
Höhere Mathematik für Physiker

Teil I

F. Krause

Kapitel 6

Analysis von Funktionen mehrerer Veränderlicher

Copyright F.Krause

Inhalt des Kapitels **Analysis von Funktionen mehrerer Veränderlicher**

- **6.0 Vorbemerkung**
- **6.1 Erster Zugang zum Verhalten der Abbildungen**
 - **6.1.0 Übersicht**
 - **6.1.1 Geschwindigkeitsfelder**
 - * **6.1.1a Analytische Darstellungen der Felder**
 - * **6.1.1b Die Rotationsbewegung eines starren Körpers**
 - * **6.1.1c Radiales Strömungsfeld**
 - **6.1.2 Inspektion: Die Typisierung der Abbildungen anhand der beteiligten Dimensionen**
 - **6.1.3 Inspektion: Allgemeine Hilfsmittel**
 - **6.1.4 Parametrisierungen**
 - * **6.1.4a Ebene Polarkoordinaten**
 - * **6.1.4b Zylinderkoordinaten**
 - * **6.1.4c Kugelkoordinaten oder räumliche Polarkoordinaten**
 - **6.1.5 Das Umsetzen geometrischer Bedingungen mit Hilfe von Parametrisierungen**
 - **6.1.6 Quantifizierung des Werteraumes mit Hilfe von Basisfeldern**
 - * **6.1.6a Das begleitende Zweibein einer Kurve**
 - **6.1.7 Koordinatendarstellungen der Abbildungen**
 - **6.1.8 Symmetriebetrachtungen**

- **6.2 Normierte Räume / Grenzwert und Stetigkeit**
 - **6.2.1 Vorbemerkung**
 - **6.2.2 Normierte Vektorräume**
 - **6.2.3 Vektorfolgen**
 - **6.2.3a Zwei Beispiele zur Bedeutung der Grenzwertkonstruktion**
 - **6.2.3b Die Konvergenz von Abbildungsfolgen unter der Supremumsnorm**
 - **6.2.4 Stetigkeit von Abbildungen**
 - **12.2.4a Offene Teilmenge**
 - **(Kompakte Teilmengen)**

- Kap. 6.3: Die totale Ableitung
 - 6.3.0 Das Szenenbild
 - 6.3.1 Der allgemeine Formalismus
 - 6.3.2 Quantifizierung und analytische Darstellung der totalen Ableitung
 - * 6.3.2a Quantifizierung der Totalen Ableitung über die Vorgabe eines konstanten Basisfeldes im Urbildraum
 - * 6.3.2b Der Übergang zu Funktionen der Koordinaten
 - * 6.3.2c Wahl einer Basis in Urbildraum und Werteraum (Konstante Basisfelder)
 - * 6.3.2d Bekannte Spezialfälle der totalen Ableitung
 - * 6.3.2e Das Potential eines mathematischen Dipols
 - 6.3.3 Lokale Abbildungsgeometrie um den Aufpunkt $(x_0, f(x_0))$
 - Der Schluß von der Rechnung auf die Geometrie
 - * 6.3.3a Präzisierung des Tangentialraumbegriffes
 - * 6.3.3b Die Bestimmung von Tangentialräumen mit Hilfe der totalen Ableitung
 - * 6.3.3c Das Rangkriterium
 - 6.3.4 Restriktionsgeometrie und krummlinige Koordinaten
 - * 6.3.4a Koordinatenpartitionen des Konfigurationsraumes
 - * 6.3.4b Koordinatenbasisfelder
 - * 6.3.4c Darstellungen mit Hilfe der Koordinatenbasisfelder
 - * 6.3.4d Verallgemeinerungen und Modifikationen
 - 6.3.5 Über den Nutzen des Formalismus / Zusammenfassung

- 6.4 Kurven
 - – 6.4.0 Vorbemerkung
 - 6.4.1 Höhere Ableitungen von Kurven
 - * 6.4.1a Entfaltung des Glattheitsbegriffes im Kurvenfall
 - * 6.4.1b Zur Bedeutung der Beweise
 - * 6.4.1c Polynomapproximationen
 - * 6.4.1d Der Satz vom beschränkten Zuwachs.
 - 6.4.2 Stammfunktion und Integral
 - * 6.4.2a Das Integral als Grenzwert von Summen
 - * 6.4.2b Die Komponentendarstellung des Integrales
 - 6.4.3 Die Taylorentwicklung der Kurven
 - 6.4.4 Funktionsdefinition mit Hilfe von Reihen

- * 6.4.4a Potenzreihen
- 6.4.5 "e-hoch Matrix"
 - * 6.4.5a Durch e^{tM} erzeugte Entwicklungsgruppen
- 6.4.6 Weg- und Arbeitsintegrale
- 6.4.7 Übersicht
- 6.4.8 Die Corioliskraft

Kapitel 6

Analysis von Funktionen mehrerer Veränderlicher

6.0 Vorbemerkung

In diesem Kapitel soll die Analysis von Funktionen mehrerer reeller Veränderlicher entwickelt werden. Allgemeines Ziel ist die **Verhaltensanalyse von glatten Abbildungen** $f:V \rightarrow W$, wobei V und W endlichdimensionale reelle Vektorräume sind. Es geht darum, einen einheitlichen Formalismus zu entwickeln, der Kurven, Felder und geometrische Transformationen ebenso wie gewöhnliche reelle Funktionen erfaßt. Dieser Formalismus ist von größtem Nutzen, da in den Anwendungen meist Funktionen mehrerer Veränderlicher auftreten. Ebenso ist der zu entwickelnde Formalismus Grundlage weiterführender Mathematik, wie der Differentialgeometrie.

Ein zentrales Problem besteht darin, sich eine **ganzheitliche Vorstellung** vom Verhalten der jeweiligen Abbildung zu verschaffen. Dem stehen zwei Komplikationen entgegen: Zum einen die Anzahl der beteiligten Raumdimensionen nebst den vielfältigen Möglichkeiten, diese Dimensionen miteinander zu kombinieren. Zum anderen die zugelassene Nichtlinearität der Abbildungen, die eine enorme Vielfalt komplexesten Verhaltens bewirkt. Die Analysis entwickelt zur Behandlung beider Problemkreise ein mächtiges Werkzeug, die totale Ableitung. Dieses Werkzeug erlaubt die lokale, auf einen Punkt bezogene, Rückführung der Abbildung auf lineare Abbildungen. Und lineare Abbildungen kann man mit den Methoden der linearen Algebra behandeln. Man erhält die totale Ableitung, indem man eine Tangentenerlegung für den betrachteten Funktionstyp entwickelt.

Das Kapitel gliedert sich in mehrere Teilkapitel mit folgendem Inhalt:

- Kap. 6.1: **Erster Zugang**. Der erste Zugang zum Verhalten allgemeiner Abbildungen $f:V \rightarrow W$ kann und muss auf unterschiedliche Weise angegangen werden. Die Vielzahl der Möglichkeiten erschwert den Einstieg, erscheint vielen zunächst als verwirrende Barriere. Andererseits gibt es eine Reihe von Denkrichtungen, die sich - gepaart mit etwas Urteilsvermögen - fast immer als nützlich erweisen, so dass man von ihnen ausgehen sollte. Im ersten Teil des Kapitels beschäftigen wir uns einleitend mit dieser Frage des ersten Einstiegs in die Verhaltensanalyse.
- Kap. 6.2: **Normierte Räume**: Der Begriff des Abstandes von Vektoren muß formalisiert werden, um Glattheits- und Grenzwerteigenschaften überhaupt behandeln zu können. Die dazu benötigten topologischen Grundbegriffe werden eingeführt.
- Kap. 6.3: **Die totale Ableitung**. Der Ableitungsformalismus für Funktionen der betrachteten Art wird entwickelt. Mit Hilfe des (im Vorkurs entwickelten) Konzeptes der Tangentenerlegung ergibt sich eine einfache Anwendung der linearen Algebra, deren Ausführung dann wichtige Resultate liefert.
- Kap. 6.4: **Anwendung auf Kurven**. Für das nachfolgende Kapitel 7 über lineare Differentialgleichungen benötigen wir einige weitergehende Resultate über Kurven, die in diesem Teilkapitel bereitgestellt werden. Unabhängig davon werden die Resultate über Kurven für zahlreiche Anwendungen benötigt. Höhere Ableitungen und Integrale lassen sich für glatte Kurven problemlos einführen.

Der allgemeine Ableitungsbegriff aus Kapitel 6.3 wird sich - wie gesagt - als wichtiges Anwendungsbeispiel der linearen Algebra erweisen. Dabei ist "Anwendung" in einem weiten Sinne gemeint: Nicht einfach Konkretisierung auf einen speziellen Fall, sondern Anwendung auf eine neue Problemsituation. Die Vektorraumaxiome müssen dazu durch Einbau eines Längen- und Abstandsbegriffes ergänzt werden. Schließlich müssen einige spezifische Begriffsbildungen und Ideen der Analysis eingebracht bzw. entwickelt werden. Dann allerdings lassen sich die in der linearen Algebra entwickelten Konzepte fruchtbar und sehr effizient anwenden.

Achtung: Im gesamten Kapitel entscheidet vorrangig die Mengenangabe, ob eine Größe Vektor oder Skalar ist. Fettdruck wird teilweise zur zusätzlichen Betonung des Vektorcharakters einer Größe eingesetzt. Normale Buchstaben können aber durchaus Vektoren sein.

Kap.6.1 Erster Zugang zum Verhalten der Abbildungen

6.1.0 Übersicht

Im vorliegenden Teilkapitel besprechen wir eine Reihe von Leitideen oder Leitmodellen, mit deren Hilfe man sich einen ersten meist schnellen Zugang zum Verhalten der Abbildungen verschafft, die aber auch dazu dienen, später hinzukommende Einsichten und Informationen in das Gesamtbild zu integrieren.

Die folgenden Hauptrichtungen und Methoden für diesen ersten Zugang werden behandelt:

- 6.1.1 **Geschwindigkeitsfelder:** Ein Beispiel einer Prozessinterpretation von Abbildungen. (Vgl. Kap 1.2.7, speziell 1.2.7a)
- 6.1.2 **Typisierung der Abbildungen nach beteiligten Dimensionen** und Analogien mit einfachen Abbildungstypen. (Vgl. Kap. 1.2.6)
- 6.1.3 Hilfsmittel, die aus der Verallgemeinerung der Methoden der elementaren **Kurvendiskussion** entstehen. (Vgl. Vorkurs)
- 6.1.4 **Parametrisierung und Quantifizierung des Konfigurationsraumes**
- 6.1.5 Die **Umsetzung geometrischer Bedingungen in Formeln.**
- 6.1.6 **Quantifizierung des Werteraumes - Basisfelder**
- 6.1.7 **Koordinatendarstellung** der Abbildungen
- 6.1.8 **Symmetriebetrachtungen**

6.1.1 Geschwindigkeitsfelder

(1.1.1) Es gibt eine Reihe von Vorgängen, deren Verlauf sich viele Menschen intuitiv gut vorstellen können. Beispiele sind Bewegungsvorgänge von Körpern oder auch Flüssigkeitsströmungen. Andererseits werden diese Vorgänge durch Abbildungen der hier interessierenden Art beschrieben. Dann kann man sich u.U. das Abbildungsverhalten mit Hilfe einer geeigneten Prozeßinterpretation zugänglich machen, auch dann, wenn die Abbildung zunächst in einem anderen Zusammenhang ohne eine solche Interpretation entsteht. Wir betrachten als Beispiel die Interpretation eines Vektorfeldes als **Strömungsfeld**, die sich als grundlegend für die geometrische Behandlung der Differentialgleichungen erweist. Man geht dabei vom Feldstandpunkt aus und **gibt den den einzelnen Punkten des Konfigurationsraumes zugeordnete Feldwerten die zusätzliche Interpretation einer Strömungsgeschwindigkeit.** Solche Prozessinterpretationen lassen sich im Rahmen der analytischen Geometrie dann auf andere Konfigurationsräume übertragen.

(1.1.2) Als Konfigurationsraum wählen wir eine geeignete Teilmenge $G \subset V_0^3$. In dieser Menge sollen sich alle betrachteten physikalischen Vorgänge abspielen. Wir stellen uns das folgende ideale System vor: Der gesamte Konfigurationsraum $G \subset V_0^3$ sei mit beweglichen materiellen Punkten angefüllt. Dann hat jeder dieser Punkte zu jedem Zeitpunkt einen bestimmten Ort und eine vektorielle Geschwindigkeit. Man kann einzelne dieser Punkte identifizieren (anfärben) und ihre Bahn verfolgen. Die Bewegung jedes Punktes wird durch eine (seine!) Bahnkurve beschrieben.

Bitte immer auseinanderhalten: Den Materiepunkt, der sich im Konfigurationsraum bewegt und den geometrischen Konfigurationsraumpunkt als möglichen Ort eines Materiepunktes. Ersterer wird mathematisch durch eine Bahnkurve, letzterer durch ein Element aus G beschrieben. Jeder Materiepunkt könnte im Prinzip noch weitere individuelle Eigenschaften besitzen, von denen wir hier jedoch absehen.

Das Ganze ist offensichtlich ein Modell einer idealisierten Flüssigkeits- oder Gasströmung, bei der man von der atomaren Struktur der Materie absieht und stattdessen eine immer weiter gehende Teilbarkeit

annimmt. Ebenso wird Glattheit der Strömung erwartet: Benachbarte Punkte sollen also auch benachbarte Geschwindigkeiten besitzen. Im Kapitel 6.2 werden die intuitiven Vorstellungen von benachbart und glatt präzisiert. Solange man außerhalb des Bereichs atomarer Dimensionen bleibt, ist das eine gute Idealisierung, deren Nützlichkeit darin besteht, dass alles durch glatte Abbildungen vom Feldtyp beschrieben werden kann. Vielfach hat man intuitive ausgeprägte ganzheitliche Vorstellungen über derartige Strömungsprozesse bzw. entwickelt solche relativ leicht.

(1.1.3) Zwei miteinander verwandte Veranschaulichungsverfahren für Geschwindigkeitsfelder bieten sich an, die beide vom Feldstandpunkt ausgehen:

- **1. Veranschaulichung:** Wir wählen zu einem Zeitpunkt N geometrische Punkte zufällig oder nach einem bestimmten System aus. An jedem dieser Punkte befindet sich dann ein materieller Punkt, dessen Bahn wir einen von uns wählbaren Zeitraum Δt verfolgen. Das Bild der entstehenden Bahnen ergibt eine Veranschaulichung des Strömungsverhaltens im Anschluß an den gewählten Zeitpunkt.
- **2 Veranschaulichung:** Man kann anstatt die Bahn eines individuellen Materiepunktes zu verfolgen - auch an einem festen geometrischen Punkt $\vec{x} \in G$ verbleiben und beobachten, wie nacheinander eine ganze Schar der materiellen Punkte durch ihn hindurchgeht, jeder mit einer bestimmten Geschwindigkeit. Dann kann man zu jedem Zeitpunkt t die vektorielle Geschwindigkeit des gerade passierenden Massenpunktes beobachten und bestimmen. Wir erhalten die Zuordnung $t \mapsto \vec{v}(t, \vec{x})$. Wählt man erneut N Punkte geeignet, so erhält jeder Punkt einen u.U. zeitlich veränderlichen Geschwindigkeitspfeil, was sich in Form einer Filmanimation vorstellen läßt.

Im ersten Fall beobachtet man die Bahn einer bestimmten Person in der Menge, im zweiten betrachtet man eine feste Stelle und bestimmt die Geschwindigkeit der jeweils passierenden Personen. (Beides läßt sich in der Ebene gut mit dem Computer simulieren.)

In der zweiten Veranschaulichung können wir zusätzlich den Beobachtungspunkt \vec{x} variieren, ihn per Rollenwechsel vom äußeren Parameter in eine unabhängige Variable umwandeln und erhalten so das allgemeine **zeitabhängige** Geschwindigkeitsfeld

$$\vec{v} = (G \times I, (\vec{x}, t) \mapsto \vec{v}(\vec{x}, t), V^3)$$

Klar entsteht eine Abbildung vom Feldtyp mit einem vierdimensionalen Urbildraum, dem man meist auch die Rolle eines Konfigurationsraumes gibt. (Kap 1.1.2)

(1.1.4) Wir wollen uns hier auf den Spezialfall beschränken, bei dem die Geschwindigkeit an jedem geometrischen Punkt \vec{x} nicht von der Zeit abhängt. Dann spricht man von einem *stationären Geschwindigkeitsfeld*. Dieses läßt sich durch die einfachere Abbildung $\vec{v} = (G, \vec{x} \mapsto \vec{v}(\vec{x}), V^3)$ beschreiben. Hier ist der Konfigurationsraum der Feldabbildung gleich dem Konfigurationsraum des Systems: Die Geschwindigkeit ein und derselben Person kann von Ort zu Ort wechseln, nur beim Passieren derselben Stelle haben alle Personen stets dieselbe Geschwindigkeit. D.h., dass bei der zweiten Veranschaulichung die angehefteten Geschwindigkeitsvektoren unveränderlich bleiben, der Film kann durch ein Foto ersetzt werden.

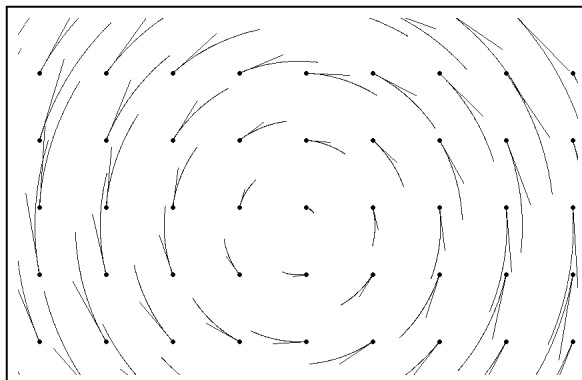
(1.1.5) Wir betrachten in der Regel nur glatte Abbildungen. Das heißt intuitiv, dass für ausreichend benachbarte Punkt \vec{x}_1 und \vec{x}_2 sich auch die zugehörigen Geschwindigkeiten nicht allzusehr unterscheiden dürfen. Die Geschwindigkeiten sollen nicht wild von Punkt zu Punkt springen. Unter diesen Umständen ist es möglich, **von endlich vielen Beobachtungen und Messungen mit einer gewissen Ungenauigkeit auf die Zwischenwerte zu schließen**. Wir werden diese intuitive Vorstellung der Glattheit im Laufe des Kapitels präzisieren und formalisieren, den Begriff entfalten. Aber Achtung: Es kommen durchaus Anwendungen vor, in denen nicht glatte Abbildungen auftauchen, unsere Annahmen nicht erfüllt sind.

Die Glattheitsannahme erlaubt die übliche Veranschaulichung vom Feldstandpunkt: Wir wählen im Konfigurationsraum ein Gitter aus und tragen an jeden Gitterpunkt den zugehörigen Geschwindigkeitsvektor geeignet ab. Für die zwischen den Gitterpunkten liegenden Punkte interpoliert dann die Anschauung den Feldwert näherungsweise. Dabei ist es meist sinnvoll, für die Feldvektoren einen anderen Maßstab als für die Ortsvektoren des Urbildraumes zu verwenden. Bei sehr großer Längenveränderlichkeit ist es vorteilhaft, alle Feldvektoren auf dieselbe Länge zu normieren, also zum Richtungsfeld überzugehen.

□ Konzipieren sie ein Computerprogramm, das für Abbildungen $\mathbb{R}^2 \rightarrow \mathbb{R}^2$ eine derartige Veranschaulichung produziert.

(1.17) **Wie hängt die Veranschaulichung durch Geschwindigkeitsvektoren mit der anderen durch Bahnstücke zusammen?** Wenn wir das Zeitintervall Δt ausreichend klein wählen, dann gilt in

Tangentenapproximation $\Delta\vec{x} = \vec{v}\Delta t$. Also ist $\vec{v}\Delta t$ aus der 2. Veranschaulichung näherungsweise ein kurzes Bahnstück $\Delta\vec{x}$ der ersten Veranschaulichung. Für ausreichend kleine ortsunabhängige Δt geht daher die zweite Darstellung in die erste über! Das nachfolgende Bild gibt hierzu ein Beispiel. Gewählt ist ein Gitter von Punkten. An jeden dieser Punkte ist einmal die momentane Geschwindigkeit als Strecke angeheftet und zum anderen ist die zugehörige Bahn ein Stück weit (für ein festes Zeitintervall) verfolgt. Die Bahn ist kreisförmig, so dass jeweils ein Stück eines Kreises gezeichnet ist. Die Geschwindigkeit nimmt von innen nach außen zu, wie deutlich zu erkennen ist.



(1.1.8) Wählen wir im stationären Fall für die erste Veranschaulichung Δt nicht klein, sondern groß, so dass wir die Bahnen der ausgewählten Punkte über den gesamten Beobachtungsbereich verfolgen können, dann erhalten wir eine andere wichtige Darstellung, nämlich die durch **Feldlinien oder Trajektorien**. Wir nehmen also eine typische Auswahl geeigneter Materiepunkte und verfolgen die Bahnen der zugehörigen materiellen Punkte, etwa indem wir uns vorstellen, dass jeder Punkt eine Farbspur hinterläßt. Von der Bahnkurve betrachten wir demnach das Bild, das im stationären Fall von der Wahl der Startzeit unabhängig ist. (Dieser Sachverhalt wird später exakt bewiesen!) Die wesentliche Eigenschaft der so entstehenden Feldlinien ist, dass in jedem Punkt einer Feldlinie die dort herrschende Geschwindigkeit tangential zur Feldlinie verläuft. Auch dies werden wir später präzisieren und genauer diskutieren. Die Feldlinien zeigen die Richtung der Feldvektoren, über ihre relative Stärke sagen sie zunächst noch nichts aus.

(1.1.9) Hat man ein inhaltlich zunächst anders interpretiertes Feld, etwa ein Kraftfeld, dann kann man sich immer vorstellen, dass der Feldwert gleich dem eines Geschwindigkeitsfeldes sei und die zugehörigen ganzheitlichen Veranschaulichungen entwickeln. Das liefert einen Einstieg in die Verhaltensanalyse. Im Rahmen des Programmes der analytischen Geometrie (Kap. 4.1.5) sind auch andersdimensionale Konfigurationsräume zulässig, ja wünschenswert. Wichtig ist dabei nur, dass $\dim V = \dim W$ sein sollte.

□ Es sei $f: \mathbb{R}^2 \rightarrow \mathbb{R}^3$. Welche Veranschaulichung läßt sich übernehmen, welche nicht?

6.1.1a Analytische Darstellungen der Felder

(1.1.10) Meist erhält man die Felder analytisch über einen Rechenausdruck. D.h. man geht von einer Formel für den Feldvektor aus nebst einer Rollenverteilung für die beteiligten Größen. Dabei wird der Ortsvektor \vec{x} als unabhängige Variable spezifiziert. Der Rechenausdruck baut daraus mit algebraischen Methoden den Feldvektor $\vec{v}(\vec{x})$ auf. Das ergibt die Zuordnung $\vec{x} \mapsto \vec{v}(\vec{x})$. In diesem Rechenausdruck können neben \vec{x} weitere Buchstaben mit der Rolle von Konstanten oder äußeren Parametern auftauchen. Einige Beispiele derartiger Konstruktionen:

$$\vec{x} \mapsto \vec{v}_C(\vec{x}) = \frac{a\vec{x}}{|\vec{x}|^3} \quad \vec{x} \mapsto \vec{v}_S(\vec{x}) = \vec{\omega} \times \vec{x} \quad \vec{x} \mapsto (\vec{a} \cdot \vec{x})\vec{x} + \vec{x}^2\vec{a}.$$

Hierbei sind a , $\vec{\omega}$ und \vec{a} äußere Parameter. Durch kleine Transformationen (Translationen, Rotationen ...Vgl. Kap 3.3.5d) sowie Superposition (Kap. 4.1.6) kann man aus diesen Feldern viele weitere bilden. Etwa $\vec{w}(\vec{x}) = \vec{\omega} \times \vec{x} + a \frac{\vec{x} - \vec{b}}{|\vec{x} - \vec{b}|^3}$. Wichtig: Beherrscht man das Verhalten der Ausgangsfelder, dann hat man meist auch Zugang zu dem der neu gebildeten!

Die übliche Feldveranschaulichung sieht so aus, dass man einen Ortsvektor \vec{x} wählt, dazu $\vec{v}(\vec{x})$ berechnet und diesen Pfeil am Ende von \vec{x} abträgt (und sich dann vorstellt, dass ein Geschwindigkeitsfeld vorliegt). Das ergibt die zweite unserer beiden Veranschaulichungen.

Ein wichtiges Problem besteht darin, aus dem Rechenausdruck möglichst viel über das zugehörige Feldverhalten zu erschließen. Eine Interpretation als Geschwindigkeitsfeld ist dazu hilfreich, da man sich den Prozeß einer Strömung als Ganzes gut vorstellen kann, selbst in Fällen, in denen es die entsprechende Strömung in der Realität sich nicht gibt. Überprüfen Sie das einmal am Beispiel der Feldlinienbilder aus Kap. 4.1.6.

Zwei Felder wollen wir nachfolgend als grundlegende Beispiele genauer besprechen. Genauer: Dies sind zwei Felder, deren Verhalten man sicher beherrschen sollte, von denen man vielfach ausgehen sollte, um sich neue Felder durch Vergleich zugänglich zu machen.

6.1.1b Die Rotationsbewegung eines starren Körpers

(1.1.11) Wir betrachten einen starren Körper. D.h. die materiellen Punkte sind starr miteinander verbunden, so dass alle skalaren Abstände und Winkel zwischen ihnen während jeder Bewegung erhalten bleiben. Die Punkte können sich nur als Ganzes gegenüber dem Konfigurationsraum V_0^3 bewegen. Bei einer Flüssigkeit dürfen sie sich dagegen auch gegeneinander bewegen.

Speziell wollen wir eine Drehung um eine feste Achse mit Richtungsvektor \vec{e} formalisieren. Wir legen den Ursprung auf die Achse. Dann bewegen sich alle Punkte auf Kreisbahnen senkrecht zur Drehachse mit Mittelpunkt auf der Drehachse. Genauer ist stets $\vec{x}(t) = z\vec{e} + \vec{r}(t)$. Hierbei steht $\vec{r}(t)$ senkrecht auf \vec{e} und schreibt sich nach Einführung einer geeigneten Basis in der Ebene senkrecht zu \vec{e} als

$$\vec{r}(t) = r(\vec{e}_1 \cos(\varphi(t)) + \vec{e}_2 \sin(\varphi(t))).$$

Dabei ist $t \mapsto \varphi(t)$ sei die zugehörige Polarwinkelfunktion. Kennt man sie, so kennt man auch die Bahnkurve. Die gegebene Formel erlaubt die Berechnung des jeweiligen Ortsvektors. Eine kleine weitergehende Besinnung ist angebracht: Soeben haben wir die Bahnkurven aller materiellen Punkte des Körpers vektoriell beschrieben. Zur Festlegung der Bahn eines Punktes benötigten wir das Tupel $(\vec{e}, r, z, \varphi(t))$. Was ändert sich bei diesen Beschreibungsgrößen, wenn man zur Bahn eines anderen materiellen Punktes des starren Körpers übergeht? \vec{e} bleibt fest, r und z können sich ändern und das neue und das alte φ werden sich höchstens durch eine additive Konstante unterscheiden: $\varphi_N(t) = \varphi_A(t) + \varphi_0$.

D.h., kennt man eine Bahnkurve und den Ort irgendeines weiteren Punktes für einen festen Zeitpunkt, etwa $t=0$, dann kann man auch dessen Bahnkurve angeben!

(1.1.12) Zurück zur Bahnkurve selbst. Differenzieren der Bahnkurve ergibt für die vektorielle Geschwindigkeit die Formel

$$\dot{\vec{r}}(t) = r\dot{\varphi}(t)(-\vec{e}_1 \sin(\varphi(t)) + \vec{e}_2 \cos(\varphi(t))).$$

Die Richtung ist immer tangential. Der Betrag wird durch $r\dot{\varphi}(t)$ gegeben. Beim Differenzieren fällt aber die eventuelle additive Konstante φ_0 in $\varphi_N(t)$ fort, so dass die Größe $\dot{\varphi}(t)$ **für alle Punkte des starren Körpers (auß erhalb der Drehachse) den gleichen Wert hat**. $\dot{\varphi}(t)$ ist die momentane Winkelgeschwindigkeit (der Bewegung). Die Einheit ist *Winkel pro Zeit*. Man bezeichnet sie üblicherweise mit $t \mapsto \omega(t) = \dot{\varphi}(t)$. (ω glesen "Omega").

(1.1.13) Es zeigt sich, dass es sinnvoll ist, die beiden Größen \vec{e} und $\omega(t)$, die beide die Drehbewegung unabhängig von der Wahl des Punktes charakterisieren, zu einem Vektor zusammenzufassen:

$t \mapsto \vec{\omega}(t) = \vec{e}\omega(t)$	$\omega(t) = \dot{\varphi}(t)$	Die momentane Winkelgeschwindigkeit der Drehbewegung
--	--------------------------------	--

Beachten Sie: Die Richtung des Vektors ist nach unseren Annahmen konstant, zeitlich variabel ist höchstens der Betrag. Wir sagen: Drehbewegung um eine **feste** Drehachse.

(1.1.14) Jetzt können wir Beschreibungsgrößen, die sich auf die Bewegung des gesamten Körpers beziehen,

von solchen für die individuellen Massenpunkte trennen:

	<i>Die Kinematik der Drehbewegung (mit fester Drehachse)</i>
Bahnkurven der Körperpunkte	$\vec{r}(t) = \vec{e}_z + r(\vec{e}_1 \cos(\varphi(t)) + \vec{e}_2 \sin(\varphi(t)))$
Die momentane Winkelgeschwindigkeit	$\vec{\omega}(t) = \vec{e}\dot{\varphi}(t)$ ist für alle Körperpunkte konstant
Zur Bahnfestlegung eines individuellen Punktes	benötigt man zusätzlich z,r und φ_0 .
Ableiten gibt die momentane Geschwindigkeit	$\vec{v}(t) = r\dot{\varphi}(t)(-\vec{e}_1 \sin(\varphi(t)) + \vec{e}_2 \cos(\varphi(t))).$

(1.1.15) Das eigentliche Objekt unseres Interesses ist das **Geschwindigkeitsfeld einer Drehbewegung**. Hierfür besitzen wir noch keine analytische Formel. $\vec{x} \mapsto \dot{\vec{x}}(t)$ sieht hingeschrieben aus wie eine Zuordnungsverfahren, ist aber keines! Gibt man den Punkt \vec{x} vor, so ist man mit dieser Information allein nicht in der Lage $\dot{\vec{x}}(t)$ auszurechnen! Die Geschwindigkeit in einem gegebenen Punkt kann noch beliebig sein.

Wie erhalten wir die Zuordnung analytisch? Offensichtlich ist das Skalarprodukt zwischen \vec{v} und \vec{x} zu jedem Zeitpunkt Null. D.h. Orts- und Geschwindigkeitsvektor stehen stets aufeinander senkrecht. Dies läßt vermuten, dass man die Geschwindigkeit auch über ein Vektorprodukt darstellen kann. Inspektion und Nachrechnen zeigen, dass $\dot{\vec{x}}(t) = \vec{\omega}(t) \times \vec{x}(t)$ gilt, was ein erstes Indiz für die Bedeutung des eingeführten Vektors $\vec{\omega}$ ergibt. (Beide Größen $\dot{\vec{x}}(t)$ und $\vec{x}(t)$ sind ja gegeben. Also kann man das Kreuzprodukt bilden und die Formel verifizieren!) Die gewählte Basis ist aus dieser Formel wieder herausgefallen. Sie hatte nur Katalysatorfunktion. Ebenso steht rechts keine Ableitung, nur der Bahnort selbst geht ein.

Damit haben wir aber unser gesuchtes Geschwindigkeitsfeld gefunden:

$(t, \vec{x}) \mapsto \vec{\omega}(t) \times \vec{x}$	=momentane vektorielle Geschwindigkeit, die der in \vec{x} befindliche materielle Punkt zur Zeit t besitzt!
---	---

Hier ist noch eine Zeitabhängigkeit zugelassen. Das Feld wird stationär, wenn wir $\omega(t) = \text{const.}$ verlangen. Dann folgt $\varphi(t) = \omega t + \varphi_0$ mit zeitlich konstantem ω . Eine Drehbewegung mit konstantem ω heißt auch gleichförmige Drehbewegung mit der (vektoriellen) Winkelgeschwindigkeit $\vec{\omega}$. Beachten Sie: Die Richtung \vec{e} von $\vec{\omega} = \vec{e}\omega$ gibt die Richtung der Drehachse an und der Betrag ω gibt an, welcher Winkel pro Zeiteinheit überstrichen wird (*Winkelgeschwindigkeit*). Jetzt - festes $\vec{\omega}$ - erhalten wir ein stationäres Geschwindigkeitsfeld, das es mit den eingangs beschriebenen Methoden zu veranschaulichen gilt:

$\vec{x} \mapsto \vec{v}_S(\vec{x}) = \vec{\omega} \times \vec{x}$	Geschwindigkeitsfeld einer stationären starren Drehbewegung.
--	--

Die Figur in (1.1.7) gehört zu einem derartigen Feld.

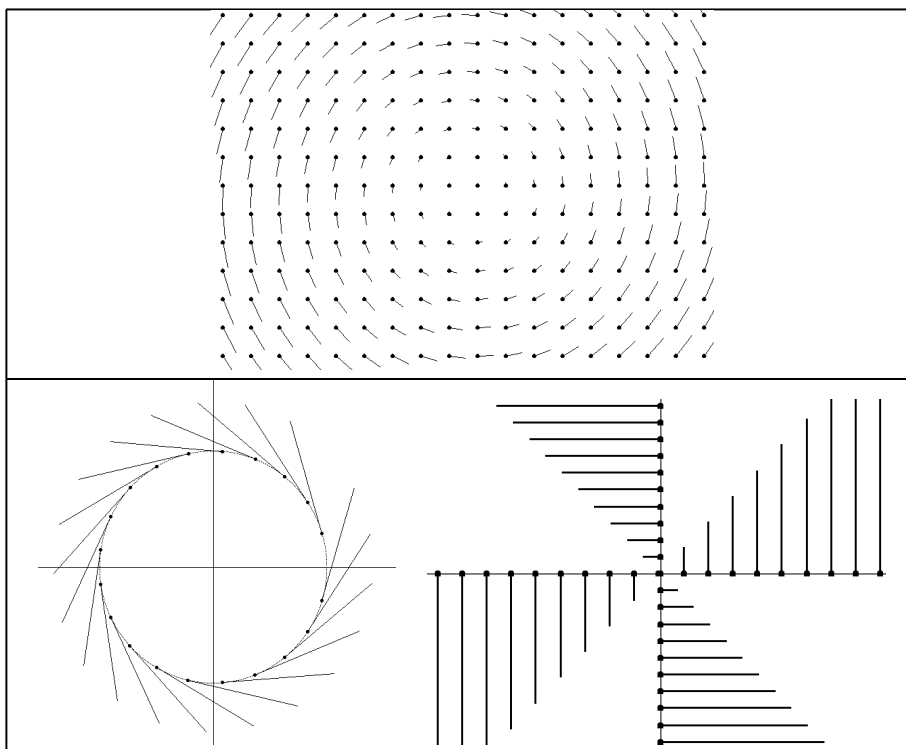
(1.1.16) Die Formel liefert nicht nur die vektorielle Geschwindigkeit für jeden Ort, sondern ergibt auch per Inspektion wichtige Eigenschaften der Feldkonfiguration. So genügt es, das Verhalten des Feldes in einer einzigen Ebene senkrecht zu $\vec{\omega}$ zu studieren. Alle anderen Werte folgen durch Parallelverschiebung der Konfiguration (in Richtung der Drehachse \vec{e}). Die nachfolgende kleine Rechnung belegt das:

$$\vec{v}_S(\vec{x} + \lambda \vec{e}) = \vec{\omega} \times (\vec{x} + \lambda \vec{e}) = \vec{\omega} \times \vec{x} = \vec{v}_S(\vec{x}) \quad \text{da} \quad \vec{\omega} \times \vec{e} = \vec{0}.$$

Die Feldlinien sind in diesem Fall Kreise.

(1.1.17) Nochmals drei Bilder zur Verdeutlichung des Feldverhaltens vom Feldstandpunkt aus. Gezeichnet

wird immer eine Ebene senkrecht zur Drehachse.



(1.1.18) Abschließend nehmen wir gezielt einige Änderungen an unserem Ausgangsfeld vor. Angenommen wir wählen $\vec{v}(\vec{x}) = f(|\vec{x}|)\vec{\omega} \times \vec{x}$. Dann bleibt die **Richtung** der Feldvektoren unverändert. Die Feldlinien sind immer noch Kreise, aber Punkte mit unterschiedlichem Ursprungsabstand können sich mit einer Geschwindigkeit bewegen, die nicht zu der eines starren Körper gehört. Zu Punkten mit unterschiedlichem Achsabstand gehören dann in der Regel unterschiedliche Umlaufzeiten! Stellen Sie sich als Beispiel etwa $f(r)=\sin(kr)$ vor.

(1.1.19) Physikalisch relevant ist der Fall des Gravitationsgesetzes. Betrachten wir Planeten, die sich auf Kreisbahnen mit unterschiedlichem Radius um eine Sonne bewegen. Das dritte Keplersche Gesetz ergibt hier $f(r)=\beta r^{-\frac{5}{3}}$. Die weiter entfernten Planeten haben ein längeres Jahr. Im Falle des starren Körpers wären alle Jahre gleich lang.

- Beispiel einer nicht stationären Drehbewegung: $\vec{\omega}(t) = \vec{e}_\alpha \sin(\beta t)$. Was für ein Bewegungsvorgang wird hierdurch beschrieben?
- Sei $\vec{\varphi} = \varphi \vec{n}$ und $R(\vec{\varphi})$ der lineare Operator, der eine Drehung um die durch \vec{n} bestimmte Achse mit dem Drehwinkel φ beschreibt. Dann gibt es für $\vec{x} \mapsto R_{\vec{n}}(\varphi).\vec{x}$ eine Formel, nämlich

$$R_{\vec{n}}(\varphi).\vec{x} = \vec{x} \cos \varphi + \vec{n}(\vec{n} \cdot \vec{x})(1 - \cos \varphi) + (\vec{n} \times \vec{x}) \sin \varphi.$$

Die Formel wurde in Kap. 4.3.11d bewiesen. Leiten Sie hieraus erneut $\vec{v} = \vec{\omega} \times \vec{x}$ ab.

6.1.1c Radiales Strömungsfeld

(1.1.20) Als zweites grundlegendes Beispiel betrachten wir eine Kugel vom Radius R , aus der radial und symmetrisch eine Flüssigkeit konstanter Dichte austritt. Die Austrittsmenge sei zeitlich konstant, so dass ein Dauerzustand fortströmender Flüssigkeit entsteht. Durch jede außen liegende Kugeloberfläche strömt pro Zeiteinheit dieselbe Gesamtmenge an Flüssigkeit. Da mit zunehmender Entfernung von der Austrittskugel immer größere Raumbereiche gefüllt werden müssen, wird entsprechend die Geschwindigkeit immer weiter abnehmen! Uns interessiert das Geschwindigkeitsfeld dieser Strömung. Aus Symmetriegründen verläuft die Richtung immer radial nach außen und der Betrag der Geschwindigkeit ist konstant für gleichen Abstand von der Kugel. Oder $\vec{v}(\vec{x})=f(r)\frac{\vec{x}}{r}$ mit $r=|\vec{x}|$. Gesucht ist die Funktion $f(r)=|\vec{v}(\vec{x})|$.

(1.1.21) Sei $r > R$. Dann strömt pro Zeiteinheit durch die Kugeloberfläche vom Radius r gerade das Volumen $f(r) \cdot 4\pi \cdot r^2$ an Flüssigkeit. Und diese Menge muß gleich dem entsprechenden Auströmungsvolumen $f(R) \cdot 4\pi \cdot R^2$ sein. Es folgt $f(r) = \frac{f(R)R^2}{r^2} = \frac{\alpha}{r^2}$. Dabei ist α eine Systemkonstante, die durch die Austrittsmenge pro Zeiteinheit festgelegt ist.

(1.1.22) Unser gesuchtes Strömungsfeld ist folglich:

$$\vec{x} \mapsto \vec{v}(\vec{x}) = \frac{\alpha}{r^2} \frac{\vec{x}}{r} \quad \text{für } r = |\vec{x}| > R$$

(1.1.23) Die Feldlinien sind in diesem Fall die vom Ursprung ausgehenden Halbgeraden für $r > R$. Die Bewegung auf ihnen wird mit zunehmendem Abstand von der Kugel immer langsamer und zwar so, dass die Dichte überall konstant bleibt. Die Bilder verdeutlichen das. Den Feldlinienbildern kann man die Verlangsamung nur indirekt entnehmen: In Form der Ausdünnung der Linien.

(1.1.24) Was benötigt man in diesem Fall, um das gesamte Feld anschaulich zu beherrschen? Nur die Feldkonfiguration entlang einer einzigen Halbgeraden! Denn durch Drehung kann man diese Konfiguration einfach in die Feldkonfiguration einer beliebigen anderen Halbgeraden überführen.

Wenn man α fest wählt, hängt das Feld nicht mehr vom Radius R der stromerzeugenden Kugel ab. Man kann diesen daher idealisierend gegen Null gehen lassen und erhält so schließlich ein Feld, das für alle $\vec{x} \neq \vec{0}$ erklärt ist.

(1.1.25) Betrachten wir zum Vergleich das Feld $\vec{v}(\vec{x}) = \beta \vec{x}$, das wir auch wieder als Auströmung interpretieren wollen. Aber da jetzt die Geschwindigkeit zunimmt, muß die Dichte abnehmen, die Materie muß sich ausdünnen. Aus Symmetriegründen hängt die zugehörige Dichte ρ nur von r ab, so dass die Stoffmenge $\rho(r)v(r) \cdot 4\pi r^2$ pro Zeiteinheit durch die Kugeloberfläche vom Radius r strömt. Jetzt ist $v(r) = \beta r$, womit die Bilanz $\rho(R)\beta R \cdot 4\pi R^2 = \rho(r)\beta r \cdot 4\pi r^2$. Also $\rho(r) = \frac{\rho(R)R^3}{r^3} = \frac{\gamma}{r^3}$. Die Dichte muß hier also wie $1/r^3$ abnehmen. Die Feldlinien sind erneut Halbgeraden. Die Vergrößerung der Feldstärke mit wachsendem Ursprungsabstand kann man durch Vergrößerung der Linienzahl andeuten. Wir werden später hierfür ein quantitatives Maß entwickeln. Offensichtlich erzwingt die Forderung nach Dichtekonstanz (Inkompressibilität) der ausströmenden Substanz gerade das $1/r^2$ -Gesetz.

- Wie muß man das Modell ändern, wenn man einerseits das Feld als Geschwindigkeitsfeld interpretieren will, aber andererseits auf konstanter Dichte besteht?
- Versuchen sie das Modell auf andere Dimensionen zu verallgemeinern. Was benötigt man hierzu noch?

6.1.2 Inspektion: Die Typisierung der Abbildungen anhand der beteiligten Dimensionen

(1.2.1) Ein einfacher, aber trotzdem zentraler Schritt bei der Inspektion einer zu verstehenden Abbildung $V \rightarrow W$ ist die Klassifikation nach den beiden beteiligten Dimensionen $n = \dim V$ und $m = \dim W$. Die nachfolgende Tabelle zeigt die Vielfalt der dabei entstehenden Abbildungstypen. Einordnung in dieses Matrixschema ist zur ersten Orientierung immer nützlich, da es einem hilft, die Abbildungen in ihren häufig doch sehr unterschiedlichen Verkleidungen wiederzuerkennen und ihnen auch zugleich erste mathematische Eigenschaften unmittelbar zuzuordnen.

$\downarrow n \backslash$	$m=1$	2	3	4...
1	reelle Funktionen $x \mapsto x^2$ F, G	$t \mapsto (t, t^2)$ G, P	$t \mapsto (t, t^2, \sin t)$ P	Bahnk... Kurven
2	$\begin{pmatrix} x \\ y \end{pmatrix} \mapsto x^2 + y^2$ Skalarfeld F	$\begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} x+y \\ xy \end{pmatrix}$ Transf. /krumml.K. F, T	$\begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} x+2y \\ xy \\ x-1 \end{pmatrix}$ Flächenparam. F, T, P	Flächen
3	$\begin{pmatrix} x \\ y \\ z \end{pmatrix} \mapsto x^2 + y^2 - z$ Skalarfeld F	$\begin{pmatrix} x \\ y \\ z \end{pmatrix} \mapsto \begin{pmatrix} x \\ y \end{pmatrix}$ Projektion F, T	$\begin{pmatrix} x \\ y \\ z \end{pmatrix} \mapsto \begin{pmatrix} x+y \\ xy \\ x-y \end{pmatrix}$ Feld/Trans. F, T	...
4...	Skalarfelder

Jeder Typ wird durch eine typische zugehörige Zuordnung illustriert.

(1.2.2) Diese unendliche Matrix ist auch im Rahmen unseres Programms der analytischen Geometrie zu sehen: So gibt die erste Zeile die Ausdehnung des Kurven- und Bahnkurvenbegriffs auf höhere Dimensionen. Die erste Spalte die des Skalarfeldes. In der Diagonalen finden wir Vektorfelder ebenso wie Transformationsabbildungen. Unterhalb der Diagonalen liegen die Projektionen, oberhalb die Injektionen. Ganz oben in der Ecke als Keimzelle alles Übrigen findet man schließlich die gewöhnlichen reellen Funktionen. Die Einordnung in dieses Schema gibt Hinweise, in welcher Richtung man jeweils nach Analogieschlüssen suchen sollte. Das Stichwort Projektion etwa weist darauf hin, dass Eigenschaften zu erwarten sind, die analog oder ähnlich zu denen vertrauter geometrischer Projektionen sind wie

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} \mapsto \begin{pmatrix} x \\ y \end{pmatrix}.$$

(1.2.3) Unten sind für jeden Fall die bevorzugten Veranschaulichungsstandpunkte eingetragen, wie sie in Kapitel 1.2.7 eingeführt worden sind. **F** für *Feldstandpunkt*, usw. In einigen Fällen sind mehrere Standpunkte sinnvoll, wobei einige weniger bedeutsam sind. Dies ist durch kleinere Buchstaben kenntlich gemacht. Beachten Sie insbesondere die drei Fälle, in denen der Graphenstandpunkt der Anschauung zugänglich ist. (Was besagt der Graphenstandpunkt speziell für die ebenen Skalarfelder?) Generell läßt sich sagen, dass vorzugsweise der Raum mit der höheren Dimension die Konfigurationsraumrolle erhält. Das schränkt in den meisten Fällen den Standpunkt stark ein.

(1.2.4) An veranschaulichenden Prozessen haben wir die Punktbewegungen für die erste Zeile und die Strömungsfeldinterpretation für die Diagonale. Für die Skalarfelder fehlt eine Prozeßinterpretation. Statt dessen besteht der Zugang über die Niveaumengenpartition. (Vgl. Vorkurs und Kap. 1 (3.3.7))

6.1.3 Inspektion: Allgemeinen Hilfsmittel

(1.3.1) Wie kann man sich das Verhalten von Abbildungen $\mathbb{R}^n \rightarrow \mathbb{R}^m$ oder allgemeiner $V \rightarrow W$ veranschaulichen? Wie erkennt man zugehörige wichtige Eigenschaften? Neben der Methode, die Abbildung als Darstellung eines gut vorstellbaren Prozesses - etwa einer Strömung - zu interpretieren, gibt es weitere nützliche Verfahren. Sie entstehen als naheliegende Verallgemeinerung der Methoden, die für das Kurvendiskussionsproblem für reelle Funktionen $\mathbb{R} \rightarrow \mathbb{R}$ entwickelt wurden. Die Anknüpfung an Vertrautes erleichtert den Lernprozeß für die neue Situation beträchtlich.

Z.T. kommen neuartige Elemente hinzu sowie andersartige Betonungen: Insbesondere gewinnen Restriktionen und Symmetriebetrachtungen eine viel größere Bedeutung als im eindimensionalen Bereich.

(1.3.2) Der typische durch eine Anwendungssituation gegebene Einstieg erfolgt über eine **Formel mit zugehöriger Rollenzuweisung**. Dabei kann mehr als eine unabhängige Veränderliche auftreten. Hieraus ergibt sich zunächst eine Zuordnung $\vec{x} \mapsto \vec{F}(\vec{x})$ der in diesem Kapitel betrachteten Art. Situation und Rechenausdruck zusammen bestimmen den Definitionsbereich oder auch eine ganze Klasse zulässiger Bereiche, so dass schließlich eine Abbildung des zu behandelnden Typs entsteht. Die Zuordnung kann koordinatenfrei geometrisch oder in Koordinaten zahlorientiert dargestellt sein. Die Natur stellt zunächst einmal in der Regel etwas Koordinatenfreies zur Verfügung, das zur genaueren Beherrschung jedoch quantifiziert, also in Koordinaten dargestellt werden muß. Die Freiheiten bei der Wahl der Quantifizierung werden zur Erleichterung der rechnerischen Probleme genutzt.

Nun sollte man versuchen, durch Inspektion und Bearbeitung des Rechenausdrucks möglichst viel über das Verhalten der Abbildung herauszubekommen und die gefundenen Einzelheiten zu einem anschaulichen Gesamtbild zusammenzusetzen. Schließlich sollte man in der Lage sein, dieses Gesamtbild verständlich und prägnant zu beschreiben.

(1.3.3) Die einfachste Verhaltensdiskussion verläuft für Funktionen $V \rightarrow W$, insbesondere $\mathbb{R}^n \rightarrow \mathbb{R}^m$, weitgehend **analog zur üblichen Kurvendiskussion!** Die hierzu im Vorkurs eingeführten Techniken sind entsprechend zu verallgemeinern und durch einige Punkte zu ergänzen. Ein großer Teil dieser Überlegungen kann bereits vor Einführung des Ableitungsformalismus besprochen werden. Ein wichtiger Unterschied ergibt sich daraus, dass man zur Veranschaulichung **meist nicht den Graphenstandpunkt einnehmen kann**. Vielmehr muß man einen anderen Standpunkt wählen, vielfach den Feldstandpunkt. Hinzu kommt die beständige Notwendigkeit, **den zu diskutierenden Urbildbereich durch Restriktionsbildung geeignet zu verkleinern, um eine anschauliche Darstellung überhaupt zu ermöglichen**.

(1.3.4) Sofern möglich, sollte man mit der Inspektion des Rechenausdrucks beginnen. Dabei geht man im Prinzip wie bei den reellen Funktionen vor. Gehen wir die zugehörigen Stichworte nacheinander durch:

- (1.3.5) **Bau des Rechenausdruckes.** Neu ist, dass in den Termen Vektoren vorkommen. Ein nützliches Hilfsmittel bilden zugehörige Verlaufsdiagramme (Automatenstandpunkt, Kap.1(1.2.7)). Dabei treten neue Bauteile auf, die den Übergang zwischen Vektoren sowie den zwischen Vektoren und Zahlen regeln. Auf diese müssen wir unser Augenmerk richten. Im Falle von Vektor \rightarrow Zahl haben wir $\vec{x} \mapsto |\vec{x}|$ und $\vec{x} \mapsto (\vec{a} \cdot \vec{x})$ fundamentale Bauteile, deren Verhalten wir beherrschen sollten (Vorkurs). Im Falle Vektor \rightarrow Vektor sind es die Felder $\vec{x} \mapsto \vec{a} \times \vec{x}$ und $\vec{x} \mapsto s(\vec{x})\vec{x}$, wobei s irgendein Skalarfeld sein darf. Wichtig ist auch die Translation $\vec{x} \mapsto \vec{x} - \vec{a}$ als Spezialfall einer kleinen Transformation.
- (1.3.6) **Absolute Form und Koordinatenform.** Die Abbildungen können in absoluter koordinatenfreier Form oder aber in Koordinatendarstellung auftreten.

$$s(\vec{x}) = |\vec{x}| \quad \text{bzw} \quad .s^K(x, y, z) = x^2 + y^2 + z^2.$$

Dabei sind nicht nur die üblichen Basisdarstellungen der linearen Algebra (lineare Koordinaten), sondern auch krummlinige Koordinaten wie Polarkoordinaten zulässig. Wir werden in 6.1.7 einige Arbeit investieren, um diese beiden Formen ineinander umzuwandeln. Das ergibt eine Verallgemeinerung der Koordinatendarstellungstheorie der linearen Algebra aus Kap. 4.4.

Bei Funktionen $\mathbb{R} \rightarrow \mathbb{R}$ bildet die Darstellung des Graphen mit Hilfe logarithmischer Skalen ein Vorläufer dieses Problemkreises krummliniger Koordinaten..

- (1.3.7) **Umformungen des Rechenausdrucks** erweisen sich erneut als wichtiges Hilfsmittel, um bestimmte Eigenschaften leichter per Inspektion zu erkennen. Charakteristisch für die Vektorrechnung ist die Möglichkeit, Vektoren als Linearkombination darzustellen und die zugehörigen Komponenten herauszupräparieren. Unter dem Stichwort Basisfelder im Werteraum werden wir hierauf eingehen. Bei Vektorfeldern sollte man auch an die einfache, aber immer wieder wichtige **Aufspaltung der Feldwerte in richtungsbestimmenden Einheitsvektor und Betrag** zu denken.
- (1.3.8) **Dominanzargumente.** In geeigneten Bereichen der Urbildmenge versucht man den Rechenausdruck so zu vereinfachen, dass die Werte dort nur unwesentlich verändert werden, dass das vom vereinfachten Ausdruck erzeugte Verhalten dagegen vertraut oder zugänglich ist. Typische Kandidaten derartiger Bereiche sind die Umgebung des Ursprungs, aber auch Bereiche mit sehr großem

$|\vec{x}|$, eventuell im Anschluss an kleine Transformationen.

In den Interferenzbereichen der Terme dagegen dominiert kein Teil und man muss versuchen, zu interpolieren und alternative Methoden der Verhaltensklärung heranzuziehen. Bei der Superpositionen von Coulombfeldern beispielsweise verhält sich das resultierende Feld in großen Entfernungen näherungsweise so, als wäre die Gesamtladung im Ursprung verteilt. Und das Verhalten dieser Zuordnung ist vertraut. Oder nehmen wir das Skalarfeld $(x,y) \mapsto x+y + \frac{\alpha}{x^2+y^2}$. In der Nähe des Ursprungs dominiert der dritte Term, entfernt man sich dagegen ausreichend weit vom Ursprung, dann dominieren die ersten beiden Terme.

- (1.3.9) **Suche nach speziellen Punkten, maximaler Definitionsbereich.** Für welche Punkte $\vec{x} \in G$ ist der Rechenausdruck nicht definiert? Wo gibt es Nullstellen? Nullstellen sind besonders für Vektorfelder wichtig. Interpretieren wir des Feld als Kraftfeld: Dann handelt es sich um Punkte, in denen die Kraft den Wert Null hat. Bemerkenswerterweise sind das in vielen wichtigen Fällen wieder **isolierte** Punkte, nicht etwa Mengen mit inneren Freiheitsgraden. Hinzu kommen **Sattelpunkte** und singuläre Punkte, für die der Wert nicht sinnvoll erklärbar ist (singuläre Punkte wie $\vec{x}=\vec{0}$ bei $\vec{x} \mapsto \frac{\vec{x}}{|\vec{x}|}$).
- (1.3.10) **Restriktionsbildung.** Bei zu großem oder unübersehbarem Urbildraum wird man die Abbildung geeignet restringieren. Etwa, indem man bestimmte Koordinaten konstant setzt oder indem man auf besondere Figuren wie den Einheitskreis einschränkt oder gar, indem man die Werte für einzelne Punkte gezielt ausrechnet. Insbesondere wenn eine Koordinatenform der Abbildung vorliegt, sollte man das Nullsetzen einzelner Koordinaten ziemlich zu Beginn versuchen, da man so häufig ohne Aufwand viel Information zum Verhalten erhält.
- (1.3.11) **Einführung von Hilfsgrößen.** Das erweist sich als ein besonders wichtiger Weg, um Ausdrücke mit mehreren Veränderlichen in den Griff zu bekommen. Tritt im Rechenausdruck für das Feld beispielsweise mehrfach der Term $x^2+y^2+z^2$ auf, so liegt es nahe, ihn durch die Hilfsgröße r^2 zu ersetzen. Und das deutet wieder auf eine Polarparametrisierung hin. Also: Günstige Hilfsgrößen legen jeweils die Wahl bestimmter Parametrisierungen nahe und umgekehrt. Zusätzlich wird man untersuchen, was Konstantsetzen derartiger Hilfsgrößen bewirkt. Wir werden den zugehörigen Formalismus für die Hilfsgrößen in zwei größeren Abschnitten formalisieren, in Kap. 6.1.4 und in Kap. 6.1.6. Zwei Abschnitte, weil es angebracht ist, Urbildraum und Werteraum gesondert zu behandeln. Die Resultate der beiden Abschnitte lösen zugleich das daran anschließend behandelte Problem des Überganges zwischen absoluter Form und Koordinatenform der Abbildungen.
- (1.3.12) **Niveaumengen.** Das sind Mengen konstanter Werte. Oder formal die Werte der Potenzmengenenerweiterung $\overset{-1}{\underline{f}}$. Besonders für Skalarfelder ($m=1$) bilden diese Niveaumengen wichtige Hilfsmittel zur Verhaltensfassung. Man kann versuchen, Niveaumengen (also "Flächen konstanter Temperatur" usw.) durch die Einführung geeigneter Hilfsgrößen (im Urbildraum) besonders gut in den Griff zu bekommen. Vgl. Kap.1 (3.3.7).
- (1.3.13) **Symmetriebetrachtungen.** wie bereits angedeutet, erweisen sich Symmetriebetrachtungen im mehrdimensionalen Fall als besonders wichtig, da es mit ihrer Hilfe vielfach gelingt, den zu diskutierenden Urbildbereich auf ein zugängliches Maß zu verkleinern, eventuell sogar das Problem der Verhaltensdiskussion einer Abbildung auf den Fall der einer Funktion $\mathbb{R} \rightarrow \mathbb{R}$ zurückzuführen.

Die Idee ist einfach: Man schränkt das Feld auf einen geeigneten, möglichst kleinen Teilbereich (des Urbildraumes) ein, derart, dass man den Rest vermittels einer geeigneten Gruppenoperation erhält. Diese ist vielfach vom Typ einer der Anschauung unmittelbar zugänglichen kleinen Transformationen.

Ein gutes Beispiel bietet das Feld $\vec{x} \mapsto \vec{x} \times (\vec{a} \times \vec{x})$.

Im Fall der in 6.1.1b besprochenen Drehbewegung des starren Körpers benötigt man nur einen einzigen Feldwert außerhalb der Drehachse. Der Rest folgt durch Skalentransformation, Drehung und Translation!

Einfache Beispiele für Symmetrieüberlegungen lassen sich ohne größeren Aufwand problemlos direkt diskutieren. Ein allgemeineres Verständnis verlangt jedoch einige mathematische Vorbedingungen: Den oben genannten Formalismus zur Einführung von Hilfsgrößen und den Formalismus der Gruppenoperationen. Dafür erhält man mit den Symmetriebetrachtungen aber auch ein sehr wichtiges Handwerkszeug. Enthalten ist hierbei die Verallgemeinerung des nützlichen Konzeptes der kleinen Transformationen, das wir im Vorkurs für gewöhnliche Funktionen eingeführt und in Kap. 3 sehr verallgemeinert haben. Als besonders wichtig erweisen sich die Translationen $\vec{x} = x \mapsto \vec{x} - \vec{a}$ des Urbildbereiches und die Skalentransformationen $\vec{x} \mapsto \frac{1}{h}\vec{x}$ des Urbild- und $\vec{y} \mapsto h\vec{y}$ des Wertebereiches. Wir gehen darauf genauer in Kap. 6.1.8 ein.

(1.3.14) Zwei **Beispiele**:

$$\left(V_0^3 - \{\vec{0}\}, \vec{x} \mapsto \vec{a} + \frac{\alpha \vec{x}}{|\vec{x}|^3}, V^3 \right) \quad \alpha, \vec{a} \text{ äußere Parameter ungleich Null.}$$

Vom Typ liegt ein Vektorfeld vor (m=n=3). Eine einfache Umformung des Rechenusdrucks bietet sich an:

$$\vec{a} + \frac{\alpha}{|\vec{x}|^2} \frac{\vec{x}}{|\vec{x}|}.$$

In der Nähe des Ursprungs dominiert der zweite Summand, weg vom Ursprung der erste. Beide Einzelsummanden haben direkt angebbares Verhalten. Im Interferenzbereich kann es eine Nullstelle \vec{x}_0 des Feldes geben. Nullsetzen des Feldwertes gibt die Gleichung: $\vec{a} + \alpha \frac{\vec{x}_0}{|\vec{x}_0|^3} = \vec{0}$. Betrag und Richtung müssen getrennt übereinstimmen.

Wir setzen $\vec{a} = a\vec{e}$ mit $|\vec{e}|=1$. Die Richtungsbedingung gibt $\vec{x}_0 = \varepsilon |\vec{x}_0| \vec{e}$ mit $\varepsilon = +1$ wenn $\alpha < 0$ und $\varepsilon = -1$ wenn $\alpha > 0$. Nun folgt für den Betrag:

$$|\vec{a}| - |\alpha| \frac{|\vec{x}_0|}{|\vec{x}_0|^3} = 0 \quad \text{also} \quad |\vec{x}_0| = \sqrt{\frac{|\alpha|}{|\vec{a}|}}.$$

Endergebnis: **Das Feld hat eine Nullstelle bei**

$$\vec{x}_0 = \varepsilon \frac{\vec{a}}{|\vec{a}|} \sqrt{\frac{|\alpha|}{|\vec{a}|}}.$$

Für \vec{x} -Werte in der Nähe von \vec{x}_0 sind auf jeden Fall beide Summanden wichtig. (Interferenzbereich.)

