Title
A Simple Traffic Analysis Procedure

Author
Daganzo, Carlos F.

Publication Date
1997

Institute of Transportation Studies
University of California at Berkeley

A Simple Traffic Analysis Procedure

Carlos F. Daganzo

WORKING PAPER
UCB-ITS-WP-1997-4

Publications in the working paper series are issued for discussion and are not considered final reports. Comments are invited.

December 1997
ISSN 0192 4141

A SIMPLE TRAFFIC ANALYSIS PROCEDURE

Carlos F. Daganzo
Department of Civil Engineering and
Institute of Transportation Studies
University of California, Berkeley CA 94720
(December 5, 1997)

Abstract

This paper presents a simple approximate procedure for traffic analysis that can be described geometrically without calculus. The procedure, which is graphically intuitive, operates directly on piecewise linear approximations of the N-curves of cumulative vehicle count. Because the N-curves are both readily observable and of direct interest for evaluation purposes (e.g., they yield the total vehicle-hours and vehicle-miles of travel in a time interval, and the vehicular accumulation as a function of time) the predictions made with this method should be practical and easy to test.

Queued traffic is treated first. Predictions of cumulative counts for this traffic regime are based on three mild assumptions stating the relationship between the linearized input and output curves of vehicle count. These assumptions are a simple twist on those of the kinematic wave (KW) model of Lighthill and Whitham (1955) and Richards (1956); thus, the part of the proposed procedure that deals with queued traffic is just an efficient way of solving the KW model based on readily observable data. The method is somewhat more complicated but also more general than that in Newell (1993). The paper shows that the particular way in which the data are linearized prior to the procedure is not important provided one works within certain tolerances, and that errors in the model parameters also have a limited effect. These results can help discriminate between model errors and calculation errors in a validation effort.

The second part of the paper examines traffic streams that include queued and unqueued traffic, as well as bottlenecks. Here, unqueued traffic is allowed to obey any reasonable model, e.g., as could be the result of a simulation with multiple vehicle classes, drivers that wish to travel at different speeds and certain rules for passing. The unqueued model, however, cannot include vehicle classes with speeds lower than the maximum possible in queued traffic. The bottlenecks are assumed to have a well-defined capacity and always to allow the maximum possible flow consistent with: (i) the availability of upstream traffic, (ii) the presence of a downstream queue and (iii) the capacity. The capacity can be time-dependent and endogenous. The paper presents the theory, some examples and a computational framework.

1. BACKGROUND AND MOTIVATION

The kinematic wave (KW) model of Lighthill and Whitham (1955) and Richards (1956), is the simplest theory of traffic dynamics that recognizes that vehicles take space. However, the KW theory has two (relatively minor) deficiencies: (i) it does not model unqueued (light) traffic properly because it does not account for passing, and (ii) it models traffic inside queues as being smooth-flowing when the reality is sometimes different. The KW theory is also difficult to test because it is based on variables such as density that are difficult to measure and even define. In view of this, it seems that a slight modification of the theory that would alleviate some of these deficiencies while being based on readily observable data could be more reliable and easier to test.

Newell's simplification of the KW model (Newell, 1993), which assumed a concave flow-density relation and used waves to operate on curves of cumulative count was a first step in this direction.[1] The queue analysis methodology in Lawson et al (1997) for time-independent bottlenecks, which does away with waves, density functions and local flow-density relations altogether, represents yet another step; its predictions are consistent both with KW theory and (approximately) with certain forms of traffic instability. A combination of the ideas in these two references has been applied to the time-dependent case with concave flow-density relations (Erera et al, 1997).

The present paper relaxes the requirement of a concave flow-density relation and extends the ideas in these references further, so as to address some of the above-mentioned concerns. It presents an approximate macroscopic model of traffic behavior for an inhomogeneous highway with bottlenecks that overcomes deficiency (i) but not (ii). Of course, a model's inability to capture the details of traffic behavior inside queues (ii) does not mean that it cannot predict accurately important measures of performance such as total delay, time-dependent vehicle trip times and accumulations. Thus, the paper suggests possible tests for this hypothesis. It is also shown that the model predictions of these relevant measures are insensitive to input data errors.

The paper is divided into two main parts: Sec. 2 which focuses on queued traffic, and Sec. 3 which extends the results to mixed traffic. Section 2 shows how to predict the cumulative N-curve of vehicle count at a particular location from downstream data at another location when

[1] Validation work is now proceeding to see whether waves behave as predicted in this theory--see Cassidy and Windover (1995,1996).

traffic is queued. The method is just a fast and intuitive way of solving a kinematic wave (KW) problem. The effect of this approximation, e.g., on calculation errors and the interpretation of field tests, is examined and discussed. Section 3 then extends the model to mixed (queued/unqueued) traffic for highways with time-dependent bottlenecks. A generic, non-KW, model of unqueued traffic including multiple vehicle classes is used. This section shows how the "N-curves" of vehicle count, both by class and aggregated, can be predicted when input counts are given at the upstream location. Examples are provided.

2. TRAFFIC BEHAVIOR INSIDE QUEUES: THE 2-DETECTOR PROBLEM

Our system is a freeway section between two detectors (U-upstream and D-downstream) in which traffic flow is queued as a result of some unspecified downstream restriction(s). The system is assumed to be sufficiently far from any exits or entrances, as would occur for example in the middle of a long tunnel or bridge, so that the vehicular trip times, τ, on all its lanes would be similar. Our goal is predicting approximately the time-dependent traffic accumulation and vehicular trip times between "U" and "D" when the "schedule" of departures is given at "D"; i.e., to predict the N-curve at "U" from that at "D".

2.1 Preliminaries.

N-curves - The vehicle number function N(t, x), originally proposed in Moskowitz (1965), later refined in Makigami et al (1971) and more recently introduced to KW theory in Newell (1993), shall be used throughout this paper to summarize the traffic stream features of interest.

Moskowitz's function is most easily described in terms of imaginary numbered labels that are carried by vehicles. It is assumed that each vehicle carries one and only one label at all times, and that the labels have been numbered consecutively at time t = 0, increasing in unit increments in the upstream direction along the line of cars. It is also assumed that labels are exchanged among passing vehicles so as to ensure that labels do not pass one another even if vehicles do. Thus, the number on each label indicates at all times the position of its current vehicle in the traffic stream. Of course, if there is no passing labels also identify vehicles. In any case, since there is one and only one label with every vehicle at all times, a count of labels, or label-miles, or label-hours in a region of the (t, x)-plane, which can be easily derived from the N(t, x) function, always matches the corresponding count for vehicles. If traffic flows in the direction of increasing x, as is assumed in this paper, then N(t, x) increases with t and decreases with x. This means that N(t, x) is a function of t for a given x and also a function of x for a given t.

Note that a particular Moskowitz function describes a geometric surface in (t, x, n) space. The intersection of this surface and the plane x = x' , corresponding to a specific location x' , yields a (t, n) curve that will be denoted by a capital letter (usually "N") subscripted by an identifier of the location; e.g., $N_{x'}$. The capital letter represents the particular geometric surface (traffic instance) from which the curve comes. When a location is identified by means of a subscript, e.g., x_p , the subscript may be used to identify the corresponding (t, n) curve, e.g., N_p. A collection of "N-curves" at different locations along a highway { x_p } is useful for evaluation purposes because, as is well known, the horizontal separation between two N-curves for a given ordinate, n , (or their vertical separation for a given abscissa, t) represents the trip time of the n^{th} label between the two locations (or the vehicle accumulation at time t). The N-curves readily yield information about vehicle-miles and vehicle-hours of travel as well. They will be the object of our analysis.

Tolerances - Because the procedures presented in this paper are based on approximations to the data and to the parameters of the model, we introduce now the notion of tolerances. The paper will show that the numerical error in N due to the approximations is uniformly bounded. In particular, it will show that if N_D and N'_D are two input data curves within a tolerance ϵ (vehicles) of each other, then the output curves at any upstream location N_U and N'_U will also be within the same tolerance. That is,

$$\| N_D - N'_D \| \leq \epsilon \;\Rightarrow\; \| N_U - N'_U \| \leq \epsilon, \tag{1}$$

where the vertical bars signify the maximum absolute difference in count between the curves. The use of these and related ideas in validation experiments will be discussed.

2.2 The KW Theory Revisited.

This subsection presents a simple method for predicting N_U from N_D. The procedure is based on three postulates equivalent to those of KW theory. It will be described in steps, one postulate at a time, while at the same time introducing the (limited) scenarios that can be studied without the postulates not yet presented. This incremental approach is useful because if the theory is found not to match reality then we can see which of the assumptions (if any) can be salvaged in an effort to construct a better theory and which scenarios can still be studied despite the flaws.

> **Postulate 1: Reproducibility of stationary conditions inside a queue.** If curve N_D becomes straight with slope q, and remains so as time advances, then curve N_U should become parallel to it and remain to the left of N_D by an amount $\tau(q)$. Alternatively, we can require curve N_U to stabilize reproducibly $m(q) = q\tau(q)$ vehicle positions above curve N_D.**#**

This postulate means that the real curve $\tilde{N}_U$ should fluctuate within a specified tolerance of N_U if the real $\tilde{N}_D$ is within a specified tolerance of N_D. This should be reproduced on every observation instance independent of the location and/or cause of the obstruction that is generating the queue. Note that we have not required an absence of stop-and-go oscillations and have not said anything about the geometry of the highway between the detectors, which can include lane drops, changes in grade, etc. The only condition imposed is that the observed flow should be strictly less than the maximum possible stationary flow through the section, q_{max}, and that the section is queued. We assume that postulate 1 holds for any pair of locations x_i, x_j ($x_i < x_j$) inside the queue and therefore that the translations at consecutive locations are additive; i.e., that $m_{ij}(q) + m_{jk}(q) = m_{ik}(q)$. From now on the location-specific subscripts may be omitted when there is no room for confusion.

