Das isoperimetrische Problem

Thomas Peters
Thomas’ Mathe-Seiten
www.mathe-seiten.de

18. Oktober 2003

Das isoperimetrische Problem, auch bekannt als das Problem der Dido, ist es, unter allen geschlossenen ebenen Kurven gleichen Umfangs diejenige zu finden, welche die größte Fläche umschließt. Es zeigt sich, dass die Lösung dieses Problems der Kreis ist. In diesem Artikel werden drei verschiedene Beweise dieser Tatsache vorgestellt.

Einleitung

Beim isoperimetrischen Problem zeigt sich beispielhaft, welche einfache Gestalt mathematisch tiefe Ergebnisse annehmen können bzw. umgekehrt, dass zur Lösung eines in wenigen einfachen Worten formulierbaren Problems durchaus fortgeschrittene mathematische Methoden angewendet werden müssen. Wir können es hier nur behandeln, weil wir erstens annehmen, dass eine Lösung dieses Extremalproblems überhaupt existiert, und zweitens bei der analytischen Behandlung für die Lösungskurve gewisse Voraussetzungen machen. Diese Annahmen können wir hier leider nicht beweisen.

Beweis über Geometrie

Angenommen, die ebene Kurve K löse das isoperimetrische Problem. Man macht sich leicht klar, dass K sich dann nicht selbst überschneiden darf. Ansonsten könnte man ja die entstehenden Schleifen einfach weglassen und erhielte immer noch eine geschlossene Kurve, die denselben Flächeninhalt einschlïeßen würde, aber kürzer wäre, was der Voraussetzung widerspräche, K löse das isoperimetrische Problem.

Desweiteren ist K konvex, d. h. für alle Punkte P, Q auf K liegt die Verbindungsstrecke $\overline{PQ}$ im Inneren von K. Denn nehmen wir einmal an, K wäre nicht konvex. Dann gäbe es Punkte P_1 und P_2, so dass die Strecke $\overline{P_1P_2}$ außerhalb von K liegt. Dann könnte man den Teil von K zwischen P_1 und P_2 an $\overline{P_1P_2}$ spiegeln und erhielte so eine Kurve gleicher Länge, aber größeren Flächeninhalts, was erneut ein Widerspruch ist.

Mit demselben Argument muss jede Strecke zwischen Punkten P, Q auf K, die K in zwei Teilkurven gleicher Länge teilt, auch den Flächeninhalt in zwei gleiche Hälften teilen. Ansonsten könnte man nämlich die größere Hälfte an $\overline{PQ}$ spiegeln und erhielte eine Kurve gleicher Länge, aber erneut größeren Flächeninhalts.

Sei nun $\overline{PQ}$ eine Strecke mit der Eigenschaft, dass sie K in zwei gleich lange Hälften teilt. Wir betrachten eine dieser Teilkurven $\widetilde{K}$ und einen Punkt R auf $\widetilde{K}$. Wichtig für unsere Überlegungen ist nun das Dreieck $\triangle(PQR)$. Wir stellen uns vor, die Kurvenstücke von P nach R sowie von R nach Q wären in ihrer Form fest, aber die Punkte P und Q könnte man auseinander ziehen bzw. zusammen schieben. Dann hängt die Fläche unter $\widetilde{K}$ nur vom Dreieck $\triangle(PQR)$ ab, und an diesem ist lediglich die Strecke $\overline{PQ}$ variabel. Bezeichnet h das Lot von P auf $\overline{RQ}$, so ist die Fläche dieses Dreiecks

$$A = \frac{1}{2}\,\overline{RQ}\,h.$$

Nennen wir den Winkel bei R, der $\overline{PQ}$ gegenüber liegt, α, so gilt $\sin\alpha = h/\overline{PR}$, d. h.

$$A = \frac{1}{2}\,\overline{PR}\,\overline{RQ}\,\sin\alpha.$$

In diesem Ausdruck sind nun alle Größen fest außer α. Offenbar wird die Fläche A maximal für $\alpha = 90^\circ$. Da der Punkt R auf $\widetilde{K}$ aber beliebig war, folgt aus der Umkehrung des Satzes von Thales sofort, dass $\widetilde{K}$ ein Halbkreis, K also ein Kreis, sein muss.

Beweis über Variationsrechnung

Wir können uns nach den Symmetrieüberlegungen im letzten Abschnitt darauf reduzieren, eine Kurve $y(x)$ der Länge $L > 2a$ mit $y(-a) = y(a) = 0$ zu finden, die mit der Strecke $-a \leq x \leq a$ eine größtmögliche Fläche einschließt. In der *Variationsrechnung* lernt man, dass das Integral

$$\int_{-a}^{a} y(x)\,\mathrm{d}x$$

unter der Nebenbedingung

$$\int_{-a}^{a} \sqrt{1+y'(x)^2}\,\mathrm{d}x = L$$

dann extremal wird, wenn die *Euler-Lagrange-Gleichung*

$$\frac{\mathrm{d}}{\mathrm{d}x}\frac{\partial(y(x)+\lambda\sqrt{1+y'(x)^2})}{\partial y'} - \frac{\partial(y(x)+\lambda\sqrt{1+y'(x)^2})}{\partial y} = 0$$

erfüllt ist, wobei λ ein Parameter ist, der aus den Nebenbedingungen bestimmt werden muss.

