1996

Convex functions

Susanna Maria Zagar

CONVEX FUNCTIONS

A Project

Presented to the

Faculty of

California State University,

San Bernardino

In Partial Fulfillment

of the Requirements for the Degree

Master of Arts

in

Mathematics

by

Susanna Maria Zagar

March 1996

CONVEX FUNCTIONS

A Project

Presented to the

Faculty of

California State University,

San Bernardino

by

Susanna Maria Zagar

March 1996

Approved by:

[redacted]

Dr. Hajrudin Fejzic, Chair, Mathematics

03-14-96

Date

[redacted]

Dr. Dan Rinne

[redacted]

Dr. Chetan Parkash

ABSTRACT

This paper will analyze convex functions. In particular, it will investigate criteria for convexity. The investigation will list the criteria from the weakest to the strongest based on theorems, definitions, propositions, and various examples.

The theory of convex functions is based on the theory of real-valued functions of a real variable. The purpose of this thesis is to analyze convex functions.

We begin with the analytic and geometric representation of convex functions. Then we establish some properties of convex functions. Namely, we show that a convex function on an open interval has to be continuous and it has to be differentiable with the possible exception of a countable set. We provide and prove two different versions of Jensen's inequality for convex functions, and we use this inequality to establish the familiar inequality between arithmetic and geometric means.

In the second chapter, we give sufficient criteria for convexity, from the familiar criteria involving Second derivatives to criteria involving Schwartz and Peano derivatives.

The bibliography lists several books which deal with convex functions and related material.

TABLE OF CONTENTS

LIST OF FIGURES

CHAPTER I

1 Definitions and Examples

Definition 1.1 A function f is said to be convex on [a,b] when

$$f(\alpha x+(1-\alpha)x') \le \alpha f(x)+(1-\alpha)f(x')$$

for all x, x' $(a<x<x'<b)$ and $0\le\alpha\le 1$.

Example 1.1 The function $f(x)=x^2$ is convex on (0,1).

To show that f is convex, we have to verify that the following inequality holds.

$$f(\alpha x+(1-\alpha)x') \le \alpha f(x)+(1-\alpha)f(x')$$

Let $0<x<x'<1$ and $0\le\alpha\le 1$. Then

$$\begin{aligned}
&f(\alpha x+(1-\alpha)x') \\
&=(\alpha x+(1-\alpha)x')^2 \\
&=\alpha^2x^2+2\alpha x(1-\alpha)x'+(1-\alpha)^2(x')^2 \\
&=\alpha^2x^2+2\alpha xx'-2\alpha^2xx'+(1-2\alpha+\alpha^2)(x')^2
\end{aligned}$$

$$\begin{aligned}
&\alpha f(x)+(1-\alpha)f(x') \\
&=\alpha x^2+(1-\alpha)(x')^2 \\
&=\alpha x^2+(x')^2-\alpha(x')^2
\end{aligned}$$

Hence, we have to check that

$$\alpha^2x^2+2\alpha xx'-2\alpha^2xx'+(x')^2-2\alpha(x')^2+\alpha^2(x')^2-\alpha x^2-(x')^2+\alpha(x')^2 \le 0$$

$$\vdots$$

$$\alpha x^2(\alpha-1)+2\alpha xx'(1-\alpha)+\alpha(x')^2(\alpha-1) \le 0$$

$$(\alpha-1)\left[\alpha x^2-2\alpha xx'+\alpha(x')^2\right] \le 0$$

$$\alpha(\alpha-1)\left(x^2-2xx'+(x')^2\right) \le 0$$

$$\alpha(\alpha-1)(x-x')^2 \le 0$$

$$-\alpha(1-\alpha)(x-x')^2 \le 0,$$

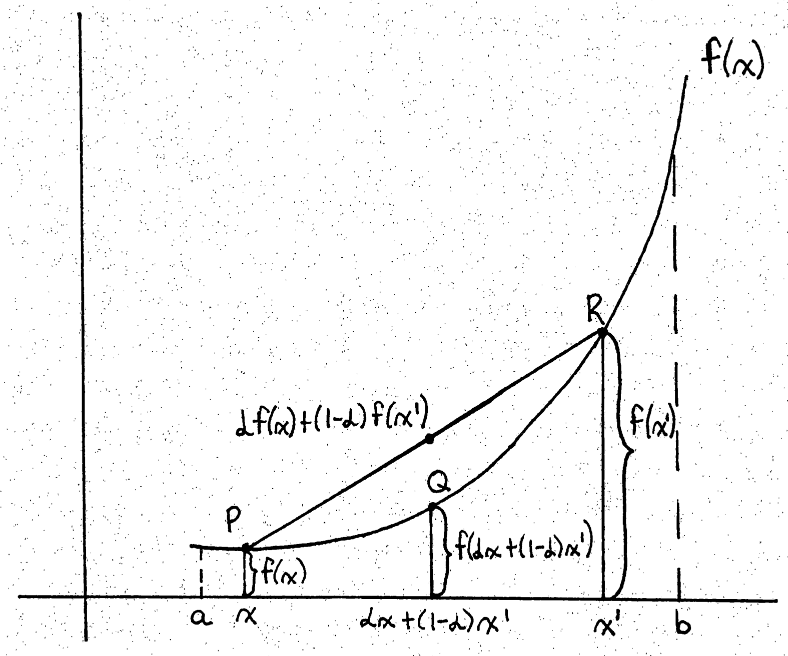
and the last inequality holds since $0\le\alpha\le 1$.

Therefore, $f(\alpha x+(1-\alpha)x') \le \alpha f(x)+(1-\alpha)f(x')$.

Hence, $f(x)=x^2$ is convex.

Fig. 1 displays a geometric interpretation, which will be very useful in establishing results about convex functions. Geometric arguments will be used in proofs wherever possible, so it is useful to keep this figure in mind.

Fig. 1



Notice that as α ranges from 0 to 1, the point $t = \alpha x + (1-\alpha)x'$, ranges from x to x' or from x' to x. Therefore, $Q = (t, f(t))$ is always below or on the line connecting $P = (x, f(x))$, and $R = (x', f(x'))$. Therefore, the slope of the line connecting points R and Q is always greater than or equal to the slope of the line connecting points P and R. Analytically, this says that whenever $x < t < x'$, then

$$\frac{f(x') - f(x)}{x' - x} \le \frac{f(x') - f(t)}{x' - t}. \qquad (1)$$

Also, R is always on or above the line connecting points P and Q. See Figures 2 and 3 below.

