# A CONTINUUM MODEL FOR TWO-DIRECTIONAL TRAFFIC FLOW*

BY

J. H. BICK** AND G. F. NEWELL†

*Brown University*

**Abstract.** The flow of traffic in two directions of an undivided highway is investigated using the equations of continuity and assumed empirical relations between the average velocities and densities in both lanes. These lead to a pair of quasi-linear partial differential equations. Even if the velocity in one lane depends only very weakly on the density in the other lane, it is found that for a certain small range of densities the equations are of elliptic rather than the expected hyperbolic type. For densities outside this range, solutions of the equations can be found for various special types of initial conditions.

**1. Introduction.** A number of papers have now been published treating the theory of traffic flow from the standpoint of the dynamics of individual cars, statistics and continuum mechanics. Lighthill and Whitham [1], Richards [2], and De [3] have studied unidirectional flow of traffic by assuming that the empirical relation between the average velocity of cars and the density of cars, determined under steady state conditions, applies also to non-steady flows. They then use the one dimensional equation of continuity to describe how the densities vary with time for various types of disturbances.

Here we shall attempt to extend their work to include two lanes of traffic flowing in opposite directions. We use two equations of continuity and assume that the average velocities are functions of the densities in both lanes. If we let $p$ and $q$ denote the densities of cars and $U$ and $V$ the average velocities in the right-moving lane ($p$-lane) and the left-moving lane ($q$-lane), respectively, the equations of continuity for the two lanes are

$$p_t + (pU)_x = 0, \qquad q_t + (qV)_x = 0, \tag{1}$$

in which subscripts denote partial differentiation with respect to the subscript variable.

We will assume that $p$, $q$, $U$ and $V$ satisfy the following conditions: (A) $U$ and $V$ are functions of $p$ and $q$ only; i.e., the highway is uniform in both space and time; (B) $p$ and $q$ are never negative; (C) $U$ is non-negative and $V$ non-positive; (D) $U$ and $V$ have continuous first derivatives with respect to $p$ and $q$; (E) $U$ is a monotone decreasing and $V$ a monotone increasing function of both $p$ and $q$; (F) for each value of $q$, there is a value of $p$ at which $U = 0$ and for each $p$ there is a value of $q$ at which $V = 0$; (G) $U(p, q) = -V(q, p)$; i.e., both lanes are physically identical. Property (F) in conjunction with (A) and (B) implies that the allowable domain of solutions is finite and bounded by the four curves $p = 0$, $q = 0$, $U = 0$ and $V = 0$ in the $(p, q)$ plane.

If shocks develop, they must satisfy the integrated equations of continuity

$$A[U(A, B) - W] = p[U(p, q) - W], \qquad B[V(A, B) - W] = q[V(p, q) - W], \tag{2}$$

*Received Sept. 15, 1959.

**Now with Atomics International, Canoga Park, Calif. The work of J. H. B. was supported by the International Business Machine Corporation of New York City.

†Alfred P. Sloan Research Fellow.

where $W$ is the shock velocity, $A$ and $B$ represent the densities in the $p$- and $q$-lanes respectively on the left side of the shock and $p$, $q$ the corresponding densities on the right side of the shock. It follows from this, that if there is a shock in one lane, there must also be a contiguous shock in the other. If, for example, $A$ were equal to $p$ (no shock in the $p$-lane), Eq. (2) gives $U(A, B) = U(A, q)$. By property ($E$), this implies that $B = q$, so there can be no shock in the $q$-lane either.

Several methods exist for solving quasi-linear differential equations such as Eq. (1). First, there is the method of characteristics in the physical or $(x, t)$ plane. The characteristics of a system of equations are those curves in the $(x, t)$ plane along which disturbances will propagate and a set of two first order partial differential equations is called hyperbolic if there are two real, one-parameter families of such curves. If, for hyperbolic equations, initial conditions are specified along a non-characteristic curve, the solution, can be obtained by numerical integration along the characteristics [4, 5].

A second method of solution results from exchanging the roles of dependent and independent variables. The technique is well known in the theory of compressible fluids and is referred to as the hodograph transformation [6]. The equations of continuity in the $(p, q)$ or hodograph plane are,

$$\begin{aligned} x_q - (U + pU_p)t_q + pU_q t_p &= 0, \\ x_p - (V + qV_q)t_p + qV_p t_q &= 0. \end{aligned} \tag{3}$$

Since this system is linear, the characteristics, unlike those in the physical plane, are independent of initial conditions. The initial conditions, given by specifying $p$ and $q$ as functions of $x$ at time zero, describe a parametric representation of a curve, $p = p(x, 0)$, $q = q(x, 0)$, in the hodograph plane along which $x$ and $t$ ($t = 0$) are specified. Such data, if it does not lie along a hodograph characteristic can be uniquely extended to define a solution $x = x(q, p)$, $t = t(p, q)$.

Although one would expect the equations of continuity to be of hyperbolic type, the only type which gives a unique solution for initial data $p(x, 0)$ and $q(x, 0)$, Eq. (1) is not always hyperbolic. The physical plane characteristics are defined by the velocities

$$2w \equiv 2\frac{dx}{dt} = (pU)_p + (qV)_q \pm \{[(pU)_p - (qV)_q]^2 + 4qpU_qV_p\}^{1/2}. \tag{4}$$

For every fixed value of $q$, $pU$ is zero when $p$ is zero and also when $p$ has its maximum value, i.e., when $U$ is zero. It is positive between these points, and therefore, by virtue of property (D), it has a maximum at which $(pU)_p$ is zero. Similarly, $qV$ has at least one negative minimum in the allowable range of $q$ for every fixed value of $p$. Consequently, there exist values of $p$ and $q$ for which $[(pU)_p - (qV)_q]^2$ vanishes. By properties (B), (D), and (E), the term $4qpU_qV_p$ is negative for all $q$ and $p$ greater than zero. For a certain range of $p$ and $q$, the characteristics are therefore complex, and the equations of continuity are of elliptic type. Outside this region the equations are hyperbolic.

