8. FUNKTIONALABLEITUNG

8.1. Physikalische Bedeutung

In den letzten Abschnitten haben wir den Umgang mit linearen Operatoren bzw. linearen Funktionalen gelernt. Nun spielen aber sehr oft in der Physik nicht-lineare Operatoren eine Rolle. Als Beispiel haben wir in § 1.4. den Boltzmannschen Stoßoperator kennengelernt; weitere Beispiele sind:

a) Extremalprinzipien für Bewegungsgleichungen.

Man berechnet etwa in der klassischen Punktmechanik die W i r k u n g s - f u n k t i o n $\mathfrak{L}$ eines Teilchens, welches die Bahn $x(t)$ mit einer Geschwindigkeit $\dot{x}(t)$ zurücklegt:

$$\mathfrak{L} = \int_{t_1}^{t_2} L(t, x(t), \dot{x}(t))\, \mathrm{d}t \,. \tag{8.1}$$

Die Wahl der L a g r a n g e - F u n k t i o n L charakterisiert das speziell vorgegebene System. Allgemein gesehen vermitteln L und auch $\mathfrak{L}$ eine Abbildung einer Funktion, nämlich der Bahn $x(t)$, auf eine reelle Zahl, d.h. $\mathfrak{L}(x)$ ist ein Funktional! Es ist allerdings ein nicht-lineares Funktional. (Z.B. für ein Teilchen im Potential $V(x)$ lautet bekanntlich $L = \frac{m}{2}\dot{x}^2 - V(x)$.)

Nun braucht man aber nicht das ganze Funktional $\mathfrak{L}$, sondern die "Stelle" $x(t)$, an der es extremal ≡ stationär wird. Eben diese Bahn $x(t)$ ist diejenige, die in der Natur realisiert wird. Es liegt also nahe, das Funktional $\mathfrak{L}$ um diese Stelle zu "entwickeln", ähnlich wie auch im Falle des nicht-linearen Boltzmannschen Stoßoperators eine Entwicklung um die Gleichgewichtsverteilung $f_E(\vec{p})$ zweckmäßig war. Eine solche Entwicklung führt man bei Funktionen mit Hilfe der Ableitung durch, eventuell sogar einer mehrgliedrigen Taylorreihe. Somit erhebt sich die Frage nach einer V e r a l l g e m e i n e r u n g des Begriffes der Ableitung einer Funktion auf den der A b l e i t u n g e i n e s F u n k t i o n a l s .

b) Als weiteres Beispiel eines nicht-linearen Funktionals sei die thermodynamische Freie Energie pro Teilchen f erwähnt. Sie hängt von der Wechselwirkungsenergie $w(\vec{r})$ zwischen den das System konstituierenden Teilchen ab, wobei alle möglichen Abstände $\vec{r} = \vec{r}_i - \vec{r}_j$ der Teilchenpaare i,j vorkommen. Also ist $f = f(w)$ ein (nicht-lineares) Funktional der Wechselwirkung; natürlich auch eine Funktion von Temperatur T und Dichte $n \equiv v^{-1}$. Bei Anwendung der klassischen Physik ist

$$f = f_{\text{ideal}}(T, n) - \kappa\, T \ln \frac{1}{V^N} \int_V e^{-\frac{1}{\kappa T} \sum_{i<j} w(\vec{r}_i - \vec{r}_j)}\, d\vec{r}_1 \ldots d\vec{r}_N\,. \tag{8.2}$$

Der erste Summand, f_{ideal}, beschreibt ein ideales Gas und hängt folglich nicht von $w(\vec{r})$ ab; κ ist die Boltzmannkonstante, V das Volumen.

Wiederum interessiert man sich für die Veränderungen von f bei Veränderung seiner Variablen. Sie haben alle eine wichtige physikalische Bedeutung: $-\partial f/\partial T$ ist die Entropie pro Teilchen, $-\partial f/\partial V$ ist der Druck und $\partial f/\partial w(\vec{r})$ ist die Paarverteilungsfunktion. Wie aber hat man die "Ableitung nach einer Funktion", $w(\vec{r})$, zu verstehen? Gerade in der statistischen Physik braucht man sie häufig!

c) Nicht nur nicht-lineare Funktionale, also Abbildungen nach K, spielen eine Rolle in der Physik. Auch nicht-lineare Operatoren, also Abbildungen in allgemeinere Räume $\mathfrak{M}$, treten auf. Schon der Boltzmannsche Stoßoperator war hierfür ein Beispiel. Er beschreibt die Stoß-Übergänge an der Stelle $\vec{r}, \vec{p}$ im Phasenraum, bildet also auf eine Funktion von $\vec{r}, \vec{p}$ ab als Funktional der Verteilungsfunktion $f(\vec{r}', \vec{p}', t)$. Oder: In der Quantenmechanik berechnet man Wirkungsquerschnitte mit der T-Matrix, den Matrixelementen $\langle p_1 p_2 | T | p_1' p_2' \rangle$ des Operators $T(z)$ zwischen den Impulszuständen p_1, p_2 vor bzw. p_1', p_2' nach dem Stoß bei der vor und nach dem Stoß gleichen Energie $z = E$. Der T-Operator wird nicht-linear vom Wechselwirkungsoperator V bestimmt ($H = H_{\text{kin}} + V$, Hamiltonoperator):

$$T = V - V \frac{1}{H - z} V = V - V \frac{1}{H_{\text{kin}} - z} T\,. \tag{8.3}$$

