Title
Theory of highway traffic flow: 1945 to 1965

Permalink
https://escholarship.org/uc/item/20s9h43s

Journal
Course Notes, 1995(1)

Author
Newell, Gordon F.

Publication Date
1995
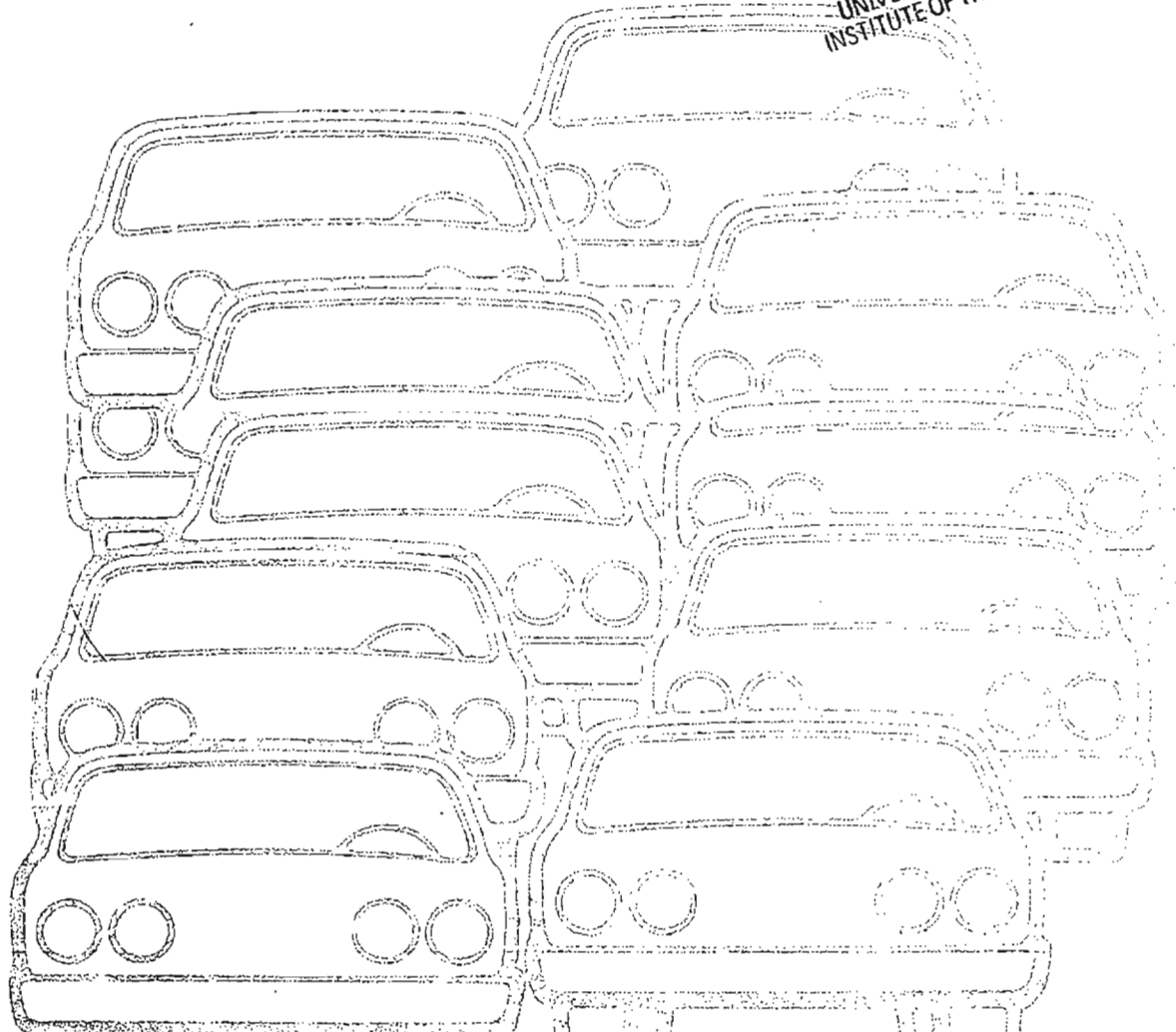
THEORY OF HIGHWAY TRAFFIC FLOW 1945-1965

G.F. NEWELL

COURSE NOTES UCB-ITS-CN-95-1

INSTITUTE OF TRANSPORTATION STUDIES
UNIVERSITY OF CALIFORNIA, BERKELEY

THEORY OF HIGHWAY TRAFFIC FLOW

1945 TO 1965

G.F. Newell

Institute of Transportation Studies

University of California. Berkeley
Institute of Transportation Studies

University of California, Berkeley

PREFACE

The following is a slightly edited version of lecture notes on traffic flow theory composed originally between about 1964 and 1966.

