4. Burger's equation

In Lecture (2) we remarked that if the coefficients in

(1) $$u(x,y,\theta)\frac{\partial\theta}{\partial x} + v(x,y,\theta)\frac{\partial\theta}{\partial y} = 0$$

depend not only on (*x*,*y*) but also on θ, then the characteristics may cross and the solutions become multi-valued. In this lecture we consider the special case

(2) $$\frac{\partial\theta}{\partial t} + c(\theta)\frac{\partial\theta}{\partial x} = 0, \qquad -\infty < x < \infty$$

with initial condition

(3) $$\theta(x,0) = f(x)$$

where $f(x)$ is a given function. In (2) the prescribed coefficient depends only on the dependent variable θ.

Equations like (2) are typically derived from conservation laws in the form

(4) $$\frac{\partial\theta}{\partial t} + \frac{\partial}{\partial x}Q(\theta) = 0$$

where $Q(\theta)$ is the flux of θ in the *x*-direction. For example, if θ is the number of cars per unit distance, then $Q(\theta)$ is cars per unit time, and $c(\theta) = Q'(\theta)$.

The characteristic equations corresponding to (2), namely

(5) $$\frac{dt}{ds} = 1, \quad \frac{dx}{ds} = c(\theta), \quad \frac{d\theta}{ds} = 0$$

describe a trajectory in *x*-*t*-θ space along which θ and $x - c(\theta)t$ are constants. Thus the slope of each characteristic is a constant determined by the value of θ at the point where the characteristic intersects the *x*-axis. Let *C* be the characteristic passing through the *x*-axis at $x = x_0$. Then *C* is described by the 2 equations

(6) $$x - c(\theta)t \qquad \text{and} \qquad \theta = f(x_0).$$

Thus *C* is given by

(7) $$x = x_0 + F(x_0)t, \qquad \text{where } F(x_0) \equiv c(f(x_0)).$$

Eqn. (7) describes the whole family of characteristics, with each member corresponding to a diffrerent value of x_0. The solution of (2-3) is $\theta = f(x_0)$ along the curve given by (7). To check this, we compute

(8) $$\theta_t = f'(x_0)\frac{\partial x_0}{\partial t}, \qquad \theta_x = f'(x_0)\frac{\partial x_0}{\partial x}.$$

Taking the *t*- and *x*-derivatives of (7), we obtain

(9) $$0 = \frac{\partial x_0}{\partial t} + t F'(x_0)\frac{\partial x_0}{\partial t} + F(x_0)$$

and

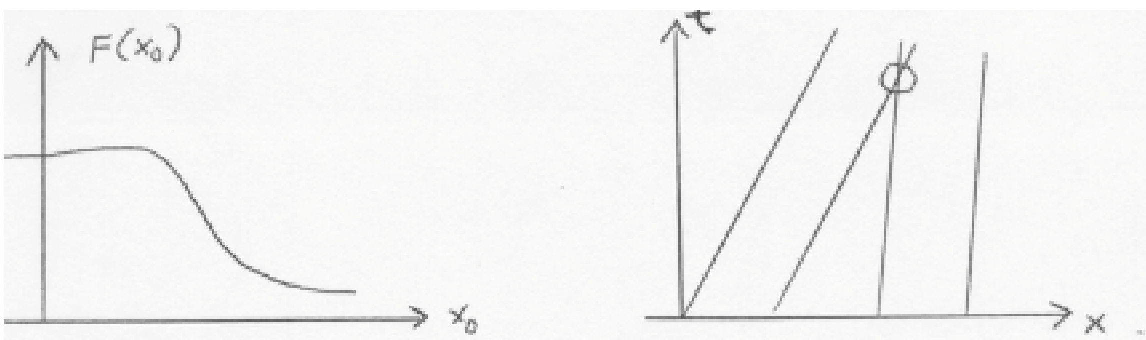
(10) $$1 = \frac{\partial x_0}{\partial x} + t F'(x_0)\frac{\partial x_0}{\partial x}.$$

