Kapitel 8

Hamiltonsches Prinzip

In diesem Kapitel lernen wir eine neues Prinzip der klassischen Mechanik kennen, welches sich den bisher diskutierten Prinzipien von NEWTON und D'ALEMBERT als zumindest ebenbürtig erweist. Die Gesetze der klassischen Mechanik lassen sich aus zwei Typen von *Variationsprinzipien* ableiten. Beim differentiellen Prinzip von D'ALEMBERT wird ein *momentaner* Zustand des Systems mit kleinen virtuellen Verrückungen aus diesem Zustand verglichen. Beim integralen Prinzip von HAMILTON wird eine tatsächlich durchlaufene Bahn des Systems mit einer kleinen virtuellen Abweichung von dieser Bahn verglichen. Wie beim Prinzip von D'ALEMBERT ist das Ergebnis auch hier die Bewegungsgleichung.

8.1 Variationsrechnung

In der Differentialrechnung besteht eine einfache Aufgabe darin, die stationären Punkte einer Funktion $y(x)$ zu bestimmen. Die notwendige Bedingung für das Vorliegen eines stationären Punktes an der Stelle $x = a$ ist $y'(a) = 0$. Hinreichende Bedingungen dafür, daß es sich um ein Minimum oder Maximum handelt, sind $y''(a) > 0$ bzw. $y''(a) < 0$. Die Variationsrechnung beschäftigt sich mit einem ähnlichen, allerdings schwierigeren Problem: Gesucht ist eine Funktion $y(x)$, für die ein bestimmtes Integral über eine Funktion dieser Funktion (ein Funktional) einen extremalen Wert annimmt.

Der Name *Variationsrechnung* wurde zum ersten Mal von L. EULER im Jahre benutzt. Damit bezeichnete er die neue Methode, aus der er selbst sehr virtuos mögliche Folgerungen zog. Heute wird der Begriff Variationsrechnung in einem breiteren Sinn verwendet. Gegenstand der Variationsrechnung sind das Aufsuchen von Minima, Maxima und Sattelpunkten (also kritischer Punkte) einer Funktion

$$F : M \longrightarrow \mathbb{R}, \qquad \mathbb{R} = \{\text{reelle Zahlen}\}.$$

In den Anwendungen ist M eine Menge von Zahlen, Funktionen, Wegen, Kurven, Flächen, Feldern, usw. Extremalprobleme spielen nicht nur in der theoretischen Physik eine wichtige Rolle, sondern zum Beispiel auch in der Wirtschaft, Regelungstechnik oder Spieltheorie. Eines der klassischen Probleme der Variationsrechnung war die isoperimetrische Aufgabe, d.h. das Aufsuchen derjenigen geometrischen Figur größter Fläche bei gegebenen Umfang. Eine wichtige Anwendung findet die Variationrechnung im FERMATschen Prinzip, nach dem ein Lichtstrahl denjenigen Weg nimmt, auf dem die benötigte Zeit minimal ist. Das FERMATsche Prinzip wurde von ERNST ABBE geschickt benutzt, um optische Geräte zu berechnen.

Die Geburt der modernen Variationsrechnung wird gewöhnlich auf jenen Tag des Juni 1696 gelegt, als das Problem der *Brachystochrone* durch JOHANN BERNOULLI gestellt wurde (siehe unten). Seit 1732 setzte sich EULER systematisch mit Extremalproblemen auseinander. Er fand notwendige Bedingungen für ein Extremum von einfachen Funktionalen und erhob die Variationsrechnung zu einer eigenständigen mathematischen Disziplin. Er stellte das *Prinzip der kleinsten Wirkung* auf eine fundierte Grundlage. LAGRANGE entwickelte die Methoden zur systematischen Behandlung einer großen Klassen von Variationsproblemen. Diese wurden von LEGENDRE und JACOBI fortentwickelt und weiter verallgemeinert. Das Prinzip der kleinsten Wirkung spielte in den späteren Arbeiten von HAMILTON eine wesentliche Rolle.

Für eine spezielle Klasse von Funktionalen, das heißt von Funktionen, deren Argumente Funktionen sind, entdeckte L. EULER eine erste notwendige Bedingung, welche ein Extremum des Funktionals erfüllen muß. Sie trägt heute den Namen EULER-Gleichung. Diese Gleichung entspricht der Bedingung $f'(x) = 0$ für das Extremum x einer Funktion $f : \mathbb{R} \to \mathbb{R}$. EULER betrachtete folgende *Aufgabenstellung der Variationsrechnung:* Suche eine Funktion $y(x)$, für die das Integral

$$F[y] = \int_{x_1}^{x_2} f\big(x, y, y'\big)dx \tag{8.1}$$

extremal ist. Der Integrand f ist eine Funktion der unabhängigen Variablen x, der abhängigen Variablen y und deren Ableitung $y' = dy/dx$. An den Grenzen x_1 und x_2 sind die Werte y_1 und y_2 vorgeschrieben. Das Integral F nimmt für verschiedene Kurven $y(x)$ zwischen $P_1 = (x_1, y_1)$ und $P_2 = (x_2, y_2)$ im allgemeinen verschiedene Werte an. Wir nehmen nun an, das Funktional habe einen extremalen Wert für $y(x)$, d.h. die Kurve $y(x)$ sei ein Maximum, Minimum oder Sattelpunkt von F. Nun vergleichen wir mit dem Wert des Funktionals für benachbarte Kurven $y(x)+\delta y(x)$, wobei $\delta y(x)$ infinitesimal klein sein soll für alle x zwischen x_1 und x_2, siehe Abbildung (8.1). Wir definieren

$$\delta f = f(x, y + \delta y, y' + \delta y') - f(x, y, y'). \tag{8.2}$$

Das Symbol δ bedeutet *Variation.* δf ist der Zuwachs von f wenn man an der festen Stelle x von der Kurve $y(x)$ zur Vergleichskurve $y(x) + \delta y(x)$ übergeht. Offenbar ist dann $\delta x = 0$. Weiterhin gilt

$$\delta y' = \frac{d}{dx}\big(y + \delta y\big) - \frac{dy}{dx} = \frac{d}{dx}\delta y, \tag{8.3}$$

und die Symbole δ und d/dx vertauschen. Da δy infinitesimal klein ist, folgt aus (8.2)

$$\delta f = \frac{\partial f}{\partial y}\delta y + \frac{\partial f}{\partial y'}\delta y'.$$