Modeling of highway traffic had progressed at a rapid pace during the late 1950s and early 1960s. In 1963, I gave a short course on traffic flow theory for honours mathematics students at the University of Adelaide in Australia at the invitation of Professor Renfrey Potts. The following year I gave a more extensive special topics course for graduate students in applied mathematics at Brown University. Most of these lecture notes were written for the latter course.

In 1965 I came to the University of California, Berkeley to teach in the transportation engineering graduate program. Although I taught courses on "traffic flow theory" and "highway traffic control" a few times in the early 1970s, the former based mostly on these lecture notes, enrollment in both of these classes was small and these courses were discontinued in the mid 1970s. An "advanced course" in traffic flow theory has not been given at UC Berkeley since then (as of 1994). At no time since 1965 have I supervised any students doing research on traffic flow theory and, until recently, very little of my own research was directed toward this subject.

students doing research on traffic flow theory and, until recently, very little of my own research was directed toward this subject.

Not only did I abandon these subjects, but so did most of the other people who had contributed to their development in the 1950s and 60s. The journals, Transportation Science and Transportation Research, which were both initiated by people who were active in the development of traffic flow theory and which were expected to be a vehicle for publication of

i

new developments in transportation theory, did not start until 1967. Most of the key literature on traffic theory, which appeared before 1967, is scattered over a variety of journals, symposium proceedings, and books. Much of this literature has been ignored by newcomers to the transportation field.

My own attempt to revive some of the lost theory started in 1984 when some students asked me to teach a special topics course on highway traffic control (after about a 10 year lapse). Although there had been little improvement in the theory (particularly on highway traffic signals) during the previous 10 years, there had been substantial advances in techniques for approximating queueing delays, and the analysis of queueing delays did not depend on an accurate theory for the dynamics of traffic flow (the delay to a driver caused by a traffic signals does not depend on <u>where</u> the driver waits). It was possible not only to revive some of the older works on traffic signals, but to write a fairly comprehensive analysis. This culminated in publication of a 450-page treatise, "Theory of Highway Traffic Signals," UCB-ITS-CN-89-1 in 1989.

Publication of the present notes was also inspired by a proposal to give a special topics course on traffic flow theory (after about a 20 year lapse). As compared with the theory of highway traffic signals, the status of traffic flow theory is quite different. There have been few significant developments over the last 20 years and much of what has been theory of highway traffic signals, the status of traffic flow theory is quite different. There have been few significant developments over the last 20 years and much of what has been done is even less realistic than the theories which existed 30 years ago. Many attempts to "improve" the theories have only made them worse.

The modeling of light traffic for which cars interact only occasionally and, at most, only two at a time, is fairly straightforward but of little concern to practical traffic engineers.

A theory of light traffic was essentially complete 30 years ago, to the extent that one would care to analyze it. If traffic is so congested that cars can seldom pass each other, a theory of traffic flow should, in principle, be straightforward if one knew precisely how one driver follows another. Unfortunately, we still do not know how drivers behave well enough to construct realistic models of "car-following." Indeed, we do not seem to understand it much better now than 30 years ago. The modeling of moderately dense traffic with clusters of cars, lane changing, etc. is extremely difficult and attempts to do so have not been very successful.

In contrast with the theory of highway traffic signals, there is not really much that I can add to what was written 30 years ago even though these notes end very abruptly. I wrote what I knew at the time expecting that new experimental observations would soon resolve some of the deficiencies of existing theories, and that I could add a concluding chapter to the part on dense traffic. The chapter on moderately dense traffic ends abruptly because I lost interest in pursuing something that appeared to be going nowhere.

