Mat1062: Introductory Numerical Methods for PDE

Mary Pugh

March 3, 2009

1 Ownership

These notes are the joint property of Rob Almgren and Mary Pugh.

2 Numerical methods

For more on numerical methods for hyperbolic conservation laws see "Numerical Methods for Conservation Laws" by R.J. LeVeque. It's in the Engineering, Gerstein, and the Math/Stats Library: QA377 .L4157.

3 Consistent Conservative Schemes

Our focus now turns to approximate schemes for equations that are in conservative form. Our goal is to construct discrete schemes that will be conservative themselves.

As usual, let h be the space step and k be the time step. We want to shift our focus from standard finite difference methods, where we interpreted u_j^n as the value of u *at* the node point (x_j, t_n). Now we want to interpret it as the *average* of $u(x, t_n)$ over the interval the interval of length h, centered at $x_j = jh$. That is

$$u_j^n \;=\; \frac{1}{h}\int_{x_{j-\frac{1}{2}}}^{x_{j+\frac{1}{2}}} u(x,t_n)\,dx$$

where $x_{j-1/2} = (j - 1/2)h$ and $x_{j+1/2} = (j + 1/2)h$. We recall the integral form of the conservation law

$$\int_\alpha^\beta u(x,t_1)\;dx = \int_\alpha^\beta u(x,t_0)\;dx + \int_{t_0}^{t_1} f(u(\alpha,t))\;dt - \int_{t_0}^{t_1} f(u(\beta,t))\;dt.$$

and we use the cell $[x_{j-1/2}, x_{j+1/2}] \times [t_n, t_{n+1}]$

$$\begin{aligned}&\int_{x_j-\frac{1}{2}}^{x_j+\frac{1}{2}} u(x,t_{n+1})\;dx \\ &= \int_{x_j-\frac{1}{2}}^{x_j+\frac{1}{2}} u(x,t_n)\;dx + \int_{t_n}^{t_{n+1}} f(u(x_{j-1/2},t))\;dt - \int_{t_n}^{t_{n+1}} f(u(x_{j+1/2},t))\;dt\end{aligned}$$

which we write as

$$u_j^{n+1} - u_j^n = -\frac{k}{h}\left(F^n_{j+\frac{1}{2}} - F^n_{j-\frac{1}{2}}\right), \qquad (1)$$

in which $F^n_{j-1/2}$ is defined by

$$F^n_{j-\frac{1}{2}} = \frac{1}{k}\int_{t_n}^{t_{n+1}} f\big(u(x_{j-1/2},t)\big)\;dt. \qquad (2)$$

That is, $F^n_{j-1/2}$ is the time average of the flux into the interval $[x_{j-1/2}, x_{j+1/2}]$ across the point $x = x_{j-1/2}$. We take (1) as our discrete update formula, and the only remaining question is how to approximate (2) based on the values of u_j^n and the flux function f.

However we choose to approximate (2), conservation will be preserved:

$$\sum_{j=0}^N u_j^{n+1} = \sum_{j=0}^N u_j^n - \frac{k}{h}\sum_j^N \left(F^n_{j+\frac{1}{2}} - F^n_{j-\frac{1}{2}}\right) = \sum_{j=0}^N u_j^n - \frac{k}{h}\left(F^n_{N+\frac{1}{2}} - F^n_{0-\frac{1}{2}}\right).$$

The only change in the amount of "stuff" is due to the flux at the boundaries at $x_{0-1/2}$ and $x_{N+1/2}$.

It remains to decide how to determine $F^n_{j\pm1/2}$ from the flux function f and the values $\{u_j^n\}$. For finite difference schemes for linear PDE we needed to show that the scheme was consistent and stable to know that it was convergent. We certainly cannot use the same definition of consistency now — the finite-difference definition relied on doing Taylor series expansions and smoothness. As a result, a scheme which is consistent for $u_t + (u^2/2)_x = 0$ would also be consistent for $(u^2)_t + (2/3\;u^3)_x = 0$. This means that it couldn't converge to a weak solution — we've already seen that these two conservation laws have different weak solutions.

We'll state things for the simplest case in which $F^n_{j+1/2}$ depends only on the cell averages on each side of $x_{j+1/2}$ (that is on u^n_j and u^n_{j+1}):

$$F^n_{j-1/2} = F(u^n_{j-1}, u^n_j) \qquad F^n_{j+1/2} = F(u^n_j, u^n_{j+1})$$

(In principle, they could be taken to be a function of other cell averages as well.) And so now we see that all we need to do is specify a function $F(u_L, u_R)$ of two variables in such a way that the resulting discrete scheme converges to a weak solution (when it converges at all).