Fig. 2

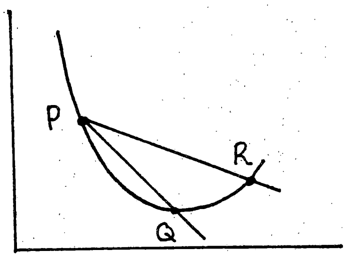
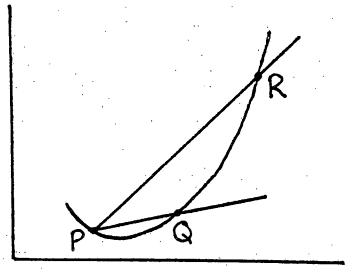


Fig. 3



Therefore, the slope of the line connecting points P and R is always greater than or equal to the slope of line connecting points P and Q. Analytically, this says that whenever $x < t < x'$, then

$$\frac{f(t)-f(x)}{t-x} \le \frac{f(x')-f(x)}{x'-x}. \tag{2}$$

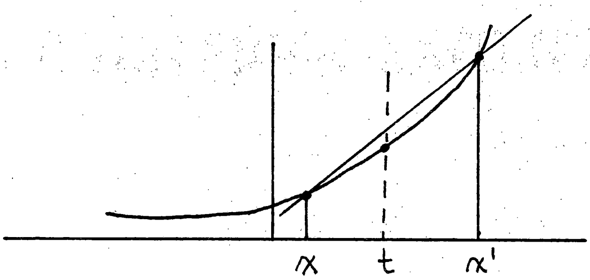
By combining (1) and (2), we obtain the following useful proposition.

Proposition 1.1 Let f be convex on (a,b). Then for $a < x < t < x' < b$, $P=(x,f(x))$, $Q=(t,f(t))$, $R=(x',f(x'))$, we have the following inequality

$$slope\ PQ = \frac{f(t)-f(x)}{t-x} \le slope\ PR = \frac{f(x')-f(x)}{x'-x} \le slope\ RQ = \frac{f(x')-f(t)}{x'-t}.$$

Example 1.2 The function $f(x)=e^x$ is convex on $(-\infty,\infty)$.

Fig. 4



By looking at the graph of $f(x)=e^x$, it is obvious that for any $-\infty<x<t<x'<\infty$, the point $(t,f(t))$ is below the line connecting the points $(x,f(x)),(x',f(x'))$. We will give a proof of this fact later.

Continuity

In this section, we investigate continuity of convex functions. The main theorem is the following result.

Theorem 1.1 If a function, f, is convex on (a,b), then f is continuous on (a,b).

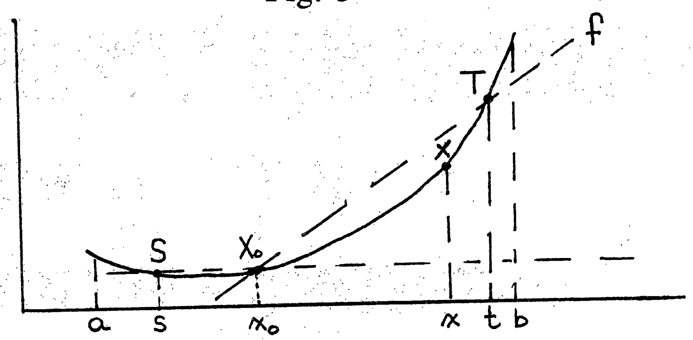
Proof. Let $a<s<x_0<x<t<b$, and

$$S=(s,f(s))$$
$$X_0=(x_0,f(x_0))$$
$$X=(x,f(x))$$
$$T=(t,f(t))$$

We will show that f is continuous at x_0 from the right, for the continuity from the left is established in an analogous manner.

Since f is convex, from Figure 5 below, X_0 is on or below the line SX which implies X is on or above the line SX_0. Also X is on or below the line X_0T.

Fig. 5



Therefore, $(x,f(x))$ is between the lines SX_0, and X_oT.
Analytically,

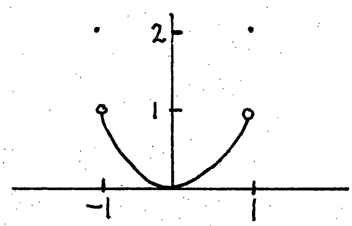
$$\frac{f(s)-f(x_0)}{s-x_0}(x-x_0)+f(x_0)\le f(x)\le\frac{f(t)-f(x_0)}{t-x_0}(x-x_0)+f(x_0) \quad (3)$$

As $x \to x_0^+$, the left-hand side of (3); and the right-hand side of (3) converge to $f(x_0)$. Therefore, by the Squeeze Theorem, $\lim_{x \to x_0^+} f(x) = f(x_0)$.

A convex function on a closed interval $[a,b]$, doesn't necessarily have to be continuous at the endpoints a and b.

For example, the function

Fig. 6



$$f(x) = \begin{cases} x^2 & -1 < x < 1 \\ 2 & x = -1 \\ 2 & x = 1 \end{cases},$$

is convex on $[-1,1]$, but obviously it fails to be continuous at the endpoints.

We will use Theorem 1.1, to establish the well-known Jensen's Inequality.

Theorem 1.2 (Jensen's Inequality) Let ψ be a convex function on an interval (a,b) and f a Riemann integrable function on (a,b) with $a < f(x) < b$.

Then $\psi\left(\frac{1}{b-a}\int_a^b f(x)dx\right) \le \frac{1}{b-a}\int_a^b \psi(f(x))dx.$

The more general inequality involving Lebesgue integration holds and it is proved in [7]. Lebesgue's integration is beyond the scope of this thesis. We will use the ideas from [7] to prove the theorem. In order to prove the theorem, we will briefly review the definition of the supremum.

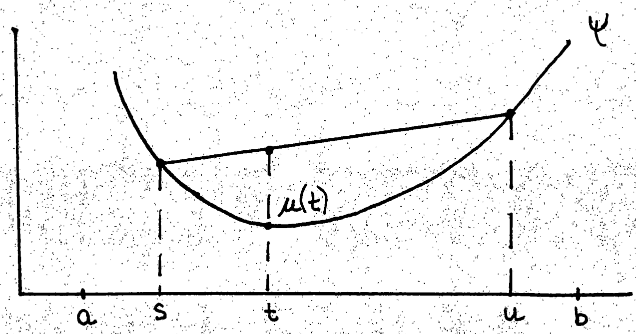
Definition 1.2 Let A be an ordered set and X a subset of A. An element $b \in A$ is called the least upper bound (or supremum) of X if b is an upper bound for X and there is no upper bound b' for X that is less than b. We denote the supremum of X by sup X.