Wie weit können wir das Feld restringieren? Das Feld ist symmetrisch bezüglich der durch \vec{a} erzeugten Geraden durch den Ursprung. Ist E eine Ebene, die diese Gerade enthält, und kennt man die Feldkonfiguration in dieser Ebene, dann erhält man die gesamte Konfiguration einfach durch Drehung. Obige Nullstelle bleibt bei dieser Drehung invariant. Legt man ein Koordinatensystem so, dass $\vec{a}^K = (0,0,a)$ gilt, dann erhält man die folgende Koordinatendarstellung des Feldes:

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} \mapsto \begin{pmatrix} 0 \\ 0 \\ a \end{pmatrix} + \frac{\alpha}{r^3} \begin{pmatrix} x \\ y \\ z \end{pmatrix} \quad \text{mit } r = \sqrt{x^2 + y^2 + z^2}.$$

x=0 ergibt dann eine naheliegende Restriktion. Ausklammern von a zeigt schließlich, dass die gesamte Konfiguration nichttrivial nur von dem einen äußeren Parameter $q = \frac{\alpha}{|\vec{a}|}$ abhängt. Die Nullstelle liegt dann immer bei $\vec{x}_0^K = \frac{\varepsilon}{\sqrt{q}}(0,0,1)$.

□ Zeichnen Sie diese Konfiguration für einige q-Werte mit Hilfe eines Computeralgebraprogrammes.

(1.3.15) Als **zweites Beispiel** betrachten wir das Feld

$$\vec{x} \mapsto \vec{F}(\vec{x}) = \vec{x}^2 \cdot \vec{x} + \alpha \frac{\vec{x}}{\vec{x}^2}.$$

Wir haben Dominanz mit vertrautem Verhalten bei $\vec{0}$ und weit weg von $\vec{0}$. Setzt man $\vec{x} = r\vec{e}$ mit $r = |\vec{x}|$, so folgt $\vec{F}(\vec{x}) = \vec{e}(r^3 + \frac{\alpha}{r})$. Der Betrag ist somit auf Kugelschalen um den Ursprung konstant! Für $\alpha < 0$

verschwindet der Betrag bei $r_0 = |\alpha|^{\frac{1}{4}}$, also auf einer ganzen Kugelschale! Die Richtung des Feldes ist stets radial. Das bedeutet: Kennt man das Feld entlang der positiven 1-Achse, dann erhält man das Feld in allen anderen Richtungen einfach durch Drehung der x-Achse mitsamt Feldkonfiguration in die neue Richtung.

$$\boxed{\text{Ist } \vec{x} = \vec{e}_1 x \text{ und R Drehung mit } \vec{e} = R(\vec{e}_1), \text{ dann ist } \vec{F}(\vec{x}) = R(\vec{F}(\vec{e}_1 x)).}$$

Somit genügt die Kenntnis der Feldkonfiguration entlang der 1-Achse!

- Zeichnen Sie die Konfiguration für einige α -Werte.

6.1.4 Parametrisierungen

(1.4.1) Parametrisierungen der Urbildmenge erweisen sich als wichtiges Hilfsmittel der Verhaltensveranschaulichung von Abbildungen $V \rightarrow W$, sobald $\dim V > 1$ ist. Dieses Hilfsmittel ist neu in dem Sinne, dass es für Funktionen $\mathbb{R} \rightarrow \mathbb{R}$ kaum benutzt und benötigt wird. Wir gehen daher hier etwas genauer auf das Thema ein.

(1.4.2) Unter einer Parametrisierungsabbildung haben wir in Kap. 1.2.6e generell eine Abbildung einer Parametermenge in einen Konfigurationsraum verstanden. Hier benötigen wir einen spezielleren, an der Physik orientierten Typ von Parametrisierungsabbildungen.

(1.4.3) Unser Konfigurationsraum sei der n-dimensionale reelle Vektorraum V und die Parametermenge sei eine Teilmenge G_P des \mathbb{R}^k mit $k \leq n = \dim V$. Der Vektorraum V enthält die Urbildmenge der betrachteten Feldabbildungen.

(1.4.4) Was den Konfigurationsraum angeht, so ist unser zentrales Anliegen die Beschreibung und mathematische Modellierung geometrischer Objekte und physikalischer Phänomene in diesem Konfigurationsraum. An den Punkten dieses Raumes finden Ereignisse statt, können sich Objekte und Figuren befinden und in ihnen können physikalische Messungen vorgenommen werden. Bisher haben wir im Rahmen des Vektorraumformalismus zwei Leistungen stillschweigend identifiziert: Die **Benennung** der Punkte des Konfigurationsraumes und ihre **mathematische Darstellung**. Ein Ortsvektor stellte den zugehörigen Endpunkt mathematisch als Vektor (=gebundener geometrischer Pfeil) dar. Und wenn wir diesen Punkt irgendwie benennen, mit seiner Hilfe eine Rechnung ausführen wollten, dann geschah das vorzugsweise auch über diesen Ortsvektor.

(1.4.5) **Eine Parametrisierung trennt die beiden Aspekte.** Bei der Parametrisierung der Punkte einer Geraden im Raum erfolgt die Darstellung des Punktes immer noch über seinen Ortsvektor $\vec{x}_g(\alpha)$. Aber die Benennung erfolgt über den Parameterwert α . Dieser ist einfacher und als Zahl auch mathematisch besser zu handhaben.

Eine solche Trennung ist generell möglich und nützlich, gleichgültig, ob es um geometrische Objekte wie Flächen oder Kurven oder um physikalische Bahnen oder um Felder auf dem Konfigurationsraum geht.

(1.4.6) Besonders günstig ist es, wenn die Bezeichnung - der Name - zugleich inhaltlichen Aufschluß über den interessierenden Prozeß, die betrachtete geometrische Struktur usw. gibt.

Zur Verdeutlichung: Stellen Sie sich vor, Sie hätten es mit einer größeren Gruppe von Personen zu tun, wobei Herkunftsort und Beruf für Ihre Tätigkeit wichtig ist. Aber anstelle üblicher Namen wie "Peter Müller" hätte man Bezeichnungen wie "Busfahrer Wuppertal" oder "Bäcker Düsseldorf". Das wäre für die Arbeit unter den gewählten Aspekten sicher sehr nützlich. **Entsprechendes leisten die Parametrisierungen von V tatsächlich.**

(1.4.7) Als orientierendes Alltagsbeispiel kann man die Parametrisierung der Orte der Erdoberfläche durch Längen- und Breitengrade wählen. Die dabei benutzten Namen (=Zahlangaben) liefern unmittelbar inhaltliche Information über den Ort des Punktes.

(1.4.8) Bisher haben wir die **Trennung von Darstellung und Benennung** bei Bedarf immer auf eine einzige recht primitive Weise durchgeführt: Wir haben eine feste Basis gewählt und alle Objekte bezüglich dieser Basis zerlegt. Die entstehenden Koordinatentupel, die Koordinatenvektoren waren dann unsere Beschreibungsgrößen.

(1.4.9) Es zeigt sich jedoch, dass dieses Vorgehen für viele Problemsituationen unbequem und nicht optimal ist und wir erweitern unseren Formalismus jetzt gezielt.

Definition: Es sei V Vektorraum mit $\dim V = n$ und $G \subset V$ Konfigurationsraum.
 Weiter sei $k \leq n$ und $G_P \subset \mathbb{R}^k$ eine geeignete Teilmenge und $\pi : G_P \rightarrow G$.
 Dann heißt π eine *Parametrisierung* von $\text{Bild}(\pi) \subset G$.

(1.4.10) Bitte immer an die Zutat $\text{Bild}\pi$ denken! Diese Teilmenge (=Figur) von G wird parametrisiert und nicht immer ganz G . Das ist insbesondere der Fall, wenn $k < n$ ist. Ist $\text{Bild}\pi = G$, dann liegt eine Parametrisierung aller Punkte des Konfigurationsraumes vor und man spricht auch von allgemeinen, eventuell krummlinigen Koordinaten für die Punkte von G . Die Abbildung π muß keineswegs linear sein, so dass $\text{Bild}\pi$ auch kein Teilraum sein muß. Ebensovwenig ist injektiv verlangt! (Ein Punkt darf mehrere Namen haben!)

(1.4.11) Die Rollen sind klar: Das Tupel $\vec{\alpha} \in G_P \subset \mathbb{R}^k$ dient zur Bezeichnung der Elemente von $\text{Bild}\pi$. Dieses Parametertupel steht für zugehörige Formeln und Rechnungen zur Verfügung. Und $\pi(\vec{\alpha}) \in G \subset V$ ergibt eine korrekte vektorielle Darstellung dieses Konfigurationsraumpunktes, die mit Hilfe der Abbildung π jederzeit realisiert werden kann. Nochmals: π muß weder injektiv noch surjektiv sein. Einzelne Punkte können mehrere Namen haben oder auch gar keinen.

(1.4.12) Als Konkretisierung kann man an die Polarparametrisierung der komplexen Zahlen denken: $(\mathbb{R}_+ \times \mathbb{R}, (r, \alpha) \mapsto r e^{i\alpha}, \mathbb{C})$. Der Name ist das Tupel $\vec{\alpha} = (r, \alpha)$. Der Zahlwert der ersten Komponente gibt den Abstand des Punktes vom Ursprung, der der zweiten den Winkel zur positiven x-Achse, sofern dieser Winkel existiert. Wie versprochen enthalten die Namen unmittelbar inhaltliche Information!

(1.4.13) Wir führen jetzt eine **vorteilhafte Bezeichnungskonvention** ein. Es sei K ein Index, der die Parametrisierung π charakterisiert. Dann bezeichnen wir die Abbildung π von jetzt ab mit $\vec{x}^K : G_K \rightarrow V$. Die Elemente von G_K bezeichnen wir mit $\vec{\alpha}$, so dass $\vec{y} = \vec{x}^K(\vec{\alpha})$ immer ein Element von V , genauer von $\text{Bild}(\vec{x}^K)$ ist. Manchmal schreiben wir auch \vec{x}_K statt \vec{x}^K . Typischerweise beinhaltet $\vec{x}^P(\vec{\alpha}) = \vec{x}^P(r, \theta, \varphi)$ für $\vec{\alpha} = (r, \theta, \varphi)$ eine Polarkoordinatendarstellung eines Ortsvektors aus dem Raum V_0^3 . Das entspricht dem Brauch, eine Flugparabel mit \vec{r}_P oder eine Ebenenparametrisierung mit \vec{x}_E zu bezeichnen.

Im Falle von kartesischen Koordinaten mit Basis e haben wir den Koordinatenvektor $\vec{\alpha} = (x, y, z)$. Ist \vec{x} der geometrische Ortsvektor, so können wir $\vec{x} = \vec{x}^K(\vec{\alpha}) = \vec{x}^K(x, y, z) = \vec{e}_1 x + \vec{e}_2 y + \vec{e}_3 z$ schreiben. (Also: Unserer früheres \vec{x}^K ist jetzt $\vec{\alpha}$ und \vec{x}^K ist die zur Basis gehörige Linearkombinationsabbildung x_e)

(1.4.15) Eine Formel der Form $\vec{x}^K(\vec{\alpha}) = \dots$ liest sich wie folgt: der Ortsvektor in V , der zum Parametertupel $\vec{\alpha}$ gehört, wird mit $\vec{x}^K(\vec{\alpha})$ bezeichnet und ist durch konkret berechenbar. Parametrisierungen werden üblicherweise durch derartige Formeln vorgegeben. **D.h. auch, dass man sich gerade diese Formeln merken sollte.** Geht man von einer solchen Formel aus, kann man den gesamten Rest meist leicht rekonstruieren

(1.4.16) Das neben den kartesischen Parametrisierungen wichtigste Beispiel bildet die (nichtlineare) Polarparametrisierung. Wir geben die zugehörige Parametrisierungsabbildung und zwar zunächst für den Fall der Ebene. Dann folgen die beiden darauf aufbauenden wichtigsten räumlichen Parametrisierungen, die Zylinderkoordinaten und die Kugel- oder räumlichen Polarkoordinaten. Polarkoordinaten heißt immer "Polarkoordinaten relativ zu einem vorgegebenen kartesischen Koordinatensystem K !

6.1.4a Ebene Polarkoordinaten

(1.4.17) Die zugehörige Parametrisierungsabbildung sieht im Detail wie folgt aus:

$$\vec{x}^P = (G_P, \vec{r}^P = \begin{pmatrix} r \\ \theta \end{pmatrix}) \mapsto \vec{x}^P(r, \theta), V_0^2) \quad \text{mit } \vec{x}^P(r, \theta) = \vec{e}_1 r \cos \theta + \vec{e}_2 r \sin \theta$$

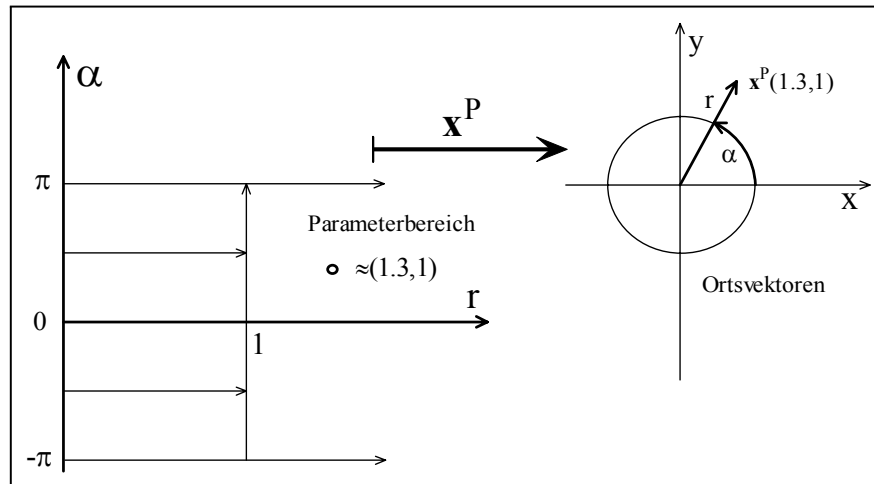
Oder als suggestive "Merkform" $x = r \cos \theta$ $y = r \sin \theta$

Oder vektoriell $\vec{x}^P(r, \theta) = r \vec{e}_r(\theta)$ mit $\vec{e}_r(\theta) = \vec{e}_1 \cos \theta + \vec{e}_2 \sin \theta$

Dabei ist (\vec{e}_1, \vec{e}_2) die kartesische Basis des vorgegebenen Koordinatensystems.

(1.4.18) Formal korrekt sollte man den Parametervektor wie in diesem Beispiel immer als Spaltenvektor interpretieren. Vielfach werden wir diesen als transponierte Zeile darstellen, also durch Schreibweisen wie

$t(x,y)$. Die mit $GP=G_P$ bezeichnete Urbildmenge der Polarkoordinatenabbildung wird fallspezifisch gewählt. Die Figur zeigt die wichtigste Wahl mit $r>0$ und $-\pi < \theta < \pi$.



(1.4.9) Die Veranschaulichung in der Figur erfolgt vom Transformationsstandpunkt (Kap. 1.2.7): Links befindet sich mit den Achsen r und θ der Parameterraum G_P . Dies ist ein Streifen, der durch $r=0$ und $\theta = -\pi$ berandet wird. Eingezeichnet ist noch die Strecke $s = \{(r, \theta) | r = 1, -\pi \leq \theta \leq \pi\}$ sowie ein Parameterpunkt $g = t(1.3, 1)$. Also: r sollte nie negativ werden und der Polarwinkel zwischen $-\pi$: und $+\pi$: liegen.

Rechts befindet sich mit den Koordinatenachsen x und y der zu parametrisierende V_0^2 mit der kartesischen Basis. Darin ist der Einheitskreis eingezeichnet als Bild der Strecke s sowie der Ortsvektor $\vec{x}^P(g)$ als Bild von g . Die Figur zeigt, wie der geometrische Pfeil aus den Parametern r und θ zu konstruieren ist und weshalb wir den Urbildraum G_P in der angegebenen Weise gewählt haben: Will man erreichen, dass die Darstellungsabbildung bijektiv wird, dann sollte man G_P wie beschrieben wählen und zwar ohne den Rand. Denn die gesamte Achse $r=0$ geht bei der Parametrisierung in den Nullpunkt über und die Halbgeraden $\theta = \pi$ und $\theta = -\pi$ werden beide auf die negative x -Achse abgebildet. G enthält (bei der bijektiven Wahl) den Nullpunkt und die negative x -Achse nicht, ist also nicht die gesamte Ebene. Man bezeichnet diese Menge gerne als *die entlang der negativen x -Achse aufgeschnittene Ebene*.

(1.4.21) Für den bijektiven Bereich kann man den Rechenausdruck für die **Umkehrabbildung** sofort angeben:

Die Umkehrabbildung $\begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} r(x,y) \\ \theta(x,y) \end{pmatrix}$ $r(x,y) = \sqrt{x^2 + y^2}$ $\theta(x,y) = \text{atn}(y,x)$	mit $\text{atn}(y,x) = \begin{cases} \text{atn}(\frac{y}{x}) & \text{für } x > 0 \\ \text{atn}(\frac{y}{x}) + \pi \text{sgn}(y) & \text{für } x < 0 \\ \text{atn}(y,0) = \pi \text{sign}(y). & \end{cases}$
--	---

Hierbei ist $\text{sgn}(y)$ das Vorzeichen der reellen Zahl y . Beachten Sie, dass $x < 0$ und $y = 0$ nicht vorkommt, so dass die Definition immer sinnvoll ist.

(1.4.22) Benötigt man bijektiv nicht, kann man den Urbildbereich vergrößern. Zunächst einmal für beliebige Winkel θ , und eventuell auch für negative r vermöge der Beziehung $\vec{x}^P(-r, \theta) = \vec{x}^P(r, \theta + \pi)$. In der Darstellung durch komplexe Zahlen ist das die Gleichung $-re^{i\theta} = re^{i(\theta + \pi)}$.

(1.4.23) Eine solche Polarpametrisierung der Ebene wird man immer in Betracht ziehen, wenn die anstehende Problemsituation irgendwelche radialen Symmetrien aufweist. Für Probleme mit rechtwinkliger Symmetrie ist sie weniger nützlich.

- Formulieren Sie versuchsweise einmal die übliche Vektoraddition des \mathbb{R}^2 in Polarkoordinaten.
- Sei $t \mapsto r(t)$ eine reelle Funktion. Restringieren Sie die Polarabbildung aus (1.4.17) wie folgt: $t \mapsto \vec{r}(t) = \vec{x}^P(r(t), t) = r(t)\vec{e}_r(t)$. Wie ist Bild geometrisch zu interpretieren? Wählen Sie speziell für $r(t)$ die folgenden 3 Konkretisierungen: $r_1(t) = t$, $r_2(t) = 2R \sin(t/2)$ und $r_3(t) = \frac{1}{1+t^2}$. Was für Figuren ergeben sich? Wie erhalten Sie derart beschriebene Figuren mit Hilfe eines Computeralgebraprogramms?
- Man kann kurvenartige Figuren der Ebene auf unterschiedliche Weisen darstellen. Wir nennen 4 wichtige Möglichkeiten:

1. Als Graph einer reellen Funktion, eventuell auch mehrerer.
2. Als Bild einer kartesischen Parametrisierung $t \mapsto \vec{e}_1 x(t) + \vec{e}_2 y(t)$. Oder
3. als Bild einer polaren Parametrisierung $t \mapsto r(t) \vec{e}_r(t)$ und
4. Als Lösungsmenge einer Gleichung zwischen den Koordinaten x und y .

Konkretisieren Sie die 4 Möglichkeiten für den Fall eine Ursprungskreises und einer einfachen Parabel.

- Was für eine Figur wird durch die folgende gestückelte Zuordnung beschrieben?

$$t \mapsto \vec{r}(t) \quad \text{mit} \quad \vec{r}(t) = \begin{cases} t\vec{e}_1 + t^2\vec{e}_2 & \text{für } -1 \leq t \leq 1 \\ \vec{e}_2 + (2-t)\vec{e}_1 & \text{für } 1 \leq t \leq 2 \\ \vec{e}_1 - (t-3)\vec{e}_2 & \text{für } 2 \leq t \leq 3 \end{cases}$$

(1.4.24) als nächstes geben wir die Parametrisierungsabbildungen der beiden wichtigsten krummlinigen **räumlichen** Koordinatensysteme an, die beide auf der Polarparametrisierung der Ebene aufbauen. Nochmals zur Erinnerung: Der darstellende Ortsvektor ist als Funktion der beschreibenden Parameter anzugeben.

6.1.4b Zylinderkoordinaten

(1.4.25) Diese Koordinaten erfassen den gesamten V_0^3 . Man gibt dazu ein kartesisches Koordinatensystem mit Basis e vor. Die von \vec{e}_3 erzeugte Gerade heißt Polachse. In der 1-2-Ebene wählt man ebene Polarkoordinaten (r, θ) . Weiter sei $h = x_3$. Dann bilden (r, θ, h) die Zylinderkoordinaten und ergeben die folgende zugehörige Ortsvektordarstellung:

$$\begin{aligned} {}^t(r, \theta, h) &\mapsto \vec{x}^Z(r, \theta, h) = r(\vec{e}_1 \cos \theta + \vec{e}_2 \sin \theta) + \vec{e}_3 h \\ \text{Injektiver Bereich: } &r > 0, 0 \leq \theta < 2\pi, h \text{ beliebig} \quad (\text{Die Polachse fehlt im Bild!}) \end{aligned}$$

- Was ist zu merken? Fertigen Sie selbst eine Skizze.

(1.4.26) Einige Restriktionen dieser Parametrisierung:

$$\begin{aligned} (\theta, h) &\mapsto \vec{x}^Z(R, \theta, h) && \text{Zylindermantel} \\ ((\theta, h) &\mapsto \vec{x}^Z(mh, \theta, h) && \text{Kegelmantel} \\ t &\mapsto (\theta(t), h(t)) = (t, \sin(2\pi nt)) \mapsto \vec{x}^Z(R, \theta(t), h(t)) \end{aligned}$$

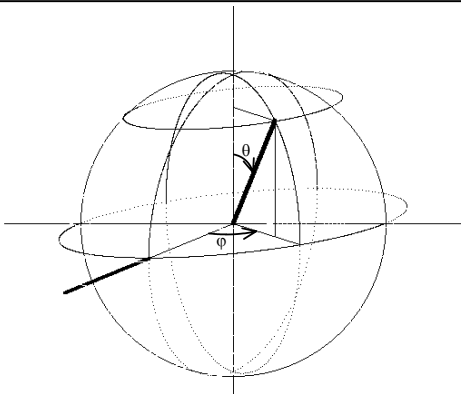
Die letzte Zuordnung ergibt für ganzes n eine geschlossene Kurve auf dem Zylindermantel!

- Was ergibt sich als Bild der letzten Kurve, wenn man n irrational wählt?

6.1.4c Kugelkoordinaten oder räumliche Polarkoordinaten

(1.4.27) Auch diese Koordinaten erfassen den gesamten V_0^3 . Man gibt wieder ein kartesisches Koordinatensystem mit Basis e vor. e_3 erzeugt eine Polachse. Der Winkel zwischen positiver Polachse und dem Ortsvektor ist der Breitenwinkel θ . Er läuft von 0 bis π . Die Länge des Ortsvektors sei r . Projiziert man den Ortsvektor in die 1-2-Ebene (= Äquatorebene), so hat die Projektion die Länge $r \cdot \sin(\theta)$. Diese Projektion wird in ebenen Polarkoordinaten dargestellt, was den Längenwinkel φ ergibt, der von 0 bis 2π läuft. Winkel zwischen der Äquatorprojektion von x und der \vec{e}_1 -Achse mit Umlaufssinn von \vec{e}_1 nach \vec{e}_2 . (1.4.28) Zusammen

erhält man die folgende Parametrisierung des Ortsvektors:

${}^t(r, \theta, \varphi) \mapsto \vec{x}^P(r, \theta, \varphi)$ $= r \sin \theta (\vec{e}_1 \cos \varphi + \vec{e}_2 \sin \varphi) + \vec{e}_3 \cdot r \cos \theta$ <p>Injektiver Bereich :</p> $r > 0, \quad 0 < \theta < \pi \quad 0 < \varphi < 2\pi.$	
--	--

□ Überlegen Sie sich, was zu merken ist.

(1.4.29) Und erneut einige Restriktionen der Parametrisierung, die interessante geometrische Figuren erfassen:

$$\begin{aligned}
 (\theta, \varphi) &\mapsto \vec{x}^P(R, \theta, \varphi) && \text{Kugeloberfläche} \\
 (r, \varphi) &\mapsto \vec{x}^P(r, \theta_0, \varphi) && \text{Kegelmantel} \\
 t &\mapsto \vec{x}^P(r, \theta, t) && \text{Bewegung auf Breitenkreis.} \\
 t &\mapsto \vec{x}^P(R, \sin t, \sin(nt)) && n \in \mathbb{N}. \text{ Geschl.Bew. auf der Kugel.}
 \end{aligned}$$

Die Formeln für die Umkehrabbildungen (der injektiven Bereiche) geben wir für beide räumliche Darstellungen nicht an. Man kann Sie sich relativ leicht herleiten. Diese Umkehrabbildungen haben typischerweise immer kompliziertere Berechnungsterme als die Parametrisierungen selbst.

(1.4.30) Als Abschluß der Beispiele sei nochmals daran erinnert, dass kartesische Koordinaten einen Spezialfall des allgemeinen Formalismus bilden. Die Parametrisierungsabbildung ist unsere frühere Linearkombinationsabbildung:

$${}^t(x_1, x_2, x_3) \mapsto \vec{x}^K(x_1, x_2, x_3) = \sum_i \vec{e}_i x_i.$$

□ Entwickeln Sie jetzt selbst eine Parametrisierung durch Toruskoordinaten. (Ein Torus ist ein Körper in Form eines Rettungsringes). Das ist recht einfach und ein guter Verständnistest für das Gesagte.

6.1.5 Das Umsetzen geometrischer Bedingungen mit Hilfe von Parametrisierungen

(1.5.1) Wie wandelt man geometrische Bedingungen und Einschränkungen für Figuren und Vorgänge in möglichst einfache, mathematisch gut handhabbare Formeln um? Eine Kurve soll nicht beliebig im Raume verlaufen sondern auf der Oberfläche einer bestimmten Figur, etwa eines Paraboloiden. Eine andere Kurve soll als Schnitt zweier gegebener Fläche entstehen. Eine Fläche soll einerseits eine bestimmte Form haben, sagen wir Ellipsoidform, und andererseits gewisse Bedingungen erfüllen, sagen wir einen vorgegebenen kürzesten Abstand zu einer zweiten Fläche besitzen. Usw.

All dieses läßt sich im Prinzip in kartesischen Koordinaten rechnen, aber die Formeln werden rasch kompliziert und schlecht handhabbar.

(1.5.2) Die Wahl einer günstigen Parametrisierung ergibt vielfach eine enorme Verbesserung der Handhabbarkeit. Dabei nutzt man den folgenden elementaren Sachverhalt: Hat man einmal eine Parametrisierung, dann erhält man durch Restriktion oder durch Vorschalten einer weiteren Parametrisierung erneut eine Parametrisierung und zwar von geometrischen Objekten, die gerade gewisse Zusatzbedingungen erfüllen. Formal sollte man dabei auf die beteiligten Dimensionen (=Zahl der inneren Freiheitsgrade) achten. Zu einer Parametrisierung $x^K:G_K \rightarrow V$ gehören ja immer zwei Dimensionen, nämlich die zu G_K gehörige Dimension k , die angibt, wieviele freie Komponenten der Parameter $\vec{\alpha}$ hat und die Dimension $n = \dim V$. Wir haben $n \geq k$. Ist $n = k$, so liegen typischerweise krummlinige Koordinaten vor. Einschränkende geometrische Bedingungen sind meist an eine Verkleinerung von k gekoppelt, d.h die Zahl der Parameter (=Freiheitsgrade) wird

reduziert und das läßt sich formal besonders einfach durch Restriktionen der Parametrisierung oder durch Vorschalten einer anderen Parametrisierung erreichen.

(1.5.3) Eine Flächenparametrisierung wie $\vec{\pi} = (\mathbb{R}^2, t(\alpha, \beta) \mapsto (\alpha, \beta, 2 + \alpha^2 + 3\beta^2, \mathbb{R}_K^3)$ ist ein typisches Beispiel für den Fall $k < n$. ($k=2, n=3$). Das zugehörige Bild stellt eine Parabolfläche im Raum dar. Die Koordinaten beschreiben inhaltlich die Projektion der Flächenpunkte auf die Horizontalebene. Schalten wir jetzt eine weitere Parametrisierung vor, sagen wir $\vec{\varphi} : t \mapsto (\alpha(t), \beta(t)) = (t \cos(\omega t), t \sin(\omega t))$, so erhalten wir eine Parametrisierung einer Kurve auf der Parabolfläche! Hier die einer zunehmende Spirale (für $t > 0$). Die mathematische Erfassung eines derartigen geometrischen Gebildes Kurve auf einer vorgegebenen Fläche ist ohne Zuhilfenahme einer Parametrisierung recht mühsam, mit aber weitgehend unproblematisch. Die resultierende Parametrisierung des Beispiels zeigt das:

$$\vec{\pi} \circ \vec{\varphi} = (]0, \infty[, t \mapsto (t \cos(\omega t), t \sin(\omega t), 2 + t^2(1 + \sin^2(\omega t))), \mathbb{R}_K^3)$$

(1.5.4) Ein anderes Beispiel: Durch Restriktion der Kugelkoordinaten auf $r=R$ erhalten wir eine Parametrisierung der Kugeloberfläche. Dabei wird $k=3$ zu $k=2$. Vorschalten der Abbildung $t \mapsto (R, \theta(t), h(t))$ vor die Zylinderkoordinaten liefert einen Bewegungsprozeß auf dem Zylindermantel. Oder alternativ die Parametrisierung einer Kurve auf dem Zylindermantel. Dabei wird $k=3$ zu $k=1$.

(1.5.6) Die Koordinaten der Parametrisierung sind also bei geeigneter Wahl nicht nur Namensgeber, sondern der Name gibt zugleich geeignete quantifizierte geometrische Systemeigenschaften wieder, die für den Formelbau benötigt werden. Damit lassen sich dann geometrische Bedingungen (etwa: "Ursprungsabstand soll R sein") unmittelbar in Formeln umwandeln ($r=R$). Die entstehenden Parametrisierungen können mathematisch weiterverarbeitet werden.

(1.5.7) Im Rahmen unseres Alltagsmodelles: Hat man Bezeichnungen, die Beruf und Herkunft geben, dann erhält man per Restriktion eine Namensgebung für die Berufe allein. Restringiert man im Personenbeispiel auf alle Düsseldorfer, so ergibt sich eine Berufsparametrisierung aller erfaß ten Düsseldorfer Personen.

(1.5.8) Gehen wir für den Anschauungsraum V_0^3 einmal die vier möglichen Werte für k durch, um konkret zu verdeutlichen, was alles an Typen erfaßt wird:

- Ist $G_K \subset \mathbb{R}^3$ ein Quader, dann ist (bei injektiver Parametrisierung) Bild $\vec{\pi}$ ein Körper im Raume und $\vec{\pi}$ gibt eine Parametrisierung seiner Punkte. Hier ist $k=3$. Die Parametrisierung bestimmt ein (vielfach krummliniges) Koordinatensystem auf K .

Beispiel: $(r, \theta, h) \mapsto \vec{x}^Z(r, \theta, h) = r(\vec{e}_1 \cos \theta + \vec{e}_2 \sin \theta) + \vec{e}_3 h$ parametrisiert den Ortsvektor durch Zylinderkoordinaten. Wird K durch die Ungleichungen $0 \leq h \leq H$ und $0 \leq r \leq R$ erzeugt, dann ist das Bild ein Zylinder.

- Ist $G_K \subset \mathbb{R}^2$ ein Rechteck, dann ist (bei injektiver Parametrisierung) Bild $\vec{\pi}$ eine Fläche im Raume und $\vec{\pi}$ gibt eine zugehörige Parametrisierung. Wir haben $k=2$. Natürlich kann $\vec{\pi}$ auch als Koordinatensystem auf dieser Fläche interpretiert werden.

Beispiel: Restringiert man die Zylinderkoordinaten auf das Rechteck $r=R$, so ergibt

$$(\theta, h) \mapsto \vec{x}^Z(R, \theta, h) = R(\vec{e}_1 \cos \theta + \vec{e}_2 \sin \theta) + \vec{e}_3 h$$

eine Parametrisierung des Zylindermantels zum Radius R .

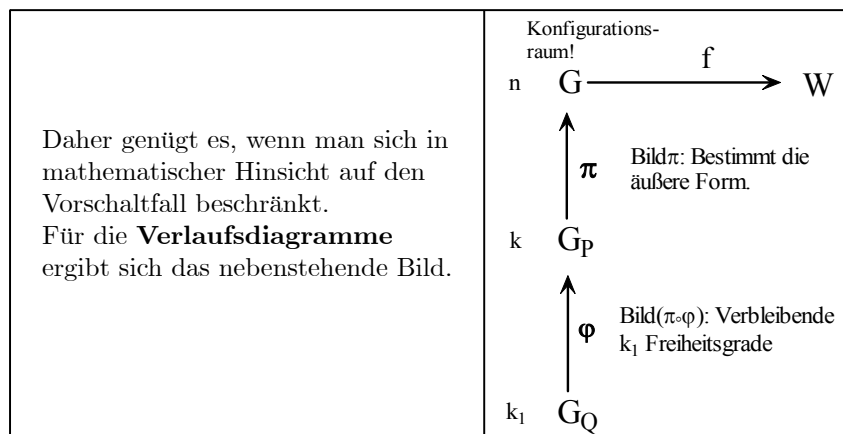
- Ist $G_k = I$ ein reelles Intervall, also $k=1$, dann ist Bild $\vec{\pi}$ ein geometrischer Kurvenzug oder Weg. Etwa eine Strecke oder ein Parabelstück usw., der durch $\vec{\pi}$ parametrisiert wird. (Kurve = Parametrisierung eines Weges).

Beispiel: $t \mapsto \vec{x}^Z(R, \theta(t), h(t))$ gibt die Bewegung eines Punktes auf dem Zylindermantel.

- $k=0$ schließlich parametrisiert einen einzigen Punkt. $\vec{a} = \vec{x}^P(2, \frac{\pi}{2}, \pi)$ etwa liefert über die räumlichen Polarkoordinaten einen ganz bestimmten Punkt in der Äquatorebene.

(1.5.9) **Fassen wir zusammen:** Mit Hilfe elementarer mengentheoretischer Operationen wie **Restriktion und Zusammensetzung**, die man mit **Parametrisierungsabbildungen** durchführt, lassen sich allgemeine geometrische Zusatzbedingungen günstig in Formeln umwandeln und **quantifizieren**.

(1.5.10) Anmerkung: Formal kann man eine Restriktion immer auch durch Vorschalten einer Injektionsabbildung realisieren.



- Sei $(x,y) \mapsto \vec{x}^P(x,y) = \vec{e}_1 x + \vec{e}_2 y + \vec{e}_3 \alpha(x^2 + y^2)$. Was für eine Figur wird durch Bild \vec{x}^P beschrieben? Was durch die Bilder der folgenden Abbildungen: $t \mapsto \vec{x}^P(\alpha \cos(\omega t), \alpha \sin(\omega t))$ und $t \mapsto \vec{x}^P(\alpha \cos(\omega t), \beta \sin(\omega t))$ und $t \mapsto \vec{x}^P(\alpha \sin(\omega t), \beta \sin(\omega t))$. Die äußeren Parameter und die Definitionsbereiche sollten Sie jeweils geeignet ergänzen. Wie sieht jeweils die der Parametrisierung vorgeschaltete Abbildung aus?

6.1.6 Quantifizierung des Werteraumes mit Hilfe von Basisfeldern

(1.6.1) Bisher haben wir den Konfigurationsraum $G \subset V$ parametrisiert, den Urbildraum unserer Abbildungen $f: V \rightarrow W$. Etwa die Vektoren der Ebene durch Polarkoordinaten. Vgl. die Figur in (1.5.10). Wie steht es aber mit dem Werteraum W bei Feldabbildungen $f: V \rightarrow W$? Wir wissen bereits aus der Theorie der linearen Gleichungen in Kap. 5, dass es auch für den Werteraum auf die Wahl einer geschickten Basis ankommen kann. Dieser Frage wenden wir uns jetzt zu.

(1.6.3) Die Idee sieht wie folgt aus: Angenommen man hat in der Ebene ein Vektorfeld zu einem Problem mit Kreissymmetrie. Dann gibt es zu jedem Punkt zwei ausgezeichnete Richtungen, die radiale und die tangentielle. Es liegt nahe, in jedem Punkt von V , eine kartesische Basis (\vec{e}_r, \vec{e}_t) mit diesen beiden Richtungen einzuführen und die Feldvektoren in die zugehörigen Komponenten zu zerlegen. Diese Komponenten wird man *radiale* und *tangentiale* Komponente nennen. Die Basen an verschiedenen Punkten von V gehen jedoch nicht einfach durch Parallelverschiebungen auseinander hervor, sie können von Punkt zu Punkt unterschiedliche Richtung haben. Daher haben wir es nicht wie bisher mit einer einzigen festen Basis zu tun, sondern mit einem (glatten) **Basisfeld** $\vec{x} \mapsto (\vec{e}_r(\vec{x}), \vec{e}_t(\vec{x}))$. Jedem Punkt wird eine eigene, zugehörige Basis zugeordnet. Vgl. die Figur in (1.6.15).

(1.6.4) Wir verallgemeinern und abstrahieren:

Es sei W Vektorraum und $G \subset W$ eine Teilmenge.
Ein **Basisfeld auf G** ist eine Abbildung $e = (G, \vec{y} \mapsto e(\vec{y}), \mathfrak{B})$, die jedem $\vec{y} \in G$ eine zugehörige Basis $\vec{e}(\vec{y})$ zuordnet. Dabei sei \mathfrak{B} die Menge aller Basen von W .
Das Basisfeld heißt **glatt**, wenn diese Abbildung glatt ist, d.h. genauer, wenn die einzelnen Basisvektoren entsprechend glatt vom Punkt \vec{y} abhängen.

Man geht also von **einer festen Basis für alle Punkte zu einer eigenen Basis für jeden Punkt über**, eben einem Basisfeld.

(1.6.5) Häufig ist es so, dass die Punkte $\vec{y} \in G$ durch eine Parametrisierung gegeben sind. Sagen wir: $\vec{y}^P : \vec{\alpha} \mapsto \vec{y}^P(\vec{\alpha})$. Oder: $\vec{\alpha}$ ist der Name des Punktes $\vec{y}=\vec{y}^P(\vec{\alpha})$. Dann ist es naheliegend, dass man auch die Basis, die zu \vec{y} gehört, durch diesen Namen charakterisiert. Bereits bei der Arbeit mit Partitionen fanden wir, dass man vielfach besser mit einer Parametrisierung der Partition arbeitet als mit der Partition selbst.

(1.6.6) Wir definieren daher entsprechend:

Es **SEI** W Vektorraum und $\vec{y}^P = (G_P, \vec{\alpha} \mapsto \vec{y}^P(\vec{\alpha}), G)$ eine Parametrisierung von Bild $\vec{y}^P \subset G \subset W$. Weiter sei \mathfrak{B} die Menge aller Basen von W .
Dann heißt eine Abbildung $e=(G_P, \vec{\alpha} \mapsto e(\vec{\alpha}), \mathfrak{B})$ ein **Basisfeld über der Parametrisierung \vec{y}^P** .

(1.6.8) Also: Jedem $\vec{\alpha}$ wird eine Basis $e(\vec{\alpha})$ zugeordnet, die zum Punkte $\vec{y}^P(\vec{\alpha})$ gehört. Wir werden sehen, dass sich das als nützliche Konstruktion erweist. Ein Grund dafür ist, dass sich die Basisvektoren häufig mit Hilfe der Parameter einfacher angeben lassen als mit den üblichen Koordinaten. Die Parametrisierung trennt ja -wie besprochen - Bezeichnung (oder bezeichnete Eigenschaft) von der Darstellung der Vektoren. Der unten in (1.6.15) besprochene Fall der Polarkoordinaten wird das illustrieren.

(1.6.9) Sei jetzt $G=\text{Bild}(\vec{y}^P)$. Ist \vec{y}^P injektiv, dann kann man aus einem Basisfeld über \vec{y}^P ein Basisfeld auf $G=\text{Bild}(\vec{y}^P)$ machen. Dies geht wie folgt: Sei $\vec{\alpha}_P=(G, \vec{y} \mapsto \vec{\alpha}_P(\vec{y}), G_P)$ die zu \vec{y}^P inverse Abbildung. Also $\vec{\alpha}_P \circ \vec{y}^P(\vec{\alpha}) = \vec{\alpha}$. Dann setzen wir $f(\vec{y})=e(\vec{\alpha}_P(\vec{y}))$. Das gibt offenbar das gewünschte Basisfeld.

(1.6.10) Ist \vec{y}^P nicht injektiv, dann kann es zu ein und demselben Punkt \vec{y} mehrere Basen geben. Denn ist $\vec{y}^P(\vec{\alpha}) = \vec{y}^P(\vec{\beta})$ mit $\vec{\alpha} \neq \vec{\beta}$, dann muß keineswegs $e(\vec{\alpha})=e(\vec{\beta})$ gelten

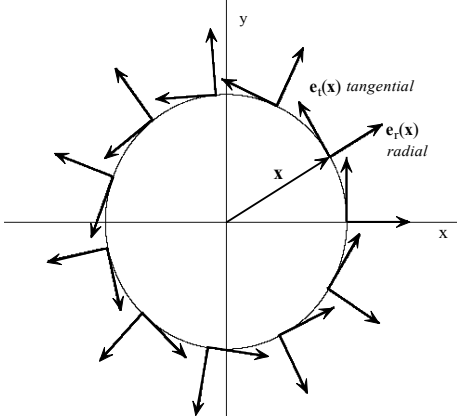
(1.6.11) In injektiven Bereichen - aber eben nur solchen - ist es daher relativ unproblematisch, zutatenfrei von *Basisfeld* zu sprechen und dieses Feld dann je nach Situation als Feld über den Konfigurationsraumpunkten \vec{y} oder als Feld über dem Parameterraum - also in $\vec{\alpha}$ - zu interpretieren. Man identifiziert die beiden Felder.

(1.6.12) Ist die Parametrisierung dagegen nicht injektiv, sollte man sorgfältig auf die Zutaten achten. Offenbar liegt wieder ein Rollenzuweisungsproblem vor. In Kap. 11.1.2 kommen wir auf die hier angedeuteten Konventionen zurück.

(1.6.13) Der bisherige Fall einer festen Basis in W ergibt sich im neuen Konzept als Spezialfall eines *konstanten Basisfeldes*.

(1.6.14) Hat man zwei Basisfelder für dasselbe $G \subset W$, dann läuft der Basiswechselformalismus nach dem in Kap.4.5 beschriebenen Schema ab. Nur dass die Transformationsmatrix T jetzt als ein Matrixfeld $\vec{y} \mapsto T(\vec{y})$ zu interpretieren ist. Im nachfolgenden Beispiel (1.6.15) wird ein Wechsel von der konstanten kanonischen Basis zum Polarbasisfeld besprochen.

(1.6.15) Führen wir das Beispiel der ebenen Polarkoordinaten aus. Es entsteht ein Basisfeld über der Polarparametrisierung \vec{x}^P aus (1.4.16). Die Transformationsmatrix führt von der kanonischen Basis (alt) zum Polarbasisfeld (neu).

<p>Das Polarbasisfeld der Ebene: $\vec{e}_r^P(r, \theta) = \vec{e}_1 \cos \theta + \vec{e}_2 \sin \theta$ $\vec{e}_t^P(r, \theta) = -\vec{e}_1 \sin \theta + \vec{e}_2 \cos \theta$</p>	<p>Transformationsmatrix $T^{-1}(r, \theta) = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}$</p>
	<p>In radialer Richtung sind die Basen einfach parallel zu verschieben. Im Ursprung ist das Basisfeld über \mathbb{R}^2 nicht definiert. Das Basisfeld über \vec{x}^P dagegen ist es mit unendlich vielen Werten.</p>

- Wir schränken die Polarabbildung nicht auf den injektiven Bereich aus (1.4.18) ein. Dann bestimmt die Polarparametrisierung dennoch ein eindeutiges Vektorfeld außerhalb des Ursprunges! Begründen Sie das. Kurz: Das Polarisfeld über \vec{x}^P ist auf der punktierten Ebene definiert, wogegen die Polarparametrisierung selbst nur auf der (entlang der negativ reellen Achse) geschnittenen Ebene glatt definiert ist.
- Stellen Sie das Polarisfeld in der punktierten Ebene als Funktion der kartesischen Koordinaten dar.
- Wie sehen rein geometrisch die Basisfelder aus, die zu den räumlichen Polar- bzw. Zylinderkoordinaten gehören?

(1.6.16) Was leistet das Basisfeld aus (1.6.15)? Offenbar gehen in unserem Beispiel die Basen für verschiedene Punkte durch Drehung auseinander hervor, nicht durch Translation wie bisher. Genauer: Sei $\vec{x} \in \mathbb{R}(\alpha)$ eine Drehung um den Winkel α bezüglich des Ursprunges. Wenn also $\vec{u} = R.\vec{x}$ der Punkt ist, der aus $\vec{x} = \vec{x}^P(r, \theta)$ durch diese Drehung hervorgeht, dann ist $R\vec{e}_i^P(r, \theta) = \vec{e}_i^P(\theta + \alpha)$. D.h. man kann die Wirkung der Gruppenoperation "Drehung" auf einfache Weise mit Hilfe der Parameter beschreiben. Die Drehung des Ortsvektors entspricht der Drehung der Basisvektoren. (Dies kommt dem allgemeinen Konzept entgegen, alles mit Hilfe der Parameter zu berechnen!). In r dagegen ist das Basisfeld konstant.

(1.6.17) Sei also $(\vec{x}, \vec{F}(\vec{x}))$ ein Ortsvektor mit zugeordnetem Feldwert. Dann ist $(R.\vec{x}, R.\vec{F}(\vec{x}))$ derjenige Feldwert am gedrehten Ort $R.\vec{x}$, den man durch Mitdrehen des Feldwertes $\vec{F}(\vec{x})$ erhält. Häufig möchte man diesen gedrehten Wert mit dem dort herrschenden Feldwert $\vec{F}(\vec{u}) = \vec{F}(R.\vec{x})$ vergleichen. Das ist normalerweise rechnerisch aufwendig. Mit Hilfe unseres Basisfeldes wird die Wirkung der Drehung automatisch durch den Parameterwechsel $\theta \mapsto \theta + \alpha$ mit erfaßt.

(1.6.18) Man sollte in diesem Zusammenhang sorgfältig die folgenden beiden Probleme auseinanderhalten: Ein Feldvektor wird am gleichen Ort gedreht. Also aus $(\vec{x}, \vec{F}(\vec{x}))$ wird $(\vec{x}, R.\vec{F}(\vec{x}))$. Und: Man geht vom Orte \vec{x} zum Orte $\vec{u} = R.\vec{x}$ und nimmt den Feldvektor mit, hier indem man ihn starr mitdreht. Aus $(\vec{x}, \vec{F}(\vec{x}))$ wird dann $(R.\vec{x}, R.\vec{F}(\vec{x})) = (\vec{u}, R.\vec{F}R^{-1}(\vec{u}))$. Am neuen Ort \vec{u} herrscht dagegen der Feldvektor $(R.\vec{x}, \vec{F}(R.\vec{x})) = (\vec{u}, \vec{F}(\vec{u}))$.

(1.6.19) In der nachfolgenden Beispielrechnung vergleichen wir die Polardarstellung mit der kartesischen. Dabei lassen wir nicht benötigte Größen wie P und r in der Bezeichnung fort. Als Argument wählen wir jeweils passend \vec{x} oder θ , von denen je nach Situation eine durch die andere Größe ausgedrückt werden kann.

(1.6.20) Wir entwickeln den Feldwert $\vec{F}(\vec{x})$ einmal nach der kartesischen Basis und einmal nach dem polaren Basisfeld. Dann drehen wir den Feldwert und vergleichen den Formalismus.

1) Entwicklung nach den beiden Basen: $\vec{F}(\vec{x}) = \vec{e}_1 F_1(\vec{x}) + \vec{e}_2 F_2(\vec{x}) = \vec{e}_r(\theta) F_r(r, \theta) + \vec{e}_t(\theta) F_t(r, \theta).$
2) Drehung um den Winkel α $R.\vec{F}(\vec{x}) = \vec{e}_1 (\sum R_{1k} F_k(\vec{x})) + \vec{e}_2 (\sum R_{2k} F_k(\vec{x})) = \vec{e}_r(\theta + \alpha) F_r(r, \theta) + \vec{e}_t(\theta + \alpha) F_t(r, \theta)$
3) Vergleich der Komponenten bei kartesischer Darstellung: $\sum_k R_{ik} F_k(\vec{x})$ mit $F_i(R\vec{x})$ $i=1,2.$
4) Vergleich der Komponenten bei Basisfelddarstellung: $F_j(r, \theta)$ mit $F_{ji}(\theta + \alpha)$ vergleichen.

Ist insbesondere der Feldvektor $\vec{F}(R.\vec{x})$ am Orte $R.\vec{x}$ durch Mitdrehen aus dem Feldvektor $\vec{F}(\vec{x})$ am Orte \vec{x} hervorgegangen, dann muß $F_j(\theta) = F_j(\theta + \alpha)$ gelten für $j=r, t$ (radial und tangential). Kurz die beiden Komponenten sind konstant. Halten Sie also immer gut auseinander: Die beiden kartesischen Feldkomponenten F_1 und F_2 sowie die beiden polaren F_r und F_t .

(1.6.21) Versuchen wir eine erste Zusammenfassung:

Vom Nutzen der Basisfelder (1): Basisfelder können bestimmte Symmetrieeigenschaften des Systems repräsentieren. Feldabbildungen mit entsprechenden Symmetrieeigenschaften besitzen dann bezüglich des Basisfeldes besonders einfache Komponenten, aus denen man die Symmetrien durch Inspektion entnimmt.

(1.6.25) Einige zugehörige Beispiele - immer zum V_0^2 :

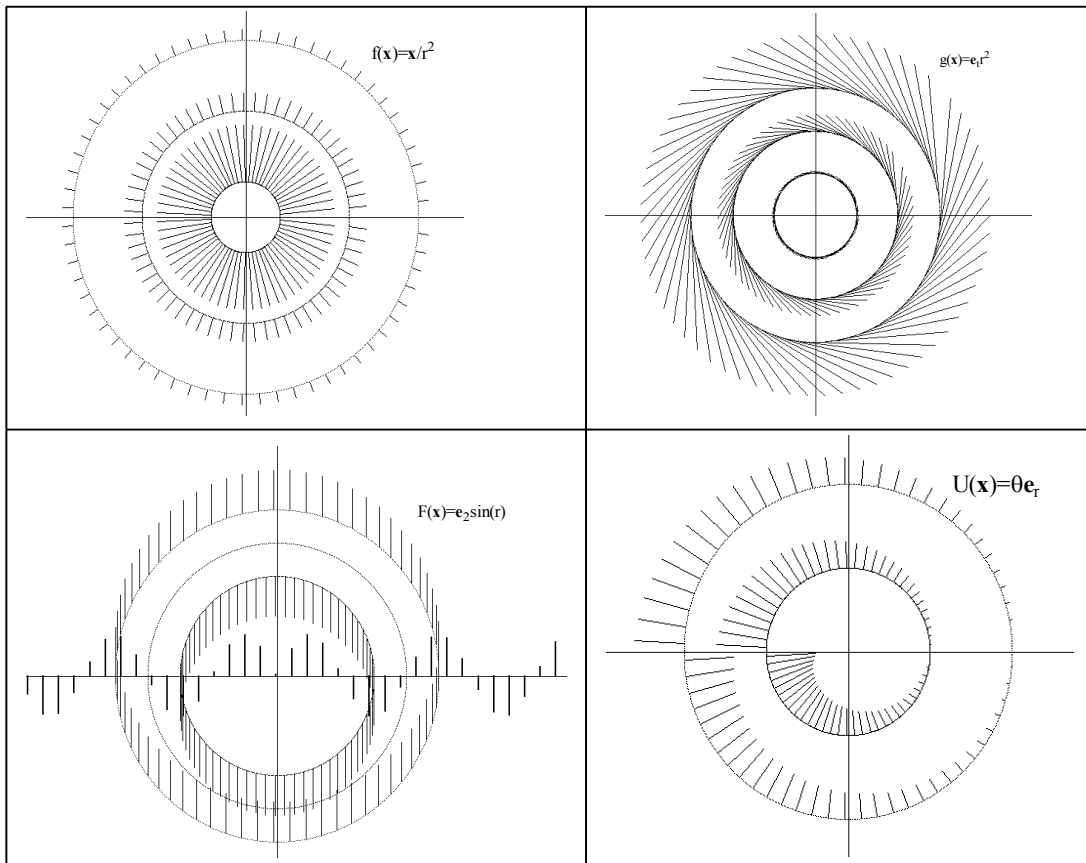
- Ein *radial gerichtetes Feld* ist offenbar eines mit $F_t(r, \theta) = 0$. Etwa $\vec{F}(\vec{x}) = \frac{\vec{x}}{x^2} = \vec{e}_r(\theta) \frac{1}{r}$. Hier ist $F_t = 0$ und $\vec{F}_r(r, \theta) = 1/r$ unabhängig von θ . Wir haben eine radiale Symmetrie. Dagegen ist $g(\vec{x}) = \vec{x}(\vec{a} \cdot \vec{x})$ radial gerichtet, aber nicht radial symmetrisch. Sei $\vec{a} = \vec{e}_1 a$. Dann folgt $g_r(r, \theta) = r^2 a \cos(\theta)$ und $g_t(r, \theta) = 0$.

- $g(r, \theta) = \vec{e}_t^P(r, \theta)r^2$ hat $g_t(r, \theta) = r^2$ und $g_r(r, \theta) = 0$. Das Feld ist tangential und nicht radial gerichtet, hat aber radiale Symmetrie, da g_t konstant in θ ist.
- Zum Vergleich: $\vec{F}(\vec{x}) = \vec{a} \sin(|\vec{x}|)$. Dies Feld ist auf Kreisen (um den Ursprung) konstant, es ist invariant gegen Drehungen, aber es ist nicht radial symmetrisch. Der Betrag des Feldvektors ist vom Radius abhängig.
- $\vec{U}(\vec{x}) = \vec{e}_r(\theta)\theta$. Also $U_r(r, \theta) = \theta$. Hier beobachten wir ein wichtiges Phänomen: Während

$$(-\pi, \pi[, \theta \mapsto U_r(r, \theta), \mathbb{R})$$

eine völlig glatte Funktion ist, ist dies beim zugehörigen $\vec{U}(\vec{x})$ nicht der Fall. Auf der negativen x-Achse entwickelt sich ein Sprung. In radialer Richtung bleibt der Feldvektor konstant.

(1.6.26) Die Figuren zeigen das Verhalten dieser Felder.



□ Kommentieren Sie den Brauch mancher Lehrbücher - auch des schulischen Bereichs - Felder vom Coulombtyp als "das radialsymmetrische Feld" zu bezeichnen.

(1.6.27) Der Nutzen der Basisfelder ist noch allgemeiner:

Vom Nutzen der Basisfelder (2): Mit ihrer Hilfe lassen sich geometrische, anschaulich oder inhaltlich gegebene Bedingungen leicht in Formeln für die Parameter oder die Feldkomponenten umwandeln. Sie leisten für die Felder das, was die Parametrisierungen für Figuren leisten.

Man hat also ein wichtiges Hilfsmittel, um verbal-inhaltlich gegebene Bedingungen einer Aufgabensituation in den Formalismus einzubauen. Umgekehrt werden geometrische Eigenschaften eines Feldes so der Inspektion unmittelbar zugänglich.

Zur Erläuterung einige weitere Beispiele:

(1.6.28) Wir betrachten eine stationäre Strömung in einem langen Zylinder. Der Zylinderradius sei R . Was liefert eine geeignete Darstellung über das Verhalten des zugehörigen Geschwindigkeitsfeldes?

Wir verwenden das zur Zylindersymmetrie gehörige Basisfeld $\vec{e}_r, \vec{e}_\theta, \vec{e}_z$, dessen Geometrie sich selbst erklären sollte. Erneut schreiben wir $\vec{e}_r(\theta)$ statt $\vec{e}_r(r, \theta, z)$ usw. Weiter sei $\vec{x} \mapsto \vec{v}(\vec{x})$ unser Geschwindigkeitsfeld, das wir jetzt nach diesem Basisfeld entwickeln. \vec{v}^Z bezeichne die Darstellung dieses Feldes in Zylinderkoordinaten. Diese hat folgende Form:

$$\vec{v}^Z(r, \theta, z) = \vec{e}_r(\theta)v_r(r, \theta, z) + \vec{e}_\theta(\theta)v_\theta(r, \theta, z) + \vec{e}_z v_z(r, \theta, z)$$

Die Formel zeigt, weshalb man üblicherweise alle Argumente fortläßt, sofern das irgend möglich ist. Wir lassen alle Argumente fort, in denen die jeweilige Funktion konstant ist.

Mit Hilfe der angegebenen Darstellung kann man unmittelbar naheliegende Situationsbedingungen in Formeln darstellen. Einige Beispiele des Typs *von der Situation zur Formel*:

- Durch den Zylindermantel soll keine Strömung hindurchgehen. Dies bedeutet einfach $v_r(R, \theta, z) = 0$, wenn R der Zylinderradius ist. Man sagt auch, *die Normalkomponente der Geschwindigkeit auf dem Zylindermantel muß verschwinden*. Versuchen Sie das einmal kartesisch in einer einzigen Formel zu formulieren.
- Angenommen die Strömung ist drehsymmetrisch zur Zylinderachse. Dann müssen alle drei Komponenten unabhängig von θ sein. Beachten Sie: Eine Komponente $v_\theta \neq 0$ ist dann durchaus möglich. Sie beschreibt eine Verdrillung der Gesamtströmung um die Achse. Wählt man $v_r = v_\theta = 0$ und $v_z = \omega r$, so erhält man die Drehbewegung eines starren Körpers um die Zylinderachse.
- Angenommen die Strömung ist drehsymmetrisch und achsenparallel. Dann gilt $v_r = v_\theta = 0$ und das Geschwindigkeitsfeld hat die Form $\vec{e}_z v_z(r, z)$. Üblicherweise wird v , nach außen hin abnehmen und vielfach nicht mehr von z abhängen.

(1.6.29) Und jetzt der Aufgabentyp *von der Formel zur Interpretation*.

- Sei $\vec{v}(\vec{x}) = \vec{e}_z v_0 (1 - \frac{r}{R})^2$. Das beschreibt eine achsenparallele Strömung, die symmetrisch zur Achse verläuft. Auf dem Zylindermantel ($r=R$) wird sie Null und nimmt von dort aus monoton nach innen zu. Auf der Zylinderachse wird der maximale Betrag v_0 der Geschwindigkeit erreicht. Der ergänzende Faktor ist einheitenfrei! Die Strömung hängt nicht von der Koordinate z ab, ist invariant gegen Verschiebungen in Richtung z -Achse. Die Feldlinien sind daher achsenparallele Geraden.
- $\vec{v}(\vec{x}) = \vec{e}_z f(r \sin(\theta))$. Da $r \sin(\theta) = y$ ist, haben wir es hier mit einer Strömung zu tun, die in den Ebenen $y = \text{const}$ konstanten Wert hat.

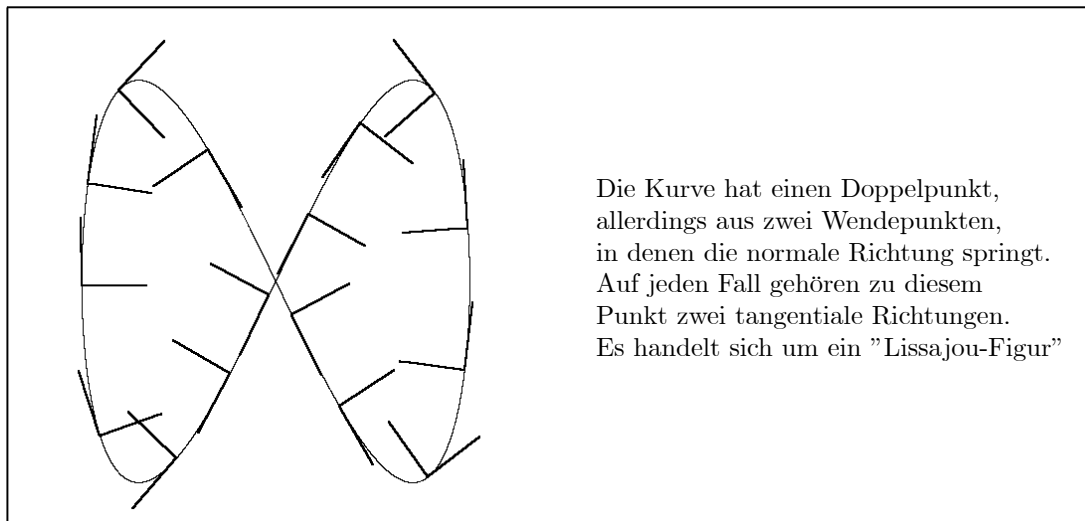
6.1.6a Das begleitende Zweibein einer Kurve

(1.6.30) In allen bisherigen Beispielen sind wir stillschweigend von einer injektiven Parametrisierung ausgegangen und haben entsprechend Basisfeld auf dem Konfigurationsraum und Basisfeld über der Parametrisierung gefahrlos identifiziert. Bei dem nächsten Beispiel sollte man das nur unter zusätzlichen Voraussetzungen, keinesfalls allgemein tun.

