Mathematik II, SS2000, Protokolle

1.Woche, 10.4.-13.4.2000

LIN. ALGEBRA: Bitte ziehen Sie auch die Literatur zu Rate: Fischer, Koecher, Klingenberg, Jänich.
Am Beispiel "Gruppe" wurde noch einmal betont, daß es sich in der Lin. Algebra meist um strukturelle Begriffe handelt; diese Begriffe bezeichnen nicht individuelle Objekte wie z.B. $\mathbb{R}$. Gruppenhomomorphismen kommen ab jetzt vor, z.B. die Determinante als Homomorphismus von der Gruppe der invertierbaren Endomorphismen in die multiplikative Gruppe K^* des K-Vektorraums V. Vorher benötigen wir die Gruppe $\mathcal{S}_n$ aller Permutationen von n Elementen und den Homomorphismus $\text{sign} : \mathcal{S}_n \to \{+1, -1\}$.
Beschreibung von Permutationen, die gut für Komposition geeignet ist:

$$\sigma \in \mathcal{S}, \quad \sigma : \ \{1, \dots, n\} \to \{1, \dots, n\}, \ \begin{pmatrix} 1 & 2 & & n \\ \sigma(1), & \sigma(2), & \dots, & \sigma(n) \end{pmatrix}.$$

Mit der Zykeldarstellung werden individuelle Eigenschaften einer Permutation deutlicher, Beispiel:

$$(1\ 7\ 2\ 5)\ (3\ 4)\ (6), \text{ bezeichnet } \begin{pmatrix} 1\,2\,3\,4\,5\,6\,7 \\ 7\,5\,4\,3\,1\,6\,2 \end{pmatrix}.$$

Eine Permutation, die genau zwei Elemente vertauscht, heißt *Transposition*, Beispiel: (1 2) oder $(n_1\ n_2)$ – Einerzykeln, (6), werden oft nicht hingeschrieben. Da man Zykeln leicht als Komposition von Transpositionen schreiben kann, zeigt die Zykelbeschreibung: *Jede Permutation ist Komposition (auch: Produkt) von Transpositionen.*
Satz. Für jede Permutation $\sigma \in \mathcal{S}_n$ kann man ein Vorzeichen $\text{sign}\,(\sigma)$ definieren: Sei a die *Anzahl* der Transpositionsfaktoren in einer Darstellung von σ als Produkt von Transpositionen, dann gilt unabhängig von der Darstellung $\text{sign}\,(\sigma) = (-1)^a$. Ferner gilt:

$$\text{sign}\,(\sigma) = \prod_{j<k} \frac{k-j}{\sigma(k) - \sigma(j)}.$$

Beweis. Das Produktzeichen $\prod_{j<k}$ bedeutet, daß das Produkt über alle geordneten Paare $j, k \in \{1, \dots, \}, j < k$ genommen werden soll. Alle Zähler sind daher positiv. Wir sprechen von einer "Fehlstellung", falls $\sigma(k) < \sigma(j)$ ist. Ich habe gezeigt: Das Produkt ist -1 für jede Transposition; das Produkt ändert sein Vorzeichen, wenn σ durch $\tau \circ \sigma$, τ eine Transposition, ersetzt wird. Das beweist den Satz. 10.4.

ANALYSIS: Bitte ziehen Sie auch die Literatur zu Rate: Barner-Flohr, Königsberger. Außerdem die Texte zum letzten Semester.
Überblick über das vergangene Semester: Differenzieren von Polynomen (und rationalen Funktionen) mit Hilfe expliziter Fehler vor Kenntnis von Grenzwerten oder reellen Zahlen. Schrankensatz mit Archimedes Argument (wegen gleichmäßiger Fehlerkonstanten). Für die Differentiationsregeln gnügt es, Hauptterme und Fehlerterme übersichtlich zu sortieren.

— Die nächste Aufgabe ist die Konstruktion neuer Funktionen (exp, sin,...). Zur Definition benötigt man die Begriffe *Intervallschachtelung, Konvergenz, Cauchyfolge* und die Vollständigkeit von $\mathbb{R}$. Zur Kontrolle der Grenzfunktionen verwenden wir die Schrankensatz-Anwendung: $|f_n''| \leq B \Rightarrow |f_n(x) - f_n(a) - f_n'(a)(x-a)| \leq B \cdot |x-a|^2$ zusammen mit dem Archimedes Argument. Zum Auffinden von Approximationen f_n sind Taylorpolynome ein guter Versuch. Ein zweites Werkzeug ist die geometrische Majorisierung, $|x_{n+1} - x_n| \leq q \cdot |x_n - x_{n-1}|$ oder $|x_{n+1} - x_n| \leq q^n \cdot C$. Diese ist bei kontrahierenden Abbildungen und bei Potenzreihen besonders erfolgreich. *Vergleiche erstes Aufgabenblatt, immer mal wieder.* Integralrechnung: Der Schrankensatz erlaubt, Riemannsummen von $f = F'$ bis auf kleine Fehler aus $F(b) - F(a)$ zu berechnen; Stammfunktionen findet man für alle Funktionen $f : [a,b] \to \mathbb{R}^d$, die sich *gleichmäßig* durch stückweise lineare Funktionen approximieren lassen; Integrale werden als Grenzwerte von Riemannsummen definiert; alles zusammen gibt den Hauptsatz: $\int_a^b F'(x)\, dx = F(b) - F(a)$. Interessante Anwendung: Die Bogenlänge stetig differenzierbarer Kurven $f : [a,b] \to \mathbb{R}^d$ – **definiert** als Supremum von Sehnenlängen – kann als Integral $\int_a^b |f'(t)|\, dt$ berechnet werden. Bleibt die Frage: Warum ist die Berechnung der Fläche unter dem Graphen von Funktionen keine lehrreiche Anwendung? 11.4.

Analysis: Bei der Vokabel "Flächeninhalt" denkt man automatisch verschiedene Eigenschaften mit:
1.) Teilmengen von $\mathbb{R}^2$, die man sich vorstellen kann, "haben" einen Flächeninhalt.
2.) $G_1 \subset G_2 \Rightarrow Fl(G_1) \leq Fl(G_2)$.
3.) $Fl(G) = Fl(\text{Drehung}(G))$.
4.) $G_1 \subset G_2 \subset \ldots \subset G_n \ldots, \quad G := \cup_n G_n \Rightarrow \ Fl(G) = \lim_n Fl(G_n)$.
Es ist wesentlich schwieriger als bei der Bogenlänge, eine Definition des Flächeninhalts anzugeben, die all diese Eigenschaften hat. Infolgedessen hat die Aussage *"Der Flächeninhalt unter dem Graphen von f ist $\int_a^b f(x)\, dx$"* keine mathematische Bedeutung. Andererseits kann man leicht einsehen, daß keine andere Zahl außer $\int_a^b f(x)\, dx$ für den Flächeninhalt unter dem Graphen von f in Frage kommt, vgl. das Manuskript.
Wiederholung des Hauptsatzes der Differential- und Integralrechnung:
1. Definition des Integrals $\int_a^b f(x)\, dx$mit Hilfe von Riemannsummen $\mathfrak{RS}(f)$,
2. stetige Funktionen $f : [a,b] \to \mathbb{R}^d$ haben Stammfunktionen $F : [a,b] \to \mathbb{R}^d,\ F' = f$,
3. Zu jedem $\epsilon > 0$ gibt es ein $\delta > 0$, so daß für alle Einteilungen von $[a,b]$, die feiner als δ sind, gilt: $|F(b) - F(a) - \mathfrak{RS}(f)| \leq \epsilon$.

Daraus folgt $F(b) - F(a) = \int_a^b f(x)\, dx$ und für stetig differenzierbare $F : [a,b] \to \mathbb{R}^d$ hat man die Rekonstruktion aus der Ableitung:
$$F(x) = F(a) + \int_a^x f(s)\, ds.$$

Dieser Hauptsatz der Differential- und Integralrechnung ist zusammen mit der kontinuierlichen Dreiecksungleichung $|\int_a^b f(x)\, dx| \leq \int_a^b |f(x)|\, dx$ in der Lage, Taylorabschätzungen sehr übersichtlich herzuleiten. Voraussetzung $|f^3(x)| \leq S_3$:

$$|f(x)-f(a)-f'(a)(x-a)-\frac{1}{2}f''(a)(x-a)^2| = |\int_a^x\int_a^w\int_a^v f^{(3)}(u)\;du\;dv\;dw|$$
$$\leq \int_a^x\int_a^w\int_a^v |f^{(3)}(u)|\;du\;dv\;dw \leq \frac{S_3}{3!}|x-a|^3,\quad \text{usw.}$$

Lin. Algebra: Wiederholung des Steinitzschen Austauschsatzes und Vergleich mit dem Lösen linearer Gleichungssysteme; es handelt sich um verschiedene Ansichten desselben Problems. Induktionsbeweis. In diesen geht zweimal der richtige Umgang mit linear unabhängig ein, erstens falls die Vektoren $w_j := \sum_{k=1}^n a_{kj}\cdot v_k$ linear unabhängig sind, so ist mindestens einer der Koeffizienten a_{1j} ungleich 0, z.B. $a_{1j^*}\neq 0$; daher kann man v_1 durch w_{j^*} ersetzen, d.h. $v_1 = (1/a_{1j^*})\cdot(w_{j^*}-\sum_{k=2}^n a_{kj^*}\cdot v_k)$ und dies in die übrigen $n-1$ Linearkombinationen einsetzen:

$$\tilde{w}_j := w_j - a_{1j}/a_{1j^*}\cdot w_{j^*} = \textstyle\sum_{k=2}^n(a_{kj}-a_{1j}/a_{1j^*}\cdot a_{kj^*})\cdot v_k),\ j\neq j^*.$$

Hier kann die Induktionsvoraussetzung benutzt werden, sobald man eingesehen hat, daß die $n-1$ Vektoren $\tilde{w}_j$ wieder linear unabhängig sind — aber andernfalls hätte man eine nichttriviale Linearkombination auch der $w_j,\ j=1\dots n$. 13.4.

2.Woche, 17.4.-20.4.2000

Lin. Algebra: Beendigung der Gegenüberstellung von "Umkehren von Linearkombinationen", "Lösen von Gleichungssystemen", "Invertieren von Matrizen". Demorechnung: Bei dreistelliger Rundung aller Zwischenergebnisse hängt das Ergebnis der Gaußelimination u.U. **stark** von der gewählten Reihenfolge der Gleichungen (also von Vertauschungen der Zeilen der Matrix) ab. *Bearbeiten Sie alle Aufgaben zu "linear un-abhängig" besonders sorgfältig, die angestrebte Routine ist noch längst nicht erreicht. Trotzdem ist dies Thema für die Vorlesung beendet.*

Eigenwerte und Eigenvektoren. Wiederholung des Satzes über die Diagonalisierbarkeit *symmetrischer* Endomorphismen (d.h. $\langle Av,w\rangle = \langle v,Aw\rangle$), invarianter Unterräume (d.h. $A\cdot U\subset U$), orthogonaler Abbildungen (d.h. $\langle Lv,Lw\rangle = \langle v,w\rangle$), (auch: "skalarprodukterhaltend", längentreu", "isometrisch"). Die symmetrischen Endomorphismen bilden einen Vektorraum, die orthogonalen bilden unter Komposition eine Gruppe. Beide gehören eng zu den Skalarprodukten, und es gilt die für das Herstellen von Normalformen wichtige Eigenschaft (gut für Induktionsbeweise):

$$M\cdot U\subset U \Rightarrow M\cdot U^\perp\subset U^\perp,\ M \text{ symmetrisch oder orthogonal.}$$

Hat man einen Eigenwert λ, so findet man durch Lösen eines linearen Gleichungssystems den Unterraum E_λ aller Eigenvektoren zum Eigenwert λ. (*Beachte*: Eigenvektoren werden immer $\neq 0$ vorausgesetzt, trotzdem erlaubt man $0\in E_\lambda$.) Aus dem Dimensionssatz folgt $\dim(E_\lambda) = n - \ker(A-\lambda\cdot\mathrm{id})$. Außerdem hatten wir in Dimension 2 die möglichen Eigenwerte als Nullstellen des Polynoms $P_A(\lambda) = (a_{11}-\lambda)(a_{22}-\lambda) - a_{12}a_{21}$ bestimmt.

Beispiele. Für die Matrix $A = \left(\begin{smallmatrix}0 & -1\\ 1 & 0\end{smallmatrix}\right)$ der 90°-Drehung hat das Polynom $P_A(\lambda) = \lambda^2+1$ in $\mathbb{R}$ keine Nullstellen (und die 90°-Drehung natürlich keinen Eigenvektor), während es in $\mathbb{C}$ die

Nullstellen $\pm\mathbf{i}$ hat, es also auch Eigenvektoren von A gibt, z.B. $\left(\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right) \cdot \left(\begin{smallmatrix} 1+\mathbf{i} \\ 1-\mathbf{i} \end{smallmatrix}\right) = \mathbf{i} \cdot \left(\begin{smallmatrix} 1+\mathbf{i} \\ 1-\mathbf{i} \end{smallmatrix}\right)$.
Für die Matrix $A = \left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right)$ ist $P_A(\lambda) = (\lambda - 1)^2$, also 1 der einzige mögliche Eigenwert; da jedoch $\mathrm{rang}(A - 1 \cdot \mathrm{id}) = 1$ ist, gibt es nur einen **eindimensionalen** Eigenraum, d.h. man kann keine Basis aus Eigenvektoren finden, obwohl $P_A(\lambda)$ in Linearfaktoren zerfällt.

Wir hatten im zweidimensionalen den Ausdruck $a_{11}a_{22} - a_{12}a_{21}$ als Flächeninhalt des Parallelograms aus den Spaltenvektoren interpretiert. Dies wird zur Definition der Determinante führen und zwar als Volumen von

$$\mathrm{spat}(v_1, \ldots, v_d) := \{v \in \mathbb{R}^d;\ v = \textstyle\sum_{j=1}^d x_j \cdot v_j,\ 0 \le x_j \le 1\}.$$

Vorher müssen noch die Fragen zu Permutationen geklärt werden, Wiederholung vom 10.4., vor allem des Homomorphismus $\mathrm{sign} : \mathbf{S}_n \to \{+1, -1\}$, der auf Transpositionen den Wert -1 hat. 17.10.

Analysis: Meine Zusammenfassung zu den Grundbegriffen folgt dem Manuskript Argumente der Analysis II:
Axiome für Normen, beschränkte Mengen, dehnungsbeschränkte Abbildungen, Nullfolgen, konvergente Folgen, Cauchyfolgen, vollständige Mengen, bei $a \in A$ stetige Abbildungen $F : A \subset V \to W$ (folgenstetig bei a, ϵ-δ-stetig bei a, stetig an jeder Stelle $a \in A$, gleichmäßig stetig auf A).
Leichte Eigenschaften: Summe, (Skalar-)Produkt und Komposition dehnungsbeschränkter bzw. stetiger Abbildungen ist dehnungsbeschränkt bzw. stetig.
Hauptsätze: (H1) Äquivalenz von folgenstetig und ϵ-δ-stetig.
(H2) Für stetige Abbildungen von *beschränkten und vollständigen* Mengen, $F : A \to W$, ist $F(A)$ *beschränkt.*
(H3) Stetige Abbildungen von beschränkten und vollständigen Mengen, $F : A \to W$, sind auf A *gleichmäßig stetig.*
Die definierten Eigenschaften ändern sich nicht, wenn eine Norm $|\ |_1$ durch eine äquivalente Norm $|\ |_2$ ersetzt wird (äquivalent: Es gibt C so daß $\frac{1}{C}|v|_1 \le |v|_2 \le C|v|_1$ für alle $v \in V$).
Beispiele : Eine $|\ |_1$-Nullfolge ist auch $|\ |_2$-Nullfolge,
eine $|\ |_1$-beschränkte Menge ist auch $|\ |_2$-beschränkt,
eine $|\ |_1$-stetige Abbildung ist auch $|\ |_2$-stetig.

Wir streben den Satz an: *Auf* $\mathbb{R}^d$ *sind je zwei Normen äquivalent.*

Erstes Teilergebnis: Ist $|\ |_P$ eine Pythagoras-Norm und $|\ |_V$ eine beliebige Norm auf einem d-dimensionalen Vektorraum V, so gibt es ein C so daß

$$v \in V \Rightarrow |v|_V \le C \cdot |v|_P, \ \text{wir sagen: } |\ |_P \, \textit{dominiert} \ |\ |_V.$$

Beweis mit ON-Basen für $|\ |_P$, vergleiche Argumente der Analysis II, S.4.

Für die umgekehrte Ungleichung benötigen wir:
Die dominierte Norm ist als Funktion $|\ |_V : V \to \mathbb{R}$ bezüglich der Pythagoras-Norm dehnungsbeschränkt (und damit stetig), d.h.

$$\big|\, |v|_V - |w|_V \,\big|_{\mathbb{R}} \le |v - w|_V \le C \cdot |v - w|_P. \qquad 18.4.$$

ANALYSIS: Auf dem Vektorraum der Polynome sind die folgenden beiden Normen nicht äquivalent: $|P|_1 := \max\{|P(x)|; 0 \le x \le 1\}$, $|P|_2 := \sqrt{\int_0^1 P(x)^2 dx}$.
Die erste dominiert: $|P|_2 \le |P|_1$, aber $|X^n|_1 = 1$, $|X^n|_2 = 1/\sqrt{2n+1}$, eine Nullfolge.

Ende des Beweises zu: Je zwei Normen auf $\mathbb{R}^d$ sind äquivalent. Das Schlußargument ist: Die auf der $|\ \ |_P$-Einheitssphäre **stetige** Funktion $v \to 1/|v|_V$ ist dort beschränkt, also $|v|_P = 1 \Rightarrow 1/|v|_V \le M$, oder $\frac{1}{M}|v|_P \le |v|_V$.

Damit sind die Grundbegriffe (noch einmal: *beschränkte Menge, dehnungsbeschränkte Abbildung; Nullfolge - konvergente Folge - Cauchyfolge; folgenstetig - ϵ-δ-stetig; vollständig)* unabhängig davon, welche Norm auf $\mathbb{R}^d$ verwendet wird. Die Weiterentwicklung dieser abstrakten Begriffe wird unterbrochen für handwerkliche Analysiskenntnisse.
Newtonverfahren. Urbilder unter der Funktion f des Wertes A können als Fixpunkte der Newton-Abbildung $N(z) = z - (f(z) - A)/f'(z)$ bestimmt werden. Wegen $z^* = N(z^*) \Rightarrow N'(z*) = 0$ ist die Newton-Abbildung immer in einer kleinen Umgebung von z^* kontrahierend, also $z_{n+1} := N(z_n)$ konvergent gegen einen Fixpunkt – **falls** man die Startwerte in der genannten kleinen Umgebung des Fixpunktes wählen kann. Dies wird zuerst reell an Skizzen diskutiert, einschließlich des gelegentlich sehr viel einfacheren Verfahrens $\tilde{N}(z) = z-(f(z)-f(a))/f'(a)$, bei dem durch die Ableitung im Zielpunkt dividiert wird. (Beispiel für die Berechnung von π als Nullstelle von sin: $\tilde{N}(x) = x+\sin(x), x_0 = 3$.) Danach werden noch einmal Bilder komplexer Abbildungen besprochen, die besonders an Nullstellen der Ableitung dasselbe lokale Verhalten wie $z \to z^2$ zeigen. Danach werden die Iterationsbilder für zwei Newtonverfahren, die gegen dritte Einheitswurzeln konvergieren, gezeigt. Mit einem endlich anziehenden Fixpunkt und ∞ als zweitem Attraktor: $\tilde{N}(z) = z + 1/3 - (z^3)/3$. Und mit drei endlichen attraktiven Fixpunkten und ∞ als expandierendem Fixpunkt: $N(z) = \frac{2}{3}z + \frac{1}{3}z^{-2}$. 20.4.

