MAT320 Final Practice Problems Answers

12/16/2023

Below are answers to the practice problems.

Problem 1. Consider the functions

$$f_n(x) = \frac{1}{1 + |x|^{\frac{n^3}{n^2+2024}}}$$

for $n = 100, 101, 102, \ldots$ Find

$$\lim_{n\to\infty} \int_0^\infty f_n(x)dx$$

where we take the limit over all $n > 100$. You may use the fact that

$$\frac{1}{1 + |x|^N}$$

is integrable over $\mathbb{R}$ for all $N \geq 2$.

Proof. We want to exchange the limit with the integral. First, let's assume that we can. We may assume everywhere that $x \geq 0$. We have that

$$g(x) := \lim_{n\to\infty} f_n(x) = \begin{cases} 1 & 0 \leq x < 1 \\ \frac{1}{2} & x = 1 \\ 0 & x > 1. \end{cases}$$

Since the Lebesgue integral is unchanged upon modifying the integrand arbitrarily on a measure zero set,

$$\int_0^\infty g(x)dx = \int_0^\infty \chi_{[0,1]} = \int_0^1 1 = 1.$$

Here $\chi_{[0,1]}$ is the indicator function of $[0, 1]$, and $g - \chi_{[0,1]} = 0$ almost everywhere.

Now we need to argue that

$$\lim_{n\to\infty} \int_0^\infty f_n(x) = \int_0^\infty \lim_{n\to\infty} f_n(x) = \int_0^\infty g(x) = 1.$$

To see this we use Lebesgue dominated convergence. We have that $1/(1 + |x|^2)$ is integrable over $\mathbb{R}$. Now for all $n \geq 100$, we have that

$$f_n(x) \leq 1/(1 + |x|^2) \text{ if } x \geq 1.$$

The inequality actually goes the other way if $0 < x < 1$, but that doesn't mater, because we also have that

$$f_n(x) \leq 1 \text{ if } 0 \leq x \leq 1.$$

Thus choosing

$$h(x) = \begin{cases} 1 & \text{if } 0 \leq x < 1/2 \\ \frac{1}{1+|x|^2} & \text{if } x > 1 \end{cases}$$

we see that h is integrable over $\mathbb{R}$ and $|f_n(x) \leq h(x)$ for all $x \in [0, \infty)$ and all $n = 100, 101, \ldots$; thus Lebesgue dominated convergence applies with dominating function h, and we can exchange the limit and the integral. $\square$

Problem 2. Given a closed interval $[a, b] \subset \mathbb{R}$ with $b \geq a$ we write ℓ for its length:

$$\ell([a, b]) = (b - a).$$

If $b < a$ we set $\ell([a, b]) = \ell(\emptyset) = 0$.

For any $\alpha > 0$, and any subset $A \subset \mathbb{R}$, define

$$\mu_\alpha^*(A) = \liminf_{\delta \to 0^+} \left\{ \sum_{i=1}^{\infty} \ell([a_i, b_i])^\alpha \mid [a_i, b_i] \subset \mathbb{R} \text{ for each } i = 1, \ldots \infty, A \subset \bigcup_{i=1}^{\infty} [a_i, b_i], \ell([a_i, b_i]) < \delta \text{ for all } i = 1, \ldots, \infty \right\}.$$

Here the lim inf is taken as δ approaches zero from above.

a) Show that if $A \subset B \subset \mathbb{R}$ then $\mu_\alpha^*(A) \leq \mu_\alpha^*(B)$.

b) Show that when $\alpha = 1$ then $\mu_\alpha^*(A)$ agrees with the usual outer measure of A.

c) Show that if $\alpha > 1$ then $\mu_\alpha^*(A) = 0$ for any $A \subset \mathbb{R}$.

d) (This part is harder.) Show that for any collection of sets $B_i \subset \mathbb{R}$, $i = 1, \ldots, \infty$, we have

$$\mu_\alpha^* \left(\bigcup_{i=1}^{\infty} B_i \right) \leq \sum_{i=1}^{\infty} \mu_\alpha^*(B_i).$$

Thus, μ_α^* generalizes the outer measure.