(1.6.31) Basisfelder über einer Parametrisierung gibt es keineswegs nur für den Fall $k=n=m$. Wir besprechen ein illustratives Beispiel für den Fall einer ebenen Kurve ($k=1, n=m=2$). Unsere Parametrisierung sei $\vec{r} = (J, t \mapsto \vec{r}(t) = \vec{e}_1 x(t) + \vec{e}_2 y(t), V_0^2)$. Das ist eine nicht notwendig injektive Parametrisierung von $\text{Bild}(\vec{r})$. Ortsabhängige Basis heißt hier Basisfeld über der Kurve \vec{r} . Für jedes t aus J gehört also zum zugehörigen Ortsvektor $\vec{r}(t)$ eine Basis. Nicht injektive Punkte des Bildes von \vec{r} können mehrere Basen für unterschiedliche Zeitpunkte erhalten. Für Punkte außerhalb des Bildes hat man dagegen keine Basis.

(1.6.32) Wir brauchen für jedes t zwei Basisvektoren. Einen wählen wir tangential in Bewegungsrichtung und den anderen senkrecht dazu. Beide sollen normiert sein und der senkrechte in Richtung des Krümmungsmittelpunktes, also nach innen zeigen. Diese Bedingungen legen die beiden Richtungen weitgehend fest. Ausnahmen sind Wendepunkte oder geradlinige Kurvenstücke. Ein solches Basisfeld (aus den beiden Einheitsvektoren $\vec{e}_T(t)$ und $\vec{e}_N(t)$) nennt man suggestiv das (die Kurve) begleitende Zweibein. Wir werden

im Rahmen der Differentialgeometrie zeigen, wie man solche Basisfelder rechnerisch beherrscht, auch für Kurven in höheren Dimensionen.



Die Kurve hat einen Doppelpunkt, allerdings aus zwei Wendepunkten, in denen die normale Richtung springt. Auf jeden Fall gehören zu diesem Punkt zwei tangentielle Richtungen. Es handelt sich um ein "Lissajou-Figur"

Alternativ kann man den zweiten Vektor so wählen, dass beide Vektoren zusammen immer ein Rechtssystem bilden. Dann tritt das Problem mit den Wendepunkten nicht auf. (Wie sieht das in obiger Figur aus?)

(1.6.33) Für den einfachen ebenen Fall können wir den zugehörigen Formelapparat aufstellen, wenn wir beachten, dass die gesuchte tangentielle Einheitsrichtung gleich der Richtung der vektoriellen Geschwindigkeit ist. Nur das Vorzeichen der Normalenrichtung verbleibt hier noch fallspezifisch zu klären. Der Einschub zeigt Vorgehen und Resultate.

(1.6.34) **Formeln für das ebene begleitende Zweibein.**

Gegeben sei $\vec{r}(t) = \vec{e}_1 x(t) + \vec{e}_2 y(t)$.
 Ableiten, und das Ergebnis in Betrag und Richtung aufspalten.
 $\dot{\vec{x}}(t) = \vec{v}(t) = \vec{e}_1 \dot{x}(t) + \vec{e}_2 \dot{y}(t) = \vec{e}_T(t) v(t)$
 mit $v(t) = \sqrt{\dot{x}^2(t) + \dot{y}^2(t)}$ und $\vec{e}_T(t) = \frac{1}{v(t)}(\vec{e}_1 \dot{x}(t) + \vec{e}_2 \dot{y}(t))$
 Die übliche Konstruktion liefert bis auf ein Vorzeichen ε den dazu senkrechten Vektor $\vec{e}_N(t) = \frac{\varepsilon}{v(t)}(\vec{e}_1(-\dot{y}(t)) + \vec{e}_2 \dot{x}(t))$.

(1.6.25) Als einfache Konkretisierungen kann man den Fall eines Kreises oder einer Parabel mit $x(t)=t$ und $y(t)=\alpha t^2$ rechnen. $\vec{r}^K(t)=(\sin(t), \sin(2t))$ dagegen ergibt die oben abgebildete Lissajoufigur mit dem Doppelpunkt im Ursprung.

□ Rechnen Sie die angegebenen Beispiele durch.

(1.6.36) Wir wollen noch kurz auf eine physikalische Interpretation derartiger begleitender Zweibeine hinweisen: Angenommen \vec{r} beschreibt die (physikalische) Bahn eines Massenpunktes, der unter dem Einfluß eines Kraftfeldes \vec{F} steht. Dann wirkt zum Zeitpunkt t die Kraft $\vec{F}(\vec{r}(t))$ auf diesen Massenpunkt. Wir bilden die Zerlegung der Kraft nach diesem Basisfeld: $\vec{F}(\vec{r}(t)) = \vec{e}_T(t) F_T(t) + \vec{e}_N(t) F_N(t)$. Dann ist F_T die jeweilige Kraftkomponente, die tangential zur Bewegungsrichtung wirkt, also den Betrag der Geschwindigkeit zu ändern sucht und F_N ist die dazu senkrechte Komponente, die nur die Richtung der Geschwindigkeit ändern kann, die Normalkomponente der Kraft. D.h., mit Hilfe des Zweibeins lassen sich diese wichtigen Größen sofort angeben und zur Parametrisierung der Gesamtkraft verwenden.

□ Gegeben sei die kanonische Parametrisierung einer Parabel, also $\vec{r}(t) = (t, \alpha t^2)$ und eine Kraft $\vec{F} = (0, -g)$. Zerlegen Sie \vec{F} in die zugehörige parallele und tangentielle Komponente. Was ist, wenn \vec{F} ein Kraftfeld ist, etwa $\vec{F}(x,y)=(x,y)(x^2+y^4)$?

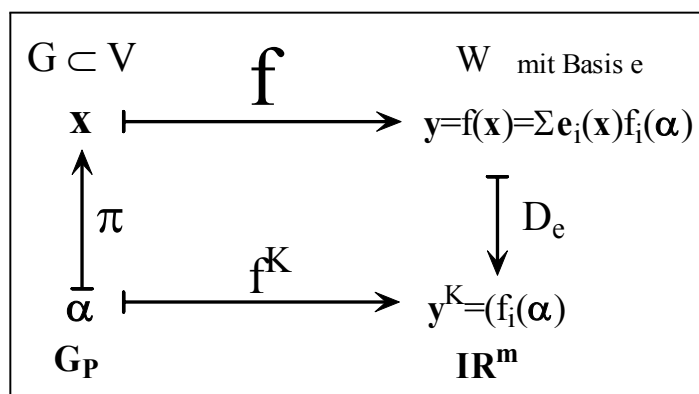
6.1.7 Koordinatendarstellungen der Abbildungen

(1.7.1) Mit Hilfe eines Basisfeldes kann man die geometrischen Vektoren $f(\mathbf{x})$ durch Zahltuple, also Koordinatenvektoren beschreiben. Diese können jeweils geeignete oder gewünschte Eigenschaften der Konfiguration besonders gut sichtbar machen. Insgesamt sind wir in der Lage, eine vollständige Koordinatendarstellung unserer absoluten Abbildung $f:G \rightarrow W$ zu geben. Wir müssen nur noch die Komponenten bezüglich des Basisfeldes zum Tupel, zum Koordinatenvektor zusammenfassen. Es geht hier also um die Verallgemeinerung der Matrixdarstellung linearer Abbildungen!

(1.7.2) **Wir lösen das Problem dadurch, dass wir das im Kapitel 4.4.3 für lineare Abbildungen entwickelte Diagramm an die neue Situation anpassen.** Das erweist sich als relativ leicht, wobei man beachten sollte

- Die Parametrisierung $\pi = \vec{x}^\pi$ muss nicht invertierbar sein
- Weiter verwenden wir gleich ein Basisfeld e über G_P .

Der zweite Punkt ist delikat, weil er besagt: Es genügt nicht, den Ort \vec{x} zu kennen, um die Basisdarstellung zu konstruieren. Vielmehr benötigt man den Parameterwert $\vec{\alpha}$ mit $\vec{x}^\pi(\vec{\alpha}) = \vec{x}$, um die Basis festzulegen. Ist \vec{x}^π nun nicht injektiv, dann kann s zu einem Ort \vec{x} durchaus mehrere Basen geben.



f ist die gegebene zu quantifizierende Abbildung und π eine Parametrisierung des Konfigurationsraumes. Weiter sei $\vec{x} \mapsto \vec{e}(\vec{x})$ ein Basisfeld über der Abbildung f . D.h. jedem Punkt \vec{x} wird eine Basis $\vec{e}(\vec{x})$ von W (nicht von V !) zugeordnet. Es entsteht eine zugehörige Koordinatendarstellungsabbildung $D_e(\vec{x})$ mit äußerem Parameter \vec{x} . Damit wird der absolute Vektor $\vec{y} = f(\vec{x})$ aus W in ein Koordinatentupel umgewandelt.

Das Ergebnis ist die Abbildung $f^K = D_e \circ f \circ \pi = (G_P, \vec{\alpha} \mapsto (f_i(\vec{\alpha})), \mathbb{R}^m)$

- Welche Vereinfachungen erhält man für den Spezialfall eines Skalarfeldes, was wird dann nicht benötigt? Zeichnen Sie das zugehörige Diagramm. Welche Identifizierung und Vereinfachung des Diagramms liegt nahe?

(1.7.3) Die Abbildung f^K ist die gesuchte zum Koordinatensystem K gehörige Quantifizierung der Ausgangsabbildung f . Jeder Wert ist ein m -tupel von Komponenten, die alle von den n Beschreibungsparametern $\vec{\alpha}$ abhängen. Damit gehört zur Quantifizierungsabbildung f^K eine Zuordnung, die aus dem k -tupel der α_j ein m -tupel von Funktionen f_i macht. Wir schreiben diese Darstellung von f einmal im Matrix- und einmal im Indexkalkül genauer auf, wobei sich erneut die Schreibökonomie des Indexkalküls zeigt.

$$\left(\begin{array}{c} \alpha_1 \\ \vdots \\ \alpha_k \end{array} \right) \xrightarrow{f^K} \left(\begin{array}{c} f_1(\alpha_1, \dots, \alpha_k) \\ \vdots \\ f_m(\alpha_1, \dots, \alpha_k) \end{array} \right) \quad \text{bzw.} \quad \alpha_j \xrightarrow{f^K} f_r(\alpha_i) \quad \begin{array}{l} i, j = 1, 2, \dots, k \\ r = 1, 2, \dots, m \end{array}$$

(1.7.4) Im linearen Fall ist f^K eine Matrixabbildung. Dann kann man auf der Werteseite per Matrixmultiplikation Abbildung und Parameter noch algebraisch trennen: $\vec{y}^K = \text{Matrix} \cdot \vec{\alpha}$. Im allgemeinen Fall ist das nicht mehr möglich.

(1.7.5) **Mathematisch sind f und f^K zwei verschiedene Abbildungen, auch wenn sie dasselbe physikalische oder geometrische Phänomen beschreiben. Nur eben einmal koordinatenfrei absolut und einmal koordinatenbezogen.**

(1.7.6) Beispiel: Wir stellen absolute und Koordinatendarstellung für eine Abbildung nebeneinander: Als Parametrisierung wählen wir Polarkoordinaten der Ebene. Unser (geometrisches) Feld der Ebene sei (mit Skalarprodukt):

$$f(\vec{x}) = \vec{e}_1 \cdot \frac{R}{r} (-\vec{e}_2 \cdot \vec{x}) + \vec{e}_2 \frac{R}{r} (\vec{e}_1 \cdot \vec{x}) \quad \text{wobei } r = \sqrt{\vec{x}^2}.$$

r bewirkt die Nichtlinearität von f . Darstellung von \vec{x} durch Polarkoordinaten ergibt

$$(r, \theta) \mapsto f(\vec{x}^P(r, \theta)) = R(-\vec{e}_1 \sin \theta + \vec{e}_2 \cos \theta) = R \cdot \vec{e}_t(\theta) + 0\vec{e}_r(\theta).$$

Daraus lesen wir sofort die gesuchte zugehörige Koordinatenabbildung ab:

$$\begin{pmatrix} r \\ \theta \end{pmatrix} \xrightarrow{f^K} \begin{pmatrix} f_r(r, \theta) \\ f_t(r, \theta) \end{pmatrix} = \begin{pmatrix} 0 \\ R \end{pmatrix}$$

Das ist eine sehr einfache konstante Funktion! Die gesamte Ortsabhängigkeit ist hier in der Ortsabhängigkeit der Basisvektoren enthalten. Die Komponenten sind infolge der "richtigen" Basisfeldwahl konstant. Vergleichen Sie damit den doch ziemlich komplizierte Bau von $f(\vec{x})$ in kartesischen Koordinaten. In den meisten physikalischen Texten erfolgt die Darstellung und Vorgabe unserer Abbildungen $f:V \rightarrow W$ fast ausschließlich über derartige Koordinatendarstellungen f^K .

- Sei $\vec{F}(\vec{x}) = \alpha \frac{\vec{x}}{|\vec{x}|^3}$ Kraftfeld vom Coulombtyp. Bestimmen Sie für dieses Feld die folgenden Koordinatendarstellungen: a) F^K für kartesische Koordinaten b) \vec{F}^P für räumliche Polarkoordinaten und c) \vec{F}^Z für Zylinderkoordinaten.

(1.7.8) Der einfachste, aber keineswegs seltene Fall ist der eines konstanten kartesischen Basisfeldes über dem Konfigurationsraum G . Das Diagramm bleibt gültig, vereinfacht sich nur dadurch, dass man an zwei Stellen das Argument \vec{x} fortlassen kann: In $\vec{e}(\vec{x})$ und $D_e(\vec{x})$. Beide Größen werden konstant. Dies erweist sich als wichtig, wenn man f Ableiten will.

(1.7.9) Was ist, wenn die Parametrisierung π nicht injektiv ist, man aber mit einem Basisfeld arbeiten möchte, das von den Parametern α_r abhängt? Erinnert sei an den Fall einer Kurve mit Doppelpunkten. Dann ist eine leichte Modifikation des Diagrammes aus (1.7.2) erforderlich. Vergleichen Sie dazu (1.6.4) und (1.6.6): Dagegen kann man die Formeln aus (1.7.3) übernehmen, sofern man die sonst unproblematische Identifikation von $\vec{\alpha}$ und \vec{x} ausschließt.

- Modifizieren Sie das Diagramm (1.7.2), um die Frage aus (1.7.9) zu beantworten: Eine zusätzliche diagonale Linie ist erforderlich und einige Funktionsargumente sind abzuändern.

(1.7.9) Das Verständnis des Diagrammes (1.7.2) samt seinen Spezialisierungen und Verallgemeinerungen ist für zahlreiche Probleme von größtem Nutzen. Verstehen Sie es daher genau und prägen Sie es sich gut ein.

6.1.8 Symmetriebetrachtungen

Damit haben wir alle Vorbereitungen getroffen, um den etwas anspruchsvolleren, aber umso wichtigeren Punkt der Symmetriebetrachtungen anzugehen.

(1.8.1) Die zugehörige Idee können wir wie folgt formulieren:

■ 1)	$f:G \rightarrow W$ soll analysiert werden.
■ 2)	Wähle dazu (nachfolgend günstige) Parametrisierung $\vec{x}^K : G_K \rightarrow G$ von G .
■ 3)	Wähle weiter (günstiges) Basisfeld e auf G_K für W .
♣ 4)	Bilde die zugehörige Koordinatendarstellung $f^K : G_K \rightarrow \mathbb{R}^m$ des Feldes.
♠ 5)	Schränke die Parametrisierung auf einen möglichst kleinen Bereich $G_{K_0} \subset G_K$ ein, für den f^K noch spezifisch zu diskutieren ist.
♠ 6)	Durch eine Gruppenoperation (möglichst vom Typ der kleinen Transformationen) soll man dann in V von G_{K_0} nach G_K gelangen können, derart, dass man die Änderung des Feldvektors verfolgen kann , i.a. getrennt nach Komponenten und Basisfeld.

(1.8.2) Erneut liegt ein Beispiel einer Denkfigur vor. Im Deklarationsteil 1-3) werden die benötigten Zutaten zusammengestellt. 4) Enthält eine Hilfskonstruktion als Anwendung unseres Formalismus. 5) und

6) enthalten die Problemaufspaltung in einen verbleibenden nichtrivialen fallspezifischen **Diskussionsrest** der zu untersuchenden Abbildung **auf einer kleineren Urbildmenge** und einen geometrischen Teil. Die Abbildungseinschränkung in 5) sollte möglichst so sein, dass sie durch Konstantsetzen von Koordinaten - siehe 4)- erfolgt. D.h. der verbleibende interessante Teil des Verhaltens der Abbildung ist auf einer Niveaumenge dieser Koordinate vorzufinden.

Beispiel: Im Falle von $\vec{x} \mapsto \vec{\omega} \times \vec{x}$ genügt es, das Verhalten auf einem einzigen Ursprungskreis in der Ebene senkrecht zu $\vec{\omega}$ zu untersuchen.

(1.8.3) Schritt 6) sieht genauer wie folgt aus:

Man hat eine Transformationsgruppe T, die sowohl auf V als auch auf W operiert. Die erste Operation bezeichnen wir mit $\circ: T \times V \rightarrow V$ und die zweite mit $*$: $T \times W \rightarrow W$. Dann operiert T auch auf den Graphenpunkten der zu studierenden Abbildung vermöge $(\vec{x}, f(\vec{x})) \mapsto (g \circ \vec{x}, g * f(\vec{x}))$. G_{k_0} soll ein Vertretersystem der Bahnen von \circ in G sein. D.h. die zugehörigen Bahnen erzeugen ganz G möglichst als Resultat kleiner Transformationen. Dann beschreibt $*$, was entlang jeder Bahn mit dem Wertevektor $\vec{y} = f(\vec{x})$ geschieht. Natürlich muß insgesamt als **Symmetriebedingung für f** gelten:

$$f(g \circ \vec{x}) = g * f(\vec{x})$$

(1.8.4) Die für uns wichtigsten Beispiele für die Mitführungsoperation $*$ sind $g * \vec{y} = \vec{y}$, also Invarianz oder Parallelverschiebung des Wertes, und im Falle $V=W=V$, die radiale Symmetrie $f(R \cdot \vec{x}) = R \cdot f(\vec{x})$, wenn R eine Drehung ist. Unten werden wir unter dem Stichwort *Homogenität* ein weiteres wichtiges Beispiel einer Transformation kennenlernen.

□ Gehen Sie jetzt selbst das Programm für die beiden in 6.1 besprochenen Strömungsfelder durch.

(1.8.5) Hängt ein Skalarfeld **nicht** von einer Koordinate α ab, so kommt für G_{k_0} eine Niveaulfläche $\alpha = \alpha_0 = \text{const}$ in Frage. Die Transformationsgruppe T muß dann unter Festlassen der übrigen Parameter die möglichen Werte von α erzeugen. D.h. die Transformationsgruppe sollte zunächst im Parameterraum erzeugt und dann in den Konfigurationsraum übertragen werden. Entlang der durch die α -Variation erzeugten Bahnen hat man für das Skalarfeld dann die konstante Mitführungsoperation $*$ (Denn der Feldwert hängt ja nicht von diesem α ab).

□ Beispiel: $\vec{x} \mapsto (\vec{e}_3 \cdot \vec{x}) \vec{x}^2$ hängt in räumlichen Polarkoordinaten nicht vom Längewinkel φ ab. Die Niveaumenge $\varphi = \varphi_0$ ist eine Halbebene und T ist die Drehgruppe um die Polarachse. Was ist hier $*$? In der Niveaumenge selbst kann man noch die radiale Längenänderung betrachten. Dann verbleibt für G_{k_0} nur ein Halbkreis.

(1.8.6) All dies ist natürlich im Rahmen unserer Konfigurationsraumvorstellungen naheliegend und wie folgt zu interpretieren: Wir bewegen uns entlang einer Bahn im Konfigurationsraum und messen dabei die Feldstärke aus. Symmetrie besagt, dass sich die Meßergebnisse nach einer einfachen durch $*$ erfaßten geometrischen Gesetzmäßigkeit beschreiben und somit vorhersagen lassen.

(1.8.7) In speziellen Fällen können Teile des Programms trivial werden. Im Fall eines Skalarfeldes etwa entfällt das Problem des zu wählenden Basisfeldes e.

(1.8.8) **Homogenität.** Die betrachteten Abbildungen f sind meist nicht linear. Linearität setzt sich aus Additivität und Homogenität zusammen. Vielfach ist die Additivität verletzt, aber die Homogenität gilt noch in einer verallgemeinerten Weise und legt fest, wie sich das Feld entlang Ursprungshalbgeraden im Konfigurationsraum verändert.

Sei	$\vec{F} = (V, \vec{x} \mapsto \vec{F}(\vec{x}), W)$. Es gebe eine Zahl $k \in \mathbb{R}$ derart, dass für alle $\lambda \in \mathbb{R}$ und $\vec{x} \in V$ gilt: $\vec{F}(\lambda \vec{x}) = \lambda^k \vec{F}(\vec{x})$.
Dann	heißt \vec{F} <i>homogen vom Grade k</i> .

(1.8.9) In diesem Fall haben wir daher die Gruppenoperation $(\vec{x}, \vec{F}(\vec{x})) \mapsto (\lambda \vec{x}, \lambda^k \vec{F}(\vec{x}))$ im Graphenraum $V \times W$. (Prüfen Sie nach, dass eine Gruppenoperation für \mathbb{R}_+ vorliegt!)

Im Falle der Linearität ist $k=1$. Ist \vec{F} allgemein homogen vom Grade k, so besagt die Operation: Entlang einer Halbgeraden durch den Ursprung ändert der Feldvektor $\vec{F}(\vec{x})$ in W nur seine Länge, nicht aber seine Richtung. Und die Längenänderung erfolgt gemäß der vertrauten reellen Funktion $x \mapsto x^k$ (für $x > 0$.)

(1.8.10) Der übliche Test auf Homogenität erfolgt durch simple Inspektion. "Was geschieht, wenn ich \vec{x} bzw. alle Koordinaten x_i mit λ multipliziere? Vielfach sieht man unmittelbar, ob Homogenität vorliegt oder nicht. Dabei darf der Homogenitätsgrad durchaus Null, negativ oder nicht ganzzahlig sein.

□ Prüfen Sie, dass $(x, y) \mapsto \frac{xy}{x^2+y^2}$ homogen vom Grade Null ist. Was bedeutet das?

(1.8.11) Wir gehen das Programm für einige Beispiele durch. Vergleichen Sie das Vorgehen im ersten Beispiel mit Ihrem eigenen Versuch ihn (1.8.4)!

(1.8.12) **1. Beispiel:** Das Geschwindigkeitsfeld des rotierenden starren Körpers: $\vec{x} \mapsto \vec{\omega} \times \vec{x}$.

Für den 2. und 3. Schritt aus (1.8.1) verwenden wir eine Zylinderkoordinatenparametrisierung (r, θ, h) mit Achse in Richtung von $\vec{\omega}$. Das zugehörige Basisfeld ist wieder selbsterklärend $(\vec{e}_r(\theta), \vec{e}_t(\theta), \vec{e}_h = \vec{e}_3)$. Damit erhalten wir

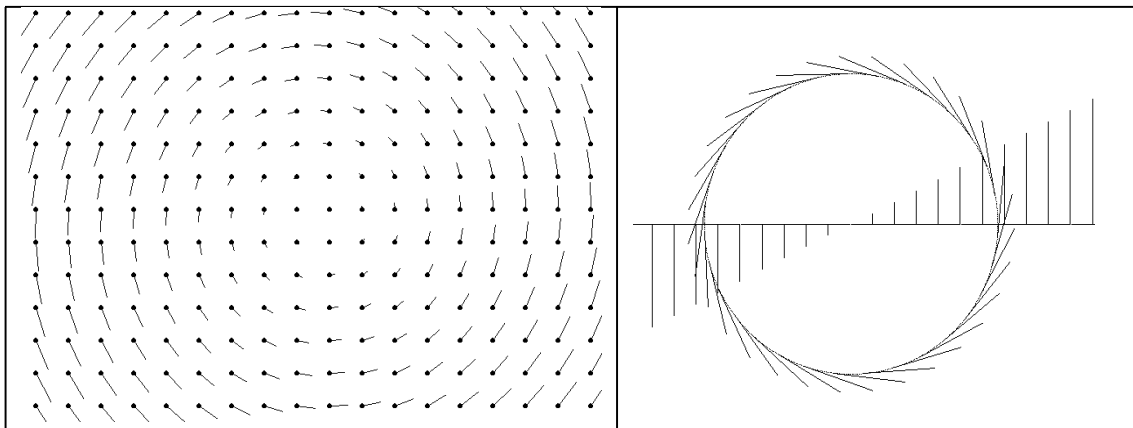
$$\vec{\omega} \times \vec{x} = \omega r \vec{e}_t(\theta) \quad \text{und in Koordinatenform} \quad {}^t(r, \theta, h) \mapsto {}^t(0, \omega r, 0),$$

wie man sofort einschließlich des Vorzeichens verifiziert. Jetzt bestimmen wir den Feldwert für einen einzigen Punkt \vec{x} , mit Parametern $(r, \theta, 0)$. außerhalb der Achse, also mit $r > 0$ in der Koordinatenebene $h=0$. (Die einpunktige Menge $\{\vec{x}_0\}$ ist unser G_{K_0} .) Das ist Punkt 5).

Es bleibt 6). Zuerst betrachten wir die Halbgerade vom Ursprung durch \vec{x}_0 . Wir können für sie sofort eine Parametrisierung $r \mapsto \vec{x}^Z(r, \theta, h)$ angeben. Der Feldwert ändert sich entlang dieser Geraden trivial, einfach durch Längenänderung wie unsere Darstellung $\omega r \vec{e}_t(\theta)$ zeigt. Denn das Feld ist überall homogen vom Grade 1.

Jetzt drehen wir die gesamte Halbgerade um den Ursprung um den Winkel α . Wegen $R\vec{e}_t(\theta) = \vec{e}_t(\theta + \alpha) =$ (tangentialer Vektor am neuen Ort) ist aber der gedrehte Feldvektor der Feldvektor am neuen Ort. Damit beherrschen wir das Feld in der Ebene $h=0$. Diese verschieben wir parallel zur Polachse. Dabei ändern sich $\vec{e}_t(\theta)$ und r nicht, so dass der verschobene Vektor nach unserer Darstellungsformel gleich dem Feldvektor am neuen Ort ist. **Damit ist der Feldwert für alle Punkte konstruiert.**

Im linken Bild ist die Feldkonfiguration im Bereich des Nullpunktes gezeichnet. Die Zeichenebene ist senkrecht zum Vektor a gewählt. Die Länge von a ist klein gewählt, damit die Feldvektoren nicht zu groß werden. Im rechten Bild ist die Feldkonfiguration auf einem Kreis um den Ursprung und entlang einer Achse gezeichnet. Hier ist a größer gewählt.



(1.8.13) **2. Beispiel:** Ein Skalarfeld in der Ebene.

$$s(x, y) = \frac{xy}{x^2+y^2} \quad \text{für } (x, y) \neq (0, 0).$$

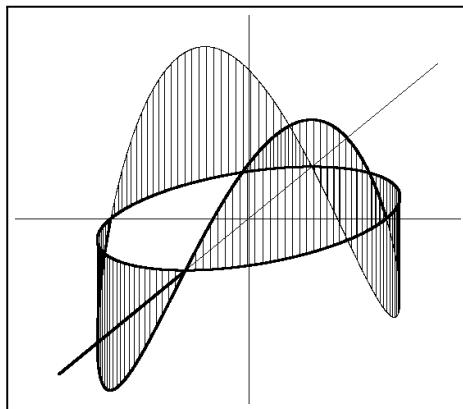
Inspektion zeigt: Das Feld ist homogen vom Grade Null. Als Parametrisierung wählen wir Polarkoordinaten. Das Basisfeld im Werteraum entfällt. Wie sieht die Koordinatendarstellung aus? Wir setzen $x=r\cos\theta$ und $y=r\sin\theta$ und erhalten:

$$S^P(r, \theta) = \frac{r \cos \theta r \sin \theta}{r^2(\cos^2 \theta + \sin^2 \theta)} = \cos \theta \sin \theta.$$

Dieser Ausdruck hängt nicht mehr von der Koordinate r ab. Daher benötigen wir den Feldwert nur für einen Kreis, sagen wir den Einheitskreis. Entlang Halbgeraden durch den Ursprung ist das Feld konstant, jede Halbgerade ist Teil einer Niveaumenge. Wir benötigen also nur die Funktion $\theta \mapsto \sin \theta \cos \theta = \frac{1}{2} \sin(2\theta)$ für $0 \leq \theta \leq 2\pi$. Diese Funktion ist uns aber wohlbekannt und damit das gesamte Feld.

Die Niveaumenge zum Wert 0 besteht hier aus den 4 Koordinatenachsen ohne den Ursprung. Entlang der Winkelhalbierenden wird das Feld maximal bzw. minimal. Der Graph dieses Feldes entsteht mit Hilfe

des folgenden Prozesses: Nehme die positive x-Achse und drehe sie um den Ursprung. Dabei ist der z-Wert, die Höhe der gesamten Halbgeraden, gleich $z(\theta) = \cos(\theta)\sin(\theta)$ anzusetzen. Die so entstehende Fläche ist der Graph. Man erkennt, dass das Feld im Ursprung unstetig ist. Im nachfolgenden Bild ist der Schnitt dieser Fläche mit einem Zylinder, dessen Achse die z-Achse ist, dargestellt. Das ergibt einfach eine Aufwicklung des Graphen von $\theta \mapsto \sin(\theta)\cos(\theta)$ auf den Zylindermantel!



Überlegen Sie sich selbst Fall des Feldes $s(x,y,z) = \frac{xyz}{(x^2+y^2+z^2)^{\frac{3}{2}}}$ im Raume.

3. Beispiel: $\vec{G}(\vec{x}) = \vec{x} \times (\vec{\omega} \times \vec{x})$ mit $\vec{\omega} \neq \vec{0}$.

Über die geometrische Definition des Vektorproduktes sieht man sofort, dass \vec{G} drehinvariant zur $\vec{\omega}$ -Achse ist. Es genügt, eine Ebene zu betrachten, die $\vec{\omega}$ enthält. $\vec{G}(\vec{x})$ ist dann auch ein Vektor dieser Ebene. $\vec{G}(\vec{x})$ steht senkrecht auf \vec{x} . Außerdem sieht man sofort: $\vec{G}(-\vec{x}) = \vec{G}(\vec{x})$. Beachten Sie, was das für die **Feldrichtung** bedeutet! Schließlich ist \vec{G} homogen vom Grade 2. Auf der von $\vec{\omega}$ erzeugten Geraden ist der Feldwert Null.

Führen wir eine kartesische Basis K ein mit $\vec{\omega} = e_3\omega$, so erhalten wir folgende Darstellung:

$$G = \left(\mathbb{R}_K^3, \begin{pmatrix} x \\ y \\ z \end{pmatrix} \mapsto \omega(\vec{e}_1(-xy) + \vec{e}_2(x^2 + z^2) + \vec{e}_3(yz)), V \right)$$

$$G^K = \left(\mathbb{R}_K^3, \begin{pmatrix} x \\ y \\ z \end{pmatrix} \mapsto \begin{pmatrix} -xy \\ x^2 + y^2 \\ yz \end{pmatrix}, \mathbb{R}_K^3 \right)$$

Führen wir dagegen räumliche Polarkoordinaten (mit zugehörigem Basisfeld) ein, so folgt:

$$\begin{aligned} \vec{G}(\vec{x}) &= (\vec{x} \times \vec{\omega}) \times \vec{x} & \vec{x} &= \vec{e}_r(r, \theta, \varphi) \\ \vec{x} \times \vec{\omega} &= \vec{e}_\varphi r \omega \sin \theta & (\text{pos. Zeichen. Wähle z.B. } \varphi = 0, \theta = \frac{\pi}{2}) \\ (\vec{x} \times \vec{\omega}) \times \vec{x} &= (\vec{e}_\varphi \times \vec{e}_r) r^2 \omega \sin \theta = \vec{e}_\theta r^2 \omega \sin \theta \end{aligned}$$

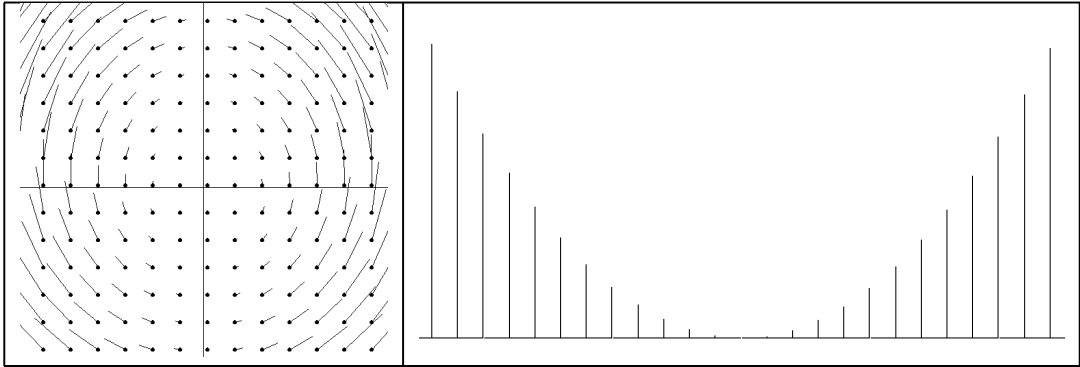
Also

$$\vec{G}(\vec{x}) = (\vec{x} \times \vec{\omega}) \times \vec{x} = \vec{e}_\theta r^2 \omega \sin \theta$$

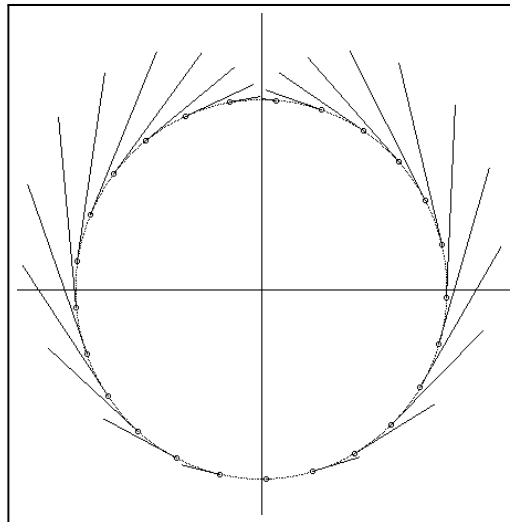
$$G^P = \left(G_P, \begin{pmatrix} r \\ \theta \\ \varphi \end{pmatrix} \mapsto \begin{pmatrix} 0 \\ r^2 \omega \sin \theta \\ 0 \end{pmatrix}, \mathbb{R}^3 \right).$$

Die Polarkoordinatendarstellung des Feldes hat die gewünschte einfache Form. Sie zeigt sofort: Das Feld ist symmetrisch bei Drehungen um die Polachse, genauer gesagt, das Feld wird mitrotiert (*-Operation). Geht man radial nach außen, so erhöht sich die Länge des Feldvektors quadratisch (Homogen vom Grade 2). Die Richtung des Feldvektors ist immer tangential zu den Kugeloberflächen vom Südpol zum Nordpol gerichtet. Insgesamt genügt die Kenntnis der Feldvektoren auf einem Halbkreis um den Ursprung, der die $\vec{\omega}$ -Achse enthält.

Zur bildlichen Darstellung restringieren wir auf die y - z -Ebene. Das ist hier unser G_{K_0} . Dann wird $G^K(0, y, z) = \omega x \begin{pmatrix} -y \\ x \\ 0 \end{pmatrix}$. D.h. auch der Bildvektor liegt in der y - z -Ebene. Die nachfolgenden Bilder verdeutlichen das Verhalten des Feldes in dieser Ebene. Das zweite Bild zeigt den Schnitt entlang der y -Achse und die Homogenität vom Grade 2.



Und abschließend die Restriktion auf den Kreis



Kap. 6.2 Normierte Räume

Grenzwert und Stetigkeit

6.2.1 Vorbemerkung

Wir führen ein Hilfsmittel ein, das uns die bisherige axiomatische Vektorrechnung nicht liefert. Es geht darum, den Begriff der "Länge" oder des "Betrages eines Vektors" zu formalisieren. Das geschieht nach dem allgemeinen Konzept des strukturellen Zugangs zur Mathematik so, dass man zu den in Kapitel 3 besprochenen Vektorraumaxiomen weitere (unabhängige) Axiome hinzunimmt, die diese Formalisierung leisten. Und zwar soll das unmittelbar geschehen, nicht mit Hilfe des Skalarproduktes, das ja mehr als nur die Länge formalisiert. Damit können wir den für uns sehr wichtigen Begriff der **Umgebung eines Punktes** einführen und quantifizieren. Die entstehende Struktur eines normierten Vektorraumes erweist sich generell als nützlich. Eine Winkelbestimmung ist jedoch immer noch nicht möglich.

6.2.2 Normierte Vektorräume

(2.2.1) Wir gehen mit unseren Überlegungen von einem Modellsystem aus, in dem die gewünschte Struktur zur Verfügung steht: dem euklidischen \mathbb{R}_K^3 . Dort wissen wir, wie die Länge eines Vektors zu bestimmen ist. Die Festlegung erfolgt dort mit Hilfe der Betragsabbildung

$$|\cdot| = \left(\mathbb{R}_K^3, \vec{x} = {}^t(x, y, z) \mapsto |\vec{x}| = \sqrt{x^2 + y^2 + z^2} \cdot \mathbb{R} \right).$$

(2.2.2) Genaueres Studium dieser Abbildung erlaubt es, einen Satz von Eigenschaften heraus zu destillieren, die für den Längenbegriff grundlegend sind. Das Ergebnis ist (2.3.3), wobei die anschließenden Folgerungen zeigen, dass tatsächlich so etwas wie die Quintessenz des Längenbegriffs gefunden wurde. Auch die Stetigkeitseigenschaft läßt sich damit erfassen. Links geben wir die Axiome, rechts sind einige unmittelbar naheliegende Kommentare angefügt.

(2.2.3) **Die Normaxiome:**

Sei	V Vektorraum über $K=\mathbb{R}$ oder \mathbb{C} . Zusätzlich zu den Vektorraumaxiomen gelte:	Körper nicht beliebig. $\dim V = \infty$ zugelassen.
(VN.0)	Eine Norm genannte Abbildung ist gegeben: $\ \cdot\ = (V, \vec{x} \mapsto \ \vec{x}\ , [0, \infty[)$	$\ \vec{x}\ $ Länge von \vec{x} nie negativ.
(VN.1)	$\ \vec{x}\ = 0 \iff \vec{x} = \vec{0}$	$\vec{x} \neq \vec{0} \implies \ \vec{x}\ \neq 0$
(VN.2)	<i>Homogenität:</i> $\ \lambda \vec{x}\ = \lambda \ \vec{x}\ $ für alle $\lambda \in K$ und $\vec{x} \in V$.	$ \lambda $ ist nur in \mathbb{R} und \mathbb{C} erklärt. $ 1+i = \sqrt{2}$
(VN.3)	<i>Dreiecksungleichung</i> $\ \vec{x} + \vec{y}\ \leq \ \vec{x}\ + \ \vec{y}\ $ für alle $\vec{x}, \vec{y} \in V$	Verbindungsstrecke ist kürz. Verbindung.
	Ein derartiger Raum heißt normierter Raum	

(2.2.4) Was folgt aus den Axiomen? Wir stellen nachfolgend etwa ein Dutzend unmittelbarer Konsequenzen zusammen, die bereits die Qualität dieser Axiome verdeutlichen. Die Beweise sind meist elementar und werden ausgelassen. **Die Resultate sollte man jedoch beherrschen.**

(2.2.5) Eine Normabbildung erlaubt die folgenden **inhaltlichen Interpretationen:**

$\|\vec{x}\|$ ist die Länge oder der Betrag des Vektors \vec{x} aus V . Und $\|\vec{x} - \vec{y}\|$ ist der (skalare) Abstand der beiden Vektoren \vec{x} und \vec{y} . Werden diese beiden Vektoren als Ortsvektoren zweier Punkte interpretiert, so ist $\|\vec{x} - \vec{y}\|$ auch der Abstand der beiden Punkte.

(2.2.6) Alle Vektoren besitzen einen definierten Abstand voneinander. Verschiedene einen Abstand ungleich Null. Oder: $\|\vec{x} - \vec{y}\| = 0 \implies \vec{x} = \vec{y}$. (Beachten Sie $\|\vec{x}\| = \|\vec{y}\|$ leistet das nicht!)

Der Abstandsbegriff kann in naheliegender Weise ausgedehnt werden auf gewisse Teilmengen von Vektoren entweder im Sinne von "kürzestmöglicher Abstand der Punkte" oder im Sinne von "Schwerpunktsabstand".

(2.2.7) Man kann mit Hilfe einer Norm 2 Vektoren \vec{x} und \vec{y} unterschiedlicher Richtung hinsichtlich ihrer Längen vergleichen, hat aber noch kein Maß, wie unterschiedlich ihre Richtungen sind, also keinen Winkel.

(2.2.8) Alle Abstände sind *translationsinvariant*, d.h. es gilt $\|\vec{x} - \vec{y}\| = \|(\vec{x} - \vec{a}) - (\vec{y} - \vec{a})\|$. Diese Eigenschaft verwendet man gern, um Rechnungen oder Überlegungen auf den Ursprung zu beziehen. Typischerweise wird $\lim_{f(\vec{x})=\vec{y}_0}$ ersetzt durch $\lim_{f(\vec{x})-\vec{y}_0}=\vec{0}$. Fortlassen dieser Eigenschaft der *Translationsinvarianz des Abstandes* führt zu einer anderen mathematischen Struktur, der der *metrischen Räume*.

(2.2.9) **Ähnliche und parallel liegende Figuren** haben wegen (VN.2) proportionale Längen.

(2.2.10) Induktiv folgen aus der Dreiecksungleichung und der Homogenität folgende wichtige **Ungleichungen** (für die kürzeste Verbindung). Summiert wird über endliche Linearkombinationen:

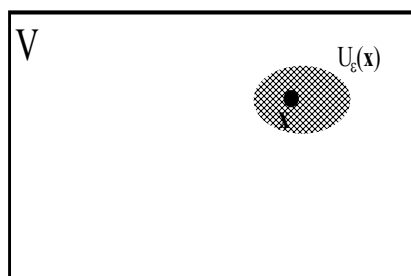
$$\boxed{\|\sum_i \vec{x}_i\| \leq \sum_i \|\vec{x}_i\| \quad \|\sum_i \vec{x}_i \lambda_i\| \leq \sum_i |\lambda_i| \|\vec{x}_i\| \quad \text{Speziell: } \|\vec{x} - \vec{y}\| \leq \|\vec{x}\| + \|\vec{y}\|}$$

- Beweisen Sie $\boxed{\left| \|\vec{x}\| - \|\vec{y}\| \right| \leq \|\vec{x} - \vec{y}\|}$. Welche geometrische Interpretation hat diese Ungleichung? (In den Körpern \mathbb{R} und \mathbb{C} haben wir den Betrag $|\alpha|$, im Körper $\mathbb{Z}/(p)$ nicht!)

(2.2.11) Ein und derselbe Vektorraum kann **mehrere Normen** haben. Ist $\dim V < \infty$, so sind diese Normen jedoch alle **topologisch gleichwertig**. D.h. insbesondere: Existenz und Wert von Grenzwerten hängen nicht von der Normwahl ab. (Hierzu geben wir keinen Beweis.) In Problemsituationen kann man dann immer eine besonders günstige Norm wählen, so wie man ein besonders günstiges Koordinatensystem wählt.

(2.2.12) Ein wichtiges Beschreibungshilfsmittel sind die mit Hilfe einer Norm definierbaren offenen ε -Umgebungen der Punkte von V .

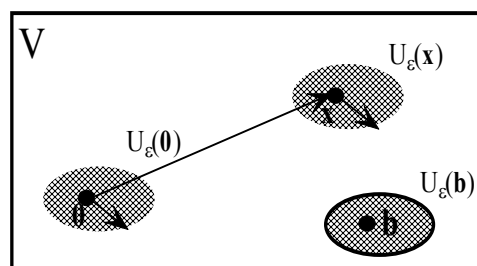
Sei V normierter Raum, $\vec{a} \in V$ und $\varepsilon > 0$. Die Menge $U_\varepsilon(\vec{a}) = \{\vec{x} \mid \vec{x} \in V, \|\vec{x} - \vec{a}\| < \varepsilon\}$ heißt (*offene*) *Kugel vom Radius ε um \vec{a}* oder auch (*offene*) ε -*Umgebung von \vec{a}* .



$S_\varepsilon(\vec{a}) = \{\vec{x} \mid \vec{x} \in V, \|\vec{x} - \vec{a}\| = \varepsilon\}$ mit $\varepsilon > 0$ ist die Oberfläche oder der Rand der Kugel. Das sollte man nicht mit der Kugel selbst verwechseln. Wir bezeichnen $S_\varepsilon(\vec{a})$ auch als *Sphäre*.

Beachten Sie unbedingt: $\varepsilon > 0$ und $\|\vec{x} - \vec{a}\| < \varepsilon$, nicht aber ≥ 0 oder $\leq \varepsilon$! Inhaltlich geometrisch interpretiert enthalten diese Mengen alle Punkte, deren Abstand zum zugehörigen Mittelpunkt \vec{a} geringer als ε ist. Dabei wird Abstand mit Hilfe der Normabbildung gemessen. Die Figuren sollten unbedingt als Gedächtnisstütze verwendet werden.

Infolge der in (2.2.8) besprochenen Translationsinvarianz lassen sich die Umgebungen immer zum Ursprung parallel verschieben. Die "Randpunkte" gehören üblicherweise nicht zur Umgebung ($< \varepsilon$). In $U_\varepsilon(\vec{b})$ der Figur sind sie mitgenommen ($\leq \varepsilon$).



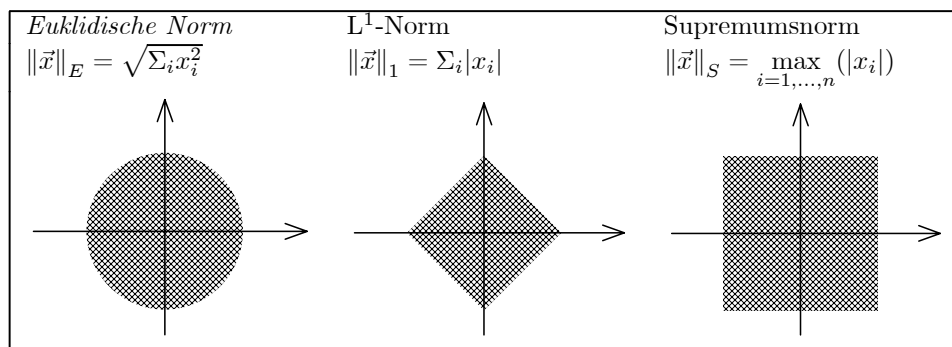
Wie bereits erwähnt, werden Überlegungen zu $U_\varepsilon(\vec{x})$ mit Hilfe der Translationsinvarianz auf solche zu $U_\varepsilon(\vec{0})$ zurückgeführt.

- Welche geometrische Beziehung besteht zwischen den beiden Mengen $U_\varepsilon(\vec{a})$ und allgemeiner $U_{\alpha\varepsilon}(\vec{a})$ für $\alpha > 0$?

(2.2.13) Für $\vec{x} \neq \vec{0}$ ergibt $\boxed{\vec{e} = \frac{\vec{x}}{\|\vec{x}\|}}$ einen "Einheitsvektor derselben Richtung wie \vec{x} ". Die Menge $S_R = S_R(\vec{0}) = \{\vec{x} \mid \|\vec{x}\| = R\}$ und $R > 0$, also die "Kugeloberfläche mit Radius R um den Ursprung" enthält alle

Vektoren der Länge R . Diese Kugeloberfläche trennt über $\|\vec{x}\| < R$ und $\|\vec{x}\| > R$ den gesamten Raum in einen Innen- und einen Außenbereich (relativ zu S_R). (Vgl. die entsprechende Eigenschaft der Hyperebenen aus Kap. 5)

(2.2.14) Die drei **wichtigsten Beispiele** topologisch gleichwertiger Normen für den \mathbb{R}^2 mit zugehörigen ε -Umgebungen (Kugeln) geben wir jetzt an. Die Formeln sind per Summenschreibweise sofort auf den \mathbb{R}^n verallgemeinerbar. Die Prüfung der Gültigkeit der Normaxiome ist jeweils erforderlich, aber weitgehend Routine. Inspizieren Sie auf jeden Fall, wie die Homogenitätsforderung durch die drei Formeln gesichert wird.



(2.2.15) Wie die benutzte Indexschreibweise verdeutlicht, ist die Verallgemeinerung auf n Dimensionen in allen drei Fällen unproblematisch. Der Fall $n=1$ ist auch enthalten. Hier stimmen alle drei Normen mit dem üblichen Betrag der reellen bzw. komplexen Zahlen überein.

- Verifizieren Sie die Normeigenschaften für die drei Abbildungen.
- Wie steht es mit der Verallgemeinerbarkeit (dieser Definitionen) bei abzählbar unendlich vielen Komponenten ("Folgenräume")?
- Wieso ergibt $\|(x, y)\|_{1/2} = (\sqrt{|x|} + \sqrt{|y|})^2$ keine Norm auf \mathbb{R}^2 ?

★ Computerprogramm für das Aussehen der Kugeln im \mathbb{R}^2 für verschiedene Normen.

(2.2.16) "Diese Kugeln sehen doch garnicht wie Kreise aus!" Das ist voreilig argumentiert. Denn sie haben alle die folgende wichtige Eigenschaft der Kugeln:

Jede vom Mittelpunkt ausgehende Halbgerade (Richtung) liegt ein Stück weit ganz in der Kugel, der Rest aber liegt außerhalb. Jede solche Gerade hat seine eigene Längenmessung und die unterschiedlichen Formen entstehen dadurch, dass die Längenmaße z.T. andere sind als das Längenmaß, das unsere Anschauung der Papieroberfläche aufprägt. Und für dieses "Stück in der Umgebung" gibt es die in den Richtungen gleichmäßige untere Schranke ε .

(2.2.17) Da die **Matrixräume** $\text{Mat}_{\mathbb{R}}(m, n)$ als \mathbb{R}^N interpretierbar sind, sind auch das normierte Räume. Für sie und ebenso die Endomorphismenräume gibt es eine weitere wichtige Norm neben den drei bereits angegebenen. Das ist die wie folgt definierte *Operatornorm*:

Sei $A: V \rightarrow V$ linear. $\|\cdot\|$ sei Norm in V
 Dann setzt man $\|A\|_{Op} = \sup_{\|\vec{x}\|=1} \|A\vec{x}\|$

D.h. man wählt alle Einheitsvektoren aus V , berechnet deren Bildvektoren, bestimmt deren Länge und sucht diejenigen heraus, die möglichst groß werden. Im endlichdimensionalen Fall ist das eine stetige Funktion der Koordinaten, so dass die maximale Länge auch angenommen wird. Oder auch: Bilde zu jedem Vektor $\neq \vec{0}$ den Bildvektor $A\vec{x}$ und von beiden die zugehörige Länge. I.a. haben \vec{x} und $A\vec{x}$ sowohl verschiedene Richtung als auch verschiedene Längen. **Die Operatornorm ist dann das größte Längenverhältnis**, das auftritt. Wegen der Linearität von A und der Homogenität der Norm genügt es, nur die Vektoren der Länge 1 zu betrachten. (Die Zuordnung $\vec{x} \mapsto \frac{\|A\vec{x}\|}{\|\vec{x}\|}$ ist homogen vom Grade Null!) Eine Operatornorm mit Wert 3 besagt (jedenfalls im endlichdimensionalen Fall), dass es mindestens einen Vektor $\vec{a} \neq \vec{0}$ gibt mit $\|A\vec{a}\| = 3\|\vec{a}\|$, aber keinen, der noch stärker verlängert wird. Beachten Sie auch, dass die Operatornorm von der Wahl der Norm in V abhängt, was wir in der Schreibweise nicht zum Ausdruck gebracht haben.

(2.2.18) Der Nachweis, dass auch hier die Normaxiome erfüllt sind, ist Routine. Aus der Definition folgt unmittelbar noch die folgende, später sehr wichtige Eigenschaft:

$$\|A\vec{x}\| \leq \|A\|_{Op} \|\vec{x}\| \quad \text{für alle } \vec{x} \in V$$

□ Beweis!

□ Wie groß ist die Operatornorm einer Drehung R in V_0^3 , (mit euklidischer Norm)?

(2.2.19) Bisher haben wir Vektorräume über \mathbb{R} betrachtet. Für den **komplexen Raum** \mathbb{C}^n können wir beispielweise $\sqrt{\sum_i |z_i|^2}$ als Norm nehmen. Sagen wir $\vec{z} = (1+i, 2-i) \in \mathbb{C}^2$. Dann ist $\|\vec{z}\| = \sqrt{2+5} = \sqrt{7}$.

(2.2.20) Ein Beispiel einer Norm für den (reellen) **unendlichdimensionalen Polynomraum** sieht wie folgt aus: $\|p\| = \sqrt{\int_{-1}^1 dx p^2(x)}$. Hierauf werden wir in Kap. 10 genauer eingehen. Jetzt genügt es, wenn Sie sich erneut davon überzeugen, dass die Normaxiome erfüllt sind.

(2.2.21) Wir beschreiben als nächstes eine Norm, die sich für Abbildungs- und Funktionsräume als besonders wichtig erweist. Man kann sie nicht für alle Abbildungen bilden, sondern nur für bestimmte. Wir bezeichnen sie kurz als *Supremumsnorm*.

(2.2.22) Sei M irgendeine Menge und W ein normierter Vektorraum mit Norm $\|\cdot\|$. Weiter sei $f: M \rightarrow W$ eine Abbildung. Wir nennen f *beschränkt*, wenn es eine Zahl $C=C_f$ gibt, so dass $\|f(x)\| \leq C$ für alle $x \in M$ gilt. (Zu verschiedenen f werden i.a. unterschiedliche konstanten C gehören!) Dann sei $\mathfrak{B} = \mathfrak{B}(M, W)$ die Menge aller beschränkten Abbildungen $M \rightarrow W$. Per Wertemengentübertragung wird daraus ein Vektorraum. Dann gilt:

Sei W normiert und $f: M \rightarrow W$ eine beschränkte Abbildung.
 Dann definiert man $\|f\|_\infty$ durch $\|f\|_\infty = \sup_{x \in M} \|f(x)\|$.
 Dann ist $f \mapsto \|f\|_\infty$ eine Norm auf $\mathfrak{B} = \mathfrak{B}(M, W)$

Natürlich müssen auch hier die Normaxiome geprüft werden.

(2.2.23) Beachten Sie, dass im letzten Beispiel der Normwert keineswegs angenommen werden muß. Nehmen wir reelle Funktionen $\mathbb{R} \rightarrow \mathbb{R}$. Dann ist die W -Norm einfach der Betrag. Wir haben $\|\sin\|_\infty = 1$ und dieser Wert wird vielfach angenommen. Für den Arcustangens gilt $\|\arctan\|_\infty = \frac{\pi}{2}$. Aber dieser Wert wird nicht angenommen! Und für \exp ist die Supremumsnorm nicht bildbar, da diese Funktion auf \mathbb{R} nicht beschränkt ist. Schränkt man ein zu $f = ([-T, T], x \mapsto \exp(x), \mathbb{R})$ so folgt $\|f\|_\infty = e^T$.

(2.2.24) Anschaulich ist $\|f\|_\infty$ der größtmögliche Abstand, den die Funktionswerte vom Nullpunkt annehmen können oder den man per Grenzprozeß erreichen kann. Das bedeutet, dass $\|f\|_\infty$ keineswegs ein Maß für den *typischen* Abstand sein muß, einzelne Ausreißer können den Wert verfälschen. Stattdessen ist es ein Abstand, der auf Nummer sicher geht, ein Abstand für den schlimmstmöglichen Fall: Ein größerer Wert kann nicht vorkommen.

□ Verifizieren Sie die Normaxiome für die Supremumsnorm.

(2.2.25) Die Beispiele zeigen bereits, auf wie vielfältige Weise man den Normbegriff realisieren kann. Und das heißt auch, auf wie viele Weisen man auf den unterschiedlichen Räumen einen Längenbegriff einführen kann. Und in jedem Fall sind die Folgerungen aus den Normaxiomen gültig.

6.2.3 Vektorfolgen

(2.3.1) Jetzt können wir daran gehen, den Grenzwertbegriff (zunächst für Folgen) zu formalisieren. D.h. Fälle zu erfassen, in denen im Rahmen der Vektorrechnung der Übergang von endlich zu unendlich vielen Schritten vorgenommen wird.

(2.3.2) Sei also V ein normierter Vektorraum und $a = (\mathbb{N}, n \mapsto a_n, V)$ eine Folge von Vektoren. Wie üblich interessiert das Verhalten dieser Abbildung. Und ein spezieller, aber besonders wichtiger Aspekt ist die Frage: **Wie verhält sich a , wenn n immer weiter wächst?** Zur Veranschaulichung eignet sich der Feldstandpunkt, wobei der mit der Rolle des Konfigurationsraumes versehene Urbildraum diskret, aber abzählbar unendlich ist.

(2.3.2) Das typische Szenenbild, das zu einer derartigen Folge führt, sieht so aus, dass wir in gleichen (zeitlichen oder räumlichen) Abständen t_n eine bestimmte vektorielle Größe - etwa eine Kraft- messen oder berechnen. Wir stellen uns vor, dass das immer wieder geschieht und fragen, ob die Folge der Ergebnisse einem Grenzwert zustrebt. Zu beachten ist, dass dieser Grenzwert nach endlich vielen Schritten nicht erreicht

sein muß, sondern man sich ihm nur beliebig genähert haben wird, wobei der Normbegriff benutzt wird, um diese Nähe zu beschreiben. Oder anders herum: Es ist möglich, dass dieser Grenzwert selbst nie erreicht wird.

(2.3.3) **Wann wird man einer Vektorfolge $a=(N, n \mapsto a_n, V)$, um deren Verhalten es geht, einen Grenzwert, der wieder ein Vektor aus V ist, zuordnen?** Verdeutlichen wir uns das Problem:

Endlicher Fall: Der gesuchte Vektor ist der letzte in einer endlichen Folge von Vektoren. (Der 27. Rechenschrift ergibt $a=3.1417$, danach ändert sich nichts mehr. Das letzte Element ist der Grenzwert.)

Unendlicher Fall: Hier gibt es keinen "letzten" Vektor. Wie legt man stattdessen einen Vektor fest, der die Rolle des letzten Vektors übernimmt?

IDEA: "Die fünfte Nachkommastelle ändert sich nach dem 49. Schritt nicht mehr, die 10. nach dem 500. Schritt usw. " Diese Formulierung ist auszubauen.

(2.3.4) Intuitiv sollen sich die Folgenglieder einem Vektor $g \in V$ beliebig nähern, ohne dass sie diesen Vektor im Grenzfall ganz erreichen müssen (sie können das, müssen aber nicht!). Wie ist dieses "sich Nähern" zu formalisieren? Wir bilden die Abstandsvektoren $a_n - g$ und die zugehörigen Abstände $\|a_n - g\|$. Wegen der Translationinvarianz der Norm ist a_n von g ebensoweit entfernt wie der Vektor $(a_n - g)$ vom Nullvektor! Wir wechseln zur Zahlfolge $(N, n \mapsto \|a_n - g\|, V)$. Hierbei kann und sollte $g \in V$ als äußerer Parameter angesehen werden.

(2.3.5) Wann ist g **kein Grenzwert** der (ursprünglichen) Folge? Dann wird der Abstand $\|a_n - g\|$ nicht Null werden. Wir stellen alle Zahlen $\|a_n - g\|$ als Dezimalbrüche dar. Nicht Null werden bedeutet offenbar, dass es irgendeine Nachkommastelle gibt, die auch nach noch so vielen Schritten n nicht auf Dauer (= mit immer weiter fortschreitendem n) Null wird. **Umgekehrt ist g Grenzwert**, wenn es für jede Nachkommastelle einen n -Wert gibt, von dem ab diese Nachkommastelle immer Null ist.

Beachten sie die Reihenfolge: Zuerst ist eine zu testende Nachkommastelle vorzugeben. Dazu zu bestimmen ist dann ein Indexwert n , der für die vorgegebene Stelle etwas sicherstellt.

(2.3.6) In leicht geänderter Formulierung können wir die Situation als Zweipersonenspiel interpretieren:

- Ein Herausforderer nennt eine Nachkommastelle.
- Und ein Verteidiger muß einen Folgenindex N nennen, von dem ab in der Abstandsfolge $\|a_n - g\|$ bis zu dieser Stelle mit Sicherheit nur noch Nullen auftreten.

Die Interpretation realisiert den folgenden wichtigen Aspekt des Überganges von endlich vielen Operationen zu unendlich vielen: Im endlichen Fall kann man alle Operationen ausführen, im unendlichen hat man nur die Möglichkeit, aus einer bestimmten Operationsmenge eine beliebige Operation auszuwählen und diese auszuführen. Man kann die Gesamtheit aller Möglichkeiten nur potentiell erfassen, man führt sie aber nicht alle aus! Und diesem Sachverhalt muß die Verhaltens- und Vorgehensbeschreibung angepaßt werden. Man realisiert das dadurch, **daß man dem Herausforderer erlaubt, seine Wahl nach Belieben zu treffen.**

(2.3.7) Was benötigt der Verteidiger (der Aussage, dass der gewählte Vektor g Grenzwert der Folge sei), um zu gewinnen? Er muß auf jede mögliche Herausforderung eine sachlich korrekte Antwort geben. D.h. aber, er benötigt eine Strategiefunktion der folgenden Art:

Kommastelle $t \mapsto$ ausreichender Folgenindex.

Nun ist die "Stellenzahl z " kein besonders schöner Parameter besser ist $\varepsilon = 10^{-z}$, also der Zahlwert, durch den die Stellenzahl bestimmt ist.

