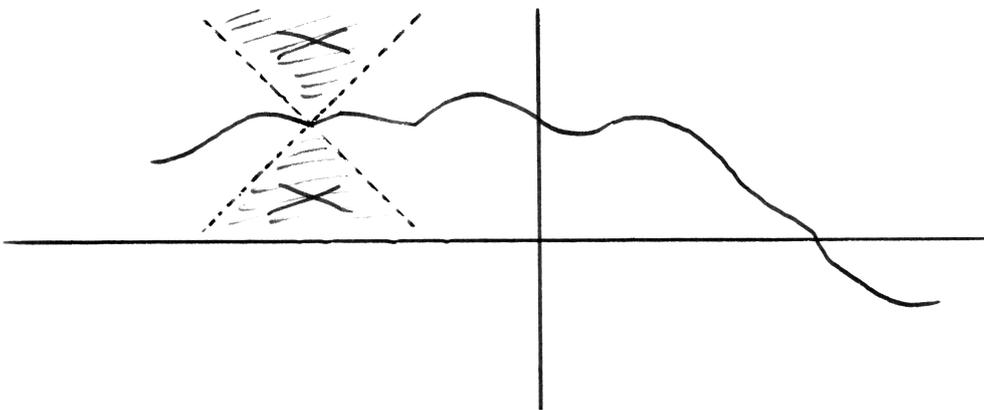
8.2 Existenz und Eindeutigkeit von Lösungen

8.2.1. Wir erinnern daran, daß nach 7.1.6 eine Abbildung f zwischen metrischen Räumen lipschitzstetig heißt genau dann, wenn es eine Konstante $L > 0$ gibt mit $d(f(x), f(y)) \leq Ld(x, y)$ für alle x, y im Ausgangsraum. Eine Abbildung zwischen metrischen Räumen heißt **lokal lipschitzstetig** genau dann, wenn jeder Punkt des Ausgangsraums eine Umgebung besitzt, auf der unsere Funktion lipschitzstetig ist.

8.2.2. Nach 4.3.5 ist jedes stetig differenzierbare Vektorfeld auf einer offenen Teilmenge eines normierten Raums lokal lipschitzstetig, deshalb folgt 1 aus der anschließenden etwas technischeren Version 1. Die Hauptlast des Beweises trägt jedoch das folgende Lemma 8.2.3.

Lemma 8.2.3 (Lokale Existenz und Eindeutigkeit). *Gegeben X ein endlich-dimensionaler reeller Raum, $U \subseteq X$ offen und $A : U \rightarrow \vec{X}$ ein beschränktes lipschitzstetiges Vektorfeld existieren zu jedem Anfangswert $p \in U$ Integalkurven von A mit offenem Definitionsbereich, und je zwei Integalkurven $\gamma : I \rightarrow U$ und $\phi : J \rightarrow U$ mit demselben Anfangswert stimmen für hinreichend kleines $\varepsilon > 0$ auf $I \cap J \cap [-\varepsilon, \varepsilon]$ überein.*

Ergänzung 8.2.4. Allgemeiner gilt das auch für stetige beschränkte zeitabhängige Vektorfelder $A : (-a, a) \times U \rightarrow \vec{X}$, die nur “partiell lipschitzstetig” sind in dem Sinne, daß es eine Konstante L gibt mit $\|A(t, x) - A(t, y)\| \leq L\|x - y\|$ für alle $t \in (-a, a)$ und $x, y \in U$. Der Beweis ist mutatis mutandis derselbe. Diese Variante ist insofern stärker, als das Lemma beim Übergang 8.1.19 von zeitabhängigen zu zeitunabhängigen Vektorfeldern dieselbe Folgerung nur liefert unter der stärkeren Annahme, daß $A : (-a, a) \times U \rightarrow \vec{X}$ nicht nur “partiell” sondern “auch in Bezug auf die erste Variable” lipschitzstetig ist.



Die Restriktion auf die negative x -Achse der hier durch ihren Graphen dargestellten Funktion ist lipschitzstetig mit Lipschitzkonstante 1, da sie an jeder Stelle den schraffierten verbotenen Bereich der entsprechen verschobenen Figur vermeidet. Die Begrenzungslinien haben darin als Steigung die Lipschitzkonstante, in diesem Fall die Steigung 1. Die Restriktion auf die positive x -Achse ist zwar lipschitzstetig, aber mit einer größeren Lipschitzkonstante.

Ergänzung 8.2.5. Das Lemma gilt auch für einen beliebigen vollständigen normierten reellen Raum X , nur benötigt in dieser Allgemeinheit unser Beweis die Integration von Funktionen mit Werten in Banachräumen 1.3.3 mitsamt der zugehörigen Variante des Hauptsatzes der Differential- und Integralrechnung 1.3.6.

Beweis. Wir betrachten für ein beliebiges halboffenes kompaktes reelles Intervall $K \subset \mathbb{R}$ mit $0 \in K$ den affinen Raum

$$\mathcal{C}_p(K, X)$$

aller stetigen Wege $\gamma : K \rightarrow X$ mit $\gamma(0) = p$ und versehen seinen Richtungsraum $\mathcal{C}_0(K, \vec{X})$ mit der Norm $\|\cdot\|_\infty$ der gleichmäßigen Konvergenz. Nach 10.5.30 erhalten wir so einen vollständigen normierten Vektorraum. Nun betrachten wir in unserem affinen Raum die offene Teilmenge $\mathcal{C}_p(K, U)$ aller in U verlaufenden Wege und die Abbildung

$$\begin{aligned} F : \mathbb{R} \times \mathcal{C}_p(K, U) &\rightarrow \mathbb{R} \times \mathcal{C}_p(K, X) \\ (\tau, \gamma) &\mapsto (\tau, \gamma - \tau \int (A \circ \gamma)) \end{aligned}$$

wobei $\int : \mathcal{C}(K, \vec{X}) \rightarrow \mathcal{C}_0(K, \vec{X})$ gegeben sei durch $(\int \psi)(t) = \int_0^t \psi(s) ds$ mit unserem vektorwertigen Integral aus 1.3.3. Bezeichne κ den konstanten Weg bei p . Unter unserer Abbildung geht aufgrund der vektorwertigen Variante 1.3.6 des Hauptsatzes der Differential- und Integralrechnung (τ, γ) nach (τ, κ) genau dann, wenn $\gamma : K \rightarrow U$ eine Integralkurve des reskalierten Feldes τA ist. Insbesondere haben wir $(0, \kappa) \mapsto (0, \kappa)$. Wir wenden nun den Umkehrsatz für stetige Funktionen 7.1.10 an und zeigen genauer, daß für $\eta > 0$ hinreichend klein und $K \subset [-1/4S, 1/4S]$ mit $S > 0$ einer oberen Schranke der Normen der Vektoren unseres Vektorfeldes A die Restriktion von $F - \text{id}$ auf $(-\eta, \eta) \times \mathcal{C}_p(K, U)$ kontrahierend ist. Dazu rechnen wir

$$\begin{aligned} \|(F - \text{id})(\sigma, \psi) - (F - \text{id})(\tau, \gamma)\| &= \left\| \sigma \int A\psi - \tau \int A\gamma \right\|_\infty \\ &\leq |\sigma - \tau| \left\| \int A\psi \right\|_\infty + |\tau| \left\| \int A\psi - A\gamma \right\|_\infty \\ &\leq |\sigma - \tau|(S/4S) + (\eta L/4S) \|\psi - \gamma\|_\infty \\ &\leq |\sigma - \tau|/4 + \|\psi - \gamma\|_\infty/4 \\ &\leq (1/2) \|(\tau - \sigma, \gamma - \psi)\| \end{aligned}$$

falls $\eta > 0$ so klein ist, daß gilt $\eta L/S < 1$. Dann liefert uns der Umkehrsatz für stetige Abbildungen 7.1.10 wegen $F : (0, \kappa) \mapsto (0, \kappa)$, daß es für $\tau > 0$ hinreichend klein genau ein Urbild (τ, γ_τ) von (τ, κ) unter F gibt, also genau eine

Integralkurve $\gamma_\tau : K \rightarrow U$ des reskalierten Vektorfelds τA . Gehen wir etwa von $K = [-\beta, \beta]$ aus, so ist $\gamma(t) := \gamma_\tau(\tau^{-1}t)$ eine auf $(-\tau\beta, \tau\beta)$ definierte Integralkurve des Vektorfelds A zu p und die Existenzaussage des Lemmas ist gezeigt. Seien andererseits $\gamma : I \rightarrow U$ und $\phi : J \rightarrow U$ Integralkurven mit demselben Anfangswert. Besteht $I \cap J$ nur aus dem Nullpunkt, so ist die Behauptung eh klar. Sonst gibt es $\alpha > 0$ mit $I \cap J \cap [-\alpha, \alpha]$ halboffen und kompakt und in $[-1/4S, 1/4S]$ enthalten, und für alle $\tau \in [0, 1]$ sind die Abbildungen $t \mapsto \gamma(\tau t)$ und $t \mapsto \phi(\tau t)$ auf $I \cap J \cap [-\alpha, \alpha]$ definierte Integralkurven zu p des reskalierten Vektorfelds τA . Für hinreichend kleines $\tau > 0$ gibt es aber nach dem, was wir gezeigt haben, nur eine derartige Integralkurve, und damit folgt auch die zweite Behauptung des Lemmas. \square

Ergänzung 8.2.6. Die allgemeinere Aussage 8.2.4 über lokale Existenz und Eindeutigkeit der Integralkurven für partiell lipschitzstetige zeitabhängige Vektorfelder folgt analog mithilfe der Abbildung

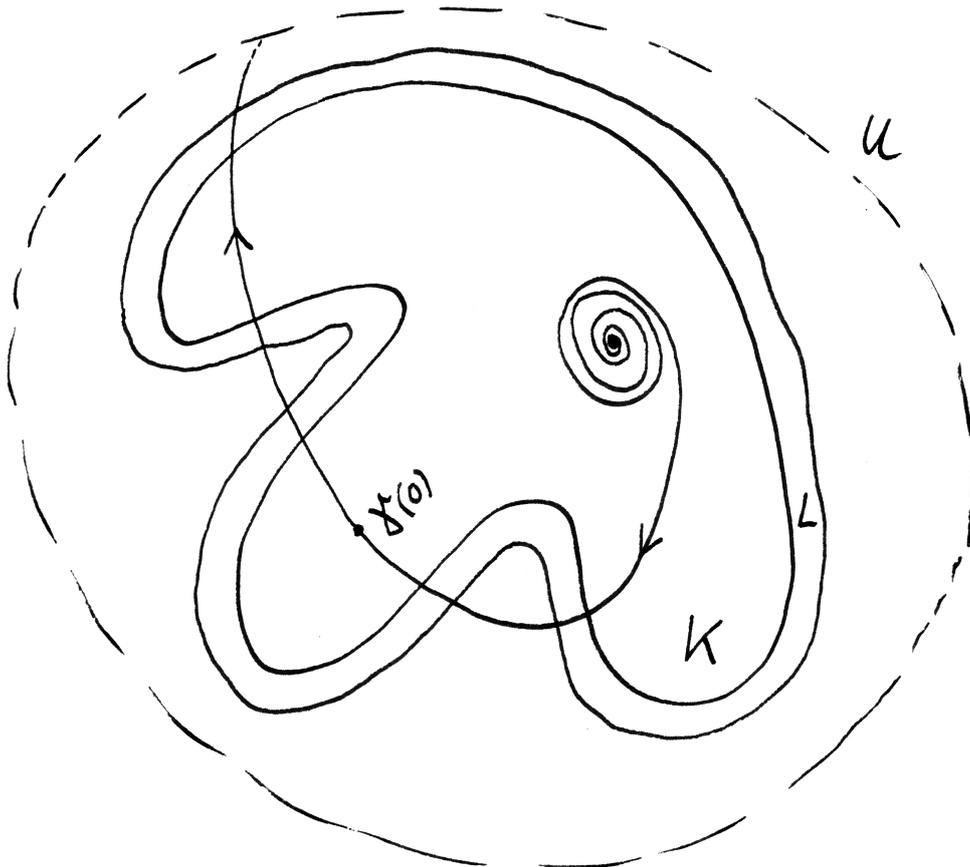
$$F : (\tau, \gamma) \mapsto \left(\tau, \gamma - \tau \int_0^t A(\tau s, \gamma(s)) ds \right)$$

Satz 8.2.7 (Globale Existenz und Eindeutigkeit). 1. *Gegeben ein lokal lipschitzstetiges Vektorfeld auf einer offenen Teilmenge eines endlichdimensionalen reellen Raums gibt es zu jedem Anfangswert genau eine größte Integralkurve.*

2. *Diese größte Integralkurve hat als Definitionsbereich ein offenes Intervall, und ist dieses Intervall nach oben beschränkt, so verläßt die fragliche Integralkurve für positive Zeiten jedes Kompaktum aus unserer offenen Teilmenge irgendwann einmal endgültig.*

8.2.8. Oft wird dieser Satz auch als Satz von **Picard-Lindelöf** zitiert. Beim Beweis zeigen wir darüber hinaus: Ist das Definitionsintervall unserer größten Integralkurve nach oben beschränkt, so kann sie nicht ab irgendeinem Zeitpunkt ganz innerhalb irgendeiner in unserem affinen Raum abgeschlossenen Teilmenge bleiben, die im Definitionsbereich unseres Vektorfelds enthalten ist und auf der unser Vektorfeld beschränkt ist.

Ergänzung 8.2.9. Allgemeiner gilt das auch für stetige zeitabhängige Vektorfelder $A : U \rightarrow \vec{X}$ auf $U \subseteq \mathbb{R} \times X$, die nur "lokal partiell lipschitzstetig" sind in dem Sinne, daß jeder Punkt von U eine offene Umgebung besitzt, in der sie im Sinn von 8.2.4 partiell lipschitzstetig sind. Der Beweis ist derselbe, man muß sich dafür nur auf 8.2.4 stützen. Des weiteren reicht es, X vollständig normiert anzunehmen: Wir müssen uns dann nur beim Beweis auf 8.2.5 stützen.



Eine maximale Integralkurve, deren Definitionsbereich nach oben beschränkt ist und die so jedes Kompaktum wie etwa K oder L irgendwann einmal engültig verläßt. In diesem Fall wäre der Definitionsbereich nach unten unbeschränkt und unsere Integralkurve würde für negative Zeiten gegen eine Nullstelle unseres Vektorfeldes konvergieren, die im Zentrum der Spirale liegt.

Beweis. Zunächst zeigen wir, daß je zwei Integralkurven γ, ψ mit demselben Anfangswert p und demselben Definitionsintervall I übereinstimmen. Wir zeigen nur, daß sie auf $I \cap [0, \infty)$ übereinstimmen, für $I \cap (-\infty, 0]$ argumentiert man analog. Stimmen aber unsere Wege auf $I \cap [0, \infty)$ nicht überein, so wäre das Supremum s über alle $t \in I$ mit $\gamma|_{[0, t]} = \psi|_{[0, t]}$ nicht das Supremum von I . Wegen der Stetigkeit der Integralkurven gälte $\gamma(s) = \psi(s)$, und nach der Eindeutigkeitsaussage in Lemma 8.2.3 muß dann auch gelten $\gamma|_{[0, t + \eta]} = \psi|_{[0, t + \eta]}$ für ein positives η , im Widerspruch zur Wahl von s . Folglich stimmen je zwei Integralkurven mit Anfangswert p auf dem Schnitt ihrer Definitionsbereiche überein und es gibt genau eine größte Integralkurve mit Anfangswert p , deren Definitionsbereich eben die Vereinigung der Definitionsbereiche aller Integralkurven zu p ist. Wäre dieser Definitionsbereich nicht offen, so enthielte er sein Supremum oder sein Infimum. Dann könnten wir jedoch um die Bilder dieser Grenzpunkte auch wieder Integralkurven mit offenem Definitionsbereich finden und “ankleben” und unsere Integralkurve wäre nicht maximal gewesen. Dieser Widerspruch zeigt, daß unsere größte Integralkurve offenen Definitionsbereich hat. Bezeichne schließlich A unser Vektorfeld und $U \subseteq X$ seinen Definitionsbereich. Ist $\gamma : [0, b) \rightarrow U$ eine Integralkurve von A , deren Bild in einem Kompaktum $M \subset U$ landet, so ist wegen $\dot{\gamma}(t) = A(\gamma(t))$ ihre Geschwindigkeit $\|\dot{\gamma}(t)\|$ beschränkt auf $[0, b)$, mithin ist γ Lipschitzstetig und besitzt nach 10.5.5 eine stetige Fortsetzung $\tilde{\gamma} : [0, b] \rightarrow M$. Die Integralform unserer Differentialgleichung zeigt dann sofort, daß auch $\tilde{\gamma}$ eine Integralkurve von A sein muß. Mithin kann eine Integralkurve mit nach oben beschränktem Definitionsbereich, die ganz in einem Kompaktum $M \subset U$ verläuft, schon einmal nicht maximal sein. Wir erklären nun noch, warum eine maximale Integralkurve mit nach oben beschränktem Definitionsbereich ab einem gewissen Zeitpunkt auch nicht mehr in ein vorgegebenes Kompaktum zurückkehren darf. Sicher besitzt unser Kompaktum M eine endliche Überdeckung durch offene Teilmengen von U , auf denen unser Vektorfeld jeweils Lipschitzstetig ist. Mit 9.7.16 finden wir auch ein $\varepsilon > 0$ derart, daß die Menge N aller Punkte von X mit Abstand $\leq \varepsilon$ zu einem Punkt von M in der Vereinigung der Mengen dieser endlichen Überdeckung enthalten ist. Da unser Vektorfeld auch auf N Lipschitzstetig ist, hat unsere Integralkurve dann an allen Stellen aus N , die sie durchläuft, eine gleichmäßig beschränkte Geschwindigkeit. Wann immer unsere maximale Integralkurve einen Punkt aus M durchläuft, muß sie also noch für eine gewisse von diesem Punkt unabhängige Zeitspanne innerhalb von N weiterlaufen. Sind wir näher als diese Zeitspanne am oberen Ende des Definitionsbereichs unserer maximalen Integralkurve, so kann unsere Integralkurve demnach keine Punkte aus M mehr durchlaufen, da sie ja wegen 8.2.8 nicht ab einem vorgegebenen Zeitpunkt in N bleiben darf, und es sonst nicht mehr schaffen könnte, N noch zu verlassen. \square

8.3 Lineare Differentialgleichungen

Satz 8.3.1 (Homogene lineare Differentialgleichungen). Gegeben ein halboffenes Intervall $I \subset \mathbb{R}$, ein endlichdimensionaler reeller Vektorraum V und eine stetig differenzierbare Abbildung $M : I \rightarrow \text{End } V$ bilden die differenzierbaren Abbildungen $\gamma : I \rightarrow V$ mit

$$\dot{\gamma}(t) = M(t)\gamma(t) \quad \forall t \in I$$

einen Untervektorraum $\mathcal{L} \subset \text{Ens}(I, V)$, den **Lösungsraum** unserer Differentialgleichung. Weiter ist für jedes $t_0 \in I$ das Auswerten bei t_0 ein Isomorphismus $\mathcal{L} \xrightarrow{\sim} V, \gamma \mapsto \gamma(t_0)$, der **Anfangswertisomorphismus**.

Ergänzung 8.3.2. Dieser Satz gilt mit fast demselben Beweis auch unter der schwächeren Voraussetzung M stetig, nur muß man sich dann beim Beweis auf die Existenz- und Eindeutigkeitsaussagen für Integralkurven lokal partiell lipschitzstetiger zeitabhängiger Vektorfelder stützen, die wir nur mit einer Beweisskizze als Ergänzung 8.2.9 erwähnt hatten. Der Satz gilt mit fast demselben Beweis auch im Fall eines beliebigen vollständigen normierten reellen Vektorraums V , wenn man statt dem Raum $\text{End } V$ aller Endomorphismen von V den normierten Raum $\mathcal{B}(V)$ aller stetigen Endomorphismen von V betrachtet. Wir müssen dann nur am Schluß etwas sorgfältiger argumentieren, etwa in dem Sinne, daß eine auf $[0, b)$ definierte Integralkurve nach den Abschätzungen im Beweis gleichmäßig stetig wäre und sich nach 10.5.5 stetig auf $[0, b]$ fortsetzen ließe. Das Bild dieser stetigen Fortsetzung ist dann das gesuchte Kompaktum, das nicht verlassen wird, im Widerspruch zu 8.2.9.

Beweis. Daß unser Lösungsraum $\mathcal{L} \subset \text{Ens}(I, V)$ ein Untervektorraum ist und das Auswerten bei t_0 linear, scheint mir beides offensichtlich. Es bleibt nur, Injektivität und Surjektivität des Auswertens zu zeigen. Ohne Beschränkung der Allgemeinheit dürfen wir dazu $t_0 = 0$ annehmen. Falls I nicht offen ist, wählen wir eine stetig differenzierbare Fortsetzung von M auf ein offenes Intervall $J \supset I$. Nun erfüllt $\gamma : I \rightarrow V$ nach 8.1.19 unsere Differentialgleichung genau dann, wenn $(\text{id}, \gamma) : I \rightarrow J \times V$ eine Integralkurve des Vektorfelds $(z, v) \mapsto (1, M(z)v)$ auf $J \times V$ ist. Dies Vektorfeld ist lokal lipschitzstetig, also besitzt es nach 1 zu jedem Anfangswert höchstens eine auf I definierte Integralkurve, und das zeigt die Injektivität. Für den Beweis der Surjektivität reicht es zu zeigen, daß jede maximale Integralkurve des Vektorfelds $(z, v) \mapsto (1, M(z)v)$ mit Anfangswert $(0, v_0)$ auf ganz J definiert ist. Sicher reicht es zu zeigen, daß sie bis zum oberen Ende von J definiert ist. Sonst gäbe es aber $b \in J$ derart, daß die Lösung nicht in positiver Richtung über $[0, b)$ hinaus fortgesetzt werden könnte, und mit der letzten Aussage von 1 folgte, daß $\gamma : [0, b) \rightarrow V$ unbeschränkt wäre. Es gibt jedoch L mit

$\|M(t)\| \leq L$ für alle $t \in [0, b]$, daraus folgt

$$\|\gamma(t)\| = \left\| v_0 + \int_0^t M(\tau)\gamma(\tau) \, d\tau \right\| \leq \|v_0\| + L \int_0^t \|\gamma(\tau)\| \, d\tau$$

und dann dem Lemma von Gronwall 8.3.3 sofort $\|\gamma(t)\| \leq \|v_0\| e^{Lt}$ und das wäre doch beschränkt, nämlich durch $\|v_0\| e^{Lb}$. \square

Lemma 8.3.3 (Gronwall). *Ist $b \in \mathbb{R}_{\geq 0}$ und $f : [0, b] \rightarrow [0, \infty)$ stetig und gibt es nichtnegative Konstanten L, C mit*

$$f(t) \leq L \int_0^t f(\tau) \, d\tau + C$$

für alle $t \geq 0$, so erfüllt f die Abschätzung $f(t) \leq C e^{Lt}$.

8.3.4. Es scheint mir von der Anschauung her ziemlich klar, daß eine differenzierbare Funktion F mit der Eigenschaft $F' \leq F$ höchstens exponentiell wachsen kann. Das Lemma von Gronwall präzisiert und verallgemeinert diese Intuition.

Beweis. Ohne Beschränkung der Allgemeinheit dürfen wir L und C positiv annehmen. Bezeichnet $F(t)$ den Wert des obigen Integrals, so folgern wir erst

$$\frac{F'(t)}{LF(t) + C} \leq 1$$

und dann durch Integrieren von 0 bis t mithilfe der Substitutionsregel weiter $L^{-1} \log(LF(t) + C) - L^{-1} \log C \leq t$ alias $\log(LF(t) + C) \leq Lt + \log C$ und durch Exponentieren und das Erinnern unserer Voraussetzungen schließlich

$$f(t) \leq LF(t) + C \leq C e^{Lt} \quad \square$$

Korollar 8.3.5 (Inhomogene lineare Differentialgleichungen). *Gegeben ein halboffenes Intervall $I \subset \mathbb{R}$, ein endlichdimensionaler reeller Vektorraum V sowie stetig differenzierbare Abbildungen $M : I \rightarrow \text{End } V$ und $f : I \rightarrow V$, bilden die differenzierbaren Abbildungen $\gamma : I \rightarrow V$ mit*

$$\dot{\gamma}(t) = M(t)\gamma(t) + f(t) \quad \forall t \in I$$

einen affinen Teilraum $\mathcal{L}_i \subset \text{Ens}(I, V)$ mit dem Lösungsraum der zugehörigen linearen Gleichung als Raum von Richtungsvektoren und für jedes $t_0 \in I$ definiert das Auswerten bei t_0 eine Bijektion $\mathcal{L}_i \xrightarrow{\sim} V$, $\gamma \mapsto \gamma(t_0)$, den **Anfangswertisomorphismus**.

Ergänzung 8.3.6. Dies Korollar gilt wieder mit fast demselben Beweis auch im Fall, daß $M : I \rightarrow \text{End } V$ und $f : I \rightarrow V$ nur stetig sind. Es gilt mit demselben Beweis auch im Fall eines beliebigen vollständigen normierten reellen Vektorraums V , wenn man statt dem Raum $\text{End } V$ aller Endomorphismen von V den normierten Raum $\mathcal{B}(V)$ aller stetigen Endomorphismen von V betrachtet.

Beweis. Die Differenz von je zwei Lösungen der inhomogenen Gleichung ist offensichtlich eine Lösung der zugehörigen homogenen Gleichung, und die Summe einer Lösungen der homogenen und einer Lösung der inhomogenen Gleichung ist offensichtlich eine Lösung der inhomogenen Gleichung. Damit bleibt nur zu zeigen, daß die inhomogene Gleichung überhaupt eine Lösung besitzt. Das folgt ähnlich wie im homogenen Fall und ohne weitere Schwierigkeiten aus unseren allgemeinen Prinzipien. Wir geben nun aber sogar eine Lösungsmethode an, die Methode der **Variation der Konstanten**. Dazu wählen wir eine Basis $\gamma_1, \dots, \gamma_n$ des Lösungsraums der homogenen Gleichung und fassen sie zusammen zu einer Lösung $X : I \rightarrow V^n = \text{Hom}(\mathbb{R}^n, V)$ der homogenen linearen Differentialgleichung

$$\dot{X}(t) = M(t)X(t)$$

für Funktionen $I \rightarrow \text{Hom}(\mathbb{R}^n, V)$. Da die Werte von $\gamma_1, \dots, \gamma_n$ an jeder Stelle eine Basis von V bilden, ist $X(t)$ an jeder Stelle ein Vektorraumisomorphismus. Nun machen wir für die Lösung unserer inhomogenen Gleichung den Ansatz $\gamma(t) = X(t)c(t)$ mit $c : I \rightarrow \mathbb{R}^n$ differenzierbar alias $\gamma(t) = c_1(t)\gamma_1(t) + \dots + c_n(t)\gamma_n(t)$ und finden

$$\begin{aligned} \dot{\gamma}(t) &= \dot{X}(t)c(t) + X(t)\dot{c}(t) \\ &= M(t)X(t)c(t) + X(t)\dot{c}(t) \\ &= M(t)\gamma(t) + X(t)\dot{c}(t) \end{aligned}$$

Unser Ansatz führt also zu einer Lösung der inhomogenen Gleichung genau dann, wenn gilt $X(t)\dot{c}(t) = f(t)$ alias $\dot{c}(t) = X^{-1}(t)f(t)$. Ein c mit dieser Eigenschaft existiert aber ganz offensichtlich, eben das Integral der rechten Seite. \square

8.3.7 (Verhalten benachbarter Integralkurven). Sei X ein endlichdimensionaler reeller Raum, $U \subseteq X$ eine offene Teilmenge und $A : U \rightarrow \vec{X}$ ein lipschitzstetiges Vektorfeld mit Lipschitz-Konstante L . Sind $\gamma_p, \gamma_q : [0, b] \rightarrow U$ Integralkurven zu Anfangswerten $p, q \in U$, so finden wir für alle $t \in [0, b]$ die Abschätzung

$$\begin{aligned} \|\gamma_p(t) - \gamma_q(t)\| &= \left\| p + \int_0^t A(\gamma_p(\tau)) \, d\tau - q - \int_0^t A(\gamma_q(\tau)) \, d\tau \right\| \\ &\leq \|p - q\| + L \int_0^t \|\gamma_p(\tau) - \gamma_q(\tau)\| \, d\tau \end{aligned}$$

und das Lemma von Gronwall 8.3.3 liefert für alle $t \in [0, b]$ die Abschätzung

$$\|\gamma_p(t) - \gamma_q(t)\| \leq \|p - q\| e^{Lt}$$

Salopp gesprochen besagt diese Abschätzung, daß zwei Integralkurven in einem Lipschitzstetigen Vektorfeld “höchstens exponentiell auseinanderlaufen können”. Man mag das Argument vom Schluß des Beweises des Satzes über homogene lineare Differentialgleichungen 8.3.1 auch etwas vergrößernd dahingehend zusammenfassen, daß sich in diesem Fall eine beliebige Lösung höchstens exponentiell von der Null-Lösung entfernt und folglich nicht in endlicher Zeit ins Unendliche entweichen kann.

8.4 Höhere Ableitungen ohne Koordinaten

8.4.1. Gegeben eine offene Teilmenge $U \subseteq \mathbb{R}^n$ heißt eine Abbildung $f : U \rightarrow \mathbb{R}^m$ **glatt** oder **beliebig differenzierbar** oder auch eine C^∞ -**Abbildung** genau dann, wenn zu allen Komponenten f_μ von f für $1 \leq \mu \leq m$ alle gemischten höheren partiellen Ableitungen, in der Multiindexschreibweise aus 5.2.3 also alle $\partial^\alpha f_\mu$ für beliebige $\alpha \in \mathbb{N}^n$, auf ganz U existieren. Existieren sie bis zum Totalgrad $|\alpha| \leq k$ und sind stetig, so spricht man von einer C^k -**Abbildung**. Wir wollen nun diese Bedingungen auf Abbildungen mit halboffenem Definitionsbereich auf beliebigen normierten reellen Räumen ausdehnen.

8.4.2. Gegeben Vektorräume V, W und $k \geq 0$ bilden wir dazu den Vektorraum $\text{Mult}^k(V, W)$ aller multilinearen Abbildungen des Produkts von k Kopien von V nach W . Im Fall $k = 0$ verstehen wir $\text{Mult}^0(V, W) = W$. Man bemerke die Isomorphismen $\text{Hom}(V, \text{Mult}^k(V, W)) \xrightarrow{\sim} \text{Mult}^{k+1}(V, W)$ durch $f \mapsto \langle f \rangle$ mit $\langle f \rangle(v_0, v_1, \dots, v_k) = (f(v_0))(v_1, \dots, v_k)$.

8.4.3. Die Notation $\langle f \rangle$ haben wir auch schon an anderer Stelle für den “von f erzeugten Untervektorraum” eingeführt. Was im Einzelfall gemeint ist, gilt es aus dem Kontext zu erschließen.

8.4.4. Gegeben normierte reelle Vektorräume V, W und $k \geq 0$ bilden wir ähnlich den normierten Vektorraum $\mathcal{B}^k(V, W)$ aller stetigen multilinearen Abbildungen des Produkts von k Kopien von V nach W , versehen mit der Norm 9.9.28. Im Fall $k = 0$ verstehen wir $\mathcal{B}^0(V, W) = W$. Man bemerke die Isomorphismen von normierten Vektorräumen $\mathcal{B}(V, \mathcal{B}^k(V, W)) \xrightarrow{\sim} \mathcal{B}^{k+1}(V, W)$ durch $f \mapsto \langle f \rangle$.

8.4.5. Sind X, Y normierte reelle Räume und ist $A \subset X$ eine halboffene Teilmenge und $g : A \rightarrow \mathcal{B}^k(\vec{X}, \vec{Y})$ eine differenzierbare Abbildung, so fassen wir dg mit der Identifikation aus 8.4.4 auf als diejenige Abbildung $dg : A \rightarrow \mathcal{B}^{k+1}(\vec{X}, \vec{Y})$, die gegeben wird durch $x \mapsto \langle d_x g \rangle$.

Definition 8.4.6. Gegeben X, Y normierte reelle Räume und $A \subset X$ eine halboffene Teilmenge und $f : A \rightarrow Y$ eine Abbildung setzen wir $d^0 f := f$ und $d^1 f := df : x \mapsto d_x f$ und definieren induktiv die **k -te Ableitung**

$$d^k f : A \rightarrow \mathcal{B}^k(\vec{X}, \vec{Y})$$

als $d^k f := d(d^{k-1} f)$, falls die $(k - 1)$ -te Ableitung existiert und differenzierbar ist auf A . Existieren alle höheren Ableitungen von f bis zur Ordnung k und sind stetig, so nennen wir f **von der Klasse \mathcal{C}^k** oder auch eine **\mathcal{C}^k -Abbildung**. Zum Beispiel bedeutet \mathcal{C}^1 nichts anderes als stetig differenzierbar und \mathcal{C}^0 nichts anderes als stetig. Ist f von der Klasse \mathcal{C}^k für alle k , so heißt die Abbildung f **glatt** oder **beliebig differenzierbar** oder **von der Klasse \mathcal{C}^∞** oder eine **\mathcal{C}^∞ -Abbildung**.

Übung 8.4.7. Die Exponentialabbildung $\exp : M(n \times n; \mathbb{C}) \rightarrow M(n \times n; \mathbb{C})$ ist glatt. Hinweis: [5.3.11](#).

Übung 8.4.8. Gegeben ein Banachraum V ist das Invertieren eine glatte Abbildung auf der Menge der invertierbaren Elemente von $\mathcal{B}(V)$. Hinweis: [4.4.11](#).

Übung 8.4.9. Seien X, Y normierte Räume und $A \subset X$ halboffen. Eine differenzierbare Abbildung $f : A \rightarrow Y$ ist von der Klasse \mathcal{C}^k genau dann, wenn die Abbildung $A \times \vec{X} \rightarrow Y \times \vec{Y}$, $(x, v) \mapsto (f(x), (d_x f)(v))$ von der Klasse \mathcal{C}^{k-1} ist.

Übung 8.4.10. Jede Verknüpfung von \mathcal{C}^k -Abbildungen ist von der Klasse \mathcal{C}^k . Jede Verknüpfung von glatten Abbildungen ist glatt. Hinweis: [4.5.10](#) und [8.4.9](#).

Übung 8.4.11. Eine Abbildung in ein Produkt von endlich vielen normierten reellen Vektorräumen ist \mathcal{C}^k genau dann, wenn ihre Komponenten \mathcal{C}^k sind.

8.4.12. Die Vertauschbarkeit der partiellen Ableitungen läßt sich koordinatenfrei dahingehend formulieren, daß unsere höheren Ableitungen $d^k f$ in den symmetrischen multilinearen Abbildungen landen, d.h. in denjenigen multilinearen Abbildungen, die ihren Wert nicht ändern, wenn man ihre Argumente untereinander vertauscht. Wir gehen darauf nicht näher ein, empfehlen aber dem Leser zu prüfen, daß sich die höheren Terme der Taylorentwicklung [5.2.5](#) koordinatenfrei in der Form $(k!)^{-1}(d^k f)(h, \dots, h)$ darstellen lassen.

Definition 8.4.13. Seien X, Y normierte reelle Räume und $A \subset X$, $B \subset Y$ halboffene Teilmengen. Eine Abbildung $f : A \rightarrow B$ heißt ein **\mathcal{C}^k -Diffeomorphismus** genau dann, wenn f bijektiv ist und sowohl f als auch seine Umkehrung $f^{-1} : B \rightarrow A$ beide \mathcal{C}^k -Abbildungen sind. Sprechen wir ohne nähere Spezifizierung von einem **Diffeomorphismus**, so meinen wir einen \mathcal{C}^∞ -Diffeomorphismus.

8.4.14. Seien X und Y vollständige normierte reelle Räume. Ist $U \subseteq X$ offen und $f : U \rightarrow Y$ eine \mathcal{C}^k -Abbildung für $1 \leq k \leq \infty$ und ist an einer Stelle $p \in U$ das Differential ein Isomorphismus mit stetigem Inversen, so induziert f

einen C^k -Diffeomorphismus von einer offenen Umgebung von p mit einer offenen Umgebung von $f(p)$. Das folgt sofort aus 7.1.4 mit den vorhergehenden Übungen, insbesondere 8.4.8.

8.5 Lösungen als Funktionen ihres Anfangswerts

Definition 8.5.1. Gegeben ein Vektorfeld auf einer offenen Teilmenge U eines normierten reellen Raums X , das zu jedem Anfangswert $q \in U$ eine größte Integralkurve $\gamma_q : I_q \rightarrow U$ besitzt, erklären wir seinen **Fluß** als die Abbildung

$$\Phi : (t, q) \mapsto \gamma_q(t)$$

von der Menge $\tilde{U} = \{(t, q) \in \mathbb{R} \times U \mid t \in I_q\}$, dem **Definitionsbereich des Flusses**, in den Definitionsbereich U unseres Vektorfelds.

Satz 8.5.2 (Lösungen als Funktionen ihres Anfangswerts). *Gegeben ein glattes Vektorfeld auf einer offenen Teilmenge eines endlichdimensionalen reellen Raums hat sein Fluß offenen Definitionsbereich und ist ebenfalls glatt.*

Ergänzung 8.5.3. Auch dieser Satz gilt mit fast demselben Beweis allgemeiner für einen vollständigen normierten und nicht notwendig endlichdimensionalen Raum. Die zusätzlich benötigten Zutaten sind wie in 8.2.5.

Beweis. Gegeben ein endlichdimensionaler reeller Raum X , eine offene Teilmenge $U \subseteq X$, ein C^k -Vektorfeld $A : U \rightarrow \vec{X}$ mit $k \geq 1$ und ein Punkt $p \in U$ wählen wir zunächst offene Umgebungen $V \subseteq U$ von p und $W \subseteq \vec{X}$ von Null mit $V + W \subset U$. Weiter wählen wir eine Norm auf \vec{X} . Dann betrachten wir für ein halboffenes kompaktes reelles Intervall $I \subset \mathbb{R}$ mit $0 \in I$ den affinen Raum

$$\mathcal{C}_p^1(I, X)$$

aller stetig differenzierbaren Wege $\gamma : I \rightarrow X$ mit $\gamma(0) = p$ und versehen seinen Richtungsraum $\mathcal{C}_0^1(I, \vec{X})$ mit der Norm $\|\varphi\|_\infty + \|\varphi'\|_\infty$ der gleichmäßigen Konvergenz von Funktion und erster Ableitung. Nach 10.5.31 erhalten wir so einen vollständigen normierten Vektorraum. Nun betrachten wir die Abbildung

$$\begin{aligned} F : \mathbb{R} \times V \times \mathcal{C}_0^1(I, W) &\rightarrow \mathcal{C}(I, \vec{X}) \\ (\tau, q, \psi) &\mapsto \psi' - \tau(A \circ (q + \psi)) \end{aligned}$$

Genau dann wird (τ, q, ψ) auf Null abgebildet, wenn $t \mapsto \gamma(t) = q + \psi(t)$ eine Integralkurve des reskalierten Vektorfelds τA zum Anfangswert $\gamma(0) = q$ ist. Nach Summenregel 4.4.4, Produktregel 4.4.5 und dem im Anschluß bewiesenen Lemma 8.5.4 ist F differenzierbar mit Differential

$$(d_{(\tau, q, \psi)} F)(h, v, \alpha) = \alpha' - h(A \circ (q + \psi)) - \tau((dA) \circ (q + \psi, v + \alpha))$$

Insbesondere gilt $(d_{(0,p,0)}F)(0,0,\alpha) = \alpha'$, und da $\alpha \mapsto \alpha'$ eine stetige und stetig umkehrbare Bijektion $\mathcal{C}_0^1(I, \vec{X}) \xrightarrow{\sim} \mathcal{C}(I, \vec{X})$ definiert und F nach unserer Formel stetig differenzierbar ist, dürfen wir den Satz über implizite Funktionen 7.2.8 anwenden. Er liefert uns ein Paar (A_1, B_1) mit $(0,p) \in A_1 \subseteq \mathbb{R} \times V$ und $0 \in B_1 \subseteq \mathcal{C}_0^1(I, W)$ derart, daß es für jedes $(\tau, q) \in A_1$ genau ein $\psi_{\tau,q} \in B_1$ gibt, für das $\gamma_{\tau,q} = q + \psi_{\tau,q}$ eine auf I definierte Integralkurve des reskalierten Vektorfelds τA mit Anfangswert q ist. Wählen wir also etwa $I = [-1, 1]$, so finden wir in A_1 eine offene Umgebung von $(0,p)$ der Gestalt $(-\eta, \eta) \times D$, und daselbst ist dann auch der Fluß definiert. Da Integralkurven unter Zeitverschiebung Integralkurven bleiben, zeigt das schon mal, daß unser Fluß einen offenen Definitionsbereich hat. Weiter ist F nach obiger Formel für sein Differential sogar von der Klasse \mathcal{C}^k im Sinne von 8.4.6. Damit zeigt der Satz über implizite Funktionen mit den Resultaten und Definitionen von 8.4 aber auch, daß die Zuordnung $(\tau, q) \mapsto \gamma_{\tau,q}$ eine \mathcal{C}^k -Abbildung $(-\eta, \eta) \times D \rightarrow \mathcal{C}^1(I, U)$ ist. Verknüpfen wir diese mit dem Auswerten an einer festen Stelle $t \in I \setminus \{0\}$, einer stetigen affinen Abbildung, und beachten $\gamma_{\tau,q}(t) = \gamma_q(\tau t)$, so folgt, daß der Fluß selbst eine \mathcal{C}^k -Abbildung ist. \square

Lemma 8.5.4. *Seien X, Y normierte Räume, $U \subseteq X$ eine offene Teilmenge und $A : U \rightarrow Y$ stetig differenzierbar. Für jedes Kompaktum K ist dann auch die Abbildung $(A \circ) : \mathcal{C}(K, U) \rightarrow \mathcal{C}(K, Y)$ differenzierbar und ihr Differential paßt in ein kommutatives Diagramm*

$$\begin{array}{ccc} \mathcal{C}(K, U) \times \mathcal{C}(K, \vec{X}) & \xrightarrow{d(A \circ)} & \mathcal{C}(K, \vec{Y}) \\ \downarrow \wr & & \parallel \\ \mathcal{C}(K, U \times \vec{X}) & \xrightarrow{(dA) \circ} & \mathcal{C}(K, \vec{Y}) \end{array}$$

8.5.5. Hier verstehen wir für jeden normierten Raum Z den Abbildungsraum $\mathcal{C}(K, Z)$ mit seiner Norm der gleichmäßigen Konvergenz und identifizieren den Richtungsraum unseres Abbildungsraums in der offensichtlichen Weise mit $\mathcal{C}(K, \vec{Z})$. In der oberen Horizontalen meinen wir die Abbildung $(\gamma, \alpha) \mapsto (d_\gamma(A \circ))(\alpha)$ und in der unteren Horizontalen meint dA entsprechend die Abbildung $dA : U \times \vec{X} \rightarrow \vec{X}$, $(x, v) \mapsto (d_x A)(v)$.

Beweis. Es reicht, an jeder Stelle $\gamma \in \mathcal{C}(K, U)$ die Differenzierbarkeit zu untersuchen und das Differential $d_\gamma(A \circ)$ zu bestimmen. Gegeben $\varepsilon > 0$ gibt es für alle $x \in \gamma(K)$ ein größtes $\eta(x) = \eta_\varepsilon(x) \in (0, 1)$ derart, daß gilt $B(x; \eta(x)) \subset U$ und

$$\|x - z\| < \eta(x) \Rightarrow \|d_x A - d_z A\| \leq \varepsilon$$

Man erkennt unschwer, daß $\eta : \gamma(K) \rightarrow (0, 1)$ stetig ist, ja sogar lipschitzstetig mit Lipschitz-Konstante Zwei. Sei $\delta = \delta_\varepsilon > 0$ das Minimum von η auf unserem

Kompaktum $\gamma(K)$. Für $x \in \gamma(K)$ und $h \in \vec{X}$ mit $\|h\| \leq \delta$ liefert dann der Schrankensatz oder vielmehr sein Korollar 4.3.5 die Abschätzung

$$\|A(x+h) - A(x) - (d_x A)(h)\| \leq \|h\|\varepsilon$$

Für jedes $\alpha : K \rightarrow \vec{X}$ mit $\|\alpha\| \leq \delta$ gilt also

$$\|A \circ (\gamma + \alpha) - A \circ \gamma - (dA) \circ (\gamma, \alpha)\| \leq \|\alpha\|\varepsilon$$

und das war im wesentlichen die Behauptung. \square

Satz 8.5.6 (Normalform eines Vektorfelds). *Gegeben ein glattes Vektorfeld auf einer offenen Teilmenge eines endlichdimensionalen reellen Raums, das an einer festen Stelle nicht verschwindet, finden wir für diese Stelle stets eine offene Umgebung mitsamt einem Diffeomorphismus besagter Umgebung auf eine offene Teilmenge eines endlichdimensionalen reellen Raums, unter dem unser Feld zu einem konstanten Feld verwandt ist.*

8.5.7. In Koordinaten gesprochen hat also jedes Vektorfeld, daß an einer vorgegebenen Stelle nicht verschwindet, in geeigneten lokalen Koordinaten x_1, \dots, x_n um diese Stelle die Gestalt $\frac{\partial}{\partial x_1}$.

Beweis. Sei X unser Raum, $U \subseteq X$ unsere offene Teilmenge, $A : U \rightarrow \vec{X}$ unser Vektorfeld und $p \in U$ die vorgegebene Stelle. Wir wählen $\vec{Y} \subset \vec{X}$ komplementär zur Geraden mit Richtungsvektor A_p und wählen einen Isomorphismus $L : \mathbb{R}^{n-1} \xrightarrow{\sim} \vec{Y}$. Gegeben ein glattes Vektorfeld A auf einer offenen Teilmenge U eines endlichdimensionalen affinen Raums schreiben wir im folgenden $A^t q$ für die Stelle $A^t q \in U$, an der der Punkt $q \in U$ landet, wenn er sich für die Zeitspanne t mit dem Fluß des Vektorfelds A treiben läßt. Man zeigt mühelos, daß für hinreichend kleines $\varepsilon > 0$ und eine hinreichend kleine Umgebung $W \subseteq \mathbb{R}^{n-1}$ des Ursprungs die Abbildung $(-\varepsilon, \varepsilon) \times W \rightarrow U, (t, \vec{w}) \mapsto A^t(p + L\vec{w})$ sinnvoll definiert und ein Diffeomorphismus der gewünschten Art ist. \square

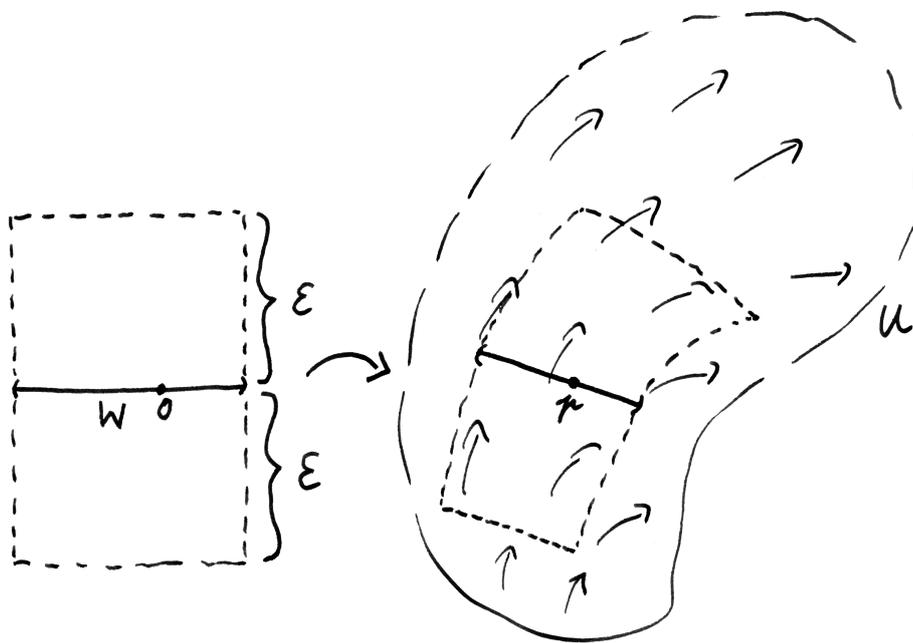


Illustration zum Beweis von Satz 8.5.6 zur lokalen Normalform eines Vektorfelds ohne Nullstelle

9 Maß und Integral

Es mag nahe liegen zu versuchen, jeder Teilmenge des \mathbb{R}^n ein “Volumen” oder “Maß” in $[0, \infty]$ so zuzuordnen, daß (1) das Verschieben von Mengen ihr Maß nicht ändert, daß (2) bei beliebigen disjunkten Vereinigungen das Maß der Vereinigung die Summe der Maße ist, und daß (3) dem Einheitswürfel $[0, 1]^n$ das Maß Eins zugeordnet wird. Kurzes Nachdenken zeigt jedoch, daß das unmöglich gelingen kann: Für solch ein Volumen müßte nämlich jeder Punkt Volumen Null haben, da ja unendlich viele verschiedene Punkte im Einheitswürfel liegen, und dann müßte auch der ganze Einheitswürfel Volumen Null haben als disjunkte Vereinigung einpunktiger Teilmengen. Um diesen Widerspruch zu vermeiden, mag man etwas schwächer statt (2) nur noch bei *abzählbaren* disjunkten Vereinigungen fordern wollen, daß das Maß der Vereinigung die Summe der Maße ist, aber auch solch einen Volumenbegriff kann es für beliebige Teilmengen von \mathbb{R}^n nicht geben, wie in 9.1.28 ausgeführt wird. Es ist jedoch möglich, gewisse Teilmengen des \mathbb{R}^n als “meßbar” auszuzeichnen derart, daß alle “einigermaßen vernünftigen” Teilmengen meßbar sind, und jeder dieser meßbaren Mengen ein Volumen so zuzuordnen, daß Bedingung (1), die abzählbare Variante von (2) sowie (3) entsprechend gelten. Im folgenden führen wir das aus und zeigen, wie davon ausgehend auch eine sehr allgemeine Integrationstheorie entwickelt werden kann, die sich sowohl in der weiteren Entwicklung der Analysis als auch bei der mathematischen Modellierung der Wahrscheinlichkeit als außerordentlich nützlich erweisen wird.

9.1 Maßräume und Maße

9.1.1. Gegeben eine Menge X erinnern wir daran, daß wir nach 2.1.11 die Menge aller ihrer Teilmengen bilden dürfen, und daß diese Menge die Potenzmenge $\mathcal{P}(X)$ von X hieß. Weiter erinnern wir daran, daß wir, gegeben eine Menge X , aus rein didaktischen Erwägungen heraus Teilmengen der Potenzmenge $\mathcal{P}(X)$ vorzugsweise als “Systeme von Teilmengen von X ” oder “Mengensysteme” ansprechen.

Definition 9.1.2. Ein System von Teilmengen $\mathcal{A} \subset \mathcal{P}(X)$ einer Menge X heißt eine **Mengenalgebra** genau dann, wenn gilt:

1. $\emptyset \in \mathcal{A}$;
2. $A, B \in \mathcal{A} \Rightarrow (A \cup B) \in \mathcal{A}$;
3. $A \in \mathcal{A} \Rightarrow (X \setminus A) \in \mathcal{A}$.

In Worten ist ein System von Teilmengen einer Menge X also eine Mengenalgebra genau dann, wenn es stabil ist unter dem Bilden von endlichen Vereinigungen und von Komplementen bezüglich X .

9.1.3. Ich will kurz diskutieren, warum unsere Definition in Formeln zu unserer Definition in Worten gleichbedeutend sind. Bei der Definition in Worten ist gemeint, daß eine Mengenalgebra die leere Menge enthalten soll als “die Vereinigung über überhaupt keine Teilmenge von X ”, vergleiche ???. Bei der Definition in Formeln folgt umgekehrt die Stabilität von \mathcal{A} unter endlichen Vereinigungen von mehr als zwei Mengen induktiv.

9.1.4. Eine Mengenalgebra ist natürlich auch stabil unter dem Bilden von endlichen Schnitten und von Differenzmengen.

Definition 9.1.5. Eine Mengenalgebra, die sogar stabil ist unter abzählbaren Vereinigungen, heißt eine σ -**Algebra**. Ein Paar (X, \mathcal{M}) bestehend aus einer Menge X und einer σ -Algebra $\mathcal{M} \subset \mathcal{P}(X)$ heißt ein **Meßraum**. Die Mengen aus \mathcal{M} heißen dann die **meßbaren** Mengen von (X, \mathcal{M}) oder kurz von X .