Damit ist die Bedeutsamkeit des Begriffes der *Funktionalableitung* klar, d.h. der Ableitung eines beliebig vorgegebenen Operators A auf $\mathfrak{D}_A \subseteq \mathfrak{M}$ nach seinem "Argument", einem Element $f \in \mathfrak{M}$; in der Regel ist A ein Funktional, daher der Name. Wir werden sie in völliger Analogie zur gewöhnlichen Ableitung definieren, um den Umgang mit ihr leicht und bequem zu machen. Es wird sich dabei zeigen, daß die Funktionalableitung eine Distribution ist, weshalb sich ihre Untersuchung gerade im Anschluß an § 7 anbietet. Beachtet man diesen Distributionscharakter, ist der Gebrauch der Funktionalableitung dem Physiker ohne mathematische Probleme an die Hand gegeben!

8.2. Definition der Funktionalableitung

Sei A auf $\mathfrak{D}_A$ ein Operator, der von $\mathfrak{M}$ nach $\mathfrak{M}'$ vermittelt. Wir setzen zur Vereinfachung voraus, $\mathfrak{M}$ und $\mathfrak{M}'$ seien Banachräume, obwohl eine Verallgemeinerung auf lineare, metrische Räume bei gewissen Eigenschaften der Metrik möglich ist. Sei $f \in \mathfrak{D}_A$; wir fragen uns, wie sich $Af \in \mathfrak{W}_A \subseteq \mathfrak{M}'$ verändert, wenn man f "etwas"

verändert. Dabei heißt "etwas" verändern, daß wir $f + h \in \mathfrak{D}_A$ mit "kleinem" $h \in \mathfrak{M}$ betrachten, d.h. $\| h \|$ klein. Wenn $\| h \| \to 0$, sollte $A(f+h) \Rightarrow Af$ sein. Sonst ist A an der Stelle f nicht einmal stetig, und dann kann man auch in der gewöhnlichen Analysis keine Ableitung definieren. Wie aber konvergiert $A(f+h)$ gegen Af? In Analogie zu differenzierbaren Funktionen könnte das in h linear sein. Da hier aber h ein Punkt aus einem Banachraum $\mathfrak{M}$ ist, heißt linear genauer: Es sollte

$$A(f+h) \approx Af + Lh, \quad \text{wenn} \quad h \approx 0\,, \tag{8.4}$$

sein, wobei L ein linearer Operator ist. Selbstverständlich hängt L von A und der Stelle f ab, weshalb wir diese Abhängigkeit durch die Schreibweise $L \equiv \frac{\delta A}{\delta f}$ symbolisieren und als *Ableitung von A an der Stelle f* bezeichnen. $\frac{\delta A}{\delta f}$ ist ein linearer Operator, ja wir fordern definitionsgemäß sogar die Linear-Beschränktheit (s. § 5.4.).

Definition: *Ein Operator A* **auf $\mathfrak{D}_A$ heißt an der Stelle** $f \in \mathfrak{D}_A$ *differenzierbar,* **wenn es einen linear-beschränkten Operator $\frac{\delta A}{\delta f}$ gibt, so daß**

$$\| A(f+h) - Af - \frac{\delta A}{\delta f} h \| \leqslant \| h \| \, \epsilon(\| h \|), \quad \textbf{für alle } h \textbf{ mit } \| h \| < \delta\,. \tag{8.5}$$

Dabei ist ϵ eine Nullfunktion, $\epsilon(x) \to 0$ mit $x \to 0$.

Die Norm auf der linken Seite ist diejenige in $\mathfrak{M}'$, $\|h\|$ ist diejenige in $\mathfrak{M}$. In linearen, metrischen Räumen hätte man $\rho(A(f+h) - Af, \frac{\delta A}{\delta f} h)$ zu untersuchen bzw. $\rho(h, 0)$. Nur solche h werden verwendet, für die $f + h \in \mathfrak{D}_A$. Doch soll $\frac{\delta A}{\delta f}$ als linear-beschränkter Operator überall definiert sein. Man beachte, daß $\frac{\delta A}{\delta f}$ von $f \in \mathfrak{M}$ abhängt und als linearer Operator über $\mathfrak{M}$ wirkt. Bezüglich f ist die Funktionalableitung selbstverständlich nicht notwendig linear! Man bezeichnet diese Ableitung auch als Variationsableitung oder *Fréchet-Ableitung*, da sie zuerst von *M. Fréchet*, 1910, eingeführt wurde (Ann. Ec. Norm. 27, 1910, 3; 193 - 216).

