Problem Set

Problem Set #2

Math 5322, Fall 2001

March 4, 2002

ANSWERS

All of the problems are from Chapter 3 of the text.

Problem 1. [**Problem 24, p. 100**] If $f \in L^1_{\mathrm{loc}}$ and f is continuous at x, then x is in the Lebesgue set of f.

Answer:
To show that x is in the Lebesgue set of f, we must show

$$(*) \qquad \lim_{r \to 0} \frac{1}{m(\mathrm{B}(r,x))} \int_{\mathrm{B}(r,x)} |f(y) - f(x)|\, dy = 0.$$

Let $\varepsilon > 0$ be given. Since f is continuous at x, there is a $\delta > 0$ such that $|f(y) - f(x)| < \varepsilon$ for all $y \in \mathrm{B}(\delta, x)$.

Hence, if $0 < r < \delta$, we have $|f(y) - f(x)| < \varepsilon$ for all $y \in \mathrm{B}(r, x)$, so

$$\frac{1}{m(\mathrm{B}(r,x))} \int_{\mathrm{B}(r,x)} |f(y) - f(x)|\, dy \leq \frac{1}{m(\mathrm{B}(r,x))} \int_{\mathrm{B}(r,x)} \varepsilon\, dy = \varepsilon.$$

Since ε was arbitrary, this shows that $(*)$ holds.

Problem 2. [**Problem 28, p. 107**] If $F \in \mathrm{NBV}$, let $G(x) = |\mu_F|((-\infty, x])$. Prove that $|\mu_F| = \mu_{T_F}$ by showing $G = T_F$ via the following steps.

a. From the definition of T_F, $T_F \leq G$.

b. $|\mu_F(E)| \leq \mu_{T_F}(E)$ when E is an interval, and hence when E is a Borel set.

c. $|\mu_F| \leq \mu_{T_F}$ and hence $G \leq T_F$. (Use Exercise 21.)

Answer:
The definition of T_F is at the top of page 102 in the text. If we have points $-\infty < x_0 < x_1 < \cdots < x_n = x$, then

$$\begin{aligned}
\sum_{j=1}^{n} |F(x_j) - F(x_{j-1})| &= \sum_{j=1}^{\infty} |\mu_F((x_{j-1}, x_j])| \\
&\leq \sum_{j=1}^{n} |\mu_F|((x_{j-1}, x_j]) \\
&= |\mu_F|((x_0, x]) \\
&\leq |\mu_F|((-\infty, x]) \\
&= G(x).
\end{aligned}$$

Taking the sup over all choices of the partition $\{ x_j \}$, we get $T_F(x) \leq G(x)$.

For an h-interval $(a,b]$, we have

$$\begin{aligned} |\mu_F((a,b])| &= |F(b)-F(a)| \\ &\le T_F(b)-T_F(a) && \text{by equation (3.24), p. 102} \\ &= \mu_{T_F}((a,b]) \end{aligned}$$

Hence, we have the inequality

$$(*) \qquad |\mu_F(E)| \le \mu_{T_F}(E)$$

when E is an h-interval. We want to show this holds when E is a Borel set. It seems to me that some work is required—the uniqueness theorem Theorem 1.14 (p. 31) doesn't seem to help, since we are not proving the equality of two measures.

A simple solution is to proceed as follows. If $E = \bigcup_n I_n$ is the countable union of a family of disjoint h-intervals then we have

$$\begin{aligned} |\mu_F(E)| &= \left|\sum_{n=1}^{\infty} \mu_F(I_n)\right| \\ &\le \sum_{n=1}^{\infty} |\mu_F(I_n)| \\ &\le \sum_{n=1}^{\infty} \mu_{T_F}(I_n) \\ &= \mu_{T_F}(E). \end{aligned}$$

Hence $(*)$ holds for $E \in \mathcal{A}$, where $\mathcal{A}$ is the algebra of all finite disjoint unions of h-intervals. The σ-algebra generated by $\mathcal{A}$ is the Borel sets, but it doesn't seem too easy to show that the collection of Borel sets that satisfy $(*)$ is a σ-algebra (the problem being to show it's closed under taking complements). We can get around this by using the Monotone Class Lemma (page 68). Let $\mathcal{C}$ be the collection of Borel sets such that $(*)$ holds. We know that $\mathcal{A} \subseteq \mathcal{C}$. We claim that $\mathcal{C}$ is a monotone class. To see this, suppose that we have an increasing family $E_1 \subseteq E_2 \subseteq E_3 \subseteq \cdots$ of elements of $\mathcal{C}$ and $E = \bigcup_n E_n$. We want to show that $E \in \mathcal{C}$. We have

