Einführung in die Variationsrechnung

Teilnehmer:

David Alcer	Käthe-Kollwitz-Oberschule, Berlin
Pauline Bimberg	John-Lennon-Gymnasium, Berlin
Robert Clausecker	Herder-Oberschule, Berlin
Cinar Fidan	Staatl. Gymn. Bergschule, Apolda
Yann Paul Hartmann	Herder-Oberschule, Berlin
Tobias Kretschmar	Immanuel-Kant-Oberschule, Berlin
Dominik Nehls	Heinrich-Hertz-Oberschule, Berlin

Gruppenleiter:

Frank Feudel	Humboldt-Universität zu Berlin

Wir haben uns mit Minimierungsproblemen beschäftigt, bei denen Funktionen minimiert werden sollen, deren Argumente selbst Funktionen sind. Das mathematische Teilgebiet, welches sich mit der Minimierung von solchen Funktionalen (d.h. von Funktionen, deren Argumente selbst Funktionen sind) beschäftigt, nennt sich Variationsrechnung.

In diesem Kurs erfolgte eine Einführung in dieses Teilgebiet. Dazu wurde auf der Basis der Kenntnisse der Differential- und Integralrechnung versucht, ein ähnliches Kalkül für Funktionale zu entwickeln. Dabei haben wir viele Gemeinsamkeiten, aber auch überraschende Unterschiede kennen gelernt. Außerdem haben wir dabei festgestellt, dass es einen engen Bezug zu Differentialgleichungen gibt.

Mit dem entwickelten Kalkül wurden Fragestellungen aus Physik und Geometrie beantwortet, wie z.B.:
Was sind die kürzesten Verbindungen auf der Kugeloberfläche?
Welche Fläche der Ebene, die von einer Kurve mit gegebener Länge umrandet wird, hat den größten Flächeninhalt?
Wie muss der Graph einer Funktion aussehen, sodass der Rotationskörper, der durch Rotation um die Abszisse entsteht, eine minimale Mantelfläche hat?
Auf welcher Bahn rollt eine Kugel am schnellsten nach unten, wenn Start und Zielpunkt vorgegeben sind? (Brachistochronenproblem)

1 Einführung

Die Variationsrechung findet in vielen (anwendungsorientierten) Problemen der Mathematik Anwendung. Im Allgemeinen versucht man eine Einsetzung $u(x)$ in eine Funktion $F(x, u(x), u'(x))$ zu finden, sodass (1.1) minimal wird.

$$J(u) := \int_a^b F(x, u(x), u'(x))\,\mathrm{d}x \tag{1.1}$$

Dabei soll u eine auf $[a; b]$ stetig differenzierbare Funktion sein, an die verschiedene Randbedingungen gestellt werden können.

Die folgenden Probleme demonstrieren die Fragestellung.

1.1 Die Brachistochrone

Gegeben seien zwei Punkte A und B in einer Ebene. Es wirke ein homogenes Gravitationsfeld, das eine konstante Fallbschleunigung g ausübe in negative y-Richtung. Sei $u : [a; b] \to \mathbb{R}$ eine Funktion, deren Graph durch die Punkte A und B geht. Nun setzt man eine ideale, punktförmige Kugel auf den Punkt A und lässt diese zum Punkt B auf dem Graphen von u rollen. Für welches $u(x)$ ist die Rolldauer minimal?
Unter Anwendung verschiedener physikalischer Formeln ergibt sich die Gleichung (1.2) für die Horizontalgeschwindigkeit v_{Hor} der Kugel

$$v_{\text{Hor}} = \sqrt{\frac{2gu(x)}{1 + u'(x)^2}}. \tag{1.2}$$

Integration über die x-Koordinate liefert die Rolldauer; diese Größe möchten wir minimieren. Dabei legen wir den Ursprung in den Beginn der Bahn. Man kann damit folgende Gleichung herleiten

$$s = \int_0^{x_0} \frac{1}{v_{\mathrm{Hor}}} \, \mathrm{d}x = \int_0^{x_0} \sqrt{\frac{1+u'(x)^2}{2gu(x)}} \, \mathrm{d}x. \tag{1.3}$$

1.2 Minimalfläche

Wir betrachten den Rotationskörper um die x-Achse einer Funktion u zwischen den Stellen a und b. Nach den Guldinschen Regeln hat dieser die Mantelfläche aus (1.4).

$$M = 2\pi \int_a^b u(x)\sqrt{1+u'(x)^2} \, \mathrm{d}x \tag{1.4}$$

Welche Funktion $u(x)$ hat die minimale Mantelfläche M, wenn ihr Graph durch bestimmte Punkte läuft?