As yet we have found no satisfactory explanation of why the equations should be elliptic nor any satisfactory suggestion as to what one should do about it. Although one could imagine hypothetical situations that would give rise to densities in the region of ellipticity, it is not obvious whether or not such situations could be realized in practice. It may, however, also be true that one is not justified in assuming that the relations between the velocities and densities, determined under stationary conditions, are also valid under non-stationary conditions, in which case much of the theory described here as well as that in reference [1, 2, 3] may collapse. One might, in fact, seriously question any theory in which the future behavior of a system is determined with no reference to the distribution of the velocities of individual cars. On the other hand, the ellipticity

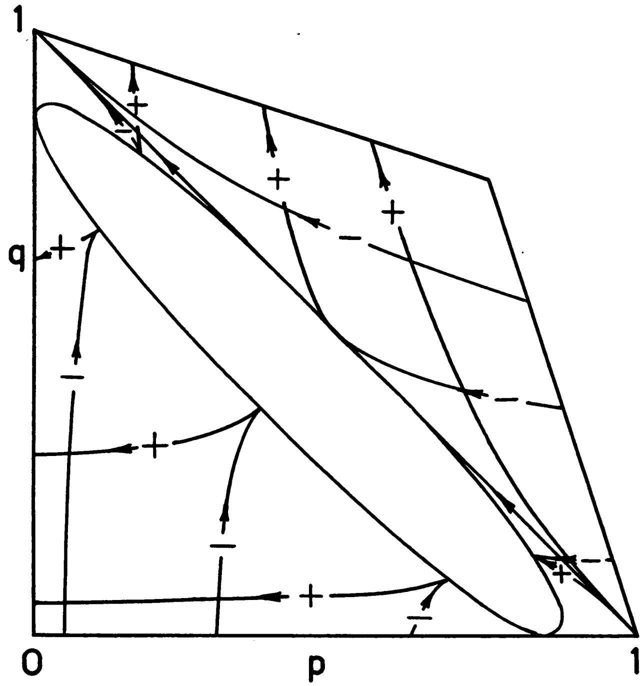


FIG. 1. The bounded domain of solutions in the hodograph plane and the region of ellipticity obtained from Eq. (5 a) are shown for $\beta = 1/8$. The two families of characteristics are distinguished by the labels, + or −, and the arrows indicate the direction of increasing wave velocity.

arises when the densities are such that, in the absence of a coupling, the wave velocities in the two lanes are nearly equal. Under such conditions, one could reasonably expect any coupling between the lanes to produce some unusual effects.

**2. A special case.** It has been found experimentally for unidirectional traffic that the average car velocity is approximated reasonably well by the linear relation $U = U_0 - \alpha p$. To illustrate some of the consequences of the present model, we consider the velocity-density relations

$$U = V_0 - \alpha p - \beta q, \qquad V = -U_0 + \alpha q + \beta p, \tag{5}$$

where $\alpha$ and $\beta$ are constants with $\alpha > \beta > 0$. These satisfy all the properties listed in the previous section and reduce to Richards velocity functions if $\beta = 0$, $p = 0$ or $q = 0$. Furthermore they have the form of the first three terms of a Taylor's series expansion, so if we know the velocities for some values of $p$ and $q$, we could use Eq. (5) for small variations in $p$ and $q$ about these values.

The velocity relations, the equations of continuity, and the shock equations, may be non-dimensionalized by the substitutions $U' = U/U_0$, $V' = V/U_0$, $p' = \alpha p/U_0$,

$q' = \alpha q/U_0$, $\beta' = \beta/\alpha$, $t' = U_0 t$, and $W' = W/U_0$. Dropping the primes for the sake of simplicity in notation, we obtain from Eq. (5)

$$U = 1 - p - \beta q, \qquad V = -1 + q + \beta p, \tag{5a}$$

whereas Eqs. (1–4) remain unchanged.

The region of the hodograph plane in which the equations of continuity are of elliptic type is the interior of an ellipse (see Fig. 1) obtained by setting the radical of Eq. (4) equal to zero. This ellipse is tangent to the two axes of the hodograph plane and to the line $p + q = 1$. It has a minor axis of width $(2)^{1/2}\beta/(2 + 2\beta)$.

The characteristics in the hodograph plane are determined by the differential equation

$$\beta p\,(dq)^2 - [2 - (2 + \beta)(p + q)]\,dp\,dq + \beta q\,(dp)^2 = 0 \tag{6}$$

the solution of which can be found by means of the Legendre transform [7]. In parametric form, the solution is

$$p = X\frac{dY(X)}{dX} - Y(X), \qquad q = \frac{dY(X)}{dX} \tag{7}$$

with

$$Y(X) = \frac{X^{\epsilon}(X + 1)^{1-2\epsilon}}{1 + \beta}\left\{\int^{X} \frac{d\lambda}{\lambda^{\epsilon}(\lambda + 1)^{2-2\epsilon}} + C\right\}, \tag{7a}$$

$$\epsilon = \beta/(2 + 2\beta), \tag{7b}$$

and $C$ is a constant of integration. For any $X$, Eq. (7a) determines $Y$ and Eq. (7) then gives both $p$ and $q$ as a function of the parameter $X$. In addition Eq. (6) has the singular solution

$$p + q = 1 \tag{8}$$

which forms an envelope of the general solutions. Some characteristics for $\beta = 1/3$ are shown in Fig. 1. A more detailed description of the hodograph characteristics particularly for $\beta \ll 1$ is given in [8].

