Chapter 5
The so-called vector calculus (in German)

Die sogenannte Vektoranalysis

Wir haben immer wieder Restfunktionen r betrachtet, die die Eigenschaft hatten, daß $\lim_{h\to 0} \frac{r(h)}{\|h\|} = 0$ galt. Dies war gleichwertig zu $r(h) = \|h\| R(h)$ mit einr Funktion R, die in 0 stetig war und $R(0) = 0$ erfüllte. Die von LANDAU für solche Funktionen einheitlich eingeführte Schreibweise ist häufig bequem. Man sagt, eine Funktion wir r sei $\mathrm{o}(\|h\|)$ („klein Oh von h“). Mit dieser Schreibweise drückt sich die Differenzierbarkeit einer Funktion in a wie folgt aus:

$$f(x) = f(a) + f'(a)(x - a) + o(\|x - a\|).$$

Man stelle sich die Übungsaufgabe, Rechenregeln für die o-Funktionen aufzustellen: Die Summe und das Produkt einer o-Funktion ist eine o-Funktion; das Produkt einer beschränkten Funktion mit einer o-Funktion ist eine o-Funktion.

In der allgemeinen Diskussion von Funktionen $f\colon X \to \mathbb{R}^m$ mit $X \subseteq \mathbb{R}^n$ haben wir immer wieder Spezialfälle betrachtet. Beispiel: $n = 1$, m beliebig: Kurven; $m = 1$, n beliebig: Höhenfunktion. Ein weiterer Sonderfall, der gegen Ende immer wieder auftrat war $n = m$ beliebig. Dann ist die Ableitung, falls sie existiert, eine lineare Abbildung $f'(a)$ von $\mathbb{R}^n$ in sich. Die Matrix von $f'(a)$ ist $((\partial_k f_j)(a))_{j,k=1,\dots,n}$ und ist somit quadratisch. Mit solchen Funktionen befaßte sich der Satz von den lokalen Inversen; auch beim Satz von den impliziten Funktionen kamen sie vor: Man kann ja eine Funktion $F\colon X \times Y \to \mathbb{R}^m$ mit $X \subseteq \mathbb{R}^p$, $Y \subseteq \mathbb{R}^m$ als eine ganze Familie von Funktionen $f_x\colon Y \to \mathbb{R}^m$ auffassen, die durch Vektoren $x \in X$ parametrisiert ist, und die f_x sind wieder von dem besagten Typ. Nun erlauben diese Funktionen $f\colon X \to \mathbb{R}^n$, $X \subseteq \mathbb{R}^n$ eine in gewissem Sinne anschauliche Interpretation, die bei vielen Problemen, etwa in der Physik sehr nützlich ist.

Wir betrachten die Menge $X \times \mathbb{R}^n$ als disjunkte Vereinigung der Mengen $T_x = \{x\} \times \mathbb{R}^n$, und jede davon ist ein Vektorraum der zu $\mathbb{R}^n$ isomorph ist, wenn wir z.B. $(x, v) + (x, w) = (x, v + w)$ und $r \cdot (x, v) = (x, r \cdot v)$ setzen. Die Funktion $f: X \to \mathbb{R}^n$ ordnet jedem Punkt $x \in X$ einen Vektor $\big(x, f(x)\big) \in T_x$ zu.

Die Argumente aus der Menge X sind wegen $X \subseteq \mathbb{R}^n$ insbesondere Punkte im $\mathbb{R}^n$, etwa
$x = (x_1, \ldots, x_n)$. Wir können uns den Vektorraum T_x als die Menge $\mathbb{R}^n$ vorstellen, aber mit dem Urspung x. Dies machen wir uns jedesmal dann zunutze machen, wenn wir einen Vektor "am Punkt x anheften und als Pfeil mit Ursprung x zeichnen".

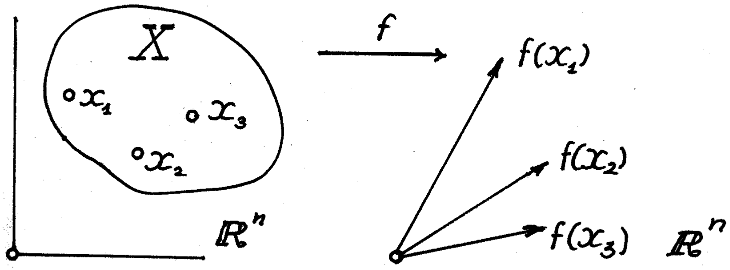

Figure 5.1

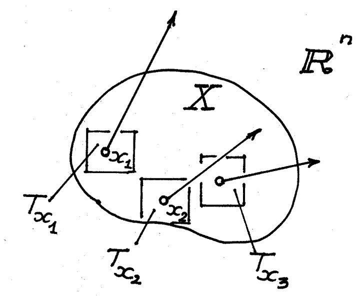

Figure 5.2

Wir können im Bildbereich $\mathbb{R}^n$—der ja mit dem umfassenden Vektorraum des Definitionsbereiches X übereinstimmt—in jedem Punkt $x \in X$ den Bildvektor $f(x)$ „anlegen“, d.h. also, dass wir ihn als Element $\big((x, f(x)\big)$ von T_x betrachten. Man eine solche Funktion $f: X \to \mathbb{R}^n$, $X \subseteq \mathbb{R}^n$ auch ein *Vektorfeld.* Man halte sich also stets vor Augen: Ein Vektorfeld ist nichts anderes als eine besondere Art von Funktion. Es ist darin nichts Geheimnisvolles verborgen. Insbesondere kann man daher ohne weiteres von stetigen, differenzierbaren, stetig differenzierbaren, n-mal differenzierbaren Vektorfeldern sprechen, denn alle diese Begriffe sind ja schon längst für Funktionen in größerer Allgemeinheit diskutiert worden. Also noch einmal: Die Ableitung $f'(a)$ eines Vektorfeldes in einem Punkt a ist eine lineare Selbstabbildung des $\mathbb{R}^n$ und ist daher durch eine quadratische Matrix $((D_k f_j)(a))_{j,k=1,\ldots,n} = \left(\frac{\partial f_j}{\partial x_k}\Big|_{x=a}\right)_{j,k=1,\ldots,n}$ beschrieben, die man auch *Funktionalmatrix* oder auch *Jacobi'sche Matrix* nennt.

Von der linearen Algebra wissen wir, daß man einer linearen Abbildung gewisse Invarianten zuordnen kann, die von der Wahl einer Basis unabhängig sind.
Erstes Beispiel: *Die Determinante.* Diese Invariante können wir sofort auf die Ableitung eines Vektorfeldes anwenden und bekommen die Zahl $\det f'(a) = \det\left[\big((\partial_k f_j)(a)\big)_{j,k=1,\ldots,n}\right]$, die wir auch schon wiederholt erwähnt haben: Sie gibt uns ja Auskunft darüber, ob $f'(a)$ invertierbar ist oder nicht. Das erstere ist ja genau dann der Fall, wenn diese Determinante nicht verschwindet. Sie hat nun auch einen besonderen Namen: $\det f'(a)$ heißt die *Funktionaldeterminante des Vektorfeldes.*

Wir versuchen uns nun die Bedeutung der Funktionaldeterminante anschaulich zu machen. Wir erinnern uns an die Einführung der Determinante in der linearen Algebra. Sind uns n Vektoren $v_1, \ldots, v_n$ in $\mathbb{R}^n$ gegeben, so ist $D(v_1, \ldots, v_n)$ das Volumen des von den v_j aufgespannten Parallelepipeds, und wenn A die Matrix ist, deren Spalten die Koeffizientenvektoren der v_j sind, so schreibt man $\det A = D(v_1, \ldots, v_n)$. Nun sei also $f\colon X \to \mathbb{R}^n$ ein Vektorfeld und a ein innerer Punkt von X, in dem das Vektorfeld differenzierbar sei. Wir betrachten die n Grundvektoren $e_1 = (1, 0, \ldots, 0)$, $e_2, \ldots, e_n$ und eine hinreichend kleine positive reelle Zahl h.