2 Abstände zwischen Funktionen und Stetigkeit des Funktionals

Um Differentialrechung betreiben zu können, benötigt man einen Ableitungsbegriff und für diesen wiederum einen Abstand. Im Folgenden gilt es, den Abstand zwischen zwei Funktionen zu berechnen. Doch was verlangen wir eigentlich von diesem Abstand?

1. Der Abstand ist eine nicht-negative, reelle Zahl, Dinge ohne Abstand sind einander gleich.
2. Der Abstand ist symmetrisch.
3. Die Gerade ist die kürzeste Verbindung.

Exakte Mathematisierung:

Ein Abstand wird durch eine Funktion $d : X \times X \to [0, \infty[$ (X ist eine Menge) mit folgenden Eigenschaften beschrieben:

1. $d(x,y) \geq 0, d(x,y) = 0 \Leftrightarrow x = y$,
2. $d(x,y) = d(y,x)$,
3. $d(x,y) \leq d(x,z) + d(z,y)$.

Für stetige Funktionen auf einem abgeschlossenen Intervall $[a,b]$ lässt sich ein natürlicher Abstand definieren.

$$d(f,g) := \max_{x \in [a,b]} \{|f(x) - g(x)|\} \tag{2.1}$$

Wir wollen jetzt die Grenzwertdefinition auf Funktionen übertragen, deren Argumente Funktionen sind, damit wir sie für Funktionen wie (1.1) verwendet können. Wir definieren:

$$\begin{aligned} \lim_{u \to u_0} J(u) = y_0 :\Leftrightarrow \forall \varepsilon > 0 \; \exists \delta > 0 \; \forall u \in C^1([a,b]) \\ 0 < d(u,u_0) < \delta \Rightarrow |J(u) - y_0| < \varepsilon. \end{aligned} \tag{2.2}$$

Gilt: $\lim_{u \to u_0} J(u) = J(u_0)$? Also: Ist J stetig? (Sonst brauchen wir über Ableitungen gar nicht zu reden.) Eine intuitive Rechnung ergibt:

$$\begin{aligned} \lim_{u \to u_0} J(u) &= \lim_{u \to u_0} \int_a^b F(x, u(x), u'(x)) \, \mathrm{d}x \\ &= \int_a^b \lim_{u \to u_0} F(x, u(x), u'(x)) \, \mathrm{d}x \\ &= \int_a^b F(x, u_0(x), u_0'(x)) \, \mathrm{d}x \quad \text{(falls } F \text{ stetig).} \end{aligned}$$

Jedoch nähern sich bezüglich unseres Abstandsoperators bei $u \to u_0$ die Ableitungen u' und u_0' nicht zwangsläufig einander an, weshalb die zweite Gleichheit nicht gesichert ist. Ein Gegenbeispiel liefert die Folge

$u_n(x) := \int_0^x t^n \, \mathrm{d}t$ auf $[0;1]$, die bezüglich d gegen die Nullfunktion konvergiert, deren Ableitung aber nicht punktweise gegen 0 geht.

Wir definieren daher einen neuen Abstand d, für den $d(u, u_0) \to 0 \Leftrightarrow u \to u_0 \wedge u' \to u_0'$ gilt:

$$d(f,g) := \max_{x\in[a,b]} \{|f(x) - g(x)|\} + \max_{x\in[a;b]} \{|f'(x) - g'(x)|\} \,. \tag{2.3}$$

Weiterhin ist zu begründen, dass man Integral und Grenzwert vertauschen kann. Im Allgemeinen ist das nicht möglich. Dazu betrachten wir folgendes Gegenbeispiel, wobei alle beteiligten Funktionen stetig auf [0;1] sind.

$$f_n(x) := \begin{cases} n^2x & \text{für } 0 \leq x \leq \frac{1}{n} \\ -n^2x + 2n & \text{für } \frac{1}{n} < x \leq \frac{2}{n} \\ 0 & \text{für } \frac{2}{n} < x \leq 1 \end{cases} .$$

Die Funktionen sind stetig und konvergieren punktweise gegen 0. Andererseits gilt:

$$\int\limits_0^1 f_n(x)dx = 1 \quad \forall n \in \mathbb{N}, n \geq 2.$$

Damit gilt nicht:

$$\lim_{n\to\infty} \int\limits_0^1 f_n(x)dx = \int\limits_0^1 \lim_{n\to\infty} f_n(x)dx.$$

In unserem speziellen Fall dürfen wir jedoch Grenzwert und Integral vertauschen. Dies ist mit dem folgenden Satz zu begründbar:

Satz. Seien $f_n : [a,b] \to \mathbb{R}$, $f : [a,b] \to \mathbb{R}$ stetig und $\max\limits_{x\in[a,b]} \{|f_n(x) - f(x)|\} \to 0$. Dann gilt:

$$\int\limits_a^b \lim_{n\to\infty} f_n(x) \, \mathrm{d}x = \lim_{n\to\infty} \int\limits_a^b f_n(x) \, \mathrm{d}x. \tag{2.4}$$

Daraus folgt die Stetigkeit von J, wenn F stetig ist.

