Real Analysis Exchange
Vol. 29(1), 2003/2004, pp. 365–379

Chris Freiling, Department of Mathematics, California State University, San Bernardino, CA, 92407. email: `cfreilin@csusb.edu`

Cynthia Pittman, Department of Mathematics, California State University, San Bernardino, CA, 92407. email: `cynthiapittman@prodigy.net`

CONVEXITY USING A COMBINATION OF DERIVATIVES

Abstract

We show how a combination of three different second derivatives can yield convexity, even though individually, none of them is sufficient.

1 Introduction

One of the most useful properties of the second derivative is the convexity property:

Theorem 1.1. *If $f''(x) \geq 0$ on an open interval, then $f(x)$ is convex (i.e., concave up) on that interval.*

This article deals with the following three generalizations of the second derivative and their relationship to convexity:

1) The forward lower second derivate:

$$\underline{f}_2^+(x) = \liminf_{h\to 0^+} \frac{f(x+2h) - 2f(x+h) + f(x)}{h^2}$$

2) The backward lower second derivate:

$$\underline{f}_2^-(x) = \liminf_{h\to 0^+} \frac{f(x) - 2f(x-h) + f(x-2h)}{h^2}$$

3) The symmetric lower second derivate:

$$\underline{f}_2^0(x) = \liminf_{h\to 0^+} \frac{f(x+h) - 2f(x) + f(x-h)}{h^2}.$$

Key Words: convexity, Schwarz derivative, generalized derivative.
Mathematical Reviews subject classification: 26A24, 26A51, 52A10.
Received by the editors April 29, 2003
Communicated by: B. S. Thomson

The last of the three has been well-studied. For example, if f is assumed to be continuous, then the condition $\underline{f}_2^0 \geq 0$ easily yields the desired convexity property. (In fact, *liminf* can be weakened to *limsup.*) This important fact, due to Schwarz, was a key ingredient in Cantor's proof of uniqueness of trigonometric series, see [2]. It is interesting to explore what happens when the continuity assumption is relaxed. Much work has been done in this area, see [3], culminating in the work of Fejzić [1] (see Section 2).

In this article we will be relaxing the assumption of continuity, but we will always assume the measurability of our function f. Under this weaker assumption, none of these derivates is strong enough to force convexity. In fact, no pair of them is enough, not even when the function is very simple. For example,

$$\mathrm{SGN}(x) = \begin{cases} 1 & \text{if } x > 0 \\ 0 & \text{if } x = 0 \\ -1 & \text{if } x < 0 \end{cases}$$

has nonnegative values for the backward and the symmetric lower derivates, yet is not convex. Observe that this example is discontinuous, as must be the case by Schwarz' Theorem. On the other hand, $y = -|x|$ has both forward and backward lower second derivates identically zero. Because this example is continuous, Schwarz' Theorem would be false if the symmetric derivate were replaced with the two one-sided derivates. In this sense, we might say that the backward and forward derivates together are not as strong as the symmetric derivate.

We will show that all three derivates, working together, do imply convexity. We give two proofs of this result. In Section 2, we prove it as a corollary of the work of Fejzić [1]. In Section 3, we give a self-contained proof. The two proofs use different roles for the three derivates, and because of this, they lead to different generalizations.

2 Proof I.

Fejzić[1], Theorem 6, proved the following theorem.

Theorem 2.1 (Fejzić). *Let I be an open interval. If f is a measurable function defined on I such that the lower second symmetric derivate is non-negative on I, then there is a dense open subset of $G \subseteq I$ such that f is convex on every component of G.*

From this, he proves the following corollary.

Corollary 2.2 (Fejzić)**.** *Let I be an open interval and f a measurable function defined on I such that the lower second symmetric derivate is non-negative on I. Let G be the dense open set from Theorem 2.1. If f is continuous on the closure of every component of G, then f is convex on I.*

Our main result will follow from this corollary. Indeed, all that remains is to show the following two lemmas.

Lemma 2.3. *Suppose f is convex on an open interval (a,b), with*

$$\limsup_{h\to 0^+} f(a+2h) - 2f(a+h) + f(a) \geq 0\,.$$

Then f is convex on $[a,b)$. Similarly, if

$$\limsup_{h\to 0^+} f(b-2h) - 2f(b-h) + f(b) \geq 0\,,$$

then f is convex on $(a,b]$.

