

MULTIPHASE SATURATION EQUATIONS, CHANGE OF TYPE AND INACCESSIBLE REGIONS

Barbara Lee Keyfitz*

Abstract. We identify a class of flux functions which give rise to conservation laws which are hyperbolic except along a codimension one subspace of state space. We show that a number of systems modelling porous medium flow can be regarded as perturbations of such systems, and describe the phenomenon of change of type for these perturbations. We also discuss a property of solutions of such systems, the existence of inaccessible regions - subsets of state space which appear to be avoided by solutions.

1. INTRODUCTION

Equations that change type appear in some models for multiphase flow, where they cause a certain amount of controversy. The context is as follows. In a fluid consisting of several phases, components or both, in a flow regime where dissipative, diffusive or dispersive effects should be negligible, one ends up with a system of quasilinear first-order partial differential equations, in space and time, which are expected to be of hyperbolic type. Details of the modelling depend on the specific problem — Darcy’s law for porous medium flow, the usual conservation of mass and momentum for compressible two-phase flow, other models for transport of solutes or sediments — but the same disconcerting behavior occurs: not only do the characteristic speeds depend on the state variables (as is usual) but also there is a region of state space where some characteristics are complex conjugates. In the simplest case, where there are two equations only and a single space variable, this means that the equation changes type from hyperbolic to elliptic. Recall that steady transonic flow also contains both supersonic (hyperbolic) and subsonic (elliptic) regimes. However, it turns out that the two sorts of problems — steady and unsteady models — have different mathematical structure; see Keyfitz [8]. In addition, issues that are important for applications — such as well-posedness of the initial-value problem — are different in the two cases. In this paper, ‘change of type’ refers to systems of first-order quasilinear equations, modelling unsteady problems, in which some characteristic speeds change from real to complex in a region of state space. We shall generally assume that the equations are in conservation form.

More background on change of type in porous medium flow can be found in the review article by Keyfitz [10].

In the next section, we describe some models where the phenomenon appears. Based on these examples, we identify a class of nonlinear flux functions with the property that the associated conservation laws are hyperbolic everywhere, but nonstrictly hyperbolic on a codimension one submanifold of phase space. For systems of two equations, we write down some explicit criteria for membership in this class.

In §3 what happens when an equation in this class is subjected to a general perturbation is discussed: change of type may occur. In the case of two equations, for example, elliptic regions arise near the original curve of nonstrict hyperbolicity. Other phenomena affecting the structure of solutions of conservation laws, such as curves of linear degeneracy, are also associated with the perturbations.

The next section, §4, contains an informal discussion of one aspect of the ill-posedness associated with change of type: the existence of so-called inaccessible regions in the flow. Difficulties associated with this are mentioned in Allen *et al.* [1]. We show that this behavior is similar to strictly hyperbolic systems which admit linear degeneracies.

2. MULTICOMPONENT SATURATION EQUATIONS

The general form of the equations we study is

$$\partial_t p_i + \partial_x(p_i v_i(P)) = 0, \quad 1 \leq i \leq n. \quad (2.1)$$

Here $P = (p_1, \dots, p_n)$ is the vector of states, and we define $F = (f_1, \dots, f_n)$, the vector of corresponding fluxes. The particular form $f_i = p_i v_i(P)$ is appropriate when each p_i is the density or relative saturation of a component or phase. Then (2.1) is the set of continuity equations for n species, and v_i , the velocity of the i th species, depends on the state vector.

A system of kinematic equations like (2.1) is part of a more complete model of a fluid system. Momentum and energy equations have been omitted and the system has been closed by the assumption that velocities depend on densities or concentrations alone. Rarely is this completely realistic. However, it is an approximation which is often taken seriously. For example, it is used to describe the so-called *miscible displacement problem* in enhanced oil recovery, in which one component of P represents the saturation of a solvent in a fluid whose concentration is given by the other component. In this case it is often assumed that all the v_i are identical (see, for example, Johansen and Winther [5]). Another example arises in three-phase *immiscible* porous medium flow, where the momentum equation is replaced by Darcy's law and it is further assumed, for a single space dimension, that the pressure equation can be solved explicitly. In this case, the v_i are complicated functions of the phase fractions, P , involving the three-phase relative permeabilities, which are usually determined by interpolation (see Allen *et al.* [1]). One classic example of a kinematic equation is a continuum model for traffic flow; the adaptation of this model to a system purportedly describing two-directional traffic flow leads to change of type, (Bick and Newell [2]).

Standard models for two-phase compressible, nonreacting flow consist of a pair of equations which represent conservation of each phase, coupled with momentum and energy transfer equations; see Stewart and Wendroff [18]. These equations reduce asymptotically to a pair of continuity equations of the form (2.1) when the faster-moving waves are ignored; see Keyfitz [9].

Finally, kinematic equations like (2.1) are important in some chemically reacting systems, such as chromatography and other adsorption processes; see Rhee *et al.* [15] and Temple [20]. Change of type and failure of strict hyperbolicity occur in some but not all of these systems.

