COMPUTATION OF DISCONTINUOUS SOLUTIONS OF HYPERBOLIC SYSTEMS WITH WEAKLY CONSERVATIVE DIFFERENCE SCHEMES*

LAN CHIEH HUANG (黄兰洁)
(Computing Center, Academia Sinica, Beijing, China)

It is well known that for the computation of discontinuous solutions of hyperbolic partial differential equations, the use of conservative difference schemes has partial theoretical justification. The theorem of Lax and Wendroff in [1] states that for a conservative difference approximation of a conservative hyperbolic system $\frac{\partial U}{\partial t}+\frac{\partial F}{\partial x}=0$, if the difference solution converges boundedly almost everywhere, then the limit function is a weak solution of the original system of partial differential equations, and hence satisfies the Rankine Hugoniot condition. Of-course the weak solution obtained may not be the unique physically relevant solution, but under normal circumstances it will be. Now, for real practical problems the partial differential equations often have nonhomogeneous terms and the computational regions usually require coordinate transformations for simplification. Therefore we consider hyperbolic systems with coefficients which depend only on the independable variables and with nonhomogeneous terms——we call such systems weakly conservative. Computational experience over the years tells us that the use of weakly conservative difference schemes derived from the weakly conservative hyperbolic systems also yields in general the correct discontinuous solutions. The reason will be stated and proven in this note.

First of all, let us observe that the Lax and Wendroff theorem holds also for equations with nonhomogeneous terms. That is, for

$$\frac{\partial U}{\partial t}+\frac{\partial F}{\partial x}+B=0, \tag{1}$$

where U, $F(x, t, U)$ and $B(x, t, U)$ are vectors, its weak solution U satisfies

$$\iint\left(\frac{\partial W}{\partial t}\cdot U+\frac{\partial W}{\partial x}\cdot F-W\cdot B\right)dx\,dt+\int W(x, 0)\cdot U(x, 0)dx=0 \tag{2}$$

for every test function W which has continuous first derivatives and which vanishes outside some bounded region. Suppose (1) has difference approximation

$$\frac{\Delta V}{\Delta t}+\frac{\Delta G}{\Delta x}+C=0, \tag{3}$$

where Δ denotes any difference operator. From consistency we have

$$G(x, t, V(x-k\Delta x, t), \cdots, V(x+l\Delta x, t)\to G_0(x, t, V, \cdots, V)=F(x, t, V),$$
$$C(x, t, V(x-m\Delta x, t), \cdots, V(x+n\Delta x, t)\to C_0(x, t, V, \cdots, V)=B(x, t, V),$$

* Received February 22, 1983.

here k, l, m, n are constants. For a given mesh, also denoted by Δ, a discrete solution can be obtained with (3) and then $V_\Delta(x, t)$ in the entire computational region can be defined by interpolation. With only slight modification of the proof of the Lax and Wendroff theorem in [1], we obtain the following result: if as Δx, $\Delta t \to 0$, $V_\Delta(x, t)$ converges boundedly almost everywhere to a function $U(x, t)$, then $U(x, t)$ is a weak solution of (1).

Now consider the coordinate transformation defined by

$$\xi=\xi(x, t), \quad \eta=\eta(x, t) \tag{4}$$

with

$$J^{-1}=\frac{\partial(\xi, \eta)}{\partial(x, t)}\neq 0, \quad J=\frac{\partial(x, t)}{\partial(\xi, \eta)}\neq 0$$

in the region under consideration. In the new variables (1) is

$$\eta_t \frac{\partial U}{\partial \eta}+\xi_t \frac{\partial U}{\partial \xi}+\eta_x \frac{\partial F}{\partial \eta}+\xi_x \frac{\partial F}{\partial \xi}+B=0, \tag{5}$$

here η_t, ξ_t, η_x, ξ_x are considered as functions of ξ and η. It has difference approximation

$$\eta_t \frac{\Delta V}{\Delta \eta}+\xi_t \frac{\Delta V}{\Delta \xi}+\eta_x \frac{\Delta G}{\Delta \eta}+\xi_x \frac{\Delta G}{\Delta \xi}+C=0. \tag{6}$$