Proof. We show the second part, the first part being identical. Choose two points $c < d$ in the interval (a,b) and subtract from $f(x)$ the line joining $(c,f(c))$ with $(d,f(d))$. Subtracting this line does not change any convexity (or lack of convexity) and also does not change the second order difference, $f(b-2h)-2f(b-h)+f(b)$. Therefore, we can assume without loss of generality that this line has already been subtracted from f. Then f passes through both $(c,0)$ and $(d,0)$ and hence by convexity, it will be non-decreasing on the interval (d,b). Therefore $\lim_{x\to b^-} f(x)$ exists as an extended real number $L \in [0,\infty]$. We complete the proof by showing that $L \leq f(b)$. Suppose not. Let $\delta = L - f(b)$ if L is finite, and $\delta = 1$ otherwise. Then for all sufficiently small $h > 0$, we have $f(b-h) - f(b) > \delta/2$. But then

$$\begin{aligned} f(b) - 2f(b-h) + f(b-2h) =&[f(b) - f(b-h)] + [f(b-2h) - f(b-h)] \\ <& -\delta/2 + 0 = -\delta/2\,. \end{aligned}$$

contradicting our hypothesis. ☐

Lemma 2.4. *Suppose f is convex on an closed interval $[a,b]$, and f is defined in a neighborhood of a and b such that*

- $\liminf_{h\to 0^+} f(a-2h) - 2f(a-h) + f(a) \geq 0$
- $\liminf_{h\to 0^+} f(a-h) - 2f(a) + f(a+h) \geq 0$
- $\liminf_{h\to 0^+} f(b+2h) - 2f(b+h) + f(b) \geq 0$

- $\liminf_{h\to 0^+} f(b+h) - 2f(b) + f(b-h) \geq 0$.

Then f is continuous on $[a,b]$.

PROOF. It is well-known that if a function is convex on an open interval, then it is also continuous on that interval. Therefore, all that needs to be shown is that the function is continuous from the left at b and continuous from the right at a. By symmetry, we only need to show continuity from the left at b.

Choose two points $c < d$ in the interval (a,b) and subtract from $f(x)$ the line joining $(c, f(c))$ with $(d, f(d))$. The remaining function will have $f(c) = f(d) = 0$. Hence by convexity, it will be non-decreasing on the interval (d,b). Therefore as in the last lemma, we may assume without loss of generality that f is non-decreasing in a left-neighborhood of b. Then $\lim_{x\to b^-} f(x) = L$ exists. We need to show that $f(b) = L$. The convexity assumption already gives us that $f(b) \geq L$.

On the other side of b, the limit may not exist. We let $M = \liminf_{x\to b^+} f(x)$. We will show that $2f(b) - L \leq M \leq f(b)$. This will finish the proof since it implies that $f(b) \leq L$. We use process of elimination.

Case 1. $M = \infty$. Then for sufficiently small $h > 0$, $f(b+2h) > f(b)+1$. If h is small enough, then by hypothesis we also have $f(b+2h)-2f(b+h)+f(b) \geq -1$. Combining, $f(b+h) \leq [f(b+2h) + f(b) + 1]/2 < f(b+2h)$. Thus the sequence $f(b+2h), f(b+h), f(b+\frac{h}{2}), \ldots$ is decreasing, contradicting that $M = \infty$.

Case 2. $M < 2f(b) - L$. Let $\delta = f(b) - \frac{M+L}{2} > 0$, if M is finite, and $\delta = 1$ otherwise. In the first case, $2f(b) - L - \delta = M + \delta$, and in either case, $2f(b) - L - \delta > M$. Then for arbitrarily small $h > 0$, $f(b+h) < 2f(b) - L - \delta$. Then, $f(b+h) - 2f(b) + f(b-h) < -L - \delta + f(b-h) \leq -\delta$, since f is non-decreasing on the left of b. But this contradicts our hypothesis.

Case 3. $f(b) < M < \infty$. We will show that for arbitrarily small $h > 0$, $f(b+h) < M - \frac{M-f(b)}{4} < M$, which immediately contradicts the definition of M. To see this, let $\delta = \frac{M-f(b)}{4} > 0$, and choose $h > 0$ such that $f(b+2h) < M + \delta$. Suppose that $f(b+h) \geq M - \delta$. Then $f(b+2h) - 2f(b+h) + f(b) < (M+\delta) - 2(M-\delta) + f(b) = -(M - f(b)) + 3\delta = -4\delta + 3\delta = -\delta$, contradicting our hypothesis. □

Combining the previous two lemmas with Corollary 2.2 immediately gives the following.