$$(**) \qquad |\mu_F(E_n)| \le \mu_{T_F}(E_n),$$

for each n. Since continuity from below works for complex measures, we have

$$\begin{aligned} \lim_{n\to\infty} \mu_F(E_n) &= \mu_F(E) \\ \lim_{n\to\infty} \mu_{T_F}(E_n) &= \mu_{T_F}(E), \end{aligned}$$

so letting n go to infinity in $(**)$ yields $|\mu_F(E)| \le \mu_{T_F}(E)$. Thus $E \in \mathcal{C}$. A similar argument shows that $\mathcal{C}$ is closed under countable decreasing intersections. Thus, $\mathcal{C}$ is a monotone class.

It $\mathcal{M}$ is the smallest monotone class containing $\mathcal{A}$, we must have $\mathcal{A} \subseteq \mathcal{M} \subseteq \mathcal{C} \subseteq \mathcal{B}_{\mathbb{R}}$. But, by the Monotone Class Lemma, the monotone class $\mathcal{M}$ generated by $\mathcal{A}$ is the same as the σ-algebra generated by $\mathcal{A}$, which is $\mathcal{B}_{\mathbb{R}}$. Thus, $\mathcal{C} = \mathcal{B}_{\mathbb{R}}$, so $(*)$ holds for all Borel sets.

We're now ready for the final step. By Exercise 21 (p. 94), if E is a Borel set, then

$$|\mu_F|(E) = \sup \left\{ \sum_{n=1}^{\infty} |\mu_F(E_n)| \mid E = \bigcup_{n=1}^{\infty} E_n,\ E_n\text{'s disjoint} \right\}.$$

So, suppose that E is the union of a disjoint family $\{ E_n \}$. We then have

$$\begin{aligned} \sum_{n=1}^{\infty} |\mu_F(E_n)| &\le \sum_{n=1}^{\infty} \mu_{T_F}(E_n) \\ &= \mu_{T_F}(E). \end{aligned}$$

Taking the sup over all Borel partitions of E, we get

$$|\mu_F|(E) \le \mu_{T_F}(E).$$

Since this holds for all Borel sets, we can set $E = (-\infty, x]$, and then

$$G(x) = |\mu_F|((-\infty, x]) \le \mu_{T_F}((-\infty, x]) = T_F(x) - T_F(-\infty) = T_F(x).$$

Thus, we have $G \le T_F$. Combined with our earlier result, this shows that $G = T_F$, and hence that $|\mu_F| = \mu_G = \mu_{T_F}$.

Problem 3. [**Problem 30, p. 107**] Construct an increasing function on $\mathbb{R}$ whose set of discontinuities is $\mathbb{Q}$.

Answer:
Lets start with a general construction. Let $A = \{ a_n \}_{n=1}^{\infty}$ be a countable set of distinct points in $\mathbb{R}$ and suppose that $\{ \alpha_n \}_{n=1}^{\infty}$ is a sequence of nonnegative numbers such that

$$\sum_{n=1}^{\infty} \alpha_n < \infty.$$

Define a function g_n by

$$g_n(x) = \begin{cases} 0, & x < a_n \\ \alpha_n, & a_n \le x. \end{cases}$$

Then we can define a function g by

$$g(x) = \sum_{n=1}^{\infty} g_n(x),$$

where the series converges for each x.

Claim. The function g is nondecreasing and is continuous at all points $x \notin A$.

Since each g_n is nondecreasing, it should be clear that g is nondecreasing. Suppose that $p \notin A$. We want to show that g is continuous at p. Let $\varepsilon > 0$ be given. There is some $N \in \mathbb{N}$ such that

$$\sum_{n=N+1}^{\infty} \alpha_n < \varepsilon.$$

Consider the finite list of points $a_1, a_2, \ldots, a_N$. Since p is not in this list, we can find a $\delta > 0$ so that none of the points in the list is in the interval $(p-\delta, p+\delta)$. Suppose that $p < x < p+\delta$. Then $g(x) - g(p)$ is the sum of the α_n's for which the corresponding point a_n is in the interval $(p, x]$. None of the points a_n for $n \leq N$ is in this interval, so

$$g(x) - g(p) = \sum_{a_n \in (p,x]} \alpha_n \leq \sum_{n=N+1}^{\infty} \alpha_n < \varepsilon.$$

Similarly, if $p - \delta < x < p$, then $g(p) - g(x)$ is the sum of the α_n's such that a_n is in $(x, p]$. By the same reasoning as above $g(p) - g(x) < \varepsilon$. This completes the proof of the claim.