We examine systems with the property that a single pair of eigenvalues coincides for some values of P , while remaining real everywhere. For concreteness, consider a pair of equations,

$$\begin{aligned} p_t + f_x &= 0, \\ q_t + g_x &= 0, \end{aligned} \quad (2.2)$$

with Jacobian matrix

$$dF = \begin{pmatrix} f_p & f_q \\ g_p & g_q \end{pmatrix}.$$

The eigenvalues of dF , which are the characteristic speeds, are real or complex according to the sign of the discriminant

$$D(p, q) = \left(\frac{f_p - g_q}{2} \right)^2 + f_q g_p. \quad (2.3)$$

System (2.2) is *strictly* hyperbolic when D is positive, and, since this is an open condition, (2.2) has this property on an open subset of \mathbf{R}^2 . For the same reason, if (2.2) is strictly hyperbolic for a flux vector F_0 in the entire region of physical interest (the positive quadrant or unit square, say), then it remains strictly hyperbolic in this entire region for all perturbations of F_0 in a sufficiently small C^1 -open set about F_0 .

However, it may happen that D is simply nonnegative everywhere. In this case, the system (2.1) or (2.2) is called *nonstrictly* hyperbolic. The condition $D \geq 0$ is *not* open: if there are no other constraints on F there will be points (p, q) with $D < 0$ in any neighborhood of a point where $D = 0$. Let us introduce a terminology for flux functions which satisfy a constraint which prevents this.

Definition 2.1 *We say that F is a nonstrictly hyperbolic (NSH) flux function for system (2.2) in a subset $R \subset \mathbf{R}^2$ if the discriminant D defined in (2.3) is nonnegative for all $(p, q) \in R$ and $D = 0$*

for at least one point in the interior of R . The coincidence locus, Σ , is the set of points (p, q) where $D = 0$.

The definition excludes strictly hyperbolic systems and also systems, such as the gas dynamics equations, which lose strict hyperbolicity at the vacuum state, on the boundary of the physically interesting region.

One well-known class of NSH flux functions consists of gradients of a potential; if $F = \nabla\Phi$ then $dF = d^2\Phi$ is symmetric and hence its eigenvalues are real. In this case,

$$D = \left(\frac{\Phi_{pp} - \Phi_{qq}}{2} \right)^2 + \Phi_{pq}^2,$$

and the system is NSH if there is a point (p, q) where the two equations $\Phi_{pp} - \Phi_{qq} = 0$ and $\Phi_{pq} = 0$ are satisfied. If Φ is an arbitrary C^2 function of two variables, then these equations constitute two independent conditions, and Σ consists of isolated points (p, q) — usually called *umbilic* points. Non-strictly hyperbolic fluxes with umbilic points have been studied extensively, beginning with Schaeffer and Shearer [16].

However, many examples that arise in modelling have a different structure: they are nonstrictly hyperbolic, but the condition $D = 0$ reduces to a single equation whose solution space Σ is a curve contained in R . An example is the class of separated potentials: a flux $F = (f(p), g(q))$ corresponding to an uncoupled system has this structure, as does the flux that results in a triangular Jacobian:

$$F = (f(p, q), g(q)). \quad (2.4)$$

In this example, eigenvalue coincidence occurs where

$$G(p, q) \equiv f_p(p, q) - g'(q) = 0,$$

and if this equation is satisfied at a point (p_0, q_0) then, by the implicit function theorem, it is satisfied on a curve through that point provided

$$\nabla G(p_0, q_0) \neq 0. \quad (2.5)$$

A two-component flow with a flux of the form (2.4) has the property that the velocity of the q -component is independent of the density of the p -component: the system is at least partially uncoupled. This may not be a realistic approximation for multiphase flows; though the two-way traffic equations, (Bick and Newell [2]), have this property when there is no interaction between the two directions of flow.

A more interesting example of a NSH flux is given by the two-component miscible displacement saturation equations. These can be written with a flux vector of the form

$$F(p, q) = (p\phi(p, q), q\phi(p, q)) \quad (2.6)$$

(see Johansen and Winther [5]); in this case

$$D = \left(\frac{p\phi_p + q\phi_q}{2} \right)^2,$$

and the system is nonstrictly hyperbolic if the equation $G \equiv p\phi_p + q\phi_q = 0$ has a solution (p_0, q_0) ; eigenvalue coincidence occurs along a curve under the nondegeneracy condition (2.5). A two-fluid model for gas chromatography, (Temple [20]), has a structure similar to (2.6), except that the second component of F is multiplied by a constant; however, in the standard model (using the Langmuir isotherm), there is no eigenvalue coincidence: $D > 0$ everywhere in the physical region.

The two NSH models above — a flow in which the velocities of both components are the same, and a flow in which one is independent of the other — seem to represent extremes in kinematic modelling. However, they have similar mathematical properties. Furthermore, general perturbations of either

model, which take them out of the class of NSH systems, also have much in common, as we shall explore in the next section.