From the definition of a supremum, it is clear that if b = sup X, then for all $a \in X$, $a \le b$. Also, if c is such that for all $a \in X$, $a \le c$, then $b \le c$.

The proof of Jensen's inequality now follows.

Proof. Let $a < s < t < u < b$.

Fig. 7



The equation of the line passing through points $(s, \psi(s))$ *and* $(u, \psi(u))$ is

$$y = \frac{\psi(u) - \psi(s)}{u - s}(x - s) + \psi(s)$$

Since ψ is convex, $\psi(t) \le y(t)$. More specifically,

$$\psi(t) \le \frac{\psi(u) - \psi(s)}{u - s}(t - s) + \psi(s)$$

and from here

$$\frac{\psi(t) - \psi(s)}{t - s} \le \frac{\psi(u) - \psi(s)}{u - s}. \tag{4}$$

By Proposition 1.1, applied to ψ with $u = x', s = x$, we have

$$\frac{\psi(u) - \psi(s)}{u - s} \le \frac{\psi(u) - \psi(t)}{u - t}. \tag{5}$$

Let $t = \frac{1}{b-a}\int_a^b f(v)dv$. Since $a < f(x) < b$, we have $a < t < b$.

Let $\beta = \sup\limits_{a<s<t} \frac{\psi(t) - \psi(s)}{t - s}$ (6)

Then from (4) and (5), β is finite and $\beta \le \frac{\psi(u)-\psi(t)}{u-t}$ for all $b>u>t$.
Hence, for all $b>u>t$, we have

$$\psi(u) \ge \psi(t)+\beta(u-t) \tag{7}$$

Also from (6), for all $a<s<t$,

$$\frac{\psi(t)-\psi(s)}{t-s} \le \beta \text{ and from here}$$

$$\psi(s) \ge \psi(t)+\beta(s-t). \tag{8}$$

Combining (7) and (8), we have that for all $a<s<b, \psi(s) \ge \psi(t)+\beta(s-t)$. In particular, the inequality is true if we replace s by $f(x)$. Then for all $x \in (a,b)$, we have $\psi(f(x))-\psi(t)-\beta(f(x)-t) \ge 0$. Since by Theorem 3.1, ψ is continuous, $\psi(f(x))$ is Riemann integrable, so integrating the inequality above, we obtain

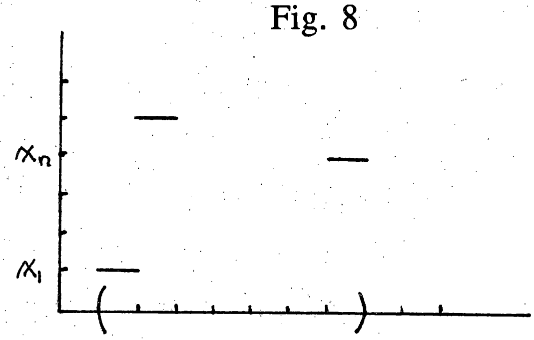
$$\int_a^b \psi(f(x))dx - \int_a^b \psi(t)dx - \beta\int_a^b f(x)dx + \beta\int_a^b t dx \ge \int_a^b 0 dx$$

$$\int_a^b \psi(f(x))dx - (b-a)\psi(t) - (b-a)\beta\cdot t + (b-a)\beta\cdot t \ge 0$$

$$\int_a^b \psi(f(x))dx \ge (b-a)\psi(t) = (b-a)\psi\left(\frac{1}{b-a}\int_a^b f(x)dx\right),$$ which is the desired inequality.

Example 1.3 Let $a<x_1,x_2,\cdots,x_n<b$.
Let $f(x)=\begin{cases} x_i & \text{for } a+\frac{(i-1)(b-a)}{n} \le x \le a+\frac{i(b-a)}{n}, \quad 1\le i\le n. \end{cases}$
and $\psi(x)=e^x$.

Fig. 8



The function ψ is convex and f is Riemann integrable. By Jensen's Inequality,

$$e^{\frac{1}{b-a}\int_a^b f(x)dx} \le \frac{1}{b-a}\int_a^b e^{f(x)}dx. \tag{9}$$

Since $\int_a^b f(x)dx = \left(\frac{b-a}{n}\right)\sum_{i=1}^{n} x_i$, and $\int_a^b e^{f(x)}dx = \frac{b-a}{n}\sum_{i=1}^{n} e^{x_i}$, the left-hand side of (9) is equal to $e^{\frac{1}{n}(x_1+\cdots+x_n)} = \left(e^{x_1}\cdot e^{x_2}\cdot\ldots e^{x_n}\right)^{\frac{1}{n}}$, while the right-hand side of (9) is $\frac{1}{n}\left(e^{x_1}+\ldots+e^{x_n}\right)$. Thus $\left(e^{x_1}\cdot e^{x_2}\cdot\ldots e^{x_n}\right)^{\frac{1}{n}} \le \frac{1}{n}\left(e^{x_1}+\ldots+e^{x_n}\right)$. Now let $y_i = e^{x_i}$. Then $y_i \in (0,+\infty)$ and by substitution, we get $(y_1y_2\ldots y_n)^{1/n} \le \frac{1}{n}(y_1+y_2+\ldots+y_n)$ which is a familiar inequality between the arithmetic and geometric means.

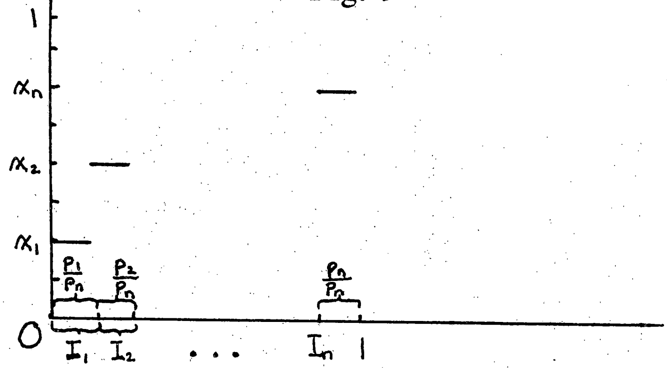
A weaker form of Jensen's inequality, Theorem 1.3, is described below. Here we give a different proof of Theorem 1.3 by using Theorem 1.1.