We can now use this idea to complete the problem. Let $\{ r_n \}_{n=1}^{\infty}$ be an enumeration of the rationals. Define functions f_n by

$$f_n(x) = \begin{cases} 0, & x < r_n \\ \frac{1}{2^n}, & r_n \leq x. \end{cases}$$

and define f by

$$f(x) = \sum_{n=1}^{\infty} f_n(x).$$

By our previous work, f is nondecreasing. In fact f is strictly increasing since if $x < y$, there is some rational r_n so that $x < r_n < y$. But then $f(y) \geq f(x) + 1/2^n > f(x)$.

By our work above, f is continuous at the irrational points. It remains to prove that f is discontinuous at the rationals. To see this, let r_k be a rational point. Define a function h by

$$h(x) = \sum_{\substack{1 \leq n < \infty \\ n \neq k}} f_n(x).$$

By our claim above, h is continuous at r_k and we have $f(x) = f_k(x) + h(x)$. Then, $f(r_k+) = f_k(r_k+) + h(r_k+) = 1/2^k + h(r_k)$. Similarly, $f(r_k-) = 0 + h(r_k)$. Thus, $f(r_k+) - f(r_k-) = (1/2^k + h(r_k)) - h(r_k) = 1/2^k$, so f has a jump discontinuity at r_k.

We've now shown that the set of discontinuities of f is exactly $\mathbb{Q}$, and the solution is complete.

Problem 4. [**Problem 34, p. 108**] Suppose that $F, G \in \mathrm{NBV}$ and $-\infty < a < b < \infty$.

a. By adapting the proof of Theorem 3.36, show that

$$(*) \quad \int_{[a,b]} \frac{F(x)+F(x-)}{2}\, dG(x) + \int_{[a,b]} \frac{G(x)+G(x-)}{2}\, dF(x) = F(b)G(b) - F(a-)G(a-).$$

b. If there are no points in $[a,b]$ where both F and G are discontinuous, then

$$(**) \qquad \int_{[a,b]} F\, dG + \int_{[a,b]} G\, dF = F(b)G(b) - F(a-)G(a-).$$

Answer:
We will use Exercise 1.28, page 39 (which was assigned last semester). In particular, $\mu_H([r,s]) = H(s) - H(r-)$.

Define $\Omega \subseteq [a,b] \times [a,b]$ by

$$\Omega = \{\, (x,y) \mid a \le x \le y \le b \,\}.$$

We calculate $(\mu_F \times \mu_G)(\Omega)$ two ways, by the Fubini-Tonelli Theorem. First we have

$$\text{(A)} \quad \begin{aligned} (\mu_F \times \mu_G)(\Omega) &= \int_{[a,b]} \mu_F([a,y])\, d\mu_G(y) \\ &= \int_{[a,b]} [F(y) - F(a-)]\, dG(y) \\ &= \int_{[a,b]} F(y)\, dG(y) - F(a-)[G(b) - G(a-)]. \end{aligned}$$

Similarily, we have

$$\text{(B)} \quad \begin{aligned} (\mu_F \times \mu_G)(\Omega) &= \int_{[a,b]} \mu_G([x,b])\, d\mu_F(x) \\ &= \int_{[a,b]} [G(b) - G(x-)]\, dF(x) \\ &= G(b)[F(b) - F(a-)] - \int_{[a,b]} G(x-)\, dF(x). \end{aligned}$$

Change that name of the variable of integration in (A) to x and equate the right-hand sides of (A) and (B). After a little algebra, this yields,

$$\text{(C)} \quad \int_{[a,b]} F(x)\, dG(x) + \int_{[a,b]} G(x-)\, dF(x) = F(b)G(b) - F(a-)G(a-).$$

If we repeat this argument exchanging the roles of F and G (i.e., compute $(\mu_G \times \mu_F)(\Omega)$), we will get

$$\text{(D)} \qquad \int_{[a,b]} F(x-)\, dG(x) + \int_{[a,b]} G(x)\, dF(x) = F(b)G(b) - F(a-)G(a-).$$

If we add the equations (C) and (D) and divide by 2, we get $(*)$.

Now, for the second part of the problem, assume that there are no points where F and G are both discontinuous. If we let D_F be the set of points of disconinuity for F, and D_G the set of points of discontinuoity of G, we have $D_F \cap D_G = \emptyset$.