The postulate can be violated in at least two ways: (i) if one finds that a reproducible relation between τ (or m) and q cannot be found even though the N-curves become parallel, and (ii) if the N-curves do not even become parallel.[2] Although the postulate can be useful by itself,

---

[2] If certain vehicles had to be on certain lanes (e.g., for the purposes of turning) then, depending on which lanes are obstructed downstream on a particular day, one may observe significantly different trip times and vehicular accumulations across lanes and days, even if the overall flow is roughly the same every day. If the overall accumulation varies across days with the same q , this would give rise to a failure of type (i). A failure of type (i) could also arise if traffic was found to be unstable on some days and stable on others (with the same q), and if the trip times were to be significantly different on stable and unstable days. A failure of type (ii) could arise if people's behavior inside the queue were to depend on how long they have waited or on other events occurring outside our

e.g., as in the model of Lawson et al (1997), more hypotheses are needed in order to treat time-dependent downstream data.

It will now be assumed that the highway section is homogeneous, in the sense that the stationary accumulation between "D" or "U" and any intermediate location "M" which is 100α% of the way toward "U" (i.e., such that the following relation holds for the distance coordinates of the three locations: $(x_D-x_M) = \alpha(x_D - x_U)$ for $0 < \alpha < 1$) satisfies $m_{DM}(q) = \alpha m_{DU}(q)$ and $m_{MU}(q) = (1-\alpha)\, m_{DU}(q)$ for all q . This means that accumulations and trip times should be evenly distributed over the segment for all flows. Clearly, homogeneity allows us to write all the $m_{x,x'}(q)$ of a highway in terms of the normalized relation between m and q that holds for a segment of unit length, K(q); i.e., $m_{x,x'}(q) = (x-x')K(q)$. Our objective in this section is to find a reasonable set of rules that will allow us to predict N_U from N_D for any queued homogeneous segment on which there is a reproducible $m_{DU}(q)$ relation.

In what follows it will be convenient to describe the performance of the algorithm on a particular highway by a set of "upstream" operators $\mathsf{U} = \{U_s\}$, with one operator for each possible segment length, s. That is, U_sN denotes the result of the algorithm, when applied to a generic curve "N" at the downstream end of a length-s segment. When referring to a specific highway segment, U will be subscripted by two variables that identify the end points of the segment (instead of the single variable "length") with the downstream identifier placed first. With this convention the output of the algorithm for a given downstream curve and particular highway segment can be compactly expressed as $N_U = U_{DU}N_D$. Note that when the two identifiers are coordinates then $U_{x,x'} = U_{x-x'}$.

The first new postulate is just an extension of postulate 1 to the time-dependent case. We shall assume that N_D is piecewise linear, changing from state "i" to state "i+1" at time t_i. The reproducibility postulate indicates that if there is a transition between two lasting stationary states at detector D, indicated by a change in the slope of N_D from q_1 to q_2 , then the same stationary states should also arise at the upstream detector. It is now further assumed that the transitions between neighboring states occur rapidly and propagate upstream as a wave. (We anticipate at this point that not all pairs of states will be allowed to be neighboring in the final version of the

---

system.

model for stability reasons.) We assume the following:

> **Postulate 1 (Time-dependent reproducibility).** Lasting stationary states are still reproduced upstream. Furthermore, transitions between neighboring states propagate sharply as a wave.**#**

The procedure implied by this postulate is illustrated in Fig. 1 for the simple case when there is only one transition. Part a of the figure shows that the construction of N_U is easy, insofar as the two segments of curve N_U must be at specific separations from the corresponding segments of N_D. The procedure can be made more intuitive with the diagram on part b of the figure. It contains both, an m(q) relation and a companion curve of τ(q) vs m(q) with q (the slope of the ray passing through the origin) as the parameter. The relationship between m (or τ) and q in queued traffic is generally expected to be declining, as shown in the figure.[3] Curve m(q) will play the role of the so-called "fundamental diagram" of KW theory. The companion curve of m vs. τ can be useful because the axes of this diagram have the same physical dimensions as those of the (t, n)-plane, and this can help visually in the construction of the N_x curves. In particular, if the diagrams have been drawn with the proper scale then the horizontal and vertical separation between any two parallel portions of the N-curves that correspond to the same state should be equal to the coordinates of the state on the companion portion of the "fundamental diagram". Furthermore, the slope of an N-curve should always be equal to the slope of the line on the companion diagram that connects the given state with the origin. Let us now return to part a of the figure.

Note from the geometry of the picture that the transition point "C" from the first segment to the second must occur with a precise delay, w_{12} , that should only be a function of the two states "1" and "2" among which the transition occurs. More specifically, note from the slopes of the sides of triangle (ABC) that this delay can be written as:

$$w_{12} = - [m_1-m_2]/[q_1-q_2] \qquad (2)$$

---

[3] This is not considered a postulate, since it can be taken as our definition of a queued state.

and that as a result w_{12} will be positive for any pair of states if m(q) decreases. A positive w means that information (causation) travels in the upstream direction, as one would expect inside queues. A graphical interpretation of (2) is possible from the fundamental diagram; namely, that the negative slope of a line joining two states on the m(q) plot is the delay in the propagation of information.[4]

If N_D changes stationary state more than once and the times $t'_i = t_i + w_{i,i+1}$ which would indicate upstream state changes satisfy:

$$t'_i < t'_{i+1} , \quad \text{for all } i , \quad \text{where } t'_i = t_i + w_{i,i+1} , \tag{3}$$

then curve N_U can be obtained by applying the graphical procedure of Fig. 1b to all the states and the result will contain all the original states with breaks at the t'_i . The time (t'_{i+1} - t'_i) is the duration of state i+1 upstream.

If Eq. (3) is not satisfied, additional assumptions are needed. The following composition rule for the upstream operator is proposed:

Postulate 2 (Transitivity). If s = s' + s" , then $U_s = U_{s'}U_{s"}$.[5] **#**

This means that if N_U is the curve that is obtained from N_D using relationship $m_{DU}(q)$, then N_U should also be the curve that is obtained in two steps: (i) by applying the procedure to curve N_D with relationship $m_{DM}(q)$, and (ii) repeating the process with the resulting curve (N_M) and relationship $m_{MU}(q)$. The postulate also implies that a highway section can be divided into many parts and studied sequentially.

The reader is encouraged to verify that the procedure as it currently stands is divisible and satisfies this postulate; e.g., that the breaks in the N-curve at any intermediate position that is 100α% of the way toward the upstream detector occur with a delay αw_{12} .

---

[4] In the parlance of KW theory, this delay is the shockwave trip time.

[5] Insofar as the composition operation is obviously closed and associative, postulate 2 implies that the set of operators and the composition operation is a semigroup that is isomorphic with the semigroup of segment lengths (the non-negative real numbers) with addition.

Postulate 2 can be used to extend the procedure to cases where (3) does not hold. If postulate (3) is not satisfied then we look for an intermediate location "M_1" with the largest possible value of the interpolation parameter α, α_1 , for which (4a), below holds:

$$t'_i \leq t'_{i+1} , \quad \text{for all } i , \quad \text{where } t'_i = t_i + \alpha w_{i,i+1} . \tag{4a}$$

Note that all the states i for which $t'_{i-1} = t'_i$ are observed for zero time at location "M_1". Thus, they have vanished. In order to find the solution for locations upstream of M_1 we use the transitivity postulate and treat these locations as if M_1 was the location of the downstream data. To express this as a recursion it is convenient to rewrite (4a) as:

$$t'_i \leq t'_j , \quad \text{for all } (i , j) \text{ such that } i < j, \quad \text{where } t'_i = t_i + \alpha w_{i,i+1} , \tag{4b}$$

and introduce S_k for the set of surviving states at location M_k after the k^{th} iteration (k = 1, for now). Since the straight lines obtained by shifting the segments of the N-curve at location M_k are the same as those obtained by shifting the original (surviving) segments from the N-curve at location "D", we can simply choose α_{k+1} as the maximum value of α ($\alpha \leq 1$) that satisfies:

$$t'_i \leq t'_j , \quad \text{for all } i, j \in S_k \text{ such that } i < j , \text{ where } t'_i = t_i + \alpha w_{i,i+1} . \tag{5}$$

The recursive process terminates when (5) is satisfied for $\alpha = 1$. This algorithm will be named "A" and its results expressed by $N_U = A_{DU}N_D$.

Although the maximum number of iterations could be as large as the number of states (if all but one state were eliminated) in actual applications just a few iterations should be needed. The procedure is particularly quick when done by hand, and this can help in the interpretation of data. Consider as an illustration Fig. 2, which shows the result when the downstream state changes from "1" to "3" with a brief sojourn in state "2", for the m(q) relation shown on the top left corner. The figure also displays the N-curve for a detector "M" which is half way between "U" and "D" ($\alpha = 0.5$).

The figure shows that the shifted lines corresponding to the three linear segments of N_D

(lines OAB, AC and ED) do not intersect in the order required by (3) because line ED is beyond the apex of angle OAC. However, if the same construction is repeated for location "M" by cutting the three shifts in half then the three lines intersect in the proper way, as shown. We see at a glance that the duration of state "2" is reduced in going from "D" to "M", and that state "2" would be completely eliminated farther upstream, somewhere between "M" and "U"; i.e., that $0.5 < \alpha_1 < 1$. Note that we do not need to find α_1 to obtain N_U, since our knowledge that only states "1" and "3" survive the trip to location "U" already allows us to use Eq. (5). That is, the final solution--line OAED-- is simply obtained by connecting the original translations of the two surviving segments. In cases with more state transitions and longer separations between detectors more steps may be needed to determine which states are eliminated but procedure "A" remains simple.