Nun ist $y(x)$ ein stationärer 'Punkt' von F, wenn das Integral längs y in erster Näherung gleich demjenigen längs $y + \delta y$ ist,

$$\delta F = \int_{x_1}^{x_2} \delta f \, dx \overset{(8.3)}{=} \int_{x_1}^{x_2} \Big[\frac{\partial f}{\partial y}\delta y + \frac{\partial f}{\partial y'}\frac{d}{dx}(\delta y)\Big]dx = 0.$$

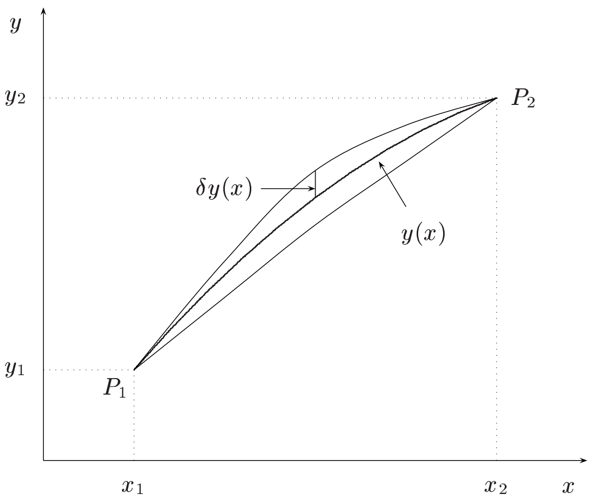

Abbildung 8.1: *Unter allen differenzierbaren Kurven, die von P_1 nach P_2 laufen, wird diejenige gesucht, die das Integral F extremal macht.*

Der zweite Summand läßt sich durch partielle Integration umformen,

$$\delta F = \int_{x_1}^{x_2} \left[\frac{\partial f}{\partial y} - \frac{d}{dx}\frac{\partial f}{\partial y'}\right]\delta y\, dx + \left[\frac{\partial f}{\partial y'}\delta y\right]_{x_1}^{x_2} . \tag{8.4}$$

Hier verschwinden die Randterme, weil an den beiden Endpunkten x_1 und x_2 nach Voraussetzung $\delta y = 0$ ist. Die Stationarität von F bedingt also

$$\delta F = \int_{x_1}^{x_2} \left[\frac{\partial f}{\partial y} - \frac{d}{dx}\frac{\partial f}{\partial y'}\right]\delta y(x) dx = 0. \tag{8.5}$$

Die Gleichung muß für beliebige Deformationen δy der Kurve y verschwinden. Dies impliziert, daß der Ausdruck in eckigen Klammern Null sein muß,

$$\frac{\partial f}{\partial y(x)} - \frac{d}{dx}\frac{\partial f}{\partial y'(x)} \equiv \frac{\delta F}{\delta y(x)} = 0. \tag{8.6}$$

Diese Gleichung stellt die gesuchte notwendige Bedingung an $y(x)$ dar. Eine Funktion y, die diese Differentialgleichung erfüllt, heißt Extremale. In der Menge der Extremalen ist dann die Minimal- oder Maximalkurve enthalten, falls sie existiert. Die Gleichung (8.6) heißt die zum Variationsproblem gehörige EULER-*Gleichung*.

8.1.1 Geodätische Linien

Wir wollen den oft als selbstverständlich angesehenen Sachverhalt beweisen, daß die gerade Linie die kürzeste Verbindung zwischen zwei Punkten in der *Ebene* darstellt. Das Linienelement ist in kartesischen Koordinaten durch $ds^2 = dx^2 + dy^2$ gegeben, also ist die Länge der

Kurve $x \to y(x)$ gleich

$$s = \int_{x_1}^{x_2} ds = \int_{x_1}^{x_2} \sqrt{1 + y'^2}\, dx.$$

Soll y ein Minimum von s sein, so muß die EULER-Gleichung (8.6) mit $f = (1 + y'^2)^{1/2}$ erfüllt sein, also

$$\frac{d}{dx}\Big[\frac{\partial}{\partial y'}\sqrt{1 + y'^2}\Big] = 0,$$

oder

$$\frac{y'}{\sqrt{1 + y'^2}} = \text{const.},$$

gelten, was y' =const. nach sich zieht. Die minimierende Kurve ist eine Gerade durch die Punkte (x_1, y_1) und (x_2, y_2).

Das Linienelement auf der *Oberfläche einer Kugel* vom Radius R ist

$$ds = R\sqrt{d\vartheta^2 + \sin^2\vartheta d\varphi^2}. \tag{8.7}$$

Wir wollen φ als Funktion von ϑ so bestimmen, daß s stationär ist. Also müssen wir die EULER-Gleichung für $f = R(1 + \varphi_\vartheta^2 \sin^2\vartheta)^{1/2}$ lösen, d.h.

$$\frac{d}{d\vartheta}\Big[\frac{\varphi_\vartheta \sin^2\vartheta}{(1 + \varphi_\vartheta^2 \sin^2\vartheta)^{1/2}}\Big] = 0.$$

Der Ausdruck in eckigen Klammern ist gleich einer Konstanten c. Die resultierende Differentialgleichung für $\varphi(\vartheta)$ hat die Lösung

$$\varphi = a - \arcsin(k\cot\vartheta) \quad \text{mit} \quad c = -\frac{k}{\sqrt{1 + k^2}} \Longrightarrow Rk\cot\vartheta = R\sin(a - \varphi).$$

Zur Interpretation dieser Bedingung schreiben wir sie in kartesische Koordinaten um. Nach Multiplikation mit $\sin\vartheta$ ergibt sich

$$Rk\cos\vartheta = R\sin\vartheta(\sin a\cos\varphi - \cos a\sin\varphi) \quad \text{bzw.} \quad kz = x\sin a - y\cos a.$$

Dies ist die Gleichung einer Ebene durch den Mittelpunkt der Kugel, die folglich die Kugeloberfläche in einem Großkreis schneidet. Die kürzeste Verbindung zwischen zwei Punkten auf der Kugeloberfläche ist einer der beiden Bögen auf dem Großkreis durch diese Punkte.

