

Die Jensensche Ungleichung

von Alfred Schreiber und Emese Vargyas

1 Einleitung

Ist f eine reellwertige, auf einem Intervall $I \subseteq \mathbb{R}$ konvexe Funktion, so gilt für alle $x_1, x_2, \dots, x_n \in I$ und beliebige nichtnegative reelle Zahlen p_1, p_2, \dots, p_n mit $p_1 + p_2 + \dots + p_n = 1$:

$$f(p_1 x_1 + p_2 x_2 + \dots + p_n x_n) \leq p_1 f(x_1) + p_2 f(x_2) + \dots + p_n f(x_n). \quad (1)$$

Diese nach [Jensen 1906] benannte Ungleichung¹ (die wir mit JU(n) abkürzen, $n \geq 2$) ist ein ebenso einfaches wie produktives Werkzeug mit vielseitigen inner- und außermathematischen Anwendungen, z. B. in der Geometrie, Stochastik, Ökonomie und Physik. Eine zentrale Rolle spielt die JU in der Theorie der Mittelwerte, lassen sich doch durch spezielle Wahl der Funktion f mit geringem Aufwand die bekannten Ungleichungen zwischen den quasiarithmetischen Mittelwerten gewinnen².

Im vorliegenden Beitrag wollen wir einen elementaren und anschaulichen Zugang zur JU entwickeln, ohne dabei auf Anwendungen einzugehen. Die darin ausgeführten Beweise sind in eine *explikative Heuristik* eingebettet. Diese folgt der Grundidee, zunächst und vorrangig den geometrischen Gehalt des Konvexitätsbegriffs zu entfalten und danach die Ergebnisse in die „Sprache der Funktionen“ zu übertragen. Die algebraisch-analytischen Mittel, die dabei – gerade in Abschnitt 2 (zur Begriffsbildung und Charakterisierung der Konvexität) – benutzt werden, sind nahe am Unterrichtsstoff der Sekundarstufen, wenn sie nicht ohnehin auf dem Lehrplan stehen. Im Abschnitt 3 behandeln wir die Verallgemeinerung und eine wichtige Spezialisierung der JU. Die betreffenden Analysen führen zu weiteren Zugängen und Beweisvarianten; speziell ergibt sich für differenzierbare Funktionen ein zwangloser Anschluss an den schulischen Analysisstoff.

2 Elementare Eigenschaften konvexer Funktionen

Wir betrachten die Begriffe „konvexe Figur“ und „konvexe Funktion“ und stellen ihren Zusammenhang her. Anschließend entwickeln wir vier Kennzeichnungen von Konvexität und erarbeiten ihre Äquivalenz. In diesem Kontext schälen sich zwei Beweise von JU(2) heraus.

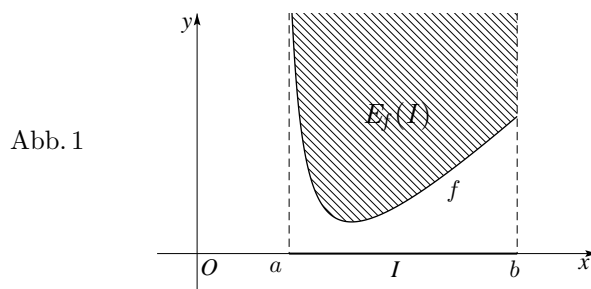
2.1 Der Begriff der Konvexität

Das Attribut „konvex“ wird in mathematischen Lehrtexten zwei Arten von Objekten beigelegt: Figuren (Punktmengen im Raum und in der Ebene) und Funktionen (hier der Einfachheit halber eingeschränkt auf reelle Funktionen einer Variablen).

Eine Figur heißt *konvex*, wenn sie zu je zwei ihrer Punkte P, Q stets auch die ganze Verbindungsstrecke \overline{PQ} enthält. Konvexe Figuren verkörpern eine anschaulich leicht fassliche Gestaltqualität und spielen in Alltag und Technik eine wichtige Rolle (wovon einiges sich bereits Grundschulkindern erschließt)³. Merkwürdig mag daher erscheinen – dies freilich nur auf den ersten Blick –, dass der Begriff „konvexe Funktion“ üblicherweise durch JU(2) definiert wird⁴. Es stellt sich hier die Frage nach einem passenden Bindeglied, das es erlaubt, die Konvexität einer Funktion $f : I \rightarrow \mathbb{R}$ ($I \subseteq \mathbb{R}$ Intervall) durch die Konvexität einer (von f abhängigen) Figur (als Teilmenge von \mathbb{R}^2) auszudrücken.

Schaut man sich einen typischen Verlauf des Funktionsgraphen an (Abb. 1), so wird schnell klar, dass die schraffierte Menge als Kandidat der gesuchten Figur in Betracht kommt. Sie wird *Epigraph von f über I* genannt und wie folgt notiert⁵

$$E_f(I) := \{(x, y) \in I \times \mathbb{R} \mid f(x) \leq y\}. \quad (2)$$



Im allgemein lässt sich $E_f(I)$ nicht beschränken, z. B. wenn (wie in Abb. 1 angedeutet) die Funktion in einem Randpunkt ihres Definitionsintervalls nicht definiert ist und divergiert (hier: $f(x) \rightarrow +\infty$ für $x \rightarrow a+0$).

Es liegt nun nahe, die Konvexität von f mit Bezug auf den Epigraphen zu reformulieren. Wir nennen f *konvex auf I* , wenn $E_f(I)$ eine konvexe Figur ist. Natürlich ist diese einfache Neudefinition (die wir für das Folgende mit KNV abkürzen) nur gerechtfertigt, wenn sie sich als gleichwertig mit JU(2) erweist. Das ist in der Tat der Fall, wie wir in 2.2 im Zusammenhang mit weiteren Charakterisierungen darlegen werden.

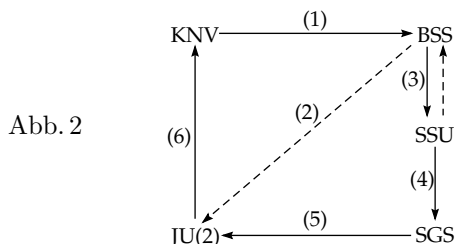
Zweifelloos bietet JU(2) einen formelhaften Ausdruck, der auf das Arbeiten mit Funktionen zugeschnitten ist und sich in der Praxis gut bewährt. Gerade deswegen lohnt es sich den Weg abzuschreiten, der ihn von der ursprünglichen einfachen Figurenkonvexität trennt. Schließlich ist JU(2) schon der wesentliche Spezialfall der Jensenschen Ungleichung; auf der Wegstrecke von KNV nach JU(2) dürfen wir daher einen von geometrischer Anschauung geleiteten Zugang zu ihr erwarten⁶.

2.2 Charakterisierungen im Zusammenhang

Markiert man auf dem Graphen einer konvexen Funktion (von der Art wie in Abb. 1) zwei Punkte, so bildet ihre Verbindungsstrecke eine Sehne, die an jeder Stelle des Intervalls oberhalb des Bogens liegt, den der Funktionsgraph zwischen den betreffenden Punkten bildet. Dieser anschaulich evidente Bogen-Sehnen-Satz (BSS) ist eine einfache und direkte Kon-

sequenz aus KNV. Aus ihm lassen sich weitere Aussagen gewinnen: eine Sehnensteigungs-Ungleichung (SSU), der Stützgeraden-Satz (SGS) und schließlich auch JU(2).

Das Diagramm in Abb. 2 zeigt eine mögliche (natürliche) Abhängigkeitsstruktur für die fünf genannten Aussagen:

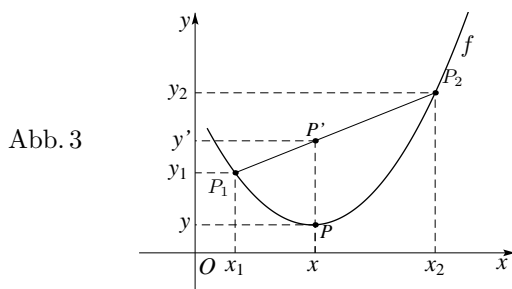


Jeder Pfeil bedeutet Herleitbarkeit, und zwar in einem doppelten Sinn, der logische Folgerung ebenso einschließt wie einen heuristischen Übergang (nahegelegt durch den begrifflich-sachlichen Gehalt). Bereits aufgrund des Ringschlusses über die durchgezogenen Pfeile ergibt sich die Äquivalenz der fünf Aussagen. Die strichlierten Pfeile sind Vorschläge für vertiefendes Üben. Zum Verständnis der Jensenschen Ungleichung trägt die Herleitung von JU(2) aus BSS in besonderem Maße bei, sofern dazu der Verlauf des Bogens unterhalb der Sehne in Koordinatenschreibweise (unter Rückgriff auf einfachste analytische Geometrie) übersetzt werden muss.