Chapter I Introduction is mostly some commentary on the connection between highway traffic and statistical mechanics or the kinetic theory of gases. It is this similarity which attracted physicists and chemists to model traffic behavior in the 1950s. Chapter II deals with very light traffic in which interactions between cars are neglected completely. Traffic is represented simply as the superposition of independent vehicle trajectories. The key deals with very light traffic in which interactions between cars are neglected completely. Traffic is represented simply as the superposition of independent vehicle trajectories. The key result here is the tendency of traffic to behave like a Poisson process (in either space or time) with statistically independent velocities. Chapter III treats weak interactions, the first order effects of interactions between cars when they are close together. The key conclusion here is that, in this second approximation, traffic has a tendency to behave like the superposition of

independent Poisson processes of single cars and interacting pairs (if one can neglect interactions involving three or more cars). The treatments in Chapters II and III are very detailed and describe just about everything one would care to say about light traffic approximations. Any extension of this theory to involve interactions among three or more cars, however, would be very tedious.

Chapter IV on dense traffic gives an exhaustive analysis of models of car-following in which every driver chooses a velocity dependent only on the spacing between himself and the car he is following (possibly with a delayed response). This class of models includes or is equivalent to most of the car-following or (first order) fluid models of traffic flow at high density which had been proposed during the 1950s.

Chapter V describes some theories related to moderately dense traffic, but it does not include all the things promised in the introductory section. This chapter was never completed. The main topic here is an introduction to the theory of stationary stochastic point processes, a subject which is described in much more detail in the literature on applied probability (although mostly in books written after these notes were first written).

In the original notes, the present Chapter V was labeled Chapter VI. Chapter V had not been written. It was to have been a follow-up to Chapter IV containing more realistic theories of dense traffic explaining instability, "stop and go" driving, etc. I had already not been written. It was to have been a follow-up to Chapter IV containing more realistic theories of dense traffic explaining instability, "stop and go" driving, etc. I had already proposed a possible structure for such a theory by 1962, but it was still rather speculative and lacking quantitative verification. I expected then (in 1965) that a more refined theory would soon emerge to complete the story (but it still has not happened).

The original notes also contained the start of a Chapter VII dealing with delays at an isolated fixed-cycle traffic signal, but the theory of highway traffic signals was in an early stage of development then. Anything of value contained in this chapter of the notes has since been absorbed in the above mentioned treatise on traffic signals.

To close out the present notes I have added a new Chapter VI "postscript" commenting on some things which I might explain differently today, some things that have happened during the last 30 years, and some theories old and new which I believe have failed(even more so than those described in the previous chapters). I have also attached a chronological bibliography on car-following and continuum theories which, I believe, is nearly complete to 1972. This at least illustrates the rapid rise and decline of activity during the 1950s and 60s.

The reproduction of these notes would not have happened without considerable prodding from Carlos Daganzo who offered some hope that new experimental techniques may soon resolve some of the deficiencies of existing theories.

Most of the huge job of typing equations was done by Ping Hale. Nadine Zalinsky typed much of the text. Reinaldo C. Garcia proofread the entire text.

General References

1. M. Beckmann, C.B. McGuire, and C.B. Winsten, "Studies in the economics of transportation," Yale University Press (1956).

2. R. Herman, ed., Proc. 1st international symposium on theory of traffic flow (1959), Elsevier Press (1961).

3. F.A. Haight, "Mathematical theories of traffic flow," Academic Press, N.Y.C. (1963).

4. D.L. Gerlough and D.G. Capelle, eds., "An introduction to traffic flow theory," Highway Research Board Special Report 79, National Research Council, Washington, D.C. (1964).

5. Joyce Almond, ed., Proc. 2nd international symposium on the theory of road traffic flow London 1963, The Organisation for Economic Co-operation and Development, Paris (1965).

6. W.D. Ashton, *The Theory of Road Traffic Flow*, Methuen Monographs, London (1966).

7. L.C. Edie, R. Herman and R. Rothery, eds., Vehicular Traffic Science, Proc. 3d. International Conference on Theory of Traffic Flow (1965), Elsevier (1967).

CONTENTS

I. INTRODUCTION

1. <u>Structure of mathematical equations</u>. If it were possible to construct a complete theory of traffic flow which in principle predicts the motion of each and every car on the highway, the equations describing this motion would probably be too difficult to solve and if solved would only produce a library of figures most of which would be of no practical value. Nevertheless the gross features of traffic flow that we do consider to be important must be a result of the collective behavior of individual cars.