8.1.2 Die Brachystochrone

In diesem von JOHANN BERNOULLI gestellten Problem ist diejenige Kurve gesucht, auf der ein Körper unter dem Einfluß der Schwerkraft und ohne Reibung von einem gegebenen Punkt zu einem zweiten Punkt in der kürzesten (brachystos) Zeit (chronos) gleitet. Die gesuchte Kurve, die sogenannte Brachystochrone, ist eine Zykloide.

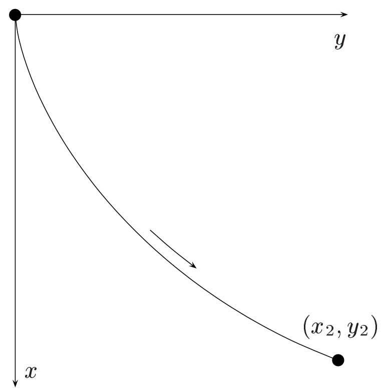

Abbildung 8.2: *Die Zykloide löst das Brachystochronenproblem von J. Bernoulli*

Der Massenpunkt soll am Anfang im Ursprung ruhen, der Endpunkt sei (x_2, y_2). Bei diesem Problem ist es bequem, die y-Achse nach rechts zu legen und x nach unten zu messen. Aus dem Energiesatz folgt

$$\frac{1}{2}mv^2 = mgx \quad \text{oder} \quad v = \sqrt{2gx},$$

wobei v die Geschwindigkeit des Massenpunktes längs seiner Bahn ist und g die Fallbeschleunigung bezeichnet. Wegen

$$vdt = ds = \sqrt{dx^2 + dy^2} = \sqrt{1 + y'^2}dx$$

schreibt sich der Energiesatz gemäß

$$\sqrt{1 + y'^2}dx = \sqrt{2gx}dt.$$

Die vom fallenden Körper benötigte Zeit ist gegeben durch das Integral

$$T = \int dt = \frac{1}{\sqrt{2g}} \int_0^{x_2} x^{-1/2} \sqrt{1 + y'^2}\, dx$$

und diese gilt es zu minimieren. Die entsprechende Euler-Gleichung lautet

$$\frac{d}{dx} \frac{y'}{\sqrt{x(1 + y'^2)}} = 0.$$

und eine erste Integration führt auf die einfache Differentialgleichung

$$\frac{y'^2}{x(1 + y'^2)} = c \quad \text{bzw.} \quad y' = \frac{x}{\sqrt{x/c - x^2}}.$$

Die weitere Integration ergibt dann

$$y = a \arccos\left(1 - \frac{x}{a}\right) - \sqrt{2ax - x^2} + c', \quad \text{wobei} \quad a = 1/2c \tag{8.8}$$

ist. Die neue Integrationskonstante c' muß Null sein, damit y bei $x = 0$ verschwindet. Diese Gleichung stellt eine *Zykloide* über der y-Achse mit einer Spitze im Ursprung dar. Die Konstante a muß so gewählt werden, daß die Zykloide durch den Punkt (x_2, y_2) geht.

8.1.3 Mehrere abhängige oder/und unabhängige Variable

In diesem Abschnitt verallgemeinern wir die bisherigen Resultate auf Systeme mit mehr als einer abhängigen Variablen y oder/und mit mehr als einer unabhängigen Variablen x sowie auf Variationsprobleme, bei denen die Funktion f von höheren Ableitungen $y^{(n)}$ von y abhängt.

Mehrere abhängige Variablen: Wir wollen voraussetzen, daß der Integrand f des Integrales

$$F = \int_{x_1}^{x_2} f(x, y_1, \ldots, y_N, y_1', \ldots, y_N')dx, \tag{8.9}$$

welches ein Maximum oder Minimum annehmen soll, eine Funktion einer unabhängigen, aber mehrerer abhängiger Variablen ist. In der Mechanik wäre die unabhängige Variable die Zeit t und die abhängigen Variablen die (verallgemeinerten) Koordinaten q^i. Wir suchen nun Funktionen $y_1(x), \ldots, y_N(x)$, für die das Funktional mit festen Randbedingungen

$$y_i(x_1) = y_{i1} \quad \text{und} \quad y_i(x_2) = y_{i2}, \qquad i = 1, \ldots, N \tag{8.10}$$

stationär wird. Die Bedingung dafür ist wie oben

$$\delta F = \int_{x_1}^{x_2} \delta f \, dx = 0, \tag{8.11}$$

nur ist jetzt

$$\delta f = \frac{\partial f}{\partial y_1}\delta y_1 + \frac{\partial f}{\partial y_2}\delta y_2 + \ldots + \frac{\partial f}{\partial y_1'}\delta y_1' + \frac{\partial f}{\partial y_2'}\delta y_2' + \ldots = \frac{\partial f}{\partial y_i}\delta y_i + \frac{\partial f}{\partial y_i'}\delta y_i',$$

wobei wir wieder von der EINSTEINschen Summenkonvention Gebrauch machten. Bei der Berechnung des Integrales (8.11) formen wir wieder die Summanden der zweiten Gruppe durch partielle Integration um, und erhalten

$$\delta F = \int_{x_1}^{x_2} \left[\frac{\partial f}{\partial y_i} - \frac{d}{dx}\frac{\partial f}{\partial y_i'}\right]\delta y_i \, dx + \Big[p_i \delta y_i\Big]_{x_1}^{x_2}, \quad \text{wobei} \quad p_i = \frac{\partial f}{\partial y_i'}$$

gesetzt wurde. Bei festgehaltenen Anfangs- und Endpunkten verschwindet der letzte Term. Das Funktional ist stationär bezüglich beliebiger Variationen δy_i der Funktionen y_i, wenn die EULER-*Gleichungen*

$$\frac{\delta F}{\delta y_i(x)} = \frac{\partial f}{\partial y_i(x)} - \frac{d}{dx}\frac{\partial f}{\partial y_i'(x)} = 0 \tag{8.12}$$

gelten. Es handelt sich um N gewöhnliche Differentialgleichungen zweiter Ordnung für die N gesuchten Funktionen $y_1(x), \ldots y_N(x)$, die unter Berücksichtigung der *Randbedingungen* (8.10) zu lösen sind. Diese Verallgemeinerung ist für die Beschreibung mehrerer Massenpunkte notwendig.