Wir überzeugen uns zunächst, daß die Definition der Ableitung eines Operators sinnvoll ist. Wenn (!) es eine gibt, so ist sie nämlich eindeutig. Gäbe es zwei lineare Operatoren L_1, L_2, die der Definition genügen, so müssen sie gleich sein:

$$\| L_1 h - L_2 h \| \leqslant \| L_1 h - A(f+h) + Af \| + \| L_2 h - A(f+h) + Af \| \leqslant$$
$$\leqslant \| h \| \, (\epsilon_1(\| h \|) + \epsilon_2(\| h \|))$$

$$\curvearrowright \| L_1 - L_2 \| \leqslant \epsilon_1(\| h \|) + \epsilon_2(\| h \|), \curvearrowright \| L_1 - L_2 \| = 0\,.$$

Ein differenzierbarer Operator ist erst recht stetig. Denn aus (8.5) folgt

$$\| A(f+h)-Af \| \leqslant \| h \| \, (\epsilon(\| h \|) + \| \frac{\delta A}{\delta f} \|) \to 0 \quad \text{mit} \quad \| h \| \to 0 .$$

Sofern A auf $\mathfrak{D}_A$ auf die komplexen Zahlen abbildet, ist es ja Funktional, $\phi(f)$. Deshalb bezeichnen wir die Operatorableitung (8.5) auch als *Funktionalableitung*. Als linear-stetiges Funktional hat folglich $\frac{\delta\phi}{\delta f}$ seine Bedeutung als Anzuwendendes auf die Elemente aus $\mathfrak{M}$, $\langle \frac{\delta\phi}{\delta f} | h \rangle$. Speziell über Funktionenräumen $\mathfrak{M}$ schreiben wir das Innere Produkt mit $h \in \mathfrak{M}$ und $\frac{\delta\phi}{\delta f} \in \mathfrak{M}^*$ als (eventuell formales!) Integral:

$$\langle \frac{\delta\phi}{\delta f} | h \rangle = \int \frac{\delta\phi}{\delta f(\xi)} \, h(\xi) \, d\xi . \tag{8.6}$$

(8.6) verleiht der Schreibweise $\frac{\delta\phi}{\delta f(x)}$ für die Funktionalableitung einen Sinn, gelesen als Änderung des Funktionals $\phi(f)$ bei Änderung der Funktion f an der Stelle x. Diese Deutung lesen wir aus (8.6) ab, indem wir $h(\xi)$ nur in der Umgebung von x von Null verschieden wählen, also $f + h$ nur an der Stelle x gegenüber f verändert ist.

8.3. Beispiele

a) Betrachten wir die Wirkung $\mathfrak{L}$ gemäß (8.1) als Funktional der Bahnkurve $x(t) \in C^1(t_1, t_2)$, so können wir die Funktionalableitung $\frac{\delta \mathfrak{L}}{\delta x(t)}$ leicht bestimmen.

$$\mathfrak{L}(x+\xi) - \mathfrak{L}(x) = \int_{t_1}^{t_2} L(t, x(t)+\xi(t), \dot{x}(t)+\dot{\xi}(t)) \, dt - \int_{t_1}^{t_2} L(t, x(t), \dot{x}(t)) \, dt =$$

$$= \int_{t_1}^{t_2} \left[\frac{\partial L}{\partial x}(t) \, \xi(t) + \frac{\partial L}{\partial \dot{x}}(t) \, \dot{\xi}(t) + \tfrac{1}{2} \xi^2 \frac{\partial^2 L}{\partial x^2}(t) + \ldots + \tfrac{1}{2} \dot{\xi}^2 \frac{\partial^2 L}{\partial \dot{x}^2}(t) \right] dt .$$

Die Lagrangefunktion L ist als zweimal stetig differenzierbar vorausgesetzt worden. Die Schreibweise $\frac{\partial L}{\partial x}(t)$ heißt genauer $\left. \frac{\partial L(t,y,z)}{\partial y} \right|_{y=x(t),\, z=\dot{x}(t)}$ usw. In den quadratischen Gliedern wähle man y, z an einem geeigneten Zwischenwert; sie sind offenbar von der Größenordnung $\| \xi \|^2$. Also

$$\mathfrak{L}(x+\xi) - \mathfrak{L}(x) = \int_{t_1}^{t_2} \left[\frac{\partial L}{\partial x}(t) - \frac{d}{dt} \frac{\partial L}{\partial \dot{x}}(t) \right] \xi(t) \, dt + \xi(t) \frac{\partial L}{\partial \dot{x}}(t) \Big|_{t_1}^{t_2} + \sim \| \xi \|^2 .$$

Die Funktionalableitung lautet also - einfach aus $\int \ldots \xi(t) \, dt$ abzulesen -

$$\frac{\delta \mathfrak{L}}{\delta x(t)} = \frac{\partial L}{\partial x}(t, x(t), \dot{x}(t)) - \frac{d}{dt} \frac{\partial L}{\partial \dot{x}}(t, x(t), \dot{x}(t)) + [\delta(t-t_2) - \delta(t-t_1)] \frac{\partial L}{\partial \dot{x}}(t, x(t), \dot{x}(t)). \tag{8.7}$$

Beschränkt man sich auf Änderungen $\xi(t) \in C_0^1(t_1, t_2)$, die am Rande des Zeitintervalls $[t_1, t_2]$ Null sind, so fallen die Randterme weg:

$$\frac{\delta \mathfrak{L}}{\delta x(t)} = \frac{\partial L}{\partial x}(t, x(t), \dot{x}(t)) - \frac{d}{dt} \frac{\partial L}{\partial \dot{x}}(t, x(t), \dot{x}(t)) . \tag{8.7'}$$