Theorem 1.3 (Jensen's Inequality) Let ψ be a convex function on $(0,1)$, $0 \le x_i \le 1$, $0 < p_i$, *and for* $i = 1,2,\ldots,n$, define $P_n = \sum_{i=1}^{n} p_i$. Then

$$\psi\left(\frac{1}{P_n}\sum_{i=1}^{n} p_i x_i\right) \le \frac{1}{P_n}\sum_{i=1}^{n} p_i \psi(x_i).$$

Proof. Let ψ be a convex function on $(0,1)$, $0 \le x_i \le 1$ and $0 < p_i$ for $i = 1,2,\ldots,n$. Define $f(x) = \{x_i \;\; on\, I_i$ where I_i is an interval of the length $\frac{p_i}{P_n}$, see Figure 9.

Fig. 9



Clearly, $f(x)$ is between 0 and 1 and it is Riemann integrable. Therefore by Theorem 1.2, $\psi\left(\int_0^1 f(x)dx\right) \le \int_0^1 \psi(f(x))dx$. Finally since $\int_0^1 f(x)dx = \sum_{i=1}^n x_i \cdot \frac{p_i}{p_n}$ and $\int_0^1 \psi(f(x))dx = \sum_{i=1}^n \psi(x_i)\cdot\frac{p_i}{p_n}$, we obtain $\psi\left(\frac{1}{P_n}\sum_{i=1}^n p_i x_i\right) \le \frac{1}{P_n}\sum_{i=1}^n p_i \psi(x_i)$.

Differentiability

Now we examine differentiability properties of convex functions. Our goal is to show that convex functions are differentiable with the possible exception of a countable number of points. We will need some background on one-sided derivatives and monotonic functions.

Definition 1.3 The right-hand derivative $f_+'(x)$ and the left-hand derivative $f_-'(x)$ are defined to be as follows:

$$f_+'(x) = \lim_{y \downarrow x} \frac{f(y)-f(x)}{y-x} \tag{10}$$

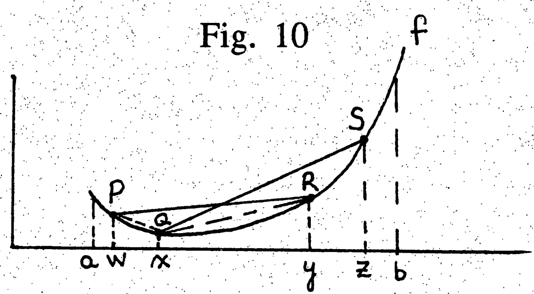
and

$$f_-'(x) = \lim_{y \uparrow x} \frac{f(y)-f(x)}{y-x} \tag{11}$$

Theorem 1.4 Let f be convex on (a,b). Then
a) f_+' and f_-' exist on (a,b),
b) f_+' and f_-' are monotone increasing, and
c) for w in (a,b), $\lim_{x\uparrow w} f_+'(x) \le f_-'(w) \le f_+'(w) \le \lim_{x\downarrow w} f_+'(x)$.

Proof. Let $y \in (a,b)$. Consider the points $w < x < y < z < t$ in (a,b) with P, Q, R, and S, the corresponding points on the graph of the function f. See Figure 10.

Fig. 10



Since f is convex, by Proposition 1.1,
$slope\,PQ \le slope\,PR \le slope\,QR \le slope\,QS \le slope\,RS$.
Since $w<x<y$, implies $slope\,PR \le slope\,QR$, it follows that slope QR increases as $x \uparrow y$. Similarly, slope RS decreases as $z \downarrow y$. Therefore, the left-hand side of the inequality
$slope\,QR = \dfrac{f(x)-f(y)}{x-y} \le \dfrac{f(z)-f(y)}{z-y} = slope\,RS$ increases as $x \uparrow y$, and the right-hand side decreases as $z \downarrow y$. Therefore, $f'_-(y) = \lim_{x\uparrow y} \dfrac{f(x)-f(y)}{x-y}$ and $f'_+(y) = \lim_{z\downarrow y} \dfrac{f(z)-f(y)}{z-y}$ exist. Since $y \in (a,b)$ was arbitrary, part (a) of the theorem is proved. Moreover, the monotonicity of the slopes PR and QR implies that whenever $x<y<z$,

$$\frac{f(x)-f(y)}{x-y} \le f'_-(y) \le f'_+(y) \le \frac{f(z)-f(y)}{z-y}. \tag{12}$$

To prove (b), let $w<y$ be two points in (a,b). If $w<x<y$, then from (12) we have $f'_-(w) \le f'_+(w) \le \dfrac{f(x)-f(w)}{x-w}$, and $\dfrac{f(x)-f(y)}{x-y} \le f'_-(y) \le f'_+(y)$. On the other hand, from Proposition 2.1, we have $\dfrac{f(x)-f(w)}{x-w} \le \dfrac{f(x)-f(y)}{x-y}$,
and hence $f'_-(w) \le f'_+(w) \le f'_-(y) \le f'_+(y)$, (13)
establishing the monotonic nature of f'_- and f'_+.
Finally to prove (c), let $w \in (a,b)$. For $y>w$ from (13) we have
$f'_+(w) \le f'_+(y)$. (14)
Since f'_+ is monotone increasing, letting $y \downarrow w$ in (14), we get $f'_+(w) \le \lim_{y\downarrow w} f'_+(y)$.
On the other hand, for $x<w$, from (13) we have $f'_+(x) \le f'_-(w)$. (15)
Letting $x \uparrow w$ in (15), we get $\lim_{x\uparrow w} f'_+(x) \le f'_-(w)$.

In order to prove that the set E where the derivative of a convex function fails to exist is countable, we need Theorem 1.2 below regarding continuity of monotonic functions.

Definition 1.4 If g is an increasing function on (a,b), then $g_-(x)$ is defined to be $\lim_{h\to 0^-} g(x+h)$. Similarly, $g_+(x)$ is defined to be $\lim_{h\to 0^+} g(x+h)$. Since g is an increasing function, $g_-(x)$ and $g_+(x)$ exist, and $g_-(x) \le g(x) \le g_+(x)$.