Wir behandeln die einzelnen Herleitungen in der Reihenfolge, die an den Pfeilen vermerkt ist. Schrittweise werden dabei die benötigten Sprachmittel (Bezeichnungen, Begriffe) entwickelt. Entsprechend präzisieren wir die in Rede stehenden Eigenschaften bzw. Aussagen, wann und wieweit es die jeweilige Beweislage erfordert.

2.2.1 BSS: Bogen-Sehnen-Satz

Wir betrachten eine im Intervall I konvexe Funktion f und zwei Punkte $P_1 = (x_1, y_1)$, $P_2 = (x_2, y_2)$ auf ihrem Graphen, dabei ist $x_1, x_2 \in I$. Abb. 3 zeigt an einer Stelle $x \in I$, $x_1 \leq x \leq x_2$, den zugehörigen Punkt $P = (x, y)$ auf dem Bogen $\widehat{P_1 P_2}$, $y = f(x)$, und den entsprechenden Punkt $P' = (x, y')$ auf der Sehne $\overline{P_1 P_2}$.



Der eingangs in bloßer Wortfassung gegebene BSS besagt demnach, dass P unterhalb von P' liegt oder mit P' zusammenfällt⁷. Dies ergibt sich unmittelbar aus der Konvexität (KNV)

von f , denn es ist $P' \in \overline{P_1 P_2} \subseteq E_f(I)$, d.h.

$$y = f(x) \leq y'. \quad (3)$$

2.2.2 JU(2): Jensensche Ungleichung

Im wesentlichen ist (3) bereits die JU; es fehlt nur noch eine „Aufschlüsselung“ der Ordinaten y und y' nach den Abszissen x_1, x_2 . Als technisches Rüstzeug wird dazu lediglich das Bestimmen von Geraden(gleichungen) durch zwei gegebene Punkte benötigt.

Im ersten Schritt drücken wir die Zwischenlage von $x \in [x_1, x_2]$ durch Koeffizienten $p_1, p_2 \geq 0$ mit $p_1 + p_2 = 1$ aus: $x = p_1 x_1 + p_2 x_2$. Dies wird einsichtiger durch Umnotation zu $x = (1-p)x_1 + px_2$ (mit $p = p_2$ und $1-p = 1-p_2 = p_1$). Durchläuft nämlich p , bei 0 beginnend, das Intervall $[0, 1]$, so durchläuft x , bei x_1 beginnend, das Intervall $[x_1, x_2]$ ⁸. Damit ergibt sich für den Punkt P (in Abb. 3) die Darstellung:

$$y = f(x) = f(p_1 x_1 + p_2 x_2). \quad (4)$$

Nun zu P' : Als Punkt auf der Geraden durch P_1, P_2 hat er Koordinaten (x, y') , welche die Geraden-Gleichung erfüllen:

$$y' = g(x) = mx + c. \quad (5)$$

Die Steigung m und der y -Achsenabschnitt c sind durch die Bedingungen $g(x_1) = y_1$, $g(x_2) = y_2$ festgelegt (was bei der Schülerschaft – zumindest aus Formelsammlungen – als bekannt vorauszusetzen ist). Für die Berechnung von y' muss man die Werte m, c nicht kennen. Wir setzen $x = p_1 x_1 + p_2 x_2$ in (5) ein und erhalten

$$\begin{aligned} y' &= g(p_1 x_1 + p_2 x_2) \\ &= m(p_1 x_1 + p_2 x_2) + c \\ &= p_1 m x_1 + p_2 m x_2 + (p_1 + p_2)c \\ &= p_1(m x_1 + c) + p_2(m x_2 + c) \\ &= p_1 g(x_1) + p_2 g(x_2) \\ &= p_1 y_1 + p_2 y_2 \\ &= p_1 f(x_1) + p_2 f(x_2). \end{aligned} \quad (6)$$

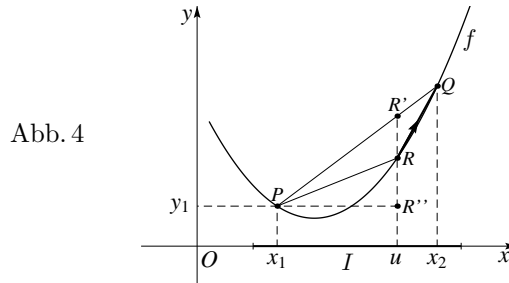
Nimmt man (3), (4) und (6) zusammen, so steht die JU(2) sofort da:

$$f(p_1 x_1 + p_2 x_2) \leq p_1 f(x_1) + p_2 f(x_2). \quad (7)$$

Bemerkung: In ähnlich einfacher Weise gewinnen wir in 2.2.6 aus JU(2) die geometrische Konvexität zurück. Zum Nachweis der Äquivalenz von JU(2) und KNV ist somit kaum mehr nötig als die Bogen-Sehnen-Eigenschaft in Koordinaten auszudrücken. Nichtsdestoweniger bilden die im Folgenden näher untersuchten Eigenschaften konvexer Funktionen eine optimale, sinnvolle Ergänzung. Die etwas anspruchsvolleren Aussagen SSU und SGS tragen zum vertieften geometrischen Verständnis der Konvexität bei.

2.2.3 SSU: Sehnensteigungs-Ungleichung

Aus dem BSS lassen sich auf anschauliche Weise Erkenntnisse über die Steigung von Sehnen gewinnen. Wir betrachten dazu zwei Sehnen mit einem gemeinsamen Endpunkt $P = (x_1, y_1)$ (auf dem Graphen einer konvexen Funktion $f : I \rightarrow \mathbb{R}$; Abb. 4):



Für die Steigung m_{PQ} der Sehne \overline{PQ} gilt

$$m_{PQ} = \frac{f(x_2) - f(x_1)}{x_2 - x_1} =: m_f(x_1, x_2), \quad (8)$$

ein Ausdruck, der offensichtlich in x_1, x_2 symmetrisch ist:

$$m_{PQ} = m_f(x_1, x_2) = m_f(x_2, x_1) = m_{QP}. \quad (9)$$

Sei nun R ein beliebiger Punkt auf dem Bogen \widehat{PQ} (des Funktionsgraphen). Durch diese Wahl liegt R , ebenso wie Q , rechts von P und bildet mit P die Sehne \overline{PR} . Dem Anschein nach ist zu vermuten, dass die Steigung von \overline{PR} die Steigung von \overline{PQ} nicht übertrifft:

$$m_{PR} \leq m_{PQ}. \quad (10)$$

(Dies ist natürlich klar, wenn $m_{PR} \leq 0$ und $m_{PQ} \geq 0$ ist; anderenfalls erhalten wir eine Situation wie in Abb. 4 oder einen analogen Fall mit beiden Steigungen ≤ 0 .)