Although we lack a complete theory of the motion of individual cars, there are many simple facts that even the most inexperienced driver knows and there are others which we could determine through experiment if we thought these facts were worth the effort required to find them. The lack of such a theory however, should not deter us from constructing a framework of possible theories consistent with what is known and seeing if such an incomplete theory can give any useful information about the gross aspects of traffic. Also by considering various hypothetical motions of individual cars one may draw conclusions regarding what features of the motion of single cars are relevant to the large scale behavior of traffic.

The above situation is similar to that which faced physicists almost a century ago. They traffic.

The above situation is similar to that which faced physicists almost a century ago. They knew many laws regarding the bulk properties of matter. They were also convinced that matter was composed of atoms and that there were certain laws of motion for the individual atoms. The problem was to find the connection between one set of laws and the other and to

1

use experimental results obtained on large systems to infer properties of small systems or vice versa. The study of these problems generated the branch of physics called statistical mechanics.

Cars are driven by people and do not satisfy the laws of classical mechanics or quantum mechanics as do atoms, but in constructing any theory of traffic one must start from the conjecture that drivers do behave according to some pattern. It would be incorrect to assume that all drivers behave in the same way or that even a single driver will always react the same in a given situation but we will assume that there exists some probability distribution of behaviors. If a single driver or different drivers with similar desires are repeatedly confronted with the same surroundings, they will respond in a given way a certain fraction of the time. Whether or not we actually find these probabilities is not at issue. We only assume that they can in principle be found or inferred if we were willing to expend sufficient effort to find them.

Despite the fact that statistical mechanics deals with large aggregates of particles which specifically obey equations of particle dynamics, much of the mathematics or logic of statistical mechanics can be translated into a corresponding theory of traffic flow, namely those parts of statistical mechanics that deal only with the interrelation between macroscopic and microscopic laws. The similarity, however, goes even further than this. One important those parts of statistical mechanics that deal only with the interrelation between macroscopic and microscopic laws. The similarity, however, goes even further than this. One important feature of atomic forces is that they are usually of "short range," i.e., two atoms which are sufficiently far apart do not influence one another. The same feature is true of the interaction between cars. Two drivers who are out of visible range of each other are not expected to influence one another. In statistical mechanics this fact that there are short range forces is the

key to the theory of an ideal gas and in traffic theory one can construct simple models of flow at arbitrarily low density in an analogous way.

In physics the behavior of bulk systems changes drastically as one varies the density of atoms from the ideal gas limit where the average distance between atoms is large compared with the range of force to the limit of close packed structures (solids) where the distance between atoms is comparable with or less than the range of force. The corresponding situation is true of highway traffic also.

The development of traffic theory is even following a pattern similar to the historical development of statistical physics. In physics the theory of gases based upon nearly independent motions of atoms developed early and has advanced rapidly, as has the theory of solids based upon a highly ordered motion of atoms, but the theory of intermediate systems, particularly the theory of the liquid state, has progressed very slowly. In traffic also, a theory for low density and a theory for high density are off to a good start but the intermediate densities present difficult problems.

Despite the similarities between traffic and statistical physics, there are obvious differences in addition to the fact that cars do not satisfy the laws of dynamics. Cars are constrained to move on highways which are parts of a two-dimensional space and in most cases can be considered essentially networks of one-dimensional spaces. This should be a constrained to move on highways which are parts of a two-dimensional space and in most cases can be considered essentially networks of one-dimensional spaces. This should be a simplification over the inherently three-dimensional nature of particle motion. On the other hand the most annoying difference is that one seldom has the occasion to consider more than a relatively small number of cars at a time, perhaps 10 or 100 or even 1000 and the drivers of these cars are not all the same. Statistical fluctuations in observations are therefore quite

large. In physics a typical system is likely to contain something like 10^{20} or 10^{30} particles all with identical properties. Whereas typical fractional fluctuations in density for example of a physical system are of order 10^{-10}, those of traffic may be anything from a few percent to 30 or 40 percent.

The logical foundations of statistical mechanics have been the subject of heated debate for more than 75 years. In developing the mathematical formalism one still must make heuristic arguments, try to argue away certain paradoxes and finally say that despite the loopholes in the arguments, the theory must be correct most of the time because it gives the answers we wanted. Unfortunately the parts of statistical mechanics which we wish to mimic in traffic theory are just those which are so controversial. We can only hope that what seems plausible will again prove to be correct most of the time. For example, in studying the behavior of 100 drivers on the road, can one give some logical argument for selecting these 100 cars at random from a population of possible drivers when it seems clear that the behavior of these cars does depend upon which drivers are selected and there is no grand roulette machine in nature which picks the drivers which are to drive on the highway each day?