Lemma 1.1 Let an increasing function, g, be defined on (a,b). Let $x_1, x_2, \cdots, x_n$ be arbitrary points lying in (a,b). Then

$$[g_+(a) - g(a)] + \left(\sum_{k=1}^{n} [g_+(x_k) - g_-(x_k)]\right) + [g(b) - g_-(b)] \le g(b) - g(a) \qquad (16)$$

Proof. We may assume that $a < x_1 < x_2 < \cdots < x_n < b$. Let $x_0 = a$, $x_{n+1} = b$. Choose points $y_0, y_1, \cdots, y_n$ such that $x_k < y_k < x_{k+1}$ $(k = 0,1,\ldots,n)$ Since g is an increasing function, $g_+(x_k) \le g(y_k)$ and $g_-(x_k) \ge g(y_{k-1})$. Combining these inequalities, we obtain

$$g_+(x_k) - g_-(x_k) \le g(y_k) - g(y_{k-1}) \quad (k = 1,2,\cdots,n)$$
$$g_+(a) - g(a) \le g(y_0) - g(a) \text{ and}$$
$$g(b) - g_-(b) \le g(b) - g(y_n)$$

Adding the left and right side of the inequalities yields

$$\sum_{k=1}^{n} \left(g_+(x_k) - g_-(x_k)\right) + g_+(a) - g(a) + g(b) - g_-(b) \le \sum_{k=1}^{n} \left(g(y_k) - g(y_{k-1})\right) + g(y_0) - g(a) + g(b) - g(y_n)$$

By simplifying the right-hand side of this inequality, we obtain (16).

Corollary 1.1 An increasing function, g, defined on (a,b) can have only a finite number of points of discontinuity at which $g_+(x) - g_-(x)$ is greater than a given positive number σ.

Proof. If the points $x_1, x_2, \ldots, x_n \in (a,b)$ are points of discontinuity with $g_+(x_i) - g_-(x_i)$ greater than σ, then from (15),
$n\sigma \le \sum_{k=1}^{n} [g_+(x_k) - g_-(x_k)] \le g(b) - g(a)$, and hence $n \le \frac{g(b) - g(a)}{\sigma}$. Therefore, there can be only finitely many $x_1, x_2, \ldots, x_n \in (a,b)$ for which $g_+(x_i) - g_-(x_i)$ is greater than σ.

Theorem 1.5 The set of points of discontinuity of an increasing function, g, defined on (a,b) is at most countable.

Proof. Let H be the set of all points of discontinuity of the function g. Let H_k be the set of those points of discontinuity of this function, at which $g(x)-g_-(x)$ *or* $g_+(x)-g(x)$ is greater than $\frac{1}{k}$. Clearly, if $x \in H$, then there is an integer k such that $g(x)-g_-(x) \geq \frac{1}{k}$ or $g_+(x)-g(x) \geq \frac{1}{k}$. Thus, $x \in H_k$, and hence, $H = \bigcup_{k=1}^{\infty} H_k$. If $x \in H_k$, then $g_+(x)-g_-(x) = g_+(x)-g(x)+g(x)-g_-(x) \geq \frac{1}{k}$, therefore by Corollary 1.1, each H_k is finite. Hence H is at most countable.

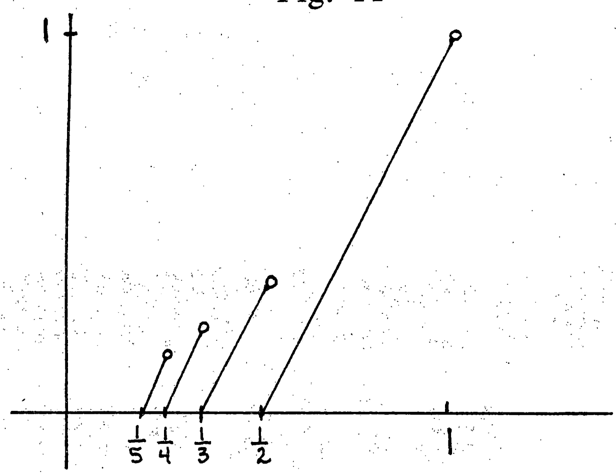
Now we are ready to prove our main result of this section.

Theorem 1.6 If f is convex on (a,b), then the set where f' fails to exist is countable.

Proof. By Theorem 1.4, f'_+ is an increasing function, and for $w \in (a,b)$, $\lim_{x \uparrow w} f'_+(x) \leq f'_-(w) \leq f'_+(w) \leq \lim_{x \downarrow w} f'_+(x)$. Therefore, if w is a point of continuity of f'_+, then $f'_-(w) = f'_+(w)$ i.e. f is differentiable at w. By Theorem 1.5, the set of points where f'_+ is not continuous is at most countable. Hence, the set of points where f is not differentiable is at most countable. The converse is false, as the following example shows.

$$f(x) = \left\{ x - \frac{1}{h} \; for \; \frac{1}{h} < x \leq \frac{1}{h-2}, \quad h = 2,3,\ldots \right.$$

Fig. 11



Then $f'(x)$ exists on (0,1) except at the sequence $\left\{\frac{1}{h}\right\}_{h=1}^{\infty}$. Also it is clear that $f'(x)=1$ where it exists, but $f(x)$ is not convex.

CHAPTER II

In this chapter, we give criteria for convexity. We divide these criteria into two groups, and we present them from the weakest to the strongest. We begin by the first set of criteria for convexity which is based on the differentiability of a function.

2 First Set of Criteria for Convexity

The following is a useful criterion to check for convexity.

Proposition 2.1 Suppose $f''(x) \geq 0$. Then f(x) is convex.

Proof. Suppose f is not convex. Then there is $a<x<b$ such that $(x,f(x))$ is above the line connecting the points $(a,f(a))$ and $(b,f(b))$. See Figures 12 and 13 below.

Fig. 12

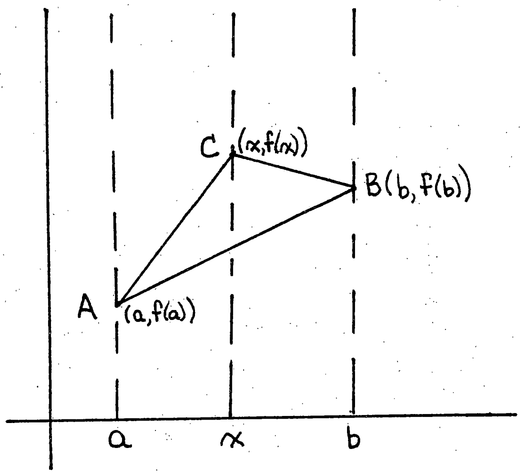
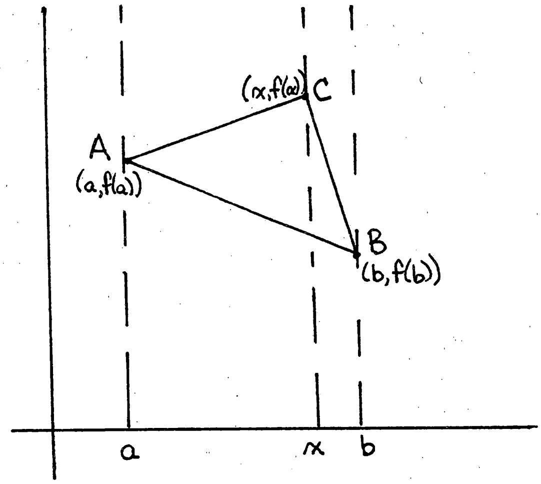


