Title

A Variational Formulation of Kinematic Wave Theory

Author

Daganzo, Carlos F.

Publication Date

2003-05-01

Institute of Transportation Studies
University of California at Berkeley

A Variational Formulation of Kinematic Wave Theory

Carlos F. Daganzo

RESEARCH REPORT
UCB-ITS-RR-2003-3

May 2003
ISSN 0192 4095

A Variational Formulation of Kinematic Wave Theory

Carlos F. Daganzo
Institute of Transportation Studies and Department of Civil and Environmental Engineering
University of California, Berkeley CA 94720
(May 23, 2003)

ABSTRACT

This paper proves that kinematic wave (KW) problems with concave (or convex) equations of state can be formulated as calculus of variations problems. Every well-posed problem of this type, no matter how complicated, is reduced to the determination of a shortest tree in a relevant region of space-time where "cost" is predefined. A duality between KW theory and "least cost networks" is thus unveiled.

In the new formulation space-time curves that constrain flow, such as sets of moving bottlenecks, become space-time shortcuts. These shortcuts become part of the network and affect the nature of the solution but not the speed with which it can be obtained. Complex boundary conditions are naturally handled in the new formulation as constraints/shortcuts of this type.

1. INTRODUCTION

This paper is concerned with problems, e.g., traffic flow, that are described by the following first order partial differential equation for an unknown function N of two arguments (t, x), usually associated with time and space:

$$N_t = Q(-N_x , t, x). \tag{1}$$

The subscripts t and x denote partial derivatives, and Q is a piecewise differentiable and concave function in its first argument.

In applications, N is often interpreted as a cumulative count of objects: cars, fluid units, etc. The rate at which the item count changes with time is the flow, $q = N_t$, and the rate at which it changes with distance the negative of the density, $k = -N_x$. Thus, Eq. (1) is often written in terms of q and k as:

$$q = Q(k , t, x) \tag{2a}$$

where q and k are linked by a "conservation equation" which merely expresses the equality $N_{xt} = N_{tx}$:

$$k_t + q_x = 0. \tag{2b}$$

A third way of writing (1) in terms of the density alone is obtained by substituting (2a) for q in (2b):

$$k_t + Q_k k_x + Q_x = 0. \tag{3}$$

Equations (1), (2) and (3) are equivalent. Formulation (1) is natural for certain geophysics problems (see e.g., Luke, 1973) and formulations (2) and (3) for fluids (see e.g., Lighthill and Whitham, 1955, and Richards, 1956). Formulations (2) and (3), called "conservation laws," have been extensively studied; see Lax (1973) and LeVeque (1992) for background.

Newell (1993) has shown the considerable practical advantage of using (1) for solving traffic flow problems. Independently, he and Luke proposed that if N has been defined on a boundary $\boldsymbol{D}$, then the value of N at a point P is given by the rule:

$$N_P = \min \{B(\mathscr{P}) + \Delta(\mathscr{P}) : \ \forall \mathscr{P} \in \mathbf{W} \cap \mathbf{P}_P\} \tag{4}$$

In this expression and elsewhere in this paper $\mathscr{P}$ is a space-time path, $x(t)$, $\mathbf{P}_P$ is the set of all paths from $\boldsymbol{D}$ to P and $\mathbf{W}$ is the set of all wave paths. The functionals $B(\mathscr{P})$ and $\Delta(\mathscr{P})$ respectively give the N-value at the beginning of the path (a data point) and the predicted change in N along the wave path. The formula for the latter is:

$$\Delta(\mathscr{P}) = \int_{t_B}^{t_P} R(x', t, x)dt \qquad \text{if } \mathscr{P} \in \mathbf{W} \tag{5}$$

where a prime denotes differentiation, t_B and t_P are the times at the two ends of $\mathscr{P}$, and R is a known function that can be derived from Q. [In traffic flow, this function gives the rate at which vehicles pass an observer traveling with the wave, at speed x'.]

Recall from the theory of conservation laws that a wave path is defined by the property $x'(t) = Q_k$ and that if a problem is physically well-posed (no contradicting data) a wave path can be drawn from at least one boundary point to P. Thus, the set $\mathbf{W} \cap \mathbf{P}_P$ is not empty. There can be more than one wave reaching point P, however. Luke (1973) and Newell (1993) proposed the minimum operation as a way of selecting the correct wave but provided no mathematical proof.

While (4) represents a significant advance over previous methods to solve (1)-(3) it is still cumbersome for general problems because identifying the relevant set of paths $\mathbf{W} \cap \mathbf{P}_P$ is not easy, except in special cases. [Both Luke and Newell successfully applied the minimum principle to homogeneous, time-independent problems where $Q(k, t, x) = Q(k)$, because in these instances waves are straight lines and the passing rate R is constant along each wave. But application of the principle to general problems is tedious, as illustrated by the solutions of Lighthill and Whitham's bottleneck problem in Newell (1999).]