3.Woche, 25.4.-27.4.2000

LIN. ALGEBRA. Ziel: Wir wollen diejenigen λ bestimmen, für die der Rang von $(A - \lambda \cdot \mathrm{id})$ kleiner als n ist, so daß also Eigenvektoren zu λ existieren. Dazu suchen wir eine Funktion der Spaltenvektoren, die das Volumen des von den Spaltenvektoren aufgespannten Spates liefert, wie es für $n = 2$ durch $a_{11}a_{22} - a_{12}a_{21}$ geleistet wird.
Erstens sollte diese Funktion der n Vektorargumente *homogen* sein, also

$$f(a_{\cdot 1}, \ldots, r \cdot a_{\cdot j}, \ldots, a_{\cdot n}) = r \cdot f(a_{\cdot 1}, \ldots, a_{\cdot j}, \ldots, a_{\cdot n}),\ r \in \mathbb{R}$$

weil sich Volumina bei Verlängerung einer Kante des Spates so verhalten – offensichtlich zunächst für $r \ge 0$, aber wegen des zweidimensionalen Beispiels $a_{11}a_{22} - a_{12}a_{21}$ verlangen wir die Eigenschaft für alle $r \in \mathbb{R}$.
Zweitens sagt das Prinzip von Cavalieri, daß sich das Volumen eines Stapels Papier nicht ändert, wenn wir die Blätter gegeneinander verschieben, d.h. wenn wir zu einem Kantenvektor des Spats eine beliebige Linearkombination der übrigen addieren:

$$f(a^{\vdots}{}_1,\dots,a^{\vdots}{}_j,\dots,a^{\vdots}{}_n) = f(a^{\vdots}{}_1,\dots,a^{\vdots}{}_j,\dots,a^{\vdots}{}_n + \textstyle\sum_{j=1}^{n-1} x_j\cdot a^{\vdots}{}_j).$$

Drittens können wir zwei Spate, die sich nur in einem Kantenvektor unterscheiden, aufeinander stellen, und, wieder nach Cavalieri, sollte das Volumen dasselbe bleiben, wenn wir die geknickte Kante durch den Summenvektor ersetzen, also:

$$f(a^{\vdots}{}_1,\dots,a^{\vdots}{}_j,\dots,a^{\vdots}{}_n) + f(a^{\vdots}{}_1,\dots,a^{\vdots}{}_j,\dots,b^{\vdots}{}_n) = f(a^{\vdots}{}_1,\dots,a^{\vdots}{}_j,\dots,a^{\vdots}{}_n + b^{\vdots}{}_n).$$

Etwas gestraffter zusammengefaßt sagen die erste und die dritte Eigenschaft, daß die partiellen Abbildungen unserer gesuchten Funktion alle linear sein sollen:

$$a^{\vdots}{}_j \mapsto f(a^{\vdots}{}_1,\dots,a^{\vdots}{}_j,\dots,a^{\vdots}{}_n) \text{ ist } \textbf{linear} \text{ für jedes } j\in\{1\dots n\}.$$

Aus den ersten beiden Eigenschaften folgt, daß sich beim Vertauschen von irgend zwei Argumenten das Vorzeichen ändert:

$$f(a^{\vdots}{}_1,\dots,a^{\vdots}{}_j,\dots,a^{\vdots}{}_k,\dots,a^{\vdots}{}_n) = -f(a^{\vdots}{}_1,\dots,a^{\vdots}{}_k,\dots,a^{\vdots}{}_j,\dots,a^{\vdots}{}_n).$$

Der Name für eine Abbildungen mit diesen beiden Eigenschaften ist *alternierende n-Form.* Wir haben bisher großzügig Wünsche aneinandergereiht, nun folgt aus dem Vorzeichen-Homomorphismus von Permutationen, $\sigma\mapsto \mathrm{sign}\,(\sigma)$, daß es wirklich eine Funktion mit diesen Eigenschaften gibt. Das habe ich zuerst für $n=3$ durch Aufzählen aller Permutationen explizit vorgeführt, weil im $n=3$ Fall auch eine sehr einfache Merkregel für das Berechnen von 3×3-Determinanten herauskommt. Insbesondere sind die Zahlen λ, für die der Rang von $(A-\lambda\cdot\mathrm{id})$ kleiner als 3 ist, die Nullstellen eines *kubischen Polynoms*, $P(\lambda)=\det(A-\lambda\cdot\mathrm{id})$. Der Fall $n>3$ ist nicht wirklich schwieriger, sobald man den Vorzeichen-Homomorphismus für Permutationen kennt. — Wenn Sie über Determinanten Bücher lesen, vergessen Sie bitte nicht, daß die Eigenschaften der Determinante von n Vektoren motiviert sind durch das, was Sie von dem Volumen des (von den Vektoren aufgespannten) Spates erwarten. 25.4.

Analysis: Als unmittelbare Ziele muß ich handwerkliche Fähigkeiten zu (komplexer) Differential- und Integralrechnung üben und Verständnis für die Hauptsätze über stetige Funktionen wecken.

Zur Erinnerung an die Leibnizreihen habe ich die Irrationalität von $1/e$ bewiesen: Die Leibnizreihe liefert $0<|\sum_{k=0}^{k=n}(-1)^k/k! \;-\; 1/e\,|<1/(n+1)!$, außerdem erlaubt die Annahme $1/e=P/Q,\ P,Q\in\mathbb{N}$, in der Abschätzung $n=Q$ zu nehmen, um den Widerspruch $1/Q!\le|\sum_{k=0}^{k=Q}(-1)^k/k! \;-\; 1/e\,|<1/(Q+1)!$ zu bekommen.

Die Definition *f ist stetig an jeder Stelle $a\in A\subset\mathbb{R}^d$* und die Definition *$f$ ist gleichmäßig stetig in $A\subset\mathbb{R}^d$* wurden noch einmal mit Beispielen gegenübergestellt. Dazu die Warnung: Stetige Abbildungen bilden i.a. **nicht** Cauchyfolgen auf Cauchyfolgen ab. Gleichmäßig stetige F aber haben diese angenehme Eigenschaft, Beweis durch (trickloses) Einsetzen der Definitionen. Ziel ist jetzt, zu verstehen, daß stetige Abbildungen mit manchen Sorten Mengen besonders gut zusammengehen, vgl. Aufg. 3.8 zu *offenen Mengen*; bei den Hauptsätzen sind die *vollständigen und beschränkten Mengen*, die auch *kompakt* genannt werden, der Angelpunkt. Ich habe den Beschränktheitssatz (Manuskript S.5, WS S.56) und den Satz

über gleichmäßige Stetigkeit auf *kompakten* Mengen noch einmal bewiesen (WS Manuskript S.57). (Angenommen $F : A \to \mathbb{R}^e$ sei nicht gleichmäßig stetig; dann gibt es zu einem $\epsilon * > 0$ kein δ mehr, d.h. zu jedem $\delta_n := 2^{-n}$ gibt es ein Ausreißerpaar $a_n, b_n \in A,\ |a_n - b_n| \le \delta_n$ mit $|F(a_n) - F(b_n)| \ge \epsilon *$. Da A *beschränkt* ist, liegt A in einem Quader Q_1; durch Halbieren der Seitenlängen findet man 2^d Teilquader und für mindestens einen, genannt Q_2, gibt es zu $\epsilon * > 0$ schon in $A \cap Q_2$ kein δ mehr; daher kann man die a_n als Cauchyfolge wählen. Da A *vollständig* ist, konvergiert $\{a_n\}$ gegen $a_\infty \in A$. Da F in $a_\infty \in A$ *stetig* ist, gibt es zu $\epsilon := \epsilon^*/3$ ein $\delta > 0$ mit $x \in A,\ |x - a_\infty| \le \delta \Rightarrow |F(x) - F(a_\infty)| \le \epsilon^*/3$. Schließlich können wir n so groß wählen, daß a_n, b_n die Voraussetzungen an x erfüllen. Das gibt den Widerspruch $|F(a_n) - F(b_n)| \le \frac{2}{3}\epsilon *$ zur Ausreißereigenschaft $|F(a_n) - F(b_n)| \ge \epsilon *$.)
Sie sollten die Begriffe “stetig” und “gleichmäßig stetig” jetzt sicher unterscheiden können.

Zurück zum Handwerklichen. Ich habe begonnen, die beiden für **affine** Funktionen $f : [a, b] \to \mathbb{R}^d,\ f(x) := f(a) + f'(a) \cdot (x - a)$ “richtigen” Riemannsummen für $\int_a^b f(x)\ dx$, nämlich $ST(f) := (f(a) + f(b))/2 \cdot (b - a)$ und $TT(f) := f((a + b)/2) \cdot (b - a)$, für allgemeinere f auf ihre Fehler zu untersuchen. Für $d = 1$ vergleiche Manuskript S.7-8.Unter der Voraussetzung $|f''| \le B$ haben wir auch für $d > 1$ die Schrankensatzkontrolle über die Abweichung von der Tangente: $|f(x) - f(c) - f'(c)(x - c)| \le \frac{1}{2}B \cdot |x - c|^2$. Mit $c = (a + b)/2$ folgt hieraus mit der kontinuierlichen Dreiecksungleichung:

$$\Big|\int_a^b f(x)\ dx - TT(f)\Big| \le \frac{B}{24}|b - a|^3.$$

Vergleiche Aufgabe 4.6. 27.4.

4.Woche, 2.5.-4.5.2000

Lin. Algebra. Die Behandlung der Determinanten wurde weitgehend beendet: Falls $f : V \times \ldots \times V \to K$ n-linear und alternierend (d.h. Vorzeichenänderung beim Vertauschen von zwei Argumenten) so folgte aus dem Vorzeichen-Homomorphismus für Permutationen eine Formel, die erstens mindestens eine Abbildung mit allen gewünschten Eigenschaften definiert und die zweitens zeigt, daß *zwei solche Abbildungen proportional sind.* Die Formel:
Sei $\{e_1, \ldots, e_n\}$ eine Basis von V und $v_j := \sum_{k=1}^n a_{kj} \cdot e_k,\ j = 1 \ldots n$. Dann gilt:

$$f(v_1, \ldots, v_n) = \sum_{\sigma \in \mathfrak{S}_n} a_{\sigma(1)1} \cdot \ldots \cdot a_{\sigma(n)n} \cdot \mathrm{sign}\,(\sigma)$$

Der Laplace’sche Entwicklungssatz ergibt sich, indem man die Terme mit erstem Faktor a_{k1} zusammenfaßt; die Summe der restlichen Faktoren ist die Determinante der komplementären Matrix, bis auf das Vorzeichen $(-1)^{k+1}$. (Begründung: $\sigma : \{2, \ldots, n\} \to \{1, \ldots, n\} \setminus \{k\}$ kann als Permutation $\hat{\sigma} \in \mathfrak{S}_{n-1}$ aufgefaßt werden; die Berechnung von $\mathrm{sign}\,(\sigma) = (-1)^{k-1}\mathrm{sign}\,(\hat{\sigma})$ gelingt mit der Formel, die Fehlstellungen zählt.)
Ich habe zwar die Bezeichnung Determinante schon verwendet, bisher wäre jedoch n-lineare alternierende Abbildung $f : V \times \ldots \times V \to K$ besser gewesen, um einen deutlichen Unterschied zur Determinante von Endomorphismen zu machen. Sei jetzt f multilinear und

alternierend und $A \in \mathrm{End}(V)$; definiere

$$f_A(v_1, \ldots, v_n) := f(A \cdot v_1, \ldots, A \cdot v_n).$$

Offenbar ist f_A multilinear und alternierend, also *proportional zu* f; der Proportionalitätsfaktor heißt Determinante des Endomorphismus, also

$$f(A \cdot v_1, \ldots, A \cdot v_n) = f_A(v_1, \ldots, v_n) = \det(A) \cdot f(v_1, \ldots, v_n).$$

Die Definition liefert direkt den wichtigen

Determinanten-Multiplikationssatz. Sind $A, B \in \mathrm{End}(V), \mathrm{rang}(B) = n$ so gilt

$$\begin{aligned}\det(A \circ B) &= f(A \circ B \cdot v_1, \ldots, A \circ B \cdot v_n)/f(v_1, \ldots, v_n) = \\ &= f(A \circ B \cdot v_1, \ldots, A \circ B \cdot v_n)/f(B \cdot v_1, \ldots, B \cdot v_n) * f(B \cdot v_1, \ldots, B \cdot v_n)/f(v_1, \ldots, v_n) \\ &= \det(A) \cdot \det(B).\end{aligned}$$

Und für $\mathrm{rang}(B) < n$ heißt die Behauptung $0 = \det(A) \cdot 0$.

Damit kann $\det(A)$ aus *jeder* Matrix für A berechnet werden, außerdem ändert Konjugation die Determinante nicht: $\det(A) = \det(S \circ A \circ S^{-1})$. Auch die Bestimmung von Eigenwerten wird jetzt beantwortet: $\mathrm{rang}(A - \lambda \cdot \mathrm{id}) < n$ genau dann wenn λ Nullstelle des *charakteristischen Polynoms* $P_A(\lambda) := \det(A - \lambda \cdot \mathrm{id})$ *ist.* Natürlich sind auch die Koeffizienten des charakterischen Polynoms unter Konjugation invariant, insbesondere der Koeffizient $\sum_j a_{jj}$ von $(-\lambda)^{n-1}$, genannt spur A, vgl. Aufgabe 3.4.

Mit Hilfe von Skalarprodukten $\langle\ \ ,\ \rangle$ und Orthonormalbasen $\{e_1, \ldots, e_n\}$ hatten wir die Koeffizienten von Linearkombinationen $v = \sum_{k=1}^{n} x_k \cdot e_k$ leicht berechnen können als $x_k = \langle v, e_k \rangle$. Die Determinanten liefern eine Formeln ohne spezielle Basen:

$x_k = \det(e_1, \ldots, e_{k-1}, v, e_{k+1}, \ldots, e_n)/\det(e_1, \ldots, e_n)$. 2.5.

ANALYSIS: Die Definition des Integrals der "Kurve" $f : [a, b] \to \mathbb{R}^e$ mit Hilfe von Riemannsummen wurde (zum letzten Mal) wiederholt. Die bequeme Anwendung der kontinuierlichen Dreiecksungleichung $|\int_a^b f(x)\ dx| \le \int_a^b |f(x)|\ dx$ wurde an numerischen Berechnungen vorgeführt: Sei $tang(x) := f(\frac{a+b}{2}) + f'(\frac{a+b}{2})(x - \frac{a+b}{2})$ die Tangente im Intervallmittelpunkt; dann gilt unter der Voraussetzung $|f''| \le B$:

$$\begin{aligned}&|\int_a^b f(x)\ dx - f(\frac{a+b}{2})(b-a)| = |\int_a^b (f(x) - tang(x))\ dx| \\ &\le \int_a^b |f(x) - tang(x)|\ dx \le \int_a^b \frac{B}{2}(x - \frac{a+b}{2})^2\ dx \le \frac{B}{24}(b-a)^3\end{aligned}$$

Diese Rechnung ist unabhängig von der Dimension des Bildraums $\mathbb{R}^e$ der Kurve. Für $e = 1$ habe ich dasselbe mit der Sehne, $sehne(x) := \big((f(a)(b-x) + f(b)(x-a)\big)/(b-a)$, gemacht und dazu die Abschätzung $|f''| \le B \Rightarrow |f(x) - sehne(x)| \le \frac{B}{2}(x-a)(b-x)$ wiederholt, deren Beweis aus nicht verallgemeinerbaren 1-dimensionalen Argumenten besteht. Fast dieselbe Rechnung wie eben liefert:

$$|\int_a^b f(x)\ dx - (f(a) + f(b))\frac{b-a}{2}| \le \frac{B}{12}(b-a)^3.$$

Kommentare. Erstens bleibt die Frage, ob man den Unterschied zwischen Kurve und Sehne auch bei $e > 1$ durch $|f''| \le B$ kontrollieren kann. Zweitens verweise ich auf das Manuskript, in dem diskutiert ist, wie sich diese Formeln verhalten, wenn man das Intervall erst unterteilt. Drittens kann man die beiden Näherungen mitteln, mit Gewichten $2:1$, umgekehrt wie die Größe ihrer Fehler, um eine Näherung zu bekommen, die bis zu *kubischen* Polynomen den richtigen Wert des Integrals liefert, die Simpsonregel:

$$simpson(f) := \frac{2}{3} f(\frac{a+b}{2})(b-a) + \frac{1}{3}(f(a)+f(b))\frac{b-a}{2} = \frac{b-a}{6}\big(f(a)+4f(\frac{a+b}{2})+f(b)\big).$$

Eine nicht vorgeführte viermalige Anwendung des Schrankensatzes liefert

$$\Big|\int_a^b f(x)\;dx - simpson(f)\Big| \le \frac{\max|f''''|}{2880}(b-a)^5.$$

Zur Fortsetzung der Diskusion der Hauptsätze (über stetige Funktionen) benötigen wir den Begriff **gleichmäßig konvergente Funktionenfolge**. Bisher war uns *gleichmäßig* begegnet bei *gleichmäßig stetig, gleichmäßig stückweise linear approximierbar* (bei der Konstruktion von Stammfunktionen) und *gleichmäßige Fehlerschranken bei Tangentenapproximation* (bei den ersten Beweisen des Schrankensatzes und bei Differenzierbarkeitsbeweisen von Grenzfunktionen). Jetzt benötigen wir die

Definition. Eine Funktionenfolge $f_n : A \subset \mathbb{R}^d \to \mathbb{R}^e$ heißt **gleichmäßig konvergent**, falls es zu jedem $\epsilon > 0$ ein n_ϵ gibt, ausdrücklich *unabhängig* von $x \in A$, so daß gilt

$$x \in A,\; m, n \ge n_\epsilon \Rightarrow |f_m(x) - f_n(x)| \le \epsilon.$$

Hauptsatz Gleichmäßig konvergente Folgen stetiger Funktionen $f_n : A \to \mathbb{R}^e$ haben eine **stetige** Grenzfunktion $f_\infty : A \to \mathbb{R}^e$. (vgl. Manuskript)

Beweis. a) Die Grenzfunktion existiert, weil die $\{f_n(x)\}_n$ Cauchyfolgen in $\mathbb{R}^e$ sind.
b) Zu $\epsilon > 0$ wähle ein n_ϵ so daß für alle $x \in A$ und $m, n \ge n_\epsilon$ gilt $|f_n(x) - f_m(x)| \le \epsilon/3$, also auch $|f_\infty(x) - f_m(x)| \le \epsilon/3$. Sodann wähle ein $m^* \ge n_\epsilon$ und wähle mit der Stetigkeit von f_{m^*} ein $\delta > 0$ so daß $x, a \in A,\; |x-a| \le \delta \Rightarrow |f_{m^*}(x) - f_{m^*}(a)| \le \epsilon/3$. Danach ergibt die Dreiecksungleichung $x, a \in A,\; |x-a| \le \delta \Rightarrow$
$|f_\infty(x) - f_\infty(a)| \le |f_\infty(x) - f_{m^*}(x)| + |f_{m^*}(x) - f_{m^*}(a)| + |f_\infty(a) - f_{m^*}(a)| < 3\epsilon/3$,
also die Stetigkeit von f_∞.
Diese bisher noch nicht besprochene Beweisstrategie werde ich als $\epsilon/3$-Argument zitieren.

5.Woche, 9.5.-11.5.2000

ANALYSIS. Wiederholung zu Definitionen und zum Hauptsatz über gleichmäßig konvergente Folgen stetiger Funktionen. Beispiele:
Der Weierstraß-Oszillator ist eine *stetige* Funktion, weil die Approximationen $f_n(x) := \sum_{k=1}^n 2^{-k}\text{säg}\,(4^k x)$ gleichmäßig konvergieren (geometrisch majorisiert durch $\{2^{-k}\}$).

Falls eine Potenzreihe $P_n(z) := \sum_{k=0}^n a_k z^k$ bei $z = z_0 \neq 0$ konvergiert, so konvergiert $\{P_n\}$gleichmäßig in der Scheibe $A_q := \{z \in \mathbb{C}; |z| \leq q|z_0|\}$, $0 < q < 1$ (geometrisch majorisiert durch $\{S \cdot q^k\}$, da die Nullfolge $\{a_k z_0^k\}$ beschränkt ist, z.B. $\leq S$).
Umformulierung des Hauptsatzes mit Hilfe der Supremum-Norm auf dem Vektorraum $V := \{f : A \to \mathbb{R}^e; f \text{ stetig, beschränkt}\}$,
Definition: Sei $f \in V$, setze $\|f\| := \sup\{|f(x)|_{\mathbb{R}^e};\ x \in A\}$, also $x \in A \Rightarrow |f(x)| \leq \|f\|$.
Überprüfung der Axiome für eine Norm, vgl. Manuskript.

Satz, Umformulierung des Hauptsatzes über gleichmäßige Konvergenz: Der Vektorraum V ist $\|\ \ \|$-vollständig, d.h. jede Cauchyfolge $\{f_n\} \subset V$ konvergiert gegen eine Funktion $f_\infty \in V$.