(2.3.8) Das Spiel kann daher wie folgt formuliert werden:

Der Herausforderer nennt ein $\varepsilon > 0$, also eine Genauigkeitsschranke, und der Verteidiger (des behaupteten Grenzwertes) muß mit einem $N(\varepsilon)$ kontern, so dass mit der geforderten Genauigkeit Folgenglieder und Grenzwert ab diesem N nicht mehr zu unterscheiden sind!

(2.3.9) Wir nennen eine Funktion $\varepsilon \mapsto N(\varepsilon)$ mit deren Hilfe der Verteidiger mit Sicherheit gewinnt, eine *Strategiefunktion (der Folge a für den Grenzwert g)*.

(2.3.10) Kann man eine solche Strategiefunktion angeben, **ist g Grenzwert** der Folge a. Gibt es eine solche Strategiefunktion nicht, ist g auch nicht Grenzwert. Gibt es eine Strategiefunktion, kann der Verteidiger damit immer gewinnen, gleichgültig, was der Herausforderer versucht. Gibt es eine solche Funktion nicht, kann der Herausforderer mit Geschick gewinnen.

(2.3.11) **Die Festlegung oder Angabe einer Strategiefunktion kann durchaus mit endlich vielen Schritten erfolgen! Und damit beherrschen wir die unendlich vielen Möglichkeiten des Herausforderers bzw. eigentlich der Folge.**

(2.3.12) **Beispiel:** Sei $n \mapsto \vec{a}_n = (\frac{1}{n}, \frac{2}{n^2}, \frac{2}{3n}) \in \mathbb{R}_K^3$. und $n \geq 1$. Diese Folge hat den Grenzwert $\vec{0}$. Wir müssen gemäß Rezept bilden:

$$n \mapsto \|\vec{a}_n - \vec{0}\| = \sqrt{\frac{1}{n^2} + \frac{4}{n^4} + \frac{4}{9n^2}} = \frac{1}{n} \sqrt{1 + \frac{4}{n^2} + \frac{4}{9}} < \frac{\sqrt{3}}{n} < \frac{\sqrt{3}}{N} \quad \text{mit } N \leq n.$$

Und das soll kleiner als ε sein! D.h. $\frac{9}{N} < \varepsilon$ soll gelten. Als Strategiefunktion können wir $N(\varepsilon) = \frac{9}{\varepsilon}$ wählen. Gleichgültig, welche Genauigkeit der Herausforderer verlangt, mit dieser Strategiefunktion gewinnt wegen der angegebenen Abschätzung der Verteidiger. (Da n darin äußerer Parameter ist, erfasst die eine Rechnung unendlich viele Fälle!)

(2.3.13) Das Beispiel zeigt bereits, wie man üblicherweise vorgeht, um eine Strategiefunktion zu gewinnen. Schematisieren wir zunächst das allgemeine Vorgehen:

Vorgehen bei der Konvergenzdiskussion einer Folge:

Sei $(\mathbb{N}, n \mapsto a_n, V)$ Vektorfolge. Ist $g \in V$ Grenzwert der Folge?
(1) Bilde neue Vektorfolge $n \mapsto a_n - g$. Eventueller Grenzwert ist $0 \in V$.
(2) Bilde Zahlfolge $n \mapsto \ a_n - g\ $. Eventueller Grenzwert ist $0 \in \mathbb{R}$
(3) Suche eine Strategiefunktion $\varepsilon \mapsto N(\varepsilon)$ für $\varepsilon > 0$ mit der Eigenschaft: $\ a_n - g\ < \varepsilon$ für $n > N(\varepsilon)$.

(2.3.14) Und jetzt das übliche Vorgehen bei der Bestimmung einer Strategiefunktion:

Rezept	zur Bestimmung einer Strategiefunktion
(3a)	Man schätzt $\ a_n - g\ $ vorsichtig nach oben ab. $\ a_n - g\ \leq \dots \leq f(n)$ n äußerer Parameter
(3b)	Vergrößert weiter zu $f(n) \leq g(N)$ für alle $n \geq N$ in der Regel über Monotonie
(3c)	Nun formuliert man die Ungleichung $g(N) < \varepsilon$ mit Rollenwechsel, so dass N Unbestimmte, $\varepsilon > 0$ äußerer Parameter. Suche $N(\varepsilon)$, so dass $g(N) < \varepsilon$ für $\varepsilon > 0$.
(3d)	Die Ungleichung versucht man zu lösen, entweder durch Umstellen nach N oder Raten einer Lösung. Gelingt das, ist $\varepsilon \mapsto N(\varepsilon)$ eine Strategiefunktion.

(2.3.15) Das Rezept läßt sich verwenden, um zahlreiche konkrete Beispiele zu diskutieren. aber auch, um die üblichen Grenzwertsätze zu beweisen.

(2.3.16) Wir geben nur ein Beispiel eines solchen Grenzwertsatzes:

- Die Zahlfolge a habe den Grenzwert $g > 0$ und es gelte $a_n > 0$ für alle n. Dann hat die Folge $n \mapsto \frac{1}{a_n}$ den Grenzwert $\frac{1}{g}$.

(2.3.17) **Beweis:** Bei einer Zahlfolge ist die Norm gleich dem Betrag. Sei $N_a(\varepsilon)$ eine zu a gehörige Strategiefunktion, die es bei Konvergenz ja gibt!. Dann gilt $|a_n - g| < g/2$ für $n \geq N_0 = N_a(\frac{g}{2})$. Also ist $a_n > \frac{g}{2} > 0$ für $n \geq N_a(\frac{g}{2})$. Jetzt benötigen wir eine Strategiefunktion für b. Durch Verkleinern des Nenners folgt für $n \geq N$:

$$\left| b_n - \frac{1}{g} \right| = \left| \frac{1}{a_n} - \frac{1}{g} \right| = \left| \frac{g - a_n}{ga_n} \right| < \frac{|g - a_n|}{g^2/2} < \varepsilon$$

Die Ungleichung wird erfüllt durch $N_b(\varepsilon) = \max(N_a(\varepsilon g^2/2), N_0)$. Also ist dies N_b eine Strategiefunktion dafür, dass die Folge b den Grenzwert $1/g$ hat.

□ Wie lauten die weiteren üblicherweise benötigten Grenzwertsätze? Beweisen Sie davon einige.

(2.3.18) Bei der Konstruktion von Strategiefunktionen sollte man beachten, dass keineswegs ein "bestmögliches" Ergebnis gefordert ist, sondern nur **irgendeine** Funktion mit den gesuchten Eigenschaften. Die betrachteten Nachkommastellen müssen ab $N(\varepsilon)$ Null sein, sie dürfen es aber auch schon vorher sein. Ist also etwa $N(\varepsilon)$ eine Strategiefunktion und gilt $M(\varepsilon) \geq N(\varepsilon)$, so ist auch $M(\varepsilon)$ erst recht eine solche. Unverlangte Suche nach nach einer besonders guten Abschätzung kann viel unnötige Arbeit machen.

(2.3.19) Beispiel: Wir betrachten die Zahlfolge $n \mapsto a_n = \frac{3 \log(n)}{2 + \log(n)}$. Wir vermuten $g=3$ als Grenzwert. Wir bilden $a_n - 3 = \frac{-6}{2 + \log(n)}$. Normbildung -hier wieder der Betrag- und vorsichtiges Vergrößern gemäß (2.2.14):

$$|a_n - 3| = \frac{6}{2 + \log(n)} < \frac{6}{\log(n)} < \frac{6}{\log(N)} \quad \text{Nach (3c): } \frac{6}{\log(N)} < \varepsilon \quad \text{gibt } 10^{\frac{6}{\varepsilon}} < n$$

Dabei folgt $\frac{6}{\log(n)} < \frac{6}{\log(N)}$ sofort über die Monotonie der Logarithmusfunktion.

Setze $N(\varepsilon) = \lceil 10^{\frac{6}{\varepsilon}} \rceil =$ "größte ganze Zahl $\leq 10^{\frac{6}{\varepsilon}}$ ". Dann ist $\frac{6}{\log(N)} < \varepsilon$ für alle $n \geq N(\varepsilon)$ wie gefordert.

Demnach ist $N(\varepsilon) = \lceil 10^{\frac{6}{\varepsilon}} \rceil$ Strategiefunktion, hier sogar die bestmögliche. Beachten Sie, wie stark N wächst, wenn ε kleiner wird, also mehr korrekte Stellen verlangt werden. Zur graphischen Veranschulichung mit Hilfe des Computers ist es meist gut, $\log(N)$ gegen $z = -\log(\varepsilon)$ aufzutragen.

□ Weitere Beispiele:

(2.3.20) Das bisherige Vorgehen enthält ein großes Defizit: Entweder muß man den Grenzwert bereits kennen, um das Verfahren durchziehen zu können, oder man muß g als äußeren Parameter ansehen und dann das Verfahren für viele g testen.

(2.3.21) Besser wäre es, mit den Folgegliedern allein zu arbeiten zu können. Also ihnen bereits anzusehen, ob die Folge einen Grenzwert hat oder nicht, ohne dass man diesen Grenzwert bereits kennen muß. Was hätten wir in (2.3.17) und (2.3.19) tun sollen, wenn wir keinen Kandidaten für den jeweiligen Grenzwert gehabt hätten?

(2.3.22) Die folgende Definition leistet das Gewünschte mit Hilfe einer nun naheliegenden, aber wie sich zeigt, genialen Verallgemeinerung obiger Überlegungen:

(2.3.23) **Definition Cauchyfolge:**

Sei V normierter Raum und $n \mapsto a_n$ eine Vektorfolge.
 Es gebe ein Strategiefunktion $\varepsilon \mapsto N(\varepsilon)$ für $\varepsilon > 0$ der folgenden Art: $\|a_n - a_m\| < \varepsilon$ für alle $n, m \geq N(\varepsilon)$
 Dann heißt a eine *Cauchyfolge in a* .

(2.3.24) D.h. oberhalb von $N(\varepsilon)$ unterscheiden sich alle Folgenglieder höchstens um ε voneinander (anstelle vom Grenzwert). In Kommastellen stimmen sie ab $N(\varepsilon)$ sicher bis zur Stelle $z = \log(\varepsilon)$ überein. Oder: Bis zu dieser Stelle haben sich die Meßresultate mit Sicherheit stabilisiert, wenn man im Konfigurationsraum \mathbb{N} weiter voranschreitet.

(2.3.25) **Jede konvergente Folge mit Grenzwert g ist Cauchyfolge.**

Beweis: Es sei $\varepsilon \mapsto N_g(\varepsilon)$ eine Strategiefunktion für den Grenzwert g . Also $\|a_n - g\| < \varepsilon$ für $n > N_g(\varepsilon)$. Wenn der Herausforderer $\varepsilon/2$ fordert, dann ist $n > N_g(\frac{\varepsilon}{2})$ zu wählen, wie man weiß. Also $\|a_n - g\| < \frac{\varepsilon}{2}$ für $n > N_g(\frac{\varepsilon}{2})$. Über die Dreiecksungleichung der Norm folgt aber

$$\|a_n - a_m\| = \|(a_n - g) - (a_m - g)\| \leq \|a_n - g\| + \|a_m - g\| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

Damit ist $N(\varepsilon) = N_g(\frac{\varepsilon}{2})$ eine (gesuchte) Strategiefunktion für die Cauchyfolgeeigenschaft.

□ Die soeben benutzte "Teleskopmethode" ist sehr fruchtbar. Beweisen Sie analog, dass eine Folge (im normierten Raum) höchstens einen Grenzwert haben kann.

(2.3.26) **Hat umgekehrt eine Cauchyfolge immer einen Grenzwert?** Diese Frage darf man nicht voreilig mit ja beantworten. Die Antwort hängt nämlich vom gewählten Raum ab! Nehmen wir eine Folge rationaler Zahlen, die gegen die irrationale Zahl $\sqrt{2}$ konvergiert. Dann ist das eine Cauchyfolge, da sich die Stellen gegen die von $\sqrt{2}$ stabilisieren. Wählt man als Raum \mathbb{Q}^1 über \mathbb{Q} , mit der üblichen Abstandsnorm, so hat die Folge keinen Grenzwert in \mathbb{Q}^1 . Denn $\sqrt{2}$ ist ja nicht rational.

(2.3.27) \mathbb{R} selbst entsteht in der üblichen Konstruktion aus der Menge aller Cauchyfolgen rationaler Zahlen über die Äquivalenzrelation: "Die Cauchyfolgen a und b sind äquivalent, wenn $a-b$ Nullfolge ist".

(2.3.28) Und in \mathbb{R} gilt dann (letztlich infolge dieser Konstruktion): **Jede Cauchyfolge reeller Zahlen hat genau einen Grenzwert in Form einer reellen Zahl.**

(2.3.29) Diese Eigenschaft überträgt man auf Vektorräume:

Sei V normierter Vektorraum über \mathbb{R} oder \mathbb{C} .
 Dann heißt V *Banachraum*, falls jede Cauchyfolge aus V einen Grenzwert in V hat.

(2.3.30) \mathbb{R}^n und \mathbb{C}^n mit den zugehörigen Normen sind Banachräume. (Ohne Beweis hier! Man benötigt die Vollständigkeit von \mathbb{R}). Der Polynomraum \mathfrak{P} mit der in (2.2.20) gegebenen Norm ist nicht vollständig.

□ Begründen Sie die letzte Behauptung.

□ Zur Verdeutlichung des Problems: E seien a und g zwei positive Zahlen. Setzen Sie $a_0 = a$ und $g_0 = g$ sowie rekursiv $a_{n+1} = \frac{1}{2}(a_n + g_n)$ und $g_{n+1} = \sqrt{a_n g_n}$. Damit sei $\vec{m}_n = (a_n, g_n)$. Ist die Folge $n \mapsto \vec{m}_n$ konvergent? Was können sie darüber herausfinden, was beweisen?

(2.3.31) **Jeder normierte Raum läßt sich zu einem Banachraum vergrößern** (" vervollständigen", so wie \mathbb{Q} zu \mathbb{R} erweitert wurde. Ohne Beweis hier.)

(2.3.32) Wir werden meist mit Banachräumen arbeiten, so dass die definierende Eigenschaft benutzt werden kann: **D.h. eine Cauchyfolge hat im Falle eines Banachraumes immer einen Grenzwert in V** . Einige zugehörige Schreibweisen:

Ist $a : n \mapsto a_n$ Vektorfolge aus V und $g \in V$ Grenzwert dieser Folge, so schreiben wir für diesen Sachverhalt symbolisch: $g = \lim_{n \rightarrow \infty} a_n = g$ oder $g = \lim_n a_n$.
 Und wir sagen: *Die Folge a ist konvergent mit Grenzwert g .*

(2.3.33) Zu Beginn von Kap.1.2 wurde gesagt: Alles Wichtige ist Beziehung und Beziehungen sind vornehmlich Zuordnungen. Die Grenzwertkonstruktion ist wichtig, stellt sie doch ein geistiges Instrument dar, mit dessen Hilfe man sich Zugang zu einer Welt verschafft, in der man Erfahrungen über unendliche Operationsabläufe sammeln kann. Eine Welt, die uns physisch völlig unzugänglich ist. Liegt also hier eine Zuordnung, eine Abbildung vor? Ja, und zwar eine vom Typ der Funktionale. Einer Abbildung $a: \mathbb{N} \rightarrow V$ wird ein Element aus V zugeordnet. Sei also $\mathfrak{K} = \mathfrak{K}(\mathbb{N}, V)$ die Menge aller konvergenten Folgen $\mathbb{N} \rightarrow V$. Dann haben wir es mit der folgenden Abbildung zu tun:

$$\left(\mathfrak{K}(\mathbb{N}, V), a \mapsto g = \lim_n a_n, V \right)$$

Der Wert g gibt eine wichtige Verhaltenseigenschaft der Folge a wieder. Das Zuordnungsverfahren die Bestimmung des Grenzwertes kann sehr schwer zu realisieren sein. Die Zuordnung selbst dagegen ist wieder einfach.

6.2.3a Zwei Beispiele zur Bedeutung der Grenzwertkonstruktion

(2.3.34) Es folgen zwei Beispiele, die die Bedeutung der Grenzwertkonstruktion im Rahmen konkreter Problemsituationen verdeutlichen.

(2.3.35) Erstes Beispiel: **Der Differentialquotient.**

Sei $n \mapsto a_n$ eine *Nullfolge von Zahlen*, d.h. eine Folge mit Grenzwert Null. Etwa $a_n = \frac{1}{n}$ für $n > 0$ oder $a_n = 10^{-n}$. Für alle n sei $a_n \neq 0$. Die zweite Folge $n \mapsto 10^{-n}$ strebt schneller nach Null und bildet für uns das anschaulich bessere Beispiel.

Weiter die $t \mapsto f(t)$ eine glatte reelle Funktion, etwa $f(t) = e^{-t}$ oder $f(t) = \frac{1}{1+t^2}$. Wir wählen einen festen Aufpunkt t_0 (äußerer Parameter) und bilden damit die beiden neuen Folgen $n \mapsto \Delta t_n = a_n$ und $\Delta y_n = (f(t_0 + \Delta t_n) - f(t_0))$. Die erste Folge ist einfach nur umbenannt. Beide Folgen zeigen ein einfaches Verhalten, sind Nullfolgen! Aber nicht dieses Verhalten interessiert, sondern das der weiteren Folge $n \mapsto q_n = \frac{\Delta y_n}{\Delta t_n}$. Da alle $a_n \neq 0$ waren, sind alle q_n bildbar.

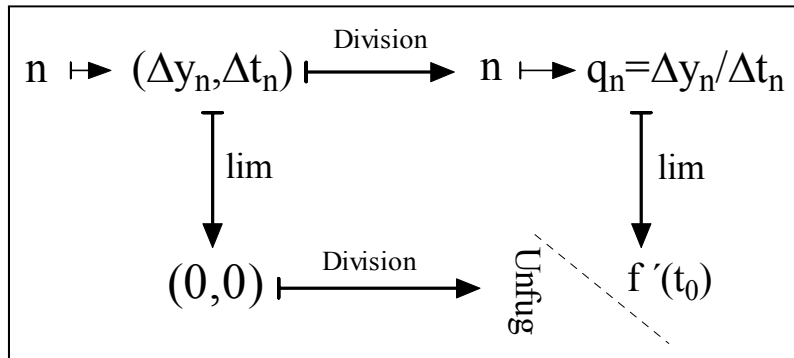
Welches Verhalten zeigt nun die Folge q ? **In der Regel liegt keine Nullfolge mehr vor.** Ist f in der Nähe von t_0 ausreichend glatt, dann strebt die Quotientenfolge einem Grenzwert $f'(t_0)$ zu, dem

Ableitungswert von f in t_0 . Für das Beispiel $a_n=10^{-n}$ und $f(t)=\frac{1}{1+t^2}$ läßt sich der Grenzwert leicht bestimmen. Der Grenzwert selbst wird typischerweise nie angenommen, muß aus dem Verhalten der unendlich vielen Folgeglieder erschlossen werden, so dass insgesamt eine Zuordnung vom Funktionstyp entsteht. Also (2.3.33). Und beachten Sie: Dieser "Grenzwert der Quotienten" ist keineswegs gleich dem "Quotienten der Grenzwerte", das wäre ja der undefinierte Ausdruck $\frac{0}{0}$.

Wählt man $f(t)=t\sin(\frac{1}{t})$ und $t_0 = 0$, so wird $q_n = \sin(\frac{1}{\Delta t_n})$. Hier ist f nicht ausreichend glatt und die Folge der q_n springt unkontrollierbar zwischen -1 und $+1$ herum. Sie strebt keinem Grenzwert zu!

- Berechnen Sie den Grenzwert der q , für $f(t)=\frac{1}{1+t^2}$. Zeichnen sie mit Hilfe einesm Computeralgebraprogramms das typische Verhalten von q , für $f(t)=t\sin(\frac{1}{t})$.
- Was ändert sich, wenn man statt einer reellen Funktion $t \mapsto f(t)$ eine vektorwertige Bahnkurve $t \mapsto \vec{f}(t)$ nimmt. Gehen Sie damit (2.3.35) erneut durch.

(2.3.36) Die gesamte Konstruktion läßt sich gut als Verlaufsdiagramm darstellen.



Im Gegensatz zu den typischen Diagrammen der Algebra (aus Kap.3-5) ist dieses Diagramm **nicht kommutativ**. Der zweite Weg bietet sich an, erscheint attraktiv einfach, führt aber in die Irre. Es gibt keinerlei Grund, weshalb ein derartiges Diagramm kommutativ sein sollte. Historisch gesehen hat die unreflektierte Annahme, ein solches Diagramm müsse kommutativ sein, zu viel Verwirrung geführt.

(2.3.37) **Zweites Beispiel: Die Krümmung.**

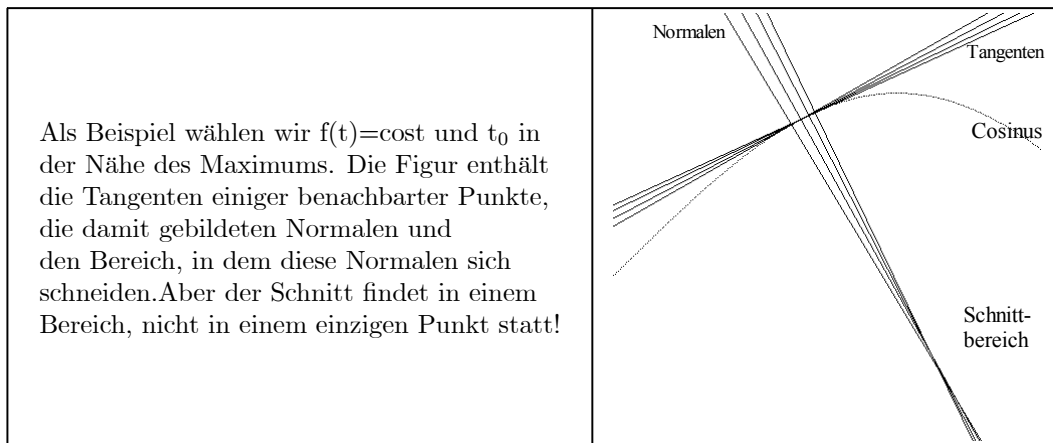
(2.3.38) Jetzt sei $r = (\mathbb{R} \mapsto \vec{r}(t).V_0^2)$ eine glatte Bahnkurve in der Ebene. t_0 sei wieder äußerer Parameter. Wir machen noch zwei Voraussetzungen:

- (*) In einer Umgebung von t_0 sei die Geschwindigkeit ungleich Null und
- (**) für jedes t_1 aus dieser Umgebung habe $\dot{\vec{r}}(t_1)$ eine andere Richtung als $\dot{\vec{r}}(t_0)$.

Weiter sei $n \mapsto a_n$, eine Nullfolge wie im vorausgegangenen Beispiel (2.3.35). Etwa $a_n=10^{-n}$. Mit $a_n \neq 0$ für alle n . Nun bilden wir für $t_n=t_0+a_n$ zunächst die Tangente an Bild \vec{r} zum Punkte $\vec{r}(t_n)$ und dann die zugehörige Normale, die wir mit $N(t_n)$ bezeichnen. Denken Sie daran: Tangente und Normale sind beides Geraden. $N(t_n)$ ist die Normale, die zu $t_n=t_0 + a_n$, eigentlich zu $\vec{r}(t_n)$ gehört.

Damit kommen wir zu der eigentlichen Konstruktion: **Wir bilden den Schnitt der Normalen** $N(t_n) = N(t_0 + a_n)$ mit der normalen $N(t_0)$. Infolge unserer Bedingung (*) und (**) ist dieser Schnitt (zweier Geraden) immer ein einzelner Punkt, den wir mit S_n bezeichnen der zugehörige Ortsvektor sei \vec{s}_n .

Uns interessiert die Vektorfolge $n \mapsto \vec{s}_n$ und deren Grenzverhalten.



Die konkrete Berechnung der \vec{s}_n , ist weitgehend unproblematisch, ein einfaches Schnittproblem zweier Geraden. Die Folge der Schnittpunkte wird i.a. einen Grenzwert haben und das ist **der Mittelpunkt des Krümmungskreises an Bild \vec{r} zum Punkte $\vec{r}(t_0)$** . Also der Kreis, der Bild \vec{r} lokal am besten approximiert. So wie die Tangente die bestapproximierende Gerade ist. Damit wird auch die Bedeutung der Voraussetzung (***) verstehen. Bei geradliniger Bewegung haben die Normalen keinen Schnitt. Bei Wendepunkten existiert der Grenzwert nicht. Der Schnitt wandert nach unendlich.

- Berechnen Sie die Schnittpunktfolge und bestimmen Sie den Grenzwert. Wie in (1.6.24) ein festes Koordinatensystem verwenden! Leiten Sie eine Formel für den Radius des Krümmungskreises her.

(2.3.39) Wie im ersten Beispiel kommt es auf die Reihenfolge des Vorgehens an. Bildet man erst den Grenzwert in n , dann erhält man zweimal dieselbe Normale $N(t_0)$. Der Schnitt ist die gesamte Gerade. Während im ersten Beispiel der falsche Weg zu überhaupt keiner Lösung führte, erhalten wir hier zu viele.

6.2.3b Die Konvergenz von Abbildungsfolgen unter der Supremumsnorm

(2.3.40) Sei W normierter Vektorraum mit Norm $\|\cdot\|_W$ und M eine Menge. Dann haben wir für beschränkte Abbildungen $M \rightarrow W$ in (2.2.22) die Supremumsnorm eingeführt. Der Raum $\mathfrak{B} = \mathfrak{B}(M, W)$ dieser Abbildungen wird per Wertemengenübertragung ein Vektorraum. Die Elemente von \mathfrak{B} können daher immer in zwei Rollen auftreten: Entweder als Abbildungen oder als Vektoren eines normierten Raumes! Im Rahmen der Vektorrolle können wir Folgen (Vektoren) bilden und Grenzwertbetrachtungen anstellen.

(2.3.41) "Eine Folge f_n beschränkter Abbildungen $M \rightarrow W$ konvergiert gegen eine Grenzfunktion $f : M \rightarrow W$ " besagt daher: Die Zahlfolge $\|f_n - f\|_S$ konvergiert nach Null. Und das heißt laut Definition der Supremumsnorm, dass die Zahlfolge $n \mapsto \sup_{x \in M} \|f_n(x) - f(x)\|_W$ eine Nullfolge bildet. Für jeden Punkt $x \in M$ haben wir die Vektorfolge der Funktionswerte $n \mapsto f_n(x)$. Wir erwarten, dass diese Folge gegen das Element $f(x)$ aus W konvergiert. Tatsächlich gilt ja für jedes $y \in M$ die Beziehung

$$\|f_n(y) - f(y)\|_W \leq \sup_{x \in M} \|f_n(x) - f(x)\|_W = \|f_n - f\|_S$$

Und das heißt, dass für jedes $y \in M$ die Folge $n \mapsto f_n(y)$ gegen $f(y)$ konvergiert, wobei man immer - für jedes y - dieselbe Strategiefunktion wählen kann. Nämlich eine, die die für die vorausgesetzte Konvergenz gegen g in der Supremumsnorm existieren muß. Man nennt dieses Verhalten suggestiv **gleichmäßige Konvergenz der Funktionen gegen f** .

Eine Alternative des Verhaltens solch einer Funktionsfolge ist, dass man zwar für jeden Punkt Konvergenz hat, dass es aber nicht möglich ist, eine **gemeinsame Strategiefunktion** anzugeben. Jeder Punkt benötigt dann seine eigene maßgeschneiderte Strategiefunktion. Nehmen wir die Funktionen $f_n = ([0, 1], x \mapsto x^n, \mathbb{R})$. Diese Funktionsfolge konvergiert in jedem Punkt gegen die (unstetige) Grenzfunktion g mit $g(x)=0$ für $0 \leq x < 1$ und $g(1)=1$. Man ist jedoch nicht in der Lage eine gemeinsame Strategiefunktion für alle x zu finden, da es immer noch kurz vor 1 liegende Punkte gibt, die jeweils langsamer konvergieren.

Oder auch: bezüglich der Supremumsnorm bildet dies Funktionsfolge keine Cauchyfolge. Eine derartige Konvergenz nennt man punktweise Konvergenz. Aus gleichmäßiger Konvergenz folgt, wie wir gesehen haben, punktweise. Aber aus punktweise muß nicht gleichmäßig folgen.

- Zeigen Sie, dass die eingeführten Funktionen $f_n = ([0, 1], x \mapsto x^n, \mathbb{R})$ keine Cauchyfolge bezüglich der Supremumsnorm bilden!

(2.3.43) Veranschaulichen wir uns den Unterschied mit Hilfe des Zweipersonenspiels. Wie ist das für Funktionsfolgen auszugestalten? Der Herausforderer gibt einen Punkt $a \in M$ vor und nennt ein $\varepsilon > 0$. Im Falle punktweiser Konvergenz muß der Verteidiger sich gleichsam nach a begeben, dort eine spezifische Strategiefunktion konstruieren und damit reagieren. Im Falle gleichmäßiger Konvergenz dagegen besitzt er eine für alle a gültige Funktion. Um die spezifischen Verhältnisse in a muß er sich nicht kümmern.

Nehmen wir als Beispiel s_n , mit $s_n(x) = \frac{1}{n} \sin(nx)$. Diese Folge konvergiert gleichmäßig gegen die Nullfunktion auf ganz \mathbb{R} . Denn wir haben $\|s_n - 0\|_S = \frac{1}{n}$. Daher können wir $N(\varepsilon) = \frac{2}{\varepsilon}$ wählen. Und das ist eine Strategiefunktion für $n \mapsto s_n$ für jedes a .

- Zeichnen Sie die beiden Funktionsfolgen f_n und s_n geeignet mit Hilfe eines Computeralgebrasystems. Wie stellt sich dabei das Vorhandensein einer gemeinsamen Strategiefunktion dar?

(2.3.44) Die Überlegungen aus (2.3.41) legen eine Reihe von Anschlussfragen nahe: Was ist, wenn man man eine Cauchyfolge $n \mapsto f_n$ von Abbildungen aus $\mathfrak{B}(M, W)$ für die Supremumsnorm hat? Konvergiert die gegen eine Grenzfunktion? Wenn ja, dann sollte $\mathfrak{B}(M, W)$ vollständig sein, also ein Banachraum. Das ist tatsächlich der Fall, wenn W auch Banachraum ist. **Ist W Banachraum, dann ist der Raum $\mathfrak{B}(M, W)$ auch Banachraum.** Insbesondere ist das immer der Fall, wenn W endlichdimensional ist.

- Beweisen Sie analog zur Rechnung aus (2.3.41), dass f_n in der Supremumsnorm gegen eine Grenzfunktion f konvergiert. Jetzt muss man noch eine Sache zusätzlich beweisen, um zu zeigen, dass ein Banachraum vorliegt. Was nämlich? Tun Sie das.

6.2.4 Stetigkeit von Abbildungen

(2.4.1) Der - neben der Konvergenz von Folgen - zweite wichtige Begriff, den wir mit Hilfe des Normbegriffs einführen, ist der der Stetigkeit (einer Abbildung in einem Punkt). Hierbei handelt es sich um eine erste Präzisierung und Verallgemeinerung des intuitiven Glattheitsbegriff für Funktionen. Weitere werden folgen.

(2.4.2) Genauer haben wir es immer mit den folgenden Zutaten zu tun:

$f: G \rightarrow W$	mit $G \subset V$ und V, W normierte Vektorräume
$x_0 \in G$	Der "Aufpunkt". Dann ist f im Punkte x_0 in zu präzisierender Weise entweder stetig oder unstetig.
$\ \cdot\ _V$ und $\ \cdot\ _W$	V und W können gleich oder verschieden sein. Ihre Normen werden bei Bedarf mit $\ \cdot\ _V$ und $\ \cdot\ _W$ bezeichnet.

Diese Zutaten sollten Sie sich immer bewußt machen. Beachten Sie insbesondere, dass G hierbei eine beliebige Teilmenge von V sein darf.

- Wieso ist das Szenenbild aus (2.3.40) zum Stichwort *beschränkte Abbildungen* sehr viel allgemeiner als das aus (2.4.2)? Welche wichtige zusätzliche Bedingung muss jetzt erfüllt ein?

(2.4.3) Der Stetigkeitsbegriff erfaßt die folgende Problemsituation: Von unserer Abbildung f seien alle Werte bekannt mit Ausnahme von $f(x_0)$, dem Wert im Aufpunkte x_0 . **Kann man dann diesen Wert mit Hilfe der übrigen Werte vorhersagen oder nicht?** Offenbar ist eine solche Vorhersage nur möglich, wenn f gewisse Glattheitseigenschaften hat, nicht aber, wenn man $f(x_0)$ beliebig und unabhängig von den sonstigen Werten vorgibt **oder** wenn G so gestaltet ist, dass man innerhalb von G nicht beliebig nahe an x_0 herankommt. Das ist in dieser Allgemeinheit ein anspruchsvolles Problem.

(2.4.4) Interpretieren wir die Situation vom Feldstandpunkt:

$G \subset V$ ist unser Konfigurationsraum. $x_0 \in G$ hat den Wert $y_0 = f(x_0)$. Wir nähern uns mit einem Wertemessgerät dem Punkte x_0 , können überall in G messen, nur nicht in x_0 selbst. Wenn sich dann auch $f(x_0)$ dem Vektor y_0 , nähert, dann nähert sich $f(x) - y_0$ dem Nullvektor und $\|f(x) - y_0\|_W$ nähert sich der Zahl Null. Oder auch: In der Dezimaldarstellung der Zahl $\|f(x) - y_0\|_W$ muß jede Nachkommastelle Null werden, wenn nur x nahe genug an x_0 herankommt.

Das ist weitgehend analog zur Situation beim Folgengrenzwert, nur dass zu präzisieren ist, was "nahe genug bei x_0 " bedeutet. Hierzu benötigen wir die Normen und die damit in (2.2.7) eingeführten ε -Kugeln. Mit diesen können wir den qualitativen Begriff der Nähe quantifizieren. Sie gibt mit Hilfe des Abstandes ein Maß dafür, wie benachbart die Punkte jeweils sind. Wieder ist die Vorstellung eines Zweipersonenspieles nützlich: Der Herausforderer gibt ein Genauigkeitsmaß $\varepsilon > 0$ an. Der Verteidiger muß eine Kugelumgebung von x_0 , nennen, in der $f(x)$ und $f(x_0)$ nicht mehr als ε voneinander abweichen. Also $\|f(x) - y_0\|_W < \varepsilon$ gilt. Der Verteidiger nennt die zulässige Umgebung durch Angabe eines geeigneten Kugelradius in V , den man traditionellerweise mit δ bezeichnet. Dieser Radius wird i.a. von ε abhängen, so dass wir jetzt eine Strategiefunktion $\varepsilon \mapsto \delta(\varepsilon)$ benötigen.

(2.4.5) Etwa: Der Herausforderer verlangt eine Genauigkeit von $\varepsilon=10^{-7}$. Der Verteidiger befragt seine Strategiefunktion und sagt: Wenn der Abstand des Meßwagens von x , geringer ist als $\delta = \delta(10^{-7})=10^{-9}$, dann ist der Unterschied zwischen $f(x)$ und y , sicherlich geringer als der geforderte Wert ε . Welche Genauigkeit verlangt wird, legt der Herausforderer fest. Er darf einen beliebig kleinen Fehler fordern, nur exakte Übereinstimmung darf er nicht verlangen, da dieser möglicherweise erst im unzugänglichen Punkt x_0 erreicht wird. Und: Der Herausforderer darf für jeden Punkt der zugelassenen Konfigurationsraumumgebung eine Messung verlangen und damit die Genauigkeitsschranke testen.

(2.4.6) Alles in Formeln und formalisiert. **Die Stetigkeitsdefinition:**

<p style="text-align: center;">Ist f in x_0 stetig? $y_0 = f(x_0)$.</p> <p>1.) Gebe $\varepsilon > 0$ vor.</p> <p>2.) Bestimme dazu $\delta = \delta(\varepsilon)$ derart, dass gilt:</p> <p>3.) $\ f(x) - y_0\ _W < \varepsilon$ für alle $x \in G$ mit $\ x - x_0\ _V < \delta$</p>
--

1) ist der Zug des Herausforderers. 2) der Gegenzug des Verteidigers und 3) der Versuch des Herausforderers, den Verteidiger zu bezwingen.

(2.4.7) Eine Funktion $\varepsilon \mapsto \delta(\varepsilon)$, die das für jedes $\varepsilon > 0$ leistet, nennen wir erneut eine Strategiefunktion zum Punktepaar (x_0, y_0) des Graphenraumes. Kann man zu $(x_0, f(x_0))$ eine Strategiefunktion finden, so heißt f in x_0 stetig. Ist f in X_0 nicht definiert und kann man ein y_0 , finden mit einer Strategiefunktion für (x_0, y_0) , so sagen wir, dass sich "f stetig nach x_0 fortsetzen läßt mit Wert $f(x_0)=y_0$ ".

(2.4.8) **Beachten Sie die Forderung** $x \in G$ in der Stetigkeitsdefinition! Sie legt fest, welche eventuellen Punkte außerhalb von x_0 bei der Meßwerterhebung überhaupt zulässig sind. Nun kann es vorkommen, dass der Meßbereich G so eingeschränkt ist, dass man nicht ausreichend an x_0 herankommt. Sagen wir $V=\mathbb{R}^2$ und $G=\{(x,y) \mid x>0\} \cup (-1,0)$. Dann ist der Funktionswert im isolierten Punkt $x_0=(-1,0)$ des Definitionsbereiches durch die anderen Funktionswerte nicht festgelegt, nicht vorhersagbar. Die Stetigkeitsdefinition verarbeitet einen solchen Fall dadurch, dass sie ihn als stetig klassifiziert, unabhängig vom Wert y_0 in x_0 . (Als Strategiefunktion können Sie $\delta(\varepsilon)=\frac{1}{2}$ wählen. Da damit nur x_0 selbst noch zulässig ist - es kein weiteres $x \in G$ mit $\|x - x_0\|_V < \frac{1}{2}$ gibt, gewinnt der Verteidiger stets.) Stetigkeit schließt also generell nur inkonsistente Meßergebnisse aus, sichert nicht in jedem Fall auch tatsächliche eindeutige Vorhersagbarkeit des Wertes aus den übrigen.

(2.4.9) Wieso fordert man nicht generelle Vorhersagbarkeit? Nun dann benötigt man offensichtlich zusätzliche einschränkende Annahmen zum Definitionsbereich G von f . Und wenn diese möglichst allgemein sein sollen, wird das recht kompliziert und unangenehm. Die oben abstrahierte Form erweist sich als günstiger und nützlicher.

(2.4.10) Zwei einfache Beispiele:

$$f = (\mathbb{R}^2, (x, y) \mapsto f(x, y), \mathbb{R}) \quad \text{mit} \quad f(x, y) = \begin{cases} \frac{xy}{x^2+y^2} & \text{für } (x, y) \neq (0, 0) \\ 0 & \text{für } (x, y) = (0, 0) \end{cases}$$

$$F = (\mathbb{R}^2, (x, y) \mapsto F(x, y), \mathbb{R}) \quad \text{mit} \quad F(x, y) = \begin{cases} \frac{xy}{\sqrt{x^2+y^2}} & \text{für } (x, y) \neq (0, 0) \\ 0 & \text{für } (x, y) = (0, 0) \end{cases}$$

f ist im Ursprung unstetig, wie man sofort prüft. (Vgl. die Figur in (1.8.13)). F dagegen ist dort stetig, wie man sofort prüft: Wir wählen die euklidische Norm. Im ersten Fall nähert man sich entlang $x=y$ der Null. Dann gewinnt der Herausforderer mit jedem ε unterhalb $\frac{1}{2}$. Denn welches $\delta(\varepsilon)$ der Verteidiger auch versucht, der Herausforderer bestimmt den Wert im Punkt $\frac{1}{3}(\delta, \delta)$ mit Abstand $< \delta$ vom Ursprung. Der Feldwert ist 1 und $1-0 > \varepsilon$. Die geforderte Genauigkeit wird nicht erreicht. Im Fall von F sei $r^2 = x^2 + y^2$.

Damit folgt

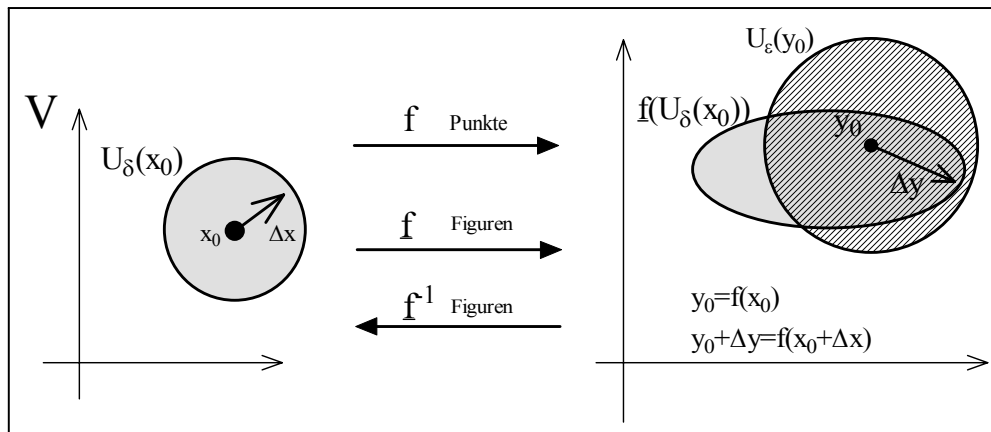
$$\|F(x, y) - F(0, 0)\| \leq \frac{r^2}{r} = r = \|(x, y) - (0, 0)\|$$

Also kann der Verteidiger beispielsweise $\delta(\varepsilon) = \frac{\varepsilon}{2}$ wählen, was für ihn eine Gewinnstrategie ergibt. Ein Schema zum Bestimmen einer Gewinnstrategie besprechen wir unten.

- Veranschaulichen Sie sich die beiden Funktionen f und F mit Hilfe eines Computeralgebraprogramms vom Graphenstandpunkt.

(2.4.11) Wir betrachten das Glattheitsproblem jetzt anstatt vom Feldstandpunkt vom Zuordnungsstandpunkt aus, interpretieren also beide Räume als Konfigurationsräume. Das ist für Stetigkeits- und Ableitungsprobleme immer ein sehr nützlicher Standpunkt, wenn es darum geht, die Übersicht über den Formalismus zu wahren.

Die Figur verdeutlicht das entstehende Szenenbild.



(2.4.12) Wähle x_0 im Urbildraum. Dazu gehört $f(x_0)=y_0$ im Werteraum. Jetzt geht man im Urbildraum von x_0 nach $x=x_0+\Delta x$. Dann geht der zugehörige Wert von y_0 nach $f(x_0+\Delta x)=y_0+\Delta y$. D.h. Δy ist die Wertänderung. **Kann man über Kenntnis der Werte $f(x_0+\Delta x)$ den Wert y_0 vorhersagen?**

(2.4.13) Durch \underline{f} werden die Figuren in V zu Figuren in W . Insbesondere wird aus einer δ -Kugel in V eine Teilmenge von W , die y_0 enthält. Dies ist im allgemeinen **keine Kugel** in W . Alle Punkte y aus W , die in der ε -Kugel $U_\varepsilon(y_0)$ liegen, unterscheiden sich im Abstand von y_0 um weniger als ε . Man kann daher ε als Meßgenauigkeitsgrenze interpretieren. Umgekehrt wird durch \underline{f}^{-1} aus einer ε -Kugel in W eine Teilmenge von V .

(2.4.14) Wie sieht jetzt das Stetigkeitsspiel aus?

- 1) Der Herausforderer gibt eine Meßgenauigkeit $\varepsilon > 0$ vor. Diese bestimmt $U_\varepsilon(y_0)$ in W . Start im Werteraum!
- 2) Der Verteidiger benennt über $\delta = \delta(\varepsilon)$ eine Umgebung $U_\delta(x_0)$ in der der Herausforderer messen darf. Findet der Herausforderer ein Meßresultat, das um mehr als ε abweicht, hat er gewonnen. Oben in der Figur ist das der Fall! Denn $\underline{f}(U_\delta(x_0))$ liegt nicht ganz in der vorgegebenen ε -Kugel. D.h. es gibt Punkte, für die, der Funktionswert zu stark abweicht. Vermutlich kann der Verteidiger seine Niederlage durch Wahl eines kleineren δ vermeiden. Unstetigkeit heißt: Jede δ -Wahl führt zu einer Niederlage.
- 3)

(2.4.15) Formal muß S so gewählt werden, dass gilt:

$$\underline{f}(U_\delta(x_0)) \subset U_\varepsilon(y_0).$$

Das kann man als kurze zusammenfassende Formel der Stetigkeitsbedingung ansehen.

(2.4.16) Man sieht jetzt auch, wann der Herausforderer das Spiel gewinnt: Er nennt ein beliebig kleines $\varepsilon > 0$. Der Verteidiger antwortet mit einem $\delta > 0$. Und jetzt muß der Herausforderer ein x_g mit $\|x_g - x_0\| < \delta$

benennen, für das sich der Funktionswert in der gegebenen Genauigkeit von y_0 unterscheidet, für den also $\|f(x_g) - f(x_0)\| > \varepsilon$ gilt. Und er muß dazu für jede δ -Wahl in der Lage sein.

(2.4.17) Beachten Sie: Die Funktion $\varepsilon \mapsto \delta(\varepsilon)$ geht von W nach V , hat also die entgegengesetzte Richtung von f . Würde man umgekehrt $\delta \mapsto \varepsilon = \varepsilon(\delta)$ vorgeben, käme u.U. nur eine monotone Abnahme des Fehlers heraus, aber dieser müßte nicht auf Null schrumpfen.

(2.4.18) Die typische **Konstruktion einer Strategiefunktion** verläuft analog zum Fall der Folgen. Wir gehen die einzelnen Schritte kurz durch:

1)	Man interpretiert δ als äußeren Parameter und betrachtet alle x mit $\ x - x_0\ _V < \delta$
2)	Für diese x schätzt man die Größe $\ f(x) - f(x_0)\ _W$ vorsichtig nach oben ab. Das Ergebnis sei $\ f(x) - f(x_0)\ _W \leq a(\delta, x_0)$. Beachten Sie: Die rechte Seite ist unabhängig von x .
3)	Jetzt formuliert man über Rollenwechsel die Bedingungsungleichung $a(\delta, x_0) < \varepsilon$ wobei δ Unbestimmte und ε äußerer Parameter.
4)	Kann man die Gleichung durch Umstellen lösen oder auf andere Weise eine Lösung $\delta = \delta(\varepsilon) > 0$ angeben, so hat man eine Strategiefunktion.

Wir müssen unser **Begriffssystem** zum Stichwort "stetig" noch ausbauen. Genauer definiert man:

- " $f: D \rightarrow W$ ist stetig in $x_0 \in D$ " wie eingeführt.
- " f ist stetig" bedeutet: f ist in jedem Punkt $x_0 \in D$ stetig. (D.h. man kann zu jedem Punkt $x_0 \in D$ eine zugehörige Strategiefunktion $\delta = \delta(\varepsilon, x_0)$ angeben.)
- " f ist gleichmäßig stetig". Dies bedeutet, dass es eine gemeinsame Strategiefunktion für alle $x_0 \in D$ gibt. Oder ausformuliert:
 - Zu jedem $\varepsilon > 0$ kann man ein $\delta = \delta(\varepsilon)$ angeben, derart dass $\|f(x_1) - f(x_2)\|_W < \varepsilon$ gilt, sofern nur $\|x_1 - x_2\|_V < \delta$ erfüllt ist!
- Die bisher beschriebene Art von Stetigkeit nennen wir "Umgebungsstetigkeit". Es gibt nämlich eine zweite Art von Stetigkeit, die wir "Folgenstetigkeit" nennen wollen:
 - $f: D \rightarrow W$ heißt "in $x_0 \in D$ **folgenstetig**", falls immer gilt: Wenn $n \mapsto a_n$ eine Folge von Punkten aus D ist, die in V gegen x_0 konvergiert, dann konvergiert die Folge $n \mapsto f(a_n)$ in W gegen $f(x_0)$.

Zwischen den beiden Definitionen besteht der folgende Zusammenhang: Umgebungsstetig impliziert immer folgenstetig. (Das ist leicht zu beweisen). Und folgenstetig impliziert im Rahmen unserer Räume immer umgebungsstetig. D.h. für uns sind die beiden Begriffe gleichwertig. Hat man die eine Eigenschaft, dann darf man die andere immer als gültig annehmen.

Stetigkeit und Linearität

(5) Satz: V, W normiert. $L: V \rightarrow W$ sei linear und weiter sei $n = \dim V < \infty$. **Dann ist L stetig.**

Beweis: Sei a Basis von V . Dann ist $L(\vec{x}) = \sum L(\vec{a}_i)x_i$ und daher

$$\begin{aligned} L(\vec{x}_0 + \Delta x) - L(\vec{x}_0) &= L(\Delta \vec{x}) = \sum L(\vec{a}_i)\Delta x_i \\ \|L(\vec{x}_0 + \Delta x) - L(\vec{x}_0)\|_W &\leq \sum \|L(\vec{a}_i)\|_W |\Delta x_i| \leq A \sum |\Delta x_i| \leq CnA \|\Delta \vec{x}\|_V \end{aligned}$$

Mit $A = \max_i \|L(\vec{a}_i)\|_W$. Die letzte Abschätzung folgt mit dem nachfolgenden Hilfssatz. Die resultierende Ungleichung liefert die gesuchte Stetigkeit. Etwa $\delta(\varepsilon) < \frac{\varepsilon}{CnA}$.

Hilfssatz: Es gibt $C > 0$, so dass für alle $\|\Delta x_i\| \leq C \|\Delta \vec{x}\|_V$ mit $\Delta \vec{x} = \sum \vec{a}_i \Delta x_i$

Beweis:

1) Es genügt, die Ungleichung für Vektoren mit $\sum \Delta x_i^2 = 1$ zu beweisen. (Homogenität, Ungleichung durch $\sqrt{\sum \Delta x_i^2}$ dividieren bzw. damit multiplizieren)

2) Für diese Vektoren ist $\Delta x_i \leq 1$.

3) a Basis und $\Delta \vec{x} = \sum \vec{a}_i \Delta x_i$. Also $\Delta \vec{x} = \vec{0}$ nur wenn alle $\Delta x_i = 0$.

3) Sei $S^{n-1} \subset \mathbb{R}^n$ die n -Sphäre. Sie ist kompakt. Betrachte die Abbildung:

$$N_a = (S^{n-1}, \Delta \vec{x}^K \mapsto \|\Delta \vec{x}\| = \|\sum \vec{a}_i \Delta x_i\|, \mathbb{R})$$

4) N_a ist stetig mit kompaktem Urbild, nimmt also ihr Wertinfimum an. Etwa in $\Delta \vec{x}_0$. Also

$$\|\Delta \vec{x}_0\| = K \leq \|\Delta \vec{x}\| \quad \text{für alle } \Delta \vec{x} \text{ mit } \Delta \vec{x}^K \in S^{n-1}.$$

5) $\boxed{\text{Es muss } K > 0 \text{ gelten}}$. Wäre $K=0$, wäre $\Delta \vec{x}_0 = \vec{0} \notin S^{n-1}$ und die Familie a linear abhängig! a ist aber Basis!

6) Wegen $K > 0$ kann man $C > 0$ so bestimmen, dass $CK = C \|\Delta \vec{x}_0\| > 1$ gilt. Dann folgt für alle $\Delta \vec{x}^K \in S^{n-1}$

$$|\Delta x_i| \leq 1 < C \|\Delta \vec{x}_0\| \leq C \|\Delta \vec{x}\|.$$

Wegen der Homogenitätseigenschaft der Norm gilt die Ungleichung dann für alle Vektoren.

Gilt nicht, wenn V unendlichdimensional ist!!!! Beispiel: \mathcal{P} Raum der Polynome mit Legendrenorm $\|p\|_L^2 = \int_{-1}^1 dx p^2(x)$. Dann ist die Polynomfolge $n \mapsto p_n = h_n$ eine Nullfolge. Es gilt ja $\|p_n\|^2 = \frac{2}{2n+1}$. Die Ableitung ist eine lineare Abbildung $D: \mathcal{P} \rightarrow \mathcal{P}$. Aber die Folge der Ableitungen $n \mapsto p'_n = nh_{n-1}$ hat die Normen $\|p'_n\|^2 = n^2 \int_{-1}^1 dx x^{2n-2} = \frac{2n^2}{2n-1} \rightarrow \infty$.

$$p_n \rightarrow 0 \quad \text{aber} \quad Dp_n \text{ konvergiert nicht!}$$

12.2.4a Offene Teilmenge

(2.4.19) Wann darf man sicher sein, dass die Werte in den Punkten $\neq x_0$ den Wert in x_0 im stetigen Fall **eindeutig** bestimmen? Wir haben gesehen, dass in diesem Fall Einschränkungen an den Definitionsbereich erforderlich waren. Wir formulieren jetzt eine solche Einschränkung, die gut zu handhaben ist und die für die meisten praktischen Zwecke ausreicht.

Vom Feldstandpunkt aus gesehen müssen wir in der Lage sein, uns aus jeder Richtung her dem kritischen Punkt x_0 zu nähern, wobei es erlaubt ist, die Annäherung beliebig nahe bei diesem Punkt zu beginnen.

(2.4.20) Eine Menge, die sicherlich alle Richtungen enthält, ist eine ε -Kugel in V um x_0 . Wir verlangen nun, dass mit x_0 auch immer mindestens eine ε -Kugel im Definitionsbereich G von f liegen soll. Dabei darf der Radius ε beliebig klein sein, muß aber natürlich $\varepsilon > 0$ erfüllen.

(2.4.21) Wir erhalten folgende **wichtige Definition**

Es sei V normierter Raum.
Eine Teilmenge $G \subset V$ heißt eine *offene Teilmenge von V* , wenn es zu jedem $x \in G$ ein $\varepsilon > 0$ gibt, so dass die Kugel $U_\varepsilon(x)$ ganz in G liegt. Also $U_\varepsilon(x) \subset G$.

(2.4.22) "Offenheit einer Teilmenge" ist natürlich jeweils nachzuprüfen. So sind die offenen Intervalle $]a, b[$ des \mathbb{R}^1 tatsächlich offen. Ist $G \subset \mathbb{R}^n$ bezüglich einer Norm offen, so auch bezüglich jeder anderen Norm dieses Raumes.

(2.4.23) Ist U Teilmenge von V , dann sollte man zwischen "offen in U " und "offen in V " sorgfältig unterscheiden. Wir gehen hier auf den ersten Begriff nicht weiter ein.

(2.4.24) Ein abgeschlossener Kreis wie $K = \{(x, y) | x^2 + y^2 \leq 4\}$ ist nicht offen. Beispielsweise liegt $(2, 0)$ in dieser Menge. Jeder noch so kleine ε -Kreis um diesen Punkt liegt aber etwa zur Hälfte außerhalb von K . Unsere Forderung ist verletzt. $(2, 0)$ ist ein "Randpunkt". Ganz V dagegen ist offen. **Wir werden üblicherweise -meist stillschweigend - fordern, dass die Definitionsbereiche unserer Abbildungen $f: V \rightarrow W$ offene Teilmengen von V sind.** Dann darf man für jeden Punkt voraussetzen, dass mit

ihm auch eine kleine ε -Umgebung im Definitionsbereich liegt. Die Radien dieser Kugeln für verschiedene Punkte dürfen (natürlich) verschieden sein.

(2.4.25) Abschließend ein Beweis der Behauptung: Ist $f:G \rightarrow W$ stetig in x_0 und G offen und V vollständig, dann ist der Wert $f(x_0)$ als Grenzwert vorhersagbar.

Angenommen man kann f für $x=x_0$ auf zwei Weisen stetig ergänzen. Seien f_1 und f_2 diese beiden Ergänzungen mit $f_1(x_0)=y_1$ und $f_2(x_0)=y_2$. Für alle anderen x gilt $f_1(x)=f_2(x)$. Wir müssen zeigen: $y_1 = y_2$. Angenommen $y_1 \neq y_2$. Dann ist $\|y_1 - y_2\|_W = a > 0$. Dies folgt aus dem ersten Normaxiom. Wir setzen $\varepsilon = \frac{a}{3} > 0$. Falls f_1 in x_0 stetig ist, kann man ein zugehöriges δ_1 bestimmen. Entsprechend ein δ_2 für f_2 . Wir wählen $\delta = \min(\delta_1, \delta_2)$. Dies ist ein δ für beide Funktionen. Sei jetzt $x \in U_\delta(x_0)$ mit $x \neq x_0$, so dass $f_1(x)=f_2(x)$ gilt. Dann rechnen wir:

$$\begin{aligned} a &= \|y_1 - y_2\|_W = \|f_1(x_0) - f_2(x_0)\|_W = \|(f_1(x_0) - f_1(x)) - (f_2(x_0) - f_2(x))\|_W \\ &\leq \|f_1(x_0) - f_1(x)\|_W + \|f_2(x_0) - f_2(x)\|_W < \varepsilon + \varepsilon \end{aligned}$$

Das ist aber Unsinn, da $a=3\varepsilon > 2\varepsilon > 0$ war. Also muß $y_1=y_2$ gelten!

(2.4.26) Das Vorgehen ist typisch für die Arbeit mit der Stetigkeitsbedingung: Das teleskopartige Einfügen eines Zwischenpunktes und anschließende Anwendung der Dreiecksungleichung.

(2.4.27) Der gesamte Beweis geht nur, wenn es immer einen Punkt $x \in G$ mit den geforderten Eigenschaften gibt. Andernfalls läßt sich die Ungleichung $a < 2\varepsilon$ nicht folgern. Ist der Definitionsbereich offen, dann gibt es immer solche Vektoren, da man die Längen von Δx durch Umskalieren beliebig klein machen kann.

Kompakte Teilmengen

Definition

Kompakt und Stetigkeit:

Stetiges Bild von kompakt ist kompakt

Auf kompakt ist stetig automatisch gleichmäßig stetig

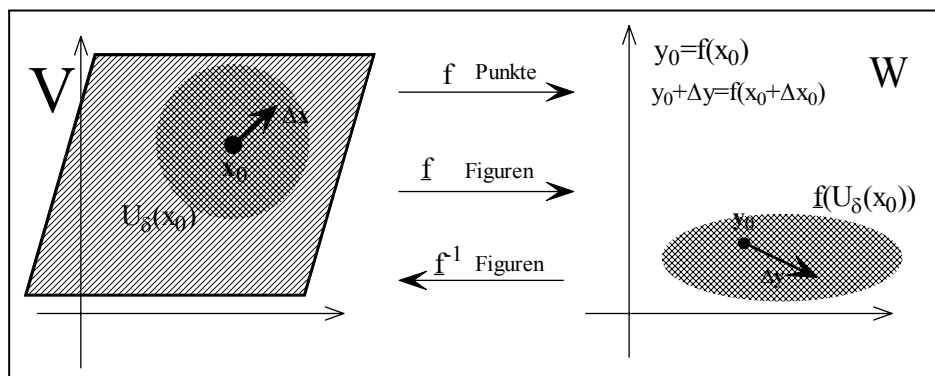
$f: \text{kompakt} \rightarrow \mathbb{R}$ stetig nimmt Extremwerte an.

Kap. 6.3: Die totale Ableitung

6.3.0 Das Szenenbild

(3.0.1) Wir wenden uns jetzt dem eigentlichen Thema des Kapitels zu, der Frage nach der Ableitung unserer Funktionen $f:V \rightarrow W$.

(3.0.2) Dabei müssen wir von dem folgenden Szenenbild ausgehen, das wir nach den Vorbereitungen der vorangehenden Teilkapitel formulieren können:



V und W seien zwei Banachräume mit Normen $\|\cdot\|_V$ und $\|\cdot\|_W$ über demselben Körper (\mathbb{R} oder \mathbb{C}). Weiter sei G eine offene Teilmenge von V , das Parallelogramm der Figur. $f:G \rightarrow W$ sei eine Abbildung und $x_0 \in G$ ein Punkt, den wir den **Aufpunkt** nennen. Da G offen ist, können wir immer eine Umgebung $U_\delta(x_0)$ finden, die ganz in G liegt. Wir setzen $y_0 = f(x_0)$. Das ergibt einen Punkt im Wertekonfigurationsraum W .

Jetzt gehen wir im Urbildraum von x_0 nach $x = x_0 + \Delta x$. Auf jeden Fall soll auch x in G liegen. Uns interessieren kleine Änderungen Δx . Durch Verkleinern können wir stets erreichen, dass $x = x_0 + \Delta x$ in der Umgebung $U_\delta(x_0)$ liegt. (Im Falle unendlichdimensionaler Vektorräume benötigt man für f noch zusätzliche Stetigkeitsannahmen, worauf wir hier nicht weiter eingehen. Aber im Prinzip ist dieser Fall mit eingeschlossen.)

(3.0.3) Unser eigentliches Interesse gilt dem Funktionswert $f(x) = f(x_0 + \Delta x)$. **Wie können wir mit möglichst wenig Aufwand möglichst viel über diesen Wert aussagen?** Können wir diesen Wert geeignet approximieren, ihn näherungsweise durch einen leicht auswertbaren Rechenausdruck ersetzen?

(3.0.4) Etwas formaler: **Wir suchen eine lokale Approximation von $\Delta x \mapsto f(x_0 + \Delta x)$ um den Aufpunkt x_0 herum.**

(3.0.5) Üblicherweise liegt folgende Rollenzuweisung vor: f ist festgelegt, x_0 ist äußerer Parameter. Dadurch ist $y_0 = f(x_0)$ bestimmt. $\Delta x = x - x_0$ wird unabhängige Variable und $x = x_0 + \Delta x$ ist Hilfsgröße. Unsere Zuordnung erhält die Form $\Delta x \mapsto f(x_0 + \Delta x)$. Dabei darf Δx so gewählt werden, dass $x \in G$ gilt. Also $\Delta x \in G - \{x_0\}$.

6.3.1 Der allgemeine Formalismus

(3.1.1) Die zentrale und enorm fruchtbare Idee besteht darin, folgende Umschreibung des Rechenausdruckes vorzunehmen:

$$\boxed{f(x_0 + \Delta x) = f(x_0) + L \cdot \Delta x + \|\Delta x\| R_f(x_0, \Delta x)}$$

mit $f \in \text{Hom}_K(V, W)$ lineare Abbildung.