The third and final postulate of the proposed theory for the description of queues involves the idea of "stability". The proposed algorithm does not rule out the possibility that an infinitesimal change in N_D could trigger a finite change in N_U, and we wish to exclude such "unstable" solutions from the set of possibilities. In particular, we specify that if a continuum of transition states of infinitesimal duration are introduced at every corner of N_D, i.e., we smooth the corners of N_D, then the final solution should not be affected by the smoothing. We also require this to be true of any intermediate solution. That is, if S denotes the smoothing operator, we assume that $U_s = U_s S$ for all s; i.e.:

Postulate 3 (Stability). The operators U_s and S commute.**#**

Figure 3 is used to illustrate these ideas. Part (a) shows by means of thick solid lines a construction for N_U similar to that of Fig. 2, with algorithm "A", when a transition from state "1" to state "2" has occurred at "D". If we imagine that the transition at point V has occurred by way of an infinitesimally quick sojourn through intermediate state "k", we see that this state cannot appear into the solution because if we shift the imaginary line passing through "V" corresponding to this state, the resulting line (labeled L_k in the figure) does not intersect our test N_U curve. Thus, the test curve is not destabilized by "k". Further consideration shows that the state will not enter into the solution as long as its corresponding point "k" on the (q, m)-plane is below or on the chord joining states “1" and “2" on such plane; see figure. Clearly, the higher this point is, the larger the shift imparted to L_k . It turns out that if the point is on the chord, then line L_k passes through the vertex of N_U. However, states that lie above the chord, such as “k' ”, would enter the solution as shown by line $L_{k'}$ and they would destabilize the solution.

If instead of an increase in flow, we had experienced a decrease, then we see from the geometry that algorithm “A” introduces intermediate states into the solution if they experience a small shift. They will appear in the solution if they lie below the chord. Any solution that can be modified by means of an infinitesimal perturbation could not be expected to arise in real life and should be ruled out. Thus, algorithm “A” needs to be modified and this is done below. The desired effect is achieved by treating each change, i, in the slope of N_D as being gradual and then determining which sequence of intermediate states appears immediately upstream; this will be called the “stable transition sequence”.

The logic used earlier on part (a) of the figure reveals that a proposed transition involving a particular sequence of intermediate states and an *increase* in flow is stable if and only if the portion of the m(q) curve joining two consecutive states in a stable transition sequence does not protrude above the chord joining the two states on the (q, m)-plane. For piecewise linear m(q) relations as in Fig. 3b the relevant solution should be a sequence of m(q) break-points that can be consecutively joined by chords lying entirely on the dotted region of the (q, m)-plane.[6] It should perhaps be intuitive without the need for discussion that: (i) this procedure will also connect the two original segments of the N_U-curve by the highest possible arc that can be constructed by a succession of intermediate states, and (ii) that if the m(q) curve is convex no states propagate into the solution. (The original jump is then said to be “stable”; stable jumps correspond to “shocks” in the KW theory.)

If the transition involves a reduction in flow, then the results are reversed. One would look for the sequence of break-points that can be joined with chords lying entirely within the *shaded* area, and would find that the corner of N_U is spanned by the *lowest* possible arc of intermediate states. The original jump would be stable if the m(q) curve is *concave*.

Note that in both cases, whether there is an increase or a decrease in flow, the corner of

---

[6] Mathematicians would say that they are the vertices of the convex hull of the shaded region.

the N_U-curve is spanned by an arc of new states that is as far away from the original corner as possible; i.e. in all cases, corners are smoothed as much as possible. Furthermore, in the piecewise linear case only break-points of the m(q) curve can appear as new intermediate states. Obviously, we can imagine that these are the only states actually introduced by S, and this simplification allows us to devise a stable procedure "U" that requires just a few steps in the piecewise linear case.

Algorithm "U"

Let x_D and x_U denote the positions of the downstream and upstream detectors, and x_M the position of an intermediate trial location such that $x_U \le x_M \le x_D$. The intermediate location represents the most upstream location at which an N-curve, N_M , is known. The algorithm starts with $x_M = x_D$ and then reduces x_M iteratively until $x_M = x_U$. Each iteration has two steps: "smoothing" and "shifting":

Step 1: Smoothing. Smooth N_M by introducing the necessary intermediate states at its corners as described above. The result is denoted SN_M .

Step 2: Shifting. Find the most upstream location x*, $x_M \ge x^* \ge x_U$, for which the shift imparted to the segments of SN_M with algorithm A does not eliminate any states. (This ensures that all the intermediate states are stable.) The resulting curve is called N*. If $x^* = x_U$, then $N_U = N^*$ and the procedure terminates. Otherwise, we set $x_M = x^*$, $N_M = N^*$ and repeat the iterations.

Only one iteration may be needed to obtain the result if the study section is short and/or the data vary slowly. Alternatively, and as an approximation that is useful when solving a problem by hand, we may apply algorithm A to curve SN_D. This may introduce a small error in the solution on the order of just a few vehicles per lane per kilometer, because the approximation ignores the smoothings that might have to take place at intermediate locations; thus, $AS \approx U$. The error arising from this simplification is zero when the m(q) curve is concave or convex because then, as the reader can verify from the "chord" rules of Figure 4, the elimination of a state always introduces a jump that is already stable ($AS \equiv U$ in these cases). To obtain a set of predictions for a large system including many detectors one would then obtain the N-curves in the upstream direction, stepwise from detector to detector, i.e., "predicting" with A or U and then "smoothing" the data at each detector.

2.3 Properties of the procedure.

This section examines the effect that errors in the input data N_D and the "fundamental diagram" $m_{DU}(q)$ have on the solution. It will be shown that the maximum error in the solution cannot exceed the maximum error in the data.

Let us first consider the effect of the differential operator $dU = U_{dx}$ on two similar (piecewise linear) data curves N and N'. We are interested in determining the change in the maximum vertical separation between the curves that is induced by the operation.

In the neighborhood of the location where this maximum change takes place, one or both of the curves will bend; e.g., as shown on Fig. 4. The dashed lines on this figure represent the upstream curves. It should be clear from the geometry of the figure that at the time t* where the separation is at a maximum one or both of the curves must bend, and that the smoothed lines SN and SN' must share a common flow q* at that time. (Otherwise their separation would be changing at t = t*, which is not possible.)

We then see from the rules of Fig. 4 and the present geometry that the separation between either of the lines, e.g., N, and an intermediate data line N* with slope q* (also shown in the figure) cannot increase because otherwise there is a contradiction. That is, if the separation between dUN* and dUN was larger, as shown in the figure, then the shifted line corresponding to state q* of curve N would intersect dUN (since the shift imparted to the line with slope q* is the same as that imparted to N*), which is impossible given the rules of Fig. 3.

Since neither of the solution lines can drift away from the intermediate line, it must be true that: $\|dUN - dUN'\| \leq \|N-N'\|$. Clearly now, since U is transitive, it follows that:

$$\|U_s N - U_s N'\| \leq \|N - N'\| \quad \text{for all s, N and N'.} \qquad (6)$$

Note that the operation $\|N - N'\|$, which is the L_∞ norm, satisfies the triangle inequality.

Equation (6) is reassuring because it implies that errors in the input data do not grow into the solution. That is, if we take N and N' to be the closest possible approximations respectively to a discrete data curve $\tilde{N}$ that is subject to measurement error and to the true (unknown) data curve $\tilde{N}'$ then (6) indicates that the maximum discrepancy in the two input curves cannot increase.[7]

Equation (6) can also be used to decide how many linear segments are appropriate for approximating a few hours of data, depending on the level of approximation that one desires in the predictions. This is illustrated with the following example.

Example. Figure 5 displays by means of solid lines the periodic N-curves predicted by the theory at evenly spaced locations upstream of a pretimed oversaturated traffic signal with a "true" downstream curve N'_D when the m(q) curve for the segments between neighboring locations is as in the top of the figure. The dashed lines are the predictions obtained from an approximation N_D where the signal cycles have been smoothed out. Note how the separation between the solid and smooth lines declines as one moves upstream, and that their separation would not increase if the N_D curve was shifted but remained parallel to the original.

Note as well that the convex corners of N_D are smoothed, but the concave ones remain sharp. This is true in general: for concave relations the convex corners are smoothed and for convex relations the concave corners are smoothed. Furthermore, corners are never sharpened for any m(q).**#**

Insofar as random variations in accumulation on the order of $\delta = 10$ vehicles per lane can be expected for freeway sections comparable with a mile for the same set of downstream conditions, e.g., due to stop-and-go effects and randomness in vehicular spacings, the above suggests that raw input data does not need to be linearized beyond this level of precision. The ability to work with roughly approximated data should facilitate calculation and data interpretation, since the required effort grows with the number of state changes.

Errors in the approximation of the m(q) relation are examined next. Note first, using the same logic as above, that "dU" has the reverse effect of (6) relative to the minimum separation

---

[7] Note that the operator cannot be applied to the raw data curves because they are step functions. The statement assumes that both raw data curves can be approximated by valid piecewise linear curves (with slopes between 0 and q_{max}) to within a tolerance ϵ which can be neglected compared with the error in the data ϵ'; i.e., that the discrete and continuous curves are undistinguishable for practical purposes. If this is not the case, the discrepancy between the two raw curves could be anywhere in $[\epsilon' \pm 2|\epsilon|]$ by virtue of the triangle inequality and this amount could be (slightly) smaller than the discrepancy in the solution.

between two disjoint N-curves. That is, the minimum separation operation $s(N, N')$ satisfies:

$$s(UN, UN') \geq s(N, N'). \tag{7}$$

Inequality (7) implies that U is a monotone mapping; i.e., that:

$$N > N' \Rightarrow U_sN > U_sN' \quad \text{for all s.} \tag{8}$$

Monotonicity is reasonable; it means that if every vehicle leaves a (queued) downstream location on day “1" earlier than on day “2", then every queued vehicle will also get to pass upstream (queued) locations earlier on day “1" than on day “2".

Consider now two normalized m(q) relations for a given highway-- true and approximate --and the associated operators U’ and U. We show below that the maximum error in the solution can never exceed the maximum difference between the true and approximate accumulation curves, $\|m_s(q) - m'_s(q)\|$; i.e., that:

$$\|U_sN - U'_sN\| \leq s\Delta \quad \text{for all s and N,} \tag{9}$$

where Δ is the maximum discrepancy between the two normalized accumulation curves.