Beweis von (10): Die Parallele zur y -Achse durch R möge \overline{PQ} in R' und die durch P verlaufende Gerade ($y = y_1$) in R'' treffen. Dann gilt:

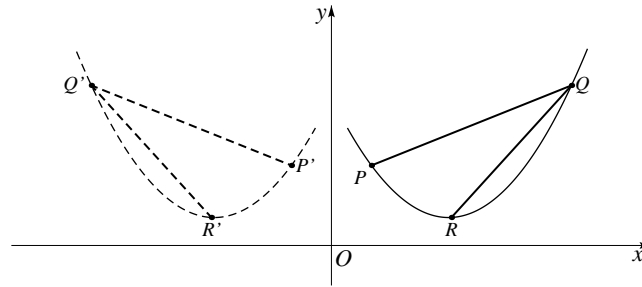
$$m_{PR} = \frac{|R''R|}{|PR''|} \leq \frac{|R''R| + |RR'|}{|PR''|}. \quad (11)$$

Nun ist (im angenommenen Fall) $m_{PR} \geq 0$, d. h. R'' liegt unterhalb von R oder fällt mit R zusammen. Ferner liegt nach dem BSS der Punkt R unterhalb von R' oder fällt mit R' zusammen. Alle drei Punkte R, R', R'' liegen auf einer Geraden ($x = u$); es gilt daher $|R''R| + |RR'| = |R''R'|$, und mit (11) erhalten wir die zu beweisende Ungleichung

$$m_{PR} \leq \frac{|R''R'|}{|PR''|} = m_{PR'} = m_{PQ}.$$

Der Sachverhalt bleibt unverändert auch in dem Fall, dass der gemeinsame Sehnenendpunkt auf der anderen Seite liegt (nämlich: Q rechts von P und R wie im ersten Quadranten von Abb. 5):

Abb. 5



Wir zeigen – durch Zurückführung auf (10) – die Ungleichung

$$m_{QP} \leq m_{QR}. \quad (12)$$

Um eine Orientierung wie in Abb. 4 zu gewinnen, werde die Koordinatenebene an der y -Achse gespiegelt. Die Bildpunkte P', R', Q' von P, R, Q liegen dann auf dem Spiegelbild des Funktionsgraphen; dieser gehört ebenfalls zu einer konvexen Funktion. Im zweiten Quadranten liefert daher (10) zunächst

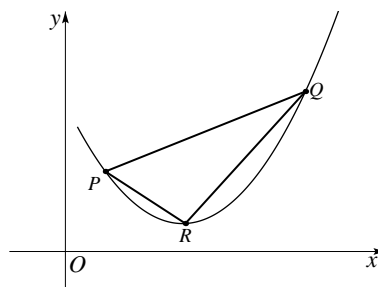
$$m_{Q'R'} \leq m_{Q'P'}.$$

Für das Übrige sind lediglich Vorzeichenwechsel unter der Spiegelung zu beachten:

$$m_{QP} = -m_{Q'P'} \leq -m_{Q'R'} = m_{QR}. \quad (13)$$

Fassen wir (10) und (13) zusammen, so erhalten wir, unter Berücksichtigung der Symmetrie (9), die Sehnensteigungs-Ungleichung für ein beliebiges Sehnendreieck PRQ , dessen Ecken auf dem Graphen einer konvexen Funktion liegen; dabei ist R ein Punkt auf dem Bogen \widehat{PQ} (Abb. 6).

Abb. 6



$$m_{PR} \leq m_{PQ} \leq m_{QR}. \quad (14)$$

Wir greifen auf (8) zurück und schreiben (mit den Bezeichnungen aus Abb. 4) die Ungleichung (14) noch einmal in Koordinaten auf:

$$m_f(x_1, u) \leq m_f(x_1, x_2) \leq m_f(x_2, u). \quad (15)$$

Diese Ungleichungskette erschließt uns bemerkenswerte Einsichten:

a) Die Ungleichung linkerhand besteht unter der Voraussetzung $u \leq x_2$ und besagt daher, dass $m_f(x_1, \cdot)$ eine im zweiten Argument monoton wachsende Funktion darstellt. Wegen (9) gilt dasselbe auch für $m_f(\cdot, x_2)$. Die SSU bringt somit zum Ausdruck, dass *die Sehnensteigung einer konvexen Funktion in beiden Argumenten (Abszissen) monoton wächst*. Dies ist bereits

das Vorstadium der Kennzeichnung konvexer Funktionen in der Differenzialrechnung (durch monotonen Wachstum der Ableitung).

b) Wählen wir im Intervall $[x_1, x_2]$ (von Abb. 4) weitere Punkte v_1, v_2 mit $v_1 < u < v_2$, so ergibt sich mittels SSU:

$$m_f(x_1, u) \leq m_f(v_1, u) \leq m_f(v_1, v_2) \leq m_f(v_2, u) \leq m_f(x_2, u).$$

Ist nun M der größere der Absolutbeträge der beiden Steigungen auf der linken und rechten Seite dieser Ungleichungskette, so erhalten wir die Abschätzung

$$|m_f(v_1, v_2)| \leq M$$

und unter Verwendung von (8):

$$|f(v_1) - f(v_2)| \leq M \cdot |v_1 - v_2|. \quad (16)$$

Diese Ungleichung ist von besonderer Bedeutung. An ihr ist direkt abzulesen: Liegen die Argumente v_1, v_2 nahe beieinander, so gilt das auch für die zugehörigen Funktionwerte⁹.

Im Kern enthält (16) bereits alles Nötige zum Nachweis einer Stützgeraden in jedem Punkt einer konvexen Kurve. Wir wollen den Sachverhalt in 2.2.4 entwickeln, uns dabei aber enger an der SSU in den Fassungen (14) und (15) orientieren.

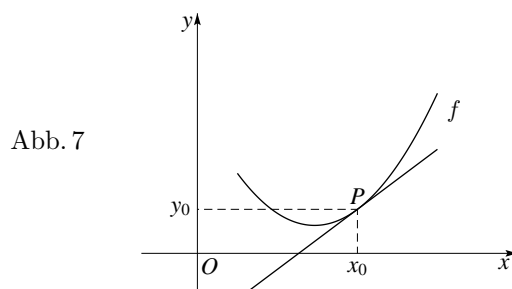
2.2.4 SGS: Stützgeraden-Satz

An eine beliebige konvexe Figur (der Ebene) lässt sich durch jeden Randpunkt eine Gerade derart legen, dass sämtliche inneren Punkte der Figur in einer der beiden Halbebenen liegen, die durch die Gerade begrenzt (getrennt) werden. Eine solche Gerade heißt *Stützgerade*¹⁰.

Wir wollen im Folgenden den Begriff der Stützgeraden für Funktionen $f : I \rightarrow \mathbb{R}$ entwickeln und einen dazu passenden Stützgeraden-Satz (SGS) herleiten. Sei $x_0 \in I$ und $P = (x_0, y_0)$ ein Punkt auf dem Graphen von f , d. h. $y_0 = f(x_0)$. Die Geraden durch P haben die Gleichung

$$g(x) = y_0 + m(x - x_0). \quad (17)$$

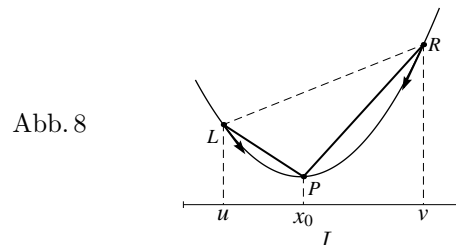
Die Steigung m soll nun variiert werden. Lässt sich m so wählen, dass $g(x) \leq f(x)$ für alle $x \in I$ gilt, so bedeutet dies, dass das Geradenstück über dem Intervall I unterhalb des Funktionsgraphen verläuft und keinen inneren Punkt von $E_f(I)$ enthält.



In diesem Fall ist (17) die Gleichung einer *Stützgeraden im Punkt x_0* (Abb. 7). Der SGS für

konvexe Funktionen besagt nun: *In jedem Punkt des Definitionsintervalls existiert eine Stützgerade.*

Wir legen uns einen Beweis zurecht, der von der SSU Gebrauch macht. Dazu wählen wir x_0 aus dem Innern¹¹ von I und betrachten links und rechts von $P = (x_0, f(x_0))$ auf dem Funktionsgraphen weitere Punkte $L = (u, f(u))$ und $R = (v, f(v))$; Abb. 8 zeigt eine solche Konstellation mit eingezeichneten Sehnen.



Offenbar kommen die von P ausgehenden Sehnen (oder Sekanten) als Stützgeraden in x_0 nicht in Frage, da sie innere Punkte von $E_f(I)$ enthalten. Wandern hingegen L und R gegen P (wie in Abb. 8 durch Pfeile angedeutet), so entstehen Grenzlagen beider Sehnen, zwischen denen die möglichen Stützgeraden verlaufen.