The shock relations, Eq. (2), consist of two equations in the five variables $A$, $B$, $p$, $q$ and $W$. If we eliminate $W$ from these equations, we obtain a quadratic equation in $p$, $q$, $A$, and $B$ which for any fixed values of $A$ and $B$ can easily be solved for $q$ as a function of $p$ or vice versa. The resulting curve in the $(p, q)$ plane is known as the shock polar corresponding to the given values of $A$ and $B$. If $(A, B)$ lies in the hyperbolic region of the hodograph plane, the shock polar will pass through the point $(A, B)$ twice and will be tangent each time to one or the other of the two hodograph characteristics through the point $(A, B)$. Some typical shock polars are shown in Fig. 2.

Although it is always possible, for some shock velocity $W$, to find a shock satisfying Eq. (2) between any denisities $(A, B)$ and densities $(p, q)$ lying on the appropriate shock polar, the resulting shock may not be stable (even for $\beta = 0$). If at zero time, we specify continuous strong discontinuities that lie on a shock polar and we wish to calculate the densities at time $\Delta t$ later, Eq. (2) gives us two equations in the five unknowns $A$, $B$, $p$, $q$ and $W$ at the time $\Delta t$. Additional relations are given by Eq. (1) applied along the physical plane characteristics provided these characteristics emanate from the line $t = 0$ along which the initial data are specified. It follows that if a strong discontinuity

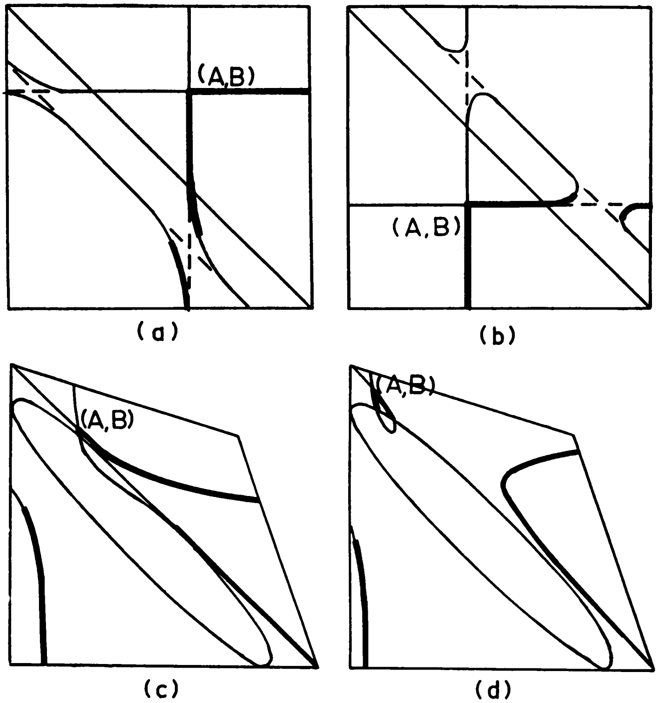


FIG. 2. The shock polars of points, $(A, B)$, above and below the ellipse are drawn for $\beta = 0$ and $\ll 1$ in (a) and (b). The dashed lines are the shock polars for $\beta = 0$, and the solid lines, for $\beta \ll 1$. Shock polars for a point, $(A, B)$ located just above and just below the line $p + q = 1$ are shown for $\beta = \frac{1}{3}$ in (c) and (d). The stable segments of the shock polars are shown by the heavy lines.

in initial data is to persist, three of the four characteristics issuing from the line $t = 0$ on either side of the shock path must intersect this shock path in the $(x, t)$ plane. If less than three characteristics intersect the shock path, we should look for wave type solutions of Eq. (1). If four intersect the shock path, the problem is overspecified. Physical plane characteristics originating from the left side of the shock will meet the shock path if their velocities are greater than the shock velocity $W$, whereas those on the right side of the shock path will meet it if their velocities are less than the shock velocity. It is possible to distinguish two types of stable shocks accordingly as two of the three characteristics that must cross a stable shock originate from the left or from the right of a shock. We refer to these as left or right shocks, respectively. If a shock is to be a stable left shock, the inequalities

$$w_+(A, B) > w_-(A, B) > W > w_-(p, q), \qquad w_+(p, q) > W \tag{9}$$

must be satisfied and to be a stable right shock, the inequalities

$$w_+(A, B) > W > w_+(p, q) > w_-(p, q), \qquad w_-(A, B) < W \tag{10}$$

must be satisfied where $w_+$ and $w_-$ represent the positive and negative roots of Eq. (4).

Since the stability criterion adopted here is based upon the properties of the physical plane characteristics, we can say nothing about the stability of shocks when either of the points $(A, B)$ or $(p, q)$ lies in the elliptic region of the hodograph plane.

**3. Discontinuous initial data.** In certain cases, as for example when a stoplight changes from red to green, we may expect the initial conditions to be discontinuous. We shall consider here solutions of the equations of continuity for initial data of the type $p(x, 0) = p_l$, $q(x, 0) = q_l$ for $x \leqq 0$ and $p(x, 0) = p_r$, $q(x, 0) = q_r$ for $x > 0$ where $p_l$, $p_r$, $q_l$, and $q_r$ are constants. Such initial data map into the hodograph plane as just two points, $(p_l, q_l)$ and $(p_r, q_r)$.