Mehrere Argumente: Wir betrachten nun den Fall, daß mehrere Argumente $x_1, \ldots, x_M$ anstelle von x vorliegen. Gesucht ist demnach eine Funktion $y = y(x_1, \ldots, x_M)$, die das Funktional

$$F\big[y(\ldots)\big] = \int_B d^M x \; f\Big(x_1, .., x_M, y(x_1, .., x_M), \frac{\partial y(x_1, .., x_M)}{\partial x_1}, \ldots, \frac{\partial y(x_1, .., x_M)}{\partial x_M}\Big) \quad (8.13)$$

extremal macht. Hierbei sei die Funktion $y = y(x_1, \ldots, x_M)$ auf dem Rande des Integrationsgebietes B fest vorgegeben. Aus der Stationaritätsbedingung

$$\frac{\delta F[y(., \ldots, .)]}{\delta y(x_1, \ldots, x_M)} = 0$$

folgt dann die EULER-Gleichung

$$\frac{\partial f}{\partial y} - \sum_{i=1}^{M} \frac{\partial}{\partial x_i} \frac{\partial f}{\partial(\partial y / \partial x_i)} = 0. \quad (8.14)$$

Dies ist eine partielle Differentialgleichung zweiter Ordnung für die gesuchte Funktion $y(x_1, \ldots, x_M)$. Eine mögliche Anwendung für diese Verallgemeinerung der Variationsrechnung wäre etwa die Ausdehnung einer dehnbaren Membran oder einer Seifenhaut im Schwerefeld der Erde, die längs einer Kurve in der $x_1 - x_2$-Ebene eingespannt ist. Hierbei konkurrieren Schwerkraft und Oberflächenspannung miteinander und die Gleichgewichtskonfiguration wird durch ein Minimum der Summe der beiden potentiellen Energien bestimmt.

Die beiden Verallgemeinerungen (8.9) und (8.13) können auch miteinander kombiniert werden. Dann enthält das Funktional mehrere Funktionen, die jeweils von mehreren Variablen abhängen. Diese relativ allgemeine Form des Funktionals,

$$F = \int_B d^M x \; f\Big(x_a, y_i, \frac{\partial y_i}{\partial x_a}\Big), \qquad y_i = y_i(x, \ldots, x_M), \qquad i = 1, \ldots, N, \quad (8.15)$$

benötigt man in Feldtheorien.

Höhere Ableitungen: Falls das Funktional höhere Ableitungen von $y(x)$ enthält, wie z.B.

$$F = \int_{x_1}^{x_2} f\big(x, y(x), y'(x), y''(x)\big)\, dx, \quad (8.16)$$

dann treten auch höhere Ableitungen in der EULER-Gleichung auf. Die Extremalbedingung

$$\frac{\delta F[y(.)]}{\delta y(x)} = 0 \quad (8.17)$$

führt zum Beispiel für das Funktional (8.16) auf die Differentialgleichung

$$\frac{\partial f}{\partial y(x)} - \frac{d}{dx} \frac{\partial f}{\partial y'(x)} + \frac{d^2}{dx^2} \frac{\partial f}{\partial y''(x)} = 0. \quad (8.18)$$

Es handelt sich hierbei offensichtlich um eine gewöhnliche Differentialgleichung vierter Ordnung für $y(x)$. Die üblichen Randbedingungen $y(x_i) = y_i$ für die Funktion y legen diese nun nicht mehr eindeutig fest. Um auf die EULER-Gleichung (8.18) zu kommen, muß man mehrfach partiell integrieren, wobei zusätzliche Randterme auftreten. Diese verschwinden, wenn die weiteren Randbedingungen

$$\left.\frac{\partial f}{\partial y(x)}\right|_{x_i} = 0, \qquad i = 1, 2,$$

gefordert werden.

8.2 Isoperimetrische Probleme

In Anwendungen tritt oft das Problem auf, bei dem ein Integral stationär sein soll, während gleichzeitig ein oder mehrere Integrale, in denen dieselben Variablen vorkommen, konstant zu halten sind. Ein bekanntes Beispiel ist die Aufgabe, diejenige ebene geschlossene Kurve vorgegebener Länge zu finden, die eine möglichst große Fläche einschließt. Nach diesem Beispiel heißen solche Probleme *isoperimetrische Probleme*. Solche Nebenbedingungen oder Zwangsbedingungen können mit der Lagrangeschen Methode der unbestimmten Multiplikatoren bestimmt werden. Man sucht einen stationären Wert von

$$F = \int f\,dx, \qquad f = f(x, y_i, y_i'), \quad y_i = y_i(x), \qquad i = 1, \ldots, N, \tag{8.19}$$

wobei die s Nebenbedingungen

$$F_1 = \int f_1 dx = c_1, \; \ldots \; , F_s = \int f_s dx = c_s \tag{8.20}$$

gelten soll. Sämtliche Integranden enthalten dieselben Variablen, und die Grenzen sind gleich bei allen Integralen. Nun führen wir s konstante LAGRANGEsche Multiplikatoren $\lambda_1, \ldots, \lambda_s$ ein, deren Werte wir vorerst unbestimmt lassen. Offenbar ist mit F auch

$$F^* = F + \lambda_1 F_1 + \ldots + \lambda_s F_s \tag{8.21}$$

stationär, und zwar wegen (8.20) für jede Wahl der Multiplikatoren λ_i. Wir stehen also vor einem ganz ähnlichen Problem wie früher, nämlich ein einziges Integral extremal zu machen, nur mit einem abgeänderten Integranden. F muß durch

$$F^* = \int f^*\,dx, \qquad f^* = f + \sum_{i=1}^{s} \lambda_i f_i \tag{8.22}$$

ersetzt werden. Eine notwendige Bedingung für die Stationarität von F^* ist

$$\sum_{i=1}^{N} \int \left(\frac{d}{dx} \frac{\partial f^*}{\partial y_i'(x)} - \frac{\partial f^*}{\partial y_i(x)} \right) \delta y_i(x) dx = 0. \tag{8.23}$$