Proof. a) We have that $\mu^*\alpha(A) = \liminf_{\delta \to 0} S_\delta^\alpha(A)$ where S_δ^α is the set of real numbers in the definition of $\mu_\alpha^*(A)$. If $A \subset B \subset \mathbb{R}$ then clearly $S_\delta^\alpha(A) \supset S_\delta^\alpha(B)$, which immediately implies that $\inf S_\delta^\alpha(A) \leq \inf S_\delta^\alpha(B)$ for every δ and thus the claim.

b) When $\alpha = 1$ the definitions are very similar: we have that

$$\mu^*(A) = \inf S(A), S(A) = \{\sum_{i=1}^{\infty} \ell([a_i, b_i]) | [a_i, b_i] \subset \mathbb{R}, A \subset \cup_{i=1}^{\infty} [a_i, b_i].\}$$

(Technically the book uses open rather than closed intervals, but we discussed in class why one can switch the open interval definition with the closed interval definition.) Thus the only difference between $S_\delta^{\alpha=1}(A)$ and $S(A)$ is the extra condition for the former that each of the intervals involved has length at most δ; so $S(A) \supset S_\delta^{\alpha=1}$. But on the other hand, choosing any δ' such that $0 < \delta' < \delta$, given any interval $[a_i, b_i]$ then we can write it as a finite union

$$[a_i, b_i] = \cup_{j=-\infty}^{\infty} [a_i, b_i] \cap [j/\delta', (j+1)/\delta']$$

of intervals of length less than δ. this is a countable union a priori, but all but finitely many terms of the union above are empty since since the interval $[a_i, b_i]$ is finite. Moreover we have that

$$\ell([a_i, b_i]) = \sum_{j=-\infty}^{\infty} \ell([a_i, b_i] \cap [j/\delta', (j+1)/\delta'])$$

where again this is actually a finite sum. Since a countable union of finite sets is still countable we thus have that $S(A) \subset S_\delta^{\alpha=1}$ for any δ as well. Thus these sets of real numbers are actually always equal to $S(A)$ and are independent of δ, which proves the claim.

c) We use the same trick as before. Choose any countable collection

$$[a_i, b_i] \subset \mathbb{R} \text{ such that } A \subset \cup_{i=1}^{\infty} [a_i, b_i].$$

First, we use the subdivision trick to assume that $\ell([a_i, b_i]) < \delta$ for every i (this is just the statement that $S_\delta^{\alpha=1}(A) = S(A)$ above). Then for any natural number N we can write

$$[a_i, b_i] = [a_i^1, b_i^1] \cup [a_i^2, b_i^2] \cup \dots [a_i^N, b_i^N]$$

as a finite union of N intervals (so the quantities $a_i^{r_i}, b_i^{r_i}$ implicitly depend on N as well) which only overlap at their endpoints and such that the length of each interval is $\ell([a_i, b_i])/N$. We then have that

$$\ell([a_i, b_i])^\alpha \geq \sum_{j=1}^{N} \ell([a_i^j, b_i^j]) = N \left(\frac{\ell([a_i, b_i]}{N} \right)^\alpha = N^{1-\alpha} \ell([a_i, b_i]^\alpha.$$

for each i and N. But then by using the subdivisions of the fixed cover of A by the $\{[a_i, b_i]\}$ described above we conclude that that

$$\inf S_\delta^\alpha(A) \leq \liminf_{N\to\infty} \sum_i \sum_{j=1}^{N} \ell([a_i^j, b_i^j])^\alpha \leq \lim_{N\to\infty} \sum_i N^{1-\alpha} \ell([a_i, b_i])^\alpha = 0.$$

But this worked for arbitrary $\delta > 0$ and A, so we conclude that $\mu^*(A) = 0$.