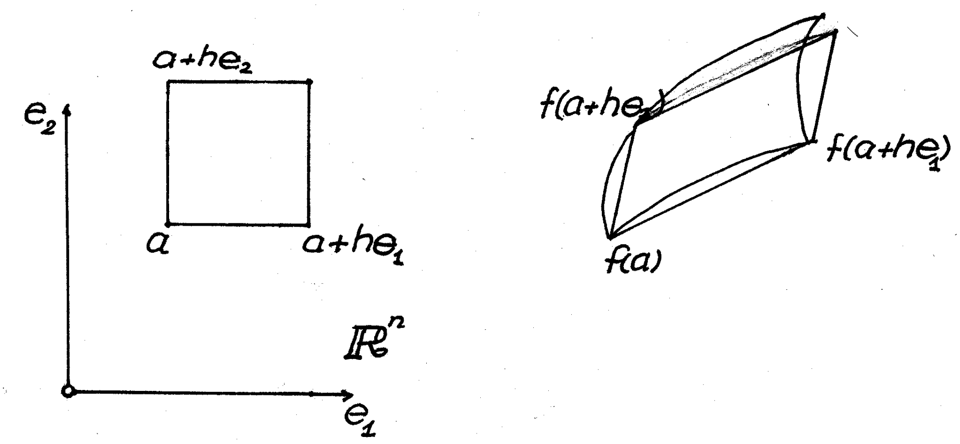

Figure 5.3

Nun bekommen wir einen n-dimensionalen Würfel in $\mathbb{R}^n$ mit den Ecken a, $a + he_1, \ldots,$ nämlich die Punktmenge $W(h) = \{a + t_1 he_1 + t_2 he_2 + \cdots + t_n he_n \colon 0 \le t_j \le 1, j = 1, \ldots, n\}$. Das Volumen dieses Würfels ist $D(he_1, \ldots, he_n) = h^n D(e_1 \ldots, e_n) = h^n$ wegen der Multilinearität und der Normierung der Determinante. Ist h klein genug, so ist $W(h)$ in X enthalten.

Nun wenden wir die Abbildung f auf die Ecken des Würfels $W(h)$ an. Die Vektoren $f(a + he_1) - f(a), \ldots, f(a + he_n) - f(a)$ spannen ein Parallelepiped auf welches für kleine h von dem Bild $f\big(W(h)\big)$ nur geringfügig abweichen wird. Auf alle Fälle können wir nun das Volumen dieses Parallelepipeds $W_f(h)$ wie folgt berechnen: Volumen $W_f(h) = D(f(a+he_1) - f(a), \dots, f(a+he_n) - f(a))$. Nun gilt wegen der Differenzierbarkeit der Funktion f im Punkte a die Beziehung $f(a+he_j) - f(a) = h\cdot f(a)(e_j) + \mathrm{o}(h)$. Wir erinnern uns: $\mathrm{o}(h)$ ist eine Funktion—hier vektorwertig—für die $\lim \frac{1}{h}\,\mathrm{o}(h) = 0$ gilt bei $h \to 0$! Dies setzen wir ein und bekommen Volumen $W_f(h) = D(hf'(a)(e_1) + \mathrm{o}(h), \dots, hf'(a)(e_n) + \mathrm{o}(h)) = h^n D(f'(a)(e_1), \dots, f'(a)(e_n)) + \mathrm{o}(h)$.

Nun bemerken wir aber, daß $f'(a)(e_j)$ nichts anderes ist als die j-te Spalte der Funktionalmatrix $((D_k f_i)(a))_{i,k=1,\dots,n}$. Damit erkennen wir nun

$$D(f'(a)(e_1), \dots, f'(a)(e_n)) = \det f'(a).$$

Somit haben wir das folgende Ergebnis gefunden:

Bemerkung 5.1. Ist $W(h)$ der Würfel $\{a + \sum_{j=1}^n t_j h e_j \colon (t_1, \dots, t_n) \in [0,1]^n\}$ und $W_f(h)$ der Würfel $\{f(a) + \sum_{j=1}^n t_j (f(a+he_j) - f(a)) \colon (t_1, \dots, t_n) \in [0,1]^n\}$ so gilt $\mathrm{Vol}\big(W_f(h)\big) = \mathrm{Vol}\big(W(h)\big)\cdot \det f'(a) + \mathrm{o}(h) = h^n \det f'(a) + \mathrm{o}(h)$. $\square$

Mit anderen Worten: Bis auf einen sehr kleinen Fehler ist die Funktionaldeterminante $\det f'(a)$ das Verhältnis $\mathrm{Vol}\big(W_f(h)\big) : \mathrm{Vol}\big(W(h)\big)$. Da die Abweichung $(f(a+t_j h_j) - f(a)) - t_j h f'(a)(e_j)$ gleich $\mathrm{o}(h)$ ist, dürfen wir $W_f(h)$ als eine gute Approximation des Bildes $f\big(W(h)\big)$ ansehen. Wenn es erlaubt wäre, vom Volumen des Bildes des Würfels $W(h)$ bei der Abbildung f zu sprechen, so müßte dieses in der Nähe von $W_f(h)$ zu finden sein. (Volumenbestimmung von allgemeinen Mengen im $\mathbb{R}^n$ ist aber ein Problem für sich, dem man sich bei der Integrationstheorie in mehreren Variablen zuwenden muss.) In diesem Sinne ist nun die Funktionaldeterminante von f im Punkte a der *Verzerrungsfaktor des Volumens sehr kleiner Würfel bei a.*

Wir sind immer noch bei der Betrachtung von Größen, die einer linearen Abbildung $L\colon \mathbb{R}^n \to \mathbb{R}^n$ invariant zugeordnet sind und haben uns eben mit der Determinante befaßt.

Zweites Beispiel: Die *Spur* $\mathrm{Tr}\, L$ (man schreibt $\mathrm{Tr}\, L$ für "trace", kann aber auch die Schreibweise $\mathrm{Sp}\, L$ lesen). Die Spur von L berechnet sich aus der Matrix $(a_{jk})_{j,k=1,\dots,n}$ von L sofort als die Summe der Diagonalelemente $\mathrm{Tr}\, L = a_{11} + a_{22} + \cdots + a_{nn}$. Die Spur ist aber von der Matrixdarstellung unabhängig, da sie auch zugleich die Summe der mit ihrer Vielfachheit gezählten Eigenwerte von L ist, und die Eigenwerte sind ja basisunabhängig. Die Determinante $\det(L - \lambda\cdot \mathrm{id})$ ist ein Polynom n-ten Grades in λ, das sogenannte *charakteristische Polynom.* Dieses ist offenbar basisunabhängig, da die Determinante basisunabhängig ist. Da man sehr rasch

$$\det(L - \lambda\cdot \mathrm{id}) = (-1)^n \lambda^n + (-1)^{n-1}\,\mathrm{Tr}(L)\cdot \lambda^{n-1} + \cdots + \det L$$

nachprüft, hat man somit einen alternativen Beweis für die Basisunabhängigkeit der Spur. (Die Menge aller Eigenwerte von L heißt auch das *Spektrum von L*; deswegen vermeiden wir lieber die Bezeichnung $\mathrm{Sp}\, L$ für die Spur!) Im Einzelnen werden alle diese Begriffe in der linearen Algebra besprochen. Wir weisen hier nur noch darauf hin, daß wir die Spur mit Hilfe des Skalarproduktes auf $\mathbb{R}^n$ und mit Hilfe der Grundvektoren $e_1 = (1, 0, \dots, 0), \dots, e_n$ wie folgt angeben können:

$$\operatorname{Tr} L = (L(e_1) \mid e_1) + \cdots + (L(e_n) \mid e_n),$$

denn es ist ja in der Tat $(L(e_j) \mid e_j) = a_{jj}$, $j = 1, \dots, n$. Die Vorschrift $L \mapsto \operatorname{Tr} L\colon \operatorname{Lin}(\mathbb{R}^n) \to \mathbb{R}$ ist eine Linearform auf dem Vektorraum aller linearen Abbildungen von $\mathbb{R}^n$ in sich, und sie erfüllt die Bedingungen $\operatorname{Tr} 1_{\mathbb{R}^n} = n = \dim \mathbb{R}^n$ und $\operatorname{Tr}(L_1 L_2) = \operatorname{Tr}(L_2 L_1)$ für alle $L_1, L_2 \in \operatorname{Hom}(\mathbb{R}^n, \mathbb{R}^n)$. (Übungsaufgabe!). Tatsächlich ist die Linearform „Spur“ durch diese Bedingungen schon vollständig bestimmt; es gibt also nur eine Linearform mit diesen Eigenschaften, nämlich die Spur. Leider ist eine geometrisch intuitive Interpretation der Spur nicht anzugeben. An der Wichtigkeit der Spur ist indessen kein Zweifel erlaubt.