Of course the function $F(u_L, u_R)$ needs to be related to $f(u)$. Surprisingly, the necessary conditions are very weak. The scheme is *consistent* if

- $F(\overline{u}, \overline{u}) = f(\overline{u})$ for all $\overline{u} \in \mathbb{R}$, and
- $F(u_L, u_R)$ is Lipshitz continuous around the diagonal $u_L = u_R$. Specifically, given $\overline{u}$ there is a constant $K \geq 0$ (which may depend on $\overline{u}$) such that

$$|F(u, v) - F(\overline{u}, \overline{u})| \leq K \max\{|u - \overline{u}|, |v - \overline{u}|\}$$

That is, F needs only to have the correct value along the diagonal, and "nice" behavior nearby (e.g. be slightly more than continuous).

We understand the first condition as follows: Consider a solution which is constant (equals $\overline{u}$) in two cells $[x_{j-1/2}, x_{j+1/2}]$ and $[x_{j+1/2}, x_{j+3/2}]$ during the period $[t_n, t_{n+1}]$. Then the flux at $x_{j+1/2}$ will equal $f(\overline{u})$ during the time $[t_n, t_{n+1}]$. And so its time average (2) will also be $f(\overline{u})$. That is, $F^n_{j+1/2} = f(\overline{u})$. Recall that we approximate this time-average with $F(u_j, u_{j+1})$ which equals $F(\overline{u}, \overline{u})$ in this case. And so we see that we want $F(\overline{u}, \overline{u}) = f(\overline{u})$ to hold whenever u is constant in neighboring cells for a short time. Having a solution which is constant across a group of cells for a period of time is not only reasonable but it must be prepared for: the piecewise constant weak solutions are precisely like this. And we'd like a scheme that will preserve piecewise constant behaviour for piecewise constant initial data away from the shock.

If u is continuous in a neighborhood of $(x_{j+1/2}, t_n)$ then for h and k small u will be nearly constant in the region of interest $[x_{j-1/2}, x_{j+3/2}] \times [t_n, t_{n+1}]$ and so u_L will be close to u_R. Because $F(u_L, u_R)$ is continuous around the diagonal and has the right value when $u_L = u_R$ we will get a good approximation of $F^n_{j+1/2}$.

For finite-difference schemes we had the Lax-Friedrichs theorem which said that if a scheme for a linear PDE was consistent and stable then it was convergent. In this direction, for hyperbolic conservation laws we have the Lax-Wendroff theorem which says that if a scheme is consistent and conservative and if the sequence of numerical approximations converges to a function u then u is a weak solution of the conservation law:

Lax-Wendroff Theorem *Consider a sequence of grids indexed by* $l = 1, 2, \ldots$ *with mesh parameters* $h_l, k_l \to 0$ *as* $l \to \infty$. *Let* $U_l(x,t)$ *denote the numerical approximation computed with a consistent and conservative mehtod on the* l*th grid. Suppose that* U_l *converges to a function* u *in* L^1_{loc} *and that for each* t, *the sequence of functions* $\{U_l(\cdot,t)\}$ *has uniformly bounded total variation*[1] *then* $u(x,t)$ *is a weak solution of the conservation law.*

Note a major difference: in the Law-Friedrichs theorem convergence was a result of the consistency and stability of the scheme. For conservation laws things are different: we aren't ensured convergence. But we do know that if the approximations happen to converge then they're converging to the right thing. This difference is a reflection of how much harder the analytical study of conservation laws is. The Lax-Wendroff theorem was first proven in "Systems of Conservation Laws" by P.D. Lax and B. Wendroff in *Communications in Pure and Applied Mathematics* 13(1960)217–237.

4 The Lax-Wendroff Theorem

In the January 29 notes, we proved that the finite-difference explicit Euler timestepping scheme converges to a solution. There, we had a consistent, stable scheme which implied that it was convergent. The proof relied on having sufficiently smooth solutions that one could use Taylor expansions to control errors. And all the errors were measured in the L^∞ norm — a norm which is sensitive to pointwise information (when function are continuous). Also, one of the outcomes of the proof was a convergence rate — we learnt

---

[1] For more on this theorem, including the definitions of L^1_{loc} and bounded total variation, see Section 4.

that if we chose a refinement path in which we divided h by 2 and k by 4 at each step then the errors would decrease by a factor of 4.