$\square$

Problem 3. Consider the function on the reals that computes the shortest distance from x to any integer:

$$D : \mathbb{R} \to \mathbb{R}, D(x) = d(x, \mathbb{Z}) = \min_{n \in \mathbb{Z}} |x - n|.$$

Given a pair of sequences of nonnegative real numbers

$$\{a_n\}_{n=1}^{\infty}, \{b_n\}_{n=1}^{\infty},$$

we can consider the partial sums

$$f_N(x) = \sum_{n=1}^{N} a_n D(b_n x)$$

as functions $f_N : [0, 1] \to \mathbb{R}$ and see if there is a well-defined limit

$$f(x) = \sum_{n=1}^{\infty} a_n D(b_n x)$$

in some sense.

a) Show that if we set $a_n = (1/3)^n$ then for any sequence b_n the functions $f_N(x) \in C^0([0, 1])$ are a Cauchy sequence in the C^0 norm. As such, in this case the limit function f is a well defined element of $C^0([0, 1])$.

b) Show that if we set $a_n = (1/3)^n$ and $b_n = 2^n$ then the limit function $f : [0, 1] \to \mathbb{R}$ is Lipshitz.

For either of these questions it may be useful to recall that for any $0 < x < 1$ we have that the following sum converges:

$$\sum_{n=0}^{\infty} x^n = \frac{1}{1 - x}.$$

Proof. 1. Notice that $\|D\|_{C^0([a,b])} \leq \frac{1}{2}$ since the maximum distance from any real to the nearest integer is 1/2. But then $\|D(b_n x)_{C^0([0,1])} \leq 1/2$ for any $b_n > 0$ since we are just rescaling the x axis. For $N > M$, we have that

$$\|f_N - f_M\|_{C^0} = \| \sum_{n=M+1}^{N} a_n D(b_n x)\|_{C^0} \leq \sum_{n=M+1}^{N} |a_n| \|D(b_n)x\|_{C^0} = \sum_{n=M+1}^{N} |a_n| \leq \sum_{n=M+1}^{\infty} (1/3)^n \xrightarrow{M \to \infty} 0$$

Here the second inequality is the triangle inequality; the final observation proves that the original sequence was Cauchy.

2. This is similar but we need to estimate the slopes. The Lipshitz constant of D is 1, because D is a 1-periodic function and $D|_{[-1/2,1/2]}(x) = |x|$. Thus the lipshitz constant of $a_n D(b_n x)$ is $a_n b_n$ whenever $a_n, b_n > 0$ (because a_n rescales vertically and b_n rescales horizontally). Thus

$$|f_N(x) - f_N(y)| \leq \sum_{n=1}^{N} |a_n D(b_n x) - a_n D(b_n y)| \leq \sum_{n=1}^{N} a_n b_n |x - y| = \sum_{n=1}^{N} \left(\frac{2}{3}\right)^{\alpha} |x - y| \leq \frac{(2/3)}{1 - (2/3)} |x - y|.$$

Now we take the limit on the left hand side as $N \to \infty$ and conclude that the Lipshitz constant of f is finite since it is bounded by

$$\frac{(2/3)}{1 - (2/3)} = 2.$$

Notice that if we let b_n grow faster then in fact the Lipshitz constant of f would diverge. In fact, this is a famous example which caused mathematicians a tremendous amount of stress when it was first discovered – this is a variant of the Weierstrass function (`https://en.wikipedia.org/wiki/Weierstrass_function`). If a_n and b_n are as in (b) then f is Lipshitz and thus has a derivative almost everywhere; but if b_n grows too fast, e.g. $b_n = 4^n$, then f is continuous but it is not differentiable at *any point*. The plots of functions such as f are very pretty fractals, and I suggest taking a look at the Wikipedia page for more information.

(The wikipedia page uses a construction with cosines, which is a little bit harder to work with than the construction I described.)

One can go further with this story – by choosing the a_n to be random variables $\pm(1/3)^n$ and choosing b_n appropriately, one gets that f is a random function which describes a 1-periodic version of Brownian motion! $\square$

Problem 4. Recall the Banach space

$$\ell^2 = \{(a_n)_{n=1}^{\infty} \mid \sum_{n=1}^{\infty} |a_n|^2 < \infty\},$$

$$\|(a_n)\|_{\ell^2} = \sqrt{\sum_{n=1}^{\infty} |a_n|^2}.$$

Consider the subset

$$\mathcal{H} \subset \ell^2, \mathcal{H} = \{((a_n)_{n=1}^{\infty} | a_n \in [0, 1/n]\}.$$

a) Show that $\mathcal{H}$ does *not* contain any open ℓ^2 ball of positive radius, i.e. that $\mathcal{H}$ has empty interior.