9.1.6. In Formeln ist eine σ -Algebra also eine Mengenalgebra $\mathcal{A} \subset \mathcal{P}(X)$ derart, daß aus $A_i \in \mathcal{A}$ für $i = 1, 2, \dots$ folgt $\bigcup_{i \in \mathbb{N}} A_i \in \mathcal{A}$. Gegeben ein Meßraum (X, \mathcal{M}) ist dann natürlich auch ganz X meßbar und abzählbare Schnitte meßbarer Mengen sind wieder meßbar.

Beispiel 9.1.7. In jeder Menge bilden die endlichen Teilmengen mitsamt ihren Komplementen eine Mengenalgebra, die jedoch nur dann eine σ -Algebra ist, wenn wir unsere Konstruktion in einer endlichen Menge durchführen. In jeder Menge bilden die abzählbaren Teilmengen mitsamt ihren Komplementen eine σ -Algebra.

Beispiel 9.1.8. Alle endlichen Vereinigungen von Intervallen bilden eine Mengenalgebra von Teilmengen von \mathbb{R} . Die abzählbaren Vereinigungen von Intervallen bilden *keine* σ -Algebra von Teilmengen von \mathbb{R} , da dieses Mengensystem nicht unter dem Bilden von Komplementen stabil ist: Zum Beispiel ist die Menge der rationalen Zahlen $\mathbb{Q} \subset \mathbb{R}$ eine abzählbare Vereinigung von Intervallen, genauer von einelementigen Intervallen, aber ihr Komplement $\mathbb{R} \setminus \mathbb{Q}$ ist keine abzählbare Vereinigung von Intervallen mehr.

Definition 9.1.9. Sei $X = (X, \mathcal{M})$ ein Meßraum. Ein **Maß** oder genauer ein **nichtnegatives Maß** auf X ist eine Abbildung $\mu : \mathcal{M} \rightarrow [0, \infty]$ derart, daß für jede abzählbare Familie von paarweise disjunkten meßbaren Mengen das Maß ihrer Vereinigung übereinstimmt mit der Summe der Maße der einzelnen Mengen. Wir sagen dann auch, die Abbildung μ sei σ -**additiv** und nennen das Tripel (X, \mathcal{M}, μ) einen **Maßraum**.

9.1.10. In Formeln fordern wir also für jede abzählbare Familie $(A_n)_{n \in \mathbb{N}}$ von paarweise disjunkten meßbaren Mengen die Gleichheit

$$\mu \left(\bigcup_{n \in \mathbb{N}} A_n \right) = \sum_{n \in \mathbb{N}} \mu(A_n)$$

Diese Gleichheit ist in $[0, \infty]$ zu verstehen. Über den Fall $N = \emptyset$ enthält unsere Forderung nach unseren Konventionen 3.1.17 insbesondere die Forderung $\mu(\emptyset) = 0$. Fordert man diese Bedingung separat, so braucht man die große Formel nur noch für Folgen meßbarer Mengen $(A_n)_{n \in \mathbb{N}}$ zu fordern.

Ergänzung 9.1.11. Ein Maßraum, bei dem die ganze Menge Maß Eins hat, heißt ein **Wahrscheinlichkeitsraum**. Mit diesem Wort geht allerdings eine völlig andere Motivation, Intuition und Buchstabenwahl einher: Das Ziel ist nun nicht mehr ein begrifflicher Rahmen zur Berechnung der Volumina einfacher Körper und dergleichen, also die explizite Berechnung von Maßen vorgegebener Mengen, sondern die mathematische Modellierung des Zufalls. Man notiert Wahrscheinlichkeitsräume statt (X, \mathcal{M}, μ) meist (Ω, \mathcal{A}, P) und denkt sich dabei Ω als eine völlig unstrukturierte und von der speziell untersuchten Fragestellung unabhängige Menge “von sich paarweise ausschließenden Möglichkeiten”, und $P(A)$ mit P wie “Probability” als die **Wahrscheinlichkeit** für das Eintreten einer Möglichkeit aus A . Das Interesse konzentriert sich in diesem Zusammenhang auf das Studium der Beziehungen zwischen sogenannten **Zufallsvariablen**. Ganz allgemein versteht man unter einer Zufallsvariablen eine Abbildung von Ω in einen weiteren Maßraum mit der Eigenschaft, daß die Urbilder aller meßbaren Mengen meßbar sind. So würde etwa ein gerechter Würfel modelliert durch eine Abbildung $W : \Omega \rightarrow \{1, \dots, 6\}$ mit $W^{-1}(i)$ meßbar und $P(W^{-1}(i)) = 1/6$ für $1 \leq i \leq 6$. Hier könnte Ω aus eben diesen sechs Elementen bestehen, es könnte jedoch auch viel größer sein und etwa aus allen möglichen Ausgängen eines einmaligen Würfels mit hundert Würfeln bestehen, von denen unser Würfel nur einer ist, oder Ω könnte aus allen Paaren in $\{1, \dots, 6\} \times \{\text{an}, \text{aus}\}$ bestehen, bei denen der zweite Eintrag erinnert, ob das Licht in dem Raum, in dem gewürfelt wurde, nun an oder aus war. Ganz allgemein interessieren in der Wahrscheinlichkeitstheorie im Wesentlichen nur diejenigen Konstruktionen und Definitionen, die veträglich sind mit dem Zurückholen unter **Verfeinerungen**, d.h. unter Abbildungen $\Omega' \rightarrow \Omega$ von einem weiteren Wahrscheinlichkeitsraum $(\Omega', \mathcal{A}', P')$ nach Ω mit der Eigenschaft, daß Urbilder meßbarer Mengen meßbar bleiben und dasselbe Maß behalten wie die ursprüngliche Menge. Die meßbaren Teilmengen von Ω heißen auch **Ereignisse**. Ist jede einelementige Teilmenge von Ω meßbar, so mag man die Elemente der Menge Ω auch **Elementar-Ereignisse** nennen. Man beachte jedoch, daß bei Verfeinerungen neue Ereignisse hinzukommen können und Elementar-Ereignisse im allgemeinen keine Elementar-Ereignisse bleiben werden.

Beispiele 9.1.12. Einfache Beispiele für Maße sind das **Dirac-Maß** δ_x an einem Punkt $x \in X$, gegeben durch

$$\delta_x(A) = \begin{cases} 1 & x \in A; \\ 0 & \text{sonst,} \end{cases}$$

und das **Zählmaß** $\zeta(A) = |A| \in \mathbb{N} \cup \{\infty\}$, das jeder Teilmenge die Zahl ihrer Elemente zuordnet. Diese beiden Maße sind sogar auf der gesamten Potenzmenge einer beliebigen Menge X definiert. Allgemeiner kann man für jede Menge X und jede Abbildung $f : X \rightarrow [0, \infty]$ wieder auf der gesamten Potenzmenge von X das Maß $A \mapsto \sum_{x \in A} f(x)$ betrachten, bei dem in gewisser Weise “jeder Punkt $x \in X$ mit dem Faktor $f(x)$ gewichtet wird”. Das vielleicht wichtigste Beispiel für ein Maß ist das “Lebesgue-Maß” auf den “topologisch meßbaren Mengen” oder “Borelmengen” des \mathbb{R}^n , dessen Konstruktion noch aussteht.

9.1.13. Sei X eine feste Menge. Sind $\mathcal{M}, \mathcal{N} \subset \mathcal{P}(X)$ zwei σ -Algebren, so ist auch ihr Schnitt $\mathcal{M} \cap \mathcal{N}$ eine σ -Algebra. Sogar ein beliebiger Schnitt von σ -Algebren in X ist wieder eine σ -Algebra in X .

Definition 9.1.14. Ist $\mathcal{A} \subset \mathcal{P}(X)$ irgendein System von Teilmengen einer Menge X , so betrachten wir den Schnitt aller σ -Algebren, die \mathcal{A} enthalten. Dieser Schnitt ist sicher die kleinste σ -Algebra auf X , die \mathcal{A} enthält. Er heißt die **von \mathcal{A} erzeugte σ -Algebra** und wird $\sigma(\mathcal{A})$ notiert.

9.1.15. Ich rate davon ab, nach einer expliziteren Beschreibung für die von einem Mengensystem erzeugte σ -Algebra zu suchen, und ich rate erst recht davon ab, dafür eine Anschauung entwickeln zu wollen. Es handelt sich hierbei vielmehr um ein abstraktes logisches Konstrukt, das nur dazu dient, einen begrifflichen Rahmen für unsere weiteren Überlegungen zu liefern. Um Ihnen die Schwierigkeiten einer “expliziten” Beschreibung zu zeigen, will ich kurz andeuten, wie es *nicht* geht. Man könnte versucht sein, unser Mengensystem zunächst einmal zu ergänzen durch Hinzunehmen aller abzählbaren Vereinigungen. Dann durch Hinzunehmen aller Komplementmengen. Dann wieder durch Hinzunehmen aller abzählbaren Vereinigungen, und immer so weiter. Kriegt man so in endlich vielen Schritten jede Menge der von unserem Mengensystem erzeugten σ -Algebra? Eben nicht: Denn nun muß man auch noch die Vereinigungsmengen aller Mengenfolgen dazunehmen, bei denen die erste Menge nach einem Schritt erhalten wurde, die zweite nach zwei Schritten und so weiter. Um die ganze σ -Algebra zu beschreiben, muß man stattdessen mit transfiniten Induktion arbeiten, wie etwa im Beweis von Lemma ?? ausgeführt wird.

Definition 9.1.16. Die von den offenen Mengen eines metrischen oder allgemeiner eines topologischen Raums X erzeugte σ -Algebra nennen wir die σ -Algebra der **topologisch meßbaren Teilmengen** oder auch der **Borelmengen** von X und notieren sie $\text{Borel}(X)$. Unter einem **topologischen Maß** verstehen wir ein Maß auf der σ -Algebra aller topologisch meßbaren Mengen eines topologischen Raums; unter einem **Borelmaß** ein topologisches Maß, das auf allen Kompakta endliche Werte annimmt. Das Zählmaß auf \mathbb{R} etwa ist in unserem Sinne ein topologisches

Maß, aber kein Borelmaß. Für den Begriff einer Borelmenge und eines Borelmaßes sind jedoch leider auch viele andere Bedeutungen in der Literatur verbreitet.

Ergänzung 9.1.17. Auch ohne die Kontinuumshypothese vorauszusetzen kann man zeigen, daß jede überabzählbare Borelmenge bereits in Bijektion zu ganz \mathbb{R} ist. Wie das genau geht, können Sie etwa in der Mengenlehre lernen.

Übung 9.1.18. Man zeige, daß die offenen Intervalle mit rationalen Endpunkten die σ -Algebra aller Borelmengen der reellen Zahlengeraden erzeugen.

9.1.19. Natürlich sind mit unserer Definition auch alle abgeschlossenen Mengen topologisch meßbar, und für jede Borelmenge $A \subset \mathbb{R}^n$ und beliebiges $a \in \mathbb{R}^n$ ist auch die verschobene Menge $a + A$ eine Borelmenge. Unser nächstes Ziel sind die beiden folgenden Sätze.

Satz 9.1.20 (Charakterisierung des Lebesguemaßes). *Es gibt genau ein topologisches Maß λ auf dem \mathbb{R}^n , das dem Einheitswürfel das Maß Eins zuordnet, in Formeln $\lambda([0, 1]^n) = 1$, und das außerdem **translationsinvariant** ist, d.h. die Eigenschaft hat, daß für beliebiges $a \in \mathbb{R}^n$ und jede Borelmenge $A \subset \mathbb{R}^n$ gilt*

$$\lambda(a + A) = \lambda(A)$$

9.1.21. Dieses Maß λ heißt das **Lebesgue-Maß** auf dem \mathbb{R}^n . Wenn wir zum Ausdruck bringen wollen, welches n gemeint ist, notieren wir es auch λ^n . Die erste Bedingung an unser Lebesguemaß nennen wir seine **Normierung**. In dieser Terminologie können wir also das Lebesgue-Maß charakterisieren als das eindeutig bestimmte normierte translationsinvariante Maß auf den Borelmengen des \mathbb{R}^n . Anschaulich ordnet λ jeder Borelmenge $A \subset \mathbb{R}^n$ ihr **Volumen** oder **Maß** $\lambda(A) \in [0, \infty]$ zu. Der Nachweis der Eindeutigkeit wird dem Leser als Übung 9.2.25 und 9.2.26 überlassen. Die Existenz zeigen wir für $n = 1$ im folgenden Abschnitt in Bemerkung 9.2.14 und für beliebiges n in 9.6.4.

9.1.22. Viele Autoren verstehen unter dem Lebesguemaß auch erst die “Vervollständigung” des hier beschriebenen Maßes im Sinne von 9.2.32. Präziser sollte man das im Satz 9.1.20 beschriebene Maß also vielleicht das “Lebesguemaß auf den Borelmengen” oder das **Lebesgue-Borel-Maß** nennen.

Satz 9.1.23 (Regularität des Lebesguemaßes). *Für das Lebesguemaß λ auf dem \mathbb{R}^n und jede Borelmenge $A \subset \mathbb{R}^n$ gilt*

$$\lambda(A) = \inf_{\substack{U \supset A \\ U \text{ offen in } \mathbb{R}^n}} \lambda(U) = \sup_{\substack{K \subset A \\ K \text{ kompakt}}} \lambda(K)$$

9.1.24. Dieser Satz wird in 9.7.1 gezeigt. Er deutet eine mögliche Konstruktion des Lebesgue-Maßes λ auf dem \mathbb{R}^n an: Um das Lebesgue-Maß einer Borelmenge

$A \subset \mathbb{R}^n$ zu bestimmen können wir beginnen mit dem Fall endlicher disjunkter Vereinigungen von Produkten von Intervallen. Solche "Quadermengen" haben noch ein anschauliches Volumen. Dann wird für U offen der Wert $\lambda(U)$ definiert als das Supremum über die Volumina aller in U enthaltenen Quadermengen, und schließlich erhält man das Maß $\lambda(A)$ einer beliebigen Borelmenge $A \subset \mathbb{R}^n$ als Infimum von $\lambda(U)$ über alle offenen U , die A umfassen. Diese Beschreibung von $\lambda(A)$ ähnelt unserer definitiven Konstruktion des Lebesguemaßes. Die wesentliche Schwierigkeit ist, zu zeigen, daß die so konstruierte Abbildung von den Borelmengen in die um ∞ erweiterten nichtnegativen reellen Zahlen σ -additiv ist.

9.1.25. Wir wollen uns zur besseren Motivation sofort überlegen, daß es schon im Fall $n = 1$ keinen vernünftigen Volumenbegriff für beliebige Teilmengen des \mathbb{R}^n geben kann. Wir beginnen mit dem Nachweis einiger Eigenschaften von Maßen, die wir auch an anderer Stelle noch oft benötigen werden.

9.1.26. Für jedes Maß und meßbare A, B gilt $A \subset B \Rightarrow \mu(A) \leq \mu(B)$, es gilt ja sogar genauer $\mu(B) = \mu(A) + \mu(B \setminus A)$. Ist weiter $A = \bigcup_{n=0}^{\infty} A_n$ eine abzählbare aufsteigende Vereinigung meßbarer Mengen, also $A_n \subset A_{n+1} \forall n \in \mathbb{N}$, so gilt

$$\mu(A) = \lim_{n \rightarrow \infty} \mu(A_n)$$

In der Tat, schreiben wir A als die disjunkte Vereinigung der $B_n = A_n \setminus A_{n-1}$, so haben wir $\mu(A_n) = \mu(B_n) + \dots + \mu(B_0)$ und die Behauptung folgt aus der Definition. Ist schließlich $A = \bigcup_{n=0}^{\infty} A_n$ ein beliebige abzählbare Vereinigung meßbarer Mengen, so können wir A auch als disjunkte Vereinigung der kleineren Mengen $B_n = A_n \setminus (A_{n-1} \cup \dots \cup A_0)$ schreiben und erhalten so die Abschätzung

$$\mu(A) \leq \sum_{n=0}^{\infty} \mu(A_n)$$

Übung 9.1.27. Ist $A = \bigcap_{n=0}^{\infty} A_n$ ein abzählbarer absteigender Schnitt meßbarer Mengen, also $A_n \supset A_{n+1} \forall n \in \mathbb{N}$, so folgt $\mu(A) = \lim_{n \rightarrow \infty} \mu(A_n)$ erst unter der zusätzlichen Voraussetzung $\mu(A_n) < \infty$ für mindestens ein n .

Lemma 9.1.28. *Es gibt kein Maß $\lambda : \mathcal{P}(\mathbb{R}) \rightarrow [0, \infty]$ auf der σ -Algebra aller Teilmengen von \mathbb{R} , das translationsinvariant und normiert ist.*

Beweis. Wir werden eine Teilmenge $A \subset \mathbb{R}$ und Folgen r_n und q_n reeller Zahlen konstruieren derart, daß gilt:

1. Die Mengen $r_n + A$ sind paarweise disjunkt und alle in $[0, 3]$ enthalten.
2. Die Mengen $q_n + A$ überdecken \mathbb{R} .

Für unsere Menge A müßte also gleichzeitig gelten

$$\begin{aligned}\sum_{n=0}^{\infty} \lambda(A) &= \sum_{n=0}^{\infty} \lambda(r_n + A) \leq \lambda([0, 3]) \leq 3\lambda([0, 1]) = 3 \\ \sum_{n=0}^{\infty} \lambda(A) &= \sum_{n=0}^{\infty} \lambda(q_n + A) \geq \lambda(\mathbb{R}) \geq \sum_{n=0}^{\infty} \lambda(2n + [0, 1]) = \infty\end{aligned}$$

und dieser Widerspruch zeigt dann das Lemma. Um unsere Menge A zu konstruieren, wählen wir mithilfe von ?? eine Teilmenge $I \subset \mathbb{R}$ derart, daß $\{1\} \cup I$ eine \mathbb{Q} -Basis von \mathbb{R} ist, und betrachten den von I erzeugten \mathbb{Q} -Untervektorraum $\langle I \rangle_{\mathbb{Q}} \subset \mathbb{R}$ und die Menge $A = \langle I \rangle_{\mathbb{Q}} \cap [0, 2]$. Für jede Folge r_n von paarweise verschiedenen rationalen Zahlen aus $[0, 1]$ sind dann die $r_n + A$ paarweise disjunkt und in $[0, 3]$ enthalten. Andererseits finden wir auch für alle $n \in \mathbb{Z}$ ein $b_n \in \langle I \rangle_{\mathbb{Q}} \cap [n - 1, n]$, es folgt $\langle I \rangle_{\mathbb{Q}} = \bigcup b_n + A$ und dann

$$\mathbb{R} = \bigcup_{q \in \mathbb{Q}, n \in \mathbb{Z}} q + b_n + A \quad \square$$

9.1.29. Es kommt sogar noch schlimmer: Nach Banach und Tarski ist es möglich, jedenfalls wenn man das Auswahlaxiom akzeptiert, die abgeschlossene Einheitskugel so in sechs paarweise disjunkte Teilmengen zu zerlegen, daß sich diese Teilmengen geeignet verschoben und im Raum gedreht ohne Überlappungen zu zwei Einheitskugeln zusammenfügen lassen. Das ist das sogenannte **Banach-Tarski-Paradoxon**. Die fraglichen sechs Teilmengen sind dann natürlich nicht alle meßbar.

Übung 9.1.30. Sei X eine Menge und $\mathcal{M} \subset \mathcal{P}(X)$ eine Mengenalgebra. Man zeige: (1) Genau dann ist \mathcal{M} eine σ -Algebra, wenn \mathcal{M} stabil ist unter abzählbaren *disjunkten* Vereinigungen. (2) Genau dann ist \mathcal{M} eine σ -Algebra, wenn \mathcal{M} stabil ist unter abzählbaren *aufsteigenden* Vereinigungen im Sinne von 9.1.26.

Übung 9.1.31. In einem metrischen Raum mit einer abzählbaren dichten Teilmenge erzeugen die offenen Bälle bereits die σ -Algebra der Borelmengen.

Übung 9.1.32. Konstruieren Sie in \mathbb{R} eine offene dichte Teilmenge von endlichem Lebesguemaß.

Ergänzende Übung 9.1.33. Sei X ein Meßraum und μ_i eine Folge von endlichen Maßen derart, daß für jedes meßbare $A \subset X$ die Folge der $\mu_i(A)$ monoton wachsend und beschränkt ist. So wird durch die Formel $\mu(A) = \lim_{i \rightarrow \infty} \mu_i(A)$ ein weiteres Maß auf X erklärt.

9.2 Konstruktion des Lebesguemaßes auf \mathbb{R}

Definition 9.2.1. Ein System von Teilmengen einer gegebenen Menge heißt ein **Mengenring** genau dann, wenn es stabil ist unter dem Bilden von endlichen Vereinigungen und von Differenzmengen. In Formeln ausgedrückt ist ein System von

Teilmengen $\mathcal{A} \subset \mathcal{P}(X)$ einer Menge X also ein Mengerring genau dann, wenn gilt:

1. $\emptyset \in \mathcal{A}$;
2. $A, B \in \mathcal{A} \Rightarrow (A \cup B) \in \mathcal{A}$;
3. $A, B \in \mathcal{A} \Rightarrow (B \setminus A) \in \mathcal{A}$.

9.2.2. Daß unsere Definition in Worten und unsere Definition in Formeln gleichbedeutend sind, erkennt man wie in 9.1.3. Ein Mengensystem $\mathcal{A} \subset \mathcal{P}(X)$ ist eine Mengenalgebra genau dann, wenn \mathcal{A} ein Mengerring ist und wenn zusätzlich gilt $X \in \mathcal{A}$.

Ergänzung 9.2.3. Identifiziert man in der offensichtlichen Weise $\mathcal{P}(X)$ mit der Menge $\text{Ens}(X, \mathbb{F}_2)$ aller Abbildungen von X in den zweielementigen Körper, so entsprechen unsere Mengerringe $\mathcal{A} \subset \mathcal{P}(X)$ genau den “nicht-unitären Teilringen” oder in unserer Terminologie Teilringen von $\text{Ens}(X, \mathbb{F}_2)$ unter der punktweisen Addition und Multiplikation und unsere Mengenalgebren den “unitären Teilringen”, die wir in unserer Terminologie ?? schlicht Teilringe nennen.

Übung 9.2.4. Gegeben Mengen X und Y sowie Mengerringe $\mathcal{A} \subset \mathcal{P}(X)$ und $\mathcal{B} \subset \mathcal{P}(Y)$ ist auch das System aller endlichen Vereinigungen von paarweise disjunkten Mengen der Gestalt $A \times B$ mit $A \in \mathcal{A}$ und $B \in \mathcal{B}$ ein Mengerring in $\mathcal{P}(X \times Y)$.

Definition 9.2.5. Ist X eine Menge und $\mathcal{A} \subset \mathcal{P}(X)$ ein Mengerring und $\mu : \mathcal{A} \rightarrow [0, \infty]$ eine Abbildung, die unsere Bedingung

$$\mu \left(\bigcup_{n \in \mathbb{N}} A_n \right) = \sum_{n \in \mathbb{N}} \mu(A_n)$$

für abzählbare Familien $(A_n)_{n \in \mathbb{N}}$ paarweise disjunkter Mengen aus \mathcal{A} erfüllt wann immer die Vereinigung der A_n wieder zu \mathcal{A} gehört, so nennen wir die Abbildung μ ein **Prämaß** auf dem Mengerring \mathcal{A} und sagen wieder, μ sei **σ -additiv**. Wieder ist unsere auch äquivalent zu den beiden Forderungen $\mu(\emptyset) = 0$ und $\mu \left(\bigcup_{n \in \mathbb{N}} A_n \right) = \sum_{n \in \mathbb{N}} \mu(A_n)$ für jede Folge $(A_n)_{n \in \mathbb{N}}$ paarweise disjunkter Mengen aus unserem Mengerring, bei der die Vereinigung wieder zu unserem Mengerring gehört.

Lemma 9.2.6. Auf dem Mengerring $\mathcal{I} \subset \mathcal{P}(\mathbb{R})$ aller endlichen Vereinigungen von beschränkten Intervallen gibt es genau ein Prämaß λ derart, daß für jedes nichtleere beschränkte Intervall $I \subset \mathbb{R}$ gilt

$$\lambda(I) = \sup I - \inf I$$

Beweis. Gegeben $A \in \mathcal{I}$ betrachten wir die bis auf Reihenfolge eindeutige Darstellung $A = J_1 \cup \dots \cup J_r$ von A als disjunkte Vereinigung der maximalen darin enthaltenen Intervalle und müssen setzen

$$\lambda(A) = \lambda(J_1) + \dots + \lambda(J_r)$$

Das zeigt die Eindeutigkeit. Es gilt damit nur noch zu zeigen, daß für $A = \bigsqcup_{n \in \mathbb{N}} A_n$ eine disjunkte Vereinigung mit $A, A_n \in \mathcal{I}$ gilt

$$\lambda(A) = \sum_{n \in \mathbb{N}} \lambda(A_n)$$

Offensichtlich gilt schon einmal $\lambda(B \cup C) = \lambda(B) + \lambda(C)$ für $B, C \in \mathcal{I}$ disjunkt. Wir setzen nun $B_n = A \setminus (A_0 \cup \dots \cup A_n)$. Natürlich sind dann auch die B_n endliche Vereinigungen beschränkter Intervalle, es gilt $B_0 \supset B_1 \supset \dots$ und $\bigcap_{n \in \mathbb{N}} B_n = \emptyset$, und es reicht, wenn wir zeigen

$$\lim_{n \rightarrow \infty} \lambda(B_n) = 0$$

Sei $\varepsilon > 0$ beliebig. Wir finden für jedes n eine kompakte Menge $C_n \subset B_n$, die endliche Vereinigungen kompakter Intervalle ist und für die gilt

$$\lambda(B_n \setminus C_n) \leq 2^{-n} \varepsilon$$

Jetzt betrachten wir $D_n = C_0 \cap \dots \cap C_n$. Auch die D_n sind endliche Vereinigungen kompakter Intervalle, es gilt $D_n \subset C_n \subset B_n$, und zusätzlich haben wir $D_0 \supset D_1 \supset D_2 \dots$. Wir zeigen nun $\lambda(B_n \setminus D_n) \leq 2\varepsilon$ für alle n . In der Tat gilt ja

$$B_n \setminus D_n = \bigcup_{k=0}^n B_n \setminus C_k \subset \bigcup_{k=0}^n B_k \setminus C_k$$

und folglich

$$\lambda(B_n \setminus D_n) \leq \sum_{k=0}^n \lambda(B_k \setminus C_k) \leq \sum_{k=0}^n 2^{-k} \varepsilon \leq 2\varepsilon$$

Nun folgt aber aus $\bigcap_{n \in \mathbb{N}} D_n = \emptyset$ und der Kompaktheit der D_n und 9.7.15 schon $D_N = \emptyset$ für ein N , und damit ergibt sich $\lambda(B_n) \leq 2\varepsilon$ für $n \geq N$. \square

Ergänzung 9.2.7. Ist allgemeiner $f : \mathbb{R} \rightarrow \mathbb{R}$ monoton wachsend und **linksseitig stetig**, d.h. $\lim_{y \nearrow x} f(y) = f(x)$ für alle $x \in \mathbb{R}$, so zeigt man ähnlich, daß es auf dem Mengenring aller endlichen Vereinigungen beschränkter Intervalle der Gestalt $[a, b)$ genau ein Prämaß df gibt mit $(df)([a, b)) = f(b) - f(a)$ für alle $a, b \in \mathbb{R}$ mit $a < b$. Man muß dann nur ein wenig feiner argumentieren, C_n aus dem besagten Mengenring wählen mit $\tilde{C}_n \subset B_n$, und $\bigcap_{n \in \mathbb{N}} D_n = \emptyset$ aus $\bigcap_{n \in \mathbb{N}} \tilde{C}_n = \emptyset$ folgern.

Ergänzung 9.2.8. Ist noch allgemeiner $f : \mathbb{R}^n \rightarrow \mathbb{R}$ monoton wachsend und linksseitig stetig in jeder Variablen, so zeigt man in derselben Weise, daß es auf dem Mengerring aller endlichen Vereinigungen beschränkter Quader der Gestalt $[a_1, b_1) \times \dots \times [a_n, b_n)$ genau ein Prämaß μ_f gibt derart, daß für $a_i, b_i \in \mathbb{R}$ mit $a_i < b_i$ der Wert $\mu_f([a_1, b_1) \times \dots \times [a_n, b_n))$ auf dem Quader die “alternierende Summe der Werte von f auf den Ecken” ist. Die Vorzeichen sind hierbei in der Weise zu wählen, daß für keine zwei durch eine Kante verbundenen Ecken dasselbe Vorzeichen gewählt wird und daß der Wert an der Ecke (b_1, \dots, b_n) mit positivem Vorzeichen eingeht.

Definition 9.2.9. Eine Teilmenge eines Raums mit Prämaß heißt **σ -endlich** genau dann, wenn sie sich durch eine Folge von Mengen endlichen Maßes aus dem entsprechenden Mengerring überdecken läßt. Ein Prämaß heißt **σ -endlich** genau dann, wenn der ganze Raum in diesem Sinne σ -endlich ist.

Satz 9.2.10 (Maßfortsetzungssatz von Caratheodory). Gegeben eine Menge X , ein Mengerring $\mathcal{A} \subset \mathcal{P}(X)$ und ein σ -endliches Prämaß $\mu : \mathcal{A} \rightarrow [0, \infty]$ existiert genau eine Fortsetzung von μ zu einem Maß auf der von \mathcal{A} erzeugten σ -Algebra $\sigma(\mathcal{A})$.

9.2.11. Sowohl die Existenz als auch die Eindeutigkeit von Maßfortsetzungen unter den gegebenen Voraussetzungen sind zentrale Aussagen der Maßtheorie.

Ergänzung 9.2.12. Ist unser Prämaß nicht σ -endlich, so besitzt es zwar stets die gleich in 9.2.13 konstruierte **kanonische Fortsetzung** zu einem Maß auf $\sigma(\mathcal{A})$, aber möglicherweise ist das nicht seine einzige Fortsetzung auf diese σ -Algebra. Als Beispiel betrachte man den Mengerring aller endlichen Teilmengen von \mathbb{R} und darauf das Prämaß, das jeder Menge die Null zuordnet. Die von unserem Mengerring erzeugte σ -Algebra besteht aus allen Teilmengen von \mathbb{R} , die entweder abzählbar sind oder abzählbares Komplement haben, und die möglichen Fortsetzungen unseres Prämaßes sind alle die Abbildungen, die allen abzählbaren Mengen die Null zuordnen und allen Komplementen abzählbarer Mengen ein beliebiges aber festes Element von $[0, \infty]$. Die kanonische Fortsetzung kann im übrigen auch charakterisiert werden als die eindeutig bestimmte Fortsetzung, die allen den Mengen von $\sigma(\mathcal{A})$ das Maß Unendlich zuordnet, die nicht in einer abzählbaren Vereinigung von Mengen endlichen Maßes aus \mathcal{A} enthalten sind, vergleiche ??.

Proposition 9.2.13 (Konstruktion von Maßfortsetzungen). Gegeben eine Menge X , ein Mengerring $\mathcal{A} \subset \mathcal{P}(X)$ und ein Prämaß $\mu : \mathcal{A} \rightarrow [0, \infty]$ erhält man eine Fortsetzung von μ zu einem Maß auf der von \mathcal{A} erzeugten σ -Algebra $\sigma(\mathcal{A})$, indem man für alle $M \in \sigma(\mathcal{A})$ setzt

$$\mu(M) = \inf \left\{ \sum_{n=0}^{\infty} \mu(A_n) \mid (A_n)_{n \in \mathbb{N}} \text{ ist eine Folge in } \mathcal{A} \text{ mit } M \subset \bigcup A_n \right\}$$

9.2.14. Ich erinnere hier an unsere Konvention, nach der das Infimum der leeren Menge als ∞ zu verstehen ist. Zusammen mit Lemma 9.2.6 folgt aus dem Satz über Maßfortsetzungen 9.2.10 die in 9.1.20 behauptete Existenz eines normierten translationsinvarianten Maßes auf den topologisch meßbaren Teilmengen der reellen Zahlengeraden. Den Nachweis der Eindeutigkeit in 9.1.20 überlassen wir dem Leser als Übung 9.2.25. Der Beweis des Maßfortsetzungssatzes wird nach einigen Vorbereitungen im Anschluß an den Beweis des Zerlegerlemmas 9.2.24 geführt werden.

Ergänzung 9.2.15. Zusammen mit 9.2.7 zeigt unser Satz über Maßfortsetzungen sogar, daß es für jede linksseitig stetige monoton wachsende Funktion $f : \mathbb{R} \rightarrow \mathbb{R}$ genau ein topologisches Maß df auf \mathbb{R} gibt mit $(df)([a, b)) = f(b) - f(a)$ für beliebige $a, b \in \mathbb{R}$ mit $a < b$. Unser Lebesguemaß kann man in dieser Notation auch schreiben als das Maß dx , mit x als alternativer Bezeichnung für die Identität $\text{id}_{\mathbb{R}} : \mathbb{R} \rightarrow \mathbb{R}, x \mapsto x$ auf \mathbb{R} .

Ergänzende Übung 9.2.16. Die Zuordnung $f \mapsto df$ aus 9.2.15 liefert eine eindeutige Entsprechung zwischen der Menge aller linksseitig stetigen monoton wachsenden Funktionen $f : \mathbb{R} \rightarrow \mathbb{R}$ mit $f(0) = 0$ und der Menge aller Borelmaße auf \mathbb{R} . Hinweis: Die Umkehrabbildung kann explizit angegeben werden als $\mu \mapsto f_{\mu}$ mit $f_{\mu}(x) = \mu[0, x)$ falls $x \geq 0$ und $f_{\mu}(x) = \mu[x, 0)$ falls $x \leq 0$.

Ergänzung 9.2.17. Die Zuordnung $f \mapsto \mu_f$ aus 9.2.8 zusammen mit dem Satz über Maßfortsetzungen liefert eine eindeutige Entsprechung zwischen der Menge aller in jeder Variablen monoton wachsenden und linksseitig stetigen Funktionen $f : \mathbb{R}^n \rightarrow \mathbb{R}$, die auf allen Koordinatenhyperebenen verschwinden, und der Menge aller Borelmaße auf \mathbb{R}^n .

Ergänzung 9.2.18. Im Fall von Wahrscheinlichkeitsmaßen μ auf \mathbb{R} mit seiner σ -Algebra der topologisch meßbaren Mengen ist es üblich, eine Variante der Konstruktion aus Übung 9.2.16 zu betrachten und die **Verteilungsfunktion** des Wahrscheinlichkeitsmaßes μ zu definieren durch die Vorschrift $V_{\mu}(x) = \mu(-\infty, x)$ und allgemeiner $V_{\mu}(x_1, \dots, x_n) = \mu((-\infty, x_1) \times \dots \times (-\infty, x_n))$ für Wahrscheinlichkeitsmaße auf \mathbb{R}^n . Damit erhalten wir dann eine eindeutige Entsprechung zwischen der Menge aller Wahrscheinlichkeitsmaße auf \mathbb{R} und der Menge aller linksseitig stetigen monoton wachsenden Funktionen $V : \mathbb{R} \rightarrow \mathbb{R}$ mit Infimum Null und Supremum Eins, in Formeln $\inf(V(\mathbb{R})) = 0$ und $\sup(V(\mathbb{R})) = 1$, oder allgemeiner zwischen der Menge aller Wahrscheinlichkeitsmaße auf \mathbb{R}^n und der Menge aller in jeder Variablen linksseitig stetigen monoton wachsenden Funktionen $V : \mathbb{R}^n \rightarrow \mathbb{R}$ mit Infimum Null und Supremum Eins. Andere Quellen erklären die Verteilungsfunktion abweichend als $F_{\mu}(x) = \mu(-\infty, x]$ und erhalten dann analog eine eindeutige Entsprechung zwischen der Menge aller Wahrscheinlichkeitsmaße auf \mathbb{R} und der Menge aller rechtsseitig stetigen monoton wachsenden Funktionen $F : \mathbb{R} \rightarrow \mathbb{R}$ mit Infimum Null und Supremum Eins.

Definition 9.2.19. Sei X eine Menge und $\mathcal{N} \subset \mathcal{P}(X)$ eine σ -Algebra. Ein **äußeres Maß** auf \mathcal{N} ist eine Abbildung $\mu^* : \mathcal{N} \rightarrow [0, \infty]$ derart, daß aus $Y \subset Z$ folgt $\mu^*(Y) \leq \mu^*(Z)$ und daß für jede abzählbare Familie $(Y_n)_{n \in \mathbb{N}}$ von Mengen aus unserer σ -Algebra \mathcal{N} gilt

$$\mu^* \left(\bigcup_{n \in \mathbb{N}} Y_n \right) \leq \sum_{n \in \mathbb{N}} \mu^*(Y_n)$$

Die letzte Bedingung nennen wir die σ -**Subadditivität**. Sie spezialisiert für $N = \emptyset$ zur Bedingung $\mu^*(\emptyset) = 0$.

Lemma 9.2.20. Gegeben X eine Menge, $\mathcal{A} \subset \mathcal{P}(X)$ ein Mengensystem und $\mu : \mathcal{A} \rightarrow [0, \infty]$ eine Abbildung erhalten wir ein äußeres Maß auf $\mathcal{P}(X)$ durch die Vorschrift

$$\mu^*(Y) = \inf \left\{ \sum_{n=0}^{\infty} \mu(A_n) \mid (A_n)_{n \in \mathbb{N}} \text{ ist eine Folge in } \mathcal{A} \text{ mit } Y \subset \bigcup A_n \right\}$$

Beweis. Es ist klar, daß μ^* die erste Eigenschaft eines äußeren Maßes erfüllt. Es bleibt, die σ -Subadditivität zu zeigen. Sei dazu $\varepsilon > 0$ beliebig. Wir finden für jedes n eine Folge A_n^i in \mathcal{A} mit $Y_n \subset \bigcup_{i=0}^{\infty} A_n^i$ und

$$\mu^*(Y_n) \leq \sum_{i=0}^{\infty} \mu(A_n^i) \leq \mu^*(Y_n) + \varepsilon/2^n$$

Dann gilt aber $\bigcup Y_n \subset \bigcup_{i,n} A_n^i$ und aus der Definition von μ^* folgern wir

$$\mu^* \left(\bigcup Y_n \right) \leq \sum_{i,n} \mu(A_n^i) \leq \sum_{n=0}^{\infty} \mu^*(Y_n) + 2\varepsilon$$

Da das nun gilt für alle $\varepsilon > 0$, erfüllt μ^* auch die zweite Bedingung an ein äußeres Maß. \square

Lemma 9.2.21. Gegeben X eine Menge, $\mathcal{A} \subset \mathcal{P}(X)$ ein Mengenring und $\mu : \mathcal{A} \rightarrow [0, \infty]$ ein Prämaß stimmt das in 9.2.20 konstruierte äußere Maß μ^* auf \mathcal{A} mit μ überein, in Formeln $\mu^*(A) = \mu(A) \quad \forall A \in \mathcal{A}$.

Beweis. Die Ungleichung $\mu^*(A) \leq \mu(A)$ ist offensichtlich. Wir müssen also nur noch $\mu(A) \leq \mu^*(A)$ zeigen. Dazu reicht es, wenn wir zeigen $\mu(A) \leq \mu^*(A) + \varepsilon \quad \forall \varepsilon > 0$. Für jedes $\varepsilon > 0$ finden wir aber eine Folge A_n in \mathcal{A} mit $A \subset \bigcup A_n$ und $\sum_{n=0}^{\infty} \mu(A_n) \leq \mu^*(A) + \varepsilon$, und indem wir A_n verkleinern zu $A_n \setminus (A_{n-1} \cup \dots \cup A_0)$ dürfen wir hier sogar die Fonglieder paarweise disjunkt annehmen. Wegen $A = \bigsqcup (A \cap A_n)$ erhalten wir dann wie gewünscht

$$\mu(A) = \sum \mu(A \cap A_n) \leq \sum \mu(A_n) \leq \mu^*(A) + \varepsilon \quad \square$$

9.2.22. Gegeben $A \subset X$ verwenden wir im folgenden für sein Komplement die Abkürzung $X \setminus A = A^c$

Definition 9.2.23. Sei X eine Menge und μ^* ein äußeres Maß auf $\mathcal{P}(X)$. Eine Teilmenge $A \subset X$ heißt μ^* -**meßbar** oder auch ein **Zerleger** genau dann, wenn für jede Teilmenge $Y \subset X$ gilt

$$\mu^*(Y) = \mu^*(Y \cap A) + \mu^*(Y \cap A^c)$$

Lemma 9.2.24 (Zerleger-Lemma). Ist X eine Menge und μ^* ein äußeres Maß auf der Potenzmenge $\mathcal{P}(X)$ von X , so ist das System $\mathcal{M} \subset \mathcal{P}(X)$ der μ^* -meßbaren Mengen eine σ -Algebra und μ^* ist ein Maß auf \mathcal{M} .

Beweis. Zunächst einmal zeigen wir, daß \mathcal{M} eine Mengenalgebra ist. Sicher gilt $\emptyset \in \mathcal{M}$, und aus $A \in \mathcal{M}$ folgt $A^c \in \mathcal{M}$. Wir müssen nur noch zeigen, daß aus $A, B \in \mathcal{M}$ folgt $A \cap B \in \mathcal{M}$. Sei dazu $Y \subset X$ beliebig. Es gilt, für μ^* -meßbare A und B zu zeigen

$$\mu^*(Y) = \mu^*(Y \cap A \cap B) + \mu^*(Y \cap (A \cap B)^c)$$

Da A und B schon μ^* -meßbar sind, finden wir aber in der Tat

$$\begin{aligned} \mu^*(Y \cap (A \cap B)^c) &= \mu^*(Y \cap (A \cap B)^c \cap A) + \mu^*(Y \cap (A \cap B)^c \cap A^c) \\ &= \mu^*(Y \cap B^c \cap A) + \mu^*(Y \cap A^c) \\ &= \mu^*(Y \cap A) - \mu^*(Y \cap A \cap B) + \mu^*(Y \cap A^c) \\ &= \mu^*(Y) - \mu^*(Y \cap (A \cap B)) \end{aligned}$$

Also ist \mathcal{M} schon mal eine Mengenalgebra. Sind $A, B \in \mathcal{M}$ disjunkt, so gilt $\mu^*(A \cup B) = \mu^*(A) + \mu^*(B)$, ja es gilt sogar

$$\mu^*(Y \cap (A \cup B)) = \mu^*(Y \cap A) + \mu^*(Y \cap B)$$

für beliebiges $Y \subset X$, denn unter der Voraussetzung $A \cap B = \emptyset$ können wir schreiben $Y \cap A = Y \cap (A \cup B) \cap A$ und $Y \cap B = Y \cap (A \cup B) \cap A^c$. Induktiv folgt für $A_0, \dots, A_n \in \mathcal{M}$ paarweise disjunkt und $Y \subset X$ beliebig

$$\mu^*(Y \cap (A_0 \cup \dots \cup A_n)) = \sum_{i=0}^n \mu^*(Y \cap A_i)$$

Haben wir also in \mathcal{M} eine Folge (A_i) von paarweise disjunkten Teilmengen mit Vereinigung A gegeben, so gilt für jede Teilmenge $Y \subset X$ die Abschätzung

$$\begin{aligned} \mu^*(Y) &= \mu^*(Y \cap (A_0 \cup \dots \cup A_n)) + \mu^*(Y \cap (A_0 \cup \dots \cup A_n)^c) \\ &\geq \sum_{i=0}^n \mu^*(Y \cap A_i) + \mu^*(Y \cap A^c) \end{aligned}$$

und im Grenzwert $n \rightarrow \infty$ ergibt sich mit der σ -Subadditivität unseres äußeren Maßes

$$\mu^*(Y) \geq \mu^*(Y \cap A) + \mu^*(Y \cap A^c)$$

Die andere Ungleichung gilt eh, nach unseren Annahmen an ein äußeres Maß, also haben wir Gleichheit, und A ist auch μ^* -messbar. Nach Übung 9.1.30 ist damit \mathcal{M} eine σ -Algebra, und für eine Folge A_k von paarweise disjunkten Teilmengen aus \mathcal{M} mit Vereinigungsmenge A folgt aus

$$\sum_{k=0}^n \mu^*(A_k) = \mu^*\left(\bigcup_{k=0}^n A_k\right) \leq \mu^*(A) \leq \sum_{k=0}^{\infty} \mu^*(A_k)$$

im Grenzübergang $n \rightarrow \infty$ die σ -Additivität von μ^* . Also ist μ^* in der Tat ein Maß auf \mathcal{M} . \square

Beweis der Proposition 9.2.13. Seien \mathcal{A} ein Mengenring, $\mu : \mathcal{A} \rightarrow [0, \infty]$ ein Prämaß auf \mathcal{A} und $\mu^* : \mathcal{P}(X) \rightarrow [0, \infty]$ das in 9.2.21 konstruierte äußere Maß. Um die Proposition aus dem Zerlegerlemma 9.2.24 abzuleiten, müssen wir nur noch zeigen, daß \mathcal{A} aus μ^* -messbaren Mengen besteht. Für jedes äußere Maß auf $\mathcal{P}(X)$ und beliebige $A, Y \subset X$ gilt per definitionem

$$\mu^*(Y) \leq \mu^*(Y \cap A) + \mu^*(Y \cap A^c)$$

Wir müssen für $A \in \mathcal{A}$ und beliebiges $Y \subset X$ auch die andere Ungleichung zeigen. Das ist nur im Fall $\mu^*(Y) < \infty$ problematisch. Unter dieser Voraussetzung finden wir aber für beliebiges $\varepsilon > 0$ eine Folge (B_n) in \mathcal{A} mit $Y \subset \bigcup_n B_n$ und

$$\begin{aligned} \mu^*(Y) + \varepsilon &\geq \sum_{n=0}^{\infty} \mu(B_n) \\ &\geq \sum_{n=0}^{\infty} \mu(B_n \cap A) + \mu(B_n \cap A^c) \\ &\geq \mu^*(Y \cap A) + \mu^*(Y \cap A^c) \end{aligned}$$

Da das für alle $\varepsilon > 0$ gilt, folgt die andere Ungleichung

$$\mu^*(Y) \geq \mu^*(Y \cap A) + \mu^*(Y \cap A^c)$$

und damit die Gleichheit. Also besteht \mathcal{A} in der Tat aus μ^* -messbaren Mengen und die Proposition 9.2.13 ist bewiesen. \square

Beweis des Maßfortsetzungssatzes. Die Existenz einer Maßfortsetzung haben wir bereits als Proposition 9.2.13 gezeigt und nur die Eindeutigkeit ist noch zu zeigen. Sei dazu ν eine zweite Fortsetzung. Es gilt zu zeigen $\mu(C) = \nu(C)$ für alle $C \in \mathcal{M}$. Aus der Konstruktion von μ in 9.2.13 folgt bereits $\nu(C) \leq \mu(C)$. Da wir unser Prämaß σ -endlich angenommen hatten, gibt es jedoch eine aufsteigende

Folge $A_0 \subset A_1 \subset A_2 \subset \dots$ in \mathcal{A} mit $\bigcup A_n \supset C$ und $\mu(A_n) < \infty \quad \forall n$. Wir müssen nur für alle n die Gleichungen

$$\mu(C \cap A_n) = \nu(C \cap A_n)$$

zeigen, dann ergibt sich $\mu(C) = \nu(C)$ im Grenzwert $n \rightarrow \infty$. Wie bereits erwähnt gilt jedoch $\nu(C \cap A_n) \leq \mu(C \cap A_n)$ und ganz genauso auch $\nu(C^c \cap A_n) \leq \mu(C^c \cap A_n)$, und da die Summe dieser Ungleichungen die Gleichung $\nu(A_n) = \mu(A_n)$ liefert, müssen unsere Ungleichungen beide schon Gleichungen gewesen sein. \square

Übung 9.2.25. Zeigen Sie, dass es höchstens ein normiertes translationsinvariantes topologisches Maß λ auf \mathbb{R} geben kann. Hinweis: Zeigen Sie zunächst $\lambda(\{a\}) = 0$, für alle $a \in \mathbb{R}$, und anschließend, dass für alle $n \in \mathbb{N}$ gilt: $\lambda([0, 1/n]) = 1/n$. Erweitern Sie als nächstes die Aussage auf Intervalle mit rationalen Endpunkten und schließlich auf beliebige Intervalle. Wenden Sie dann den Satz über Maßfortsetzungen an.

Übung 9.2.26. Zeigen Sie, dass es höchstens ein normiertes translationsinvariantes topologisches Maß λ auf \mathbb{R}^n geben kann. Hinweis: [9.2.25](#).

Ergänzende Übung 9.2.27 (Benford's Gesetz). Zeigen Sie, daß es auf jedem nichtleeren kompakten Intervall $I = [a, b]$ genau ein topologisches Maß μ gibt, das dem ganzen Intervall das Maß Eins zuweist und das "partiell translationsinvariant" ist in dem Sinne, daß für jede Borelmenge $A \subset I$ und jedes $a \in \mathbb{R}$ mit $a + A \subset I$ gilt $\mu(A) = \mu(a + A)$. Zeigen Sie, daß es auf jedem nichtleeren kompakten Intervall $I = [\alpha, \beta] \subset \mathbb{R}_{>0}$ genau ein topologisches Maß μ gibt, das dem ganzen Intervall das Maß Eins zuweist und das "partiell skaleninvariant" ist in dem Sinne, daß für jede Borelmenge $A \subset I$ und jedes $c \in \mathbb{R}_{>0}$ mit $cA \subset I$ gilt $\mu(A) = \mu(cA)$, und daß dieses Maß ein Vielfaches von $x^{-1} dx$ ist. Gegeben ein derartiges Maß mit Gesamtmasse Eins und I so groß, daß gilt $\beta > 10^n \alpha$ für $n \geq 1$, wird dann für jede Ziffer $i \in \{1, \dots, 9\}$ das Maß der Menge M_i aller $x \in I$, die als Dezimalbruch mit erster von Null verschiedener Ziffer i geschrieben werden können, von $(\log(i+1) - \log(i))/\log(10)$ um weniger als $1/(n+1)$ abweichen. Diese Verteilung der Anfangsziffern "zufälliger" Zahlenreihen tritt in der Wirklichkeit häufig auf und heißt **Benford's Gesetz**. Es wird unter anderem eingesetzt, um Steuerbetrug zu entlarven, da von Menschen willkürlich hingeschriebene Zahlenreihen typischerweise eine andere Verteilung von Anfangsziffern haben.

Ergänzende Übung 9.2.28 (Gleichverteilung im Folgenraum). Man zeige: Auf dem Raum $\text{Ens}(\mathbb{N}, \{W, Z\})$ aller Folgen in der zweielementigen Menge $\{W, Z\}$ mit der in [9.7.17](#) erklärten Metrik gibt es genau ein Borelmaß, das für jeden n -gliedrigen Folgenanfang der Menge aller Folgen mit diesem Anfang das Maß 2^{-n} zuordnet. Man zeige weiter, daß die durch die dyadische Entwicklung gegebene

Surjektion $\text{Ens}(\mathbb{N}, \{W, Z\}) \rightarrow [0, 1]$ stetig ist und daß hier das Maß des Urbilds einer Borelmenge gerade das Lebesguemaß der ursprünglichen Menge ist. Hinweis: Man mag eine Teilmenge unseres Folgenraums “ n -vernünftig” nennen genau dann, wenn sie mit einer Folge auch alle anderen Folgen enthält, die sich von dieser frühestens im n -ten Folgenglied unterscheiden. Man mag eine Teilmenge unseres Folgenraums “vernünftig” nennen genau dann, wenn sie n -vernünftig ist für mindestens ein n . Man mag von der Erkenntnis ausgehen, daß die vernünftigen Teilmengen einen Mengering bilden, und verwenden, daß alle vernünftigen Teilmengen sowohl offen als auch abgeschlossen und damit nach 9.7.17 kompakt sind. Jede Überdeckung einer vernünftigen Teilmenge durch vernünftige Teilmengen besitzt folglich eine endliche Teilüberdeckung.

Ergänzende Übung 9.2.29. Wir betrachten die **Cantor-Menge** C , die aus dem Einheitsintervall $C_0 = [0, 1]$ entsteht, indem wir das mittlere Drittel $(1/3, 2/3)$ herausnehmen, dann aus den beiden so entstehenden kompakten Intervallen wieder jeweils das offene mittlere Drittel und so weiter, und schließlich als C den Schnitt über alle Mengen C_n nehmen, die wir in dieser Weise in n Schritten erhalten. Man zeige, daß die Cantor-Menge das Lebesgue-Maß $\lambda(C) = 0$ Null hat und überabzählbar ist. Hinweis: Man kann die Cantor-Menge auch beschreiben als die Menge aller Zahlen, die sich in der Basis Drei mit einer Null vor dem Komma und ohne die Ziffer Eins ausdrücken lassen, in Formeln

$$C = \left\{ \sum_{i=1}^{\infty} a_i 3^{-i} \mid a_i \in \{0, 2\} \right\}$$

Ergänzende Übung 9.2.30. Die Menge aller reellen Zahlen, die sich darstellen lassen durch einen unendlichen Dezimalbruch, in dem die Ziffer 6 nicht vorkommt, bilden eine abgeschlossene Teilmenge von \mathbb{R} vom Lebesgue-Maß Null.