Und so kommt der Spur der Ableitung $f'(a)$ eines Vektorfeldes auch eine besondere Bedeutung zu. Sie ist also invariant und basisunabhängig definiert und heißt die *Divergenz* des Vektorfeldes f im Punkte a und wird $\div_a f$ geschrieben. Also

$$\begin{aligned} \div_a f &= \operatorname{Tr} f'(a) = (D_1 f_1)(a) + \cdots + (D_n f_n)(a) \\ &= \left.\frac{\partial f_1}{\partial x_1}\right|_{x=a} + \cdots + \left.\frac{\partial f_n}{\partial x_n}\right|_{x=a} \\ &= (f'(a)(e_1) \mid e_1) + \cdots + (f'(a)(e_n) \mid e_n). \end{aligned}$$

Über die Bedeutung dieser Divergenz wird im Zusammenhang mit den großen Integralsätzen die Rede sein müssen. Es gibt da noch eine Merkregel, der im Übrigen mathematisch keine sonderliche Bedeutung beizumessen ist. Faßt man das n-tupel der Operationsbefehle

$$(D_1\{\cdot\})(a), \dots, (D_n\{\cdot\}(a)) = \partial/\partial x_1|_{x=a}, \dots, \partial/\partial x_1|_{x=a}$$

als einen "Vektor" auf und schreibt dafür $\bigtriangledown_a$ („Nabla“) so ist die Divergenz eine Art „Skalarprodukt“ $(\bigtriangledown_a \mid f)$. Man klammere sich nicht zu sehr an diese Sache; dies ist bestenfalls Kalkül, trägt aber schwerlich zum Verständnis des Inhaltes bei. Wir halten fest: Wir haben die Divergenz definiert und zu ihrem Hintergrund in der linearen Algebra einige vorbereitende Bemerkungen gemacht; wir haben aber bis jetzt mit der Divergenz noch nichts angefangen.

Gradientenvektorfelder

Wir wollen uns ein paar Gedanken machen, wo Vektorfelder sich natürlich aus anderen Daten ergeben. Ein kurzer Rückblick auf unsere Diskussion der Höhenfunktionen $p\colon X \to \mathbb{R}$, $X \subseteq \mathbb{R}^n$ gibt uns gewiß schon einmal eine ganze Klasse von Vektorfeldern. Ist nämlich X offen und p überall differenzierbar, so ist die Ableitung $p'\colon X \to \mathbb{R}^n$ mit $p'(x) = \operatorname{grad}_x p$ ein Vektorfeld. Gradienten sind also Vektorfelder. Wir wissen: Dieses Vektorfeld liefert uns in jedem Punkt x von X die Richtung des größten Anwachsens der Funktion p und die Wachstumsrate. Die Vektoren des Vektorfeldes $f\colon X \to \mathbb{R}^n$, $f(x) = \operatorname{grad}_x p$ stehen in jedem Punkt auf den Höhenflächen (bzw. Höhenlinien) der Funktion p senkrecht. (Wir sagen

„Höhenflächen“, meinen aber dabei natürlich „Höhenmengen“, wenn $n > 2$ ist. Diese Dinge sind in aller Exaktheit durchdiskutiert, zuletzt mit dem Satz 3.30

Eine nicht von der Hand zu weisenden Frage erhebt sich an dieser Stelle: Sind vielleicht alle Vektorfelder $f: X \to \mathbb{R}^n$ von der Form $f(x) = p'(x)$ für eine Höhenfunktion $p: X \to \mathbb{R}$, die man in diesem Zusammenhang gelegentlich auch ein *skalares Feld* nennt. Überlegen wir: Wenn wir $f = p'$ haben und einmal annehmen, daß das Vektorfeld f noch stetig differenzierbar ist, dann bedeutet dies, daß das skalare Feld p zweimal stetig differenzierbar ist. Wir wissen, daß $f'(x) = p''(x)$ eine lineare Abbildung $\mathbb{R}^n \to \mathbb{R}^n$ ist, die wir uns als n mal n-Matrix mit den Koeffizienten $D_k f_j(x) = (D_k D_j p)(x)$, $j, k = 1, \ldots, n$ denken dürfen. Wenn wir nun f' und damit p'' als stetig vorausgesetzt denken, dann können wir jetzt den Satz 4.1 anwenden und schließen, daß $\partial_k f_j(x) = (\partial_k \partial_j p)(x) = (\partial_j \partial_k p)(x) = \partial_j f_k(x)$ sein muß. Mit anderen Worten:
Eine notwendige Bedingung dafür, daß ein Vektorfeld f der Gradient eines Skalarfeldes p ist, ist die Symmetrie der Ableitung $f'(x)$—d.h. die Symmetrie der Matrix $(\partial_k f_j(x)_{j,k=1,\ldots,n} = (\partial_k \partial_j p)(x)$—für alle $x \in X$.

Selbstverständlich wollen wir jetzt auch wissen, ob diese Bedingung auch hinreichend ist, d. h. ob wir zu einem gegebenen Vektorfeld $f: X \to \mathbb{R}^n$ auf einer offenen Menge X des $\mathbb{R}^n$ mit symmetrischer Ableitung f' auch immer ein Skalarfeld p auf X so finden, daß $f = p'$ gilt. Bei $n = 1$ ist dies nichts anderes als die Frage nach einer Stammfunktion, die für stetige f durch den Hauptsatz der Differential und Integralrechnung 5.18 in Analysis I erledigt ist. (Die Symmetrievoraussetzung ist hier gegenstandslos.) Es ist also naheliegend, daß eine Integration auch in höheren Dimensionen nötig sein wird. Da wir die Integrationstheorie im Falle mehrerer Veränderlicher erst noch aufrollen müssen, soll die Antwort hier zunächst noch vorläufig bleiben.

Als erstes machen wir eine Voraussetzung über die offene Menge X: Jeder Punkt x sei mit dem festen Punkt a durch ein gerades Streckensegment innerhalb X verbindbar, d. h. die sämtlichen Punkte $a + t(x - a)$ mit $t \in [0, 1]$ seien in X enthalten. (Beispiel: Die Kugelumgebung $U_1(0)$ vom Radius 1 um 0. Gegenbeispiel: Die Schalenmenge $U_1(0) \backslash B_{1/2}(0) = \{x \in \mathbb{R}^n : \frac{1}{2} < \|x\| < 1\}$.) Nun betrachten wir das Vektorfeld auf diesem Geradensegment und projizieren den Vektor $f(a + t(x - a))$ auf die von dem Streckensegment aufgespannte Gerade, indem wir das Skalarprodukt $(f(a + t(x - a)) \mid x - a)$ bilden. Die Funktion $t \mapsto (f(a + t(x - a)) \mid x_a)$ ist stetig und daher integrierbar. Wir definieren $p(x) = \int_0^1 (f(a + t(x - a)) \mid x_a) dt$. Wir versuchen nun, die Funktion $p: X \to \mathbb{R}$ partiell zu differenzieren, falls dies möglich sein sollte. So befassen wir uns z.B. mit der ersten partiellen Ableitung ∂_1. Der Integrand hängt (neben vielen anderen Variablen) von x_1 ab, ist also von der Form $g(x_1, t)$. Wir stoßen auf das Problem, das Integral $\int_0^1 g(x_1, t) dt$ nach x_1 abzuleiten. Dieses Problem liegt sehr nahe bei der in Satz 1.14 diskutierten. Dort wurde eine gleichmäßig gegen $f: [0, 1] \to \mathbb{R}$ konvergente Funktionenfolge $f_n: [0, 1] \to \mathbb{R}$ betrachtet und festgestellt, daß sich Integration und Limesbildung vertauschen ließen: $\lim_n \int_0^1 f_n =$