In the remainder of this section, we give a brief description of models with the property that Σ is a curve. This is motivated by the observation that special multiphase kinematic flows, as in the examples given above, appear to have this structure, rather than the umbilic structure of a NSH flux which derives from a potential. We shall refer to this class of NSH fluxes as *coincidence-line* fluxes.

In some multiphase saturation models, such as three-phase porous medium flow, the flux vectors must either be NSH or show change of type (Shearer and Trangenstein [17]), and this has motivated studying perturbations of umbilic points. Given a model with an elliptic region, one can embed it in a family of models in which the elliptic region is shrunk to a point or to a line, and it is to some extent a matter of taste which one chooses to do. For example, in Vinod [21], the elliptic region occurring in a quadratic model for the two-way traffic equations is shrunk to a line — the limit of zero interaction — while in Holden and Holden [3] an equivalent quadratic model is considered as a perturbation of an umbilic. Nonstrictly hyperbolic systems whose eigenvalues coincide along a line have simpler solutions than hyperbolic systems with an umbilic point. It would be interesting to be able to show that the constraints which force a NSH flux to be of coincidence line type have some physical basis.

We make the following obvious remark.

Proposition 2.1 *Let $F \in C^2$ be a NSH flux vector in the sense of Definition 2.1 and let D be the corresponding discriminant. Then at any point where $D = 0$ we also have $\nabla D = 0$.*

The converse holds as well under some additional conditions on F ; e. g., $F \in C^3$ and nondegeneracy conditions on d^2D . Informally, we might say that D needs to be a perfect square if the flux is to be NSH, and Proposition 2.1 shows that this condition is not likely to be met by an arbitrary pair of functions f and g . Furthermore, in order for a NSH flux vector to be a coincidence-line flux (rather than umbilic), the three equations $D = \nabla D = 0$ must be equivalent to a single equation in the two variables p and q , to yield a curve Σ .

The proposition does not suggest any useful way of characterizing these fluxes. But looking at the geometry of conservation laws in the plane gives some insight. The Jacobian matrix, dF , of a NSH flux vector has one real eigenvector at every point in R . Denote the eigenvector by $\zeta(U)$, where $U = (p, q)$, and suppose it can be chosen to depend smoothly on U . The integral curves of the line field generated by ζ give a foliation of R ; if the curves are written in the form $\chi(U) = c$ then χ is a Riemann invariant.

Now, F is a mapping from R to a subset of \mathbf{R}^2 , and by translating F by a constant vector (which does not affect the conservation law system) and possibly scaling by a constant factor, we can assume that $F(R) \subset R$. (For our purposes, we may assume R has compact closure, and then the additional assumption made here is that F is bounded.) In this case, the equation

$$(dF(U))\zeta(U) = \lambda(U)\zeta(U) \tag{2.7}$$

has the interpretation that F maps each Riemann invariant curve to a translate of that curve. (This follows from differentiating along the curve. At corresponding points U and $F(U)$, the tangent vectors, ζ , to the curve are parallel, as a consequence of equation (2.7).) Let us consider the special case that F respects the foliation: that is, the image of a Riemann invariant curve is a Riemann invariant curve. (This does not seem a particularly natural assumption, but it holds for a number of physical models, including the ones mentioned above.) Now, either F maps every curve to itself, or F maps each curve to a different curve. The condition on tangents implies that a Riemann invariant curve that is mapped to itself must be a straight line segment. On the other hand, if curves are mapped into each other by F , then the tangency condition means that all the curves are, effectively, translates of each other, so that χ is of the form $q - r(p)$, at least locally.

Thus we state

Proposition 2.2 *Let R be a bounded subset of \mathbf{R}^2 , and let ζ be a smooth line field defined on R satisfying one of the two constraints: the integral curves of ζ are straight lines, or the integral curves are translates of a single curve. Then there is a smooth flux function, F , defined on R , with eigenvector $\zeta(U)$ at each point U .*

Proof: Temple [20] has shown how to construct a flux F corresponding to any line field with straight-line integral curves. Writing $\zeta = (1, h(U))$, the condition that ζ have straight-line integral curves is

$$\nabla h \cdot \zeta = 0,$$

so h is any smooth solution of $h_p + hh_q = 0$; given h , Temple shows that the form of F is (up to inessential normalizations)

$$F(U) = \phi(U) \begin{pmatrix} 1 \\ h(U) \end{pmatrix} + \begin{pmatrix} 0 \\ H(h) \end{pmatrix}. \quad (2.8)$$

Here ϕ is an arbitrary function of U and H an arbitrary function of h . Temple's construction begins with the fixed solution $h(U)$; however, it is clear that from any foliation of R by straight lines, one can construct ζ (up to the condition that the first component be nonzero), and F . This generalizes the flux of (2.6), for which ζ is just the field of radial lines: $\zeta = (1, q/p)$ with Temple's choice of normalization. The eigenvalue corresponding to ζ is $\lambda(U) = \phi_p + h\phi_q = \nabla\phi \cdot \zeta = \partial_\zeta\phi$. The other eigenvalue is $\text{tr}(dF) - \lambda = \phi h_q + H_q = (\phi + H'(h))h_q$.