Definition 9.2.31. Eine Teilmenge eines Maßraums, die in einer meßbaren Menge vom Maß Null enthalten ist, heißt eine **Nullmenge**. Ein Maßraum $X = (X, \mathcal{M}, \mu)$ heißt **vollständig** genau dann, wenn jede Nullmenge bereits meßbar ist, d.h. zu \mathcal{M} gehört.

Proposition 9.2.32 (Vervollständigung von Maßräumen). *Gegeben ein Maßraum (X, \mathcal{M}, μ) gibt es genau eine Fortsetzung von μ zu einem Maß μ^* auf der von \mathcal{M} und den μ -Nullmengen erzeugten σ -Algebra \mathcal{M}^* , und der so entstehende Maßraum $(X, \mathcal{M}^*, \mu^*)$ ist vollständig.*

Beweis. Erweitern wir μ zu einem äußeren Maß μ^* auf $\mathcal{P}(X)$ wie im Lemma 9.2.21 und wenden auf dieses äußere Maß das Zerleger-Lemma 9.2.24 an, so folgt, daß alle μ -Nullmengen bereits μ^* -meßbar sind und daß mithin μ^* ein Maß ist auf \mathcal{M}^* . Das zeigt die Existenz von μ^* . Für die Eindeutigkeit prüft man, daß \mathcal{M}^*

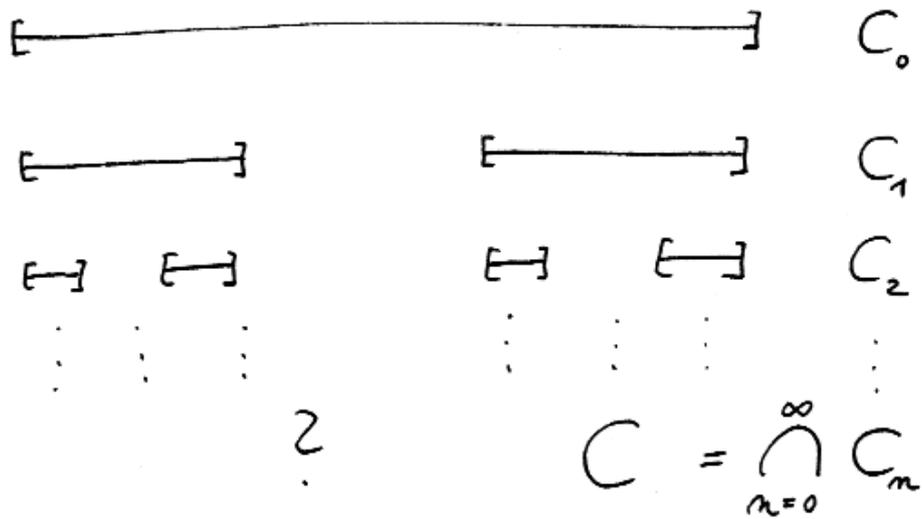


Illustration zur Cantor-Menge

genau aus allen Teilmengen $E \subset X$ besteht derart, daß es $A, B \in \mathcal{M}$ gibt mit $A \subset E \subset B$ und $\mu(B \setminus A) = 0$. In der Tat bilden nämlich alle diese E eine σ -Algebra. Jedes erweiterte Maß μ^* nimmt aber auf einer solchen Teilmenge E notwendig den Wert $\mu^*(E) = \mu(A)$ an. \square

Übung 9.2.33. Man arbeite den Schluß des vorhergehenden Beweises aus und zeige, daß gegeben ein Maßraum (X, \mathcal{M}, μ) das Mengensystem aller $E \subset X$ derart, daß es $A, B \in \mathcal{M}$ gibt mit $A \subset E \subset B$ und $\mu(B \setminus A) = 0$, eine σ -Algebra ist.

9.2.34. Der Maßraum $(X, \mathcal{M}^*, \mu^*)$ heißt die **Vervollständigung** des Maßraums (X, \mathcal{M}, μ) . Die bezüglich der Vervollständigung des Lebesgue-Maßes meßbaren Teilmengen von \mathbb{R} bzw. \mathbb{R}^n nennt man die **Lebesgue-meßbaren** Teilmengen oder kurz **Lebesgue-Mengen**. Es ist nicht ganz einfach, eine Lebesgue-Menge in \mathbb{R} explizit anzugeben, die nicht topologisch meßbar ist. Genauer gesagt wüßte ich selber nicht, wie ich das machen sollte, und müßte einen Logiker um Hilfe bitten. Man kann jedoch zeigen, daß es im Sinne der Mengenlehre “mehr” Lebesgue-Mengen in \mathbb{R} gibt als topologisch meßbare Teilmengen, vergleiche etwa ??.

Übung 9.2.35. Eine Teilmenge der reellen Zahlengeraden ist eine Nullmenge in Bezug auf das Lebesgue-Maß genau dann, wenn sie sich für jedes $\varepsilon > 0$ durch eine Folge von kompakten Intervallen $[a_n, b_n]$ überdecken läßt mit $\sum_{n=0}^{\infty} (b_n - a_n) < \varepsilon$. Hinweis: 9.2.10 und 9.2.13.

9.3 Meßbare Abbildungen

9.3.1. Bis jetzt haben wir uns nur mit dem Messen von Mengen beschäftigt. Wir haben gesehen, daß das Messen ganz beliebiger Teilmengen der reellen Zahlengerade problematisch ist, konnten jedoch gewisse Mengen als meßbar auszeichnen und solchen Mengen sinnvoll ein Maß zuordnen. Nun wollen wir reellwertigen Funktionen auf Maßräumen ein Integral zuordnen. Wieder ist das für ganz beliebige Funktionen problematisch, aber wieder können wir gewisse Funktionen als “meßbar” auszeichnen und zumindest allen nichtnegativen meßbaren Funktionen sinnvoll ein Integral zuordnen. In einem zweiten Schritt geben wir dann auch eine Definition für das Integral beliebiger meßbarer reellwertiger Funktionen mit der Eigenschaft, daß ihr Betrag ein endliches Integral hat.

Definition 9.3.2. Seien (X, \mathcal{M}) und (Y, \mathcal{N}) Meßräume. Eine Abbildung $f : X \rightarrow Y$ heißt **meßbar** genau dann, wenn das Urbild jeder meßbaren Menge meßbar ist, in Formeln $V \in \mathcal{N} \Rightarrow f^{-1}(V) \in \mathcal{M}$.

9.3.3. Die folgende Übung zeigt, daß nicht alle Abbildungen $\mathbb{R} \rightarrow \mathbb{R}$ meßbar sind. Ganz im Gegenteil sind die wenigsten Abbildungen $\mathbb{R} \rightarrow \mathbb{R}$ meßbar, obwohl es nicht ganz leicht ist, Beispiele anzugeben. Die Lage ist ähnlich wie bei

den transzendenten Zahlen: Obwohl alle reellen Zahlen bis auf abzählbar viele Ausnahmen transzendent sind, ist es doch vergleichsweise schwierig, explizite Beispiele für transzendente Zahlen anzugeben.

Übung 9.3.4. Man gebe eine Abbildung $\mathbb{R} \rightarrow \mathbb{R}$ an, die nicht meßbar ist in Bezug auf die Borel'schen σ -Algebren. Hinweis: Man beachte 9.1.28 und seinen Beweis.

Ergänzung 9.3.5. Eine meßbare Abbildung von einem Wahrscheinlichkeitsraum in einen Meßraum heißt auch eine **Zufallsvariable** auf unserem Wahrscheinlichkeitsraum mit Werten in besagtem Meßraum oder auch ein **zufälliges Element** von besagtem Meßraum.

Lemma 9.3.6. *Jede Verknüpfung meßbarer Abbildungen ist meßbar.*

Beweis. Das folgt unmittelbar aus der Definition mit demselben Argument wie wir es in 9.5.18 für die Stetigkeit der Verknüpfung gesehen hatten. \square

9.3.7. Wenn nichts anderes gesagt ist, denken wir uns einen metrischen oder allgemeiner topologischen Raum stets mit der durch die Borel'sche σ -Algebra gegebene Struktur eines Meßraums versehen. Im Fall des topologischen Raums $\overline{\mathbb{R}}$ kann die Borel'sche σ -Algebra auch beschrieben werden als die von allen Intervallen erzeugte σ -Algebra, und wer mag, kann im folgenden auch diese Beschreibung als Definition nehmen.

Proposition 9.3.8. *Alle stetigen Abbildungen sind meßbar.*

Beweis. Da die Urbilder offener Mengen unter stetigen Abbildungen nach 9.5.17 stets offen sind, folgt das unmittelbar aus dem anschließenden Lemma 9.3.9. \square

Lemma 9.3.9. *Seien (X, \mathcal{M}) und (Y, \mathcal{N}) zwei Meßräume und sei die σ -Algebra \mathcal{N} erzeugt von einem Teilsystem $\mathcal{S} \subset \mathcal{N}$. Genau dann ist $f : X \rightarrow Y$ meßbar, wenn gilt*

$$V \in \mathcal{S} \Rightarrow f^{-1}(V) \in \mathcal{M}$$

Beweis. Es reicht zu zeigen, daß für eine beliebige Abbildung $f : X \rightarrow Y$ die Menge $f_*\mathcal{M} = \{V \subset Y \mid f^{-1}(V) \in \mathcal{M}\}$ eine σ -Algebra ist. Das ist jedoch klar. Im Übrigen heißt die σ -Algebra $f_*\mathcal{M}$ das **Bild unter f** der σ -Algebra \mathcal{M} . \square

Übung 9.3.10. Alle monotonen Abbildungen $\mathbb{R} \rightarrow \overline{\mathbb{R}}$ sind meßbar.

9.3.11. Gegeben ein Meßraum (X, \mathcal{M}) betrachten wir auf jeder Teilmenge $A \subset X$ die **induzierte σ -Algebra** $\mathcal{M}|_A$, die gerade aus allen Schnitten mit A von meßbaren Mengen in X besteht, $\mathcal{M}|_A = \{Z \cap A \mid Z \in \mathcal{M}\}$. Wir machen unsere erweiterten reellen Zahlen $\overline{\mathbb{R}} = [-\infty, \infty]$ zu einem Meßraum, indem wir darauf die von allen Intervallen erzeugte σ -Algebra betrachten. Für die natürliche Topologie auf $\overline{\mathbb{R}}$ im Sinne von 9.5.8 ist das genau die σ -Algebra der topologisch meßbaren Mengen.

Ergänzende Übung 9.3.12. Gegeben eine Abbildung $f : X \rightarrow Y$ und eine σ -Algebra $\mathcal{N} \subset \mathcal{P}(Y)$ erhalten wir eine σ -Algebra $f^*\mathcal{N} \subset \mathcal{P}(X)$ durch die Vorschrift $f^*\mathcal{N} = \{f^{-1}(N) \mid N \in \mathcal{N}\}$. Wir nennen sie das **Urbild unter f** der σ -Algebra \mathcal{N} . Zum Beispiel ist unsere induzierte σ -Algebra aus 9.3.11 das Urbild der ursprünglichen σ -Algebra unter der Einbettung.

Ergänzung 9.3.13. Ein topologischer Raum heißt **separabel** oder gleichbedeutend **zweitabzählbar** genau dann, wenn es darin ein abzählbares System von offenen Teilmengen gibt derart, daß jede offene Menge als Vereinigung eines Teilsystems dieses abzählbaren Systems geschrieben werden kann. Ein metrischer Raum ist separabel in diesem Sinne genau dann, wenn er eine abzählbare dichte Teilmenge besitzt. In der Tat können wir aus einem abzählbaren System von offenen Teilmengen wie eben leicht eine abzählbare dichte Teilmenge erhalten, indem wir aus jeder nichtleeren Teilmenge unseres Systems einen Punkt auswählen. Umgekehrt erhalten wir aus einer abzählbaren dichten Teilmenge ein abzählbares System offener Mengen mit der gewünschten Eigenschaft als das System aller offenen Bälle mit rationalen Radien um besagte Punkte.

Ergänzung 9.3.14. Manche Autoren, so etwa Jameson oder Rudin, verwenden eine andere Terminologie, die auf Fréchet zurückzugehen scheint und nach der ein topologischer Raum “separabel” heißt genau dann, wenn er eine abzählbare dichte Teilmenge besitzt. Bourbaki führt diesen Begriff nur für metrische Räume ein. In der Bedeutung dieses Textes wird der Begriff “separabel” auch im Buch von Halmos zur Maßtheorie verwendet.

Ergänzende Übung 9.3.15. Jeder kompakte metrische Raum ist separabel.

Proposition 9.3.16 (Komponentenregel). *Seien (X, \mathcal{M}) ein Meßraum, Y, Z separable metrische Räume und $f : X \rightarrow Y$ und $g : X \rightarrow Z$ Abbildungen. So ist $(f, g) : X \rightarrow Y \times Z$ meßbar genau dann, wenn f und g meßbar sind.*

Ergänzung 9.3.17. Leser, denen die Produkt-Topologie vertraut ist, werden unschwer erkennen, daß der Satz mit fast demselben Beweis allgemeiner für beliebige separable topologische Räume Y und Z gilt.

Beweis. Mit (f, g) sind natürlich auch $f = \text{pr}_1 \circ (f, g)$ und $g = \text{pr}_2 \circ (f, g)$ meßbar. Sind andererseits f und g meßbar, so gilt $f^{-1}B(y; \delta) \in \mathcal{M}$ und $g^{-1}B(z; \delta) \in \mathcal{M}$ für alle $y \in Y, z \in Z, \delta > 0$, und daraus folgt

$$(f, g)^{-1}B((y, z); \delta) = f^{-1}B(y; \delta) \cap g^{-1}B(z; \delta) \in \mathcal{M}$$

Mit 9.1.31 folgt aber, daß die offenen Bälle $B((y, z); \delta)$ in $Y \times Z$ schon die σ -Algebra der topologisch meßbaren Mengen in $Y \times Z$ erzeugen. Damit folgt die Proposition aus 9.3.9. \square

Korollar 9.3.18. *Die Summe und das Produkt reellwertiger meßbarer Funktionen sind wieder meßbar.*

Beweis. Sei X ein Meßraum und seien $f, g : X \rightarrow \mathbb{R}$ meßbare Funktionen. Nach Proposition 9.3.16 ist dann $(f, g) : X \rightarrow \mathbb{R}^2$ meßbar, und damit auch die Verknüpfung von (f, g) mit der stetigen und daher meßbaren Addition bzw. Multiplikation $\mathbb{R}^2 \rightarrow \mathbb{R}$. \square

Satz 9.3.19 (Meßbarkeit von Grenzwerten von Funktionenfolgen). *Sei (X, \mathcal{M}) ein Meßraum.*

1. *Eine Funktion $f : X \rightarrow \overline{\mathbb{R}}$ ist genau dann meßbar, wenn für jedes $a \in \mathbb{R}$ die Menge $\{x \mid f(x) > a\} = f^{-1}(a, \infty]$ meßbar ist in X .*
2. *Für jede Folge meßbarer Funktionen $f_n : X \rightarrow \overline{\mathbb{R}}$ sind auch das Supremum und das Infimum $s, i : X \rightarrow \overline{\mathbb{R}}$, $s(x) = \sup f_n(x)$ bzw. $i(x) = \inf f_n(x)$ meßbar.*
3. *Konvergiert eine Folge meßbarer Funktionen $f_n : X \rightarrow \overline{\mathbb{R}}$ punktweise gegen $f : X \rightarrow \overline{\mathbb{R}}$, so ist auch f meßbar.*

Beweis. 1. Die Intervalle der Form $(a, \infty]$ erzeugen die σ -Algebra der Borel-Mengen in $\overline{\mathbb{R}}$.

2. Es gilt $s^{-1}(a, \infty] = \bigcup_{n=0}^{\infty} f_n^{-1}(a, \infty]$ und ganz analog haben wir auch $i^{-1}[-\infty, a) = \bigcup_{n=0}^{\infty} f_n^{-1}[-\infty, a)$.

3. Nach 2 sind auch die Funktionen $s_N(x) = \sup_{n \geq N} f_n(x)$ meßbar, und dann auch die Funktion $g : X \rightarrow \overline{\mathbb{R}}$, $g(x) = \inf_N s_N(x)$. Diese Funktion bezeichnet man auch mit $g = \limsup f_n$. Falls die f_n punktweise gegen eine Grenzfunktion f konvergieren, so gilt insbesondere $f = \limsup f_n$, mithin ist dann auch f meßbar. \square

Lemma 9.3.20. *Ist X ein Meßraum und Y ein metrischer Raum und konvergiert eine Folge meßbarer Funktionen $f_n : X \rightarrow Y$ punktweise gegen eine Grenzfunktion $f : X \rightarrow Y$, so ist auch die Grenzfunktion f meßbar.*

9.3.21. Wir können den dritten Teil von Satz 9.3.19 auch als Spezialfall dieses Lemmas erhalten, wenn wir etwa beachten, daß unsere Topologie auf $\overline{\mathbb{R}}$ auch als eine metrische Topologie erhalten werden kann. Das Lemma gilt im übrigen mit demselben Beweis für einen beliebigen Hausdorffraum mit der Eigenschaft, daß jede seiner abgeschlossenen Mengen als der Schnitt einer abzählbaren Familie offener Mengen geschrieben werden kann.

Beweis. Nach 9.3.9 reicht es, für alle abgeschlossenen Teilmengen $A \subseteq Y$ zu zeigen, daß ihr Urbild unter f meßbar ist. Nun gibt es jedoch eine absteigende Folge offener Mengen $U_0 \supset U_1 \supset \dots$ mit Schnitt A . Dann ist $f(x) \in A$ gleichbedeutend dazu, daß es für jedes $i \in \mathbb{N}$ ein $N = N(x, i)$ gibt mit $f_n(x) \in U_i$ für $n \geq N(x, i)$. Damit können wir $f^{-1}(A)$ wie folgt beschreiben: Wir bilden zunächst für jedes i die Menge

$$V_i := \{x \in X \mid \exists N \text{ mit } n \geq N \Rightarrow f_n(x) \in U_i\} = \bigcup_{N \geq 0} \bigcap_{n \geq N} f_n^{-1}(U_i)$$

und erhalten dann $f^{-1}(A) = \bigcap_{i \geq 0} V_i$. Diese Darstellung zeigt jedoch, daß mit den $f_n^{-1}(U_i)$ auch $f^{-1}(A)$ meßbar sein muß. \square

Übung 9.3.22. Gegeben eine meßbare Abbildung $\phi : X \rightarrow Y$ von Meßräumen und ein Maß μ auf X erklärt man das **Bildmaß** $\phi_*\mu$ auf Y dadurch, daß man für jede meßbare Menge $A \subset Y$ setzt

$$(\phi_*\mu)(A) = \mu(\phi^{-1}A)$$

Man zeige, daß diese Vorschrift in der Tat ein Maß auf Y liefert. Man zeige auch für eine weitere meßbare Abbildung $\psi : Y \rightarrow Z$ von Meßräumen die Formel $\psi_*(\phi_*\mu) = (\psi \circ \phi)_*\mu$ und für die Identität auf X die Formel $\text{id}_*\mu = \mu$.

Ergänzung 9.3.23. Ist (Ω, P) ein Wahrscheinlichkeitsraum, so heißt das Bildmaß des Wahrscheinlichkeitsmaßes unter einer Zufallsvariable $X : \Omega \rightarrow Y$ die **Verteilung der Zufallsvariable** und wird P^X notiert. Im Fall einer reellwertigen Zufallsvariable $Y = \mathbb{R}$ ist dann P^X ein Wahrscheinlichkeitsmaß auf \mathbb{R} , und seine Verteilungsfunktion im Sinne von 9.2.18 heißt die **Verteilungsfunktion** oder präziser die **kumulative Verteilungsfunktion** unserer Zufallsvariablen. Das Bildmaß des Dirac-Maßes auf einer einpunktigen Menge unter einer Abbildung in einen Meßraum ist das Dirac-Maß am Bildpunkt.

9.3.24. Manchmal scheint mir die äquivalente Terminologie der “Verwandtschaft” transparenter, die ich nun einführe. Gegeben eine meßbare Abbildung $\phi : X \rightarrow Y$ von Meßräumen und ein Maß μ auf X und ein Maß ν auf Y heißen die beiden Maße ϕ -**verwandt** und wir schreiben

$$\phi : \mu \rightsquigarrow \nu$$

genau dann, wenn gilt $\nu(A) = \mu(\phi^{-1}A)$ für jede meßbare Teilmenge $A \subset Y$. Gleichbedeutend ist per definitionem $\nu = \phi_*\mu$. Jedes Maß hat also unter jeder meßbaren Abbildung genau einen “Vorwärtsverwandten”. Das mag den konzeptionellen Unterschied zwischen Maßen und Funktionen deutlich machen, die im Gegensatz dazu stets genau einen “Rückwärtsverwandten” haben.

Ergänzung 9.3.25. Eine Funktion auf einem Maßraum (X, \mathcal{M}, μ) , die meßbar ist auf dem in Bezug auf das Maß μ vervollständigten Maßraum, nennen wir **μ -meßbar**. Insbesondere heißt eine Abbildung $\mathbb{R} \rightarrow \overline{\mathbb{R}}$ **Lebesgue-meßbar** genau dann, wenn sie **λ -meßbar** ist für λ das Lebesgue-Maß. Wir werden nach Möglichkeit versuchen, ohne diese Begrifflichkeit auszukommen. Die Beziehung dieses Begriffs zur Meßbarkeit klärt 9.4.22.

9.4 Das Integral von nichtnegativen Funktionen

Definition 9.4.1. Eine Funktion, die nur endlich viele Werte annimmt, nenne ich eine **Stufenfunktion**.

9.4.2. Ist X eine Menge und $A \subset X$ eine Teilmenge, so ist zum Beispiel die **charakteristische Funktion** $\chi_A = [A] : X \rightarrow \{0, 1\}$ von A , definiert durch die Vorschrift

$$[A](x) = \begin{cases} 1 & x \in A \\ 0 & x \notin A \end{cases}$$

eine Stufenfunktion. Ist (X, \mathcal{M}) ein Maßraum, so ist jede reellwertige meßbare Stufenfunktion $s : X \rightarrow \mathbb{R}$ von der Form

$$s = \sum_{i=1}^n c_i [A_i]$$

für eine Zerlegung $X = \bigsqcup_{i=1}^n A_i$ von X in endlich viele paarweise disjunkte meßbare Mengen und geeignete $c_i \in \mathbb{R}$, und die reellwertigen meßbaren Stufenfunktionen bilden einen Untervektorraum im Raum aller reellwertigen Funktionen auf X .

Definition 9.4.3 (Integral nichtnegativer reeller Stufenfunktionen). Gegeben ein Maßraum (X, \mathcal{M}, μ) und eine nichtnegative meßbare reellwertige Stufenfunktion $s : X \rightarrow [0, \infty)$ erklärt man das **Integral** $\int s = \int_X s = \int_X s \mu \in [0, \infty]$ von s über X durch die Formel

$$\int_X s \mu = \sum_{c \in s(X) \setminus 0} c \cdot \mu(s^{-1}(c))$$

9.4.4. Ich haben in dieser Formel den Summanden für $c = 0$ weggelassen, um den Ausdruck $0 \cdot \infty$ zu vermeiden. Im folgenden erweist es sich jedoch als bequemer, diesen Summanden zuzulassen und mit der Konvention $0 \cdot \infty = 0$ zu arbeiten. Weiter habe ich nur $c \in s(X)$ statt $c \in \mathbb{R}$ geschrieben, um eine endliche Summe zu erhalten. Da aber die Summanden für $c \notin s(X)$ eh Null sind, hätten wir ohne etwas am Resultat zu ändern auch über alle $c \in \mathbb{R}$ summieren können.

9.4.5. Natürlich gilt $\int \alpha s = \alpha \int s$, $\forall \alpha \in (0, \infty)$, und ist $t : X \rightarrow [0, \infty)$ eine zweite meßbare Stufenfunktion, so gilt $\int s + t = \int s + \int t$ und mithin auch $s \leq t \Rightarrow \int s \leq \int t$. In der Tat, schreiben wir $X_{a,b} = s^{-1}(a) \cap t^{-1}(b)$, so ergibt sich mit der Additivität des Maßes unmittelbar

$$\begin{aligned} \int s + t &= \sum_c c \cdot \mu \left(\bigcup_{c=a+b} X_{a,b} \right) = \sum_{a,b} (a+b) \cdot \mu(X_{a,b}) \\ \int s &= \sum_a a \cdot \mu \left(\bigcup_b X_{a,b} \right) = \sum_{a,b} a \cdot \mu(X_{a,b}) \\ \int t &= \sum_b b \cdot \mu \left(\bigcup_a X_{a,b} \right) = \sum_{a,b} b \cdot \mu(X_{a,b}) \end{aligned}$$

Für $s = \sum_{i=1}^n c_i [A_i]$ mit $c_i \in [0, \infty)$ wird also das Integral gegeben durch die Formel $\int s = \sum_i c_i \mu(A_i)$.

Definition 9.4.6 (Integral nichtnegativer Funktionen). Gegeben ein Maßraum $X = (X, \mathcal{M}, \mu)$ und eine meßbare Abbildung $f : X \rightarrow [0, \infty]$ definieren wir ein Element $\int f \mu$ aus $[0, \infty]$, das **Integral von f über X** , als

$$\int f \mu = \sup_{s \leq f} \int_X s \mu$$

wobei das Supremum gebildet wird über alle reellwertigen nichtnegativen meßbaren Stufenfunktionen $s : X \rightarrow [0, \infty)$ mit $s(x) \leq f(x)$ für alle $x \in X$. Ist f bereits selbst eine reelle Stufenfunktion, so wird das fragliche Supremum bei $s = f$ angenommen und wir erhalten unser Integral von Stufenfunktionen 9.4.3 für $s = f$.

9.4.7. Wir verwenden für dies Integral auch die Notationen

$$\int f = \int_X f \mu = \int_X f(x) \mu \langle x \rangle$$

Die eckigen Klammern sollen andeuten, daß mit $\mu \langle x \rangle$ nicht der Wert einer etwaigen Funktion μ an einer Stelle x gemeint ist. Vielmehr wird dieser Ausdruck erst in Verbindung mit dem Integralzeichen sinnvoll. In der Literatur findet man meist die Notation $\int_X f \, d\mu$. Diese leider allgemein übliche Notation scheint mir jedoch im Lichte der ursprünglichen Bedeutung des Symbols d unter dem Integralzeichen völlig abwegig, um nicht zu sagen irreführend: Das Differential d macht aus einer Funktion ein Maß, wie etwa in 9.2.7 erklärt, aber wo bereits ein Maß steht, hat es nichts mehr zu suchen.

9.4.8. Unsere Definition des Integrals 9.4.6 ist sogar sinnvoll für nicht notwendig meßbare Funktionen $f : X \rightarrow [0, \infty]$. Das Supremum heißt dann das **Unterintegral** von f . Der Beweis des folgenden Satzes über monotone Konvergenz zeigt, welche Rolle die Meßbarkeit von f spielt, und der Beweis des gleich anschließenden Satzes 9.4.11 zeigt dann weiter, wie die Meßbarkeit in den Beweis der Formel $\int f + g = \int f + \int g$ eingeht, die für nicht meßbare Funktionen f und g im Allgemeinen nicht mehr richtig ist.

Satz 9.4.9 (über monotone Konvergenz). Ist (X, \mathcal{M}, μ) ein Maßraum und (f_n) eine monoton wachsende Folge meßbarer Funktionen $f_n : X \rightarrow [0, \infty]$, d.h. $0 \leq f_0 \leq f_1 \leq \dots$, so ist auch der punktweise Grenzwert $f : X \rightarrow [0, \infty]$, $f(x) = \lim_{n \rightarrow \infty} f_n(x)$ der f_n meßbar und es gilt

$$\int_X f \mu = \lim_{n \rightarrow \infty} \int_X f_n \mu$$

9.4.10. Das Bild zu 8.1.14 zeigt eine Folge stetiger Funktionen, die punktweise gegen die Nullfunktion konvergieren, ohne daß ihre Integrale deshalb gegen Null streben. Die Annahme der Monotonie unserer Folge ist also wesentlich.

Beweis. Die Meßbarkeit von f folgt aus 9.3.19. Die Abschätzung \geq ist evident. Es gilt, \leq zu zeigen. Dafür reicht es, wenn wir für jede Stufenfunktion s mit $s \leq f$ und jedes $\eta \in (0, 1)$ die Abschätzung

$$\eta \int s \leq \lim_{n \rightarrow \infty} \int f_n$$

zeigen. Nun haben wir ja $s = \sum_{i=0}^r c_i [A_i]$ für geeignete paarweise disjunkte meßbare A_i und $c_i \in (0, \infty)$. Setzen wir $A_i^n = \{x \in A_i \mid f_n(x) \geq \eta c_i\}$, so sind auch die A_i^n meßbar und es gilt $A_i^0 \subset A_i^1 \subset A_i^2 \subset \dots$ sowie $\bigcup A_i^n = A_i$, nach 9.1.26 also

$$\lim_{n \rightarrow \infty} \mu(A_i^n) = \mu(A_i)$$

Betrachten wir die Stufenfunktionen $s_n = \sum_i \eta c_i [A_i^n]$, so gilt mithin

$$\lim_{n \rightarrow \infty} \int s_n = \eta \int s$$

Andererseits haben wir aber auch nach Konstruktion $s_n \leq f_n$ und folglich $\int s_n \leq \int f_n$. Bilden wir nun auf beiden Seiten den Grenzwert für $n \rightarrow \infty$, so ergibt sich damit $\eta \int s \leq \lim_{n \rightarrow \infty} \int f_n$ wie gewünscht. \square

Satz 9.4.11. Gegeben f, g nichtnegative meßbare Funktionen mit Werten in $[0, \infty]$ auf einem Maßraum gilt

1. $f \leq g \Rightarrow \int f \leq \int g$;
2. $\int cf = c \int f \quad \forall c \in (0, \infty)$;
3. $\int f + g = \int f + \int g$.

Beweis. Nur der dritte Punkt braucht einen Beweis. Sind f und g reelle Stufenfunktionen, so haben wir die Behauptung schon in 9.4.5 gezeigt. Um den allgemeinen Fall daraus abzuleiten, brauchen wir folgendes

Lemma 9.4.12. Sei (X, \mathcal{M}) ein Maßraum und $f : X \rightarrow [0, \infty]$ eine meßbare Funktion. So gibt es eine monotone Folge von meßbaren Stufenfunktionen $0 \leq \varphi_0 \leq \varphi_1 \leq \dots$ mit Werten in $[0, \infty)$, die punktweise gegen f konvergiert.

Beweis. Wir konstruieren φ_n zum Beispiel wie folgt: Sei $0 = a_0 < a_1 < \dots < a_r = n$ die äquidistante Einteilung von $[0, n]$ in Stücke der Länge 2^{-n} , also $r = n2^n$ und $a_i = i2^{-n}$. Wir setzen $A_i = f^{-1}[a_i, a_{i+1})$ für $0 \leq i < r$ sowie $A_r = f^{-1}[n, \infty)$ und bilden $\varphi_n = \sum_{i=0}^r a_i \mathbb{1}_{A_i}$. Es ist offensichtlich, daß wir so eine monotone Folge von Stufenfunktionen erhalten, die punktweise gegen f konvergiert. \square

Jetzt schreiben wir f und g als punktweise Grenzwerte von monotonen Folgen meßbarer Stufenfunktionen, $0 \leq \varphi_0 \leq \varphi_1 \leq \dots$ und $0 \leq \psi_0 \leq \psi_1 \leq \dots$, und folgern mit dem Satz 9.4.9 über monotone Konvergenz

$$\int f + g = \lim_{n \rightarrow \infty} \int \varphi_n + \psi_n = \lim_{n \rightarrow \infty} \int \varphi_n + \lim_{n \rightarrow \infty} \int \psi_n = \int f + \int g \quad \square$$

Übung 9.4.13. Sei $(X, \mathcal{M}^*, \mu^*)$ ein Maßraum, $\mathcal{M} \subset \mathcal{M}^*$ eine σ -Unteralgebra und $\mu = \mu^*|_{\mathcal{M}}$ das darauf induzierte Maß. Man zeige: Für jede bezüglich \mathcal{M} meßbare Funktion $f : X \rightarrow [0, \infty]$ gilt $\int f \mu = \int f \mu^*$.

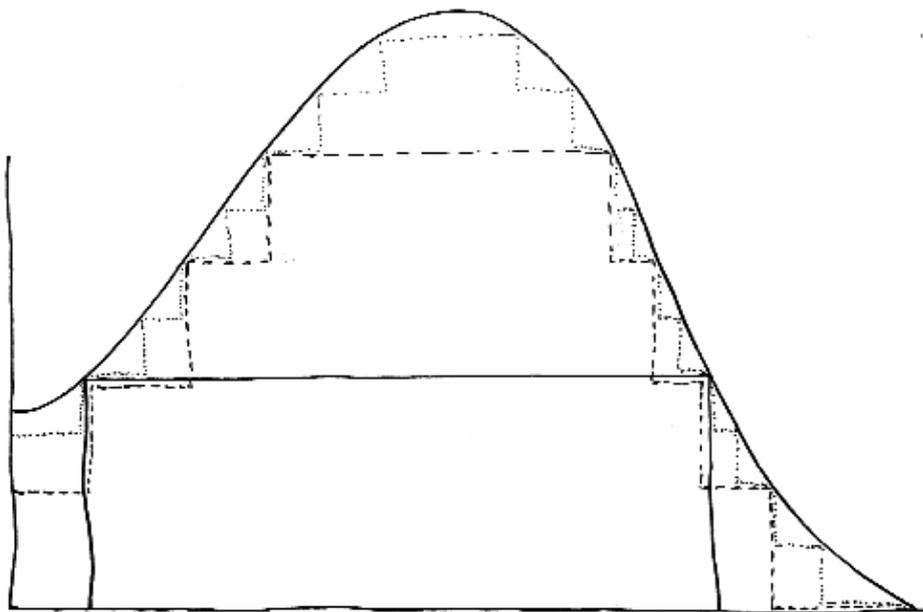
Übung 9.4.14. Sei (X, \mathcal{M}) ein Maßraum. Man zeige: Die Summe zweier Maße μ, ν auf \mathcal{M} ist wieder ein Maß $\mu + \nu$ auf \mathcal{M} und für jede meßbare Funktion $f : X \rightarrow [0, \infty]$ gilt $\int f(\mu + \nu) = \int f \mu + \int f \nu$.

9.4.15 (Restriktion von Maßen). Ist (X, \mathcal{M}, μ) ein Maßraum und $A \subset X$ eine meßbare Teilmenge, so erhalten wir in offensichtlicher Weise einen weiteren Maßraum $(A, \mathcal{M}|_A, \mu|_A)$. Meist kürzen wir die Restriktion $\mu|_A$ von unserem Maß auf die induzierte σ -Algebra $\mathcal{M}|_A = \{B \in \mathcal{M} \mid B \subset A\}$ mit μ ab. Integrale in Bezug auf diesen Maßraum notieren wir $\int_A f \mu$ oder auch $\int_A f$. Ist $f : X \rightarrow [0, \infty]$ meßbar, so haben wir offensichtlich

$$\int_A f = \int_X [A]f$$

wobei links das Integral der Restriktion von f auf A gemeint ist und rechts das Integral des Produkts von f mit der charakteristischen Funktion von A , gebildet mit der Konvention $0 \cdot \infty = 0$.

Ergänzung 9.4.16. Ich habe nicht durchdacht, ob für $j : A \hookrightarrow X$ eine injektive meßbare Abbildung von Maßräumen mit der Eigenschaft, daß die Bilder meßbarer Mengen wieder meßbar sind, die Notation $j^! \mu$ für das mittels j eingeschränkte Maß sinnvoll sein könnte. Es ist ja schon so, daß derartige Abbildungen das meßbare Analogon von étalen Abbildungen sind, und in kartesischen Diagrammen in der Kategorie der Maßräume gilt durchaus $\psi_* \varphi^! \mu = \varphi^! \psi_* \mu$.



Die ersten Glieder unserer monotonen Folge von nichtnegativen Stufenfunktionen, die punktweise gegen eine gegebene nichtnegative Funktion konvergieren.

Übung 9.4.17. Gegeben eine meßbare nichtnegative Funktion g auf einem Maßraum (X, \mathcal{M}, μ) mit $\int_X g \mu < \infty$ gibt es für jedes $\varepsilon > 0$ ein $\alpha = \alpha_\varepsilon > 0$ derart, daß für alle $A \in \mathcal{M}$ gilt

$$\mu(A) < \alpha \Rightarrow \int_A g \mu < \varepsilon$$

Hinweis: Es gibt sicher eine meßbare Stufenfunktion $h : X \rightarrow [0, \infty)$ mit $h \leq g$ und $\int g \mu \leq \int h \mu + \varepsilon/2$.

Übung 9.4.18. Gegeben eine meßbare nichtnegative Funktion g auf einem Maßraum (X, \mathcal{M}, μ) mit $\int_X g \mu < \infty$ gibt es für jedes $\varepsilon > 0$ ein $\beta = \beta_\varepsilon > 0$ mit

$$\int \inf(g, \beta) \mu < \varepsilon$$

Hinweis: Es gibt sicher eine meßbare Stufenfunktion $h : X \rightarrow [0, \infty)$ mit $h \leq g$ und $\int g \mu \leq \int h \mu + \varepsilon/2$.

Übung 9.4.19. Man zeige: Gegeben ein Maßraum X und eine Folge (f_n) nichtnegativer meßbarer Funktionen $f_n : X \rightarrow [0, \infty]$ ist auch ihre Summe $\sum f_n$ meßbar und es gilt

$$\int \sum f_n = \sum \int f_n$$

Übung 9.4.20 (Produkte von Maßen mit Funktionen). Ist (X, μ) ein Maßraum und $g : X \rightarrow [0, \infty]$ meßbar, so erhalten wir ein neues Maß $g\mu$ auf X durch die Vorschrift $(g\mu)(A) = \int_A g \mu$ und für jede weitere meßbare Funktion $f : X \rightarrow [0, \infty]$ gilt mit der Konvention $0 \cdot \infty = 0 = \infty \cdot 0$ die Identität von Maßen $f(g\mu) = (fg)\mu$. Ist ähnlich $f : X \rightarrow \mathbb{R}$ gegeben, so ist f integrierbar in Bezug auf $(g\mu)$ genau dann, wenn unter der Konvention $0 \cdot \infty = 0$ die Funktion fg integrierbar ist in Bezug auf μ , und unter diesen Voraussetzungen gilt

$$\int (fg) \mu = \int f (g\mu)$$

Übung 9.4.21. Ist (X, μ) ein σ -endlicher Maßraum und sind $f, g : X \rightarrow [0, \infty]$ meßbar, so gilt die Gleichheit von Maßen $f\mu = g\mu$ genau dann, wenn f und g außerhalb einer meßbaren Menge vom Maß Null übereinstimmen. Man gebe auch ein Gegenbeispiel im Fall nicht σ -endlicher Maßräume. Hinweis: Man ziehe sich auf den Fall $\mu(X) < \infty$ zurück und betrachte dann zunächst die Mengen $\{x \mid n > f(x) > g(x) + 1/n\}$.

Ergänzung 9.4.22. Sei (X, \mathcal{M}, μ) ein Maßraum. Für jede μ -meßbare Funktion $f : X \rightarrow \overline{\mathbb{R}}$ im Sinne von 9.3.25 gibt es eine meßbare Funktion $\tilde{f} : X \rightarrow \overline{\mathbb{R}}$, die

außerhalb einer μ -Nullmenge mit f übereinstimmt. In der Tat können wir f nach 9.4.12 schreiben als punktweisen Grenzwert einer Folge von μ -meßbaren Stufenfunktionen $f(x) = \lim s_n(x)$. Dann verkleinern wir die Grundflächen aller Stufen von Null verschiedener Höhe zu meßbaren Mengen so, daß sich das Maß der Stufen nicht ändert, und erhalten eine Folge von meßbaren Stufenfunktionen \tilde{s}_n , die außerhalb einer meßbaren Nullmenge A punktweise gegen f konvergiert, und betrachten den punktweisen Grenzwert $\tilde{f}(x) = \lim_{n \rightarrow \infty} [A](x) \tilde{s}_n(x)$.

9.5 Integrierbare Funktionen und ihr Integral

Definition 9.5.1 (Integral einer integrierbaren Funktion). Gegeben ein Maßraum $X = (X, \mathcal{M}, \mu)$ heißt eine Funktion $f : X \rightarrow \mathbb{R}$ **integrierbar** oder genauer **meßbar und integrierbar** genau dann, wenn sie meßbar ist und wenn zusätzlich gilt $\int |f| < \infty$ im Sinne des in 9.4.6 erklärten Integrals nichtnegativer Funktionen. Wir definieren das **Integral** $\int f \in \mathbb{R}$ einer integrierbaren Funktion $f : X \rightarrow \mathbb{R}$ durch die Vorschrift

$$\int f = \int f^+ - \int f^-$$

wobei $f^+, f^- : X \rightarrow [0, \infty)$ den **positiven** bzw. **negativen Anteil** von f bezeichnen, der gegeben wird durch $f^\pm(x) = \sup(\pm f(x), 0)$. Die Menge aller integrierbaren Funktionen $f : X \rightarrow \mathbb{R}$ notieren wir je nach der gewünschten Präzision $\mathcal{L}_{\mathbb{R}}^1(X) = \mathcal{L}_{\mathbb{R}}^1(X; \mu) = \mathcal{L}_{\mathbb{R}}^1(X; \mathcal{M}, \mu)$.

9.5.2. Wir haben nun genau genommen zwei Integrale definiert: Erst ein Integral für meßbare nichtnegative Funktionen mit Werten in $[0, \infty]$, das Werte in $[0, \infty]$ annimmt, und dann ein Integral für integrierbare reellwertige Funktionen, das reelle Werte annimmt. Offensichtlich stimmen im Fall einer nichtnegativen reellwertigen integrierbaren Funktion diese beiden Integrale überein. Auf dem Schnitt ihrer Definitionsbereiche liefern unsere beiden Varianten des Integralbegriffs in anderen Worten dasselbe. Es ist deshalb sinnvoll, für beide Konzepte dasselbe Symbol zu verwenden. Es gilt jedoch zu beachten, daß man einer beliebigen meßbaren reellwertigen Funktion im Allgemeinen nicht mehr sinnvoll ein Integral zuordnen kann: Das gelingt nur bei meßbaren Funktionen mit nichtnegativen Werten und wenn man ∞ als Wert des Integrals zuläßt. Wie wir gesehen haben, gelingt dann ja das Integrieren sogar allgemeiner für alle meßbaren Funktionen mit Werten in $[0, \infty]$. Man kann hier sogar noch ein wenig mehr herauskratzen und analog wie oben auch meßbaren Funktionen mit Werten in $[-\infty, \infty]$ sinnvoll ein Integral in $(-\infty, \infty]$ zuordnen, wenn nur ihr Negativteil f_- integrierbar ist, aber in dieser Allgemeinheit werde ich das Integral nie verwenden.

9.5.3. Allgemeiner mag man auch solche Funktionen $f : X \rightarrow \mathbb{R}$ noch “integrierbar” nennen wollen, die im Sinne der vorhergehenden Definition meßbar und

integrierbar sind in Bezug auf den vervollständigten Maßraum. Diese Terminologie entspricht vielleicht noch besser dem Sprachempfinden, führt jedoch leicht zu technischen Verkrampfungen. In den Fällen, in denen die Integrierbarkeit in Bezug auf den vervollständigten Maßraum gemeint ist, werde ich stets gesondert darauf hinweisen.

Übung 9.5.4. Auf einem topologischen Raum mit einem Borelmaß ist jede stetige reellwertige Funktion mit kompaktem Träger integrierbar.

Ergänzung 9.5.5. Gegeben eine reellwertige integrierbare Zufallsvariable X auf einem Wahrscheinlichkeitsraum (Ω, P) heißt ihr Integral auch der **Erwartungswert** der Zufallsvariable und wird $E(X) := \int_{\Omega} X(\omega)P(d\omega)$ notiert.

Beispiel 9.5.6 (Integrierbarkeit und absolute Konvergenz). Ist I eine abzählbare Menge, so ist eine Funktion $f : I \rightarrow \mathbb{R}$ integrierbar für das Zählmaß genau dann, wenn für eine oder gleichbedeutend jede "Abzählung" von I die Reihe $\sum_{i \in I} f(i)$ absolut konvergiert. Das Integral unserer Funktion ist in diesem Fall genau der Grenzwert der Reihe. Ist I eine beliebige Menge, so ist eine Funktion $f : I \rightarrow \mathbb{R}$ integrierbar für das Zählmaß genau dann, wenn die Familie der $f(i)$ summierbar ist im Sinne von 5.5.24, was ja nach 5.5.26 auch im Wesentlichen absolute Konvergenz bedeutet.

Übung 9.5.7. Man zeige, daß für jede integrierbare Funktion die Menge der Punkte, auf denen sie nicht den Wert Null annimmt, σ -endlich sein muß. Lösung: ??.

9.5.8. Aus der Definition erhalten wir für f, g integrierbar sofort $|\int f| \leq \int |f|$ und $f \leq g \Rightarrow \int f \leq \int g$.

Satz 9.5.9 (Linearität des Integrals). Sei (X, \mathcal{M}, μ) ein Maßraum. Der Raum $\mathcal{L}_{\mathbb{R}}^1(X)$ aller integrierbaren Funktionen ist ein Untervektorraum im Raum aller Funktionen $X \rightarrow \mathbb{R}$, und das Integral ist eine lineare Abbildung

$$\int : \mathcal{L}_{\mathbb{R}}^1(X) \rightarrow \mathbb{R}$$

Beweis. Wir überlassen den Nachweis der ersten Aussage dem Leser und zeigen nur die Linearität des Integrals. Zunächst zeigen wir die Additivität

$$\int f + g = \int f + \int g$$

Seien $f = f^+ - f^-$, $g = g^+ - g^-$ und $f + g = h = h^+ - h^-$ die Zerlegungen in den positiven und negativen Anteil. Wir folgern durch Einsetzen $f^+ + g^+ + h^- = f^- + g^- + h^+$ und mit 9.4.11 ergibt sich $\int f^+ + \int g^+ + \int h^- = \int f^- + \int g^- + \int h^+$, woraus mit der Definition dann wieder $\int f + \int g = \int f + g$ folgt. Nun zeigen wir noch die Verträglichkeit mit der Multiplikation mit Skalaren

$$\int cf = c \int f$$

Für $c = -1$ folgt das aus den Definitionen, für $c \geq 0$ folgt es aus 9.4.11, und der allgemeine Fall ergibt sich aus diesen beiden Spezialfällen. \square

Satz 9.5.10 (über dominierte Konvergenz). Sei (X, \mathcal{M}, μ) ein Maßraum und $f_n : X \rightarrow \mathbb{R}$ eine Folge meßbarer Funktionen, die punktweise gegen eine Funktion $f : X \rightarrow \mathbb{R}$ konvergiert. Gibt es eine integrierbare Funktion $g : X \rightarrow \mathbb{R}$ mit $|f_n| \leq g$ für alle n , so sind alle f_n und auch f integrierbar und es gilt

$$\int f = \lim_{n \rightarrow \infty} \int f_n$$

Ergänzung 9.5.11. Verschärfungen dieses Satzes zeigen wir in ??, dort fordern wir statt punktweiser Konvergenz nur “stochastische Konvergenz”, und noch weitergehend in ??, dort fordern wir außerdem statt der Dominiertheit nur noch die “gleichgradige Integrierbarkeit”.

Ergänzung 9.5.12. Eine eher unwesentliche Verallgemeinerung erhält man, wenn man allgemeiner nur eine Domination der Konvergenz durch eine meßbare Funktion $g : X \rightarrow [0, \infty]$ mit $\int g < \infty$ voraussetzt: Aus dieser Annahme folgt nämlich, daß g außerhalb einer Nullmenge doch wieder reelle Werte annehmen muß, und schwupps finden wir uns im bereits behandelten Fall wieder.

Beweis. Aus unseren Annahmen folgt $\int |f_n| \leq \int g < \infty$, also sind die f_n integrierbar. Weiter ist f auch meßbar als punktweiser Grenzwert meßbarer Funktionen und dann ist mit demselben Argument auch f integrierbar. Um die Vertauschbarkeit des Grenzwerts mit dem Integral zu zeigen, betrachten wir nun die Funktionenfolgen

$$\begin{aligned} i_n(x) &= \inf\{f_n(x), f_{n+1}(x), \dots\} \\ s_n(x) &= \sup\{f_n(x), f_{n+1}(x), \dots\} \end{aligned}$$

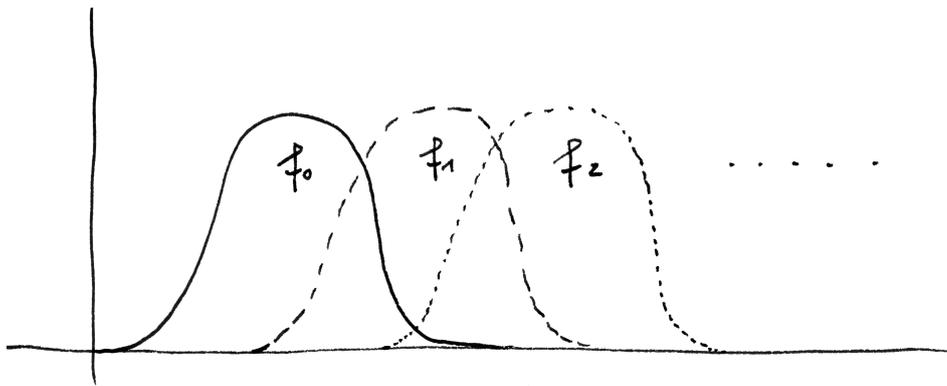
Sie bestehen aus meßbaren Funktionen, beide Folgen konvergieren punktweise gegen f , und es gilt

$$-g \leq i_0 \leq i_1 \leq \dots \leq f \leq \dots \leq s_1 \leq s_0 \leq g$$

Mit dem Satz über monotone Konvergenz erhalten wir also $\lim_{n \rightarrow \infty} \int g + i_n = \int g + f$ und $\lim_{n \rightarrow \infty} \int g - s_n = \int g - f$, also

$$\lim_{n \rightarrow \infty} \int i_n = \int f = \lim_{n \rightarrow \infty} \int s_n$$

Da aber nach Definition gilt $i_n \leq f_n \leq s_n$, folgt die Behauptung aus dem Quetschlemma 5.1.31. \square



Eine Folge integrierbarer Funktionen auf der reellen Zahlengeraden, die zwar punktweise gegen die Nullfunktion konvergiert, deren Integrale jedoch keine Nullfolge bilden. In diesem Fall können wir auch offensichtlich keine alle Funktionen unserer Folge dominierende integrierbare Funktion g finden.

Korollar 9.5.13. *Jede stetige reellwertige Funktion auf einem kompakten reellen Intervall ist integrierbar im Sinne der vorhergehenden Definition und ihr Riemann-Integral stimmt mit ihrem Lebesgue-Integral überein.*

Beweis. Jede stetige reellwertige Funktion f auf einem kompakten reellen Intervall $[a, b]$ ist meßbar und beschränkt. Aus $|f| \leq M$ folgt dann sofort $\int |f| \leq M(b - a) < \infty$ und damit die Integrierbarkeit von f . Bilden wir Stufenfunktionen f_r , indem wir $[a, b]$ äquidistant unterteilen durch Zwischenpunkte $a = a_0 < a_1 < \dots < a_r = b$ und f_r auf $[a_{i-1}, a_i)$ konstant den Wert $f(a_i)$ geben und bei b den Wert $f(b)$, so konvergieren die f_r wegen der gleichmäßigen Stetigkeit von f punktweise gegen f . Andererseits sind ihre Integrale offensichtlich genau unsere Riemannsummen $S^r(f)$ aus 6.5.5, in Formeln $S^r(f) = \int f_r$, und für $r \rightarrow \infty$ strebt die linke Seite nach 6.5.5 gegen das Riemannintegral und die Rechte nach dem Satz über dominierte Konvergenz 9.5.10 gegen das Lebesgueintegral von f . \square

Übung 9.5.14 (Vertauschen von Integration und Ableitung). Sei (X, μ) ein Maßraum und $I \subset \mathbb{R}$ halboffen und $f : X \times I \rightarrow \mathbb{R}$ eine Abbildung derart, daß $x \mapsto f(x, t)$ integrierbar ist für alle $t \in I$ und $t \mapsto f(x, t)$ differenzierbar für alle $x \in X$. Existiert eine integrierbare Abbildung $g : X \rightarrow \mathbb{R}$ mit $g(x) \geq |\partial_t f(x, t)|$ für alle x und t , so ist $x \mapsto \partial_t f(x, t)$ integrierbar für alle t und es gilt

$$\partial_t \int f(x, t) \mu\langle x \rangle = \int \partial_t f(x, t) \mu\langle x \rangle$$

Hinweis: Dominierte Konvergenz 9.5.10 und Mittelwertsatz.