Theorem 2.5. *Let I be an open interval and f a real valued measurable function defined on I such that the forward, backward, and symmetric lower second derivates are all non-negative on I. Then f is convex on I.*

3 Proof II.

The reader may have noticed that in the last proof, the bulk of the work was accomplished by the symmetric derivate (Corollary 2.2), while the one-sided derivates were only used for the "clean-up" work (Lemmas 2.3 and 2.4). This in underscored by the fact that the denominators of $\underline{f}_2^-$ and $\underline{f}_2^+$ were never used. (The denominator of $\underline{f}_2^0$ was crucial in the proof of Theorem 2.1.) In fact, we actually proved the following generalization of Theorem 2.5.

Theorem 3.1. *Let I be an open interval and f a real valued measurable function defined on I such that the symmetric lower second derivate is non-negative on I. Assume, also, that $lim_{h\to 0} f(x+2h)-2f(x+h)+f(x) \geq 0$ on I. Then f is convex on I.*

In this section we will reverse the roles. We will give a proof in which the bulk of the work is done by the one-sided derivatives and relegate the "clean-up" work to the symmetric derivate. The reader might wonder for a minute whether this means that we can eliminate the denominator in the definition of $\underline{f}_2^0$. This is not possible, the function $y = -|x|$ being a counterexample. However, we will be able to reduce the degree of the denominator from h^2 to h. When this is done for the second symmetric derivative, that is, when

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

the function is called *smooth*. (The notion is due to Riemann; the terminology is due to Zygmund, see [3]). Here we have a one-sided derivate, so we make the following definition.

Definition 3.2. A function f will be called *semi-smooth* at x if

$$\liminf_{h\to 0^+} \frac{f(x+h)-2f(x)+f(x-h)}{h} \geq 0.$$

A function will be called semi-smooth on an open interval if it is semi-smooth at each point of the interval.

The main goal of this section is to prove the following theorem which also generalizes Theorem 2.5.

Theorem 3.3. *Suppose f is measurable on an open interval I, f is semi-smooth on I, and $\underline{f}_2^-$ and $\underline{f}_2^+$ are non-negative on I. Then f is convex on I.*

We first observe that there is an equivalent formulation with strict inequalities.

Theorem 3.4. *Suppose f is measurable on an open interval I, f is semi-smooth on I, and $\underline{f}_2^-$ and $\underline{f}_2^+$ are positive on I. Then f is convex on I.*

Theorem 3.4 is more convenient to prove. This is because if the *liminf* of an expression is positive, then the expression itself is eventually positive, a fact which is not necessarily true if "positive" is replaced by "non-negative". However, a standard trick easily shows that Theorem 3.3 follows from Theorem 3.4.

PROOF. [Proof of Theorem 3.3 from Theorem 3.4.] Suppose f is measurable and semi-smooth on I. Suppose also that $\underline{f}_2^+ \geq 0$, $\underline{f}_2^- \geq 0$ on I, but f is not convex on I. Then there exists $x < y < z$ in I where $(y, f(y))$ is above the line joining $(x, f(x))$ and $(z, f(z))$. Let $g(x) = f(x) + \epsilon x^2$. For small enough $\epsilon > 0$, $(y, g(y))$ is still above the line joining $(x, g(x))$ and $(z, g(z))$. But $\underline{g}_2^+ = \underline{f}_2^+ + 2\epsilon$ and $\underline{g}_2^- = \underline{f}_2^- + 2\epsilon$ are now strictly positive on I. Also, since f is semi-smooth, g will be semi-smooth. Applying Theorem 3.4 to g gives the desired contradiction. □

The same sort of argument can be used to show that the "convex" in Theorem 3.4 can be replaced with "strictly convex".

The rest of this section will be devoted to the proof of Theorem 3.4. We will work as far as we can using only the one-sided derivates. Only toward the end of the proof will we invoke the condition of *semi-smooth*.

3.1 Boundedness on an Interval.

In this section we show that if either of the one-sided derivates is positive, then this is enough to force the function to be bounded on a subinterval.