When $\chi = q - r(p)$, then $\zeta = (1, r'(p))$; it is easy to verify that

$$F(U) = \begin{pmatrix} a(p) - q \\ b(p) \end{pmatrix} \quad (2.9)$$

is a flux vector which gives rise to this eigenvector as long as b , a and r are related by $b' = r'(a' + r')$. The corresponding eigenvalue is $\lambda = a' - r'$, and the other eigenvalue is r' . Notice that there is no condition imposed on r other than sufficient smoothness. The region R will, of course, depend on r . ■

So far, we have not required that the fluxes be NSH. However, for both the families we have constructed, it is straightforward to impose the condition that eigenvalues coincide on a subset of R ; under suitable nondegeneracy conditions, this occurs along a curve for both types of models.

Proposition 2.3 *Let a flux function F be defined by (2.8) or (2.9) on a domain R . Then F is a nonstrictly hyperbolic flux vector if there is a point (p_0, q_0) where, in the first case,*

$$G(p, q) \equiv \phi_p + h\phi_q - \phi h_q - H_q = 0,$$

and, in the second,

$$G(p) \equiv a'(p) + 2r'(p) = 0.$$

Furthermore, $U_0 = (p_0, q_0)$ lies on a nondegenerate curve, Σ , of coincident eigenvalues in the first case if $\nabla G(U_0) \neq 0$. In the second case, Σ includes the line $p = p_0$, and this is an isolated coincidence line if $G'(p_0) \neq 0$.

Proof: This is a straightforward calculation using the expressions for the eigenvalues given in Proposition 2.2. The existence and nondegeneracy of Σ follow from the implicit function theorem applied to G . ■

Another interpretation of equation (2.7) is as follows. The flux corresponding to (2.4) produces a nondegenerate curve as a line of coincident eigenvalues as long as (2.5) holds. Again suppose that F maps the domain R into R , and take any smooth coordinate transformation, T , from R to itself; then the composition

$$\hat{F} = T \circ F \circ T^{-1} \quad (2.10)$$

defines a new flux function on R . (It is not particularly natural, on the basis of the physical problems, to regard F as a mapping of R to itself; however, it is a reasonable approach if one wants to study the structure of eigenvalues of dF .) Defining $V = T(U)$, one has another system

$$\partial_t V + \partial_x(\hat{F}(V)) = 0. \quad (2.11)$$

There need not be any relation between the eigenvalues of dF and those of $d\hat{F}$. In fact, defining $S = T^{-1}$ as an abbreviation,

$$d\hat{F}(V) = (dT)_{F \circ S(V)}(dF)_S(V)(dS)_V \quad (2.12)$$

where the subscripts on the right indicate where the Jacobians are to be evaluated. By contrast, if one begins with the original conservation law (2.1) and applies the change of coordinates to U , one obtains the quasilinear system

$$\partial_t V + [(dT)_V(dF)_U(dS)_V] \partial_x V = 0 \quad (2.13)$$

which cannot, in general, be put in conservation form. Comparing the Jacobian in (2.12) with the matrix in (2.13), we see that they are the same only if

$$(dT)_{F \circ S(V)} = (dT)_V. \quad (2.14)$$

In this case, (2.11) is the same as the original system, but written in new coordinates. (The two systems are not equivalent when it comes to weak solutions or shock structure, but they have the same characteristic speeds, and the corresponding eigenvectors transform to each other under T .) However, (2.14) is a very restrictive condition.

Nonetheless, one can generate conservation laws (2.11) with a structure that is qualitatively like a given one, as follows. Suppose that, with $U = (p, q)$, F is of the form (2.4), and so dF is upper triangular. In that case, so is $C(dF)$, where

$$C = \begin{pmatrix} 1 & e(U) \\ 0 & c(U) \end{pmatrix}$$

is an upper triangular matrix. If now we replace (2.14) by

$$(dT)_{F \circ S(V)} = (dT)_V C \quad (2.15)$$

then (2.11) is a conservation law whose eigenvalues are those of $C(dF)$ — that is f_p and cg_q . In particular, they are real everywhere in R and coincide along the curve Σ where $f_p(U) = c(U)g_q(U)$. Thus we have proved

Proposition 2.4 *Let F be any smooth flux of the form (2.4) which maps a subset R of \mathbf{R}^2 to itself. Let $T(U) = V$ be a nonsingular coordinate change in R and C any smooth upper triangular matrix. Then, provided the compatibility condition (2.15) is satisfied, the system (2.11), with flux defined by (2.10), is hyperbolic in R , with eigenvalues f_p and cg_q . Equation (2.11) has a NSH flux vector \hat{F} if there is a point, U_0 , where*

$$G(U) \equiv f_p(U) - c(U)g'(q) = 0,$$

and this equation is satisfied on a curve Σ if $\nabla G(U_0) \neq 0$.