The equations of continuity and these initial conditions are invariant under the continuous group transformation $x' = \gamma x$, $t' = \gamma t$, with $\gamma$ any arbitrary constant and so we may expect to find solutions in which the dependent variables, $p$ and $q$, are functions of $x/t$ only [9]. In this case, an increase in time corresponds only to a rescaling of the $x$ axis. Thus, if a solution exists, its form will be the same for any positive time. Substitution of $y = x/t$ into Eq. (1) gives,

$$(1 - y - 2p - \beta q) \frac{dp}{dy} = \beta p \frac{dq}{dy}, \qquad (1 + y - 2q - \beta p) \frac{dq}{dy} = \beta q \frac{dp}{dy}. \tag{11}$$

Since Eq. (11) is homogeneous and linear in the derivatives, either $dp/dy = dq/dy = 0$, in which case the solutions, $p$ and $q$, are constants or the determinant of coefficients must vanish i.e.,

$$y = (1 - \beta/2)(q - p) \pm \{[1 - (1 + \beta/2)(p + q)]^2 - \beta^2 pq\}^{1/2}. \tag{12}$$

In the latter case, substitution of Eq. (12) into Eq. (11) gives the differential equation for the hodograph plane characteristics, Eq. (6), which has been solved in Sec. 2. Thus, for each hodograph characteristic there exists a special solution $p(y)$, $q(y)$ obtained by assigning to each point $(p, q)$ along this characteristic the value of $y$ given by Eq. (12). Whether one chooses the $+$ or $-$ sign in Eq. (12) depends upon which of the two families of characteristics the one in question belongs. The characteristics in Fig. 1 are labeled $+$ or $-$ accordingly as one should use the $+$ or $-$ sign in Eq. (12) and we shall refer to the corresponding solutions of Eq. (11) as plus or minus waves, respectively.

The complete solution of Eq. (11) is obtained now by considering all possible ways in which one can decompose the $y$-axis into intervals on each of which $p(y)$ and $q(y)$ are either constants or form plus or minus waves and at the boundaries of which $p(y)$ and $q(y)$ are either continuous or have stable shocks. For any shock or wave, $y$ is finite, so there must exist a finite $y$ to the left of which $p$ and $q$ are constants, $p = p_l$ and $q = q_l$, and another $y$ to the right of which $p = p_r$ and $q = q_r$.

The values of $y$ given by Eq. (12) are the velocities of the physical plane characteristics corresponding to given values of $p$ and $q$, as one can see by comparing Eq. (12) with Eq. (4) and (5a). The conditions for stability equations (9) and (10), therefore, take the special form in which $w_+ (p, q)$ is replaced by $y = y_+ (p, q)$ evaluated from Eq. (12) with the positive root, etc., and $W$ is the value of $y$ for the shock.

There are only a few combinations of waves and/or shocks that can be used to construct a permissible solution without violating the shock stability conditions or the requirement that $p$ and $q$ be single valued. Suppose, for example, that as $y$ increases, a plus wave were to terminate at some point where the densities are $(A, B)$ and the velocity $y$ is $y_+ (A, B)$. One could attach to the right of the plus wave, $y > y_+ (A, B)$, a region of constant densities but according to Eqs. (9) and (10), a shock from the densities $(A, B)$ to any other densities $(p, q)$ would have a velocity $W$ which is less than $y_+ (A, B)$, i.e., it would occur to the left of $y_+ (A, B)$. A minus wave starting at a point where the densities are $(A, B)$ would also have a velocity $y_-(A, B) < y_+(A, B)$, to the left, of $y_+(A, B)$, unless per chance $(A, B)$ lies on the ellipse where $y_- = y_+$. If $(A, B)$ does lie on the ellipse, there is the possibility that a plus wave can be followed by a minus wave, provided that the plus wave characteristic through $(A, B)$ approaches $(A, B)$ with $y$ increasing and the minus wave leaves $(A, B)$ with $y$ increasing. Figure 1 shows the direction of increasing $y$ for the various characteristics and one can see that the only characteristic having this property is the singular characteristic, Eq. (8), and the only point $(A, B)$ where this can happen is the point $(1/2, 1/2)$. We conclude from this that the only thing that can appear to the right of a plus wave is a region of constant densities unless the plus wave belongs to the singular characteristic and terminates with densities $(1/2, 1/2)$.

If one has a stable right shock from any density $(A, B)$ to densities $(p, q)$, it can be followed by nothing except a region of constant density because, by Eqs. (9) and (10) again, the shock velocity is larger than either $y_+(p, q)$ or $y_-(p, q)$. Similarly one can show that from a left shock, one may go to a region of constant densities which in turn may be followed by either a plus wave or a right shock but not another minus wave or left shock.

It follows from this that the only combinations of waves and shocks that can form permissible solutions are (1) a single wave or shock going from $(p_l, q_l)$ to $(p_r, q_r)$ with regions of constant density on either side, (2) a minus wave of left shock going from $(p_l, q_l)$ to some densities $(p_i, q_i)$ followed by a region of constant densities $(p_i, q_i)$ and