Thus

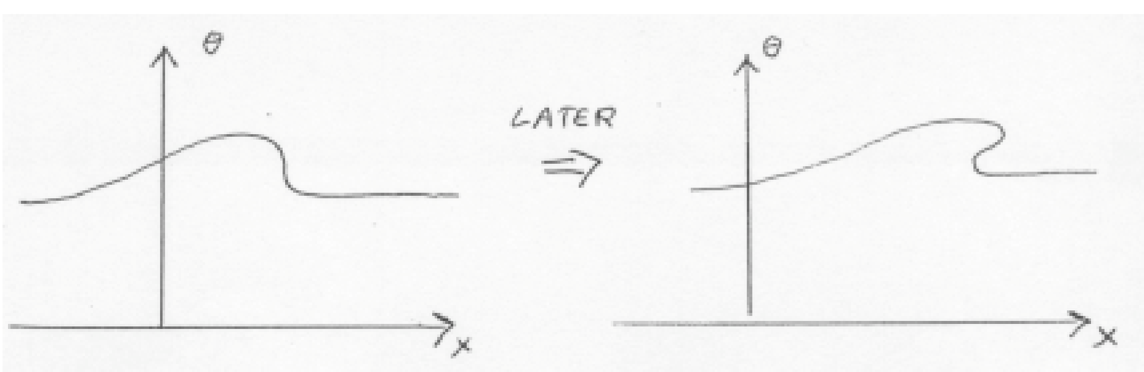
(11) $$\theta_t = f'\left[\frac{-F}{1+tF'}\right] \qquad \text{and} \qquad \theta_x = f'\left[\frac{1}{1+tF'}\right]$$

and hence $\theta_t + F\theta_x = 0$.

If $F'(x_0) < 0$, then the slopes of (7) increase to the right, and the characteristics eventually cross:



When and where does this first occur? When the characetristics cross, the solution becomes multivalued:



This first occurs when $|\theta_x| = \infty$. By (11) we see that both θ_x and θ_t become infinite when

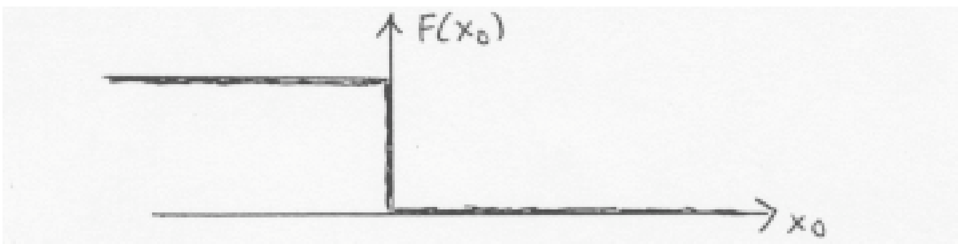
(12) $$1 + tF'(x_0) = 0, \quad i.e. \quad t = \frac{-1}{F'(x_0)}.$$

This first happens on the characteristic with the largest *negative* value of $F'(x_0)$.

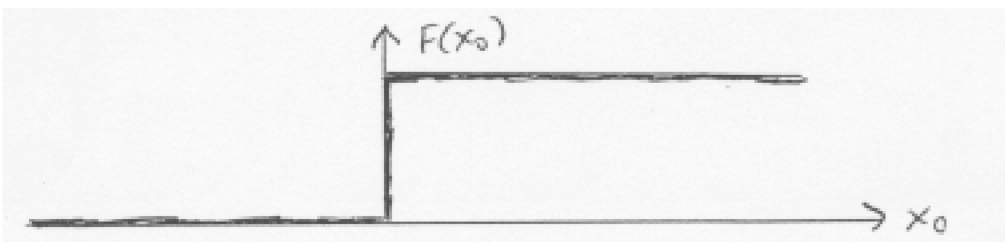
If we regard the initial condition as a line in *x*-*t*-θ space defined by

(13) $$t = 0 \quad \text{and} \quad \theta = f(x),$$

and if every point on this line "moves" according to the trajectory equations (5), then we obtain a surface in *x*-*t*-θ space. Multivalued $\theta(x,t)$ corresponds to a fold in this surface. If



then the solution is *immediately* multi-valued. On the other hand, if



then the solution is *never* multi-valued. What does it look like in that case?

Multivalued $\theta(x,t)$ signals a breakdown in the physics used to derive the original equations (2). The breakdown tells us that we need a more complete physical formulation. There are 2 approaches to introducing the additional physics. One is comprehensive and the other is quite minimal. However, both approaches lead to virtually the same answer provided that they respect the same conservation laws.