This proposition relates several of the examples above — for example, the transformation from Cartesian to polar coordinates turns the triangular flux function, (2.4), into the flux vector for the miscible displacement problem, (2.6). In [11], we show that a flux generated this way produces a discriminant D which is a perfect square, as in Proposition 2.1. One class of such fluxes is given by

$$\hat{F}(u, v) = \varphi(u, v) \begin{pmatrix} t(Q) \\ r(Q) \end{pmatrix} + \begin{pmatrix} s(Q) \\ a(Q) \end{pmatrix}, \quad (2.16)$$

where φ is any function of $V = (u, v)$; t , r , s and a are arbitrary functions satisfying the constraint that

$$\begin{aligned} u &= pt(q) + s(q), \\ v &= pr(q) + a(q), \end{aligned}$$

is a locally invertible change of coordinates, and $Q(u, v)$ is the solution of

$$ur(Q) - vt(Q) = b(Q) \equiv r(Q)s(Q) - t(Q)a(Q). \quad (2.17)$$

Since any smooth solution of the quasilinear equation

$$t(b' - ur')Q_u + r(b' + vt')Q_v = 0$$

with $b(Q(0, 0)) = 0$ satisfies (2.17) and so generates a flux of the form (2.16), this generalizes (2.8). Finally, we have expressions for the eigenvalues of the Jacobian $d\hat{F}$ (from Proposition 2.4):

$$\lambda_1 \equiv f_p = \varphi_u t + \varphi_v r$$

and

$$\lambda_2 \equiv cg_q = \frac{\varphi(r't - rt') + a't - s'r}{ur' - vt' - (rs + at)'},$$

which yield a nondegenerate coincidence line when Q is nonconstant. Derivation of (2.16) will be found in [11].

3. PERTURBATIONS OF NONSTRICTLY HYPERBOLIC FLUXES

Suppose F is a NSH flux with eigenvalue coincidence along Σ , a nondegenerate curve given by

$$\delta(p, q) = 0$$

where $\delta^2 = D$ and $\nabla\delta \neq 0$. (Here D is given by (2.3).) We consider a one-parameter family of perturbed fluxes

$$\tilde{F} = \begin{pmatrix} f + \epsilon\phi \\ g + \epsilon\psi \end{pmatrix}.$$

In general, the new flux will not be NSH; in fact (tildes denote the perturbed quantities), the discriminant of \tilde{F} is

$$\tilde{D}(p, q) = D(p, q) + \epsilon P(p, q, \epsilon) = \delta^2 + \epsilon P$$

where

$$P(p, q, \epsilon) = \frac{(f_p - g_q)}{2}(\phi_p - \psi_q) + \phi_q g_p + \psi_p f_q + \epsilon \left[\left(\frac{\phi_p - \psi_q}{2} \right)^2 + \phi_q \psi_p \right].$$

Proposition 3.1 *Let \tilde{F} be a smooth perturbation of a smooth coincidence-line NSH flux with $\Sigma = \{(p, q) \mid \delta(p, q) = 0\}$. Suppose that $P(p, q, 0) > 0$ on Σ . Then, for sufficiently small $\epsilon > 0$, \tilde{F} is strictly hyperbolic near Σ , while for $\epsilon < 0$ there will be a nonhyperbolic strip near Σ whose width is of order $\sqrt{\epsilon}$. On the other hand, if P changes sign on Σ , let*

$$P(p(\tau), q(\tau), 0) = \tau \hat{P}(p(\tau), q(\tau)) \quad (3.1)$$

where τ parameterizes Σ and $\hat{P} > 0$ there. Then, $d\tilde{F}$ has nonreal eigenvalues inside a region whose boundary is, to a first approximation, a narrow parabola with its vertex near $U_0 = (p(0), q(0))$, opening toward $\tau > 0$ or $\tau < 0$ as $\epsilon < 0$ or $\epsilon > 0$, respectively.

Proof: The result when $P \neq 0$ on Σ follows immediately from applying the implicit function theorem to the two equations $\delta \pm \sqrt{-\epsilon P} = 0$: if $\epsilon < 0$, there are two solution curves for small ϵ and \tilde{D} is negative between them.

If P changes sign along Σ , then equation (3.1) implies that the change of sign is nondegenerate. One can now solve

$$\delta(p(\tau, \epsilon), q(\tau, \epsilon)) \pm \sqrt{-\epsilon \tau \hat{P}} = 0$$

for each fixed τ , again using the implicit function theorem, since $\nabla\delta \neq 0$. The geometry of the nonhyperbolic region follows. ■

Perturbations of this type also have a connection with genuine nonlinearity. Curves of linear degeneracy (corresponding to isolated local extrema of the characteristic speeds along the eigenvectors of dF) appear when certain nonstrictly hyperbolic systems are perturbed so that they become strictly hyperbolic [6]. For these systems, the solution of the Riemann problem for the nonstrictly hyperbolic equation is qualitatively like the solution of the perturbed, strictly hyperbolic problem, which approaches it in the limit as $\epsilon \rightarrow 0$.