Fig. 13



For both cases, f(a)≤ f(b) and f(a)≥ f(b), it is obvious that the slope of AC > the slope of AB, and also that the slope of BC < the slope of AB. From here, we find that the slope of AC > the slope of BC. Since f is differentiable, by the Mean Value Theorem, there exists a number c in (a,x) such that

$$f'(c)=\frac{f(x)-f(a)}{x-a} = \text{slope of AC.}$$

And also, there exists a number d in (x,b) such that

$$f'(d) = \frac{f(b)-f(x)}{b-x} = \text{slope of BC.}$$

It follows that $f'(c) > f'(d)$.
Since c < d, f' decreases.
But $f''(x) \geq 0$ implies f' is non-decreasing, a contradiction.
Hence, f must be convex.

We use this proposition to provide several examples of convex functions.

Example 2.1 $f(x) = e^x$

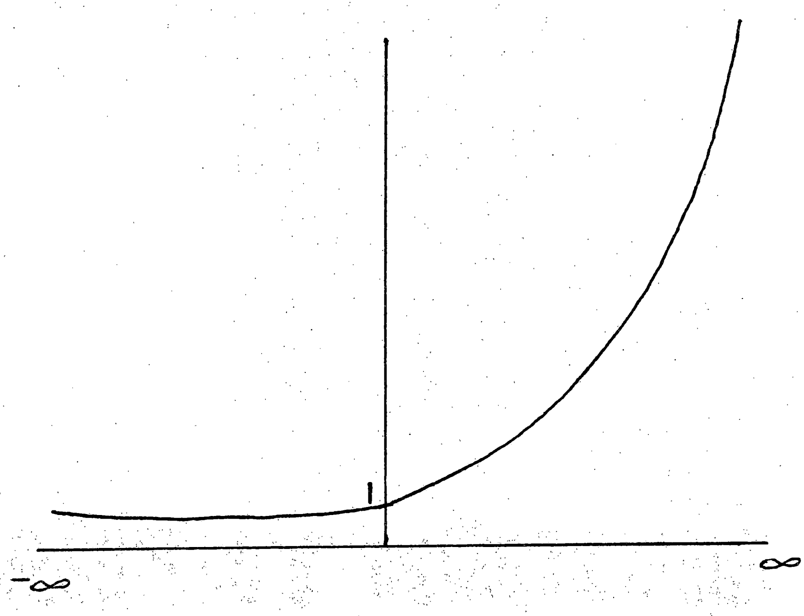
$$f'(x) = e^x$$

$$f''(x) = e^x$$

Since $e^x > 0$, the function, $f(x) = e^x$, is convex by Proposition 2.1.

Fig. 14

Graph of $f''(x)$



Example 2.2 $f(y)=\dfrac{-1}{1+e^{-y}}$

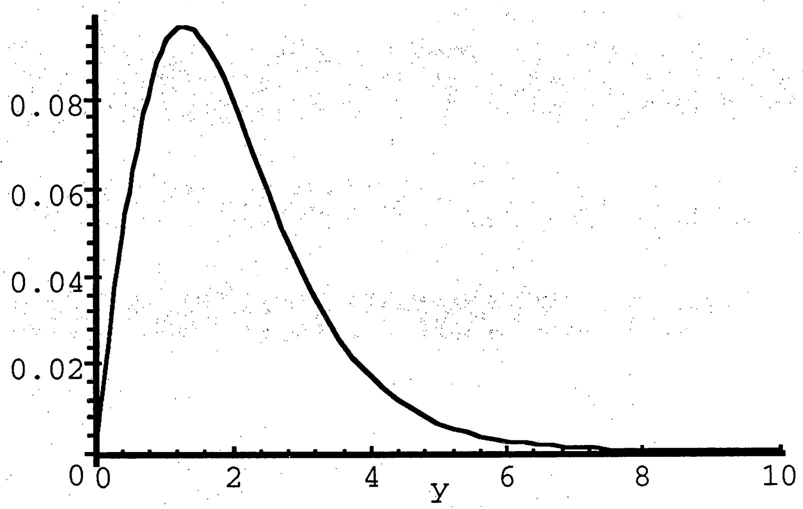
$$f'(y)=\frac{-e^{-y}}{\left(1+e^{-y}\right)^{2}}$$

and
$$f''(y)=\frac{-e^{-2y}+e^{-y}}{\left(1+e^{-y}\right)^{3}}$$

From the graph of $f''(y)$ below, we see that $f''(y)>0$ for $0<y<\infty$. Therefore, by the proposition, f is convex on $(0,\infty)$.

Fig. 15

Graph of $f''(y)$



Previously it had been shown that, for a differentiable function, convexity implies an increasing derivative which leads into Proposition 2.2 which shows the converse to be true, i.e. an increasing derivative implies that f(x) is convex.

Proposition 2.2 Suppose f'(x) is non-decreasing. Then f(x) is convex.

Proof. Let f'(x) be increasing. Suppose that f is not convex. Then there is $x_1 < x < x_2$ such that $(x, f(x))$ is above the line connecting $(x_1, f(x_1)), (x_2, f(x_2))$. As in the proof of Proposition 2.1, we would obtain two points $c < d$ such that $f'(c) > f'(d)$, which is a contradiction since $f'(x)$ is increasing. Therefore, f is convex.

The next example shows that Proposition 2.2 is stronger than Proposition 2.1.

Example 2.3 $f(x) = \begin{cases} x^2 & for\ x \geq 0 \\ 0 & for\ x < 0 \end{cases}$

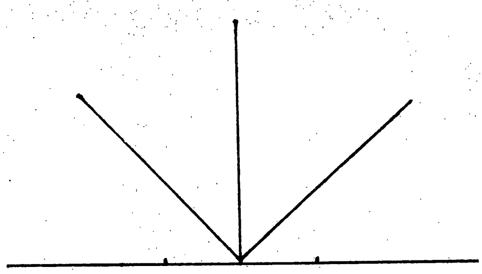
Then $f'(x) = \begin{cases} 2x & for\ x \geq 0 \\ 0 & for\ x < 0 \end{cases}.$

$f'(x)$ is non-decreasing, therefore $f(x)$ is convex, but obviously $f''(0)$ doesn't exist.