Beweis. Wegen: Zu jedem $\epsilon > 0$ gibt es n_ϵ so daß $m, n \geq n_\epsilon \Rightarrow \|f_m - f_n\| \leq \epsilon$ folgt aus der Definition dieser Norm auch $|f_m(x) - f_n(x)| \leq \epsilon$ für jedes $x \in A$. Daher ist die Folge *gleichmäßig konvergent* gegen eine *stetige* Grenzfunktion f_∞. Wegen $|f_\infty(x)| \leq |f_\infty(x) - f_m(x)| + |f_m(x)| \leq \epsilon + \|f_m\|$ ist die Grenzfunktion auch beschränkt und damit $f_\infty \in V$. Schließlich folgt mit Archimedes $m \geq n_\epsilon \Rightarrow |f_m(x) - f_\infty(x)| \leq \epsilon$, also auch $\|f_m - f_\infty\| \leq \epsilon$. $\{\|f_m - f_\infty\|\}$ ist damit Nullfolge also $\{f_n\}$ $\|\ \ \|$-konvergent gegen f_∞ in V.
Ich habe bei dieser Gelegenheit auch verschiedene Standardargumente wiederholt und betrachte das Kapitel stetige Abbildungen zunächst als abgeschlossen.

Fundamentalsatz der Algebra:

Jedes nicht konstante komplexe Polynom zerfällt in Linearfaktoren.

Beweis. (*Schritt* 1) Wegen des Abspaltens von Nullstellen, $P(z) - P(a) = (z - a)P_1(z)$, genügt es, zu **jedem** nicht konstanten Polynom eine Nullstelle zu finden, dann Induktion.
(*Schritt* 2) Sei $P(z) = (\sum_{k=0}^{n-1} a_k z^k) + z^n$. Dann braucht man außerhalb der Scheibe vom Radius $R := 1 + 3 \cdot \sum_{k=0}^{n-1} |a_k| \geq 1$ nicht nach Nullstellen zu suchen, denn dort ist

$$|P(z)| \geq \tfrac{2}{3}|z|^n \geq \tfrac{2}{3}R > 2|a_0|.$$

(*Schritt* 3) Indirekte Annahme: "P habe **keine** Nullstelle". Dann ist zunächst $1/P : \mathbb{C} \to \mathbb{C}$ stetig. Für $|z| \geq R$ zeigt *Schritt 2 :* $1/|P(z)| \leq 3/(2R)$, und auf der vollständigen und beschränkten (=**kompakten**) Scheibe $B_R := \{z; |z| \leq R\}$ *nimmt die stetige Funktion* $1/|P|$ *ihr Maximum* $M := 1/|P(m)| \geq 1/|P(0)| \geq 1/|a_0| > 3/R$ *an.* Damit ist M das Maximum von $1/P$ nicht nur auf $|z| \leq R$ sondern auf $\mathbb{C}$.
(*Schritt* 4) In Aufg. 4.7 wurde der Umkehrsatz begonnen: Ist f komplex differenzierbar, $f(a) = A$, $f'(a) \neq 0$ so existiert auf einer Umgebung von A eine Umkehrfunktion zu f, und daraus folgt: Eine geeignete Umgebung von A besteht nur aus Bildpunkten von f. — Dies kann in Schritt 3 auf $f = 1/P, a = m$ angewandt werden, *falls* $P'(m) \neq 0$; in dem Fall besteht eine Umgebung von $1/P(m)$ aus Bildpunkten von $1/P$ und $M = 1/|P(m)|$ kann **nicht** das Maximum von $1/|P|$ sein.
(*Schritt* 5) Die letzte Folgerung, nämlich daß eine Umgebung von $1/P(m)$ ganz aus Bildpunkten von $1/P$ besteht, gilt für *komplex differenzierbare* f *auch* **ohne** *die Voraussetzung*

$P'(m) \neq 0$. Daher bekommen wir auch in diesem Fall den Widerspruch aus Schritt 4 zu der indirekten Annahme in Schritt 3. Der Fundamentalsatz ist bewiesen.
In Schritt 5 wird eine sehr wichtige Eigenschaft komplex differenzierbarer Funktionen wenigstens für rationale f bewiesenen. 8.5.

LIN. ALGEBRA. Als Beispiel zu unerwarteten Eigenschaften von Determinanten habe ich gezeigt, daß die zweireihigen Unterdeterminanten eines Produktes $A \cdot B$ bilinear aus den zweireihigen Unterdeterminanten von A und B berechnet werden können. Auch die Beweismethode betrachte ich als lehrreich. Es genügt, zu zwei Paaren von Vektoren, a_1, a_2 und b_1, b_2, die Determinante $\langle a_1, b_1\rangle \cdot \langle a_2, b_2\rangle - \langle a_1, b_2\rangle \cdot \langle a_2, b_1\rangle$ aus den Unterdeterminanten $(a_{1j}a_{2k} - a_{1k}a_{2j})$ und $(b_{1j}b_{2k} - b_{1k}b_{2j})$ zu berechnen.

$$
\begin{aligned}
\langle a_1, b_1\rangle\langle a_2, b_2\rangle - \langle a_1, b_2\rangle\langle a_2, b_1\rangle &= (\sum_j a_{1j}b_{1j}) \cdot (\sum_k a_{2k}b_{2k}) - (\sum_j a_{1j}b_{2j}) \cdot (\sum_k b_{1k}a_{2k}) \\
&= \sum_{j,k} a_{1j}a_{2k} \cdot \big(b_{1j}b_{2k} - b_{1k}b_{2j}\big) \\
\text{und ebenso:} \qquad &= (\sum_k a_{1k}b_{1k}) \cdot (\sum_j a_{2j}b_{2j}) - (\sum_k a_{1k}b_{2k}) \cdot (\sum_j b_{1j}a_{2j}) \\
&= \sum_{j,k} -a_{1k}a_{2j} \cdot \big(b_{1j}b_{2k} - b_{1k}b_{2j}\big) \\
\text{also schließlich:} \qquad &= \frac{1}{2}\sum_{j,k} \big(a_{1j}a_{2k} - a_{1k}a_{2j}\big) \cdot \big(b_{1j}b_{2k} - b_{1k}b_{2j}\big).
\end{aligned}
$$

Wiederholt habe ich auch, daß ein Skalarprodukt auf V einen bijektiven Homomorphismus $V \to \mathrm{Hom}(V, \mathbb{R})$, $a \in V \mapsto l_a := \langle a, \ \rangle$ liefert. Dies wird bei festem $a, b \in V = \mathbb{R}^3$ angewandt auf $l(c) := \det(a, b, c)$, also $\det(a, b, c) = \langle \phi(a, b), c\rangle$. Da die Determinante linear in jedem Argument und alternierend ist, ist ϕ bilinear und schief, $\phi(a, b) = -\phi(b, a)$. Jede solche bilineare Abbildung $V \times V \to V$ können wir "Produkt" nennen, diese hier heißt **Kreuzprodukt in** $\mathbb{R}^3$, $\det(a, b, c) = \langle a \times b, c\rangle$. Ich wollte den Physikern später eine wichtige Anwendung auf den Drehimpuls vorführen: die Symmetrie des Trägheitstensors .

Die Auswirkung von Basiswechseln auf zusammengehörige Vektorräume

Zu V gehörige Vektorräume sind $\mathrm{Hom}(V, K)$, $\mathrm{End}(V)$, $\mathrm{Bil}(V), \ldots$. Ich habe am Beispiel der Polynome wiederholt, daß *jede Basiswahl in V in ganz natürlicher Weise (*d.h. wie wir es ohne besondere Hervorhebung schon immer gemacht haben) *zu einer Basiswahl in den anderen Vektorräumen führt.* In $\mathrm{Hom}(V, K)$ wählten wir die Dualbasis, also diejenigen linearen Abbildungen, die genau einen der Basisvektoren $v_j \in V$ auf 1, alle anderen auf $0 \in K$ abbilden; in $\mathrm{End}(V)$ wählten wir die $(\dim V)^2$ Endomorphismen, die genau einen Basisvektor v_j auf einen (meist anderen) Basisvektor abbilden und alle anderen Basisvektoren auf $0 \in V$. **Deshalb führt ein Basiswechsel in V automatisch zu Basiswechseln in diesen anderen "assoziierten" Vektorräumen. Insbesondere sind die Formeln**

für Koordinatenwechsel in den gleichdimensionalen Räumen V und $\mathrm{Hom}(V,K)$ verschieden. Dies besitzt bei konsequenter Verwendung der Matrixmultiplikation ein übersichtlicheres Erscheinungsbild als bei Verwendung von Σ-Zeichen. — Ich weise noch auf ein sprachliches Problem hin: Die Indices von Spaltenvektoren sind offenbar die Nummern der Zeilen; die deutsche Sprache erlaubt im Prinzip die Benutzung zusammengesetzter Substantive wie Spaltenindices und Zeilennummern, aber der Satz: " Die Spaltenindices sind die Zeilennummern" verwirrt sicher nicht nur mich, ich werde diese Komprimierung nicht verwenden. — Ich werde eine zusätzliche Information mitteilen: In allen Gebieten, in denen Lineare Algebra und Analysis eng vermischt sind, ist die Unterscheidung der Elemente von V und $\mathrm{Hom}(V,\mathbb{R})$ sehr wichtig, so wichtig, daß die Unterscheidung auch funktionieren muß, wenn Basen gewählt sind und nur die Komponenten der Vektoren bezüglich der Basen mitgeteilt werden. Daher hat sich die *Konvention* durchgesetzt, die Komponenten der Vektoren aus V mit *oberen* Indices zu schreiben, die der Vektoren aus $\mathrm{Hom}(V,\mathbb{R})$ mit *unteren* Indices. Diese Unterscheidung darf in dieser Vorlesung ignoriert werden, sie hilft aber, eine wirklich wichtige Unterscheidung nicht zu übersehen.

Jetzt seien $\{v_1,\ldots,v_d\}$ die "alte" Basis von V und $\{w_1,\ldots,w_d\}$ die "neue". Ferner sei $A = (a_j^k)$ die Transformationsmatrix von der alten auf die neue Basis, d.h. $w_j = \sum_k a_j^k v_k$. Für die Umkehrung $v_k = \sum_l b_k^l w_l$ folgt aus der linearen Unabhängigkeit der $\{w_l\}$: $\sum_k b_k^l a_j^k = 0$ außer für $j = l$, oder als Matrixmultiplikation: $B \cdot A = (\mathrm{id})$. – Wir schreiben $x \in V$ als $x = \sum_k \alpha^k \cdot v_k = \sum_l \beta^l \cdot w_l$ und finden wie gehabt $\beta^l = \sum_l b_k^l \cdot \alpha^k$ oder $(\beta) = B \cdot (\alpha)$. Mit anderen Worten: Die alten Komponenten α werden mit der **inversen** Matrix $B = A^{-1}$ in die neuen umgerechnet, invers zu der Matrix A, mit der die alte Basis in die neue umgerechnet wird. *Daher heißen in der Physik die Elemente von V* **kontravariante** *Vektoren.*
Nun sei $\{\ell^1,\ldots,\ell^d\}$ die Basis von $\mathrm{Hom}(V,K)$, die dual zu der alten Basis ist (also $\ell^j(v_k) = 1$ falls $j = k$ und $= 0$ sonst). Wir schrieben schon früher $\phi \in \mathrm{Hom}(V,K)$ als Linearkombination $\phi = \sum_j \omega_j \cdot \ell^j$ und berechneten $\phi(x) = \sum_j \omega_j \alpha^j = \text{Zeilenvektor}\times\text{Spaltenvektor}$. Wegen $\sum_j \omega_j \alpha^j = \omega(B^{-1} \cdot B)\alpha = (\omega \cdot A) \cdot \beta$ rechnen sich die alten Komponenten von ϕ in die neuen Komponenten (bezüglich der Dualbasis von $\{w_1,\ldots,w_d\}$) mit **derselben** Matrix A um, wie die Basis $\{v_1,\ldots,v_d\}$ in $\{w_1,\ldots,w_d\}$. *Daher heißen in der Physik die Elemente des Dualraums* $\mathrm{Hom}(V,K)$ **kovariante** *Vektoren.*
Die Verwendung hoch- und tief-gestellter Indices unterstützt also die **Unterscheidung** *der Elemente von V und* $\mathrm{Hom}(V,K)$*; sie teilt außerdem mit, daß für den Matrizenkalkül die Komponenten ω_j als Zeilenmatrix, die Komponenten α^k als Spaltenmatrix gelesen werden sollen.* 9.5.2000

Lin. Algebra. Kommentare zu den Aufgaben 3.3, 3.5, 3.7. In 3.5a) hätte es "konvex und *im endlich dimensionalen Fall* vollständig" heißen müssen; ich hatte nicht damit gerechnet, daß Sie nach dem Äquivalenzsatz für Normen auf *endlich dimensionalen* Vektorräumen und vor der Behandlung der Supremum-Norm für stetige Funktionen über unendlichdimensionale Vektorräume nachdenken würden — dieser Fehler tut mir leid.

Kurze Wiederholung vom 9.5., Fortsetzung mit Endomorphismen $L \in \mathrm{End}(V)$, deren Matrix nach Wahl der alten Basis $\{v_1, \ldots, v_d\}$ für V definiert ist durch $L \cdot v_j = \sum_{k=1}^{d} l^k{}_j \cdot v_k$, woraus mit $x = \sum_j \alpha^j \cdot v_j$ folgt $L \cdot x = \sum_j \alpha^j v_j$ oder ganz im Matrixkalkül $(Lx) = (L) \cdot (\alpha)$. Der Basiswechsel $w_j = \sum_k a^k{}_j v_k$ führt zu $x = \sum_l \beta^l \cdot w_l$ und daher zur "kontravarianten" Umrechnung der alten in die neue Koordinaten: $(\beta) = B \cdot (\alpha)$. Dies angewandt auf Lx ergibt:

$$(Lx)_{neu} = B \cdot (Lx)_{alt} = B \cdot (L)_{alt} \cdot (\alpha) = B \cdot (L)_{alt} \cdot B^{-1} \cdot (\beta)$$

oder

$$(L)_{neu} = B \cdot (L)_{alt} \cdot B^{-1}.$$

Dies Transformationsverhalten heißt in der Physik *einfach kovariant* (wg B^{-1}) *und einfach kontravariant* (wg B); hervorgehoben wird dies dadurch, daß die Matrix $(L) = (l^k{}_j)$ mit einem hochgestellten und einem tiefgestellten Index geschrieben wird.
Als letzten "assoziierten" Vektorraum betrachten wir die Bilinearformen $\mathrm{Bil}(V)$. Ein $\phi \in \mathrm{Bil}(V)$ wird auf zwei Vektoren $x = \sum_j x^j \cdot v_j,\ y = \sum_k y^k \cdot v_k \in V$ angewandt: $\phi(x, y) = \sum_{j,k} x^j y^k \phi(v^j, v^k)$. Daher wird $(f_{jk})_{alt} = \phi(v^j, v^k)$ als Matrix von ϕ bezüglich der alten Basis bezeichnet. Man erkennt, daß sich nicht beide Summen (über j, k) als Matrixprodukt von Spalten $(x), (y)$ mit der quadratischen Matrix (f_{jk}) ansehen lassen — eigentlich klar, man kann ja auch Bilinearformen ϕ_1, ϕ_2 **nicht** komponieren wie Endomorphismen $(L_1 \circ L_2)$. Schließlich, setzen wir die neuen Basisvektoren $w_l = \sum_j a^j{}_l v_j,\ \ w_m = \sum_k a^k{}_m v_k$ in ϕ ein, so finden wir:

$$(\phi)_{neu} = \phi(w_l, w_m) = \textstyle\sum_{j,k} a^j{}_l a^k{}_m \phi(v^j, v^k),$$

d.h. zum Umrechnen der alten Matrix in die neue wird **zweimal** die Matrix A benötigt. Daher heißen die Elemente von $\mathrm{Bil}(V)$ *zweifach kovariant*, und wer das hervorheben will, schreibt die Matrix mit zwei unteren Indices.
Ich hoffe, das überzeugt Sie davon, daß die Verwendung von hoch- und tief-gestellten Indices geeignet ist, vor Fehlern bei Basiswechseln zu warnen.

ANALYSIS. Besprechung von Aufg. 3.8. – Zusammenfassende Wiederholung und Beendigung des Beweises, daß komplexe Polynome in Linearfaktoren zerfallen. Uns fehlt noch: P sei ein Polynom, $P(a) \neq 0$, dann ist die in einer Umgebung von a differenzierbare Abbildung $1/P$ auch dann surjektiv auf eine Umgebung von $1/P(a)$, wenn die Ableitung $P'(a) = 0$ ist. — In Aufg. 4.7 haben wir die Sache mit den kontrahierenden Abbildungen benutzt, um unter wesentlicher Benutzung von $f'(a) \neq 0$ eine Umkehrabbildung auf einer Scheibe um den Bildpunkt $f(a)$ zu definieren. Diese Umkehrabbildung wird auf Blatt 6 als dehnungsbeschränkt nachgewiesen. Danach liefert dasselbe Argument wie im ersten Semester, daß die Umkehrabbildung differenzierbar ist. — Dies wird zunächst auf $f(z) = z^k$ z.B. bei $z = 1$ angewandt, um auf kleinen Umgebungen die k-te Wurzel als **differenzierbare** Funktion zu haben. Betrachte nun

$$1/P(z) - 1/P(a) = (P(a) - P(z))/(P(a) \cdot P(z)) = ((z - a)^k \cdot P_1(z))/(P(a) \cdot P(z))$$

mit k **maximal**, also $P_1(a) \neq 0$. Nach dem eben über z^k gesagten finden wir in einer Umgebung von a eine **differenzierbare** Funktion g mit $g(z)^k = P_1(z))/(P(a) \cdot P(z))$, also

$$1/P(z) - 1/P(a) = \big((z-a)\cdot g(z)\big)^k.$$

Damit ist $z \to 1/P(z)$ surjektiv auf eine Umgebung von $1/P(a)$, denn $z \to (z-a)\cdot g(z)$ hat bei a Ableitung $\neq 0$, ist also nach dem Umkehrsatz surjektiv auf eine Umgebung von 0 und die anschließende k-te Potenz ist explizit als surjektiv auf eine Umgebung von 0 bekannt.

11.5.2000

6.Woche, 15.5.-18.5.2000

ANALYSIS. Wiederholung über *geometrische Majorisierung und kontrahierende Abbildungen* $f: A \to A,\ A$ vollständig, f q-kontrahierend. Fixpunkt $f(a_f) = a_f \in A$ als Grenzwert des Standardverfahrens $a_0 \in A$ beliebig, $a_{n+1} := f(a_n)$, vgl. die Aufgaben über Inverse (4.7, 6.4) und zum Newtonverfahren. Zusätzlich: Hat man zwei q-kontrahierende Abbildungen $f, g : A \to A$ mit bekanntem Supremum-Abstand $\|f-g\|_\infty$, so kontrolliert man auch den Abstand ihrer (nicht **explizit** bekannten) Fixpunkte: $|a_f - a_g| \le \|f-g\|_\infty/(1-q)$. *Riemannsummen und Integrale*: Auch wenn Sie es in Aufg 2.8 vermieden haben, oft hilft es, Summen als Riemannsummen zu interpretieren, z.B. $\sum_{k=3}^n 1/(n\log n)$ ist Riemann(ober)summe von $\int_3^{n+1} 1/(x\log x)dx = \log(\log(x))\big|_3^{n+1} = \log\log(n+1) - \log\log 3$. Die durch das Integral gegebene untere Abschätzung wächst *unbeschränkt*, daher konvergiert die Summe nicht.

Hauptthema: Verallgemeinerung der Konstruktion neuer Funktionen mit Hilfe von Reihen $F_\infty(x) := \sum_k^\infty f_k(x)$ auf analoge Konstruktionen mit Integralen, also

$F(x) := \int_a^b f(t,x)dt$, wobei im Falle $b = \infty$ mit $\int_a^\infty$ gemeint ist: $\lim_{R\to\infty}\int_a^R$.

Wir werden die Sätze am Beispiel der Eulerschen Gammafunktion illustrieren, also der Funktion:

$\Gamma(x) := \int_0^\infty t^{x-1}\cdot e^{-t}\,dt,\ x \ge 1$, um Schwierigkeiten bei $t = 0$ zu vermeiden.