Curves of linear degeneracy also bifurcate from distinguished points on the boundary between an elliptic and a hyperbolic region in state space [7]. This bifurcation may occur generically at points like U_0 in Proposition 3.1. As we shall discuss in the next section, there is also some resemblance between the solutions in the two cases corresponding to $\epsilon > 0$ and $\epsilon < 0$ of Proposition 3.1.

4. INACCESSIBLE REGIONS

A disturbing feature of change of type in saturation equations like (2.1) is that there appear to be open sets in phase space, corresponding to physically feasible saturation vectors, which are nevertheless inaccessible because the Jacobian is nonhyperbolic there, (Allen *et al.* [1]). There are two related questions here. The first is whether some indefensible assumption in the physical model has resulted in a saturation vector which is linearly unstable. Unlike simplistic models for phase transitions which change type and are nonhyperbolic precisely for the physically unstable range of the order parameter (see Pego and Serre [14] for an example), the flows discussed in this paper are not expected to contain unstable states.

The second point is a mathematical one. The nonhyperbolic region has some special properties. Specifically, solutions to the Riemann problem, when the data are in the hyperbolic region, avoid the nonhyperbolic states. This has been demonstrated analytically for some systems and is believed on the basis of numerical experience for others (see Allen *et al.* [1] and Pego and Serre [14]); there are no counterexamples, to the best of my knowledge. On the other hand, for more general initial data (Cauchy data) the hyperbolic region is not invariant: data in the hyperbolic region for which the solution enters the nonhyperbolic region are given in Holden *et al.* [4] and Pego and Serre [14]. The fact that Riemann data and Cauchy data behave so differently raises additional questions, which we leave aside.

In this section, we try to shed a bit of light on the second, mathematical, question by calling attention to an analogous phenomenon in strictly hyperbolic systems which contain hypersurfaces (curves, in the case of two equations) of linear degeneracy in state space. We shall concentrate on Riemann problems, since this is where the behavior is seen in flows which change type. There are some implications about the modelling of flows, because the analogy, even if only in Riemann problems, with a strictly hyperbolic system, suggests strongly that the appearance of inaccessible regions is linked more to wrinkles in the nonlinear dependence of the fluxes than it is to change of type in the equations. The fact that both examples occur as perturbations of coincidence-line fluxes suggests that there may be a relationship between them.

Riemann solutions for conservation laws without convexity (genuine nonlinearity) assumptions were first given by Liu [12] by constructing a solution separately in each wave family and superimposing waves from different families. We formulate a result for a scalar equation, where we can give an explicit description, and then indicate the generalization.

Definition 4.1 *Let $u_t + f(u)_x = 0$ be a scalar conservation law. We say the Riemann problem with data $\{a, b\}$ avoids states in a nonempty subset $\Sigma \subset (a, b)$ if neither the solution to the Riemann problem*

$$u(x, 0) = \begin{cases} a, & x < 0, \\ b, & x \geq 0, \end{cases} \quad (4.1)$$

nor the solution to

$$u(x, 0) = \begin{cases} b, & x < 0, \\ a, & x \geq 0, \end{cases} \quad (4.2)$$

takes values in Σ .

We have the following result.

Proposition 4.1 *Let $f \in C^1$, and let f' be strictly monotone on open intervals with no accumulation point. If f is strictly convex (or concave) on $[a, b]$, then no states in (a, b) are avoided. However, any interval on which f' is not weakly monotone will contain at least one subinterval of avoided states.*

Proof: Osher [13] presents a formula which gives a closed-form solution to the Riemann problem. It is equivalent to the following construction.

The *lower convex hull* of f on $[a, b]$ is

$$\underline{f}(u; a, b) = \inf \left\{ f(c) + (u - c) \frac{f(d) - f(c)}{d - c} \right\}$$

where the *inf* is taken over c and d with $a \leq c \leq u \leq d \leq b$. Similarly the *upper convex hull* of f on $[a, b]$ is

$$\overline{f}(u; a, b) = \sup \left\{ f(c) + (u - c) \frac{f(d) - f(c)}{d - c} \right\}$$

with the *sup* taken over the same domain. Write the Riemann problem as

$$u(x, 0) = \begin{cases} u_L, & x < 0, \\ u_R, & x \geq 0. \end{cases}$$

For problem (4.1), where $u_L < u_R$, solve the Riemann problem for $u_t + \underline{f}(u)_x = 0$. For problem (4.2), with $u_L > u_R$, use $u_t + \overline{f}(u)_x = 0$. The Riemann problem for a (nonstrictly) convex function f is solved as follows. The centered solution $u(\zeta)$, $\zeta = x/t$, satisfies

$$[-\zeta + f'(u(\zeta))] u'(\zeta) = 0$$

in the sense of distributions, and can be written

$$u(\zeta) = g^{-1}(\zeta), \tag{4.3}$$

where $g \equiv f'$ is monotone. If g is constant on an interval I , then g^{-1} is discontinuous and the interior of I is not in its range. If g is increasing then (4.3) provides a solution for $u_L < u_R$; if g is decreasing, then (4.3) is a solution with $u_L > u_R$.