A function can still be convex without $f'(x)$ existing at a point.

Example 2.4 Let $f(x) = |x|$.

Fig. 16



Clearly the function is convex, but it is also clear that $f'(0)$ doesn't exist. Moreover, there is a convex function for which f' doesn't exist at countable many points.

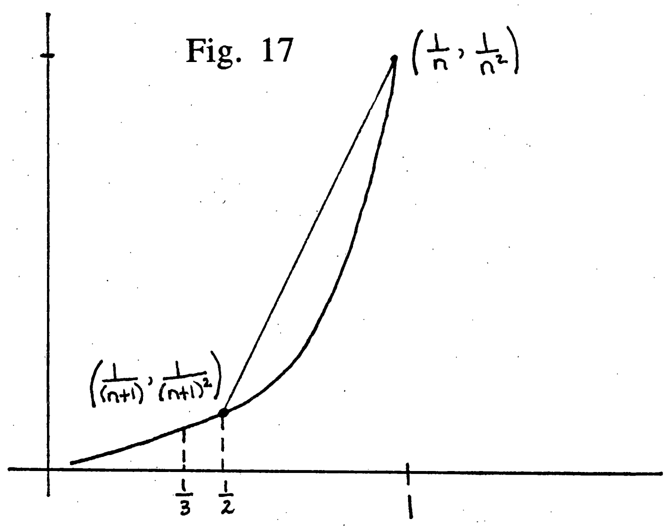
From Theorem 1.4(b), we see that for convex functions where $f'(x)$ exists, it has to be increasing.

Example 2.5 Consider

$$f(x)=\begin{cases}\dfrac{(2n+1)x}{n(n+1)}-\dfrac{1}{n(n+1)} & \textit{for}\quad \dfrac{1}{n+1}<x<\dfrac{1}{n}\\[2ex] 0 & \textit{for}\quad x\le 0\end{cases}$$

which are segments with endpoints $\left(\dfrac{1}{n+1},\dfrac{1}{(n+1)^2}\right)$ and $\left(\dfrac{1}{n},\dfrac{1}{n^2}\right)$ for n=1, 2, 3, ...

Fig. 17



Therefore,

$$f'(x)=\begin{cases}\dfrac{2n+1}{n(n+1)} & \textit{for}\quad \dfrac{1}{n+1}<x<\dfrac{1}{n}\\[2ex] \textit{doesn't exist} & \textit{for}\quad x=\dfrac{1}{n}\end{cases}$$

Now it will be shown that f'(x) is increasing.

In order to show that f'(x) is increasing, it will be shown below that

$$\frac{2m+1}{m(m+1)} > \frac{2n+1}{n(n+1)} \qquad \textit{for } \frac{1}{m} > \frac{1}{n}.$$

For $\dfrac{2m+1}{m(m+1)} > \dfrac{2n+1}{n(n+1)}$,

$$\frac{2m+1}{m(m+1)} - \frac{2n+1}{n(n+1)} > 0 \qquad \textit{for } \frac{1}{m} > \frac{1}{n} \textit{ implies } n-m>0$$

$$\frac{(n^2+n)(2m+1)-(2n+1)(m^2+m)}{mn(m+1)(n+1)} > 0$$

$$\vdots$$

Simplifying, we obtain

$$\frac{(n-m)(2mn+n+m+1)}{mn(m+1)(n+1)} > 0$$

We know that $(n-m)>0$ since $\dfrac{1}{m} > \dfrac{1}{n}$, and $2mn+n+m+1>0$.
Also, $mn(m+1)(n+1)>0$.

Therefore,

$$\frac{2m+1}{m(m+1)} > \frac{2n+1}{n(n+1)} \qquad \textit{for } \frac{1}{m} > \frac{1}{n}.$$

Hence, f'(x) is increasing but doesn't exist at countably many points. Convexity of f(x) is obvious from the graph.

Second Set of Criteria for Convexity

The second set of criteria is based on some generalized derivatives, in particular on Peano and Schwartz derivatives.

We begin with definitions of Peano and Schwartz derivatives.

Definition 2.1 Second Peano Derivative

If $\lim\limits_{h\to 0}\frac{1}{2!}\frac{f(x+h)-f(x)-hf'(x)}{h^2}$ exists, then the limit is denoted $f_2(x)$ and it is called the Second Peano Derivative.

If f"(x) exists, then by Taylor's formula $f_2(x)$ exists, and $f_2(x)$=f"(x). If $f_2(x)$ exists, f" doesn't necessarily exist as the following example shows.

Example 2.6 $f(x)=\begin{cases} x^2\sin\frac{1}{x} & x\neq 0 \\ 0 & x=0 \end{cases}$

Here $f_2(x)=0$, but $f''(x)$ doesn't exist.
By definition, if $f_2(x)$ exists, then $f'(x)$ must exist.

Definition 2.2 The Schwartz Derivative is defined as

$$f_s''(x)=\lim_{h\to 0}\frac{f(x-h)-2f(x)+f(x+h)}{h^2}$$

If $f_2(x)$ exists, then $f_s''(x)$ exists, as it can be easily verified.
If $f_s''(x)$ exists, then $f_2(x)$ doesn't have to exist.

Theorem 2.1 If $f_2\geq 0$ on (a,b), then f is convex.

Proof. First we prove the special case when $f_2>0$. Suppose that f is not convex. Then there are $a<x_0<x_1<x_2<b$ such that $(x_1,f(x_1))$ is above the graph of the function between the points x_0 and x_2. Let $g(x)=\frac{x-x_0}{x_2-x_0}f(x_2)$ + $\frac{x_2-x}{x_2-x_0}f(x_0)$ be the line connecting points $(x_0,f(x_0))$ and $(x_2,f(x_2))$, and consider the function $h(x)=f(x)-g(x)$. Then it is easy to check that $h(x_0)=h(x_2)=0$, $h(x_1)>0$; and since $g''(x)=0$ $h_2(x)=f_2(x)>0$. Since h is continuous on $[x_0,x_2]$, there is $x_0<\bar{x}<x_2$ such that h attains its maximum at $\bar{x}$, and hence $h'(\bar{x})=0$. From $h_2(\bar{x})=\lim\limits_{t\to 0}\frac{h(\bar{x}+t)-h(\bar{x})-t\cdot h'(\bar{x})}{t^2}=\lim\limits_{t\to 0}\frac{h(\bar{x}+t)-h(\bar{x})}{t^2}$, and $h_2(\bar{x})>0$ for all sufficiently small t, we have $\frac{h(\bar{x}+t)-h(\bar{x})}{t^2}>0$, and from here it follows

$h(\bar{x}+t)>h(\bar{x})$ which is a contradiction to the choice of $\bar{x}$. Therefore f is convex.