$\int_0^1 f = \int_n^1 \lim f_n$. Da die Differentiation von $g(x_1, t)$ nach dem Parameter x_1 die Limesbildung $\lim_{h \to 0} \frac{1}{h}(g(x_1 + h, t) - g(x_1, t)) = (partial_{x_1} g)(x_1, t)$ ist, können wir Differentiation und Integration vertauschen, falls die Differenzenquotienten *gleichmäßig* gegen die Ableitung konvergieren. Wir nehmen hier an, daß dies gesichert sei und beschäftigen uns später eingehend mit dieser Frage. Wir gehen also davon aus, daß $(\partial_1 p)(x) = \partial_{x_1} \int_0^1 g(x_1, t)dt = \int_0^1 (\partial_{x_1} g)(x_1, t)dt$ ist. Damit bleibt uns also die Berechnung von $\frac{\partial}{\partial u_1}\Big|_u = x\big(f(a + t(u - a)) \mid u - a\big)$. Nun ist $\big(f(a+t(x-a)) \mid x-a\big) = (x_1 - a_1)f_1(a+t(x-a)) + \sum_{j=2}^n (x_j - a_j)f_j(a+t(x-a))$. Die Ableitung der ersten Summanden nach x_1 ist nach der Produktregel und Kettenregel $f_1(a + t(x - a)) + (x_1 - a_1)(\mathrm{grad}_{a+t(x-a)} f_1 \Big| \frac{\partial}{\partial u_1}\Big|_{u=x} (a + t(u - a))) = f_1(a+t(x-a)) + t(x_1 - a_1)(\partial_1 f_1)(a+t(x-a))$. Die Ableitung der nachfolgenden Summanden nach x_1 ist einfacher: Sie ist $(x_j - a_j)(\partial_1 f_j)(a + t(x - a))$ für $j > 1$. Aber nun kommt unsere Hauptvoraussetzung ins Spiel: Es ist nämlich $\partial_1 f_j = \partial_j f_1$. Damit aber bekommen wir $(\partial_1 p)(x) = \int_0^1 f_1(a+t(x-a))dt + \int_0^1 \sum_{j=1}^n t(\partial_j f_1)(x + t(x-a))(x_j - a_j) = \int_0^1 f_1(a+t(x-a))dt + \int_0^1 t(\mathrm{grad}_{a+t(x-a)} f_1 \mid x-a)dt$. Setzen wir vorübergehend $F(t) = f_1(a + t(x - a))$, dann gilt $F'(t) = (\mathrm{grad}_{a+t(x-a)} f_1 \mid x - a)$ und mit Hilfe der partiellen Integration 5.29 in Analysis I bekommen wir $\int_0^1 t(\mathrm{grad}_{a+t(x-a)} f_1 \mid x - a)dt = \int_0^1 tF'(t)dt = tF(t)|_0^1 - \int_0^1 F = F(1) - \int_0^1 f_1(a + t(x - a))dt$. Es ergibt sich $(\partial_1 p)(x) = F(1) = f_1(x)$, da sich die übrigen Summanden gerade wegheben. In genau der gleichen Weise finden wir ganz allgemein $(\partial_j p)(x) = f_j(x)$ und somit $p'(x) = f(x)$. Bis auf die noch später zu füllende Beweislücke über Vertauschbarkeit von Integration und Differentiation nach einem Parameter haben wir den folgenden Satz ermittelt:

Satz 5.2. *Sei X ein offenes Gebiet im $\mathbb{R}^n$, in welchem jeder Punkt mit einem festen Punkt a geradlinig verbindbar ist. Für ein stetig differenzierbares Vektorfeld $f\colon X \to \mathbb{R}^n$ sind die beiden folgenden Bedingungen gleichwertig:*

(1) $f'(x)$ ist symmetrisch, d. h. $\partial_k f_j = \partial_j f_k$.
(2) Es gibt eine zweimal stetig differenzierbare Funktion $p\colon X \to \mathbb{R}$ mit $p' = f$, d.h. mit $f(x) = \mathrm{grad}_x\, p$ für alle $x \in X$. □

Eine Funktion p mit der in (2) angegebenen Eigenschaft wird auch ein Potential für f genannt und ein Vektorfeld f, das die Bedingung (2) erfüllt, wird auch *Gradientenfeld* genannt. Mit jeder Potentialfunktion p für ein Gradientenfeld f ist auch die Funktion $x \mapsto p(x) + c$ für jede feste Zahl c ein Potential. Zwei Potentiale für ein Gradientenfeld unterscheiden sich nur durch eine Konstante, denn ihre Differenz ist eine Höhenfunktion mit der Ableitung 0. Nach Satz 2.10 (und die nachfolgende Beobachtung folgt daraus sofort das Verschwinden einer solchen Funktion.

Wir erinnern uns an die Drehung A der Ebene $\mathbb{R}^2$ um 90° mit der Matrix $\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$. Auf der offenen Menge $X = \mathbb{R}^2$ ist das stetig differenzierbare Vektorfeld $f(x) = \|x\|^{-2}Ax$ definiert. Es erfüllt (1) (Übungsaufgabe). Wir werden später zeigen, daß dennoch (2) nicht erfüllt sein kann; in der Zwischenzeit zeichne man eine Skizze dieses Vektorfeldes und versuche sich klarzumachen, wie ein Potential dazu aussehen müßte, dessen Gradient gerade gleich f ist. (Höhenlinien? Anstieg?) (S.Figur). Auf der anderen Seite ist das durch $f(x) = -\|x\|^{-m-1}x$ gegebene Vektorfeld $f\colon \mathbb{R}^n\backslash\{0\} \to \mathbb{R}^n$ allemal ein Gradientenfeld für $m = 1, 2, 3\ldots$: Man finde ein Potential!

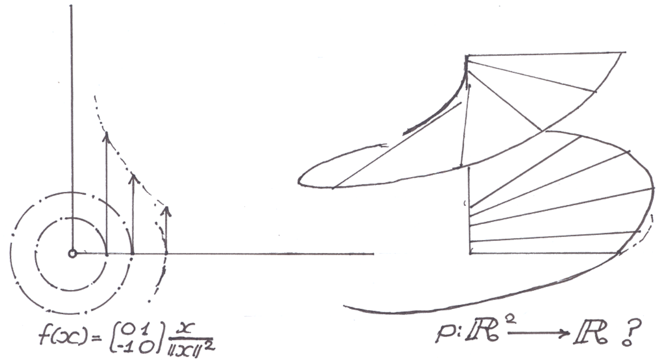

Figure 5.4

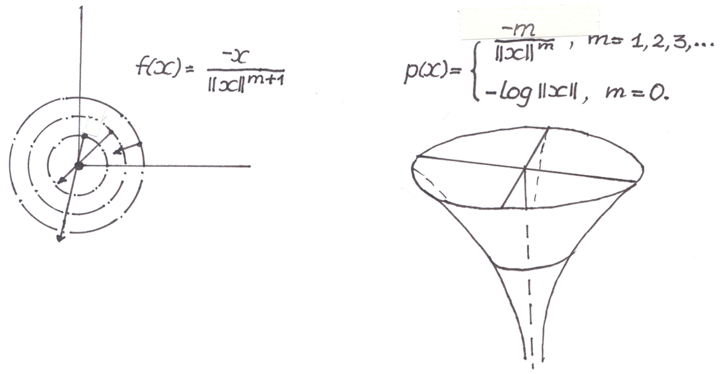

Figure 5.5

Man erkennt an den Beispielen schon ein wenig die Rolle der geometrischen Struktur des Definitionsbereiches des gegebenen Vektorfeldes f. Im Falle des durch $f(x) = A(\frac{x}{\|x\|^2})$ auf $\mathbb{R}^2\backslash\{0\}$ gegebenen Vektorfeldes kann man nach dem Satz 5.2 zwar Potentiale auf Teilgebieten finden, so etwa auf einem Teilgebiet der Form $\{(x, y) : y = 0 \Rightarrow x > 0\}$, aber diese lokalen Potentiale lassen sich nicht zu einem auf $\mathbb{R}^2\backslash\{0\}$ definierten Potential zusammenfügen. Auf der anderen Seite gibt es aber auch auf „Lochgebieten" wie etwa $\mathbb{R}^n \setminus \{0\}$ globale Potentiale für gewisse Vektorfelder, wie etwa die Felder $f(x) = \frac{-x}{\|x\|^{m+1}}$. Diese haben natürlich in der Physik eine große Bedeutung. Für $m = 2$ bekommen wir ein Vektorfeld, das bis auf einen Proportionalitätsfaktor das Feld einer Kraft beschreibt, die in $\mathbb{R}^n\{0\}$ auf den Ursprung 0 hin gerichtet ist, und deren Größe dem Abstandsquadrat umgekehrt proportional ist. Darunter fallen Kräfte wie die Gravitation oder die elektrostatische Anziehung entgegengesetzter Ladungen. Das dazugehörige Potential heißt daher auch das Schwerepotential oder das elektrostatische Potential.