If f is strictly convex on (a, b) , then $\underline{f} = f$ on (a, b) and g is strictly increasing. Then (4.3) yields a continuous solution (rarefaction wave) if $u_L < u_R$, and the range of g is the entire interval $[u_L, u_R]$. (In this case, \overline{f} is a straight line segment

$$\overline{f}(u) = f(a) + (u - a) \frac{f(b) - f(a)}{b - a} \equiv f(a) + (u - a)s,$$

and $\overline{f}' = \overline{g} \equiv s$ in (a, b) . The solution to the Riemann problem with $u_L = b$ and $u_R = a$ is $u(\zeta) = \overline{g}^{-1}(\zeta)$: u is piecewise constant with a discontinuity at $\zeta = s$. The range of \overline{g}^{-1} is the two values $\{u_L, u_R\}$.)

If f is strictly concave, then $\overline{f} = f$ and \underline{f} is a line segment, and the result is the same: this time $f' = g$ is strictly decreasing and has a continuous inverse, and no states are avoided.

Finally, if f is neither convex nor concave, then both \underline{f} and \overline{f} differ from f , and both contain line segments. In fact, let there be a point c in (a, b) such that f' is strictly increasing on $(c - \epsilon, c)$ and strictly decreasing on $(c, c + \epsilon)$ for some $\epsilon > 0$. Then f is strictly convex on the first interval and strictly concave on the second, so if $a < c - \epsilon$ and $b > c + \epsilon$ then there is an open interval $(c - \delta, c + \delta)$, $0 < \delta < \epsilon$, on which $\underline{f} < f < \overline{f}$. From the construction of the convex hulls, \underline{f} and \overline{f} are affine functions on $(c - \delta, c + \delta)$, and hence this interval is not in the range either of \overline{g}^{-1} or of \underline{g}^{-1} . ■

For a system of conservation laws, Liu [12] constructs a curve $\gamma(U_L)$ in state space which is locally a shock (part of the Hugoniot locus) or a rarefaction. This provides a higher-dimensional analogue to \underline{f} and \overline{f} and a solution like the scalar one can be constructed. The construction works whenever F has isolated hypersurfaces of linear degeneracy (which play the same role as the isolated extrema of f') and the eigenvalues of dF are separated. If $U_M \in \gamma_1(U_L)$ for a 1-wave curve, say, then some subintervals

of that curve will not be in the range of the Riemann solution. Further, for U_R in a neighborhood of such a U_M , there will continue to be open sets of the interval that are missed in solving the Riemann problem for $\{U_L, U_R\}$. It is no longer the case that $U_M \in \gamma_1(U_L) \Rightarrow U_L \in \gamma_1(U_M)$. However, for small-amplitude waves, U_L will be near $\gamma(U_M)$, and a generalization of the idea of *avoided states* can be given in several ways. One formulation is to fix a left state U_L and consider right states in an open ball \mathcal{B} of radius r centered at a point U_0 . For each $U_R \in \mathcal{B}$, let $\mathcal{S}(U_R) \subset \mathbf{R}^n$ be the range of the Riemann solution with data $\{U_L, U_R\}$, and define

$$\mathcal{S}(\mathcal{B}) = \bigcup_{U_R \in \mathcal{B}} \mathcal{S}(U_R).$$

If F is genuinely nonlinear, $\mathcal{S}(\mathcal{B})$ has at most $n + 1$ connected components. Also, if U_0 is sufficiently close to U_L and F is genuinely nonlinear at U_L , then $\mathcal{S}(\mathcal{B})$ has at most $n + 1$ components. For a system with linear degeneracies, as $|U_0 - U_L|$ grows, $\mathcal{S}(\mathcal{B})$ will develop more than $n + 1$ components at some U_0 . The introduction of new components into $\mathcal{S}(\mathcal{B})$ corresponds to production of additional discontinuities in g^{-1} in the scalar case. Thus it is associated with avoided states. The qualitative change in $\mathcal{S}(\mathcal{B})$ also suggests a lack of continuous dependence of the solution on the data.

Avoided states differ from the inaccessible regions in systems that change type, as they are not precisely the complements of invariant regions. Similar behavior is observed in the two cases, since sampling data repeatedly in the exterior of a convex set, E , produces a solution which never enters a convex set C contained in E . The difference is that for Riemann data in systems which change type, one can apparently take E to be the elliptic region and C to coincide with E . For hyperbolic equations with linear degeneracies, C is strictly smaller than E . Nonetheless, the similarities are striking.

In a recent paper, Temple [19] contrasts Riemann problems (their relation to stability and asymptotics) in genuinely nonlinear strictly hyperbolic problems to their role in a model NSH system with a coincidence-line flux. The results (summarized in the title of the paper) are unexpected. We conjecture that NSH coincidence-line models may provide good prototypes for mathematical properties and qualitative behavior of a larger class of problems, including some models for porous media flow.