To prove the general case $f_2(x)\geq 0$ implies $f(x)$ is convex, we will assume that f is not convex. Therefore, there are $x_0<x_1<x_2$ such that $(x_1,f(x_1))$ is above the chord connecting $(x_0,f(x_0))$ and $(x_2,f(x_2))$. As in the special case, consider $h(x)=f(x)-g(x)$. Then $h(x_0)=h(x_2)=0$, $h(x_1)>0$ and $h_2(x)=f_2(x)\geq 0$.

Let $H(x)=h(x)-\dfrac{h(x_1)}{2}\dfrac{(x-x_0)(x_2-x)}{(x_1-x_0)(x_2-x_1)}$. Then $H(x_0)=H(x_2)=0$, and

$H(x_1)=h(x_1)-\dfrac{h(x_1)}{2}=\dfrac{h(x_1)}{2}>0$ and therefore H is not convex. On the other hand, $H_2(x)=h_2(x)+\dfrac{h(x_1)}{(x_1-x_0)(x_2-x_1)}>0$ and by the special case, H is convex. A contradiction. Therefore, f is convex.

Theorem 2.2 If f is continuous and the Schwartz derivative $f_s''(x)\geq 0$ on (a,b), then f is convex.

Proof. First we prove the special case when $f_s''(x)>0$. Suppose that f is not convex. Let $x_0<x_1<x_2$, g and h be as in the proof of the special case of Theorem 2.1. Then we have that h_s'' exists and $h_s''>0$. Since h is continuous, there is $x_0<\bar{x}<x_2$ such that h attains its maximum at $\bar{x}$. From $h_s''(\bar{x})=\lim\limits_{t\to 0}\dfrac{h(\bar{x}+t)-2h(\bar{x})+h(\bar{x}-t)}{t^2}>0$ for all sufficiently small t, $h(\bar{x}+t)+h(\bar{x}-t)>2h(\bar{x})$. But this can happen only if $h(\bar{x}+t)>h(\bar{x})$ or $h(\bar{x}-t)>h(\bar{x})$. In either case, h doesn't attain a maximum at $\bar{x}$, which is a contradiction.

The general case can be proved in exactly the same way as the general case of Theorem 2.1.

In the proofs of Theorems 2.1 and 2.2, we used basically the same idea which we briefly describe below:

We assume that a function f is not convex. By adding a suitable function g, we obtain the function h with the same original hypothesis as the function f, but for which there are three points

$x_1 < x_2 < x_3$ such that $h(x_1) = h(x_3) = 0$, and $h(x_2) > 0$. Since h is continuous there is $x_1 < \bar{x} < x_3$ such that the function h attains its maximum at $\bar{x}$. Finally the assumption on the function contradicts the choice of $\bar{x}$. We use the same idea to prove the next two theorems but first we introduce Schwartz derivates:

Let f be a function, the following expression $\underline{f_s'}(x) = \lim\inf_{t\to 0} \frac{f(x+t)-2f(x)+f(x-t)}{t^2}$ is called the lower Schwartz derivate of f at a point x while $\overline{f_s'}(x) = \lim\sup_{t\to 0} \frac{f(x+t)-2f(x)+f(x-t)}{t^2}$ is called the upper Schwartz derivate of f at a point x. Note that although neither Peano nor Schwartz derivative has to exist, the Schwartz derivates always exist.

Theorem 2.3 If f is continuous and the lower Schwartz derivate is positive, then f is convex.

Proof. Suppose that f is not convex. Let $x_0 < x_1 < x_2$, g and h be as in the proof of the special case of Theorem 2.1. Then we have that $\underline{h_s'} \geq 0$. Since h is continuous, there is $x_0 < \bar{x} < x_2$ such that h attains its maximum at $\bar{x}$. As in proofs of Theorems 2.1 and 2.2 above, it is enough to consider only the special case $\underline{h_s'} > 0$. From $\underline{h_s'}(\bar{x}) = \lim\inf_{t\to 0} \frac{h(\bar{x}+t)-2h(\bar{x})+h(\bar{x}-t)}{t^2} > 0$, for sufficiently small t, $h(x+t)-2h(x)+h(x-t) > 0$. But that can happen only if $h(\bar{x}+t) > h(\bar{x})$ or $h(\bar{x}-t) > h(\bar{x})$. In either case, h doesn't attain a maximum at $\bar{x}$, which is a contradiction.

Theorem 2.4 If f is continuous and the upper Schwartz derivate is positive, then f is convex.

Proof. Suppose that f is not convex. Let $x_0 < x_1 < x_2$, g and h be as in the proof of the special case of Theorem 2.1. Then we have that $\overline{h_s'} \geq 0$ Since h is continuous, there is $x_0 < \bar{x} < x_2$ such that h attains its maximum at $\bar{x}$. From $\overline{h_s'}(\bar{x}) = \lim\sup_{t\to 0} \frac{h(\bar{x}+t)-2h(\bar{x})+h(\bar{x}-t)}{t^2} > 0$, there is a sequence $\{t_n\}$ converging to 0 such that $h(\bar{x}+t_n)-2h(\bar{x})+h(\bar{x}-t_n)>0$. But that can happen only if $h(\bar{x}+t_n)>h(\bar{x})$ or $h(\bar{x}-t_n)>h(\bar{x})$. In either case, h doesn't attain a maximum at $\bar{x}$, which is a contradiction.

BIBLIOGRAPHY

[1] R. Benson, *Euclidean Geometry and Convexity,* McGraw-Hill, New York, 1966.

[2] H. G. Eggleston, *Problems in Euclidean Space: Application of Convexity,* Cambridge University, New York, 1957.

[3] S. G. Krantz, *Real Analysis and Foundations,* CRC Press, New York, 1991.

[4] I. P. Natanson, *Theory of Functions of a Real Variable,* Vol. II, Frederick Ungar Publishing, New York, 1960, 36-47.

[5] A. W. Roberts and D. E. Varberg, *Convex Functions,* Academic Press, New York, 1973.

[6] R. T. Rockafellar, *Convex Analysis,* Princeton University, Princeton, New Jersey, 1970.

[7] W. Rudin, *Real and Complex Analysis,* McGraw-Hill, New York, 1974.

[8] A. Zygmund, *Trigonometric Series*, Cambridge University Press, Cambridge and New York, 1990.