Übung 9.5.15 (Integrale unter Bildmaßen). Man zeige, daß das Integrieren Verwandtschaft respektiert. Sind genauer $\phi : X \rightarrow Y$ eine meßbare Abbildung von Meßräumen, μ, ν verwandte Maße, in Formeln $\phi : \mu \rightsquigarrow \nu$, und g, f verwandte meßbare Funktionen nach $[0, \infty]$, in Formeln $\phi : g \rightsquigarrow f$, so gilt

$$\int_X g(x) \mu\langle x \rangle = \int_Y f(y) \nu\langle y \rangle$$

Wir können dieselbe Aussage auch wie folgt formulieren: Gegeben Meßräume X, Y , eine meßbare Abbildung $\phi : X \rightarrow Y$, eine meßbare Abbildung $f : Y \rightarrow [0, \infty]$ und ein Maß μ auf X gilt

$$\int_X f(\phi(x)) \mu\langle x \rangle = \int_Y f(y) (\phi_*\mu)\langle y \rangle$$

Weiter ist eine reellwertige meßbare Funktion $f : Y \rightarrow \mathbb{R}$ integrierbar in Bezug auf $\phi_*\mu$ genau dann, wenn $f \circ \phi$ integrierbar ist in Bezug auf μ , und unter dieser Voraussetzung gilt dieselbe Gleichung in \mathbb{R} .

Übung 9.5.16 (Satz von Beppo Levi). Sei f_n eine monoton wachsende Folge integrierbarer Funktionen. Ist die Folge ihrer Integrale beschränkt, so ist die Menge N aller $x \in X$ mit $\lim_{n \rightarrow \infty} f_n(x) = \infty$ meßbar vom Maß Null und die Funktion $f = \lim_{n \rightarrow \infty} f_n : (X \setminus N) \rightarrow \mathbb{R}$ ist integrierbar mit Integral $\int f = \lim_{n \rightarrow \infty} \int f_n$.

Übung 9.5.17. Sei $Q = [a, b] \times [c, d] \subset \mathbb{R}^2$ ein nichtleerer kompakter zweidimensionaler Quader und $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ eine stetige Funktion mit Träger in Q . Sei weiter μ ein Borelmaß auf \mathbb{R}^2 . Für $r \geq 1$ definieren wir dann die r -te **Riemannsumme** $S^r(f; \mu)$ von f wie folgt: Wir betrachten die äquidistanten Unterteilungen

$$a = a_0 \leq a_1 \leq \dots \leq a_r = b$$

$$c = c_0 \leq c_1 \leq \dots \leq c_r = d$$

der Kanten unseres Rechtecks, erhalten r^2 klitzekleine halboffene Rechtecke $Q_{i,j}^{\perp} = [a_i, a_{i+1}) \times [c_j, c_{j+1})$ und setzen

$$S^r(f; \mu) = \sum_{i,j=0}^{r-1} f(a_i, c_j) \mu(Q_{i,j}^{\perp})$$

Man zeige, daß unter unseren Annahmen diese Riemannsummen gegen das Integral streben, in Formeln

$$\int_Q f \mu = \lim_{r \rightarrow \infty} S^r(f; \mu)$$

Übung 9.5.18. Ist (X, μ) ein Maßraum und $g : X \rightarrow [0, \infty]$ meßbar und $g\mu$ das in 9.4.20 konstruierte Maß, so zeige man für $f : X \rightarrow \mathbb{R}$ meßbar, daß f integrierbar ist in Bezug auf $(g\mu)$ genau dann, wenn unter der Konvention $0 \cdot \infty = 0$ die Funktion fg integrierbar ist in Bezug auf μ , und daß unter diesen Voraussetzungen gilt

$$\int (fg) \mu = \int f (g\mu)$$

9.6 Integration auf Produkträumen

Satz 9.6.1 (Produktmaß). Gegeben σ -endliche Maßräume (X, \mathcal{M}, μ) und (Y, \mathcal{N}, ν) gibt es auf der von allen Produkten $A \times B$ mit $A \in \mathcal{M}$ und $B \in \mathcal{N}$ erzeugten **Produkt- σ -Algebra** $\mathcal{M} \boxtimes \mathcal{N} \subset \mathcal{P}(X \times Y)$ genau ein Maß $\mu \boxtimes \nu$ derart, daß für alle $A \in \mathcal{M}$ und $B \in \mathcal{N}$ gilt

$$(\mu \boxtimes \nu)(A \times B) = \mu(A)\nu(B)$$

9.6.2. Hier verwenden wir unsere Konvention $0 \cdot \infty = \infty \cdot 0 = 0$. Sind unsere Maßräume nicht σ -endlich, so wird das Produktmaß durch die angegebene Bedingung noch nicht eindeutig festgelegt. Eine mögliche Definition in dieser Allgemeinheit geben wir in ???. In der Literatur wird für die Produkt- σ -Algebra und Produktmaße meist das Symbol \otimes verwendet, aber mir gefällt das Symbol \boxtimes hier besser, da es sich beim Produkt von Maßen eher um eine Art “externes Produkt” und jedenfalls nicht um ein Tensorprodukt handelt.

Übung 9.6.3. Gegeben metrische Räume X, Y mit abzählbaren dichten Teilmengen fällt das Produkt der σ -Algebren der jeweils topologisch meßbaren Mengen zusammen mit der σ -Algebra der topologisch meßbaren Mengen des Produkts, in Formeln

$$\text{Borel}(X) \boxtimes \text{Borel}(Y) = \text{Borel}(X \times Y)$$

Für den Beweis mag man sich an 9.3.16 orientieren, das vom logischen Aufbau der Theorie her eigentlich besser als Korollar der in der vorstehenden Formel enthaltenen Aussage bewiesen worden wäre. Genauer zeigt man \supset wie dort und verwendet für \subset die Borel-Meßbarkeit der Projektionen. Leser mit Kenntnissen in Topologie werden unschwer erkennen, daß diese Formel auch allgemeiner für beliebige separable topologische Räume X und Y gilt.

Beweis. Die Gesamtheit aller endlichen disjunkten Vereinigungen von Quadern $A \times B$ mit $A \in \mathcal{M}$ und $B \in \mathcal{N}$ bildet sicher einen Mengenring \mathcal{E} , und in 9.2.4 haben Sie das sich auch zur Übung bereits überlegt. Wir zeigen zunächst, daß es auf \mathcal{E} ein Prämaß $\mu \times \nu$ gibt mit $(\mu \times \nu)(A \times B) = \mu(A)\nu(B)$. Es ist jedoch klar, daß für $C \in \mathcal{E}$ und beliebiges $y \in Y$ die Abbildung $x \mapsto [C](x, y)$ eine meßbare Stufenfunktion $X \rightarrow \{0, 1\}$ ist und daß wir weiter mit $y \mapsto \int_X [C](x, y) \mu\langle x \rangle$ eine meßbare Stufenfunktion $Y \rightarrow [0, \infty]$ erhalten. Wir können also definieren

$$(\mu \times \nu)(C) = \int_Y \left(\int_X [C](x, y) \mu\langle x \rangle \right) \nu\langle y \rangle$$

und die σ -Additivität von $\mu \times \nu$ folgt dann aus der Additivität der Integrale 9.4.11 zusammen mit dem Satz über monotone Konvergenz 9.4.9. Unser Satz zum Produktmaß folgt damit aus dem Maßfortsetzungssatz 9.2.10. \square

9.6.4. Mit diesem Satz können wir durch Induktion über n das Lebesgue-Maß auf dem \mathbb{R}^n aus dem Lebesgue-Maß auf \mathbb{R} konstruieren und so die Existenz in 9.1.20 im allgemeinen zeigen.

Übung 9.6.5. Gegeben meßbare Abbildungen $f : X \rightarrow X'$ und $g : Y \rightarrow Y'$ ist auch ihr Produkt $f \times g : X \times Y \rightarrow X' \times Y'$ meßbar. Sind zusätzlich μ ein σ -endliches Maß auf X und ν ein σ -endliches Maß auf Y mit σ -endlichen Bildmaßen, so ist das Bildmaß ihres Produkts das Produkt der Bildmaße, in Formeln

$$(f \times g)_*(\mu \boxtimes \nu) = (f_*\mu) \boxtimes (g_*\nu)$$

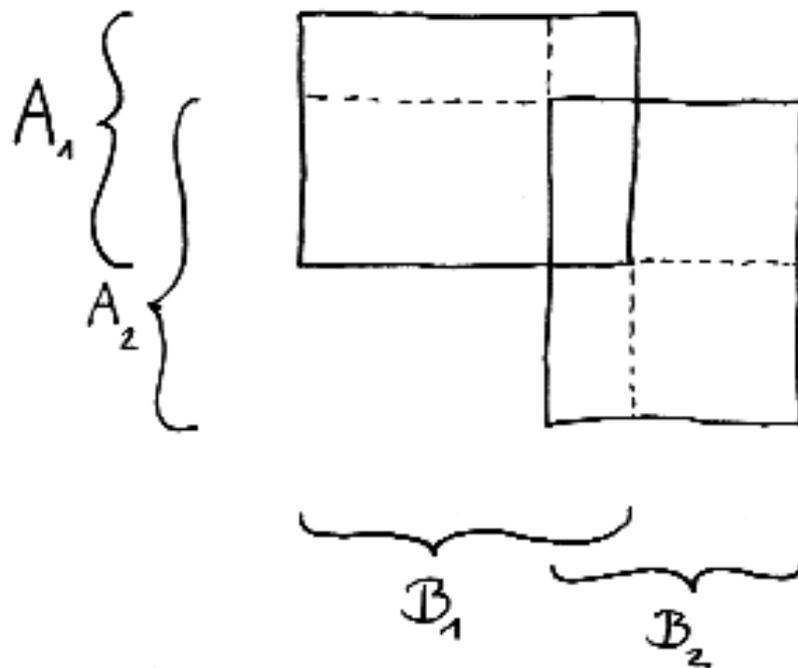


Illustration der Tatsache, daß die Vereinigung zweier Quader auch als die disjunkte Vereinigung von sieben Quadern geschrieben werden kann.

Übung 9.6.6. Man präzisiere und zeige die “Assoziativität von Produkten” bei Maßräumen.

Ergänzende Übung 9.6.7. Gegeben ein endlichdimensionaler Raum X induziert das Bilden des Produkts mit dem Lebesguemaß eine Bijektion zwischen der Menge aller Borelmaße auf X und der Menge aller Borelmaße auf $X \times \mathbb{R}$, die invariant sind unter Translation in der zweiten Komponente. Hinweis: Gegeben ein in dieser Weise translationsinvariantes Borelmaß μ auf $X \times \mathbb{R}$ beachte man, daß für $A \subset X$ mit kompaktem Abschluss die Vorschrift $B \mapsto \mu(A \times B)$ ein translationsinvariantes Borelmaß auf \mathbb{R} definiert.

Satz 9.6.8 (positiver Fubini). Gegeben σ -endliche Maßräume (X, μ) und (Y, ν) sowie eine meßbare Funktion $f : X \times Y \rightarrow [0, \infty]$ ist $x \mapsto f(x, y)$ für alle $y \in Y$ eine meßbare Funktion $X \rightarrow [0, \infty]$ und das partielle Integral $y \mapsto \int f(x, y) \mu \langle x \rangle$ ist eine meßbare Funktion $Y \rightarrow [0, \infty]$ und es gilt

$$\int_{X \times Y} f(x, y) (\mu \boxtimes \nu) \langle x, y \rangle = \int_Y \left(\int_X f(x, y) \mu \langle x \rangle \right) \nu \langle y \rangle$$

9.6.9. Unser Satz impliziert insbesondere, daß unter den gegebenen Voraussetzungen die partiellen Integrale vertauscht werden dürfen. Bezeichnet in der Tat $\tau : X \times Y \rightarrow Y \times X$ das Vertauschen, so haben wir die offensichtliche Verwandtschaft von Maßen $\tau : \mu \boxtimes \nu \rightsquigarrow \nu \boxtimes \mu$ und die ebenso offensichtliche Verwandtschaft von Funktionen $\tau : f \rightsquigarrow \tilde{f}$ mit $\tilde{f}(y, x) = f(x, y)$ und damit unmittelbar und formal nach 9.5.15 die Gleichheit

$$\int_{X \times Y} f(x, y) (\mu \boxtimes \nu) \langle x, y \rangle = \int_{Y \times X} \tilde{f}(y, x) (\nu \boxtimes \mu) \langle y, x \rangle$$

9.6.10. Dieser Satz und verschiedene seiner Varianten werden auch oft als **Satz von Tonelli** zitiert. Daß die partiellen Integrale bei nicht notwendig σ -endlichen Maßräumen im allgemeinen nicht mehr vertauscht werden dürfen, zeigt das folgende Beispiel: Sei $X = Y = [0, 1]$ versehen mit der σ -Algebra \mathcal{B} der topologisch meßbaren Mengen und dem Lebesgue-Maß λ beziehungsweise dem Zählmaß ζ . Die Diagonale Δ ist dann meßbar, für ihre charakteristische Funktion $[\Delta]$ gilt jedoch

$$\int_Y \left(\int_X [\Delta](x, y) \lambda \langle x \rangle \right) \zeta \langle y \rangle = 0 \neq 1 = \int_X \left(\int_Y [\Delta](x, y) \zeta \langle y \rangle \right) \lambda \langle x \rangle$$

Beweis. Für jedes $y \in Y$ ist die y -Horizontale $X \rightarrow X \times Y, x \mapsto (x, y)$ meßbar nach Lemma 9.3.9, da die Urbilder von Erzeugern der σ -Algebra der meßbaren Mengen des Produkts meßbar sind. Also ist auch $x \mapsto f(x, y)$ meßbar auf X als die Verknüpfung von f mit der y -Horizontalen. Um die anderen Aussagen

des Satzes zu zeigen, müssen wir weiter ausholen. Zunächst einmal dürfen wir annehmen, daß X und Y endliches Maß haben: Sonst schreiben wir X bzw. Y als aufsteigende Vereinigungen von Teilmengen X_n bzw. Y_m endlichen Maßes und erhalten die Meßbarkeit des partiellen Integrals über X_n und

$$\int_{X_n \times Y_m} f(x, y) (\mu \boxtimes \nu)\langle x, y \rangle = \int_{Y_m} \left(\int_{X_n} f(x, y) \mu\langle x \rangle \right) \nu\langle y \rangle$$

Im Grenzwert $n \rightarrow \infty$ ergibt sich dann auf der linken Seite nach dem Satz 9.4.9 über monotone Konvergenz $\int_{X \times Y_m} f(x, y) (\mu \boxtimes \nu)\langle x, y \rangle$, und auf der rechten streben die meßbaren Funktionen $y \mapsto \int_{X_n} f(x, y) \mu\langle x \rangle$ punktweise monoton gegen $y \mapsto \int_X f(x, y) \mu\langle x \rangle$. Mithin ist diese Funktion auch meßbar und hat das Integral

$$\int_{Y_m} \left(\int_X f(x, y) \mu\langle x \rangle \right) \nu\langle y \rangle = \lim_{n \rightarrow \infty} \int_{Y_m} \left(\int_{X_n} f(x, y) \mu\langle x \rangle \right) \nu\langle y \rangle$$

Bilden wir dann schließlich den Grenzwert für $m \rightarrow \infty$, so folgt wie gewünscht $\int_{X \times Y} f = \int_Y \int_X f$. Wir dürfen also ohne Beschränkung der Allgemeinheit voraussetzen, daß X und Y endliches Maß haben. Wir zeigen nun den Satz zunächst für Funktionen der Gestalt $f = [C]$ mit $C \in \mathcal{M} \boxtimes \mathcal{N}$. Dazu brauchen wir einen neuen Begriff.

Definition 9.6.11. Sei Z eine Menge. Ein System $\mathcal{A} \subset \mathcal{P}(Z)$ von Teilmengen von Z heißt **monoton** genau dann, wenn die beiden folgenden Aussagen gelten:

1. Liegen $A_0 \subset A_1 \subset A_2 \subset \dots$ alle in \mathcal{A} , so auch $\bigcup A_n$;
2. Liegen $B_0 \supset B_1 \supset B_2 \supset \dots$ alle in \mathcal{A} , so auch $\bigcap B_n$.

Lemma 9.6.12. Ist Z eine Menge und $\mathcal{R} \subset \mathcal{P}(Z)$ eine Mengenalgebra, so kann die von \mathcal{R} erzeugte σ -Algebra $\sigma(\mathcal{R})$ auch beschrieben werden als das kleinste monotone System $\mathcal{A} \subset \mathcal{P}(Z)$, das \mathcal{R} umfaßt, alias als der Schnitt aller monotonen Systeme, die \mathcal{R} umfassen.

Beweis. Offensichtlich gilt $\mathcal{R} \subset \mathcal{A} \subset \sigma(\mathcal{R})$. Wir müssen also nur zeigen, daß \mathcal{A} eine σ -Algebra ist. Dazu reicht es nach 9.1.30 zu zeigen, daß \mathcal{A} eine Mengenalgebra ist. Zunächst zeigen wir, daß \mathcal{A} stabil ist unter dem Bilden von Komplementen. Mit \mathcal{A} ist jedoch auch $\mathcal{A}^c := \{A \in \mathcal{A} \mid A^c \in \mathcal{A}\}$ ein monotones System und wegen $\mathcal{R} \subset \mathcal{A}^c$ folgt $\mathcal{A}^c = \mathcal{A}$ aus der Minimalität von \mathcal{A} . Ähnlich ist für jede Teilmenge $Y \subset Z$ mit \mathcal{A} auch $\mathcal{A}_Y := \{A \in \mathcal{A} \mid A \cap Y \in \mathcal{A}\}$ ein monotones System. Für $Y \in \mathcal{R}$ gilt natürlich $\mathcal{R} \subset \mathcal{A}_Y$ und daraus folgt mit der Minimalität von \mathcal{A} sofort $\mathcal{A} = \mathcal{A}_Y$. Damit haben wir gezeigt:

$$A \in \mathcal{A} \text{ und } Y \in \mathcal{R} \Rightarrow A \cap Y \in \mathcal{A}$$

Mit dieser Erkenntnis lassen wir nun dasselbe Argument noch einmal laufen: Für $X \in \mathcal{A}$ wissen wir damit nämlich, daß gilt $\mathcal{R} \subset \mathcal{A}_X$, und daraus folgt $\mathcal{A} \subset \mathcal{A}_X$, also $\mathcal{A} = \mathcal{A}_X$. Damit haben wir gezeigt

$$A \in \mathcal{A} \text{ und } X \in \mathcal{A} \Rightarrow A \cap X \in \mathcal{A}$$

Also ist \mathcal{A} eine Mengenalgebra, und da es auch stabil ist unter abzählbaren aufsteigenden Vereinigungen, ist es dann wie bereits erwähnt sogar eine σ -Algebra. \square

Mit diesem Lemma können wir nun den Satz im Fall $f = [C]$ zeigen: Da wir uns nämlich bereits auf den Fall $\mu(X) < \infty, \nu(Y) < \infty$ zurückgezogen haben, ist die konstante Funktion 1 integrierbar auf $X \times Y$, und natürlich dominiert diese Funktion die charakteristischen Funktionen aller Teilmengen von $X \times Y$. Aus dem Satz über dominierte Konvergenz 9.5.10 folgt also, daß das System aller $C \in \mathcal{M} \boxtimes \mathcal{N}$, für deren charakteristische Funktion der Satz gilt, ein monotones System ist. Dies monotone System enthält aber offensichtlich alle $C \in \mathcal{M} \times \mathcal{N}$, mithin besteht es nach dem Lemma aus allen meßbaren Mengen $C \in \mathcal{M} \boxtimes \mathcal{N}$. Damit ist der Satz nun auch für meßbare Stufenfunktion $f : X \times Y \rightarrow [0, \infty)$ klar. Für beliebiges meßbares f folgert man die Aussage dann, indem man f mithilfe von 9.4.12 als punktweisen Grenzwert einer monoton wachsenden Folge meßbarer Stufenfunktionen schreibt und beachtet, daß nach dem Satz über monotone Konvergenz 9.4.9 auf beiden Seiten Integral und Grenzwert vertauscht werden dürfen. \square

9.6.13. Nimmt man im vorherigen Satz für f die charakteristische Funktion einer σ -endlichen meßbaren Menge $C \in \mathcal{M} \boxtimes \mathcal{N}$, so ergibt sich, daß für $C_x = i_x^{-1}(C)$ das Urbild von C unter der x -Vertikalen $i_x : Y \rightarrow X \times Y, y \mapsto (x, y)$ die Abbildung $X \rightarrow [0, \infty], x \mapsto \mu(C_x)$ meßbar ist und daß gilt

$$(\mu \boxtimes \nu)(C) = \int_X \nu(C_x) \mu \langle x \rangle$$

Kippen wir das in unserer Vorstellung, so folgt insbesondere das **Prinzip von Cavalieri**, nach dem zwei meßbare Mengen $C, D \subset \mathbb{R}^3$ dasselbe Volumen haben, wenn ihre horizontalen Schnitte in jeder Höhe dieselbe Fläche haben. Weiter können wir so beweisen, daß das Integral einer nichtnegativen meßbaren Funktion auf \mathbb{R} tatsächlich die zwischen ihrem Graphen und der x -Achse eingeschlossene Fläche ist, wie im folgenden Korollar ausgeführt wird.

Korollar 9.6.14. Sei (X, μ) ein σ -endlicher Maßraum und $f : X \rightarrow [0, \infty]$ eine meßbare Funktion. So ist auch die Menge $F \subset X \times \mathbb{R}$ gegeben als $F = \{(x, y) \mid 0 \leq y < f(x)\}$ meßbar und es gilt $(\mu \boxtimes \lambda)(F) = \int_X f \mu$.

Beweis. Um zu zeigen, daß F meßbar ist, schreiben wir F als abzählbare Vereinigung von Quadraten $F = \bigcup_{q \in \mathbb{Q}_{>0}} f^{-1}([q, \infty]) \times [0, q)$. Die Formel für das Volumen von F folgt sofort aus 9.6.13. \square

Bemerkung 9.6.15. Man findet bei Lebesgue und auch in vielen Texten die Bemerkung, das Lebesgue-Integral unterscheide sich vom Riemannintegral dadurch, daß die Fläche unter dem Graphen der Funktion in horizontale statt in vertikale Streifen aufgeschnitten werde, deren Flächen man dann addiert. Ich kann das nur bedingt nachvollziehen, nach Cavalieri liefert ja beides dasselbe Integral. Der wesentliche Schritt ist meines Erachtens vielmehr der Übergang vom Messen reeller Intervalle zum Messen beliebiger “meßbarer Mengen”. Ich gebe aber zu, daß die horizontalen Streifen im Gegensatz zu den vertikalen Streifen eben keine Intervalle und dadurch näher an allgemeinen meßbaren Mengen sind.

Beispiel 9.6.16. Es folgt sofort, daß für $k < n$ die Teilmenge $\mathbb{R}^k \subset \mathbb{R}^n$ eine Nullmenge ist. Es folgt weiter mit 10.6.10, daß das Lebesgue-Maß einer Kreisscheibe D vom Radius r in der Tat gegeben wird durch $\lambda^2(D) = \pi r^2$. Ist schließlich $I \subset \mathbb{R}$ ein Intervall oder allgemeiner eine meßbare Teilmenge und $f : I \rightarrow [0, \infty]$ stetig oder allgemeiner meßbar, so folgt für das **Volumen des Rotationskörpers** $R = \{(x, y, z) \in \mathbb{R}^3 \mid z \in I, x^2 + y^2 \leq f(z)^2\}$ die Formel

$$\lambda^3(R) = \pi \int_I f(z)^2 dz$$

Übung 9.6.17. Man zeige, daß die Einheitskugel das Volumen $4\pi/3$ hat.

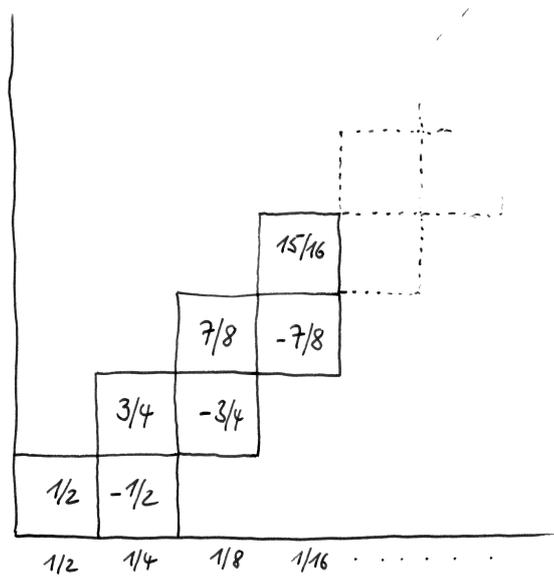
Satz 9.6.18 (Fubini). Gegeben σ -endliche Maßräume (X, μ) und (Y, ν) und eine integrierbare Funktion $f : X \times Y \rightarrow \mathbb{R}$ ist die Menge N aller $y \in Y$, für die $x \mapsto f(x, y)$ nicht integrierbar ist, meßbar vom Maß $\nu(N) = 0$, und die Funktion $Y \setminus N \rightarrow \mathbb{R}, y \mapsto \int_X f(x, y) \mu\langle x \rangle$ ist integrierbar mit Integral

$$\int_{Y \setminus N} \left(\int_X f(x, y) \mu\langle x \rangle \right) \nu\langle y \rangle = \int_{X \times Y} f(x, y) (\mu \boxtimes \nu)\langle x, y \rangle$$

9.6.19. Will man diesen Satz in der Praxis anwenden, so wird man in der Regel zuerst den positiven Fubini 9.6.8 benutzen, um die Integrierbarkeit von f nachzuweisen.

Beispiel 9.6.20. Die Funktion $f : \mathbb{R}^2 \rightarrow \mathbb{R}$, die außerhalb der x -Achse verschwindet und auf der x -Achse bei $(x, 0)$ jeweils den Wert x annimmt, ist integrierbar auf \mathbb{R}^2 . Jedoch ist $x \mapsto f(x, y)$ nur integrierbar für $y \neq 0$.

Beispiel 9.6.21. Unser Satz sagt insbesondere, daß wir unter gewissen Umständen “die Integrationsreihenfolge vertauschen dürfen”. Das folgende Beispiel zeigt, welche Probleme beim Vertauschen der Integrationsreihenfolge im allgemeinen auftreten können. Sei ζ das Zählmaß auf \mathbb{N} . Wir finden induktiv eine Funktion $f :$



Eine meßbare Funktion auf \mathbb{R}^2 wie in 9.6.21 derart, daß die partiellen Integrale existieren und selbst wieder integrierbar sind, das Endresultat jedoch von der Integrationsreihenfolge abhängt. Der “positive Fubini” greift hier nicht, da unsere Funktion auch negative Werte annimmt, der “Fubini” greift auch nicht, da unsere Funktion nicht integrierbar ist.

$\mathbb{N} \times \mathbb{N} \rightarrow \mathbb{R}$ mit Träger in der “treppenförmigen” Menge $\{(i, j) \mid 0 \leq i - j \leq 1\}$ mit $f(0, 0) = 1/2$ und

$$\begin{aligned} f(n, n) + f(n+1, n) &= 0 \\ f(n+1, n) + f(n+1, n+1) &= 2^{-(n+2)} \end{aligned}$$

für alle $n \geq 0$. Die beiden partiellen Integrale von f existieren und sind integrierbar. Ihre Integrale sind jedoch verschieden, genauer gilt

$$\int \left(\int f(n, m) \zeta\langle n \rangle \right) \zeta\langle m \rangle = 0 \neq 1 = \int \left(\int f(n, m) \zeta\langle m \rangle \right) \zeta\langle n \rangle$$

Indem wir unsere Funktion etwas “verschmieren” erhalten wir auch eine stetige Funktion auf \mathbb{R}^2 mit entsprechenden Eigenschaften, und durch geeignete Transformation sogar eine stetige reellwertige Funktion auf dem offenen Einheitsquadrat derart, daß die partiellen Integrale existieren und selbst wieder integrierbar sind, das Endresultat jedoch von der Integrationsreihenfolge abhängt.

Beweis von Satz 9.6.18. Ist f nichtnegativ, so folgt die Behauptung aus dem positiven Fubini 9.6.8, denn aus $\int_Y \left(\int_X f(x, y) \mu\langle x \rangle \right) \nu\langle y \rangle < \infty$ folgt, daß die Menge N aller $y \in Y$ mit $\int_X f(x, y) \mu\langle x \rangle = \infty$ Maß Null hat. Im allgemeinen folgt die Behauptung dann mit der Zerlegung $f = f^+ - f^-$. \square

Beispiel 9.6.22. Wir integrieren die Funktion y über die durch eine Parabel und die Gerade $y = 0$ begrenzte Fläche $P = \{(x, y) \mid 0 \leq y \leq 1 - x^2\}$ und erhalten

$$\begin{aligned} \int_P y &= \int_{-1}^1 \left(\int_0^{1-x^2} y \, dy \right) dx = \int_{-1}^1 \frac{(1-x^2)^2}{2} dx \\ &= \int_{-1}^1 \frac{1}{2} - x^2 + \frac{x^4}{2} dx = \frac{x}{2} - \frac{x^3}{3} + \frac{x^5}{10} \Big|_{-1}^1 = 1 - \frac{2}{3} + \frac{1}{5} = \frac{8}{15} \end{aligned}$$

Teilen wir noch durch die Gesamtfläche

$$\int_P 1 = \int_{-1}^1 (1 - x^2) dx = x - \frac{x^3}{3} \Big|_{-1}^1 = 2 - \frac{2}{3} = \frac{4}{3}$$

so ergibt sich die Höhe des Schwerpunkts unserer abgeschnittenen Parabelfläche zu $\frac{2}{5}$. Hier haben wir den Satz von Fubini, und zwar die positive Variante, angewandt auf das Produkt der Funktion y mit der charakteristischen Funktion $[P]$ unserer Fläche P . Die Funktion y ist meßbar, weil sie stetig ist, die Funktion $[P]$ ist meßbar als charakteristische Funktion einer meßbaren da abgeschlossenen Menge, und das Produkt dieser beiden meßbaren Funktionen ist damit auch meßbar nach 9.3.18.

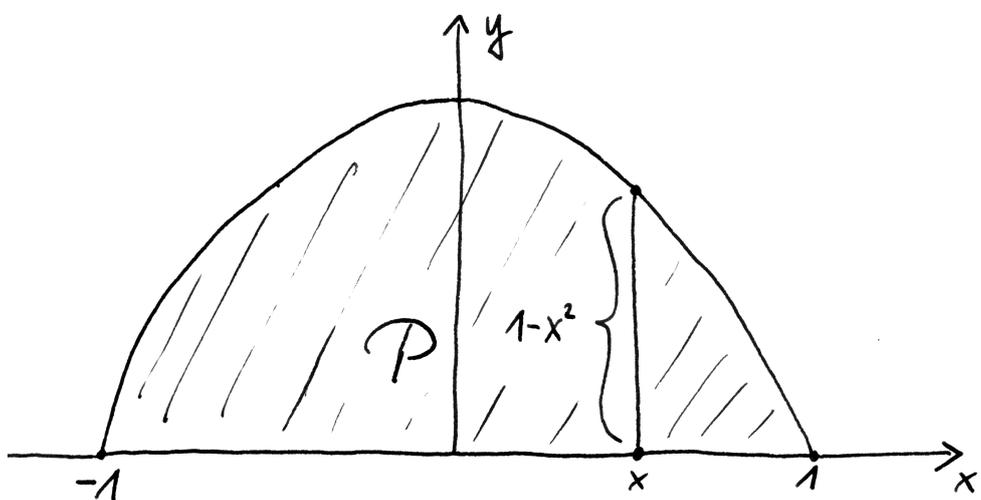


Illustration zur Anwendung des Satzes von Fubini in [9.6.22](#).

Übung 9.6.23. Jede stetige reellwertige Funktion auf einem kompakten Quader im \mathbb{R}^n ist integrierbar und ihr Riemann-Integral nach 5.1.1 stimmt mit ihrem Lebesgue-Integral überein. Hinweis: 9.5.13.

Übung 9.6.24. Zeige: Die Menge $\{x \in \mathbb{R}^n \mid 0 \leq x_1 \leq \dots \leq x_n \leq 1\}$ hat das Volumen $(n!)^{-1}$.

Übung 9.6.25. Man diskutiere den Zusammenhang zwischen dem Satz von Fubini und dem Satz über das Produkt von Reihen 5.6.11.

Proposition 9.6.26 (Partielle Integration). Gegeben reelle Zahlen $a < b$ und integrierbare Funktionen $f, g : [a, b] \rightarrow \mathbb{R}$ mit “Stammfunktionen” $F, G : [a, b] \rightarrow \mathbb{R}$ gegeben durch $F(y) = \int_a^y f(x) dx$ und $G(x) = \int_a^x g(y) dy$ gilt

$$\int_a^b Fg = FG|_a^b - \int_a^b fG$$

Dieselbe Formel gilt auch allgemeiner, wenn wir F oder G jeweils noch um eine additive Konstante abändern.

Beweis. Die zweite Aussage folgt leicht aus der ersten. Um die erste Aussage zu zeigen, berechnen wir das Integral der Funktion $f(x)g(y)$ über das Quadrat $[a, b]^2$ und finden mit Fubini $F(b)G(b)$. Andererseits können wir dies Integral auch schreiben als das Integral über das dreieckige Gebiet unterhalb der Diagonalen plus das Integral über das dreieckige Gebiet oberhalb der Diagonalen. Diese Integrale ergeben sich aber wieder mit Fubini leicht zu $\int_a^b Fg$ und $\int_a^b fG$. \square

9.7 Regularität von Borelmaßen

Satz 9.7.1 (Regularität von Borelmaßen auf \mathbb{R}^n). Gegeben ein Borelmaß λ auf dem \mathbb{R}^n gelten für jede Borelmenge $A \subset \mathbb{R}^n$ die Formeln

$$\lambda(A) = \inf_{\substack{U \supset A \\ U \text{ offen in } \mathbb{R}^n}} \lambda(U) = \sup_{\substack{K \subset A \\ K \text{ kompakt}}} \lambda(K)$$

9.7.2. Das “umgekehrte” Approximieren durch Kompakta von außen oder offene Mengen von innen ist nicht möglich, wie zum Beispiel die Fälle der Mengen A aller rationalen bzw. aller irrationalen Punkte in $[0, 1]$ zeigen.

Übung 9.7.3. Analoges gilt auch für Borelmaße auf beliebigen offenen Teilmengen $W \subseteq \mathbb{R}^n$ und wird in der Allgemeinheit in ?? benötigt. Um es zu zeigen, mag man etwa W als Vereinigung einer Folge offener Quader W_n schreiben und A als disjunkte Vereinigung topologisch meßbarer Teilmengen $A_n \subset W_n$ und die analoge Aussage für alle $A_n \subset W_n$ verwenden. Diese Aussage folgt nämlich direkt aus 9.7.1, da jeder offene Quader in einer die Topologien respektierenden Weise mit dem \mathbb{R}^n identifiziert werden kann.

Ergänzung 9.7.4. Ein topologischer Raum heißt **lokal kompakt** genau dann, wenn sich für jeden Punkt jede Umgebung zu einer kompakten Umgebung desselben Punktes verkleinern läßt. Diese Terminologie ist nicht unumstritten. Viele Autoren schreiben auch lokalkompakt zusammen und definieren diesen Begriff als das, was wir “lokal kompakt und Hausdorff” nennen. In Übung ?? zeigen Sie die Aussage des Satzes ganz allgemein für jedes Borelmaß auf einem beliebigen separablen lokal kompakten Hausdorffraum.

Beweis. Für beschränkte Intervalle $I_1, \dots, I_n \subset \mathbb{R}$ betrachten wir den Quader $I_1 \times \dots \times I_n \subset \mathbb{R}^n$. Die Gesamtheit aller endlichen Vereinigungen derartiger beschränkter Quader ist ein Mengenring $\mathcal{Q} \subset \mathcal{P}(\mathbb{R}^n)$, der die σ -Algebra der topologisch meßbaren Mengen erzeugt. Wir können also die Proposition über die Konstruktion von Maßfortsetzungen 9.2.13 in Verbindung mit der Eindeutigkeitsaussage aus dem Satz von Caratheodory 9.2.10 anwenden und folgern für jede topologisch meßbare Teilmenge $A \subset \mathbb{R}^n$ die Gleichung

$$\lambda(A) = \inf \left(\sum_{\nu=0}^{\infty} \lambda(Q_{\nu}) \right)$$

wobei das Infimum gebildet wird über alle Folgen Q_{ν} in \mathcal{Q} mit $A \subset \bigcup_{\nu} Q_{\nu}$. Für jedes $\varepsilon > 0$ finden wir demnach eine Folge Q_{ν} von beschränkten Quadern wie oben mit

$$\lambda(A) \leq \sum_{\nu=0}^{\infty} \lambda(Q_{\nu}) \leq \lambda(A) + \varepsilon$$

Nun finden wir für jeden dieser Quader Q_{ν} einen offenen Quader $B_{\nu} \supset Q_{\nu}$ mit $\lambda(Q_{\nu}) \leq \lambda(B_{\nu}) \leq \lambda(Q_{\nu}) + 2^{-\nu}\varepsilon$, da nämlich jeder unserer Quader der Schnitt einer absteigenden Folge offener beschränkter Quader ist, und für die offene Menge $U = \bigcup_{\nu} B_{\nu}$ folgt dann

$$\lambda(A) \leq \lambda(U) \leq \sum_{\nu=0}^{\infty} \lambda(B_{\nu}) \leq \sum_{\nu=0}^{\infty} \lambda(Q_{\nu}) + 2^{-\nu}\varepsilon \leq \lambda(A) + 3\varepsilon$$

Da das für alle $\varepsilon > 0$ gilt, ist das Maß von A in der Tat das Infimum über die Maße aller offenen Mengen, die A umfassen. Um die zweite Behauptung zu zeigen, wählen wir eine Folge $L_0 \subset L_1 \subset L_2 \subset \dots$ kompakter Teilmengen des \mathbb{R}^n mit $\bigcup L_i = \mathbb{R}^n$. Natürlich gilt $\lambda(A) = \lim_{i \rightarrow \infty} \lambda(A \cap L_i)$ und es reicht folglich, die zweite Gleichung für alle $A \cap L_i$ zu zeigen. In anderen Worten dürfen wir annehmen, daß es ein Kompaktum L gibt mit $A \subset L$. Nach dem schon bewiesenen Teil gilt natürlich erst recht

$$\lambda(L \setminus A) = \inf_{\substack{U \supset L \setminus A \\ U \text{ offen}}} \lambda(U \cap L)$$

Jetzt beachten wir die Gleichungen $\lambda(L \setminus A) = \lambda(L) - \lambda(A)$ und $\lambda(U \cap L) = \lambda(L) - \lambda(L \setminus U)$ und erhalten

$$\lambda(A) = \sup_{\substack{U \supset L \setminus A \\ U \text{ offen}}} \lambda(L \setminus U)$$

Aber es gilt ja $\{L \setminus U \mid U \supset (L \setminus A), U \text{ offen}\} = \{K \subset A \mid K \text{ kompakt}\}$, und damit ist auch die zweite Behauptung bewiesen. \square

Übung 9.7.5. Eine Teilmenge eines \mathbb{R}^n ist eine Nullmenge in Bezug auf das Lebesgue-Maß genau dann, wenn sie sich für jedes $\varepsilon > 0$ durch eine Folge von kompakten Quadern Q_n überdecken läßt mit $\sum_{n=0}^{\infty} \text{vol } Q_n < \varepsilon$.

9.8 Rechnen mit dem Lebesgue-Integral

Satz 9.8.1 (Transformationsformel). Sind $U, V \subseteq \mathbb{R}^n$ offene Teilmengen und ist $\phi : U \rightarrow V$ ein C^1 -Diffeomorphismus, so gilt für jede meßbare Funktion $f : V \rightarrow [0, \infty]$ in $[0, \infty]$ die Gleichheit

$$\int_V f = \int_U (f \circ \phi) |\det d\phi|$$

Ist alternativ $f : V \rightarrow \mathbb{R}$ eine integrierbare Funktion, so ist auch die Funktion $(f \circ \phi) |\det d\phi| : U \rightarrow \mathbb{R}$ integrierbar und dieselbe Formel gilt in \mathbb{R} .

9.8.2. Steht x für eine Veränderliche des \mathbb{R}^k , so benutzen wir die Notation $\int f(x) d^k x$ auch für Integrale bezüglich des Lebesguemaßes. In diesem Zusammenhang hat also $d^k x$ dieselbe Bedeutung wie $\lambda^k \langle x \rangle$. Wir können die Aussage des Satzes mit dieser Notation auch interpretieren als die Verwandtschaft von Maßen

$$\phi : |\det d\phi| d^k x \rightsquigarrow d^k y$$

9.8.3. Wir kennen unsere Formel aus 7.4.6 bereits für stetige Funktionen f mit kompaktem Träger. Dem eigentlichen Beweis schicken wir ein Lemma voraus.

Lemma 9.8.4. Ist $A \subseteq \mathbb{R}^n$ eine offene Teilmenge, so gibt es eine monoton wachsende Folge von stetigen, ja sogar von glatten nichtnegativen Funktionen mit kompaktem, in A enthaltenem Träger, die punktweise gegen die charakteristische Funktion $[A]$ von A strebt.

Beweis. Man schreibe A als Vereinigung einer Folge offener Quader Q_k mit kompaktem Abschluß, wähle etwa mithilfe von 7.2.11 für jedes Q_k eine glatte Funktion g_k auf \mathbb{R}^n , die auf Q_k positiv ist und außerhalb von Q_k verschwindet, und betrachte die Folge der Funktionen $f_k = g_1 + \dots + g_k$. Des weiteren wähle man

eine Folge von monoton wachsenden glatten Funktionen $h_k : \mathbb{R} \rightarrow \mathbb{R}$ derart, daß h_k unterhalb von $1/(k+1)$ verschwindet und oberhalb von $1/k$ konstant den Wert 1 annimmt. Die $h_k \circ f_k$ bilden dann eine Folge von Funktionen der gewünschten Art. \square

Beweis von 9.8.1. Wir zeigen nun zunächst die erste Behauptung. Mit Lemma 9.8.4 folgt unsere Formel für die charakteristischen Funktionen $f = [O]$ von offenen Teilmengen $O \subseteq V$, indem wir $[O]$ als punktweisen monotonen Grenzwert einer Folge f_n aus $\mathcal{C}_c(V)$ schreiben und den Satz über monotone Konvergenz 9.4.9 verwenden und erinnern, daß wir die Transformationsformel für stetige Funktionen mit kompaktem Träger bereits als 7.4.6 gezeigt haben. Dann folgt sie für die charakteristischen Funktionen $f = [K]$ von kompakten Teilmengen $K \subset V$, indem wir eine offene Menge O endlichen Maßes finden mit $K \subset O \subset V$ und $[K] = [O] - [O \setminus K]$ schreiben. Dann folgt sie für die charakteristischen Funktionen $f = [A]$ von meßbaren Teilmengen $A \subset V$, indem wir mithilfe von 9.7.1 eine absteigende Folge offener Mengen und eine aufsteigende Folge kompakter Mengen finden mit $O_0 \supset O_1 \supset \dots \supset A \supset \dots \supset K_1 \supset K_0$ und $\lim_{n \rightarrow \infty} \lambda(O_n) = \lambda(A) = \lim_{n \rightarrow \infty} \lambda(K_n)$. Dann folgt sie für meßbare reellwertige Stufenfunktionen aus der Additivität des Integrals 9.4.11. Schließlich folgt sie für beliebige meßbare Funktionen $f : V \rightarrow [0, \infty]$ mit dem Satz über monotone Konvergenz, indem wir f mit 9.4.12 als punktweisen monotonen Grenzwert meßbarer Stufenfunktionen schreiben. Der Fall integrierbarer Funktionen folgt ohne Mühe aus dem Fall meßbarer Funktionen mit Werten in $[0, \infty]$. \square

Übung 9.8.5. Liefern zwei Borelmaße auf dem \mathbb{R}^n dasselbe Integral für alle glatten Funktionen mit kompaktem Träger, so stimmen sie überein. Hinweis: 9.8.4 und 9.7.1. Verwendet man 9.7.3, so folgt dieselbe Aussage sogar für Borelmaße auf beliebigen offenen Teilmengen eines \mathbb{R}^n .

Proposition 9.8.6 (Nützliche Nullmengen). *Ist $U \subseteq \mathbb{R}^k$ offen, $k < n$ und $\varphi : U \rightarrow \mathbb{R}^n$ stetig differenzierbar, so ist $\varphi(U)$ eine Nullmenge in \mathbb{R}^n .*

Ergänzung 9.8.7. Der **Satz von Sard** besagt, daß auch für $k \geq n$ und φ mindestens $(k-n+1)$ -mal stetig partiell differenzierbar auf U das Bild unter φ der Menge aller Stellen $p \in U$, an denen $d_p\varphi$ nicht surjektiv ist, eine Lebesgue-Nullmenge sein muß. Wir werden das nicht zeigen.

Beweis. Nach dem technischen Lemma 9.8.8, das wir im Anschluß beweisen, können wir U schreiben als abzählbare Vereinigung über eine Folge von offenen Quadern Q_ν mit \bar{Q}_ν kompakt und $\bar{Q}_\nu \subset U$. Folglich dürfen wir ohne Beschränkung der Allgemeinheit voraussetzen, daß \bar{U} selbst ein kompakter Quader ist und daß $|d\varphi|$ beschränkt ist auf U . Nach einer affinen Koordinatentransformation dürfen wir zusätzlich sogar annehmen $U = (0, 1)^k$. Wir arbeiten wie immer mit der

Maximumsnorm auf \mathbb{R}^n und \mathbb{R}^k , die Bälle $B(x, \delta)$ sind also offene Würfel und ihre Abschlüsse $\bar{B}(x; \delta)$ abgeschlossene Würfel. Ist C eine Schranke für $|\mathrm{d}\varphi|$, so gilt nach dem Schrankensatz $\varphi(U \cap \bar{B}(x; \delta)) \subset \bar{B}(\varphi(x); C\delta)$ für alle $x \in U$. Für $r \in \mathbb{N}$, $r \geq 1$ finden wir nun eine Überdeckung von $(0, 1)^k$ durch r^k abgeschlossene Würfelchen der Gestalt $\bar{B}(x; 1/r)$, also finden wir eine Überdeckung von $\varphi(U)$ durch r^k abgeschlossene Würfelchen der Gestalt $\bar{B}(y; C/r)$ mit Gesamtvolumen $r^k(2C/r)^n$. Dies Gesamtvolumen strebt aber gegen Null für $r \rightarrow \infty$, mithin ist $\varphi(U)$ eine Nullmenge. \square

Lemma 9.8.8. *Jede offene Überdeckung einer Teilmenge des \mathbb{R}^n besitzt eine abzählbare Teilüberdeckung.*

Beweis. Wir betrachten in \mathbb{R}^n alle offenen Bälle mit rationalem Radius und rationalem Mittelpunkt. Diese offenen Bälle lassen sich abzählen und bilden dann eine Folge U_0, U_1, \dots von Teilmengen des \mathbb{R}^n . Ist $V \subseteq \mathbb{R}^n$ offen und $p \in V$ ein Punkt, so finden wir stets ein ν mit $p \in U_\nu \subset V$. In der Tat gibt es $\varepsilon > 0$ rational mit $B(p; \varepsilon) \subset V$ und einen Punkt mit rationalen Koordinaten in $B(p; \varepsilon/2)$ und dann gilt $p \in B(q; \varepsilon/2) \subset V$. Ist nun $A \subset \mathbb{R}^n$ überdeckt durch eine Familie V_i von offenen Teilmengen des \mathbb{R}^n , so betrachten wir in der Folge der U_ν die Teilfolge $U_{\nu(0)}, U_{\nu(1)}, \dots$ aller U_ν , die in einem der V_i liegen, und wählen für jedes $r \in \mathbb{N}$ ein $i(r)$ mit $V_{i(r)} \supset U_{\nu(r)}$. Die $V_{i(r)}$ bilden dann die gesuchte abzählbare Teilüberdeckung. \square

Übung 9.8.9. Jede diskrete Teilmenge eines \mathbb{R}^n ist abzählbar.

Übung 9.8.10. Man zeige: Gegeben eine meßbare Teilmenge $A \subset \mathbb{R}^n$ und $c \in \mathbb{R}$ gilt $\lambda(cA) = |c|^n \lambda(A)$. Zum Beispiel hat eine Kugel vom doppelten Radius das achtfache Volumen.

Beispiel 9.8.11. Wir berechnen die Höhe des Schwerpunkts der massiven Halbkugel $H = \{(x, y, z) \in \mathbb{R}^3 \mid x^2 + y^2 + z^2 \leq 1 \text{ und } z \geq 0\}$. Per definitionem ist das diejenige Zahl $h \in \mathbb{R}$, für die gilt $\int_H (z - h) = 0$, so daß wir unter Zuhilfenahme von 9.6.17 erhalten

$$h \frac{2\pi}{3} = h \int_H 1 = \int_H z$$

Durch Übergang zu Kugelkoordinaten nach 7.4.13 folgt

$$\begin{aligned} \int_H z &= \int_0^{\pi/2} \int_0^{2\pi} \int_0^1 (r \cos \vartheta) r^2 \sin(\vartheta) \, dr \, d\vartheta \, d\varphi \\ &= 2\pi \left(\int_0^1 r^3 \, dr \right) \left(\int_0^{\pi/2} \cos(\vartheta) \sin(\vartheta) \, d\vartheta \right) \\ &= 2\pi \cdot \left(\frac{1}{4} \right) \cdot \frac{1}{2} \int_0^{\pi/2} \sin(2\vartheta) \, d\vartheta = \frac{\pi}{4} \cdot \frac{-\cos(2\vartheta)}{2} \Big|_0^{\pi/2} = \frac{\pi}{4} \end{aligned}$$

womit sich die Höhe des Schwerpunkts ergibt zu $h = 3/8$.

Beispiel 9.8.12. Wir berechnen das Integral der Funktion $x^2 + y^2$ über die Einheitskugel $K = \{(x, y, z) \mid x^2 + y^2 + z^2 \leq 1\}$. Physikalisch Gebildete erkennen, daß wir eigentlich das Trägheitsmoment der Einheitskugel um die z -Achse suchen, aber das spielt in unserer Rechnung keine Rolle. Durch Übergang zu Kugelkoordinaten 7.4.13 und mit 1.1.5 erhalten wir

$$\int_K x^2 + y^2 = \int_0^{2\pi} \int_0^\pi \int_0^1 r^4 \sin^3(\vartheta) dr d\vartheta d\varphi = \frac{8\pi}{15}$$

Proposition 9.8.13 (Fläche unter der Gauß'schen Glockenkurve). Für die Fläche unter der Gauß'schen Glockenkurve gilt die Formel $\int_{-\infty}^{\infty} \exp(-x^2) dx = \sqrt{\pi}$.

9.8.14. Die Gauss'sche Glockenkurve spielt eine zentrale Rolle in der Wahrscheinlichkeitstheorie, wie in ?? ausgeführt wird.