REFERENCES

- [1] Allen M. B., Behie G. A., Trangenstein J. A. *Multiphase flow in porous media: mechanics, mathematics, and numerics*. Lecture Notes in Engineering 34. Springer-Verlag, New York, 1988.
- [2] Bick J. H., Newell G. F. A continuum model for two-directional traffic flow. *Quart. Appl. Math.*, XVIII:191–204, 1960.
- [3] Holden H., Holden L. On the Riemann problem for a prototype of a mixed type conservation law, II. In *Current Progress in Hyperbolic Systems: Riemann Problems and Computations*, Contemporary Mathematics 100, pages 331–367. Amer. Math. Soc., Providence, 1989. B. Lindquist, ed.
- [4] Holden H., Holden L., Risebro N. H. Some qualitative properties of 2×2 systems of conservation laws of mixed type. In *Nonlinear Evolution Equations that Change Type*, IMA Volumes in Mathematics and its Applications 27, pages 67–78. Springer, 1990. B. Keyfitz and M. Shearer, eds.
- [5] Johansen T., Winther R. The solution of the Riemann problem for a hyperbolic system of conservation laws modelling polymer flooding. *SIAM J. Math. An.*, 19:541–566, 1988.
- [6] Keyfitz B. L. Some elementary connections among nonstrictly hyperbolic conservation laws. In *Nonstrictly Hyperbolic Conservation Laws*, Contemporary Mathematics 60, pages 67–77. Amer. Math. Soc., Providence, 1987. B. Keyfitz and H. Kranzer, eds.
- [7] Keyfitz B. L. A criterion for certain wave structures in systems that change type. In *Current Progress in Hyperbolic Systems: Riemann Problems and Computations*, Contemporary Mathematics 100, pages 203–213. Amer. Math. Soc., Providence, 1989. B. Lindquist, ed.

- [8] Keyfitz B. L. Shocks near the sonic line: a comparison between steady and unsteady models for change of type. In *Nonlinear Evolution Equations that Change Type*, IMA Volumes in Mathematics and its Applications 27, pages 89–106. Springer, 1990. B. Keyfitz and M. Shearer, eds.
- [9] Keyfitz B. L. Change of type in simple models of two-phase flow. In *Viscous Profiles and Numerical Approximation of Shock Waves*, pages 84–104. SIAM, Philadelphia, 1991. M. Shearer, ed.
- [10] Keyfitz B. L. Conservation laws that change type and porous medium flow: a review. In *Modeling and Analysis of Diffusive and Advective Processes in Geosciences*, pages 122–145. SIAM, Philadelphia, 1992. W. E. Fitzgibbon and M. F. Wheeler, eds.
- [11] Keyfitz B. L. A method for generating nonstrictly hyperbolic fluxes with eigenvalue coincidence along a line. In preparation.
- [12] Liu T.-P. The Riemann problem for general 2×2 conservation laws. *Amer. Math. Soc. Trans.*, 199:89–112, 1974.
- [13] Osher S. J. Riemann solvers, the entropy condition, and difference approximations. *SIAM Jour. Numer. Anal.*, 21:217–235, 1984.
- [14] Pego R. L., Serre D. Instabilities in Glimm’s scheme for two systems of mixed type. *SIAM Jour. Numer. Anal.*, 25:965–988, 1988.
- [15] Rhee H.-K., Aris R., Amundson N. R. *First-Order Partial Differential Equations: Volume I, Theory and Application of Single Equations*. Prentice-Hall, Englewood Cliffs, 1986.
- [16] Schaeffer D. G., Shearer M. The classification of 2×2 systems of non-strictly hyperbolic conservation laws, with application to oil recovery. *Comm. Pure Appl. Math.*, 40:141–178, 1987.
- [17] Shearer M., Trangenstein J. A. Loss of real characteristics for models of three-phase flow in a porous medium. *Transport in Porous Media*, 4:499–525, 1989.
- [18] Stewart H. B., Wendroff B. Two-phase flow: models and methods. *Jour. Comp. Physics*, 56:363–409, 1984.
- [19] Temple J. B. The L^1 -norm distinguishes the strictly hyperbolic from a nonstrictly hyperbolic theory of the initial value problem for systems of conservation laws. In *Nonlinear Hyperbolic Equations – Theory, Computational Methods and Applications*, Notes Numer. Fluid Mech. 24, pages 608–616. Aachen, 1988; Vieweg, Braunschweig, 1989.
- [20] Temple J. B. Systems of conservation laws with coinciding shock and rarefaction curves. In *Nonlinear Partial Differential Equations*, Contemporary Mathematics 17, pages 143–151. American Mathematical Society, Providence, 1983. J. A. Smoller, ed.
- [21] Vinod V. *Structural stability of Riemann solutions for a multiphase kinematic conservation law model that changes type*. PhD thesis, University of Houston, 1992.