Das Extremalprinzip der Punktmechanik besagt, daß diejenige Bahn $x(t)$ realisiert wird, für die die Wirkung extremal ist, d.h. $\frac{\delta \mathfrak{L}}{\delta x(t)} = 0$. Folglich gilt gemäß (8.7') die Lagrangesche Gleichung

$$\frac{\partial L}{\partial x} - \frac{d}{dt} \frac{\partial L}{\partial \dot{x}} = 0 . \tag{8.8}$$

Sie ist genaugenommen eine Distributionsgleichung, was aber bei den physikalisch interessierenden Lagrangefunktionen L deshalb nicht auffällt, weil reguläre Distributionen vorliegen.

b) Sei $\phi(f)$ die Zuordnung $\phi(f) =: f(y) \in K$. Dann ist $\phi(f+h) - \phi(f) = h(y) \equiv \equiv \int \frac{\delta\phi}{\delta f(x)} \, h(x) \, dx$. Folglich existiert die Funktionalableitung, und es gilt die einfache aber nützliche Formel für $\frac{\delta\phi}{\delta f(x)} \equiv \frac{\delta f(y)}{\delta f(x)}$:

$$\boxed{\frac{\delta f(y)}{\delta f(x)} = \delta(x-y)} \; . \tag{8.9}$$

Wenden wir am Beispiel (a) die Kettenregel an sowie (8.9), finden wir sogleich (8.8). Ob dieses einfache Verfahren erlaubt ist, überlegen wir uns in § 8.5.2. Da die Handhabung jedoch so zweckmäßig und leicht ist, üben wir sie im nächsten Beispiel.

c) Feldgleichungen aus dem Extremalprinzip.

$$\mathfrak{L}(f) = \int d^4x \, L(x, f(x), f_{|\mu}(x)). \qquad \frac{\delta \mathfrak{L}}{\delta f(x)} = 0 \curvearrowright \frac{\partial L}{\partial f} - \frac{\partial}{\partial x^\mu} \frac{\partial L}{\partial f_{|\mu}} = 0 . \tag{8.10}$$

c1) Für ein reelles, skalares Feld setzen wir für L:

$L =: \frac{1}{2}\,(\varphi_{|\nu}\,\varphi^{|\nu} - m^2\varphi^2)$. Folgt $(\Box + m^2)\,\varphi = 0$, Klein-Gordon-Gleichung. (8.11)

c2) Komplexes Spinorfeld; nach ψ und $\overline{\psi}$ ableiten.

$L = -\frac{1}{2}\,\overline{\psi}(-i\gamma^\mu\partial_\mu + m)\,\psi - \frac{1}{2}\,(i\partial_\mu\overline{\psi}\gamma^\mu + m\overline{\psi})\,\psi$.

$$\curvearrowright \frac{\partial L}{\partial\overline{\psi}} = (\tfrac{1}{2}\,i\gamma^\mu\partial_\mu - m)\,\psi \quad \text{und} \quad \frac{\partial L}{\partial\overline{\psi}_{|\mu}} = -\tfrac{1}{2}\,i\gamma^\mu\,\psi$$

$\curvearrowright$ Dirac-Gleichung: $(-i\gamma^\mu\partial_\mu + m)\,\psi = 0$. (8.12)

c3) Komplexes (geladenes), skalares Klein-Gordon-Feld φ, angekoppelt an ein elektromagnetisches Potentialfeld A^μ mit dem Feldstärketensor $F_{\mu\nu} = A_{\nu|\mu} - A_{\mu|\nu}$.

$L =: -\frac{1}{4}\,F^{\mu\nu}F_{\mu\nu} - m^2\varphi^*\varphi + (\varphi^*_{|\mu} + ieA_\mu\varphi^*)\,(\varphi^{|\mu} - ieA^\mu\varphi)$.

Ohne φ-Feld ergeben sich aus (8.10) die Maxwellschen Gleichungen ohne Ladungen oder Ströme. Da $\partial L/\partial A_\mu = 0$, ist nämlich

$$0 = -\frac{\partial}{\partial x^\mu}\,\frac{\partial L}{\partial A_{\nu|\mu}} = \ldots = A^{\nu|\mu}{}_{|\mu} - A^{\mu|\nu}{}_{|\mu} = F^{\mu\nu}{}_{|\mu} \,. \qquad (8.13)$$

Berücksichtigt man φ, so erhält man

$$\frac{\partial L}{\partial A_\nu} = ie\,[\varphi^*(\varphi^{|\nu} - ieA^\nu\varphi) - (\varphi^{*|\nu} + ieA^\nu\varphi^*)\varphi] =: -j^\nu \,,$$

d.h.

$$F^{\mu\nu}{}_{|\mu} = \Box A^\nu = j^\nu \,. \qquad (8.14)$$

d) Der Zusammenhang (1.22) zwischen statischer elektrischer Ladungsverteilung $\rho(\vec{r}\,')$ und dem elektrostatischen Potential $\varphi(\vec{r})$ ist linear.

$$\phi(\rho) \equiv \varphi(\vec{r}) = \int \frac{1}{|\vec{r} - \vec{r}\,'|}\,\rho(\vec{r}\,')\,d\vec{r}\,' \,. \qquad (8.15)$$