Wie üblich schreiben wir $L \cdot \Delta x$ statt $L(\Delta x)$. Inspizieren wir die Umschreibungsformel genauer:

(3.1.2) Die Zerlegung ist **sinnvoll** für jede lineare Abbildung $L:V \rightarrow W$. Dabei muß jeder Summand typmäßig ein Element aus W sein. D.h. insbesondere, dass $\Delta x \mapsto R_f(x, \Delta x)$ eine Abbildung $V \rightarrow W$ sein muß, genauer eine Abbildung von der Teilmenge zulässiger Δx , die durch $(x_0 + \Delta x) \in G$ bestimmt ist nach W .

(3.1.3) Für $\Delta x \neq 0$ und jedes L ist $R = R_f^L$ durch die Gleichung eindeutig festgelegt: Man kann ja nach R auflösen. Δx darf wegen G offen stets in einer kleinen δ -Umgebung von 0 variieren. $\|\Delta x\|_V$ ist eine positive Zahl (ungleich Null), durch die dividiert werden kann.

(3.1.4) Die Zerlegung wird eindeutig gemacht über eine zusätzliche Bedingung an R . Die Funktion R nennen wir den *Restterm von f um x_0* und die Bedingung, die die Eindeutigkeit gewährleistet, die *Resttermbedingung*. Der letzte Summand $\|\Delta x\|_V R_f(x_0, \Delta x)$ ist der *absolute Fehler (der linearen Näherung)*. Zur Formulierung der Bedingung für R benötigen wir die Resultate des vorangegangenen Teilkapitels:

Ist in obiger Zerlegung L so gewählt, dass $R = R_f^L$ im Ursprung $\Delta x = 0$ stetig ergänzbar ist mit Grenzwert Null, dann sagen wir, die Zerlegung erfülle **die Resttermbedingung**. Anders gesagt, es muß gelten:

$\lim_{\Delta x \rightarrow 0, \Delta x \neq 0} R_f(x_0, \Delta x) = 0 \text{ in } W.$	Also:	$\lim_{\Delta x \rightarrow 0, \Delta x \neq 0} \ R_f(x_0, \Delta x)\ _W = 0.$
--	-------	--

(3.1.5) Hiermit beweist man wie im eindimensionalen Fall in der üblichen Weise die Eindeutigkeit: **Es gibt höchstens ein L , für das die Resttermbedingung erfüllt ist.**

□ Führen Sie den Nachweis.

Dieses ausgezeichnete L -sofern es existiert- ist abhängig von f und x_0 . Der Buchstabe L sollte zunächst daran erinnern, dass L typmäßig eine lineare Abbildung $V \rightarrow W$ ist, also ein Vektorraumhomomorphismus, für den der gesamte in der linearen Algebra entwickelte Apparat zur Verfügung steht. Wir werden gleich für L eine passendere Bezeichnung einführen.

(3.1.5) Natürlich wird es Fälle geben, in denen eine solche Entwicklung nicht existiert. Aber wenn es eine gibt, dann ist sie unter den gegebenen Umständen die einzig mögliche.

(3.1.6) Noch einige Bezeichnungen:

Existiert eine Zerlegung mit Resttermeigenschaft, so nennen wir diese Zerlegung die *Tangentenzerlegung von f um x_0* . Die Zerlegung stellt eine Umschreibung des Rechenausdrucks dar. Das zugehörige L nennen wir *die totale Ableitung von f in x_0* und bezeichnen sie mit $Df(x_0)$. Das ist eine lineare Abbildung $Df(x_0) = (V, \Delta x \mapsto Df(x_0).\Delta x, W)$ mit x_0 als äußerem Parameter oder Konstante! Man kann dafür auch $f'(x_0)$ schreiben. Man sagt, **f sei im Punkte x_0 differenzierbar.**

(3.1.7) Bei all diesen Begriffen und Formulierungen sollte man darauf achten, sich die jeweils richtigen oder benötigten Zutaten bewußt zu machen. Für $Df(x_0)$ darf man Δx natürlich beliebig in ganz V wählen. Die Beschränkung auf $G - \{x_0\}$ entfällt für die totale Ableitung selbst, da diese eine lineare Abbildung ist. Die von-Klammer in den Werten wird durch einen Punkt ersetzt, d.h. wir schreiben wieder $Df(x_0).\Delta x$ anstelle $Df(x_0)(\Delta x)$ oder gar $(Df(x_0))(\Delta x)$. Aber dieser Punkt bezeichnet i.a. kein gewöhnliches Produkt! Darauf ist zu achten.

(3.1.8) Die Verwendung der totalen Ableitung erfolgt meist im Rahmen einer der folgenden beiden Denkfisuren unter Nutzung der Resultate der linearen Algebra.

Die erste Denkfigur: Weist man für eine Zerlegung mit einem bestimmten linearen L die Resttermeigenschaft nach, so kann man daraus die (dann existierende) zugehörige totale Ableitung ablesen. Die Zerlegung liefert ja alle Werte $L.\Delta x$. Kennt man alle Werte, so hat man auch die Zuordnung! Das ist eine Folge der Eindeutigkeit aus (3.1.5).

Die zweite Denkfigur: Weiß man, dass f in x_0 differenzierbar ist und kennt man die totale Ableitung $Df(x_0)$, dann kann man die Tangentenzerlegung als gültige Gleichung hinschreiben, wobei der ansonsten nicht näher spezifizierte Restterm die Resttermbedingung erfüllt. Was aus Stetigkeitsgründen bedeutet, dass sein Beitrag lokal nur eine relativ zu den beiden ersten Termen kleine Korrektur bildet.

(3.1.9) Die zweite Denkfigur ist der Grund der großen Anwendungsrelevanz des Formalismus.

Die Tangentenzerlegung stellt **eine geschickte Umcodierung** der in der Zuordnung von f enthaltenen Information dar, derart, dass der letzte Summand in der Nähe von x_0 einen möglichst kleinen Beitrag liefert, wogegen die ersten beiden wertemäßig relevanten Summanden als affine lineare Abbildung mathematisch besonders einfach und leicht zugänglich sind.

Zu dem Begriffssystem der totalen Ableitung gehören weiterhin die folgenden Definitionen:

Läßt man den Restterm fort, erhält man die Tangentenapproximation von f um x_0.
Unter dem totalen Differential df (von f in x_0) versteht man den Wert $df=Df(x_0).\Delta x$. Diese Größe wird besonders in der Physik benutzt.
Ist U ein Teilraum von V , so versteht man unter der partiellen Ableitung von f in Richtung U die Restriktion von $Df(x_0)$ auf U . Also $D_U f(x_0) = (Df(x_0)) _U$. Besonders wichtig ist eindimensionales U .

Auf den Fall des eindimensionalen U werden wir in Kap. 6.3.2a ausführlich zurückkommen. Das totale Differential gibt eine Näherung für die Änderung von f , wenn man von x_0 nach $x_0 + \Delta x$ geht. Also $df \approx f(x_0 + \Delta x) - f(x_0)$.

(3.1.11) Wir skizzieren kurz die Schritte des weiteren **routinemäßigen Vorgehens**, die wir hier nicht alle ausführen.

- Mit Hilfe der beider Denkfiguren beweist man die üblichen Ableitungsregeln:
 - **Linearität:** $D(\alpha f + \beta g)(x_0) = \alpha Df(x_0) + \beta Dg(x_0)$.
 - **Kettenregel:** $D(g \circ f)(x_0) = Dg(y_0) \circ Df(x_0)$. (Unter geeigneten Voraussetzungen. $y_0 = f(x_0)$)
 - **Produktregeln** unter geeigneten Bedingungen.

□ Beweisen Sie Linearität und Kettenregel.

□ Beweisen Sie die folgende Produktregel: $F : V \rightarrow W$ Vektorfeld und $s : V \rightarrow \mathbb{R}$ Skalarfeld. Beide seien in $x_0 \in V$ differenzierbar. Dann ist auch $sF = (V, x \mapsto s(x)F(x), W)$ in x_0 differenzierbar mit totalem Differential

$$D(sF)(x_0).\Delta x = (Ds(x_0).\Delta x)F(x_0) + s(x_0)(DF(x_0).\Delta x).$$

- Mit Hilfe der ersten Denkfigur leitet man die totalen Ableitungen der Abbildungen einer Grundausstattung des jeweiligen Abbildungstyps her. Bei Skalarfeldern sind das die beiden Felder $\vec{x} \mapsto \vec{x}^2$ und $\vec{x} \mapsto (\vec{a} \cdot \vec{x})$. Bei den üblichen Vektorfeldern ist das $\vec{x} \mapsto \vec{x}$ und $\vec{x} \mapsto \vec{a} \times \vec{x}$. Dann arbeitet man sich rekursiv mit Hilfe der Ableitungsregeln hoch, bestimmt die Ableitungen der interessierenden Rechenausdrücke.

□ Beweisen Sie: $D(\vec{x}_0^2).\Delta \vec{x} = 2\vec{x}_0 \cdot \Delta \vec{x}$ und $D(\vec{a} \cdot \vec{x}).\Delta \vec{x} = \vec{a} \cdot \Delta \vec{x}$.

- Unter Verwendung der so bestimmbareren Ableitungen nutzt man die 2. Denkfigur zur Problemlösung, wobei man gerne unter dem Stichwort Linearisierung die Tangentenzerlegung durch die Tangentenapproximation ersetzt.

Im Zusammenhang mit der ersten Denkfigur ist die Resttermbedingung nachzuweisen. Hierzu stehen einem vornehmlich die folgenden Hilfsmittel zur Verfügung:

Nützliche Hilfen zum Nachweis der Resttermbedingung:

- (R1) *Sortieren nach Potenzen von Δx_i .* Man ersetzt in Gedanken $\Delta \vec{x}$ durch $\alpha \Delta \vec{x}$ und sortiert nach Summanden mit der Potenz α^0 , α^1 und α^k mit $k > 1$. Hierzu nimmt man geeignete zulässige Umformungen des Rechenausdrucks vor. Der Beitrag, der zu α^k mit $k > 1$ führt, ist dann der Kandidat für den absoluten Fehler. Division durch $\|\Delta \vec{x}\|_V$ gibt den Kandidaten für den Restterm.
- (R2) *Vorsichtiges Abschätzen des Resttermkandidaten.* Man schätzt vorsichtig nach oben ab. Etwa mit $|\Delta x_i| \leq \|\Delta \vec{x}\|_V$ oder $\frac{1}{a^2 + \|\Delta \vec{x}\|_V^2} \leq \frac{1}{a^2}$, so dass der entstehende vereinfachte Term gegen Null geht. (Vergleichskriterium).
- (R3) *Einsetzen der Tangentenzerlegung von Bausteinen des Rechenausdrucks für $f(x_0 + \Delta x)$.* Man setzt bereits verfügbare Tangentenzerlegungen ein und benutzt die zugehörige (dann gültige) Resttermbedingung, um die gesuchte Resttermbedingung zu beweisen. (Analog zum Beweis der Ableitungsregeln.)
- (R4) *Grenzwertsätze.* Man nutzt die allgemeinen Grenzwertsätze

□ Beweisen Sie: Es sei $\Delta \vec{x} \in \mathbb{R}^n$ mit einer der drei üblichen Normen. Dann gilt für $i=1,2,\dots,n$ $|\Delta x_i| \leq \|\Delta \vec{x}\|_V$.

(3.1.12) **Beispiel.** Es sei $F = \left(\mathbb{R}^2, \begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} x + y^2 \\ (x + y^2)^2 \end{pmatrix}, \mathbb{R}^2 \right)$. Das ist eine gut rechenbare Abbildung vom Feldtyp. Wir verwenden die erste Denkfigur, bilden $F(x + \Delta x, y + \Delta y)$, sortieren nach Potenzen in Δx und Δy und bestimmen einen Kandidaten für den Restterm:

$$F(x + \Delta x, y + \Delta y) = F(x, y) + \begin{pmatrix} 1 & 2y \\ 2(x + y^2) & 4y(x + y^2) \end{pmatrix} \begin{pmatrix} \Delta x \\ \Delta y \end{pmatrix} + \|\Delta \vec{x}\| \begin{pmatrix} 1 & \Delta y^2 \\ \|\Delta \vec{x}\| & s(\Delta x, \Delta y) \end{pmatrix}$$

wobei $s(\Delta x, \Delta y) = \Delta x^2 + 4y\Delta x\Delta y + \Delta y^4 + 2(x + y^2)\Delta y^2 + 2(\Delta x + 2y\Delta y)\Delta y^2$ gesetzt ist.

Das ist die angestrebte Form! Zum Nachweis der Resttermeigenschaft wählen wir die L^1 -Norm in beiden Räumen. Mit $r = \|\Delta \vec{x}\|_1 = \|\Delta(\Delta x, \Delta y)\|_1 = |\Delta x| + |\Delta y|$ folgt

$$\|R(\vec{x}, \Delta \vec{x})\|_1 = \frac{1}{\|\Delta \vec{x}\|_1} (|\Delta y^2| + |s(\Delta x, \Delta y)|)$$

Zunächst wird jeder Summand des Zählers durch seinen Betrag ersetzt (Dreiecksungleichung): $|s(\Delta x, \Delta y)| \leq |\Delta x^2| + 4|y||\Delta x||\Delta y| + \dots$. Anschließend werden im Zähler durch $|\Delta x| \leq r$ und $|\Delta y| \leq r$ alle Δx und Δy beseitigt. Das gibt

$$\|R(\vec{x}, \Delta \vec{x})\|_1 \leq \frac{1}{r} (r^2 + r^2 + 4|y|r^2 + \dots) = r + r + 4|y|r + \dots$$

Und das hat offensichtlich die gewünschte Nullstelle in r .

□ Lesen Sie jetzt mit Hilfe der ersten Denkfigur unmittelbar das totale Differential dF und die totale Ableitung $DF(\vec{x})$ ab.

(3.1.13) **Beispiel:** $q = (V^3, \vec{x} \mapsto \vec{x}^2, \mathbb{R})$. Das ist ein grundlegendes Skalarfeld auf dem euklidischen Raum. Wir rechnen $q(\vec{x}_0 + \Delta \vec{x}) = (\vec{x}_0 + \Delta \vec{x})^2 = \vec{x}_0^2 + (2\vec{x}_0) \cdot \Delta \vec{x} + \Delta \vec{x}^2$. Die Resttermbedingung ist erfüllt, da $\Delta \vec{x}^2 = \|\Delta \vec{x}\|_E^2$ für die euklidische Norm gilt. Damit lesen wir ab:

$$Dq(\vec{x}_0) = (V^3, \Delta \vec{x} \mapsto (2\vec{x}_0) \cdot \Delta \vec{x}, \mathbb{R}), \text{ oder } dq = Dq(\vec{x}_0) \cdot \Delta \vec{x} = (2\vec{x}_0) \cdot \Delta \vec{x}$$

Schreibt sich das totale Differential eines Skalarfeldes (auf V^3 oder \mathbb{R}^n) als Skalarprodukt eines gewissen Vektors mit dem zweiten Vektor $\Delta \vec{x}$ so bezeichnet man diesen Vektor als **Gradienten des Feldes**. Im vorliegenden Fall gilt daher

$$\text{grad}(q(\vec{x}_0)) = 2\vec{x}_0 \quad \text{oder etwas leger } \text{grad}(\vec{x}_0^2) = 2\vec{x}_0.$$

(3.1.15) Im Falle eines **ebenen** Skalarfeldes steht uns die folgende schöne **Veranschaulichung der Tangenzenzerlegung** zur Verfügung, die sich in naheliegender Weise verallgemeinern läßt.

Wir legen einen Kreis um den Aufpunkt mit einer Parametrisierung $\alpha \mapsto \vec{x}_0 + r\vec{e}_r(\alpha)$, wobei $\vec{e}_r(\alpha) = \vec{e}_1 \cos \alpha + \vec{e}_2 \sin \alpha$ ist. Also $\Delta \vec{x} = r\vec{e}_r(\alpha)$. Entlang dieser Kreislinie bestimmen wir einmal ds , also die Feldänderung zwischen Aufpunkt und Kreispunkt in linearer Approximation und zum anderen die exakte Änderung der Feldwerte, also $ds_{\text{exakt}} = s(\vec{x}_0 + \Delta \vec{x}) - s(\vec{x}_0)$. Das Ergebnis sind zwei reelle Funktionen (von α), die Art und Qualität der Tangentenapproximation gut verdeutlichen. (Das Ganze ist eine gezielte Nutzung des Feldstandpunktes zur Veranschaulichung eines abstrakten Sachverhaltes.)

Die beiden Funktionen werden gegeben durch

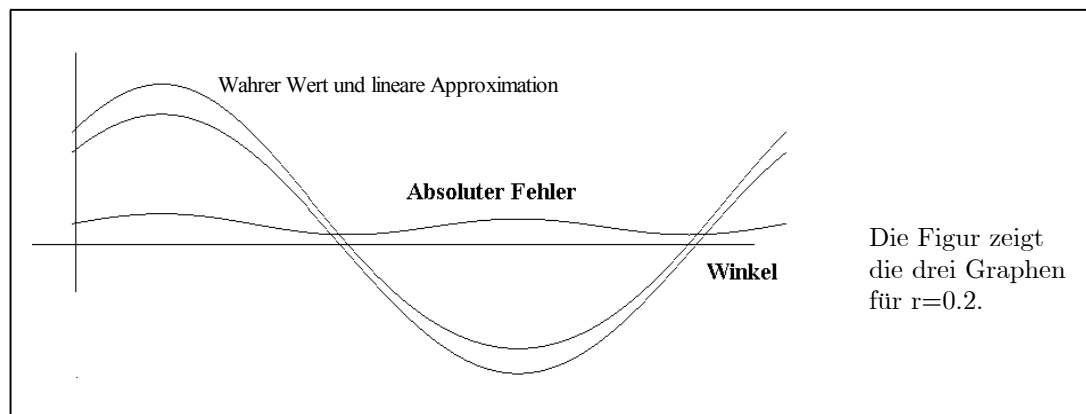
$$\begin{aligned} \alpha &\mapsto ds = Ds(\vec{x}_0) \cdot (r\vec{e}_r(\alpha)) = rDs(\vec{x}_0) \cdot \vec{e}_r(\alpha) && \text{Linearität von } Ds! \\ \alpha &\mapsto ds_{\text{exakt}} = s(\vec{x}_0 + r\vec{e}_r(\alpha)) - s(\vec{x}_0) \end{aligned}$$

Das Skalarprodukt läßt sich noch weiter auswerten zu:

$$\begin{aligned} ds &= r \text{grad}s(\vec{x}_0) \cdot \vec{e}_r(\alpha) = r |\text{grad}s(\vec{x}_0)| \cos(\varphi + \alpha) \quad \text{und} \\ ds_{\text{exakt}} - ds &= r R_s(\vec{x}_0, r\vec{e}_r(\alpha)) \end{aligned}$$

Hierbei ist φ der Winkel zwischen \vec{e}_1 und dem Gradienten. Daher ergibt die Approximation ds immer eine Funktion vom Cosinustyp: $ds(\alpha) = r |\text{grad}s(\vec{x}_0)| \cos(\varphi + \alpha)$. Je größer r wird, desto stärker werden i.a. die Unterschiede, besonders auch in der Form der Kurven.

- Wählen Sie $s(\vec{x}) = (\vec{x}^2)^2$. Also $\text{grad}s(\vec{x}) = 4\vec{x}^2\vec{x}$. Weiter sei $\vec{x}_0 = {}^t(1, 1)$, also $\varphi = \frac{\pi}{4}$. Bestimmen Sie dazu ds, ds_{exakt} und rR_s . Zeichnen Sie diese Funktionen für einige r -Werte. (Etwa 1, 0,5 und 0,1.)



- Entwickeln Sie eine entsprechende Veranschaulichung der Tangentenzersetzung für Abbildungen $\mathbb{R}^2 \mapsto \mathbb{R}^2$. Gehen Sie dabei vom Transformationsstandpunkt aus.
- Was ist die totale Ableitung einer linearen Abbildung $L:V \mapsto W$?

6.3.2 Quantifizierung und analytische Darstellung der totalen Ableitung

(3.2.1) Wie und durch wieviele Rechenausdrücke in \vec{x}_0 kann man die totale Ableitung festlegen?

Meist bestimmt man zunächst eine Formel für das totale Differential df , also für die Werte der totalen Ableitung. Formeln für das totale Differential erweisen sich als anwendungsrelevant, denn sie liefern Formeln für die Änderung von Größen und solche **Formeln für Änderungen** werden vielfach benötigt. Kennt man aber die Werte $df = Df(\vec{x}_0) \cdot \Delta\vec{x}$, dann kann man auf die Zuordnung und damit auf die totale Ableitung selbst schließen.

(3.2.2) Wir besprechen dies Problem hier in **6.3.2** für den wichtigen Fall konstanter Basisfelder und kartesischer Koordinaten. In diesem Fall liegt eine weitgehende Analogie zu der entsprechenden Behandlung des Problems für die linearen Abbildungen in Kapitel 4 vor. Unser Ziel ist, das Diagramm aus (1.7.2) im oberen Bereich durch $Df(x_0)$ und unten durch die zu bestimmende beschreibende Matrix von $Df(x_0)$ zu ergänzen.

Der Übergang zu allgemeinen Koordinaten und zu allgemeinen Basisfeldern bereitet Probleme. Wir werden den zugehörigen Formalismus voll erst in Kapitel 16 entwickeln. In Kap. **6.3.4** führen wir bereits eine erste wichtige und nützliche Verallgemeinerung ein. Sei dazu $f:G \rightarrow W$ und \vec{x}^P eine Parametrisierung von G . Vgl. Kap. 6.1.4, insbesondere (1.4.13). In **6.3.4** wird gezeigt, wie die totale Ableitung von \vec{x}^P (nicht f !) ein Basisfeld von G bzw. genauer $\text{Bild}\vec{x}^P$ erzeugt. Und das ist weitgehend das Basisfeld, das man üblicherweise auch rein geometrisch den \vec{x}^P bestimmten Koordinaten zuordnen würde. Mit Hilfe von Df kann man daraus schließlich ein ebenso nützliches Basisfeld für die Tangentialräume von $\text{Bild}f$ machen.

Hierzu müssen wir allerdings zunächst den fundamentalen Begriff des Tangentialraumes klären und definieren, was in **6.3.3** geschieht.

Also zunächst: **Kartesische Koordinaten und konstantes Basisfeld.**

6.3.2a Quantifizierung der Totalen Ableitung über die Vorgabe eines konstanten Basisfeldes im Urbildraum

Sei $e = (e_1, \dots, e_n)$ nachfolgend eine Basis im Urbildraum von f . Hierzu gehört insbesondere der Fall $G \subset \mathbb{R}^n$ mit der kanonischen Basis.

(3.2.3) Wir berechnen das totale Differential df für $\Delta x = \sum e_i \Delta x_i$ unter Verwendung der Linearität der Ableitungsoperation:

$$df = Df(x_0) \cdot \Delta x = DF(x_0) \cdot (\sum e_i \Delta x_i) = \sum_i (DF(x_0) \cdot e_i) \Delta x_i$$

(3.2.4) Die Faktoren vor den Skalaren Δx_i kann man nun auch auf eine **andere Weise** erhalten: Dazu bilden wir die Restriktionsabbildungen auf die Parallelen zur Koordinatenachse, also $t \mapsto f(x_0 + e_i t)$. Kinetisch laufen wir also im Urbildraum mit Geschwindigkeit 1 parallel zur i -ten Achse und verfolgen im Werteraum die Bahn des Bildpunktes! Es entstehen Abbildungen vom Kurventyp, die für $t=0$ über den Aufpunkt x_0 nach $f(x_0)$ gehen und die wir wie üblich differenzieren können. Wir suchen die momentane Geschwindigkeit dieser Kurve zur Zeit $t=0$.

(3.2.5) Das Ergebnis der Ableitung solcher Restriktionen sind die üblichen partiellen Ableitungen:

$$\frac{\partial f}{\partial x_i}(x_0) = \lim_{\Delta t \rightarrow 0, \Delta t \neq 0} \frac{f(x_0 + e_i \Delta t) - f(x_0)}{\Delta t} \quad \text{oder} \quad f(x_0 + \Delta t) = f(x_0) + \frac{\partial f}{\partial x_i}(x_0) \Delta t + \dots$$

(3.3.6) Diese partielle Ableitung ist als momentane Geschwindigkeit jeweils ein Vektor aus W . Andererseits können wir auch die volle Tangentenerlegung für f um x_0 auf die Parallele zur i -ten Achse restringieren: $f(x_0 + e_i \Delta t) = f(x_0) + (Df(x_0) \cdot e_i) \Delta t + \dots$. Erneut haben wir hierbei die Linearität von Df benutzt. Die Eindeutigkeit der Tangentenerlegung der Kurvenabbildung $t \mapsto f(x_0 + e_i t)$ ergibt jetzt den folgenden wichtigen Zusammenhang zwischen der totalen Ableitung und den üblicherweise benutzten partiellen Ableitungen (bezüglich des konstanten Basisfeldes e):

$$\boxed{\frac{\partial f}{\partial x_i}(x_0) = Df(x_0) \cdot e_i}$$

(3.2.7) Damit folgt sofort die gesuchte und grundlegende **Darstellungsformel für das totale Differential**

$$df = \sum_{i=1}^n \frac{\partial f}{\partial x_i}(x_0) \Delta x_i \quad \text{Man schreibt meist sogar } dx_i \text{ statt } \Delta x_i.$$

(3.2.8) Zur Bedeutung dieser Formel: Sie gibt die Änderung der Feldgröße f als Funktion der Änderung der Konfigurationsraumlage wieder. Von der Funktion f gehen dabei die partiellen Ableitungen bezüglich eines konstanten Basisfeldes, meist sogar einer festen kartesischen Basis ein. Diese Änderung df ist eine Größe, die für zahlreiche Anwendungen interessant ist. Die Formel erlaubt eine rechnerisch einfache näherungsweise Berechnung der Feldänderung. Die Tangentenapproximation erhält man über $f(x_0 + \Delta x) = f(x_0) + df$.

Beachten sie die folgende Gedächtnishilfe

$$df \quad / \quad \frac{\partial f}{\partial x} dx \quad / \quad \frac{\partial f}{\partial x_i} dx_i \quad / \quad \sum_i \frac{\partial f}{\partial x_i} dx_i$$

- Sei $n=3$ und $\vec{x} = \vec{e}_1 x + \vec{e}_2 y + \vec{e}_3 z$. Weiter sei $\vec{f}(\vec{x}) = \vec{x} + e_1 x^2 + e_3 yz$. Konkretisieren Sie hierfür (3.2.7). Dasselbe für das Skalarfeld $s(\vec{x}) = (\vec{x}^2)^2$.
- Sei $f = F \circ g$. Berechnen Sie df mit Hilfe der Kettenregel unter Verwendung der Indexschreibweise. Schreiben Sie die entstehende Summe für den Fall aus, dass alle beteiligten Räume zweidimensional sind.

6.3.2b Der Übergang zu Funktionen der Koordinaten

(3.2.6) Beim Versuch einer vollständigen Quantifizierung von $f:V \rightarrow W$ tritt das folgende Problem auf:

Auch in W werde eine konstante Basis b gegeben. Für den Funktionswert $f(\vec{x}) \in W$ soll eine Koordinatendarstellung entwickelt werden. Das ist mehr als nur eine Basisdarstellung des Vektors. Denn $f(\vec{x})$ hängt ja von $\vec{x} \in V$ ab und das ist der absolute Vektor \vec{x} , nicht aber der zur Basis e gehörige Koordinatenvektor \vec{x}^K . Man muß die Abhängigkeit im Urbildbereich wechseln, also von f zu $f \circ L_e$ übergehen, was man stets durch Gleichungen der folgenden Art erreicht.

(3.2.7) Koordinatendarstellung der unabhängigen Variablen eines Funktionswertes:

$$\boxed{f(\vec{x}) = f(\Sigma \vec{e}_r x_r) = (f \circ L_e)(x_1, x_2, \dots, x_n) = (f \circ L_e)(\vec{x}^K)}.$$

$$\begin{aligned} \text{Etwa} \quad & : \quad s(\vec{x}) = \vec{x}^2 = (\vec{e}_1 x_1 + \vec{e}_2 x_2 + \vec{e}_3 x_3)^2 = x_1^2 + x_2^2 + x_3^2 = s^K(x_1, x_2, x_3). \quad \text{Oder} \\ \vec{f}(\vec{x}) & = \vec{a}(\vec{x}) \times \vec{x} = \vec{e}_1(a_2(\vec{x})x_3 - x_2 a_3(\vec{x})) + \vec{e}_2(\dots) + \vec{e}_3(\dots) \\ & = \vec{e}_1(A_2(x_1, x_2, x_3)x_3 - x_2 A_3(x_1, x_2, x_3)) + \dots \\ & = \vec{e}_1(A_2(\vec{x}^K)x_3 - x_2 A_3(\vec{x}^K)) + \dots \end{aligned}$$

Links steht der Wert als Funktion des absoluten Vektors. In der Funktion wird der absolute Vektor \vec{x} durch seine Koordinatendarstellung $\Sigma \vec{e}_r x_r$ ausgedrückt und der entstehende Rechenausdruck eventuell vereinfacht. Das Ergebnis wird als Funktion des Koordinatentupels (=Koordinatenvektor \vec{x}^K) interpretiert. Und das ist ja $f \circ L_e$. Der Koordinatenvektor erhält die Rolle einer unabhängigen Variablen, was dann die Koordinatenfunktion ergibt. \vec{x} auf der anderen Seite der Gleichung wird dabei zur Hilfsgröße. Beachten Sie die Endgleichung $\boxed{f(\vec{x}) = f \circ L_e(\vec{x}^K)}$ mit deren Hilfe und entsprechender Rollenzuweisung man beide Formen jederzeit ineinander umwandeln kann.

(3.2.8) Vgl. auch Kap. 6.1.4. Der dort behandelte Fall ist viel allgemeiner. Wählt man als Parametrisierung $\pi = \vec{x}^\pi$ die Linearkombinationsabbildung L_e der gegebenen Basis e , dann kommt man zum augenblicklichen Fall. Dabei ist einfach $f \circ \vec{x}^\pi = f \circ L_e$. Also in Worten: $f \circ L_e(\vec{x}^K) = f(L_e(\vec{x}^K)) = f(\vec{x})$. Der Index K steht hier für Koordinatendarstellung des Urbildes, wogegen er in (1.4.27) auch die Koordinatendarstellung des Wertes einschloß.

(3.2.9) Auf der Ebene der Koordinatendarstellung sollten möglichst nur derartige Koordinatenfunktionen auftreten! Also \vec{x}^K als Argument, nicht aber \vec{x} . Bei der Quantifizierung der Ableitung ist es zumindest vorläufig empfehlenswert, auf die Identifizierung von \vec{x} und \vec{x}^K verzichten, da eine solche unangenehme Konsequenzen hat.

(3.2.10) Anwenden von (3.2.7) auf eine **Vektorfunktion** $f:G \rightarrow W$ ergibt nach Wahl einer Basis e für V und einer Basis b für W insgesamt die folgende Umrechnung, die analog zur Fundamentalidentität linearer Abbildungen ist:

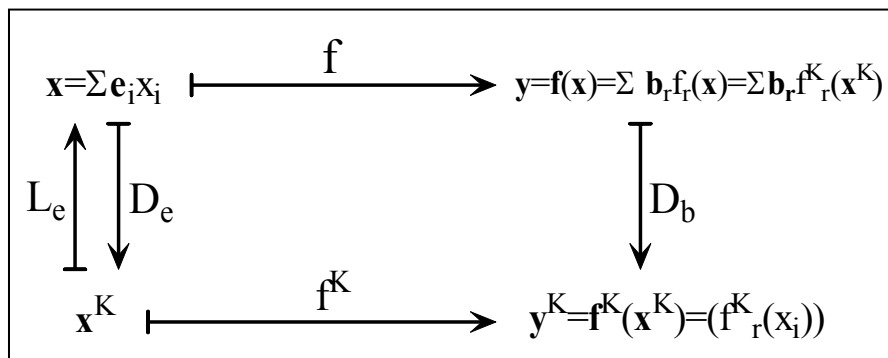
$$\vec{f}(\vec{x}) = \Sigma_{r=1}^m \vec{b}_r f_r(\vec{x}) = \Sigma_{r=1}^m \vec{b}_r (\Sigma_i \vec{e}_i x_i) = \Sigma_{r=1}^m \vec{b}_r f_r^K(x_1, \dots, x_n) = \Sigma_{r=1}^m \vec{b}_r f_r^K(\vec{x}^K).$$

(3.2.11) Nun müssen wir auch im Werteraum W zum Koordinatenvektor (=Spalte der Feldkomponenten!) übergehen und erhalten damit die voll quantifizierte Form einer Vektorabbildung, wie sie typischerweise in den Anwendungsgebieten formuliert wird (n-tupel der $x_i \mapsto m$ -tupel der y_r). Vgl. (1.7.3):

$$\vec{f}^K(\vec{x}^K) = {}^t(f_1^K(\vec{x}^K), \dots, f_m^K(\vec{x}^K)) = (f_r^K(\vec{x}^K))_{r=1,2,\dots,m} = (f_r^K(\vec{x}^K))$$

In der Koordinatendarstellung wird alles als Funktion der Koordinaten interpretiert. Das Anbringen des oberen Index K für *Koordinatendarstellung* bringt dies zum Ausdruck. Erneut sollte man sich den Sachverhalt

in der üblichen Diagrammform verdeutlichen.



(3.2.12) Jetzt können wir uns der Quantifizierung der totalen Ableitung zuwenden, also (3.2.7) fortführen. Wir benötigen den Zusammenhang zwischen den partiellen (eindimensionalen) Ableitungen von f und den Koordinatenfunktionen f^K . Es genügt, sich den Sachverhalt am Beispiel $n=3$ zu verdeutlichen. Wir beginnen mit $\vec{f}(\vec{e}_1x + \vec{e}_2y + \vec{e}_3z) = \vec{f}^K(x, y, z)$. Wir ersetzen etwa y durch $y + \Delta y$ und entwickeln beide Seiten dieser Gleichung. Es folgt die Gleichheit der beiden partiellen Ableitungen nach y ! Allgemein zeigt das die Gültigkeit der folgenden Beziehungen:

$$D\vec{f}(\vec{x}_0) \cdot \vec{e}_j = \frac{\partial \vec{f}}{\partial x_j}(\vec{x}_0) = \frac{\partial \vec{f}^K}{\partial x_j}(\vec{x}_0^K) \quad j=1,2,\dots,n.$$

(3.2.13) Beachten Sie den folgenden Sachverhalt: Wenn man die Basis abändert und zwar so, dass der i -te Basisvektor fest bleibt, dann ändert sich die linke Seite, also die partielle Ableitung der absoluten Funktion f nicht. Aber die gesamte Koordinatenfunktion f^K kann sich ändern einschließlich der Abhängigkeit von der Koordinaten x_i . Und natürlich kann auch der Aufpunktvektor \vec{x}^K seinen Wert ändern. Der Index K beschreibt diese Abhängigkeit vom **gesamten** Koordinatensystem ja gerade.

(3.2.14) Zur Verdeutlichung können Sie in der Ebene die beiden Basen (\vec{e}_1, \vec{e}_2) und $(\vec{f}_1 = \vec{e}_1, \vec{f}_2 = \vec{e}_1 + \vec{e}_2)$ einführen, für $s(\vec{x}) = \vec{x}^2$ die zugehörigen Koordinatendarstellungen ausrechnen und nach der ersten Koordinate partiell differenzieren. (Ergebnis: $s^K(x_1, x_2) = x_1^2 + x_2^2$ und $s^L(u_1, u_2) = u_1^2 + 2u_1u_2 + 2u_2^2$. Also $\frac{\partial s^K}{\partial x_1}(x_1, x_2) = 2x_1$ und $\frac{\partial s^L}{\partial u_1}(u_1, u_2) = 2u_1 + 2u_2$, obwohl $\vec{e}_1 = \vec{f}_1$ war.)

(3.2.15) In der klassischen Thermodynamik wird dieser Sachverhalt der Abhängigkeit von allen Koordinaten relevant und erklärt viele dort übliche Schreibweisen, die erforderlich werden, weil es üblich ist, auf den Koordinatenindex zu verzichten.

6.3.2c Wahl einer Basis in Urbildraum und Werteraum (Konstante Basisfelder).

(3.2.16) Nach diesen Vorbereitungen können wir die totale Ableitung vollständig quantifizieren. Nochmals das Szenenbild: Sei e eine (konstante Basis) des Urbildes V und b eine Basis des Werteraumes W . Zusammen ergibt das eine volle Quantifizierung \vec{f}^K von \vec{f} wie bereits besprochen ebenso wie eine Quantifizierung der zugehörigen totalen Ableitung $Df(x)$ für alle Punkte x aus der offenen Menge G . Das Vorgehen zur Bestimmung der zugehörigen Matrixdarstellung der totalen Ableitung ist klar. Wir finden einerseits

$$\vec{f}(\vec{x}) = \sum \vec{b}_r f_r(\vec{x}) = \sum \vec{b}_r f_r^K(\vec{x}^K),$$

andererseits berechnen wir das totale Differential mit der Fundamentalidentität

$$D\vec{f}(\vec{x}) \cdot \Delta \vec{x} = \sum_{r,i} (\vec{b}_r Df_r(\vec{x}) \cdot \vec{e}_i) \Delta x_i.$$

Jetzt verwenden wir (3.2.12) an, um zur partiellen Ableitung der Koordinatenfunktionen überzugehen.

(3.2.17) Ergebnis:

$$d\vec{f} = \sum_{r,i} \vec{b}_r (Df_r(\vec{x}) \cdot \vec{e}_i) \Delta x_i = \sum_{r,i} \vec{b}_r \frac{\partial f_r^K}{\partial x_i}(\vec{x}^K)$$

(3.2.18) Das ergibt unmittelbar die folgende beschreibende Matrix für $Df(\vec{x})$, die man üblicherweise **Jacobimatrix oder Funktionalmatrix (der Abbildung f in \vec{x}^K)** nennt:

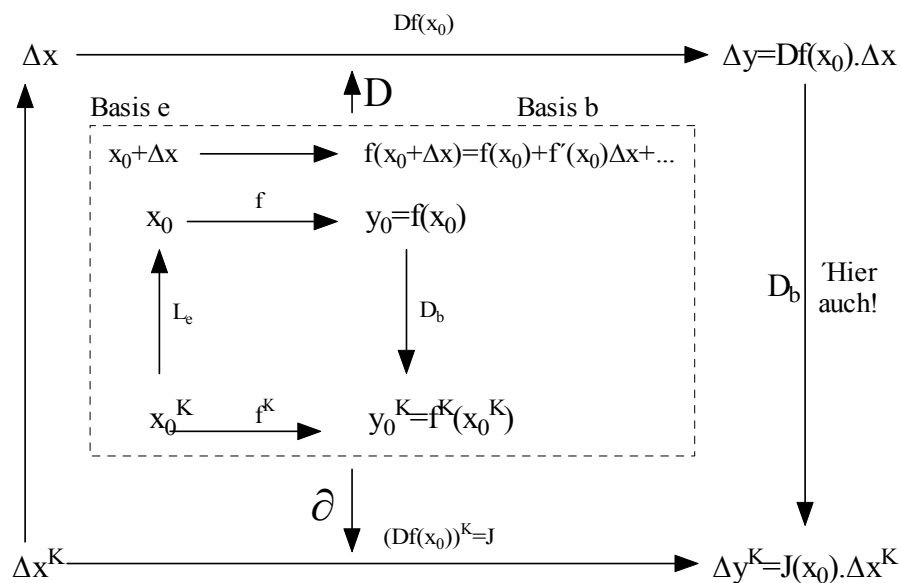
$J_{ri}(\vec{x}^K) = \frac{\partial f_r^K}{\partial x_i}(\vec{x}^K)$	Vorsicht mit der Indexreihenfolge! Zuerst die Feldkomponente.
--	--

Teilweise schreibt man ∂_i für $\frac{\partial}{\partial x_i}$. Dann hat man $J_{ri} = \partial_i f_r$. Die Jacobimatrix ist also die $m \times n$ -Matrix der partiellen Ableitungen und zugleich die beschreibende Matrix der totalen Ableitung $Df(\vec{x})$ in den gewählten Koordinaten.

(3.2.19) Fassen wir alles zusammen, wobei wir parallel das konkrete Beispiel $f = (V^2, \vec{x} \mapsto a^2 \vec{x} - 5(\vec{a} \cdot \vec{x})\vec{x}, V^2)$ geben:

1)	Im Urbildraum und im Werteraum je eine konstante Basis wählen.	In beiden Räumen \vec{e}_1, \vec{e}_2 kartesisch, so dass $\vec{a} = \vec{e}_1 a$				
2)	Von f zur Koordinatendarstellung f^K übergehen	$f^K(\vec{x}^K) = \begin{pmatrix} f_1^K(x, y) \\ f_2^K(x, y) \end{pmatrix} = \begin{pmatrix} a^2 x - 5ax^2 \\ a^2 y - 5axy \end{pmatrix}$				
3)	Die Komponentenfunktionen f_r^K partiell nach x_i ableiten.	<table border="1" style="width: 100%; border-collapse: collapse;"> <tr> <td style="padding: 2px;">$\frac{\partial f_1^K}{\partial x}(x, y) = a^2 - 10ax$</td> <td style="padding: 2px;">$\frac{\partial f_1^K}{\partial y}(x, y) = 0$</td> </tr> <tr> <td style="padding: 2px;">$\frac{\partial f_2^K}{\partial x}(x, y) = -5ay$</td> <td style="padding: 2px;">$\frac{\partial f_2^K}{\partial y}(x, y) = a^2 - 5ax$</td> </tr> </table>	$\frac{\partial f_1^K}{\partial x}(x, y) = a^2 - 10ax$	$\frac{\partial f_1^K}{\partial y}(x, y) = 0$	$\frac{\partial f_2^K}{\partial x}(x, y) = -5ay$	$\frac{\partial f_2^K}{\partial y}(x, y) = a^2 - 5ax$
$\frac{\partial f_1^K}{\partial x}(x, y) = a^2 - 10ax$	$\frac{\partial f_1^K}{\partial y}(x, y) = 0$					
$\frac{\partial f_2^K}{\partial x}(x, y) = -5ay$	$\frac{\partial f_2^K}{\partial y}(x, y) = a^2 - 5ax$					
4)	Damit hat man die Komponenten J_{ri}^K der beschreibenden Matrix von $Df(\vec{x})$ in den gewählten Koordinaten	$J^K(x, y) = (J_{ri}^K(x, y)) = \begin{pmatrix} a^2 - 10ax & 0 \\ -5ay & a^2 - 5ax \end{pmatrix}$				

Natürlich lassen sich die insgesamt schon recht komplizierten Zusammenhänge am besten wieder in einem Diagramm überblicken. Dabei handelt es sich um eine Ergänzung des Diagramms aus (3.2.11). Im Innern di im allgemeine nicht lineare Abbildng f, außen die zugehörigen Ableitungen.



□ Rechnen Sie die beiden Beispiele aus der Frage unter (3.2.8).

6.3.2d Bekannte Spezialfälle der totalen Ableitung

Für einige Abbildungstypen besitzt man eine herkömmliche Definition der Ableitung ohne Verwendung der linearen Algebra. Wie hängt diese herkömmliche Definition mit der totalen Ableitung zusammen?

(3.2.20) Beginnen wir mit Abbildungen vom Kurventyp, also $n = \dim V = 1$. Dann haben wir in $V = \mathbb{R}^1$ die kanonische Basis $e = (1)$. Jeder Zuwachsvektor schreibt sich $(\Delta t) = (1)\Delta t$. Unter Verwendung der Linearität der totalen Ableitung folgt: $Df(t_0) \cdot \Delta t = (Df(t_0) \cdot (1))\Delta t$. Vergleich mit der zugehörigen Tangenterzerlegung zeigt $f'(t_0) = \dot{f}(t_0) = Df(t_0) \cdot (1)$. Die herkömmliche Ableitung ($m=1$) bzw. der momentane Geschwindigkeitsvektor $\dot{f}(t_0)$ ist gleich dem Wert der totalen Ableitung auf dem Basisvektor. Oder auch intuitivunterstützend: **Die (vektorielle) Geschwindigkeit ist gleich der Ortsänderung df während einer Zeiteinheit $\Delta t=1$.**

Haben wir in W ein konstantes Basisfeld b , so dass $f(t) = \sum b_j f_j(t)$ gilt, so folgt

$$f'(t) = \sum b_j (Df_j(t) \cdot 1) = \sum b_j f'_j(t).$$

Das ist ein weiteres bekanntes Resultat: **Die Geschwindigkeitskomponenten sind die Ableitungen der Koordinatenfunktionen.**

(3.2.21) Den zweiten wichtigen Typ bilden die **Skalarfelder**. Hier beschreiben wir die Feldänderung üblicherweise mit Hilfe des Gradienten und das ist ein geometrischer Vektor. Für das totale Differential finden wir zunächst $ds = Ds(\vec{x}_0) \cdot \Delta \vec{x}$. Der Gradient ist nun definitionsgemäß ein Vektor, für den $ds = \text{grad}(x) \cdot \Delta \vec{x}$ gilt. Vgl. (3.1.11). Führt man kartesische Koordinaten ein, dann ist die beschreibende Matrix von $Ds(\vec{x}_0)$ ein Zeilenvektor. Macht man daraus einen Spaltenvektor, so erhält man die Koordinatendarstellung des Gradienten! Denn bei dieser Operation wird aus dem Matrixprodukt Zeile \times Spalte die Komponentenformel für das Skalarprodukt. Genauer hierzu in Kapitel 11.

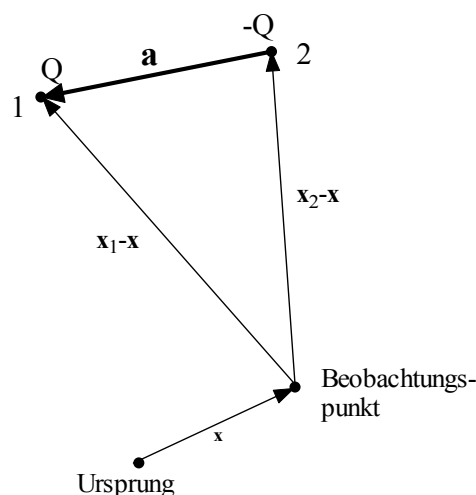
6.3.2e Das Potential eines mathematischen Dipols

(3.2.22) Das folgende Beispiel gibt eine Anwendung der zweiten Denkfigur, also des Einsetzens einer Tangenterzerlegung sowie der Verwendung von Grenzprozessen beim Erzeugen einer Idealisierung eines physikalischen Systems.

(3.2.23) Wir betrachten zwei Punktladungen mit entgegengesetzt gleichen Ladungen Q und $-Q$. Für die Ortsvektoren der beiden Ladungen gelte $\vec{x}_1 = \vec{x}_2 + \vec{a}$. Damit zeigt \vec{a} von $-Q$ nach Q . Nach dem Superpositionsprinzip entsteht folgendes resultierendes Gesamtpotential:

$$\vec{x} \mapsto V(\vec{x}) = \frac{Q}{\|\vec{x} - \vec{x}_1\|} - \frac{Q}{\|\vec{x} - \vec{x}_2\|} = \frac{Q}{\|(\vec{x} - \vec{x}_2) - (\vec{x}_1 - \vec{x}_2)\|} - \frac{Q}{\|\vec{x} - \vec{x}_2\|}$$

\vec{x} sei der Ortsvektor des Punktes, an dem wir das von den Ladungen erzeugte Feld beobachten. Uns interessiert der Fall, dass $\|\vec{x} - \vec{x}_1\|$ und $\|\vec{x} - \vec{x}_2\|$ sehr viel größer als $\|\vec{a}\|$ ist. Oder auch: Die Beobachtungsstelle ist weiter von den Ladungen, entfernt als diese voneinander entfernt sind.



Zu diesem Zweck setzen wir für den ersten Summanden dessen Tangentenerlegung um $\vec{x}-\vec{x}_2$ ein. Also $\Delta\vec{x} = \vec{a}$. Wir kennen den Gradienten von $s(\vec{x}) = \frac{1}{\|\vec{x}\|}$. Der Gradient ist $-\frac{\vec{x}}{\|\vec{x}\|^3}$. Damit erhalten wir sofort die gewünschte Tangentenerlegung:

$$\frac{1}{\|(\vec{x}-\vec{x}_2)-\vec{a}\|} = \frac{1}{\|(\vec{x}-\vec{x}_2)\|} + Ds(\vec{x}-\vec{x}_2)\cdot(-\vec{a}) + \dots \quad \text{mit} \quad Ds(\vec{x}-\vec{x}_2)\cdot(-\vec{a}) = -\frac{(\vec{x}-\vec{x}_2)\cdot(-\vec{a})}{\|(\vec{x}-\vec{x}_2)\|^3}.$$

Nach Einsetzen dieser Formel in V hebt sich der erste konstante Term heraus! Man behält:

$$V(\vec{x}) = +QDs(\vec{x}-\vec{x}_2)\cdot\vec{a} + \dots = (Q\|\vec{a}\|)Ds(\vec{x}-\vec{x}_2)\cdot\vec{e} + \dots \quad \text{mit} \quad \vec{e} = \frac{\vec{a}}{\|\vec{a}\|}$$

Der fortgelassene nur angedeutete Restterm hat die Gestalt $Q\|\vec{a}\|R(\dots,\vec{a})$, wobei der Faktor R nach Null geht für $\|\vec{a}\|$ nach Null. Wir setzen $D=Q\|\vec{a}\|$ und lassen $\|\vec{a}\|$ nach Null gehen und Q nach unendlich: Und zwar so gekoppelt, dass das Produkt D dabei fest bleibt. Dann geht der Restterm immer noch nach Null, wie wir gesehen haben!

(3.2.21) Physikalisch erfaßt die Grenzwertbildung die folgende **Idealisierung**: Die Größe beider Ladungen wächst über alle Grenzen, dafür nähern sie sich bei fester Richtung immer mehr an, so dass sich ihre Wirkungen infolge des entgegengesetzten Vorzeichens teilweise kompensieren. Wir haben es mit einem Grenzprozeß für eine Produktgröße zu tun, bei der ein Faktor nach Null geht, der andere über alle Grenzen wächst. Unsere daraus berechneten Größen Q und V haben bei diesem Prozeß jedoch einen wohlbestimmten Grenzwert. In unmittelbarer Nähe der beiden Ladungen bewirkt dieser Prozeß enorme Änderungen des Feldes, während man weiter außen rasche Konvergenz vorfindet. Als Ergebnis dieses Grenzprozesses kommt ein wohlbestimmtes und auch einfaches Potential heraus. Der Einfachheit halber legen wir das neue Objekt in den Ursprung, wählen also $\vec{x}_2=\vec{0}$. Das Objekt wird durch die Beschreibungsgröße $\vec{D} = D\vec{e}$ festgelegt, also durch eine Richtung und einen Betrag. Man bezeichnet es als *mathematischen Dipol*. Im Gegensatz dazu ist ein *physikalischer Dipol* ein System aus zwei entgegengesetzten Ladungen mit **endlichem** Abstand a. Während ein gewöhnlicher Massenpunkt ein idealisiertes System beschreibt, für das alle Richtungen gleichwertig sind, haben wir es beim mathematischen Dipol mit einem punktförmigen Objekt zu tun, das eine ausgezeichnete Richtung besitzt: Es ist die Richtung \vec{e} , die während des Grenzprozesses von der negativen zur positiven Ladung zeigte. Im Ursprung selbst (=Ort des Dipols) ist das Feld nicht definiert.

(3.2.22) Für das Potential ergibt die Konstruktion den folgenden Grenzwert:

$V_{\vec{D}}(\vec{x}) = \frac{\vec{D}\cdot\vec{x}}{\ \vec{x}\ ^3}.$	Das Potential eines mathematischen Dipols im Ursprung mit Dipolmoment \vec{D} .
---	---

Bei der Grenzwertbildung kann man \vec{x} als äußeren Parameter ansehen, so dass sicherlich punktweise Konvergenz vorliegt. Beachten Sie, das $V_{\vec{D}}(\alpha\vec{x}) = \frac{1}{\alpha^2}V_{\vec{D}}(\vec{x})$, wogegen für das Coulompotential ein Faktor $\frac{1}{\alpha}$ entsteht. (Homogen vom Grade -2 bzw. -1. Das Coulombkraftfeld dagegen ist homogen vom Grade -2. Vgl. Kap.6.1.7).

(3.2.23) Mit Hilfe von Gradientenbildung erhält man das Kraftfeld des Dipolmomentes:

$$\vec{F}_{\vec{D}}(\vec{x}) = \frac{\vec{x}^2\vec{D} - 3(\vec{D}\cdot\vec{x})\vec{x}}{\|\vec{x}\|^5} \quad \text{Kraft, die ein mathematischer Dipol im Ursprung am Ort } \vec{x} \text{ erzeugt.}$$

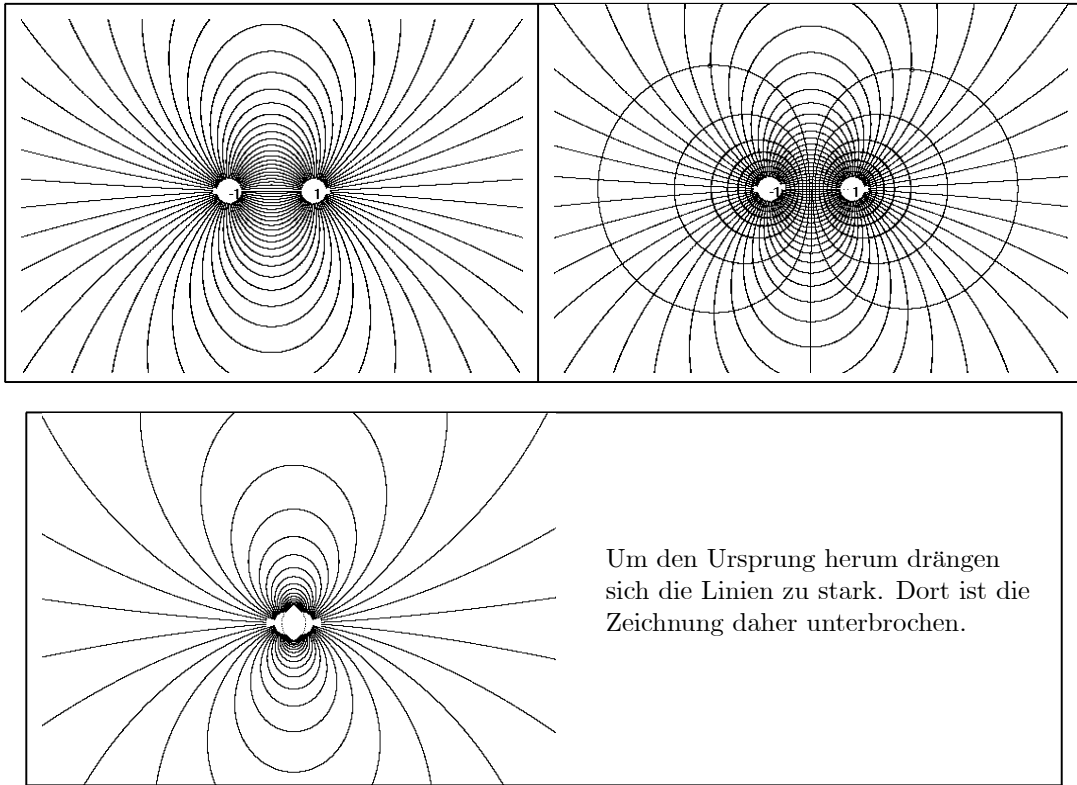
□ Wie lautet das Kraftfeld, wenn der Dipol sich am Orte \vec{A} befindet?

(3.2.24) Beide Felder - Potentialfeld und Kraftfeld - sind drehsymmetrisch bezüglich der durch \vec{D} erzeugten Achse. Daher genügt zur Veranschaulichung das Feld in einer den Vektor \vec{D} enthaltenden Ebene. Will man Koordinaten einführen, so sollte man möglichst eine Achse in die Richtung von \vec{D} legen. Nehmen wir kartesische Koordinaten und legen die z-Richtung in Richtung von \vec{D} - Also $\vec{D}=\vec{e}_3D$. Dann folgt mit $r=\sqrt{x^2+y^2+z^2}$

$V_D^K(x,y,z) = \frac{Dz}{r^3} \quad \text{und} \quad \vec{F}_D^K(x,y,z) = \frac{D(x^2+y^2-2z^2)\vec{e}_3 - 3Dz(x\vec{e}_1+y\vec{e}_2)}{r^5}$

- Stellen Sie die beiden Felder entsprechend in Polar- und Zylinderkoordinaten dar.

(3.2.25) Die beiden ersten der nachfolgenden Bilder zeigen die Feldlinien für endlichen horizontalen Abstand \vec{a} , also vor dem Grenzübergang in der Nähe des Ursprungs. Im zweiten Bild sind die Orthogonaltrajektorien mit eingezeichnet. Das dritte Bild zeigt das Feld in einer gewissen Entfernung vom Ursprung. Der unmittelbare Bereich um den Ursprung ist nicht gezeichnet, so dass das Bild wahlweise den Zustand im gezeichneten Bereich kurz vor dem Grenzübergang oder nach diesem -also für den mathematischen Dipol - wiedergibt. Der Unterschied ist im gezeichneten Bereich und bei der durch die Bildauflösung bestimmten Genauigkeit nicht zu erkennen.



6.3.3 Lokale Abbildungsgeometrie um den Aufpunkt $(\mathbf{x}_0, f(\mathbf{x}_0))$

Der Schluß von der Rechnung auf die Geometrie

Zu jeder Abbildung f gehören gewisse durch die Abbildung festgelegte Teilmengen sowohl des Urbildraumes G als auch des Werteraumes W . Im Urbildraum sind das die Niveaumengen $f^{-1}(\{y\})$ und im Werteraum die Menge $\text{Bild}f$ sowie die Bilder glatter Teilmengen von G , also Mengen des Typs $f(U)$. Der Beschreibung dieser Mengen wenden wir uns jetzt zu. Dabei hat man es immer mit zwei Problemrichtungen zu tun: (◆) Für geometrisch gegebene Mengen einen Rechenausdruck finden und (◆) einen gegebenen Rechenausdruck geometrisch zu interpretieren.

(3.3.1) Für (rechnerisch) glattes f erwarten wir, dass die durch f bestimmten Teilmengen auch geometrisch glatt sind. Was soll geometrisch glatt genauer bedeuten? Nun die betrachteten Mengen sind in der Regel nicht vom linearen Typ, sind gekrümmt. (Vgl. Kap 4 (1.5.22-28)). Aber in kleinen Umgebungen um jeden Punkt der Menge sollte man sie ohne großen Fehler durch ein Stück einer linearen Menge - also eines Teilraumes - ersetzen können, so wie man den Graphen einer glatten Funktion lokal durch ein Stück der Tangente ersetzen kann, wie man ein Stück der gekrümmten Erdoberfläche durch ein Ebenenstück idealisiert approximieren kann. Ist also M eine glatte Teilmenge von V bzw. W und $a \in M$, dann suchen wir einen Teilvektorraum T von V bzw. W derart, dass der affine Raum $a+T$ die Menge M um a herum approximiert. Wir nennen

diesen bis jetzt hypothetischen Raum *den Tangentialraum an M zum Punkte a*. Mit seiner Hilfe können wir dann lokal geometrische Probleme für die Mengen auf Probleme der linearen Algebra zurückführen.

(3.3.2) Damit entstehen aber einige (vornehmlich mathematische) Probleme: Gibt es einen solchen Tangentialraum überhaupt? Wenn ja, welche Dimension hat er? Und wie kann man ihn konkret bestimmen, insbesondere im Falle der Mengen, die in der angedeuteten Weise mit einer glatten Abbildung f zusammenhängen?

(3.3.3) Zur Orientierung sollte man an glatte Flächen im Raum denken. Dort erhält man offenbar **Tangentialebenen** (Dimension 2) als Tangentialräume. Nehmen Sie etwa eine Kugeloberfläche. Entsprechend sind für glatte Kurven die Tangentialräume Geraden.

6.3.3a Präzisierung des Tangentialraumbegriffes

(3.3.4) Man kann auch sagen: Der Tangentialraum soll die Vereinigungsmenge aller Tangenten an den betrachteten Punkt sein. Und jede Tangente ist eine Gerade. Damit sind wir bei der Frage: **Wie erklärt man allgemein, was eine Tangente an M im Punkte \vec{a} ist?** Also für beliebige Dimension und beliebige glatte Menge? Hier bietet die Physik unmittelbar eine Idee für eine Antwort an: **Tangentenvektor gleich momentaner Geschwindigkeitsvektor!** Genauer: Man bildet eine glatte Bahnkurve $\vec{r} = (\mathbb{R}, t \mapsto \vec{r}(t), M)$ mit $\vec{r}(0) = \vec{a}$. Also eine Kurve, deren Bild ganz in M verläuft und die durch \vec{a} geht. Ist diese Kurve differenzierbar mit momentaner Geschwindigkeit $\vec{r}'(0) = \vec{v}$, dann werden wir \vec{v} als einen Tangentenvektor interpretieren. Diese Konstruktion liefert im Anschauungsraum das Gewünschte und ist problemlos auf den allgemeinen Fall ausdehnbar. Denken Sie an das Beispiel einer Kugeloberfläche. Für jeden so bestimmten Vektor \vec{v} geht der Abstand zwischen Kurvenpunkt und Tangentenapproximation stärker als linear gegen Null.