This statement is proven by induction; i.e., by showing that Eq. (9) is true for length ds (fact “a”); then assuming that Eq. (9) is true for length s (“b”); and then showing that (a) and (b) imply Eq.(9) for length s+ds (fact “c”). A geometrical construction such as that of Fig. 4 reveals that (a) is true. Then, if we write N_s and N'_s for U_sN and U'_sN we just need to show that the norm of $(U_{ds}N_s - U'_{ds}N_s) = (U_{ds}N_s - U_{ds}N'_s) + (U_{ds}N'_s - U'_{ds}N'_s)$ does not exceed (s+ds)Δ. To see that this is true, note that

$$\|U_{ds}N_s - U'_{ds}N_s\| \leq \|U_{ds}N_s - U_{ds}N'_s\| + \|U_{ds}N'_s - U'_{ds}N'_s\|,$$

by virtue of the triangle inequality, and also that (i) the first term on the right side of this inequality cannot exceed sΔ by virtue of (6) and fact (b), and (ii) the second term cannot exceed dsΔ by virtue of fact (a). It then follows that the first term cannot exceed (s+ds)Δ. Q.E.D.

Taken together, Eqs. (6) and (9) imply:

$$\|U_s N - U_s N'\| \leq \|N - N'\| + s\Delta \qquad \text{for all s, N and N'},$$

again by virtue of the triangle inequality. This result is important for model validation purposes because it bounds the portion of the discrepancy between observation and prediction, $\|N_U - \tilde{N}_U\|$, that can be attributed to the combined effect of errors in the N-curve data and in the estimation of m(q).

3. LIGHT TRAFFIC AND QUEUED TRAFFIC ON AN INHOMOGENEOUS HIGHWAY

3.1 The three-detector problem

Consider now three locations on a highway $x_1 \leq x_2 \leq x_3$, and assume that the N-curves at x_1 and x_3 are known. The identifiers "1", "2" and "3" are now used for the locations instead of U, M and D, because the ideas will soon be generalized to more than 3 detectors. We also assume that traffic is queued at x_3, that the queue does not reach all the way back to x_1 (traffic is unqueued at x_1), and that the flow observed at any unqueued location over any period of time that includes a significant number of cars never exceeds the maximum flow (q_{max}) that can be dissipated by a queue. These assumptions will be relaxed later. Our immediate goal is to predict N_2 from the available information.

To this end, a general model of (uncongested) traffic with multiple vehicle classes will be specified in terms of a "downstream operator" that returns the conditions that should prevail at a downstream location from those observed upstream. The upstream data consists of a set of cumulative count curves by vehicle class $N_{1(i)}$, where the parenthetical subscript i = 1, 2, 3, ... refers to the class and the other subscript to the location, and a combined curve $N_1 = \sum_i N_{1(i)}$. The boldface symbol $\mathbf{N_1}$ will denote the complete set of curves, including the combined curve.

It will be assumed that there is no passing within each vehicle class and that passing is possible among vehicles of different classes. Note that this assumption does not reduce generality because one can always define a separate class for each individual vehicle. It is also assumed that the uncongested traffic model can be described in terms of a "downstream" operator $\mathbf{D_{12}}$ which predicts $\mathbf{N_2}$ from $\mathbf{N_1}$; i.e., $\mathbf{N_2} = \mathbf{D_{12}}\,\mathbf{N_1}$. This assumption is somewhat restrictive because it implies that the conditions found at x_1 are sufficient to describe what happens downstream, but the restriction is minor. The proposed model includes as special cases all the continuum models introduced to date, as well as most microscopic theories and simulations. A parenthetical subscript (i) and no boldface will be used for the operator that returns the curve for the i^{th} vehicle class; this subscript will be omitted for the aggregate curve.

> **Postulate 4: Queue behavior.** It will be assumed that a queue always forces individual vehicles and labels entrapped in it to travel more slowly than predicted by the light traffic theory, and that queued vehicles cannot pass. It will also be also assumed that the queue has no effect upstream of its domain of existence and that any transition from light traffic into queued traffic is rapid and stable.**#**

The first premise of this postulate excludes situations where some vehicle classes with very low desired speeds may be embedded in a faster moving queue. It also implies that a vehicle, or a label, cannot cross a highway location such as x_2 any earlier than the time allowed by light traffic theory and the time allowed by the queue. In other words, the vehicle number seen by an observer at x_2 at time t cannot not exceed the minimum of the number arising from light traffic theory (from conditions at x_1) and the number arising from the heavy traffic theory (from conditions at x_3). That is:

$$N_2 \leq \min\{U_{32}N_3\,;\;\; D_{12}\mathbf{N_1}\}. \tag{10}$$

Equation (10) should be a pure equality for the two extremes of our section, $x_2 \approx x_3$ or $x_2 \approx x_1$. This implies that reasonable curves $\{\mathbf{N_1}\,; N_3\}$ should satisfy:

$$N_3 \leq \;D_{13}\mathbf{N_1} \tag{11a}$$

and

$$U_{31}N_3 \geq N_1. \tag{11b}$$

Equation (11a) specifies that a vehicle cannot pass by x_3 any sooner than predicted by the light traffic theory and the upstream data at x_1. Conversely, (11b) specifies that a vehicle cannot pass by the upstream location any later than if it had been embedded in a queue the whole way--since only a queue can slow down a vehicle.

Stability - For $x_2 \approx x_3$ we expect (10) to be a pure equality. Thus, if the interval is divided into short segments of length dx → 0, postulate 4 would imply that the intermediate N-curves should satisfy the following relation,

$$N_x = \min\{U_{dx}N_{x+dx}\ ;\ D_{1x}\mathbf{N_1}\} = \min\{A_{dx}SN_{x+dx}\ ;\ D_{1x}\mathbf{N_1}\}. \qquad (12)$$

which allows us to calculate N_2 recursively for any intermediate location, x_2.

Note that the operator S in (12) smooths all the corners of N_{x+dx} , including those that involve a transition between a queued and unqueued state. This is logical; it essentially prevents the queue from growing so fast that perturbations introduced at its back end grow into it. [If the back of the queue propagates too fast then (12) automatically introduces queued transition states with (fast) waves that can keep up with this growth.]

Note as well that procedure (12) can be used without any changes even if the end of the queue is not always contained between x_1 and x_3 ; i.e., even if it ebbs and flows past these two locations. This is true because the particular location of the back end of the queue when it is outside our study section is irrelevant insofar as the interior of the interval is concerned. Thus, the behavior of traffic the study section (x_1, x_3) should be the same as if the end of the queue never strayed far from our interval. That is, we can assume in our calculation of N_2 that the freeway is permanently queued at a location "3+" immediately downstream of "3", so that curve N_3 (in a queued state) is observed there, and permanently unqueued at a location "1-" immediately upstream of "1", so that the curves $\mathbf{N_1}$ (in an unqueued state) are observed there. Recipe (12) is simply the result of using the data at "1-" and "3+" as an input.

3.2 The P-detector problem.

Let us now assume that there are more than three detectors, located at positions x_p (p = 1, 2, 3, ... , P) , with $x_p < x_{p+1}$, that data are available for the first and last detectors and that we want to predict N_p for $1 < p < P$. If the detectors are closely spaced (12) would suggest:

$$N_p \approx \min\{ A_{p+1,p}SN_{p+1} \,;\; D_{1p}\mathbf{N_1} \} \qquad \text{for } 1 < p < P\,. \qquad (13)$$

This is a reasonable approximation because with closely spaced detectors the new transition states are always introduced close to their true birthplaces. We estimate that temporary errors in count due to this discrepancy should be on the order of just a few vehicles per lane of freeway for detector spacings comparable with 1 Km. Because errors in count depend on the square of the spacing, they can be virtually eliminated by reducing the detector spacing to a few hundred meters. Some of these ideas are illustrated in the following subsection with two examples. The subsection may be skipped without loss of generality.

3.3.- Examples.

The first example examines an uphill section of a two-lane freeway with a substantial portion of slow-moving trucks when an incident blocks the road. It is assumed that the m(q) relation is as shown on part (a) of Fig. 6. The (unspecified) units of time and vehicle count have been chosen on this diagram so as to ensure that the numerical values for flow, trip time and vehicle number that arise in the solution are small "round" numbers.[8]

The proposed behavior of the traffic stream in the absence of downstream queues is similar to that seen on California State Highway 17 just out of "Los Gatos" and toward "Santa Cruz", which, as we shall see, is inconsistent with the KW model.

Part (b) of the diagram displays the N-curves observed at a location x_1 (solid lines) for i = 1 (cars) and i = 2 (trucks). This is the data set $\mathbf{N_1}$. Note that the automobile flow reaches a maximum $q^{(1)} = 4$ for $0 < t < 1$. This flow is assumed to be close to capacity, so that both lanes of the freeway are almost fully used. At t = 1 a steady stream of trucks with $q^{(2)} = 1$ enters the road directly upstream of our location. Because trucks travel more slowly, cars are assumed to avoid the right lane so as not to lose their position in the faster moving left lane and the system capacity drops, as in California State Highway 17. It has been assumed in the example, somewhat

[8] The numbers in this example could be plausible (although very rough) for a two lane freeway if the detectors were spaced about 1 Km apart, time was measured in minutes and vehicles in groups of 24 (a couple of dozens).

drastically for illustration purposes, that the entrance of the trucks reduces the car flow by a factor of two. This reduction should generate a growing car queue upstream of x_1, but this is of no concern to us.[9] Of interest is the effect on the traffic stream downstream of x_1, which would consist of two non-interacting parallel vehicle lines moving with different speeds for $t > 1$.

Thus, the unqueued traffic model for our problem can be very simple: we propose that cars and trucks should travel at class-specific speeds that are fixed but significantly different. The dotted lines are the class-specific N-curves at location x_2 that are obtained with this simple unqueued model by shifting each curve by its specific trip time. This is assumed to be 1 unit for cars and 2 for trucks -- a large speed difference is used so as to enhance the clarity of the diagrams. The summation of the counts across both classes yields aggregate curves N_1 and $D_{12}\mathbf{N_1}$; see part (c) of the figure.