3 Die erste Variation

Nun sind wir soweit, die Ableitung von (1.1) zu definieren, die man Variation nennt. Diese schreiben wir als $DJ_u(v)$, wobei v die Richtung angibt, in welche die Ableitung bestimmt wird. Wir definieren die erste Variation analog zur h-Methode bei normalen Ableitungen:

$$\begin{aligned} &DJ_u(v) := \lim_{t\to 0} \frac{J(u+tv)-J(u)}{t} \\ &\text{mit } v \in C_0^1([a,b]) := \left\{ v \in C^1([a;b]) : v(a) = v(b) = 0 \right\}. \end{aligned}$$

Da wir nur die Funktionen betrachten, die die Randbedingungen $u(a) = A$ und $u(b) = B$ erfüllen, kann $u(a)+tv(a) = u(a) = A$ nur gelten, falls $v(a) = 0$ ist. Gleiches gilt für $v(b)$.
Den Grenzwert kann man wie folgt berechnen:

$$\begin{aligned} &\lim_{t\to 0} \frac{J(u+tv)-J(u)}{t} \\ &= \lim_{t\to 0} \int_a^b \frac{J(F(x,u(x)+tv(x),u'(x)+tv'(x))) - F(x,u(x),u'(x)))}{t} dx \\ &= \int_a^b \lim_{t\to 0} \frac{J(F(x,u(x)+tv(x),u'(x)+tv'(x))) - F(x,u(x),u'(x)))}{t} dx. \end{aligned}$$

Den letzten Grenzwert können wir mit Hilfe des Mittelwertsatzes und der Dreiecksungleichung berechnen. Für das Vertauschen von Grenzwert und Integral wird dieselbe Argumentation wie für die Stetigkeit benötigt. Daraus lässt sich die erste Variation berechnen:

$$DJ_u(v) = \int_a^b \left[\frac{\partial F}{\partial y}(u,u(x),u'(x))v(x) + \frac{\partial F}{\partial z}(u,u(x),u'(x))v'(x) \right] \mathrm{d}x$$

$\forall v \in C_0^1([a,b])$.

Nun können wir mit der gerade definierten Ableitung Minima der Funktion J bestimmen. Die notwendige Bedingung für ein Minimum ist bekanntlich, dass die Ableitung gleich null ist. In diesem Fall gibt es jedoch unendlich viele Ableitungen, für alle Funktionen $v \in C_0^1([a,b])$ jeweils eine. Die notwendige Bedingung ist in diesem Fall logischerweise, dass für alle Funktionen $v \in C_0^1([a;b])$ die Ableitung gleich null ist, d. h. wenn für alle $v \in C_0^1([a;b])$ $DJ_u(v) = 0$ gilt.

4 Die Euler-Lagrange-Gleichungen

Die notwendige Bedingung führt zu einer Differentialgleichung. Ist u lokales Minimum von (1.1) so gilt die notwendige Bedingung (4.1).

$$DJ_u(v) = 0 \quad \forall v \in C_0^1([a,b]). \tag{4.1}$$

Durch Einsetzung in die o. g. Formel für die erste Variation ergibt sich:

$$\int_a^b \left(\frac{\partial F}{\partial y}(x, u(x), u'(x))v(x) + \frac{\partial F}{\partial z}(x, u(x), u'(x))v'(x) \right) \mathrm{d}x = 0$$

für alle $v \in C_0^1([a;b])$. Nach partieller Integration des zweiten Terms mit Aufleitung von v' erhält man, da $v(a) = v(b) = 0$:

$$\int_a^b \left[\frac{\partial F}{\partial y}(x, u(x), u'(x))v(x)\mathrm{d}x - \frac{\mathrm{d}}{\mathrm{d}x}\frac{\partial F}{\partial z}(x, u(x), u'(x)) \right] v(x)\ \mathrm{d}x = 0.$$

Jetzt benutzen wir das sogenannnte Fundamentallemma der Variationsrechnung, welches wie folgt lautet.