Beweis. Wir rechnen das Integral über die Ebene der nichtnegativen Funktion $\exp(-(x^2 + y^2))$ auf zwei Weisen aus, einmal direkt mit Fubini und ein zweites Mal, indem wir mithilfe von 9.8.6 die Ebene längs der negativen x -Achse aufschneiden und mit 9.8.1 zu Polarkoordinaten übergehen. Ein Vergleich der Resultate liefert die Behauptung. Genauer rechnen wir

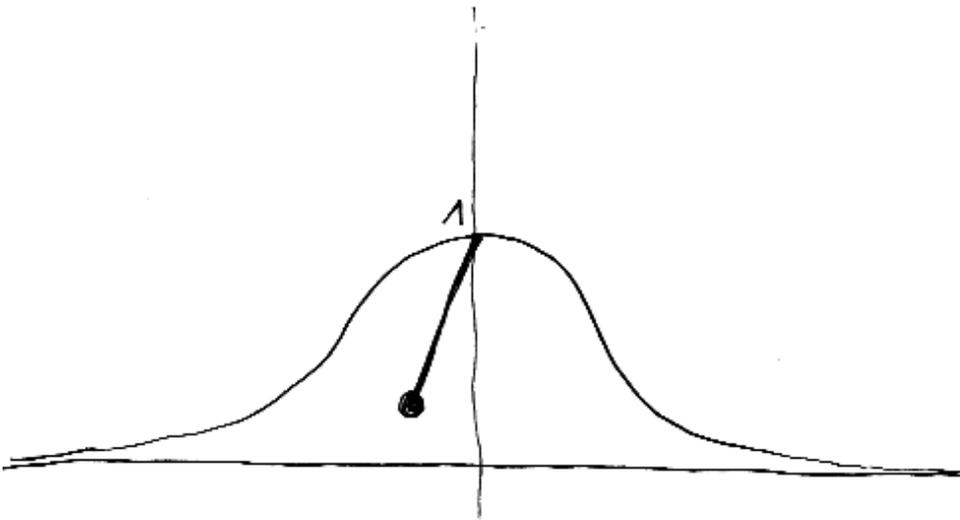
$$\begin{aligned} \int_{\mathbb{R}^2} \exp(-(x^2 + y^2)) &= \int_{\mathbb{R}} \int_{\mathbb{R}} \exp(-x^2) \exp(-y^2) dx dy \\ &= \left(\int_{-\infty}^{\infty} \exp(-x^2) dx \right)^2 \end{aligned}$$

$$\begin{aligned} \int_{\mathbb{R}^2} \exp(-(x^2 + y^2)) &= \int_{\mathbb{R}^2 \setminus \{(x,0) \mid x \leq 0\}} \exp(-(x^2 + y^2)) dx dy \\ &= \int_{(0,\infty) \times (-\pi,\pi)} \exp(-r^2) r dr d\theta \\ &= \int_0^\infty \int_{-\pi}^\pi \exp(-r^2) r dr d\theta \\ &= -\pi \exp(-r^2) \Big|_0^\infty \\ &= \pi \end{aligned} \quad \square$$

Beispiel 9.8.15. Es sollte wohl irgendwann einmal gezeigt werden, daß mit der in ?? definierten Interpolation $\Gamma : \mathbb{R}_{\geq 1} \rightarrow \mathbb{R}$ der Zuordnung $n \mapsto (n-1)!$ und der Konvention $x! := \Gamma(x+1)$ gilt

$$(\text{Volumen der Einheitskugel im } \mathbb{R}^n) = \frac{\pi^{n/2}}{(n/2)!}$$

Übung 9.8.16. Gegeben ein von Null verschiedenes Polynom $P \in \mathbb{R}[x_1, \dots, x_n]$ hat seine Nullstellenmenge $P^{-1}(0) \subset \mathbb{R}^n$ Lebesgue-Maß Null. Hinweis: Induktion über den Grad des Polynoms. Außerhalb der kritischen Stellen ist $P^{-1}(0)$ eine Untermannigfaltigkeit.



Skizze der Gauß'schen Glockenkurve alias dem Graphen von $x \mapsto \exp(-x^2)$ mit zusätzlich eingezeichnetem Klöppel.

9.9 Flächenmaß

Satz 9.9.1. Auf jeder Untermannigfaltigkeit $M \subset \mathbb{R}^n$ gibt es genau ein topologisches Maß $\sigma = \sigma_M$ derart, daß für jede Karte $\varphi : W \rightarrow M$ und jede topologisch meßbare Menge $A \subset \varphi(W)$ gilt

$$\sigma(A) = \int_{\varphi^{-1}(A)} \sqrt{\det(d_x \varphi)^\top (d_x \varphi)} \, d^k x$$

mit dem Lebesgue-Integral über $\varphi^{-1}(A) \subset W \subseteq \mathbb{R}^k$ auf der rechten Seite. Dieses Maß heißt das **Flächenmaß** von M . Es ist ein Borelmaß. Jede in M enthaltene Mannigfaltigkeit N echt kleinerer Dimension ist für das Flächenmaß von M eine Nullmenge.

9.9.2. Die hinter diesen Definitionen stehende Anschauung wurde bereits in 7.5.1 diskutiert. Man sieht leicht ein, daß das dort für stetige Funktionen mit kompaktem Träger erklärte Integral mit ihrem Integral in Bezug auf das hier erklärte Flächenmaß übereinstimmen muß. Der Buchstabe σ steht für englisch und französisch “surface”. Die Bezeichnung suggeriert zwar die Vorstellung zweidimensionaler Mannigfaltigkeiten, aber wir benutzen sie auch im allgemeinen. Gegeben eine integrierbare Funktion $f : M \rightarrow \mathbb{R}$ notieren wir ihr Integral bezüglich des Flächenmaßes

$$\int_M f = \int_M f \sigma = \int_M f(x) \sigma \langle x \rangle$$

In der Literatur ist es üblich, $d\sigma$ hinter das Integral zu schreiben, und man findet auch die Notationen dS und in der deutschen Literatur $d\omega$ oder dO mit ω oder O wie “Oberfläche”.

Ergänzung 9.9.3. Gegeben ein endlichdimensionaler reeller affiner Raum X und ein Skalarprodukt auf seinem Richtungsraum mit Einheiten in einem orientierten eindimensionalen Vektorraum L im Sinne von ?? liefern die analogen Definitionen auf jeder k -dimensionalen Untermannigfaltigkeit das Flächenmaß in Gestalt eines topologischen Maßes mit Werten in $L^{\otimes k}$. Unter einem **Maß mit Werten in einem eindimensionalen orientierten Vektorraum T** verstehen wir dabei eine σ -additive Abbildung von einer σ -Algebra nach $T_{\geq 0} \sqcup \{\infty\}$. Wie bereits erwähnt messen sich also auch in der Mathematik Längen in Metern, Flächen in Quadratmetern und Volumen in Kubikmetern.

9.9.4. Gegeben eine Karte $\varphi : W \rightarrow M$ ist die Einschränkung des Oberflächenmaßes auf ihr Bild $\varphi(W)$ also genau das Bild unter φ des Produkts des Lebesgue-Maßes auf W mit der Funktion $\text{vol}(d_x \varphi) = \sqrt{\det(d_x \varphi)^\top (d_x \varphi)}$ in der Notation aus 7.5.6. Gegeben eine meßbare Funktion $f : M \rightarrow [0, \infty]$, die außerhalb von $\varphi(W)$ verschwindet, haben wir nach 9.5.15 in $[0, \infty]$ die Identität

$$\int_M f = \int_W f(\varphi(x)) \text{vol}(d_x \varphi) \, d^k x$$

Ist ähnlich eine Funktion $f : M \rightarrow \mathbb{R}$ gegeben, die außerhalb von $\varphi(W)$ verschwindet, so ist f integrierbar über M genau dann, wenn die Funktion unter dem Integral auf der rechten Seite unserer Gleichung integrierbar ist über W , und dann gilt unsere Formel ganz genauso in \mathbb{R} .

Beispiel 9.9.5. Ist $\varphi : [a, b] \rightarrow \mathbb{R}^n$ eine Kurve derart, daß das Bild des offenen Intervalls $\varphi((a, b))$ eine 1-Mannigfaltigkeit M ist und $\varphi : (a, b) \rightarrow M$ eine Karte von M , so ist das Integral einer stetigen Funktion $f : \varphi([a, b]) \rightarrow \mathbb{R}$ über M genau das Kurvenintegral im Sinne von 10.3.6 der Funktion f längs der Kurve φ .

Beweis von Satz 9.9.1. Nach 9.8.8 können wir eine Folge von Karten (W_n, φ_n) finden, deren Bilder unsere Mannigfaltigkeit überdecken. Gegeben eine solche Folge zerlegen wir jede topologisch meßbare Menge $A \subset M$ in die disjunkte Vereinigung gewisser Teilmengen A_n , die jeweils aus allen Punkten im Bereich der n -ten Karte bestehen sollen, die nicht bereits im Bereich von Karten mit kleinerem Index liegen. In Formeln bilden wir also $M_n = \bigcup_{\nu \leq n} \varphi_\nu(W_\nu)$ und setzen $A_n = (A \setminus M_{n-1}) \cap \varphi_n(W_n)$. Dann erklären wir ein Maß σ auf M durch die Vorschrift

$$\sigma(A) = \sum_{n=0}^{\infty} \int_{\varphi_n^{-1}(A_n)} \text{vol}(d_x \varphi_n) \, d^k x$$

und überlassen dem Leser den Nachweis, daß diese Vorschrift auch wirklich ein Maß auf M definiert. Nun müssen wir für (W, φ) eine weitere Karte von M und $A \subset \varphi(W)$ meßbar zeigen

$$\sigma(A) = \int_{\varphi^{-1}(A)} \text{vol}(d_x \varphi) \, d^k x$$

Sicher reicht es aus, das für alle die eben zu A gebildeten A_n zu zeigen. Wir zeigen gleich allgemeiner für jede im Bild zweier Karten (W, φ) und (V, ψ) enthaltene topologisch meßbare Menge $B \subset \varphi(W) \cap \psi(V)$ die Formel

$$\int_{\psi^{-1}(B)} \text{vol}(d_y \psi) \, d^k y = \int_{\varphi^{-1}(B)} \text{vol}(d_x \varphi) \, d^k x$$

Ohne Beschränkung der Allgemeinheit dürfen wir hierfür $\varphi(W) = \psi(V)$ annehmen. Der Kartenwechsel $g = \psi^{-1} \circ \varphi : W \xrightarrow{\sim} V$ ist dann ein Diffeomorphismus mit $\psi \circ g = \varphi$, also gilt $d_{g(x)} \psi \circ d_x g = d_x \varphi$. Wir erhalten mit der Multiplikativität der Determinante also

$$\text{vol}(d_x \varphi) = |\det d_x g| \text{vol}(d_{g(x)} \psi)$$

und folgern die behauptete Gleichheit der Integrale aus der Transformationsformel 9.8.1 angewandt auf die Funktion $f(y) = \text{vol}(d_y \psi)[\psi^{-1}(B)](y)$ mit dem

transformierenden \mathcal{C}^1 -Diffeomorphismus $\phi = g : W \xrightarrow{\sim} V$ alias der Substitution $y = g(x)$. Daß es kein anderes Maß mit der geforderten Eigenschaft geben kann, ist eh klar. Daß jede in M enthaltene Untermannigfaltigkeit echt kleinerer Dimension für unser Oberflächenmaß eine Nullmenge ist, folgt leicht aus 9.8.6. Um zu zeigen, daß unser Flächenmaß endlich ist auf Kompakta $K \subset M$, wählen wir zunächst mithilfe von 9.10.3 eine endliche Überdeckung von K durch offene Mengen $V_1, \dots, V_r \subseteq \mathbb{R}^n$, in denen M jeweils plättbar ist, so daß insbesondere gilt $(M \cap V_i) \sqsubseteq V_i$ für $i = 1, \dots, r$. Dann betrachten wir eine Teilung der Eins nach 7.4.10 und finden also $\alpha_i : \mathbb{R}^n \rightarrow [0, 1]$ stetig mit kompaktem in V_i enthaltenem Träger und $\sum_{i=1}^r \alpha_i(x) = 1$ für alle $x \in K$. Es reicht also zu zeigen $\int_M \alpha_i(x) \sigma \langle x \rangle < \infty$ für alle i . Ist aber $\varphi : W \xrightarrow{\sim} M \cap V_i$ die zu einer entsprechenden Plättung gehörige Karte, so haben wir

$$\int_M \alpha_i(x) \sigma \langle x \rangle = \int_W (\alpha_i \circ \varphi)(x) \operatorname{vol}(d_x \varphi) d^k x < \infty$$

da ja $(\operatorname{supp} \alpha_i) \cap M \cap V_i$ und damit auch $\operatorname{supp}(\alpha_i \circ \varphi)$ kompakt ist. \square

Beispiel 9.9.6 (Oberfläche der Einheitskugel). Wir machen nun unsere heuristische Argumentation aus 7.5.5 präzise und zeigen für die Oberfläche der Einheitskugel die Formel

$$\int_{S^2} \sigma = 4\pi$$

Lassen wir aus der Kugelschale S^2 den Äquator weg, also alle Punkte (x, y, z) mit $z = 0$, und dazu noch einen halben Großkreis von Pol zu Pol, sagen wir alle Punkte (x, y, z) mit $y = 0$ und $x \leq 0$, so ändert sich nach 9.9.1 das Integral nicht. Der Rest ist die disjunkte Vereinigung von zwei geschlitzten offenen Hemisphären $U_+ \cup U_-$ und U_\pm und unsere Rechnung aus 7.5.5 zeigt bereits $\int_{U_\pm} \sigma = 2\pi$.

Übung 9.9.7 (Oberfläche eines Rotationskörpers). Sei $I \subset \mathbb{R}$ ein offenes Intervall und $f : I \rightarrow (0, \infty)$ stetig differenzierbar. So ist die **Mantelfläche** $M = \{(x, y, z) \in \mathbb{R}^3 \mid x^2 + y^2 = (f(z))^2\}$ eine zweidimensionale Untermannigfaltigkeit des \mathbb{R}^3 mit der Oberfläche

$$\int_M \sigma = 2\pi \int_I f(z) \sqrt{1 + (f'(z))^2} dz$$

Allgemeiner zeige man für das Bildmaß des Oberflächenmaßes unter der orthogonalen Projektion $p : M \rightarrow I$ unserer Mantelfläche auf die z -Achse die Formel $p_* \sigma = 2\pi f(z) \sqrt{1 + (f'(z))^2} dz$. Ist speziell M die Einheitskugel, so zeige man $p_* \sigma = 2\pi dz$ und berechne nochmals die Oberfläche der Einheitskugel.

9.9.8. Die anschauliche Bedeutung unserer Formel für die Oberfläche eines Rotationskörpers erkennt man, wenn man unsere Rotationsfläche durch eine Vereinigung von dünnen Bändern der Gestalt “unten abgeschnittenener Eiswaffeln” approximiert.

Übung 9.9.9 (Zwiebelformel). Ist $S^{n-1} = \{x \in \mathbb{R}^n \mid \|x\| = 1\}$ die Einheitskugel mit ihrem Flächenmaß σ , so ist unter der Multiplikationsabbildung $\text{mult} : \mathbb{R}_{>0} \times S^{n-1} \xrightarrow{\sim} \mathbb{R}^n \setminus \{0\}$ das Produktmaß $r^{n-1} dr \boxtimes \sigma$ verwandt zum Lebesgue-Maß auf $\mathbb{R}^n \setminus \{0\}$, in Formeln

$$\text{mult} : r^{n-1} dr \boxtimes \sigma \rightsquigarrow \lambda^n$$

Hinweis: Man rechne mit einer beliebigen Karte von S^{n-1} und erweitere sie zu einer Karte von $\mathbb{R}^n \setminus \{0\}$. Man beachte, daß für eine differenzierbare Kurve, die ganz in der Einheitskugel verläuft, der Geschwindigkeitsvektor stets auf dem Ortsvektor senkrecht steht.

Übung 9.9.10. Genau dann ist die Funktion $\mathbb{R}^n \rightarrow [0, \infty], x \mapsto \|x\|^\alpha$ für gegebenes $\alpha \in \mathbb{R}$ integrierbar auf dem Komplement eines und jedes offenen Balls um den Ursprung, wenn gilt $\alpha < (-n)$. Hinweis: Zwiebelformel 9.9.9.

Übung 9.9.11. Sei $\Gamma \subset \mathbb{R}^n$ ein Gitter, d.h. das Gruppenerzeugnis einer Basis von \mathbb{R}^n als \mathbb{R} -Vektorraum. Genau dann konvergiert $\sum_{\omega \in \Gamma \setminus \{0\}} \|\omega\|^\alpha$, wenn gilt $\alpha < (-n)$. Hinweis: 9.9.10.

Übung 9.9.12. Berechnen Sie das Integral der Funktion $(xyz)^2$ über die Einheitskugel in \mathbb{R}^3 .

10 Der Satz von Stokes

Im vorigen Abschnitt haben wir unser Kurvenintegral aus 10.3.6 verallgemeinert zum Integral einer Funktion über eine Unterannigfaltigkeit eines \mathbb{R}^n . In diesem Abschnitt werden wir unser Wegintegral aus 6.3, d.h. das Integral eines Kovektorfelds auf einem endlichdimensionalen reellen Raum längs eines Weges verallgemeinern zum Integral einer “ k -Form” auf einem endlichdimensionalen reellen Raum über eine “orientierte” k -dimensionale Untermannigfaltigkeit. Als Spezialfälle enthält diese Konstruktion insbesondere die Definition des “Flusses eines Vektorfelds in \mathbb{R}^3 durch eine orientierte Fläche in \mathbb{R}^3 ”. Unser eigentliches Ziel ist dann der sogenannte “allgemeine Satz von Stokes” 10.8.1, der den Hauptsatz der Differential- und Integralrechnung 7.5.1 auf höhere Dimensionen verallgemeinert.

10.1 Multilineare Algebra und Dachprodukt

Definition 10.1.1. Sei k ein Körper. Gegeben ein k -Vektorraum V und eine natürliche Zahl $p \in \mathbb{Z}$ bilden wir den Raum der **alternierenden p -Multilinearformen** oder kurz **p -Formen**

$$\text{Alt}^p V := \{\omega : V \times \dots \times V \rightarrow k \mid \omega \text{ ist multilinear und alternierend}\}$$

Hier meint alternierend, daß $\omega(v_1, \dots, v_p)$ verschwindet, wenn es $i \neq j$ gibt mit $v_i = v_j$. Hat unser Körper nicht die Charakteristik 2, so ist es gleichbedeutend zu fordern, daß $\omega(v_1, \dots, v_p)$ sein Vorzeichen ändert wenn man zwei Einträge v_i und v_j vertauscht, daher die Bezeichnung “alternierend”. Unter Nullformen verstehen wir Skalare, in Formeln setzen wir also $\text{Alt}^0 V = k$. Einsformen sind Elemente des Dualraums, wir haben also $\text{Alt}^1 V = V^\top$. Gegeben $f_1, \dots, f_p \in V^\top$ definieren wir ein Element $\text{alt}(f_1, \dots, f_p) \in \text{Alt}^p V$ durch die Vorschrift

$$\text{alt}(f_1, \dots, f_p)(v_1, \dots, v_p) = \det(f_i(v_j))$$

Sind Linearformen $f_1, \dots, f_n \in V^\top$ gegeben und ist $I \subset \{1, \dots, n\}$ eine Teilmenge mit p Elementen, so setzen wir

$$f_I = \text{alt}(f_{i_1}, \dots, f_{i_p}) \in \text{Alt}^p V$$

für $i_1 < \dots < i_p$ die der Größe nach gereihten Elemente von I . Für $I = \emptyset$ vereinbaren wir $f_\emptyset = 1$.

Ergänzung 10.1.2. Im Rahmen unserer Diskussion des Tensorprodukts werden die Begriffsbildungen dieses Abschnitts auch noch unter einem anderen Gesichtspunkt besprochen. Genauer konstruieren wir in ?? einen kanonischen Isomorphismus zwischen dem hier definierten Raum $\text{Alt}^p V$ der alternierenden Multilinearformen auf V und dem Dualraum $(\bigwedge^p V)^\top$ seiner dort definierten p -ten äußeren

Potenz $\bigwedge^p V$. Zusätzlich erklären wir in ?? für V von endlicher Dimension kanonische Isomorphismen $(\bigwedge^p V)^\top \xrightarrow{\sim} \bigwedge^p(V^\top)$ zwischen den Dualräumen der äußeren Potenzen und den äußeren Potenzen des Dualraums und erhalten so zusammen einen kanonischen Isomorphismus

$$\text{Alt}^p V \xrightarrow{\sim} \bigwedge^p(V^\top)$$

Proposition 10.1.3. *Ist V ein Vektorraum und f_1, \dots, f_n eine Basis von V^\top , so bilden die f_I mit $|I| = p$ eine Basis von $\text{Alt}^p V$.*

Beweis. Ist v_1, \dots, v_n die duale Basis von V und ist auch $J = \{j_1, \dots, j_p\} \subset \{1, \dots, n\}$ gegeben mit $j_1 < \dots < j_p$, so gilt offensichtlich

$$f_I(v_{j_1}, \dots, v_{j_p}) = \begin{cases} 1 & I = J; \\ 0 & \text{sonst.} \end{cases}$$

Das zeigt die lineare Unabhängigkeit der f_I . Andererseits ist klar, daß eine alternierende Multilinearform schon festgelegt wird durch ihre Werte auf den p -Tupeln $(v_{j_1}, \dots, v_{j_p})$ mit $j_1 < \dots < j_p$. Das zeigt, daß die f_I auch $\text{Alt}^p V$ erzeugen. \square

Proposition 10.1.4. *Sei V ein endlichdimensionaler k -Vektorraum und seien $p, q \geq 0$. So gibt es genau eine bilineare Abbildung, das **Dachprodukt***

$$\begin{aligned} \text{Alt}^p V \times \text{Alt}^q V &\rightarrow \text{Alt}^{p+q} V \\ (\omega, \eta) &\mapsto \omega \wedge \eta \end{aligned}$$

derart, daß für alle $f_1, \dots, f_{p+q} \in V^\top$ gilt

$$\text{alt}(f_1, \dots, f_p) \wedge \text{alt}(f_{p+1}, \dots, f_{p+q}) = \text{alt}(f_1, \dots, f_p, f_{p+1}, \dots, f_{p+q})$$

10.1.5. Mit 10.1.3 folgern wir unmittelbar die **Assoziativität des Dachprodukts**

$$(\omega \wedge \eta) \wedge \xi = \omega \wedge (\eta \wedge \xi)$$

Damit brauchen wir auch bei längeren Dachprodukten keine Klammern zu setzen und unsere Notation “alt” wird schon wieder obsolet, denn offensichtlich folgt aus der Proposition auch

$$\text{alt}(f_1, \dots, f_p) = f_1 \wedge \dots \wedge f_p$$

Ein natürlichere Konstruktion des Dachprodukts besprechen wir im Rahmen der multilinearen Algebra in ?. Sie mögen zur Übung zeigen, daß unter unserem Isomorphismus 10.1.2 das Dachprodukt aus ?? genau unserem Dachprodukt aus 10.1.4 entspricht, vergleiche auch ?. In der Tat reicht es angesichts der Assoziativität beider Dachprodukte, diese Behauptung im Fall des Dachprodukts zweier Linearformen zu prüfen, und in diesem Fall ist sie schnell nachgerechnet.

Beweis. Die Eindeutigkeit folgt sofort aus 10.1.3 und nur die Existenz ist noch zu zeigen. Wir betrachten dazu die Menge $\mathcal{S}_{p,q} \subset \mathcal{S}_{p+q}$ aller Permutationen, die die Reihenfolge der ersten p Einträge und die der letzten q Einträge unverändert lassen. Stellen wir uns unsere Permutationen als Mischvorschriften für ein Spiel von $p+q$ Karten vor, so heben wir also p Karten ab und schieben die beiden so gebildeten Stapel von p bzw. q Karten irgendwie ineinander. Solche Permutationen heißen auch (p, q) -**Shuffles**, in Formeln haben wir

$$\mathcal{S}_{p,q} = \{\sigma \in \mathcal{S}_{p+q} \mid \sigma(1) < \dots < \sigma(p) \text{ und } \sigma(p+1) < \dots < \sigma(p+q)\}$$

Weiter betrachten wir in \mathcal{S}_{p+q} die Untergruppe $\mathcal{S}_p \times \mathcal{S}_q$ aller Permutationen, die die ersten p Einträge unter sich vertauschen und die letzten q Einträge ebenso. Die Verknüpfung von Permutationen liefert dann offensichtlich eine Bijektion

$$\mathcal{S}_{p,q} \times (\mathcal{S}_p \times \mathcal{S}_q) \xrightarrow{\sim} \mathcal{S}_{p+q}$$

Jetzt definieren wir für ω und η wie oben eine Multilinearform $\omega \wedge \eta$ durch die Vorschrift

$$(\omega \wedge \eta)(v_1, \dots, v_{p+q}) = \sum_{\sigma \in \mathcal{S}_{p,q}} \text{sgn}(\sigma) \omega(v_{\sigma(1)}, \dots, v_{\sigma(p)}) \eta(v_{\sigma(p+1)}, \dots, v_{\sigma(p+q)})$$

Betrachten wir andererseits unsere Definition

$$\text{alt}(f_1, \dots, f_n)(v_1, \dots, v_n) = \sum_{\tau \in \mathcal{S}_n} \text{sgn}(\tau) f_1(v_{\tau(1)}) \dots f_n(v_{\tau(n)})$$

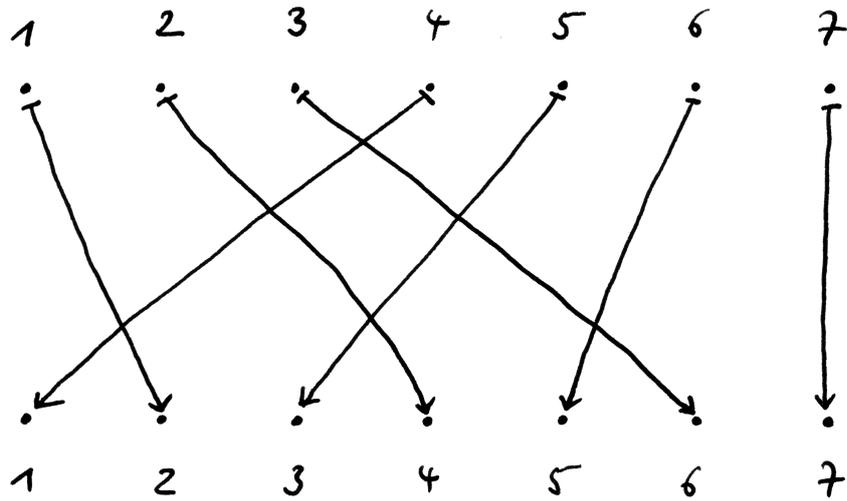
für $n = p, q$ und setzen in die Definition von \wedge ein, so ergibt sich mithilfe unserer Zerlegung $\mathcal{S}_{p,q} \times (\mathcal{S}_p \times \mathcal{S}_q) \xrightarrow{\sim} \mathcal{S}_{p+q}$ wie gewünscht

$$\text{alt}(f_1, \dots, f_p) \wedge \text{alt}(f_{p+1}, \dots, f_{p+q}) = \text{alt}(f_1, \dots, f_p, f_{p+1}, \dots, f_{p+q})$$

Die Bilinearität von \wedge zeigt dann weiter, daß die Multilinearform $\omega \wedge \eta$ auch im allgemeinen alternierend ist, so daß unsere Formel für \wedge in der Tat eine Abbildung $\text{Alt}^p V \times \text{Alt}^q V \rightarrow \text{Alt}^{p+q} V$ mit den geforderten Eigenschaften liefert. \square

Lemma 10.1.6 (Graduierte Kommutativität des Dachprodukts). Sei V ein endlichdimensionaler Vektorraum. Für beliebige $\omega \in \text{Alt}^p V$ und $\eta \in \text{Alt}^q V$ gilt $\omega \wedge \eta = (-1)^{pq} \eta \wedge \omega$. Bezeichnet $|\omega|$ den Grad von ω , also $|\omega| = p$ für $\omega \in \text{Alt}^p$, so können wir diese Regel auch schreiben in der Gestalt

$$\omega \wedge \eta = (-1)^{|\omega||\eta|} \eta \wedge \omega$$



Ein (3, 4)-Shuffle

Beweis. Aus 10.1.4 folgt sofort $f_{\sigma(1)} \wedge \dots \wedge f_{\sigma(n)} = (\text{sgn } \sigma) f_1 \wedge \dots \wedge f_n$ für jede Permutation $\sigma \in \mathcal{S}_n$ und alle $f_1, \dots, f_n \in V^\top$. Die Permutation $\sigma \in \mathcal{S}_{p+q}$, die die ersten p Einträge an den Schluß schiebt und die letzten q Einträge an den Anfang, hat aber nach ?? das Signum $\text{sgn}(\sigma) = (-1)^{pq}$. Das Lemma folgt so zunächst für ω, η iterierte Dachprodukte und dann auch im Allgemeinen. \square

Definition 10.1.7. Zu jeder linearen Abbildung $L : V \rightarrow W$ bilden wir ihre transponierte Abbildung $L^\top : W^\top \rightarrow V^\top$, $f \mapsto f \circ L$ und allgemeiner auch die linearen Abbildungen $L^\top : \text{Alt}^p W \rightarrow \text{Alt}^p V$, $\omega \mapsto \omega \circ (L \times \dots \times L)$.

10.1.8. Aus den Definitionen folgen leicht die Formeln $\text{id}^\top = \text{id}$ und $(L \circ M)^\top = M^\top \circ L^\top$ für die transponierten Abbildungen sowie die Verträglichkeit mit dem Dachprodukt

$$L^\top(\omega \wedge \eta) = (L^\top \omega) \wedge (L^\top \eta)$$

Lemma 10.1.9 (Dachprodukt und Determinante). Gegeben ein n -dimensionaler Vektorraum V und eine lineare Abbildung $L : V \rightarrow V$ gilt

$$L^\top = (\det L) : \text{Alt}^n V \rightarrow \text{Alt}^n V$$

Beweis. Ist V ein n -dimensionaler Vektorraum, so ist $\text{Alt}^n V$ eindimensional. Für $L : V \rightarrow V$ linear muß also $L^\top : \text{Alt}^n V \rightarrow \text{Alt}^n V$ die Multiplikation mit einem Skalar aus dem Grundkörper sein. Ist v_1, \dots, v_n eine Basis von V und f_1, \dots, f_n die duale Basis von V^\top , so ist $f_1 \wedge \dots \wedge f_n$ eine Basis von $\text{Alt}^n V$ und das Lemma folgt mit expliziter Rechnung, für $(\det L)$ die Determinante der Matrix von L in der gewählten Basis. Daß die fragliche Determinante von der Wahl der Basis gar nicht abhängt und deshalb in der Tat $(\det L)$ notiert werden darf, erhält man als Konsequenz. \square

10.1.10. Nehmen wir 10.1.8 und 10.1.9 zusammen, so ergibt sich unmittelbar die Multiplikationsformel für Determinanten ??.

10.1.11. Gegeben endlichdimensionale Vektorräume V, W und Formen $\omega \in \text{Alt}^p V$ und $\eta \in \text{Alt}^q W$ kürzen wir die $(p+q)$ -Form $(\text{pr}_1^\top \omega) \wedge (\text{pr}_2^\top \eta)$ auf $V \times W$ auch gerne mit $\omega \wedge \eta$ ab und hoffen, daß der Leser aus dem Kontext erschließen kann, wann \wedge dieses ‘‘äußere Dachprodukt’’ meint und wann das ‘‘innere Dachprodukt’’ aus 10.1.4.

10.1.12. Für einen Vektorraum V der Dimension $\dim V = n$ liefert das Dachprodukt $V^\top \times \text{Alt}^{n-1} V \rightarrow \text{Alt}^n V$ eine nichtausgeartete Paarung, als da heißt, jeder Isomorphismus $\text{Alt}^n V \xrightarrow{\sim} \mathbb{R}$ liefert vermittels unserer Paarung einen Isomorphismus $\text{Alt}^{n-1} V \cong V$.

10.2 Differentialformen höheren Grades

Definition 10.2.1. Sei X ein endlichdimensionaler reeller Raum und $U \subset X$ eine Teilmenge. Eine **relative p -Form** auf U ist eine Abbildung

$$\begin{aligned} \omega : U &\rightarrow \text{Alt}^p \vec{X} \\ x &\mapsto \omega_x \end{aligned}$$

die also ausgeschrieben jedem Punkt $x \in U$ eine alternierende p -Multilinearform $\omega_x : \vec{X} \times \dots \times \vec{X} \rightarrow \mathbb{R}$ zuordnet. Den Raum aller relativen p -Formen auf U notieren wir $\Omega_X^p(U)$. Wenn wir hoffen, daß die genaue Bedeutung aus dem Kontext hervorgeht, sprechen wir auch oft abkürzend schlicht von p -Formen.

10.2.2. Später werden wir ganz allgemein Differentialformen auf Mannigfaltigkeiten erklären als Zuordnungen, die jedem Punkt eine alternierende Multilinearform auf dem Tangentialraum am entsprechenden Punkt zuordnen. Im Fall einer eingebetteten Mannigfaltigkeit $U \subset X$ positiver Kodimension ist das natürlich etwas anderes, als jedem Punkt eine alternierende Multilinearform auf dem Richtungsraum des umgebenden affinen Raums zuzuordnen. Das ist der Grund, aus dem ich das hier eingeführte Konzept eine “relative Differentialform” genannt habe.

10.2.3. Unglücklicherweise kann mit der bis hierher eingeführten Terminologie eine “ p -Form auf U ” zwei sehr verschiedene Dinge bedeuten: Entweder ist U ein k -Vektorraum und unsere p -Form ist ein Element von $\text{Alt}^p(U)$, also eine alternierende multilineare Abbildung $\omega : U \times \dots \times U \rightarrow k$, oder aber U ist Teilmenge eines endlichdimensionalen reellen Raums X und unsere p -Form ist eine Abbildung $\omega : U \rightarrow \text{Alt}^p(\vec{X})$ mit $x \mapsto \omega_x$. Eigentlich schiene es mir natürlicher, letztere Objekte als **Felder von p -Formen** oder **p -Formenfelder** anzusprechen, da sie ja zu alternierenden p -Multilinearformen in derselben Beziehung stehen wie Vektorfelder zu Vektoren. Von Formenfeldern aber redet kein Mensch, deshalb will ich auch nicht damit anfangen, und der Leser muß aus dem Kontext erschließen, welche Bedeutung im Einzelfall gemeint ist.

Beispiel 10.2.4. Eine 0-Form auf U ist eine Funktion $f : U \rightarrow \mathbb{R}$ und eine 1-Form auf einer halboffenen Teilmenge U ein Kovektorfeld im Sinne von 6.1.4. Ist $U \subset X$ eine halboffene Teilmenge eines endlichdimensionalen reellen Raums, so schreiben wir statt $\Omega_X^p(U)$ meist $\Omega^p(U)$, da in diesem Fall unsere relativen Differentialformen mit den üblichen Differentialformen auf abstrakten Mannigfaltigkeiten übereinstimmen, wie wir sie in ?? kennenlernen werden.

Beispiel 10.2.5. Ein magnetisches Feld ist formal betrachtet bis auf skalare Einheiten eine 2-Form auf dem Anschauungsraum. In der Tat kann man ein Magnetfeld etwa messen, indem man ein Halbleiterplättchen hineinhält, zwischen zwei

gegenüberliegenden Kanten einen Strom fließen läßt, und zwischen den zwei verbleibenden Kanten die Spannung mißt. So erhält man für jedes “orientierte Flächenelement” eine Zahl, und solch eine Zuordnung ist eben gerade eine 2-Form. Sicher kann man eine 2-Form auf dem Anschauungsraum auch als Vektorfeld ansehen, genauer erhält man eine Identifikation von Vektorfeldern mit 2-Formen, indem man jedem Vektor u die 2-Form $(v, w) \mapsto \text{vol}(u, v, w)$ zuordnet, mit $\text{vol}(u, v, w)$ dem “Volumen” des Parallelogramms mit Kanten u, v, w und einem Vorzeichen, das von der “Orientierung” unseres Tripels abhängt, oder noch genauer die 2-Form $(v, w) \mapsto \langle u, v, w \rangle \text{m}^{-3}$ gegeben durch das Spatprodukt aus ?? und eine willkürlich gewählte Längeneinheit $\text{m} \in \mathbb{L}_{>0}$. Diese Identifikation scheint mir jedoch eher künstlich, insbesondere hängt sie von der Wahl einer “Orientierung” des Anschauungsraums und der Wahl einer Längeneinheit ab.

Definition 10.2.6. Für zwei Differentialformen $\omega \in \Omega^p$ und $\eta \in \Omega^q$ definieren wir ihr **Dachprodukt** $\omega \wedge \eta \in \Omega^{p+q}$ als punktweises Dachprodukt im Sinne von 10.1.4, in Formeln $(\omega \wedge \eta)_x = \omega_x \wedge \eta_x$. Für $f \in \Omega^0$ eine Funktion schreiben wir meist $f\eta$ statt $f \wedge \eta$.

10.2.7. Ist speziell $X = \mathbb{R}^n$ und sind $x_i : \mathbb{R}^n \rightarrow \mathbb{R}$ die Koordinatenfunktionen, so läßt sich nach 10.1.3 jede p -Form $\omega \in \Omega_X^p(U)$ eindeutig schreiben in der Gestalt

$$\omega = \sum_{|I|=p} a_I dx_I$$

Hier läuft die Summe wie angedeutet über alle p -elementigen Teilmengen $I \subset \{1, \dots, n\}$, die Koeffizienten a_I sind reelle Funktionen auf U , und dx_I ist ähnlich wie in 10.1.1 eine Abkürzung für $dx_I = dx_{i_1} \wedge \dots \wedge dx_{i_p}$ mit $i_1 < \dots < i_p$ den der Größe nach geordneten Elementen von I . Diese Notation ist allerdings mit Vorsicht zu genießen, denn natürlich ist dx_I in keinster Weise das Differential einer wie auch immer gearteten Funktion x_I . Das Dachprodukt zweier in dieser Standarddarstellung gegebenen Formen ergibt sich dann leicht mittels der Regeln $dx_i \wedge dx_i = 0$ und $dx_i \wedge dx_j = -dx_j \wedge dx_i$.

Definition 10.2.8. Gegeben endlichdimensionale reelle Räume X, Y und eine stetig differenzierbare Abbildung $\phi : A \rightarrow Y$ auf einer halboffenen Teilmenge $A \subset X$ definieren wir das **Zurückholen von Differentialformen**, eine \mathbb{R} -lineare Abbildung $\phi^* : \Omega_Y^p(\phi(A)) \rightarrow \Omega^p(A)$, durch die Vorschrift

$$(\phi^*\omega)_x = (d_x\phi)^\top(\omega_{\phi(x)})$$

Hier bezeichnet $(d_x\phi)^\top : \text{Alt}^p \vec{Y} \rightarrow \text{Alt}^p \vec{X}$ die vom Differential $d_x\phi : \vec{X} \rightarrow \vec{Y}$ von ϕ an der Stelle $x \in A$ induzierte Abbildung. Alternativ könnten wir auch schreiben $(\phi^*\omega)_x = \omega_{\phi(x)} \circ (d_x\phi \times \dots \times d_x\phi)$ mit p Faktoren ganz rechts.

10.2.9. Das Zurückholen von Funktionen alias Nullformen mit einer Abbildung ist schlicht das “Vorschalten” von besagter Abbildung, in Formeln $\phi^*(g) = g \circ \phi$ für eine Funktion $g : B \rightarrow \mathbb{R}$. Das Zurückholen von 1-Formen haben wir bereits in 6.1.17 diskutiert. Wir verallgemeinern die dort eingeführte Terminologie auf den vorliegenden Fall und nennen Differentialformen η und ω **verwandt unter** ϕ und schreiben $\phi : \eta \rightsquigarrow \omega$ genau dann, wenn gilt $\phi^*(\omega) = \eta$.

Lemma 10.2.10. *Für das Zurückholen von Differentialformen gilt die Kettenregel, d.h. wir haben stets $\text{id}^* = \text{id}$ und*

$$\psi^*(\phi^*\omega) = (\phi \circ \psi)^*(\omega)$$

Beweis. Das folgt mit der üblichen Kettenregel 4.3.1 sofort aus den Definitionen. Wir können die Aussage des Lemmas auch im Sinne von 6.1.18 dahingehend verstehen, daß Verwandtschaft transitiv ist. \square

Lemma 10.2.11. *Verwandschaft alias das Zurückholen ϕ^* von Differentialformen ist verträglich mit dem Dachprodukt, in Formeln gilt also*

$$\phi^*(\omega \wedge \eta) = \phi^*(\omega) \wedge \phi^*(\eta)$$

Beweis. Dem Leser überlassen. \square

Beispiel 10.2.12. Wir erinnern 6.1.21. Für $X = \mathbb{R}^n$ mit Koordinaten x_1, \dots, x_n und $Y = \mathbb{R}^m$ mit Koordinaten y_1, \dots, y_m und $\phi = (\phi_1, \dots, \phi_n)$ eine differenzierbare Abbildung von einer halboffenen Teilmenge von \mathbb{R}^m in eine halboffene Teilmenge von \mathbb{R}^n ergibt sich $\phi^*(dx_i) = d(\phi^*x_i) = d\phi_i = \sum_j \frac{\partial \phi_i}{\partial y_j} dy_j$. Folglich kann das Zurückholen von 1-Formen in Koordinaten beschrieben werden durch die Formel

$$\phi^*\left(\sum_i a_i dx_i\right) = \sum_{i,j} (a_i \circ \phi) \frac{\partial \phi_i}{\partial y_j} dy_j$$

Beispiel 10.2.13. Ist ϕ die Polarkoordinatenabbildung

$$\begin{aligned} \phi : \mathbb{R}^2 &\rightarrow \mathbb{R}^2 \\ (r, \vartheta) &\mapsto (r \cos \vartheta, r \sin \vartheta) \end{aligned}$$

und haben wir auf \mathbb{R}^2 die 1-Form $y dx$ gegeben, so wird sie zurückgeholt zu

$$\begin{aligned} \phi^*(y dx) &= \phi^*(y)\phi^*(dx) \\ &= r \sin \vartheta d(r \cos \vartheta) \\ &= r \sin \vartheta \cos \vartheta dr - r^2 \sin^2 \vartheta d\vartheta \end{aligned}$$

und für die 2-Form $dx \wedge dy$ erhalten wir

$$\begin{aligned}\phi^*(dx \wedge dy) &= \phi^*(dx) \wedge \phi^*(dy) \\ &= d(r \cos \vartheta) \wedge d(r \sin \vartheta) \\ &= (\cos \vartheta \, dr - r \sin \vartheta \, d\vartheta) \wedge (\sin \vartheta \, dr + r \cos \vartheta \, d\vartheta) \\ &= r \, dr \wedge d\vartheta\end{aligned}$$

Lemma 10.2.14. *Für A halboffen in \mathbb{R}^n und $\phi : A \rightarrow \mathbb{R}^n$ stetig differenzierbar gilt stets*

$$\phi^*(dx_1 \wedge \dots \wedge dx_n) = (\det d\phi) dx_1 \wedge \dots \wedge dx_n$$

Beweis. Für jeden Endomorphismus L eines n -dimensionalen Vektorraums V ist die induzierte Abbildung $L^\top : \text{Alt}^n V \rightarrow \text{Alt}^n V$ nach 10.1.9 gerade die Multiplikation mit $\det L$. \square

10.3 Das Maß einer Differentialform

Proposition 10.3.1 (Das Maß einer Differentialform). *Gegeben eine k -dimensionale Untermannigfaltigkeit M eines endlichdimensionalen reellen Raums und eine meßbare k -Form ω auf M gibt es genau ein topologisches Maß $|\omega|$ auf M mit der Eigenschaft, daß für jede Karte (W, φ) von M und jede Borelmenge $A \subset \varphi(W)$ gilt*

$$|\omega|(A) = \int_{\varphi^{-1}(A)} |(\varphi^*\omega)(e_1, \dots, e_k)|$$

Jede in M enthaltene Mannigfaltigkeit echt kleinerer Dimension ist für dieses Maß eine Nullmenge und jede stetige k -Form liefert ein Borelmaß.

10.3.2. Die Notation $|\omega|$ hatten wir eigentlich bereits vereinbart für den Grad einer Differentialform, also $|\omega| = p$ im Fall einer p -Form. Es ist also a priori ungeschickt, dieselbe Notation noch für ein völlig anderes Konzept zu verwenden. Andererseits sind beide Notationen üblich, und welche Bedeutung im Einzelfall gemeint ist, kann der Leser leicht aus dem Kontext erschließen: Im Wesentlichen tritt $|\omega|$ in der Bedeutung als Grad fast nur im Exponenten von (-1) auf, und $|\omega|$ in der Bedeutung als Maß nie im Exponenten von was auch immer.

10.3.3. Auf der rechten Seite meint e_1, \dots, e_k die Standardbasis des \mathbb{R}^k , auf der also unsere k -Form $\varphi^*\omega$ an jeder Stelle ausgewertet werden soll. Für die 2-Form $\eta = y^2 \sin x \, dx \wedge dy$ auf dem \mathbb{R}^2 etwa wäre $\eta(e_1, e_2)$ die Funktion $y^2 \sin x$. Das Integral auf der rechten Seite ist als Integral in Bezug auf das Lebesgue-Maß der meßbaren reellwertigen Funktion $p \mapsto |(\varphi^*\omega)_p(e_1, \dots, e_k)|$ über die meßbare Menge $\varphi^{-1}(A) \subset W \subset \mathbb{R}^k$ zu verstehen. Mit einer k -Form meinen wir vorerst noch eine relative k -Form. Sobald wir die wirklichen k -Formen auf Mannigfaltigkeiten kennenlernen, wird dieselbe Definition jedoch auch für diese sinnvoll und richtig werden.

10.3.4. Ich will einen wesentlichen Unterschied zum in 9.9.1 eingeführten Flächenmaß hervorheben: Das Maß zu einer Differentialform können wir auf Untermannigfaltigkeiten beliebiger endlichdimensionaler Räume einführen, wohingegen wir das Flächenmaß nur auf Untermannigfaltigkeiten des \mathbb{R}^n erklärt haben und bestenfalls auf Untermannigfaltigkeiten endlichdimensionaler euklidischer Räume hätten erklären können.

Beispiel 10.3.5. Das Lebesgue-Maß auf \mathbb{R}^k ist das Maß zu $dx_1 \wedge \dots \wedge dx_k$, in Formeln $d^k x = |dx_1 \wedge \dots \wedge dx_k|$. Speziell ist $|dx|$ das Lebesgue-Maß auf \mathbb{R} . In diesem Fall erlauben wir uns aus den bereits in 9.2.15 dargelegten Gründen auch die Notation dx . Allgemeiner erhalten wir für $f : \mathbb{R}^n \rightarrow \mathbb{R}$ meßbar die Gleichheit von Maßen $|f dx_1 \wedge \dots \wedge dx_k| = |f| d^k x$.

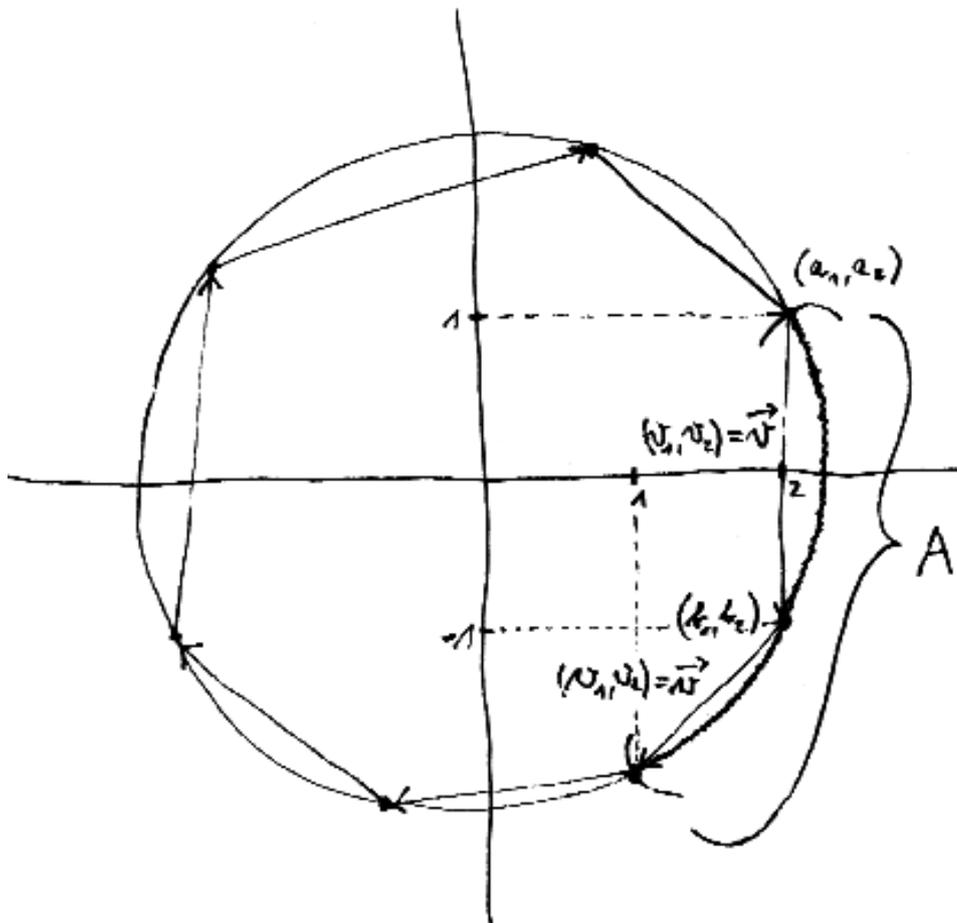
Beispiel 10.3.6. Ist M eine 0-Mannigfaltigkeit, als da heißt eine diskrete Teilmenge eines endlichdimensionalen Raums, und ω eine relative 0-Form alias eine reellwertige Funktion f auf M , so ist $|\omega|$ das Maß, das jedem Punkt $p \in M$ als Maß den Betrag des Funktionswerts $|\omega_p| = |f(p)|$ zuordnet.

Beispiel 10.3.7. Wir betrachten als Differentialform die 1-Form $\omega = dx$ auf \mathbb{R}^2 und als Untermannigfaltigkeit die Kreislinie $S^1 = \{(x, y) \mid x^2 + y^2 = 1\}$. In diesem Fall stimmt das Maß $|dx|$ auf S^1 überein mit der Summe der Bildmaße μ_{\pm} des Lebesgue-Maßes auf dem Intervall $[-1, 1]$ unter seinen vertikalen Projektionen auf den oberen bzw. auf den unteren Halbkreis. Insbesondere hätten wir also $|dx|(S^1) = 4$. In der Tat, betrachten wir etwa die Karte $\varphi : (-1, 1) \rightarrow S^1$, $t \mapsto (t, \sqrt{1-t^2})$ und eine Borelmenge A in ihrem Bild, dem offenen oberen Halbkreis, so ergibt sich mit unseren Definitionen

$$|dx|(A) = \int_{\varphi^{-1}(A)} |\varphi^*(dx)(e_1)| = \int_{\varphi^{-1}(A)} |dt(e_1)| = \int_{\varphi^{-1}(A)} 1 = \lambda(\varphi^{-1}(A))$$

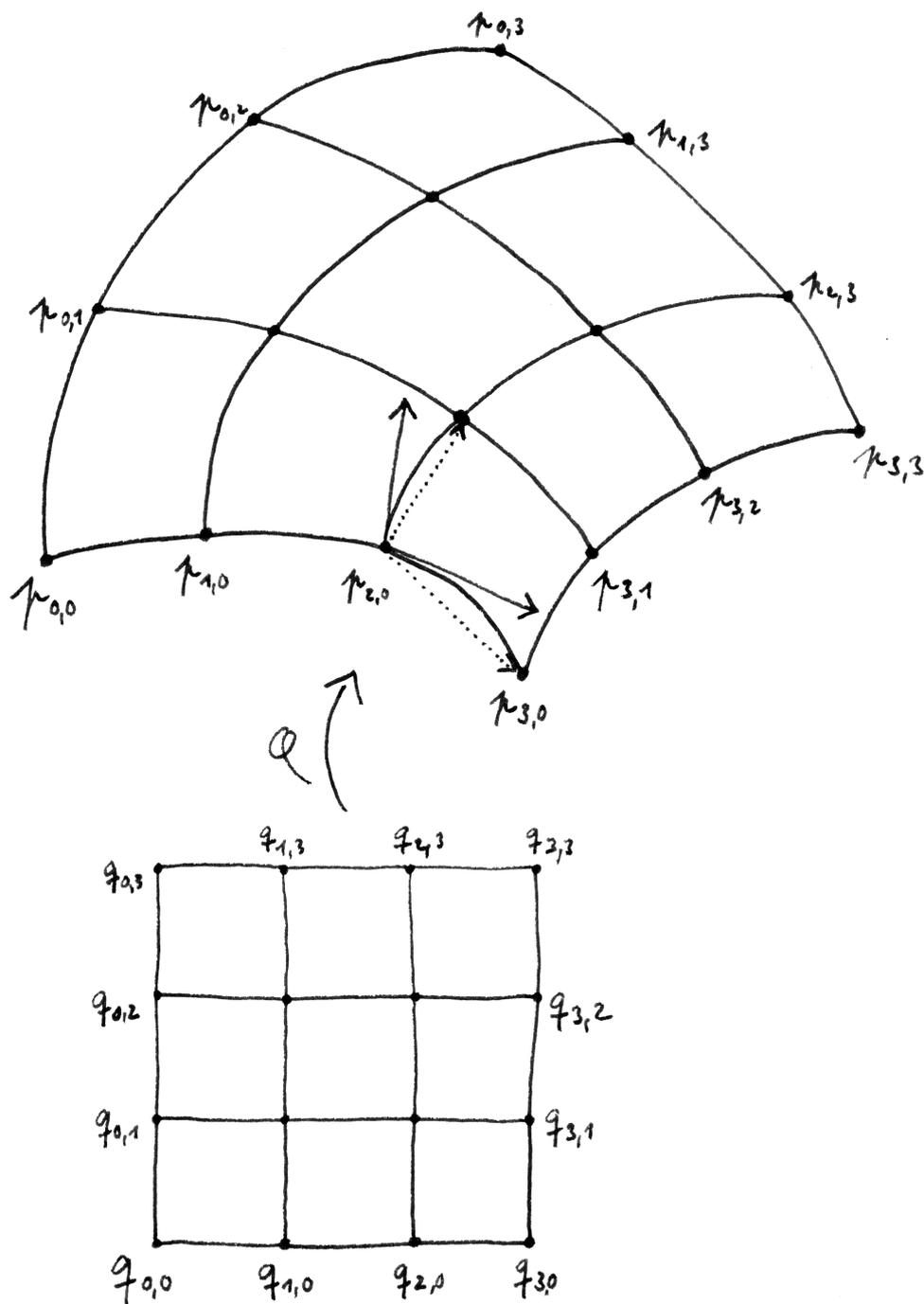
für λ das Lebesguemaß auf $[-1, 1]$.