Example. The shallow-water equations in rotating coordinates are:

(14) $$\begin{aligned} \frac{Du}{Dt} - f\,v &= -g\frac{\partial h}{\partial x} \\ \frac{Dv}{Dt} + f\,u &= -g\frac{\partial h}{\partial y} \\ \frac{\partial h}{\partial t} + \frac{\partial}{\partial x}(hu) + \frac{\partial}{\partial y}(hv) &= 0 \end{aligned}$$

where $f = f_0 + \beta y$ is the Coriolis parameter. If the flow is nearly geostrophic, then we can neglect *Du/Dt* and *Dv/Dt*. The result:

$$(15) \qquad \frac{\partial h}{\partial t} + \frac{\partial}{\partial x}\left[h\left(-\frac{g}{f}\frac{\partial h}{\partial y}\right)\right] + \frac{\partial}{\partial y}\left[h\left(\frac{g}{f}\frac{\partial h}{\partial x}\right)\right] = 0$$

which simplifies to

$$(16) \qquad \frac{\partial h}{\partial t} + c(h)\frac{\partial h}{\partial x} = 0$$

where

$$(17) \qquad c(h) = -\frac{g\beta}{f^2}h$$

in which y is simply a parameter. The waves described by (16-17) are *Rossby waves*. Since $c(h)<0$ Rossby waves always propagate to the west. If the initial condition is such that $\partial h / \partial x > 0$ then $c'(h) < 0$ and the Rossby wave will break. The breaking can be avoided by adding the simplest type of friction, a linear drag:

$$(18) \qquad \begin{aligned} -f\,v &= -g\frac{\partial h}{\partial x} - \varepsilon\, u \\ +f\,u &= -g\frac{\partial h}{\partial y} - \varepsilon\, v \end{aligned}$$

If $f >> \varepsilon$, then (18) implies

$$(19) \qquad \begin{aligned} u &= -\frac{g}{f}\frac{\partial h}{\partial y} - \frac{\varepsilon g}{f^2}\frac{\partial h}{\partial x} \\ v &= +\frac{g}{f}\frac{\partial h}{\partial x} - \frac{\varepsilon g}{f^2}\frac{\partial h}{\partial y} \end{aligned}$$

and we obtain instead of (16)

$$(20) \qquad \frac{\partial h}{\partial t} + c(h)\frac{\partial h}{\partial x} = \frac{\partial}{\partial x}\left(\frac{\varepsilon g h}{f^2}\frac{\partial h}{\partial x}\right) + \frac{\partial}{\partial y}\left(\frac{\varepsilon g h}{f^2}\frac{\partial h}{\partial y}\right)$$

The ε-terms in (20) keep $h(x,y,t)$ single-valued, but they also make the equation harder to solve. However, we are interested in the case $\varepsilon \to 0$. In that case, the solution of (20) is identical to the solution of (16) *almost everywhere*. That is, the problem reduces to patching together the solutions of (16). This situation is closely analogous to the problems discussed in Section 3.

Instead of (20) we shall consider the model equation

(21) $$\frac{\partial h}{\partial t} + h\frac{\partial h}{\partial x} = \varepsilon\frac{\partial^2 h}{\partial x^2}$$

called *Burger's equation*. Burger's equation is the simplest equation that contains both the nonlinearity needed to make characteristics (now subcharacteristics) cross, and the diffusion needed to keep the solution single-valued. We solve (21) on $-\infty < x < \infty$ with initial condition

(22) $$h(x,0) = f(x).$$

There are 3 methods of attack. In order of increasing sophistication, they are:

(1) shock fitting

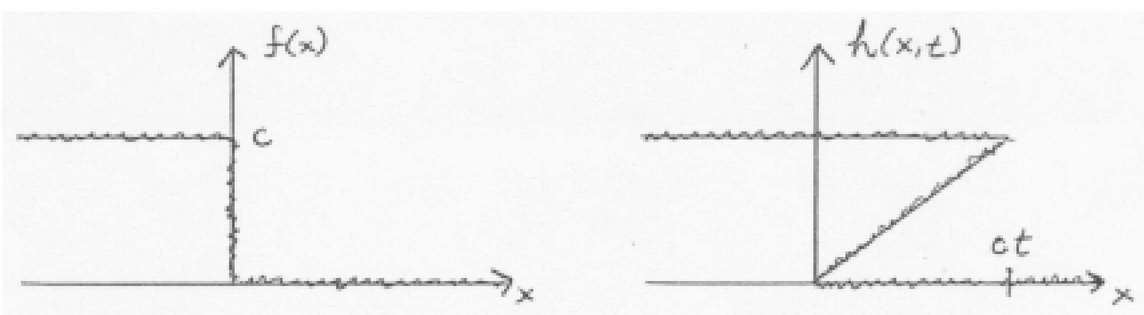
(2) boundary layer theory

(3) exact analytical solution

We start with the second method, boundary layer theory. We assume that "breaking" produces a boundary layer at the unknown location $x_b(t)$. To the left and right of this boundary layer, the solution of (21) is approximately equal to the solution of