b) Show that every sequence in $\mathcal{H}$ has a subsequence that is Cauchy in ℓ^2. You may use without proof the fact that

$$\sum_{n=1}^{\infty} \frac{1}{n^2} < \infty.$$

Proof. 1. Suppose $B(x, \epsilon) \subset \mathcal{H}$. Then $x = (x_n)_{n=1}^{\infty}$ and $x_n \in [0, 1/n]$ for each n. Pick n such that $\epsilon/2 > 1/n$. Write

$$e_n = (0, \ldots, 1, \ldots 0)$$

where the 1 is in the n-th place. Then $\epsilon/2e_n \in B(0, \epsilon)$ since e_n has length 1 with respect to the ℓ^2 metric, so $\epsilon/2e_n$ has length $\epsilon/2$. But then $x + \epsilon/2e_n \in B(x, \epsilon)$. However, the n-th coordinate of $x + \epsilon$ is $x_n + \epsilon/2 > 1/n$, which contradicts the statement that $x + \epsilon \in \mathcal{H}$ which is needed for $\mathcal{H}$ to contain $B(x, \epsilon)$.

2. Let x^k be a sequence of vectors in $\mathcal{H}$; then

$$x^k = (x_1^k, x_2^k, \ldots).$$

(We are just writing out the coordinates of each vector x^k here.) Each x_j^k, thought of as a sequence in k with fixed j, is a sequence of real numbers in $[0, 1/j]$; so there is subsequence $x_j^{k_{a_i^j}}$ where $a_1^j < a_2^j < \ldots$ is an increasing sequence of indices, such that $\lim_{i\to\infty} x_j^{k_{a_i^j}} = x_j$ for some real number $x_j \in [0, 1/j]$. Set $x^{k'} = x^{k_{a_j^j}}$ by using the Cantor diagonal argument. Let us subsequently use k to denote k', since we will be working to prove that this subsequence converges. Set $x = (x_1, \ldots)$ to be the corresponding vector in $\mathcal{H}$ with coordinates x_j. We claim that $x^k \to x$ in ℓ^2. We have that for each natural number R, the finite dimensional vectors

$$\bar{x}^{k,R} = (x_1^k, \ldots, x_R^k) \in \mathbb{R}^R$$

converge in $\mathbb{R}^R$ to

$$\bar{x} = (x_1, \ldots, x^R)$$

because each of the coordinates converge. Thus, for any $\epsilon > 0$, choose R such that

$$\sum_{n=R+1}^{\infty} \frac{1}{n^2} < \epsilon/2,$$

and then N such that $d_{\mathbb{R}^R}(\bar{x}^{k,R}, \bar{x}^R) < \epsilon/2$ for all $k > N$. (Here we use the Euclidean distance on $\mathbb{R}^R$.) We then have that for all $k > N$, We write

$$x^k = x^{k,R} + (x^{k,R} - x^{k,R})$$

where

$$x^{k,R} = (x_1^k, \ldots, x_R^k, 0, \ldots, 0 \ldots)$$

and thus $(x^k - x^{k,R})$ is just the zeroing out of x^k in the first R coordinates. Similarly we write

$$x = x^R + (x - x^R).$$

Then for $k > N$ we have that

$$\|x^k - x\|_{\ell^2} \leq \|(x^{k,R} - x^R)\|_{\ell^2} + \|(x^k - x^{k,R}) - (x - x^R)\|_{\ell_2}.$$

We have that

$$\|(x^{k,R} - x^R)\|_{\ell^2} = \|\bar{x}^{k,R} - \bar{x}^R\|_{\mathbb{R}^R}$$

so this is bounded above by $\epsilon/2$. But $(x^k - x^{k,R}) - (x - x^R)$ has its j-th coordinate lying in $[-1/j, 1/j]$ with its first R coordinates equal to zero, so

$$\|(x^k - x^{k,R}) - (x - x^R)\| \leq \sum_{n=R+1}^{\infty} \frac{1}{n^2} < \epsilon/2,$$

so

$$\|x^k - x\|_{\ell^2} < \epsilon$$

for $k > N$. We are thus done – $x^k \to x$ in ℓ^2, so x^k is a Cauchy sequence in ℓ^2, but x^k was really a subsequence $x^{k'}$ of an original arbitrary sequence $x^k \in \mathcal{H}$.