The proof of the Lax-Wendroff theorem is markedly different. First of all, it requires more advanced mathematics than Calculus. It needs that the sequence of discrete solutions is convergent to start with — convergence isn't an outcome. And frustratingly, it does not provide convergence rates.

4.1 Preliminaries from Real Analysis

You can find most of this stuff in any Real Analysis book. Here, I refer to 3rd edition of Royden's *Real Analysis* and to Kolmogorov & Fomin's *Introductory Real Analysis*) when needed.

First of all, we need some mechanism to go from values on a grid $h_l\mathbb{Z} \times k_l\mathbb{N}$ to a function defined on $\mathbb{R} \times [0,\infty)$ and vice versa. Given a collection of values on the grid $h_l\mathbb{Z} \times k_l\mathbb{N}$

$$\{u_j^n\}_{n=0,j=-\infty}^{\infty,\infty}$$

we use these to construct a piecewise constant function on $\mathbb{R} \times [0,\infty)$ as follows:

$$\mathrm{round}(x/h_l) = j, \quad \lfloor t/k_l \rfloor = n \quad \Longrightarrow \quad u(x,t) = u_k^n$$

That is, u is constant on each rectangle $[x_{j-1/2}, x_{j+1/2}) \times [t_n, t_{n+1})$ where it equals u_j^n. Similarly, given ϕ defined on $\mathbb{R} \times [0,\infty)$ we can define a piecewise constant function on $\mathbb{R} \times [0,\infty)$ via

$$\mathrm{round}(x/h_l) = j, \quad \lfloor t/k_l \rfloor = n \quad \Longrightarrow \quad \phi_l(x,t) = \phi(x_j, t_n).$$

That is, ϕ_l is constructed from the samples of ϕ on the grid $h_l\mathbb{Z} \times k_l\mathbb{N}$.

The definition of weak solution involves "test functions" — we denote the set of all such functions $C_0^\infty(\mathbb{R} \times [0,\infty))$. A function ϕ is in $C_0^\infty(\mathbb{R} \times [0,\infty))$ if ϕ has compact support in $\mathbb{R} \times [0,\infty)$ (it is zero outside of a closed and bounded subset of $\mathbb{R} \times [0,\infty)$) and all of ϕ's partial derivatives exist and are continuous on $\mathbb{R} \times (0,\infty)$.

We need to define $L^1_{loc}(\mathbb{R} \times [0,\infty))$. A function f is in L^1_{loc} if for every compact set K we have

$$\iint_K |f(x,t)| \, dx dt < \infty.$$

This is a useful space of functions for various reasons. One of which is that a function like $f(x,t) = \tanh(x - at)$ which isn't in $L^1(\mathbb{R} \times [0,\infty))$ is in $L^1_{loc}(\mathbb{R}\times[0,\infty))$. Even though it doesn't tend to zero at infinity, it's integrable if you only integrate over closed, bounded sets. Also, the definition of weak solution

$$\int_{\mathbb{R}} \int_0^\infty u(x,t)\ \phi_t(x,t)\ dx\ dt + \int_{\mathbb{R}} \int_0^\infty f(u(x,t))\ \phi_x(x,t)\ dx\ dt \tag{3}$$
$$= - \int_{\mathbb{R}} u_0(x)\ \phi(x,0)\ dx$$

involves integrating u and $f(u)$ against a functions which have compact support. Because the integrands have products which include ϕ_t and ϕ_x the integral isn't over $\mathbb{R} \times [0,\infty)$ it's over a compact set. As a result, for the first integral in (3) to make sense we need

$$\int_{\mathbb{R}} \int_0^\infty |u(x,t)\ \phi_t(x,t)|\ dtdx < \infty$$

This follows immediately if ϕ is nice enough that its time derivative is bounded (which it is):

$$\int_{\mathbb{R}} \int_0^\infty |u(x,t)\ \phi_t(x,t)|\ dtdx = \iint_{\text{supp}\{\phi\}} |u(x,t)\ \phi_t(x,t)|\ dtdx$$
$$\leq ||\phi_t||_{L^\infty} \iint_{\text{supp}\{\phi\}} |u(x,t)|\ dtdx < \infty$$

In the last step we used that u is in L^1_{loc}. Also, if f is Lipschitz then $f(u)$ is also in L^1_{loc}. (To see why f must be Lipschitz if the scheme is consistent just take $u = v$ in (5).) As a result, the second integral in (3) makes sense.