Hier können wir nicht ohne weiteres den Übergang zu den EULERschen Gleichungen machen, weil die δy_i nicht mehr beliebig sind; die Variation muß mit (8.20) verträglich sein. Doch nun kommt uns zu Hilfe, daß die Multiplikatoren λ_i noch beliebig gewählt werden können. Diese lassen sich tatsächlich so bestimmen[1], daß (8.23) die EULERschen Gleichungen

$$\frac{\partial f^*}{\partial y^i(x)} - \frac{d}{dx} \frac{\partial f^*}{\partial y_i'(x)} = 0, \qquad i = 1, \ldots, N, \tag{8.24}$$

[1] L. Page, *Introduction to Theoretical Physics*, third edition, D. van Nostrand Co., New York 1952

nach sich zieht. Nach dem Lösen dieser Gleichungen erscheinen die konstanten, aber zunächst unbekannten λ_i als Parameter in den Extremalen. Man kann sie mit Hilfe der Bedingungen (8.20) eliminieren. und dann erhalten sie oft eine unmittelbare physikalische Bedeutung.

Als Beispiel betrachten wir das klassische isoperimetrische Problem: Gesucht ist diejenige ebene geschlossene Kurve, die eine möglichst große Fläche begrenzt. Wir fragen also nach einer Funktion $r(\varphi)$ für welche die Fläche

$$A = \frac{1}{2} \int_0^{2\pi} r^2 \, d\varphi$$

maximal ist, während der Umfang

$$U = \int_0^{2\pi} \sqrt{r^2 + r'^2} d\varphi, \qquad r' = \frac{dr}{d\varphi},$$

fest vorgegeben ist. Es gilt also, das Funktional $F = \int f^* \, d\varphi$ mit

$$f^* = \frac{1}{2} r^2 + \lambda \sqrt{r^2 + r'^2}, \tag{8.25}$$

zu minimieren. Die entsprechende EULER-Gleichung lautet

$$r + \frac{\lambda r}{\sqrt{r^2 + r'^2}} - \frac{d}{d\varphi} \Big[\frac{\lambda \, r'}{\sqrt{r^2 + r'^2}} \Big] = 0.$$

Führen wir die Ableitung aus, dann folgt unmittelbar

$$\frac{r r'' - 2r'^2 - r^2}{(r^2 + r'^2)^{3/2}} = \frac{1}{\lambda}.$$

Die linke Seite dieser Gleichung ist gerade die Krümmung der Kurve. Diese muß konstant sein und folglich ist die Kurve ein Kreis mit dem Radius λ.

8.3 Hamiltonsches Prinzip

Wir stellen nun ein Variationsprinzip auf, dessen EULER-Gleichungen die LAGRANGE-Gleichungen zweiter Art sind. Hierzu betrachten wir das Zeitintegral der LAGRANGE-Funktion als Funktional der Bahnkurve

$$S[q^1(.), \ldots, q^f(.)] \equiv S[q(.)] = \int_{t_1}^{t_2} L\big(t, q^1(t), \ldots, q^f(t), \dot{q}^1(t), \ldots, \dot{q}^f(t)\big) dt. \tag{8.26}$$

Dieses Funktional wird als *Wirkung* oder *Wirkungsfunktional* bezeichnet. Da die LAGRANGE-Funktion die Dimension einer Energie hat, ist die Dimension der Wirkung

$$[S] = \text{Zeit} \cdot \text{Energie} = \text{kg m}^2/\text{s}.$$

Wir untersuchen nun das HAMILTONsche Prinzip, nach dem für die klassisch erlaubten Bahnen die Wirkung stationär sein soll

$$\frac{\delta S[q(.)]}{\delta q^i(t)} = 0, \qquad i = 1, \ldots, f. \tag{8.27}$$

Hierbei sind die Variationen dadurch eingeschränkt, daß Anfangs- und Endpunkte festgehalten werden,

$$q^i(t_1) = q_1^i \quad \text{und} \quad q^i(t_2) = q_2^i, \qquad i = 1, \dots, f. \tag{8.28}$$

Nach Abschnitt (8.1.3) führt das HAMILTONsche Prinzip auf die EULER-Gleichungen

$$\frac{\partial L}{\partial q^i(t)} - \frac{d}{dt}\frac{\partial L}{\partial \dot{q}^i(t)} = 0, \qquad i = 1, \dots, f, \tag{8.29}$$

welche gleich den LAGRANGE-Gleichungen zweiter Art sind. Deshalb bezeichnet man (8.29) oft auch als EULER-LAGRANGE-Gleichungen.

Das Prinzip von HAMILTON besagt demnach, daß aus der Menge aller möglichen Bahnkurven diejenige realisiert ist, welche die Wirkung (8.26) stationär macht. Es spielt dabei keine Rolle, ob es sich beim Extremum um ein Minimum, ein Maximum oder einen Sattelpunkt handelt. In der Regel wird die Wirkung minimal. Daher kommt der Name *Prinzip der kleinsten Wirkung*. Gelegentlich aber ist S nicht minimal.

Zur Festlegung der Lösungen der EULER-LAGRANGE-Gleichungen braucht es $2f$ Integrationskonstanten. Dies können die f Anfangsorte und f Anfangsgeschwindigkeiten sein oder aber die Koordinaten zur Zeit t_1 und zur Zeit t_2, wie in (8.28). Im ersten Fall spricht man von einem *Anfangswertproblem*, im zweiten Fall von einem *Randwertproblem*. Das HAMILTONsche Prinzip führt auf ein Randwertproblem. Dies macht es auf den ersten Blick etwas unanschaulich. Es widerspricht sogar unserem Kausalgefühl, da die Bewegung nicht nur aus einem Anfangszustand, sondern aus Vergangenheit und *Zukunft* abgeleitet wird. Die Äquivalenz mit den anderen Prinzipien der Mechanik beweist aber die Kausalität.