Lemma 3.5. *Suppose $\underline{f}_2^+(x) > 0$ on a non-empty open interval I on which f is measurable. Then f is bounded on a non-empty open subinterval of I.*

PROOF. For each x in I, let $\delta(x)$ denote the largest positive value of δ such that for all $h \in (0, \delta)$, we have $f(x+2h) - 2f(x+h) + f(x) > 0$. Choose δ small enough so that the set $A = \{x \in I | \delta(x) > \delta\}$ has positive outer measure. Let J be a closed subinterval of I with non-empty interior and with size less than δ such that the relative outer measure of A in J (i.e., $\frac{\lambda^*(A \cap J)}{\lambda(J)}$)is more than 99%. We partition J into five equal-length subintervals, labeled left to right as K_1, K_2, K_3, K_4, K_5. Then A has relative outer measure more than 95% in each K_i. Using that f is measurable, choose M so that at least 99% of the function values in J are between $-M$ and M. Let B denote the set of $x \in J$ with $f(x) < M$. Then B must include at least 95% of each K_i and therefore the relative outer measure of $A \cap B$ in each K_i must be at least 90%. We will show that for $b \in K_3$, $f(b) \in [-3M, M]$.

Suppose $b \in K_3$ with $f(b) > M$. Let $x \in A \cap B \cap K_2$ and let $h = b - x$. Then, using the definitions of A and δ,

$$\begin{aligned} f(x+2h) - 2f(x+h) + f(x) &> 0 \\ f(2b-x) - 2f(b) + f(x) &> 0 \\ f(2b-x) &> 2f(b) - f(x) > M. \end{aligned}$$

Also, $2b - x \in J$. But this is true for a set of x of outer measure at least 90% of the size of K_2, which is 18% of the size of J, contradicting the choice of M.

Next, suppose $b \in K_3$ with $f(b) < -3M$. Let $x \in A \cap B \cap K_1$. Then $z = \frac{x+b}{2}$ will also be in J. Let $h = \frac{b-x}{2}$. Then, using the definitions of A and δ,

$$\begin{aligned} f(x+2h) - 2f(x+h) + f(x) &> 0 \\ f(b) - 2f(\frac{x+b}{2}) + f(x) &> 0 \\ f(\frac{x+b}{2}) < \frac{f(b)+f(x)}{2} &< \frac{-3M+M}{2} = -M. \end{aligned}$$

But this is true for a set of x of outer measure at least 90% of the size of K_1, which is 18% of the size of J. Therefore, $f(z) < -M$ for a set of z of outer measure at least 9% of the size of J, contradicting the choice of M. □

3.2 Convexity on an Interval.

The main goal of this section is to show that if both one-sided derivates are positive, then f is convex on a subinterval.

Lemma 3.6. *Suppose $\underline{f}_2^+(x) > 0$ and $\underline{f}_2^-(x) > 0$ on a non-empty open interval I on which f is measurable. Then f is convex on a non-empty open subinterval of I.*

Proof. By the previous lemma we may assume that f is bounded on I. For each $x \in I$ let $\delta(x)$ denote the largest positive value of δ such that for all $h \in (0, \delta)$, we have both $f(x+2h) - 2f(x+h) + f(x) > 0$ and $f(x-2h) - 2f(x-h) + f(x) > 0$. For each $\delta > 0$ we define $A_\delta = \{x \in I | \delta(x) > \delta\}$. By the Baire Category Theorem, for some $\delta > 0$ and some non-empty open subinterval $J \subseteq I$, A_δ is dense in J. By shrinking J if necessary we may also assume that the size of J is less than δ. We will show that f is convex on J. Suppose not. Then there is some $a < b < c$ in J with $(b, f(b))$ above the line joining $(a, f(a))$ with $(c, f(c))$. By subtracting this line, we may assume without loss of generality that $f(a) = f(c) = 0$. Let $M > 0$ be the least upper bound of the function values $f(b)$ for $b \in (a, c)$.