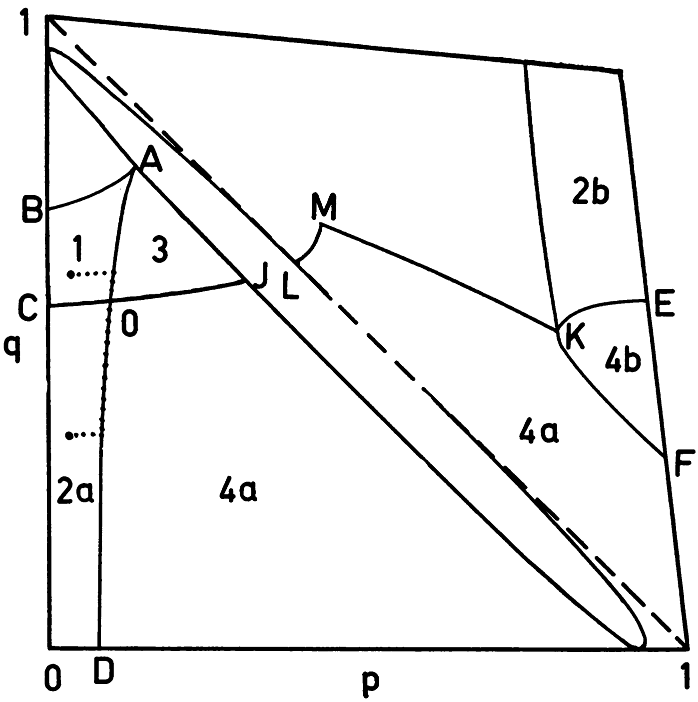


FIG. 3. Points in numbered regions, shown for $\beta = 0.1$, can be reached by various allowed solution sequences starting from some point 0 below the ellipse. The dotted curves represent the maps of sample solutions.

then either a plus wave or a right shock to $(p_r, q_r)$ (3) a minus wave or left shock going from $(p_l, q_l)$ to a point on the singular characteristic followed by a region of constant density, then the plus wave to minus wave sequence described above, another region of constant densities, and then possibly also a right shock or plus wave to $(p_r, q_r)$.

Figure 3 shows the various domains in the hodograph plane that can be reached by these allowable combinations of waves and/or shocks from a point $(p_l, q_l)$ designated in Fig. 3 as the point 0 located below the ellipse. The line $0A$ is the image of a minus wave, a hodograph characteristic through 0 in the direction of increasing $y-$, and the line $0C$ represents a plus wave, a hodograph characteristic through 0 in the direction of increasing $y_+$. The lines $0D$ and $EK$ depict those segments of the shock polar which are the loci of stable left shocks from the point 0, and the lines $0J$ and $LM$ represent stable right shocks from 0. Points on these lines can be reached by a single wave or shock from 0.

Any point in region 1 of Fig. 3 can be reached from 0 by a combination of a minus wave along $0A$ followed by a plus wave into region 1. The boundary $AB$ of this region is the plus characteristic through $A$. A sample path is sketched in Fig. 3, and the physical plane solution corresponding to it is shown in Fig. 4(a). All points in regions 2a in Fig.

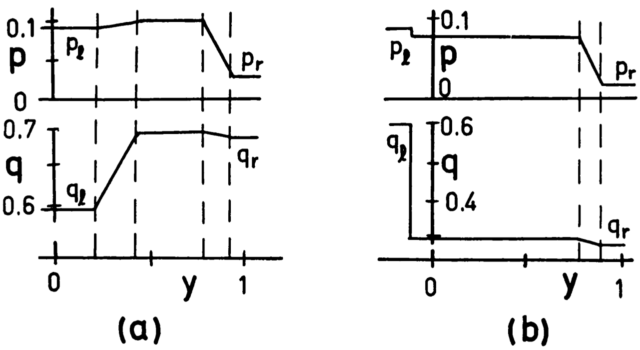


FIG. 4. The two solutions whose hodograph maps are shown in Fig. 3 are represented by plotting $p$ and $q$ as functions of $y = x/t$. Figure 4 (a) shows a minus wave followed by a plus wave and Fig. 4(b) a left shock followed by a plus wave.

3 can be reached by a combination of a left shock along the line $0D$ followed by a plus wave into region 2a. A sample path is shown in Fig. 3, and the corresponding physical plane solution in Fig. 4(b). All points in the region 2b can be reached with a left shock from 0 to the line $EK$ followed by a plus wave into region 2b and all points in region 3 can be reached with a minus wave followed by a right shock.

Points in region 4a can be reached with a left shock along the line $0D$ followed by a right shock, and points in 4b by a left shock to the line $EK$ followed by a right shock. The line $MK$ is the locus of the end points of the stable shock polars for right shocks from $0D$. As one approaches any point on the boundary $KF$ between these regions, one finds that the velocities of the two shocks used to reach this point become equal and the two shocks become a single shock. The line $KF$ is, therefore, part of the shock polar determined by 0 but it is an unstable part because four physical plane characteristics cross the shock path rather than the usual three.

If we set initial data with $(p_l, q_l)$ at the point 0 as in Fig. 3, we may set $(p_r, q_r)$

anywhere in regions 1 to 4 and still obtain a stable solution. For any given $(p_l, q_l)$, however, stable solutions are not assured for all values of $(p_r, q_r)$. Regions 1 to 4 do not cover the elliptic region or even all of the hyperbolic regions of the hodograph plane.

When there is no coupling between lanes ($\beta = 0$) the densities are independent of each other, and every point in the hodograph plane may be reached from every other point by the solutions of two Richards-type problems. That there exist regions in the hodograph plane that cannot be reached from an arbitrary point $(p_l, q_l)$ by any allowable combination of two shocks and or waves for $\beta > 0$ is due to the fact that these particular waves cannot be extended through the ellipse.

The diagrams analogous to Fig. 3 for various other positions of the point 0, such as above the ellipse or between the ellipse and the singular characteristic, will have quite different forms. Whereas there was no opportunity to use the singular characteristic in any solution sequence of Fig. 3, it does appear for some other positions of the point 0. That one can not reach every point in the hodograph plane with an allowable sequence of waves and shocks from 0 is, however, true for all positions of 0.