(23) $$\frac{\partial h}{\partial t} + h\frac{\partial h}{\partial x} = 0 .$$

Suppose that the initial condition is as on the left:



Then after time *t* the solution of (23) is as on the right. This solution is triple-valued between *x*=0 and *x*=*ct*, where *c* is defined in the picture. Thus we anticipate that $0 < x_b(t) < ct$.

Let $\eta = x - x_b(t)$ be the boundary layer coordinate. For definiteness, we take $\eta = 0$ at $h = c/2$. Then transforming Burger's equation (21) from (*x,t*) coordinates to (η,*t*) coordinates, we obtain

(24) $$\frac{\partial h}{\partial t} + \left(h - \dot{x}_b(t)\right)\frac{\partial h}{\partial \eta} = \varepsilon\frac{\partial^2 h}{\partial \eta^2}.$$

In the boundary layer *h* changes rapidly with η, but $\partial h/\partial t$ is negligible. Thus

(25) $\quad \left(h - \dot{x}_b(t)\right)\dfrac{\partial h}{\partial \eta} = \varepsilon \dfrac{\partial^2 h}{\partial \eta^2}$

which is an ODE depending parametrically on t. The solution must approach h=0 as $\eta / \varepsilon \to \infty$, and h=c as $\eta / \varepsilon \to -\infty$. The first integral of (25) is

(26) $\quad \dfrac{1}{2}h^2 - \dot{x}_b h = \varepsilon \dfrac{\partial h}{\partial \eta} + A(t).$

However, since h and $\partial h / \partial \eta$ vanish as $\eta \to \infty$, $A(t) \equiv 0$. Thus

(27) $\quad \dfrac{1}{2}h^2 - \dot{x}_b h = \varepsilon \dfrac{\partial h}{\partial \eta}.$

However as $\eta \to -\infty$, $h \to c$ and $\partial h / \partial \eta \to 0$. Thus we must have

(28) $\quad \dot{x}_b(t) = c / 2$

and (27) becomes

(29) $\quad \dfrac{1}{2}h^2 - \dfrac{c}{2}h = \varepsilon \dfrac{\partial h}{\partial \eta}.$

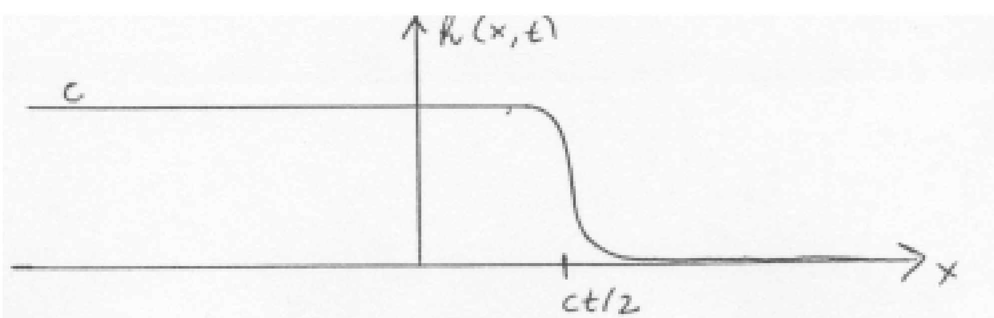
The solution of (29) with $h = c / 2$ at η=0 (as arbitrarily assumed) is

(30) $\quad h = \dfrac{c}{1 + \exp\left(\dfrac{\eta c}{2\varepsilon}\right)}.$

That is,

(31) $\quad h = \dfrac{c}{1 + \exp\left(\dfrac{\left(x - x_b(t)\right)c}{2\varepsilon}\right)}.$

The approximation (31) is uniformly valid. The solution is a "smooth step" propagating to the right at speed c/2. h changes from c to 0 in a boundary layer of thickness ε.