Den Schlüssel zur Durchführung dieser Beweisidee liefert die SSU in der Form (15):

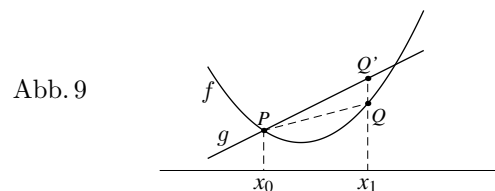
$$m_f(u, x_0) \leq m_f(u, v) \leq m_f(v, x_0).$$

Wir lassen nun L fest und R gegen P wandern. Dabei geht $v \rightarrow x_0$, und $m_f(v, x_0)$ durchläuft eine monoton fallende Folge (v nimmt ab!). Diese Folge ist durch $m_f(u, x_0)$ nach unten beschränkt, sie strebt daher gegen einen Grenzwert¹² m_+ . Wir erhalten somit $m_f(u, x_0) \leq m_+$. Völlig analog ergibt sich bei $u \rightarrow x_0$ (Annäherung von L an P von links) eine durch $m_f(v, x_0)$ (besser noch: durch m_+) nach oben beschränkte, monoton wachsende Folge von Werten $m_f(u, x_0)$ mit einem entsprechenden Grenzwert m_- . Es gilt:

$$m_- \leq m_+. \tag{18}$$

Zum Nachweis einer Stützgeraden in x_0 genügt es, ein einzelnes $m \in \mathbb{R}$ anzugeben, das $g(x) = f(x_0) + m(x - x_0) \leq f(x)$ für alle $x \in I$ erfüllt, etwa $m = \frac{m_- + m_+}{2}$, das bei [Kirsch 1993, 137] genannt wird.

Tatsächlich liefert aber jedes m mit $m_- \leq m \leq m_+$ eine Stützgerade (17). Wir zeigen dies indirekt und nehmen dazu $g(x_1) > f(x_1)$ für ein $x_1 \in I$ an. Ohne Einschränkung sei $x_1 > x_0$ (der Fall $x_1 < x_0$ verläuft analog, und $x_1 = x_0$ ist nicht möglich). Mit den Bezeichnungen aus Abb. 9 liegt dann $Q' = (x_1, g(x_1))$ oberhalb von $Q = (x_1, f(x_1))$.



Damit wird $m_{PQ} < m_{PQ'} = m \leq m_+$, obwohl doch nach SSU für alle Punkte Q rechts von P gilt: $m_+ \leq m_{PQ}$.

Bemerkung: Ist f in x_0 differenzierbar, so ist $m_- = m_+ (= m)$, und man hat in der Tangente (17) mit der Steigung m die in diesem Fall eindeutig festgelegte Stützgerade durch den Kurvenpunkt $(x_0, f(x_0))$. Tritt in (18) Ungleichheit ein, so ergibt sich eine Situation wie z. B. in Abb. 1. (an der Stelle $x = a$) oder Abb. 11.

2.2.5 JU(2) aus dem Stützgeraden-Satz

Wir geben nun einen zweiten Beweis der Jensenschen Ungleichung ($n = 2$), diesmal mit dem SGS. Dazu werden $x_1, x_2 \in I$ (Definitionsintervall der konvexen Funktion f) beliebig gewählt und irgendein Zwischenpunkt $u = p_1 x_1 + p_2 x_2 \in I$ betrachtet, wobei $p_1, p_2 \geq 0$ und $p_1 + p_2 = 1$.

Nach dem SGS existiert eine Stützgerade in u , d. h. (abhängig von u) ein $m \in \mathbb{R}$ derart, dass $f(x) \geq f(u) + m(x - u)$ für alle $x \in I$. Insbesondere gilt

$$f(x_1) \geq f(u) + m(x_1 - u),$$

$$f(x_2) \geq f(u) + m(x_2 - u).$$

Multipliziert man diese Ungleichungen beziehentlich mit p_1 und p_2 und addiert sie seitenweise, so ergibt sich

$$p_1 f(x_1) + p_2 f(x_2) \geq (p_1 + p_2) f(u) + m(p_1 x_1 + p_2 x_2) - m u (p_1 + p_2)$$

und damit (wegen $p_1 + p_2 = 1$ und $u = p_1 x_1 + p_2 x_2$):

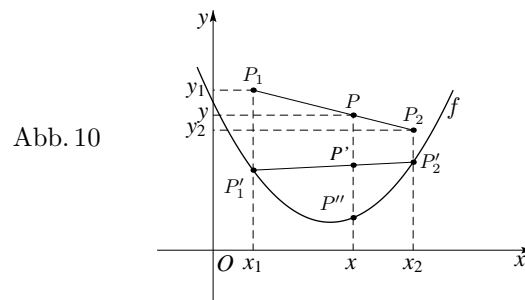
$$p_1 f(x_1) + p_2 f(x_2) \geq f(u) = f(p_1 x_1 + p_2 x_2),$$

mithin JU(2).

2.2.6 KNV: Konvexität der Funktion

Schließlich bleibt zu zeigen, dass JU(2) die Konvexität der Funktion f (im Sinne der in 2.1 gegebenen Definition) nach sich zieht.

Wir haben die Konvexität des Epigraphen $E_f(I)$ nachzuweisen, d.h. für irgend zwei Punkte $P_1, P_2 \in E_f(I)$ die Inklusion $\overline{P_1 P_2} \subseteq E_f(I)$. Zu beliebigem Punkt $P = (x, y) \in \overline{P_1 P_2}$ zeigen wir $P \in E_f(I)$, d. h. $f(x) \leq y$. Die Parallelen zur y -Achse durch P_1 und P_2 treffen den Graphen von f in den Punkten $P'_1 = (x_1, y'_1)$ und $P'_2 = (x_2, y'_2)$ (siehe Abb. 10).



Die Parallele zur y -Achse durch P schneide die Sehne $\overline{P'_1 P'_2}$ in $P' = (x, y')$ und den Bogen $\overline{P'_1 P'_2}$ in $P'' = (x, y'')$. Dabei ist $y'' = f(x)$; wir haben demnach $y'' \leq y$ zu zeigen. Da nach Voraussetzung die Punkte P_1, P_2 zu $E_f(I)$ gehören, ist $y'_1 = f(x_1) \leq y_1$ und $y'_2 = f(x_2) \leq y_2$. Wir stellen x wie üblich als Zwischenpunkt dar: $x = p_1 x_1 + p_2 x_2$ (mit $p_1, p_2 \geq 0$, $p_1 + p_2 = 1$) und erhalten mittels JU(2):

$$\begin{aligned} y'' &= f(x) = f(p_1 x_1 + p_2 x_2) \\ &\leq p_1 f(x_1) + p_2 f(x_2) \\ &= p_1 y'_1 + p_2 y'_2 \\ &\leq p_1 y_1 + p_2 y_2. \end{aligned}$$

In Abb. 10 erkennt man sofort an der darin enthaltenen Ähnlichkeitsfigur, dass x die Strecke zwischen x_1 und x_2 im selben Verhältnis teilt wie y die Strecke zwischen y_1 und y_2 . Mithin ist $p_1 y_1 + p_2 y_2 = y$.

Dieser Beweis komplettiert den Ringschluss (Abb. 2). Die Äquivalenz der Aussagen KNV, BSS, SSU, SGS, JU(2) ist damit gezeigt.

3 Verallgemeinerung und Spezialisierung

Im letzten Abschnitt wollen wir die Verallgemeinerung der Jensenschen Ungleichung behandeln und dabei auch auf Zugänge und weitere Beweise unter der Voraussetzung der Differenzierbarkeit eingehen. Den Abschluss bildet eine Erörterung der (spezielleren) Schwerpunktform der JU sowie eine kurze Diskussion des Gleichheitsfalls. – Natürlich erhält man durch Wahl spezieller Funktionen eine unübersehbare Fülle von Sonderfällen, die eigentlich Anwendungen der JU sind. Auf einige elementare Beispiele wird in den anderen Beiträgen dieses Heftes eingegangen.