This part of the figure also contains the predicted curve $D_{13}\mathbf{N_1}$ for a location x_3 that is twice as far downstream, assuming that the highway is homogeneous. The resulting (t, n) plot is divided into four regions of constant flow "q" and constant average trip time "τ", as indicated. The values of the accumulation "m" between detectors that would arise in these regions if conditions did not change with time (i.e., the vertical separation between the sets of parallel lines) is also shown. These four stationary states are also displayed on part (a) of the figure and are labeled "a, b, c, d", in chronological order of occurrence. Note that a "fundamental diagram" cannot be defined for this problem in unqueued traffic because more than one value of m arises for $q = 3$. Therefore the KW theory does not apply.

Let us now assume that an incident that completely blocks the road occurs at point $(t, x) = (2, x_3)$ and examine with the help of Eq. (13) (with $p = 2$ and $P = 3$) the effect of the queue on N_2. Since $D_{12}\mathbf{N_1}$ has already been obtained, the only missing data item is curve N_3, which will readily yield the first term of (13): $A_{32}SN_3$. In order not to clutter the diagram these steps are taken on part (d) of the figure which contains copies of curves N_1, $D_{12}\mathbf{N_1}$ and $D_{13}\mathbf{N_1}$.

We hypothesize that curve N_3 matches the predicted curve $D_{13}\mathbf{N_1}$ up to $t = 2$ and, since the incident does not let any vehicles through, we also assume that N_3 adopts a slope of zero (state "f") for $t > 2$, as shown. (The behavior of the bottleneck, albeit reasonable, is not part of the

---

[9] Interestingly, note that this queue could be eliminated by forcing all cars to travel at the speed of the trucks; e.g., by *reducing* the speed limit !

proposed theory.) The operations $A_{32}S$ consist of a "smoothing", which introduces state "e" into N_3 at t = 2 (see dotted line), followed by the shifts and truncations arising from algorithm A. In our example the result is the curve labeled $A_{32}SN_3$, as the reader can verify from the data in part (a) of the figure. As per Eq. (13), N_2 is the lower envelope of the two curves calculated for location 2. This solution is interesting because it involves a transition from state "d" to state "f" at the end of the queue (where $D_{12}\mathbf{N_1} = A_{32}SN_3$) which turns out to be unstable.

That is, if curve N_2 were to be used to make predictions farther upstream (e.g., at some location x, such that $x_1 < x < x_2$) then break-point state "e" would have to be introduced as part of the smoothing step for N_2 , as shown in the figure. Consideration shows that the new state would indeed be propagated upstream.

Part (e) of the figure contains the exact stable solution of our problem on a (t, x)-plot. It displays the (t, x)-domain for each state arising in our problem. The figure is the result of applying (13) with infinitely close detectors. Curve N_2 can be obtained from it by integration of the flows observed along the line $x = x_2$. One can see at a glance that the only difference between the exact curve and that on part (c) of the figure is that the latter should have been slightly beveled due to the appearance of state "e" somewhere between $t \approx 7$ and $t \approx 9$. The largest discrepancy in count is only about ½ vehicle unit, which converts to about 12 vehicles for the two lanes (given the 1 Km separation between detectors). The approximate curve, which is otherwise exact, does introduce state "e" at t = 8. Thus, the discrepancy in count would not grow in magnitude into the solution for $x < x_2$. Instead, it could eventually disappear as occurs with the earlier episode of state "e", which has no lasting effect.

Another example, presented in Fig. 7, is the well-known lead-vehicle problem of KW theory for a homogeneous highway in which the flow-density relation is consistent with the m(q) curve of part (a). This curve corresponds to a section of length "L", so that the density m/L. The example demonstrates that the results in the two theories are consistent with each other.

Part (b) of the figure shows the unique solution with stable shocks that arises in KW theory when a slow vehicle (pace = 16) enters a steady traffic stream that is in state "a". The reference locations x_p shown on the figure are assumed to be spaced L distance units apart. The entrance of the vehicle at point (t_0, x_3) induces a two-state queue upstream of it that encompasses states "d" and "e"; state "c" does not appear into the solution, nor do any other queued states. This is the stable solution in KW theory.

Part (c) of the figure shows the N_3 curve for the downstream location, which can be justified outside the proposed theory. The figure also displays curve $D_{12}N_1$, and curve RSTPQE, which is the result of applying the operator pair $A_{32}S$ to N_3 .[10] The minimum rule finally yields N_2 = curve OPQE. Note how the flows of this curve and the times at which the transitions occur precisely match those in part (b) of the figure.

3.4. The inhomogeneous highway with over saturated flows: time-dependent bottlenecks.

It is assumed here that the freeway can be modeled as a series of homogeneous sections with boundaries at a series of points $\{x_p\}$. Each section is characterized by an $m_{p,p+1}(q)$ relation with a maximum stationary flow and an operator family: $U_{p,p+1}$ (or $A_{p,p+1}$). Each boundary is characterized by a time-dependent capacity $C_p(t)$ which is assumed never to exceed the maximum stationary flows of the two sections it separates. This means that $m_{p,p+1}(C_p)$ and $m_{p-1,p}(C_p)$ are always defined.

Only "transitive" light traffic theories, i.e., such that $\mathbf{D_{p,p+2} = D_{p+1,p+2}D_{p,p+1}}$, are considered. In other words, models for which knowledge of the N-curve *by vehicle class* at a location is sufficient to describe the behavior downstream without any further information from upstream. Thus, each $\mathbf{D_{p,p+1}}$ is a property of a section (p, p+1), which acts on vehicles independent of what other sections have done.

In order to make predictions when flows larger than capacity arise at some locations it is necessary to introduce one last assumption with regard to the boundaries between sections:

> **Postulate 5: Bottleneck behavior.** The flow through a bottleneck x_p is always as large as possible but cannot exceed any of the following: 1) the bottleneck capacity, 2) the predicted aggregate flow across classes from upstream data and 3) the prediction from downstream data.**#**

This bottleneck behavior will prevent flows larger than the maximum possible to be put through

---

[10] Note that state "a" of N_3 was treated as if it was queued, " a' ".

any section. According to postulate 5 the aggregate flow upto time t should be given by:

$$N_p\,(t) = \min\{\ (U_{p+1,p}N_{p+1})(t)\ ;\ (D_{p-1,p}\mathbf{N_{p-1}})(t)\ ;\ N_p(t\text{-}dt) + C_p(t)dt\ \}, \qquad (14)$$

which is a simple extension of (13) that recognizes the capacity condition. Note that the transitive property of the light traffic model has also been used in (14).

The capacity $C_p(t)$ can be either: 1) exogenously defined or 2) exclusively dependent on system conditions prior to t. (It can depend for example on how long people currently passing through the bottleneck have been delayed, how long the bottleneck has been active, etc...) Under these conditions, $N_p(t)$ can be evaluated for all t by stepping through time with Eq. (14). This is true because the first two terms on the right side of (14) only depend on the neighboring N-curves for times prior to t-dt. This calculation strategy is very efficient for computer implementation because in the final analysis it turns out that one only needs to carry forward a very small amount of information from one step to the next. A description of these details is beyond the scope of this paper, however, and is not given here.

Equation (14) is sufficient for prediction purposes if the practical situation only involves one vehicle class. Otherwise, the algorithm needs to be completed by providing a recipe that gives the disaggregate solution $\mathbf{N_p}(t)$ based on (14).

If the unqueued theory involves no vehicle passing, e.g., as in Newell (1993), this can be done by ensuring that vehicles from all classes should share the same trip time if they arrive simultaneously at the upstream detector. This rule can be expressed in terms of the N-curves as follows:

$$\mathbf{N_p}(t) = \mathbf{N_{p-1}}(t') \ , \qquad \text{for some } t' < t. \qquad (15a)$$

In other words, we should find the only t' for which $N_{p-1}(t') = N_p(t)$, which is possible since we are stepping through time and $t' < t$, and then set $N_{p(i)}(t) = N_{p-1(i)}(t')$. Note that t' is the time when label $N_p(t)$ arrived at the detector directly upstream fom p. Equation (15a) expresses the FIFO rule because vehicles do not pass labels under FIFO and as a result all the vehicles that crossed the upstream detector prior to t' must have crossed the downstream detector prior to t.

For non-FIFO models the following expression is exact when there is no queueing:

$$\mathbf{N_p}(t) = (\mathbf{D_{p-1,p}N_{p-1}})(t)\ ; \qquad \text{if } N_p(t) = (D_{p-1,p}N_{p-1})(t). \tag{15b}$$

Of course, (15a) continues to hold when the section (x_{p-1}, x_p) is entirely queued. Thus, we propose to approximate $\mathbf{N_p}$ (t) by (15b) if its no-queueing condition holds and by (15a) otherwise. Thus, the complete algorithm is given by (14) and (15). To be sure, the use of (15a) when a section is only partially queued introduces a small error into the solution, but for typical applications where vehicular speed differences are on the order of 10 or 20 Km/hr and detectors are 1 or 2 Kms apart, the discrepancy between (15a) and (15b) is small relative to other sources of noise. These errors can be reduced by shortening the detector spacing. The errors can also be reduced with a more elaborate calculation of the time and location where each vehicle joins/leaves the queue. This calculation can be streamlined in some cases, e.g., with the procedure suggested in Lawson et.al. (1997), but a discussion of this somewhat minor issue is beyond the scope of this paper.

4. CONCLUSION

The ideas presented in this paper can be extended to situations including classes of vehicles, i , that space themselves differently in a queue if it is reasonable to assume that a generic m(q) relation in terms of *"passenger car equivalents"*, α_i , exists; i.e., if after assigning α_i tags to each vehicle of class i (for all classes) we find that a reproducible m(q) relation *for tags* exists that is independent of the queue composition. If this is true, we can simply view each class as being composed of convoys of consecutive vehicles that carry exactly one label; i.e., such that $1/\alpha_i$ vehicles of type i = 1 convoy. Since these convoys would behave in agreement with postulates 1-4, the results obtained up to Sec. 3.3 apply to the convoys. [If one prefers to work with vehicles rather than convoys then it should be remembered that the overall count of labels is related to the vehicle counts by: $N_1 = \sum_i \alpha_i N_{1(i)}$.] For the results of Sec. 3.4 to apply, a reproducible relation $C_p(t)$ in terms of tags would have to exist.