Fix some $b \in (a, c)$ with $f(b) > .99M$. Assume without loss of generality that $b \in (a, \frac{a+c}{2})$. Choose $x \in A_\delta \cap (a, a + \delta(a)) \cap (a, b)$. Then $2b - x$ is in (a, c) and so is $2x - a$. Since $x \in A_\delta$, we have

$$\begin{aligned} f(2b-x) - 2f(b) + f(x) &> 0 \\ f(x) + f(2b-x) &> 2f(b) > 1.98M. \end{aligned}$$

But by the definitions of $\delta(a)$ and M, $f(2x - a) - 2f(x) + f(a) > 0$

$$\begin{aligned} M \geq f(2x-a) &> 2f(x) \\ f(x) &< M/2\,. \end{aligned}$$

But then $f(2b - x) + M/2 > 1.98M$, so $f(2b - x) > 1.48M$, contradicting the choice of M. □

3.3 One-Dimensional Structure.

For the rest of this paper, we will be dealing with a counter-example to Theorem 3.4. That is, we have a function f that is measurable and semi-smooth on an open interval I, which is not convex, but for all $x \in I$, $\underline{f}_2^-(x) > 0$ and $\underline{f}_2^+(x) > 0$. In this section we will explore some of the structure that such a function must exhibit, structure that will eventually lead to a contradiction.

By Lemma 3.6, we know there will be non-empty subinterval on which f is convex. By Lemma 2.3 f will be convex on the closure of this interval in I. Choose a maximal such interval P and apply the lemma again on each side of P. By successive such applications, we obtain a dense open subset of I such that f is convex on the closure (relative to I) of each component of this set. Furthermore, no component can be expanded to a larger interval where f is convex. We call this dense open set *Pink.* Any point in I that is not Pink will be called *Green.* We will use the adjectives *pink* and *green* to refer to membership in Pink and Green respectively.

We begin with a lemma on the components of Pink. This establishes that Green is a perfect set and is the first time in the proof where we invoke the notion of *semi-smooth.*

Lemma 3.7. *No two components of Pink are adjacent.*

PROOF. Suppose (a, b) and (b, c) are components of Pink. We will show that f is convex on (a, c). Suppose not. Then there are points $x < y < z$ in (a, c) with $(y, f(y))$ above the line joining $(x, f(x))$ with $(z, f(z))$. Assume without loss of generality that $f(x) = f(z) = 0$ and hence $f(y) > 0$. Since f is convex on $(a, b]$ and on $[b, c)$, it must be that $x < b < z$. But one of these pieces must contain y and so $0 = f(z) = f(x) < f(y) \le f(b)$. Therefore, f attains a maximum on $[x, z]$ at b. Let $m_1 = \frac{f(b)}{b-x} > 0$, $m_2 = \frac{f(b)}{b-z} < 0$. Then for sufficiently small $h > 0$, convexity on $[x, b]$ and $[b, z]$ implies that

$$\frac{f(b) - f(b-h)}{h} \ge m_1 \text{ and } \frac{f(b+h) - f(b)}{h} \le m_2.$$

Subtracting in reverse order, $\frac{f(b+h)-2f(b)+f(b-h)}{h} \le m_2 - m_1 < 0$, contradicting the semi-smoothness of f at b. □

For each $x \in$Green, let $\delta(x)$ denote the largest positive value of δ such that for all $h \in (0, \delta)$, we have both $f(x + 2h) - 2f(x + h) + f(x) > 0$ and $f(x - 2h) - 2f(x - h) + f(x) > 0$. For each $\delta > 0$ we define $A_\delta = \{x \in \text{Green} \mid \delta(x) > \delta\}$. By the Baire Category Theorem, for some $\delta > 0$ and some subinterval $J \subseteq I$, we have both that $J\cap$ Green $\neq \emptyset$ (i.e., f is not convex on J) and A_δ is dense in $J\cap$ Green. By shrinking J if necessary, we may also assume that the size of J is less than δ. We fix one such pair J, δ. Then every $x \in J \cap A_\delta$ is both *left-good* and *right-good* according to the following definition.

Definition 3.8. Let $x \in J \cap Green$. Then x is *right-good* (resp. *left-good*) if and only if whenever $y > x$ (resp. $y < x$) and $2y - x \in J$, then $f(2y - x) - 2f(y) + f(x) \ge 0$.

Definition 3.9. Let $x \in J \cap Green$. Then x is *good on the right* (resp. *good on the left*) if and only if there are right-good points $y \ge x$ (resp. left-good points $y \le x$) arbitrarily close to x.