3.1 Übergang zu mehr als zwei Variablen

Es liegt nahe, die Zwei-Punkte-Kombination $p_1 x_1 + p_2 x_2$ in JU(2) durch eine n -Punkte-Kombination $p_1 x_1 + p_2 x_2 + \dots + p_n x_n$ zu ersetzen ($p_1, p_2, \dots, p_n \geq 0$, $p_1 + p_2 + \dots + p_n = 1$ und $x_1, x_2, \dots, x_n \in I$). Auf diese Weise erhalten wir die verallgemeinerte JU(n) (wie in (1)):

$$f(p_1 x_1 + p_2 x_2 + \dots + p_n x_n) \leq p_1 f(x_1) + p_2 f(x_2) + \dots + p_n f(x_n). \quad (19)$$

Üblicherweise wird (19) durch vollständige Induktion bewiesen; JU(2) dient dabei als Induktionsanfang. Da solche Induktionen häufig mit unmotivierten „Kunstgriffen“ leichtgängig (aber deswegen nicht verständlicher) gemacht werden, ziehen wir es hier vor, eine *heuristische* (aber gleichwohl strenge) Überlegung für den Übergang von JU(2) zu JU(3) zu entwickeln. Sie lässt sich ohne weiteres auf den allgemeinen Induktionsschritt übertragen. Nebenbei wird deutlich, dass über JU(2) hinaus nichts Neues in die Verallgemeinerung(en) einfließt.

Wir zeigen nun JU(3)

$$\begin{aligned} f(p_1 x_1 + p_2 x_2 + p_3 x_3) &\leq p_1 f(x_1) + p_2 f(x_2) + p_3 f(x_3) \\ (p_1, p_2, p_3 \geq 0, p_1 + p_2 + p_3 = 1, x_1, x_2, x_3 \in I \text{ beliebig}) \end{aligned} \quad (20)$$

durch Zurückführung auf JU(2). Aus dem Argument von f linkerhand machen wir zunächst eine Summe aus zwei Summanden $u = p_1 x_1 + p_2 x_2$ und $p_3 x_3$. Um JU(2) anwenden zu können, muss in u ein Koeffizient abgespalten werden, der mit p_3 summiert 1 ergibt. Dies ist natürlich $1 - p_3$, d. h. wir schreiben für das Argument von f :

$$(1 - p_3) \frac{u}{1 - p_3} + p_3 x_3.$$

Wir setzen dies auf der linken Seite von (20) ein und wenden JU(2) an:

$$\begin{aligned} f(p_1 x_1 + p_2 x_2 + p_3 x_3) &= f\left((1 - p_3) \frac{u}{1 - p_3} + p_3 x_3\right) \\ &\leq (1 - p_3) f\left(\frac{u}{1 - p_3}\right) + p_3 f(x_3). \end{aligned} \quad (21)$$

Dazu bleibt noch $f\left(\frac{u}{1 - p_3}\right)$ auszuwerten. Betrachten wir das Argument ein wenig genauer:

$$\frac{u}{1 - p_3} = \frac{p_1}{1 - p_3} x_1 + \frac{p_2}{1 - p_3} x_2 \in I.$$

Offenbar ist $\frac{p_1}{1 - p_3} + \frac{p_2}{1 - p_3} = 1$ (wegen $p_1 + p_2 + p_3 = 1$) und wir können JU(2) ein zweites Mal anwenden:

$$f\left(\frac{u}{1 - p_3}\right) \leq \frac{p_1}{1 - p_3} f(x_1) + \frac{p_2}{1 - p_3} f(x_2).$$

Setzen wir dies in (21) ein, so steht JU(3) da.

Bemerkung: Wir erachten diese Beweisvariante dem Niveau der S II als angemessen. Eine ähnliche Darstellung gibt [Rüthing 1995]. Eine geringfügig abweichende Induktion findet man bei [Fichtenholz 1975, 280], wo die $(n + 1)$ -gliedrige Summe im Argument von f durch Zusammenfassung der letzten beiden Summanden in eine n -gliedrige Summe umgeschrieben wird.

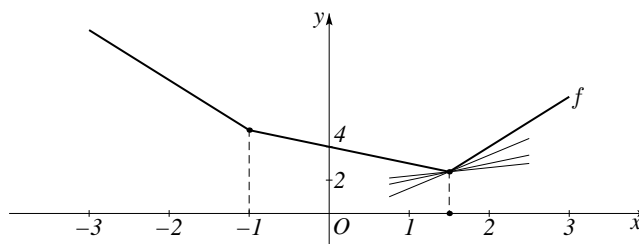
3.2 Differenzierbarkeitsannahmen

Nach den geometrisch-anschaulich geprägten Studien zur JU wird man das Thema natürlich auch aufs analytische Gleis setzen. Dies geschieht durch die Voraussetzung, dass die in Rede stehenden (konvexen) Funktionen $f: I \rightarrow \mathbb{R}$ im Innern I° ihres Definitionsintervalls differenzierbar sind. Das ist keine sehr einschneidende Annahme. Denn zum einen besitzt eine auf I konvexe Funktion in jedem Punkt von I° eine links- und eine rechtsseitige Ableitung (was sich bei genauerem Hinsehen aus den Darlegungen unter 2.2.4, insbesondere (18), ergibt). Zum anderen ist die Menge der Punkte, an denen f' nicht existiert, höchstens abzählbar¹³.

Beispiele von Funktionen dieser Art können bzw. sollten im S II-Unterricht ohnehin einen Platz haben – zumindest solche von der Bauart

$$f(x) = |x + 1| + |2x - 3|, \quad -3 \leq x \leq 3,$$

Abb. 11



die an endlich vielen Stellen keine Tangente, (dafür aber eine unendliche Schar von Stützgeraden) besitzen (Abb. 11). Zwischen den Knickstellen hat man dann wieder differenzierbare Funktionen.

Für den Rest dieses Abschnitts 3.2 sei $f : I \rightarrow \mathbb{R}$ als eine in I konvexe Funktion vorausgesetzt, deren Ableitung $f'(x)$ für alle $x \in I^\circ$ existiert.

Diese Annahme wirkt sich in leicht überschaubarer Weise auf SSU und SGS aus. Laut SSU ist die Sehnen- bzw. Sekantensteigung $m_f(\cdot, \cdot)$ eine in beiden Argumenten monoton wachsende Funktion. Für $x_1, x, x_2 \in I^\circ$, $x_1 < x < x_2$ hat man daher nach (15):

$$m_f(x_1, x) \leq m_f(x_1, x_2) \leq m_f(x_2, x).$$

Für $x \rightarrow x_1$ wird aus dem linken Teil

$$f'(x_1) \leq m_f(x_1, x_2) \tag{22}$$

und entsprechend für $x \rightarrow x_2$ aus dem rechten Teil

$$m_f(x_1, x_2) \leq f'(x_2). \tag{23}$$

Aus (22) und (23) folgt sofort $f'(x_1) \leq f'(x_2)$, was besagt, dass f' eine auf I° monoton wachsende Funktion ist. – Damit hat man die Anbindung des Themas an den üblichen Analysisstoff.

Im Folgenden wollen wir unterschiedliche Beweise der JU erörtern; dabei wird wesentlicher Gebrauch von der Differenzierbarkeit der Funktion gemacht.

1.) In 2.2.5 haben wir die JU aus dem SGS hergeleitet. Wir beweisen nun den SGS direkt aus der Differenzierbarkeit und erhalten damit einen analytischen Beweis der JU. In [Kirsch 1993] wird dieser Idee wegen ihrer Einfachheit und Anschaulichkeit der Vorzug gegeben.

Sei $x_0 \in I^\circ$. Wir zeigen: Die Tangente $t(x) := f(x_0) + f'(x_0)(x - x_0)$ an den Punkt $(x_0, f(x_0))$ des Funktionsgraphen ist Stützgerade: $t(x) \leq f(x)$ ($x \in I$). Es ist $t(x) \leq f(x)$ äquivalent zu

$$f'(x_0)(x - x_0) \leq f(x) - f(x_0). \tag{24}$$

Für $x < x_0$ ist (24) gleichwertig zu $f'(x_0) \geq m_f(x_0, x)$, eine bereits bewiesene Aussage (vgl. (23)). Im Fall $x > x_0$ gilt (24) genau dann, wenn $f'(x_0) \leq m_f(x_0, x)$, was ebenfalls gültig ist (vgl. (22)). Für $x = x_0$ hat man sofort $t(x) = f(x)$.