Die Gesamtheit aller möglichen Tangentenvektoren (= Geschwindigkeitsvektoren) soll unseren geometrischen Tangentialraum bilden! Das ist ein Teilvektorraum des einbettenden Raumes, den wir mit $T_{\vec{a}}$ bezeichnen, bei Bedarf auch $T_{\vec{a}}(M)$, um die zugehörige Menge anzugeben. (Hier findet eine typische Mengendefinition im Sinne von Kapitel 1.1 statt! Dass zumindest bei glatten Mengen ein Teilvektorraum vorliegt, ist noch nachzuweisen. Genauer in Kap. 16). Der geometrische Tangentialraum selbst entsteht durch Parallelverschiebung, ist also $\vec{a} + T_{\vec{a}}$. Beide Objekte sollte man nicht verwechseln. Für $\vec{a} \neq \vec{0}$ ist $\vec{a} + T_{\vec{a}}$ i.a. kein Vektorraum.

(3.3.5) Damit wäre abstrakt geklärt, was ein Tangentenvektor und der aus diesen gebildete Tangentialraum sein sollen, aber eine tatsächliche Bestimmung erscheint zunächst schwierig. Für Mengen M , die in der beschriebenen Weise mit einer glatten Abbildung f zusammenhängen, erweist sich nun aber eine Bestimmung der Tangentialräume über die totale Ableitung als vielfach möglich, ja als leicht. Dieser Schluß von der Rechnung auf die Geometrie soll nachfolgend besprochen werden.

Zu beachten ist dabei stets, dass der Aufpunkt \vec{a} als äußerer Parameter in die Konstruktion eingeht. D.h. die Konstruktion ist für jeden Punkt vorzunehmen und das Ergebnis wird von \vec{a} abhängen. Die tangentialen Richtungen können sich von Punkt zu Punkt ändern, wie das Beispiel der Kugeloberfläche sofort zeigt. Formal bildet $\vec{a} \mapsto T_{\vec{a}}$ daher ein *Teilraumfeld!*

6.3.3b Die Bestimmung von Tangentialräumen mit Hilfe der totalen Ableitung.

(3.3.6) Im Urbildraum legt die Abbildung f die **Niveaumengen** $N(\vec{x}_0) = \{\vec{x} | f(\vec{x}) = f(\vec{x}_0)\}$ fest. Dabei ist \vec{x}_0 für $f: G \rightarrow W$ ein beliebiges Element aus G . D.h. inhaltlich: $N(\vec{x}_0)$ ist die Menge aller Punkte, in denen dieselbe Feldstärke herrscht wie in \vec{x}_0 selbst. Diese Mengen bilden üblicherweise eine glatte, aber häufig komplizierte geometrische Figur in V , deren Form und Größe interessiert. Ein typisches zugehöriges Anwendungsproblem sieht so aus: In welche Richtungen $\Delta\vec{x}$ kann / muß man von \vec{x}_0 aus gehen, so dass sich die Feldstärke nicht ändert? Das sind Richtungen, für die näherungsweise $f(\vec{x}_0 + \Delta\vec{x}) = f(\vec{x}_0)$ gilt. Wertet man die linke Seite in Tangentenapproximation aus, folgt $Df(\vec{x}_0) \cdot \Delta\vec{x} = \vec{0}$. Und d.h. gerade: **Die gesuchten Richtungen werden durch den Kern von $Df(\vec{x}_0)$ bestimmt!** Hat man eine Basis von $\text{Kern} Df(\vec{x}_0)$, so hat man eine Basis für den jeweiligen Tangentialraum an die zugehörige Niveaumenge. Genauer: Einen Kandidaten für eine Basis.

(3.3.7) Wie hängen die Elemente $\Delta\vec{x}$ aus Kernf und die geometrischen Tangentenvektoren an die Niveaumenge zusammen? Eine Richtung (also eine Mengeninklusion, Kap. 1) ist einfach: Sei $(\mathbb{R}, t \mapsto \vec{\gamma}(t), N)$ eine ganz in N verlaufende Kurve, die einen geometrischen Tangentenvektor $\vec{h} = \dot{\vec{\gamma}}(0)$ in $\vec{\gamma}(0) = \vec{a}$ erzeugt. Dann ist $f(\vec{\gamma}(t)) = c$, dem konstanten Niveaufwert. Die totale Ableitung liefert

$$f(\vec{\gamma}(0 + \Delta t)) = f(\vec{\gamma}(0)) + Df(\vec{\gamma}(0)) \cdot \vec{h} \Delta t + \dots$$

Da $\vec{\gamma}$ aber in der Niveaumenge von f verläuft, heben sich die ersten beiden Beiträge fort. Division durch Δt und anschließender Grenzwert Δt nach Null gibt wegen $\vec{a} = \vec{\gamma}(0)$ unmittelbar $Df(\vec{\gamma}(0)) \cdot \vec{h} = \vec{0}$. D.h. der geometrische Tangentenvektor ist automatisch im Kern gelegen. Oder $T \subset \text{KernDf}(\vec{a})$.

Sei N Niveaumenge einer glatten Abbildung f mit $a \in N$. Dann ist $T_a \subset \text{KernDf}(a)$

Die Bestimmungsgleichung für den Kern schreibt sich $Df(x_0) \cdot \Delta x = 0$. Oder mit dem totalen Differential $df=0$. Das ist merktechnisch günstig.

Die umgekehrte Mengeninklusion dagegen ist problematisch. Offenbar genügt es, zu zeigen, dass $\dim T_{\vec{a}} = \dim \text{KernDf}(a)$ gilt. Wir werden darauf zurückkommen.

(3.3.8) In der Wertemenge W bestimmt f **Teilmengen in Form von Bildern** wie $\text{Bild}(f)$ und $\text{Bild}(f \circ r)$, wenn r eine vorgeschaltete glatte Abbildung, etwa eine Restriktionsabbildung ist. Ist f nicht surjektiv, dann bilden diese Bilder für glattes $f \circ r$ typischerweise eine glatte geometrische Figur in W. Der Aufpunkt \vec{x}_0 bestimmt mit $\vec{y}_0 = f(\vec{x}_0)$ einen Punkt auf dieser Figur. Hier stellt sich die Frage: Für welche Richtungen $\Delta\vec{y}$ in W verbleibt man dann (zumindest näherungsweise) in $\text{Bild}(f)$ bzw. $\text{Bild}(f \circ r)$, für welche verläßt man das Bild?. Erstere Richtungen ergeben die Richtungen geometrischer Tangenten der jeweiligen Figur in \vec{y}_0 .

Alle diese Fälle können wir einheitlich wie folgt zusammenfassen: $\pi: G \rightarrow W$ sei eine glatte Abbildung vom Parametrisierungstyp, die die Menge $M = \text{Bild}\pi$ parametrisiert. G sei dabei offene Teilmenge eines geeigneten Vektorraumes U. (f und $f \circ r$ sind beide von diesem Typ mit eventuell unterschiedlichem U.)

Nachfolgende Überlegungen sollte man sich vom Transformationsstandpunkt aus veranschaulichen: Man bewegt sich vorgebar im Urbildraum nicht im Werteraum, verfolgt aber die zugehörige Bewegung in $\text{Bild}\pi$.

(3.3.9) **Was läßt sich über Tangenten von M sagen?** Die Tangentenapproximation liefert uns den Teilvektorraum $\text{BildD}\pi(\vec{a})$ von W mit Elementen, die alle die Tangentenerlegung für π erfüllen. (Bitte die Menge $\text{Bild}\pi \subset M$ und den Teilvektorraum $\text{BildD}\pi(\vec{a})$ nicht verwechseln!) Sei $t \mapsto \vec{c}(t)$ eine glatte Kurve im Urbildraum mit $c(0) = a$ und $\dot{c}(0) = h$. Wir bilden die zusammengesetzte Kurve $\pi \circ c$. Durch Differenzieren folgt

$$\pi(c(0 + \Delta t)) = \pi(a) + D\pi(a) \cdot h \Delta t + \dots$$

Also hat $\pi \circ c$ die momentane Geschwindigkeit $D\pi(a) \cdot h$. Aber $\pi \circ c$ ist eine Kurve, die unsere Bedingung für eine geometrische Tangente an $M = \text{Bild}\pi$ in $b = \pi(a)$ erfüllt: Sie verläuft in der Menge $M = \text{Bild}\pi$. Weiter geht sie durch den Punkt $\pi(a)$ mit Geschwindigkeit $D\pi(a) \cdot h$ und für diesen Vektor ist die Approximation an die Kurve stärker als linear! D.h. $D\pi(a) \cdot h$ ist geometrischer Tangentenvektor an M in b. Oder $\text{BildD}\pi(a) \subset T_b$.

Sei $\pi: G \rightarrow W$ glatte Parametrisierung von $M = \text{Bild}\pi$ sowie $b = \pi(a) \in W$.

Dann gilt $\text{BildD}\pi(a) \subset T_b$. Dabei ist $T_b = T_b(M)$ der Tangentialraum an M in b.

Erneut erweist sich die umgekehrte Inklusionsrichtung als problematischer.

(3.3.10) Beachten Sie die Praktikabilität dieser Resultate: Sei $f: V \rightarrow W$ gegeben mit $b = f(a)$. Bilde $Df(a)$ und dazu Kern und Bild. Dann liegt jeder Tangentenvektor an die Niveaumenge $N(a)$ in $\text{KernDf}(a)$ und jeder Vektor aus $\text{BildDf}(a)$ ist Tangentenvektor an $\text{Bild}f$ in b.

Schöner wäre natürlich, wenn man so **genau alle** Tangentenvektoren erhielte.

(3.3.11) Betrachten wir als erstes Beispiel die Feldfunktion $F(x,y) = \begin{pmatrix} y + x^2 \\ (y + x^2)^2 \end{pmatrix}$. Für dies F rechnet man die Niveaumengen ebenso wie das Bild sofort aus:

$$N_c = \left\{ \begin{pmatrix} x \\ y \end{pmatrix} \mid y = c - x^2 \right\} \quad \text{Eine Schar parallel verschobener Parabeln}$$

$$M = \text{Bild}(F) = \left\{ \begin{pmatrix} u \\ v \end{pmatrix} \mid v = u^2 \right\} \quad \text{Auch eine Parabel.}$$

Durch jeden Punkt ${}^t(x_0, y_0)$ des Urbildraumes geht genau eine Niveauparabel. Der Kern der oben bestimmten totalen Ableitung ist stets eindimensional, aber aufpunktabhängig. Man prüft sofort: Der in den Aufpunkt verschobene Kern ist tatsächlich die Tangente an die Niveauparabel in diesem Punkt. Die Rechnung ergibt in beiden Fällen korrekt den geometrischen Tangentialraum. D.h. hier gelten nicht nur die oben allgemein bewiesenen Inklusionen, sondern Gleichheit.

(3.3.12) Beachten Sie auch, dass der Dimensionssatz für Homomorphismen erfüllt sein muß: $\dim\text{Kern} + \dim\text{Bild} = \dim V$. In unserem Fall $1+1=2$. D.h. die beiden Tangentialräume sind immer per Dimension aneinander gekoppelt. Und die Dimensionen sind in der Regel nicht 1 sondern können auch größer sein: Ein dreidimensionaler Teilraum in einem fünfdimensionalen V . Usw. Auch Dimension Null ist wichtig. Ist etwa $\dim\text{Kern}(\pi) = 0$, liegt ein trivialer Kern vor und deutet auf ein lokal injektives f hin.

(3.3.13) Leider ist die soeben skizzierte Methode, mit Hilfe der totalen Ableitung geometrische Eigenschaften von Niveaumengen und Bildern zu bestimmen, nicht für alle Punkte anwendbar. Die Mengen können sich als verschieden erweisen.

Rechnen Sie als Beispiel etwa ${}^t(x, y, z) \mapsto x^2 + y^2$. Die z -Achse ist hier Niveaumenge (zum Wert Null). Die z -Achse ist als Gerade in jedem Punkt ihr eigener geometrischer Tangentialraum. Aber die totale Ableitung ist dort überall die Nullabbildung, hat also einen zu großen Kern! Oder nehmen Sie die Kurve $t \mapsto \vec{r}(t) = (t^3, t^3)$. Das Bild ist die von $(1,1)$ erzeugte Gerade, die erneut überall gleich ihrem Tangentialraum ist. Aber für $t=0$ ist $\text{Bild}D\vec{r}(0)$ zu klein, nur der Nullpunkt. Oder auch: $\text{Bild}D\vec{r}(0)$ ist nulldimensional, nicht eindimensional wie es die Geometrie verlangt.

6.3.3c Das Rangkriterium

(3.3.14) Man muß nach einer zusätzliche Forderung suchen, die die gutartigen Punkte von den Problempunkten trennt. Dabei sind keineswegs solche Punkte gemeint, in denen die Ableitung nicht existiert. Dort versagt das Verfahren selbstverständlich. Ohne Beweis geben wir ein Kriterium dafür an, wann die erwünschte Gleichheit gilt:

Kriterium zur Trennung der gutartigen Schafe von den Böcken unter den Punkten:
 Es sei $G \subset V$ offen und zusammenhängend und $f: G \rightarrow W$ differenzierbar, also für jedes $x_0 \in G$ differenzierbar.
 * Bestimme für jeden Punkt $x_0 \in G$ den Rang von $Df(x_0)$, also $\dim\text{Bild}Df(x_0)$.
 * Suche darunter den maximal auftretenden Wert r . Natürlich ist $r \leq \min(m, n)$.
 * Bilde die Menge $G_0 \subset G$ aller Punkte, in denen der Rang diesen maximalen Wert hat.
 G_0 offen. **Dann** ist G_0 die Menge aller gutartigen Punkte, für die man von der analytischen Rechnung auf den geometrischen Tangentialraum schließen darf.
 Also die Punkte, für die $T_a = \text{Kern}Df(a)$ und $T_{f(a)} = \text{Bild}Df(a)$ gilt.

Die Menge der Punkte aus G , die **nicht** in G_0 liegen, ist meist in dem Sinne klein, dass sie durch weniger Parameter als G selbst parametrisiert wird. Vielfach handelt es sich sogar um isolierte Punkte. Das lokale geometrische Verhalten für diese Punkte ist bei Bedarf gesondert zu analysieren. Man nennt die Menge der Punkte, die nicht in G_0 liegen, auch kritische Punkte von f und solche, die in G_0 liegen, reguläre Punkte von f .

(3.3.15) Ein **Beispiel:** Wir betrachten $Q = (\mathbb{R}^2, {}^t(x, y) \mapsto (x^2, y^2), \mathbb{R}^2)$. Direkte Tangenzenzerlegung gibt $DQ(x, y) = \begin{pmatrix} 2x & 0 \\ 0 & 2y \end{pmatrix}$. Der Rang dieser Abbildung ist außerhalb der Koordinatenachsen 2 und das ist der maximal auftretende Wert. Diese Punkte erfüllen das Kriterium. Die Niveaumengen bestehen aus einzelnen Punkten (Dimension Null). $\text{Bild}Q$ hat für all diese Punkte zweidimensionale Tangentialräume. Die Abbildungsgeometrie von Q ist leicht zu verstehen. Man faßt dazu die Zuordnung als Transformation auf. Die Urbildmenge wird um beide Achsen gefaltet, so dass alle Werte in den ersten Quadranten der Wertemenge fallen. Dort wird dann durch das Bilden der Quadrate noch eine Verzerrung vorgenommen (Zuordnungsstandpunkt). Auf den Achsen des Urbildraumes ist der Rang der Ableitung bis auf den Ursprung gleich 1. Im Ursprung ist er sogar Null. Dies sind gerade die kritischen Problempunkte. Die Achsenpunkte haben alle einen eindimensionalen Kern und entsprechend ein eindimensionales Bild, was der Faltungsprozedur entspricht. Die Niveaumengen bestehen trotzdem aus einzelnen Punkten, ihre Geometrie wird daher nicht richtig erfaßt. Im Ursprung ist der Kern zweidimensional, trotz der zugehörigen nulldimensionalen Niveaumenge.

- Prüfen Sie jetzt selbst die beiden Beispiele aus (3.3.13). Überzeugen Sie sich, dass es auch dort gerade die kritischen Punkte waren, in denen die Ableitung die geometrischen Tangentialräume nicht bestimmte.

Eine Beweis für das Kriterium findet man....

(3.3.16) **Zusammenfassung:** f sei in x_0 differenzierbar. Die Ableitung $Df(x_0)$ habe im Vergleich zu den übrigen Punkten von G maximalen Rang, d.h. x_0 sei kein kritischer Punkt. $N \subset G$ sei die Niveaumenge, die durch x_0 bestimmt wird. Dann besitzen die beiden Mengen N in x_0 und $\text{Bild}(f)$ in $y_0=f(x_0)$ je einen geometrischen Tangentialraum in Form eines affinen (=parallelverschobenen) Teilraumes. Diese beiden Räume erhält man mit Hilfe der totalen Ableitung wie folgt:

Der Tangentialraum von N zu x_0 ist $x_0 + \text{Kern}Df(x_0)$ D.h. löse $Df(x_0) \cdot \Delta x$ in Δx Kurz: $df=0$.	Der Tangentialraum von $\text{Bild}f$ in y_0 wird parametrisiert durch $\Delta x \mapsto f(x_0) + Df(x_0) \cdot \Delta x$ Also durch die Tangentenapproximation
--	--

Es sei $r=\ell=\text{Rang}Df(x_0)$ und $k=\dim\text{Kern}Df(x_0)$. Nach dem Dimensionssatz gilt stets: $n=\dim V=k+\ell$. Die Dimensionen der beiden Teilräume hängen daher zusammen. Insbesondere ist $k=0$ möglich. Dann bestehen die Niveaumengen aus isolierten Punkten. Oder auch:

Ist angegebene Rangbedingung erfüllt, der Punkt x_0 regulär und $Df(x_0)$ injektiv, dann verhält sich f selbst in der Nähe von x_0 näherungsweise wie die lineare Abbildung $Df(x_0)$. Alle Methoden und Resultate der linearen Algebra können eingesetzt werden. Insbesondere kann über eine geeignete Basiswahl der Tangentialraum quantifiziert werden.

(3.3.17) Möchte man daher zu einer vorgegebenen glatten Menge $M \subset V$ den Tangentialraum T_a zum Punkte a bestimmen, so geht man günstig wie folgt vor: Man sucht sich eine glatte Parametrisierung π von M , für die $a=\pi(c_0)$ ist und c_0 kein kritischer Punkt. Dann bildet man die totale Ableitung $D\pi(c_0)$ und erhält den gesuchten Raum über $T_a=\text{Bild}(D\pi(c_0))$. Damit sieht man auch sofort, dass T_a ein Teilvektorraum von V ist. Jedes Erzeugendensystem von $\text{Bild}D\pi(c_0)$ ist eines von T_a . Hat man insbesondere eine Koordinatendarstellung, dann liefern die Spalten der Jakobimatrix ein Erzeugendensystem. Und ist $D\pi(c_0)$ injektiv, dann bilden diese Spalten sogar eine Basis.

(3.3.18) **Beispiel: Der Tangentialraum an den Graphen.** Im Falle von Funktionen $f : \mathbb{R} \rightarrow \mathbb{R}$ erhält man mit Hilfe der Ableitung die Tangente an den Graphen der Funktion. Wie verallgemeinert sich das? Bisher wurden Tangenten an Niveaumengen und an Bilder behandelt. Das sieht nach einem wesentlichen Unterschied aus. Wir haben den Graphen nicht benutzt, weil der Graphenraum i.a. eine zu hohe Dimension hat, um noch anschaulich zugänglich zu sein. Aber es ist kein Problem, ihn formal als Teilmenge von $V \times W$ zu bilden. Wir haben für den Graphen jeder Abbildung die kanonische Parametrisierung nämlich $F=(V, x \mapsto (x, f(x)), V \times W)$. Und hierfür gilt $\text{Graph}(f)=\text{Bild}F$.

Nun können wir den Tangentialraum an die Graphenpunkte von f bilden, indem wir den Tangentialraum an $\text{Bild}F$ konstruieren. Das gibt die Verallgemeinerung der klassischen Tangente für Funktionen $\mathbb{R} \rightarrow \mathbb{R}$ an den Graphen. Die Tangentenerlegung von F erhalten wir über $F(x)=(x, f(x))$ sofort aus $F(x+\Delta x)=F(x)+(\Delta x, 0) + (0, Df(x) \cdot \Delta x) + \dots$

Der erste Beitrag zum totalen Differential - also $(\Delta x, 0)$ - zeigt, dass der Rang überall den maximal überhaupt möglichen Wert $n=\dim V$ hat ($n=\min(n, n+m)$)! Daher ist hier der Schluß von der Rechnung auf die Geometrie zulässig: Mit Hilfe von $Df(x)$ erhalten wir daher den geometrischen Tangentialraum an den Graphen von f . Wie bei den reellen Funktionen erhalten wir mit $\Delta x \mapsto (\Delta x, Df(x) \cdot \Delta x)$ sogar eine Parametrisierung. Es liegt nahe, das wieder im Sinne eines Steigungsfeldes zu interpretieren: **Geht man im Urbildraum um Δx weiter, dann geht man im Graphen näherungsweise um $(\Delta x, Df(x) \cdot \Delta x)$ weiter.**

Die Tangentenapproximation von F , also $\Delta x \mapsto (x, f(x)) + (\Delta x, Df(x) \cdot \Delta x)$ gibt uns überall den geometrischen Tangentialraum an den Graphen der Abbildung! Kritische Punkte treten nicht auf. Für Abbildungen $\mathbb{R} \rightarrow \mathbb{R}^2$ $\mathbb{R}^2 \rightarrow \mathbb{R}$ lassen sich Graph und die zugehörigen Tangentialräume an den Graphen noch problemlos in drei Dimensionen veranschaulichen.

(3.3.19) **Das Energiefeld.**

Sei $U=(V_0^3, \vec{x} \mapsto U(\vec{x}), \mathbb{R})$ ein glattes Skalarfeld, das wir *potentielle Energie* nennen wollen. Wir bilden den neuen Raum $V_0^3 \times V^3$, den wir *Phasenraum* nennen, und dazu das folgende Skalarfeld (auf dem Phasenraum):

$E = (V_0^3 \times V^3, (\vec{x}, \vec{v}) \mapsto \frac{m}{2} \vec{v}^2 + U(\vec{x}), \mathbb{R})$	Energiefeld
---	-------------

Die Bezeichnung ist selbsterklärend: Der Feldwert entsteht aus der üblichen Formel für die Energie (eines Massenpunktes mit Masse m), sofern man verlangt, eine einfache Zuordnung zu erhalten, die **ohne die Grenzwertprozedur des Ableitens** auskommt. Dann muß man die Geschwindigkeit \vec{v} kennen, d.h. hier vorgeben, um den Energiewert berechnen zu können. Mit Hilfe der Zuordnungsformel für E können wir für das Urbildelement (\vec{x}, \vec{v}) einen eindeutig bestimmten Energiewert berechnen. Mit \vec{x} allein ist das nicht möglich. In der Feldzuordnung sind beide Komponenten \vec{x} und \vec{v} frei vorgebar. Der Urbildraum, also der Phasenraum, ist sechsdimensional. Wir betrachten eine Niveaumenge. Diese wird durch die Gleichung $E(\vec{x}, \vec{v}) = E_0$ festgelegt. Eine solche Niveaumenge wird typischerweise eine fünfdimensionale Hyperfläche sein, eine Energiehyperfläche. Variiert man E_0 , entsteht eine Partition des Urbildraumes. Eine Vergrößerung dieser Partition zerlegt den Phasenraum in die Bereiche gleicher, größerer und kleinerer Energie. Differenzieren von E durch direkte Zerlegung gibt $dE = m\vec{v} \cdot \Delta\vec{v} + \text{grad}U(\vec{x}) \cdot \Delta\vec{x}$. Und $dE=0$ gibt die Richtungen näherungsweise konstanter Energie.

Ersetzen wir den ursprünglichen Konfigurationsraum V_0^3 durch einen eindimensionalen Raum V_0^1 , so läßt sich der zugehörige Phasenraum problemlos als Ebene veranschaulichen. Für $U(\vec{x}) = \frac{k}{2}\vec{x}^2$ sind die Niveaumengen Ellipsen.

Wieder allgemein setzen wir $\vec{F}(\vec{x}) = -\text{grad}U(\vec{x})$ und interpretieren \vec{F} als Kraftfeld. Dann erfüllt eine physikalische Bahn $t \mapsto \vec{r}(t)$ eines Massenpunktes in diesem Kraftfeld die Bewegungsgleichung $m\dot{\vec{v}}(t) = \vec{F}(\vec{r}(t)) = -\text{grad}U(\vec{r}(t))$. Diese Gleichung multiplizieren wir skalar mit $\vec{v}(t)$. Achtung. Hier ist $\vec{v}(t)$ nicht frei wählbar, sondern ist durch Ableiten von $t \mapsto \vec{r}(t)$ festgelegt. Um $\vec{v}(t_0)$ durch Ableiten zu bestimmen benötigt man $\vec{r}(t)$ in einer Umgebung von t_0 , d.h. man benötigt unendlich viele Werte! Vgl. die einleitenden Bemerkungen zur Festlegung von E . Nun wenden wir auf beiden Seiten der resultierenden Gleichung die Kettenregel (in umgekehrter Richtung) an. Etwa $\frac{d}{dt}\vec{v}^2(t) = 2\vec{v}(t) \cdot \dot{\vec{v}}(t)$. Dann folgt:

$$\frac{m}{2} \frac{d}{dt} \vec{v}^2(t) = -\frac{d}{dt} U(\vec{r}(t)) \quad \text{Oder auch:} \quad \frac{d}{dt} E(\vec{r}(t), \vec{v}(t)) = 0$$

Das ergibt den üblichen Energiesatz: Entlang einer physikalischen Bahn $t \mapsto \vec{r}(t)$ bleibt die Gesamtenergie erhalten, ist eine zeitlich konstante Größe! Geometrisch besagt unser Ergebnis: Hat man eine physikalische Bahn $t \mapsto \vec{r}(t)$, so bilde man damit die Phasenraumkurve $t \mapsto (\vec{r}(t), \vec{v}(t))$. Dann muß das Bild dieser Phasenraumkurve ganz in einer zugehörigen Energiehyperfläche verlaufen! Und der zugehörige Geschwindigkeitsvektor $t \mapsto (\dot{\vec{r}}(t), \ddot{\vec{r}}(t))$ muß stets tangential zu der Energiehyperfläche liegen! Die Bewegungsgleichung stellt diesen tangentialen Verlauf (über obige Rechnung!) gerade sicher. In Kapitel 7 werden wir sehen, welche Bedeutung diese Phasenraumkurve im Bereich der Differentialgleichungen hat.

Allerdings gibt es auch Kurven, die ganz in der Energiehyperfläche verlaufen und die nicht physikalisch sind! Der Energiesatz allein reicht nicht, die physikalischen Bahnen festzulegen. Am Ende von Kapitel 6.4 werden wir auf obige Rechnung im Zusammenhang mit dem Arbeitsintegral zurückkommen.

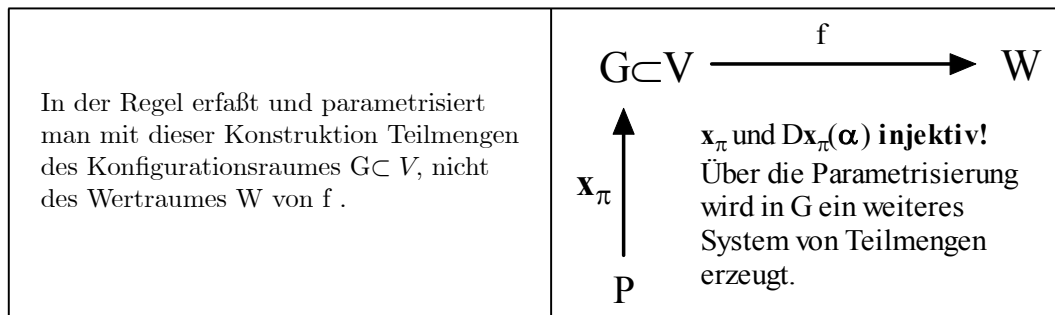
Im Sinne begrifflicher Sorgfalt sollte man zwischen dem *Energiefeld* $(\vec{x}, \vec{v}) \mapsto E(\vec{x}, \vec{v})$ und der *Energiekurve* $t \mapsto E(\vec{r}(t), \vec{v}(t))$ unterscheiden, was vielfach leider nicht üblich ist. Genauer mit allen Zutaten sollte es heißen: **Energiefeld eines Massenpunktes im Potentialfeld U .**

6.3.4 Restriktionsgeometrie und krummlinige Koordinaten

Zwischen Urbildraum und Werteraum besteht in folgender Hinsicht Ungleichheit: Im Urbildraum bilden die Niveaumengen ein ganzes System von Teilmengen. Diese Teilmengen erhält man als Lösungsmengen von Gleichungen, was für den rechnerischen Umgang nicht immer vorteilhaft ist. $\{(x, y) | x^2 + y^2 = 1\}$ etwa stellt die Kreispunkte als Lösungsmenge einer Gleichung dar. Im Werteraum erhalten wir Mengen als Bilder von geeigneten Teilmengen von V , also durch Abbildungen vom Parametrisierungstyp. Und solche Parametrisierungen sind für die analytisch-rechnerische Arbeit meist günstiger. $\{(x, y) | x = \cos\theta, y = \sin\theta, \theta \in \mathbb{R}\}$ ist die übliche Parametrisierung der Kreispunkte. Insbesondere können wir vor f weitere Abbildungen vorschalten, etwa solche vom Kurventyp. $\text{Bild}(f \circ \gamma)$ ist dann ein Kurvenzug in W .

(3.4.1) Durch Vorschalten geeigneter Abbildungen kann man auch Teilmengen des Urbildraumes V parametrisieren. Genauer werden wir sehen, dass eine injektive Parametrisierung ein kanonisches System von Teilmengen von $\text{Bild}\vec{x}_\pi \subset V$ erzeugt und zugleich parametrisiert.

Ein Spezialfall der Parametrisierungskonstruktion ist hierfür von größter Bedeutung: Unter dem Stichwort *krummlinige Koordinaten* wird auf einen Schlag ein ganzes System miteinander verbundener Parametrisierungen erzeugt. Wir beschreiben nachfolgend den zugehörigen Formalismus.



Man geht dazu von einer Abbildung $\vec{x}_\pi: P \rightarrow G$ vom Parametrisierungstyp aus, für die G ja Wertemenge ist. Notfalls muß man sich per Rollenwechsel an diese Stelle unseres Hauptszenenbildes begeben. Beachten Sie die beiden Voraussetzungen für \vec{x}_π .

6.3.4a Koordinatenpartitionen des Konfigurationsraumes

(3.4.2) Die Parametrisierung $\vec{x}_\pi: P \rightarrow V$ sei injektiv, notfalls nach Restriktion des Urbildes. Damit hat jeder Punkt aus Bild \vec{x}_π ein eindeutig bestimmtes Namenstupel $\vec{\alpha} \in P$. Formal bedeutet das, dass die Restriktion von \vec{x}_π auf das Bild invertierbar ist. **Dann ist das Bild einer Partition in P erneut eine Partition in Bild \vec{x}_π .** Diesen Sachverhalt werden wir benötigen. Weiter sei die Parametrisierung ausreichend glatt, auf jeden Fall differenzierbar. Wir haben $k \leq n$. Der Rang von \vec{x}_π soll überall den maximal möglichen Wert k annehmen. D.h. $D\vec{x}_\pi(\vec{\alpha}_0)$ ist stets injektiv. Damit ist die in (3.3.15) geforderte Rangbedingung erfüllt. **Wir dürfen von der Rechnung auf die Geometrie schließen.**

(3.4.3) Wir haben $P \subset \mathbb{R}^k$ und $\vec{\alpha}_0 \in P$. Sei J Indexmenge der Koordinaten, etwa $J = \{1, 2, \dots, k\}$. Weiter sei $K_r \subset J$ eine Teilmenge mit genau r Elementen. Dann setzen wir $\alpha_i = \alpha_{0i} + t_i$ für alle $i \in K_r$ und $\alpha_i = \alpha_{0i}$ für alle i , die nicht in K_r liegen. Variieren wir \vec{t} , so variieren wir genau die Koordinaten, die in K_r liegen. $(\mathbb{R}^r, \vec{t} \mapsto \vec{\alpha}_0 + \vec{t}, \mathbb{R}^k)$ beschreibt daher einen r -dimensionalen Teilraum, der parallel zu den durch K_r festgelegten Koordinatenachsen verläuft. (Zur Veranschaulichung an den Würfel denken: $r=1$ ergibt Kanten, $r=2$ Oberflächenstücke.)

Variiert man jetzt die α_i , deren Index nicht in K_r liegt, dann wird die parametrisierte Teilmenge parallel verschoben. Sagen wir $\alpha_j = \alpha_{0j} + s_j$ für alle $j \notin K_r$. Insgesamt entsteht eine Partition des gesamten Raumes. Schneidet man mit dem Parameterraum P , so erhält man eine entsprechende Partition von P . Und da \vec{x}_π injektiv sein sollte, wird daraus durch die Potenzmengenerweiterung \vec{x}_π eine Partition von Bild \vec{x}_π . Diese Partition wollen wir $Z_\pi(K_r)$ nennen. Für jede Klasse dieser Partition haben wir auch eine Parametrisierung in Form einer besonders einfachen Restriktion der vorgegebenen Gesamtparametrisierung: $(\mathbb{R}^r, \vec{t} \mapsto \vec{x}_\pi(\vec{\alpha} + \vec{s} + \vec{t}), G)$. Und \vec{s} kann als Indizierung der Klassen angesehen werden. Fixiert man \vec{t} , so liefert $\vec{s} \mapsto \vec{x}_\pi(\vec{\alpha} + \vec{s} + \vec{t})$ ein Vertretersystem der Klassen.

Denken Sie daran: Zu jeder Teilmenge K_r von J gehört eine eigene Partition der beschriebenen Art. Für $r=2$ und $k=3$ etwa gibt es drei Partitionen.

Nochmals inhaltlich: Die Klassen $Z_\pi(K_r)$ bestehen aus Niveaumengen folgender Art: Alle Namensparameter α_i , deren Index i nicht in K_r liegt, sind für die Elemente der Klasse konstant, die Parameter mit Index aus K_r dürfen sich ändern. Bei den Klassen handelt es sich um Teilmengen des Konfigurationsraumes V . Es sind i. a. nichtlineare Bilder der affinen Teilräume von P und daher **Mengen nichtlinearer Form**. Sie repräsentieren jeweils bestimmte geometrische Formen.

(3.4.4) $r=0$ und $r=k$ liefern triviale Fälle. $r=1$ ist interessant und wichtig. Es gibt k Möglichkeiten, den Index i für das eine veränderliche α zu wählen, also ebensoviele Teilmenge $K_i \subset J$. D.h. es gibt auch k derartige Partitionen. Für jede dieser Partitionen haben die Klassen Kurventyp: **Kurven, für die alle Parameterwerte fest sind, bis auf den einen ausgezeichneten.** Durch jeden Punkt von Bild \vec{x}_π , gehen k solcher Kurven, von denen jede einen der Parameterwerte bestimmt. Im Falle ebener Polarkoordinaten entsteht das bekannte Netz aus Ursprungskreisen $r = \text{const.}$ und radialen Halbgeraden $\theta = \text{const.}$

(3.4.5) Als Illustrationsbeispiel für den Raum nehmen wir räumliche Polarkoordinaten mit $\vec{\alpha}=(R,\theta,\varphi)$. Um Injektivität zu erlangen, sind auf jeden Fall die Punkte der Polachse herauszunehmen. Für $r=1$ erhalten wir drei Möglichkeiten: $Z_P(\{R\})$, $Z_P(\{\varphi\})$ und $Z_P(\{\theta\})$. Wie sehen die zugehörigen Partitionen aus? Genauer, wie sehen die jeweiligen Klassen aus? Bei $Z_P(\{R\})$ darf nur R variiert werden, d.h. es entstehen Halbgeraden, die vom Ursprung ausgehen. Dieser Punkt ist herauszunehmen. Wir werden Ausnahmepunkte nachfolgend nicht mehr gesondert aufführen, sie lassen sich bei Bedarf leicht angeben. Die Klassen für $Z_P(\{\theta\})$ sind Halbkreise um den Ursprung. Die Klassen von $Z_P(\{\varphi\})$ sind Kreise mit Mittelpunkt auf der Polachse. Alles sind jeweils Bilder achsenparalleler Strecken in $PC \mathbb{R}^3$.

Für $r=2$ erhält man Mengen vom Flächentyp (zwei freie Parameter). Wieder gibt es drei Möglichkeiten: $Z_P(\{R,\theta\})$, $Z_P(\{R,\varphi\})$ und $Z_P(\{\theta,\varphi\})$. Im dritten Fall $Z_P(\{\theta,\varphi\})$ erhält man Kugeloberflächen als Klassen, im zweiten Kegelmantel und im ersten Halbebenen deren Rand die Polachse ist. **Auch hier hat man automatisch wieder einfache Parametrisierungen dieser Figuren.**

- Beschreiben Sie die geometrische Form der 6 relevanten Klassentypen für den Fall der Zylinderkoordinaten.

(3.4.6) Jetzt können wir abstrahieren, was Parametrisierungen des besprochenen Typs leisten:

Hat man irgendeine einfache geometrische Figur mit Kugelsymmetrie bezüglich des Ursprungs, gleichgültig mit welcher Zahl von Freiheitsgraden, dann bestehen gute Aussichten, sie als eine der Klassen zu identifizieren. Und ist das der Fall, hat man automatisch eine einfache Parametrisierung in Form einer Restriktion von \vec{x}_π mit deren Hilfe man die geometrische Menge in analytische Rechnungen einbeziehen kann! Da der Rang von $D\vec{x}_\pi(\vec{\alpha})$ überall den maximalen Wert k haben soll, dürfen wir vom analytischen Tangentialraum auf den geometrischen schließen. Und weiter hat man eine Indizierung der Mengen, die automatisch durch den äußeren Parameter \vec{s} in dieser Beschreibung gegeben ist. Kurz: **Man erhält einfache Parametrisierungen und Indizierungen für eine Art Grundausstattung von Figuren eines bestimmten durch die Koordinaten festgelegten Formtyps!** Wie üblich wird man dann komplexere Figuren dieses Typs in einfacher Weise aus diesen Grundtypen aufbauen.

Hat man ein Skalarfeld mit Niveaumengen, die zu dieser Figurengrundausstattung gehören, so wird man auch das Feld mit Hilfe dieser Parametrisierung darstellen und dabei Unabhängigkeit von gewissen Koordinaten erhalten. Ein Klassiker dieser Argumentation ist in der folgenden Gleichung zusammengefaßt:

$$U(\vec{x}) = U^P(r, \theta, \varphi) = u(r).$$

Das Skalarfeld U hängt in der Polardarstellung nur von der Koordinate r ab und die zugehörige dies beschreibende Funktion wird mit u bezeichnet. Die Niveaumengen des Feldes sind dann Kugelschalen, deren jeweiliger Feldwert durch u bestimmt wird.

6.3.4b Koordinatenbasisfelder

(3.4.7) Mit Hilfe der gewonnenen Parametrisierung ist es leicht, für alle Figuren von $Z_\pi(K_r)$ die zugehörigen geometrischen Tangentialräume zu konstruieren. Sei $\vec{x}_0 = \vec{x}_\pi(\vec{\alpha}_0)$ und $M \in Z_\pi(K_r)$ eine Klasse, die durch \vec{x}_0 verläuft. Wir haben die totale Ableitung $D\vec{x}_\pi(\vec{\alpha}_0)$. Sie ist als injektiv vorausgesetzt. Dann ist auch jede Restriktion injektiv mit konstantem Rang. Insbesondere gilt dies für Restriktionen, die durch die Koordinatenauswahl K_r bestimmt sind. Wir können somit mit Hilfe der totalen Ableitung auf den geometrischen Tangentialraum von M in \vec{x}_0 schließen.

(3.4.8) Ist $r=1$ mit freier i -ter Koordinate, so ergibt sich die Restriktion von $D\vec{x}_\pi(\vec{\alpha}_0)$ durch die partielle Ableitung (in die gegebene Richtung). Für diesen Fall führen wir eine besondere Bezeichnung ein:

$$\vec{\partial}_i = \vec{\partial}_i(\vec{\alpha}_0) = \frac{\partial \vec{x}_\pi}{\partial \alpha_i}(\vec{\alpha}_0) = D\vec{x}_\pi(\vec{\alpha}_0) \cdot \vec{E}_i.$$

Um Verwechslungen zu vermeiden, bezeichnen wir die Vektoren der kanonischen Basis von P bzw. \mathbb{R}^k mit \vec{E}_i . Der große Buchstabe soll an den Parameterbereich erinnern. Zutatenkorrekter wäre die Schreibweise $\vec{\partial}_i^\pi(\vec{\alpha}_0)$. Aber solange wir nur eine feste Parametrisierung betrachten, werden wir auf den Index π verzichten.

Der letzte Ausdruck der definierenden Formel stellt den Bezug zur totalen Ableitung her: Restriktion der totalen Ableitung auf den durch \vec{E}_i erzeugten eindimensionalen Teilraum von \mathbb{R}^k . Der vorletzte Ausdruck

zeigt, wie man diese Größe als Ableitung der restringierten Parametrisierung $t \mapsto \vec{x}_\pi(\vec{\alpha}_0 + t\vec{E}_i)$ erhält, die ja vom Kurventyp ist. Die ersten beiden Ausdrücke stellen die neu eingeführten **Bezeichnungen** dar, die daran erinnern sollen, dass diese Tangentenvektoren durch partielles Ableiten entstehen. Das Ergebnis hängt vom Parametertupel $\vec{\alpha}_0$ ab, ist also ein Vektorfeld auf dem Parameterraum. Im Sinne von (1.6.6) können wir sagen: **Ein Vektorfeld über der Parametrisierung \vec{x}_π .**

Die gegebene Definition der $\vec{\partial}_i$ als partielle Ableitung erlaubt eine nützliche inhaltliche Interpretation dieser Vektoren: $\vec{\partial}_i$ gibt die Änderung (=momentane Geschwindigkeit) des Ortsvektors \vec{x} wieder, die erfolgt, wenn man im Parameterraum mit Geschwindigkeit 1 parallel zur i-ten Koordinatenachse läuft!

(3.4.9) Und mit Hilfe der Vektorfelder $\vec{\partial}_i$ kann man nun Basisfelder für die Tangentialräume aller Figuren der Partitionen $Z_\pi(K_r)$ bilden. Man muß nur die Felder mit den durch K_r festgelegten Indizes zusammensetzen. Genauer: Für jedes $\vec{\alpha}$ erhält man eine Basis des Tangentialraumes von $\vec{x}_\pi(\vec{\alpha})$. In Verallgemeinerung von (1.6.6) nennen wir die Abbildung $\vec{\alpha} \mapsto \partial(\vec{\alpha})$ *das kanonische Basisfeld über der Parametrisierung \vec{x}_π .* Beachten Sie: Für $\vec{\alpha} \neq \vec{\beta}$ ergeben sich i.a. auch verschiedene Tangentialräume. D.h die beiden Basen $\partial(\vec{\alpha})$ und $\partial(\vec{\beta})$ können verschiedene Räume aufspannen.

(3.4.10) Für $r=k$ erhält man ein Basisfeld für den Tangentialraum zu $\text{Bild}\vec{x}_\pi$. Ist sogar $r=k=n$, so erhält man ein Basisfeld für den gesamten Konfigurationsraum $G \subset V$! Sonst spannen diese Basisfelder natürlich immer nur k -dimensionale Teilräume von V auf.

(3.4.11) Beispiel: Räumliche Polarkoordinaten. Durch partielles Ableiten der Darstellungsformel $\vec{x}_P(r, \theta, \varphi) = \dots$ aus (1.4.27) erhält man sofort die drei Vektorfelder. **Beachten Sie, wie leicht das geht:**

$\begin{aligned} \vec{x}_P(r, \theta, \varphi) &= r(\vec{e}_1 \cos \varphi + \vec{e}_2 \sin \varphi) \sin \theta + \vec{e}_3 r \cos \theta \\ \vec{\partial}_r(r, \theta, \varphi) &= (\vec{e}_1 \cos \varphi + \vec{e}_2 \sin \varphi) \sin \theta + \vec{e}_3 \cos \theta \\ \vec{\partial}_\theta(r, \theta, \varphi) &= r(\vec{e}_1 \cos \varphi + \vec{e}_2 \sin \varphi) \cos \theta - \vec{e}_3 \sin \theta \\ \vec{\partial}_\varphi(r, \theta, \varphi) &= r(-\vec{e}_1 \sin \varphi + \vec{e}_2 \cos \varphi) \sin \theta \end{aligned}$
--

$\vec{\partial}_\theta, \vec{\partial}_\varphi$ bildet eine Basis des Tangentialraumes an die Kugelflächen. $\vec{\partial}_r, \vec{\partial}_\varphi$ eine Basis an die Tangentialflächen der Kegelmäntel usw. Wie versprochen ergibt sich das alles problemlos. Die Vektoren $\vec{\partial}$ sind keineswegs immer Einheitsvektoren und man sollte sie auch nicht ohne besonderen Grund normieren (wie das immer wieder geschieht). Dagegen stehen die Vektoren $\vec{\partial}_i$ (im Beispiel) am selben Ort paarweise aufeinander senkrecht.

- Rechnen Sie die entsprechenden Basisfelder über den Zylinderkoordinaten aus.

(3.4.12) Kehren wir zum allgemeinen Fall zurück:

Für $k=n$ und $r=n-l$ erhält man Klassen vom **Hyperflächentyp**. Erneut gibt es n zugehörige Partitionen. Durch jeden Punkt gehen genau n Klassen, wobei jede Klasse Niveaumenge zu einer der Koordinaten ist. Jede solche Hyperfläche hat einen Tangentialraum, der durch die $(n-l)$ Vektoren $\vec{\partial}_i$ mit $i \in K_{n-l}$ aufgespannt werden. Nur der verbleibende letzte Vektor zeigt nicht in tangentialer Richtung! Vielfach - aber nicht immer - steht er senkrecht auf dem Tangentialraum der Hyperfläche zum gewählten Punkt.

- Führen Sie Polarkoordinaten in vier Dimensionen ein. Gehen Sie dafür das Hyperflächenproblem durch.

6.3.4c Darstellungen mit Hilfe der Koordinatenbasisfelder

(3.4.13) Die Parametrisierung der Figuren eines bestimmten geometrischen Typs samt ihrer geometrischen Tangentialräume ist noch nicht alles, was die eingeführten Parametrisierungen leisten: Die Koordinatenbasisfelder $\vec{\partial}_i$ repräsentieren auch automatisch die Symmetrien und Formen, die durch die Koordinaten bestimmt sind, leisten gerade das, was in 6.1.6 als Nutzen von Basisfeldern herausgearbeitet wurde.

(3.4.14) Viele physikalische Feldabbildungen sind vom Typ $f : V_0^3 \rightarrow V^3$. In diesem Fall verfügen wir über eine kanonische Identifizierungsabbildung $(V^3, \vec{x} \mapsto \vec{x}, V_0^3)$ zwischen freien und gebundenen Vektoren: Man verschiebt den freien Vektor parallel, bis der Pfeil im Ursprung beginnt. Dann kann man auch die Werte $f(\vec{x})$ nach den Basisfeldern $\vec{\partial}_i$ entwickeln:

$$\vec{\alpha} \mapsto f(\vec{x}_\pi(\vec{\alpha})) = \sum \vec{\partial}_i(\vec{\alpha}) f_i(\vec{\alpha}).$$

Im Sinne von (1.6.6) liegt ein Basisfeld über $f \circ \vec{x}_\pi$ vor.

Beispiel: Ein Feld in einem Zylinder! Durch partielles Ableiten der früher gegebenen Zylinderparametrisierung finden wir sofort mit $\vec{\alpha} = (r, \theta, h)$ und mit Fortlassen der Argumente:

$$\vec{\partial}_r = \vec{e}_r(\theta) = \vec{e}_1 \cos \theta + \vec{e}_2 \sin \theta \quad \vec{\partial}_\theta = r(-\vec{e}_1 \sin \theta + \vec{e}_2 \cos \theta) \quad \vec{\partial}_h = \vec{e}_3.$$

Das sind Richtungsfelder, die der Zylindersymmetrie entsprechen: Radial, tangential und axial. Beachten Sie erneut, dass die Vektoren im euklidischen Raum zwar überall aufeinander senkrecht stehen, aber keineswegs immer die Länge 1 haben. Das tangentiale Feld etwa wächst mit r an! Und diese Vektoren sind auch keine reine Funktion des Ortes, sondern hängen vom Parametertupel ab! Auf der Achse r=0 sind zwei der Felder ungleich Null. Und $\vec{\partial}_r$ hat dort für jedes θ einen anderen Wert, ist auf der Achse also kein Feld auf G.

Jetzt nehmen wir an, \vec{v} sei ein Geschwindigkeitsfeld. Wir können schreiben - unter erneutem Auslassen der meisten Argumente:

$$\vec{v}(\vec{x}) = \vec{v}(\vec{x}_Z(r, \theta, h)) = \vec{\partial}_r(r, \theta, h)v_r(r, \theta, h) + \vec{\partial}_\theta v_\theta + \vec{\partial}_h v_h.$$

Dann ist v_r überall die Geschwindigkeitskomponente in radialer Richtung zur Polachse usw. Alle drei Komponenten haben unmittelbaren Bezug zur Zylindersymmetrie wie früher bereits besprochen!

(3.4.15) **Teilweise entstehen derartige Basisdarstellungen auch unmittelbar durch Differenzieren unter Verwendung der Kettenregel.**

Wir betrachten eine Bahnbewegung auf dem Zylindermantel. Eine solche wird durch eine vor die Parametrisierung geschaltete Abbildung gegeben $t \mapsto \vec{\alpha}(t) = (R, \theta(t), h(t))$ mit R=Zylinderradius. Die Zuordnung der Bahnkurve lautet:

$$t \mapsto \vec{r}(t) = \vec{x}_Z(\vec{\alpha}(t)) = \vec{x}_Z(R, \theta(t), h(t))$$

Also: Der Ortsvektor entsteht, indem man zuerst die Zylinderparameter zur Zeit t festlegt und dann damit in die Darstellungsformel der Zylinderkoordinaten eingeht. Vgl. 6.1.4.

Die momentane Geschwindigkeit der Bewegung erhalten wir, wenn wir unter Verwendung der Kettenregel differenzieren. Die äußeren partiellen Ableitungen sind aber gerade unsere Basisfelder. Wegen $\dot{R} = 0$ verschwindet ein Beitrag. Es bleibt ($\vec{\partial}_\theta = \frac{\partial \vec{x}_Z}{\partial \theta}$ usw!!!):

$$\vec{v}(t) = \dot{\vec{r}}(t) = \vec{\partial}_\theta(R, \theta(t), h(t))\dot{\theta}(t) + \vec{\partial}_h(R, \theta(t), h(t))\dot{h}(t)$$

Ein bemerkenswertes Resultat! Die beiden Ableitungen der Parameterkurve sind die Komponenten der wahren physikalischen Geschwindigkeit bezüglich des Koordinatenbasisfeldes. In Koordinatenvektorform: $\vec{v}^Z(t) = \dot{\theta}(t)\vec{\partial}_\theta + \dot{h}(t)\vec{\partial}_h$.

(3.4.16) Ein weiteres **Beispiel**. Wir betrachten die Parametrisierung einer Parabelfläche:

$$\vec{p} = \left(\mathbb{R}^2, (\rho, \theta) \mapsto \begin{pmatrix} \rho \cos \theta \\ \rho \sin \theta \\ \rho^2 \end{pmatrix}, \mathbb{R}_K^3 \right)$$

Dann ist Bild(\vec{p}) eine Rotationsparabel bezüglich der z-Achse. Die totale Ableitung - in Form der Jakobimatrix - hat in den **Spalten** gerade die beiden Basisfelder:

$$J(\rho, \theta) = (\vec{\partial}_\rho(\rho, \theta), \vec{\partial}_\theta(\rho, \theta)) = \begin{pmatrix} \cos \theta & -\rho \sin \theta \\ \sin \theta & \rho \cos \theta \\ 2\rho & 0 \end{pmatrix}$$

Grund: J_{ik} entsteht als Ableitung der i-ten Ortsvektorkomponente nach dem k-ten Parameter! i ist aber der Zeilenindex! Als Gedächtnisstütze: $\vec{\partial}_k = \frac{\partial \vec{x}_\pi}{\partial \alpha_k}$.

(3.4.17) Das gilt offenbar allgemein:

In einer Koordinatendarstellung ergeben die **Spalten der Jakobimatrix** gerade die geometrischen Tangentenvektoren $\vec{\partial}_i$.

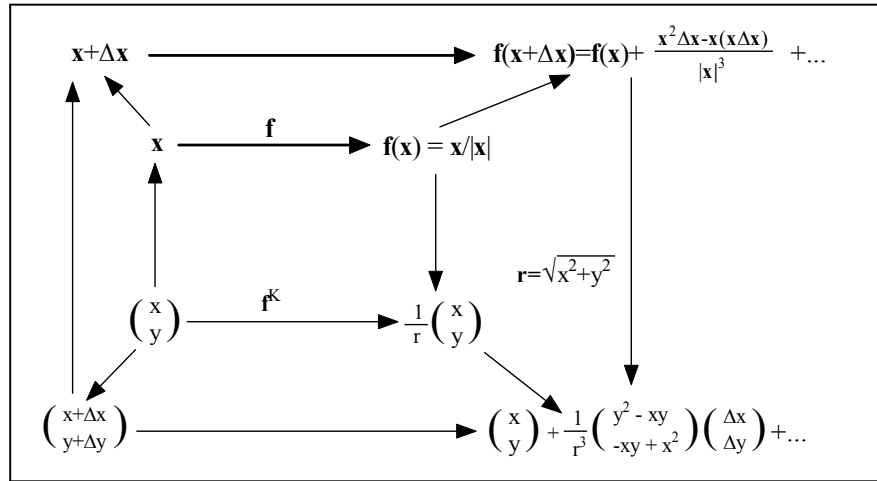
- Verifizieren Sie die Gültigkeit dieser Aussage mit Hilfe von (3.4.8). Konkretisieren Sie die Aussage dann noch für räumliche Polarkoordinaten.
- Sei $f: G \rightarrow W$ glatte Abbildung und \vec{x}_π eine Parametrisierung von G. Wir bilden $F = f \circ \vec{x}_\pi$. Welche geometrische Interpretation haben die Spalten der Funktionalmatrix von F? Welche Bedingungen müssen erfüllt sein, damit man ein Basisfeld über F erhält? Wann erhält man Basen von W?

6.3.4d Verallgemeinerungen und Modifikationen

Wir möchten die Gelegenheit nutzen, eine wichtige geistige Leistung samt ihrer Förderung durch Mathematik zu erläutern. Die Ausführungen über die krummlinigen Koordinaten und die damit verbundenen Eigenschaften erfassen bereits relativ komplexe Sachverhalte. Zur Herleitung der Resultate haben wir eine Reihe von Voraussetzungen benötigt und benutzt. Jetzt stößt man nicht selten auf Situationen, in denen diese Voraussetzungen nicht alle erfüllt sind, für die man jedoch dieselben oder ähnliche Resultate haben möchte. Verständnis der Herleitung bedeutet dann auch, analysieren zu können welche Voraussetzungen jeweils benötigt wurden und ob eine Verallgemeinerung möglich ist. Wir propagieren hier nicht die vielfach in der Mathematik geschätzte Methode, die Aussage möglichst allgemein zu beweisen, so dass die neue Situation als Spezialfall erscheint, sondern halten es für ebenso wichtig, zu lernen, auf die neue Situation verständlich einzugehen.

Hier an dieser Stelle stellen wir als größerer Übung die Frage: Was bleibt von den vorangegangenen Überlegungen gültig, wenn man die Parametrisierung \vec{x}_π ersetzt durch die Abbildung $f \circ \vec{x}_\pi$? Wenn man also hinter die Parametrisierung noch die glatte Feldabbildung $f: G \rightarrow W$ schaltet, also eine Abbildung des Typs, das Hauptthema dieses Kapitels ist. Was bleibt, was ändert sich worauf muß man achten? Wie sehen zugehörige illustrierende Beispiele aus? →

(3.4.18) Stellen wir dazu für ein Beispiel - ein ebenes Vektorfeld - nochmals alle beteiligten Größen im üblichen Schema zusammen. Als Koordinaten wählen wir kartesische Koordinaten.



Ganz oben und ganz unten die beiden Tangentzerlegungen. Oben absolut und unten in der Koordinatenform. In diesem Fall finden wir $\vec{\partial}_x = \frac{1}{r^3}y \begin{pmatrix} y \\ -x \end{pmatrix} = -\frac{y}{r^2}\vec{e}_t$ und $\vec{\partial}_y = \frac{1}{r^3}x \begin{pmatrix} -y \\ x \end{pmatrix} = \frac{x}{r^2}\vec{e}_t$. Beide Vektoren haben dieselbe Richtung, sind abhängig. Dabei ist $r\vec{e}_t = \vec{e}_1(-y) + \vec{e}_2x$ wie üblich. Das gibt für das totale Differential:

$$df = \vec{\partial}_x \Delta x + \vec{\partial}_y \Delta y = \frac{1}{r^2}(-y\Delta x + x\Delta y)\vec{e}_t(x, y).$$

Die in der oberen Zeile des Schemas gegebene absolute Darstellung von $Df(x)$ erhält man durch Einsetzen der Tangentzerlegung von $\frac{1}{\|\vec{x} + \Delta \vec{x}\|}$ und dem üblichen Sortieren nach Potenzen von $\Delta \vec{x}$.

Im Ursprung ist die Funktion f nicht definiert. Außerhalb sind die Niveaulinien Halbgeraden. Das Bild ist der Einheitskreis. Der Rang ist überall konstant gleich 1, wogegen $m=n=2$ gilt. Der Schluß auf die Geometrie ist zulässig. $\text{Kern}Df$ und Bild sind beide eindimensional und ergeben die geometrischen Tangentialräume wie man unmittelbar sieht.

- Was ist Bildf? Was ergibt das Rangkriterium (6.3.14)? Wieso ergibt $(\vec{\partial}_x, \vec{\partial}_y)$ keine Basis?
- Wie erkennt man hier unmittelbar, dass man in diesem Fall nicht von der Rechnung auf die Geometrie schließen kann? Der Grund läßt sich übrigens auch unmittelbar aus der Abbildung selbst herauslesen.

6.3.5 Über den Nutzen des Formalismus / Zusammenfassung

Wir nennen drei sehr allgemeine und perspektivreiche Problembereiche, für deren Behandlung der einheitliche Ableitungsformalismus wichtig ist:

1. Die totale Ableitung als Hilfsmittel beim Rechnen : Man erhält eine lokale Näherung der Zuordnung der Abbildung. Diese Näherung ist der rechnerischen Analyse gut zugänglich. Der Zugang erfolgt meist über das totale Differential. Die Tangentenerlegung liefert eine Umformung des Rechenausdrucks, die diese Näherung isoliert.
Insbesondere erhält man so Zugang zu Systemen, bei denen man über Kenntnisse von Änderungen von Größen verfügt bzw. sich für diese Änderungen interessiert. (Stichwort: Differentialgleichungen. Bisheriges Beispiel: Der Dipol.) .
2. Die totale Ableitung als Werkzeug der analytischen Geometrie: Man erhält Zugang zu den Tangentialräumen der geometrischen Mengen, die mit der Abbildung verbunden sind. Dies sind die Niveaumengen im Urbildraum und das Bild im Werteraum. Per Umkehrung algebraisiert man geometrische Probleme: Man beschreibt geometrisch interessierende Mengen als Niveaumengen oder Bilder geeigneter Abbildungen löst die geometrischen Probleme dann rechnerisch - analytisch.
3. Die totale Ableitung als Mittel der Verbindung geometrischer und prozessualer Strukturen: Man kann Prozesse analysieren, die mit glatten Teilmengen des Konfigurationsraumes verbunden sind. Dabei koppelt der Formalismus über die Wahl einer geeigneten Parametrisierung des Konfigurationsraumes die Geometrie der Menge und die Parameter des Prozesses in angemessener Weise. Beispiele sind Bewegungen, die durch Zwangskräfte an glatte Flächen gebunden sind, Felder mit bestimmten Symmetrieeigenschaften

Kapitel 6.4 Kurven

6.4.0 Vorbemerkung

(4.0.1) Wir behandeln die Abbildungen $\mathbb{R} \rightarrow W$ des Kurventyps gesondert und etwas ausführlicher, da wir im nachfolgenden Kapitel 7 eine Reihe zugehöriger Resultate benötigen. Das Besondere an den Abbildungen dieses Typs ist, dass problemlos höhere Ableitungen eingeführt werden können, die alle erneut von demselben Abbildungstyp sind. Man erhält so eine weitgehende Analogie zum vertrauten Fall $\mathbb{R} \rightarrow \mathbb{R}$ der reellen Funktionen. Auf das nicht so einfache Problem der Bildung höherer Ableitungen in den anderen Fällen gehen wir kurz in Kap.10 ein.

□ Was fällt Ihnen vergleichsweise zum Stichwort "2. Ableitung eines Skalarfeldes" ein?

Es geht also nachfolgend um Abbildungen $f=(I, t \mapsto f(t), W)$, wobei I ein (nicht einpunktiges) Intervall sein soll. W habe endliche Dimension und sei normiert. f selbst sei ausreichend glatt, typischerweise beliebig oft differenzierbar.

(4.0.2) Zur Sprechweise dieses Kapitelteiles: Eine Abbildung vom Typ f nennen wir eine *Kurve in W* . Dagegen nennen wir die Menge $\text{Bild } f$ einen *Weg in W* . Zu ein und demselben Weg haben wir viele mögliche Kurven. Eine Kurve r ist also eine Parametrisierung des Weges $\text{Bild } r \subset W$.

6.4.1 Höhere Ableitungen von Kurven

(4.1.1) Wie entstehen die höheren Ableitungen?