Terminologie. Bei festem x heißt die Funktion $t \to f(t,x)$ partielle Funktion zur ersten Variablen von f, bei festem t heißt $x \to f(t,x)$ partielle Funktion von f zur zweiten Variablen. Die Ableitungen der partiellen Funktionen heißen **partielle Ableitungen** von f nach x bzw. nach t. Bezeichnung: $\frac{\partial}{\partial x}f(t,x),\ \frac{\partial}{\partial t}f(t,x)$. Natürlich gilt der Schrankensatz für die partiellen Funktionen, also z.B.

$$|\tfrac{\partial}{\partial x}f(t,x)| \le L(t) \Rightarrow |f(t,x) - f(t,y)| \le L(t)\cdot|x-y|.$$

Mit der kontinuierlichen Dreicksungleichung ergibt sich sofort eine Dehnungsschranke für die neue Funktion F:

$$|\tfrac{\partial}{\partial x}f(t,x)| \le L(t) \Rightarrow |F(x)-F(y)| = |\int_a^b (f(t,x)-f(t,y))dt| \le \big(\int_a^b L(t)dt\big)\cdot|x-y|.$$

Für die Eulersche Gammafunktion zerlegen wir das Integral in zwei Teile, $\int_0^\infty = \int_0^1 + \int_1^\infty$, und müssen in jedem Teilintervall eine brauchbare Schranke für die partielle Ableitung finden. Wir differenzieren:

$$\tfrac{\partial}{\partial x}(t^{x-1}\cdot e^{-t}) = \tfrac{\partial}{\partial x}(\exp((x-1)\log t\ - t)) = \log t\cdot(t^{x-1}\cdot e^{-t})$$

und finden zunächst für $1 \le t \le \infty,\ 1 \le x \le n$ wegen $0 \le \log t \le t$:

$$|\tfrac{\partial}{\partial x}(t^{x-1}\cdot e^{-t})| \le t^n\cdot e^{-t};$$

und dann für $0 < t \le 1,\ 2 \le x$ wegen $0 \le -\log t = \log(1/t) \le 1/t$:

$$|\frac{\partial}{\partial x}(t^{x-1}\cdot e^{-t})| \leq (t^{x-2}\cdot e^{-t}) \leq 1.$$

Der vorher vorgeführte Beweis liefert also eine Dehnungsschranke für Eulers Funktion:

$$2 \leq x, y \leq n \Rightarrow |\Gamma(x) - \Gamma(y)| \leq (1 + n!)\cdot |x - y|,$$

wobei die 1 vom linken Teilintervall kommt und $n!$ vom rechten. 15.5.

Lin. Algebra. Zu den Grundstrukturen *Gruppe, Körper, Vektorraum* habe ich die **Ringe** besprochen. Zuerst die Liste der Axiome und einige Beispiele: $\mathbb{Z}, \mathbb{Z} \bmod n, K[X]$, $\mathrm{End}(V), \mathbb{Z} + \mathbf{i}\cdot\mathbb{Z}$, Matrizen mit Koeffizienten in einem Ring, Polynome mit Koeffizienten in einem Ring. Wie bei Gruppen haben die Kerne von Ringhomomorphismen $\phi : R \to S$ spezielle Eigenschaften, hier: $r \in R, k \in \ker(\phi) \Rightarrow r \cdot k \in \ker(\phi)$ – daher liegt die $1 \in R$ nur in $\ker(\phi)$ in dem uninteressanten Fall $\phi : R \to 0 \in S$.
Ein wichtiger Ringhomomorphismus ist die Auswertung (für uns $\dim(V) < \infty$):

$$W_A : K[X] \to \mathrm{End}(V), \quad P \mapsto W_A(P) := P(A) \in \mathrm{End}(V).$$

Wir wollen zeigen, daß der Kern von W_A aus den Vielfachen eines Polynoms, aus *den Vielfachen des Minimalpolynoms von A* besteht. Außerdem sollen möglichst viele Analogien zwischen $\mathbb{Z}$ und $K[X]$ hervorgehoben werden.

Wiederholung wie $\mathbb{Q}$ aus $\mathbb{Z}$ konstruiert wird, und *wörtlich ebenso* der Körper der Polynomquotienten aus dem Ring der Polynome: Auf der Menge der Paare $\{(P, Q);\ P, Q \in K[X], Q \neq 0\}$ werden
Äquivalenzklassen definiert durch $(P_1, Q_1) \sim (P_2, Q_2)$ genau dann wenn $P_1Q_2 = P_2Q_1$,
"repräsentantenweise" eine Addition: $(P_1, Q_1) + (P_2, Q_2) := (P_1Q_2 + P_2Q_1,\ Q_1Q_2)$,
und eine invertierbare Multiplikation: $(P_1, Q_1)\cdot(P_2, Q_2) := (P_1P_2,\ Q_1Q_2)$.

Weiterhin ist die Teilbarkeitslehre in $\mathbb{Z}$, $\mathbb{Z}+\mathbf{i}\cdot\mathbb{Z}$ und in $K[X]$ analog. Definiere in den ersten beiden Ringen die "Größe" eines Elementes durch den üblichen Betrag und die "Größe" eines Polynoms durch seinen Grad. Dann gibt es eine *Division mit Rest*:

Zu g_1, g_2 aus einem dieser Ringe gibt es q_2, g_3 aus demselben Ring so daß $g_1 = q_2 \cdot g_2 + g_3$, mit g_3 **kleiner** als g_2. Offenbar ist *jeder* gemeinsame Teiler von g_1, g_2 auch Teiler von g_3.

Seit Euklid wird dies Verfahren wiederholt (mit $g_2, g_3 \ldots$), bis es nach endlich vielen Schritten mit $g_n \neq 0, g_{n+1} = 0$ abbricht. Dann ist jeder gemeinsame Teiler von g_1, g_2 auch Teiler von g_n, aber auch umgekehrt: g_n ist Teiler von g_{n-1}, daher von g_{n-2} usw., schließlich von g_1, g_2. Damit ist g_n größter gemeinsamer Teiler von g_1, g_2, geschrieben g.g.T(g_1, g_2)=g_n. Euklid's Verfahren liefert aber auch, daß g_n *Linearkombination* von g_1, g_2 ist, zunächst $g_n = g_{n-2} - q_{n-1}\cdot g_{n-1}$, dann durch Einsetzen von $g_{n-1} = g_{n-3} - q_{n-2}\cdot g_{n-2}$ aus der vorhergehenden Zeile, usw. und schließlich $g_n = r_1g_1 + r_2g_2$.
Dies kann angewandt werden auf je zwei Polynome $P, Q \in \ker(W_A)$; da jede Linearkombination dieser Polynome in $\ker(W_A)$ liegt, so auch g.g.T(P, Q). Da die Dimension von $K[X]$ größer als die von $\mathrm{End}(V)$ ist, ist $\ker(W_A) \neq 0$, enthält also ein Polynom niedrigsten

Grades — und nach dem Bewiesenen sind alle anderen im Kern Vielfache von diesem.

16.5.

ANALYSIS. Überblick über Ziele und Vorgehensweise der Analysis, dabei handelt es sich vor allem um Untersuchung von Funktionen. Geschrieben habe ich das schon oft genug; ich hoffe, es hat genützt, noch einmal darüber geredet zu haben.

Abschluß der Diskussion *parameterabhängiger Integrale* $F(x) := \int_a^b f(t,x)dt$ (nicht zu verwechseln mit Stammfunktionen $H(x) := \int_a^x h(t)dt$). Aus Voraussetzungen über f sind Behauptungen über F zu folgern. Wiederholung vom 15.5. (Dehnungsschranken):

$$|\tfrac{\partial}{\partial x} f(t,x)| \leq L(t) \Rightarrow |f(t,y) - f(t,x)| \leq L(t)|y-x| \Rightarrow |F(y) - F(x)| \leq \int_a^b L(t)dt \cdot |y-x|.$$

Weniger quantitativ: Voraussetzung $(t,x) \to f(t,x)$ sei stetig und $[a,b]$ beschränkt. Nach dem Hauptsatz über *gleichmäßige Stetigkeit* ist dann f für $(t,x) \in [a,b] \times [x_0 - r, x_0 + r]$ gleichmäßig stetig, und daraus folgt die **Stetigkeit** von F: Zu $\epsilon > 0$ gibt es $\delta > 0$ so daß $|x - x_0| \leq \delta \Rightarrow |f(t,x) - f(t,x_0)| \leq \epsilon$ (für jedes t). Integration liefert $|F(x) - F(x_0)| = |\int_a^b (f(t,x) - f(t,x_0))dt| \leq \epsilon|b-a|$, was zu zeigen war.

Die **Differenzierbarkeit** von F hat ebenfalls eine quantitative Version, Voraussetzung:

$$|f(t,x) - f(t,x_0) - \tfrac{\partial}{\partial x} f(t,x_0)(x-x_0)| \leq B(t)|x-x_0|^2,$$

sowie eine qualitative Version, Voraussetzung:

$[a,b]$ beschränkt und $(t,x) \mapsto \frac{\partial}{\partial x} f(t,x)$ sei stetig, also gleichmäßig stetig.

Integration der **ersten** Voraussetzung und Einsetzen der Definition von F ergibt

$$\Big| F(x) - F(x_0) - \big(\int_a^b \tfrac{\partial}{\partial x} f(t,x_0)dt\big) \cdot (x-x_0)\Big| \leq \int_a^b B(t)dt \cdot |x-x_0|^2,$$

und diese Abschätzung sagt, daß die *Ableitung von F durch Differenzieren unter dem Integral* berechnet werden kann, oder mit anderen Worten: Im Grenzwert des Differenzenquotienten können Limes und Integral vertauscht werden, $\lim_{x\to x_0} \int_a^b \ldots = \int_a^b \lim_{x \to x_0} \ldots$.

Ist statt der quantitativen ersten Voraussetzung die qualitative **zweite** Voraussetzung bekannt, so liefert zunächst der Hauptsatz der Differential- und Integralrechnung (Voraussetzung: stetiger Integrand):

$$f(t,x) - f(t,x_0) - \tfrac{\partial}{\partial x} f(t,x_0)(x-x_0) \;=\; \int_{x_0}^x (\tfrac{\partial}{\partial x} f(t,x) - \tfrac{\partial}{\partial x} f(t,x_0))\, dx,$$

danach ergibt die gleichmäßige Stetigkeit von $\frac{\partial}{\partial x} f$ zu jedem $\epsilon > 0$ ein $\delta > 0$ so daß

$$|x - x_0| \leq \delta \Rightarrow |\tfrac{\partial}{\partial x} f(t,x) - \tfrac{\partial}{\partial x} f(t,x_0)| \leq \epsilon,$$

also mit der vorhergehenden Zeile

$$|f(t,x) - f(t,x_0) - \tfrac{\partial}{\partial x} f(t,x_0)(x-x_0)| \;=\; \int_{x_0}^x |\tfrac{\partial}{\partial x} f(t,x) - \tfrac{\partial}{\partial x} f(t,x_0)|\, dx \leq \epsilon \cdot |x-x_0|,$$

und mit der Definition von F und Integration über t die Behauptung:

$$|x - x_0| \leq \delta \Rightarrow |F(x) - F(x_0) - \int_a^b \tfrac{\partial}{\partial x} f(t,x_0)\; dt \cdot (x-x_0)| \leq \epsilon|b-a| \cdot |x-x_0|.$$

Nun wenden wir uns höherdimensionalen Abbildungen $F: \mathbb{R}^d \to \mathbb{R}^d$ zu. Eine wichtige Interpretation solcher Abbildungen ist, sie als **Geschwindigkeitsfelder** oder **Vektorfelder** zu veranschaulichen. Ich habe als Beispiele die Oberfläche des Rheins und eine dreidimensionale Drehbewegung beschrieben. Formaler drücken wir uns so aus: An jeder Stelle $x \in \mathbb{R}^d$ wird der Wert $F(x)$ interpretiert als Vektor, als Differenzvektor von x nach $x + F(x)$.

18.5.

7.Woche, 22.5.-25.5.2000

ANALYSIS. Weitere Erläuterungen, um anschauliche Vorstellungen zum Begriff *Vektorfeld* zu erzeugen. Speziell *lineare Vektorfelder*, beschrieben durch lineare Abbildungen: $p \mapsto A \cdot p$. Eine zentrale Frage ist: Kann man z.B. aus den (als gemessen vorgestellten) Windgeschwindigkeiten die Bahnkurven der Luftteilchen rekonstruieren? Für lineare Vektorfelder wird dies durch die Exponentialreihe für Matrizen gelöst:

$$\text{Definition: } \exp(t \cdot A) := \lim_{n\to\infty} (\mathrm{id} + \sum_{k=1}^{\infty} \frac{t^k A^k}{k!}), \qquad \text{und: } |t - t_0| \le 1 \Rightarrow$$

$$\| \exp(t \cdot A) - \exp(t_0 \cdot A) - A \cdot \exp(t_0 \cdot A)(t - t_0) \| \le \frac{\|A\|^2}{2} \exp(\|A\|(|t_0| + 1))|t - t_0|^2.$$

Dies bedeutet, daß man die Exponentialreihe für Matrizen, $\exp(t \cdot A)$, wie bei der eindimensionalen Exponentialfunktion nach t differenzieren kann. Die Einzelheiten der Rechunung sind im Manuskript II, S.16, aufgeschrieben. Strategie: Definitionen einsetzen; Dreiecksungleichung und Produktabschätzung von Normen: $\| \sum A^k \| \le \sum \|A^k\| \le \sum \|A\|^k$; Schluß wie in den Beweisen in $\mathbb{R}$. Da die beiden ersten Schritte leicht zu lernen sind, hat man also für diese Beweise über Matrixfunktionen *genau dieselben technischen Probleme wie in* $\mathbb{R}$. Beachten Sie bitte, wie oft es vorkommt, daß in neuer Umgebung die alten Argumente funktionieren! (*Tip*: Zweimaliger Schrankensatz für Potenzen:

$$|t - t_0| \le r \;\Rightarrow\; |t^k - t_0^k - k \cdot t_0^{k-1}(t - t_0)| \;\le\; k(k-1)/2 \cdot (|t_0| + r)^{k-2} \cdot |t - t_0|^2.)$$

22.5.

LIN. ALGEBRA. Wiederholung zum Minimalpolynom als Anwendung des ggT in $K[X]$. Wiederholung spezieller Endomorphismen bei Auswahl eines Skalarproduktes: symmetrisch, schiefsymmetrisch , orthogonal. Ziele von Normalformen: einfache Berechnung von Matrixfunktionen wie $\det(A), A^{-1}, \exp(A)$. Früher erledigt: Zu symmetrischen $A \in \mathrm{End}(V)$ gibt es eine Basis aus Eigenvektoren, die zugehörige Matrix von A ist diagonal.
Normalform schiefsymmetrischer $A \in \mathrm{End}(V)$: Wegen $\langle Av, v\rangle = -\langle v, Av\rangle$ sind Eigenwerte notwendig $= 0$. Zuerst wird V zerlegt in die invarianten Teilräume $U := \ker(A), U^\perp$. Auf $U^\perp$ ist A injektiv; offenbar ist A^2 symmetrisch und ebenfalls injektiv. Daher besitzt $U^\perp$ eine Eigenbasis $\{f_1, \ldots, f_k\}$ zu A^2 mit Eigenwerten $-(\mu_j)^2 < 0$. Nun ist $U_j := \mathrm{span}\{f_j, Af_j\}$ A-invariant wegen $A\{f_j, Af_j\} = \{Af_j, -(\mu_j)^2 f_j\}$ und $v_1 := f_j/|f_j|, v_2 := Av_1/\mu_j$ ist eine ON-Basis bezüglich der A die Matrix: $\begin{pmatrix} 0 & -\mu_j \\ \mu_j & 0 \end{pmatrix}$ hat. Die gewünschte Normalform von einer Matrix für A in $U^\perp$ besteht aus lauter solchen Zweierblöcken längs der Hauptdiagonale.
Beachte: *Es ist bei dieser Diskussion wesentlich, daß zu jedem invarianten Unterraum U ein* **komplementärer invarianter Unterraum** *existiert, nämlich* $U^\perp$.
Daher habe ich auf $\mathbb{R}^2$, mit Standardbasis, durch Matrizen Endomorphismen mit *eindimensionalem* Eigenraum U angegeben, die **keinen** zu U komplementären Unterraum besitzen: Ein solcher müßte eindimensional sein, also Eigenraum, aber es gibt keinen zweiten Eigenwert. Die Matrizen: $\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$, $\begin{pmatrix} a & 1 \\ 0 & a \end{pmatrix}$, mit charakteristischen Polynomen: $\chi_1(X) = X^2$, $\chi_2(X) = (a - X)^2$.

Was für Normalformen kann man im Licht solcher Beispiele wünschen? — Sei B eine Matrix mit Nullen auf der Hauptdiagonalen und darunter; dann hat B^2 auch die erste obere Paralleldiagonale voller Nullen, usw bis $B^d = 0$. Daher kann man für Matrizen $A := \lambda\cdot\mathrm{id}+B$ alle Potenzen, und damit auch Potenzreihen ausrechnen! Wir werden Normalformen finden, die längs der Diagonale solche Blöcke haben und sonst Nullen. 23.5.

Analysis. Noch ein Versuch, Vektorfelder anschaulich zu machen: Betrachten Sie eine Landkarte; die Punkte haben zwar *Koordinaten* in $\mathbb{R}^2$, trotzdem macht Köln+Köln oder 2.5·Köln nicht viel Sinn. Betrachten Sie zusätzlich die *Geschwindigkeiten* von Wanderern, ebenfalls mit $\mathbb{R}^2$ beschrieben; aber hier beschreibt die Addition in $\mathbb{R}^2$ die Vektoraddition der Geschwindigkeiten in der beabsichtigten Weise. Vor diesem Hintergrund soll man die Definition verstehen: *Ein Vektorfeld auf dem Punktraum $\mathbb{R}^d$ ist eine Abbildung in den Vektorraum $\mathbb{R}^d$ mit der Interpretation: der Wert $V(p)$ wird interpretiert als Vektor mit Fußpunkt p.* Oder: Geschwindigkeiten in Köln sind Vektoren mit Fußpunkt Köln.
Wiederholung: Bahnkurven linearer Vektorfelder $p \mapsto A\cdot p$ werden durch die Exponentialreihe gegeben: $t \mapsto \exp(t\cdot A)\cdot p$. Computerbilder dazu mit verschiedenen Matrizen A; erinnern Sie sich vor allem an die Spiralen und die "Sattelbilder".
Die Exponentialreihe liefert zu jedem Startpunkt p **eine** Lösungskurve — gibt es weitere? Da mit zwei Lösungskurven $t \mapsto p(t), q(t)$ mit $\dot p = A\cdot p,\ \dot q = A\cdot q$ auch $r := p - q$ Lösung ist, betrachten wir die Hilfsfunktion

$$h(t) := \langle r(t), r(t)\rangle \cdot \exp(-2\|A\|\cdot t).$$

Behauptung: $\dot h \le 0$, also $h(t) \le h(0)$ oder $|r(t)| \le |r(0)|\cdot\exp(+\|A\|\cdot t)$.
Bemerkung: Diese Ungleichung zeigt nicht nur die gewünschte Eindeutigkeit, sie sagt auch, daß Fehler höchstens exponentiell (genauer: wie $\exp(+\|A\|\cdot t)$ wachsen.

Beweis. Die Produktregeln für $\cdot$ und $\langle\ ,\ \rangle$ liefern:

$$\dot h(t) = \big(2\langle \dot r(t), r(t)\rangle - 2\|A\|\langle(r(t), r(t)\rangle\big)\cdot\exp(-2\|A\|\cdot t);$$

Einsetzen von $\dot r(t) = A\cdot r(t)$ und Cauchy-Schwarz ($|\langle v,w\rangle| \le |v|\cdot|w|$) ergeben in der großen Klammer 0, also $\dot h \le 0$.

Wir haben dies Resultat in eine andere Sprache übersetzt, indem wir die Matrixgleichung $\dot p = A\cdot p$ (im dreidimensionalen Fall) ausgeschrieben haben als

$$\begin{aligned}\dot p_1(t) &= a_{11}\cdot p_1(t) + a_{12}\cdot p_2(t) + a_{13}\cdot p_3(t)\\ \dot p_2(t) &= a_{21}\cdot p_1(t) + a_{22}\cdot p_2(t) + a_{23}\cdot p_3(t)\,.\\ \dot p_3(t) &= a_{31}\cdot p_1(t) + a_{32}\cdot p_2(t) + a_{33}\cdot p_3(t)\end{aligned}$$

Solche Gleichungen tragen bei anderen Gelegenheiten den Namen:
Lineare Differentialgleichungssyteme erster Ordnung. Die letzte Behauptung ist daher auch ein Eindeutigkeits- und Fehlerkontroll-Satz für solche Differentialgleichungen. Und unsere Ableitungsregel $(t \mapsto \exp(t\cdot A))^\bullet = A\cdot\exp(t\cdot A)$ besagt, daß wir mit Hilfe der Exponentialreihe Lösungen für alle solchen Differentialgleichungen finden können. Daher suchen wir in der Linearen Algebra *Normalformen* für A, für die die Berechnung von $\exp(t\cdot A)$ leicht ist.