Note that since A_δ is dense in $J \cap Green$, every $x \in J \cap Green$ is either good on the right or it is a left endpoint of a Pink component. Similarly, every $x \in J \cap Green$ is either good on the left or it is a right endpoint of a Pink component.

Definition 3.10. Let $x \in J \cap Green$. Then x is *right-convex* (resp. *left-convex*) if and only if whenever $y \in J$ and $y > x$ (resp. $y < x$) and z is any point between x and y, then $(z, f(z))$ is on or below the line joining $(x, f(x))$ and $(y, f(y))$.

We will now proceed to show the equivalence of the three previous definitions. But first, a lemma.

Lemma 3.11. *Suppose x is good on the right and $y > x$ is good on the left. Let z be any point in $[x, y]$. Then $(z, f(z))$ is on or below the line joining $(x, f(x))$ and $(y, f(y))$.*

PROOF. Suppose not. Assume without loss of generality that $f(x) = f(y) = 0$. Let $M > 0$ be the least upper bound of the values of $f(z)$ for $z \in (x, y)$. Fix $z \in (x, y)$ with $f(z) > \frac{3}{4}M$. We may assume by symmetry that $z \in (x, \frac{x+y}{2}]$. Choose a right-good point $p \in [x, x + \delta(x)) \cap [x, z)$. If $f(p) \leq M/2$, then $f(2z - p) \geq 2f(z) - f(p) > M$, contradicting the choice of M. But if $f(p) > M/2$, then $p \neq x$ so $p \in (x, x+\delta(x))$, from which it follows that $f(2p-x) > M$. This also contradicts the choice of M. □

Lemma 3.12. *For each $x \in J \cap Green$ the following are equivalent:*
(i) x is right (resp. left)-good
(ii) x is good on the right (resp. left)
(iii) x is right (resp. left)-convex

PROOF. We show the right version, the left being identical.
(i) $\Rightarrow$ (ii). This is by definition. The sequence of points can be chosen to be the constant sequence, $x_i = x$.
(iii) $\Rightarrow$ (i). This also follows immediately, noticing that $f(2y - x) - 2f(y) + f(x) \geq 0$ is equivalent to saying that $(y, f(y))$ is on or below the line joining $(x, f(x))$ and $(2y - x, f(2y - x))$.
(ii) $\Rightarrow$ (iii). Let x be good on the right. Choose $y > x$ in J. We will show that for any z between x and y, $(z, f(z))$ is on or below the line joining $(x, f(x))$ with $(y, f(y))$. Suppose this is not the case. If y is good on the left we are done by the previous lemma. Otherwise, y is either Pink or at the right endpoint of a component of Pink. In either case, let w be the left endpoint of that component. Then f is convex on $[w, y]$. If $x = w$ we are done, so assume $x < w$. By Lemma 3.7 w is good on the left. Assume without loss of generality that $f(x) = f(y) = 0$, and that for some $z \in (x, y)$, $f(z) > 0$. Let $M > 0$ be the least upper bound of f on (x, y).

We first claim that $f(w) = M$. If not, then some $z \in (x, y)$ has $f(z) > f(w)$ and $f(z) > 0$. If $z < w$, then because w is good on the left we have a contradiction to the previous lemma. But if $z > w$, then this contradicts the convexity of f on $[w, y]$. Therefore, $f(w) = M$.

Now choose $z \in (x, w) \cap A_\delta$, close enough to w so that $2w - z < y$. By the previous lemma applied to x and w, $f(z) < f(w)$. But then since z is right-good, we get $f(2w - z) > f(w)$, which is a contradiction. □

We are now able to give the main result of this section.

Corollary 3.13. *Every $x \in J \cap Green$ that is not a left endpoint of Pink, is right-convex. Similarly, every $x \in J \cap Green$ that is not a right endpoint of Pink is left-convex.*

PROOF. In the first case, using the density of A_δ, any such x is good on the right. Similarly in the second case, any such x is good on the left. Hence the corollary follows immediately from the previous lemma. □

3.4 Two Dimensional Structure.

In this section we present a 2-dimensional representation of the structure established in the last section. Let us first summarize this structure. We are dealing with a counterexample to Theorem 3.4. We pass to a subinterval J on which f is also not convex, and which can be partitioned into two sets Pink and Green with the following properties. Pink is a dense open set and Green is a perfect set. Our function f is convex on the closure of each component of Pink. At each green point that is not a left endpoint of Pink, f is right-convex. At each green point that is not a right endpoint of Pink, f is left-convex.