Für die praktische Lösung von Problemen bringt das HAMILTONsche Prinzip keine Vorteile, da dessen konkrete Anwendung wieder auf die EULER-LAGRANGE-Gleichungen (8.29) führt. Es faßt aber die Mechanik von Systemen mit holonomen Zwangsbedingungen in einer kompakten und prägnanten Form zusammen. Das HAMILTONsche Prinzip ist an Einfachheit und Ökonomie kaum mehr zu überbieten. Die Natur sucht unter allen denkbaren Bewegungen diejenige aus, die ihr Ziel mit der kleinsten Wirkung, also mit geringstem Aufwand, erreicht.

In der nichtrelativistischen Mechanik ist die LAGRANGE-Funktion die Differenz von kinetischer und potentieller Energie. Im Unterschied zur kinetischen oder potentiellen Energie ist die LAGRANGE-Funktion aber keine physikalische Meßgröße. Sie ist eine mathematische Hilfsgröße, die derart definiert wurde, um aus ihr die Bewegunggleichungen abzuleiten.

Unter Eichtransformationen (7.63) ändert sich die Wirkung gemäß

$$S' = \int_{t_1}^{t_2} dt\, L'\big(t, q(t), \dot{q}(t)\big) = S + F\big(t_2, q(t_2)\big) - F\big(t_1, q(t_1)\big). \tag{8.30}$$

Da bei der Variation der Bahnkurven deren Endpunkte sowie Anfangs- und Endzeit fest sind, ist das HAMILTONsche Prinzip invariant unter Eichtransformationen,

$$\frac{\delta S'[q(.)]}{\delta q^i(t)} = \frac{\delta S[q(.)]}{\delta q^i(t)}. \tag{8.31}$$

Die Wirkungen S und S' haben also insbesondere die gleichen stationären Bahnkurven.

Die Frage nach der Bedeutung des HAMILTONschen Prinzips läßt sich wie folgt beantworten:

- Prinzipien der stationären Wirkung sind allgemeine Prinzipien der Physik, die zum Beispiel auch in der Elektrodynamik oder Gravitationsphysik auftreten.
- Es ist als fundamentales Prinzip anzusehen, da aus ihm die LAGRANGE-Gleichungen erster und zweiter Art folgen.
- Die Variationsgleichung

  $$\delta S = \delta \int_{t_1}^{t_2} L dt = 0 \tag{8.32}$$

  ist unabhängig von den gewählten Koordinaten, hat also eine von den Koordinaten unabhängige Bedeutung.
- Für einige weitere Entwicklungen der klassischen Mechanik, zum Beispiel die Theorie von HAMILTON und JACOBI oder die Analogie zwischen Mechanik und geometrischer Optik, ist die im HAMILTONschen Prinzip auftretende Wirkung S von zentraler Bedeutung.
- Die Wirkung ist auch eine sehr wichtige Größe in der Quantenphysik. Die sogenannte Pfadintegral-Quantisierung macht wesentlichen Gebrauch von ihr.

8.4 Anhang: Differenziation in ∞-dimensionalen Räumen.

Erst sehr viel später, gegen Ende des 19. Jahrhunderts, als der Begriff der Differenziation auf unendlich-dimensionale Räume verallgemeinert wurde[2], erhielt die Variationsrechnung in der Mathematik eine solide Basis.

Banach-Räume: In der Regel legt man einen *reellen Vektorraum* E mit einem Längenbegriff oder einer Norm zugrunde. Eine *Norm* ist eine Abbildung $E \to [0, \infty)$ mit den Eigenschaften

$$\|\lambda y\| = |\lambda| \, \|y\|, \quad \|y + z\| \leq \|y\| + \|z\|, \quad \|y\| = 0 \Longrightarrow y = 0, \quad \lambda \in \mathbb{R}, \quad y, z \in E. \tag{8.33}$$

Ein Vektorraum, versehen mit einer Norm $\|\cdot\|$ heißt *normierter Raum*. Ein vollständiger normierter Raum, d.h. ein normierter Vektorraum, in dem jede CAUCHY-Folge konvergiert, heißt BANACH-*Raum*. Für die Physik relevante Beispiele von BANACH-Räumen sind:

- Die *Euklidischen Räume* $\mathbb{R}^n$ mit den Normen

  $$\|y\|_p = \Big(\sum_{a=1}^{n} |y_a|^p \Big)^{1/p}, \qquad 1 \leq p \leq \infty. \tag{8.34}$$

  Für $p = 2$ ist $\|y\|_2$ die Länge des Maßstabes y, der am Anfang der Vorlesung bei der Diskussion von Raumzeit-Strukturen wichtig war.

[2] unter anderem von VOLTERRA, HADAMARD und dessen Schüler FRÉCHET, GÂTEAUX, HILBERT

- Die *Folgenräume* $l_p(\mathbb{R})$: $l_p(\mathbb{R})$ ist der Vektorraum aller Folgen $y = (y_n)_{n\in\mathbb{N}}$ reeller Zahlen y_n, für die

$$\sum_{n=1}^{\infty} |y_n|^p < \infty, \qquad 1 \le p < \infty$$

gilt. Versehen mit den Normen

$$\|y\|_p = \Big(\sum_{n=1}^{\infty} |y_n|^p\Big)^{1/p}, \qquad 1 \le p < \infty. \tag{8.35}$$

sind die Folgenräume vollständig, also BANACH-Räume. Die Folgenräume spielen in der Matrizenmechanik von W. HEISENBERG eine ganz wichtige Rolle.