Damit ist der Fall *linearer Vektorfelder* abgehandelt. Allgemeinere Vektorfelder sind Abbildungen $V : \mathbb{R}^d \to \mathbb{R}^d$, wieder mit der *Interpretation*, daß wir die Argumente als Punkte p, die Werte als Vektoren $V(p) \in \mathbb{R}^d$ mit Fußpunkt p ansehen. Das Beispiel von Windgeschwindigkeiten zeigt deutlich, daß solche Vektorfelder sich im Laufe der Zeit ändern können; tatsächlich werden die folgenden Überlegungen nicht komplizierter, wenn wir solche *zeitabhängigen Vektorfelder* $V : \mathbb{R}^d \times \mathbb{R} \to \mathbb{R}^d,\ (p,t) \mapsto V(p,t) \in \mathbb{R}^d$ betrachten. Nun fragen wir wieder nach Kurven $t \mapsto p(t)$, deren Ableitung $\dot{p}(t)$ zu jedem Zeitpunkt durch das Vektorfeld gegeben ist: $\dot{p}(t) = V(p(t),t)$. Diese heißen *Lösungskurven, Bahnkurven oder Integralkurven des Vektorfeldes V (oder der zu dem Vektorfeld gehörenden Differentialgleichnung).*
Übersichtliche Aussagen für diese Lösungskurven gelten nur unter geeigneten
Voraussetzungen: Erstens muß $(p,t) \mapsto V(p,t)$ stetig sein; und zweitens muß für jedes einzelne t gelten, daß die Vektorfelder $p \mapsto V(p,t)$ sich nicht viel schlimmer als lineare Vektorfelder ändern, meist ausgedrückt als sogenannte

"Lipschitzbedingung": $|V(p,t) - V(q,t)| \leq L \cdot |p - q|$,

also durch die Voraussetzung einer Dehnungsschranke für jedes $V(.,t)$.
Für deren Lösungskurven gilt wieder ein Eindeutigkeits- und Fehlerkontroll-Satz:
Behauptung Sind $t \mapsto p(t), q(t)$ zwei Lösungen, so ist die Hilfsfunktion
$h(t) := \langle p(t) - q(t), p(t) - q(t)\rangle \cdot \exp(-2L \cdot t)$ fallend: $\dot{h} \leq 0$, also $h(t) \leq h(0)$, also

$$|p(t) - q(t)| \leq |p(0) - q(0)| \cdot \exp(+L \cdot t).$$

Der Beweis ist im Manuskript II, S.18, aufgeschrieben; er hat *dasselbe Muster* wie für lineare Vektorfelder.
Das eindimensionale Vektorfeld $V(x) := 2\mathrm{sign}\,(x)\sqrt{|x|}$, das bei $x = 0$ keine Lipschitzschranke hat, besitzt die Lösungen $x(t) := \pm(t-a)^2$, aber auch die Lösung $x(t) := 0$. Zum Anfangswert 0 gibt es also viele verschiedene Lösungen. 25.5.

8.Woche, 29.5.-30.5.2000

ANALYSIS. Der Satz von Picard-Lindelöf wurde bewiesen; er behauptet die Existenz von Lösungskurven zu Vektorfeldern, die eine Dehnungsschranke besitzen ("Lipschitz-Bedingung"), vgl. das Protokoll vom 25.5. für die Erklärung des Problems und Manuskript II, S.20-21 für die Beweise. Wendet man den Hauptsatz der Differential- und Integralrechnung auf eine (noch nicht nachgewiesene) Lösungskurve $t \to p(t)$ (also mit $\dot{p}(t) = V(p(t),t)$) an, so findet man: $p(t) = p(0) + \int_0^t V(p(\tau),\tau)d\tau$. Die entscheidende Beobachtung ist, daß man die rechte Seite auch dann als interessant ansehen kann, wenn man **nicht** von einer Lösungskurve sondern von einer **beliebigen stetigen** Kurve ausgeht.
Definition der Lindelöf-Abbildung $\mathfrak{L} : p(.) \to \mathfrak{L}p(.)$ stetiger Kurven auf differenzierbare Kurven:

$$\mathfrak{L} : p(.) \to \mathfrak{L}p(.), \quad \mathfrak{L}p(t) := p_0 + \int_0^t V(p(\tau),\tau)d\tau, \quad (\mathfrak{L}p)^\bullet(t) = V(p(t),t).$$

Daher sind *Fixpunkte* dieser Lindelöf-Abbildung Lösungen unseres Problems, nämlich In-

tegralkurven des Vektorfeldes V — und für das Finden von Fixpunkten haben wir schon ein Verfahren: *Kontrahierende Abbildungen iterieren.*
Die wichtigste Voraussetzung für dies Verfahren liefert der Hauptsatz über gleichmäßig konvergente Folgen stetiger Abbildungen: *Die Grenzfunktionen sind wieder stetig.* Mit Hilfe der Supremumsnorm, ($\|p(.)\| := \sup_{t\in[0,T]} |p(t)|$), hat dieser Satz die eingängigere Formulierung (hoffentlich!): $\|.\|$-*Cauchy-Folgen stetiger Abbildungen* **besitzen** *stetige Limes-Abbildungen.* Zu Beginn des Semesters habe ich Normen erklärt, insbesondere, daß äquivalente Normen (also: $\frac{1}{C}\cdot|\ \ |\le|\ \ |_1\le C\cdot|\ \ |$) *dieselben* Cauchyfolgen haben. Nun sind zwar Fixpunkte einer Abbildung Fixpunkte, ohne daß von einer Norm die Rede ist, aber die Kontraktionseigenschaften einer Abbildung hängen natürlich von der verwendeten Norm ab. Da wir für die Bahnkurven unseres Vektorfeldes schon bewiesen haben, wie Fehler zum Zeitpunkt $t=0$ im Laufe der Zeit wachsen, nämlich:

$$\dot p(t)=V(p(t),t),\ \dot q(t)=V(q(t),t)\quad\Rightarrow\quad |p(t)-q(t)|\le|p(0)-q(0)|\cdot e^{+L\cdot t},$$

wollen wir das in der Wahl der Norm berücksichtigen, indem wir die Größe der Werte zu späterer Zeit mit geringerem Gewicht berücksichtigen und definieren:

$$\big|\big|p(.)\big|\big| := \sup_{t\in[0,T]}\big(|p(t)|\cdot\exp(-L\cdot t)\big),\qquad \exp(-L\cdot t)\|.\|\le\big|\big|.\big|\big|\le\|.\|.$$

Mit dieser *äquivalenten* Norm ist der Kontraktionsbeweis für $\boldsymbol{\mathfrak{L}}$ leicht:

$$\begin{aligned}|\boldsymbol{\mathfrak{L}}p(t)-\boldsymbol{\mathfrak{L}}q(t)|e^{-Lt} &\le e^{-Lt}\int_0^t|V(p(\tau),\tau)-V(p(\tau),\tau)|d\tau\\ &\le e^{-Lt}\int_0^t L|p(\tau)-q(\tau)|e^{-Lt}e^{+Lt}d\tau\le(1-e^{LT})\cdot\big|\big|p(.)-q(.)\big|\big|\end{aligned}$$

also: $$\big|\big|\boldsymbol{\mathfrak{L}}p(.)-\boldsymbol{\mathfrak{L}}q(.)\big|\big|\le(1-e^{LT})\cdot\big|\big|p(.)-q(.)\big|\big|.$$

29.5.

Lin. Algebra. Erste Hälfte Wiederholung: Ziele, die mit der Herstellung von Normalformen (für die Matrizen von Endomorphismen) erreicht werden sollen; bereits erreichte Normalformen für Endomorphismen, die in Bezug auf eine Skalarprodukt speziell sind: symmetrisch oder schiefsymmetrisch oder orthogonal. Hier verwenden wir *immer Orthonormalbasen* und
— zusätzlich Eigenbasen für symmetrische $A\in\mathrm{End}(V)$ (liefern Diagonalmatrizen für A),
— oder bei schiefsymmetrischen $A\in\mathrm{End}(V)$ Eigenbasen für A^2 (bei Wahl in Paaren $v_1,v_2:=Av_1/|Av_1|$ entstehen schiefsymmetrische Zweierblöcke längs der Hauptdiagonalen),
— oder für orthogonale $A\in\mathrm{End}(V)$ werden erst zu den Eigenwerten (nur ±1) Eigenvektoren gewählt, dann werden auf dem orthogonalen Komplement (der Eigenvektoren von A) mit Hilfe von Eigenvektoren von $A+A^{-1}$ (symmetrisch!) in der Matrix für A Zweierblöcke aus Drehmatrizen $\left(\begin{smallmatrix}\cos\alpha & -\sin\alpha\\ \sin\alpha & \cos\alpha\end{smallmatrix}\right)$ längs der Hauptdiagonale erzeugt.
Was kann man aus Matrizen, die längs der Hauptdiagonale Blöcke $\lambda_j\cdot\mathrm{id}+N$ mit $N_{kj}\ne0$ nur für $j>k$ (oberhalb der Hauptdiagonale) haben und sonst 0 sind, was kann man aus solchen Matrizen ablesen? Erstens, zu den k Blöcken gehören invariante Unterräume $U_j:=\ker(A-\lambda_j\cdot\mathrm{id})^{m_j}$ der Dimension m_j und für jedes $v\in V$ gilt: $v=\sum_{j=1}^k v_j,\ v_j\in U_j$. Zweitens, in jedem U_j liefert der Span der ersten Basisvektoren eine aufsteigende Folge

invarianter Unterräume (jeder enthalten in $\ker(A - \lambda_j \cdot \mathrm{id})^\ell,\ 1 \le \ell \le m_j$), die aber im allgemeinen **keine** invarianten Komplementärräume (wie in erstens) besitzen.
Wenn man also derartige Normalformen herstellen will, dann muß man diese invarianten Unterräume finden.

Von jetzt an gilt die **Voraussetzung**: Das charakteristische Polynom von A zerfällt in Linearfaktoren: $\chi_A(X) = (\lambda_1 - X)^{m_1} \cdot \ldots \cdot (\lambda_k - X)^{m_k},\ \lambda_j \neq \lambda_\ell$.

Induktionsvoraussetzung: In jedem Vektorraum der Dimension $d - 1 \ge 0$ findet man zu jedem Endomorphismus mit zerfallendem charakteristischen Polynom eine Basis, so daß die Matrix obere Dreiecksgestalt hat (natürlich mit den Eigenwerten auf der Hauptdiagonalen).
Der Induktionsanfang bei $d = 1$ ist klar.
Der Induktionsschritt beginnt mit der Wahl eines Eigenvektors u_1 zu einem der Eigenwerte $\lambda_{\ldots}$ als erstem Basisvektor und zunächst einer beliebig erweiterten Basis $\{v_2, \ldots, v_d\}$ für V.
In der zu dieser Basis gehörenden Matrix für A gilt $a_{21} = 0, \ldots, a_{d1} = 0$ und daher:

$$\chi_A(X) = (\lambda_{\ldots} - X) \cdot \det(\ (a_{kj})_{k,j=2\ldots d} - X \cdot \mathrm{id}\ )$$

Der zweite Faktor zerfällt nach Voraussetzung in Linearfaktoren, aber wir wissen noch nicht, von welchem Endomorphismus dies das charakteristische Polynom ist.
Zu dem invarianten Unterraum $U = \mathrm{span}\{u_1\}$ betrachten wir den
Quotientenraum: $W := \{[v] := v + U;\ v \in V\}$ auch: $=: V/U$
und darin den *Quotienten-Endomorphismus*: $A_W : W \to W,\ A_W([v]) := [A(v)]$, — beachte: Da $AU \subset U$ ist A_W wohldefiniert.
Als Basis für W verwenden wir die Äquivalenzklassen $[v_2], \ldots, [v_d]$ der Vektoren der verwendeten Basis von V.
Aus $A \cdot v_j := a^1{}_j u_1 + \sum_{k=2}^d a^k{}_j \cdot v_k$ folgt $A_W([v_j]) := [A \cdot v_j] = \sum_{k=2}^d a^k{}_j \cdot [v_k]$ (mit Äquivalenzklassen **repräsentantenweise** rechnen!). Daher hat A_W die oben schon betrachtete "Restmatrix" $(a_{kj})_{k,j=2\ldots d}$ als Matrix bezüglich der Basis $\{[v_2], \ldots, [v_d]\}$, also zerfällt das charakteristische Polynom von A_W in Linearfaktoren. Nach Induktionsvoraussetzung gibt es eine Basis $\{[u_2], \ldots, [u_d]\}$ für W so daß die Matrix von A_W obere Dreiecksgestalt hat. Wähle Repräsentanten $u_j \in [u_j]$, dann ist $\{u_1, u_2, \ldots, u_d\}$ eine Basis für V, in der A obere Dreiecksgestalt hat.
Wählt man in dieser Konstruktion die m_1 Eigenwerte λ_1 zuerst, so sieht man an der Dreiecksgestalt, daß $U_1 := \mathrm{span}(u_1, \ldots, u_{m_1}) \subset \ker(A - \lambda_1)^{m_1}$ ein invarianter Unterraum (*sog. verallgemeinerter Eigenraum*) der Dimension m_1 ist. Dies geht für jedes λ_j ebenso.
Die Vereinigung der Basen der verallgemeinerten Eigenräume ist eine Basis von V in der die Matrix von A die gewünschte Blockgestalt hat, **denn**: Hat man eine Linearkombination des Nullvektors aus diesen Vektoren, so läßt man aus dem charakteristischen Polynom $\chi_A(X) = (\lambda_1 - X)^{m_1} \cdot \ldots \cdot (\lambda_k - X)^{m_k}$ z.B. den Faktor $(\lambda_j - X)^{m_j}$ weg; in das Restpolynom wird A eingesetzt und auf die Linearkombination angewandt: Alle Vektoren in den $U_\ell, \ell \neq j$ werden anulliert, aber U_j wird bijektiv auf sich abgebildet; insbesondere bleiben die Basisvektoren linear unabhängig und deren Koeffizienten müssen also verschwinden;

und dies für alle j. 30.5.

9.Woche, 5.6.-8.6.2000

LIN. ALGEBRA Vertretung durch Juan Souto. Wiederholung der Blockstruktur, die als Normalform von Endomorphismen erreicht werden kann, falls ihr charakteristisches Polynom in Linearfaktoren zerfällt.
Bei der Diskussion von Basiswechseln haben wir gesehen, daß die Matrizen von Bilinearformen sich anders verhalten als die von Endomorphismen. Dementsprechend anders sind die Normalformen für symmetrische Bilinearformen. Die Normalformen wurden nur für den Körper $\mathbb{R}$ hergestellt. Dazu wurde ein Skalarprodukt zu Hilfe genommen, die Bilinearform b durch einen symmetrischen Endomorphismus A repräsentiert, $b(v,w) = b(w,v) = \langle Av, w\rangle = \langle v, Aw\rangle$, die Diagonalisierbarkeit von A benutzt, durch Längenänderung der Basisvektoren noch etwas vereinfacht und schließlich gezeigt, warum die Anzahl der Nullen (Dimension des Kerns), die Anzahl der Einsen (maximale Dimension eines positiv definiten Unterraums) und die Anzahl der Minus-Einsen (maximale Dimension eines negativ definiten Unterraums) **nicht** von dem zu Hilfe genommenen Skalarprodukt abhängt. (Es genügt zu zeigen: Sind $\{v_1,\ldots,v_r\}$ linear unabhängig und $b(v_j,v_j) = +1, b(v_j,v_k) = 0$ für$j \neq k$, und sind ebenso $\{w_1,\ldots,w_s\}$ linear unabhängig und $b(w_j,w_j) = -1, b(w_j,w_k) = 0$ für$j \neq k$, so sind auch **alle** diese Vektoren gemeinsam linear unabhängig.) 5.6.

ANALYSIS. Diskussion der Definition der Differenzierbarkeit von mehrdimensionalen Funktionen: $f : \mathbb{R}^d \to \mathbb{R}$:
Definition: Zu jedem $\epsilon > 0$ gibt es eine $\delta > 0$ so daß Differenzen von Funktionswerten, $f(x+h)-f(x)$, durch eine lineare Abbildung, $h \mapsto l_x(h)$, gut approximiert werden können, d.h.

$$|h| \leq \delta \Rightarrow |f(x+h) - f(x) \ - \ l_x(h)| \leq \epsilon |h|.$$

1.) Quadratische Funktionen $x \to f(x) := \langle v, x\rangle \cdot \langle w, x\rangle$ oder $x \to f(x) := \langle Ax, x\rangle$ können ohne Benutzung von Grenzwerten sogar mit noch kleineren (nämlich quadratischen) Fehlern differenziert werden, vgl. Manuskript II, S.22.
2.) Funktionen, die aus eindimensionalen Funktionen zusammengesetzt sind, können ebenfalls ohne Benutzung weiterer Sätze differenziert werden, Beispiel:
$x = (x_1, x_2),\ h = (h_1, h_2) \in \mathbb{R}^2,\ f(x) = f(x_1, x_2) := g(x_1) \cdot h(x_2)$, z.B. $g = \sin,\ h = \exp$.
$f(x+h) - f(x) = g(x_1+h_1) \cdot (h(x_2+h_2) - h(x_2)) + (g(x_1+h_1) - g(x_1)) \cdot h(x_2)$.
Nun benutze eindimensionale Kenntnisse, z.B. mit $|g''|,\ |h''| \leq 2 \cdot B$ gilt ja
$|g(s) - g(a) - g'(a) \cdot (s-a)| \leq B|s-a|^2$, und damit folgt
$|f(x+h) - f(x) - (g'(x_1)h(x_2)h_1 + g(x_1)h'(x_2)h_2| \leq \text{const} \cdot |h|^2$.
Man sieht, daß man in diesem Beispiel die Ableitung von f mit Hilfe der uns schon bekannten Ableitungen der partiellen Funktionen $x_1 \mapsto f(x_1,x_2),\ x_2 \mapsto f(x_1,x_2)$, also mit den sogenannten *partiellen Ableitungen* $\frac{\partial}{\partial x_1} f,\ \frac{\partial}{\partial x_2} f$, berechnen kann.

Das ist **nicht** immer so. Betrachte wie in Manuskript II die Funktion: $f(x) = f(x_1, x_2) := x_1 \cdot x_2^3/(x_1^2 + x_2^4)$ für $x \neq 0$ und $f(0) := 0$. Wegen $|f(x)| \leq |x_2|/2$ ist f stetig bei 0. Längs **jeder** Geraden durch 0, $c(t) := t \cdot (\alpha, \beta)$, hat die Komposition $t \to f \circ c(t)$ bei $t = 0$ die Ableitung 0. Andererseits ist $\tilde{c}(t) := (t^2, t)$ eine Kurve, die bei $t = 0$ **dieselbe Ableitung** wie die vertikale Gerade $t \cdot (0, 1)$ hat, während die Komposition, wegen $f \circ \tilde{c}(t) = t/2$, die Ableitung $1/2 \neq 0$ hat. Mit anderen Worten: *Würden wir Differenzierbarkeit nur mit Hilfe der Existenz der partiellen Ableitungen definieren, so würde für die Funktion f die Kettenregel* **nicht** *gelten.* Obwohl wir bald viele Ableitungen mit Hilfe partieller Ableitungen bestimmen werden (und zwar gern, weil die partiellen Ableitungen einfach sind), so ist doch die am Anfang gegebene Differenzierbarkeitsdefinition als Konzept überlegen.