Bei ebenen Problemen kann es häufig zum Verständnis der geometrischen Struktur von Interesse sein, das Problem in Polarkoordinaten zu betrachten. Was heißt dies? Es sei uns eine Höhenfunktion (ein Skalarfeld) $p: X \to \mathbb{R}$ auf einer offenen Menge X von $\mathbb{R}^2$ gegeben, die z.B. die negative Hälfte der x-Achse nicht schneidet. Die Punkte $(x, y) \in X$ beschreiben wir in Polarkoordinaten: $x = r\cos t$ und $y = r\sin t$. Dies ist nichts anderes als die Angabe einer stetig differenzierbaren Bijektion $G : Y \to X$ mit $Y \subseteq \mathbb{R}^2$ wobei $G(r, t) = (r\cos t, r\sin t)$ gesetzt ist. Aufgrund der Voraussetzung über X gibt es dazu eine stetig differenzierbare Umkehrfunktion $F: X \to Y$ mit $F(x, y) = ((x^2 + y^2)^{1/2}, \arg(x, y))$ wobei $\arg(x, y)$ die wohlbestimmte Zahl $t \in]-\pi, \pi[$ mit $x = r\cos t, y = r\sin t$ und $r^2 = x^2 + y^2$ ist. Das Gebiet Y liegt also in $\{(r, t) : 0 < r, |t| < \pi\}$. Setzen wir nun $q = p \circ G: Y \to \mathbb{R}$ so können wir die Funktion p durch die Funktion q "als in Polarkoordinaten angegeben" betrachten, denn es gilt ja $q(r, t) = p(x, y)$ mit $x = r\cos t, y = r\sin t$. Umgekehrt ist dann $p = q \circ F$. Also gilt nach der Kettenregel $p'(x, y) = q'(F(x, y)) \circ F'(x, y)$. Nun ist $F(x, y)$ eine für allemal gegeben, also ist auch $F'(x, y) = G'\big(F(x, y)\big)^{-1}$ ein für allemal bekannt. Man rechne als Übungsaufgabe F' und G' in Matrizenform aus. Wir können nun ohne Schwierigkeiten den Gradienten einer in Polarkoordinaten gegebenen Funktion ausrechnen. (Übungsaufgabe: $q(r, t) = t; \operatorname{grad}_{(x,y)} p = ?$ für $p(x, y) = q(r, t) = t$.) (Übungsaufgabe: $G'(r, t) = \begin{pmatrix} \cos t & -r\sin t \\ \sin t & r\cos t \end{pmatrix}$, $F'(x, y) = \begin{pmatrix} x/r & y/r \\ -y/r^2 & x/r^2 \end{pmatrix}$, $r^2 = x^2 + y^2$, $r > 0$.)

Ist uns ein genügend glattes Vektorfeld $f: X \to E$, $E = \mathbb{R}^n$ auf einer offenen Menge M von E gegeben, so stellt sich natürlich die Frage, ob von jedem Punkt a eine differenzierbare Kurve ausgeht, etwa $u:]-\delta, \delta[\to \mathbb{R}^n$ mit der Eigenschaft, daß die Ableitung $u'(t) = f\big(u(t)\big)$ mit der Anfangswertbedingung $u(0) = a$ durch eine geeignete Kurve lösbar ist. Dieses Problem ist ein typischen Problem aus der Theorie der gewöhnlichen Differentialgleichungen, mit dem wir uns demnächst eingehend befassen werden. Wir werden dann sehen, daß dieses Problem weitgehend lösbar ist. In der Physik treten als Lösungskurven des Problems für ein gegebenes Vektorfeld, welches als Kraftfeld aufgefaßt wird, als sogenannte „Kraftlinien“ in Erscheinung. Man kann diese im Fall eines Magnetfeldes sogar mehr oder weniger sichtbar machen. Man erinnere sich dabei an die einschlägigen Versuche mit Eisenfeilspänen.

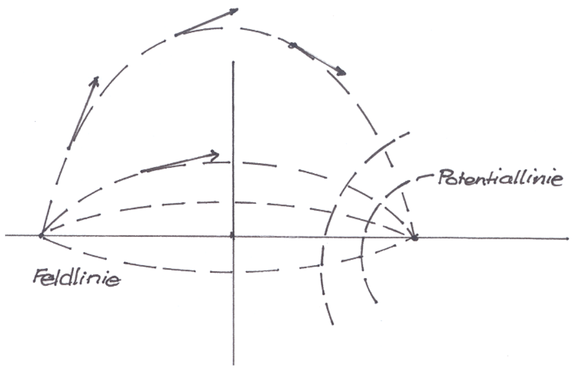

Figure 5.6

Ist das Feld ein Gradientenfeld, so stehen die „Feldlinien", deren Tangentialvektoren ja gerade die Vektoren $f(x)$ des gegebenen Feldes f sind, auf den Höhenlinien („Höhenflächen!) der zugehörigen Potentialfunktion p, für die $f = p'$ gilt, senkrecht. Man spricht daher im Zusammenhang mit dieser Interpretation des Sachverhaltes $f = p'$ auch von den *Potentiallinien* im Fall $n = 2$ und von den *Potentialflächen* im Fall $n = 3$. Man sollte sie im Fall $n > 3$ entsprechend *Potentialhyperflächen* nennen. Es ist aber gut, sich darüber Klarheit zu verschaffen, daß hier nichts Neues eingeführt wird. Die nach der Definition IV.1.65 angesprochenen und im Satz IV.1.82 nocheinmal genauer in Augenschein genommenen Höhenhyperflächen sind im Falle der Funktion $p: M \to \mathbb{R}$ mit $f = p'$ genau die Potentialhyperflächen zu dem Gradientenvektorfeld f.

Nun sei $f: M \to E = \mathbb{R}^n$ ein Gradientenvektorfeld mit der Potentialfunktion $f = p'$. wir können uns fragen, was die Divergenz des Vektorfeldes f für das Potential p bedeutet. Wir haben $\div_a f = \div_a \operatorname{grad}_a p = \operatorname{Tr}\big(p''(a)\big)$ mit der Spur $\operatorname{Tr} L$ einer linearen Abbildung $L \in \operatorname{Hom}(E, E)$. Wir können die Zahl $\operatorname{Tr}\big(p''(a)\big)$ durch die partiellen Ableitungen ausdrücken:

$$\operatorname{Tr}\big(p''(a)\big) = (\partial_1^2 p)(a) + \cdots + (partial_n^2 p)(a) = \left.\frac{\partial^2 p}{\partial_1^2}\right|_{x=a} + \cdots + \left.\frac{\partial^2 p}{\partial_n^2}\right|_{x=a}.$$

Der Operator $\div_a \operatorname{grad}_a$ hat eine eigene Abkürzung: Er heißt *Laplace-Operator* und wird $\triangle_a$ geschrieben:

$$\triangle_a p = \sum_{j=1}^{n} \left.\frac{\partial^2 p}{\partial_j^2}\right|_{x=a} = \div_a \operatorname{grad}_a p.$$

Funktionen p mit divergenzfreiem Gradientenfeld $f = p'$ sind also genau die Funktionen $p: M \to R$, die auf der offenen Menge M zweimal stetig differenzierbar sind und dort die Gleichung $\triangle p = 0$ erfüllen—etwas expliziter:

$$(\forall x \in M) \quad \triangle_x p = (\triangle p)(x) = 0.$$

Solche Funktionen heißen *harmonisch* ; man nennt sie auch *Potentialfunktionen.* (Man sollte diesen Begriff nicht mit dem Potential eines Vektorfeldes durcheinanderbringen.