In general, we have $\mu_F(\{x\}) = F(x) - F(x-)$. Thus, $\mu_F(\{x\}) = 0$ if F is continuous at x. Since F is continuous at each point of D_G, each point in D_G has μ_F-measure zero. Since D_G is at most countable (because $G \in \text{NBV}$), we have $\mu_F(D_G) = 0$. We have

$$\frac{G(x) + G(x-)}{2} = \frac{G(x) + G(x)}{2} = G(x)$$

if $x \notin D_G$. Thus, we have

$$\frac{G(x) + G(x-)}{2} = G(x), \qquad \text{for } \mu_F\text{-almost all } x,$$

and so

$$\text{(E)} \qquad \int_{[a,b]} \frac{G(x) + G(x-)}{2}\, dF(x) = \int_{[a,b]} G(x)\, dF(x).$$

The same argument with the roles of F and G reveresed gives

$$\text{(F)} \qquad \int_{[a,b]} \frac{F(x) + F(x-)}{2}\, dG(x) = \int_{[a,b]} F(x)\, dG(x).$$

Subsituting (E) and (F) in $(*)$ gives us $(**)$.

Problem 5. [**Problem 35, p. 108**] If F and G are absolutely continuous on $[a, b]$, then so is FG and

$$\int_a^b (F'G + FG')(x)\, dx = F(b)G(b) - F(a)G(b).$$

Answer:
An absolutely continuous function is continuous and $[a, b]$ is compact, so F and G are bounded on $[a, b]$. Thus, we can find some constant M so that $|F| \leq M$ and $|G| \leq M$ on $[a, b]$. Let $H = FG$.

Let $\varepsilon > 0$ be given. Since F is absolutely continuous, there is some $\delta_1 > 0$ such that if $\{ (a_j, b_j) \}_{j=1}^n$ is a collection of disjoint intervals in $[a, b]$,

$$\sum_{j=1}^n (b_j - a_j) < \delta_1 \implies \sum_{j=1}^n |F(b_j) - F(a_j)| < \varepsilon.$$

Similarly, there is a $\delta_2 > 0$ such that

$$\sum_{j=1}^n (b_j - a_j) < \delta_2 \implies \sum_{j=1}^n |G(b_j) - G(a_j)| < \varepsilon.$$

Let $\delta = \min(\delta_1, \delta_2)$ and suppose that $\sum_j (b_j - a_j) < \delta$. Then we have

$$\begin{aligned}
\sum_{j=1}^n |H(b_j) - H(a_j)| &= \sum_{j=1}^n |F(b_j)G(b_j) - F(a_j)G(a_j)| \\
&= \sum_{j=1}^n |F(b_j)G(b_j) - F(b_j)G(a_j) + F(b_j)G(a_j) - F(a_j)G(a_j)| \\
&\le \sum_{j=1}^n |F(b_j)G(b_j) - F(b_j)G(a_j)| + \sum_{j=1}^n |F(b_j)G(a_j) - F(a_j)G(a_j)| \\
&= \sum_{j=1}^n |F(b_j)||G(b_j) - G(a_j)| + \sum_{j=1}^n |G(a_j)||F(b_j) - F(a_j)| \\
&\le M \sum_{j=1}^n |G(b_j) - G(a_j)| + M \sum_{j=1}^n |F(b_j) - F(a_j)| \\
&< M\varepsilon + M\varepsilon = 2M\varepsilon.
\end{aligned}$$

Since ε was arbitrary, we conclude that H is absolutely continuous.

Thus, H is differentiable almost everywhere. Except for an exceptional set of measure zero, the functions H, F and G are all differentiable. At a point x where they are all differentiable, we have $H'(x) = F'(x)G(x) + F(x)G'(x)$ by the ordinary product rule. Thus, $H' = F'G + FG'$ a.e. From the Fundamental Theorem of Calculus for Lebesgue Integrals, we have

$$H(b) - H(a) = \int_a^b H'(x)\, dx = \int_a^b (F'G + FG')(x)\, dx,$$

which completes the solution of the problem.

Problem 6. [**Problem 37, p. 108**] Suppose $F\colon \mathbb{R} \to \mathbb{C}$. There is a constant M such that $|F(x) - F(y)| \le M|x - y|$ for all $x, y \in \mathbb{R}$ (that is, F is Lipschitz continuous) iff F is absolutely continuous and $|F'| \le M$ a.e.