6.6

ANALYSIS Wiederholung vom 6.6. und weitere Diskussion der höherdimensionalen Differenzierbarkeit. Von den vielen vorkommenden Bezeichnungen werde ich nur zwei benutzen: Für $f : \mathbb{R}^d \to \mathbb{R}^e,\ e = 1$ wird die Ableitung ohne Kollisionen an anderen Stellen mit df bezeichnet, also $|h| \leq \delta \Rightarrow \big|f(a+h) - f(a)\ -\ df|_a(h)\big| \leq \epsilon|h|$; für $e > 1$ sind mir die Kollisionen zu unangenehm, ich verwende die Bezeichnung, die am wenigsten kollidiert und der ich die besten Zukunftschancen gebe (zu jedem $\epsilon > 0$ gibt es ein $\delta > 0$ so daß):

$$f : \mathbb{R}^d \to \mathbb{R}^e,\ e > 1,\quad |h| \leq \delta \Rightarrow \big|f(a+h) - f(a)\ -\ Tf|_a(h)\big| \leq \epsilon|h|.$$

Beweis der Differentiationsregel: $T(\alpha_1 \cdot f_1 + \alpha_2 \cdot f_2) = \alpha_1 \cdot Tf_1 + \alpha_2 \cdot Tf_2$ mit Hilfe der Dreiecksungleichung, wie im ersten Semester ("$f \mapsto Tf$ ist linear").
Diskussion: *Wie könnte die Kettenregel lauten*? $c : [\alpha, \omega] \to \mathbb{R}^d, T(f \circ c)|_t = Tf|_{c(t)} \cdot c'(t)$.
Propaganda für die Kettenregel: *Es folgt ein guter Schrankensatz* :
Voraussetzung: c, f seien stetig differenzierbar, für die Ableitung von f sei längs c eine Schranke bekannt, $\|Tf_c\| \leq L$. Dann gilt (zunächst mit dem Hauptsatz der Differential- und Integralrechnung, danach mit der kontinuierlichen Dreiecksungleichung ($t_1 \leq t_2$):

$$|f(c(t_2)) - f(c(t_1))| = |g(t_2) - g(t_1)| = \Big|\int_{t_1}^{t_2} g'(t)dt\Big| \leq \int_{t_1}^{t_2} |g'(t)|dt \leq \int_{t_1}^{t_2} \|Tf|_{c(t)}\| \cdot |c'(t)|dt$$
$$\leq L \cdot \int_{t_1}^{t_2} \cdot |c'(t)|dt = L \cdot \text{Länge}(c).$$

Jetzt wird das letzte Beispiel vom 6.6. noch einmal diskutiert: Da eine brauchbare Differentialrechnung auf keinen Fall auf die Kettenregel verzichten kann, muß die Differenzierbarkeitsdefinition entsprechend gut sein, *es genügt nicht für die Differenzierbarkeit bei a, wenn $f \circ c$ für alle Geraden c durch den Punkt a differenzierbar ist.*

Schließlich wurde die Veranschaulichung dieser Ableitungen im Falle $f : \mathbb{R}^2 \to \mathbb{R}^1$ besprochen. Lineare Abbildungen $l : \mathbb{R}^2 \to \mathbb{R}^1$ werden durch die Höhenline null ($\ker l$) und eine weitere veranschaulicht. Falls nun auch f (wie auf Landkarten) durch ein Höhenlinienbild veranschaulicht ist, so gilt für die Ableitung $df|_a$ an der Stelle a: $\ker df|_a$ ist die Tangente an die Höhenlinie von f durch a, und der Abstand nahe benachbarter Höhenlinien von f und df_a ist ungefähr gleich.
Aus der Linearen Algebra kennen wir eine andere Beschreibung einer linearen Abbildung $l,\ l(h) = \langle v, h\rangle,\ v \in \mathbb{R}^d$. Wendet man dies auf $df|_a$ an, so heißt der repräsentierende Vek-

tor **Gradient** von f, also $df|_a(h) = \langle \operatorname{grad} f(a), h \rangle$. Nun benötigt man keine Höhenlinien von f, um einen Eindruck von den Ableitungen zu bekommen: *Man kann sogar ohne Kenntnisse über f das Vektorfeld $a \mapsto \operatorname{grad} f(a)$ zeichnen und sich damit die Ableitung veranschaulichen.* Die ϵ-δ-Ungleichung lautet dann:

$$f : \mathbb{R}^d \to \mathbb{R}^e,\ e = 1,\quad |h| \le \delta \Rightarrow \big| f(a+h) - f(a) \ - \ \langle \operatorname{grad} f(a), h \rangle \big| \le \epsilon |h|.$$

Fragen zu den gezeichneten Bildern wurden beantwortet. 8.6.

10.Woche, 19.6.-20.6.2000

ANALYSIS. Wiederholung der Differenzierbarkeitsdefinition, Formulierung mit *stetigem Fehlerterm* statt ϵ-δ-Kontrolle:

$F : \mathbb{R}^d \to \mathbb{R}^e$ heißt differenzierbar bei $a \in \mathbb{R}^d$ mit Ableitung $TF|_a$, falls es eine bei a **stetige** Funktion $x \mapsto \varphi_a(x),\ \varphi_a(a) = 0$, gibt mit

$$F(x) = F(a) + TF|_a \cdot (x-a) + \varphi_a(x) \cdot (x-a).$$

Komplex differenzierbare $f : \mathbb{C} \to \mathbb{C}$ sind auch in diesem reell 2-dimensionalen Sinn differenzierbar, weil $\mathbb{C}$ ein 2-dimensionaler $\mathbb{R}$-Vektorraum ist und weil Multiplikation mit $c \in \mathbb{C}$ eine $\mathbb{R}$-lineare Abbildung ist: $c \cdot (r_1 \cdot z_1 + r_2 \cdot z_2) = r_1 \cdot (c \cdot z_1) + r_2 \cdot (c \cdot z_2)$, also $Tf|a = f'(a)\bullet$.

Wir hatten am Beispiel $f(x) = f(x_1, x_2) := x_1 \cdot x_2^3/(x_1^2 + x_2^4)$ für $x \neq 0$ und $f(0) := 0$ gesehen, daß aus der Existenz der partiellen Ableitungen **nicht** folgt, daß bei Komposition von Funktionen nach der *Kettenregel* differenziert werden kann. Umgekehrt, falls die Kettenregel gilt, kann die Ableitung von $F : \mathbb{R}^d \to \mathbb{R}^e$ mit Hilfe der 1-dimensionalen "partiellen" Ableitungen $\frac{\partial}{\partial x^j} F(a)$ berechnet werden:

mit $g_j(t) := a + t \cdot e_j,\ j = 1 \dots d$ bezeichnen wir die Geraden durch a parallel zu den Koordinatenachsen. Dann ist $t \mapsto F \circ g_j(t) = F(a_1, \dots, a_j + t, \dots a_d)$ einerseits eine partielle Funktion und 1-dimensionales Differenzieren ergibt die partielle Ableitung $\frac{\partial}{\partial x^j} F(a)$; andererseits ergibt Differentiation nach der Kettenregel: $(F \circ g_j)'(0) = TF|_a \cdot e_j$.

Daher sind die Bilder der Basisvektoren, $e_j \mapsto TF|_a \cdot e_j$, als partielle Ableitungen erkannt, und die Matrix von $TF|_a$ ist die sogenannte **Jacobi-Matrix***, deren Spaltenvektoren die partiellen Ableitungen $\frac{\partial}{\partial x^j} F(a)$ sind.*

Der *Beweis der Kettenregel* ist in Manuskript II, S.25 mit ϵ-δ-Kontrolle aufgeschrieben und wurde in der Vorlesung mit obigem stetigen Fehlerterm vorgeführt, weil dann besonders deutlich wird, daß man nur Definitionen ineinander einzusetzen hat: $X = f(x), A = f(a)$ wird in die Differenzierbarkeitsdefinition

$F(X) = F(A) + TF|_A \cdot (X - A) + \phi_A \cdot (X - A),\ \phi_A$ stetig bei $X = A, \phi_A(A) = 0$

der äußeren Funktion eingesetzt. Dann wird die Differenzierbarkeit der inneren Funktion benutzt, um $f(x) - f(a)$ durch $Tf|a \cdot (x-a) + \varphi_a(x) \cdot (x-a)$ zu ersetzen. Sortieren ergibt

$$F(f(x)) = F(f(a)) + TF|f(a) \cdot Tf_a \cdot (x-a) + [\dots] \cdot (x-a)$$

wobei $[\dots]$ eine bei a stetige Funktion mit Wert 0 bei a bezeichnet, nämlich:

$[\dots] = TF|f(a) \cdot \varphi_a(x) + \phi_A(f(x) \cdot (Tf_a + \varphi_a(x)).$

Damit ist die Kettenregel bewiesen.

Damit sind die Grundlagen der mehrdimensionalen Differenzierbarkeit behandelt. Wir brauchen nur noch einen Satz, mit dem man den *partiellen Ableitungen* von F ansehen kann, was die Ableitung von F ist. Dieser Satz enthält zwar nicht die Worte "genau dann wenn", aber abgesehen davon ist er wirklich gut:

Satz. Wenn die partiellen Ableitungen $\frac{\partial}{\partial x^j}F(x)$ für x nahe a existieren und bei $x = a$ **stetig** sind, dann ist F bei a differenzierbar und die Matrix $\big(TF|_a\big)$ ist die Jacobimatrix aus den partiellen Ableitungen. 19.6.

Analysis. Wiederholung zur Veranschaulichung zwei- und dreidimensionaler Funktionen und ihrer Ableitungen. Wichtig ist, die Teilmengen des Definitionsbereichs anzusehen, auf denen die Funktion *konstant* ist: Höhenlinien (= Niveaulinien), Niveauflächen. Lineare Funktionen $x \mapsto l(x) = \langle v, x\rangle$ können durch Angabe von $\ker(l)$ und eine weitere (parallele) Höhenlinie mitgeteilt werden *oder* auch durch Angabe des Vektors v; wendet man dies auf die Ableitung einer Funktion $f : \mathbb{R}^d \to \mathbb{R}^1$ an, also $Tf|_a(h) = \langle \operatorname{grad} f(a), h\rangle$, so wird man die Ableitungsfunktion $a \mapsto \operatorname{grad} f(a)$ als *Vektorfeld* veranschaulichen.

Ausführlicher Beweis des Ableitungskriteriums mit Hilfe partieller Ableitungen (letzter Satz vom 19.6.), wie S.27 Manuskript II. Wichtigster Punkt ist,
aus der Stetigkeit: $|x - a| \le \delta \Rightarrow \frac{\partial}{\partial x^j}f(a) - \epsilon \le \frac{\partial}{\partial x^j}f(x) \le \frac{\partial}{\partial x^j}f(a) + \epsilon$
und dem Schrankensatz für $t \to f(b^1, b^2, \ldots, b^{j-1}, a^j + t, a^{j+1}, \ldots, a^d)$
(mit $b^k - a^k = h^k,\ k = 1 \ldots d,\ h = \sum_{k=1}^d h^k \cdot e_k, |h| \le \delta$) zu folgern:

$$\big|f(b^1, b^2, \ldots, b^j, a^{j+1}, \ldots, a^d) - f(b^1, b^2, \ldots, b^{j-1}, a^j, \ldots, a^d) - \tfrac{\partial}{\partial x^j}f(a) \cdot h^j\big| \le \epsilon|h^j|$$

Summation ergibt (unter Berücksichtigung von $\sum |h^j| = \langle(1, \ldots, 1),\ (|h^1|, \ldots, |h^d|)\rangle \le \sqrt{d} \cdot |h|$):

$$|b - a| \le \delta \Rightarrow \big|f(b) - f(a) - \textstyle\sum_{j=1}^d \frac{\partial}{\partial x^j}f(a) \cdot h^j\big| \le \epsilon\sqrt{d} \cdot |h|.$$

Dies zeigt, daß f differenzierbar und die Matrix von $Tf|_a$ die Jacobi-Matrix der partiellen Ableitungen ist.

So wichtig diese Berechnung von Ableitungen ist, so ist sie doch nicht immer angebracht: Betrachte $\mathbb{R}^{d\cdot d}$ als $\operatorname{Hom}(\mathbb{R}^d, \mathbb{R}^d)$, realisiert durch $d \times d$-Matrizen, und dazu $F : \mathbb{R}^{d\cdot d} \to \mathbb{R}^{d\cdot d},\ F(X) := X \cdot X$. Ohne darüber nachzudenken, in welcher Reihenfolge man jetzt die d^2 partiellen Ableitungen aufzählen soll, folgert man sofort aus $X = A + (X - A)$:

$$|X^2 - A^2 - A \cdot (X - A) - (X - A) \cdot A| = |(X - A)^2|.$$

Da $|(X - A)^2|$ ein quadratischer Fehler ist, und $H \mapsto A \cdot H + H \cdot A$ linear, so hat man: $TF|_A(H) = A \cdot H + H \cdot A$ gezeigt.

Ich habe begonnen, die Aufgaben 4.7, 6.4, 7.5 zum Umkehrsatz für Abbildungen $f : \mathbb{C} \to \mathbb{C}$ mit $f(0) = 0, |z| \le r \Rightarrow |f'(z) - 1| \le q < 1$ vorzuführen. 1. Schritt: Für $|w| \le (1 - q)r$ ist $F_w : \{|z| \le r\} \to \{|z| \le r\},\ F_w(z) := z - (f(z) - w)$ q-kontrahierend und man findet ein Urbild $z_*(w)$ von w als Fixpunkt von F_w. 2.Schritt: Wegen $|F_{w_2}(z) - F_{w_1}(z)| \le |w_2 - w_1|$ folgt wie S.21 Manuskript II: $|z_*(w_2) - z_*(w_1)| \le 1/(1 - q)|w_2 - w_1|$, also die Lipschitzstetigkeit der Umkehrabbildung $w \mapsto z_*(w)$. 20.6.

11.Woche, 26.6.-29.6.2000

ZWEI- UND DREIDIMENSIONALE GEOMETRIE. Wegen vieler Zeichnungen kann ich diese Stunde nicht gut protokollieren. Ich habe erklärt, welche 2- und 3-dimensionalen orthogonalen (= linearen und isometrischen) Abbildungen es gibt, wie man die Basen wählen muß, um einfache Matrizen zu bekommen, und wie man die verschiedenen Sorten allein mit Determinante und Spur unterscheidet. Dieses Wissen kann viel Arbeit sparen. – Den (linear) ausgezeichneten Nullpunkt kann man durch Konjugation mit Translationen, $T_a : x \mapsto x + a$, irgendwo hinlegen: Ist D eine Drehung um 0, so dreht $T_a \circ D \circ T_{-a}$ um a. – Die Charakterisierung der Mittelsenkrechten ("gleich weit von P, Q entfernte Punkte") liefert einen Beweis zum Schnittpunkt der Mittelsenkrechten im Dreieck, der auch auf der Sphäre funktioniert. Da zwei sich schneidende Geraden *zwei* Winkelhalbierende haben, muß man präzisieren: Genau die Punkte des Sektors zwischen zwei Halbgraden, die gleich weit von diesen entfernt sind, liegen auf der Winkelhalbierenden. Und darum schneiden sich auch die Winkelhalbierenden sphärischer Dreiecke. – Die Höhen eines euklidischen Dreiecks sind Mittelsenkrechte eines viermal so großen Dreiecks und schneiden sich deshalb; in sphärischen Dreiecken schneiden sich zwar die Höhen, aber dieser Beweis versagt. – In euklidischen Dreiecken A, B, C mit Gewichten m_1, m_2, m_3 an den Ecken schneiden sich die Schwerelinien (von $S_{AB} = (m_1 \cdot A + m_2 \cdot B)/(m_1 + m_2)$ nach C, m_3 usw.) in dem sogenannten Schwerpunkt $S := (m_1 A + m_2 B + m_3 C)/(m_1 + m_2 + m_3)$. Dieser Satz der Linearen Algebra gilt für sphärische Dreiecke nicht.

Sphärische Dreiecke beschreiben einen wichtigen Teil der 3-dimensionalen euklidischen Geometrie. Den Punkt $C \in \mathbb{S}^2$ können wir durch eine Spiegelung in den Nordpol $N = (1\ 0\ 0)$ legen, Eigenvektor zu -1 der Spiegelung ist $N - C$, also ist die Spiegelung:

$$x \mapsto x - \langle x, N - C \rangle \cdot (N - C) \cdot |N - C|^{-2}.$$

Einen zweiten Punkt B drehen wir um die z-Achse in die positive x-z-Halbebene, dann hat er die Koordinaten $B = (\sin a,\ 0,\ \cos a)$. Ich habe Teile der Aufgabe 11.5 vorgeführt, die zeigt, daß keine andere Kurve (außer dem Großkreis von C nach B) kürzer als a ist; daher heißt a der sphärische Abstand von C nach B; die Dreiecksungleichung gilt wegen 11.5. – Ein dritter Punkt A mit Abstand b von C ist dann durch den Winkel γ zwischen den Großkreisbögen CB und CA festgelegt, seine Koordinaten:

$$A = (\sin b \cdot \cos\gamma,\ \sin b \cdot \sin\gamma,\ \cos b).$$

Da wir die Punkte C, B durch eine Spiegelung Sp vertauschen können und da der Punkt $Sp \cdot A$ den Abstand c vom Nordpol hat und der Bogen $NSp \cdot A$ den Winkel β mit der x-z-Ebene bildet, gilt:

$$Sp \cdot A = \begin{pmatrix} \sin c \cdot \cos\beta \\ \sin c \cdot \sin\beta \\ \cos c \end{pmatrix} = \begin{pmatrix} -\cos a & 0 & \sin a \\ 0 & 1 & 0 \\ \sin a & 0 & \cos a \end{pmatrix} \cdot \begin{pmatrix} \sin b \cdot \cos\gamma \\ \sin b \cdot \sin\gamma \\ \cos b \end{pmatrix}.$$

Ausmultiplizieren gibt drei Sätze für sphärische Dreiecke, die in ähnlicher Weise Grundlage der sphärischen Geometrie sind, wie die gleichbenannten Formeln der ebenen Geometrie:

$$\begin{pmatrix} \text{Projektionssatz} \\ \text{Sinussatz} \\ \text{Kosinussatz} \end{pmatrix} = \begin{pmatrix} \sin c \cdot \cos\beta \\ \sin c \cdot \sin\beta \\ \cos c \end{pmatrix} = \begin{pmatrix} -\cos a \cdot \sin b \cdot \cos\gamma + \sin a \cdot \cos b \\ \sin b \cdot \sin\gamma \\ \cos a \cdot \cos b + \sin a \cdot \sin b \cdot \cos\gamma \end{pmatrix}.$$

26.6.

UMKEHRSATZ-FESTIVAL Besprechung und Beweis des Umkehrsatzes.
Wiederholungen:
1. Für lineare Abbildungen $A \in \mathrm{Hom}(\mathbb{R}^d, \mathbb{R}^d)$ hat man die drei äquivalenten Bedingungen für die Existenz einer Umkehrabbildung $A^{-1} \in \mathrm{Hom}(\mathbb{R}^d, \mathbb{R}^d)$:
(i) $\mathrm{rang}(A) = d$, (ii) $\ker(A) = \{0\}$, (iii) $\det A \neq 0$.
Mit dieser Umkehrabbildung löst man Gleichungssysteme $A \cdot x = b$, nämlich $x = A^{-1} \cdot b$.
2. Für $f : \mathbb{R}^1 \to \mathbb{R}^1$ bildet eine *stetige und streng monotone* Funktion Intervalle injektiv und surjektiv auf Intervalle ab, so daß man (sogar stetige) Umkehrfunktionen hat. – Ist f zusätzlich differenzierbar mit $f' \neq 0$, so erhält man (vgl. WS) Differenzierbarkeit und Ableitung der Umkehrfunktion, indem man die Differenzierbarkeitsdefinition von f mit $1/f'(a)$ multipliziert und die "zu jedem $\epsilon > 0$ gibt es ein $\delta \ldots$"-Worte in Ordnung bringt, s.u..
3. Für komplex differenzierbare $F : \mathbb{C} \to \mathbb{C}$ mit $F(a) = A, F'(a) \neq 0$ hat man keine Monotonieargumente mehr. In den Aufgaben 4.7, 6.5 wurde geübt, wie man mit dem Satz über kontrahierende Abbildungen (der zu meiner Überraschung im Lauf der Vorlesung noch einmal wiederholt werden mußte) die Umkehrfunktion findet. Der allgemeine Fall folgt diesem Muster. – Es gab eine unerwartete Schwierigkeit: Die meisten Leute sind bemüht, Beweise mit möglichst übersichtlichen Bezeichnungen zu führen; in diesem Fall spielen die Werte $a, F(a) = A, F'(a)$ keine wichtige Rolle, und es ist üblich die Bezeichnungen zu vereinfachen, indem man die geänderte Funktion $\tilde{F}(z) := (F(z+a) - A)/F'(a)$ betrachtet, die die einfacheren Voraussetzungen $\tilde{F}(0) = 0, \tilde{F}'(0) = 1$ erfüllt. Ich habe nicht verstanden, warum die Hörer diese Änderung nicht akzeptieren konnten.
4. Jetzt betrachten wir $F : \mathbb{R}^d \to \mathbb{R}^d$, $F(a) = A$, $\mathrm{rang}(TF|_a) = d$, und Ziel ist, in einer Umgebung von A eine differenzierbare Umkehrfunktion G zu konstruieren, also mit $G \circ F(x) = x$ und damit $TG|_{F(x)} = (TF|_x)^{-1}$. Wir machen für den Beweis wieder die Bezeichnungsvereinfachung: $\tilde{F}(x) := (TF|_a)^{-1} \cdot (F(x+a) - A)$, die die Voraussetzungen $\tilde{F}(0) = 0$, $T\tilde{F}|_0 = \mathrm{id}$ erfüllt. Wir werden in der Nähe von 0 eine Umkehrfunktion $\tilde{G}$ zu $\tilde{F}$ konstruieren – natürlich kann man die Vereinfachungen rückgängig machen, $G(y) := \tilde{G}((TF|_a)^{-1} \cdot (y - A)) + a$ ist eine Umkehrfunktion zu F.