$\square$

Problem 5. Let Σ be a σ-algebra on a set X, and let $\mu : \Sigma \to \mathbb{R}$ be a countably additive measure such that $\mu(X) = 1$. (In other words, μ is a *probability measure* on X.)

Suppose that $f_n : X \to \mathbb{R}$ is a measurable function for each $n = 1, 2, \ldots$, i.e. such that $f_n^{-1}([a, b]) \in \Sigma$ for all $n \in \mathbb{N}$ and all $a, b \in \mathbb{R}$. Say that there is another measurable function $f : X \to \mathbb{R}$ as well.

Suppose that for every $\epsilon > 0$,

$$\mu(\{x \in X || f_n(x) - f(x)| > \epsilon\}) \to 0 \text{ as } n \to \infty.$$

(You do not have to prove that the sets $\{x \in X || f_n(x) - f(x)| > \epsilon\}$ are in Σ.)

a) Show that for every $\delta_1, \delta_2 > 0$, there exists an n such that for all $n > N$,

$$\mu(f_n^{-1}([a, b])) \leq \mu(f^{-1}([a - \delta_1, b + \delta_1])) + \delta_2.$$

b) Show that for every $[a, b] \in \mathbb{R}$,

$$\lim_{n\to\infty} \mu(f_n^{-1}([a, b])) = \mu(f^{-1}([a, b])).$$

You may use without proof the continuity of measure, i.e. that for any sequence of sets $A_i \in \Sigma$, we have

$$\mu\left(\bigcup_{i=1}^{\infty} A_i\right) = \lim_{n\to\infty} \mu\left(\bigcup_{i=1}^{n} A_i\right),$$

$$\mu\left(\bigcap_{i=1}^{\infty} A_i\right) = \lim_{n\to\infty} \mu\left(\bigcap_{i=1}^{n} A_i\right),$$

Proof. a) Suppose that $|f_n(x) - f(x)| < \epsilon$ for all x; then if $f_n(x) \in [a, b]$ then $f(x) \in [a - \epsilon, b + \epsilon]$, so $(f_n^{-1}([a, b]))(f^{-1}([a - \epsilon, b + \epsilon])$ and so the measure of the left hand side is at most the measure of the right hand side in this inequality. Of course, what we just said works when we assume that f and f_n are defined on an arbitrary measurable set. Now, setting $\epsilon = \delta_1$, choose N such that, writing

$$E_n = \{x \in X || f_n(x) - f(x)| > \epsilon\}$$

we have

$$\mu(E_n) < \delta_2.$$

Then for $n > N$,

$$\mu(f_n^{-1}([a, b])) = \mu(f_n^{-1}([a, b]) \cap E_n^c) + \mu(f_n^{-1}([a, b]) \cap E_n)$$
$$\leq \mu(f^{-1}([a - \delta_1, b + \delta_1]) \cap E_n^c) + \mu(E_n) \leq \mu(f^{-1}([a - \delta_1, b + \delta_1]) + \delta_2.$$

b) The previous section gives that

$$\limsup_{n\to\infty} \mu(f_n^{-1}([a, b]) \leq \mu(f^{-1}([a - \delta_1, b + \delta_1]) + \delta_2.$$

Here $\delta_1, \delta_2 > 0$ are arbitrary; thus the same holds when $\delta_2 = 0$. Continuity of measure then lets us conclude the same when $\delta_1 = 0$ because

$$f^{-1}([a, b]) = \bigcap_{\delta_1 > 0} f^{-1}([a - \delta_1, b + \delta_1]).$$

We now want to show that

$$\liminf_{n\to\infty} \mu(f_n^{-1}([a, b])) \geq \mu(f^{-1}([a, b]))$$

to conclude the statement. To do this, we re-run the argument for (a) but exchanging the roles of f and f_n; then for every $\delta_1, \delta_2 > 0$ there exists an n such that for all $n > N$,

$$\mu(f^{-1}([a, b])) \leq f_n^{-1}([a - \delta_1, b + \delta_1]) + \delta_2. \tag{1}$$