To turn this into a two-dimensional structure, we first review some definitions and basic facts used in [4] (see also [3]).

Definition 3.14. A *2-interval* is an ordered quadruple (a, b, c, d) with $a+d = b+c$. The center of the 2-interval is the point $(a+d)/2 = (b+c)/2$. If $b = c$, then this is called a *Schwarz interval* with center b and radius $b - a$.

Definition 3.15. If F is a real function and I is a 2-interval, (a, b, c, d), then we define $F(I) = F(a) - F(b) - F(c) + F(d)$.

The notion of a 2-interval allows us to view the convexity of a real function in a two-dimensional setting. Given any rectangle with sides parallel to the coordinate axes, we can project the rectangle along a slope of -1 to obtain a 2-interval on the x-axis (see Figure 1). When a rectangle R projects to a 2-interval I we further extend the definition of F so that $F(R) = F(I)$. The function F now defined on rectangles is an additive function. That is, if a rectangle R can be partitioned into rectangles $R_1, \dots R_n$, then $F(R) = F(R_1) + \cdots + F(R_n)$.

Lemma 3.16. *Let F be a real measurable function and I a closed interval. The following are equivalent.*
(1) F is concave up on I
(2) F takes on a non-negative value on every Schwarz interval in I.
(3) F takes on a non-negative value on every 2-interval in I.

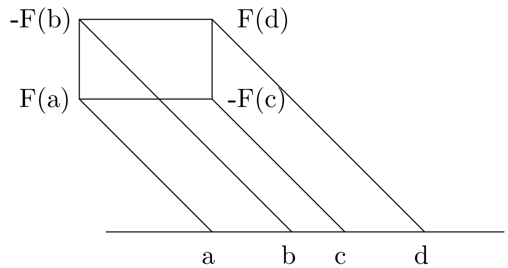

Figure 1: A 2-interval is the projection of a rectangle.

PROOF. (1)⇒(3) and (3)⇒(2) are immediate. To see that (2)⇒(1), suppose that F is not concave up on I. Then there exists $a < c < b$ in I such that $(c, F(c))$ lies above the line joining $(a, F(a))$ with $(b, F(b))$. Assume without loss of generality that $F(a) = F(b) = 0$ and $F(c) > 0$. Suppose first that almost every x between a and b has $F(x) \leq 0$. Choose any such $x < c$ that is close enough to c so that $2c - x$ is between a and b. Then using the Schwarz interval $(x, c, c, 2c - x)$ we get that $F(2c - x) - 2F(c) + F(x) \geq 0$, so that $F(2c - x) \geq 2F(c) - F(x) > 0$. Therefore, F is positive on a set of positive measure, which is a contradiction. Therefore, we may assume that $F(x) > 0$ for a subset of (a,b) that has positive measure. For each positive real number y let $B_y = \{x \,|\, a < x < b \text{ and } F(x) \geq y\}$. We will now show that $B_{2y} \geq B_y$ in measure. Let L_y be the intersection of B_y with the left half of (a, b). Assume without loss of generality that L_y has at least half the measure of B_y, the other case being symmetrical. Each x in L_y is closer to a than to b, so $2x - a$ is also between a and b. Using the Schwarz interval $(a, x, x, 2x - a)$, we get $F(2x - a) \geq 2F(x)$ so $2x - a$ is in B_{2y}. Therefore, $g(x) = 2x - a$ maps L_y into B_{2y}. Furthermore, the image of L_y under this map is twice the measure of L_y, and therefore at least as big as B_y, establishing that $B_{2y} \geq B_y$ in measure. Therefore, for each $y > 0$, $B_y \setminus B_{2y}$ has measure zero. This completes the contradiction because then B_0 is a countable union of measure zero sets and hence also has measure zero. □

We can now transform the summary given at the beginning of this section to a two-dimensional setting. If f is a counterexample to Theorem 3.4, then there is a square J with $f(J) < 0$. We consider this square to be made up of line segments of slope -1. There is a dense open set of "pink" strips with slope -1 with the following property. If $R \subseteq J$ is any rectangle with pink interior, then $f(R) \geq 0$. The perfect set of green lines that make up the complement