As a specific illustration of a solution of the above form one might consider what happens when any type of bottleneck, a slow car, a road block, traffic light, etc. is suddenly removed at time 0. Suppose, for example, that a traffic light has been red for a sufficiently long period of time that long queues have developed in both lanes, and that a steady flow of traffic is entering the highway from the side roads. If at time zero, the light turns green, the initial conditions on the left of the stop-light will be represented by some point along the right-hand boundary of the hodograph plane, and the initial conditions to the right of the stop-light will map into a point on the upper boundary. The solution for this type of initial data is shown in Fig. 5 and consists of a minus wave to the singular characteristic, a wave solution along the singular characteristic and then a plus wave, each separated by regions of constant density. This solution will be valid, however, only until such time when the light turns red again, the queues of traffic become

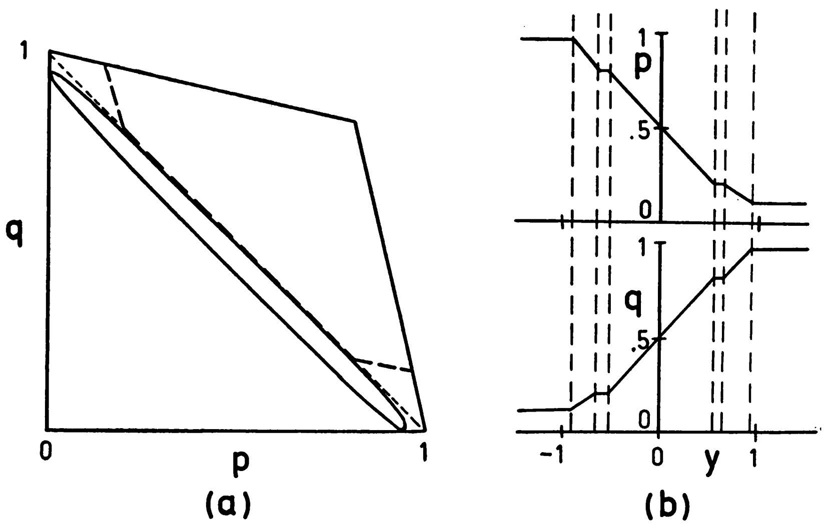


FIG. 5. The hodograph map shown by the broken line in (a) for $\beta = 0.1$ and the physical plane solution ($p$ and $q$ vs $y$) in (b) represent the flow of traffic immediately after a stop light turns green.

empty, or interactions occur between the waves resulting from successive cycles of the traffic light.

**4. Other types of initial data.** There are many types of initial data other than the discontinuous type described in Sec. 3 that yield relatively simple solutions. We consider here a few other examples. We will assume, however, that the initial data are such that the solutions lie entirely within the hyperbolic region of the hodograph plane. We have not been able to determine what should happen if the initial data lie partly in the elliptic region or if the solution touches the boundary of the elliptic region at some later time.

One simple class of solutions known as simple waves [6] arises if the image of the initial data in the hodograph plane lie entirely along a single hodograph characteristic. In this case, one can show that unless shocks develop the solution maps into this same hodograph characteristic for all times and that $p$ and $q$ are constant along the physical plane characteristics. The velocity $w$ of the physical plane characteristics, Eq. (4), being a function of only $p$ and $q$ will also be constant and to obtain $w$, one chooses the plus or minus sign in Eq. (4) accordingly as the wave lies on a plus or minus hodograph characteristic.

To construct the solution from the initial data $p(x, 0)$, $q(x, 0)$, one draws through each point $x$ at time 0, the straight line physical plane characteristic with the velocity $w$ determined by the value of $p(x, 0)$, $q(x, 0)$ and assigns the same $p$ and $q$ to all points on this characteristic. If none of these physical plane characteristics intersect, the complete solution is thereby uniquely determined. If some characteristics do intersect, a shock develops at the earliest time this happens and the shock path must, in most cases, be determined numerically.

Suppose now we assume that the initial data do not lie on a hodograph characteristic but that $p(x, 0)$ and $q(x, 0)$ are constants for $x \leqq 0$ and for $x \geqq x_0$ with values $(p_1, q_1)$ and $(p_2, q_2)$ respectively, and that for $0 < x < x_0$, the map of the initial data into the hodograph plane lies on an arc between $(p_1, q_1)$ and $(p_2, q_2)$ which is never tangent to a hodograph characteristic and on which $x(p, q)$ is single valued. Initial data of this type and its image in the hodograph plane are illustrated in Figs. 6(a) and (b).

To find the solution in this case one proceeds in the following way. From Eq. (3) we calculate the values of the partial derivatives, $t_p$ and $t_q$, on the initial data curve and from these deduce which side of the initial data curve corresponds to positive times. If one of the hodograph characteristics drawn through the point $(p_1, q_1)$ intersects one of the characteristics drawn through $(p_2, q_2)$ and the intersection, $(p_3, q_3)$, lies on the side of the initial data curve corresponding to positive time, these three curves will enclose a domain of dependence in the hodograph plane as illustrated in Fig. 6(b). Such intersections are not always assured since the characteristics may end at the perimeter of the ellipse without intersecting each other. If such a domain does exist, however, the values of $t(p, q) \geqq 0$ and $x(p, q)$ can be determined for all points inside this domain using the method of characteristics [10], or some alternative scheme.

This solution can be mapped into the physical plane by drawing the curves of constant time in the hodograph plane. Along these curves $p$ and $q$ may be considered as functions of $x$ and be plotted as such for various times. The limiting values of $x$ and $t$ obtained as we approach the boundary of the hodograph domain of dependence from its interior define two curves in the $(x, t)$ plane which emanate from the points $(0, 0)$ and $(x_0, 0)$ and have an intersection at the value of $x$ and $t$ corresponding to the point

$(p_3, q_3)$ in Fig. 6(b). These curves are the physical plane characteristics and define the physical plane domain of dependence illustrated by the area $0ax_0$, region 3, in Fig. 6(c). The solution in the hodograph domain of dependence determines the solution everywhere in the physical plane domain of dependence.