The ideas in this paper can also be extended to multi-commodity networks where physically meaningful boundary conditions have been defined at the nodes. To do this properly one needs to allocate vehicles going to different destinations to different classes since the composition of the traffic stream by destination is a determinant of turning percentages. Newell (1993) and Daganzo (1995) discuss possible boundary conditions at simple junctions.

REFERENCES

Cassidy, M. and J. Windover (1995), "Methodology for assessing dynamics of freeway traffic flow", *Trans. Res. Rec.* 1484, 73-79.

Cassidy, M. and J. Windover (1996), "Driver memory: Motorist selection and retention of individualized highways in highway traffic", Research Report UCB-ITS-RR-96-4, Institute of Transportation Studies, Univ. of California, Berkeley, CA.

Daganzo, C.F. (1995) "The cell transmission model. Part II: Network traffic" *Trans. Res.* 29B(2), 79-93.

Erera, A.L., Lawson, T.W. and Daganzo, C.F. (1997) "A simple generalized method for analysis of traffic upstream of a bottleneck", Accepted for presentation at the January 1998 annual meeting of the Trans. Res. Board, Washington D.C.

Lawson, T.W., Lovell, D.J. and Daganzo, C.F. (1997). "Using the input-output diagram to determine the spatial and temporal extents of a queue upstream of a bottleneck" Trans. Res. Board presentation, Washington, D.C., January 1997. (*Trans. Res. Rec.* in press).

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.

Makigami, Y., G.F. Newell, and R. Rothery (1971), "Three-dimensional representation of traffic flow", *Trans. Sci.*, 5, 302-313.

Moskowitz, K. (1965), "Discussion of 'freeway level of service as influenced by volume and capacity characteristics' by D.R. Drew and C. J. Keese", *Highway Res. Rec.* 99, 43-44.

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, 281-313.

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

LIST OF FIGURES

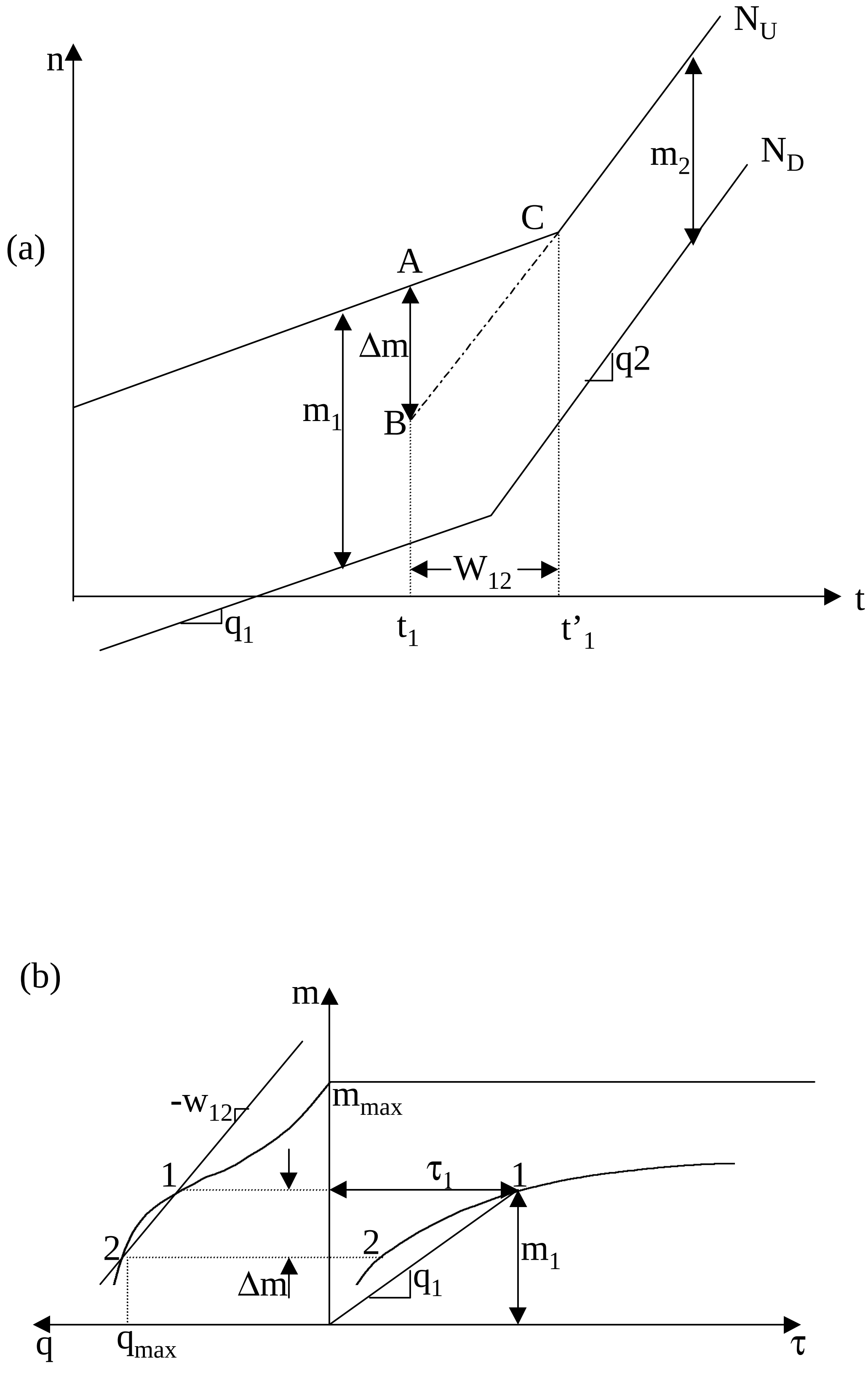

(a)
n
N_U
N_D
m_2
C
A
Δm
q2
m_1
B
W_12
t
q_1
t_1
t'_1
(b)
m
-w_12
m_max
1
τ_1
1
2
2
m_1
Δm
q_1
q
q_max
τ