Fundamentallemma der Variationsrechnung. Sei $f : [a,b] \to \mathbb{R}$ stetig und es gelte:

$$\int_a^b f(x)v(x)\ \mathrm{d}x = 0 \quad \forall v \in C_c^\infty(]a,b[).$$

Dann gilt: $f \equiv 0$.

Dabei bezeichnet $C_c^\infty(]a, b[)$ die Menge der unendlich oft differenzierbaren Funktionen auf $]a, b[$, die außerhalb eines abgeschlossenen Teilintervalls innerhalb von $]a, b[$ gleich 0 sind.

Damit kann man v aus obigen Termen eliminieren und erhält:

$$\frac{\partial F}{\partial y}(x, u(x), u'(x)) - \frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\partial F}{\partial z}(x, u(x), u'(x))\right) = 0.$$

Diese Gleichung nennt man *Euler-Lagrange-Gleichung*, die man in den meisten Fällen nach $u(x)$ (analytisch) auflösen kann.

5 Hinreichende Bedingungen für Minimierer

Wie bei normalen Funktionen wollen wir untersuchen, ob die zweite Ableitung ($D^2 J_u(v)$) ein hinreichendes Kriterium für ein Minimum liefert.

Falls u ein Minimum ist, so gilt: $D^2 J_u(v) \geq 0 \quad \forall v \in C_0^1([a; b]$.

Aus $DJ_u(v) = 0$ und $D^2 J_u(v) > 0, \forall v \neq 0$, folgt aber nicht, dass u ein lokales Minimum ist.

Es ist zum Beispiel:

$$\begin{aligned}
J(u) &:= \int_{-1}^{1} x^2(u'(x))^2 + x(u'(x))^3 \, \mathrm{d}x \\
DJ_u(v) &= \int_{-1}^{1} [2u'(x)x^2 + 3(u'(x))^2 x]v'(x) \, \mathrm{d}x \\
D^2 J_u(v) &= \int_{-1}^{1} [2x^2 + 6u'(x)x]v'(x)^2 \mathrm{d}x.
\end{aligned}$$

Betrachten wir $u_0 \equiv 0$, so ist $DJ_{u_0}(v) = 0$ und $D^2 J_{u_0}(v) > 0, \forall v \neq 0$, aber u ist kein lokales Minimum.

Wir betrachten jetzt $u_{\varepsilon,h}$ mit: $u_{\varepsilon,h}(x) = \begin{cases} -\varepsilon x + \varepsilon h & x \in [0;h] \\ \varepsilon x + \varepsilon h & x \in [-h;0] \\ 0 & \text{sonst} \end{cases}$.

Für $h = \frac{3}{4}\varepsilon$ ist $J(u_{\varepsilon,h}) < 0$. $(J(u_{\varepsilon,h}) = \frac{2}{3}h^2(h - \frac{3}{2}\varepsilon))$
Lassen wir $\varepsilon \to 0$ gehen, so gehen auch $u, u' \to 0$. Also geht die Folge gegen 0 bezüglich unseres Abstandes und hat bei Einsetzen in J Werte kleiner als 0, was begründet, dass 0 kein lokales Minimum ist.

Aber:
Wenn $D^2J_u(v) > 0, \forall v \in C_0^1([a;b])$, $v \neq 0$ und für jedes $u \in C^1([a;b])$, welches die passenden Randbedingungen erfüllt, so ist jeder Punkt u mit $DJ_u(v) = 0$, $\forall v \in C_0^1([a;b])$, ein lokales Minimum.

6 Beispiele

6.1 Lösung des Brachistochronenproblems

Nachdem wir nun die Euler-Lagrange-Gleichung hergeleitet haben, können wir versuchen, damit das anfänglich behandelte Problem der Brachistochrone zu lösen. Dazu ist die Euler-Lagrange-Gleichung auf

$$T(u) = \int_0^{x_0} \sqrt{\frac{1 + u'(x)}{2gu(x)}} dx$$

anzuwenden. Wenn man aber $\frac{\partial F}{\partial z}(x, u(x), u'(x))$ berechnet,

$$\left(\frac{\partial F}{\partial z}(x, u(x), u'(x)) = \frac{u'(x)}{\sqrt{u(x)u'(x) + u(x)}} \right)$$

wird schnell klar, dass wir nach Ableiten dieses Termes nach x die zweite Ableitung $u''(x)$, die erste $u'(x)$ sowie auch $u(x)$ in der Euler-Lagrange-Gleichung haben, was eine analytische Lösung fast unmöglich macht. Daher wird eine Vereinfachung der Euler-Lagrange-Gleichung verwendet. Wenn nur $u(x)$ sowie $u'(x)$ in dem Funktional J vorkommen, was in $T(y)$ der Fall ist, und $u(x)$