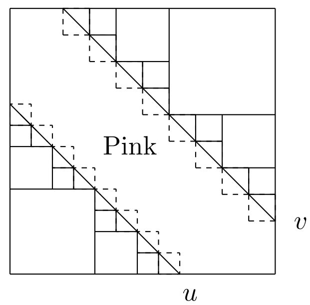

Figure 2: Partitioning of J.

of the pink strips has the following property. If $R \subseteq J$ is any square whose upper-right vertex is on a green line and this green line is not the upper-right end of a pink strip, then $f(R) \geq 0$. Similarly, if the lower left vertex of the square R is on a green line that is not the lower-left end of a pink strip, then $f(R) \geq 0$.

3.5 Conclusion of the Proof.

We have used the condition of *semi-smooth* in the proof of Lemma 3.7. Now we will use it a second time. Given any particular green line, and any square $R \subseteq J$ whose main diagonal lies on this line, the semi-smooth condition gives us that $\liminf_{h \to 0^+} f(R)/h \geq 0$, where h is the length of a side of R. This extra condition now allows us to complete the proof of Theorem 3.4.

PROOF. [Conclusion to the proof of Theorem 3.4.] We begin by choosing two green lines, u and v and a number $\epsilon > 0$. If the main diagonal of J is in a pink strip or at the edge of a pink strip, we let u be the lower-left edge of that strip and v be the upper-right edge, see Figure 2. It may turn out that one of these is outside of J because one entire half of J is pink. In this case we just ignore it. The only other possibility is that the main diagonal of J lies on a green line that is a bilateral limit of other green lines. In this case we let $u = v$ be the main diagonal.

We will partition the square J into a finite number of rectangles R_i. The first set of rectangles will be squares with upper-right vertex on u, or lower-left vertex on v. Hence, these will all have non-negative values. The squares with upper-right vertex on u can be described as follows. First choose the largest such square. Then choose the next largest square that does not overlap the first one; there will be two of these. Then choose the next four to be the largest ones that don't overlap the squares already chosen, etc. These are shown as the solid squares in Figure 2. After n stages of this procedure are carried out, the line segment u will be partitioned into 2^n equal pieces. We will choose n, the number of stages, to be a large number which will be determined later. The same procedure is carried out for the set of squares with lower-left vertex on v. These squares are illustrated with solid borders in Figure 2.

The second set of rectangles will be squares with centers on u. These will be chosen so that their main diagonals are the intervals along u that were partitioned by the first stage. They all have the same size, h which can be made arbitrarily small by choosing n large in the first stage. Thus, h will be made small enough so that each of these squares, R has $f(R) \geq -\epsilon h$, using the semi-smoothness of f. If H is the length of one side of the square J, then there are at most H/h of these second-stage squares, giving a total value of at least $-H\epsilon$. Similarly, there will be a third set of squares with centers on v with total value at least $-H\epsilon$. In the case where $u = v$, that is, the case where the main diagonal of J is bilateral limit of green lines, these last two sets of squares are really the same and we use only one copy of them. Otherwise, we can choose them small enough so that there is no overlap between them. In any case, the total value is at least $-2H\epsilon$. These squares are illustrated in Figure 2 as dashed boxes.

To form the last set of rectangles, partition the remaining region, whose interior is Pink, into any finite number of rectangles. It does not matter how this is carried out or how many rectangles are used. Since the interior of each of them is pink, f will assign each rectangle a non-negative value. This completes the partition.

The total value of all of the rectangles in the partition is at least $-2H\epsilon$. Since $\epsilon > 0$ is arbitrary, the value of $f(J)$ is non-negative, a contradiction. $\square$

References

[1] H. Fejzić, *Convex functions and Schwarz derivatives*, Proc. Amer. Math. Soc., **123**(8) (August 1995), 2473–2477.

[2] A. Zygmund, *Trigonometric Series*, 2nd ed., **1** and **2**, Cambridge University Press, Cambridge, (1959) reprinted (1993).

[3] B. S. Thomson, *Symmetric Properties of Real Functions*, *Monographs and Textbooks in Pure and Applied Mathematics*, **183**, Marcel Dekker, Inc., New York (1994).

[4] J. M. Ash, C. Freiling, and D. Rinne, *Uniqueness of rectangularly convergent trigonometric series*, *Annals of Math.*, **137**, (1993) 145–166.