Fig. 1

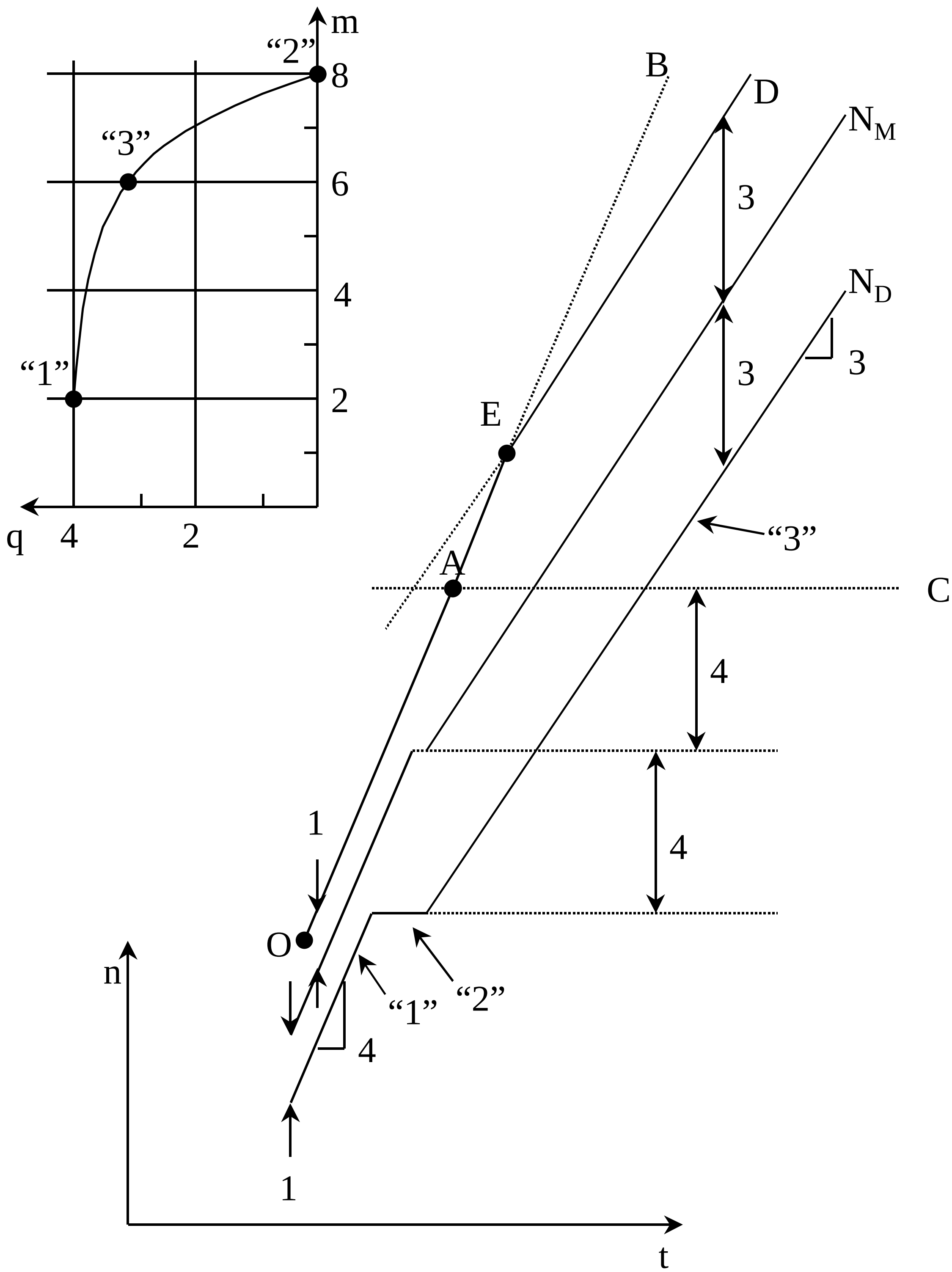


Fig. 2

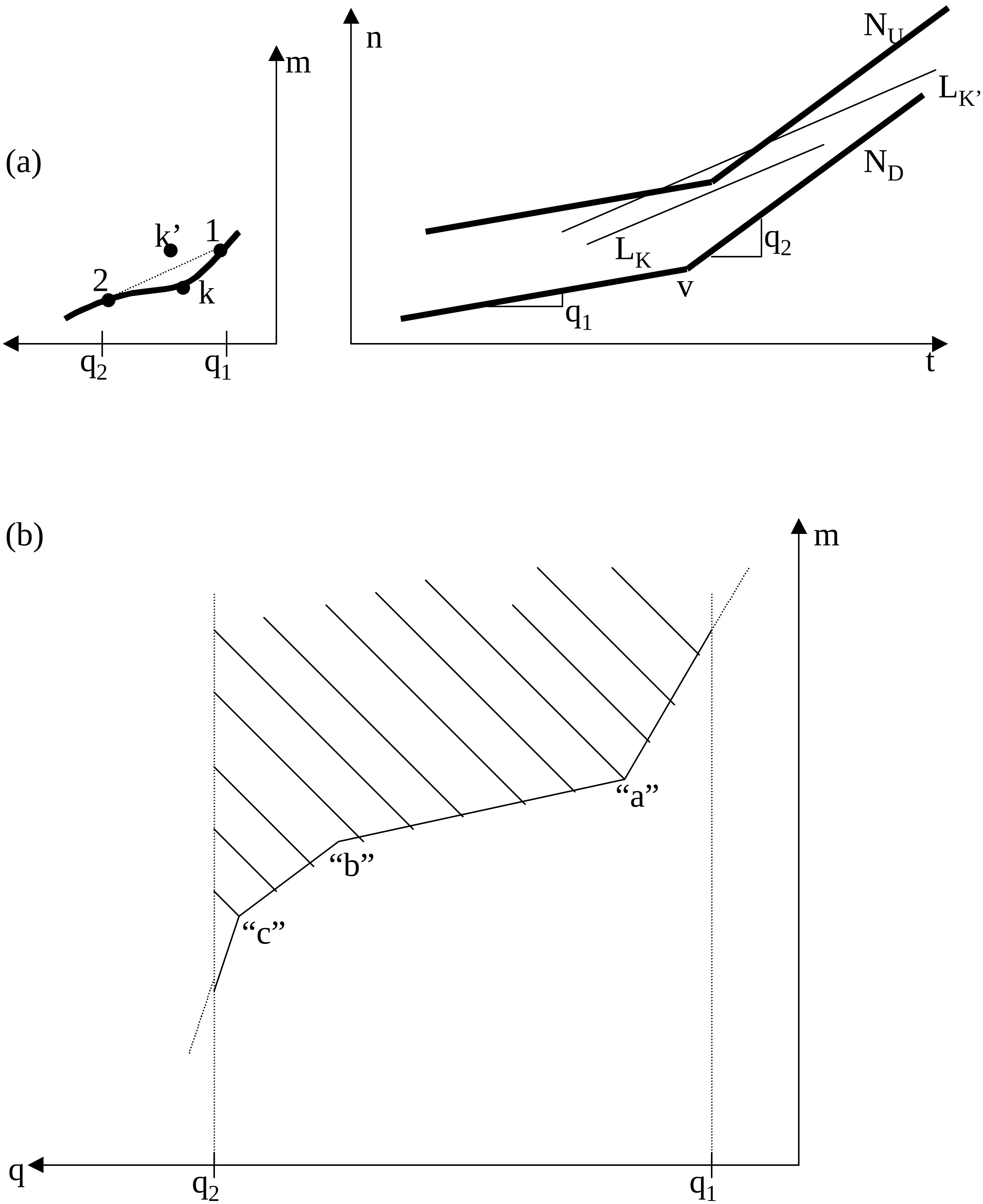


Fig. 3

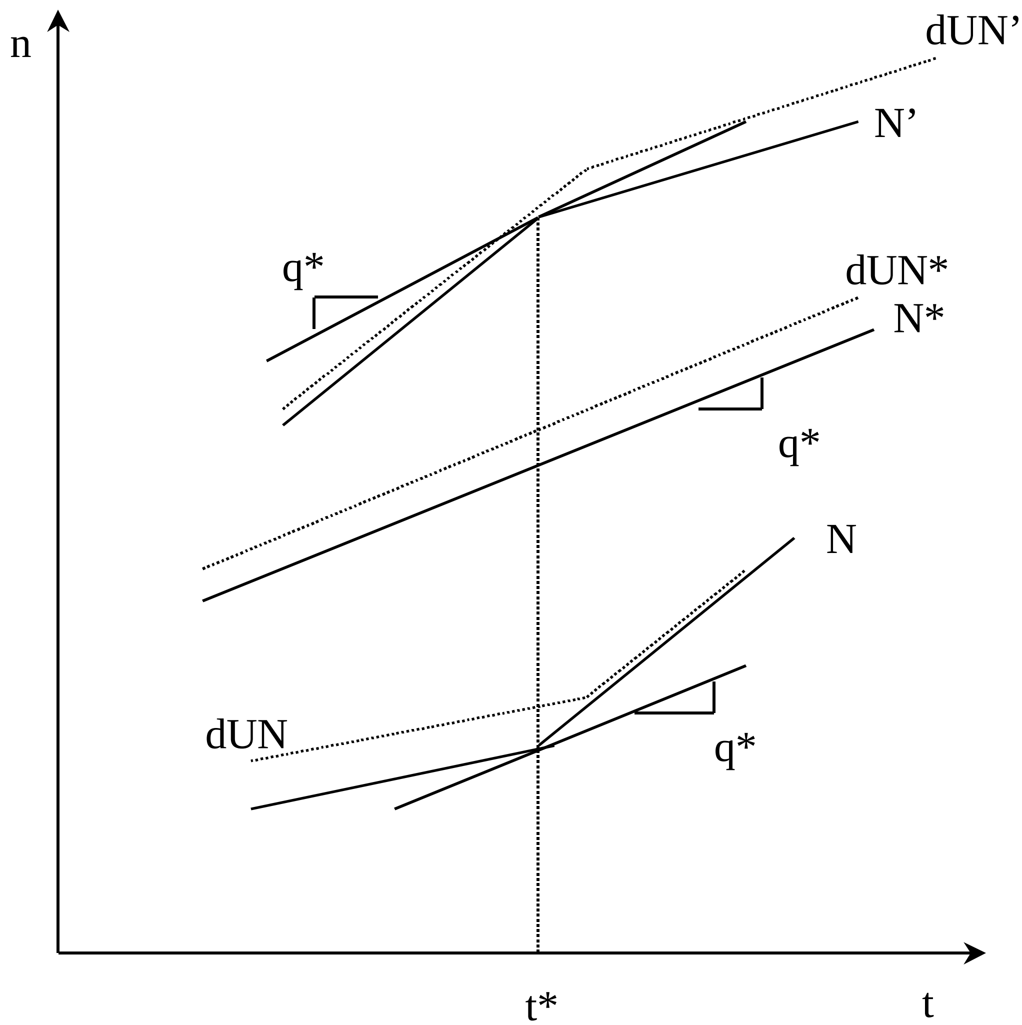


Fig. 4

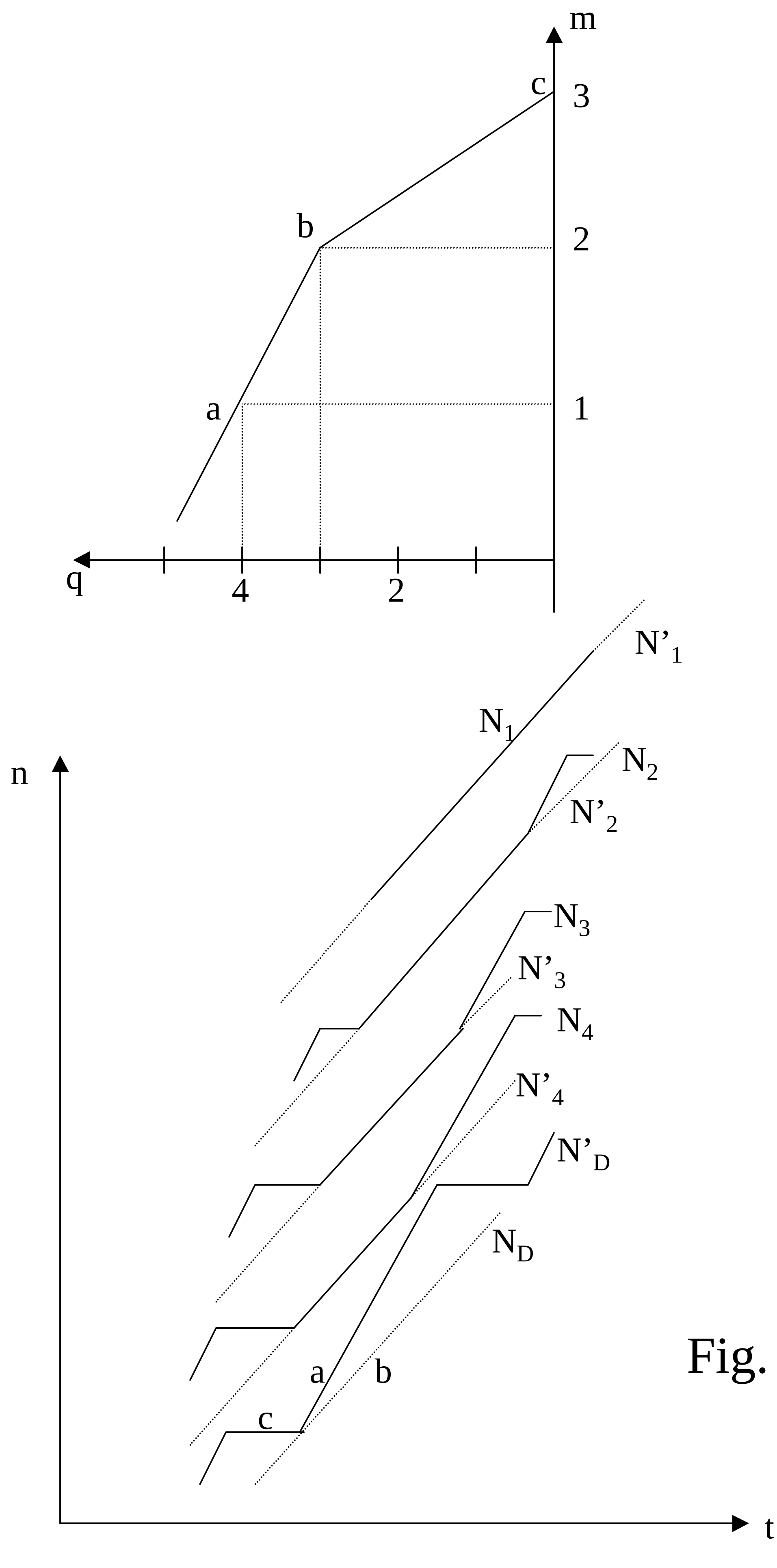
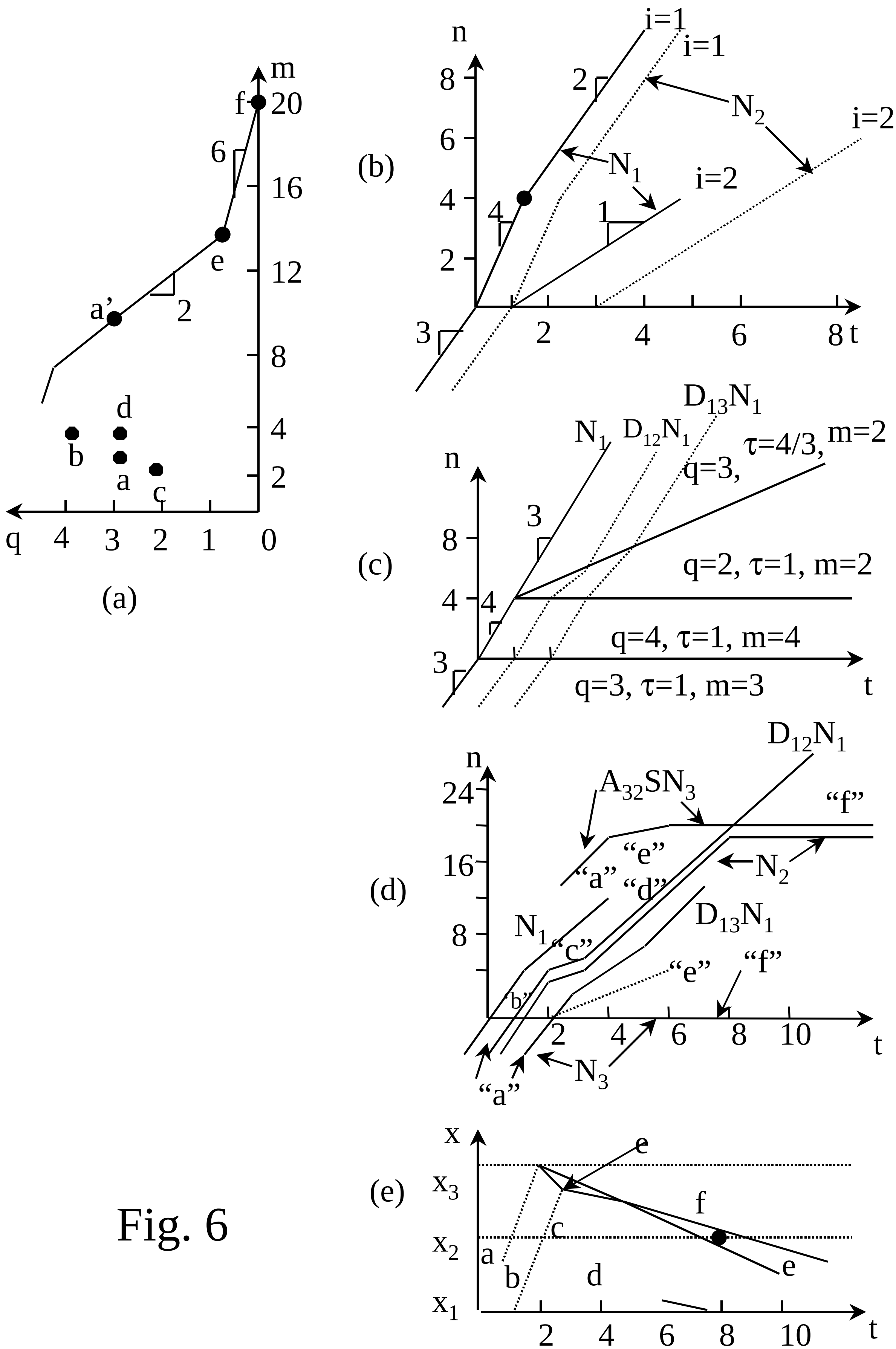


Fig. 5



m
f
20
6
16
e
12
a'
2
8
d
4
b
a
2
c
q
4
3
2
1
0
(a)
(b)
n
i=1
i=1
8
2
N2
i=2
6
N1
i=2
4
4
1
2
3
2
4
6
8
t
(c)
D13N1
N1
D12N1
q=3, τ=4/3, m=2
n
3
8
q=2, τ=1, m=2