eine Lösung der Euler-Lagrange-Gleichung ist, gilt für

$$E(y,z) := z\frac{\partial F}{\partial z} - F(y,z),$$

dass $\frac{d}{dx}E(u(x),u'(x)) = 0$ und damit $E(u(x),u'(x)) = c$ ist, wobei c eine Konstante ist. Dies lässt sich mit der Kettenregel recht leicht herleiten. Eingesetzt ergibt sich

$$\begin{aligned}
\frac{u'^2(x)}{\sqrt{2gu(x)(u'^2(x)+1)}} - \sqrt{\frac{u'^2(x)+1}{2gu(x)}} &= -c \\
\frac{u'^4(x)}{2gu(x)(u'^2(x)+1)} + c\frac{u'^2(x)}{\sqrt{2gu(x)(u'^2(x)+1)}} + c^2 &= \frac{1+u'^2(x)}{2gu(x)} \\
2u'^2(x)\sqrt{2gu(x)(u'^2(x)+1)} + c^2 &= (1+u'^2(x))^2 - u'^4(x).
\end{aligned}$$

Es lässt sich leicht einsehen, dass diese Gleichheit dann und nur dann gilt, wenn $u(x)(u'(x)+1) = C$ gilt, wobei C eine andere Konstante darstellt. Wenn man allerdings versucht, diese Differentialgleichung elementar zu lösen, merkt man schnell, dass sich diese Differentialgleichung so nicht lösen lässt. Deshalb schreiben wir den Graphen als parametrisierte Kurve. Damit kommt man über recht komplizierte Zwischenschritte zu der Differentialgleichung

$$\dot{y}(t) = \sqrt{\frac{2g}{C}}\sqrt{Cy(t) - y^2(t)},$$

welche elementar lösbar ist. Das Ergebnis mit der Randbedingung $y(0) = 0$ lautet

$$y(t) = \frac{C}{2}\left(1 - \cos\sqrt{\frac{2g}{C}}t\right).$$

Aus der Kettenregel und der Aufspaltung der Geschwindigkeit in zwei Richtungen ergibt sich außerdem

$$y(t)(\dot{x}^2(t) + \dot{y}^2(t)) = C\dot{x}^2(t).$$

Aus diesen beiden Gleichungen lässt sich nun folgende Parametrisierung für die gesuchte Bahnkurve finden:

$$\gamma(t) = \left(\frac{C}{2}\left(\sqrt{\frac{2g}{C}}t - \sin(\sqrt{\frac{2g}{C}}t)\right), \frac{C}{2}\left(1 - \sin(\sqrt{\frac{2g}{C}}t)\right)\right).$$

Dies ist ein Teil aus einer Zykloide, welche entsteht, wenn die Bewegung eines Punktes auf dem Rand eines Kreises verfolgt wird, wenn der Kreis mit konstanter Drehgeschwindigkeit entlang der x-Achse gerollt wird. Dabei kann es sogar passieren, dass die Kugel bei optimaler Bahn am Ende sogar ein Stück „hochlaufen" muss. Es ist natürlich noch zu überprüfen, ob wirklich ein Minimum vorliegt.

6.2 Lösung des Minimalflächenproblems

Wie bereits in der Einleitung beschrieben ist die Oberfläche eines Rotationskörpers um die x-Achse in den Grenzen von a bis b durch folgende Formel gegeben:

$$J(u) = 2\pi \int\limits_a^b u(x)\sqrt{1 + (u'(x))^2}dx.$$

Wieder ergibt die Euler-Lagrange Gleichung eine Gleichung, in der mehrere Ableitungen von u vorkommen. Da aber in J wieder nur u und u' vorkommen, nutzen wir wieder unsere Hilfsfunktion E mit $E(u(x), u'(x)) = c_1$. Mithilfe dessen erhält man die Differentialgleichung:

$$u'(x) = \sqrt{\frac{(u(x))^2 - (c_1)^2}{(c_1)^2}}.$$

Diese Differentialgleichung gilt es nun zu lösen und dies ist durch Trennung der Variablen und Integration durch Substitution möglich. Man erhält dann für unsere Funktion u die folgende allgemeine Struktur:

$$u(x) := c_1 \cdot \cosh\left(\frac{x + c_2}{c_1}\right).$$

Mit dieser Formel lässt sich nun berechnen, welche Funktion, die durch zwei feste Punkte geht, durch Rotation um die x-Achse, einen Rotationshyperboloiden mit minimaler Oberfläche erzeugt. c_1 und c_2 sind Parameter, welche sich durch ein Gleichungssystem für den individuellen Fall berechnen lassen.