Answer:
Suppose first that F is Lipschitz, with $|F(x) - F(y)| \leq M|x - y|$. To prove that F is absolutely continuous, let $\varepsilon > 0$ be given. Choose $\delta > 0$ so small that $M\delta < \varepsilon$. Suppose that $\{ (a_j, b_j) \}_{j=1}^n$ is a finite collection of disjoint intervals so that

$$\sum_{j=1}^{n}(b_j - a_j) < \delta.$$

Then we have

$$\begin{aligned}\sum_{j=1}^{n}|F(b_j) - F(a_j)| &\leq \sum_{j=1}^{\infty} M|b_j - a_j| \\ &= M\sum_{j=1}^{n}(b_j - a_j) \\ &< M\delta < \varepsilon.\end{aligned}$$

Thus, F is absolutely continuous. Let p be a point where F is differentiable. Then

$$F'(p) = \lim_{x \to p} \frac{F(x) - F(p)}{x - p},$$

but

$$\left| \frac{F(x) - F(p)}{x - p} \right| \leq \frac{M|x - p|}{|x - p|} = M,$$

so $|F'(p)| \leq M$.

For the second part of the proof, suppose that F is absolutely continuous and $|F'| \leq M$ a.e. Suppose that $x < y$. By the Fundamental Theorem of Calculus for Lebesgue integrals,

$$F(y) - F(x) = \int_x^y F'\, dm,$$

and so

$$\begin{aligned}|F(x) - F(y)| &= \left| \int_x^y F'\, dm \right| \\ &\leq \int_x^y |F'|\, dm \\ &\leq \int_x^y M\, dm \\ &= M(y - x) \\ &= M|x - y|,\end{aligned}$$

so F is Lipschitz.

Problem 7. [**Problem 42, p. 109**] A function $F\colon (a,b) \to \mathbb{R}$ ($-\infty \le a < b \le \infty$) is called **convex** if

$$F(\lambda s + (1-\lambda)t) \le \lambda F(s) + (1-\lambda)F(t) \tag{$*$}$$

for all $s, t \in (a,b)$ and $\lambda \in (0,1)$ Actually, you can say all $\lambda \in [0,1]$]. ...

a. F is convex iff for all $s, t, s', t' \in (a,b)$ such that $s \le s' < t'$ and $s < t \le t'$,

$$\frac{F(t)-F(s)}{t-s} \le \frac{F(t')-F(s')}{t'-s'}. \tag{$**$}$$

b. F is convex iff F is absolutely continuous on every compact subinterval of (a,b) and F' is increasing (on the set where it is defined).

c. If F is convex and $t_0 \in (a,b)$, there exists $\beta \in \mathbb{R}$ such that $F(t) - F(t_0) \ge \beta(t-t_0)$ for all $t \in (a,b)$.

d. (**Jensen's Inequality**) If $(X, \mathcal{M}, \mu)$ is a measure space with $\mu(X) = 1$, $g\colon X \to (a,b)$ is in $L^1(\mu)$ and F is convex on (a,b), then

$$F\left(\int g\,d\mu\right) \le \int F \circ g\,d\mu.$$

Answer:
We begin with the first part of the problem by supposing that F is convex. We are given s, t, s', t' with $s \le s' < t'$ and $s < t \le t'$. In other words, $s', t \in [s, t']$ and no ordering between s' and t is specified, except that they can't be equal to the endpoint with the other letter. To derive ($**$), we proceed as follows.

Since $t \in [s, t']$ we can find a λ_1 so that

$$t = (1-\lambda_1)s + \lambda_2 t',$$

since $s \ne t$, we have $\lambda_1 \in (0,1]$. Solving for λ_1 gives

$$\lambda_1 = \frac{t-s}{t'-s}.$$

Now, by ($*$), we have

$$\begin{aligned} F(t) &= F((1-\lambda_1)s + \lambda_1 t') \\ &\le (1-\lambda_1)F(s) + \lambda_1 F(t') \\ &= F(s) - \lambda_1 F(s) + \lambda_1 F(t'), \end{aligned}$$

which yields

$$F(t) - F(s) \le \lambda_1 [F(t') - F(s)].$$

Plugging in the value of λ_1 yields

$$F(t) - F(s) = \frac{t-s}{t'-s}[F(t') - F(s)].$$

Dividing both sides by the postive number $t - s$, we get

$$\text{(A)} \qquad \frac{F(t) - F(s)}{t - s} \le \frac{F(t') - F(s)}{t' - s}$$

Similarly, we can write

$$s' = \lambda_2 s + (1 - \lambda_2)t'$$

for some $\lambda_2 \in (0, 1]$. A little algebra yields

$$\lambda_2 = \frac{s' - t'}{s - t'}.$$

Thus, we have

$$\begin{aligned} F(s') &= F(\lambda_2 s + (1 - \lambda_2)t') \\ &\le \lambda_2 F(s) + (1 - \lambda_2)F(t') \\ &= \lambda_2 F(s) + F(t') - \lambda_2 F(t'), \end{aligned}$$