4
4
q=4, τ=1, m=4
3
q=3, τ=1, m=3
t
(d)
D12N1
n
A32SN3
24
"f"
"e"
16
"a"
N2
"d"
N1
D13N1
8
"c"
"e"
"f"
"b"
2
4
6
8
10
t
N3
"a"
(e)
x
e
x3
f
c
x2
a
e
b
d
x1
2
4
6
8
10
t

Fig. 6

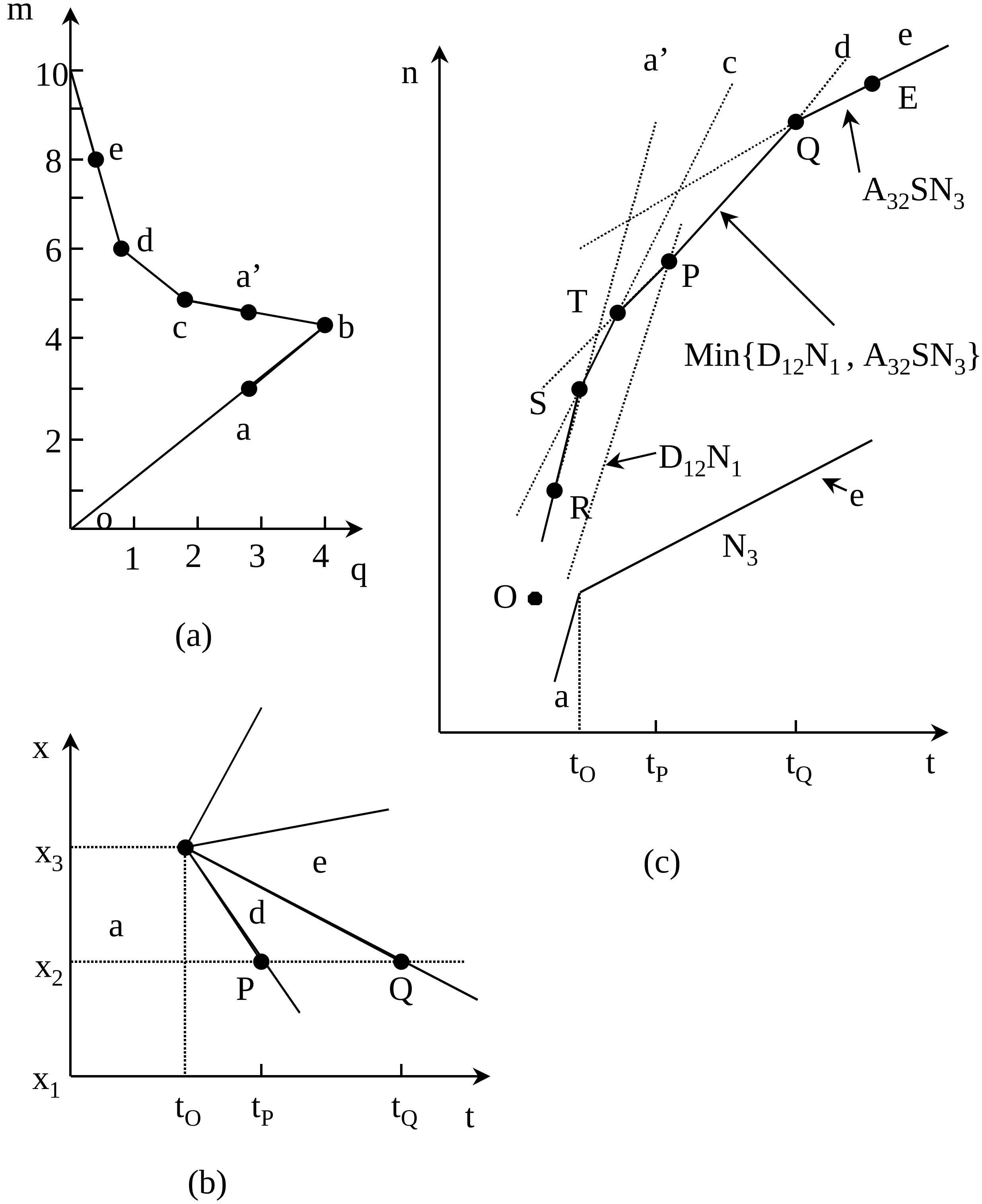

m
10
8
6
4
2
e
d
a’
c
b
a
o
1
2
3
4
q
(a)
x
x_3
x_2
x_1
a
e
d
P
Q
t_O
t_P
t_Q
t
(b)
n
a’
c
d
e
E
Q
$A_{32}SN_3$
P
T
Min{$D_{12}N_1$, $A_{32}SN_3$}
S
$D_{12}N_1$
R
e
N_3
O
a
t_O
t_P
t_Q
t
(c)

Fig. 7