To find the physical plane solutions outside the domain of dependence, we make use of the fact described above that if the map of a physical plane solution coincides with a hodograph characteristic, the solution is a simple wave. Since the solutions along

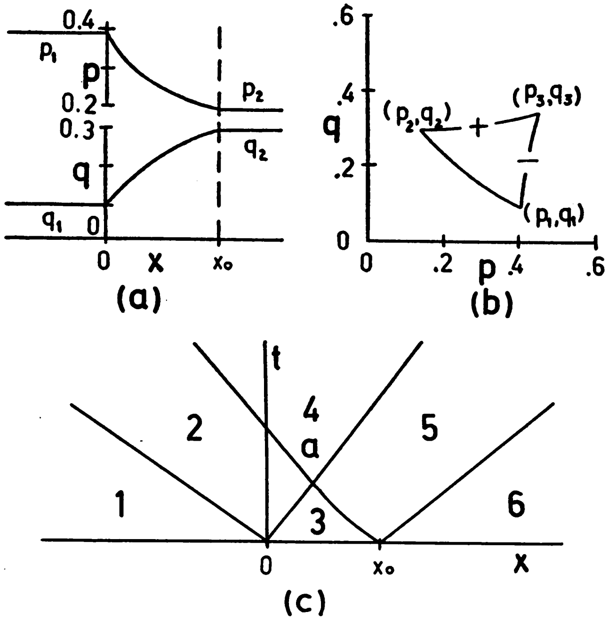


FIG. 6. An example of continuous initial data, together with their hodograph and physical plane domains of dependence are shown in (a), (b) and (c) respectively. The solutions, $p$ and $q$, are constants in regions 1, 4 and 6 of (c) and simple waves in regions 2 and 5.

$0a$ and $x_0a$ of Fig. 6(c) map into the hodograph characteristics, $(p_1, q_1) - (p_3, q_3)$ and $(p_2, q_2) - (p_3, q_3)$, respectively, it is possible to continue the physical plane solution to points outside the region 3 by means of simple waves. From any point on $0a$, for example, at which $p$ and $q$ have been found, one can draw the appropriate straight line characteristic and assign to every point on this characteristic outside region 3 the same values of $p$ and $q$. This gives the solution in region 2 of Fig. 6(c). Similarly the solution in region 5 is obtained from a simple wave originating from $x_0a$. In regions 1, 4 and 6, $p$ and $q$ will be constant and have the values $(p_1, q_1)$, $(p_3, q_3)$ and $(p_2, q_2)$, respectively.

Since the solution in the hodograph domain of dependence can be found by means of the method of characteristics, we can always obtain an approximate solution for this type of initial data at least by numerical means. For certain types of initial data, however, one can also obtain the solution analytically. If $p(x, 0)$ and $q(x, 0)$ are linear in $x$ for $0 < x < x_0$, there exist solutions of Eq. (1) having the form

$$p(x, t) = a_1(t) + b_1(t)x, \qquad q(x, t) = a_2(t) + b_2(t)x. \tag{13}$$

If one substitutes this into Eq. (1), one obtains the four ordinary differential equations

$$b_1 - 2a_1b_1 - \beta(a_1b_2 + a_2b_1) = -da_1/dt, \tag{14a}$$

$$b_2 - 2a_2b_2 - \beta(a_1b_2 + a_2b_1) = da_2/dt, \tag{14b}$$

$$2b_1(b_1 + \beta b_2) = db_1/dt, \qquad 2b_2(b_2 + \beta b_1) = -db_2/dt, \tag{14c, d}$$

for $a_1(t)$, $b_1(t)$, $a_2(t)$ and $b_2(t)$ whose values at $t = 0$ are known from the initial conditions. It is even possible to find the complete analytic solution of Eq. (14). If we divide Eq. (14c) by (14d), the resulting equation is homogeneous of degree zero and can be integrated to give a relation between $b_1(t)$ and $b_2(t)$. From this one can then find the solutions $b_1(t)$ and $b_2(t)$ of Eqs. (14c, d). Substitution of this into Eqs. (14a, b) then gives a pair of simultaneous linear differential equations which also can be integrated.

In most cases, the solutions of Eq. (14) can not be expressed in terms of elementary functions, but there are two special cases in which the solutions are very simple. If the hodograph map of the initial data is parallel with the singular characteristic, $b_1(0) = -b_2(0)$, or parallel with either axis, $b_1(0) = 0$ or $b_2(0) = 0$, then it remains this way $b_1(t) = -b_2(t)$, $b_1(t) = 0$, or $b_2(t) = 0$, respectively, for all time. The complete solutions for these cases are given in [8].

The above solutions apply only within the physical plane domain of dependence, region 3 of Fig. 6(c), or equivalently in the hodograph domain of dependence determined by the two characteristics through $(p_1, q_1)$ and $(p_2, q_2)$. The solutions outside these domains are again found by using simple waves. The solutions for two examples with $b_2(t) = 0$ are illustrated in Figs. 7 and 8.