and so

$$F(s') - F(t') \le \lambda_2[F(s) - F(t')].$$

Putting in the value of λ_2 gives

$$F(s') - F(t') \le \frac{s' - t'}{s - t}[F(s) - F(t')].$$

Dividing both sides of this inequality by the *negative* number $s' - t'$ gives

$$\frac{F(s') - F(t')}{s' - t'} \ge \frac{F(s) - F(t')}{s - t'}.$$

Using the symmetry of the difference quotients, we can rewrite this as

$$\text{(B)} \qquad \frac{F(t') - F(s)}{t' - s} \le \frac{F(t') - F(s')}{t' - s'}.$$

Combining (A) and (B) yields $(**)$.

To see the geometric significance of the inequalities set

$$\begin{aligned} m_1 &= \frac{F(t) - F(s)}{t - s} \\ m_2 &= \frac{F(t') - F(s)}{t' - s} \\ m_3 &= \frac{F(t') - F(s')}{t' - s'}, \end{aligned}$$

so we proven that $m_1 \le m_2 \le m_3$, and consider Figure 1

Conversely, suppose that $(**)$ holds. Let $s < t'$ be two points in (a, b) and let $\lambda \in (0, 1)$ be given. Apply $(**)$ with

$$t = s' = \lambda s + (1 - \lambda)t'.$$

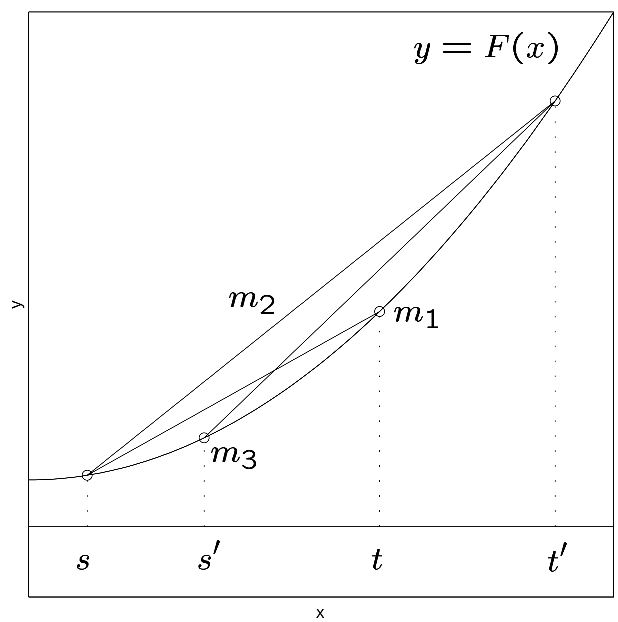

Figure 1: Slope inequalities for a convex function

A little algebra yields

$$t - s = (1-\lambda)(t' - s)$$
$$t' - s' = \lambda(t' - s),$$

and so (∗∗) becomes

$$\frac{F(t) - F(s)}{(1-\lambda)(t' - s)} \leq \frac{F(t') - F(t)}{\lambda(t' - s)}.$$

Since λ, $1 - \lambda$ and $t' - s$ are positive, cross multiplication gives

$$\lambda[F(t) - F(s)] \leq (1-\lambda)[F(t') - F(t)]$$

which simplifies easily to

$$F(t) \leq \lambda F(s) + (1-\lambda)F(t').$$

Thus, F is convex on (a, b).

Next, we do the second part of the problem. First, suppose that F is convex on (a, b); we want to show that F is absolutely continuous on every compact subinterval of (a, b) and F' is increasing.

Let $[q, r]$ be a compact subinterval of (a, b). We can choose points p and s with $a < p < q < r < s < b$. Suppose that $x, y \in [q, r]$ and $x < y$. With two applications of (∗∗) (First: $s = p$, $t = q$, $s' = x$, $t' = y$; Second: $s = x$, $t = y$, $s' = r$, $t' = s$) we get

$$\frac{F(q) - F(p)}{q - p} \leq \frac{F(y) - F(x)}{y - x} \leq \frac{F(s) - F(r)}{s - r}.$$