One thing we'll need is the following fact about products of convergent sequences. Assume that f_l converges to f in L^1 (or in L^1_{loc}.) and g_l converges to g in L^∞. This implies that $f_n g_n$ converges to fg in L^1 (or in L^1_{loc}.) This follows from the triangle inequality and from $||fg||_{L^1} \leq ||f||_{L^1} ||g||_{L^\infty}$. Specifically:

$$\begin{aligned} ||f_l g_l - fg||_{L^1} &\leq ||f_l g_l - f_l g||_{L^1} + ||f_l g - fg||_{L^1} \\ &\leq ||f_l||_{L^1} ||g_l - g||_{L^\infty} + ||g||_{L^\infty} ||f_l - f||_{L^1} \\ &\leq C||g_l - g||_{L^\infty} + C||f_l - f||_{L^1} \end{aligned}$$

Since the right-hand side goes to zero as $l \to \infty$ we see that $f_l g_l$ converges to fg. In the last step above, I used the fact that a convergent sequence is a bounded sequence.

Finally, we need to understand some stuff about bounded variation. First of all, fix an interval $[a, b] \subset \mathbb{R}$. Let $a = x_0 < x_2 < \cdots < x_i < \cdots < x_N = b$ be any collection of $N + 1$ points in $[a, b]$ then f is of bounded variation on $[a, b]$ if

$$TV_a^b(f) = \sup \sum_{i=1}^{N} |f(x_i) - f(x_{i-1})| < \infty$$

The supremum is taken over all collections of $N + 1$ points and over all $N \geq 1$. (All pairs of points, all triplets of points, etc.) Functions of bounded variation are reasonably nice — for example, if u if of bounded variation on $[a, b]$ then $f'(x)$ exists for almost all $x \in [a, b]$ (Corollary 6, page 104 of Royden).

Strange as it may seem, just because a function has a derivative almost everywhere this doesn't mean that it will satisfy the fundamental theorem of Calculus. There are functions which have bounded variation on $[a, b]$ for which

$$f(x_1) = f(x_0) + \int_{x_0}^{x_1} f'(s)\, ds \tag{4}$$

does not hold. We'll need something like this in the proof of the Lax-Wendroff theorem. Fortunately, the fundamental theorem of calculus does hold for absolutely continuous functions and it turns out that functions of bounded variation can be well approximated by absolutely continuous functions. So things are okay after all.

A real-valued function f defined on $[a, b]$ is *absolutely continuous* on $[a, b]$ if given $\epsilon > 0$ there exists a $\delta > 0$ such that

$$\sum_{i=1}^{N} |f(x_i') - f(x_i)| < \epsilon$$

for every finite collection $\{(x_i, x_i')\}$ of nonoverlapping intervals whose total length is less than δ

$$\sum_{i=1}^{N} |x_i' - x_i| < \delta$$

Absolutely continuous functions satisfy (4) for all $x_0, x_1 \in [a, b]$. (Corollary 15, page 110 of Royden.)

Finally, we need:

Lemma: If f has bounded variation on $[a,b]$ then there is a sequence of functions $\{f_n\}$ that are absolutely continuous, that converge pointwise almost everywhere to f, and satisfy

$$TV_a^b(f_n) \leq TV_a^b(f)$$

Proof: Because f has bounded variation on $[a,b]$ it is discontinuous at at most countably many points in $[a,b]$. (Problem 7, page 104 of Royden.) And so we can choose D a countable dense subset of $[a,b]$ that doesn't include any of these points of discontinuity. We add the endpoints a and b to D, if needed. From D we construct finite nested subsets

$$D_n = \{x_k^n\}_{k=0}^n, \quad \text{such that} \quad a = x_0^n < \cdots < x_i^n < \cdots < x_n^n = b$$

and $D_n \subset D_{n+1}$ for all $n \in \mathbb{N}$ and $\cup_n D_n = D$. For each n let f_n be the piecewise linear function that interpolates the points $\{(x_i^n, f(x_i^n))\}$. By construction, f_n is absolutely continuous and there is a upper bound on its total variation:

$$TV_a^b(f_n) \leq TV_a^b(f).$$

This upper bound is independent of n. Also, because f has bounded variation, f is bounded on $[a,b]$. As a result each function f_n is bounded on $[a,b]$ and we have a upper bound

$$||f_n||_{L^\infty([a,b])} \leq ||f||_{L^\infty([a,b])}.$$

This upper bound is also independent of n. These n-independent upper bounds on the functions and on their total variation allows us to apply Helly's selection theorem (Theorem 5, page 372, Kolmogorov & Fomin). This tells us there is a subsequence, call it $\{f_{n_k}\}$, which converges pointwise to some function, $\tilde{f}$, and $TV_a^b(\tilde{f}) \leq TV_a^b(f)$. By construction, $\tilde{f}$ and f agree on the set D. Because D is dense in the set of points of continuity of f, it follows that f and $\tilde{f}$ agree at all points of continuity of f. Therefore they agree almost everywhere and therefore we have a sequence $\{f_{n_k}\}$ that converges pointwise to f almost everywhere in $[a,b]$. This finishes the proof.

4.2 Statement and Proof of the Lax-Wendroff theorem

The scheme

$$u_j^{n+1} = u_j^n - \frac{k}{h}\left(F(u_j^n, u_{j+1}^n) - F(u_{j-1}^n, u_j^n)\right)$$

is conservative, by construction. It is consistent if: 1) $F(\overline{u}, \overline{u}) = f(\overline{u})$ for all $\overline{u} \in \mathbb{R}$, and 2) For each $\overline{u} \in \mathbb{R}$ there is a constant $K \geq 0$ (which may depend on $\overline{u}$) such that

$$|F(u, v) - F(\overline{u}, \overline{u})| \leq K \max\{|u - \overline{u}|, |v - \overline{u}|\} \tag{5}$$

We're now ready to state and prove the Lax-Wendroff theorem.

Lax-Wendroff Theorem *Consider a sequence of grids indexed by* $l = 1, 2, \ldots$ *with mesh parameters* $h_l, k_l \to 0$ *as* $l \to \infty$. *Each fixed pair* h_l *and* k_l *generates a collection of grid values* $\{u_j^n\}$ *using the consistent and conservative scheme on the* l*th grid. For each* l *let* $U_l(x, t)$ *be the piecewise constant function on* $\mathbb{R} \times [0, \infty)$ *created from these grid values. Assume that* U_l *converges to a function* u *in* $L^1_{loc}(\mathbb{R} \times [0, \infty))$ *and that for each* $T > 0$ *there is some* R *which may depend on* T *but not on* l *such that*

$$TV_a^b(U_l(\cdot, t)) \leq R, \qquad \forall\, l, \quad \forall\, 0 \leq t \leq T, \quad \forall\, a < b.$$

Then u *is a weak solution — it satisfies*

$$\int_{\mathbb{R}} \int_0^\infty u(x,t)\ \phi_t(x,t) + f(u(x,t))\ \phi_x(x,t)\ dx\ dt = -\int_{\mathbb{R}} u_0(x)\ \phi(x,0)\ dx$$

for all $\phi \in C_0^\infty(\mathbb{R} \times [0, \infty))$.

Proof: Fix a test function ϕ. It is zero outside some compact set. And so there is a $T > 0$ and a $L > 0$ such that $\phi(x,t) = 0$ for all $t \geq T$ and all $|x| \geq L$. Note that ϕ need not vanish at $t = 0$. (Indeed, if all the test functions vanished at $t = 0$ then the weak solution definition would be independent of the initial data u_0. In order to know that the initial data is achieved we need test functions that don't vanish at $t = 0$.)

Fix the index l and for the moment we write h instead of h_l and k instead of k_l. We know

$$u_j^{n+1} - u_j^n = -\frac{k}{h}\left(F(u_j^n, u_{j+1}^n) - F(u_{j-1}^n, u_j^n)\right)$$

and so

$$\phi(x_j,t_n)\left(u_j^{n+1}-u_j^n\right)=-\frac{k}{h}\,\phi(x_j,t_n)\,\left(F(u_j^n,u_{j+1}^n)-F(u_{j-1}^n,u_j^n)\right)$$

We now sum in n and j. Because ϕ has compact support, we need only sum over $-N\le j\le N$ where $N=\lceil L/h\rceil$ and over $0\le n\le M$ where $M=\lceil T/k\rceil$.