Die Existenz der Funktionalableitung ist leicht zu sehen. Es gilt

$$\frac{\delta\varphi(\vec{r})}{\delta\rho(\vec{r}\,')} = \frac{1}{|\vec{r} - \vec{r}\,'|} \,. \qquad (8.16)$$

Da diesmal ϕ linear in ρ ist, hängt $\frac{\delta\phi}{\delta\rho}$ nicht mehr von ρ ab. Die Distribution (8.16) ist für $\vec{r} \neq \vec{r}\,'$ mit einer Funktion zu identifizieren. Die Ableitung der Distribution (8.16) haben wir in § 7.5., Beispiel 7, gelernt.

$$\Delta_r\varphi(\vec{r}) = \Delta_r \int \frac{\delta\varphi(\vec{r})}{\delta\rho(\vec{r}\,')}\,\rho(\vec{r}\,')\,d\vec{r}\,' =$$

$$= \int \Delta_{r'}\,\frac{1}{|\vec{r} - \vec{r}\,'|}\,\rho(\vec{r}\,')\,d\vec{r}\,' = -4\pi \int \delta(\vec{r} - \vec{r}\,')\,\rho(\vec{r}\,')\,d\vec{r}\,' = -4\pi\rho(\vec{r}) \,;$$

im Distributionssinn also eine einfache Rechnung.

8.4. Die Richtungsableitung von Operatoren

In manchen Anwendungen ist es zweckmäßig, die Funktionalableitung auf die gewöhnliche Ableitung dadurch zurückzuführen, daß man $A(f + \tau h)$ als Funktion des Zahlenparameters τ auffaßt. Dann kann man das *Richtungsdifferential des Operators A an der Stelle f* definieren durch

$$\partial A(f;\,h) =: \lim_{\tau\to 0} \frac{A(f + \tau h) - A(f)}{\tau} \,, \qquad (8.17)$$

sofern dieser Limes existiert. $\partial A(f;\,h)$ ist eine zweifache Abbildung von $\mathfrak{M}$ nach $\mathfrak{M}'$. Sie hängt nicht nur von "der Stelle" f ab, sondern auch von "der Richtung" h.

Über die Abhängigkeit des Richtungsdifferentials von der Richtung $h \in \mathfrak{M}$ kann man allgemein nichts sagen. Sie könnte linear sein, muß es aber nicht. Wenn (!) $\partial A(f;\,h)$ bezüglich h durch einen linearen Operator zu beschreiben ist,

$$\partial A(f;\,h) \overset{!}{=} \frac{\partial A}{\partial f}\,h \equiv \langle \frac{\partial A}{\partial f} | h \rangle \,, \qquad (8.18)$$

so nennen wir den linearen Operator $\frac{\partial A}{\partial f}$ über $\mathfrak{M}$ *die Richtungsableitung.*

Der Zusammenhang mit der allgemeinen Funktionalableitung wird durch folgende Aussagen klargestellt.

Satz: a) Wenn (!) A an der Stelle f differenzierbar ist, so existiert erst recht die Richtungsableitung, und beide sind gleich:

$$\frac{\delta A}{\delta f} = \frac{\partial A}{\partial f} \,, \quad \textbf{linear-beschränkt} \,. \qquad (8.19)$$

b) Existiert (!) das Richtungsdifferential $\partial A(f;\,h)$ in einer Umgebung $\| f - f' \| \leqslant r$ von f für alle h und ist es in h stetig sowie in f gleichmäßig (bezüglich f, h) stetig, dann ist A in der genannten Umgebung differenzierbar; konsequenterweise, wegen (8.19), stimmt die Ableitung mit der Richtungsableitung überein.

Durch Prüfung der Zusatz-Eigenschaften des Richtungsdifferentials kann man also das in der Physik oft benutzte Vorgehen rechtfertigen und die Funktionalableitung recht einfach berechnen. Eventuell genügt auch schon die schwächere

Eigenschaft der Richtungs-Differenzierbarkeit in einer physikalisch ausgezeichneten Richtung.

Beweis, zu a) Es möge $\frac{\delta A}{\delta f}$ existieren. Dann ist gemäß (8.5) $A(f+\tau h) = Af + \frac{\delta A}{\delta f}\tau h + g$ mit $\|g\| \leqslant \tau \|h\| \,\epsilon(\tau \|h\|)$. Folglich existiert $\lim_{\tau \to 0} \frac{1}{\tau} g = 0$ und deshalb $\partial A(f, h) = \frac{\delta A}{\delta f} h$, wie man direkt nachrechnet.

Zu b) Zuerst einmal muß geprüft werden, ob das Richtungsdifferential überhaupt bezüglich h linear ist. Dies garantieren die Zusatzeigenschaften in der Tat. Man sieht es so:

$$\partial A(f+\tau h, h) = \lim_{\tau' \to 0} \frac{A(f+\tau h+\tau' h) - A(f+\tau h)}{\tau'} = \frac{d}{d\tau'} A(f+\tau h+\tau' h)\Big|_{\tau'=0} = \frac{d}{d\tau} A(f+\tau h).$$

$$\curvearrowright A(f+th) - Af = \int_0^t d\tau\, \partial A(f+\tau h, h) \quad (\pm \partial A(f,h))$$