2.) In [Treiber 1991] wird ein Beweis vorgestellt, der von der Lösungsmethode für Extremwertaufgaben Gebrauch macht. Bewiesen wird freilich ein Spezialfall (die JU(n) in Schwerpunktform) und dabei die Extremwertbestimmung im Induktionsschritt durchgeführt. Die

entscheidende Idee tritt aber deutlicher hervor (und ist dann auch einfacher anzuwenden), wenn man direkt JU(2) beweist.

Dazu nehmen wir ein beliebiges $x_1 \in I$ und $p_1, p_2 > 0$ mit $p_1 + p_2 = 1$ an. Abhängig von f definieren wir eine Differenzfunktion $h: I \rightarrow \mathbb{R}$ durch

$$h(x) := p_1 f(x_1) + p_2 f(x) - f(p_1 x_1 + p_2 x).$$

Um JU(2) zu zeigen, genügt offenbar der Nachweis von $h(x) \geq 0$ ($x \in I$). Wir erreichen dies, indem wir ermitteln, ob und an welcher Stelle h ein Minimum besitzt. Tatsächlich nimmt $h(x)$ für $x = x_1$ seinen kleinsten Wert an, weshalb gilt: $h(x) \geq h(x_1) = f(x_1) - f(x_1) = 0$ für alle $x \in I$.

Um das zu zeigen, berechnen wir zunächst die erste Ableitung von h :

$$h'(x) = p_2 f'(x) - p_2 f'(p_1 x_1 + p_2 x).$$

Demnach gilt:

$$\begin{aligned} h'(x) \geq 0 &\iff f'(x) \geq f'(p_1 x_1 + p_2 x) && \text{(beachte: } p_2 > 0) \\ &\iff x \geq p_1 x_1 + p_2 x && (f' \text{ monoton wachsend)} \\ &\iff p_1 x = (1 - p_2)x \geq p_1 x_1 \\ &\iff x \geq x_1 && \text{(beachte: } p_1 > 0) \end{aligned}$$

Ebenso ergibt sich $h'(x) \leq 0$ gdw. $x \leq x_1$ und $h'(x) = 0$ gdw. $x = x_1$. Daraus wird ersichtlich, dass $h(x)$ bei $x = x_1$ ein Minimum besitzt. Die bisher ausgeschlossenen Werte $p_1 = 0$ oder $p_2 = 0$ liefern sofort $h(x) = 0$ ($x \in I$).

3.) Bereits vor Jensen hatte Hölder die Ungleichung (1) veröffentlicht (siehe Anm. 1). Im Beweis wird der Mittelwertsatz der Differenzialrechnung verwendet und die Konvexität von f durch das monotone Wachstum von f' ausgedrückt. Wir skizzieren hier Hölders Überlegung für JU(2).

Seien $x_1, u, x_2 \in I$ mit $x_1 < u < x_2$. Der Zwischenpunkt u werde wie üblich in der Form $u = p_1 x_1 + p_2 x_2$ dargestellt ($p_1, p_2 > 0$, $p_1 + p_2 = 1$). Dann erhalten wir nach dem Mittelwertsatz geeignete Werte $\xi_1 \in (x_1, u)$ und $\xi_2 \in (u, x_2)$, für die gilt:

$$f'(\xi_1) = \frac{f(x_1) - f(u)}{x_1 - u} \text{ und } f'(\xi_2) = \frac{f(u) - f(x_2)}{u - x_2}.$$

Damit ergibt sich

$$\begin{aligned} p_1 f(x_1) + p_2 f(x_2) &= f(u) + p_1 f'(\xi_1)(x_1 - p_1 x_1 - p_2 x_2) + p_2 f'(\xi_2)(x_2 - p_1 x_1 - p_2 x_2) \\ &= f(u) + p_1 p_2 (x_2 - x_1) (f'(\xi_2) - f'(\xi_1)). \end{aligned}$$

Hier ist der zweite Summand nichtnegativ (wegen $x_1 < x_2$, $\xi_1 < \xi_2$ und aufgrund des monotonen Wachstums von f'), woraus $f(u) \leq p_1 f(x_1) + p_2 f(x_2)$ folgt.

3.3 Die Schwerpunktform der JU

Ein einfacher Sonderfall der JU(n) entsteht durch Wahl konstanter Gewichtskoeffizienten $p_1 = p_2 = \dots = p_n = \frac{1}{n}$:

$$f\left(\frac{x_1 + x_2 + \dots + x_n}{n}\right) \leq \frac{1}{n}(f(x_1) + f(x_2) + \dots + f(x_n)). \quad (25)$$

Wir wollen diese Gestalt der JU als ihre *Schwerpunktform* bezeichnen. Ungleichung (25) ist eine unmittelbare Folgerung aus JU(n); sie lässt sich aber auch unabhängig davon beweisen.

In der Originalarbeit [Jensen 1906] beweist der Autor (25) durch Induktion. Dabei legt er als Definition von „ f konvex“ den Fall $n = 2$ von (25) zu Grunde (zugleich Induktionsanfang). Im Induktionsschritt wird die Anzahl der Summanden verdoppelt. Die so zwischen den Zweierpotenzen $n = 2, n = 4, \dots, n = 2^m$, usw. entstehenden Lücken lassen sich dann durch eine Rückwärtsinduktion à la Cauchy schließen.

Ein ebenso einfacher wie aufschlussreicher Beweis, der zudem ganz im Rahmen der Schulmathematik bleibt, ergibt sich aus der Schwerpunkteigenschaft des arithmetischen Mittels. Betrachten wir dazu irgendwelche Punkte $P_1 = (x_1, y_1), P_2 = (x_2, y_2), \dots, P_n = (x_n, y_n)$ der Ebene sowie die Mittelwerte $\bar{x}_{(n)} := \frac{x_1 + x_2 + \dots + x_n}{n}$ und $\bar{y}_{(n)} := \frac{y_1 + y_2 + \dots + y_n}{n}$ ihrer Abzissen bzw. Ordinaten. Dann ist $S_n := (\bar{x}_{(n)}, \bar{y}_{(n)})$ der Schwerpunkt der Punktmenge $\{P_1, P_2, \dots, P_n\}$. Wir zeigen dies und darüberhinaus für eine im Intervall I konvexe Funktion f :

Sind P_1, P_2, \dots, P_n Punkte auf dem Funktionsgraphen von f , dann gilt $S_n \in E_f(I)$, d. h. ihr Schwerpunkt liegt im Epigraphen der Funktion¹⁴.

Diese Aussage drückt die JU in Schwerpunktform geometrisch aus. Denn $S_n \in E_f(I)$ bedeutet nach Definition (2):

$$f(\bar{x}_{(n)}) \leq \bar{y}_{(n)}. \quad (26)$$

Da die Punkte P_1, P_2, \dots, P_n auf dem Funktionsgraphen liegen, ist $y_i = f(x_i)$ ($i = 1, 2, \dots, n$), und wir erhalten nach Einsetzen der y_i in (26) sofort (25).

Wir führen nun einen Beweis durch Induktion; als Hilfsfigur dient Abb. 12. Ohne Einschränkung werde $x_1 < x_2 < \dots < x_n$ vorausgesetzt.

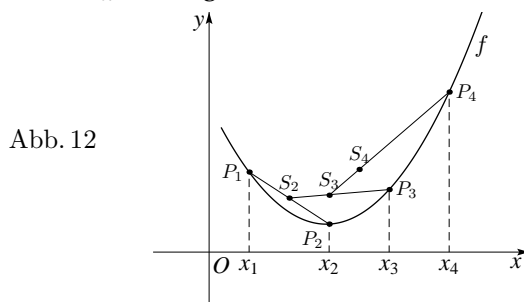


Abb. 12

Für $n = 1$ ist nichts zu zeigen. Im Fall $n = 2$ ist $S_2 = (\bar{x}_{(2)}, \bar{y}_{(2)}) = (\frac{x_1 + x_2}{2}, \frac{y_1 + y_2}{2})$ der Mittelpunkt der Sehne $\overline{P_1 P_2}$, zugleich ihr Eckenschwerpunkt. Nun ist f konvex (im Sinne der Definition KNV aus 2.1), d. h. $E_f(I)$ ist eine konvexe Figur, die $\overline{P_1 P_2}$ und damit auch S_2 enthält.