This paper will present a mathematical proof that (4) holds, putting the L-N minimum principle on a firm foundation. But more importantly, it will prove that (4) can be replaced by:

$$N_P = \min \{B(\mathscr{P}) + \Delta(\mathscr{P}) : \forall \mathscr{P} \in \mathbf{V} \cap \mathbf{P}_P\} \qquad (6)$$

where $\mathbf{V}$ is the set of all "valid" paths. A path is valid if it is continuous, piecewise differentiable and such that x' is everywhere in the range of possible wave speeds. Note that $\mathbf{V} \supset \mathbf{W}$. The significance of enlarging the set of paths is that, although $\mathbf{V} \cap \mathbf{P}_P$ is large it is convex. Convexity opens the door to variational methods, which cannot be used with (4). In essence, the new formulation reduces even the most complicated KW problems to simple shortest path problems. Section 2, below, proves (6) and Sec. 3 discusses its implications.

2. THE "LEAST ACTION" PRINCIPLE FOR KINEMATIC WAVES

The solution of (2) at a given point (t, x) is characterized by two values of k and q that satisfy (2a). Associated with these are two derived quantities: the wave speed,

$$u = Q_k(k, t, x), \qquad (7)$$

which is a monotonic function of k, and the (passing) rate at which items overtake an observer moving with the wave,

$$r = q - ku. \qquad (8)$$

We see from (2a) and (7) that r is a function of k:

$$r = Q(k, t, x) - kQ_k(k, t, x), \tag{9a}$$

and from (7) that the latter is also an implicit function of u:

$$r = R(u, t, x). \tag{9b}$$

We can write the change in item number along a wave path from B to P, $N_P - N_B$, by integrating (9b) over time. The result is (5), as expected. If $\mathscr{P} \notin \mathbf{W}$, however, $N_P - N_B$ may differ from $\Delta(\mathscr{P})$. The formula is obtained by integrating the passing rate, $q - kx'$, along the path:

$$N_P - N_B = \int_{t_B}^{t_P} [Q(k,t,x) - kx']dt \tag{10}$$

We now show that the difference between the change in item number (10) and $\Delta(\mathscr{P})$ can never be positive.

<u>Lemma</u>: If $\mathscr{P} \in \mathbf{V}$ goes from B to P then, $\Delta(\mathscr{P}) \geq N_P - N_B$.■

<u>Proof:</u> We first show that for any given speed x', the relative flow $Q(k, t, x) - kx'$ in the argument of (10) is maximized for a density with wave speed, $u = x'$. This is true because the relative flow is a concave function of k and therefore a sufficient condition for a *global* maximum is $Q_k = x'$. Note next that if we now substitute Q_k for x' in the relative flow expression we find that the result (the maximum possible passing rate) is the right side of (9a). Since (7) holds this is equal to the right side of (9b). Since the optimum wave speed is $u = x'$, we can write the maximum as $R(x', t, x)$. Thus, for any k and any x', $R(x', t, x) \geq Q(k, t, x) - kx'$. Since the integral of the left side of this inequality is $\Delta(\mathscr{P})$ and the integral of the right side is $N_P - N_B$, it follows that $\Delta(\mathscr{P}) \geq N_P - N_B$.■

It is well known that if the boundary data are well posed then there is a wave that reaches every point in the solution domain; see e.g., Lax (1973). The main result of this paper can now be presented.

<u>Theorem</u>: Equations (6) and (4) hold for well-posed problems.■

<u>Proof</u>: The lemma states that $N_P \leq B(\mathscr{P}) + \Delta(\mathscr{P})$ for all $\mathscr{P} \in \mathbf{V} \cap \mathbf{P}_P$. Thus, to prove (6) it suffices to show that there is a critical path $\mathscr{P} \in \mathbf{V} \cap \mathbf{P}_P$ for which $N_P = B(\mathscr{P}) + \Delta(\mathscr{P})$. Since the problem is well posed, there is at least one path $\mathscr{P}^*$ in $\mathbf{W} \cap \mathbf{P}_P$. For any such path, $N_P = B(\mathscr{P}^*) + \Delta(\mathscr{P}^*)$; see (5). Since $\mathbf{W} \cap \mathbf{P}_P \subset \mathbf{V} \cap \mathbf{P}_P$, $\mathscr{P}^*$ is a critical path in $\mathbf{V} \cap \mathbf{P}_P$ and (6) must hold. Since the critical path is also in $\mathbf{W} \cap \mathbf{P}_P$, (4) must obviously hold as well.■

Note that the lemma and the theorem do not apply to problems with non-concave Q.

3. DISCUSSION

<u>Uniqueness and stability</u>: If we use Δ_{BP} to denote the minimum of $\Delta(\mathscr{P})$ across all valid paths from B to P (which are fixed quantities that can be interpreted as "costs") then (6) becomes:

$$N_P = \min \{ N_B + \Delta_{BP} : \ \forall B \in \boldsymbol{D} \}. \tag{11}$$

This expression is useful in cases where the Δ_{BP} can be easily calculated. Then the calculus of variations problem becomes an ordinary optimization problem. More importantly, however, (11) implies that if $\{N_B\}$ and $\{N'_B\}$ are two data sets, then the two solutions must satisfy:

$$|N_P - N'_P| \ \leq \ \max \{ |N_B - N'_B| ; \forall B \in \boldsymbol{D} \}. \tag{12}$$

This inequality shows that two solutions can never deviate from each other at any point any more than they deviate somewhere on the boundary. This means that perturbations to the data cannot grow into the solution and that the solution to (6) must be *unique* and *stable*. This proof of uniqueness and stability is considerably more concise than the conventional one; see Lax (1973).