Wir beweisen den Satz in drei Schritten (und lassen im Beweis die ˜ weg), Strategie:
S1: Mit dem Kontraktionssatz wird für alle genügend kleinen y ein F-Urbild konstruiert.
S2: Die Umkehrabbildung G aus S1 ist dehnungsbeschränkt, weil wenig verschiedene kontrahierende Abbildungen auch wenig verschiedene Fixpunkte haben.

S3: Die Differenzierbarkeit der Umkehrabbildung G folgt, indem zuerst die Differenzierbarkeitsdefinition von F mit $(TF|_x)^{-1}$ multipliziert wird und dann mit S2 die ϵ-δ-Argumentation in Ordnung gebracht wird.

Beweis zu S1. Definiere $\Psi_y(x) := x - (F(x) - y)$. Dann ist erstens ein Fixpunkt x^* von Ψ_y ein F-Urbild von y und zweitens hat Ψ_y bei $x = 0$ die Ableitung $T\Psi_y|_{x=0} = 0 \in \mathrm{End}(\mathbb{R}^d)$. Weil F als *stetig* differenzierbar vorausgesetzt ist (sonst geht es nicht), gibt es $r > 0$ so daß $|x| \le r \Rightarrow \|T\Psi_y|_x\| \le 1/2$. Damit verkleinert Ψ schon mal Abstände mindestens um den Faktor 2 (Schrankensatz); wir brauchen noch eine Menge, die in sich abgebildet wird. Offenbar ist $\Psi_y(x) = \Psi_0(x) + y$ und für *vollständige* Kugeln $B_\rho := \{x \mid |x| \le \rho\}$ gilt: $\Psi_0(0) = 0,\ \Psi_0 : B_r \to B_{r/2}$. Daher folgt aus $|y| \le r/2$ das erste Ziel: $\Psi_y : B_r \to y + B_{r/2} \subset B_r,\ |\Psi(x_1) - \Psi(x_2)| \le 1/2 \cdot |x_1 - x_2|$. Der Fixpunkt x^* dieser kontrahierenden Abbildung Ψ_y ist F-Urbild von y und wird daher mit $G(y)$ bezeichnet.

Beweis zu S2. Die beiden 1/2-kontrahierenden Abbildungen $\Psi_{y_1},\ \Psi_{y_2}$ sind wenig verschieden, genauer: $\|\Psi_{y_1} - \Psi_{y_2}\| = |y_1 - y_2|$. Daher folgt für den Abstand der Fixpunkte mit $q = 1/2$: $|G(y_1) - G(y_2)| \le (1/(1-q)) \cdot |y_1 - y_2|$, also die Dehnungsschranke 2 für $G : B_{r/2} \to B_r$. Dieses schon mehrfach benutzte Argument ist im Manuskript II, S.21 bewiesen, zweizeilig; es sollte bekannt sein.

Beweis zu S3. Aus der Differenzierbarkeit von F bei x_0 soll die Differenzierbarkeit von G bei $y_0 = F(x_0)$ gefolgert werden. Multiplikation der Definition (der Differenzierbarkeit von F bei x_0) mit $(TF_{x_o})^{-1}$ ergibt: Zu jedem $\epsilon > 0$ gibt es ein $\delta > 0$ so daß

$$|x - x_0| \le \delta \Rightarrow \left|(TF_{x_o})^{-1} \cdot (F(x) - F(x_0)) - (x - x_0)\right| \le \|(TF_{x_o})^{-1}\| \cdot \epsilon |x - x_0|;$$

hierin wird eingesetzt: $y = F(x), x = G(y), y_0 = F(x_0), x_0 = G(y_0)$; außerdem wird aus S2 benutzt: $|x - x_0| = |G(y) - G(y_0)| \le 2 \cdot |y - y_0|$. Das ergibt:

$$|x - x_0| \le \delta \Rightarrow \left|(TF_{x_o})^{-1} \cdot (y - y_0) - (G(y) - G(y_0))\right| \le \|(TF_{x_o})^{-1}\| \cdot \epsilon \cdot 2|y - y_0|.$$

Wir erkennen das Ziel besser, wenn wir in der Mitte die beiden Terme vertauschen: Wir wären fertig, wenn wir die zweite Ungleichnung nicht aus $|x - x_0| \le \delta$ folgern würden, sondern aus $|y - y_0| \le ?$. Aber S2 liefert $|y - y_0| \le \delta/2 \Rightarrow |x - x_0| \le \delta$, also haben wir erreicht, daß für alle $y, y_0 \in B_{r/2}$ gilt:

$$|y - y_0| \le \delta/2 \Rightarrow \left|(G(y) - G(y_0)) - (TF_{x_o})^{-1} \cdot (y - y_0)\right| \le \|(TF_{x_o})^{-1}\| \cdot \epsilon \cdot 2|y - y_0|.$$

Diese Folgerung besagt gerade: G ist bei y_0 differenzierbar und hat die Ableitung

$$TG|_{y_o} = (TF|_{x_o})^{-1} = (TF|_{G(y_o)})^{-1}. \qquad 27.6.$$

ANALYSIS Wiederholung der Behandlung des Umkehrsatzes unter Berücksichtigung der aufgetretenen Schwierigkeiten, vgl. 27.6. und Manuskript II.

Umkehrsatz und gewöhnliche Differentialgleichungen Gegeben sei $F : \mathbb{R}^d \to \mathbb{R}^d$ mit $F(a) = A$, $\det(TF|_a) \neq 0$ (d.h. die Voraussetzungen des Umkehrsatzes sind erfüllt). Frage: Kann man ein F-Urbild der Kurve $t \mapsto A + t \cdot w$ finden?

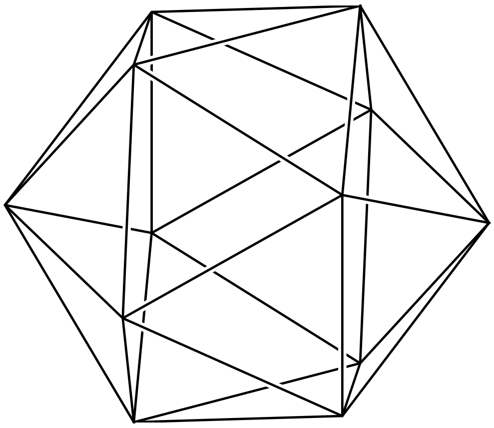
S1. Angenommen, wir hätten eine Kurve $c : [0, T] \to \mathbb{R}^d, c(0) = a$ mit $F(c(t)) = A + t \cdot w$. Dann würde die Kettenregel liefern: $TF|_{c(t)} \cdot \dot{c}(t) = w$, oder: $\dot{c}(t) = TF|_{c(t)} \cdot w$. Daher

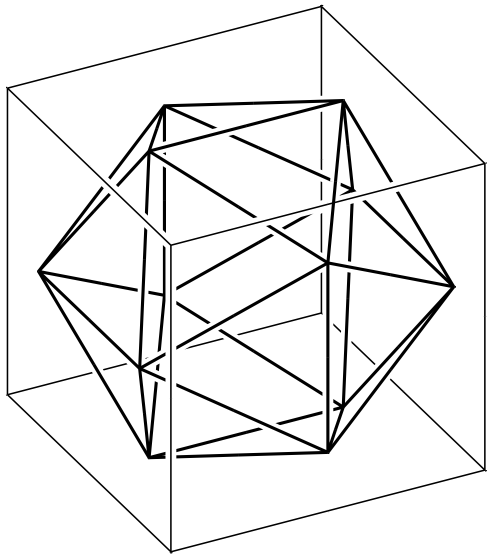
wollen wir c als Lösung einer Differentialgleichung suchen.
S2. Definiere ein Vektorfeld auf einer Umgebung von a durch: $V(p) := (TF|_p)^{-1} \cdot w$ und betrachte Lösungskurven $\gamma : t \mapsto \gamma(t)$, für die also gilt: $\dot\gamma(t) = V(\gamma(t))$. (Unser Existenzsatz liefert solche Lösungen, und sie sind auch schnell numerisch zu berechnen.)
S3. Tatsächlich löst γ unser Problem. Definiere $h(t) := F(\gamma(t)) - (A + t \cdot w)$. Dann ist erstens $h(0) = F(a) - A = 0$ und zweitens $\dot h(t) = TF|_{\gamma(t)} \cdot \dot\gamma(t) - w = w - w = 0$. Also gilt $h(t) = 0$ oder $F(\gamma(t)) = A + t \cdot w$.
Wiederholung linearer Differentialgleichungen mit *konstanten* Koeffizienten $\dot x(t) = A \cdot x(t)$ und warum sie durch die Exponentialreihe gelöst werden, Manuskript II, S.15-17. Es ist wichtig, diese Differentialgleichungen wiederzuerkennen, wenn die Matrizen ausgeschrieben sind, $A = (a_{jk})$. Die Exponentialreihe kann als Lösungsformel aufgefaßt weren, da wir sie für Matrizen in Normalform wirklich berechnen können.

Höhere Ableitungen. Grob betrachtet ist die Ableitung einer Funktion $F : \mathbb{R}^d \to \mathbb{R}^e$ eine Funktion $TF : \mathbb{R}^d \to \mathrm{Hom}(\mathbb{R}^d, \mathbb{R}^e) = \mathbb{R}^{d \cdot e}$, kann also wieder mit derselben Definition wie vorher differenziert werden. Genauer betrachtet ist schon jedes $TF|_x$ eine lineare Abbildung, etwa $v \to TF|_x(v)$. Beim Weiterdifferenzieren gibt die Approximation der Differenz $TF|_{x+h}(v) - TF|_x(v)$ ein weiteres lineares Argument h, so daß die zweite Ableitung, geschrieben $T^2F|_x$, als bilineare Abbildung aufgefaßt wird. Dies verspricht, sehr unübersichtlich zu werden, wird aber gerettet durch den Satz von Schwarz, der sagt, daß diese lineare Abbildung **symmetrisch** in v, h ist – siehe 4.7. 29.6.

12.Woche, 3.7.-6.7.2000

GEOMETRIE Wiederholung der sphärischen Geometrie vom 26.6.; die Formeln gehen beim Grenzübergang zu sehr kurzen Seitenlängen (wegen $\sin(x) \approx x,\ \cos(x) \approx 1 - x^2/2$) über in die Formeln der Euklidischen Dreiecksgeometrie, Aufg. 10.5 .
Behandlung der Platonischen Körper, einerseits vom Euklidischen Standpunkt, also einem Würfel ein- oder umbeschrieben; andererseits vom sphärischen Standpunkt: Projektion des Polyeders auf seine Umkugel gibt eine Pflasterung der Sphäre mit regelmäßigen Dreiecken, Quadraten oder Fünfecken. Da die Winkel sphärischer Dreiecke *größer* sind als die Euklidischer Dreiecke mit denselben Seitenlängen (Aufg. 11.6), haben sphärische Quadrate Winkel $> 90°$, eine Pflasterung entsteht nur, wenn der Winkel $= 120°$ ist, die sphärische Pflasterung also zum Würfel gehört. Ebenso gehört zum Dodekaeder eine Pflasterung mit sphärischen Fünfecken, deren Winkel $= 120°$ betragen, und zu Tetraeder, Oktaeder und Ikosaeder gehören Pflasterungen mit gleichseitigen Dreiecken, deren Winkel $= 120°, = 90°, = 72°$ betragen. Die Kantenlängen dieser Pflasterungen können aus den trigonometrischen Formeln berechnet werden und liefern Konstruktion und Klassifikation der Platonischen Körper.
Beim Ikosaeder und Dodekaeder benötigt man zur Euklidischen Konstruktion den goldenen Schnitt. Dieser wird auch am Rechteck und am Pentagramm (mit Irrationalitätsbeweis) besprochen.





Jedes sphärische Dreieck bestimmt einen Euklidischen Keil (Spitze im Mittelpunkt der Kugel) und umgekehrt; die Normalen auf den Keilflächen liefern das Polardreieck zum ursprünglichen sphärischen Dreieck; sind a, b, c die Seiten und α, β, γ die Winkel des ursprünglichen Deiecks, so sind $\pi-\alpha, \pi-\beta, \pi-\gamma$ die Seiten und $\pi-a, \pi-b, \pi-c$ die Winkel des Polardreiecks. Daher gilt außer dem sphärischen Kosinussatz auch der *Winkelkosinussatz*:

$$\cos\gamma = -\cos\alpha\cdot\cos\beta + \sin\alpha\cdot\sin\beta\cdot\cos c.,$$

der zum Berechnen der Seitenlängen aus den Winkeln gut geeignet ist. 3.7.

ANALYSIS Wiederholung der Differenzierbarkeitsdefinition für $F : \mathbb{R}^d \to \mathbb{R}^e$ und der Berechnung der Ableitung $TF|_a$ mit Hilfe der Jacobimatrix der eindimensionalen partiellen Ableitungen – *falls diese stetig sind.*

Noch einmal, im Falle $e = 1$ wird die Unterscheidung von $TF|_a \in \mathrm{Hom}(\mathbb{R}^d, \mathbb{R})$ und $\mathrm{grad}\, F(a) \in \mathbb{R}^d$ wichtig; der Zusammenhang ist: $TF|_a \cdot v = \langle \mathrm{grad}\, f(a), v\rangle$. Dementsprechen ist die zweite Ableitung der Funktion $F : \mathbb{R}^d \to \mathbb{R}$ ein Endomorphismus, falls der Gradient weiter differenziert wird, Bezeichnung: $T\mathrm{grad}\, F|_a = \mathrm{Hesse}\, F|_a \in \mathrm{End}(\mathbb{R}^d)$; andererseits, wird das Differential $x \mapsto TF|x$ weiterdifferenziert, so ist die zweite Ableitung eine bilineare Abbildung von $\mathbb{R}^d$ nach $\mathbb{R}$, $T^2F|_a \in \mathrm{Bil}(\mathbb{R}^d, \mathbb{R})$. Falls die Matrix dieser Bilinearform aus partiellen Ableitungen berechnet wird, heißt sie ebenfalls Hessematrix von F. Ich empfehle, vor allem wegen der Polarkoordinaten, sich anzugewöhnen, die Hesse-Bilinearform und den Hesse-Endomorphismus in den Bezeichnungen zu unterscheiden, etwa $\mathrm{hesse}\, f\ (v, w) = \langle \mathrm{Hesse}\, F(v),\ w\rangle$. Bei Verwendung von *Orthonormalbasen* stimmen die beiden zugehörigen Matrizen überein, $\mathrm{hesse}\, f\ (v, w) = \sum_{j,k} \frac{\partial}{\partial x^j}\frac{\partial}{\partial k^k} f\ v^j w^k = \langle \mathrm{Hesse}\, F(v),\ w\rangle$.

Satz von H.A.Schwarz. Voraussetzung: $F : \mathbb{R}^d \to \mathbb{R}^e$ sei zweimal *stetig* differenzierbar, insbesondere sind also die ersten und zweiten Ableitungen aus partiellen Ableitungen berechenbar.

Behauptung. Die zweite Ableitung ist symmetrisch: $T^2F|a(v)(w) = T^2F|_a(w)(v)$ oder auch $\frac{\partial}{\partial x^j}\frac{\partial}{\partial x^k}f(a) = \frac{\partial}{\partial x^k}\frac{\partial}{\partial x^j}f(a)$.
Der Beweis ist in Manuskript II, S.28 eine halbe Seite, ich habe ihn nicht gut erzählt.

Anwendung. Jede eindimensionale Funktion $f : \mathbb{R}^1 \to \mathbb{R}^1$ ist eine Ableitung, d.h. besitzt eine Stammfunktion F, $F' = f$. Im Gegensatz dazu gilt für ein sogar stetig differenzierbares Vektorfeld $V : \mathbb{R}^d \to \mathbb{R}^d$, es ist höchstens dann eine Ableitung, also $V = \operatorname{grad} f$, wenn jede Ableitung $TV|_x$ ein *symmetrischer* Endomorphismus ist. Aus diesem Grund ist die Verallgemeinerung des Hauptsatzes der Differential- und Integralrechnung auf höhere Dimensionen schwieriger und wird erst im nächsten Semester behandelt.
Nützlichkeit. Offenbar reduziert der Satz von Schwarz die Anzahl der zu berechnenden partiellen Ableitungen beinahe auf die Hälfte. Tatsächlich ist der Gewinn größer, weil oft die partiellen Ableitungen in einer Reihenfolge leichter zu berechnen sind als in der anderen.

Ich habe die Divergenz von Vektorfeldern definiert, $\operatorname{div} V(a) := \operatorname{spur} TV|_a$ und begonnen, sie anschaulich zu erläutern. 4.7.

Analysis Wiederholung des Satzes von Schwarz, grob gesagt sind die zweiten Ableitungen symmetrisch, weil wir auf zweierlei Weisen haben:

$$(F(x+h+k) - F(x+h)) - (F(x+k) - F(x)) \approx (TF|_{x+h} - TF|_x) \cdot k \approx T^2F|_x(h)(k)$$
$$= (F(x+h+k) - F(x+k)) - (F(x+h) - F(x)) \approx (TF|_{x+k} - TF|_x) \cdot h \approx T^2F|_x(k)(h).$$

Die $\approx$-Zeichen müssen mit dem Schrankensatz und mit der Stetigkeit von T^2F durch Gleichheit bis auf ausreichend kleine Fehlerschranken ersetzt werden. Die genaue Struktur dieser Abschätzung wird Ihnen vermutlich nicht (oft) wieder begegnen.