$$= t\, \partial A(f,h) + g\,,$$

wobei $g \equiv \int_0^t [\partial A(f+\tau h, h) - \partial A(f,h)]\, d\tau$. Wegen der Stetigkeit der Richtungsableitung im ersten Argument gilt für hinreichend kleines t, daß $\|g\| \leqslant t\, \epsilon(t)$, mit $\epsilon \to 0$ für $t \to 0$. Also gilt die Darstellung

$$\partial A(f,h) = \frac{1}{t}[A(f+th) - Af - g(t)]\,. \tag{8.20}$$

Hiermit folgt leicht die Additivität des Richtungsdifferentials bzgl. h:

$$\partial A(f, h_1) + \partial A(f+th_1, h_2) - \partial A(f, h_1+h_2) = \frac{1}{t}[-g_1 - g_2 + g_3] \Rightarrow 0\,,$$

wegen der gleichmäßigen Stetigkeit im ersten Argument. Aus der Additivität folgern wir mittels der Stetigkeit in h die volle Linearität $\partial A(f,h) = \frac{\partial A}{\partial f} h$. Damit ist der 1. Beweisschritt ausgeführt.

Nun noch die Fréchet-Differenzierbarkeit: (8.20) mit $t = 1$ liefert sie, sofern $\|g(1)\| \leqslant \leqslant \|h\|\, \epsilon(\|h\|)$ gezeigt werden kann.

$$\|g\| \leqslant 1 \max_{0 \leqslant \tau \leqslant 1} \left\|\frac{\partial A}{\partial(f+\tau h)} - \frac{\partial A}{\partial f}\right\| \; \|h\|\,.$$

Wiederum wegen der gleichmäßigen Stetigkeit im ersten Argument geht die Norm der Richtungsableitungs-Differenz mit $\|h\| \to 0$, q.e.d.

8.5. Rechenregeln für die Funktionalableitung

8.5.1. Höhere Ableitungen

Die Funktionalableitung $\frac{\delta A}{\delta f}$ ist ein linearer Operator über $\mathfrak{M}$, der i.a. von der Stelle f abhängt, an der man die Ableitung gebildet hat. Bei f e s t e m h ist also $\frac{\delta A}{\delta f} h \in \mathfrak{M}'$ wiederum eine Abbildung der f aus $\mathfrak{M}$ nach $\mathfrak{M}'$, die man nach der gerade gelernten Regel differenzieren kann. Durch

$$\left\|\frac{\delta A}{\delta f} h\Big|_{f+k} - \frac{\delta A}{\delta f} h\Big|_f - \frac{\delta^2 A}{\delta f^2} hk\right\| \leqslant \|k\|\, \epsilon(\|k\|) \tag{8.21}$$

wird die 2. Ableitung $\frac{\delta^2 A}{\delta f^2}$ als doppelt linear-beschränkter Operator über $\mathfrak{M}$ eindeutig definiert, sofern sie existiert.

$$\frac{\delta^2 A}{\delta f^2} hk \equiv \frac{\delta}{\delta f}\left(\frac{\delta A}{\delta f} h\right) k\,.$$

Analog ist die n. Funktionalableitung als n-facher linear-beschränkter Operator über $\mathfrak{M}$ zu verstehen, $\frac{\delta^n A}{\delta f^n} h_1 h_2 \ldots h_n \in \mathfrak{M}'$. Als Inneres Produkt geschrieben definiert dieses Distributionen von n Argumenten,

$$\frac{\delta A}{\delta f(x)},\ \frac{\delta^2 A}{\delta f(x_1)\, \delta f(x_2)},\ \ldots,\ \frac{\delta^n A}{\delta f(x_1) \ldots \delta f(x_n)}\,.$$

Interessant ist der Fall, daß die Funktionalableitungen eines Operators A gebildet werden, der selbst linear-beschränkt ist. Dann ist $A(f+h) - Af = Ah$, d.h. $\frac{\delta A}{\delta f} = A$, unabhängig von f. Die zweite Ableitung ist Null, denn $\frac{\delta A}{\delta(f+k)} h - \frac{\delta A}{\delta f} h = Ah - Ah = 0$. Von linear-beschränkten Operatoren A existiert also stets die 1. Fréchet-Ableitung, ist gleich A, und alle höheren Ableitungen sind Null. Als physikalisches Beispiel diene etwa § 8.3.d.

8.5.2. Kettenregel und Produktregel

Die aus der Analysis bekannten Ableitungen einer mittelbaren Funktion $f(g(x))$ bzw. eines Funktionenproduktes $f(x)g(x)$ sind zwei Spezialfälle von Ableitungen hintereinander auszuführender Abbildungsvorschriften. Da wir das Hintereinanderausführen als Operatorprodukt definiert haben, wollen wir jetzt die allgemeine Ableitung eines Operatorproduktes untersuchen. Damit ist sowohl die Ketten- als auch die Produktregel erfaßt. Sie lautet:

$$\frac{\delta(AB)}{\delta f} = \frac{\delta A}{\delta(Bf)}\,\frac{\delta B}{\delta f}\,. \tag{8.22}$$

Dabei ist natürlich vorausgesetzt, daß die Faktoren A und B einzeln differenzierbar sind. Die Abbildungsverhältnisse $\mathfrak{M} \to \mathfrak{M}' \to \mathfrak{M}''$ mache sich der Leser klar.