The boundary layer approach just illustrated is very powerful and can be applied to any initial condition. However, we can find the most essential property of the solution—the *location* of the boundary layer—by a much simpler method.

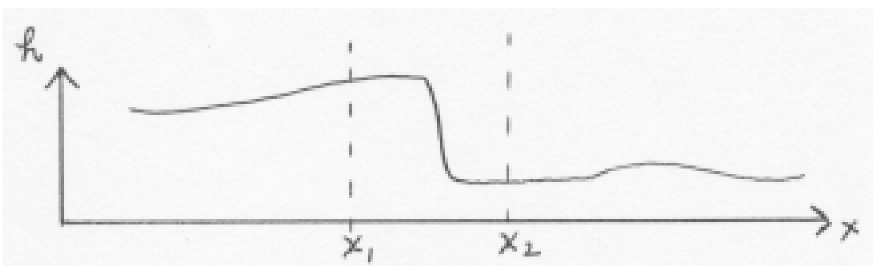
In the method of *shock fitting*, we model the boundary layer as a discontinuity and fix its location by appealing to the conservation law

(32) $$\frac{\partial h}{\partial t}+\frac{\partial}{\partial x}\left(\frac{1}{2}h^2-\varepsilon\frac{\partial h}{\partial x}\right)=0$$

implied by (21). This approach applies to any equation written in the form

(33) $$\frac{\partial h}{\partial t}+\frac{\partial Q}{\partial x}=0\,.$$

Suppose that the solution of (33) is a discontinuity moving to the right at speed U. Let x_1 (x_2) be a *fixed* location to the left (right) of the discontinuity as shown:



Let $h_1=h(x_1)$, etc. Integrating (33) from x_1 to x_2 we obtain

(34) $$\frac{\partial}{\partial t}\int_{x_1}^{x_2} h\,dx+\left(Q_2-Q_1\right)=0$$

which implies

(35) $$U\left(h_1-h_2\right)+\left(Q_2-Q_1\right)=0\,,$$

that is,

(36) $$-U[h]+[Q]=0\,,$$

where [ ] denotes the jump from 1 to 2. In the present example, $h_1=c$, $h_2=0$, and therefore $[h]=-c$. On the other hand,

(37) $$Q=\frac{1}{2}h^2-\varepsilon\frac{\partial h}{\partial x}.$$

Since the ε-term in (37) is negligible outside the boundary layer,

(38) $$[Q] = \frac{1}{2}h_2^2 - \frac{1}{2}h_1^2 = 0 - \frac{1}{2}c^2$$

Thus (36) becomes

(39) $$-U(-c) - \frac{1}{2}c^2 = 0$$

which implies U=c/2 as before. The remarkable aspect of this deduction is that it does not seem to depend on the form of the diffusion. However, *this is not really true*. Suppose we try to apply shock fitting by considering only the ε=0 equation

(40) $$\frac{\partial h}{\partial t} + \frac{\partial}{\partial x}\left(\frac{1}{2}h^2\right) = 0 .$$

As we have seen, the use of (36) in the form

(41) $$-U[h] + \left[\frac{1}{2}h^2\right] = 0$$

gives the correct answer U=c/2. However, (40) also implies the conservation law

(42) $$\frac{\partial}{\partial t}\left(\frac{1}{2}h^2\right) + \frac{\partial}{\partial x}\left(\frac{1}{3}h^3\right) = 0$$

for which (36) takes the form

(43) $$-U\left[\frac{1}{2}h^2\right] + \left[\frac{1}{3}h^3\right] = 0 \;\Rightarrow\; -U\left(-\frac{1}{2}c^2\right) + \left(-\frac{1}{3}c^3\right) = 0 \;\Rightarrow\; U = \frac{2}{3}c \;\; \textit{WRONG}!$$

which disagrees with the previous result (which we know form the boundary layer theory to be correct). What has gone wrong?

Although (42) is a conservation law for the ε=0 equation, $\int h^2 dx$ is *not* conserved when $\varepsilon \neq 0$. The result (43) would be correct if the original problem had been

(44) $$\frac{\partial h}{\partial t} + h\frac{\partial h}{\partial x} = \frac{\varepsilon}{h}\frac{\partial^2 h}{\partial x^2}$$

instead of (21). Thus the form of the diffusion term is absolutely essential in determining the conservation law on which shock fitting should be based.