Wir setzen $f'(t) = Df(t)$. Da dies Elemente aus W sind, erhalten wir die Ableitung (Ableitungsabbildung)

$$f' = (I, t \mapsto f'(t), W),$$

die erneut eine Kurve in W bildet und die wir auch *Geschwindigkeitskurve (zu f)* nennen werden. Rekursiv definieren wir nun für $n=1,2,\dots$ die höheren Ableitungen über $f^{(n+1)}(t) = (f^{(n)})'(t)$. Schließlich setzen wir noch $f^{(0)} = f$. Die übliche Beschleunigungskurve entspricht f'' . Bei eventuellen Randpunkten ist stets die zugehörige einseitige Ableitung gemeint.

Bitte beachten Sie wie immer: Die Konstruktion gilt für **alle unsere Vektorräume**, nicht nur für die vom Typ der üblichen geometrischen Pfeile.

(4.1.3) Die nachfolgenden drei Beispiele sollen die mögliche Vielfalt etwas verdeutlichen.

Beispiel a): Flugparabel $r = (\mathbb{R}, t \mapsto \frac{1}{2}\vec{g}t^2 + \vec{v}_0t + \vec{r}_0, V_0^3)$. Hier ist die erste Ableitung die Geschwindigkeit $\vec{v} = (\mathbb{R}, t \mapsto \vec{g}t + \vec{v}_0, V^3)$. Die Beschleunigung ist die konstante Kurve $t \mapsto \vec{g}$ usw.

Beispiel b): Die Kurve einer Drehbewegung in der Ebene mit ihrer Ableitung:

$$\begin{aligned} R_\omega &= \left(\mathbb{R}, t \mapsto R_\omega(t) = \begin{pmatrix} \cos \omega t & -\sin \omega t \\ \sin \omega t & \cos \omega t \end{pmatrix}, \text{End}_{\mathbb{R}}(\mathbb{R}_K^2) \right). \\ R'_\omega &= \left(\mathbb{R}, t \mapsto R'_\omega(t) = \omega \begin{pmatrix} -\sin \omega t & -\cos \omega t \\ \cos \omega t & -\sin \omega t \end{pmatrix}, \text{End}_{\mathbb{R}}(\mathbb{R}_K^2) \right) \end{aligned}$$

und $R''_\omega = -\omega^2 R_\omega$.

Beispiel c): Eine Kurve im Polynomraum \mathfrak{P} . Die nachfolgend beschriebene Kurve liegt ganz in einem dreidimensionalen Teilraum, in dem auch die Geschwindigkeitskurve liegt.

$$\begin{aligned} f &= (\mathbb{R}, t \mapsto p_t, \mathfrak{P}) \quad \text{mit } p_t = (\mathbb{R}, x \mapsto p_t(x) = (1+t^2)x^2 + x \sin t + t^3, \mathbb{R}) \\ f' &= (\mathbb{R}, t \mapsto p'_t, \mathfrak{P}) \quad \text{und } p'_t = (\mathbb{R}, x \mapsto p'_t(x) = 2(1+t^2)x + \sin t, \mathbb{R}) \end{aligned}$$

Weitere Beispiele für Kurven in Kap.6.4.7 (*Corioliskraft*) und in der Einführung von Kap. 7.

6.4.1a Entfaltung des Glattheitsbegriffes im Kurvenfall

(4.1.4) Es liegt nahe, an dieser Stelle den Glattheitsbegriff von Abbildungen stärker zu entfalten. Zur ersten Orientierung beschränken wir uns dabei auf Kurven. Als zentrales Klassifikationsmerkmal hatten wir die Stetigkeit eingeführt. Genauer: Die Stetigkeit in jedem Punkt des Urbildes.

(4.1.5) Sei $\mathfrak{F}(J,W)$ die Menge aller Kurven $J \rightarrow W$. Dabei seien $J \subset \mathbb{R}$ ein Intervall und W normierter Vektorraum. Zunächst sondern wir daraus die Teilmenge aller (in jedem Punkt von J) stetigen Kurven aus. Diese Teilmenge nennen wir $C^0(J,W)$. Der Buchstabe C erinnert an *continuous*, das englische Wort für stetig. Diese Menge enthält immer noch zahlreiche Kurven, die der Anschauung sehr fern sind und exotische Eigenschaften besitzen.

(4.1.6) Weiter bildet man die Teilmenge $C^1(J,W)$ aller stetig differenzierbaren Kurven. (D.h. differenzierbar mit stetiger Ableitung. Nicht etwa stetig und differenzierbar!). Diese Teilmenge ist eine echte, viel kleinere Teilmenge von $C^0(J,W)$. D.h. es gibt viele Kurven, die zwar stetig, aber keineswegs differenzierbar sind, oder die differenzierbar sind mit unstetiger Ableitung.

Als nächstes kommt $C^2(J,W)$, der Raum aller zweifach stetig differenzierbaren Kurven. Diese sind erneut deutlich glatter! Usw. Schließlich gelangt man zur Teilmenge $C^\infty(J,W)$ aller beliebig oft differenzierbaren Kurven. (Hier ist stetig unnötig).

(4.1.7) Damit hat man bereits eine ganze Hierarchie von Glattheitseigenschaften. Überlegen Sie sich selbst zugehörige Beispiele und gewöhnen Sie sich an, in Anwendungssituationen auf die jeweils relevante oder benötigte Glattheitseigenschaft zu achten.

(4.1.8) Manche Anwendung benötigen eine Verallgemeinerung, die dann durch das Attribut *stückweise* charakterisiert wird: eine Kurve $J \rightarrow W$ heißt stückweise stetig, wenn sie bis auf höchstens **endlich viele Ausnahmepunkte** stetig ist. Sie heißt stückweise C^1 , wenn sie bis auf endlich viele Ausnahmepunkte C^1 ist. Usw. $x \mapsto |x|$ ist ein Beispiel. Für $x=0$ ist sie nur stetig, aber nicht differenzierbar.

Bei der Idealisierung von Ein- und Ausschaltvorgängen treten solche Abbildungen auf.

- Beliebige sind auch Konstruktionen wie "stetig und stückweise C^2 ". Was bedeutet das? Ein Beispiel?
- Welche Glattheitseigenschaften haben die folgenden drei Funktionen? $x \mapsto |x|^3$, $x \mapsto |x|\sin(x)$ und $x \mapsto \sqrt{|x|^3}$
- Welche Glattheitseigenschaft hat $x \mapsto \frac{x}{\sqrt{x^2}}$?

6.4.1b Zur Bedeutung der Beweise

(4.1.9) Historischen Erfahrungen im Bereich der Analysis haben gelehrt, dass es einerseits anschaulich vertraute Sachverhalte gibt, die verallgemeinerbar und streng gültig sind, andererseits aber auch solche, bei denen die anschauliche Erwartung extrem täuscht, für die man Gegenbeispiele findet. Und das heißt, dass generell sorgfältige und strenge Beweise erforderlich sind. Weiter unten werden wir ein Beispiel eines allgemein gültigen und sehr nützlichen Sachverhaltes geben und beweisen (den "Satz vom beschränkten Zuwachs"). Andere anschaulich evidente Sachverhalte lassen sich ebenso beweisen, aber der Beweis erwies sich als sehr schwierig. Ein unerwartetes Beispiel, bei dem die Erwartung täuscht: Man kann eine Kurve der folgenden Art konstruieren: Sei $J=[0, 1]$ und $R=J \times J \subset \mathbb{R}^2$. Wir wissen bereits aus dem ersten Kapitel, dass J und R gleichmächtig sind. Aber es ist auch möglich, eine stetige **surjektive** Abbildung $J \rightarrow R$ zu geben (Peanokurve). D.h. man kann das Rechteck vollständig durch eine (stetige, nicht differenzierbare) Kurve parametrisieren. Eine solche Überdeckung des Rechteckes durch eine Linie -also einen durchlaufenden Weg - widerspricht den anschaulichen Vorstellungen außerordentlich.

6.4.1c Polynomapproximationen

(4.1.10) Die lokale Approximation einer glatten Kurve durch Polynome sollte mit zunehmender Ordnung besser werden. Die höheren Ableitungen wiederum sollten mit den zugehörigen Polynomkoeffizienten zusammenhängen. Wir werden unten sehen, dass ein solcher Zusammenhang tatsächlich besteht.

(4.1.11) Jetzt zunächst zu der Frage: Wie verallgemeinert man die Tangentenapproximation (im Kurvenfall) zu einer Polynomapproximation? Die Antwort liegt weitgehend auf der Hand:

Eine **lokale Polynomzerlegung der Ordnung N** der Kurve f um t_0 ist eine Darstellung des Kurvenwertes mit Restterm der folgenden Art:

$$\vec{f}(t_0 + \Delta t) = \sum_{k=0}^N \vec{a}_k \Delta t^k + \Delta t^N \vec{R}_f^N(t_0, \Delta t) \quad \text{mit } \Delta t = t - t_0.$$

Die zu erfüllende Resttermbedingung lautet:

$$\lim_{\Delta t \rightarrow 0} \vec{R}_f^N(t_0, \Delta t) = 0$$

Im Fehlerterm sind jetzt N Faktoren Δt abzuspalten, nicht nur einer, wie im Falle der Tangentenzerlegung. Das macht den gesamten Fehler in der Nähe von t_0 typischerweise kleiner, verbessert die Approximation. Das approximierende Polynom ist ein Polynom vom Grade N mit noch unbekanntem Koeffizienten $\vec{a}_k \in W$. (Nur $\vec{a}_0 = \vec{f}(t_0)$ liegt bereits fest!)

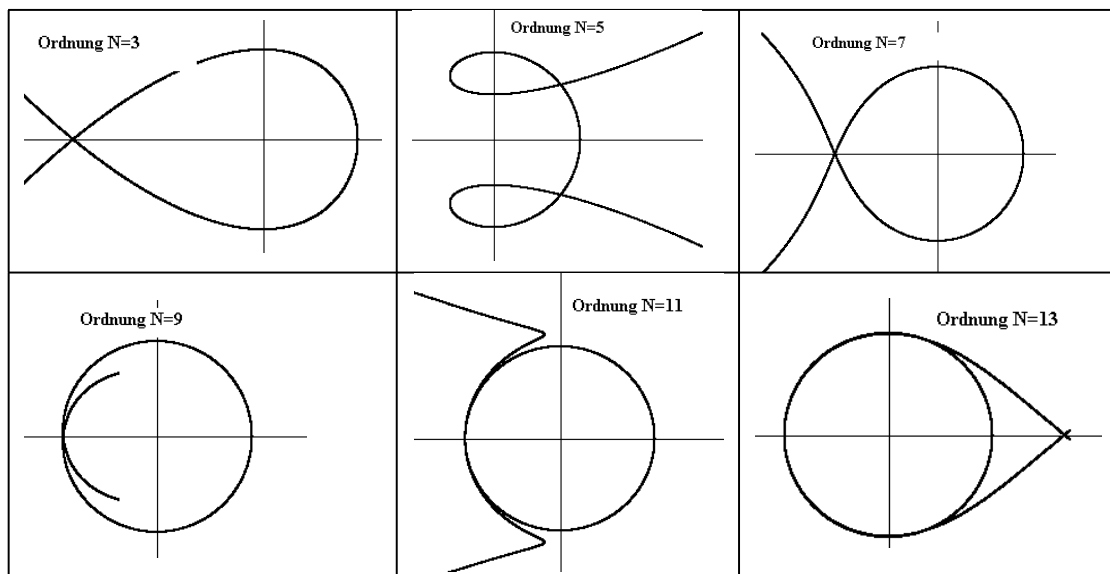
(4.1.12) Erneut ist der erste Teil des weiteren Vorgehens Routine. Man beginnt mit der Eindeutigkeit einer solchen Zerlegung. Diese folgt in der üblichen Weise:

Gleichsetzen zweier angenommener Zerlegungen derselben Ordnung. / Fortheben der Terme ohne Δt / Division durch Δt / Grenzübergang. Das gibt Gleichheit der Koeffizienten für $k=1$. / Fortheben dieser Terme in der Ausgangsgleichung / Division durch Δt^2 . Usw. Also: **Falls es überhaupt eine Zerlegung der eingeführten Art gibt, ist sie eindeutig.** Die Polynomkoeffizienten sind damit alle eindeutig bestimmt. Was noch fehlt ist der Zusammenhang mit den Ableitungen.

(4.1.13) Die nachfolgende Figur illustriert die Approximationswirkung solcher Zerlegungen am Beispiel des Kreises, insbesondere den Einfluß der Ordnung N . Dargestellt ist die Entwicklung der Kreiskurve $t \mapsto (\cos(t), \sin(t))$ um den Punkt $t_0=0$. Hier erhält man die Entwicklung leicht durch Einsetzen der entsprechenden Entwicklungen für \cos und \sin . Nacheinander wird jeweils bis zur Ordnung 3,5,7 (obere Zeile) sowie 9,11,13 (untere Zeile) entwickelt. Wegen $t_0 = 0$ ist hier $\Delta t = t - 0 = t$. Die Entwicklung bis zur Ordnung 5, die zu der zweiten Figur führt, lautet beispielsweise:

$$t \mapsto \begin{pmatrix} \cos t \\ \sin t \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} + t \begin{pmatrix} 0 \\ 1 \end{pmatrix} + \frac{t^2}{2} \begin{pmatrix} -1 \\ 0 \end{pmatrix} + \frac{t^3}{3!} \begin{pmatrix} 0 \\ -1 \end{pmatrix} + \frac{t^4}{4!} \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \dots$$

Die erste Komponente ist die Horizontalkomponente. Die größte auftretende Potenz dominiert für große t . Bei positivem zugehörigen Koeffizienten, öffnet sich die Kurve nach rechts ($x > 0$). Bei $N=5$ etwa dominiert $t^4 \frac{1}{4!}$. Für $N=10$ dominiert hier $t^9 \frac{1}{9!}$. Dieser Koeffizient ist viel kleiner, d.h. die Dominanz setzt sehr viel später ein. Für $n=13$ hat man bereits so etwas wie einen vollen Kreisumlauf. Weiter unten in (4.3.8) werden wir eine allgemeine Formel zur Berechnung derartiger Koeffizienten angeben.



(4.1.14) Die Zerlegung erlaubt erneut die Verwendung und Nutzung der beiden Denkfiguren, die in (3.1.8) für die Tangentenerlegung eingeführt wurden.

(4.1.15) Als **Anwendungsbeispiel bestimmen wir eine Formel für den Krümmungsradius**. Sei $\vec{r} \in C^2(\mathbb{R}, V_0^2)$ eine glatte ebene Kurve. In (1.3.34) haben wir hierfür bereits das begleitende Zweibein eingeführt und in (2.3.38) haben wir das Schnittpunktsproblem der Normalen angesprochen. Jetzt wollen wir eine Formel für den Krümmungsradius herleiten, wozu wir die zweite Ableitung benötigen.

(4.1.16) Differenzieren und Betragsabspaltung gibt $\dot{\vec{r}}(t) = v(t)\vec{e}_T(t)$ mit $v(t) = \left| \dot{\vec{r}}(t) \right|$. Erneutes Ableiten mit der Produktregel gibt $\ddot{\vec{r}}(t) = \dot{v}(t)\vec{e}_T(t) + v(t)\dot{\vec{e}}_T(t)$. Was ist $\dot{\vec{e}}_T(t)$? Es gilt $\vec{e}_T(t) \cdot \vec{e}_T(t) = 1$. Leitet man mit Hilfe der Produktregel ab, so folgt $\vec{e}_T(t) \cdot \dot{\vec{e}}_T(t) = 0$. D.h. aber, dass die beiden Vektoren senkrecht aufeinander stehen und das heißt, dass $\dot{\vec{e}}_T(t)$ stets dieselbe Richtung wie $\vec{e}_N(t)$ haben muß.

(4.1.17) Wir setzen (definieren):

$$\dot{\vec{e}}_T(t) = \kappa(t)v(t)\vec{e}_N(t) \quad (\text{Definierende Gleichung für } \kappa!).$$

Dass es günstig ist, den Faktor $v(t)$ vom Betrag abzuspalten, kann man auf mehrere Weisen sehen. Etwa indem man das Normalenschnittproblem bis zu Ende rechnet und am Ende diesen Faktor zur Vereinfachung benötigt. Hat r die Einheit Länge, dann hat κ die Einheit $1/\text{Länge}$. Der gesamte Betrag wird per Definition gleich κv gesetzt. $v(t) \neq 0$ ist daher wie üblich unzulässig.

(4.1.18) Gehen wir damit in die Gleichung für die Beschleunigung ein, so finden wir mit der Produktregel:

$$\ddot{\vec{r}}(t) = \dot{v}(t)\vec{e}_T(t) + v^2(t)\kappa(t)\vec{e}_N(t).$$

D.h. κ hängt mit der Normalbeschleunigung zusammen: $b_N(t) = \kappa(t)v^2(t)$.

(4.1.19) Um die geometrische Bedeutung von κ zu finden, rechnen wir den früher besprochenen Grenzwert der Normalenschnittpunkte aus. Zunächst noch etwas Vorbereitung: Wir benutzen den Trick mit dem Skalarprodukt erneut. Und zwar differenzieren wir jetzt die Gleichung $\vec{e}_N(t) \cdot \vec{e}_T(t) = 0$. Dann folgt: $\dot{\vec{e}}_N(t) \cdot \vec{e}_T(t) + \vec{e}_N(t) \cdot \dot{\vec{e}}_T(t) = 0$. Also $\dot{\vec{e}}_N(t) \cdot \vec{e}_T(t) = -\kappa(t)v(t)$. Nun hat $\dot{\vec{e}}_N(t)$ die Richtung von $\vec{e}_T(t)$, so dass folgt:

$$\dot{\vec{e}}_N(t) = \vec{e}_T(t) \left(\vec{e}_T(t) \cdot \dot{\vec{e}}_N(t) \right) = -\kappa(t)v(t)\vec{e}_T(t)$$

Im Gegensatz zu (4.1.7) ist das keine Definition, sondern eine gültige Folgerung.

(4.1.20) Zurück zum Normalenproblem. Die Normale zur Zeit t hat die Parametrisierung $\alpha \mapsto \vec{r}(t) + \alpha\vec{e}_N(t)$. Die Normale zur Zeit $t + \Delta t$ hat die Parametrisierung $\beta \mapsto \vec{r}(t + \Delta t) + \beta\vec{e}_N(t + \Delta t)$. Wir erhalten mit der 2. Denkfigur die beiden Tangentenerlegungen

$$\vec{r}(t + \Delta t) = \vec{r}(t) + v(t)\vec{e}_T(t)\Delta t + \dots \quad \text{und} \quad \vec{e}_N(t + \Delta t) = \vec{e}_N(t) + (-\kappa(t)v(t)\vec{e}_T(t))\Delta t + \dots$$

Die Resttermbeiträge haben wir wie vielfach üblich durch \dots angedeutet. Um den Schnittpunkt zu bestimmen, setzen wir die beiden Normalenparametrisierungen gleich. Das ergibt für $\Delta t = 0$ zunächst $\alpha = \beta$. Danach teilen wir wie üblich durch Δt und gehen zum Grenzübergang $\Delta t \rightarrow 0$ über. Das liefert $\beta = \frac{1}{\kappa(t)}$.

(4.1.21) **Damit haben wir den Grenzwert der Normalenschnittpunkte gefunden:**

$$\text{Der Grenzwert der Normalenschnittpunkte liegt bei } \vec{r}_K(t) = \vec{r}(t) + \frac{1}{\kappa(t)}\vec{e}_N(t)$$

Erwartungsgemäß ist das der Mittelpunkt des Krümmungskreises: Denkt man sich die Kurve in der Nähe von t durch einen approximierenden Kreis ersetzt, dann ist das der Schnittpunkt benachbarter Radien, also der Kreismittelpunkt. Damit haben wir eine rein geometrische Interpretation für die Krümmung κ gefunden.

(4.1.22) Für die Beschleunigung besagt das: Der Betrag der Normalbeschleunigung ist $\frac{v^2(t)}{R(t)}$, wenn $R(t) = \frac{1}{\kappa(t)}$ der lokale Krümmungsradius ist (für $\kappa(t) \neq 0$). Vergleichen Sie mit der in Physikeinführungsbüchern vielfach gegebenen Zerlegung der Beschleunigung in tangentielle und radiale Komponente. Wie wird dort vorgegangen?

- Hat man eine Komponentendarstellung $\vec{r}(t) = \vec{e}_1 x(t) + \vec{e}_2 y(t)$ so kann man die Krümmung durch die Ableitungen von x und y ausdrücken. Tun Sie das. Spezialisieren Sie dann auf den Fall $x(t)=t$ und $y(t)=f(t)$, also die kanonische Parametrisierung eines Graphen.
- Bestimmen sie Krümmung und Krümmungsradius für den Fall einer Parabel $y=\alpha x^2$.

6.4.1d Der Satz vom beschränkten Zuwachs

(4.1.23) Wir besprechen einen nützlichen mathematischen Satz, der ein Beispiel einer gültigen (streng beweisbaren) Verallgemeinerung eines physikalischen Sachverhaltes darstellt. Mit seiner Hilfe lassen sich wiederum zahlreiche weitere wichtige Resultate beweisen. Worum geht es?

(4.1.24) Nehmen wir an, eine Bahnkurve $t \mapsto \vec{r}(t)$ beschreibe die glatte Bewegung eines Massenpunktes im Intervall $[a, b]$. Zur Verdeutlichung interpretieren wir den Punkt als Auto. Dann ist $\vec{r}(b) - \vec{r}(a)$ der Abstandsvektor von Anfangs- und Endpunkt und $\|\vec{r}(b) - \vec{r}(a)\|$ ist der Abstand - die Entfernung - dieser beiden Punkte. Das Auto hat sich zwischenzeitlich (nicht notwendig geradlinig) mit der momentanen Geschwindigkeit $\dot{\vec{v}}(t) = \dot{\vec{r}}(t)$ bewegt und ist so vom Anfangs- zum Endpunkt gelangt. $\vec{r}(b) - \vec{r}(a)$ bedeutet einen geschlossenen Weg. $\|\dot{\vec{r}}(t)\|$ ist die zugehörige skalare Geschwindigkeit, die während der Fahrt vom Tachometer angezeigt wird.

Weiter stellen wir uns vor, wir hätten ein zweites Auto zur Verfügung, das geradlinig auf einer x -Achse fährt, mit einer Bahnkurve $t \mapsto x(t)$ und momentaner Geschwindigkeit $v(t) \geq 0$. Dieses Auto legt die Strecke $x(b) - x(a)$ zurück. Jetzt kommt die entscheidende Annahme: Das zweite Auto soll zu keinem Zeitpunkt langsamer als das erste sein. Oder $\|\dot{\vec{r}}(t)\| \geq v(t)$ für alle t mit $a \leq t \leq b$. Dann sagt uns die Lebenserfahrung, der Umgang mit realen Objekten unserer Welt, dass $x(b) - x(a)$ nie kleiner sein kann als $\|\vec{r}(b) - \vec{r}(a)\|$. Lläuft der Hase in jedem Zeitpunkt schneller als der Igel und überdies noch geradlinig, dann entfernt er sich notwendig weiter vom Startpunkt als der Igel!

(4.1.25) Dieser Sachverhalt gilt tatsächlich allgemein (für alle Kurven und Dimensionen):

Satz vom beschränkten Zuwachs:
 Sei $a, b \in \mathbb{R}$ mit $a < b$ und $I = [a, b]$. W ein Banachraum.
 Weiter seien $\vec{r}: I \rightarrow W$ und $x: I \rightarrow \mathbb{R}$ zwei stetige Abbildungen.
 Beide seien im Innern von J differenzierbar und es gelte
 $\|\dot{\vec{r}}(t)\| \leq \dot{x}(t)$ für alle t mit $a \leq t \leq b$.
 Dann gilt $\|\vec{r}(b) - \vec{r}(a)\| \leq x(b) - x(a)$.

(4.1.26) **Beweis:** Wir wählen eine Zahl $\varepsilon > 0$ und betrachten die Ungleichung

$$\|\vec{r}(t) - \vec{r}(a)\| \leq [x(t) - x(a)] + \varepsilon(t - a) + \varepsilon \quad (*)$$

Zu dieser Ungleichung definieren wir die folgende Menge $J = J_\varepsilon$:

$$J_\varepsilon = \{s \mid a \leq s \leq b \text{ und } (*) \text{ ist für alle } t \text{ mit } a \leq t \leq s \text{ erfüllt}\}.$$

◆ Wir werden zeigen: $J_\varepsilon = [a, b]$ für jedes $\varepsilon > 0$. **Das ergibt sofort die Behauptung.**

◆◆ Jetzt zur Bestimmung von $J = J_\varepsilon$. Sicher ist $a \in J$. Da beide Seiten stetig sind, ist auch ein kleines Intervall $[a, c]$ in J enthalten. Ist c nicht gleich b , so ist die Menge aller s , die $(*)$ nicht erfüllen, nicht leer und besitzt eine größte untere Schranke. Diese gehört mit zu J , denn $(*)$ ist für alle $s < c$ erfüllt und daher aus Stetigkeitsgründen auch für $s = c$. Insbesondere muß auch $c < b$ gelten. (Im anschaulichen Bild: c ist der hypothetische Zeitpunkt, nach dem der langsame Igel bis auf den ε -Beitrag zum ersten Mal weiter gekommen wäre als der schnelle Hase!)

◆◆◆ Wir zeigen, dass man ein s mit $c < s < b$ wählen kann, für das $(*)$ doch erfüllt ist, was im Widerspruch zur Konstruktion von c steht. **Und das heißt, dass $J = [a, b]$ gelten muß.**

Zur Konstruktion dieses Punktes s . Die Dreiecksungleichung gibt

$$\|\vec{r}(s) - \vec{r}(a)\| \leq \|\vec{r}(s) - \vec{r}(c)\| + \|\vec{r}(c) - \vec{r}(a)\| \quad (**)$$

Der zweite Summand erfüllt (*), ist also $\leq x(c) - x(a) + \varepsilon(c - a) + \varepsilon$. Den ersten Summanden müssen wir (etwas länglich) abschätzen. Wir haben $\left\| \dot{\vec{r}}(c) \right\| \leq \dot{x}(c)$ laut Voraussetzung. D.h. es muß $\dot{\vec{r}}(c) = \vec{u}\dot{x}(c)$ gelten für einen festen Vektor \vec{u} mit $\|\vec{u}\| \leq 1$. Damit bilden wir die Hilfsfunktion $\vec{f}(t) = \vec{r}(t) - \vec{u}x(t)$. Sie ist differenzierbar und es gilt per Konstruktion $\vec{f}(c) = \vec{0}$. Wir erhalten folgende Tangentenzersetzung von \vec{f} um c herum: $\vec{f}(s) = \vec{f}(c) + (s-c)\vec{R}(c,s-c)$. Laut Resttermbedingung können wir $s-c$ so klein wählen, dass $\left\| \vec{R}(c, t-s) \right\| < \varepsilon$ für $c \leq t \leq s$ wird. Also $\left\| \vec{f}(t) - \vec{f}(c) \right\| \leq (t-c)\varepsilon$ für alle t mit $c \leq t \leq s$. Wir haben ja $s-c > 0$ angenommen. Ausgeschrieben:

$$\|(\vec{r}(t) - \vec{r}(c) - \vec{u}(x(t) - x(c)))\| \leq (t - c)\varepsilon.$$

Mit dieser Ungleichung erhalten wir schließlich über die Dreiecksungleichung

$$\begin{aligned} \|\vec{r}(t) - \vec{r}(c)\| &= \|(\vec{r}(t) - \vec{r}(c)) - \vec{u}(x(t) - x(c)) + \vec{u}(x(t) - x(c))\| \\ &\leq \|(\vec{r}(t) - \vec{r}(c)) - \vec{u}(x(t) - x(c))\| + \|\vec{u}(x(t) - x(c))\| \\ &\leq (t - c)\varepsilon + \|\vec{u}\|(|x(t) - x(c)|). \end{aligned}$$

Nun gilt aber $\|\vec{u}\| \leq 1$ Und $x(t) > x(c)$. Letzteres, da x wegen $\dot{x}(t) \geq 0$ notwendig monoton wächst. Das gibt zusammen die gesuchte Abschätzung

$$\|\vec{r}(t) - \vec{r}(c)\| \leq (t - c)\varepsilon + (x(t) - x(c)) \quad \text{für alle } t \text{ mit } c \leq t \leq s.$$

Hiermit gehen wir in (***) ein und erhalten insgesamt:

$$\begin{aligned} \|\vec{r}(t) - \vec{r}(a)\| &\leq [(t - c)\varepsilon + (x(t) - x(c))] + [x(c) - x(a) + (c - a)\varepsilon + \varepsilon] \\ &= [x(t) - x(a)] + (t - a)\varepsilon + \varepsilon \quad \text{für alle } t \text{ mit } a \leq t \leq s. \end{aligned}$$

Beachten sie, dass wir **nie** ungerechtfertigt Betragsstriche fortgelassen haben, so dass die Beiträge mit c sich wirklich fortheben. Damit sind wir am Ziel: Das gewählte s gehört doch zur Menge J , denn die definierende Ungleichung ist ja für die t oberhalb c erfüllt und laut Konstruktion von c auch für die darunter. Der gesuchte Widerspruch ist geführt, der Satz ist bewiesen.

- Welches Detail, das im Beweis benutzt - also als bekannt und bewiesen vorausgesetzt wurde, wäre noch zu beweisen? Wie kann man in dieser Situation einen schönen "circulus vitiosus" konstruieren?

(4.1.27) Eine wichtige Folgerung die wir benötigen werden: Gilt für s mit $t_0 - \varepsilon < s < t_0 + \varepsilon$ die Ungleichung

$$\left\| \dot{\vec{f}}(s) \right\| \leq M, \text{ dann folgt}$$

$$\left\| \vec{f}(s) - \vec{f}(t_0) \right\| \leq M|s - t_0|.$$

- Beweisen Sie (4.1.27).

6.4.2 Stammfunktion und Integral.

(4.2.1) Wie im Falle $n=1$ kann man die Integration von Kurven zunächst als Umkehrung der Ableitung einführen. Mit Hilfe der Integration ist ein günstiger Zugang zur Polynomapproximation von Kurven möglich, den wir unten benutzen werden.

Definition: Sei $\vec{f}: I \rightarrow W$ stetige Kurve im nicht einpunktigem Intervall I .
Dann gelten die folgenden Bezeichnungen:

- 1) Eine Kurve $\vec{F}: I \rightarrow W$ heißt Stammfunktion zu \vec{f} , wenn $\dot{\vec{F}} = \vec{f}$ gilt. (Also: Die Ableitung einer Stammfunktion ergibt die Ausgangskurve.)
- 2) Fixiert man bei einer Stammfunktion eine Nullstelle, so erhält man eine Integralfunktion:

$\left(I, t \mapsto \int_a^t ds \vec{f}(s) = \vec{F}(t), V \right)$

 mit $\vec{F}'(t) = \vec{f}(t)$ und $\vec{F}(a) = \vec{0}$ $a \in I$

Die **Werte** der Integralfunktion heißen *bestimmte Integrale*.

(4.2.2) Ist \vec{F} im Intervall J Stammfunktion zu \vec{f} , dann ist $(W, \vec{c} \mapsto \vec{F} + \vec{c}, \mathfrak{F}(J, W))$ eine Parametrisierung aller Stammfunktionen. (Das folgt aus dem Satz vom beschränkten Zuwachs.) Hierbei bezeichnet $\vec{c} = (J, t \mapsto \vec{c}, W)$ die vom Element \vec{c} aus W erzeugte konstante Funktion. Man hat die folgende geometrische Konsequenz:

(4.2.3) **Geometrische Stammfunktionscharakterisierung:** Besitzt \vec{f} wenigstens eine Stammfunktion \vec{F} , dann geht durch jeden Punkt (t_0, \vec{w}) des Graphenraumes genau eine Stammfunktion - genauer der Graph dieser Stammfunktion. (D.h. wir erhalten eine Partition des Graphenraumes durch Kurvengraphen.)

□ Beweisen Sie (4.2.3). Das Problem ist, zu zeigen, dass höchstens eine Stammfunktion durch jeden Punkt geht. Das übliche Vorgehen: Angenommen F und G sind beides derartige Stammfunktionen. Bilden Sie $F-G$ und wenden Sie den Satz vom beschränkten Zuwachs an.

(4.2.4) Ein vertrautes Beispiel aus der Physik: Ist $t \mapsto \vec{v}(t)$ eine Geschwindigkeitskurve, dann ist die zugehörige Bahnkurve \vec{r} Stammfunktion zu \vec{v} . Gibt man jetzt den Ortsvektor für einen einzigen Zeitpunkt vor, dann legt die Geschwindigkeit die Bahnkurve für alle Zeiten fest.

(4.2.5) Eine einfache, für uns wichtige **Verallgemeinerung des Integralbegriffes** sieht nun so aus:

Sei \vec{f} stückweise auf J stetig - d.h. stetig in allen Punkten bis auf eine endliche Menge von Ausnahmepunkten. - Dann heißt \vec{F} Stammfunktion zu \vec{f} in J , wenn F stetig ist und $\vec{F}'(t) = \vec{f}(t)$ für alle Nichtausnahmepunkte gilt.

Die Bedingung der Stetigkeit von \vec{F} in den Ausnahmepunkten ist erforderlich, um die Eindeutigkeit des Integrales zu sichern!

Beispiel: Sei $\vec{f}(t) = \vec{0}$ für $t < 0$ und $f(t) = \vec{g}$ für $t \geq 0$. Bilde \vec{F} mit $\vec{F}(t) = \vec{0}$ für $t < 0$ und $\vec{F}(t) = \vec{g}t$ für $t \geq 0$. Dann ist \vec{F} Stammfunktion zu \vec{f} im Sinne der Definition! Ändert man \vec{F} für $t < 0$ ab zu $\vec{F}(t) = \vec{a}$ mit $\vec{a} \neq \vec{0}$, so ergibt sich keine Stammfunktion, da die Stetigkeit bei $t=0$ verletzt wäre. Die Ableitung dagegen bleibt in allen Punkten, in denen man differenzieren kann, korrekt. Wir hätten unendlich viele Stammfunktionen, die für $t > 0$ alle übereinstimmen.

"Stammfunktion" ist nachfolgend immer in diesem verallgemeinerten Sinne zu verstehen.

Die Stetigkeit in den Problempunkten sichert, dass die geometrische Stammfunktionscharakterisierung weiterhin gilt. D.h. durch jeden Punkt des Graphenraumes läuft genau ein Stammfunktionsgraph!

(4.2.6) Die üblichen elementaren Integrationsregeln (Linearität, Additivität....) gelten für das eingeführte Integral. Sie sind elementare Folgen des Hauptsatzes:

(4.2.7) **Hauptsatz der Differential- und Integralrechnung (Technische Form):**

Sei \vec{f} im Intervall J stückweise stetig und \vec{F} zugehörige Stammfunktion.
Dann gilt:
$$\int_a^t ds \vec{f}(s) = \vec{F}(s) \Big|_a^t = \vec{F}(t) - \vec{F}(a) \quad \text{für alle } t \in J.$$

Zum Beweis muß man nur die geometrische Stammfunktionscharakterisierung auf beide Seiten anwenden.

(4.2.8) Durch Anwenden des Satzes vom beschränkten Zuwachs auf die triviale Ungleichung $\|\vec{f}(s)\| \leq \|\vec{f}(s)\| = \dot{g}(s)$ erhalten wir folgende wichtige Beziehung,

die **Dreiecksungleichung für Integrale:** $\left\| \int_a^b ds \vec{f}(s) \right\| \leq \int_a^b ds \|\vec{f}(s)\|$

Interpretieren wir die Gleichung im Rahmen eines Bewegungsprozesses: D.h. \vec{f} sei eine Geschwindigkeitskurve und \vec{r} eine zugehörige Stammfunktion, also eine Bahnkurve. Dann ist das Integral links gleich $\vec{r}(b) - \vec{r}(a)$, also gleich dem vektoriellen Abstand von Anfangs- und Endpunkt. Und $\|\vec{r}(b) - \vec{r}(a)\|$ ist die (geradlinige) Entfernung dieser beiden Punkte, ihr skalarer Abstand. Auf der rechten Seite dagegen wird über den Betrag der Geschwindigkeit integriert, was die Weglänge ergibt! D.h. die Länge der tatsächlich entlang Bild \vec{r} gefahrenen Wegstrecke. Und die ist nie kleiner als der geradlinige Abstand.

Das ist der Inhalt obiger Ungleichung, die sich als ausgesprochen nützlich erweist.

6.4.2a Das Integral als Grenzwert von Summen

(4.2.9) Wir beschreiben ein für die Integration wichtiges Szenenbild:

Sei $J=[a,b]$ das von uns betrachtete Intervall. Wir wählen $N+1$ Punkte t_0, t_1, \dots, t_N aus J aus. Dabei soll $t_0=a < t_1 < \dots < t_{N-1} < t_N = b$ gelten. Dies gibt uns eine Zerlegung von J in N Teilintervalle $J_k = [t_{k-1}, t_k[$ für $k=1, 2, \dots, N-1$ und $J_N = [t_{N-1}, b]$. Um die Disjunktheit der Teilintervalle an den Intervallgrenzen zu sichern, kann man nicht einfach $[t_{k-1}, t_k]$ nehmen. Die Länge des Teilintervalles J_k bezeichnen wir mit Δ_k . Also $\Delta_k = t_k - t_{k-1}$.

Wir sagen kurz: " $\mathfrak{Z}_N = \{J_1, J_2, \dots, J_N\}$ ist Zerlegung von J in N Teilintervalle". Weiter sei $\varepsilon = \varepsilon(\mathfrak{Z}_N)$ das Maximum der zugehörigen Δ_k , also die Länge des größten Teilintervalles. Die Intervalle können gleichlang sein, müssen es aber nicht. Im ersteren Fall gilt $\varepsilon = \frac{1}{N}(b-a)$.

□ Fertigen Sie eine Skizze, die die Größen des Szenenbildes veranschaulicht und die Sie nachfolgend immer zur Orientierung verwenden können und sollten.

(4.2.10) Jetzt wählen wir aus jedem J_k einen Vektor \vec{f}_k ($k=1, 2, \dots, N$) und definieren $\vec{f}(t) = \vec{f}_k$ für $t \in J_k$. Dies definiert uns eine stückweise glatte (sogar stückweise konstante) Funktion $\vec{f}: J \rightarrow W$. Wir wollen f integrieren! Die ursprüngliche Definition erlaubt das nicht, mit der erweiterten Definition (4.2.5) ist das jedoch zulässig! Mit Hilfe der Additivität folgt sofort:

$$\int_a^b ds \vec{f}(s) = \sum_{k=1}^N \int_{t_{k-1}}^{t_k} ds \vec{f}(s) = \sum_{k=1}^N \int_{t_{k-1}}^{t_k} ds \vec{f}_k = \sum_{k=1}^N \vec{f}_k \Delta_k.$$

(4.2.11) Mit Hilfe diese Zwischenresultates klären wir die **grundlegende Beziehung zwischen Integral und Summe**, die besagt, dass ein Integral so etwas ist wie eine "unendliche Summe aus lauter unendlich kleinen Summanden". Beachten Sie den Unterschied: Bei einer unendlichen Reihe $\sum a_n$, werden die Summanden a_n mit zunehmendem n immer kleiner. Hier aber sollen alle Summanden klein sein. Wir wollen zeigen, dass ein solches Konzept durchaus realisierbar ist.

(4.2.12) Wir kehren zu unserem Szenenbild (4.2.9) zurück. \vec{f} sei jetzt eine stetige Kurve und \mathfrak{Z}_N die eingeführte Intervallzerlegung. Für jedes k wählen wir ein $u_k \in J_k$, eventuell auch aus $\bar{J}_k = [t_{k-1}, t_k]$ und bilden den Zwischenwert $\vec{f}_k = \vec{f}(u_k)$. Das ist so etwas wie ein typischer Kurvenwert im Bereich von J_k . Schließlich setzen wir $\vec{f}(t) = \vec{f}_k + \vec{r}_k(t)$ für $t_{k-1} \leq t \leq t_k$. Die Funktionen \vec{r}_k mit $k=1, 2, \dots, N$ beschreiben daher die Abweichungen der Kurve in J_k vom gewählten typischen Wert. Oder auch: den absoluten Fehler. Diese Interpretation sollte man im Auge behalten. Jetzt berechnen wir analog zu (4.2.10) das Integral:

$$\int_a^b dt \vec{f}(t) = \sum_{k=1}^N \int_{t_{k-1}}^{t_k} ds \vec{f}(t) = \sum_{k=1}^N \int_{t_{k-1}}^{t_k} ds \left(\vec{f}_k + \vec{r}_k(s) \right) = \sum_{k=1}^N \vec{f}_k \Delta_k + \sum_{k=1}^N \int_{t_{k-1}}^{t_k} ds \vec{r}_k(s)$$

Die letzte Summe ist neu! Sie gibt den Beitrag der Abweichungen (der Funktion von den typischen Werten) zu Mintegral. Nehmen wir an, es gelingt zu zeigen, dass diese Summe mit zunehmender Verfeinerung der Zerlegung gegen Null geht. Dann besagt die Gleichung: Der feste, links stehende Integralwert ist gleich der ersten Summe plus einem immer kleiner werdenden Fehler. Die erste Summe enthält aber N Summanden, die einzeln wegen des Intervalllängenfaktors Δ_k immer kleiner werden. Die Anzahl N der Summanden geht dagegen nach unendlich. In der naiven formalen Grenzwertbeschreibung: *Unendlich viele unendlich kleine Summanden*.

(4.2.13) Vom mathematischen Standpunkt aus besteht der Wunsch, das Verschwinden des Fehlers für möglichst viele und möglichst allgemeine Funktionen zu zeigen! Darauf wollen wir verzichten und für das Verschwinden statt dessen eine hinreichende Bedingung formulieren, die aber dafür deutlich zeigt, auf welche Weise der Zusatzbeitrag verschwindet.

(4.2.14) Wir nehmen dazu an, dass die Kurve \vec{f} im Intervall differenzierbar sei mit **beschränktem Ableitungsbetrag**. D.h. wir können eine Konstante M finden, so dass $\|\vec{f}'(t)\| \leq M$ in ganz J gilt. Dann wenden wir den Satz vom beschränkten Zuwachs in der Form (4.1.27) an und finden wegen $\vec{f}_k = \vec{f}(u_k)$ für die Abweichungsfunktionen in den zugehörigen Intervallen:

$$\|\vec{r}_k(t)\| = \|\vec{f}(t) - \vec{f}(u_k)\| \leq M|t - u_k| \leq M\Delta_k \leq M\varepsilon \quad \text{für } t \in \bar{J}_k \quad \text{und} \quad \varepsilon = \max_k \Delta_k.$$

Damit gehen wir in den Ausdruck für den Rest ein unter Beachtung von (4.2.8):

$$\begin{aligned} \|\text{Rest}\| &= \left\| \sum_{k=1}^N \int_{t_{k-1}}^{t_k} ds \vec{r}_k(s) \right\| \leq \sum_{k=1}^N \int_{t_{k-1}}^{t_k} ds \|\vec{r}_k(s)\| \\ &\leq M\varepsilon \left(\sum_{k=1}^N \int_{t_{k-1}}^{t_k} ds 1 \right) = \varepsilon M(b-a). \end{aligned}$$

Im vorletzten Schritt haben wir die Abschätzung benutzt, die aus der geforderten Beschränktheit der Ableitung folgt. Lassen wir jetzt ε nach Null gehen, dann geht der Rest nach Null und das heißt, dass sich der erste Summenausdruck $\sum_{k=1}^N \vec{f}_k \Delta_k$ aus (4.2.12) immer stärker dem Integralwert nähert. Der Rest verschwindet, weil die Abweichungen in der vom Satz vom beschränkten Zuwachs gesicherten Weise beschränkt sind!

(4.12.15) Die Summen $\sum_{k=1}^N \vec{f}_k \Delta_k$ selbst lassen sich inhaltlich als **gewichtete Mittelwerte** der Datensätze $(\vec{f}_1, \vec{f}_2, \dots, \vec{f}_N)$ interpretieren! Genauer $(b-a)\vec{f} = \sum_k \vec{f}_k \Delta_k$. Das jeweilige Gewicht ist die Breite Δ_k des Intervalles, in dem der konstante Wert \vec{f}_k angenommen wird. Bei gleichlangen Intervallen wird das wegen

$\varepsilon = \Delta_k = \frac{1}{N}(b-a)$ noch deutlicher: Dann erhält man die übliche Mittelwertformel $\vec{f} = \frac{1}{N} \sum_k \vec{f}_k$. Den

(ja existierenden!) Grenzwert dieser Größen $\vec{f} = \vec{f}(\mathfrak{Z}_N)$ für $\varepsilon \rightarrow 0$ kann bzw. muß man als Mittelwert der Kurvenwerte - mittlerer Ortsvektor oder Ähnliches - interpretieren.

(4.2.15) Fassen wir zusammen:

Das Integral (über Kurven) besitzt die folgende inhaltliche Interpretation:

$$\int_a^b dt \vec{f}(t) = (b-a)\vec{f} = \sum_k \vec{f}_k \Delta_k + \text{Rest} \approx \sum_k \vec{f}_k \Delta_k.$$

\vec{f} = Mittelwert der Kurvenwerte im Integrationsbereich.

Der Rest geht bei ausreichender Verfeinerung der Zerlegung nach Null, **falls** die Kurve \vec{f} angemessene Voraussetzungen erfüllt. Die üblichen in den Anwendungen auftretenden Funktionen erfüllen solche Bedingungen.

Beispielsweise genügt eine Abschätzung $\|\vec{f}'(t)\| \leq M$ in $[a,b]$, also die Beschränktheit der Ableitung.

6.4.2b Die Komponentendarstellung des Integrales

(4.2.16) Sei $b = (\vec{b}_i)$ ein **konstantes (!)** Basisfeld von W . Dann können wir jeden Wert $\vec{f}(t)$ bezüglich dieser Basis darstellen. Da die Basisvektoren konstant sind, kann man die Darstellung per Linearität auswerten und erhält **gewöhnliche Integrale über die Komponentenfunktionen:**

$$\int_a^b dt \vec{f}(t) = \int_a^b dt \left(\sum_i \vec{b}_i f_i(t) \right) = \sum_i \left(\int_a^b dt f_i(t) \right) \vec{b}_i$$

Kurz: **Bei konstantem Basisfeld darf man einfach komponentenweise integrieren!** Im Prinzip ist damit die neue vektorwertige Integration (über Kurven) auf Integrale der üblichen Art (für die Komponenten) zurückgeführt.

- Berechnen Sie nach dieser Methode für $\vec{x}^K(t) = R \begin{pmatrix} \cos(\pi t) \\ \sin(\pi t) \\ t \end{pmatrix}$ und $\vec{\omega}^K = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$ das Integral $\int_0^1 dt \vec{\omega}^K \times \vec{x}^k(t)$

6.4.3 Die Taylorentwicklung der Kurven

(4.3.1) Eine leicht durch Induktion zu beweisende, aber natürlich nicht so leicht selbst zu findende Umformung für $N=1,2,3,\dots$ sieht wie folgt aus (s unabh. Variable, t äußerer Parameter):

$$\vec{f}'(s) = - \sum_{k=1}^N \frac{d}{ds} \left(\frac{(t-s)^k}{k!} \vec{f}^{(k)}(s) \right) + \frac{1}{N!} (t-s)^N \vec{f}^{(N+1)}(s).$$

(Zum Vertrautwerden mit der Formel: Diese für $N=1,2,3$ ausschreiben und verstehen. Der Induktionsschluß ist dann sehr einfach!). Wir nehmen an, die gegebene Identität sei bewiesen. Dann können wir sie bezüglich s integrieren. Als Grenzen wählen wir a und t. Beachten Sie: der äußere Parameter t steht dann in der Grenze, aber auch im Integranden! Zur Variablenbezeichnung: $t=a+(t-a)=a+\Delta t$.

(4.3.2) Die Integration gibt

$$\boxed{\vec{f}(t) = \sum_{k=1}^N \frac{1}{k!} \vec{f}^{(k)}(a)(t-a)^k + \frac{1}{N!} \int_a^t ds (t-s)^N \vec{f}^{(N+1)}(s).}$$

Wie üblich wird von der Kurve $t \mapsto \vec{f}(t)$ stillschweigend angenommen, dass für sie die Manipulationen der Formel sinnvoll sind.

□ Was für Glattheitsannahmen reichen beispielsweise aus?

(4.3.3) Die gefundene Formel hat für uns einen doppelten Nutzen. Zunächst liefert sie tatsächlich die in (4.1.11) eingeführte lokale Polynomapproximation jeder Ordnung! D.h. sie liefert die gesuchten Koeffizienten. Vergleich zeigt, dass $\vec{a}_k = \frac{\vec{f}^{(k)}(a)}{k!}$ zu setzen ist. Die Koeffizienten werden somit tatsächlich durch die Ableitungen festgelegt. Beachten Sie das k! im Nenner, das für höheres k stark verkleinernd wirkt. Beachten Sie auch den folgenden Punkt: Erhöht man die Ordnung N um 1, dann muß man einen Koeffizienten neu bestimmen, die übrigen bleiben jedoch unverändert, können übernommen werden. Diese nützliche Eigenschaft ist keineswegs denkwürdig! Bei Erhöhung der Ordnung könnten sich ebensogut sämtliche Koeffizienten ändern.

□ Testen Sie die Entwicklungsformel am Beispiel der Entwicklung aus (4.1.13).

(4.3.4) Natürlich ist die Restterm eigenschaft noch nachzuweisen! Das ist jetzt allgemein möglich, denn unsere Formel bestimmt den Restterm nicht einfach nur auf dem Niveau der Bezeichnungen, sondern sie liefert uns eine **Integraldarstellung für den Restterm**.

Auf den ersten Blick allerdings sieht die Integraldarstellung wenig brauchbar aus. Wir vermissen den Faktor $\Delta t^N = (t-a)^N$, der ja abzuspalten ist. Tatsächlich ist er bisher im Integral verborgen, was man leicht mit Hilfe einer kleinen Substitution sichtbar machen kann. Wir substituieren dazu $s=a+u(t-a)$. Die neue Variable u läuft zwischen 0 und 1, d.h. der Abstand von a nach t wird als neue Koordinateneinheit interpretiert.

(4.3.5) Die Substitution ergibt für den Fehler (wieder mit $t=a+\Delta t$):

$$\boxed{\vec{F}_f^N(a, \Delta t) = \frac{1}{N!} \int_a^t ds (t-s)^N \vec{f}^{(N+1)}(s) = \frac{\Delta t^{N+1}}{N!} \int_0^1 du (1-u)^N \vec{f}^{(N+1)}(a+u\Delta t)}$$

Beachten Sie, dass über die Substitution einerseits die benötigten (N+1) Faktoren Δt erscheinen, andererseits auch die Größe t aus der Grenze herausfällt.

Wann gilt die Formel überhaupt? Inspektion gibt unmittelbar eine hinreichende Bedingung: f muß (N+1)-fach stetig differenzierbar sein - also die letzte dieser Ableitungen muß für $a \leq s \leq t$ noch stetig sein.

(4.3.6) Jetzt können wir uns an das **Problem der Resttermbedingung** machen. Geht der Rest nach Division durch Δt^N noch nach Null wie wir es wünschen? Hierzu benutzen wir (4.28):

$$\|R_f^N(a, \Delta t)\|_W = \frac{1}{\Delta t^N} \|\vec{F}_f^N(a, \Delta t)\|_W \leq \frac{|\Delta t|}{N!} \int_0^1 du (1-u)^N \|\vec{f}^{(N+1)}(a+u\Delta t)\|_W.$$

Geht das nach Null? Man hat den Faktor Δt , der dies tut. Aber das Integral dahinter hängt auch noch von Δt ab und könnte ja im Grenzwert groß werden. Bleibt es beschränkt, dann ist man fertig.

Nun sollte die (N+1)-te Ableitung **stetig** sein. Dann ist sie auf dem kompakten Intervall $0 \leq u \leq 1$ aber auch **gleichmäßig stetig**. Man kann daher die Norm im Integranden für alle betrachteten u durch eine geeignete Konstante M abschätzen.

(4.3.7) Das verbleibende Integral ist elementar und gibt ein (N+1) für den Nenner, so dass schließlich folgt:

$$\boxed{\text{Es sei } \|\vec{f}^{(N+1)}(u)\|_W \leq M \text{ für } a \leq u \leq t. \text{ Dann folgt: } \|R_f^N(a, \Delta t)\|_W \leq \frac{M}{(N+1)!} |\Delta t|.$$

Und das besagt, dass (unter den gegebenen Umständen) die Resttermbedingung tatsächlich erfüllt ist!
 (4.3.8) Fassen wir zusammen:

Gegeben eine Kurve \vec{f} und ein Aufpunkt a . Dann bildet *das Taylorpolynom der Ordnung N von f um a* eine lokale Polynomapproximation der Werte $\vec{f}(a + \Delta t)$. Das Taylorpolynom ist dabei wie folgt definiert:

$$\boxed{T_f^N(a, \Delta t) = \sum_{k=0}^N \frac{\vec{f}^{(k)}(a)}{k!} \Delta t^k = \vec{f}(a) + \vec{f}'(a) \Delta t + \frac{1}{2} \vec{f}''(a) \Delta t^2 + \dots + \frac{1}{N!} \vec{f}^{(N)}(a) \Delta t^N.}$$

Es gilt $f(a + \Delta t) \approx T_f^N(a, \Delta t)$ als Verallgemeinerung der Tangentenapproximation.

Als Verallgemeinerung der Tangententerlegung erhält man *die Taylorzerlegung der Ordnung N von f um a*:

$$\boxed{\vec{f}(a + \Delta t) = T_f^N(a, \Delta t) + F_f^N(a, \Delta t) = \sum_{k=0}^N \frac{\vec{f}^{(k)}(a)}{k!} \Delta t^k + \Delta t^N R_f^N(a, \Delta t).}$$

Dabei erfüllt $\Delta t \mapsto R_f^N(a, \Delta t)$ die Resttermbedingung, entwickelt also eine Nullstelle bei $\Delta t = 0$. Für den Restterm hat man schließlich die oben gegebenen Integraldarstellungen, die eine Abschätzung des Fehlers erlauben.

(4.3.9) Die gefundene Abschätzung ist meist nicht optimal, der mögliche Fehler wird überschätzt, aber man kann sicher sein, dass der gemachte Fehler nicht schlimmer ist.

(4.3.10) Nehmen wir als Beispiel den Fall $N=1$, also die übliche Tangententerlegung.

Angenommen es gilt im betrachteten Intervall $\|\vec{f}''(s)\| \leq M$. Dann folgt für den absoluten Fehler im betrachteten Bereich:

$$\|\vec{f}(a + \Delta t) - \vec{f}(a) - \vec{f}'(a) \Delta t\| \leq \frac{M}{2} \Delta t^2.$$

Beispiel: Wie unterscheiden sich x und $\sin x$ für $0 \leq x \leq 0.5$. Man hat $M=0.46$. Also $|x - \sin(x)| \leq 0.23x^2$. Und das gibt 0.06 für $x=0.5$. Der tatsächliche Fehler für diesen Wert ist etwa 0.02.

Oder im physikalisch kinematischen Modell: Ist der (durch die wirkende Kraft festgelegte) Betrag der Beschleunigung stets kleiner als M , dann bleibt die Abweichung der tatsächlichen Bewegung von der tangentialen (im Startpunkt) stets unterhalb von $\frac{M}{2} \Delta t^2$.

(4.3.11) Im Zusammenhang mit dem Restterm muß man zwei Aussagen sorgfältig auseinanderhalten: Die von uns bereits bewiesene **Restterm-eigenschaft** und die vielfach eigentlich erwünschte **Nullkonvergenz der Fehlerfolge**. Wir stellen beide Eigenschaften einmal nebeneinander:

<i>Restterm-eigenschaft</i>	und	<i>Nullkonvergenz der Fehlerfolge</i>
$\lim_{\Delta t \rightarrow 0} R_f^N(t_0, \Delta t) = 0 \quad N \text{ fest}$		$\lim_{N \rightarrow \infty} \vec{F}_f^N(a, \Delta t) = 0 \quad \Delta t \text{ fest}$

Die Nullkonvergenz der Fehlerfolge **ist noch nicht bewiesen!** Sie gilt auch keineswegs allgemein, muß vielmehr fallspezifisch untersucht werden und zwar abhängig von der Größe von Δt , die in diesem Zusammenhang äußerer Parameter ist.

(4.3.12) Weshalb das so ist, können wir erkennen, wenn wir die Abschätzung des Fehlers nochmals unter dem Gesichtspunkt der N -Abhängigkeit inspizieren. Für die Resttermabschätzung hatten wir N ja fixiert. Die Inspektion zeigt, dass wir die oben eingeführte Abschätzkonstante M besser M_N nennen sollten. Wir benötigen für jede Ableitungsstufe ja eine eigene Abschätzung!

(4.3.13) Damit ergibt unsere alte Abschätzung für die jetzige Situation:

Sei $\|\vec{f}^{(N)}(a + s)\| \leq M_N$ für alle N und s mit $|s| \leq |\Delta t|$.

Dann folgt für den Fehler $\|\vec{F}_f^N(a, \Delta t)\| \leq \frac{M_N \Delta t^N}{(N+1)!}$

Falls jetzt M_N mit N zu stark wächst, muß die rechte Seite keineswegs nach Null gehen, so dass man nicht unbedingt folgern kann, dass der Fehler verschwindet! Und tatsächlich finden sich viele Fälle in denen der

Fehler keineswegs mit N verschwindet! Inspektion der Abschätzung zeigt aber auch: Geht der Fehler für $|\Delta t| = K_0$ nach Null, dann auch für alle Δt mit $|\Delta t| < K_0$. Denn durch Verkleinern von Δt kann sich M_N höchstens verkleinern, nicht aber vergrößern.

(4.3.14) In den Fällen jedoch, in denen man nachweisen kann, dass der Fehler verschwindet, **erhält man eine Reihendarstellung der Funktion bzw. Kurve, die Taylorreihe:**

Geht der Fehler mit N nach Null, dann erhält man die Darstellung von \vec{f} um a mit Hilfe der (zugehörigen) und dann notwendig konvergenten **Taylorreihe**

$$\Delta t \mapsto \vec{f}(a + \Delta t) = \sum_{k=0}^{\infty} \frac{\vec{f}^{(k)}(a)}{k!} \Delta t^k$$

(4.3.15) Beachten Sie bitte, dass Folgendes denkbar ist: Die Taylorreihe ist zwar konvergent, aber nicht gleich dem Funktionswert $f(a+\Delta t)$. Der Grund: Die zugehörige Fehlerfolge geht nicht nach Null sondern hat einen anderen Grenzwert. Derartige Fälle kommen vor, so dass man bei zugehörigen Argumentationen immer große Sorgfalt verwenden sollte.

Als Beispiel verwendet man gern die folgende Funktion: $f(x) = \exp(-\frac{1}{x^2})$ für $x \neq 0$ und $f(0)=0$. Man zeigt leicht, dass für $x=0$ alle Ableitungen existieren und den Wert Null haben, so dass auch die Taylorreihe um Null identisch Null ist! Also muß $f=F$ gelten, die Funktion ist gleich dem N-Grenzwert des Fehlerterms.

In der in (4.1.7) eingeführten Hierarchie von Glattheitseigenschaften $C^0, C^1, \dots, C^\infty$ gibt es demnach jenseits von unendlich noch einen besseren Fall: Man sagt, eine Kurve $J \rightarrow W$ (J offen) sei *analytisch in* $t_0 \in J$, wenn sie in jedem Punkt von J beliebig differenzierbar, also C^∞ ist und in einer kleinen Umgebung von t_0 durch die zugehörige Taylorreihe darstellbar ist. Man bezeichnet diese Eigenschaft mit C^ω . Merke $C^\omega =$ "analytisch" ist noch "glatter" als C^∞ . ($\omega=$ Omega ist der letzte Buchstabe des griechischen Alphabets!)

(4.3.16) Noch ein anderes Vorkommnis ist möglich: Die rechte Seite - also die Taylorreihe - konvergiert für ein Δt , für das $a+\Delta t$ nicht mehr im ursprünglichen Definitionsbereich von f liegt. Das kann situationspezifische Gründe haben, etwa weil der Punkt auß erhalb des Beobachtungsbereichs liegt, es kann aber auch sein, dass der ursprüngliche Rechenausdruck für \vec{f} diesen Wert nicht zuläßt. Dann hat man den interessanten Fall, dass die Taylordarstellung eine Erweiterung der ursprünglichen Abbildung liefert. Das zugehörige Stichwort lautet *analytische Fortsetzung*.

(4.3.17) Oben haben wir bereits gesehen: Gilt die Reihendarstellung für ein Δt , dann sicher auch für alle Δt mit kleinerem Betrag. Verkleinern darf man Δt daher, aber wie weit kann man Δt vergrößern? Bei der Analyse dieser Frage vermischen sich zwei Problemkreise auf unangenehme Weise: Die Konvergenz der Taylorreihe und eventuelle Begrenzungen des Definitionsbereichs der Abbildung \vec{f} .

Die Entmischung gelingt über die folgende wichtige Idee:

6.4.4 Funktionsdefinition mit Hilfe von Reihen

(4.4.1) Anstatt mit f zu starten und dabei ständig auf die komplizierten Zusammenhänge zwischen den drei Größe $f(=$ Abbildung), $T(=$ Taylorreihe) und $F(=$ Fehler) zu stoßen, was leicht in einen argumentativen Sumpf führt, starten wir umgekehrt mit T , also mit Reihen von der Form der Taylorentwicklungen. Wir definieren f über diese Darstellung, sofern die Reihe nur konvergiert! Dann ist das zugehörige F definitionsgemäß Null. Nur: Wann erhält man auf diese Weise eine Kurve und mit welchen Eigenschaften?

(4.4.2) Es zeigt sich, dass dies eine fruchtbare Idee ist, mit deren Hilfe man dem Sumpf der oben angeschnittenen Probleme vielfach entkommt.