Zum Beweis schätzen wir $\| AB(f+h) - ABf - \frac{\delta A}{\delta Bf}\frac{\delta B}{\delta f} h \|$ ab, indem wir $\pm \frac{\delta A}{\delta g} k$ einschieben, wobei $g \equiv Bf$ und $k \equiv B(f+h) - Bf$.

$$\|\ldots\| \leqslant \| A(g+k) - Ag - \frac{\delta A}{\delta g} k \| + \| \frac{\delta A}{\delta g}(k - \frac{\delta B}{\delta f} h)\|$$

$$\leqslant \| k \|\, \epsilon_1(\| k \|) + \| \frac{\delta A}{\delta g} \|\ \| h \|\, \epsilon_2(\| h \|), \text{ da } \frac{\delta A}{\delta g} \text{ linear-beschränkt.}$$

Beachten wir noch $\| k \| = \| B(f+h) - Bf \pm \frac{\delta B}{\delta f} h \| \leqslant \| h \|\, \epsilon_3(\| h \|) + \| \frac{\delta B}{\delta f} \|\ \| h \|$, so ist $\| k \| \leqslant \text{const} \| h \|$ und damit alles klar.

Beispiel: A differenzierbar, L linear-beschränkt. Dann ist

$$\frac{\delta(LA)}{\delta f} = L\frac{\delta A}{\delta f} \quad \text{und} \quad \frac{\delta(AL)}{\delta f} = \frac{\delta A}{\delta Lf} L\,. \tag{8.23}$$

Der lineare Operator L kann also wie ein konstanter Faktor behandelt werden.

8.5.3. Mittelwertsatz und Taylorentwicklung

Der große Nutzen der Taylorentwicklung von Funktionen für vielerlei Anwendungen ist bekannt. Daher wollen wir sie auch für die allgemeine Operatorableitung kennenlernen. Eine vereinfachte Vorstufe ist der Mittelwertsatz, an dem uns alles Wesentliche klar wird.

Der klassische Mittelwertsatz lautet ja: Sei $\varphi(x)$ stetig differenzierbar in $[a, b]$, dann ist für geeignetes $\tau \in [0, 1]$

$$\varphi(x+h) - \varphi(x) = h\varphi'(x+\tau h)\,.$$

Dieser Satz kann für Operatoren, die ja verallgemeinerte Funktionen sind, i.a. n i c h t stimmen, denn schon wenn $\mathfrak{M} = K$ ist, geht er verloren. Sei etwa $\varphi(z) = e^z$; dann ist $\varphi(2\pi i) - \varphi(0) = 0$, jedoch $2\pi i \varphi'(0 + \tau 2\pi i) = 2\pi i e^{2\pi i \tau} \neq 0$. Für den Absolutbetrag gilt jedoch eine schwächere Aussage, die man auf Operatoren übertragen kann:

$$|\varphi(x+h) - \varphi(x)| = |h|\ |\varphi'(x+\tau h)| \leqslant |h| \max_{0\leqslant\tau\leqslant 1} |\varphi'(x+\tau h|\,. \tag{8.24}$$

M i t t e l w e r t s a t z f ü r O p e r a t o r e n : Sei A auf der Strecke f bis $f+h$ stetig differenzierbar. Dann gilt die Abschätzung

$$\| A(f+h) - Af \| \leqslant \| h \| \max_{0\leqslant\tau\leqslant 1} \| \frac{\delta A}{\delta(f+\tau h)} \|\,. \tag{8.25}$$

Man beweist das durch Rückführung auf den abgeschwächten klassischen Mittelwertsatz (8.24). Dazu bilden wir die Operatoraussage (8.25) einfach auf den R^1 ab, indem wir $l(A(f+\tau h)) =: \varphi(\tau) \in K$ betrachten; das linear-beschränkte Funktional l wählen wir sogleich passend. $\varphi(\tau)$ ist offenbar eine bezüglich τ stetig differenzierbare Funktion, da A das ist. Wegen (8.23) gilt $\varphi'(\tau) = l(\frac{\delta A}{\delta(f+\tau h)} h)$. Nach (8.24) ist $|\varphi(1) - \varphi(0)| \leqslant \max\limits_{0\leqslant\tau\leqslant 1} |\varphi'(\tau)|$. Also

$$|\varphi(1) - \varphi(0)| = |l(A(f+h) - Af)| = \| A(f+h) - Af \| \leqslant \max_{0\leqslant\tau\leqslant 1} \| l \|\ \| \frac{\delta A}{\delta(f+\tau h)} h \|\,.$$

Dabei haben wir l so gewählt, daß es aus dem einen Vektor $A(f+h) - Af$ gerade seine Norm macht und auf ganz $\mathfrak{M}$ mit $\| l \| = 1$ fortgesetzt ist. Das ist zulässig, wie uns der Fortsetzungssatz von *Hahn* und *Banach* aus § 6.5. (insbesondere der letzte Satz) gewährleistet. Beachtet man noch die Linear-Beschränktheit von $\frac{\delta A}{\delta(f+\tau h)}$, so folgt schon (8.25).