The development of statistical mechanics usually starts from consideration of a system of N particles with known properties (mass, electric charge, etc.) and with known equations of The development of statistical mechanics usually starts from consideration of a system of N particles with known properties (mass, electric charge, etc.) and with known equations of motion (the laws of classical mechanics, for example). If at time t = 0 one specifies the positions and velocities of all particles, then the laws of motion determine, in principle, the positions and velocities at any later time. One is immediately confronted with the following fact. For "most" initial states, the macroscopic behavior of the system depends only upon a

4

few macroscopic properties of the initial state, but if one tries to choose an initial state so as to make the solution of the dynamical equations simple (for example if one takes all particles in a gas to have velocities along coordinate directions), one is likely to be unlucky enough to pick one of the rare initial states that is not "typical" in its macroscopic behavior. To avoid this difficulty one introduces some probability distribution over the initial states and investigates only the average behavior of this ensemble of initial states. This is done in such a way that the anomalous initial states have probability zero and give no contribution to the average behavior. Many of the controversies in statistical mechanics (the famous ergodic hypothesis, etc.) deal with the arguments from which one selects a reasonable probability distribution over the initial states.

The above picture literally relates only to a physical system that is <u>closed</u> and <u>isolated</u>, i.e., there is a fixed set of particles in the system, and they do not interact with other physical systems. Actually the physical systems to which one applies statistical mechanics are never really isolated and are seldom closed but one can imagine a hypothetical physical system which differs from the real one only in that the hypothetical one is put in a box with perfectly reflecting walls. For sufficiently large systems, the effect of a box upon the physical properties of interest is usually small (it varies like the ratio of the area of the box to its volume). There are, however, extensions of statistical mechanics to open systems in which particles ties of interest is usually small (it varies like the ratio of the area of the box to its volume). There are, however, extensions of statistical mechanics to open systems in which particles leave and enter the system usually according to some stationary probability law.

Traffic systems which are either completely or nearly isolated do exist. An ideal example is a collection of cars on a circular track or a collection of cars trapped between two trucks. A city or metropolitan area is also nearly isolated in the sense that most traffic is

local. One would not change the pattern of traffic very much if each car leaving the city is replaced by another car entering (a reflecting wall for cars). On a larger scale, the traffic on the continents of the world is nearly closed—and certainly the world itself is closed. Most traffic systems which we will want to study, however, are not even approximately closed. To study the traffic on a single road, for example, or a simple network one seldom can avoid considering the traffic entering and leaving the boundaries of the system.

In statistical mechanics it is common practice to consider a physical system which is a small subsystem of a large isolated system (the so-called thermal bath). In dealing with traffic it is also appropriate to consider traffic on a simple highway as a small subsystem of the traffic in an entire city or continent. The city becomes the "thermal bath" with which the subsystem exchanges cars. It is worth noting that even if we included all cars in the world as our system we would have only of the order of 10^8 cars (few cities have more than 10^6) which from the point of view of statistical mechanics would be a relatively small system.

In physics the information necessary to describe a physical system can be separated into the following categories: 1. a description of the particles in the system, their mass, electric charge, spin, etc.; 2. the forces acting in the system including forces between particles and external constraints. (This is usually described through the Hamiltonian and in a certain sense includes 1.) and 3. a specification of the initial conditions necessary to uniquely define and external constraints. (This is usually described through the Hamiltonian and in a certain sense includes 1.) and 3. a specification of the initial conditions necessary to uniquely define the solution of the equations of motion. One thinks of the initial conditions as a point in a multi-dimensional space of all possible initial conditions, the state space or phase space. The motion of the system is then represented by a trajectory, the path in the state space defined by the parametric representation of the state of the system as a function of time. A theory, in this

case the laws of dynamics, is simply a proposal that one collection of data can be inferred from another; here the state at time t is deduced from the state at time zero, if one is given all the other data above.

In an attempt to construct theories of traffic one should keep in mind that theories are not formulas that describe everything one wishes to know. One tries to find as many relations as one can between observable quantities but we would have a theory even if we could find only one relation among a very large number of things. Certainly for the dynamical system described above there is a tremendous number of physical quantities the values of which one must find from observation. We have no theory yet which tells us what the masses of particles must be. Neither is it likely that we will find any satisfactory theory which will tell us how fast a driver wants to drive. If we know how fast a driver drives and where he is, however, we can say something about where he will be a short time later.