Since the two outer quotients don't involve x and y, we conclude that there are constants m_1 and m_2 so that

$$m_1 \leq \frac{F(y) - F(x)}{y - x} \leq m_2$$

for all $x, y \in [q, r]$. If we set $M = \max(|m_1|, |m_2|)$ then we have

$$\left| \frac{F(y) - F(x)}{y - x} \right| \leq M$$

for all $x, y \in [q, r]$. Thus, we have

$$|F(y) - F(x)| \leq M|x - y| \qquad \forall x, y \in [q, r].$$

Thus, F is Lipschitz on $[q, r]$ and so absolutely continuous. Next, we want to show that the derivatives is increasing. To see this, let $s < t'$ be two points at which the derivative exits. From (∗∗) we have

$$\frac{F(t) - F(s)}{t - s} \leq \frac{F(t') - F(s')}{t' - s'}$$

where $s \le s' < t'$ and $s < t \le t'$. If we hold s' fixed and take the limit as $t \downarrow s$, we get

$$F'(s) \le \frac{F(t') - F(s')}{t' - s'}$$

and taking the limit of this as $s' \uparrow t'$ gives $F'(s) \le F'(t')$. Thus, F' is increasing.

Next, we want to do the converse. Thus, we assume that F is absolutely continuous on every compact subinterval of (a, b) and that F' is increasing (on the set where it is defined).

We need the following Lemma.

Lemma. *Let $G\colon [c, d] \to \mathbb{C}$ be absolutely continuous. Let $\varphi\colon \mathbb{R} \to \mathbb{R}\colon x \mapsto \alpha x + \beta$, where $\alpha > 0$. Suppose that φ maps $[a, b]$ onto $[c, d]$. Then $G \circ \varphi$ is absolutely continuous on $[a, b]$ and $(G \circ \varphi)'(x) = \alpha G'(\alpha x + \beta)$ a.e.*

Proof of Lemma. We want to prove that $G \circ \varphi$ is absolutely continuous. Let $\varepsilon > 0$ be given. Then there is a $\delta > 0$ so that if $\{ (c_j, d_j) \}_{j=1}^n$ is a collection of intervals in $[c, d]$, then

$$\text{(C)} \qquad \sum_{j=1}^n (d_j - c_j) < \delta \implies \sum_{j=1}^n |G(d_j) - G(c_j)| < \varepsilon.$$

Choose $\eta > 0$ so small that $\alpha\eta < \delta$. Let $\{ (a_j, b_j) \}_{j=1}^n$ be a collection of disjoint intervals in $[a, b]$. The map φ maps the interval $[a_j, b_j]$ onto the interval $[\varphi(a_j), \varphi(b_j)]$ which has length $\alpha(b_j - a_j)$. Suppose that

$$\sum_{j=1}^n (b_j - a_j) < \eta$$

Then $\{ (\varphi(a_j), \varphi(b_j)) \}$ is a collection of disjoint intervals in $[c, d]$ and

$$\sum_{j=1}^n (\varphi(b_j) - \varphi(a_j)) = \alpha \sum_{j=1}^n (b_j - a_j) < \alpha\eta < \delta.$$

Thus, from (C), we have

$$\sum_{j=1}^n |G(\varphi(b_j)) - G(\varphi(a_j))| < \varepsilon.$$

This shows that $G \circ \varphi$ is absolutely continuous. The formula for the derivatives follows from the usual chain rule at points x such that G is differentiable at $\varphi(x)$. □

To apply this Lemma to our situation, suppose that $x < y$, fix $\lambda \in (0, 1)$ and let $p = (1 - \lambda)x + \lambda y$. Consider the function $g(t) = F(x + t\lambda(y - x))$. By our

Lemma, this is absolutely continuous, and $g'(t) = F'(x + t\lambda(y - x))\lambda(y - x)$. By the Fundamental Theorem of Calculus for Lebesgue integrals, we have

$$g(1) - g(0) = \int_0^1 g'(t)\, dt,$$

which we can write as

$$F(p) - F(x) = \int_0^1 F'(x + t\lambda(y - x))\lambda(y - x)\, dt.$$

On the other hand, consider $h(t) = F(x+t(y-x))$ and apply the same reasoning. This gives

$$F(y) - F(x) = \int_0^1 F'(x + t(y - x))(y - x)\, dt.$$

Now, for $t \in [0, 1]$, we have $x + t\lambda(y - x) \leq x + t(y - x)$, since $0 < \lambda < 1$. Since F' is increasing, we have $F'(x + t\lambda(y - x)) \leq F'(x + t(y - x))$ for almost all t. Thus,

$$\begin{aligned} F(p) - F(x) &= \lambda \int_0^1 F'(x + t\lambda(y - x))(y - x)\, dt \\ &\leq \lambda \int_0^1 F'(x + t(y - x))(y - x)\, dt \\ &= \lambda[F(y) - F(x)]. \end{aligned}$$

Now we have

$$F(p) - F(x) \leq \lambda[F(y) - F(x)],$$

which is easily rearranged to give

$$F(p) \leq (1 - \lambda)F(x) + \lambda F(y).$$

This shows that $(*)$ holds, so F is convex.