Auswerten der Euler-Lagrange-Gleichung gibt

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

Dies kann integriert werden zu

$$\frac{\lambda y'(x)}{\sqrt{1+y'(x)^2}} = x + C_1$$

mit einer Integrationskonstanten C_1. Um dies weiter integrieren zu können, müssen wir zunächst nach $y'(x)$ auflösen. Quadrieren und Multiplizieren mit dem Nenner liefert

$$\lambda^2 y'(x)^2 = (x+C_1)^2(1+y'(x)^2),$$

also

$$y'(x) = \pm\frac{x+C_1}{\sqrt{\lambda^2-(x+C_1)^2}}.$$

Eine weitere Integration ergibt

$$y(x) + C_2 = \pm\sqrt{\lambda^2-(x+C_1)^2}$$

mit der Integrationskonstanten C_2. Dies lässt sich leicht umordnen zur Kreisgleichung

$$(x+C_1)^2 + (y(x)+C_2)^2 = \lambda^2.$$

Dies ist die Gleichung eines Kreises mit Mittelpunkt $(-C_1, -C_2)$ und Radius $|\lambda|$.

Wir haben bisher nur gezeigt, dass die Lösung, so sie existiert, ein Kreis ist. Dass diese Lösung nun auch wirklich vorhanden ist, zeigen wir dadurch, dass wir die Integrationskonstanten und den Parameter λ aus den Nebenbedingungen bestimmen[1]. Aus $y(-a) = y(a) = 0$ ergibt sich das Gleichungssystem

$$(C_1 - a)^2 + C_2^2 = \lambda^2$$
$$(C_1 + a)^2 + C_2^2 = \lambda^2.$$

Durch Subtraktion der Gleichungen folgt mit der dritten binomischen Formel die Bedingung

$$(C_1 - a)^2 - (C_1 + a)^2 = (-2a)(2C_1) = 0.$$

Da natürlich $a \neq 0$ ist kann nur $C_1 = 0$ sein. Damit ist aber auch $C_2 = \pm\sqrt{\lambda^2 - a^2}$ festgelegt. Die beiden Vorzeichen entsprechen den beiden zur x-Achse symmetrischen Positionen des Kreises, bei denen einmal der Betrag der Fläche oberhalb und einmal unterhalb der x-Achse maximal wird. Es bleibt noch die Bestimmung von λ. Dazu rechnen wir

$$L = \int_{-a}^{a} \sqrt{1 + y'(x)^2}\,\mathrm{d}x = \int_{-a}^{a} \sqrt{1 + \frac{x^2}{\lambda^2 - x^2}}\,\mathrm{d}x = \int_{-a}^{a} \sqrt{\frac{\lambda^2}{\lambda^2 - x^2}}\,\mathrm{d}x$$
$$= \int_{-a}^{a} \sqrt{\frac{1}{1 - (x/\lambda)^2}}\,\mathrm{d}x = \lambda\left[\arcsin\frac{x}{\lambda}\right]_{-a}^{a} = 2\lambda\arcsin\frac{a}{\lambda}.$$

Daraus folgt

$$\sin\frac{L}{2\lambda} = \frac{a}{\lambda}.$$

Dies ist nun leider eine transzendente Gleichung in λ, so dass wir nicht hoffen können, einen allgemeinen Ausdruck für λ zu finden. Wir können uns aber leicht klarmachen, dass ein λ existiert, das diese Gleichung erfüllt. Substituieren wir $u = L/(2\lambda)$, so lautet die Gleichung

$$\sin u = \frac{2a}{L}u.$$

Nun ist aber $2a/L < 1$, d. h. die Gerade rechts hat eine geringere Steigung als der Sinus links mit $\sin' 0 = \cos 0 = 1$. Daher existiert auf jeden Fall ein Schnittpunkt. Dabei können wir eine negative Lösung ignorieren, da die Kreisgleichung und die Bestimmungsgleichungen für C_2 und λ nicht vom Vorzeichen von λ abhängen, es kommen also keine neuen Lösungen hinzu. Bleibt noch auszuschließen, dass mehrere positive Schnittstellen existieren. Diese Überlegung sei dem Leser als Übungsaufgabe überlassen.