<u>"Least action"</u>: If we think of $R(x', t, x)$ as a Lagrangian and (2) as the "action" we see that kinematic waves are least action paths that can be found with calculus of variations. We then find after some manipulation of (6) and (5) that the Euler equation for the problem is indeed (2). Since the solution to the calculus of variations problem is stable and unique, (6) can be taken as the fundamental statement of the kinematic wave problem. Unlike (2) and (3), which have to be augmented with an auxiliary "entropy condition," (6) automatically rules out unstable solutions.

<u>Networks, duality and boundary constraints</u>: The above results show that KW problems can be solved approximately by first overlaying a dense but discrete network in the solution region, with a cost $\Delta_{PP'}$ for each arc PP', then connecting a fictitious origin to all points B on the boundary with a cost N_B for each arc, and finally finding the "shortest" tree from this origin to all nodes. The network solution will be exact if the network can be guaranteed to contain one of the shortest paths in the continuum. A sequel to this paper will show how this can be done for an important class of problems. In other cases we can ensure that the network contains a near-optimum path by ensuring that it is dense and that the set of links PP' incident on every node P contains a full complement of slopes, x', within the range of validity. Since networks of this type contain no cycles, the full tree for a network with L links can be found in time $O(L)$; e.g., with dynamic programming.

The tree of shortest paths from the origin to all the nodes gives the waves. Its branches end either at the boundary of the solution region or at a "shock." Shocks are those points in solution space that can be reached by more than one path. The equi-cost contours of our network are lines with the same item number—they are the item trajectories.

Formulation (6) has a further advantage over the conventional KW formulations because it provides a natural framework for the treatment of discrete constraints (e.g., due to point bottlenecks in traffic flow) without the nuisance of having to check for stability during the solution process. To clarify, suppose we are given a set of paths $\{C_i(t)\}$ along which there is a maximum possible passing rate $\{R_i(t)\}$. (Think of a snowplow passing through a traffic signal.) To solve this problem we simply add short links to the underlying network matching the $\{C_i(t)\}$ as well as possible and then assign to them unit costs (per unit time) close to $R_i(t)$. The problem (with constraints) is then solved by finding the shortest tree for the expanded network. The new low-cost links act as shortcuts through space-time. These shortcuts obviously change the solution but do not materially alter the solution effort. Note that the inclusion of shortcuts can only lower the cost of reaching a node. This is expected, since in KW theory, the inclusion of a bottleneck can only *lower* the item number reaching a node. Boundary conditions can be naturally modeled in this theory as constraints of this type. For example, constraints of the form $\{C_f(t) = x_f\,;\, R_f(t)\}$ can be used to model a road that meets a junction at location x_f.

We may ask if this solution with constraints is stable, unique, etc… but we need not worry. Since the solution continues to be the minimum of a shortest path problem it is unique and stable; it is the correct solution and no entropy conditions need to be checked. In practical applications once can allow the passing rates to depend on endogenous data (e.g., from other roads sharing a junction) according to some meaningful rule. In this case, the rules should be tested for stability. But if they are stable boundary data will remain bounded, and the solution to the continuum problem will still exist, be unique and stable.

A sequel to this paper will present a number of examples and will discuss further simplifications, with emphasis on cases where *exact* solutions can be obtained in time O(M), where M is the number of points at which the solution is sought.

REFERENCES

Lax, P.D. (1973) *Hyperbolic systems of conservation laws and the mathematical theory of shockwaves.* SIAM Regional Conference Series in Applied Mathematics, J.W. Arrowsmith Ltd., Bristol, U.K.

LeVeque, R.J. (1992), *Numerical methods for conservation laws*. (2nd edition), Birkhauser-Verlag, Boston, MA.

Lighthill, M.J. and G.B. Whitham (1955) "On kinematic waves. I flow movement in long rivers. II A theory of traffic flow on long crowded roads." *Proc. Roy. Soc. A* **229**, 281-345.

Luke, J. C. (1972) "Mathematical models for landform evolution." *J. Geophys. Res.* **77**, 2460-2464.

Newell, G.F. (1993) "A Simplified Theory of Kinematic Waves in Highway Traffic I: General Theory. II: Queuing at Freeway Bottlenecks. III. Multi-Destination Flows." *Trans. Res.* **27B**, 281-287.

Richards P.I. (1956) "Shockwaves on the highway." *Opns. Res.* **4**, 42-51.

Newell, G.F. (1999) "Flows upstream of a highway bottleneck." *Transportation and Traffic Theory* (A. Ceder, editor), pp. 125-146, Pergamon, Amsterdam, NL.