10.3.8. Etwas vage mag man sich im Fall einer 2-Mannigfaltigkeit alias Fläche im Anschauungsraum vorstellen, daß wir unsere Fläche lokal durch “parallelogrammförmige Schuppen” approximieren, denen wir mithilfe unserer 2-Form jeweils ein Maß zuordnen können, indem wir an jeder Stelle die beiden Kantenvektoren unserer parallelogrammförmigen Schuppen in die der besagten Stelle zugeordnete alternierende bilineare Abbildung einsetzen und vom Resultat den Betrag nehmen. Dann gilt es, mit immer feineren Schuppen in geeigneter Weise zum Grenzwert überzugehen. Konkreter erkläre ich nun eine Interpretation durch Riemannsummen. Sei dazu (W, φ) eine orientierte Karte einer der Einfachheit der Notation halber zweidimensionalen Untermannigfaltigkeit $M \subset X$, sei $Q = [a, b] \times [c, d] \subset W$ ein Rechteck und $P = \varphi(Q) \subset M$ sein Bild. Sei



In diesem Bild soll angedeutet werden, welches Maß auf der Kreislinie eine
 Einsform auf der Papierebene liefert. Denken wir uns etwa die Einsform $y dx$. Für
 eine meßbaren Menge wie dem eingezeichneten A dürfen wir uns das Maß so
 vorstellen: Wir approximieren durch einen Polygonzug und summieren die
 Beträge der Werte unserer Differentialform auf den Kantenvektoren. Eine

Approximation von $|y dx|(A)$ wäre also etwa
 $|a_2 v_1| + b_2 w_1 = |1 \cdot 0| + |(-1) \cdot (-1)| = 1.$



Die gepunkteten Pfeile stellen die Vektoren $p_{3,0} - p_{2,0}$ und $p_{2,1} - p_{2,0}$ dar, der Betrag des Werts von $\omega_{p_{2,0}}$ auf diesem Paar von Vektoren geht in die Riemannsumme S_P^3 ein.

weiter $\omega : M \rightarrow \text{Alt}^2(\vec{X})$ eine stetige relative 2-Form auf M . Wir betrachten für $r \geq 1$ die äquidistanten Unterteilungen $a = a_0 < a_1 < \dots < a_r = b$, $c = c_0 < c_1 < \dots < c_r = d$ der Kanten von Q in jeweils r Segmente, bezeichnen mit $q_{i,j} = (a_i, c_j)$ die Gitterpunkte im so gegebenen Raster auf Q , und mit $p_{i,j} = \varphi(q_{i,j})$ die Bilder dieser Gitterpunkte in $P \subset M$. Dann definieren wir die r -te **Riemannsumme** $S_P^r |\omega|$ durch die Formel

$$S_P^r |\omega| = \sum_{i,j=0}^{r-1} |\omega_{p_{i,j}}(p_{i+1,j} - p_{i,j}, p_{i,j+1} - p_{i,j})|$$

Natürlich hängt diese Summe von der Karte (W, φ) ab, auch wenn das in der Notation nicht zum Ausdruck kommt. Wir können nun das Maß $|\omega|(P)$ interpretieren als den Grenzwert

$$|\omega|(P) = \lim_{r \rightarrow \infty} S_P^r |\omega|$$

Den Beweis dieser Tatsache entlang der Grundlinie des Beweises von 7.5.9 überlassen wir dem Leser zur Übung.

Beweis von 10.3.1. Der Beweis von 9.9.1 kann fast Wort für Wort wiederholt werden, wobei wir nur für jede Karte (W, φ) auf W statt der Funktionen $\text{vol}(d_x \varphi)$ die Beträge der Funktionen ω_φ mit $\omega_\varphi(x) = (\varphi^* \omega)_x(e_1, \dots, e_k)$ einsetzen müssen. Der entscheidende Schritt besteht dann wieder darin, für je zwei Karten (W, φ) und (V, ψ) mit $\varphi(W) = \psi(V)$ und Kartenwechsel $g = \psi^{-1} \circ \varphi : W \xrightarrow{\sim} V$ die Formel

$$|\omega_\varphi(x)| = |\det d_x g| |\omega_\psi(x)|$$

zu zeigen. Per definitionem haben wir jedoch $\varphi^* \omega = \omega_\varphi dx_1 \wedge \dots \wedge dx_k$ und ebenso $\psi^* \omega = \omega_\psi dx_1 \wedge \dots \wedge dx_k$ und wegen $\psi \circ g = \varphi$ gilt $g^*(\psi^* \omega) = \varphi^* \omega$. Mit 10.2.14 folgt dann sofort $(\det d_x g) \omega_\psi(g(x)) = \omega_\varphi(x)$ für alle $x \in W$, und nehmen wir hier auf beiden Seiten Beträge, so steht unsere Behauptung auch schon da. \square

Übung 10.3.9. Wir betrachten auf \mathbb{R}^2 die Funktionen $f(x, y) = \sin(xy)$ und $g(x, y) = x^2 y + y$ und die 2-Form $\omega = df \wedge dg$. Man schreibe das Maß $|\omega|$ auf \mathbb{R}^2 als Vielfaches des Lebesgue-Maßes.

Übung 10.3.10. Seien U, V offene Teilmengen n -dimensionaler reeller Räume und sei $\phi : U \xrightarrow{\sim} V$ ein \mathcal{C}^1 -Diffeomorphismus. Gegeben eine meßbare n -Form ω auf V zeige man die Verwandtschaft von Maßen $\phi : |\phi^* \omega| \rightsquigarrow |\omega|$. In anderen Worten folgt also aus der Verwandtschaft von meßbaren n -Formen $\phi : \eta \rightsquigarrow \omega$ die Verwandtschaft von Maßen $\phi : |\eta| \rightsquigarrow |\omega|$. Hinweis: Das ist bei rechtem Lichte besehen nur eine Umformulierung der Transformationsformel 9.8.1.

Übung 10.3.11. Die Kovektorfelder $d\vartheta$ und dr auf \mathbb{R}^2 zu unseren Polarkoordinaten aus 6.1.23 liefern Maße auf dem Kreis $K \subset \mathbb{R}^2$ mit Radius 4 und auf der Parallelen $G \subset \mathbb{R}^2$ zur y -Achse durch den Punkt $(1, 0)$. Man finde stetige Funktionen auf diesen Mannigfaltigkeiten derart, daß ihre Produkte mit den jeweiligen Flächenmaßen die Maße zu unseren Differentialformen liefern.

10.4 Orientierung von Mannigfaltigkeiten

Definition 10.4.1. Gegeben eine Untermannigfaltigkeit $M \subset X$ eines endlichdimensionalen reellen Raums und ein Punkt $p \in M$ definieren wir den **Tangententialraum an M in p** als den Vektorraum

$$T_p M := \text{im}(d_u \varphi) \subset \vec{X}$$

für eine und jede Karte $\varphi : U \rightarrow M$ mit $\varphi(u) = p$. Wir haben also $\dim T_p M = \dim M$, und $p + T_p M$ ist der affine Teilraum von X , der anschaulich gesprochen “ M am besten approximiert bei p ”.

Übung 10.4.2. Man zeige: Gegeben eine Untermannigfaltigkeit $M \subset X$ eines endlichdimensionalen reellen Raums und ein Punkt $p \in M$ kann der Tangentialraum $T_p M$ auch beschrieben werden als die Menge aller möglichen Geschwindigkeitsvektoren bei p von in M verlaufenden und bei p differenzierbaren Wegen.

Ergänzung 10.4.3. Man will sich meist die verschiedenen Tangentialräume als paarweise disjunkt denken, ändert die obige Definition deshalb ab und setzt formal

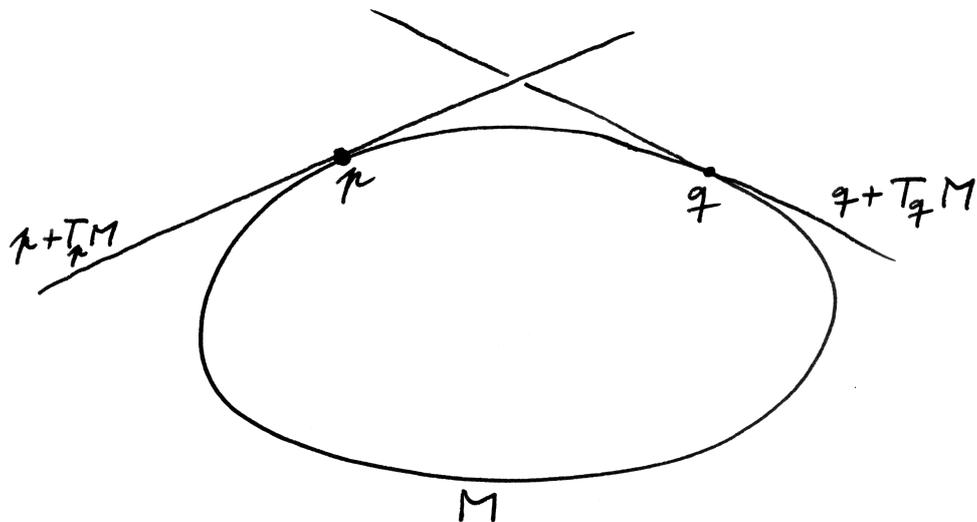
$$T_p M := \{p\} \times \text{im}(d_u \varphi) \subset \{p\} \times \vec{X}$$

So kann man dann das **Tangententialbündel** von M definieren als

$$TM := \bigcup_{p \in M} T_p M \subset M \times \vec{X}$$

Unter geeigneten zusätzlichen Differenzierbarkeitsannahmen an unsere Untermannigfaltigkeit M kann man zeigen, daß $TM \subset X \times \vec{X}$ eine Untermannigfaltigkeit der Dimension $2(\dim M)$ ist, vergleiche ???. Die einzelnen Tangentialräume erhält man als die Fasern der Projektion $\pi : TM \rightarrow M$ des Tangentialbündels auf die Mannigfaltigkeit, in Formeln $T_p M = \pi^{-1}(p)$.

10.4.4. Wir erinnern daran, daß nach ??? eine **Orientierung** eines endlichdimensionalen Vektorraums V über einem angeordneten Körper eine Vorschrift ε ist, die jeder angeordneten Basis B unseres Vektorraums ein Vorzeichen $\varepsilon(B) \in \{+1, -1\}$ zuordnet und zwar so, daß für je zwei angeordnete Basen B, B' die Determinante der Basiswechsellmatrix das Vorzeichen $\varepsilon(B)\varepsilon(B')$ hat. In ??? werden auch noch weitere Begriffsbildungen in diesem Zusammenhang formal eingeführt, deren Bedeutung Sie aber auch leicht selbst werden erraten können.



In diesem Bild habe ich zu einer eindimensionalen Untermannigfaltigkeit der Ebene zwei affine Räume eingezeichnet, deren Richtungsräume ihre Tangentialräume an den beiden fett eingezeichneten Punkten wären. Diese affinen Räume schneiden sich natürlich und ihre Richtungsräume schneiden sich desgleichen. Im bildlich dargestellten Fall besteht dieser Schnitt der Richtungsräume aus dem Nullvektor, aber im allgemeinen kann er auch größer sein. Ich habe die beiden Geraden dennoch als nicht-schneidend gemalt, um bildlich anzudeuten, daß alle diese Überschneidungen von uns bei der Definition des Tangentialbündels sozusagen wegdefiniert werden.

Definition 10.4.5. Eine **Orientierung** einer k -Mannigfaltigkeit M ist eine Vorschrift, die jedem Punkt $p \in M$ eine Orientierung im Sinne von ?? des Tangentialraums $T_p M$ zuordnet und zwar so, daß es um jeden Punkt eine Karte $\varphi : W \rightarrow M$ gibt mit der Eigenschaft, daß die Differentiale $d_x \varphi : \mathbb{R}^k \xrightarrow{\sim} T_{\varphi(x)} M$ für $x \in W$ entweder alle orientierungserhaltend oder alle orientierungsumkehrend sind.

10.4.6. Eine Orientierung einer nulldimensionalen Mannigfaltigkeit M anzugeben bedeutet schlicht, eine Abbildung $\varepsilon : M \rightarrow \{+1, -1\}$ anzugeben, deren Wert bei $p \in M$ eben das Vorzeichen der angeordneten Basis \emptyset des Tangentialraums $T_p M$ ist.

Definition 10.4.7. Unter einer **orientierten Mannigfaltigkeit** versteht man ein Paar bestehend aus einer Mannigfaltigkeit M und einer Orientierung auf M . Ich notiere orientierte Mannigfaltigkeiten oft mit einem Pfeil, etwa als \overrightarrow{M} , aber das ist nicht allgemein üblich. Eine Mannigfaltigkeit, die mindestens eine Orientierung zuläßt, nennt man eine **orientierbar**. Das ‘‘Möbiusband’’ ist ein Beispiel für eine nicht orientierbare 2-Mannigfaltigkeit in \mathbb{R}^3 .

10.4.8. Den Pfeil über einem Symbol benutze ich auch bei affinen Räumen als Notation für den zugehörigen Raum von Richtungsvektoren, vergleiche ?. Was im Einzelfall gemeint ist, muß der Leser aus dem Kontext erschließen.

Definition 10.4.9. Wir sagen, eine Karte (W, φ) einer orientierten Mannigfaltigkeit habe **die Orientierung** ε für $\varepsilon \in \{+1, -1\}$ genau dann, wenn für jeden Punkt $x \in W$ das Bild der Standardbasis unter dem Isomorphismus $d_x \varphi : \mathbb{R}^k \xrightarrow{\sim} T_{\varphi(x)} M$ die Orientierung ε hat. Nur die leere Karte kann beide Orientierungen haben. Karten, die nicht zusammenhängend sind, besitzen im allgemeinen weder die Orientierung $+1$ noch die Orientierung -1 . Eine nichtleere Karte der Orientierung $+1$ nennen wir eine **positiv orientierte Karte**.

Übung 10.4.10. Jede orientierbare zusammenhängende Mannigfaltigkeit M besitzt genau zwei Orientierungen. Hinweis: Gegeben zwei Orientierungen ist die Menge aller Punkte p , an denen sie dieselbe Orientierung von $T_p M$ liefern, ebenso offen wie die Menge aller Punkte p , an denen sie verschiedene Orientierungen von $T_p M$ liefern. Nun verwende man 6.4.16.

Ergänzende Übung 10.4.11. Seien X und Y endlichdimensionale reelle Räume, $U \subseteq X$ eine offene Teilmenge und $f : U \rightarrow Y$ eine stetig differenzierbare Abbildung mit überall surjektivem Differential. So ist für alle $c \in Y$ nach 7.3.10 das Urbild $M = f^{-1}(c)$ eine Untermannigfaltigkeit von X der Dimension $\dim X - \dim Y$. Man zeige für alle $p \in M$ die Formel $T_p M = \ker d_p f$.

10.5 Integration von Differentialformen

Definition 10.5.1. Sei M eine k -dimensionale orientierte Untermannigfaltigkeit eines endlichdimensionalen reellen Raums. Wir nennen eine meßbare k -Form ω auf M **nichtnegativ** genau dann, wenn für alle Punkte $p \in M$ und jede angeordnete Basis v_1, \dots, v_k der Orientierung ε des Tangentialraums $T_p M$ gilt $\varepsilon \omega_p(v_1, \dots, v_k) \geq 0$.

Definition 10.5.2. Wir nennen eine k -Form ω auf einer k -dimensionalen Untermannigfaltigkeit M eines endlichdimensionalen reellen Raums **integrierbar über M** genau dann, wenn sie meßbar ist und wenn für das nach 10.3.1 zugehörige Maß $|\omega|$ gilt $|\omega|(M) < \infty$.

Satz 10.5.3 (Integration von Differentialformen). Sei M eine k -dimensionale orientierte Untermannigfaltigkeit eines endlichdimensionalen reellen Raums. So bilden die über M integrierbaren k -Formen einen Untervektorraum im Raum aller k -Formen auf M , und es gibt auf diesem Untervektorraum genau eine Linearform $\omega \mapsto \int_{\vec{M}} \omega$ derart, daß für alle nichtnegativen k -Formen ω gilt

$$\int_{\vec{M}} \omega = |\omega|(M)$$

Beweis. Daß die integrierbaren k -Formen einen Untervektorraum bilden, scheint mir offensichtlich. Um die Eindeutigkeit der fraglichen Linearform zu zeigen, betrachten wir für eine beliebige meßbare k -Form ω auf M die meßbare Teilmenge

$$M^+ = \left\{ p \in M \left| \begin{array}{l} \text{Es gilt } \omega_p(v_1, \dots, v_k) > 0 \text{ für eine} \\ \text{und jede orientierte angeordnete} \\ \text{Basis } v_1, \dots, v_k \text{ von } T_p M \end{array} \right. \right\}$$

Unsere meßbare k -Form ω können wir nun schreiben als die Differenz $\omega = \omega^+ - \omega^-$ zweier nichtnegativer Formen, indem wir etwa ω^+ erklären durch $\omega_p^+ = \omega_p$ für $p \in M^+$ und $\omega_p^+ = 0$ sonst. Wir müssen also setzen

$$\int_{\vec{M}} \omega = |\omega^+|(M) - |\omega^-|(M)$$

und es bleibt nur zu zeigen, daß diese Vorschrift auch tatsächlich eine Linearform liefert. Hierbei ist nur die Additivität problematisch. Für je zwei integrierbare nichtnegative Formen ω und η gilt jedoch offensichtlich

$$\int_{\vec{M}} \omega + \eta = \int_{\vec{M}} \omega + \int_{\vec{M}} \eta$$

Im allgemeinen schreiben wir nun $\omega = \omega^+ - \omega^-$, $\eta = \eta^+ - \eta^-$ und $\omega + \eta = \rho = \rho^+ - \rho^-$ und folgern durch Einsetzen $\omega^+ + \eta^+ + \rho^- = \omega^- + \eta^- + \rho^+$. Wenden wir darauf die Additivität des Integrals für nichtnegative Formen an, so folgt sofort die Additivität des Integrals für beliebige integrierbare Formen. \square

Proposition 10.5.4 (Integration in lokalen Koordinaten). *Sei M eine k -dimensionale orientierte Untermannigfaltigkeit eines endlichdimensionalen reellen Raums, $\varphi : W \rightarrow M$ eine Karte der Orientierung ε und ω eine meßbare k -Form auf M , die außerhalb von $\varphi(W)$ verschwindet. So ist ω integrierbar genau dann, wenn die reellwertige Funktion $(\varphi^*\omega)(e_1, \dots, e_k)$ integrierbar ist auf W , und in diesem Fall gilt*

$$\int_{\vec{M}} \omega = \varepsilon \int_W (\varphi^*\omega)(e_1, \dots, e_k)$$

Beweis. Für nichtnegative Formen ω ist das im Wesentlichen die Definition, für beliebige Formen folgt es mithilfe unserer Zerlegung $\omega = \omega^+ - \omega^-$ aus dem vorhergehenden Beweis. \square

10.5.5. Unter den Annahmen und Notationen von 10.5.4 und wenn wir W mit der von \mathbb{R}^k induzierten Orientierung versehen, gilt für unsere integrierbare k -Form ω auch

$$\int_{\vec{M}} \omega = \varepsilon \int_{\vec{W}} \varphi^*\omega$$

In der Tat folgt das unmittelbar aus den Definitionen, wenn wir das Differentialformenintegral auf der rechten Seite mittels der Karte $\text{id} : W \xrightarrow{\sim} W$ berechnen. Ebenso folgt, daß ω integrierbar ist genau dann, wenn $\varphi^*\omega$ integrierbar ist.

Beispiel 10.5.6 (Funktionenintegral als Differentialformenintegral). Eine n -Form η auf einer offenen Teilmenge $W \subseteq \mathbb{R}^n$ hat die Gestalt $\eta = f dx_1 \wedge \dots \wedge dx_n$ für eine wohlbestimmte reellwertige Funktion $f : W \rightarrow \mathbb{R}$, eben der Funktion f gegeben durch $f(x) = \eta_x(e_1, \dots, e_n)$. Geben wir unserer offenen Teilmenge die von der Standardorientierung des \mathbb{R}^n induzierte Orientierung, so ist unsere n -Form η integrierbar über W genau dann, wenn die Funktion f integrierbar ist über W in Bezug auf das Lebesgue-Maß λ^n , und in diesem Fall gilt unter Verwendung unserer ganzen verschiedenen Notationen

$$\int_{\vec{W}} \eta = \int_{\vec{W}} f dx_1 \wedge \dots \wedge dx_n = \int_W f(x) d^n x = \int_W f \lambda^n$$

Beispiel 10.5.7. Wir berechnen das Integral der 2-Form $x^2 dx \wedge dy$ über die obere Hemisphäre $H = \{(x, y, z) \mid x^2 + y^2 + z^2 = 1, z > 0\}$ mit der Orientierung, für die die beiden ersten Vektoren e_1, e_2 der Standardbasis des \mathbb{R}^3 in dieser Reihenfolge eine orientierte Basis des Tangentialraums am Pol $T_{(0,0,1)}H$ bilden. Wir

betrachten das offene Rechteck $R = (0, \pi) \times (0, \pi) \subset \mathbb{R}^2$ und die orientierte Karte $\phi : R \rightarrow H$, $(\vartheta, \varphi) \mapsto (\cos \vartheta, \cos \varphi \sin \vartheta, \sin \varphi \sin \vartheta)$, anschaulich gesprochen eine “liegende Version” unserer Kugelkoordinaten aus 6.2.11, und erhalten

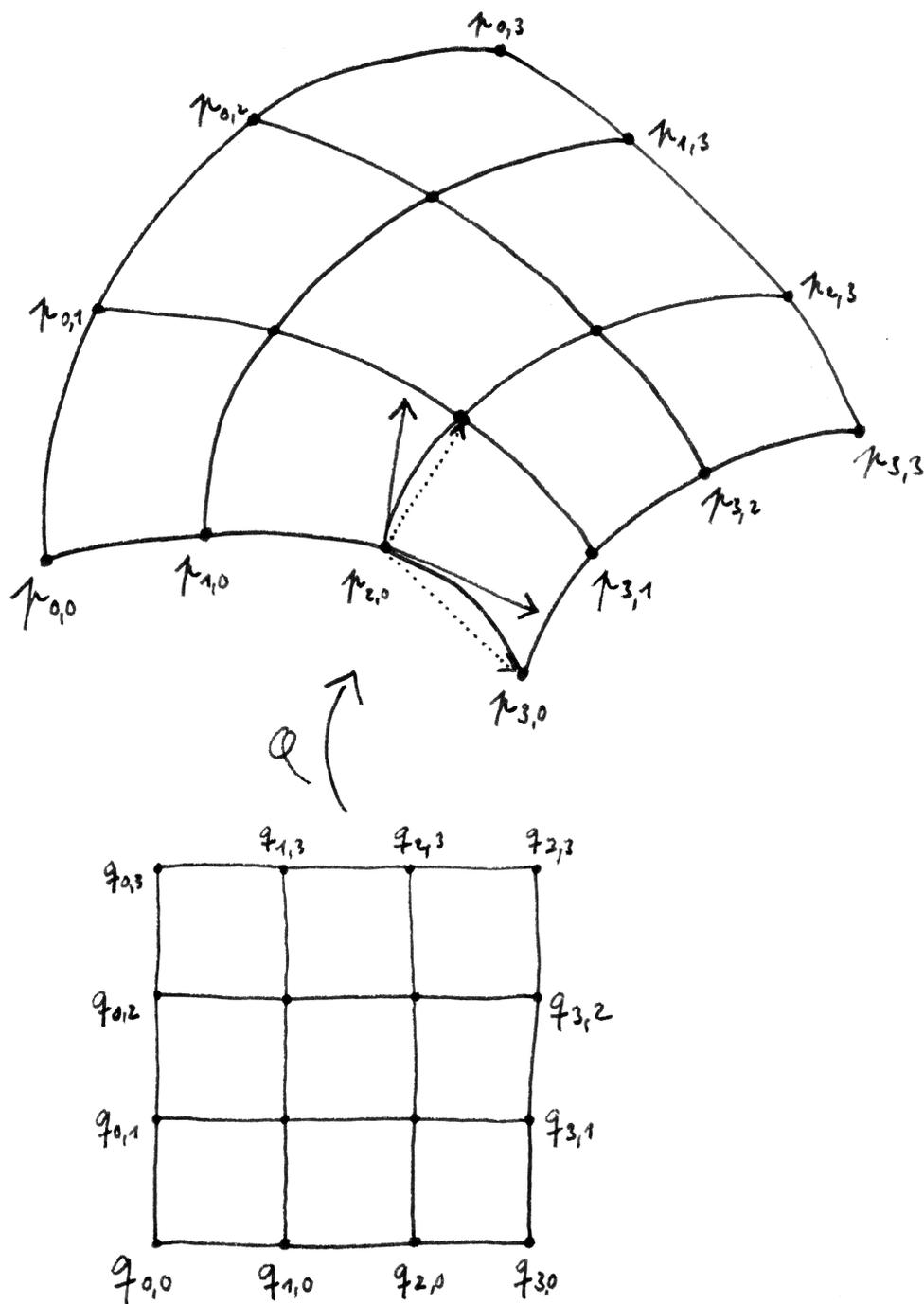
$$\begin{aligned}
 \int_{\bar{H}} x^2 dx \wedge dy &= \int_{\bar{R}} \cos^2 \vartheta d(\cos \vartheta) \wedge d(\cos \varphi \sin \vartheta) \\
 &= \int_{\bar{R}} \cos^2 \vartheta (-\sin \vartheta d\vartheta) \wedge (\cos \varphi \cos \vartheta d\vartheta - \sin \varphi \sin \vartheta d\varphi) \\
 &= \int_{\bar{R}} \cos^2 \vartheta \sin^2 \vartheta \sin \varphi d\vartheta \wedge d\varphi \\
 &= \int_R \cos^2 \vartheta \sin^2 \vartheta \sin \varphi d\vartheta d\varphi \\
 &= \int_0^\pi \int_0^\pi \cos^2 \vartheta \sin^2 \vartheta \sin \varphi d\vartheta d\varphi \\
 &= \frac{1}{4} \int_0^\pi \sin^2(2\vartheta) d\vartheta \int_0^\pi \sin \varphi d\varphi = \frac{1}{2} \int_0^\pi \frac{1}{2} - \frac{\cos 4\vartheta}{2} d\vartheta = \frac{\pi}{4}
 \end{aligned}$$

Um diese Rechnung zu rechtfertigen, müssen wir sie von hinten nach vorne lesen: Der Übergang zum Doppelintegral ist erlaubt, da der Integrand nichtnegativ ist und so der positive Fubini greift, der dann hinwiederum erst recht eigentlich die Integrierbarkeit unserer Funktion auf dem Rechteck R und damit die Integrierbarkeit unserer Differentialform auf der Hemisphäre H zeigt. Bei der Rechnung selber ist der erste Schritt Übung 10.5.5 mitsamt dem Vertauschen vom Zurückholen mit Dachprodukt und Differential 6.1.19, der zweite die Formel 6.1.10 für das Differential einer Funktion, der dritte beruht auf dem Alternieren und der Bilinearität des Dachprodukts, und der vierte auf Proposition 10.5.4. Man hätte auch mit 10.5.4 und einer Nebenrechnung direkt die vierte Gleichung erhalten können, aber das schien es mir weniger übersichtlich.

10.5.8. Um die Integration von Differentialformen anschaulich zu machen, erkläre ich ihre Interpretation durch Riemannsummen. Sei dazu (W, φ) eine orientierte Karte einer der Einfachheit der Notation halber zweidimensionalen orientierten Untermannigfaltigkeit \vec{M} eines reellen Raums X , sei $Q = [a, b] \times [c, d] \subset W$ ein Rechteck und sei $\omega : M \rightarrow \text{Alt}^2(\vec{X})$ eine stetige relative 2-Form auf M mit $(\text{supp } \omega) \cap M \subset \varphi(Q)$, wobei der Pfeil über dem X den zugehörigen Richtungsraum meint, der Pfeil über dem M dahingegen eine feste gewählte Orientierung andeutet. Wir betrachten für $r \geq 1$ die äquidistanten Unterteilungen $a = a_0 < a_1 < \dots < a_r = b$, $c = c_0 < c_1 < \dots < c_r = d$ der Kanten von Q in jeweils r Segmente, bezeichnen mit $q_{i,j} = (a_i, c_j)$ die Gitterpunkte im so gegebenen Raster auf Q , und mit $p_{i,j} = \varphi(q_{i,j})$ die Bilder dieser Gitterpunkte in M . Dann definieren wir die r -te **Riemannsumme** $S_{\vec{M}}^r(\omega)$ durch die Formel

$$S_{\vec{M}}^r(\omega) = \sum_{i,j=0}^{r-1} \omega_{p_{i,j}}(p_{i+1,j} - p_{i,j}, p_{i,j+1} - p_{i,j})$$

Natürlich hängt diese Summe von der Karte (W, φ) ab, auch wenn das in der Notation nicht zum Ausdruck kommt. Wir können nun das Integral von ω über M



Die gepunkteten Pfeile stellen die Vektoren $p_{3,0} - p_{2,0}$ und $p_{2,1} - p_{2,0}$ dar, der Wert von $\omega_{p_{2,0}}$ auf diesem Paar von Vektoren, genommen in einer durch die Orientierung gegebenen Reihenfolge, geht in die Riemannsumme $S_{\bar{M}}^3$ ein.

interpretieren als den Grenzwert

$$\int_{\vec{M}} \omega = \lim_{r \rightarrow \infty} S_{\vec{M}}^r(\omega)$$

Den Beweis dieser Tatsache entlang der Grundlinie des Beweises von 7.5.9 überlassen wir dem Leser zur Übung.

10.5.9. Unter der Voraussetzung einer quadratischen Karte, in Formeln $b - a = d - c$, betrachten wir nun die Spiegelung τ an der Hauptdiagonalen und die neue Karte $\varphi \circ \tau$. Sie ist negativ orientiert und ihre Riemannsummen sind dieselben wie die Riemannsummen von φ , wenn man nur in jedem Summanden den ersten und den zweiten Eintrag der bilinearen Abbildung ω vertauscht und das von der negativen Orientierung der Karte herrührende Vorzeichen berücksichtigt. Ist also ω alternierend, so liefert unsere neue Karte dieselben Riemannsummen und dasselbe Integral. Das soll die in unserem Satz enthaltene Aussage veranschaulichen, daß das Integral einer alternierenden Form unabhängig ist von den zur Berechnung gewählten Karten.

10.5.10. Die Integrale von Differentialformen über orientierte Untermannigfaltigkeiten eines \mathbb{R}^n der Dimensionen 0 oder 1 bzw. der Kodimensionen 0 oder 1 werden von den Anwendern oft in anderer Weise interpretiert. Besonders wichtig sind in diesem Zusammenhang die Fälle mit $n \leq 3$.

Beispiel 10.5.11 (Summation als Differentialformenintegral). Im Fall einer kompakten nulldimensionalen Mannigfaltigkeit alias einer endlichen Menge ist das Integral einer Nullform alias Funktion schlicht die Summe der Funktionswerte multipliziert mit den jeweils durch die Orientierung gegebenen Vorzeichen, in Formeln

$$\int_{\vec{M}} f = \sum_{p \in M} \text{sgn}(p) f(p)$$

Im nichtkompakten Fall ist unsere Nullform alias Funktion integrierbar genau dann, wenn die Summe ihrer Werte absolut konvergiert, und dann gilt die obige Formel entsprechend.

Beispiel 10.5.12 (Wegintegral als Differentialformenintegral). Eine 1-Form ω auf dem \mathbb{R}^n hat die Gestalt $\omega = \omega_1 dx_1 + \dots + \omega_n dx_n$. Gegeben eine orientierte 1-Mannigfaltigkeit K verstehen wir ganz allgemein unter einer **orientierten Parametrisierung von K** eine orientierte Karte $\varphi : W \rightarrow K$ mit dichtem Bild, bei der $W \subseteq \mathbb{R}$ ein offenes Intervall ist. Ist $M \subset \mathbb{R}^n$ eine orientierte eindimensionale Untermannigfaltigkeit und $\varphi : [a, b] \rightarrow \mathbb{R}^n$ ein stetig differenzierbarer Weg, dessen Einschränkung auf das offene Intervall (a, b) eine orientierte Parametrisierung von K ist, so fällt das Integral unserer 1-Form ω über unsere eindimensionale orientierte Mannigfaltigkeit K zusammen mit dem Integral der 1-Form ω über den

Weg φ und wird von Anwendern meist geschrieben als das Wegintegral des Vektorfelds $v = (\omega_1, \dots, \omega_n)^\top$ längs φ , in Formeln

$$\int_{\vec{K}} \omega = \int_{\varphi} \omega = \int_a^b \langle v, d\varphi \rangle = \int_a^b v \cdot d\varphi$$

Beispiel 10.5.13. Der Fall der Integration von Differentialformen über Hyperflächen benötigt von den hier explizit behandelten Fällen den größten begrifflichen Aufwand und wird uns bis zum Ende dieses Abschnitts beschäftigen.

Definition 10.5.14. Ist $M \subset \mathbb{R}^{n+1}$ eine Hyperfläche, so gibt es zu jedem Punkt $p \in M$ genau zwei Vektoren der Länge Eins in \mathbb{R}^{n+1} , die auf dem Tangentialraum $T_p M$ senkrecht stehen. Ist M darüber hinaus orientiert, so hat genau ein Vektor N_p von diesen beiden die Eigenschaft, daß für jede angeordnete Basis v_1, \dots, v_n von $T_p M$ der Orientierung ε die Standardorientierung der angeordneten Basis N_p, v_1, \dots, v_n des \mathbb{R}^{n+1} auch ε ist. Wir erhalten so eine stetige Abbildung

$$\begin{aligned} N : M &\rightarrow \mathbb{R}^{n+1} \\ p &\mapsto N_p \end{aligned}$$

Man nennt sie das **orientierte Normalenfeld**.

10.5.15. Wir nummerieren nun die Koordinaten auf dem \mathbb{R}^{n+1} etwas unüblich x_0, x_1, \dots, x_n und ordnen jedem Vektor $F \in \mathbb{R}^{n+1}$ eine alternierende Multilinearform $\omega_F \in \text{Alt}^n(\mathbb{R}^{n+1})$ zu durch die Vorschrift

$$\omega_F(v_1, \dots, v_n) = \det(F|v_1| \dots |v_n)$$

wo rechts die Matrix mit den entsprechenden Spaltenvektoren zu verstehen ist. In derselben Weise ordnen wir auch jedem Vektorfeld F auf \mathbb{R}^{n+1} eine n -Form ω_F zu und erkennen durch das Auswerten an Tupeln der Standardbasis, daß sie geschrieben werden kann in der Gestalt

$$\omega_F = \sum_{i=0}^n (-1)^i F_i dx_0 \wedge \dots \wedge \widehat{dx_i} \wedge \dots \wedge dx_n$$

Im \mathbb{R}^3 entspricht speziell einem Vektorfeld $F = (F_x, F_y, F_z)$ die 2-Form

$$\omega_F = F_x dy \wedge dz + F_y dz \wedge dx + F_z dx \wedge dy$$

wobei die unteren Indizes nicht als partielle Ableitungen mißverstanden werden dürfen, sondern vielmehr die Komponenten unseres Vektorfelds meinen, die wir auch hätten F_1, F_2, F_3 notieren können. Mit diesen ganzen Begriffsbildungen können wir nun formulieren:

Proposition 10.5.16 (Fluß als Differentialformenintegral). Sei $M \subset \mathbb{R}^{n+1}$ eine orientierte Hyperfläche und F ein meßbares relatives Vektorfeld auf M . So ist die zu unserem Vektorfeld F gehörige n -Form $\omega = \omega_F$ integrierbar über M genau dann, wenn das Skalarprodukt von F mit dem orientierten Normalenfeld integrierbar ist über M , und in diesem Fall gilt

$$\int_{\vec{M}} \omega_F = \int_M \langle F, N \rangle \sigma = \int_M F \cdot N \sigma$$

10.5.17. Die Mitte und die rechte Seite unterscheiden sich hier nur in der Notation für das Skalarprodukt. Das Integral des Skalarprodukts unseres Vektorfelds F mit dem orientierten Normalenfeld N heißt der **Fluß des Vektorfelds F durch die orientierte Hyperfläche M** . Dies Oberflächenintegral mag der Anschauung besser zugänglich sein als unser Integral über eine Differentialform. Für das explizite Rechnen ist jedoch die Darstellung als Integral einer Differentialform meist günstiger, wie etwa Beispiel 10.5.7 illustriert.

Beweis. Wir dürfen ohne Beschränkung der Allgemeinheit annehmen, daß es eine positiv orientierte bijektive Karte $\varphi : W \xrightarrow{\sim} M$ gibt. Der Übersichtlichkeit halber schreiben wir unser Vektorfeld F in der Form $p \mapsto F_p$, wo der Index unglücklicherweise eine völlig andere Bedeutung hat als in 10.5.15. Wir zerlegen nun unser Vektorfeld F an jedem Punkt $p \in M$ in einen orthogonalen und einen tangentialen Anteil als $F_p = \langle F_p, N_p \rangle N_p + R_p$ mit $R_p \in T_p M$ und finden für alle $x \in W$

$$\begin{aligned} (\varphi^* \omega)_x(e_1, \dots, e_n) &= \omega_{\varphi(x)}(d_x \varphi(e_1), \dots, d_x \varphi(e_n)) \\ &= \det(F_{\varphi(x)} | d_x \varphi) \\ &= \langle F_{\varphi(x)}, N_{\varphi(x)} \rangle \det(N_{\varphi(x)} | d_x \varphi) \\ &= \langle F_{\varphi(x)}, N_{\varphi(x)} \rangle \text{vol}(d_x \varphi) \end{aligned}$$

wo in der zweiten Zeile die quadratische Matrix gemeint ist, die aus der Jacobi-Matrix zu $d_x \varphi$ entsteht durch Anfügen von $F_{\varphi(x)}$ als erster Spalte. Die Gleichheit der beiden Integrale folgt nun aus den Definitionen. \square

Beispiel 10.5.18. Anschaulich kann man unser Integral aus 10.5.7 also auch als den Fluß durch die obere Hemisphäre des senkrechten Vektorfelds $x^2 e_3$ verstehen, wie wir in 10.5.16 im allgemeinen zeigen werden. In der Notation von dort hätten wir etwa

$$\int_{\vec{H}} x^2 dx \wedge dy = \int_H (x^2 e_3 \cdot N) \sigma$$

Hier meint σ das Oberflächenmaß der Einheitskugel und N das “nach außen weisende Normalenfeld”, das in unserem Fall auch das “orientierte Normalenfeld” nach 10.5.14 ist. Zur Probe rechne ich hier die rechte Seite auch noch direkt aus.

Auf der Einheitssphäre stimmen ja der Ortsvektor und der nach außen weisende Normalenvektor überein, so daß der Rückzug der Funktion $x^2 e_3 \cdot N$ bezüglich unserer Karte $\phi : R \rightarrow H$ die Funktion $\cos^2 \vartheta \sin \varphi \sin \vartheta$ ist. Um nach dem Oberflächenmaß zu integrieren, gilt es die Gram'sche Matrix zu berechnen. In unserem Fall haben wir

$$d\phi = \begin{pmatrix} -\sin \varphi & 0 \\ \cos \varphi \cos \vartheta & -\sin \varphi \sin \vartheta \\ \sin \varphi \cos \vartheta & \cos \varphi \sin \vartheta \end{pmatrix}$$

und die Matrix der Skalarprodukte der Spaltenvektoren ergibt sich zu

$$(d\phi)^\top d\phi = \begin{pmatrix} 1 & 0 \\ 0 & \sin^2 \varphi \end{pmatrix}$$

und die Wurzel aus deren Determinante zu $\sin \vartheta$, so daß wir bei demselben Doppelintegral über $\cos^2 \vartheta \sin^2 \vartheta \sin \varphi$ landen wie in 10.5.7.

Übung 10.5.19. Berechnen Sie den Fluß des Vektorfelds $F : (x, y, z) \mapsto (x, 0, 0)$ durch die Einheitssphäre, die Sie dazu mit einer Orientierung ihrer Wahl versehen mögen.

10.6 Äußere Ableitung von Differentialformen

10.6.1. Ist V ein endlichdimensionaler Vektorraum, so definieren wir für alle $k \geq 0$ eine lineare Abbildung

$$\text{alt} : \text{Hom}(V, \text{Alt}^k V) \rightarrow \text{Alt}^{k+1} V$$

Leser mit den entsprechenden Kenntnissen in multilinearer Algebra können sie unter den in ?? gegebenen Identifikationen $\bigwedge^k(V^*) \xrightarrow{\sim} \text{Alt}^k V$ verstehen als die Komposition $\text{Hom}(V, \text{Alt}^k V) \xrightarrow{\sim} V^* \otimes \bigwedge^k(V^*) \xrightarrow{\wedge} \bigwedge^{k+1}(V^*)$ des Inversen zum Standardisomorphismus $V^* \otimes W \xrightarrow{\sim} \text{Hom}(V, W)$ aus ?? mit dem Dachprodukt. Weniger Gebildete definieren diese Abbildung explizit, indem sie ähnlich wie bei der Konstruktion des Dach-Produkts setzen

$$(\text{alt } f)(v_0, v_1, \dots, v_k) := \sum_{i=0}^k (-1)^i (f(v_i))(v_0, \dots, \widehat{v}_i, \dots, v_k)$$

Hier soll die "Tarnkappe" über v_i wie üblich bedeuten, daß dieser Eintrag beim entsprechenden Summanden auszulassen ist.

Definition 10.6.2. Sei X ein endlichdimensionaler reeller Raum und $A \subset X$ halboffen. Eine Differentialform $\omega \in \Omega^k(A)$ heißt **differenzierbar** bzw. **stetig differenzierbar** genau dann, wenn sie als Abbildung $\omega : A \rightarrow \text{Alt}^k \vec{X}$ von der

halboffenen Teilmenge A des endlichdimensionalen reellen Raums X in den endlichdimensionalen reellen Vektorraum $\text{Alt}^k \vec{X}$ differenzierbar ist im Sinne von 4.2.2 bzw. stetig differenzierbar im Sinne von 4.5.2, wenn also ihr Differential auch stetig ist als Abbildung $A \rightarrow \text{Hom}(\vec{X}, \text{Alt}^k \vec{X})$ gegeben durch $x \mapsto d_x \omega$.

Definition 10.6.3. Sei X ein endlichdimensionaler reeller Raum und $A \subset X$ halboffen. Gegeben $\omega \in \Omega^k(A)$ eine differenzierbare k -Form erklären wir eine $(k+1)$ -Form $d\omega \in \Omega^{k+1}(A)$ durch die Vorschrift

$$(d\omega)_x := \text{alt}(d_x \omega)$$

wo $d_x \omega : V \rightarrow \text{Alt}^k \vec{X}$ das Differential im Sinne von 4.2.2 unserer Form $\omega : A \rightarrow \text{Alt}^k \vec{X}$ an einer Stelle $x \in A$ meint. Wir nennen $d\omega$ die **äußere Ableitung von ω** . Den Unterschied zwischen $d\omega$ und $d\omega$ bringen wir nur durch die Wahl der Schriftart zum Ausdruck. Eine Differentialform, deren äußere Ableitung verschwindet, heißt **geschlossen**.

10.6.4. Um uns die äußere Ableitung $d\omega$ zu veranschaulichen, erinnern wir zunächst an den Fall einer Nullform alias Funktion, die wir dann statt ω lieber f nennen. Deren äußere Ableitung $(df)_x$ ist schlicht das Differential $d_x f$ bei x und kann dadurch beschrieben werden, daß es jedem Richtungsvektor $\vec{v} \in \vec{X}$ die Zahl

$$(df)_x(\vec{v}) = \lim_{t \rightarrow 0} \frac{1}{t} (f(x + t\vec{v}) - f(x))$$

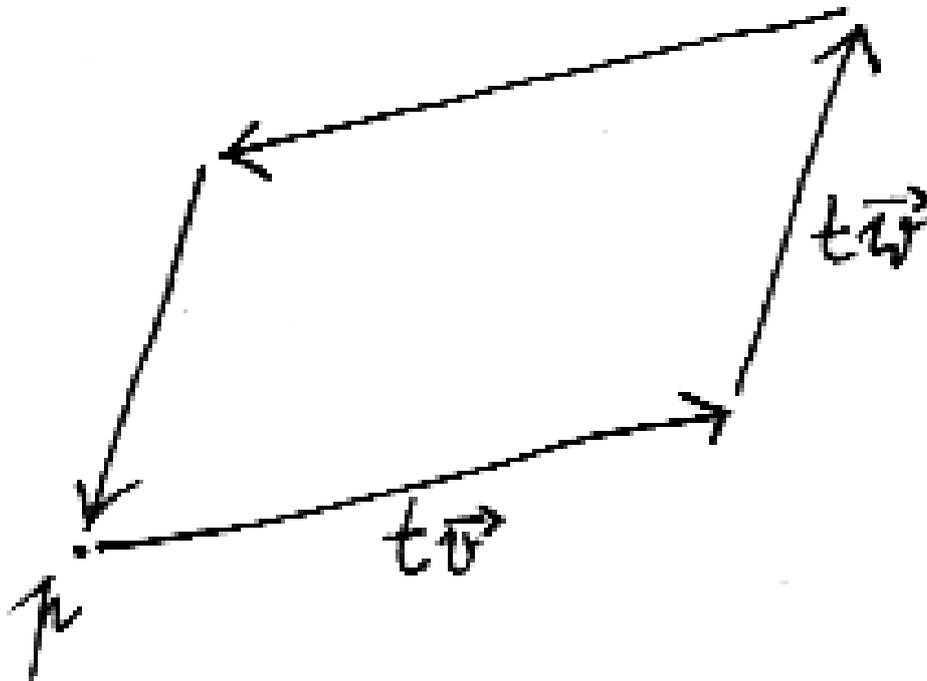
zuordnet. Im Fall einer Einsform alias eines Kovektorfelds ω kann seine äußere Ableitung $(d\omega)_x$ bei x analog dadurch beschrieben werden, daß sie jedem geordneten Paar von Richtungsvektoren $(\vec{v}, \vec{w}) \in \vec{X}^2$ die Zahl

$$(d\omega)_x(\vec{v}, \vec{w}) = \lim_{t \rightarrow 0} \frac{1}{t^2} \int_{\gamma(x, t\vec{v}, t\vec{w})} \omega$$

zuordnet mit der Notation $\gamma(x, t\vec{v}, t\vec{w})$ für den Weg, der einmal das Parallelogramm mit einer Ecke x und Kantenvektoren $t\vec{v}$ und $t\vec{w}$ umläuft, oder genauer, der stückweise linear läuft erst von x nach $x + t\vec{v}$, dann weiter nach $x + t\vec{v} + t\vec{w}$, von da nach $x + t\vec{w}$, und dann wieder zurück nach x . Möglicherweise haben Sie das bereits als Übung 6.6.8 gezeigt. Im allgemeinen Fall einer k -Form ω schließlich haben wir

$$(d\omega)_x(\vec{v}_0, \dots, \vec{v}_k) = \lim_{t \rightarrow 0} \frac{1}{t^{k+1}} \int_{F(x, t\vec{v}_0, \dots, t\vec{v}_k)} \omega$$

wobei wir uns F , zumindest für $\vec{v}_0, \dots, \vec{v}_k$ linear unabhängig, als die in geeigneter Weise orientierte Oberfläche eines Parallelepipedes mit Ecke x und Kantenvektoren



Der Weg $\gamma(p, t\vec{v}, t\vec{w})$ aus Übung 10.6.4. Mit $t \rightarrow 0$ wird er natürlich immer kleiner.

$t\vec{v}_i$ denken dürfen, über die wir dann unsere k -Form integrieren können, wenn wir etwas Mut beweisen und beim Integrieren von den Kanten einmal absehen. Es mag eine gute Übung sein, etwa für die zweite Aussage auch einmal einen Beweis auszuschreiben.

Beispiel 10.6.5. Offensichtlich ist die Zuordnung $\omega \mapsto d\omega$ linear und für Nullformen alias Funktionen f gilt $df = d f$. Ist X ein endlichdimensionaler reeller Raum, $A \subset X$ halboffen, $\omega_\circ \in \text{Alt}^k \vec{X}$ eine konstante k -Form und $f : A \rightarrow \mathbb{R}$ differenzierbar, so behaupten wir die Formel

$$d(f\omega_\circ) = df \wedge \omega_\circ$$

Leser mit entsprechenden Kenntnissen in multilinearer Algebra erkennen das auf einen Blick, die anderen müssen dumpf rechnen, können dann aber die Formel zumindest verifizieren. Im Fall $X = \mathbb{R}^n$ wird für eine Differentialform der Gestalt $\omega = \sum a_I dx_I$ insbesondere ihre äußere Ableitung $d\omega$ gegeben durch die Vorschrift

$$d\omega = \sum da_I \wedge dx_I$$

Das folgt auch formal aus den allgemeineren Aussagen, die in den drei anschließenden Lemmata formuliert werden.

Lemma 10.6.6. *Sind ω und η differenzierbare Differentialformen auf einer halboffenen Teilmenge eines endlichdimensionalen reellen Raums, so gilt für ihr Produkt die **Leibniz-Regel***

$$d(\omega \wedge \eta) = (d\omega) \wedge \eta + (-1)^{|\omega|} \omega \wedge d\eta$$

Lemma 10.6.7. *Sei ω eine Differentialform auf einer halboffenen Teilmenge eines endlichdimensionalen reellen Raums. Ist ω stetig differenzierbar und $d\omega$ auch stetig differenzierbar, so gilt*

$$d(d\omega) = 0$$

Lemma 10.6.8. *C^2 -verwandte Differentialformen haben verwandte äußere Ableitungen. Ist genauer und in Formeln $\phi : A \rightarrow B$ eine C^2 -Abbildung zwischen halboffenen Teilmengen endlichdimensionaler reeller Räume und ω eine differenzierbare Differentialform auf B , so gilt*

$$d(\phi^*\omega) = \phi^*(d\omega)$$

10.6.9. Wie Sie noch sehen werden, erlauben diese Formeln ein außerordentlich elegantes Rechnen mit Differentialformen. Der hier gegebene Formalismus geht auf Élie Cartan's Arbeiten zu Beginn des zwanzigsten Jahrhunderts zurück.

Die Verträglichkeit des äußeren Differentials mit Verwandtschaft macht die Umrechnung zwischen verschiedenen Koordinatensystemen so einfach, daß es auch bei anderen Umrechnungen oft der bequemste Weg ist, sie auf diesen Formalismus zurückzuführen. Als Beispiel bespreche ich die Umrechnung des Laplace-Operators in krummlinige Koordinaten in 10.9.15 folgende.