Both (5) and (6) are weakly conservative, but they can be written in forms (1) and (3) respectively. Equation (5) can be written as

$$\frac{\partial \tilde{U}}{\partial \eta}+\frac{\partial \tilde{F}}{\partial \xi}+\tilde{B}=0, \tag{7}$$

where

$$\tilde{U}=\eta_t U+\eta_x F, \quad \tilde{F}=\xi_t U+\xi_x F,$$
$$\tilde{B}=B-(\xi_t)_\xi U-(\eta_t)_\eta U-(\xi_x)_\xi F-(\eta_x)_\eta F; \tag{8}$$

and with

$$\tilde{V}=\eta_t V+\eta_x G, \quad \tilde{G}=\xi_t V+\xi_x G.$$

(6) can be written as

$$\frac{\Delta \tilde{V}}{\Delta \eta}+\frac{\Delta \tilde{G}}{\Delta \xi}+\tilde{C}=0, \tag{9}$$

where $\tilde{C}$ includes terms $\frac{\Delta \xi_t}{\Delta \xi} V$ etc. Since (7) and (9) are of forms (1) and (3) respectively, we have: if the difference solution $\tilde{V}$ of (9), or rather $\tilde{V}$ defined by solution V of (6), converges to $\tilde{U}$, then $\tilde{U}$ is a weak solution of (7). On the ξ, η plane, $\tilde{U}$ is a weak solution of (7) if it satisfies

$$\iint\left(\frac{\partial \tilde{W}}{\partial \eta}\cdot \tilde{U}+\frac{\partial \tilde{W}}{\partial \xi}\cdot \tilde{F}-\tilde{W}\cdot \tilde{B}\right)d\xi\, d\eta+\int \tilde{W}(\xi, 0)\cdot \tilde{U}(\xi, 0)d\xi=0 \tag{10}$$

for every test function $\tilde{W}$. Here we have assumed that $t=0$ is mapped onto $\eta=0$ and that $t>0$ corresponds to $\eta>0$, otherwise the single integral in (10) would have a minus sign in front. Let us simply call U which defines $\tilde{U}$ which satisfies (10) a weak solution of (5).

We discuss first weak solutions which are piecewise continuously differentiable in regions separated by smooth curves. The smooth parts of the solutions of (1), (5), and (7) are the same because the equations are equivalent. The discontinuity condition for (1) is

$$[U]\frac{dx}{dt}=[F], \tag{11}$$

where $\frac{dx}{dt}$ is the slope of the discontinuity on the x, t plane; the discontinuity condition for (7) is

$$[\eta_t U+\eta_x F]\frac{d\xi}{d\eta}=[\xi_t U+\xi_x F], \tag{12}$$

where $\frac{d\xi}{d\eta}$ is the slope of the discontinuity on the ξ, η plane. Using $d\xi=\xi_t dt+\xi_x dx$, $d\eta=\eta_t dt+\eta_x dx$ and $J^{-1}\neq 0$, we see readily that the two Rankine Hugoniot conditions are the same. Hence the weak solutions of (1) coincide with those of (5).