Wir diskutieren ein Beispiel: Wir setzen $X = \mathbb{R}^n \setminus \{0\}$ und definieren ein Vektorfeld $f\colon X \to \mathbb{R}^n$ mit Hilfe einer stetig differenzierbaren Funktion $s\colon \mathbb{R}^n \to \mathbb{R}$ durch $f(x) = s(x){\cdot}x$. Nach der Produktregel gilt $f'(x)(u) = s'(x)(u){\cdot}x + s(x){\cdot}u$. Also haben wird

$$\begin{aligned}\operatorname{Tr} f'(x) &= \sum_{j=1}^{n} \big(s'(x)(e_j)(x \mid e_j) + s(x)(e_j \mid e_j)\big)\\ &= \sum_{j=1}^{n} s'(x)(e_j){\cdot}(x \mid e_j) + ns(x).\end{aligned}$$

Nehmen wir insbesondere z.B. $s(x) = \|x\|^{-m-1}$ so ist

$$s'(x)(u) = \frac{-m-1}{\|x\|^{m+3}}(x \mid u)$$

und

$$\sum_{j=1}^{n}(s'(x)(e_j){\cdot}(x \mid e_j) = \frac{-m-1}{\|x\|^{m+3}}\sum_{j=1}^{n}(x \mid e_j)^2 = \frac{-m-1}{\|x\|^{m+1}},$$

so daß wir $\operatorname{Tr} f'(x) = \div_x f = \frac{-m-1}{\|x\|^{m+1}} + \frac{n}{\|x\|^{m+1}}$ bekommen. Für $m = n-1$ ist diese Funktion identisch null, d.h. f ist divergenzfrei. Nun ist aber $f(x) = \frac{x}{\|x\|^{m+1}}$ ein Gradientenfeld mit dem Potential

$$p(x) = \begin{cases} \log\|x\| & \text{falls } m = 1,\\ \frac{-1}{(m-1)\|x\|^{m-1}} & \text{falls } m \geq 2.\end{cases}$$

Zusammenfassend haben wir also die folgende Beobachtung:

Bemerkung IV.1.88. Die Funktionen $p\colon \mathbb{R}^n \setminus \{0\}\colon \mathbb{R}^n$, die durch

$$p(x) = \begin{cases} -\log\|x\| & \text{falls } m = 2,\\ \frac{1}{(m-2)\|x\|^{m-2}} & \text{falls } m \geq 3\end{cases}$$

gegeben sind sind allesamt harmonisch, d.h. sie genügen der Laplace-Gleichung $\triangle p = 0$. Die aus ihnen abgeleiteten Gradientenfelder p' sind gegeben durch

$$p'(x) = \frac{-x}{\|x\|^n} = \frac{-1}{\|x\|^{n-1}}{\cdot}\frac{x}{\|x\|}. \qquad \square$$

Das Newtonsche Schwerkraftgesetz kann also nur im dreidimensionalen Raum aus einem harmonischen Potential begeleitet werden, nicht aber in einem euklidischen Raum einer von 3 verschiedenen Dimension.

Übung E1.19. Beweise den folgenden Satz:
Satz. *Sei U eine offene Menge in $\mathbb{C}$ und $f\colon U \to \mathbb{C}$ eine holomorphe Funktion. (D.h. $f'(z) = \lim_{h\to 0}\frac{1}{h}\big(f(z+h) - f(z)\big)$ existiert für alle $z \in U$. Setzt man*

$X = \{(x,y) \in \mathbb{R}^2 : x + y{\cdot}i \in U\}$ und definiert man u, $v: X \to \mathbb{R}$ durch $f(x+y{\cdot}i) = u(x,y) + v(x,y){\cdot}i$, dann sind u und v harmonische Funktionen. $\square$

Wir werden diesen Sachverhalt später systematisch untersuchen. Jedenfalls haben wir hier eine reiche Quelle harmonischer Funktionen auf offenen Mengen des $\mathbb{R}^2$.

Vertauschung von Differentiation und Integration

Wir haben die Frage, um die es hier geht, schon im Beweis des Satzes 5.2 erkannt. Wir haben eine Funktion $f: I \times [a,b] \to \mathbb{R}$ mit einem Intervall $I \subseteq \mathbb{R}$ (also eine Höhenfunktion in zwei Variablen auf einem Rechtecksbereich der Ebene). Wir setzen $f_s(t) = f(s,t)$ und betrachten damit die Funktion f als eine parametrisierte Familie von Funktionen $f_s: [a,b] \to \mathbb{R}$ einer Variablen, die durch den „Parameter" s durchnumeriert sind, der die Menge I durchläuft, die wir hier einfacherweise als Intervall angenommen haben. Nun nehmen wir an, alle Funktionen $f_s, s \in I$ seien Riemannsch integrierbar. Wenn man diese Voraussetzung durch die Funktion f selber ausdrücken will, so heißt dies doch, daß wir voraussetzen, die Funktion f sei auf allen vertikalen Strecken des Rechtecks $I \times [a,b]$ in der zweiten Variablen integrierbar. Damit bekommen wir nun eine Funktion $F: I \to \mathbb{R}$, die durch $F(s) = \int_a^b f_s = \int_a^b f(s,t)dt$ gegeben ist.

Für diese Funktion stellen sich nun die üblichen Fragen: Ist F stetig? Ist F differenzierbar? Wenn ja, wie berechnen wir die Ableitung? Diese Fragen wollen wir im Folgenden behandeln.

Bei der Lösung dieser Fragen bewährt sich unsere Auffassung, die Menge aller Riemann-integrierbaren Funktionen $I[a,b]$ auf $[a,b]$ als Banachraum aufzufassen (s. Satz 1.14). Nun ist also $f_s \in I[a,b]$ für alle $s \in I$. Wir haben also in Wirklichkeit eine Funktion $g: I \to I[a,b]$, die durch $g(s) = f_s$ gegeben ist. (Mit dieser Bezeichnung gilt dann $g(s)(t) = f_s(t) = f(s,t)$, und weiter $F(s) = \int_a^b g(s)$.) Die Integrationsoperation $h \mapsto \int_a^b h : I[a,b] \to \mathbb{R}$ ist ihrerseits eine Funktion $J: I[a,b] \to \mathbb{R}$, und zwar eine lineare Abbildung. Wir können also nun sogar so zusammenfassen:

(1) $F = J \circ g: I \to \mathbb{R}$, wo $g(s)(t) = f(s,t)$ und $J(h) = \int_a^b h = \int_a^b h(t)dt$ gilt.

Wegen $J(h+k) = J(h)+J(k)$ ist J als lineare stetige Abbildung sogar differenzierbar, und zwar mit der Ableitung $J'(h)(k) = J(k)$. (In der Tat ist ja J stetig, denn $\|J(h) - J(k)\| = \|J(h-k)\| = \| \int_a^b (h-k)\| \leq (b-a)\|(h-k)\|$, wobei wir wie verabredet auf dem Vektorraum $I[a,b]$ die sup-Norm $\|h\| = \sup\{|h(x)| : x \in [a,b]\}$ betrachten. Wir erinnern ferner daran, daß wir schon lange bei linearen Abbildungen $L: \mathbb{R}^n \to \mathbb{R}^m$ festgestellt hatten, daß sie immer differenzierbar sind und zwar so, daß die Ableitung $L'(a)$ in einem jedem Punkt des Definitionsbereiches $\mathbb{R}^n$ gerade gleich L selbst ist.)

Differentiation in Banachräumen

Mit unserer gegenwärtigen Diskussion sind wir unversehens in einen allgemeineren Bereich vorgestoßen: Wir betrachten nämlich Funktionen $\varphi: B_1 \to B_2$ zwischen

Banachräumen: Beispielsweise ist $g: I \to I[a, b]$ eine Kurve in dem Banachraum $I[a, b]$ (einmal vorausgesetzt, daß g stetig ist), während die Funktion $J: I[a, b] \to \mathbb{R}$ vom Typ einer Höhenfunktion ist; als stetige Linearform ist diese Funktion allerdings sehr spezieller Natur.

Unsere Hauptdefinition 3.2 läßt sich, wie wir schon einmal bemerkten, sogleich für Funktionen $\varphi: X \to B_2$ von einem Bereich X in einem Banachraum B_1 in einen Banachraum B_2 aussprechen: Wir sagen, die Funktion sei in einem inneren Punkt a von X differenzierbar, falls es eine stetige lineare Abbildung $L: B_1 \to B_2$ so gibt, daß $\varphi(a+h) = \varphi(a) + L(h) + \mathrm{o}(\|h\|)$ gilt. Die lineare Abbildung L wird dann als $\varphi'(a)$ bezeichnet. Dieser Differenzierbarkeitsbegriff ist wörtlich der in der Hauptdefinition 3.2 ausgedrückte. Da sich unsere Theorie der Differenzierbarkeit nur auf die oben für den allgemeineren Fall von Banachräumen wiederholten Hauptdefinition 3.2 stützt, bleiben unsere Sätze auch für diesen allgemeineren Fall richtig, so etwa z. B. die Kettenregel 3.7 aber auch die Regeln für die Summen und Skalarprodukte 3.6, die Produktregel 3.10 IV.1.47, wo allerdings im allgemeinen Fall die Stetigkeit von B vorausgesetzt werden muß; der Mittelwertsatz in der Form 2.10. Alles was sich auf das Skalarprodukt im Bildraum bezieht, wie etwa die Interpretation der ersten Ableitung $\varphi'(a)$ einer Höhenfunktion $\varphi: X \to \mathbb{R}$, $X \subseteq B_1$ als Gradient (s. Definition IV.1.65 ff.) verlangt, daß der Raum B_1 sogar ein Hilbertraum ist (s. Definition 3.15). Andernfalls muß die erste Ableitung eben nur als stetige Linearform $\varphi'(a): B_1 \to \mathbb{R}$ aufgefaßt werden, woran die Theorie kaum, aber die geometrische Anschauung etwas leidet. Richtungsableitungen können wir nach wie vor wie in Definition 3.16 definieren, aber ihre in 3.16 angegebene Berechnung als Skalarprodukt erfordert einen Hilbertraum als B_1. Ist der Raum B_1 nur ein Banachraum, so gilt immer noch die Beziehung $\partial_{a;e}\varphi = \varphi'(a)(e)$. Mit den Richtungsableitungen, die wir nun auch im allgemeinen Fall haben, gilt auch der Satz 4.1 von der Vertauschbarkeit der Richtungsableitungen. Der Satz von den Inversen Funktionen 3.26 und der Satz von den impliziten Funktionen 3.29 gelten ebenfalls in der angedeuteten größeren Allgemeinheit mit denselben Beweisen, die eben gerade daraufhin angelegt waren, diese Allgemeinheit zu gewährleisten.