Rearranging and taking $\liminf$,

$$\liminf_{n\to\infty} \mu(f_n^{-1}([a-\delta_1, b+\delta_1])) \geq \mu(f^{-1}([a,b])) - \delta_2.$$

So we can set $\delta_2 = 0$ in the above as well. But now a, b, δ_1 were arbitrary, so choosing δ_1 small enough (e.g. $\delta_1 < (b-a)/4$ and setting $a' = a - \delta_1, b' = b + \delta_1$ we conclude that

$$\liminf \mu(f_n^{-1}([a', b'])) \geq \mu(f^{-1}([a'+\delta_1, b'-\delta_1])).$$

Continuity of measure then lets us replace $\mu(f^{-1}([a'+\delta_1, b'-\delta_1]))$ with

$$\mu\left(f^{-1}\left(\bigcup_{\delta_1>0}[a'+\delta_1, b'-\delta_1]\right)\right) = \mu(f^{-1}((a', b')).$$

So then we are done unless $\mu(f^{-1}(a)) \neq 0$ or $\mu(f^{-1}(b)) \neq 0$. To avoid this annoying issue, we notice that in (1), this would have worked if we had replaced δ_2 with $\delta_2/2$; and then using continuity of measure for f we could have concluded that for for every δ_1, δ_2 there is an N and a δ' such that

$$\mu(f^{-1}([a-\delta', b+\delta']) \leq \mu(f_n^{-1}([a-\delta_1, b+\delta_1])) + \delta_2.$$

Then as before, taking $\liminf_{n\to\infty}$, rewriting, rewriting in terms of $a' = a - \delta_1, b' = b + \delta_1$, and using continuity of measure first δ_1 and then in δ' we conclude that in fact

$$\liminf_{n\to\infty} f_n^{-1}([a', b']) \geq \mu\left(f^{-1}\left(\bigcup_{\delta'>0}[a'-\delta_1, b'+\delta']\right)\right) = \mu(f^{-1}([a', b']))$$

which lets us conclude since a', b' were indeed arbitrary.

The argument above for (b) is not maximally efficient, but hopefully it shows how one might come up with a correct argument via iteration of ideas. What you have proven in this problem is that convergence in probability implies convergence in distribution; for more details see (`https://en.wikipedia.org/wiki/Convergence_of_random_variables`). $\square$

Problem 6. Suppose that we have a sequence of measurable functions $K_n : \mathbb{R} \to \mathbb{R}$ such that for all $n = 1, 2, \ldots$,

- $K_n(x) \geq 0$ for all x,
- $$\int_{-\infty}^{\infty} K_n(x)dx = 1,$$

and also such that for all $\epsilon > 0$,

$$\lim_{n\to\infty} \int_{\{x\in\mathbb{R}||x|>\epsilon\}} K_n(x) = 0.$$

Suppose $f : \mathbb{R} \to \mathbb{R}$ is a continuous function with bounded range (i.e. there exists a constant M such that $|f(x)| < M$ for all x.)

a) Show that the product function $x \mapsto f(x)K_n(x)$ is integrable over $\mathbb{R}$ for each n. (You will get partial credit if you show this under the assumption that $\{x \in \mathbb{R}|f(x) \neq 0\} \subset [a, b]$ for some finite $a, b \in \mathbb{R}$.)

b) Show that for any $x \in \mathbb{R}$

$$\lim_{n\to\infty} \int_{-\infty}^{\infty} (f(x-y) - f(x))K_n(y)dy = 0.$$

(Hint: you should break up the integral above into several parts and bound each part separately. Remember that the continuity of f implies that for every $\epsilon > 0$ there exists a $\delta > 0$ such that if $y \in [-\delta, +\delta]$ then $|f(x-y) - f(x)| < \epsilon$.)