Ganz analog kann man nun auch höhere Restglieder abschätzen. Wir formulieren den

S a t z über die *Taylorentwicklung mit Restgliedabschätzung:*

Sei A ein $(n+1)$ mal stetig differenzierbarer Operator. Dann gilt (sofern alle $f + \tau h \in \mathfrak{D}_A$, $\tau \in [0, 1]$)

$$\| A(f+h) - [Af + \frac{1}{1!}\frac{\delta A}{\delta f} h + \frac{1}{2!}\frac{\delta^2 A}{\delta f^2} hh + \ldots + \frac{1}{n!}\frac{\delta^n A}{\delta f^n} hh \ldots h] \| \leqslant$$

$$\leqslant \frac{\| h \|^{n+1}}{(n+1)!} \max_{0\leqslant\tau\leqslant 1} \| \frac{\delta^{n+1} A}{\delta(f+\tau h)^{n+1}} \|\,. \tag{8.26}$$

$$\text{M.a.W.: } A(f+h) = Af + \frac{\delta A}{\delta f} h + \ldots + \frac{1}{n!}\frac{\delta^n A}{\delta f^n} hh \ldots h + \text{Rest}\,, \tag{8.27}$$

wobei der Rest eine kleine Norm hat, $\sim \| h \|^{n+1}$. Man kann also mit der allgemeinen Operatorabbildung völlig analog zur klassischen Analysis Taylorentwicklungen durchführen! Nicht-lineare Operatoren behandelt man somit ähnlich wie Funktionen durch lineare, quadratische etc. Näherung.

Der Beweis von (8.26) gelingt natürlich mit den beim Mittelwertsatz gelernten Hilfsmitteln. Für $\varphi(\tau) = l(A(f + \tau h))$ gilt die klassische Formel

$$|\varphi(1) - [\varphi(0) + \frac{1}{1!}\varphi'(0) + \frac{1}{2!}\varphi''(0) + \ldots + \frac{1}{n!}\varphi^{(n)}(0)]| \leqslant$$

$$\leqslant \frac{1}{(n+1)!} \max_{0 \leqslant \tau \leqslant 1} |\varphi^{(n+1)}(\tau)| .$$

l wird diesmal so gewählt, daß der in (8.26) stehende Vektor auf seine Norm abgebildet wird sowie $\| l \| = 1$.

9. DIE VERSCHIEDENEN KONVERGENZBEGRIFFE

Wir haben mehrfach vom Begriff der Konvergenz Gebrauch gemacht, sogar in den verschiedenartigsten Formen. Ursprünglich wurde Konvergenz in § 2.5.1. mittels der Metrik definiert. Später betrachteten wir die Konvergenz von Operatorfolgen, auch diejenige von linear-stetigen Funktionalen. Von letzteren benutzten wir ferner in § 7.2. zur Definition von Distributionen punktweise konvergente Folgen $\langle g_n | \varphi \rangle$. Daher ist es zweckmäßig, jetzt einmal die verschiedenen Konvergenzbegriffe übersichtlich zusammenzustellen und ihre Eigenschaften aufzuzählen. Wir werden sogleich an Hand eines physikalischen Beispiels sehen, daß sie alle in der Physik Verwendung finden.

9.1. Konvergenzbegriffe für Elemente aus allgemeinen Räumen

9.1.1. Definitionen der starken und schwachen Konvergenz

Unser ursprünglicher Konvergenzbegriff lautete so:

$$f_n \Rightarrow f, \text{ wenn } \| f_n - f \| < \epsilon, \text{ für alle } n > N(\epsilon) . \quad (9.1)$$

Wir bezeichnen (9.1) hinfort präziser als *starke Konvergenz*, manchmal auch *Normkonvergenz* im Banachraum. Starke Konvergenz setzt allein den Begriff einer Metrik voraus, ist also in metrischen, normierten oder unitären Räumen verwendbar.

Der Begriff der starken Konvergenz ist für manche physikalischen Anwendungen zu eng. Beschreibt man etwa gestreute Teilchen durch die Schrödingerwellenfunktion $\psi_t = e^{-iHt}\,\psi_o$, so sollte die Aufenthaltswahrscheinlichkeit an jedem Raumpunkt für hinreichend große Zeiten Null werden. Denn an jedem festen Raumpunkt laufen die Teilchen schließlich vorbei. Es ist aber $\| \psi_t - 0 \| = \| \psi_t \| = \| \psi_o \| \nrightarrow 0$; die Folge ψ_t ist auch gar nicht in sich konvergent. Was aber bedeutet die physikalische Aussage des Verschwindens genauer? Die Wahrscheinlichkeit, eine bestimmte feste Ortsverteilung φ vorzufinden, verschwindet mit $t \to \infty$, d.h. $|\langle \varphi | \psi_t \rangle|^2 \to 0$. Es ist also $\langle \varphi | \psi_t \rangle \to 0$ als Zahlenlimes für jede Verteilung φ. Nun ist aber allgemein gesehen $\langle \varphi | \psi_t \rangle$ ein lineares Funktional $l_\varphi(\psi_t)$. Das führt auf folgende Definition.

D e f i n i t i o n : Eine Folge $\{f_n\}$ von Banach- oder Hilbertraumelementen heißt *schwach konvergent* **gegen $f \in \mathfrak{M}$, $f_n \rightharpoonup f$, wenn die Zahlenfolge $l(f_n) \to l(f)$ konvergiert für j e d e s linear-stetige Funktional $l \in \mathfrak{M}^*$.**