$$\begin{aligned}\sum_{j=-N}^{N}\sum_{n=0}^{M}\phi(x_j,t_n)\left(u_j^{n+1}-u_j^n\right)&\\ =-\frac{k}{h}\sum_{j=-N}^{N}\sum_{n=0}^{M}\phi(x_j,t_n)\,&\left(F(u_j^n,u_{j+1}^n)-F(u_{j-1}^n,u_j^n)\right)\end{aligned}$$

We now do summation by parts (the analogue of integration by parts). In general,

$$\begin{aligned}\sum_{n=0}^{M}a_n(b_{n+1}-b_n)&=\sum_{n=0}^{M}a_nb_{n+1}-\sum_{n=0}^{M}a_nb_n=\sum_{n=1}^{M+1}a_{n-1}b_n-\sum_{n=0}^{M}a_nb_n\\&=a_Mb_{M+1}-\sum_{n=1}^{M}(a_n-a_{n-1})b_n-a_0b_0\end{aligned}$$

This results in

$$\begin{aligned}-\sum_{j=-N}^{N}\sum_{n=1}^{M}u_j^n\,(\phi(x_j,t_n)-\phi(x_j,t_{n-1}))-\sum_{j=-N}^{N}u_j^0\,\phi(x_j,0)&\\ =\frac{k}{h}\sum_{j=-N}^{N}\sum_{n=0}^{M}F(u_j^n,u_{j+1}^n)&\Big(\phi(x_j,t_n)-\phi(x_{j-1},t_n)\Big)\end{aligned}$$

Notice that all the boundary terms vanished except for the one at $t=0$. I now multiply by h, resulting in

$$\begin{aligned}-hk\sum_{j=-N}^{N}\sum_{n=1}^{M}u_j^n\,\frac{\phi(x_j,t_n)-\phi(x_j,t_{n-1})}{k}-h\sum_{j=-N}^{N}u_j^0\phi(x_j,0)&\\ =hk\sum_{j=-N}^{N}\sum_{n=0}^{M}F(u_j^n,u_{j+1}^n)\,&\frac{\phi(x_j,t_n)-\phi(x_{j-1},t_n)}{h}\end{aligned}$$

We now write these sums as integrals where the integrands involve the piecewise constant functions $U^l(x,t)$ and $\phi^l(x,t)$. Specifically, we have

$$-\int_0^T \int_{-L}^L U^l(x,t)\ \frac{\phi^l(x,t)-\phi^l(x,t-k_l)}{k_l}\ dxdt - \int_{-L}^L U^l(x,0)\phi^l(x,0)\ dx$$
$$= \int_0^T \int_{-L}^L F(U^l(x,t), U^l(x+h_l,t))\ \frac{\phi^l(x,t)-\phi^l(x-h_l,t)}{h_l}\ dxdt$$

We now want to take $l \to \infty$ and see what each of these integrals converges to. First of all, we know that $U^l \to u$ in L^1_{loc} and because ϕ is smooth and compactly supported we have

$$\frac{\phi^l(x,t)-\phi^l(x,t-k_l)}{k_l} \to \phi_t(x,t) \qquad \text{in } L^\infty$$

For this reason,

$$\int_0^T \int_{-L}^L U^l(x,t)\ \frac{\phi^l(x,t)-\phi^l(x,t-k_l)}{k_l}\ dxdt \to \int_0^T \int_{-L}^L u(x,t)\ \phi_t(x,t)\ dxdt$$

as $l \to \infty$. Similarly,

$$\int_{-L}^L U^l(x,0)\phi^l(x,0)\ dx \to \int_{-L}^L u(x,0)\phi(x,0)\ dx$$

as $l \to \infty$.

So far we haven't used that our scheme is consistent. And we haven't used anything about the total variation of U^l. These will be needed to show that

$$\int_0^T \int_{-L}^L F(U^l(x,t), U^l(x+h_l,t))\ \frac{\phi^l(x,t)-\phi^l(x-h_l,t)}{h_l}\ dxdt$$
$$\to \int_0^T \int_{-L}^L f(u(x,t))\ \phi_x(x,t)\ dxdt$$