Note that the general result implied by (41) is

(45) $$U = \frac{\frac{1}{2}h_2^2 - \frac{1}{2}h_1^2}{h_2 - h_1} = \frac{1}{2}(h_1 + h_2).$$

Cole-Hopf transformation.

It is an amazing fact that Burger's equation may be solved *exactly* using a trick discovered independently by Cole and Hopf about 1950. The trick is to change the dependent variable from $h(x,y,t)$ to $\theta(x,y,t)$ defined by

(46) $$h = -2\varepsilon \frac{\theta_x}{\theta} = \frac{\partial}{\partial x}\left(-2\varepsilon \ln\theta\right).$$

Let $\psi = -2\varepsilon \ln\theta$. Then since $h = \psi_x$, Burger's equation (21) becomes

(47) $$\psi_{xt} + \psi_x \psi_{xx} = \varepsilon\psi_{xxx} \;\Rightarrow\; \partial_x\left(\psi_t + \frac{1}{2}\psi_x^{\,2} - \varepsilon\psi_{xx}\right) = 0 \;\Rightarrow$$

(48) $$\psi_t + \frac{1}{2}\psi_x^{\,2} - \varepsilon\psi_{xx} = 0$$

(since ψ is arbitrary by a function of time). Then substituting $\psi = -2\varepsilon \ln\theta$ into (48) and simplifying, we obtain

(49) $$\theta_t = \varepsilon\theta_{xx}.$$

Thus (46) transforms Burger's equation into the heat equation, a linear equation!

The strategy is to solve (49) for θ and then apply (46) to get the solution of Burger's equation. Let $h(x,0) = f(x)$ be the initial condition for Burger's equation. The corresponding initial condition for the heat equation is obtained from

(50) $$f(x) = -2\varepsilon \frac{\partial}{\partial x}\ln\theta.$$

Thus

(51) $$\theta(x,0) = \exp\left[-\frac{1}{2\varepsilon}\int_0^x f(\eta)\,d\eta\right]$$

is the initial condition for the heat equation. (We have set $\theta(0,0)=1$ with no loss in generality.) Using the Green's function for the heat equation, we obtain the solution of (49,51) in the form

(52) $$\begin{aligned}\theta(x,t) &= \int_{-\infty}^{+\infty} dx_0\, \theta(x_0,0) \frac{1}{\sqrt{4\pi\varepsilon t}} \exp\left[-\frac{(x-x_0)^2}{4\varepsilon t}\right] \\ &= \frac{1}{\sqrt{4\pi\varepsilon t}} \int_{-\infty}^{+\infty} dx_0 \exp\left[-\frac{1}{2\varepsilon}\int_0^{x_0} f(\eta)\,d\eta - \frac{(x-x_0)^2}{4\varepsilon t}\right]\end{aligned}$$

Thus the exact solution of Burger's equation with initial condition $h(x,0) = f(x)$ is

$$(53)\qquad h(x,t) = -2\varepsilon\frac{\theta_x}{\theta} = \frac{\int_{-\infty}^{+\infty}\frac{(x-x_0)}{t}\exp\left(-\frac{G}{2\varepsilon}\right)dx_0}{\int_{-\infty}^{+\infty}\exp\left(-\frac{G}{2\varepsilon}\right)dx_0}$$

where

$$(54)\qquad G(x_0,t) \equiv \int_0^{x_0} f(\eta)\,d\eta + \frac{(x-x_0)^2}{2t}.$$

When $\varepsilon \to 0$, the integrals in (53) make their dominant contributions at the minima of $G(x_0)$, i.e. at those places where $\partial G/\partial x_0 = 0$ and $\partial^2 G/\partial x_0^{\,2} > 0$, implying

$$(55)\qquad f(x_0) + \frac{(x_0 - x)}{t} = 0 \qquad \text{and} \qquad t\,f'(x_0) > -1.$$

We recognize (55a) as the equation for the characteristic that intersects the *x*-axis at x_0.