In Fig. 7, waves start to develop at $x = 0$ and $x = x_0$ at time zero. At time $t_1$, $p$ and $q$ each show a region of constant density, a wave, a region of linearly varying density, another wave, and finally constant density again in that order as one goes from left to right. The waves develop at the expense of the linear region until at time $t_3$, the latter

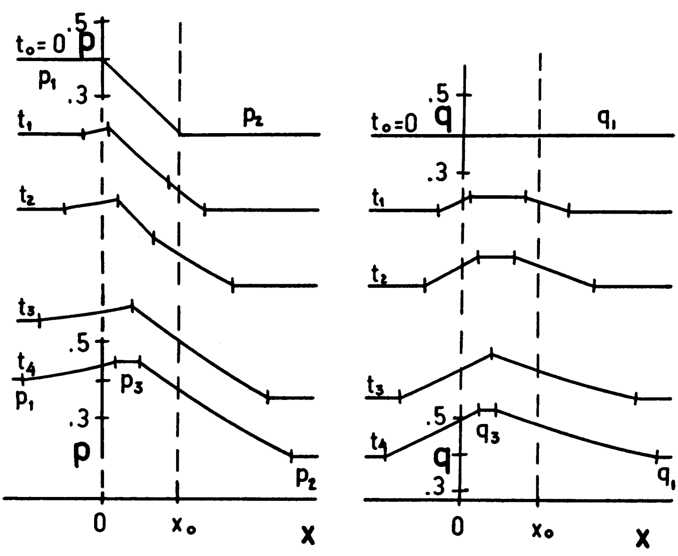


FIG. 7. The solution corresponding to the data shown for $t = 0$ is represented by plotting $p$ and $q$ as a function of $x$ for several values of the time $t_0 = 0 < t_1 \cdots < t_n$.

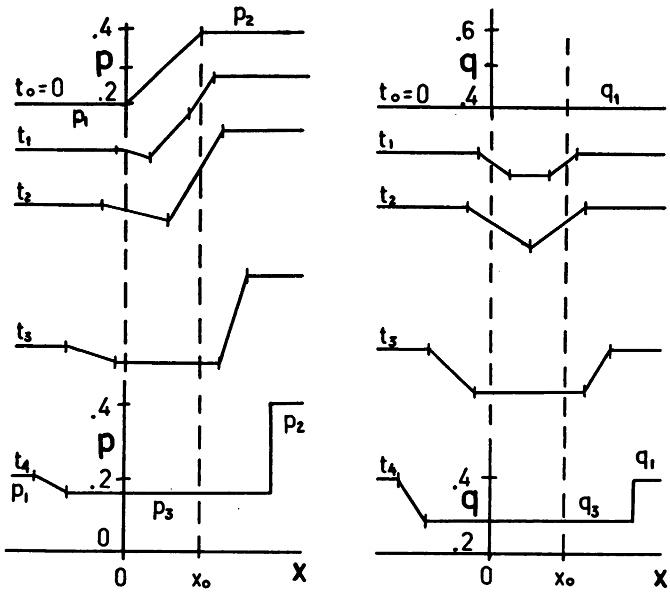


FIG. 8. This figure is the same as Fig. 7 except that the initial data lead to the formation of shocks. The final configuration (not shown) eventually becomes two diverging shocks.

region disappears. At time $t_4$, we have two waves separated by a region of constant density $(p_3, q_3)$, and for subsequent times the two waves continue to travel away from each other and spread over wider ranges of $x$.

The analogous situation but with $p(x)$ increasing with $x$ at $t = 0$ is shown in Fig. 8. At time $t_2$, the region of linearly varying density has vanished and at time $t_3$ we have two waves separated by a region of constant density $(p_3, q_3)$. The two waves, however, are of the type that develop shocks and one of them is about to form at time $t_4$. Eventually one obtains a configuration with just two shocks separated by a region of constant density. The constant densities are, however, determined by the intersection of the two shock polars for $(p_1, q_1)$ and $(p_2, q_1)$ and are not $(p_3, q_3)$ which are determined by the intersection of the characteristics through $(p_1, q_1)$ and $(p_2, q_1)$. This final state is achieved by means of waves traveling back and forth between the two shocks, decreasing in amplitude and increasing in width with each trip until the densities are nearly constant. The details of this are rather tedious to follow but for $\beta \ll 1$ these secondary waves will be of very small amplitude and are almost completely absorbed each time they overtake a shock.

**5. Acknowledgments.** The authors are indebted to Professor W. Prager for suggesting this problem and for contributing some results of his preliminary investigations of it, and to Professor R. Meyer for many very helpful suggestions.

## REFERNCES

1. M. J. Lighthill and G. B. Whitham, *On kinematic waves.* II. *A theory of traffiic flow on long crowded roads,* Proc. Roy. Soc. **A229**, 317–345 (1955)
2. P. I. Richards, *Shock waves on the highway,* J. Opns. Research Soc. Amer. **4**, 42–51 (1956)
3. S. C. De, *Kinematic wave theory of bottlenecks of varying capacity,* Proc. Cambridge Phil. Soc. **52**, 564–572 (1956)

4. K. Ostwatitsch, *Über die Charakteristikenverfahren der Hydrodynamik*, Z. angew. Math. Mech. **25**, 195–208 (1947)
5. P. D. Lax, *Non-linear hyperbolic equations*, Communs. Pure and Appl. Math. **6**, 231–258 (1953)
6. R. Courant and K. O. Friedrichs, *Supersonic flow and shock waves*, Interscience Publishers, New York, 1948, Chaps. II and III
7. E. Kamke, *Differentialgleichungen, Lösungsmethoden und Lösungen*, Akademische Verlagsgesellschaft, Leipzig, 1944, p. 32
8. J. H. Bick and G. F. Newell, *A continuum model for traffic flow on an undivided highway*, Brown University report IBM-22 (mimeographed), 1958
9. G. Birkhoff, *Hydrodynamics*, Dover Publications, New York, 1955, Chap. IV
10. I. Petrovsky, *Lectures on partial differential equations*, Interscience Publishers, New York, 1954, p. 68