Der Induktionsschritt vollzieht sich nach einer einfachen Idee: Bestimme den Schwerpunkt von $\{P_1, \dots, P_n, P_{n+1}\}$ als Schwerpunkt von $\{S_n, P_{n+1}\}$! Es mag hier genügen, dies am Beispiel des Übergangs von $n = 2$ nach $n = 3$ auszuführen. Sei S der Schwerpunkt von

$\{P_1, P_2, P_3\}$; er ist auch Schwerpunkt von $\{S_2, P_3\}$, und es gilt wegen der Konvexität von $f : S \in \overline{S_2 P_3} \subseteq E_f(I)$. In die Berechnung der Koordinaten \bar{x}, \bar{y} von S geht S_2 mit 2-fachen Gewicht ein (Gewichte von P_1 und P_2); mithin erhalten wir

$$\bar{x} = \frac{2\bar{x}_{(2)} + x_3}{2+1} = \frac{2}{3}\bar{x}_{(2)} + \frac{1}{3}x_3 = \frac{x_1 + x_2 + x_3}{3} = \bar{x}_{(3)},$$

und entsprechend $\bar{y} = \bar{y}_{(3)}$. Es ist also $S_3 = S \in E_f(I)$, was zu zeigen war.

Der Spezialfall $n = 3$ erweist sich bei näherem Hinsehen als klassisches Lehrstück der Dreiecksgeometrie: S_3 ist Schnittpunkt der Seitenhalbierenden im Dreieck $P_1 P_2 P_3$ und teilt diese bekanntlich im Verhältnis 2:1.

Verwandte Beweisüberlegungen zur JU, die ebenfalls vom Schwerpunkt der Kurven – bzw. Intervallpunkte Gebrauch machen, finden sich in [Quaisser/Sprengel 1986, 75] und [Kirsch 1993, 136].

3.4 Der Gleichheitsfall

Ist f eine im Intervall I konvexe Funktion, so liegt es nahe zu fragen, für welche Variableneinsetzungen in JU(n) das Gleichheitszeichen (anstelle von \leq) steht.

Es genügt, den Sachverhalt für JU(2) zu erörtern:

$$f(p_1 x_1 + p_2 x_2) = p_1 f(x_1) + p_2 f(x_2), \quad (27)$$

wobei $p_1, p_2 \geq 0$, $p_1 + p_2 = 1$ und $x_1, x_2 \in I$.

Offensichtlich ist (27) sogar für alle $x_1, x_2 \in I$ erfüllt, wenn f eine affine Funktion ist, d. h. von der Form

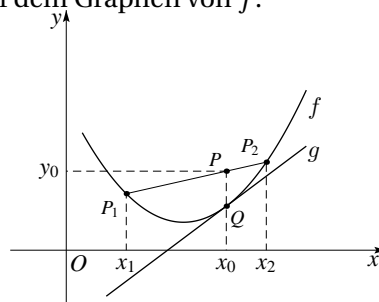
$$f(x) = mx + c \quad (x \in I). \quad (28)$$

Man rechnet dies unmittelbar nach. Der Funktionsgraph von f ist dann ein Geradenstück über dem Intervall I .

Auch die Umkehrung ist richtig: *Eine konvexe Funktion f , die (27) für alle $x_1, x_2 \in I$ erfüllt, ist affin (im Sinne von (28)).*

Zum Beweis genügt das bisher erarbeitete Rüstzeug: Wir nehmen indirekt an, f wäre nicht affin. Dann gibt es Punkte $P_1 = (x_1, f(x_1))$ und $P_2 = (x_2, f(x_2))$ auf dem Graphen von f , $x_1 < x_2$, $x_1, x_2 \in I$, zwischen denen f nicht linear verläuft; d. h. es gibt einen Punkt $P = (x_0, y_0) \in \overline{P_1 P_2}$ mit $x_1 < x_0 < x_2$ sowie $f(x_0) < y_0$ (nach dem BSS). In Abb. 13 ist $Q = (x_0, f(x_0))$ der unterhalb von P liegende Punkt auf dem Graphen von f .

Abb. 13



Nach dem SGS lässt sich eine Stützgerade g durch den Punkt Q legen, d. h. für ein geeignetes $m \in \mathbb{R}$ hat man

$$g(x) = f(x_0) + m(x - x_0) \leq f(x) \quad (x \in I). \quad (29)$$

Die Gerade kann nicht durch beide Punkte P_1 und P_2 gehen, weil dann auch P auf ihr läge (im Widerspruch zu $f(x_0) < y_0$). Wir können also etwa annehmen, dass P_1 nicht auf der Stützgeraden liegt, mithin $g(x_1) < f(x_1)$ gilt. Für den Zwischenpunkt x_0 bestimmen wir eine Darstellung $x_0 = p_1 x_1 + p_2 x_2$ mit $p_1, p_2 > 0$, $p_1 + p_2 = 1$, und schließen weiter:

$$\begin{aligned} g(x_0) &= g(p_1 x_1 + p_2 x_2) \\ &= p_1 g(x_1) + p_2 g(x_2) \quad (g \text{ ist affin}) \\ &< p_1 f(x_1) + p_2 f(x_2) \\ &= f(p_1 x_1 + p_2 x_2) \quad ((27), \text{ nach Voraussetzung}) \\ &= f(x_0). \end{aligned}$$

Dies widerspricht der Gleichheit $g(x_0) = f(x_0)$, die aus (29) für $x = x_0$ resultiert.

Bemerkung: Die Schwerpunktform von (27) (für $p_1 = p_2 = \frac{1}{2}$) impliziert (28) jedenfalls für in ganz I stetiges f ¹⁵. Ohne vorausgesetzte Stetigkeit erfordert das Problem subtilere Mittel; vgl. etwa [Aczel 1966].

Enthält der Graph einer in I konvexen Funktion f über keinem Teilintervall von I ein Geradenstück, so heißt f *streng-konvex in I* . Für keine $x_1, x_2 \in I$, $x_1 \neq x_2$, $p_1, p_2 \in (0, 1)$ kann dann der Gleichheitsfall eintreten, d. h. es gilt für alle $x_1, x_2 \in I$, $x_1 \neq x_2$, $p_1, p_2 \in (0, 1)$:

$$f(p_1 x_1 + p_2 x_2) < p_1 f(x_1) + p_2 f(x_2).$$

Im Umkehrschluss heißt das: Gilt (27) für streng-konvexes f , so ist notwendig $x_1 = x_2$. Eine analoge Aussage lässt sich auch für die allgemeine JU(n) gewinnen.

Es sei hier nur angemerkt, dass diese notwendige Bedingung für den Gleichheitsfall streng-konvexer Funktionen ein natürliches Hilfsmittel zur Lösung gewisser Extremwertaufgaben darstellt. Dazu ist es erforderlich, die Aufgabe durch eine JU zu beschreiben.

Das folgende sehr einfache Beispiel mag die prinzipielle Vorgehensweise illustrieren. Abb. 14 zeigt ein Quadrat der Seitenlänge a . Die beiden schraffierten Quadrate haben die Seitenlängen x_1 und x_2 , wobei $x_1 + x_2 = a$.

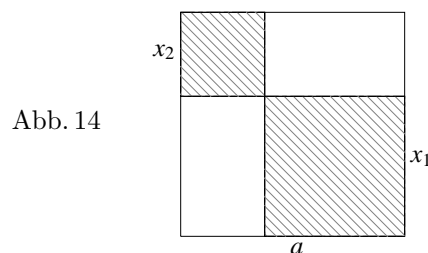


Abb. 14

Wann ist die schraffierte Fläche am kleinsten? – Den zu minimierenden Flächeninhalt $x_1^2 + x_2^2$ findet man leicht in der Ungleichung

$$\frac{a^2}{4} = \left(\frac{x_1 + x_2}{2} \right)^2 \leq \frac{x_1^2 + x_2^2}{2}$$

wieder. Sie ist ein Sonderfall der JU(2) für $f(x) = x^2$ (f streng-konvex!) mit $p_1 = p_2 = \frac{1}{2}$. Das Minimum auf der linken Seite wird nur im Gleichheitsfall angenommen, d. h. für $x_1 = x_2 = \frac{a}{2}$. Eine Vielzahl elementarer Beispiele dieses Typs werden in [Niven 1981] behandelt.