Beweis über Fourier-Reihen

Wir machen nun den Ansatz, die Lösungskurve in parametrisierter Form $(x(t), y(t))$ zu finden. Da die Kurve geschlossen ist, müssen x und y eine gewisse Periodizität haben, und wir

[1] Prinzipiell wäre es denkbar, dass es gar kein mit den Nebenbedingungen verträgliches λ gibt.

wählen die Parametrisierung so, dass die Periode gerade 2π beträgt, d. h. x und y sind Funktionen $[0, 2\pi] \to \mathbb{R}$. Dann können wir die Funktionen x und y bequem in eine Fourier-Reihe entwickeln. Dies liefert

$$x(t) = \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \cos nt + b_n \sin nt$$

und

$$y(t) = \frac{c_0}{2} + \sum_{n=1}^{\infty} c_n \cos nt + d_n \sin nt.$$

In der Analysis zeigt man, dass man den Flächeninhalt einer so parametrisierten Kurve durch die Formel

$$A = \frac{1}{2} \int_0^{2\pi} x(t)\dot{y}(t) - \dot{x}(t)y(t) \,\mathrm{d}t$$

berechnen kann, wogegen ihre Länge durch

$$L = \int_0^{2\pi} \sqrt{\dot{x}^2(t) + \dot{y}^2(t)} \,\mathrm{d}t$$

gegeben ist. Dabei bezeichnet der Punkt die Differentiation nach dem Parameter t. Wir benötigen also

$$\dot{x}(t) = \sum_{n=1}^{\infty} n(a_n \sin nt - b_n \cos nt)$$

sowie

$$\dot{y}(t) = \sum_{n=1}^{\infty} n(c_n \sin nt - d_n \cos nt).$$

Aufgrund der Orthogonalitätsrelationen für die trigonometrischen Funktionen ist nun

$$A = \frac{\pi}{2} \left(\sum_{n=1}^{\infty} n(b_n c_n - a_n d_n) - \sum_{n=1}^{\infty} n(a_n d_n - b_n c_n) \right) = \pi \sum_{n=1}^{\infty} n(a_n d_n - b_n c_n).$$

Weiter gilt

$$\int_0^{2\pi} \dot{x}^2(t) \,\mathrm{d}t = \pi \sum_{n=1}^{\infty} n^2(a_n^2 + b_n^2)$$

und

$$\int_0^{2\pi} \dot{y}^2(t) \,\mathrm{d}t = \pi \sum_{n=1}^{\infty} n^2(c_n^2 + d_n^2).$$

Bis jetzt haben wir nur den Definitionsbereich der Parametrisierung festgelegt. Wir verlangen nun zusätzlich, dass $\dot{x}^2(t) + \dot{y}^2(t) = \mathrm{const.}$ gilt. In der kinematischen Interpretation der Parametrisierung bedeutet dies, dass die Kurve mit konstanter Geschwindigkeit durchlaufen wird.

Es ist anschaulich klar, dass dies immer gehen muss, daher verzichten wir auf den nicht ganz einfachen Beweis. Aus der Formel für die Länge der Kurve folgt dann für die Konstante

$$\dot{x}^2(t) + \dot{y}^2(t) = \frac{L^2}{4\pi^2}.$$

Integration über $[0, 2\pi]$ und Einsetzen der Ableitungen ergibt

$$\pi \sum_{n=1}^{\infty} n^2(a_n^2 + b_n^2 + c_n^2 + d_n^2) = \frac{L^2}{2\pi}.$$

Damit haben wir insgesamt

$$\begin{aligned} L^2 - 4\pi A &= 2\pi^2 \sum_{n=1}^{\infty} n^2(a_n^2 + b_n^2 + c_n^2 + d_n^2) - 4\pi^2 \sum_{n=1}^{\infty} n(a_n d_n - b_n c_n) \\ &= 2\pi^2 \sum_{n=1}^{\infty} (na_n - d_n)^2 + (nb_n + c_n)^2 + (n^2 - 1)(c_n^2 + d_n^2) \end{aligned}$$

gezeigt.

Da dieser Term offensichtlich nicht negativ ist, erhalten wir die *isoperimetrische Ungleichung*

$$4\pi A \leq L^2.$$

Wir müssen nun die Fourier-Koeffizienten so wählen, dass hier Gleichheit herrscht. Dies ist genau dann der Fall, wenn alle Summanden der unendlichen Reihe verschwinden. Im Fall $n = 1$ bedeutet dies, dass $a_1 = d_1$ und $b_1 = -c_1$ sein muss. Im Fall $n > 1$ muss zunächst $c_n^2 + d_n^2 = 0$ sein, was $c_n = d_n = 0$ und ferner $a_n = b_n = 0$ impliziert. Damit haben wir die Funktionen x und y bestimmt zu

$$x(t) = \frac{a_0}{2} + a_1 \cos t + b_1 \sin t$$

und

$$y(t) = \frac{c_0}{2} - b_1 \cos t + a_1 \sin t.$$

Dies ist wegen

$$\left(x(t) - \frac{a_0}{2}\right)^2 + \left(y(t) - \frac{c_0}{2}\right)^2 = a_1^2 + b_1^2$$

gerade ein Kreis mit Mittelpunkt $(a_0/2, c_0/2)$ und Radius $R = \sqrt{a_1^2 + b_1^2}$. Wir erhalten weiter $A = \pi(a_1^2 + b_1^2) = \pi R^2$ und $L = \sqrt{2}\pi\sqrt{2(a_1^2 + b_1^2)} = 2\pi R$.

Man sieht an der isoperimetrischen Ungleichung übrigens, dass das isoperimetrische Problem zu folgendem Problem äquivalent ist: Finde unter allen Kurven mit gleichem Flächeninhalt diejenige, welche die geringste Länge hat.