Nun zurück zu unserem Problem der Vertauschung von Integration und Differentiation! Wir haben $F(s) = J(g(s)) = \int_a^b f(s,t)dt$ und wollen auf die Stetigkeit oder auch Differenzierbarkeit von F schließen. Da die Hintereinanderausführung stetiger Funktionen stetig ist genügt es, die Stetigkeit von g zu kennen, da dann die oben schon bemerkte Stetigkeit von J die Stetigkeit von F liefert. Die Stetigkeit von g besagt, daß zu jedem $s_0 \in I$ und jedem $\varepsilon > 0$ ein $\delta > 0$ so existiert, daß für alle $s \in I$ mit $|s - s_0| < \delta$ stets $\|g(s) - g(s_0)\| \leq \varepsilon$ gilt, d. h. aber doch $\sup\{|f(s,t) - f(s_0,t)| : t \in [a,b]\} \leq \varepsilon$, und dies ist gleichwertig zu $|f(s,t) - f(s_0,t)| \leq \varepsilon$ für alle $t \in [a, b]$. Dies führt zur folgenden Definition:

Definition 5.3. Sei $X \subseteq \mathbb{R}^n$ und $Y \subseteq \mathbb{R}^p$. Eine Funktion $f: X \times Y \to \mathbb{R}^m$ heißt *gleichgradig stetig in der ersten Variablen*, falls zu jedem $\varepsilon > 0$ und jedem $x_0 \in X$ ein $\delta > 0$ so existiert, daß $|x - x_0| < \delta$ und $x \in X$ allemal $\|f(x,y) - f(x_0,y)\| < \varepsilon$ zur Folge hat *für alle* $y \in Y$. □

Das Wichtige ist hierbei, daß die Abschätzung simultan für alle $y \in Y$ gilt. Wir haben vorher (in einem typischen Spezialfall) bemerkt, daß eine Funktion $f: X \times Y \to \mathbb{R}^m$ *genau dann gleichgradig stetig ist in der ersten Variablen, wenn die Funktion* $g: X \to C(Y, \mathbb{R}^m)$ *die durch* $g(x)(y) = f(x, y)$ *definiert ist, stetig ist* (*bezüglich der sup-Norm in* $C(Y, \mathbb{R}^m)$). Man kann sich nun überlegen, in welcher Weise der Satz IV.1.18 durch die folgende Konsequenz unserer Diskussion verallgemeinert wird.

Satz 5.4. *Ist* $f: X \times [a, b] \to \mathbb{R}$ *gleichgradig stetig in der ersten, und für jedes feste* x *in der zweiten Variablen integrierbar (d.h.* $t \mapsto f(x, t): [a, b] \to \mathbb{R}$ *ist integrierbar für alle* $x \in X$*). Dann ist die Funktion* $F: X \to \mathbb{R}$, $F(s) = \int_a^b f(x, t)dt$ *stetig.* □

In derselben Weise behandeln wir nun die Differenzierbarkeit. Nach der Kettenregel ist die Hintereinanderausführung differenzierbarer Funktionen differenzierbar, und ihre Ableitung berechnet sich aus der Kettenregel. Also müssen wir nun sicherstellen, daß die Funktion $g: X \to I([a, b], \mathbb{R})$, $X \subseteq \mathbb{R}^n$ differenzierbar ist; dann ist auch die Funktion $F: X \to \mathbb{R}$, $F(x) = \int_a^b f(x, t)dt = J(g(x))$ differenzierbar, und wir haben als Ableitung $F'(x) = J(g'(x))$. Wir müssen uns nun darüber klar werden, was Differenzierbarkeit von g heißt und wie man die Ableitung berechnet. Die Funktion g ist genau dann auf X differenzierbar, wenn wir für jedes $x \in X$ eine stetige lineare Abbildung $g'(x) : \mathbb{R}^n \to (I[a, b], \mathbb{R})$ so finden, daß $g(y) = g(x) + g'(x)(y - x) + o(\|y - x\|)$ gilt. Schreiben wir $f_1(x, t)(h) = g'(x)(h)(t)$, so bedeutet dies, daß $f_1(x, \cdot)(h)$ für jedes x integrierbar ist und daß $f(x, t) - f(x, t) - f_1(x, t)(y - x)$ von der Form $\|y - x\|R(y - x, t)$ mit einer Funktion R ist, für die $\lim_{h \to 0} R(h, t) = 0$ gilt, aber *gleichmäßig in* t. Wir betrachten der Einfachheithalber wieder den uns hauptsächlich interessierenden Spezialfall $X \subseteq I \subseteq \mathbb{R}$ mit einem Intervall I. Dann bedeutet dies doch

(2) $\qquad f_1(x, t) = \lim_{h \to 0} \frac{1}{h}(f(x + h, t) - f(x, t))$ gleichmäßig in $t \in I$.

Selbstverständlich zeigt (2) insbesondere, daß $f_1(x, t) = \frac{\partial f}{\partial x}(x, t)$ sein muß; das Wesentliche ist aber die Gleichmäßigkeit der partiellen Ableitung in der anderen Variablen. Wir fassen diesen Sachverhalt in der folgenden Definition zusammen:

Definition 5.5. Eine Funktion $f: I \times X \to \mathbb{R}$ heißt bezüglich der zweiten Variablen *gleichgradig differenzierbar* in der ersten Variablen falls (2) gilt, d.h. falls zu jedem $\varepsilon > 0$ und jedem $x \in I$ ein $\delta > 0$ so existiert, daß $\frac{1}{h}(f(x + h, t) - f(x, t)) - (D_1 f)(x, t) < \varepsilon$ gilt für alle $x + h \in I$ und $|h| < \delta$ für alle $t \in Y$.

Die folgende Beobachtung gibt uns nun eine hinreichende Bedingung für die gleichgradige Differenzierbarkeit:

Satz 5.6. *Die Funktion* $f: I \times Y \to \mathbb{R}$ *habe eine partielle Ableitung* $\partial_1 f: I \times Y \to \mathbb{R}$ *nach der ersten Variablen, und* $\partial_1 f$ *sei bezüglich der zweiten Variablen gleichgradig stetig in der ersten Variablen. Dann ist* f *gleichgradig differenzierbar in der ersten Variablen.*

Proof. . Zum Beweis betrachten wir den Differenzenquotienten $\frac{1}{h}(f(x+h,t)-f(x,t))$ und stellen nach dem Mittelwertsatz fest, daß eine Zahl $z(x,h,t)$ zwischen x und $x+h$ so existiert, daß dieser Differenzenquotient gerade gleich $\partial_1 f(z(x,h,t),t)$ ist. Ist nun $\partial_1 f$ gleichgradig stetig, so existiert zu $\varepsilon>0$ und x ein $\delta>0$ derart, daß für $|x-y|<\delta$ (und $y\in I$) allemal $|(\partial_1 f)(y,t)-(\partial_1 f)(x,t)|<\varepsilon$ gilt für alle $t\in Y$. Ist nun $|h|<\delta$, so ist $z(x,h,t)$ ein solches y, woraus die Behauptung folgt. □

Dieser Satz ist deswegen nützlich, weil seine Voraussetzung immer schon dann erfüllt ist, wenn $Y=[a,b]$ ist und die partielle Ableitung $\partial_1 f$ auf $I\times[a,b]$ stetig ist. Dies ist eine Folge des nachstehenden Hilfssatzes:

Hilfssatz 5.7. *Die Funktion $p\colon X\times K\to Z$ sei stetig für metrische Räume X, Z und einen kompakten metrischen Raum K. Dann ist sie auch gleichgradig stetig in der ersten Variablen bezüglich der zweiten.*

Proof. . Angenommen, dies sei nicht der Fall. Dann existiert ein $x\in X$ und ein $\varepsilon>0$ derart, daß eine Folge $(x_n,t_n)\in X\times K$ derart, daß $x=\lim x_n$, aber (*) $d\big(p(x,t_n),p(x_n,t_n)\big)\geq\varepsilon$ gilt (in der Tat wähle man zu jedem $\delta=\frac{1}{n}$ ein $x_n\in X$ mit $d(x,x_n)<\frac{1}{n}$ aber $d\big(p(x_n,t_n),p(x_n,t_n)\big)\geq\varepsilon$ für ein geeignetes t_n, welches nach der Annahme existieren muß.) Nun ist aber K kompakt, und die Folge t_n hat daher einen Häufungspunkt t (Definition III.1.31). Zu dem Punkt $(x,t)\in X\times K$ wählen wir nun aufgrund der Stetigkeit von p ein $\delta>0$ derart, daß aus $d(y,x)<\delta$ und $d(s,t)<\delta$ (und $s\in K$) allemal $d\big(p(y,s),p(x,t)\big)<\varepsilon/2$ folgt. Sei $N\in\mathbb{N}$ so gewählt, daß $n>N$ stets $d(x_n,x)<\delta$ zur Folge hat. Da t ein Häufungspunkt der Folge t_n ist, gibt es mindestens ein $m>N$ mit $d(t,t_n)<\delta$. Für dieses m gilt dann $d\big(p(x_m,t_m),p(x,t_m)\big)\leq d\big(p(x_m,t_m),p(x,t)\big)+d\big(p(x,t),p(x,t_m)\big)<\varepsilon/2+\varepsilon/2$. Dies ist aber ein Widerspruch zu (*). □

Nun fassen wir aber unsere Diskussion zusammen:

Satz 5.8. *Sei $f\colon I\times[a,b]\to\mathbb{R}$ eine Funktion, wobei $I\subseteq\mathbb{R}$ ein Intervall ist; ferner sei f bezüglich der zweiten Variablen in der ersten gleichgradig differenzierbar. Ist nun die Funktion $t\mapsto f(x,t)\colon[a,b]\to\mathbb{R}$ für jedes x Riemannsch integrierbar, so ist auch die partielle Ableitung $t\mapsto(\partial_1 f)(x,t):[a,b]\to\mathbb{R}$ für alle $x\in I$ Riemannsch integrierbar, und das Integral $F(x)=\int_a^b f(x,t)dt$ ist differenzierbar und hat die Ableitung*

$$F'(x)=\int_a^b(\partial_1 f)(x,t)dt.$$

Die Voraussetzung der gleichgradigen Differenzierbarkeit von f ist insbesondere dann erfüllt, falls die partielle Ableitung $\partial_1 f\colon I\times[a,b]\to\mathbb{R}$ stetig ist. □

Mit diesem Satz ist insbesondere die noch offen gebliebene Beweislücke von Satz 5.2 geschlossen.

Wie man ein Vektorfeld als infinitesimale Transformation auffaßt

Sei $E=\mathbb{R}^n$ und M eine offene Teilmenge von E, und sei $f\colon M\to E$ eine differenzierbres Funktion. (Man nennt eine solche Funktionen ja ein Vektorfeld). Wir wissen dann

$$f(x+h)=f(x)+f'(x)(h)+\mathrm{o}(h),\quad \lim_{h\to 0}\|h\|^{-1}\mathrm{o}(h)=0.$$

Halten wir momentan das Element x fest und setzen wir $h=f(x)$. Dann bemerken wir

$$\begin{aligned} f(x+t{\cdot}h)&=f(x)+t{\cdot}f'(x)\big(h\big)+\mathrm{o}(t),\\ \big(\exp t{\cdot}f'(x))\big)\big(h\big)&=f(x)+t{\cdot}f'(x)\big(h\big)+\mathrm{o}(t).\end{aligned}$$

Daher haben wir

$$f(x+t{\cdot}h)=\big(\exp t{\cdot}f'(x)\big)(h)+\mathrm{o}(t).$$

Die Änderung des Vektors $f(x)$ bei Ersetzung des Arguments x zu $x+t{\cdot}h=x+t{\cdot}f(x)$ ist implementiert durch die Anwendung der linearen Abbildung $\exp t{\cdot}f'(x)$ auf h.

Setzen wir $e=\|h\|^{-1}{\cdot}h$ unter der Annahme $h=f(x)\neq 0$, dann haben wir

$$\begin{aligned}\partial_{x;e}f&=\frac{d}{dt}f(x+t{\cdot}e)\bigg|_{t=0}\\ &=\frac{d}{dt}f(x+t\|h\|^{-1}{\cdot}h)\bigg|_{t=0}=\frac{d}{dt}\big(\exp t\|h\|^{-1}{\cdot}f'(x)\big)(h)\bigg|_{t=0}=f'(x)(e).\end{aligned}$$

D.h., $f'(x)$ ist die dazugehörige *infinitesimale Transformation*.

Ist $X\colon E\to E$ eine lineare Abbildung, so ist der symmetrische Anteil $X_s=\frac{1}{2}{\cdot}(X+X^\top)$ und der antisymmetrische $X_a=\frac{1}{2}{\cdot}(X-X^\top)$, und es gilt $X=X_a+X_s$. Dann ist

$$\exp t{\cdot}X=\exp t{\cdot}(X_a+X_s)=\big(\exp t{\cdot}X_a\big)\big(\exp t{\cdot}X_s\big)+\mathrm{o}(t).$$

Somit

$$f(x+t{\cdot}h)=\big(\exp t{\cdot}f'(x)_a\big)\big(\exp t{\cdot}f'(x)_s\big)(h)+\mathrm{o}(t).$$

Es existiert eine Drehung T auf $\mathbb{R}^n$ derart, daß die Matrix von $Tf'(x_s)T^{-1}$ gerade die Diagonalmatrix

$$\begin{pmatrix}\lambda_1 & 0 & \cdots & 0\\ 0 & \lambda_2 & \cdots & 0\\ \vdots & \vdots & \cdots & \vdots\\ 0 & 0 & \cdots & \lambda_n\end{pmatrix}.$$

Damit ist

$$T\big(\exp t{\cdot}f'(x)\big)T^{-1}=\exp\big(Tf'(x)T^{-1}\big)=\begin{pmatrix}e^{t\lambda_1} & 0 & \cdots & 0\\ 0 & e^{t\lambda_2} & \cdots & 0\\ \vdots & \vdots & \cdots & \vdots\\ 0 & 0 & \cdots & e^{t\lambda_n}\end{pmatrix}.$$

Ferner ist $D = \exp t{\cdot}f'(x)_a$ wegen der Antisymmetrie eine Drehung, denn $D^\top = \exp t{\cdot}f'(x)_a^\top = \exp -t{\cdot}f'(x)_a = D^{-1}$, und $\det\bigl(\exp t{\cdot}f'(x)_a\bigr) = e^{t\,\mathrm{Tr}\,f'(x)_a} = e^0 = 1$.

Man nennt daher $2f'(x)_a = f'(x) - f'(x)^\top$ auch die *Rotation* von f bei x. Bis auf den Faktor 2 ist sie der Drehanteil der infinitesimalen Transformation. Es ist eine nützliche Übungsaufgabe, die Rotation im Fall $n = 2$ und $n = 3$ explizit hinzuschreiben. Ein Vektorfeld heißt dementsprechend *rotationsfrei*, wenn die Rotation verschwindet, d.h. wenn $f'(x)$ symmetrisch ist. Der Satz 5.2 läßt sich dementsprechend auch umformulieren:

Für ein Vektorfeld auf einer offenen Menge des $\mathbb{R}^n$ ist das Vorhandensein eines Potentials hinreichend dafür, daß es rotationsfrei ist. Falls in der offenen Menge jeder Punkt mit einem festen Punkt geradlinig verbunden werden, ist die Existenz eines Potentials auch notwendig.

Wir bemerken noch

$$\det\bigl(\exp t{\cdot}f'(x)\bigr) = e^{t\,\mathrm{Tr}\,f'(x)} = e^{t \div_x f}.$$