Anmerkungen

¹ Die Ungleichung (1) findet sich bereits in der 17 Jahre zuvor erschienenen Arbeit [Hölder 1889].

² Vgl. dazu etwa [Bullen 2003], [Quaisser/Sprengel 1986] sowie die Beiträge von A. Popa-Fischer in diesem MU-Heft.

³ Die geradlinige Verbindbarkeit (Blickverbindungen, Verkehrsverbindungen, etc.) wird bevorzugt in Unterrichtsbeispielen thematisiert [Müller/Wittmann 1977, 92] wohingegen in operativ-technischen Sachsituationen häufig andere, weniger elementare Aspekte dominieren, z. B. Stabilität [Bender/Schreiber 1985, 140].

⁴ Gelegentlich auch schwächer durch JU(2) mit $p_1 = p_2 = \frac{1}{2}$, so in Jensens Originalarbeit. Für stetige Funktionen sind beide Definitionen gleichwertig. Dabei ist eine konvexe Funktion jedenfalls im Innern ihres Definitionsintervalls stetig.

⁵ Die Definition lässt sich ohne weiteres in allgemeineren Räumen nachbilden; vgl. etwa [Valentine 1964, 137].

⁶ Nachträglich könnte man JU(2) als notwendiges und hinreichendes Kriterium für KNV auffassen. In der Lehrbuchliteratur wird dieser Schritt meistens übersprungen und direkt JU(2) als Definiens der Funktionskonvexität gewählt. Praktisch spielt diese Frage allerdings keine große Rolle. – Alternativ ließe sich auch, Differenzierbarkeit vorausgesetzt, eine Funktion f konvex nennen, wenn f' monoton wächst. Natürlich ist man damit noch weiter von einer elementargeometrischen Definition entfernt, woran auch eine terminologische Bemäntelung (konvex als „linksgekrümmt“ in einem Intervall, etwa bei [Treiber 1985]) nichts ändert. Insbesondere überrascht es keineswegs, dass Treibers Definition 4 (S. 20), die wir später als Aussage BSS (Bogen-Sehnen-Satz) beweisen werden, „bei Schülern gelegentlich Erstaunen“ auslöst.

⁷ Bei *streng-konvexen* Funktionen liegt P stets unterhalb von P' .

⁸ Man kommt zum selben Ergebnis, indem man die Gleichung der Geraden durch die Punkte $(0, x_1)$ und $(1, x_2)$ mit p als unabhängiger und x als abhängiger Variablen aufstellt: $x = (x_2 - x_1)p + x_1$.

⁹ Natürlich stellt die Lipschitz-Bedingung (16) eine verschärfte Form der Stetigkeit dar. Als konvexe Funktion ist f somit in jedem abgeschlossenen Intervall $[x_1, x_2]$ im Innern von I bereits Lipschitz-stetig, damit auch absolut-stetig auf $[x_1, x_2]$; vgl. das Theorem A in [Roberts/Varberg 1973, 4]. Man vergegenwärtige sich hierzu noch einmal die Anm. 4.

¹⁰ Die Existenz von Stützgeraden wird in [Boltjanski/Jaglom 1971, 178] auf sorgfältige Weise anschaulich plausibilisiert. Auch die Umkehrung gilt: Eine Figur ist konvex, wenn sie in jedem Randpunkt eine Stützgerade besitzt (auf S. 181 ohne Beweis mitgeteilt).

¹¹ Ist x_0 ein Randpunkt von I , für den $y_0 = f(x_0)$ existiert, so ist $x = x_0$ eine Stützgerade an $E_f(I)$, allerdings nicht an den Funktionsgraphen im zuvor definierten Sinn. Eine Stützgerade in x_0 der Form (17) ergibt sich aber leicht, indem man nur für einen der beiden Nachbarpunkte L bzw. R die hier folgende Grenzwertüberlegung durchführt.

¹² In Abschnitt 2 ist es einzig diese Argumentation, die wesentlichen Gebrauch von einem Lehrsatz der elementaren Analysis macht. Eine rein geometrische (oder algebraische) Überlegung ist an dieser Stelle nicht möglich.

¹³ Für nähere Einzelheiten verweisen wir auf [Roberts/Varberg 1973].

¹⁴ Es lässt sich die (hier nicht benötigte) etwas stärkere Aussage zeigen: Das aus P_1, P_2, \dots, P_n gebildete Sehnen-vieleck (genauer: die konvexe Hülle der Punkte P_1, P_2, \dots, P_n) liegt ganz in $E_f(I)$. Ist dieses nicht zur Strecke ausgeartet, so ist sein Schwerpunkt ein innerer Punkt von $E_f(I)$.

¹⁵ Aus (27) folgt nämlich mit $x_1 = x$, $x_2 = 0$ (Intervall anpassen) zunächst $f(x) = 2f\left(\frac{x}{2}\right) - f(0)$ und weiter induktiv: $f(x) = 2^n f\left(\frac{x}{2^n}\right) - (2^n - 1)f(0)$. Damit zeigt man, dass für alle $x = \pm 2^n$ ($n \in \mathbb{N}$) $f(x)$ die Form $mx + c$ hat. Die Punkte zwischen den Zweierpotenzen lassen sich durch sukzessives Halbieren einschachteln und $f(x) = mx + c$ ($x \in I$) durch Stetigkeitsargumente gewinnen.

Literatur

- Aczel, J.: *Lectures on Functional Equations and their Applications*. Academic Press: New York 1966.
- Bender, P.; Schreiber, A.: *Operative Genese der Geometrie*. Hölder-Pichler-Tempsky: Wien und B. G. Teubner: Stuttgart 1985.
- Boltjanski, W. G.; Jaglom, I. M.: Konvexe Figuren und Körper. *Enzyklopädie der Elementarmathematik*, Bd. V: Geometrie, 173–257. VEB Deutscher Verlag der Wissenschaften: Berlin 1971.
- Bullen, P. S.: *Handbook of Means and Their Inequalities*. Kluwer Academic Publishers: Dordrecht, Boston, London 2003.
- Fichtenholz, G. M.: *Differential- und Integralrechnung I*. VEB Deutscher Verlag der Wissenschaften: Berlin 1975.
- Hölder, O.: Über einen Mittelwerthsatz. *Nachr. v. d. k. Gesellsch. d. Wiss. zu Göttingen*: 1889, 38–47.
- Jensen, J. L. W. V.: Sur les fonctions convexes et les inégalités entre les valeurs moyennes. *Acta Math.* **30** (1906), 175–193.
- Kirsch, A.: Zur Jensenschen Ungleichung. Ein „erklärender“ statt nur „beweisender“ Beweis. *Praxis der Mathematik* **35/3** (1993), 136–138.
- Müller, G.; Wittman, E.: *Der Mathematikunterricht in der Primarstufe*. Vieweg-Verlag: Braunschweig 1977.
- Niven, I.: *Maxima and Minima without Calculus*. The Dolciani Mathematical Expositions, Nr. 6. The Mathematical Association of America, 1981.
- Quaisser, E.; Sprengel, H.-J.: *Extrema*. VEB Deutscher Verlag der Wissenschaften: Berlin 1986.
- Roberts, A. W.; Varberg, D. E.: *Convex Functions*. Academic Press: New York and London 1973.
- Rüthing, D.: Zu konvexen Funktionen und zur Ungleichung von Jensen. *Praxis der Mathematik* **37/6** (1995), 247–250.
- Treiber, D.: Zur Links- und Rechtskrümmung von Funktionsgraphen. *MNU* **38/1** (1985), 18–20.
- Treiber, D.: Zur Jensenschen Ungleichung. *Praxis der Mathematik* **33/2** (1991), 71–73.
- Valentine, F. A.: *Konvexe Mengen*. Bibliographisches Institut: Mannheim 1964.