This result for more general weak solutions can also be obtained by direct transformation of the weak solution equations. Suppose U is a weak solution of (5), then for every test function $\widetilde{W}$, (10) is valid. We prove that for an arbitrary test vector W, (2) is also valid. Let

$$\widetilde{W}=JW(x(\xi,\ \eta),\ t(\xi,\ \eta));$$

obviously $\widetilde{W}$ is a test function, so (10) is true. We show first that the single integral in (10) is equal to the single integral in (2). Because $t=0$ is mapped onto $\eta=0$, on this initial line a change in ξ does not result in a change in t, so $t_\xi=0$, and from $J\eta_x=-t_\xi$, we also have $\eta_x=0$. Therefore

$$\int [JW\cdot(\eta_t U+\eta_x F)\xi_x]_{t=0}\,dx=\int[(x_\xi t_\eta)W\cdot(\eta_t U)\xi_x]_{t=0}\,dx=\int W(x,\ 0)\cdot U(x,\ 0)\,dx.$$

Here we have used $d\xi=\xi_x dx$ when $dt=0$ and $1=[JJ^{-1}]_{t=0}=(x_\xi t_\eta)(\xi_x\eta_t)$.

As for the double integral in (10), we have that its integrand satisfies

$$\left(\frac{\partial\widetilde{W}}{\partial\eta}\cdot\widetilde{U}+\frac{\partial\widetilde{W}}{\partial\xi}\cdot\widetilde{F}-\widetilde{W}\cdot\widetilde{B}\right)J^{-1}=\frac{\partial W}{\partial t}\cdot U+\frac{\partial W}{\partial x}\cdot F-W\cdot B. \tag{13}$$

Indeed,

$$\begin{aligned}
&J\left(\frac{\partial W}{\partial t}\cdot U+\frac{\partial W}{\partial x}\cdot F-W\cdot B\right)\\
&=J\left(\frac{\partial W}{\partial\eta}\eta_t+\frac{\partial W}{\partial\xi}\xi_t\right)\cdot U+J\left(\frac{\partial W}{\partial\eta}\eta_x+\frac{\partial W}{\partial\xi}\xi_x\right)\cdot F-JW\cdot B\\
&=\left(J\frac{\partial W}{\partial\eta}+J_\eta W\right)\cdot(\eta_t U+\eta_x F)+\left(J\frac{\partial W}{\partial\xi}+J_\xi W\right)\cdot(\xi_t U+\xi_x F)\\
&\quad-JW\cdot\left[B+\frac{1}{J}(J_\eta\eta_t+J_\xi\xi_t)U+\frac{1}{J}(J_\eta\eta_x+J_\xi\xi_x)F\right].
\end{aligned} \tag{14}$$

From $J\xi_t=-x_\eta$, $J\eta_t=x_\xi$ we have

$$(J\xi_t)_\xi+(J\eta_t)_\eta=0$$

or

$$J_\xi\xi_t+J_\eta\eta_t=-J[(\xi_t)_\xi+(\eta_t)_\eta].$$

Similarly

$$J_\xi\xi_x+J_\eta\eta_x=-J[(\xi_x)_\xi+(\eta_x)_\eta].$$

Substituting these into (14) and noting (8), we get (13) immediately. So the double integral in (10) is equal to the double integral in (2), here tracing the x axis and the ξ axis in the positive direction, both regions of integration lie on the left.

We see that given any test function W, (10) is valid for $\widetilde{W}=JW$, and the single and double integrals in (10) are equal to the single and double integrals in (2) respectively, hence (2) is also valid. That is, a weak solution U of (5) is a weak solution of (1). Clearly the converse is also true.

In summary, we have proven that if a conservative hyperbolic system becomes weakly conservative upon coordinate transformation, the difference scheme derived from the transformed system, though weakly conservative, can be used to compute the discontinuous solutions of the original conservative system, in the sense that if the difference solution converges, then the limit function is a weak solution of the original system.

Acknowledgement. In the process of writing this short article, the author held discussions with many colleagues, especially Zhang Yao-ke and Zhang Guan-quan of Computing Center, Academia Sinica, Shi He and Guan Chu-quan of Institute of Systems Science and Institute of Mathematics, Academia Sinica. The author is most grateful for their valuable help and suggestions.

Reference

[1] P. D. Lax, B. Wendroff, Systems of conservation laws, *Comm. Pure Appl. Math.*, 23 (1960), 217—237.