Proof. You should think of K_n as a 'bump' which has integral 1 and is getting concentrated on a very small region near zero as $n \to \infty$.

a) The point is that $f(x)$ is bounded in absolute value everywhere by some constant M independent of x, while K_n is bounded in integral (but its maximum absolute value actually blows up as $n \to \infty$, i.e. $\|K_n\|_{C^0} \xrightarrow{n\to\infty} \infty$) so their product is bounded in integral. We have that $f(x)K_n(x)$ is integrable iff $|f(x)K_n(x)|$ is; to see that the latter is integrable we use monotone convergence:

$$\int_{\mathbb{R}} \chi_{[-a,a]}|f(x)K_n(x)| \leq M \int_{-a}^{a} K_n(x) \leq M;$$

taking $a \to \infty$ and using Monotone convergence shows that $|f(x)K_n(x)|$ is integrable.

b) Let's use the hint I gave. Given $\epsilon > 0$, pick δ as in the hint. We have

$$\int_{\mathbb{R}} |f(x-y)-f(y)|K_n(y)dy = \int_{\{x\in\mathbb{R}:|x|>\delta\}} |f(x-y)-f(y)|K_n(y)dy + \int_{\{x\in\mathbb{R}:|x|<\delta\}} |f(x-y)-f(y)|K_n(y)dy.$$

Now take limits as $n \to \infty$. We have that

$$\lim_{n\to\infty} \int_{\{x\in\mathbb{R}:|x|>\delta\}} |f(x-y) - f(y)|K_n(y)dy \leq \lim_{n\to\infty} 2M \int_{\{x\in\mathbb{R}:|x|>\delta\}} K_n(y)dy = 0$$

by monotonicity of integration the triangle inequality

$$|f(x-y) - f(x)| \leq |f(x-y)| + |f(x)| \leq M + M,$$

and the squeeze theorem for limits. (The limit on the left is a sequence of nonnegative integers, so it is squeezed by the sequence $0, 0, 0, \ldots$ which has limit zero, and the limit on the right.) On the other hand

$$\lim_{n\to\infty} \int_{\{x\in\mathbb{R}:|x|<\delta\}} |f(x-y) - f(y)|K_n(y)dy \leq \lim_{n\to\infty} \epsilon \int_{\{x\in\mathbb{R}:|x|<\delta\}} K_n(y)dy \leq \lim_{n\to\infty} \epsilon \int_{\mathbb{R}} K_n(y)dy = \epsilon.$$

So we have that

$$\limsup_{n\to\infty} \int_{\mathbb{R}} |f(x-y) - f(x)|K_n(y)dy \leq \epsilon$$

for every $\epsilon > 0$, so in fact this lim sup is zero since the quantity inside the lim sup is nonnegative. But

$$\limsup_{n\to\infty} \int_{\mathbb{R}} |f(x-y) - f(x)|K_n(y)dy \leq \liminf_{n\to\infty} \int_{\mathbb{R}} (f(x-y) - f(x))K_n(y)dy$$

$$\leq \limsup_{n\to\infty} \int_{\mathbb{R}} (f(x-y) - f(x))K_n(y)dy \leq \int_{\mathbb{R}} |f(x-y) - f(x)|K_n(y)dy$$

So in fact the lim sup and the lim inf in the middle agree, and the limit is zero as claimed.

Here is another way to think about this last question, which uses idea that go a little bit beyond what we have covered in detail. First notice that

$$\int_{\mathbb{R}} f(x)K_n(y)dy = f(x).$$

So the problem follows if we show that

$$\lim_{n\to\infty} \int_{\mathbb{R}} f(x-y)K_n(y)dy = \lim_{n\to\infty} K_n(y-x)f(y)dy = f(x)$$

where the first equality is the same kind of change of variables that occurred in Problem 2 of the Fourier Analysis problem set. But now we know that $K_n(y-x)$, as a function of y, converges to a bump of 'probability 1' supported at $y = x$; so it makes sense that the limit converges to picking out the value of f at x. One should think of the integrals above as taking the *convolution* (`https://en.wikipedia.org/wiki/Convolution`) of K_n with f; the idea is that the K_n converge pointwise to a function that is zero almost everywhere, but as probability distributions they converge to the probability distribution which concentrates all of its probability mass at zero, known as the *delta function* (`https://en.wikipedia.org/wiki/Dirac_delta_function`) $\square$