If, for given (x,t), only one value of x_0 satisfies (55a), then we are in the region in which the ε=0 solution is single-valued. Let $x_1(x,t)$ be the single-valued solution of (55a). Then, for any smooth $g(x_0)$,

$$\begin{aligned}
&\int_{-\infty}^{+\infty} g(x_0)\exp\left(-\frac{G(x_0)}{2\varepsilon}\right)dx_0 \\
(56)\qquad &\approx \int_{-\infty}^{+\infty} g(x_1)\exp\left(-\frac{G(x_1)}{2\varepsilon} - \frac{G''(x_1)}{4\varepsilon}(x-x_1)^2\right)dx_0 \\
&= g(x_1)\exp\left(-\frac{G(x_1)}{2\varepsilon}\right)\sqrt{\frac{4\pi\varepsilon}{G''(x_1)}}
\end{aligned}$$

Thus (53) becomes $h \approx (x - x_1)/t$ where x_1 is given by $x = x_1 + f(x_1)t$. That is,

$$(57)\qquad h \approx f(x_1) \qquad \text{where} \qquad x = x_1 + f(x_1)t.$$

This is the same as the solution obtained using characteristic methods.

If (55) has 2 solutions, say x_1 and x_2, then we are in the region in which the characteristic solution is triple-valued. [The condition (55b) excludes the "middle branch" of the characteristic solution.] In this region,

$$(58)\qquad h \approx \frac{\left(\frac{x-x_1}{t}\right)\frac{e^{-G_1/2\varepsilon}}{\sqrt{G_1''}} + \left(\frac{x-x_2}{t}\right)\frac{e^{-G_2/2\varepsilon}}{\sqrt{G_2''}}}{\frac{e^{-G_1/2\varepsilon}}{\sqrt{G_1''}} + \frac{e^{-G_2/2\varepsilon}}{\sqrt{G_2''}}}$$

As $\varepsilon \to 0$, either the G_1-contribution or the G_2-contribution dominates unless $G_1 = G_2$. The condition

$$(59)\qquad G\big(x_1(x,t)\big) = G\big(x_2(x,t)\big),$$

that is,

$$(60)\qquad \frac{(x-x_1)^2}{2t} - \frac{(x-x_2)^2}{2t} = \int_{x_1}^{x_2} f(\eta)\,d\eta$$

along with the characteristic equations

$$(61)\qquad x = x_1 + f(x_1)t, \qquad x = x_2 + f(x_2)t$$

determine the location $x(t)$ of the shock. We will show that its prediction is consistent with

$$(62)\qquad \left[\frac{1}{2}h^2\right] = [h]\dot{x}$$

(cf. (41)) obtained by shock-fitting. First we rewrite (62) in the form

$$(63)\qquad \frac{1}{2}f(x_2)^2 - \frac{1}{2}f(x_1)^2 = \big(f(x_2) - f(x_1)\big)\dot{x},$$

that is,

$$(64)\qquad \dot{x} = \frac{1}{2}\big(f(x_1) + f(x_2)\big)$$

and use (61) to rewrite (60) as

$$(65)\qquad \frac{1}{2}\big(f(x_1) + f(x_2)\big)(x_2 - x_1) = \int_{x_1}^{x_2} f(\eta)\,d\eta.$$

Eliminating x between (61) we obtain

$$(66)\qquad t = -\frac{(x_1 - x_2)}{f(x_1) - f(x_2)}$$

and taking the average of the time derivatives of (61) we obtain

(67) $$\dot{x} = \frac{1}{2}\left\{\left(1 + t\,f'\left(x_1\right)\right)\dot{x}_1 + \left(1 + t\,f'\left(x_2\right)\right)\dot{x}_2\right\} + \frac{1}{2}\left(f\left(x_1\right) + f\left(x_2\right)\right)$$

By (64) the first and last terms in (67) cancel. Substituting (66) into what remains we obtain

(68) $$\left\{\left(f_1 - f_2\right) - \left(x_1 - x_2\right)f_1'\right\}\dot{x}_1 + \left\{\left(f_1 - f_2\right) - \left(x_1 - x_2\right)f_2'\right\}\dot{x}_2 = 0$$

which agrees with the time derivative of (65). This proves that (65) is consistent with the result of shock fitting. In the form (65), the shock condition has an illuminating "equal area" interpretation.

Exact solutions of nonlinear partial differential equations are extremely rare. For this reason, the Cole-Hopf transformation is extremely important. On the other hand, one cannot help but notice that Cole-Hopf requires much more work than either of the two approximate methods—shock fitting and boundary layer theory—and, unlike the approximate methods, it is inapplicable to other problems.

References. Whitham, chapters 1-4, especially chapter 2 pp 19-46 and chapter 3 pp 96-106.