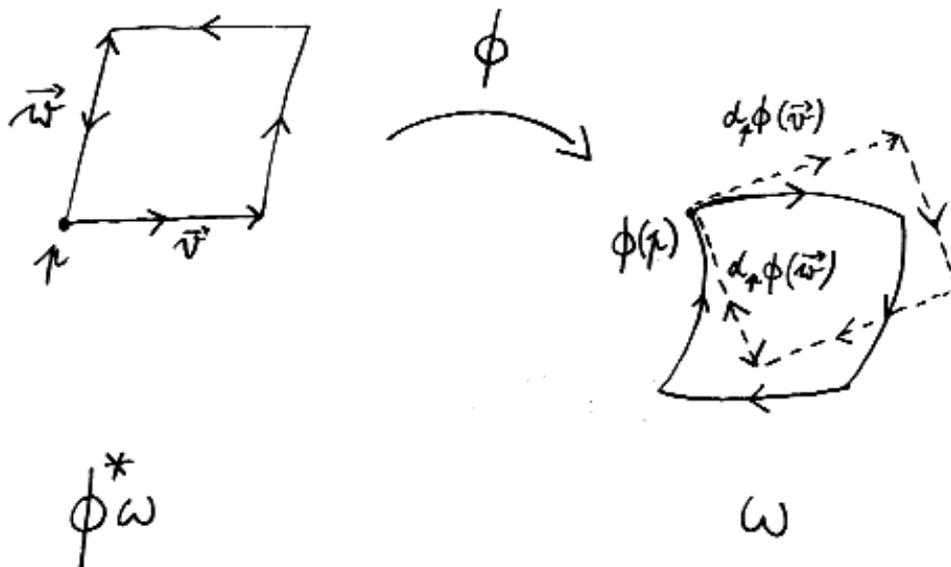
10.6.10. Man kann diese drei Lemmata durch explizite Rechnung in Koordinaten der Reihe nach beweisen. Mir schien jedoch ein anderes Vorgehen übersichtlicher, bei dem im Anschluß an einen Beweis des ersten Lemmas die beiden anderen in einer Art Kaminkletterei abwechselnd in wachsender Allgemeinheit gezeigt werden. Die letzte Regel 10.6.8 können wir auch $\phi : \eta \rightsquigarrow \omega \Rightarrow \phi : d\eta \rightsquigarrow d\omega$ schreiben. Der Leser sei ermutigt, sich das im Lichte unserer Anschauung 10.6.4 bildlich klarzumachen. Die Regel $dd\omega = 0$ ist zumindest für Nullformen im Lichte unserer Anschauung 10.6.4 leicht einzusehen, da das Integral des Differentials einer Funktion über jeden geschlossenen Integrationsweg verschwindet. Für Kovektorfelder sollte die Identität $dd\omega = 0$ zumindest aus dem Stokes'schen Satz mit Ecken 10.8.22 heraus klar werden: Er besagt, daß das Integral von $d\omega$ über eine Fläche unseres Parallelepipedes auch als Integral von ω über dessen Rand geschrieben werden kann, und die Summe aller Randintegrale über die sechs Flächen unseres Parallelepipedes ist offensichtlich wieder Null.

Beweis von 10.6.6. Wir können ω und η schreiben als Summen von Formen der Gestalt $a\omega_o, b\eta_o$ mit ω_o, η_o konstant und a, b differenzierbaren Funktionen. Es reicht also, die Behauptung für $\omega = a\omega_o, \eta = b\eta_o$ zu prüfen. Im Fall von Funktionen liefert die Produktregel, wie bereits in 6.1.10 erwähnt, unmittelbar $d(ab) = (da)b + a(db)$. Dann gilt nach 10.6.5 aber $d\omega = da \wedge \omega_o, d\eta = db \wedge \eta_o$ und damit $d(\omega \wedge \eta) = ((da)b + a(db))\omega_o \wedge \eta_o$. Da zusätzlich gilt $db \wedge \omega_o = (-1)^{|\omega|}\omega_o \wedge db$, folgt die Leibniz-Regel. \square

Beweis von 10.6.7 im Fall $X = \mathbb{R}^n$. Für eine Nullform alias eine Funktion $\omega = a$ auf einer offenen Teilmenge eines \mathbb{R}^n können wir ganz explizit rechnen

$$\begin{aligned} da &= \sum_{i=1}^n \frac{\partial a}{\partial x_i} dx_i \\ dda &= \sum_{i,j=1}^n \frac{\partial^2 a}{\partial x_j \partial x_i} dx_j \wedge dx_i \\ &= \sum_{i < j} \left(\frac{\partial^2 a}{\partial x_j \partial x_i} - \frac{\partial^2 a}{\partial x_i \partial x_j} \right) dx_j \wedge dx_i \\ &= 0 \end{aligned}$$

wo wir die Vertauschbarkeit der partiellen Ableitungen 5.1.10 verwendet haben, die hinwiederum aus der Annahme der Stetigkeit der zweiten Ableitungen folgt. Für eine k -Form ω auf einer offenen Teilmenge des \mathbb{R}^n , sagen wir $\omega = \sum_{|I|=k} a_I dx_I$, erhalten wir damit sofort $d(d\omega) = \sum d(da_I) \wedge dx_I = 0$. Für eine k -Form ω auf



Versuch einer anschaulichen Interpretation der Verträglichkeit zwischen der äußeren Ableitung und dem Zurückholen von Kovektorfeldern. Gegeben ist ein Kovektorfeld ω rechts und ein Punkt p mit zwei Richtungsvektoren \vec{v}, \vec{w} links.

Das Wegintegral von $\phi^*\omega$ über den kleinen Parallelogrammweg links approximiert $(d(\phi^*\omega))_p(\vec{v}, \vec{w})$. Es stimmt nach 6.3.9 überein mit dem Wegintegral des Kovektorfelds ω über seinen Bildweg rechts, eingezeichnet als durchgezogener Rundweg aus vier krummen Stücken. Dahingegen approximiert das Wegintegral über den kleinen gestrichelten Parallelogrammweg rechts $(d\omega)_{\phi(p)}(d_p\phi(\vec{v}), d_p\phi(\vec{w}))$. Die Anschauung soll uns nun sagen, daß im Grenzwert $t \rightarrow 0$ wie in 10.6.4 die entsprechenden beiden Wegintegrale rechts nach Teilen durch t^2 gegen denselben Wert streben. In der Tat werden ja nicht nur die beiden Rundwegsintegrale klein von zweiter Ordnung, sondern die beiden Wege werden sich bei $t \rightarrow 0$ auch sehr ähnlich, und das sorgt dafür, daß die Differenz ihrer Rundwegsintegrale für $t \rightarrow 0$ sogar von dritter Ordnung verschwindet.

einer halboffenen Teilmenge des \mathbb{R}^n folgt unsere Behauptung dann aus der Stetigkeit von $d(d\omega)$. \square

Beweis von 10.6.8 für ϕ affin. Gilt unsere Formel für ω und η , so nach der Produktregel auch für $\omega \wedge \eta$. Es reicht also, unsere Formel für Funktionen alias Nullformen und für konstante 1-Formen zu zeigen. Der Fall von Funktionen ist 6.1.19. Für eine konstante 1-Form ω_\circ und ϕ affin ist natürlich $\phi^*\omega_\circ$ auch eine konstante 1-Form, mithin gilt wie gewünscht $d(\phi^*\omega_\circ) = 0 = \phi^*(d\omega_\circ)$. \square

Beweis von 10.6.7 im Allgemeinen. Ist $\phi : \mathbb{R}^n \xrightarrow{\sim} X$ ein Isomorphismus von affinen Räumen, so folgt $\phi^*(dd\omega) = dd(\phi^*\omega) = 0$ und mithin $dd\omega = 0$. \square

Beweis von 10.6.8 im Allgemeinen. Gilt unsere Formel für ω und η , so nach der Produktregel auch für $\omega \wedge \eta$. Es reicht also, unsere Formel für Funktionen alias Nullformen und für konstante 1-Formen zu zeigen. Der Fall von Funktionen ist 6.1.19. Für eine konstante 1-Form ω_\circ und ϕ beliebig haben wir hinwiederum $\omega_\circ = da$ für eine geeignete Funktion a , genauer für jede affine Abbildung a von unserem affinen Raum nach \mathbb{R} mit linearem Anteil $\vec{a} = \omega_\circ$, und damit ergibt sich $d\phi^*\omega_\circ = d\phi^*da = dd\phi^*a = 0 = \phi^*0 = \phi^*d\omega_\circ$, wo wir im mittleren Schritt verwenden, daß uns die Regel $\phi^*da = d\phi^*a$ für Funktionen a ja bereits aus 6.1.19 zur Verfügung steht. \square

Übung 10.6.11. Prüfen Sie für die Differentialform $x^2 dx \wedge dy - 4e^y dx \wedge dz$, daß erst die äußere Ableitung bilden und dann auf Kugelkoordinaten übergehen dasselbe Resultat liefert wie erst auf Kugelkoordinaten übergehen und dann die äußere Ableitung bilden.

Übung 10.6.12. Zeigen Sie, daß für eine stetig differenzierbare k -Form ω auf dem \mathbb{R}^3 mit $k \geq 1$ die Bedingung $d\omega = 0$ gleichbedeutend ist zur Bedingung, daß es eine stetig differenzierbare $(k-1)$ -Form η auf dem \mathbb{R}^3 gibt mit $\omega = d\eta$.

Ergänzende Übung 10.6.13. Bezeichnen wir die Koordinaten des \mathbb{R}^4 mit x, y, z, t und betrachten auf dem \mathbb{R}^4 eine allgemeine glatte 2-Form

$$F = E^1 dx \wedge dt + E^2 dy \wedge dt + E^3 dz \wedge dt + B^1 dy \wedge dz + B^2 dz \wedge dx + B^3 dx \wedge dy$$

So ist die Gleichung $dF = 0$ äquivalent zu den beiden Gleichungen

$$\operatorname{div} B = 0 \quad \text{und} \quad \operatorname{rot} E = -\frac{\partial B}{\partial t}$$

für rot der Rotation wie in 6.6.12 und $\operatorname{div} B$ der ‘‘Divergenz’’ alias der Summe der partiellen Ableitungen nach x, y und z wie in 10.8.11. Leser mit physikalischer Vorbildung erkennen für $H = cB$ die beiden ersten Maxwell’schen Gleichungen

im Vakuum. Der Formalismus der Verwandtschaft von Differentialformen sagt uns dann, in welcher Weise solch ein elektromagnetisches Feld F in anderen Koordinaten geschrieben werden muß, und daß zumindest die beiden ersten Maxwell'schen Gleichungen nicht von der Wahl der Koordinaten abhängen. Wie man sogar alle vier Maxwell'schen Gleichungen im Vakuum ähnlich elegant formulieren kann, wird in ?? erklärt.

10.7 Berandete Untermannigfaltigkeiten

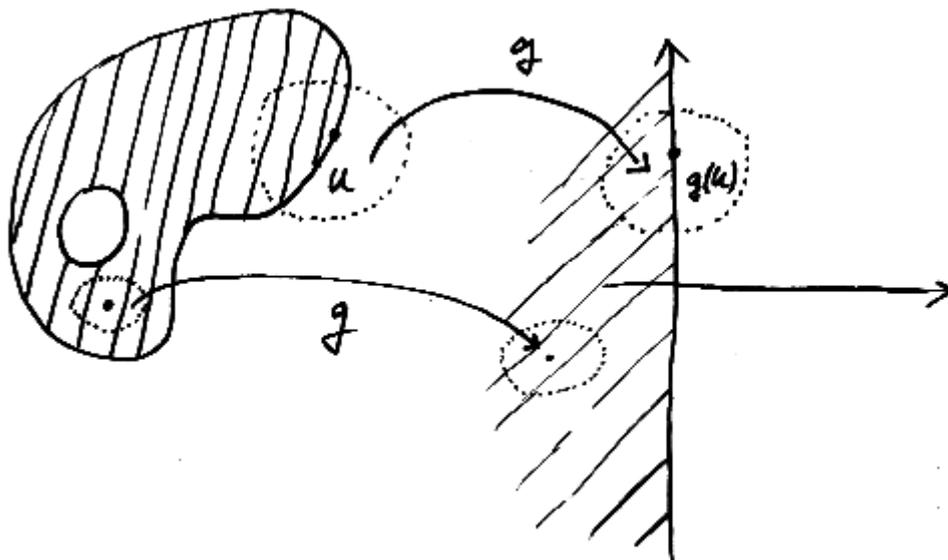
10.7.1. Wir werden im folgenden Gebilde betrachten wollen wie etwa die Halbkugelschale mit Äquator $\{(x, y, z) \in \mathbb{R}^3 \mid x^2 + y^2 + z^2 = 1, z \geq 0\}$. Das ist im Sinne unserer Definition 7.3.2 keine Untermannigfaltigkeit des \mathbb{R}^3 , da wir um Punkte auf dem Äquator keine Plättung im dort erklärten Sinne finden können. Derartige Gebilde sind aber gerade besonders wichtig für die Formulierung höherdimensionaler Verallgemeinerungen des Hauptsatzes der Differential- und Integralrechnung, weshalb wir sie nun auch in unseren allgemeinen begrifflichen Rahmen aufnehmen und dazu den Begriff einer "berandeten Untermannigfaltigkeit" einführen. Im selben Schritt diskutieren wir auch die in diesem Zusammenhang benötigten stärkeren Differenzierbarkeitsvoraussetzungen an Mannigfaltigkeiten.

Definition 10.7.2. Sei X ein reeller Raum endlicher Dimension und seien $k \geq 1$ und $1 \leq l \leq \infty$ gegeben. Eine Teilmenge $M \subset X$ heißt eine **k -dimensionale berandete \mathcal{C}^l -Untermannigfaltigkeit** oder kurz **berandete Untermannigfaltigkeit** von X genau dann, wenn es für jeden Punkt $p \in M$ ein Paar (U, g) gibt aus einer offenen Umgebung $U \Subset X$ von p und einem \mathcal{C}^l -Diffeomorphismus $g : U \xrightarrow{\sim} g(U)$ von U auf eine offene Teilmenge $g(U) \Subset \mathbb{R}^n$ derart, daß gilt:

$$g(U \cap M) = g(U) \cap (\mathbb{R}_{\leq 0} \times \mathbb{R}^{k-1} \times 0)$$

Ein derartiges Paar (U, g) nenne ich eine **Plättung als berandete \mathcal{C}^l -Untermannigfaltigkeit** oder kurz **Randplättung** von M um p . Im Fall $k = 0$ vereinbaren wir, daß eine berandete Untermannigfaltigkeit der Dimension Null dasselbe sein soll wie eine diskrete Teilmenge. Statt \mathcal{C}^∞ sagt man auch oft **glatt**. Sprechen wir ohne nähere Spezifizierung von berandeten Untermannigfaltigkeiten, so sind im Zweifelsfall stets glatte berandete Untermannigfaltigkeiten gemeint.

10.7.3. Eine Randplättung darf natürlich auch in $\mathbb{R}_{<0} \times \mathbb{R}^{n-1}$ landen und ist dann sogar eine Plättung als Untermannigfaltigkeit im Sinne unserer Definition 7.3.2. Jede \mathcal{C}^1 -Untermannigfaltigkeit im Sinne unserer Definition 7.3.2 ist insbesondere auch eine berandete \mathcal{C}^1 -Untermannigfaltigkeit im Sinne von 10.7.2: Haben wir in der Tat im Fall $k \geq 1$ eine Plättung (U, g) als \mathcal{C}^1 -Untermannigfaltigkeit um einen Punkt p , so finden wir durch Verschieben auch eine Plättung als Untermannigfaltigkeit (U, h) , unter der p auf einen Punkt mit negativer erster Koordinate



Eine zweidimensionale berandete Untermannigfaltigkeit der Papierebene mit Plättungen um zwei ausgewählte Punkte.

abgebildet wird, und verkleinern wir dann U zu $V = h^{-1}(\mathbb{R}_{<0} \times \mathbb{R}^{n-1})$, so ist (V, h) eine Randplättung um p . Ich finde die Terminologie insofern unbefriedigend, als der Zusatz “berandet” den Begriff einer Untermannigfaltigkeit erweitert, anstatt ihn einzuschränken. Wir werden in 10.7.7 den “Rand” von berandeten Untermannigfaltigkeiten definieren und aus 10.7.10 wird folgen, daß berandete \mathcal{C}^1 -Untermannigfaltigkeiten mit leerem Rand in der Tat dasselbe sind wie unsere \mathcal{C}^1 -Untermannigfaltigkeiten im Sinne von 7.3.2. Wenn wir besonders betonen wollen, daß wir Untermannigfaltigkeiten im Sinne von 7.3.2 meinen, sprechen wir folgerichtig von **Untermannigfaltigkeiten ohne Rand**.

10.7.4. Statt mit $(\mathbb{R}_{\leq 0} \times \mathbb{R}^{k-1} \times 0)$ arbeiten viele Autoren mit $(\mathbb{R}^{k-1} \times \mathbb{R}_{\geq 0} \times 0)$. Das liefert natürlich eine äquivalente Definition, die sich jedoch im weiteren Verlauf deshalb als weniger geeignet erweist, da sie zu mehr Vorzeichen führt.

10.7.5. Eine berandete \mathcal{C}^l -Untermannigfaltigkeit der Kodimension Null in einem endlichdimensionalen reellen Raum heißt auch eine **\mathcal{C}^l -berandete Teilmenge**. Unser Bild von eben stellt zum Beispiel eine glatt berandete Teilmenge der Papierebene dar.

Beispiele 10.7.6. Alle echten Intervalle in \mathbb{R} sind glatt berandete Teilmengen. Die abgeschlossene Vollkugel $\{(x, y, z) \mid x^2 + y^2 + z^2 \leq 1\}$ ist eine glatt berandete Teilmenge des \mathbb{R}^3 .

Definition 10.7.7. Sei X ein endlichdimensionaler reeller Raum. Ein Punkt einer berandeten Untermannigfaltigkeit $M \subset X$ positiver Dimension $k \geq 1$, der unter *mindestens einer* Randplättung nach $0 \times \mathbb{R}^{k-1} \times 0$ abgebildet wird, heißt ein **Randpunkt** unserer berandeten Untermannigfaltigkeit. Die Menge aller Randpunkte von M notieren wir ∂M und nennen sie den **Rand von M** . Für Untermannigfaltigkeiten M der Dimension $k = 0$ vereinbaren wir $\partial M = \emptyset$.

10.7.8. Das Symbol ∂ ist ein griechisches d . Die Notation ∂M für den Rand ist wohl darauf zurückzuführen, daß sich das Bilden des Randes nach dem Satz von Stokes 10.8.1 als eine in gewisser Weise “duale Operation” zum Differenzieren auffassen läßt und in jedem Falle dazu eng verwandt ist. Dieser Begriff von Rand fällt etwa im Fall von glatt berandeten Teilmengen mit dem in der Topologie verwendeten Begriff von Rand ?? zusammen, ist im allgemeinen aber davon verschieden.

10.7.9. Das folgende Lemma 10.7.10 zeigt, daß ein Randpunkt unter *jeder* Randplättung nach $0 \times \mathbb{R}^{k-1} \times 0$ abgebildet wird, daß es also um einen Randpunkt keine Plättung im Sinne von 7.3.2 geben kann. Insbesondere sind unsere \mathcal{C}^1 -Untermannigfaltigkeiten im Sinne von 7.3.2 genau unsere berandeten \mathcal{C}^1 -Untermannigfaltigkeiten im Sinne von 10.7.2 mit leerem Rand im Sinne der vorhergehenden Definition 10.7.7.

Lemma 10.7.10 (Randmannigfaltigkeit). *Ist M eine berandete Untermannigfaltigkeit positiver Dimension $k \geq 1$ in einem endlichdimensionalen reellen Raum und (U, g) eine Randplättung von M , so gilt stets $g(U \cap \partial M) = g(U) \cap (0 \times \mathbb{R}^{k-1} \times 0)$. Insbesondere ist der Rand ∂M einer berandeten Untermannigfaltigkeit der Dimension $k \geq 1$ stets eine Untermannigfaltigkeit ohne Rand der Dimension $(k - 1)$.*

Beweis. Gegeben ein Randpunkt $p \in U \cap \partial M$ gibt es per definitionem eine Randplättung (V, h) von M um p mit $h(p) \in 0 \times \mathbb{R}^{k-1} \times 0$. Ohne Beschränkung der Allgemeinheit dürfen wir $U = V$ annehmen. Dann ist also $h \circ g^{-1}$ ein Diffeomorphismus zwischen offenen Teilmengen $g(U), h(U)$ des \mathbb{R}^n , der die jeweiligen Schnitte G, H besagter Teilmengen mit $(\mathbb{R}_{\leq 0} \times \mathbb{R}^{k-1} \times 0)$ miteinander identifiziert. Mithin definiert $\varphi = h^{-1} \circ g$ eine stetig differenzierbare Abbildung mit stetig differenzierbarer Umkehrung zwischen diesen Schnitten $\varphi : H \xrightarrow{\sim} G$, die ja halboffen sind in $\mathbb{R}^k \times 0$. Die Komplemente H°, G° von $(0 \times \mathbb{R}^{k-1} \times 0)$ in H, G sind dann die größten in $\mathbb{R}^k \times 0$ offenen Teilmengen von H, G und aus dem Satz über die Umkehrabbildung 7.1.2 folgt $\varphi(H^\circ) = G^\circ$. Also muß unsere Abbildung auch die Schnitte beider Mengen mit $(0 \times \mathbb{R}^{k-1} \times 0)$ identifizieren. \square

Beispiele 10.7.11. Der einzige Randpunkt der glatt berandeten Teilmenge $[a, b) \subset \mathbb{R}$ ist a . Die abgeschlossene Einheitskreisscheibe in der Ebene ist auch als Teilmenge des Raums aufgefaßt eine zweidimensionale berandete Untermannigfaltigkeit mit dem Einheitskreis als Rand.

10.7.12. Eine Teilmenge eines endlichdimensionalen reellen Raums ist in der Zusammenschau der vorhergehenden Definitionen also eine **C^l -Untermannigfaltigkeit ohne Rand** oder kurz **C^l -Untermannigfaltigkeit** für $1 \leq l \leq \infty$ genau dann, wenn es um jeden Punkt unserer Teilmenge eine Plättung im Sinne von 7.3.2 gibt, die sogar ein C^l -Diffeomorphismus ist. Eine **C^∞ -Untermannigfaltigkeit** nennen wir auch eine **glatte Untermannigfaltigkeit**. Die Beschreibungen 7.3.10 und 7.3.13 von Untermannigfaltigkeiten als Urbilder bzw. als Bilder gelten analog, wenn man an den entsprechenden Stellen die Bedingung “stetig differenzierbar” zur Bedingung “von der Klasse C^l ” verstärkt. Sprechen wir ohne nähere Spezifizierung von Untermannigfaltigkeiten, so sind im Zweifelsfall stets glatte Untermannigfaltigkeiten ohne Rand gemeint.

Übung 10.7.13. Seien X und Y endlichdimensionale reelle Räume, $U \subseteq X$ eine offene Teilmenge und $f : U \rightarrow Y$ eine stetig differenzierbare Abbildung mit überall surjektivem Differential. So ist für jede berandete Untermannigfaltigkeit $C \subset Y$ ihr Urbild $M = f^{-1}(C)$ eine berandete Untermannigfaltigkeit von X der Dimension $\dim X - \dim Y + \dim C$ mit Rand $\partial M = f^{-1}(\partial C)$. Man erkennt so zum Beispiel, daß alle Vollkugeln berandete Untermannigfaltigkeiten sind. Hinweis: 7.3.10 und 7.3.12.

Proposition 10.7.14 (Berandete Untermannigfaltigkeiten als Bilder). Sei X ein endlichdimensionaler reeller Raum der Dimension $\dim_{\mathbb{R}} X = n$ und sei $k \geq 1$. Eine Teilmenge $M \subset X$ ist eine k -dimensionale berandete \mathcal{C}^l -Untermannigfaltigkeit genau dann, wenn es für jeden Punkt $p \in M$ eine \mathcal{C}^l -Abbildung $\varphi : W \rightarrow X$ von einer offenen Teilmenge $W \subseteq (\mathbb{R}_{\leq 0} \times \mathbb{R}^{k-1})$ nach X gibt derart, daß gilt:

1. $\varphi(W)$ ist offen in M und enthält p ;
2. $d_x \varphi$ ist injektiv für alle $x \in W$;
3. φ ist injektiv und $\varphi^{-1} : \varphi(W) \rightarrow W$ ist stetig.

10.7.15. Ein Paar (W, φ) wie in der Proposition nenne ich eine **Randkarte** der berandeten Untermannigfaltigkeit M , obwohl $\varphi(W)$ den Rand von M keineswegs treffen muß.

Beweis. Mutatis mutandis derselbe Beweis wie im randlosen Fall, den wir in 7.3.13 behandelt haben. Übung 4.5.9 mag auch helfen. \square

Definition 10.7.16. Sind $(W_\alpha, \varphi_\alpha)$ und (W_β, φ_β) zwei Randkarten einer berandeten Untermannigfaltigkeit M , so setzen wir $W_{\alpha\beta} = \varphi_\alpha^{-1}(\varphi_\beta(W_\beta))$ und nennen die Abbildung

$$\varphi_{\beta\alpha} = \varphi_\beta^{-1} \circ \varphi_\alpha : W_{\alpha\beta} \rightarrow W_{\beta\alpha}$$

den **Kartenwechsel** zwischen unseren beiden Randkarten.

Lemma 10.7.17. Kartenwechsel zwischen Randkarten sind stets Diffeomorphismen. Ist (W, φ) eine Randkarte einer berandeten Untermannigfaltigkeit M positiver Dimension $k \geq 1$, so gilt stets $\varphi^{-1}(\partial M) = W \cap (0 \times \mathbb{R}^{k-1})$.

Beweis. Mutatis mutandis derselbe wie der Beweis von 7.3.21. \square

10.7.18. Genau wie in 9.9.1 erklärt man auch das Flächenmaß einer berandeten Untermannigfaltigkeit eines \mathbb{R}^n und zeigt, daß es ein Borelmaß ist, für das Untermannigfaltigkeiten kleinerer Dimension Nullmengen sind. Das kann insofern nützlich sein, als man so Endlichkeitsaussagen erhält: Ist zum Beispiel M eine kompakte berandete Mannigfaltigkeit, so hat M und damit dann auch $M \setminus \partial M$ endliches Flächenmaß. Genau wie in 10.3.1 erklärt man weiter für eine meßbare relative k -Form ω auf einer berandeten Untermannigfaltigkeit eines endlichdimensionalen reellen Vektorraums das zugehörige Maß $|\omega|$, zeigt, daß Untermannigfaltigkeiten kleinerer Dimension Nullmengen sind, und daß für stetiges ω das zugehörige Maß $|\omega|$ ein Borelmaß ist. Vollständig analog wie im randlosen Fall erklärt man auch, was eine **Orientierung** einer berandeten Mannigfaltigkeit sein soll und wie k -Formen über berandete orientierte Mannigfaltigkeiten zu integrieren sind. Das ist aber fast überflüssig, da beim Integrieren der Rand eh nichts zum Integral beiträgt und höchstens beim Nachweis der Integrierbarkeit helfen kann.

Beispiel 10.7.19. Sei $\varphi : [a, b] \rightarrow \mathbb{R}^n$ eine C^1 -Injektion mit nirgends verschwindendem Differential, die einen Homöomorphismus auf ihr Bild induziert. Aus 10.7.14 folgt leicht, daß dann das Bild von φ eine eindimensionale berandete C^1 -Mannigfaltigkeit M ist, und diese 1-Mannigfaltigkeit besitzt genau eine Orientierung, für die $\varphi|_{(a,b)}$ eine orientierte Karte ist. Unsere Begriffe sind nun gerade so erklärt, daß in diesem Fall für jede stetige relative 1-Form ω auf M ihr Integral über M übereinstimmt mit ihrem Wegintegral über φ , in Formeln

$$\int_{\vec{M}} \omega = \int_{\varphi} \omega$$

Ergänzende Übung 10.7.20. Jede Orientierung des Komplements des Randes in einer Mannigfaltigkeit läßt sich eindeutig zu einer Orientierung der ganzen Mannigfaltigkeit fortsetzen.

Definition 10.7.21. Gegeben eine Randkarte (W, φ) einer berandeten $(k + 1)$ -Mannigfaltigkeit M definieren wir die **induzierte Karte** $(\bar{W}, \bar{\varphi})$ des Randes ∂M durch die Vorschrift

$$(\bar{W}, \bar{\varphi}) = (i^{-1}(W), \varphi \circ i)$$

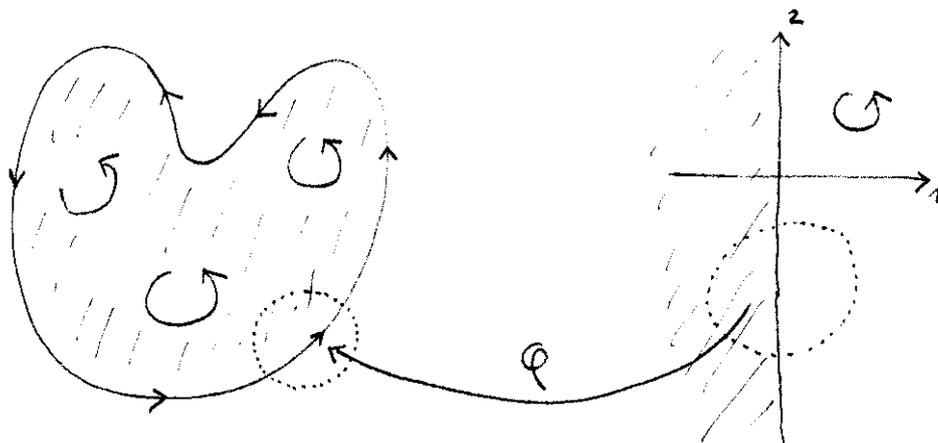
mit $i : \mathbb{R}^k \rightarrow \mathbb{R}_{\leq 0} \times \mathbb{R}^k$ der Einbettung $x \mapsto (0, x)$.

Lemma 10.7.22. Gegeben eine orientierte berandete $(k + 1)$ -Mannigfaltigkeit M gibt es genau eine Orientierung ihres Randes ∂M derart, daß für jede Randkarte der Orientierung ε auch die induzierte Karte des Randes die Orientierung ε hat. Wir nennen sie die **induzierte Orientierung des Randes**.

Beweis. Seien $(W_\alpha, \varphi_\alpha)$ und (W_β, φ_β) zwei Randkarten von M . Der Kartenwechsel $\varphi_{\beta\alpha} : W_{\alpha\beta} \xrightarrow{\sim} W_{\beta\alpha}$ identifiziert $W_{\alpha\beta} \cap (0 \times \mathbb{R}^k)$ mit $W_{\beta\alpha} \cap (0 \times \mathbb{R}^k)$ und kann dort durch den Kartenwechsel $\bar{\varphi}_{\beta\alpha}$ der auf dem Rand induzierten Karten ausgedrückt werden als $\text{id}_0 \times \bar{\varphi}_{\beta\alpha}$. Gegeben $y \in \bar{W}_\alpha \cap \bar{W}_\beta$ hat die Jacobi-Matrix $d_{(0,y)}\varphi_{\beta,\alpha}$ des Kartenwechsels die Gestalt

$$d_{(0,y)}\varphi_{\beta,\alpha} = \left(\begin{array}{c|c} \frac{\partial(\varphi_{\beta\alpha})_1}{\partial x_1}(0, y) & * \\ \hline 0 & d_y\bar{\varphi}_{\beta\alpha} \end{array} \right)$$

und ihr Eintrag oben links alias die partielle Ableitung in $(0, y)$ der ersten Komponente des Kartenwechsels nach der ersten Variablen ist offensichtlich nicht negativ. Mithin hat in jedem Randpunkt die Funktionaldeterminante eines Kartenwechsels zweier Randkarten von M dasselbe Vorzeichen wie die Funktionaldeterminante des Kartenwechsels der auf dem Rand induzierten Karten. \square



Eine orientierte berandete zweidimensionale Mannigfaltigkeit mit der induzierten Orientierung auf ihrem Rand und einer Randkarte

Beispiel 10.7.23. Im Fall der in 10.7.19 besprochenen berandeten orientierten 1-Mannigfaltigkeit M besteht der Rand aus den beiden Punkten $\partial M = \{\varphi(a), \varphi(b)\}$ und die induzierte Orientierung gibt dem Ersten dieser Punkte ein negatives Vorzeichen und dem Zweiten ein positives. Im höherdimensionalen Fall bedeutet unsere Definition anschaulich, daß die orientierten Basen der Tangentialräume des Randes diejenigen angeordneten Basen sind, die orientierte Basen der Tangentialräume der Mannigfaltigkeit liefern, wenn man noch einen Vektor davorschreibt, der tangential an die Mannigfaltigkeit ist und an unserem Randpunkt “aus der Mannigfaltigkeit heraus zeigt”. Ist speziell $M = \partial K$ der Rand einer glatt berandeten Teilmenge K mit der von einer Orientierung des umgebenden Raums induzierten Orientierung, so nennt man das orientierte Normalenfeld auch das **äußere Normalenfeld**, da dann anschaulich gesprochen N_p stets “aus K heraus zeigt”.

10.8 Der Satz von Stokes

Satz 10.8.1 (Stokes’scher Integralsatz). *Sei M eine kompakte orientierte berandete glatte Untermannigfaltigkeit der Dimension $(k + 1)$ in einem endlichdimensionalen reellen Raum und sei ω eine stetig differenzierbare k -Form auf einer offenen Teilmenge unseres Raums, die M umfaßt. Versehen wir den Rand ∂M von M mit der induzierten Orientierung, so gilt*

$$\int_{\vec{M}} d\omega = \int_{\partial\vec{M}} \omega$$

10.8.2. Der Beweis wird zeigen, daß wir statt der Kompaktheit unserer Mannigfaltigkeit schwächer nur voraussetzen brauchen, daß der Träger der Differentialform unsere Mannigfaltigkeit in einem Kompaktum trifft. Weiter reicht es anzunehmen, daß unsere Differentialform auf einer halboffenen Menge definiert ist, die unsere Mannigfaltigkeit umfaßt, und statt der Bedingung C^∞ an unsere Mannigfaltigkeit M reicht auch die Bedingung C^2 .

10.8.3. Konkrete Spezialfälle des vorhergehenden Satzes werden ab 10.8.10 diskutiert. Bereits hier sei bemerkt, daß für eine kompakte Mannigfaltigkeit ohne Rand alias mit leerem Rand das Integral auf der linken Seite verschwinden muß, in Formeln $\partial M = \emptyset \Rightarrow \int_{\vec{M}} d\omega = 0$.

Ergänzung 10.8.4. Später werden wir lernen, was eine “abstrakte Mannigfaltigkeit” und eine “Differentialform auf einer abstrakten Mannigfaltigkeit” ist und wie man solche Differentialformen ableitet und k -Formen über orientierte k -Mannigfaltigkeiten integriert. In dieser Allgemeinheit gilt dann dieselbe Formel für eine beliebige stetig differenzierbare k -Form ω mit kompaktem Träger auf einer beliebigen orientierten $(k + 1)$ -Mannigfaltigkeit M .

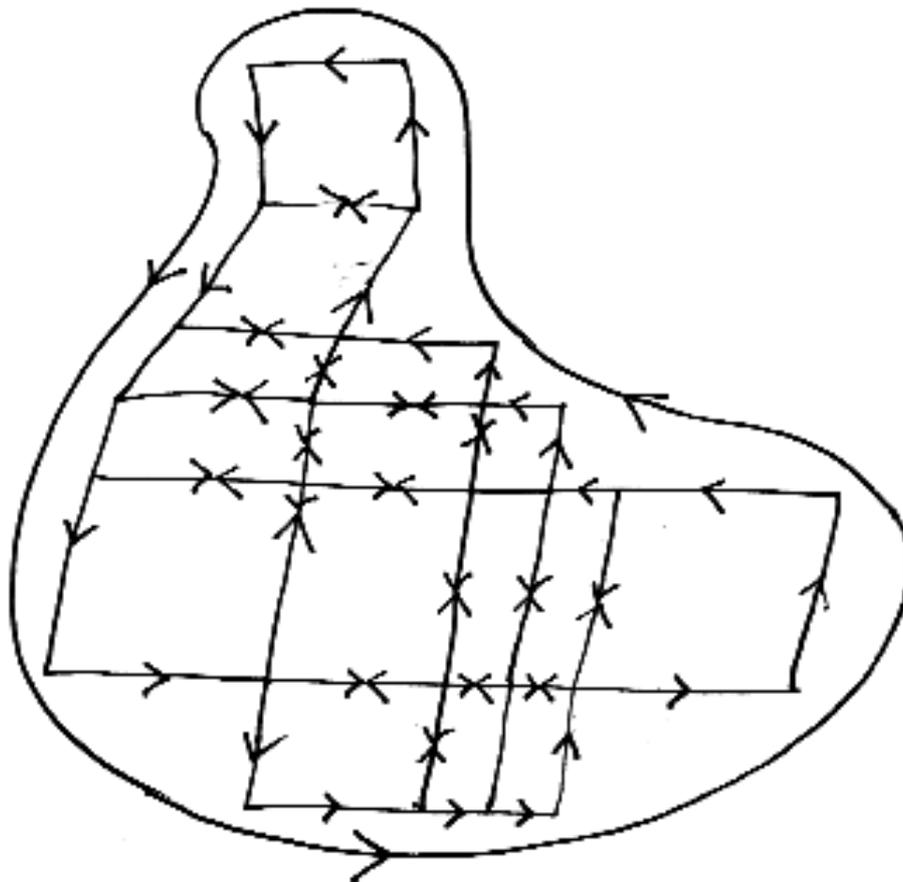


Illustration zum Stokes'schen Satz. Gegeben ein Kovektorfeld ω erinnern wir uns dazu daran, daß nach 10.6.4 seine äußere Ableitung $(d\omega)_p(\vec{v}, \vec{w})$ ausgewertet auf Richtungsvektoren \vec{v}, \vec{w} eine Approximation des Wegintegrals von ω über den Rundweg von p erst nach $p + \vec{v}$, dann weiter nach $p + \vec{v} + \vec{w}$, von dort nach $p + \vec{w}$ und zurück nach p ist. Es sollte nun anschaulich klar sein, daß die Summe über viele derartige kleine Rundwegsintegrale das Randintegral über den ganzen Bereich approximiert. Der Satz von Stokes formalisiert diese Anschauung.

Beispiel 10.8.5. Jedes echte kompakte reelle Intervall $M = [a, b]$ ist eine eindimensionale berandete Untermannigfaltigkeit von \mathbb{R} und erbt von \mathbb{R} eine Orientierung. Sein Rand ist die nulldimensionale Mannigfaltigkeit $\partial M = \{a, b\}$ und die induzierte Orientierung darauf gibt dem Punkt a das Vorzeichen -1 und dem Punkt b das Vorzeichen $+1$. Eine Nullform ω auf M ist schlicht eine Funktion f , ihr Differential ist $d\omega = df = f'(x) dx$, und wir erkennen, daß unser Satz von Stokes 10.8.1 in diesem Fall zum Hauptsatz der Integral- und Differentialrechnung 7.5.3 spezialisiert.

Beispiel 10.8.6. Wir kommen nochmal auf unser Integral über die obere Hemisphäre H der 2-Form $x^2 dx \wedge dy$ aus 10.5.7 zurück, wobei unsere Orientierung der oberen Hemisphäre unter der Projektion auf die Ebene die übliche Orientierung des \mathbb{R}^2 entsprach. Sicher können wir auch $x^2 dx \wedge dy = -d(x^2 y dx)$ schreiben. Der Rand von \vec{H} ist dann der im Gegenuhrzeigersinn orientierte Einheitskreis in der xy -Ebene $S = \{(x, y, z) \mid x^2 + y^2 = 1, z = 0\}$ und aus dem Satz von Stokes folgt

$$\int_{\vec{H}} x^2 dx \wedge dy = \int_{\vec{S}} -x^2 y dx$$

Zur Sicherheit machen wir noch die Probe und landen mit

$$\int_{\vec{S}} -x^2 y dx = \int_0^{2\pi} -\cos^2 \varphi \sin \varphi d(\cos \varphi) = \int_0^{2\pi} \cos^2 \varphi \sin^2 \varphi d\varphi$$

im wesentlichen bei demselben Integral wie dem, das wir bereits in 10.5.7 berechnet hatten. Genauer wird der fehlende Faktor 2 von $\int_0^\pi \sin \varphi d\varphi$ in der Rechnung dort hier dadurch ausgeglichen, daß das Integral bis 2π läuft.

Beweis. Gilt die Aussage für ω und ω' , so auch für $\omega + \omega'$. Wir können also nach 9.10.3 und 7.4.10 ohne Beschränkung der Allgemeinheit annehmen, daß es eine zusammenhängende Randkarte (W, φ) von M gibt, die eine Orientierung hat und derart, daß gilt $(\text{supp } \omega \cap M) \subset \varphi(W)$. Ist ε die Orientierung unserer Randkarte, so gilt per definitionem und da die äußere Ableitung vertauscht mit dem Zurückholen

$$\int_{\vec{M}} d\omega = \varepsilon \int_W \varphi^*(d\omega) = \varepsilon \int_W d(\varphi^*\omega)$$

Bezeichnet $(\vec{W}, \bar{\varphi})$ wie in 10.7.21 die induzierte Karte des Randes und $i : \mathbb{R}^k \rightarrow \mathbb{R}_{\leq 0} \times \mathbb{R}^k, x \mapsto (0, x)$ die offensichtliche Einbettung, so gilt wegen $\bar{\varphi} = \varphi \circ i$ und $\vec{W} = i^{-1}(W)$ auch

$$\int_{\partial \vec{M}} \omega = \varepsilon \int_{\vec{W}} \bar{\varphi}^*\omega = \varepsilon \int_{i^{-1}W} i^*(\varphi^*\omega)$$

Bezeichnen wir mit η die Fortsetzung durch Null von $\varphi^*\omega$ auf den ganzen Halbraum, so reduziert sich der Satz auf einen Spezialfall, den wir im Anschluss als eigenständiges Lemma formulieren und beweisen. \square

Lemma 10.8.7. *Bezeichne $i : \mathbb{R}^k \rightarrow \mathbb{R}_{\leq 0} \times \mathbb{R}^k$ wie zuvor die offensichtliche Einbettung und sei η eine stetig differenzierbare k -Form mit kompaktem Träger auf $\mathbb{R}_{\leq 0} \times \mathbb{R}^k$. So gilt*

$$\int_{\mathbb{R}^k} i^* \eta = \int_{\mathbb{R}_{\leq 0} \times \mathbb{R}^k} d\eta$$

Beweis. Wir nennen unsere Koordinaten hier ausnahmsweise x_0, x_1, \dots, x_k und können schreiben

$$\eta = \sum_{\nu=0}^k \eta_\nu dx_0 \wedge \dots \wedge \widehat{dx_\nu} \wedge \dots \wedge dx_k$$

für stetig differenzierbare Funktionen η_ν mit kompaktem Träger. Es ergibt sich $i^* \eta = \eta_0 dx_1 \wedge \dots \wedge dx_k$, die linke Seite ist also schlicht $\int_{\mathbb{R}^k} \eta_0$. Auf der rechten Seite erhalten wir

$$d\eta = \sum_{\nu=0}^k (-1)^\nu \frac{\partial \eta_\nu}{\partial x_\nu} dx_0 \wedge \dots \wedge dx_k$$

und für alle $\nu \neq 0$ verschwindet beim entsprechenden Summanden das ν -te partielle Integral, da die Stammfunktion η_ν kompakten Träger hat und von $-\infty$ bis ∞ integriert wird. Nur der erste Summand liefert also einen Beitrag, und der ist

$$\int_{\mathbb{R}^k} \left(\int_{-\infty}^0 \frac{\partial \eta_0}{\partial x_0} \right) = \int_{\mathbb{R}^k} \eta_0 \quad \square$$

10.8.8. Lassen wir den Beweis des Stokes'schen Satzes nocheinmal Revue passieren, so fällt auf, daß er ziemlich kurz ist. Das liegt daran, daß seine Formulierung in der Sprache der Differentialformen so gut mit Koordinatenwechseln verträglich ist, daß wir uns sofort auf einen sehr einfachen Spezialfall zurückziehen können. In gewisser Weise haben wir also mit der Entwicklung der Sprache der Differentialformen die Hauptarbeit bereits geleistet. Als wesentliche nichttriviale Aussage möchte ich dabei insbesondere die Verträglichkeit der äußeren Ableitung mit Koordinatenwechseln hervorheben, die sich auch in anderen Zusammenhängen noch als starkes Hilfsmittel erweisen wird.

10.8.9. Ich formuliere nun einige Spezialfälle des allgemeinen Stokes'schen Satzes 10.8.1 in klassischer Notation, um die Lektüre älterer Texte zu erleichtern. Ich hoffe jedoch, daß sich der für explizite Rechnungen und theoretische Überlegungen gleichermaßen bestens geeignete Formalismus der Differentialformen mit der Zeit auch bei den Anwendern durchsetzen wird.

Beispiel 10.8.10 (Wegintegral über ein Gradientenfeld). Sei $\varphi : [a, b] \rightarrow \mathbb{R}^n$ eine stetig differenzierbare Injektion mit nirgends verschwindendem Differential, die einen Homöomorphismus auf ihr Bild induziert. Aus 10.7.14 folgt leicht, daß dann das Bild von φ eine berandete 1-Mannigfaltigkeit M ist, und diese 1-Mannigfaltigkeit besitzt genau eine Orientierung, für die $\varphi|_{(a,b)}$ eine orientierte Karte ist. Gegeben eine Nullform alias Funktion f auf einer offenen Umgebung von M haben wir $df = \langle \text{grad } f, \cdot \rangle = (\text{grad } f) \cdot$ und der Satz von Stokes erhält nach 10.7.19 und 10.7.23 und 10.5.11 die Gestalt

$$\int_a^b \langle \text{grad } f, d\varphi \rangle = \int_a^b (\text{grad } f) \cdot d\varphi = f(\varphi(b)) - f(\varphi(a))$$

In Worten stimmt also das Wegintegral des Gradientenfelds einer Funktion überein mit der Differenz zwischen den Werten besagter Funktion am Anfangs- und Endpunkt des Integrationsweges.

Beispiel 10.8.11 (Satz von Gauss). Gegeben eine kompakte glatt berandete Teilmenge $K \subset \mathbb{R}^n$ und ein im Sinne von 4.5.2 stetig differenzierbares Vektorfeld F auf K bilden wir wie in 10.5.15 die zugehörige $(n - 1)$ -Form $\omega = \omega_F$ und finden

$$d\omega = (\text{div } F) dx_1 \wedge \dots \wedge dx_n$$

für $(\text{div } F) : K \rightarrow \mathbb{R}$ die sogenannte **Quelldichte** oder auch **Divergenz** unseres Vektorfeldes, die auf dem “Inneren” $K \setminus \partial K$ von K gegeben wird durch die Vorschrift

$$\text{div } F = \frac{\partial F_1}{\partial x_1} + \dots + \frac{\partial F_n}{\partial x_n}$$

Damit spezialisiert der allgemeine Satz von Stokes zum **Satz von Gauss**

$$\int_K \text{div } F = \int_{\partial K} F \cdot N$$

für $N : \partial K \rightarrow \mathbb{R}^n$ das äussere Normalenfeld. In Worten ist also der Fluß eines Vektorfelds durch den Rand eines glatt berandeten Kompaktums im \mathbb{R}^n gleich dem Integral seiner Quelldichte über besagtes Kompaktum. Anschaulich mag man sich im Fall $n = 2$ die Oberfläche K eines ebenen Moores denken, in dem Wasser nach oben dringt und über das Moor an den Rand des Moores fließt. Nehmen wir das Geschwindigkeitsfeld dieses Flusses als unser Vektorfeld, so wäre die Divergenz eben die Quelldichte in unserem Moor, das Randintegral mißt die Wassermenge, die pro Zeiteinheit am Rand unseres Moores herausläuft, und unser Satz besagt, daß sie gleich der Wassermenge sein muß, die pro Zeiteinheit im Inneren emporquillt.

Beispiel 10.8.12. Ein homogener, als da heißt überall gleich dichter schwerer Körper K wird an einem Seil ins Wasser gelassen. Wir wollen uns überlegen, daß auch im Wasser der Schwerpunkt unseres Körpers in der Vertikalen unter dem Aufhängepunkt bleibt. Für inhomogene Körper gilt das im allgemeinen natürlich nicht! Wir denken uns unseren Körper als kompakte glatt berandete Teilmenge $K \subset \mathbb{R}^3$ mit Schwerpunkt auf der z -Achse, also $\int_K x = \int_K y = 0$. Die Wasseroberfläche möge die Ebene $z = 0$ sein. Der Wasserdruck steigt linear mit der Tiefe, auf ein Oberflächenelement der Fläche $\sigma\langle p \rangle$ um $p \in \partial K$ wirkt also die Kraft $z(p)N_p\sigma\langle p \rangle$ mit N_p dem orientierten Normalenvektor bei p . Befindet sich der Aufhängepunkt etwa in der Höhe $h < 0$, so wird das Drehmoment um diesen Aufhängepunkt das Oberflächenintegral

$$\int_{\partial K} z(p)(N_p \times (p + h e_3))\sigma\langle p \rangle$$

Die Komponenten dieses Vektors bei $p = (x, y, z)$ mit $N_p = (N_1, N_2, N_3)$ sind $z(N_2(z+h) - N_3y)$, $z(N_3x - N_1(z+h))$ und $z(N_1y - N_2x)$ und können auch dargestellt werden als die Skalarprodukte von N_p mit den Vektorfeldern $v_1(x, y, z) = (0, z^2 + hz, -zy)$, $v_2(x, y, z) = (-z^2 - hz, 0, zx)$ und $v_3(x, y, z) = (zy, -zx, 0)$, so daß es gilt $\int_{\partial K} (N \cdot v_i)\sigma = 0$ zu zeigen. Mit dem Satz von Gauss können wir diese Integrale verwandeln in die Integrale $\int_K \operatorname{div} v_i$ und wegen $\operatorname{div} v_1 = -y$, $\operatorname{div} v_2 = x$ und $\operatorname{div} v_3 = 0$ verschwinden sie in der Tat alle drei.

Beispiel 10.8.13 (Klassischer Satz von Stokes). Sei $M \subset \mathbb{R}^3$ eine kompakte orientierte berandete Fläche und $\varphi : [a, b] \rightarrow \partial M$ eine orientierte Parametrisierung ihres Randes. Sei $F : U \rightarrow \mathbb{R}^3$ ein stetig differenzierbares Vektorfeld auf einer offenen Teilmenge $U \subset \mathbb{R}^3$, die M umfaßt, und bezeichne $\eta = \langle F, \rangle$ die zugehörige 1-Form. So finden wir $d\eta = \omega_R$ in der Notation von 10.5.15 für $R : U \rightarrow \mathbb{R}^3$ dasjenige Vektorfeld $\operatorname{rot} F$ auf U , das definiert wird durch die Vorschrift

$$\operatorname{rot} F = \left(\frac{\partial F_3}{\partial x_2} - \frac{\partial F_2}{\partial x_3}, \frac{\partial F_1}{\partial x_3} - \frac{\partial F_3}{\partial x_1}, \frac{\partial F_2}{\partial x_1} - \frac{\partial F_1}{\partial x_2} \right)$$

Dies Vektorfeld ist die Rotation unseres Vektorfelds F , wie wir sie in 6.6.12 eingeführt haben. Unser allgemeiner Satz von Stokes 10.8.1 spezialisiert in dieser Situation zum **klassischen Satz von Stokes**

$$\int_M N \cdot (\operatorname{rot} F)\sigma = \int_a^b F \cdot d\varphi$$

Hier bedeutet $N : M \rightarrow \mathbb{R}^3$ wieder das durch die Orientierung von M festgelegte Normalenfeld. In Worten ist also das Wegintegral eines Vektorfeldes über den Rand einer Fläche gleich dem Fluß der Rotation des Vektorfeldes durch besagte Fläche.

10.8.14. Bei Anwendern, die hauptsächlich im \mathbb{R}^3 arbeiten, ist eine andere symbolische Schreibweise für grad, rot und div sehr beliebt: Sie betrachten den sogenannten **Nabla-Operator** ∇ , den man sich denkt als den “Vektor von Symbolen” $(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z})$, und schreiben

$\nabla f = \text{grad } f$, zu verstehen als symbolisches Produkt des Nabla-Vektors mit einer skalaren Funktion;

$\nabla \cdot F = \text{div } f$, zu verstehen als symbolisches Skalarprodukt des Nabla-Vektors mit einer vektorwertigen Funktion; das Skalarprodukt wird von diesen Anwendern meist $v \cdot w$ notiert statt wie bei uns $\langle v, w \rangle$;

$\nabla \times F = \text{rot } F$, zu verstehen als symbolisches Vektorprodukt des Nabla-Vektors mit einer vektorwertigen Funktion, wo eben das Vektorprodukt $v \times w = (v_2 w_3 - v_3 w_2, v_3 w_1 - v_1 w_3, v_1 w_2 - v_2 w_1)$ aus der Geometrie des Raums ?? zugrundegelegt wird.