Also: *Wann produziert eine Reihe von Funktionen durch Grenzwertbildung erneut eine vernünftige Funktion oder Abbildung?* Man benötigt wie üblich den Wert für jedes x des Urbildes, wird also x als äußeren Parameter ansetzen und fordern, dass $f_n(x)$ in n einen Grenzwert g hat, den man dann als Wert $f(x)$ der neuen Grenzfunktion f interpretiert. Und das für jedes x . Das ist die "punktweise Konvergenz einer Funktionsfolge". Der Nachteil dieses naheliegenden und einfachen Ansatzes ist, dass die so entstehenden Funktionen für die meisten Zwecke zu beliebig sind. Sie haben zu allgemeine Eigenschaften, als dass sie brauchbar wären. Typischerweise: Sind die Funktionen der Folge sämtlich stetig, dann muß die Grenzfunktion dies keineswegs

sein. $f_n(x)=x^n$ auf dem Intervall $[0,1]$ ist hierfür ein Paradebeispiel. Man muß daher mehr Bedingungen hineinstecken, um handhabbare Grenzfunktionen zu erhalten.

(4.4.3) Der folgende Satz (von Weierstraß) liefert eine Art zentralen Einstieg in diesen Problemkreis. Er gibt ein Resultat, von dem aus man sich dann in weitere Richtungen vorarbeiten kann. Zur Supremumsnorm $\|\cdot\|_\infty$ vgl. 6.2.3b.

Satz: Seien V und W Banchräume. $U \subset V$ und $f_n : U \rightarrow W$.
 Für $f:U \rightarrow W$ sei $\|f\|_\infty = \sup_{x \in U} \|f(x)\|_W$
 (*) Die f_n seien beschränkt. D.h. für jedes n gebe es $M_n < \infty$ mit $\|f_n(x)\|_W \leq M_n$ für alle $x \in U$. Daher ist auch $\|f_n\|_\infty < \infty$.
 (**) $\sum_n M_n$ sei konvergent.
Dann ist $\sum f_n$ absolut und gleichmäßig konvergent.
 (***) Ist überdies jedes f_n in U gleichmäßig stetig,
dann ist auch $f = \sum_n f_n$ in U stetig.

(4.4.4) Beweis -1.Teil. Wir interpretieren t als äußeren Parameter und fragen, ob $n \mapsto \|f_n(t)\|_W$ eine Cauchyfolge bildet. Dabei muß die Strategiefunktion $\varepsilon \mapsto N(\varepsilon)$ unabhängig vom äußeren Parameter t wählbar sein, um die Gleichmäßigkeit der Konvergenz zu sichern. Folgende Kette von Abschätzungen ist möglich und ergibt das gewünschte Resultat:

$$\|\sum_{k=n}^m f_k(t)\|_W \leq \sum_{k=n}^m \|f_k(t)\|_W \leq \sum_{k=n}^m \|f_k\|_\infty \leq \sum_{k=n}^m M_k \leq \varepsilon.$$

Der erste Schritt benutzt die Dreiecksungleichung. Der zweite vergrößert jeden Term bis zur zugehörigen Supremumsnorm. Laut Definition ist ja $\|f_k(t)\|_W \leq \|f_k\|_\infty$ für jedes zulässige t . Damit fällt die t -Abhängigkeit heraus. Die Voraussetzungen des Satzes erlauben die nächste Vergrößerung. Das Ergebnis ist ein Restbeitrag der konvergenten Reihe $\sum M_k$. Für diese gibt es daher (konvergent!) eine Strategiefunktion, für die die Summe kleiner als ε wird. Die Ungleichungskette zeigt, dass das dann auch eine Strategiefunktion für die interessierende Reihe ist. Sie ist absolut konvergent und die Abschätzung ist gleichmäßig in t .

(4.4.5) Beweis - 2.Teil: Ist die Grenzfunktion F stetig, wenn die f_n dies sind? Wir müssen eine Strategiefunktion $\delta = \delta(\varepsilon)$ angeben, die die folgende Ungleichung sicherstellt:

$$\|F(t) - F(s)\|_W < \varepsilon \quad \text{sofern nur } |t-s| < \delta \text{ gilt.}$$

Die folgende Rechnung zeigt, wie eine solche Argumentation typischerweise abläuft. Erneut ist die Dreiecksungleichung entscheidendes Hilfsmittel:

$$\begin{aligned} \|F(t) - F(s)\|_W &= \|(F(t) - \sum_{k=0}^N f_k(t)) + \sum_{k=0}^N (f_k(t) - f_k(s)) + \sum_{k=0}^N (f_k(s) - F(s))\|_W \\ &\leq \|F(t) - \sum_{k=0}^N f_k(t)\|_W + \|\sum_{k=0}^N (f_k(t) - f_k(s))\|_W + \|\dots\|_W. \end{aligned}$$

Wegen der gleichmäßigen Konvergenz der Funktionsfolge können wir eine Strategiefunktion $N = N(\varepsilon)$ angeben, so dass der erste Summand und der dritte Summand jeweils kleiner als $\varepsilon/3$ werden. N wählen wir entsprechend groß. Jedes f_k ist stetig. Dann ist auch die **endliche** Summe der f_k aus dem mittleren Term stetig und laut Voraussetzung gleichmäßig stetig. Also kann man eine Strategiefunktion $\delta = \delta(\varepsilon)$ angeben, so dass die mittlere Summe $< \varepsilon/3$ bleibt, sofern nur $|t-s|$ kleiner als δ ist (t fest, $s = t + \eta t$). Insgesamt haben wir damit die gesuchte Ungleichung konstruiert. Der Satz ist bewiesen.

(4.4.6)

Die gleichmäßige Konvergenz stellt also unter den genannten Umständen sicher, dass der Grenzwert stetiger Funktionen erneut eine stetige Funktion ergibt. Man kann das wie üblich als Vorliegen eines **kommutativen Diagramms** interpretieren.

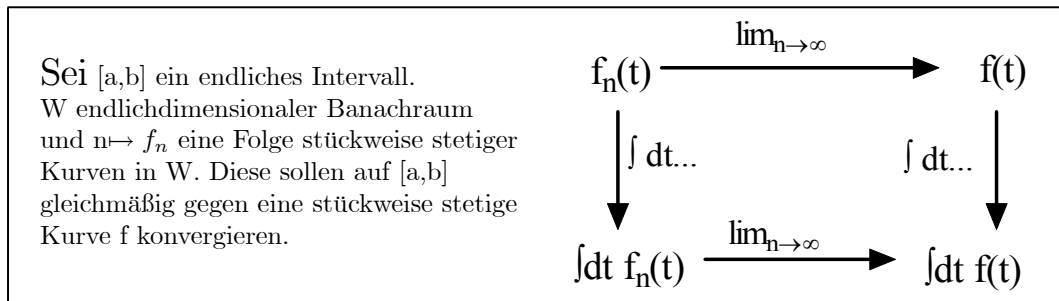
$$\begin{array}{ccc} f_n(x) & \xrightarrow{\lim_{n \rightarrow \infty}} & f(x) \\ \downarrow \lim_{x \rightarrow x_0} & & \downarrow \lim_{x \rightarrow x_0} \\ f_n(x_0) & \longrightarrow & f(x_0) = f(x_0) \\ & & ? \end{array}$$

(4.4.7) Und das ist dann generell die Hauptfrage, die man sich bei Funktionsfolgen zu stellen hat: **Welche Eigenschaften besitzt die Grenzfunktion?** Aber auch umgekehrt: **Wie kann man sicherstellen, dass die Grenzfunktion eine bestimmte, einen interessierende Eigenschaft besitzt?**

(4.4.8) Zwei vorrangige Fragen dieses Typs sind: **Kann man die Grenzfunktion ableiten?** Und wenn ja, was lässt sich über die Ableitung sagen? Und analog: **Kann man die Grenzfunktion integrieren?** Wie lässt sich das Integral berechnen?

Hierzu gibt es zwei Resultate, die üblicherweise angegeben werden und mit denen sich gut arbeiten lässt

(4.4.9) **Satz über die termweise Integration:**



Dann kann man **Integration und Grenzwert vertauschen**. D.h. es gilt:

$$\int_a^b dt \vec{f}(t) = \lim_{n \rightarrow \infty} \int_a^b dt f_n(t)$$

Beachten Sie: Unter den gemachten Annahmen sind alle Integrale bildbar. Die hier benötigte gleichmäßige Konvergenz wird beispielsweise durch den Weierstraß-Satz (4.4.3) geliefert.

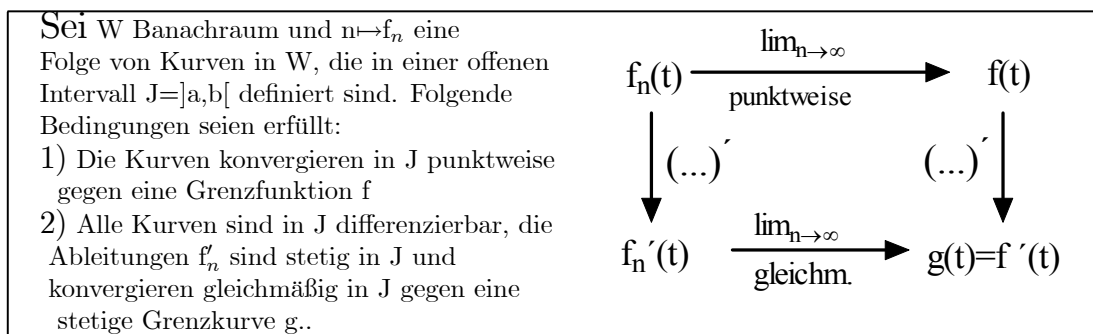
(4.4.10) Der Beweis benutzt die Dreiecksungleichung der Integration:

$$\left\| \int_a^b dt (f(t) - f_n(t)) \right\|_W \leq \int_a^b dt \|f(t) - f_n(t)\|_W \leq (b-a) \frac{\varepsilon}{b-a} = \varepsilon.$$

Wegen der Gleichmäßigkeit der Konvergenz kann man n so groß machen, dass der Integrand im gesamten Integrationsbereich kleiner als $\frac{\varepsilon}{b-a}$ wird. Damit folgt am Ende die gewünschte Ungleichung für die Integrale.

□ Kann man den Beweis so abändern, dass man eine Aussage über das unbestimmte Integral erhält? Erhält man gleichmäßige Konvergenz?

(4.4.11) **Satz über die termweise Ableitung:**



Dann dann gilt $f' = g$ in J .

(4.4.12) Beweis: Die Ableitungen sind stetig, also integrierbar. Der Hauptsatz gibt uns für $t_0 \in]a,b[$ die Gleichung:

$$f_k(t) - f_k(t_0) = \int_{t_0}^t ds f'_k(s) \quad \text{Grenzübergang gibt} \quad f(t) - f(t_0) = \int_{t_0}^t ds g(s)$$

Hierbei haben wir den vorangegangenen Satz über die Vertauschbarkeit von Grenzübergang und Integration benutzt. Wir sind im Integranden zur Grenze g übergegangen. (Die Voraussetzungen des Satzes sind so zusammengestellt, dass dies zulässig ist! Prüfen Sie das). Die rechte Seite besitzt aber nach dem Hauptsatz eine Ableitung (in t). Also ist auch die linke differenzierbar. Ableitung der Gleichung gibt wie behauptet $f'(t) = g(t)$.

6.4.4a Potenzreihen

(4.4.13) Ein weiterer Weg, auf dem man vorankommt, besteht darin, die Form der Funktionsfolgen einzuschränken, zu spezialisieren. Die bisherigen Sätze verlangen noch keinerlei Einschränkung. Was ist, wenn man etwa nur Ausdrücke des Typs zulässt, wie sie oben durch die Taylorpolynome bestimmt werden? Hierdurch gewinnt man sehr viel, wie wir sehen werden. In Kap.10 werden wir weitere wichtige Einschränkungen der Funktionsform kennen lernen.

(4.4.14) Zunächst abstrahieren wir die allgemeine durch die Taylorpolynome erfaßte Form von Funktionsfolgen:

Es sei $a = (\mathbb{N}, n \mapsto \vec{a}_n, W)$ eine Vektorfolge aus W . Für jedes $n \in \mathbb{N}$ bilden wir dazu das Polynom $\vec{s}_N(t) = \vec{a}_0 + \vec{a}_1 t + \dots + \vec{a}_N t^N$. Das ist die N -te *Partialsumme* (der *Potenzreihe*). Es ist $\vec{s}_N = (\mathbb{R}, t \mapsto \vec{s}_N(t), W)$. Die Folge $n \mapsto \vec{s}_n$ nennen wir **die (von der Koeffizientenfolge a erzeugte) Potenzreihe**. Hierfür schreiben man auch gerne kurz $\sum_k \vec{a}_k t^k$

(4.4.15) Je nach Situation kann das Symbol $\sum_k \vec{a}_k t^k$ unterschiedliche Bedeutung haben: Es kann die Prozedur der durch die Folge a festgelegten Grenzwertbildung bedeuten oder das Ergebnis derselben oder auch den speziellen zu gegebenem t gehörigen Wert der Grenzfunktion! Im Idealfall - also bei Konvergenz - ergibt eine Potenzreihe der beschriebenen Art wieder eine Kurve in W . Denken Sie dazu an das Eingangsbeispiel der Kreisapproximation.

(4.4.16) Wie üblich lassen sich auch die Potenzreihen durch kleine Transformationen modifizieren, wobei sich die zugehörigen Eigenschaften in trivialer Weise mittransformieren. Bei den Potenzreihen ist vornehmlich die Verschiebung des Ursprungs relevant. D.h. man ersetzt t durch $(t-t_0)$, was die neue Potenzreihe $\sum_k \vec{a}_k (t-t_0)^k$ ergibt.

(4.4.17) Nochmals zur Orientierung: Ist zunächst eine glatte Kurve f gegeben, dann kann man über Ableitungsbildung die zugehörigen Taylorkoeffizienten bestimmen. Das ergibt eine gewisse vorgegebene Folge a . Zu dieser kann man die zugehörige Potenzreihe bilden, die gleich der Taylorreihe von f ist. Alle allgemeinen Resultate über Potenzreihen gelten natürlich auch für diesen Fall. Alternativ verschafft man sich irgendeine Folge a , bildet dazu die Potenzreihe und fragt nach der eventuellen Grenzkurve f .

(4.4.19) Die wichtigste Eigenschaft der Potenzreihen ist, dass zu jeder von ihnen eine charakteristische Zahl gehört, der sog. *Konvergenzradius der Potenzreihe*, der ihre allgemeinen Eigenschaften weitgehend festlegt. Dieser Konvergenzradius ist eine nicht negative Zahl oder ∞ . Kurz ein Element des Intervalle $[0, \infty]$. Beachten Sie, dass hier die symbolische Größe ∞ zulässiges Element der Menge ist. Als Zuordnung

$$R = (\mathfrak{F}(\mathbb{N}, W), a \mapsto R(a), [0, \infty]).$$

Es liegt also im Sinne von Kap.1.2 eine Abbildung vom Funktionaltyp vor.

(4.4.20) Die wichtigen Eigenschaften von R sind: Ist $|t| < R$, dann konvergiert die Potenzreihe gegen eine Grenzfunktion mit vielen gesicherten Eigenschaften. Ist $|t| > R$, dann divergiert die Reihe. Für $|t| = R$ kann man keine allgemein gültige Aussage machen. Wählt man eine Zahl ρ mit $0 < \rho < R$, dann ist die Konvergenz für $|t| < \rho$ sogar gleichmäßig. (Unser Körper ist \mathbb{R} oder \mathbb{C} . Im Falle von \mathbb{C} definiert $|t| < R$ eine offene Kreisscheibe in der komplexen Ebene!)

(4.4.21) Gehen wir einige mögliche Fälle durch: Für $R=0$ erhält man für $t \neq 0$ keine Konvergenz. Allerdings ist die Potenzreihe für $t=0$ selbst sicher und trivial konvergent. Für alle anderen t ist sie divergent. Ist dagegen $R=\infty$, dann ist die Reihe für **jedes** t konvergent. Aber Sie muß nicht in ganz \mathbb{R} gleichmäßig konvergent sein. Ist $R=1$, so ist sie für $|t| < 1$ konvergent, für $|t| > 1$ divergent. Für $|t| < 0.99$ etwa ist die Konvergenz gleichmäßig.

(4.4.22) Wir **beweisen** die aufgestellten Behauptungen:

Sei r eine reelle Zahl >0 . Wir bilden die Zahlfolge $k \mapsto r^k \|\vec{a}_k\|_W$. Jetzt gibt es zwei Möglichkeiten. Entweder ist diese Folge nach oben beschränkt oder sie ist es nicht. Angenommen sie ist beschränkt, d.h. es gibt M mit $r^k \|\vec{a}_k\|_W \leq M$ für alle k . Dann können wir wie folgt abschätzen:

$$\left\| \sum_{k=n}^m \vec{a}_k t^k \right\|_W \leq \sum_{k=n}^m \|\vec{a}_k\|_W |t|^k = \sum_{k=n}^m (r^k \|\vec{a}_k\|_W) \frac{|t|^k}{r^k} \leq M \sum_{k=n}^m \left(\frac{|t|}{r} \right)^k.$$

Ganz rechts steht aber die Restsumme einer geometrischen Reihe. Und diese konvergiert für $|t| < r$. Wählen wir $|t| \leq qr$ mit festem $q < 1$, dann können wir den Satz (4.4.3) anwenden und erhalten gleichmäßige Konvergenz.

Jetzt bilden wir die Menge B aller r , für die obige Folge beschränkt ist. Das ergibt eine Teilmenge von $[0, \infty[$. Beachten Sie: Ist $r \in B$ und gilt $0 \leq s \leq r$, dann ist auch $s \in B$. Entweder ist $B \subset \mathbb{R}$ nach oben unbeschränkt, dann setzen wir $R = \infty$. Denn wir haben Konvergenz für jedes r . Oder r ist nach oben beschränkt. Dann setzen wir R gleich der kleinsten oberen Schranke für B . Diese Zahl leistet das Behauptete: Konvergenz für alle kleineren Werte und Divergenz für alle darüber. (Denn gäbe es irgendein t oberhalb von R mit Konvergenz der Reihe, so wäre R keine obere Schranke mehr für B).

(4.4.21) Folgerung: **Hat die zu a gehörige Potenzreihe den Konvergenzradius R , dann haben die beiden Reihen, die man durch formale termweise Integration bzw. formale termweise Ableitung erhält, denselben Konvergenzradius.**

Beispiel: $\sum 1t^k$ sei die Ausgangsreihe. Formales termweises Ableiten gibt $\sum kt^{k-1}$. Beachten Sie, dass der wegen t^{-1} kritische Term für $k=0$ fortfällt! Integrieren ergibt $\sum \frac{1}{k+1} t^{k+1}$. Diese Reihe hat keinen konstanten Term mehr. (D.h. $a_0 = 0$).

(4.4.22) Man kann die Sätze (4.4.9) und (4.4.11) anwenden und findet das für die Arbeit mit Potenzreihen außerordentlich wichtige Resultat:

Sei $\sum a_k t^k$ Potenzreihe mit Konvergenzradius $R > 0$. Die in diesem Bereich $|t| < R$ entstehende Funktion werde mit f bezeichnet.

Dann haben die Potenzreihen $\sum k a_k t^{k-1}$ und $\sum \frac{a_k}{k+1} t^{k+1}$ beide denselben Konvergenzradius R . Erstere konvergiert gegen die Ableitung f' und letztere gegen die Stammfunktion $\int_0^t ds f(s)$.

6.4.5 "e-hoch Matrix"

Wir kommen jetzt zu der Anwendung der vorangegangenen Resultate, die wir für die Behandlung der linearen Differentialgleichungen in Kapitel 7 benötigen.

(4.5.1) Sei dazu V endlichdimensionaler Vektorraum und $M: V \rightarrow V$ ein Endomorphismus von V . Mit Hilfe der Ringstruktur von $\text{End}(V)$ können wir zunächst beliebige Potenzen M^n mit $n \in \mathbb{N}$ bilden und daraus wieder Matrixpolynome wie $3M^2 + 4M - 7\text{id}$. Hierbei soll $M^0 = \text{id}_V$ sein. Mit V ist aber auch $\text{End}(V)$ Banachraum, so dass man die Frage nach konvergenten Reihen $\sum_n a_n M^n$ zur Matrix M stellen können. Dies ist ein sinnvolles Problem. Jede Cauchyfolge ist jedenfalls konvergent und der Grenzwert ist erneut ein Endomorphismus. Natürlich kann man auch Matrixdarstellungen dieser Endomorphismen betrachten. Ist $\dim V = n$, so handelt es sich stets um quadratische $n \times n$ -Matrizen.

(4.5.2) Ebenso können wir Kurven mit Wert im Vektorraum $\text{End}(V)$ bilden, insbesondere Potenzreihen. Uns interessiert vornehmlich ein spezieller Fall, nämlich die folgende Reihe:

$$t \mapsto \sum_k \frac{t^k}{k!} M^k = \text{id}_V + tM + \frac{t^2}{2} M^2 + \frac{t^3}{3!} M^3 + \dots$$

Ist diese Reihe konvergent, so ist das Ergebnis der Grenzwertkonstruktion erneut ein Endomorphismus, bei Matrixdarstellung eine $n \times n$ -Matrix.

- Berechnen Sie zum Vertrautwerden für $M = \frac{1}{2} \begin{pmatrix} 1 & 2 \\ -1 & 1 \end{pmatrix}$ die ersten Partialsummen von $\sum M^k$, von $\sum t^k M^k$ und von $\sum \frac{t^k}{k!} M^k$. Verwenden Sie möglichst ein Computeralgebraprogramm.

(4.5.3) Die folgende Abschätzung der N-ten Partialsumme (also $\sum_{k=0}^N$) zeigt die Konvergenz:

$$\left\| \sum_{k=0}^N \frac{t^k}{k!} M^k \right\| \leq \sum_{k=0}^N \frac{|t|^k}{k!} \|M^k\| \leq \sum_{k=0}^N \frac{|t|^k}{k!} \|M\|^k \leq e^{|t|\|M\|}$$

Bei dem entscheidenden mittleren Schritt haben wir eine Norm für die Endomorphismen oder Matrizen benutzt, die die Eigenschaft $\|AB\| \leq \|A\| \|B\|$ Die L^1 -Norm und die Operatornorm leisten dies.

Die letzte Summe rechts ist die Partialsumme der üblichen reellen Exponentialfunktion an der Stelle $x=|t|\|M\|$. Beachten Sie, x ist eine Zahl! Die Reihe für e^x ist absolut konvergent, was die absolute Konvergenz der linksseitigen uns interessierenden Reihe sichert. Gleichmäßige Konvergenz erhält man nur unter der Einschränkung $|t| \leq T$ für ein festes T . Dieses T muß man fest und endlich wählen. Je größer man es wählt, desto größer wird auch die gemeinsame Abschätzungsschranke. Es ist nicht möglich auf ganz \mathbb{R} gleichmäßig abzuschätzen. Da \exp monoton wachsend ist, kann man zur Abschätzung immer den Funktionswert beim maximalen t -Wert T nehmen.

(4.5.4) Nochmals: Die Potenzreihe definiert uns für Variables t eine Kurve in $\text{End}(V)$. Dieser Kurve möchten wir einen prägnanten Namen geben - neben der Summenbezeichnung. Wir wählen $\exp(tM)$ bzw. e^{tM} . Wir werden sehen, dass das eine sinnvolle und nützliche Bezeichnung ist. Ist $\dim V=1$, so sieht man sofort, dass die Reihe mit der Reihe der üblichen Exponentialfunktion identifiziert werden kann. Und für $\dim V > 1$ zeigt sich, dass viele - nicht alle! - der vertrauten Regeln in naheliegender Verallgemeinerung gelten. So sieht man durch Inspektion, dass $\exp(0M) = \text{id}_V$ gilt. Nochmals: $\exp(tM)$ ist zunächst einmal eine zweckmäßige Bezeichnung für den durch die Reihe definierten Endomorphismus. Für gegebenes M ist eine Berechnung mit Hilfe der Reihe meist mühsam. Aber wir wissen, dass das Resultat existiert, auch wenn man nicht gleich über eine zugehörige einfache Rechenformel verfügt. Und mit diesem Wissen werden wir weitere allgemeine Eigenschaften herleiten und später in Kap.14 auch Methoden, die einfachere Berechnungen ermöglichen.

(4.5.5) Mit Hilfe der gegebenen Resultate über Potenzreihen folgt, dass man termweise differenzieren darf. Tut man das, so sieht man, dass man aus der (absolut konvergenten) Reihe ein gemeinsames M ausklammern darf mit dem Ergebnis: Ist $f(t) = \exp(tM)$, dann ist $f'(t) = M \exp(tM) = \exp(tM)M$.

(4.5.6) Eine sehr wichtige Eigenschaft, die wir unten beweisen werden, ist die folgende:

Für $s, t \in \mathbb{R}$ gilt $e^{sM} \circ e^{tM} = e^{(s+t)M}$. Hieraus ergibt sich sofort: **$\exp(tM)$ ist immer invertierbar und die inverse Abbildung ist $\exp(-tM)$.** Invertierbarkeit bedeutet, dass $\exp(tM)$ immer in $\text{Aut}(V)$ liegt, der Automorphismengruppe von V . Man sieht bereits jetzt, wie nützlich zumindest als Gedächtnishilfe die eingeführte Exponentialbezeichnung ist. Wie üblich lassen wir das Zusammensetzungssymbol \circ bei der Endomorphismenmultiplikation nachfolgend meist aus.

(4.5.7) All diese Resultate können wir formal wie folgt zusammenfassen:

Satz: Sei $M \in \text{End}(V)$ und $(\mathbb{R}, +)$ die additive Gruppe der reellen Zahlen.
Dann ist $((\mathbb{R}, +), t \mapsto e^{tM}, \text{Aut}(V))$ ein injektiver Gruppenhomomorphismus.
 Oder auch: Eine Darstellung von $(\mathbb{R}, +)$ in $\text{Aut}(V)$.

Die Bedeutung dieses Satzes wird klar, wenn wir an das Kapitel über Gruppenoperationen zurückdenken. Offenbar erhalten wir über diesen Satz eine Vielzahl von Entwicklungsgruppen (für den kontinuierlichem Konfigurationsraum \mathbb{R}) auf dem Vektorraum V .

(4.5.8) Beweis: Wir müssen die Homomorphierelation $e^{tM} e^{sM} = e^{(t+s)M}$ beweisen. Der Rest folgt daraus trivial. Dies prüfen wir per Tunnelmethode. Da alle beteiligten Reihen absolut konvergent sind, können wir damit rechnen wie mit endlichen Summen. Wir benutzen die Summenschreibweise.

$$e^{(t+s)M} = \sum_k \frac{(t+s)^k}{k!} M^k = \sum_k \frac{1}{k!} \left(\sum_{r=0}^k \binom{k}{r} t^{k-r} s^r \right) M^k = \sum_{k,r} \frac{t^{k-r} s^r}{(k-r)! r!} M^k.$$

Man hat sich zu merken, dass hier r zwischen 0 und k läuft. Genauer wird über die Elemente der folgenden Menge summiert: $J = \{(k,r) \mid k,r \in \mathbb{N}, 0 \leq r \leq k\}$.

Jetzt starten wir von der anderen Tunnelseite und finden

$$e^{tM} e^{sM} = \left(\sum_p \frac{t^p}{p!} M^p \right) \left(\sum_q \frac{s^q}{q!} M^q \right) = \sum_{p,q} \frac{t^p s^q}{p! q!} M^{p+q}.$$

Hierbei wird über $J' = \{(p,q) \mid p,q \in \mathbb{N}\}$ summiert. Setzt man $p=r$ und $k=p+q$, so gehen die Summanden ineinander über. Da überdies $(J', (p,q) \mapsto ((k,r) = (p+q,q), J))$ bijektiv ist, folgt die gewünschte Gleichheit.

(4.5.9) Bisher sieht es so aus, als gäbe es zwischen $\exp(M)$ und der gewöhnlichen Exponentialfunktion von Seite der Rechenregeln her keinerlei Unterschied. Das ist falsch! Seien $M, N \in \text{End}(V)$. Damit können wir $\exp(M)$ und $\exp(N)$ bilden fragen, ob das Produkt $\exp(M)\exp(N)$ gleich $\exp(M+N)$ ist. Wir müssen versuchen, obigen Beweis auf die neue Situation auszudehnen, was heißt, dass wir überall tM durch M und sM durch N ersetzen müssen. Das führt unmittelbar auf große Probleme: Wir stoßen in der ersten Zeile der Rechnung auf den Ausdruck $(M+N)^k$ und dieser soll nach der binomischen Formel ausgewertet werden. Aber das ist nur zulässig, wenn M und N miteinander kommutieren, was in der Regel nicht der Fall ist. So ist $(M+N)^2 = M^2 + MN + NM + N^2$ und das reicht für die weitere Rechnung nicht. Denn die Struktur des Termes auf der andern Tunnelseite $\exp(M)\exp(N)$ verlangt, dass alle Faktoren M links von allen Faktoren N stehen. Also sind die beiden Seiten i.a. nicht gleich! Nur wenn die beiden Endomorphismen (Matrizen) miteinander vertauschen, geht der Beweis durch, gilt die erhoffte Rechenregel.

Ein wichtiges Beispiel: Sei $N = \text{Id}_V$. Dann vertauscht N mit jedem M . D.h. es gilt $e^{s\text{Id}_V} e^{tM} = e^{(s\text{Id}_V + tM)}$. Man rechnet sofort (über die Reihendefinition) nach, dass $e^{s\text{Id}_V} = e^s \text{Id}_V$ gilt. Also folgt $e^{(s\text{Id}_V + tM)} = e^s e^{tM}$.

- Welche Resultate zu e^{tM} sollte man sich merken und wie geht man günstig vor? Fertigen Sie sich eine Übersicht an.
- Sei $\gamma = \sum_n a_n \varphi^n$ eine absolut konvergente Reihe für den Endomorphismus $\varphi : V \rightarrow V$ und $M = M_\varphi$ eine Matrixdarstellung von φ . Wieso ist die Matrixdarstellung M_γ des Grenzwertes gleich dem Grenzwert der Matrixdarstellungen (der Partialsummen der Reihe)?
- Ein Endomorphismus $\varphi : V \rightarrow V$ heißt "nilpotent (von der Ordnung k)", wenn $\varphi^k = 0_V$ gilt, aber $\varphi^{k-1} \neq 0_V$. Bestimmen Sie einige einfache Beispiele nilpotenter 2×2 - und 3×3 Matrizen. Welche Vereinfachung ergibt sich für nilpotentes φ für $e^{t\varphi}$ gegenüber dem allgemeinen Fall?

6.4.5a Durch e^{tM} erzeugte Entwicklungsgruppen

(4.5.10) Zu jeder Matrix M gehört die Operationsabbildung $(\mathbb{R}, t \mapsto e^{tM}, \text{Aut}(V))$, die als Bild in $\text{Aut}(V)$ eine zu \mathbb{R} isomorphe Untergruppe aussondert, welche wiederum einen Entwicklungsvorgang in V beschreibt. Dieser Entwicklungsvorgang ist einparametrig, ist so etwas wie die zeitliche Bewegung eines Massenpunktes im Konfigurationsraum.

(4.5.11) Im Gegensatz zu den in 4.4.5c beschriebenen diskreten Entwicklungsprozessen liegt hier ein kontinuierlicher Prozess vor. Derartige Entwicklungsprozesse werden wir genauer in den Kapiteln 6 und 7 betrachten. Die Bahnen sind Kurven in V , wobei zu jedem Bildpunkt ein oder eventuell mehrere Parameterwerte gehören.

(4.5.12) Beachten Sie, dass die Abbildung $(\text{Aut}(V), \lambda \mapsto e^\lambda, \text{Aut}(V))$ injektiv ist. D.h. liegt φ im Bild dieser Abbildung, so läßt sich eindeutig ein zugehöriges λ bestimmen mit $\varphi = e^\lambda$. Formal können wir auch $\lambda = \ln(\varphi)$ schreiben. Es zeigt sich, dass λ meist einfacher zu handhaben ist als φ . In Kapitel 6 werden wir zeigen, wie man e^λ in vielen Fällen ausrechnen kann. In Kap.12 wird dieses Problem dann allgemein behandelt.

- Wie sieht die Stabilisatoruntergruppe von \mathbb{R} aus, die zu einer periodischen Bahn führt?

6.4.6 Weg- und Arbeitsintegrale

Wir greifen das Thema in Kap.1.2.13a auf und wenden unsere jetzigen Kenntnisse über die Kurvenintegration darauf an.

(4.6.1) Es sei $G \subset V_0^3$ unser Beobachtungsgebiet, also unser Konfigurationsraum. G sei offene Teilmenge.

In G sei ein glattes Feld \vec{F} gegeben. $\vec{F} = (G, \vec{x} \mapsto \vec{F}(\vec{x}), V^3)$. Physikalisch stellen wir uns wie immer vor, dass wir die Feldstärke $\vec{F}(\vec{x})$ am Orte \vec{x} lokal beobachten und messen können.

Weiter seien Bahnkurven $\vec{r}: J \mapsto G$ gegeben, die Bewegungsabläufe in G beschreiben. Etwa Expeditionen, auf denen die jeweilige Feldstärke vermessen wird. Dann ist $\vec{F} \circ \vec{r} : J \rightarrow V^3$ **die Kurve der entlang \vec{r} beobachteten Feldstärke!** $\vec{F}(\vec{r}(t))$ ist der zur Zeit t am Orte $\vec{r}(t)$ vorgefundene Feldstärkewert. $J \subset \mathbb{R}$ ist unser Beobachtungszeitintervall.

Denkbar sind sehr viele Bahnkurven. Sicherlich sollten sie stetig sein. Wir machen hier aber noch stärkere technische Annahmen. Der Einfachheit setzen wir alle Bahnkurven als beliebig oft differenzierbar voraus. Auch dürfen sie unterschiedliche Beobachtungsintervalle haben. Allgemeiner kann man sie stetig und stückweise glatt annehmen. Derartige Bahnkurvenmengen werden häufig benötigt.

(4.6.2) Sei $G \subset V_0^3$ offen und

$$\mathfrak{B} = \{\vec{r}'|\vec{r} : J \rightarrow G, J \subset \mathbb{R} \text{ echtes Intervall, } \vec{r} \text{ beliebig oft differenzierbar}\}$$

Das ist unser Raum der betrachteten denkbaren Bahnkurven. Beachten Sie: Da J in der Mengenklammer steht, sind unterschiedliche J zugelassen. Unter "echt" verstehen wir: J soll weder leer noch einelementig sein. Im Gegensatz dazu ist G fest, ein äußerer Parameter der Menge. Da in \mathfrak{B} Abbildungen mit unterschiedlichem Urbildbereich auftreten, hat \mathfrak{B} keine unserer üblichen algebraischen Strukturen. Es ist für uns eine reine Menge.

(4.6.3) Ist \vec{F} - das auf G definierte Feld - ein Kraftfeld, so enthält \mathfrak{B} eine für die Physik sehr. wichtige Teilmenge: Bringt man nämlich Massenpunkte nach G , auf die nur die von \vec{F} ausgehende Kraft wirkt, so entstehen ganz bestimmte Bahnkurven, die wir *bezüglich \vec{F} physikalische Bahnkurven* genannt haben (Kap.1.2.13). Ist \vec{F} etwa konstant, so entstehen nur zugehörige Flugparabeln, nie andere doch durchaus denkbare Bahnen.

Die Menge aller bezüglich \vec{F} physikalischen Bahnkurven bezeichnen wir mit \mathfrak{B}_p . Also :

$$\mathfrak{B}_p = \mathfrak{B}_p(\vec{F}) = \{\vec{r}'|\vec{r} \in \mathfrak{B}, \vec{r} \text{ ist physikalisch bezüglich } \vec{F}\}.$$

Die Newtonsche Bewegungsgleichung erlaubt eine genaue Bestimmung dieser zunächst rein begrifflich festgelegten Menge. $\mathfrak{B}_p(\vec{F})$ besteht gerade aus allen Lösungen der Differentialgleichung $m \ddot{\vec{r}}(t) = \vec{F}(\vec{r}(t))$.

□ Wie sieht \mathfrak{B}_p aus, wenn \vec{F} das Nullfeld ist?

(4.6.4) Wie steht es mit den Bahnkurven, die nicht in \mathfrak{B}_p liegen? Auch sie können als Bewegungen vorkommen, nur muß ihre Bewegung noch andere Ursachen als die Kraft \vec{F} haben. Diese anderen Ursachen spezifiziert man meist nicht weiter. Wir sagen dann auch, dass eine entlang \vec{r} geführte Bewegung vorliege. Ursache einer solchen Bewegung ist ein zusätzlicher Führungseinfluß, der meist auch als Kraft beschreibbar ist, während bei einer physikalischen Bewegung das Kraftfeld \vec{F} allein die physikalische Bewegung bewirkt (Kap. 1.2.13a).

(4.6.5) Unter den skizzierten Umständen sind zwei besonders wichtige Integrale bildbar. Beide benötigen als Zutaten nur eine Kurve $\vec{r} \in \mathfrak{B}$ mit Intervall J mit Endpunkten a und b sowie das Kraftfeld \vec{F} .

- Das **vektorielle** Integral $\int_a^b dt \vec{F}(\vec{r}(t))$. Das Ergebnis ist ein Vektor. Teilt man diesen Vektor durch die Länge des Zeitintervalles, so erhält man den **Mittelwert der Feldstärke während des Beobachtungszeitraumes**.

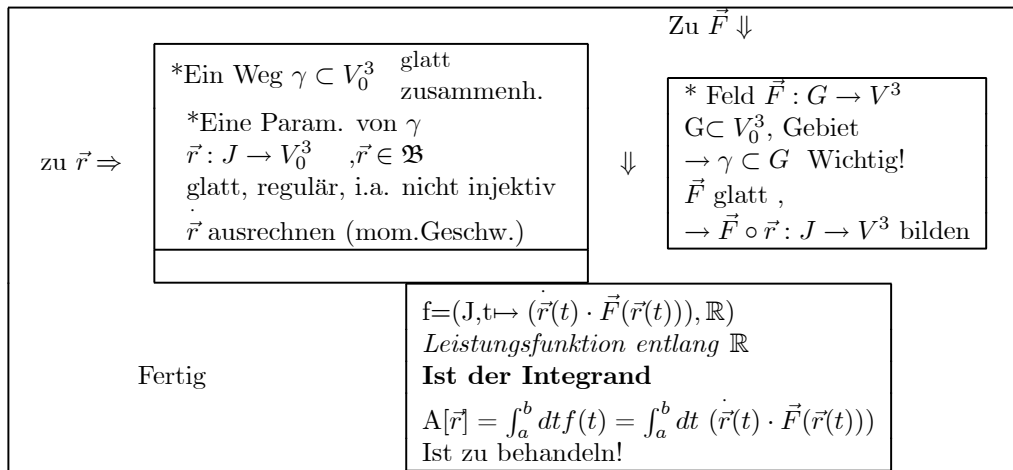
- Das zahlwertige Integral $\int_a^b dt \vec{r}'(t) \cdot \vec{F}(\vec{r}(t))$ vom Arbeitstyp. Es hat eine Reihe wichtiger Eigenschaften, die gleich besprochen werden sollen. Unter anderem hängt es weitgehend nur vom Weg $\gamma = \text{Bild } \vec{r}$ ab, nicht aber von der gewählten Parametrisierung. Es wird daher auch als Wegintegral $\int_\gamma d\vec{r} \cdot \vec{F}$ bezeichnet.

(Merke: $d\vec{r} = dt \frac{d\vec{r}}{dt}$)

Physikalisch ergibt dies Integral die Arbeit, die von einer Führungskraft geleistet werden muß, welche den Massenpunkt gegen den Einfluß von \vec{F} auf der Bahnkurve \vec{r} führt. Wir nennen diesen Integraltyp daher "Arbeitsintegral".

- (4.6.6) **Wie berechnet man ein Arbeitsintegral?** Das bereitet nicht selten Schwierigkeiten. obwohl solche bei genauer Inspektion der definierenden Formel nicht auftreten sollten. Wir geben hier noch einmal die Zutaten - also \vec{r} und \vec{F} - und zeigen, wie daraus das Integral entsteht und welche

Zusammenhänge zu beachten sind.



Da der Integrand eine gewöhnliche Funktion ist, kann die Integral wie üblich berechnet werden.

Alles soeben zum Arbeitsintegral Gesagte ist nur eine geringfügige Ergänzung zu 1.2.13a. Jetzt sollen aber einige besonders auch **physikalisch wichtige mathematische Resultate** für diesen Integraltyp hergeleitet werden.

(4.6.7) **Erste allgemeine Eigenschaft des Arbeitsintegrals :**

Invarianz gegen Umparametrisierungen.

Wir nehmen eine Umparametrisierung des Weges vor. Dies erfolgt durch die Vorgabe einer Abbildung $t = (I, s \mapsto t(s), J)$, die den alten Parameter t als Funktion des neuen s ausdrückt. (Machen Sie sich klar, dass es sinnvoll ist, diese Funktion vorzugeben! Diese Funktion t soll glatt und streng monoton wachsend sein. Daraus folgt, dass stets $t'(s) > 0$ gilt. Weiter seien $J = [a, b]$ das alte und $I = [A, B]$ das neue Intervall. Speziell folgt dann $t(a) = A$ und $t(b) = B$. Jetzt benutzen wir die **Substitutionsregel** der Integration und finden:

$$\int_a^b dt (\dot{\vec{r}}(t) \cdot \vec{F}(\vec{r}(t))) = \int_A^B ds t'(s) \left(\dot{\vec{r}}(t(s)) \cdot \vec{F}(\vec{r}(t(s))) \right) = \int_A^B ds \left(\dot{\vec{x}}(s) \cdot \vec{F}(\vec{x}(s)) \right)$$

Dabei ist $\vec{x}(s) = \vec{r}(t(s))$ gesetzt. Das ist die **neue Parametrisierung des Weges**. Also $\text{Bild } \vec{r} = \text{Bild } \vec{x}$. Es folgt $\dot{\vec{x}}(s) = t'(s) \dot{\vec{r}}(t(s))$, was in der letzten Umformung benutzt wurde.

(4.6.8) Wir sehen: **Der Wert des Integrales hängt unter den beschriebenen Umständen nicht von der benutzten Parametrisierung ab**, diese wird nur für den Rechnung benötigt.

Dabei darf man allerdings nicht auf die angegebenen Voraussetzungen an die Parametrisierung verzichten.

Zur Rechnung selbst: Der Faktor $t'(s)$ erscheint links über die Substitutionsregel durch $dt = ds t'(s)$. Im zweiten Schritt verschwindet er wieder über die Kettenregel für $\dot{\vec{x}}(s)$! Das ist ein wichtiges Element der angegebenen Umformung.

(4.6.8) Das Resultat rechtfertigt weitgehend den Sprachgebrauch, derartige Arbeitsintegrale *Wegintegrale* zu nennen und mit Symbolen wie $\int_{\gamma} d\vec{x} \vec{F}(\vec{x})$ zu bezeichnen. Ganz korrekt ist das aber nur mit gewissen Einschränkungen: Der Weg sollte orientiert sein, d.h Anfangs- und Endpunkt sollten festliegen **und die Parametrisierung sollte injektiv sein**.

Nehmen wir an, γ wäre ein Kreis. Dann macht es in der Regel für das Integral einen Unterschied, ob dieser Kreis einmal oder zweimal durchlaufen wird. Eine Umparametrisierung obiger Art, bei der ein einfacher Umlauf in einen zweifachen umgewandelt wird, kann es nicht geben, der Gleichheitsbeweis ist nicht führbar und das Ergebnis i.a. auch verschieden. Der Weg γ ist jedoch für beide Parametrisierungen derselbe, nämlich der Kreis.

(4.6.9) **Zusammenfassung:** Ein Arbeitsintegral über eine injektive Parametrisierung ändert bei orientierungserhaltender Umparametrisierung seinen Wert nicht und wird daher gerne als Wegintegral bezeichnet.

(4.6.11) **Zweite Eigenschaft des Arbeitsintegrals:**

Wegunabhängigkeit für konservative Felder .

Neben der Invarianz gegen Umparametrisierung gibt es eine zweite noch stärkere Eigenschaft, die aber nur für gewisse Felder gilt.

(4.6.12) Das Kraftfeld F heißt konservativ, wenn man es im betrachteten Gebiet als Gradient eines glatten Skalarfeldes schreiben kann. Wenn es also ein Skalarfeld V gibt, für das $\vec{F}(\vec{x}) = -\text{grad}V(\vec{x})$ gilt. Das negative Zeichen ist eine aus physikalischen Gründen eingeführte Konvention, um den üblichen Zusammenhang mit dem Energiesatz zu sichern. Unter diesen Umständen gilt die nachfolgende Rechnung:

Sei $\vec{F}(\vec{x}) = -\text{grad}V(\vec{x})$ ein **konservatives Kraftfeld in (ganz) G**

Dann ist nach der Kettenregel ($t \mapsto \vec{r}(t) \mapsto V(\vec{r}(t))$):

$$f(t) = -(\dot{\vec{r}}(t) \cdot \text{grad}V(\vec{r}(t))) = -\frac{d}{dt}V(\vec{r}(t))$$

Oder über die totale Ableitung gerechnet:

$$f(t) = -DV(\vec{r}(t)) \cdot \dot{\vec{r}}(t) = -DV(\vec{r}(t)) \cdot (D\vec{r}(t) \cdot 1) = -D(V \circ \vec{r})(t) \cdot 1 = -\frac{d}{dt}V(\vec{r}(t)).$$

Einsetzen ins Arbeitsintegral gibt

$$A[\vec{r}] = \int_a^b dt f(t) = - \int_a^b dt \frac{d}{dt}V(\vec{r}(t)) = V(\vec{r}(a)) - V(\vec{r}(b))$$

Der Integralwert ist gleich der Differenz der Potentialwerte. Das Integral ist mithin "wegunabhängig". Nur Anfangs- und Endpunkt von $\gamma = \text{Bild}\vec{r}$ werden benötigt und gehen ein. (4.6.13) Der Wert des Arbeitsintegrals hängt jetzt also nur noch von Anfangs- und Endpunkt des (ganz in G verlaufenden) Weges ab, wie man es üblicherweise vom physikalischen Arbeitsbegriff her gewöhnt ist. Die Bedingung, dass der Weg ganz in G liegen muß und dass G dort glatt sein muß, ist wichtig. Andernfalls muß die Wegunabhängigkeit des Arbeitsintegrals nicht gelten. Aber innerhalb von G darf man sich weitgehend beliebig (etwa stetig und stückweise glatt) vom Startpunkt zum Endpunkt bewegen.

Oder physikalisch: Man kann einen langen Weg unter geringem Krafteinfluß gehen oder einen kurzen unter großem. Die zu erbringende Arbeit bleibt dieselbe.

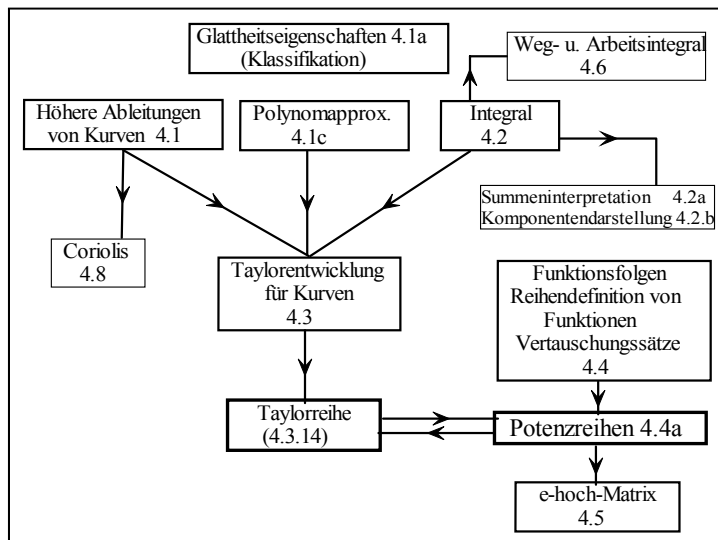
(4.6.14) In der in Kap.6.(3.3.19) besprochenen Herleitung des Energiesatzes haben wir unter derselben Voraussetzung wie eben den Energiesatz hergeleitet. Dabei erhielt V die Rolle der potentiellen Energie. D.h. wir können zusammenfassend sagen:

- Im Falle eines konservativen Kraftfeldes gilt einerseits der Energiesatz und andererseits ist das Arbeitsintegral im beschriebenen Sinne wegunabhängig.
- Stimmen insbesondere Anfangs- und Endpunkt überein, so muß das zugehörige Integral Null ergeben. Zu einem geschlossenen Weg gehört dann keine Arbeitsleistung.
- Hat man umgekehrt ein glattes Kraftfeld vorliegen, für das Arbeitsintegral nicht wegunabhängig ist, dann kann dieses nicht konservativ sein.

(4.6.15) Als Beispiel betrachten wir $\vec{F}(\vec{x}) = (\vec{a}\vec{x})\vec{x}$. Für $\vec{a} \neq \vec{0}$ ergeben sich **wegabhängige** Arbeitsintegrale wie man sofort überprüft. (Beachten Sie, dass man in diesem radial gerichteten Feld auf Kreisen um den Ursprung keine Arbeit leistet). Mit einem derartigen Kraftfeld könnte man ein Perpetuum mobile konstruieren: Laufen Sie etwa vom Ursprung ein Stück in Richtung \vec{a} , dann auf einem Viertelkreis zum Äquator und von dort zurück zum Ursprung. Jeder derartige Umlauf liefert Ihnen Energie.

6.4.7 Übersicht

Das nachfolgende Diagramm stellt wichtige Zusammenhänge des Teilkapitels 6.4 dar. Ausgelassen sind die allgemeinen Resultate wie der Satz vom beschränkten Zuwachs und die Konvergenzsätze. Es ist zu erkennen, dass eine Reihe neuer Ideen eingeführt wurde.



Θ

6.4.8 Die Corioliskraft

Die Corioliskraft liefert eine instruktive Anwendung für den entwickelten Formalismus.

(4.8.1) Wir betrachten einen starren Körper, der mit der Winkelgeschwindigkeit $t \mapsto \vec{\omega}(t)$ um eine feste Achs \vec{n} rotiert. Also $\vec{\omega}(t) = \omega(t)\vec{n}$ und $|\vec{n}| = 1$. An einem Punkt P dieses Körper sei eine Basis $e = (\vec{e}_1, \vec{e}_2, \vec{e}_3)$ angeftet. D.h. diese Vektoren sind fest mit dem Körper verbunden. **Ein vom Körper mitbewegter Beobachter sieht dann den Koordinatenvektor bezüglich e!** Das ist wichtig.

Weiter sei V der ruhende Konfigurationsraum, für den die üblichen Gesetze der Physik gelten, insbesondere also die Newtonsche Bewegungsgleichung. Der Ursprung O des mitbewegten Systems habe die Bahnkurve $\vec{R} = (\mathbb{R}, t \mapsto \vec{R}(t), V_0^3)$.

(4.8.2) Für jeden Punkt Q des starren Körpers mit Bahnkurve $\vec{s} = (\mathbb{R}, t \mapsto \vec{s}(t), V_0^3)$ gilt die in 6.1.1b besprochene Beziehung für die vektorielle Geschwindigkeit, also $\dot{\vec{s}}(t) = \vec{\omega}(t) \times \vec{s}(t)$.

(4.8.3) Ein Massenpunkt, der **nicht** mit dem Körper verbunden sein muß, habe die Bahnkurve $\vec{r} = (\mathbb{R}, t \mapsto \vec{r}(t), V_0^3)$. Diesen Vektor zerlegen wir bezüglich des mitbewegten Systems. Also

$$\vec{r}(t) = \vec{R}(t) + \vec{x}(t) = \vec{R}(t) + \sum \vec{e}_i(t) x_i(t).$$

Der Koordinatenvektor des (sich auf dem starren Körper befindenden) Beobachters ist als Tupel $\vec{x}^K(t) = (x_j(t)) = \sum \vec{E}_j x_j(t)$. Dabei sei E die kanonische Basis des \mathbb{R}_K^3 . Dagegen gibt $\vec{x}^B(t) = \sum \vec{e}_j x_j(t)$ den geometrischen Pfeil, den der mitbewegte Beobachter als Bahnkurve wahrnimmt. B für "bewegt". \vec{r} dagegen ist die Bahnkurve, wie sie vom Inertialsystem V aus wahrgenommen wird.

(4.8.4) Das Newtonsche Bewegungsgesetz gilt absolut im Inertialsystem V in der Form $m\ddot{\vec{r}}(t) = \vec{F}(\vec{r}(t))$. Dabei ist $\vec{x} \mapsto \vec{F}(\vec{x})$ das wirkende äußere Kraftfeld, das (bei kräftefreier Bewegung) auch Null sein kann. Die Ableitungen sind also immer relativ zu diesem System auszuführen.

(4.8.5) Uns interessiert $\ddot{\vec{x}}^B(t) = \sum \ddot{\vec{e}}_i(t) x_i(t)$. Das ist die Beschleunigung, die der mitbewegte Beobachter wahrnimmt (=relative Beschleunigung des Massenpunktes gegenüber dem körperfesten Bezugssystem e, nicht gegenüber dem Inertialsystem!) und die er über Newton auf Kräfte oder Scheinkräfte als Ursachen zurückführt! Die erste Ableitung von \vec{r} , der physikbezogenen Größe, gibt:

$$\dot{\vec{r}}(t) = \dot{\vec{R}}(t) + \sum \dot{\vec{e}}_j(t) x_j(t) + \sum \vec{e}_j(t) \dot{x}_j(t) \quad \dot{\vec{e}}_j(t) = \vec{\omega}(t) \times \vec{e}_j(t) \quad \text{nach (4.7.2)}$$

Die zweite dieser Gleichungen gilt ja für **jeden** mit dem Körper mitbewegten Punkt, also insbesondere für die drei Endpunkte der Basisvektoren und damit für diese. Einsetzen und erneute Ableitung gibt:

$$\begin{aligned} \ddot{\vec{r}}(t) &= \ddot{\vec{R}}(t) + \Sigma(\dot{\vec{\omega}}(t) \times \vec{e}_j(t) + \vec{\omega}(t) \times \dot{\vec{e}}_j(t))x_j(t) + \Sigma(\vec{\omega}(t) \times \vec{e}_j(t))\dot{x}_j(t) \\ &\quad + \Sigma(\dot{\vec{\omega}}(t) \times \vec{e}_j(t))\dot{x}_j(t) + \Sigma\dot{\vec{e}}_j(t)\ddot{x}_j(t) \end{aligned}$$

(4.8.6) Oder, wenn wir zusammenfassen:

$$\ddot{\vec{r}}(t) = \ddot{\vec{R}}(t) + \dot{\vec{\omega}}(t) \times \vec{x}^B(t) + \vec{\omega}(t) \times (\vec{\omega}(t) \times \vec{x}^B(t)) + 2\vec{\omega}(t) \times \dot{\vec{x}}^B(t) + \ddot{\vec{x}}^B(t)$$

(4.8.7) Die mit dem oberen Index B versehenen Vektoren sind die Vektorpfeile, die vom mitbewegten System aus beobachtet werden: $\vec{x}^B(t) = \Sigma \vec{e}_j(t) x_j(t)$. Für den mitbewegten Beobachter erscheinen die \vec{e}_j natürlich zeitunabhängig! Demnach ist $\dot{\vec{x}}^B(t) = \Sigma \dot{\vec{e}}_j(t) x_j(t)$ die Relativgeschwindigkeit des Punktes relativ zum System e. Schließlich gibt der letzte Term $\ddot{\vec{x}}^B(t)$ die Beschleunigung des Punktes relativ zu e!

Nach Einsetzen in die Bewegungsgleichung sind die übrigen Terme der rechten Seite als Scheinkräfte zu interpretieren, die als Ursachen für Relativbewegungen zum körperfesten System anzusehen sind. Da die Kräfte auf der anderen Seite der Bewegungsgleichung stehen, ist noch das Vorzeichen zu wechseln, was man durch Vertauschen der Faktoren im Kreuzprodukt auffängt. Und natürlich ist noch mit der Masse m des beobachteten Massenpunktes zu multiplizieren.

(4.8.8) Wir besprechen kurz die so entstehenden Beiträge. Der vorletzte Term ergibt die Corioliskraft. Man beachte den Faktor 2 und seine rechnerische Herkunft. Der Term hat die Form eines durch die Relativgeschwindigkeit zum bewegten Bezugssystem festgelegten Kraftfeldes. Es beschreibt, die Änderungen der Koordinaten für e, die durch die Bewegung von e relativ zu V verursacht werden.

$$\begin{aligned} \vec{F}_{Coriolis}^B(\vec{v}^B, t) &= 2m\vec{v}^B \times \vec{\omega}(t) \\ \text{wobei } \vec{v}^B &= \Sigma \dot{\vec{e}}_j(t) x_j(t) \text{ Relativgeschw. des Massenpunktes bez. e} \end{aligned}$$

Die Corioliskraft ist eine Scheinkraft. Die Corioliskraft hat besonders im Bereich der Meteorologie erstaunliche Konsequenzen bzw. eine große Erklärungsfähigkeit.

(4.8.9) Der Beitrag $\ddot{\vec{R}}$ beschreibt die Beschleunigung des bewegten Ursprungs, ist also vom Zentrifugalkrafttyp. Da der Ursprung von e am starren Körper fixiert ist, hätten wir diesen Beitrag alternativ auch wie folgt ausrechnen können:

$$\begin{aligned} \dot{\vec{R}}(t) &= \vec{\omega}(t) \times \vec{R}(t). \quad \text{Also} \\ \ddot{\vec{R}}(t) &= \dot{\vec{\omega}}(t) \times \vec{R}(t) + \vec{\omega}(t) \times \dot{\vec{R}}(t) = \dot{\vec{\omega}}(t) \times \vec{R}(t) + \vec{\omega}(t) \times (\vec{\omega}(t) \times \vec{R}(t)) \end{aligned}$$

Wir können diesen Beitrag daher den beiden anderen noch verbliebenen Termen zuschlagen. Nur für die Corioliskraft ergibt sich ein entsprechender Beitrag nicht! Wir erinnern uns, dass $\vec{r}(t) = \vec{R}(t) + \vec{x}^B(t)$ gilt. Dann ergeben die vier restlichen Terme die folgenden beiden Scheinkräfte:

(4.7.10)	$m(\dot{\vec{R}}(t) + \dot{\vec{x}}^B(t)) \times \dot{\vec{\omega}}(t) = m\dot{\vec{r}}(t) \times \dot{\vec{\omega}}(t)$	Entfällt bei konstanter Winkelgeschwindigkeit
(4.7.11)	$-m\dot{\vec{\omega}}(t) \times (\vec{\omega}(t) \times (\vec{R}(t) + \vec{x}^B(t)))$ $= -m\dot{\vec{\omega}}(t) \times (\vec{\omega}(t) \times \vec{r}(t))$	Zentrifugalkraft

Die inhaltliche Interpretation ist offensichtlich. Beide Terme haben die Form von Kraftfeldern. Dabei ist immer zu beachten, dass jeder Term als "Kraft" in der Gleichung

$$m\ddot{\vec{x}}^B(t) = \vec{F}_{eff}(\vec{x}^B(t), \dot{\vec{x}}^B(t), t)$$

zu interpretieren ist. **Die von e aus beobachteten Bewegungen müssen diese Gleichung erfüllen.** In dieser resultierenden oder effektiven Kraft \vec{F}_{eff} (=Ursache der beobachteten relativen Beschleunigung)

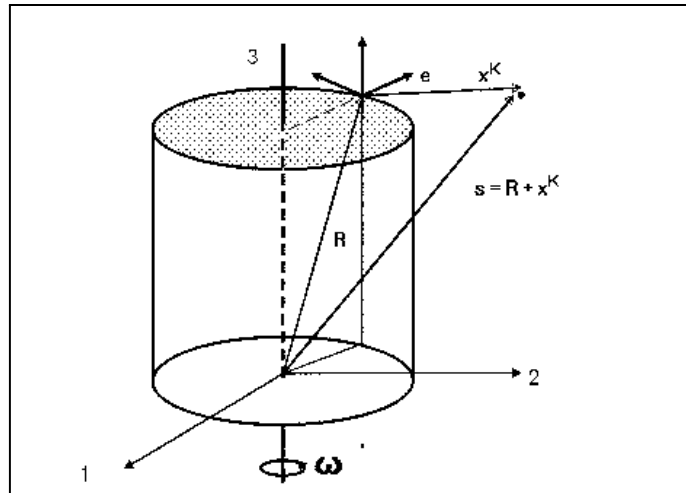
ist die absolute physikalische Kraft nur ein Anteil. Und $\vec{r} = \vec{R} + \vec{x}^B$ ist zum Merken der Formeln günstige Hilfsgröße, da $t \mapsto \vec{R}(t)$ in dieser Differentialgleichung für \vec{x}^B ja als äußerer Parameter anzusehen ist

(4.8.9) Als zusammengefaßte Formel ergibt sich die folgende Bewegungsgleichung für die Festlegung der Bewegung relativ zum mitbewegten System e:

$$m\ddot{\vec{x}}^B(t) = \vec{F}(\vec{x}(t), t) + 2m\dot{\vec{x}}^B(t) \times \vec{\omega}(t) + m\vec{\omega}(t)(\vec{r}(t) \times \vec{\omega}(t)) + m\vec{r}(t) \times \dot{\vec{\omega}}(t)$$

mit $\vec{r}(t) = \vec{R}(t) + \vec{x}^B(t)$.

(4.9.10) Die Figur fasst die Verhältnisse zusammen. Überlegen Sie sich selbst die Richtungen der drei Scheinkräfte am gezeichneten Punkt!



- Die Drehachse der Erde sei die Polachse eines Polarkoordinatensystems. Der Nordpol liege auf der positiven 3-Achse. Welches Vorzeichen hat dann ω aus $\vec{\omega} = \vec{e}_3\omega$ in folge der Erdrotation? Zerlegen Sie $\vec{\omega}$ im Ursprung des mitbewegten Systems in eine radiale und eine tangentielle Komponente. Wieso betrachtet man im Zusammenhang mit der Erdrotation üblicherweise immer die radiale Komponente der Corioliskraft? Welche tangentielle Scheinkraft entsteht?. Verwenden Sie das Polarkoordinatenbasisfeld