- Die LEBESGUE-Räume: Für jede reelle Zahl $p \ge 1$ sei

$$\mathcal{L}_p(\mathbb{R}^n) = \{f : \mathbb{R}^n \to \mathbb{R} \,\big|\, f \text{ meßbar}, |f|^p \text{ summierbar}\}.$$

$\mathcal{L}_p$ ist ein reeller Vektorraum und

$$\|f\|_p = \Big(\int_{\mathbb{R}^n} |f(x)|^p dx\Big)^{1/p} \tag{8.36}$$

eine Halbnorm auf $\mathcal{L}_p$. Sie definiert eine Norm auf $L_p = \mathcal{L}_p/N$, wo N den Unterraum aller Funktionen von $\mathcal{L}_p$ bezeichnet, die fast überall Null sind. Für $n = 1$ sind die Elemente von $\mathcal{L}_p(\mathbb{R})$ Funktionen $\mathbb{R} \to \mathbb{R}$ und können als Bahnen von Punktteilchen in $\mathbb{R}$ interpretiert werden. Dies deutet bereits die Relevanz der BANACH-Räume L_p für die klassische Mechanik an. Diese Räume spielen in der Wellenmechanik von E. SCHRÖDINGER eine herausragende Rolle.

Fréchet-Ableitung: Nach diesem kurzen Exkurs über unendlich-dimensionale Vektorräume kehren wir zu den Ableitungen zurück. Die Verallgemeinerung des Begriffs des totalen Differentials führt zum Begriff der FRÉCHET-Ableitung einer Funktion von einem Banachraum in einen anderen Banachraum. In den meisten Anwendungen ist der Bildraum gleich $\mathbb{R}$:

Definition: *Es seien $(E, \|.\|)$ ein Banach Raum und $M \subset E$ eine offene, nichtleere Teilmenge. Dann heißt eine Funktion $F : M \longrightarrow \mathbb{R}$ genau dann in einem Punkt $y \in M$ differenzierbar, wenn es eine stetige lineare Abbildung $F'_y : E \to \mathbb{R}$ gibt, so daß gilt*

$$F(y+h) - F(y) = F'_y(h) + o(y,h) \tag{8.37}$$

für alle $h \in E, y+h \in M$. Dabei hat die Funktion $h \to o(y,h) \in \mathbb{R}$ die Eigenschaften

$$o(y,0) = 0 \quad \text{und} \quad \lim_{h\to 0} \frac{|o(y,h)|}{\|h\|} = 0. \tag{8.38}$$

Die FRÉCHET-Ableitung F'_y ist ein Element aus $\mathcal{L}(E,\mathbb{R})$, das heißt aus dem Dualraum E' von E. Falls F in allen Punkten von M differenzierbar ist, so heißt F auf M differenzierbar. Man nennt die Abbildung F', welche jedem Punkt $y \in M$ die *Ableitung F'_y von F* im Punkt y zuordnet, die Ableitung von F. Falls $F' : M \to E'$ eine stetige Abbildung ist, so heißt F *stetig differenzierbar auf M* oder von der Klasse $\mathcal{S}^1$.

Man kann zeigen, daß es höchstens eine stetige lineare Abbildung $F'_y \in E'$ gibt, so daß die Gleichung (8.33) gilt. Für $E = \mathbb{R}^n$ kann man die Ableitung F'_y in kanonischer Weise mit der JACOBI-Matrix der Abbildung F im Punkte y identifizieren,

$$F'_y = \left(\frac{\partial F}{\partial y_a}\right)_y .$$

Die FRÉCHET-Ableitung ist eine *lineare Operation*,

$$(\alpha F + \beta G)' = \alpha F' + \beta G', \qquad \alpha, \beta \in \mathbb{R}. \tag{8.39}$$

Die Ableitung erfüllt die *Kettenregel*: Es seien $U \subset E$ und $V_1, V_2 \subset \mathbb{R}$ offene, nicht-leere Mengen und $F : U \to V_1$ sowie $G : V_1 \to V_2$ stetig differenzierbar. Dann ist auch

$$G \circ F : U \longrightarrow V_2 \subset \mathbb{R}$$

stetig differenzierbar, und es gilt für alle $y \in U \subset E$

$$(G \circ F)'_y = G'_{F(y)} \circ F'_y. \tag{8.40}$$

Wir wollen dies beweisen: F differenzierbar heißt

$$F(y+h) - F(y) = F'_x(h) + o_1(x,h), \qquad \forall h \in B_{r_1}, \qquad y + h \subset U_1,$$

und entsprechend, da G in $z = F(y)$ differenzierbar ist,

$$G(z+k) - G(z)) = G'_z(k) + o_2(z,k), \qquad \forall k \in B_{r_2}, \qquad z + k \in U_2.$$

Hier ist B_{r_1} die 'Kugel' mit Radius r_1 um den Ursprung in E und B_{r_2} die offene 'Kugel' mit Radius r_2 um den Ursprung in $\mathbb{R}$. Da F insbesondere stetig ist, können wir zu gegebenem $r_2 > 0$ ein $r_1 > 0$ so wählen, daß die erste Relation und überdies $F(B_{r_1}) \subset B_{r_2}$ gelten. Dann erhalten wir für alle $h \in B_{r_1}$:

$$\begin{aligned}
(G \circ F)(y+h) - (G \circ F)(y) &= G\big[F(y+h)\big] - G\big[F(y)\big] \\
&= G[F(y) + F'_y(h) + o_1(y,h)] - G[F(y)] \\
&= G'_{F(y)}\big(F'_y(h) + o_1(y,h)\big) + o_2\big\{F(y), F'_y(h) + o_1(y,h)\big\} \\
&= \big(G'_{F(y)} \circ F'_y\big)(h) + G'_{F(y)} \circ o_1(y,h) + o_2\big\{F(y), F'_y(h) + o_1(y,h)\big\}.
\end{aligned}$$

Eine kurze Rechnung zeigt, daß

$$o(y,h) \equiv G'_{F(y)} \circ o_1(y,h) + o_2\big\{F(y), F'_y(h) + o_1(y,h)\big\}$$

die Eigenschaften (8.38) besitzt, da o_1 und o_2 diese Eigenschaften haben und weil $G'_{F(y)}$ und F'_y stetige lineare Abbildungen sind. Dies beweist die Differenzierbarkeit von $G \circ F$ und bestimmt den Wert der Ableitung zu

$$(G \circ F)'_y = G'_{F(y)} \circ F'_y \in E'.$$

Gâteaux-Ableitung: Die Verallgemeinerung des Begriffs der partiellen Ableitung einer Funktion $F : \mathbb{R}^n \to \mathbb{R}$ führt dagegen zum Begriff des GÂTEAUX-Differentials $\delta F_y(h)$:

Definition: *Es sei E ein normierter Raum. Gilt für die Abbildung $F : E \longrightarrow \mathbb{R}$ an der Stelle y:*

$$\lim_{\epsilon \to 0} \left| \frac{F(y+\epsilon h) - F(y)}{\epsilon} - \delta F(y,h) \right| = 0 \qquad \forall h \in E$$

mit einer Abbildung $\delta F(y,\cdot) : E \to \mathbb{R}$, die bezüglich der 2. Variablen weder stetig noch linear sein muß, so nennt man $\delta F(y,h)$ GÂTEAUX-*Differential von F an der Stelle y in Richtung von h. Ist $\delta F(y,h)$ zusätzlich linear und stetig in h, also ein Element des Dualraumes E' von E, so schreibt man*

$$\delta F(y,h) = \delta_y F(h) \tag{8.41}$$

und nennt $\delta_y F$ die GÂTEAUX-*Ableitung von F an der Stelle y.*

In endlichdimensionalen Räumen folgt aus der Existenz aller stetigen partiellen Ableitungen die Existenz des totalen Differentials. Ganz ähnlich ist es auch für Ableitungen in unendlich-dimensionalen Räumen,

Lemma: *Sei $F : E \longrightarrow \mathbb{R}$. Existiert in einer Umgebung U von y die* GÂTEAUX-*Ableitung $\delta_y F$ und ist sie stetig, so gilt*

$$\delta_y F = F'_y. \tag{8.42}$$

Eine stetige GÂTEAUX-Ableitung ist automatisch eine FRÉCHET-Ableitung. Umgekehrt gilt

Lemma: *Ist F an der Stelle y* FRÉCHET-*differenzierbar, so ist F auch* GÂTEAUX-*differenzierbar, und die beiden Differentiale sind gleich.*

Wir wollen im Folgenden annehmen, daß die Funktionale FRÉCHET-differenzierbar sind und werden deshalb nicht mehr zwischen FRÉCHET- und GÂTEAUX-Ableitung unterscheiden.

Höhere Variationen: Eine bequeme Methode, ein Funktional auf einem BANACH-Raum zu untersuchen, besteht darin, die Funktion $F(y+sh)$ der reellen Veränderlichen s für beliebige, aber feste $y, h \in E$ zu studieren. Informationen über das Verhalten in der Umgebung von $s = 0$ erhält man aus dem TAYLORschen Satz:

$$F(y+sh) = F(y) + \sum_{n=1}^{N} \frac{s^n}{n!} \Delta^n F(y,h) + R_N, \tag{8.43}$$

Ist $F(y+sh)$ auf $(-s_0, s_0)$ N-mal differenzierbar, dann gilt für das Restglied

$$\frac{R_N}{s^N} \longrightarrow 0 \quad \text{für} \quad s \longrightarrow 0.$$

Die *n-te Variation* von F im Punkte $y \in M$ in Richtung $h \in E$ ist definitionsgemäß

$$\Delta^n F(y,h) = \frac{d^n F(y+sh)}{ds^n}\Big|_{s=0}. \tag{8.44}$$

Für genügend reguläre Funktionale ist die erste Variation gleich der GÂTEAUX-Ableitung,

$$\Delta F(y,h) = \delta_y F(h). \tag{8.45}$$

Die zweite Variation unterscheidet zwischen Maxima und Minima eines Funktionals, ähnlich wie bei Funktionen im $\mathbb{R}^n$.

Als Beispiel betrachten wir die TAYLOR-Entwicklung des Funktionals (8.1). Wir fordern, daß $h(x)$ an den Enden des Intervalls $[x_1, x_2]$ verschwindet. Die Entwicklung (8.43) hat die folgende Form

$$\begin{aligned} F(y+sh) &= \int f(x,y,y') + s\int \Big(\frac{\partial f}{\partial y(x)}h(x) + \frac{\partial f}{\partial y'(x)}h'(x)\Big) \\ &+ \frac{s^2}{2}\int \left(\frac{\partial^2 f}{\partial y^2(x)}h^2(x) + \frac{2\partial f}{\partial y(x)\partial y'(x)}h(x)h'(x) + \frac{\partial^2 f}{\partial y'^2(x)}h'^2(x)\right) + o(s^3). \end{aligned}$$

Wir dürfen partiell integrieren, wobei wegen $h(t_1) = h(t_2) = 0$ keine Randterme auftreten, und erhalten folgende Variationen von F:

$$\begin{aligned} \Delta F(y,h) &= \int \Big(\frac{\partial f}{\partial y(x)} - \frac{d}{dx}\frac{\partial f}{\partial y'(x)}\Big)h(x) \\ \Delta^2 F(y,h) &= \int h(x)\left(\frac{\partial^2 f}{\partial y^2(x)} - \Big(\frac{\partial f}{\partial y(x)\partial y'(x)}\Big)' - \Big(\frac{\partial^2 f}{\partial y'^2(x)}\Big)'\frac{d}{dx} - \Big(\frac{\partial^2 f}{\partial y'^2(x)}\Big)\frac{d^2}{dx^2}\right)h(x). \end{aligned}$$

Fordern wir das Verschwinden der ersten Variation für beliebige $h(x)$, so erhalten wir wieder die EULER-LAGRANGE Gleichungen. Die zweite Variation enthält Information über die Stabilität der Lösungen. Ist zum Beispiel $\Delta^2 F(y,h)$ positiv für alle $h \neq 0$, dann ist y ein (lokales) Minimum, ist es negativ für alle $h \neq 0$, dann ist y ein (lokales) Maximum. Ein Sattelpunkt liegt vor, wenn $\delta^2 F(y,h)$ als Funktion von h positive und negative Werte annimmt.