Die Bedeutung der **Divergenz** von Vektorfeldern. Wiederholung: Vektorfelder (am anschaulichsten als Geschwindigkeitsfelder), Bahnkurven, Lösungen von Differentialgleichungen (explizit nur in speziellen Fällen). Lineare Vektorfelder $V(p) := A \cdot p$ haben mit der Exponentialreihe angebbare Lösungsformeln: $c(t) = \exp(t \cdot A) \cdot c(0)$. Für einen festen Zeitschritt t transportiert die *lineare* Abbildung $p \mapsto \exp(t \cdot A) \cdot p$ jeden Punkt p auf der Bahnkurve durch p um den Zeitschritt t weiter. Das Volumenverhältnis zwischen einem Startwürfel und der Bildfigur zum Zeitpunkt t ist

$$v(t) := \det(\exp(t \cdot A)).$$

Satz: $$\det(\exp(t \cdot A)) = \exp(t \cdot \operatorname{spur} A).$$

Vor dem Beweis dieses Satzes einiges zu seiner Bedeutung. Erstens ist offenbar die rechte Seite *viel* leichter auszurechnen als die linke. Zweitens ist für lineare Vektorfelder: $V(p) := A \cdot p$ deren Ableitung *konstant*, $TV|_p(h) \approx V(p+h) - V(p) = A \cdot h$, also $TV|_p = A$. Die Divergenz ist als Spur der Ableitung definiert, also $\operatorname{div} V(p) := \operatorname{spur} TV|p = \operatorname{spur} A$. Daher besagt der Satz, daß die Divergenz eines (linearen) Vektorfeldes V die Wachstumsrate $v'(t)/v(t)$ der Volumenexpansion des "Flusse" ist, d.h. der Lösungsabbildung, die jeden Punkt auf seiner Bahnkurve um t vorrückt. — Diese Volumenexpansion eines Vektorfeldes spielt für den Divergenzsatz, einer höherdimensionalen Verallgemeinerung des Hauptsatzes

der Differential- und Integralrechnung, eine wesentliche Rolle. Daher ist die hier an den linearen Vektorfeldern erläuterte Interpretation der Divergenz sehr wichtig.
Beweis des Satzes. Die Behauptung gilt für $t = 0$ und die rechte Seite ist Lösung der Differentialgleichung $f'(t) = \mathrm{spur}\,(A) \cdot f(t)$. Wegen des Eindeutigkeitssatzes für Differentialgleichungen sind wir fertig, wenn wir dieselbe Differentialgleichung für die linke Seite nachweisen. Wir benutzen
$(d/dt)f(t)|_t = (d/d\tau)f(t+\tau)|_{\tau=0}$,
$\exp((t+\tau)A) = \exp((\tau)A) \cdot \exp((t)A)$,
$\det(A \cdot B) = \det(A) \cdot \det(B)$
und erhalten:
$(d/dt)\det(\exp(t \cdot A)) = (d/d\tau)\det(\exp(\tau \cdot A))|_{\tau=0} \cdot \det(\exp(t \cdot A)) = \mathit{const} \cdot \det(\exp(t \cdot A))$.
Damit ist auch die linke Seite eine Exponentialfunktion und wir müssen nur noch die Konstante identifizieren. Das beruht erstens auf der Kettenregel: Ist $t \mapsto c(t)$ eine Kurve mit Tangente $\gamma(t) = c(0) + t \cdot v$, so gilt
$(F \circ c)'(t)|_{t=0} = (F \circ \gamma)'(t)|_{t=0} = TF|_{c(0)} \cdot v$.
Die Tangente der Kurve $\tau \mapsto \exp((\tau)A)$ ist $\tau \mapsto \mathrm{id} + \tau \cdot A$, daher gilt
$\mathit{const} = (d/d\tau)\det(\exp(\tau \cdot A))|_{\tau=0} = (d/d\tau)\det(\mathrm{id} + \tau \cdot A)|_{\tau=0}$.
Und zweitens wird hier beinahe das charakteristische Polynom differenziert, $\chi_A(\lambda) := \det(A - \lambda \cdot A)$, dessen zweithöchster Koeffizient gerade $\mathrm{spur}\,(A) = \sum_{j=1}^{d} a_{jj}$ ist. Daher gilt $\det(\mathrm{id} + \tau \cdot A) = 1 + \tau \cdot \mathrm{spur}\,(A) + \tau^2 \cdot \ldots$. Dies ist leicht bei $\tau = 0$ zu differenzieren und liefert $\mathit{const} = \mathrm{spur}\,(A)$, wie gewünscht.

Die Bestimmung von Extremwerten höherdimensionaler Funktionen wurde begonnen, vgl. Manuskript S.29/30. (Ich werde dies Thema beenden, über $Tf|_a$ und $\mathrm{grad}\,f(a)$ in Polarkoordinaten reden und mindestens die letzte Stunde auf Kurvenintegrale verwenden.)

13.(letzte) Woche, 10.7.-13.7.2000

GEOMETRIE. Auf Archimedes' Grabstein sind eine Halbkugel, der einbeschriebene Kegel und der umbeschriebene Zylinder abgebildet, sowie das Volumenverhältnis $1 : 2 : 3$. Das Verhältnis $1 : 3$ zwischen Kegel und Zylinder wird schon auf der Schule bestimmt. Bohrt man den Kegel von oben aus dem Zylinder heraus, so hat der Restkörper in der Höhe h einen Kreisring mit den Radien $h \le R$ als Querschnitt, dessen Flächeninhalt $\pi \cdot (R^2 - h^2)$ ist derselbe wie bei der Halbkugel, die in der Höhe h einen Kreis vom Radius $\sqrt{R^2 - h^2}$ als Querschnitt hat. So hat Archimedes das Halbkugelvolumen zu $\frac{2}{3}\pi R^3$ bestimmt.
Pflastert man die Kugeloberfläche mit kleinen Dreiecken, so kann man sich die Vollkugel als Vereinigung von Pyramiden der Höhe R über den Dreiecken vorstellen. Daß die Grundfläche nicht eben ist, spielt eine um so kleinere rolle, je kleiner die Dreiecke der Pflasterung sind. Also folgt – natürlich mit dem Archimedes Axiom –
Kugelvolumen = *Oberfläche* $\times$ *Radius/3*, oder *Oberfläche* $= 4\pi R^2$.
Daraus folgt, daß der Flächeninhalt eines sphärischen Dreiecks sich aus den Winkeln berechnen läßt, Behauptung: *Dreiecksfläche* $= (\alpha + \beta + \gamma - \pi) \cdot R^2$.

Beweis. Zunächst ist die Fläche eines Großkreiszweiecks proportional zum Winkel α, also $\alpha/(2\pi)\cdot$*Kugelfläche* $= 2\alpha$. Weiter bestimmt ein sphärisches Dreieck an jeder Ecke zwei Großkreise, die sich unter dem Dreieckswinkel schneiden, also jeweils zwei Großkreiszweiecke mit den Winkeln α, β, γ. Diese sechs Zweiecke überdecken das sphärische Dreieck sowie seine Antipodenkopie *dreimal*, alle übrigen Teile der Sphäre *einmal*. Daher gilt tatsächlich:

$$4 \cdot (\alpha + \beta + \gamma) = 4\pi + 4\cdot \textit{Dreiecksfläche}.$$

Mit diesem Tribut an Archimedes beende ich die sphärische Geometrie.

Perspektive Bilder. Auf Zeichnungen Dürers wird erklärt, wie Punkte P der abzubildenden Gegenstände durch "Sehstrahlen" mit dem Auge Z verbunden werden; die Bildpunkte $\bar{P}$ entstehen als Schnitt der Sehstrahlen mit der Bildebene σ. Auch Photoapparate stellen solche Bilder her. – Bildet man alle Punkte einer Geraden g ab, so liegen die Sehstrahlen in der *Ebene* $(g \vee Z)$, d.h. die Menge der Bildpunkte von g ist die Schnittgerade $(g \vee Z) \cap \sigma$. Bildet man zu g parallele Geraden $g_1, g_2, \ldots$ ab so enthalten die projizierenden Ebenen $(g_j \vee Z), j = 1, 2, \ldots$ alle eine gemeinsame Gerade, nämlich die Parallelverschiebung g_Z von g durch Z. Der Bildpunkt $g_Z \cap \sigma$ liegt daher auf allen Bildgeraden $\bar{g}_j, j = 1, 2, \ldots$ und heißt daher *Fluchtpunkt* der zu g parallelen Geradenschar. Man vergleiche eine Photographie einer geraden Eisenbahnstrecke. Schließlich brauchen wir, daß die Fluchtpunkte aller Geraden aus einer Ebene ϵ wieder auf einer Geraden liegen, der sogenannten Fluchtgeraden der Ebene, nämlich $\epsilon_Z \cap \sigma$. (Die zu Geraden g oder Ebenen ϵ parallel durch Z verschobenen Geraden oder Ebenen werden durch $g_Z, \epsilon_Z, \ldots$ bezeichnet.)

Nach diesen Erläuterungen wurde ein Viereck als perspektives Bild eines Parallelograms betrachtet und mit Hilfe der Fluchtpunkte der Seiten, damit der Fluchtgeraden der Parallelogramebene, damit der Diagonalenfluchtpunkte das *eine* gegebene Bild zum Bild einer Parallelogrampflasterung fortgesetzt. Außerdem wurde die euklidische Konstruktion, (mit Hilfe paralleler Geraden) eine Strecke in äquidistante Stücke zu teilen, unter Benutzung von Fluchtpunkten ins perspektive Bild übertragen; dazu wird erstens benutzt, daß äquidistante Teilungen auf Strecken parallel zu σ auf äquidistant geteilte Bildstrecken abgebildet werden, und zweitens, daß eine Gerade g in einer Ebene ϵ dann parallel zu σ ist, wenn g parallel zur Fluchtgeraden $\epsilon_Z \cap \sigma$ von ϵ ist. – Schließlich soll das gegebene Viereck als perspektives Bild eines horizontalen Quadrates interpretiert werden. Das führt zur vollständigen Rekonstruktion der Daten der perspektiven Abbildung:

Die Fluchtgerade der Quadratebene heißt jetzt *Horizont*. Man denke sich das Zentrum Z zusammen mit wichtigen *horizontalen* Geraden durch Z um den Horizont gedreht bis sie (als Hilfszeichnung, nicht als Teil des perspektiven Bildes) in σ liegen. Insbesondere sieht man alle Winkel dieser umgeklappten Geraden in natürlicher Größe. Mit den Quadratseiten und mit den Quadratdiagonalen haben wir zwei Paare orthogonaler Richtungen, deren *Fluchtpunkte* wir schon kennen. Mit zwei Thaleskreisen wird der umgeklappte Augpunkt $Z_{geklappt}$ rekonstruiert. Damit kennt man auch die Fluchtpunkte von unter 45° in vertikalen Ebenen ansteigenden Geraden und kann das gegebene Quadratbild zum Bild eines Würfels ergänzen. Daran sieht man, daß man nun im Prinzip jeden Punkt abbilden kann, also ist die Rekonstruktion der perspektiven Abbildung (die durch die Interpretation eines gegebenen Vierecks *als Bild eines horizontalen Quadrates* **definiert** *ist*) beendet. Und der Geometrieteil dieser Vorlesung auch. 10.7.2000

ANALYSIS Wiederholung über Extremwertbestimmung für höherdimensionale Funktionen (Manuskript II, S.29); beachten Sie, daß der 1-dimensionale Fall als bekannt vorausgesetzt wird und mit seiner Hilfe *notwendige* Bedingungen hergeleitet werden, die Extrempunkte erfüllen müssen, zunächst $\operatorname{grad} f(p) = 0$.
Wie im 1-dimensionalen gibt es hinreichende Bedingungen für lokale Maxima mit Hilfe der 2. Ableitung. Zunächst lautet längs der Hilfskurven $c(t) = p + \cdot v$ das 1-dimensionale hinreichende Kriterium:
$0 > (f \circ c)''(0) = \langle \operatorname{grad} f(p), c''(0)\rangle + \langle \text{Hesse}\, f|_p \cdot c'(0), c'(0)\rangle = \langle \text{Hesse}\, f|_p \cdot c'(0), c'(0)\rangle.$
Das bedeutet insbesondere, daß die Eigenwerte $\lambda_1, \ldots, \lambda_d$ des symmetrischen Endomorphismus $\text{Hesse}\, f|_p$ alle negativ sein müssen (denn sonst gilt längs Kurven in Richtung von Eigenvektoren nicht einmal das 1-dimensionale hinreichende Kriterium). Sind diese Eigenwerte negativ, so gilt für beliebige Richtungen $v = \sum x_j \cdot, e_j$ mit $\{e_j; j = 1 \ldots d\}$ einer ON-Basis aus Eigenvektoren: $\langle \text{Hesse}\, f|_p \cdot v, v\rangle = \sum_{j=1}^{d} \lambda_j \cdot (x_j)^2 < \max \lambda_j < 0$. Mit anderen Worten: Bei *negativen* Hesse f Eigenwerten ist längs jeder Kurve durch p die 1-dimensionale *hinreichende* Bedingung für ein lokales Maximum erfüllt. Dies kann ausgebaut werden zu dem Rezept: $\operatorname{grad} f|_p = 0$ und *Eigenwerte von* $\text{Hesse}\, f|_p < 0$ impliziert: f hat in P ein *lokales* Maximum.

Die so weit entwickelte Theorie war schon bei unserem ersten Beispiel, nämlich Eigenvektoren eines symmetrischen Endomorphismus A durch Maximieren der quadratischen Funktion $x \mapsto \langle A \cdot x, x\rangle$ *auf der Einheitskugel* zu finden, **nicht anwendbar**. Die Extrempunkte müssen am Rand liegen, und durch solche Punkte kann mann man nicht beliebige Hilfskurven $c(t) = p + T \cdot v$ legen. Das verbleibende Problem heißt: *Extremwertbestimmung unter Nebenbedingungen*, vgl. Manuskript II, S.35. Anderes ausgedrückt: f soll auf einer Teilmenge des Definitionsbereichs maximiert werden. Die meisten Teilmengen des $\mathbb{R}^d$ sind für das folgende Verfahren ungeeignet, aber oft genug treten doch Teilmengen $N := \{x \in \mathbb{R}^d;\ h(x) = const,\ h : \mathbb{R}^d \to \mathbb{R}\}$ auf, sogenannte Niveaumengen von h.
Ich habe mit dem Umkehrsatz vorgeführt, daß kleine Umgebungen von Punkten p auf solchen Niveaumengen dann besonders übersichtlich als Graphen $\{(x, y) \in \mathbb{R}^d \times \mathbb{R};\ y = g(x)\}$ beschrieben werden können, wenn die *Voraussetzung* $\operatorname{grad} h|_p \neq 0$ erfüllt ist. An dieser Beschreibung sieht man, warum die Niveaumengen auch “Flächen” oder “Niveau-Flächen” genannt werden. *Nur solche Teilmengen sind gemeint, wenn von Extrema mit Nebenbedingungen gesprochen wird.*
Nun werden auf dieselbe Weise wie vorher *notwendige* Bedingungen für Extrempunkte p hergeleitet. Unsere Hilfskurven c müssen zwar die Nebenbedingungen erfüllen, d.h. $h(c(t)) = const,\ c(0) = p$, aber dann gilt wieder notwendig:

$$0 = (f \circ c)'(0) = \langle \operatorname{grad} f|_p, c'(0) \rangle.$$

Um zu verstehen, was das für $\operatorname{grad} f|_p$ bedeutet, muß man wissen, was die möglichen Anfangsrichtungen $c'(0)$ unserer Hilfskurven sind. Das wird durch die Beschreibung als Graph geliefert: **Jeder** Vektor $v \perp \operatorname{grad} h|_p$ ist Anfangsvektor einer Kurve c mit $c(0) = p$ und $h(c(t)) = const.$ (Diese Vektoren v heißen übrigens Tangentialvektoren der Niveaufläche in p.) Damit haben wir erreicht, daß sowohl $\operatorname{grad} h|_p$ als auch $\operatorname{grad} f|_p$ **senkrecht** zu der Menge aller Tangentialvektoren in p sind, oder: $\operatorname{grad} h|_p$ und $\operatorname{grad} f|_p$ sind **proportional**. Wir können also formulieren:

Rezept: In einem Extrempunkt p von f auf der Niveaufläche $N := \{x \in \mathbb{R}^d;\ h(x) = const\}$ gelten notwendig die $d+1$ Gleichungen:

$$h(p) = const,\ \operatorname{grad} f|_p = \lambda \cdot \operatorname{grad} h|_p$$

für die $d+1$ Unbekannten $\lambda, (p^1, \ldots, p^d) = p$. Die Extrempunkte müssen also unter den Lösungen dieses Gleichungssystems zu finden sein.

Hierzu ist in Manuskript II, S.35 ein Beispiel vorgerechnet; außerdem sind die Formeln für mehr als eine Nebenbedingung aufgeschrieben.

Hiermit ist die Vorlesung so weit beendet, wie ich den Stoff zum **Vordiplom** rechne. In der letzten Stunde rede ich über Kurvenintegrale und wie der Hauptsatz der Differential- und Integralrechnung anfängt, sich auf höhere Dimensionen zu verallgemeinern. 11.7.

ANALYSIS Ich habe nach Manuskript II, S.36-39.5, über Kurvenintegrale gesprochen. Ich hoffe, die Parametrisierungsunabhängigkeit ist als wichtige Invarianzeigenschaft dieser Integrale deutlich geworden. Die erste Verallgemeinerung des Hauptsatzes ("Stetig differenzierbare Funktionen sind als Integrale ihrer Ableitung rekonstruierbar") lautet:

Für stetig differnzierbare $H : \mathbb{R}^d \to \mathbb{R}$ und für beliebige stetig differenzierbare Kurven
$c : [a, b] \to \mathbb{R}^d$ mit $c(a) = p$ ein fester Startpunkt und $c(b) = x$ ein beliebiger Punkt gilt:

$$H(x) - H(p) = \int_a^b \langle\ \operatorname{grad} H|_{c(t)}, c'(t)\ \rangle dt.$$

Weiter habe ich diskutiert, wie stark derartige Kurvenintegrale vom Weg abhängen. Es sei $X : \mathbb{R}^d \to \mathbb{R}^d$ ein stetig differenzierbares Vektorfeld und $c_0, c_1 : [a, b] \to \mathbb{R}^d$ seien zwei Kurven mit gleichen Endpunkten $c_0(0) = c_1(0) = p,\ c_0(1) = c_1(1) = q$. Weiter wähle eine "Kurvenschar" $c_s(t) := (1-s) \cdot c_0(t) + s \cdot c_1(t),\ 0 \le s \le 1$, die die beiden Kurven c_0, c_1 verbindet. Wir definieren das Kurvenintegral des Vektorfeldes X längs der Scharkurve c_s als Hilfsfunktion:

$$h(s) := \int_a^b \langle\ X \circ c(t),\ \frac{\partial}{\partial t} c_s(t)\ \rangle dt.$$

Wir interessieren uns für die Differenz der beiden Kurvenintegrale $h(1), h(0)$ und schreiben mit dem Hauptsatz: $h(1) - h(0) = \int_0^1 h'(s)\ ds$.

Bei der Behandlung parameterabhängiger Integrale, Manuskript II, S.12-13, haben wir bewiesen, daß h durch Differenzieren des Integranden berechnet werden kann. Produkt- und Kettenregel führen auf

$$h'(s) := \int_a^b \left(\langle\, X \circ c(t),\ \frac{\partial}{\partial s}\frac{\partial}{\partial t} c_s(t)\,\rangle + \langle\, TX|_{c(t)} \cdot \frac{\partial}{\partial s} c_s(t),\ \frac{\partial}{\partial t} c_s(t)\,\rangle \right) dt.$$

Hier stört der erste Term, weil wir den Unterschied der beiden Kurvenintegrale durch *Integration der Ableitung* TX berechnen wollen. Dazu betrachten wir $s \to c_s(t)$ als Kurvenschar mit Scharparameter t und wiederholen die vorhergehenden Argumente. Allerdings wissen wir diesmal, daß $s \to c_s(0) = p$ und $s \to c_s(1) = q$ *konstante* Kurven sind, deren Kurvenintegrale (für jedes Vektorfeld) 0 sind. Wir erhalten also:

$$0 - 0 = \int_{t=a}^{t=b} \left(\int_{s=0}^{s=1} \left(\langle\, X \circ c(t),\ \frac{\partial}{\partial t}\frac{\partial}{\partial s} c_s(t)\,\rangle + \langle\, TX|_{c(t)} \cdot \frac{\partial}{\partial t} c_s(t),\ \frac{\partial}{\partial s} c_s(t)\,\rangle \right) ds \right) dt.$$

Dies wird von $h(1) - h(0) = \int_0^1 h'(s)ds$ abgezogen. Dazu zitieren wir erstens noch einmal die parameterabhängigen Integrale: Wir können die Reihenfolge der Integration über s und t vertauschen; und zweitens zitieren wir den Satz von Schwarz über die Symmetrie der zweiten Ableitungen, hier: $\frac{\partial}{\partial t}\frac{\partial}{\partial s} c_s(t) = \frac{\partial}{\partial s}\frac{\partial}{\partial t} c_s(t)$. Deshalb heben sich die beiden Terme, die diese zweiten Ableitungen (und das störende X) enthalten, heraus und es bleibt als Differenz der Kurvenintegrale übrig:

$$h(1) - h(0) =$$
$$\int_{t=a}^{t=b} \left(\int_{s=0}^{s=1} \left(\langle\, TX|_{c(t)} \cdot \frac{\partial}{\partial s} c_s(t),\ \frac{\partial}{\partial t} c_s(t)\,\rangle - \langle\ \frac{\partial}{\partial s} c_s(t), TX|_{c(t)} \cdot \frac{\partial}{\partial t} c_s(t)\,\rangle \right) ds \right) dt.$$

Dies ist (wie bei Gradientenfeldern offensichtlich) jedenfalls 0, wenn jedes $TX|_{c(t)}$ ein *symmetrisches* Endomorphismenfeld ist; und im allgemeinen wird die Differenz der Kurvenintegrale aus einem *2-dimensionalen Integral* über die Ableitung des Vektorfeldes X berechnet, genauer aus dem Integral über den *schiefsymmetrischen Anteil* der Ableitung TX.
Dies ist ein erster Blick auf eine höherdimensionale Version des Hauptsatzes der Differential- und Integralrechnung!

13.7.