In dieser Notation wird dann unsere Formel $dd\omega = 0$ für ω eine Funktion bzw. eine 1-Form auf dem \mathbb{R}^3 verstanden als formal-symbolische Konsequenz der Formeln $v \times v = 0$ bzw. $v \cdot (v \times w) = 0$ aus der Geometrie des Raums.

Beispiel 10.8.15 (Green’sche Formel). Sei $G \subset \mathbb{R}^2$ eine kompakte glatt berandete Teilmenge und $\varphi : [a, b] \rightarrow \mathbb{R}^2$ eine orientierte Parametrisierung ihres Randes, anschaulich “ein im Gegenuhrzeigersinn auf dem Rand umlaufender geschlossener Weg”. Gegeben ein stetig differenzierbares Vektorfeld $v = (v_1, v_2)$ auf einer offenen Umgebung von G betrachten wir die 1-Form $\langle v, \cdot \rangle = \eta = v_1 dx_1 + v_2 dx_2$ mit ihrem Differential $d\eta = (\text{rot } v) dx_1 \wedge dx_2$ für

$$\text{rot } v = \left(\frac{\partial v_2}{\partial x_1} - \frac{\partial v_1}{\partial x_2} \right)$$

die in 6.6.12 erklärte skalare Rotation eines Vektorfelds in der Ebene, und der allgemeine Satz von Stokes 10.8.1 spezialisiert zur **Green’schen Formel**

$$\int_G \text{rot } v = \int_a^b v \cdot d\varphi$$

10.8.16. Dieselbe Formel hatten wir in 6.6.18 schon für G ein Rechteck kennengelernt, nur ist ein Rechteck natürlich nicht glatt berandet. In 10.8.22 werden wir jedoch einen “Satz von Stokes mit Ecken” kennenlernen, der dann auch diese Formel für Rechtecke als Spezialfall enthält.

Beispiel 10.8.17 (Fläche eines ebenen Gebiets). Sei $G \subset \mathbb{R}^2$ wie in 10.8.15 eine kompakte glatt berandete Teilmenge und $\varphi : [a, b] \rightarrow \mathbb{R}^2$ eine orientierte Parametrisierung ihres Randes. Betrachten wir die 2-Form $\omega = x dy$ mit Differential

$d\omega = dx \wedge dy$ und $\varphi^*\omega = \varphi_1(t)\varphi_2'(t) dt$, so spezialisiert der allgemeine Satz von Stokes 10.8.1 zu einer Formel für die Fläche des Gebietes G , genauer zu der Regel

$$\int_G 1 = \int_a^b \varphi_1(t)\varphi_2'(t) dt$$

10.8.18. Ich selber finde die alternative Interpretation dieser Formel mithilfe des Gauss'schen Integralsatzes besonders anschaulich: Quillt in einem Moor überall gleichviel Wasser hoch, so können wir seine Fläche bestimmen, indem wir messen, wieviel Wasser in einem Graben um unser Moor abläuft. Wie genau das Wasser auf unserem Moor zum Randgraben läuft, ist dabei völlig unerheblich. Statt mit $\omega = x dy$ könnten wir also irgendein beliebiges ω mit $d\omega = dx \wedge dy$ nehmen und so weitere Formeln für die Fläche eines ebenen Gebiets erhalten.

Definition 10.8.19. Sei X ein reeller Raum endlicher Dimension und seien $k \in \mathbb{N}$ und $1 \leq l \leq \infty$. Eine Teilmenge $M \subset X$ heißt eine **k -dimensionale \mathcal{C}^l -Untermannigfaltigkeit mit Ecken** oder kurz **Untermannigfaltigkeit mit Ecken** von X genau dann, wenn es für jeden Punkt $p \in M$ ein Paar (U, g) gibt aus einer offenen Umgebung $U \subseteq X$ von p und einem \mathcal{C}^l -Diffeomorphismus $g : U \xrightarrow{\sim} g(U)$ von U auf eine offene Teilmenge $g(U) \subseteq \mathbb{R}^n$ derart, daß gilt:

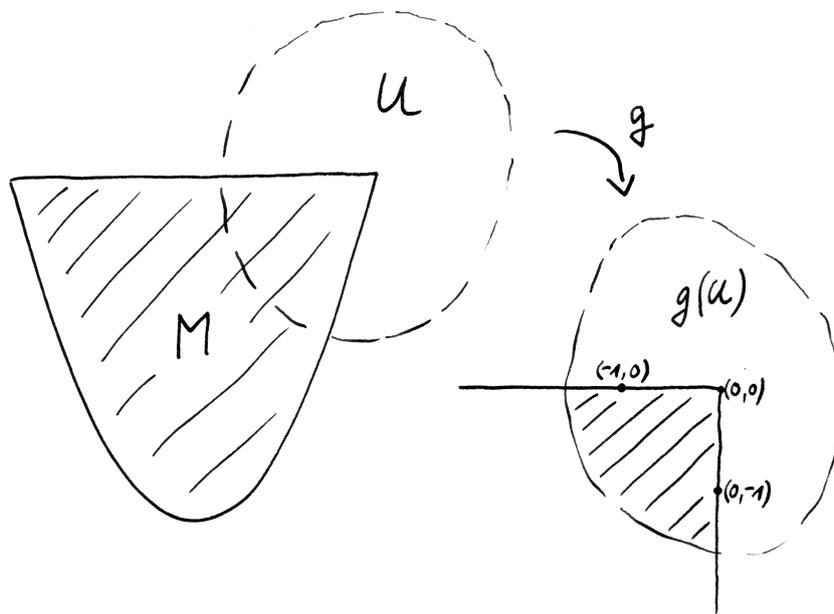
$$g(U \cap M) = g(U) \cap ((\mathbb{R}_{\leq 0})^k \times 0)$$

Ein derartiges Paar (U, g) nenne ich eine **Plättung als \mathcal{C}^l -Untermannigfaltigkeit mit Ecken** oder kurz **Eckenplättung** von M um p .

Beispiele 10.8.20. Eine nulldimensionale Untermannigfaltigkeit mit Ecken ist dasselbe wie eine diskrete Teilmenge. Eine eindimensionale Untermannigfaltigkeit mit Ecken ist dasselbe wie eine eindimensionale Untermannigfaltigkeit mit Rand. Erst in höheren Dimensionen erhalten wir etwas Neues: So wäre zum Beispiel das "Innere eines ebenen Vielecks zusammen mit seinem Rand" eine zweidimensionale Untermannigfaltigkeit mit Ecken der Ebene, aber keine zweidimensionale berandete Untermannigfaltigkeit der Ebene.

10.8.21. Alle Punkte einer Untermannigfaltigkeit M mit Ecken, um die es eine Randplättung gibt, bilden eine berandete Untermannigfaltigkeit M_r unseres endlichdimensionalen reellen Raums, die wir den **regulären Teil von M** nennen.

Satz 10.8.22 (Stokes'scher Integralsatz mit Ecken). Sei M eine kompakte glatte Untermannigfaltigkeit mit Ecken der Dimension $(k + 1)$ in einem endlichdimensionalen reellen Raum. Sei der reguläre Teil M_r von M orientiert und sei ω eine stetig differenzierbare k -Form auf einer halboffenen Teilmenge unseres Raums,



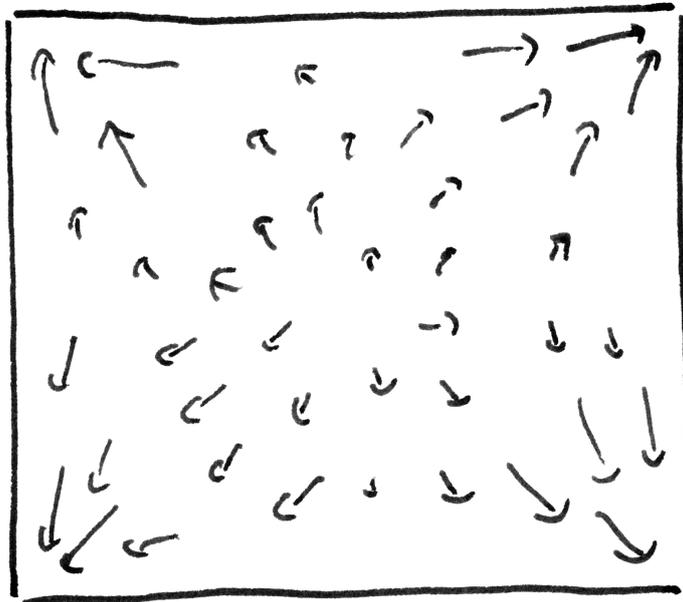
Eine kompakte 2-dimensionale Untermannigfaltigkeit M mit Ecken der Papierebene mitsamt einer Eckenplättung in die dafür in geeigneter Weise mit \mathbb{R}^2 zu identifizierende Papierebene.

die M umfaßt. So existieren die Integrale von ω über $\partial\vec{M}_r$ und von $d\omega$ über \vec{M}_r und es gilt

$$\int_{\vec{M}_r} d\omega = \int_{\partial\vec{M}_r} \omega$$

10.8.23. Im allgemeinen gilt der Satz von Stokes keineswegs für nichtkompakte berandete Mannigfaltigkeiten wie etwa unser M_r . Ist zum Beispiel Q ein Quadrat in der Ebene ohne die Ecken, so können wir auf einer offenen Menge, die unser eckenloses Quadrat umfaßt, ein Vektorfeld konstruieren, das den Fluß eines expandierenden Gases beschreibt, das “durch die Löcher in den Ecken entweicht” aber dessen Fluß durch die Randkanten des Quadrats verschwindet. In dieser Allgemeinheit gälte der Satz von Stokes also nicht. Allerdings müßte unser Gas “mit unendlicher Geschwindigkeit durch die Ecken pfeifen” und sein Geschwindigkeitsfeld könnte nicht stetig auf besagte Ecken fortgesetzt werden, weshalb auch die Voraussetzungen für unseren Satz von Stokes mit Ecken in diesem Fall nicht erfüllt wären. Es gibt noch sehr viel allgemeinere Versionen des Stokes’schen Satzes mit Ecken, vergleiche etwa [?], mit denen sich zum Beispiel auch der Fluß durch die Oberfläche eines Ikosaeders direkt diskutieren ließe. Der hier besprochene Fall scheint mir jedoch für die meisten Anwendungen ausreichend und hat den Vorteil, daß sowohl seine Formulierung als auch sein Beweis nur wenig begrifflichen Aufwand benötigen. Den Fall eines Ikosaeders kann man daraus im übrigen auch noch erhalten, etwa indem man ihn etwa in Dreieckspyramiden mit einer Ecke im Ursprung zerlegt.

Beispiel 10.8.24. Ich will noch einmal auf das schon in 10.8.6 besprochene Beispiel 10.5.7 zurückkommen, in dem wir den Fluß des Vektorfeldes $x^2 e_3$ durch die obere Hemisphäre H alias das Integral der 2-Form $x^2 dx \wedge dy$ über eben diese Hemisphäre berechnet hatten. Die Länge der Vektoren unseres Vektorfeldes $x^2 e_3$ hängt von der Höhe z gar nicht ab. Es scheint mir deshalb offensichtlich, daß sein Fluß durch die obere Hemisphäre H derselbe ist wie durch die Einheitskreisscheibe in der xy -Ebene $D = \{(x, y) \mid z = 0, x^2 + y^2 < 1\}$. Formal können wir das wegen $d(x^2 dx \wedge dy) = 0$ auch aus dem Satz von Stokes mit Ecken 10.8.22 folgern, indem wir ihn auf die massive obere Halbkugel $M = \{(x, y, z) \mid x^2 + y^2 + z^2 \leq 1, z \geq 0\}$ anwenden. Deren regulärer Teil M_r besteht aus dem Komplement der Kreislinie $\{(x, y, z) \mid x^2 + y^2 = 1, z = 0\}$, und der Rand des regulären Teils ∂M_r ist die Vereinigung der oberen Hemisphäre H aus 10.5.7 mit der offenen Einheitskreisscheibe in der xy -Ebene D . Versehen wir unsere massive Halbkugel mit der von \mathbb{R}^3 induzierten Orientierung, so erbt die obere Hemisphäre H die bereits in 10.5.7 beschriebene Orientierung, die Einheitskreisscheibe D jedoch die nicht von der Einbettung in \mathbb{R}^2 induzierte Orientierung. Schreiben wir \vec{D} für D mit der von der Einbettung in \mathbb{R}^2 induzierten Orientierung,



Ein expandierendes Gas, das durch die Ecken entweicht, als Beispiel dafür, daß die Kompaktheitsbedingung beim Satz von Stokes notwendig ist.

so erhalten wir für $\omega = x^2 dx \wedge dy$ wegen $d\omega = 0$ folglich nach Stokes in der Tat

$$0 = \int_{\vec{M}_r} d\omega = \int_{\partial\vec{M}_r} \omega = \int_{\vec{H}} x^2 dx \wedge dy - \int_{\vec{D}} x^2 dx \wedge dy$$

Beweis. Ganz genau wie beim Beweis des Stokes'schen Satzes 10.8.1 ziehen wir uns zurück auf den Fall einer stetig differenzierbaren k -Form η auf $(\mathbb{R}_{\leq q})^{k+1}$ mit kompaktem Träger. Wieder nennen wir unsere Koordinaten x_0, x_1, \dots, x_k , betrachten nun aber für $0 \leq \nu \leq k$ die Einbettungen

$$i_\nu : (\mathbb{R}_{\leq 0})^k \hookrightarrow (\mathbb{R}_{\leq 0})^{k+1}$$

die durch Einfügen einer Null an der ν -ten Stelle entstehen. Der Stokes'sche Integralsatz mit Ecken reduziert sich dann auf die Behauptung

$$\sum_{\nu=0}^k (-1)^\nu \int_{(\mathbb{R}_{\leq 0})^k} i_\nu^* \eta = \int_{(\mathbb{R}_{\leq 0})^{k+1}} d\eta$$

und das kann wie beim Beweis von 10.8.7 explizit nachgerechnet werden. \square

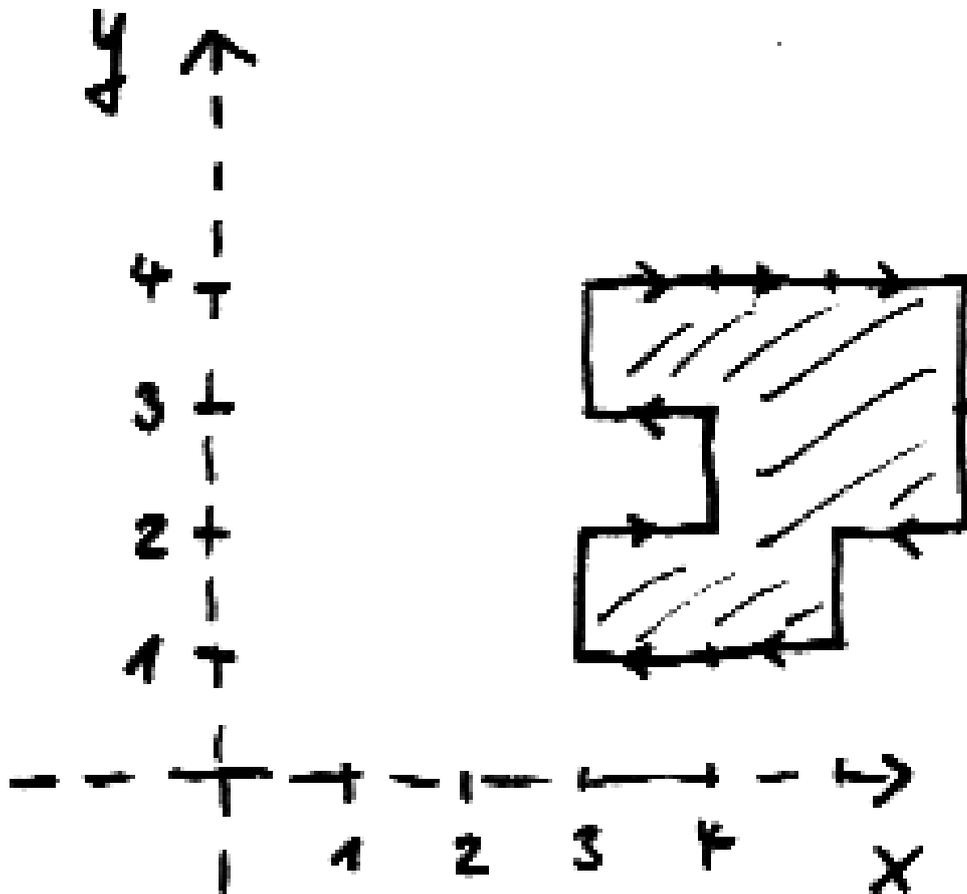
10.8.25. Wir können nun auch einen besonders kurzen Beweis für die Homotopieinvarianz von Wegintegralen in geschlossenen Kovektorfeldern 6.6.10 geben unter der stärkeren Voraussetzung, daß es zwischen unseren beiden stetig differenzierbaren Wegen $\gamma, \psi : [0, 1] \rightarrow A$ sogar eine zweimal stetig differenzierbare Homotopie $h : [0, 1]^2 \rightarrow A$ gibt. Wir nehmen genauer A offen in einem reellen Raum X an und $\omega : A \rightarrow \vec{X}^*$ ein stetig differenzierbares Kovektorfeld. Die Behauptung in 6.6.10 besagt ja gerade, daß aus $d\omega = 0$ folgt $\int_\gamma \omega = \int_\psi \omega$. Aber nun finden wir

$$\int_\gamma \omega - \int_\psi \omega = \int_{\partial([0,1]^2)} h^* \omega = \int_{[0,1]^2} d(h^* \omega) = \int_{[0,1]^2} h^*(d\omega) = 0$$

nach der Definition einer Homotopie, dem Satz von Stokes mit Ecken, der Verträglichkeit des Zurückholens von Formen mit dem äußeren Differential 10.6.8 und unserer Annahme $d\omega = 0$.

10.9 Divergenz und Laplace in krummen Koordinaten*

10.9.1. Die folgenden Argumente bauen nicht auf dem Stokes'schen Integralsatz auf, vielmehr geht es um Anwendungen des Kalküls der Differentialformen aus 10.6.



Die Formel 10.8.17 für die Fläche eines ebenen Gebiets gilt nun natürlich ebenso für "Gebiete mit Ecken". Diese Formel kann etwa angewandt werden, um ein GPS-Gerät so zu programmieren, daß es einem die Fläche des Gebiets berechnet, das man bei einem Rundweg umrundet hat. Im Spezialfall eines Gebiets, das von einem den Kanten eines Rechenpapiers folgenden Weg im Uhrzeigersinn umrundet wird, ergibt sich, wenn wir Stokes auf die Form $y dx$ anwenden, die Fläche als die Höhe des Schwerpunkts der Menge der horizontalen Kanten, wenn wir jede Kante nach rechts mit ihrer Höhe gewichten und jede Kante nach links mit dem Negativen ihrer Höhe. Für die Fläche des obigen Gebiets ergibt sich so

$$3 \times 4 + 2 - 2 \times 1 - 2 - 3 = 7$$

10.9.2. Gegeben ein orientierter n -dimensionaler reeller Vektorraum V mit einem Skalarprodukt oder allgemeiner einer nichtausgearteten symmetrischen Bilinearform t kann man im n -dimensionalen Raum $\text{Alt}^n(V)$ aller sogenannten **Volumenformen auf V** ein von Null verschiedenes Element ω , die **kanonische Volumenform**, auszeichnen durch die Bedingung, daß gilt

$$\omega(v_1, \dots, v_n) = 1$$

für jede orientierte Orthonormalbasis im positiv definiten Fall bzw. jede orientierte Basis v_1, \dots, v_n mit $|t(v_i, v_j)| = \delta_{ij}$ im allgemeinen Fall. In der Tat erfüllt die Basiswechsellmatrix A zwischen zwei derartigen Basen eine Gleichung der Gestalt $A^\top J A = J'$ mit $J = J' = I$ der Einheitsmatrix im Fall eines Skalarprodukts und $\det J = \det J' \neq 0$ im allgemeinen, so daß der Multiplikationssatz für Determinanten $\det A = \pm 1$ liefert, und die Orientiertheit beider Basen zeigt dann sogar $\det A = 1$. Damit aber folgt

$$\omega(v_1, \dots, v_n) = \omega(w_1, \dots, w_n)$$

für jede n -Form ω und je zwei Basen wie oben, etwa indem wir 10.1.9 auf den Automorphismus von V mit $v_i \mapsto w_i$ anwenden.

Definition 10.9.3. Gegeben ein orientierter n -dimensionaler reeller Vektorraum V mit einer nichtausgearteten symmetrischen Bilinearform t erklärt man für jede Zerlegung $n = p + q$ den **Hodge*-Operator**, einen Isomorphismus

$$* = *_t : \text{Alt}^p V \xrightarrow{\sim} \text{Alt}^q V$$

wie folgt: Man geht aus von der durch das Dachprodukt gegebenen nichtausgearteten Paarung

$$\text{Alt}^p V \times \text{Alt}^q V \rightarrow \text{Alt}^n V$$

und verknüpft sie mit dem Isomorphismus $\text{Alt}^n V \xrightarrow{\sim} \mathbb{R}$, der die kanonische Volumenform ω auf die Eins wirft. Die so erhaltene Paarung kann als ein Isomorphismus

$$\text{Alt}^p V \xrightarrow{\sim} (\text{Alt}^q V)^*$$

interpretiert werden, und der kanonische Isomorphismus $(\text{Alt}^q V)^* \xrightarrow{\sim} \text{Alt}^q(V^*)$ aus ?? zusammen mit dem von $\text{can}_t : V \xrightarrow{\sim} V^*$ induzierten Isomorphismus $\text{Alt}^q(V^*) \xrightarrow{\sim} \text{Alt}^q(V)$ liefert dann in der Verknüpfung schließlich unseren Hodge-Operator.

10.9.4. Gegeben ein Isomorphismus $\varphi : V \xrightarrow{\sim} W$ von endlichdimensionalen orientierten reellen Vektorräumen und nichtausgeartete symmetrische Bilinearformen

t auf V und s auf W und eine Zerlegung $n = p + q$ der Dimension n unserer beiden Vektorräume kommutiert offensichtlich das Diagramm

$$\begin{array}{ccc} \text{Alt}^p W & \xrightarrow{*s} & \text{Alt}^q W \\ \downarrow & & \downarrow \\ \text{Alt}^p V & \xrightarrow{*t} & \text{Alt}^q V \end{array}$$

10.9.5. Die in der obigen Definition 10.9.3 versteckten und in gewisser Weise zufälligen Wahlen von Vorzeichen werden wie auch hier üblicherweise so getroffen, daß im Fall eines Skalarproduktes t für alle α gilt

$$\alpha \wedge * \alpha \in \mathbb{R}_{\geq 0} \omega$$

Wir werden das gleich explizit sehen. Es wäre auch nicht besser oder schlechter, die Vorzeichen so zu wählen, daß das “umgekehrte Dach-Produkt” in diesem Sinne “positiv definit” wäre, aber an dieser Stelle muß man sich eben nun einmal auf eine Konvention einigen.

10.9.6 (**Explizite Formeln für den Hodge*-Operator**). Sei zunächst t ein Skalarprodukt, v_1, \dots, v_n eine orientierte Orthonormalbasis von V und f_1, \dots, f_n die duale Basis von V^* . Wir folgern $\omega = f_1 \wedge \dots \wedge f_n$ und gegeben I, J mit $|I| = p$ und $|J| = q$ haben wir

$$f_I \wedge f_J = \begin{cases} \varepsilon_I \omega & I \sqcup J = \{1, \dots, n\}; \\ 0 & \text{sonst,} \end{cases}$$

wobei ε_I das Vorzeichen der Permutation meint, die alle Elemente von I an den Anfang schiebt, ihre Reihenfolge untereinander aber ebenso wie die Reihenfolge der Elemente ihres Komplements unverändert läßt. Unsere Abbildung

$$\text{Alt}^p V \rightarrow (\text{Alt}^q V)^*$$

macht also die Basisvektoren f_I bis auf Vorzeichen zu den Vektoren der dualen Basis, genauer haben wir $f_I \mapsto \varepsilon_I f_{\bar{I}}^*$ für \bar{I} das Komplement von I . Unter $(\text{Alt}^q V)^* \xrightarrow{\sim} \text{Alt}^q(V^*)$ entspricht dieser Vektor dann $\varepsilon_I v_{\bar{I}}$ und unter can_t wiederum $\varepsilon_I f_{\bar{I}}$, woraus wir folgern

$$*_t f_I = \varepsilon_I f_{\bar{I}} \quad \text{für } \bar{I} = \{1, \dots, n\} \setminus I.$$

Insbesondere gilt also im Fall eines Skalarproduktes und der dualen Basis zu einer Orthonormalbasis die Formel $f_I \wedge *_t f_I = \omega$, die in diesem Fall auch sofort $*(\alpha) = (-1)^{pq} \alpha$ für alle $\alpha \in \text{Alt}^p(V)$ liefert. Ist allgemeiner im symmetrischen nicht ausgearteten Fall v_1, \dots, v_n orientiert und orthogonal, aber haben wir etwa

$t(v_i, v_i) = \eta_i = \pm 1$, so müssen wir nur ganz am Schluß noch ein Vorzeichen ergänzen und erhalten mit der Notation $\eta_{\bar{I}} = \prod_{j \in \bar{I}} \eta_j$ die Formel

$$*_t f_I = \varepsilon_I \eta_{\bar{I}} f_{\bar{I}} \quad \text{für } \bar{I} = \{1, \dots, n\} \setminus I.$$

Sei nun noch allgemeiner t symmetrisch nicht ausgeartet, v_1, \dots, v_n eine orientierte Orthogonalbasis von V mit $t(v_i, v_i) = \eta_i c_i^2$ mit $c_i > 0$ und f_1, \dots, f_n die duale Basis von V^* . Gegeben I, \bar{I} mit $|I| = p$ erhalten wir dann mit der Notation $c_I = \prod_{i \in I} c_i$ durch Reskalierung die Formel

$$*_t f_I = \frac{\varepsilon_I \eta_{\bar{I}} c_{\bar{I}}}{c_I} f_{\bar{I}} \quad \text{für } \bar{I} = \{1, \dots, n\} \setminus I.$$

Definition 10.9.7. Gegeben eine offene Teilmenge $U \subseteq X$ eines endlichdimensionalen reellen Raums X und eine Riemann'sche Metrik t auf U und ein differenzierbares Vektorfeld $v : U \rightarrow \vec{X}$ definieren wir die **Divergenz** unseres Vektorfelds als die Funktion

$$\operatorname{div}_t(v) = (*_t \circ d \circ *_t \circ \operatorname{can}_t)(v)$$

Obwohl der $*$ -Operator von einer zu wählenden Orientierung abhängt, ist die Divergenz wegen des doppelten Auftretens unseres $*$ -Operators davon unabhängig.

Ergänzende Übung 10.9.8. Man zeige, daß die Divergenz eines stetig differenzierbaren Vektorfelds auf \mathbb{R}^n genau die "lokale Volumenänderung unter dem Fluß des besagten Vektorfelds" beschreibt, daß genauer für jede stetige Funktion mit kompaktem Träger f gilt

$$\left. \frac{d}{dt} \right|_{t=0} \int f \circ X^t = - \int f \operatorname{div} X$$

Hier ist zu beachten, daß auf jedem Kompaktum der Fluß für eine positive Zeitspanne existiert. Hinweis: Man schränke sich auf den Fall von glatter f ein, ziehe die zeitliche Ableitung unter das Integral, und beachte, daß das Integral über ganz \mathbb{R}^n jeder partiellen Ableitung einer stetig differenzierbaren Funktion mit kompaktem Träger verschwindet.

Beispiel 10.9.9. Sei $X = \mathbb{R}^3$ mit der Standardorientierung und dem Standardskalarprodukt $t = s$ versehen. Gegeben ein differenzierbares Vektorfeld der Gestalt $v = a\partial_x + b\partial_y + c\partial_z$ mit differenzierbaren Funktionen $a, b, c : \mathbb{R}^3 \rightarrow \mathbb{R}$ finden wir die übliche Formel $\operatorname{div} v = a_x + b_y + c_z$, indem wir rechnen

$$\begin{aligned} v &= a\partial_x + b\partial_y + c\partial_z \\ \operatorname{can}_s(v) &= adx + bdy + cdz \\ *_s(\operatorname{can}_s(v)) &= ady \wedge dz - bdx \wedge dz + cdx \wedge dy \\ d(*_s(\operatorname{can}_s(v))) &= a_x dx \wedge dy \wedge dz - b_y dy \wedge dx \wedge dz + c_z dz \wedge dx \wedge dy \\ *_s(d(*_s(\operatorname{can}_s(v)))) &= a_x + b_y + c_z \end{aligned}$$

Hier wäre es zwar in der Tat sehr viel einfacher gewesen, schlicht diese letzte Formel hinzuschreiben. Unsere neue Interpretation verträgt sich jedoch besser mit der Verwandtschaft, insbesondere da die äußere Ableitung d sich so gut mit Verwandtschaft verträgt, und ermöglicht so eine übersichtliche Darstellung in anderen orthogonalen Koordinaten. Um etwa die Divergenz in Polarkoordinaten zu bestimmen, erinnern wir uns daran, daß nach 6.2.9 unter der Polarkoordinatenabbildung P die Standardmetrik $s = dx^{\otimes 2} + dy^{\otimes 2}$ auf der xy -Ebene verwandt ist zum 2-Tensor $g = dr^{\otimes 2} + r^2 d\vartheta^{\otimes 2}$ und rechnen

$$\begin{aligned} v &= a\partial_r + b\partial_\vartheta \\ \text{can}_g(v) &= adr + br^2d\vartheta \\ *_g(\text{can}_g(v)) &= ar d\vartheta - br dr \\ d(*_g(\text{can}_g(v))) &= (a_r r + a)dr \wedge d\vartheta + b_\vartheta r dr \wedge d\vartheta \\ *_g(d(*_g(\text{can}_g(v)))) &= a_r + b_\vartheta + r^{-1}a \end{aligned}$$

Übung 10.9.10. Wir betrachten wieder Kugelkoordinaten wie in 6.2.11. Man zeige, daß für das zum Vektorfeld $a\partial_r + b\partial_\vartheta + c\partial_\varphi$ verwandte Feld auf dem xyz -Raum die Divergenz verwandt ist zur Funktion $a_r + b_\vartheta + c_\varphi + 2r^{-1}a + b \cot \vartheta$.

Definition 10.9.11. Gegeben $U \subseteq \mathbb{R}^n$ und eine zweimal stetig differenzierbare Funktion $f : U \rightarrow \mathbb{R}$ setzen wir

$$\Delta f = \frac{\partial^2 f}{\partial x_1^2} + \dots + \frac{\partial^2 f}{\partial x_n^2}$$

und nennen Δ den **Laplaceoperator**.

10.9.12 (**Anschauung für den Laplaceoperator**). Der Wert von Δf an einer Stelle x mißt die Abweichung des Funktionswerts von f bei x vom Durchschnitt der Funktionswerte in einer kleinen Umgebung von x . In einer Veränderlichen gilt zum Beispiel für jede zweimal stetig differenzierbare Funktion

$$f''(x) = \lim_{\varepsilon \rightarrow 0} \frac{2}{\varepsilon^2} \left(\frac{f(x + \varepsilon) + f(x - \varepsilon)}{2} - f(x) \right)$$

wie der Leser mithilfe der Taylorentwicklung leicht nachprüfen kann. In mehreren Veränderlichen gilt in derselben Weise für jede zweimal stetig differenzierbare Funktion mit der Notation e_i für die Vektoren der Standardbasis

$$(\Delta f)(x) = \lim_{\varepsilon \rightarrow 0} \frac{2^n}{\varepsilon^2} \left(\frac{1}{2^n} \left(\sum_{i=1}^n f(x + \varepsilon e_i) + f(x - \varepsilon e_i) \right) - f(x) \right)$$

Übung 10.9.13 (Mehr Anschauung für den Laplaceoperator). Man zeige, daß der Laplaceoperator invariant ist unter Drehungen. Ist genauer $A \in O(n)$ eine orthogonale Matrix und bezeichnet $A : \mathbb{R}^n \rightarrow \mathbb{R}^n$ die zugehörige lineare Abbildung, so zeige man für jede zweimal stetig differenzierbare Funktion $f : \mathbb{R}^n \rightarrow \mathbb{R}$ die Formel $\Delta(f \circ A) = (\Delta f) \circ A$. Man folgere die Formel

$$(\Delta f)(x) = \lim_{\varepsilon \rightarrow 0} \frac{2^n}{\varepsilon^2} \left(\frac{\int_{\|y-x\|=\varepsilon} f(y) \sigma\langle y \rangle}{\int_{\|y-x\|=\varepsilon} \sigma\langle y \rangle} - f(x) \right)$$

auf deren rechter Seite nach dem Faktor $2^n/\varepsilon^2$ bis auf ein Vorzeichen die Differenz zwischen dem Funktionswert $f(x)$ und dem Durchschnitt der Funktionswerte auf einer Kugelschale mit Zentrum in x und Radius ε steht. Hinweis: Man middle [10.9.12](#). Die Taylorentwicklung oben liefert in einer Veränderlichen sogar präziser die Darstellung

$$\frac{2}{\varepsilon^2} \left(\frac{f(x+\varepsilon) + f(x-\varepsilon)}{2} - f(x) \right) = (f''(\xi^+) + f''(\xi^-))/2$$

mit $\xi^+ \in (x, x+\varepsilon)$ und $\xi^- \in (x-\varepsilon, x)$.

Ergänzende Übung 10.9.14. Die polynomialen Funktionen $D \in \mathbb{C}[X_1, \dots, X_n]$ auf dem \mathbb{R}^n , die invariant sind unter allen Drehungen $A \in SO(n)$, sind genau alle Polynome im quadrierten Abstand vom Nullpunkt, in Formeln

$$\mathbb{C}[X_1, \dots, X_n]^{SO(n)} = \mathbb{C}[(X_1^2 + \dots + X_n^2)]$$

Die Differentialoperatoren $D \in \mathbb{C}[\partial_1, \dots, \partial_n]$ mit konstanten Koeffizienten auf dem \mathbb{R}^n , die invariant sind unter allen Drehungen $A \in SO(n)$, sind genau alle Polynome im Laplace-Operator, in Formeln

$$\mathbb{C}[\partial_1, \dots, \partial_n]^{SO(n)} = \mathbb{C}[\Delta]$$

10.9.15 (Laplace-Operator in anderen Koordinaten). Um den Laplaceoperator Δ in anderen Koordinaten auszudrücken, kann man von der Darstellung

$$\Delta f = *_s d *_s df$$

ausgehen, mit s der üblichen Riemann'schen Metrik auf \mathbb{R}^n und $*_s$ dem zu dieser Metrik und der Standard-Orientierung gehörenden Hodge-*-Operator. Gegeben ein endlichdimensionaler reeller Raum X und eine offene Teilmenge $V \subseteq X$ und ein differenzierbare Abbildung $\phi : V \rightarrow U$ mit bijektivem Differential an jeder Stelle und eine zur Standard-Metrik ϕ -verwandte Riemann'sche Metrik t auf V haben wir dann Verwandtschaft von $\Delta f = *_s d *_s df$ und $*_t d *_t d(f \circ \phi)$. Ist speziell etwa ϕ die Polarkoordinaten- oder die Kugelkoordinatenabbildung, so läßt sich das auch sehr konkret und explizit berechnen.

Beispiel 10.9.16. Wir berechnen den Laplace-Operator auf einer Funktion f in Polarkoordinaten und finden ähnlich wie in [10.9.9](#) der Reihe nach

$$\begin{aligned} df &= f_r dr + f_\vartheta d\vartheta \\ *_g(df) &= f_r r d\vartheta - r^{-1} f_\vartheta dr \\ d(*_g(df)) &= (f_{rr} r + f_r + r^{-1} f_{\vartheta\vartheta}) dr \wedge d\vartheta \\ *_g(d(*_g(df))) &= f_{rr} + r^{-1} f_r + r^{-2} f_{\vartheta\vartheta} \end{aligned}$$

Übung 10.9.17. Man zeige, daß der Laplace-Operator auf einer Funktion f in den Kugelkoordinaten aus [6.2.11](#) gegeben wird durch die Formel

$$\Delta f = f_{rr} + 2r^{-1} f_r + r^{-2} f_{\vartheta\vartheta} + r^{-2} f_\vartheta \cot \vartheta + (r \sin \vartheta)^{-2} f_{\varphi\varphi}$$

Hinweis: Statt das direkt zu rechnen, kann man auch von [6.2.13](#) und [10.9.10](#) ausgehen.

Index

- [A] charakteristische Funktion, 223
- ⊗
 - Produkt- σ -Algebra, 234
 - Produktmaß, 234
- χ_A charakteristische Funktion, 223
- $\langle f \rangle$ bei multilinearen Abbildungen, 195
- $\mu, _A$ Restriktion eines Maßes 226
- ⊗
 - Notation für Bilinearform, 95
- \wedge
 - äußeres Dachprodukt, 259
 - Dachprodukt, 256
- \rightsquigarrow verwandt
 - 2-Tensoren, 97
 - Differentialformen, 262
 - Funktionen, 89
 - Maße, 222
 - Vektorfelder, 89
 - Wege, 105
- *
- Hodge-*-Operator, 305
- Ableitung
 - höhere, koordinatenfrei, 196
 - komplexe, 20
 - nach Vektorfeld, 87
 - partielle, 46, 65
 - Richtungs-, 50
 - von Brüchen, komplex, 21
 - von Umkehrfunktion
 - komplex, 22
- additiv
 - σ -additiv, 202, 208
- äußere Ableitung, 279
- äußeres Normalenfeld, 292
- Algebra, 37
 - σ -Algebra, 202
- alt Alternator, 255
- Alt
 - Raum alternierender Formen, 255
- alternierende p -Form, 255
- Amplitude, 27
- analytisch
 - auf \mathbb{R}^n , 111
- Anfangswert
 - von Integralkurve, 177
- Anfangswertisomorphismus, 29, 192, 193
 - bei reeller Schwingungsgleichung, 25
 - bei Schwingungsgleichung, 27
- antisymmetrisch, 99
- Arbeit
 - gegen Kraftfeld, 108
- auswerten
 - Kovektorfeld auf Vektorfeld, 85
- Banach'scher Fixpunktsatz, 131
- Banach-Tarski-Paradoxon, 207
- Benford's Gesetz, 215
- Beppo Levi, 234
- Bil Bilinearformen, 95
- Bild
 - von σ -Algebra, 219
- Bildmaß, 222
- Borel(X) Borelmengen von X , 204
- Borelmaß, 204
- Borelmenge, 204
- Borromäische Ringe, 149
- $\mathcal{C}(X)$ stetige komplexwertige Funktionen auf X , 40
- \mathcal{C}^1 -Abbildung, 129
- \mathcal{C}^1 -Diffeomorphismus, 129
- \mathcal{C}^∞ -Abbildung
 - zwischen affinen Räumen, 196
- \mathcal{C}^k -Abbildung

zwischen affinen Räumen, 196
 \mathcal{C}^k -Diffeomorphismus, 196
 $\mathcal{C}_c(X, \mathbb{R})$ stetige Funktionen mit kompaktem Träger, 156
 Cantor-Menge, 216
 Cavalieri, 239
 χ_A charakteristische Funktion, 223
 charakteristische Funktion einer Menge, 223
 cis, 6
 dx_I , 261
 ∂M Rand
 der berandeten Untermannigfaltigkeit M , 287
 ∂_i Vektorfeld, 87
 ∂_x Vektorfeld, 87
 D_v Richtungsableitung, 50
 Dachprodukt, 256
 df Differential von f , 85
 dicht, 37
 Diffeomorphismus
 \mathcal{C}^1 -Diffeomorphismus, 129
 \mathcal{C}^k -Diffeomorphismus, 196
 Differential, 50
 totales, 53
 von \exp auf Matrizen, 71
 von bilinearer Abbildung, 59
 differential equation
 ordinary, 175
 partial, 175
 Differentialform
 erster Ordnung, 82
 relative, 260
 Differentialgleichung
 gewöhnliche, 175
 partielle, 175
 Differentialquotient, 56
 differenzierbar
 beliebig, 195
 mehrere Veränderliche, 50
 partiell, 51
 total, 51
 Differenzieren unter dem Integral, 70
 Dirac-Maß, 203
 directional derivative, 50
 diskret
 in topologischem Raum, 110
 relativ, 110
 Divergenz, 296, 307
 dominierte Konvergenz, 231
 Dualraum, 82
 $E(X)$ Erwartungswert der Zufallsvariable X , 230
 Eckenplättung, 299
 Eigenschwingungen, 31
 einfach
 zusammenhängend, wegweise, 112
 einfach zusammenhängend, 114
 Einheitswurzel
 in \mathbb{C} , 8
 Eins-Element
 einer Algebra, 37
 Eins-Form, 82
 einsetzen
 Vektorfeld in Kovektorfeld, 85
 Element
 zufälliges, 219
 Elementar-Ereignis, 203
 Energie, 107
 Ereignis, 203
 Erwartungswert
 einer Zufallsvariable, 230
 erzeugt
 σ -Algebra, 204
 Euler'sche Formel, 6
 explizit
 gewöhnliche Differentialgleichung, 175
 Extrema
 in mehreren Veränderlichen, 77

unter Nebenbedingungen, 154
 Fixpunkt, 129
 Fläche, 148
 Flächenmaß, 251
 Fluß, 277
 von Vektorfeld, 197
 1-Form, 82
 Fourierreihe, 35
 Fréchet-differenzierbar, 50
 Fubini, 240
 Fundamentalform, erste, 97
 Fundamentalsatz
 der Algebra, 10
 Funktion
 implizite, 138
 Funktionaldeterminante, 160
 Funktionentheorie, 24
 Gauss
 Integralsatz von, 296
 Gauss'sche Glockenkurve, 249
 Gedämpfte Schwingungen, 25
 geschlossen
 Differentialform, 279
 Kovektorfeld, 117
 Geschwindigkeit
 vektorielle, 56
 glatt, 285
 Abbildung
 koordinatenfrei, 196
 Abbildung nach \mathbb{R}^m , 195
 Untermannigfaltigkeit
 von affinem Raum, 288
 grad_g , 96
 Grad
 einer polynomialen Abbildung, 74
 eines Polynoms
 in mehreren Veränderlichen, 72
 Gradient, 48, 93
 Gram'sche Matrix, 169
 Green'sche Formel, 298
 halboffen
 in reellem affinen Raum, 53
 Hauptzweig des Logarithmus, 17
 Hermite-Lindemann, 8
 Hesse-Matrix, 81
 Hodge-*-Operator, 305
 holomorph, 24
 homotop, 112
 mit festen Endpunkten, 112
 Wege, 112
 Homotopie
 von Wegen, 112
 Hyperfläche, 148
 implizit
 Funktion, 138
 gewöhnliche Differentialgleichung,
 175
 indefinit, 79
 induzierte σ -Algebra, 219
 induzierte Karte, 290
 induzierte Orientierung, 290
 Integral
 erstes, einer Differentialgleichung,
 178
 nichtnegative meßbare Funktion
 über Maßraum, 224
 stetige reelle Funktion
 über kompakten Quader, 66
 stetige vektorwertige Funktion
 über kompaktes Intervall, 13
 Integralkurve, 177
 maximale, 177, 179
 Integration
 partielle, 244
 Integrationsweg, 114
 integrierbar, 229
 Differentialform, 271
 isotop

Verschlingungen, 148
 Jacobi-Matrix, 51
 kanonische Fortsetzung
 von Prämaß, 210
 Karte, 150
 Kartenwechsel, 154, 289
 Kettenregel
 in einer Veränderlichen
 komplex, 21
 in mehreren Veränderlichen, 57
 Knoten, 148
 Kodimension
 einer Untermannigfaltigkeit, 148
 komplex differenzierbar, 19
 komplexe Exponentialfunktion, 4
 Komponente
 Wegzusammenhangskomponente, 108
 Komponentenregel, 58
 kontrahierend
 Abbildung metrischer Räume, 129
 kontravariant, 82
 Konvergenz
 im quadratischen Mittel, 43
 Koordinaten, 146
 Koordinatensystem, 146
 kovariant, 82
 Kovektorfeld, 82
 vektorwertiges, 107
 kritische Stelle, 77
 Kugelkoordinaten, 99
 Kugelvolumen, 240
 Kurve, 148

 λ^n Lebesgue-Maß auf \mathbb{R}^n , 205
 Lagrange'sche Multiplikatoren, 154
 Laplaceoperator
 im \mathbb{R}^n , 308
 Lebesgue-Borel-Maß, 205
 Lebesgue-Mengen, 218
 Lebesgue-meßbar, 223

 Lebesgue-meßbaren, 218
 Leibniz-Regel
 für Differentialformen, 281
 für komplexe Funktionen, 21
 linksseitig stetig, 209
 Lipschitz-Konstante, 129
 Lipschitz-stetig, 129
 lipschitzstetig
 lokal, 186
 Lösungsraum
 einer linearen Differentialgleichung,
 29
 einer linearen Differentialgleichung
 allgemeiner Fall, 192
 einer Schwingungsgleichung, 25, 26
 Logarithmus
 Hauptzweig des komplexen, 17
 komplexer, 22
 lokal kompakt
 topologischer Raum, 245
 lokale Koordinaten, 152
 lokales Koordinatensystem, 152

 Mannigfaltigkeit
 berandete Untermannigfaltigkeit von
 affinem Raum, 285
 Untermannigfaltigkeit mit Ecken von
 affinem Raum, 299
 Untermannigfaltigkeit von affinem
 Raum, 146
 Mantelfläche, 253
 Maß, 202
 einer Differentialform, 263
 mit Werten in einem orientierter Ge-
 rade, 251
 nichtnegatives, 202
 topologisches, 204
 Maßraum, 202
 maximal
 Integralkurve, 177
 Maximum

- isoliertes lokales, 77
- lokales, 77
- Mengenalgebra, 201
- Mengenring, 207
- meßbar
 - μ -meßbar, Funktion, 223
 - μ^* -meßbar, Menge, 213
 - Abbildung, 218
 - Menge, 202
- meßbar und integrierbar, 229
- Meßraum, 202
- Minimum
 - isoliertes lokales, 77
 - lokales, 77
- monotone Konvergenz, 225
- Mult^k multilineare Abbildungen, 195
- Multiindex, 72
- Nabla-Operator, 298
- negativ definit, 79
- nichtnegativ
 - Differentialform, 271
- Normalenfeld
 - orientiertes, 276
- normiert
 - Weg, 112
- Normierung, 205
- Nullmenge, 216
- ⊗
 - Notation für Bilinearform, 95
- ODE ordinary differential equation, 175
- Ordnung
 - einer gewöhnlichen Differentialgleichung, 175
- orientierbar
 - C^1 -Mannigfaltigkeit, 270
- orientiert
 - C^1 -Mannigfaltigkeit, 270
- Orientierung
 - von C^1 -Mannigfaltigkeit, 270
- von Vektorraum, 268
- Parametrisierung
 - orientierte, 275
- partiell
 - Ableitung, 46, 65
 - differenzierbar, 46
 - Integration, 244
- PDE partial differential equation, 175
- Periode, 35
- Pfaff'sche Form, 82
- Phase, 27
- Picard-Lindelöf, 189
- Plättung, 146
- Polarkoordinaten
 - Gradient in, 95
- Polygonzug, 108
- Polynom, 72
 - trigonometrisches, 40
- polynomial
 - Abbildung $\mathbb{R}^n \rightarrow \mathbb{R}^m$, 74
- positiv definit, 79
- positiv semidefinit, 79
- Potential
 - eines Kraftfelds, 108
- Prämaß, 208
- Prinzip von Cavalieri, 239
- Produkt- σ -Algebra, 234
- Produktmaß, 234
- Produktregel, 21
- Quader, 66
- quadratische Form, 79
- Quelldichte, 296
- Quotientenregel
 - im Komplexen, 21
- Rand
 - einer berandeten Untermannigfaltigkeit, 287
- Randkarte, 289
- Randplättung, 285

Rechteck, 66
 regulär
 Abbildung $\mathbb{R}^n \rightarrow \mathbb{R}^m$, 74
 Teil eines Randes, 299
 Resonanz, 34
 Resonanzkatastrophe, 34
 Richtungsableitung, 50
 Riemann
 Riemann'sche Metrik, 95
 Riemannsumme
 für Funktion auf Mannigfaltigkeit, 170
 für Funktion auf Rechteck, 66
 für Integral einer Volumenform, 273
 für Integral nach Maß, 234
 für Maß einer Differentialform, 267
 für vektorwertige Funktion, 12
 für Wegintegrale, 102
 Ringalgebra, 37
 Rotation, 119
 Rotationskörper, 240
 Rückzug
 von Funktionen, 89
 von Kovektorfeld, 89

 $\sigma(\mathcal{A})$, von \mathcal{A} erzeugte σ -Algebra, 204
 S^n die n -Sphäre, 150
 σ -Subadditivität, 212
 σ -additiv, 202, 208
 σ -Algebra, 202
 erzeugt von Mengensystem, 204
 Sard
 Satz von, 247
 σ -endlich, 210
 separabel
 topologischer Raum, 220
 Separation der Variablen, 183, 184
 Shuffle, 257
 skalare Rotation
 eines ebenen Vektorfeldes, 119
 Stammfunktion, 16

 stetig
 linksseitig, 209
 stimmen ueberein bis zur Ordnung d , 75
 Stokes
 Integralsatz von
 allgemeiner, 292
 klassischer, 297
 Stokes, mit Ecken, 299
 Stone-Weierstraß, 37
 Stufenfunktion, 223
 Summenregel, 21, 59
 support, 156
 symmetrisch, 99

 TM , *siehe* Tangentialbündel
 T_pM , *siehe* Tangentialraum
 Tangentialbündel
 im eingebetteten Fall, 268
 Tangentialraum
 im eingebetteten Fall, 268
 Taylorentwicklung
 in mehreren Veränderlichen, 71
 Teilung der Eins, 165
 Tensor
 2-Tensor, 95
 Tonelli, Satz von, 237
 topologisch
 Maß, 204
 meßbar, 204
 Totalgrad
 eines Polynoms
 in mehreren Veränderlichen, 72
 Träger
 einer Funktion, 156
 Transformationsformel
 bei kompaktem Träger, 157
 für das Lebesgue-Integral, 246
 translationsinvariant, 205
 trennt die Punkte, 37
 Treppenfunktion, 13
 trigonometrisches Polynom, 40

Umkehrsatz, 129
 Unteralgebra, 37
 Untermannigfaltigkeit
 der Klasse C^l , von affinem Raum, 288
 glatte, von affinem Raum, 288
 von affinem Raum, 146
 Unterringalgebra, 37
 Urbild
 von σ -Algebra, 220

 Variablentrennung, 184
 Variation der Konstanten, 32, 194
 Vektorfeld, 82, 177
 vektorwertig
 Kovektorfeld, 107
 Verfeinerung
 von Wahrscheinlichkeitsraum, 203
 Verschlingung, 148
 Vertauschen
 von partiellen Ableitungen, 69
 von partiellen Integrationen, 69
 Verteilung
 einer Zufallsvariable, 222
 Verteilungsfunktion
 eines Wahrscheinlichkeitsmaßes, 211
 von Zufallsvariable, 222
 Vervollständigung
 von Maßraum, 218
 verwandt
 2-Tensoren, 97
 Differentialformen, 262
 Funktionen, 89
 Kovektorfelder, 89
 Maße, 222
 Vektorfelder, 89
 Wege, 105
 vollständig
 Maßraum, 216
 Volumen, 169
 Volumenform, 305

 kanonische, 305
 Wahrscheinlichkeitsraum, 203
 Weg
 geschlossen, 112
 normierter, 112
 stückweise linear, 108
 zusammenziehbar, 112
 Wegintegral, 99
 für Kovektorfeld, 99
 für Vektorfeld, 103
 vektorwertiges, 107
 wegweise einfach zusammenhängend, 112
 Wegzusammenhang, 108
 Wegzusammenhangskomponente, 108
 Weierstraß'scher Approximationssatz, 38
 Winkelgeschwindigkeit, 27
 Wirbeldichte, 119

 \boxtimes
 Produkt- σ -Algebra, 234
 Produktmaß, 234

 \otimes
 Produkt- σ -Algebra, 235

 Zählmaß, 204
 Zerleger, in der Maßtheorie, 213
 Zirkulation, 103
 zufälliges Element, 219
 Zufallsvariable, 203, 219
 zusammenhängend
 einfach, 114
 wegweise, 108
 wegweise einfach, 112
 zusammenziehbar
 geschlossener Weg, 112
 Zweig des Logarithmus, 22
 zweitabzählbar
 topologischer Raum, 220