Next, consider the third part of the problem. Thus, we assume that F is convex on (a, b). Let $t_0 \in (a, b)$. Let s and t' be points such that $s < t_0 < t'$ ($s, t' \in (a, b)$). Apply $(**)$ with $s' = t = t_0$. This gives

$$\text{(D)} \qquad \frac{F(t_0) - F(s)}{t_0 - s} \leq \frac{F(t') - F(t_0)}{t' - t_0}.$$

The set

$$\left\{ \frac{F(t_0) - F(s)}{t_0 - s} \mid a < s < t_0 \right\}$$

is bounded above by the right-hand side of (D) for any $t' > t_0$. Thus, if we define

$$\beta_1 = \sup \left\{ \frac{F(t_0) - F(s)}{t_0 - s} \mid a < s < t_0 \right\}$$

We have

$$\beta_1 \leq \frac{F(t') - F(t_0)}{t' - t_0}$$

for all $t' > t_0$. Thus, if we set

$$\beta_2 = \inf \left\{ \frac{F(t') - F(t_0)}{t' - t_0} \mid t_0 < t' < b \right\},$$

we have $\beta_1 \leq \beta_2$. Choose β so that $\beta_1 \leq \beta \leq \beta_2$.

If $t > t_0$, then the quotient

$$\frac{F(t) - F(t_0)}{t - t_0} \tag{E}$$

is an element of the set in the definition of β_2, so we have

$$\beta \leq \frac{F(t) - F(t_0)}{t - t_0},$$

which implies

$$\beta(t - t_0) \leq F(t) - F(t_0).$$

On the other hand, if $t < t_0$, then the quotient (E) is an element of the set defining β_1, so we have

$$\beta \geq \frac{F(t) - F(t_0)}{t - t_0}.$$

which gives

$$\beta(t - t_0) \leq F(t) - F(t_0),$$

since $t - t_0 < 0$. This completes the solution of the third part of the problem.

Finally, we prove Jensen's inequality. Thus we assume that F is convex on (a, b), $g\colon X \to (a, b)$ is in $L^1(\mu)$ and $\mu(X) = 1$. Since $a < g(x) < b$ for all x, we can integrate this to get

$$a = a\mu(X) = \int a \, d\mu \leq \int g \, d\mu \leq \int b \, d\mu = b\mu(X) = b.$$

Thus, if we set $t_0 = \int g \, d\mu$, we have $t_0 \in (a, b)$. From the previous part of the problem, there is some β so that

$$F(t) - F(t_0) \geq \beta(t - t_0).$$

Since $g(x) \in (a, b)$, we can set $t = g(x)$ in this inequality to get

$$F(g(x)) - F(t_0) \geq \beta(g(x) - t_0).$$

Integrate this inequality, using $\mu(X) = 1$. This gives

$$\int F(g(x)) \, d\mu(x) - F(t_0) \geq \beta \left[\int g(x) \, d\mu(x) - t_0 \right] = 0,$$

since $t_0 = \int g\,d\mu$. Thus, we have

$$F\left(\int g\,d\mu\right) = F(t_0) \le \int F(g(x))\,d\mu(x),$$

which is Jensen's inequality and we're done. Well, almost done. Why is $F \circ g$ measurable? Does the integral $\int F \circ g\,d\mu$ make sense?

Remark. Some people tried to do Exercise 37 and Exercise 42 using the Mean Value Theorem. This would be exactly the right thing to do *if* our functions were differentiable. Unfortunately, the Mean Value Theorem does not work for absolutely continuous functions that are only differentiable almost everywhere. For a counterexample, consider $F\colon \mathbb{R} \to \mathbb{R}\colon x \mapsto |x|$. By the triangle inequality we have

$$|F(x) - F(y)| = ||x| - |y|| \le |x - y|$$

so F is Lipschitz, and hence absolutely continuous. It's differentiable except at 0, and hence differentiable *almost* everywhere, but not everywhere. Consider the cord of the graph that connects the two points $(-1, 1)$ and $(1, 1)$. The slope of this cord is zero, but there is no point on the graph where the derivative is zero (or even anywhere close to zero). Thus, it is *not* true that

$$F(1) - F(-1) = F'(c)(1 - (-1))$$

for some point c in $(-1, 1)$. The workaround for the lack of the Mean Value Theorem is to use the Fundamental Theorem of Calculus for Lebesgue Integrals instead.