Again, because ϕ is smooth and compactly supported we have

$$\frac{\phi^l(x,t)-\phi^l(x-h_l,t)}{h_l} \to \phi_x(x,t) \qquad \text{in } L^\infty$$

and so we need to show that

$$F(U^l(x,t), U^l(x+h_l,t)) \to f(u(x,t)) \qquad \text{in } L^1_{loc}$$

Because the scheme is consistent,

$$
\begin{aligned}
&|F(U^l(x,t),U^l(x+h_l,t)) - f(u(x,t))| \\
&\qquad \leq K\max\{|U^l(x,t)-u(x,t)|, |U^l(x+h_l,t)-u(x,t)|\} \\
&\qquad \leq K\max\{|U^l(x,t)-u(x,t)|, |U^l(x+h_l,t)-U^l(x,t)| + |U^l(x,t)-u(x,t)|\} \\
&\qquad = K\Big(|U^l(x+h_l,t)-U^l(x,t)| + |U^l(x,t)-u(x,t)|\Big)
\end{aligned}
$$

And so,

$$
\begin{aligned}
&\int_0^T\int_{-L}^L |F(U^l(x,t),U^l(x+h_l,t)) - f(u(x,t))|\,dxdt \\
&\leq K\int_0^T\int_{-L}^L |U^l(x+h_l,t)-U^l(x,t)|\,dxdt + K\int_0^T\int_{-L}^L |U^l(x,t)-u(x,t)|\,dxdt
\end{aligned}
\tag{6}
$$

We know that $U^l \to u$ in L^1_{loc} and so the second integral in the right-hand side of (6) goes to zero as $l \to \infty$. And so all we need to do is show that the first integral on the right-hand side of (6)

$$\int_0^T\int_{-L}^L |U^l(x+h_l,t)-U^l(x,t)|\,dxdt \to 0$$

as $l \to \infty$. For this, we use the total variation. We know there is an R such that

$$TV_{-L}^L(U^l(\cdot,t)) \leq R \qquad \text{for all } t \leq T$$

If U^l is absolutely continuous in x then

$$U^l(x+h_l,t) - U^l(x,t) = \int_0^{h_l} U^l_x(x+s,t)\,ds.$$

(Corollary 15 on page 110 of Royden.) And so

$$
\begin{aligned}
\int_{-L}^L |U^l(x+h_l,t)-U^l(x,t)|\,dx &= \int_{-L}^L \left|\int_0^{h_l} U^l_x(x+s,t)\,ds\right|\,dx \\
&\leq \int_{-L}^L\int_0^{h_l} |U^l_x(x+s,t)|\,ds\,dx \\
&= \int_0^{h_l}\int_{-L}^L |U^l_x(x+s,t)|\,dx\,ds \\
&= \int_0^{h_l} TV_{-L}^L(U^l(\cdot,t))\,ds \\
&= h_l\,TV_{-L}^L(U^l(\cdot,t)) \leq h_l\,R
\end{aligned}
\tag{7}
$$

(The step where we replace the integral of $|U^l_x|$ with the total variation of U^l is exercise 13 in Royden.) And therefore

$$\int_0^T \int_{-L}^{L} |U^l(x+h_l,t) - U^l(x,t)| \, dxdt \leq h_l \, R \, T. \tag{8}$$

Since $h_l RT \to 0$ as $l \to \infty$ this shows that the first integral on the right-hand side of (6) goes to zero, as desired.

And so, we would be done if we knew that U^l were absolutely continuous in x. But we don't know this. We see that we don't need the fundamental theorem of calculus to hold for U^l — what we need is to show that inequality (7) holds if $U^l(\cdot,t)$ has bounded variation on $[-L,L]$. By the lemma we know that we can find a sequence of absolutely continuous functions $f_n(x)$ that converge pointwise almost everywhere in $[-L,L]$ to $U^l(x,t)$. By bound (7) we know

$$\int_{-L}^{L} |f_n(x+h_l) - f_n(x)| \leq h_l R.$$

If we could take $n \to \infty$ in this inequality, we would get

$$\int_{-L}^{L} |U^l(x+h_l,t) - U^l(x,t)| \leq h_l R.$$

Integrating in time would then lead to (8) and we'd be done. Are we allowed to take this limit? Yes, by the Lebesgue Dominated Convergence theorem (Theorem 1, page 303, Kolmogorov & Fomin). NB: the Lebesgue Dominated Convergence theorem requires that $U^l(\cdot,t)$ and $\{f_n\}$ be in $L^1([-L,L])$ and that there be an n-independent upper bound for their L^1 norms. These things follow from the lemma and found the bounded variation.

This finishes the proof.