Because of various structural similarities between traffic and particles, one can also classify much of the data relating to traffic in categories analogous to those listed above even though as yet we have not proposed a theory which will be the analogue of the laws of dynamics.

In the first category belong properties associated with an individual car or driver which are more or less independent of time. There are potentially so many of these that one

In the first category belong properties associated with an individual car or driver which are more or less independent of time. There are potentially so many of these that one would not want to list all that one can imagine. The object is rather to list as few as are necessary to describe any particular theory. Some properties which will enter into theories discussed here are the origin, destination, possible routes and starting times of a driver's trips (these properties which indicate an objective have no obvious counterpart in particle physics),

his desired speed on various highways, his notion of safe driving distance behind another car, his willingness to accept passing opportunities, etc.

In the second category belong the "forces" of interaction between cars if not already included in the first category. The external forces or constraints of a physical system, however, have an obvious analogue in the geometry of the road system, traffic lights, etc. Finally in the third category belong the time dependent state variables such as the positions of the cars at any time whose evolution we hope to describe by some theory analogous to the equations of dynamics.

Although we have given only a very fuzzy indication of what might enter into a microscopic theory, we turn now to some of the statistical questions. Since we know that for physical systems much of the microscopic behavior is irrelevant to the macroscopic behavior, we postpone any further discussion of the microscopic behavior of traffic until we have some better indication of what features of the microscopic behavior have the greatest influence on the gross properties.

In statistical mechanics there are three types of arguments used to explain why certain details of the microscopic motion are irrelevant.

1. Most physical measurements can be made only in a time which is large compared with the time interval between microscopic events. For example, a measurement of tempera-
1. Most physical measurements can be made only in a time which is large compared with the time interval between microscopic events. For example, a measurement of tempera-ture with a thermometer takes several seconds during which time many particles collide with each other and with the thermometer. In effect one is measuring a time average property of the system. If the time of measurement is large compared with the time for the system to reach "thermodynamic equilibrium," then for all practical purposes the time average over a

8

finite time is equivalent to an average over an infinite time. Mathematically one can represent this in the following way. If the coordinates of the state space are represented by a collection of numbers or vectors $\{x_j\}$, then a trajectory of the system is a set of functions $\{x_j(t)\}$ that gives the state of the system at time t. If an instantaneous measurement would give a quantity $f(x_1(t), x_2(t), \cdots)$ some function of the state of the system at time t, then the quantity which one actually observes is

$$\frac{1}{T}\int_0^T dt\, f(x_1(t),\, x_2(t),\, \dots)\ . \tag{1.1}$$

We expect that this will be much less sensitive to the initial state $\{x_j(0)\}$ than the instantaneous value of f at any time t.

2. The physical quantities one thinks of as macroscopic variables all seem to have the form

$$\sum_j f_j\,(x_j) \quad or \quad \sum f_{j,k}\,(x_j, x_k) \tag{1.2}$$

where here the x_j stands for some possibly vector valued property associated with the jth particle and the sums are over all particles or pairs of particles respectively. Furthermore if all particles are identical as in a monoatomic gas then the functions f_j do not depend upon the particle number j. For example, the total momentum of the system is the sum of the momenta of all the particles and the total potential energy is usually the sum of the energies of particle number j. For example, the total momentum of the system is the sum of the momenta of all the particles and the total potential energy is usually the sum of the energies of interaction between pairs. The number of particles in a region D is obtained by choosing f_j to be 1 if x_j is a point in D and zero otherwise. The average density in D is then defined as this number divided by the volume of D. Since the sums extend over a very large number of particles one expects considerable cancellation of the fluctuation and a value for this sum

9

which is not very sensitive to the state of the system. The above sums are in effect averages over all particles.

3. If the above arguments fail and the physical observations do depend upon certain detailed properties of the microscopic state, it is likely that these observations are so sensitive to small changes in this state that it is impossible in practice to reproduce the same state with sufficient accuracy to guarantee the same outcome of an experiment. A coin, for example, is a rigid body that supposedly obeys the laws of classical dynamics. If we specify the initial velocity, angular velocity, position etc. of the coin and describe the geometry of the region into which it is thrown, we should be able to determine with certainty whether the coin will land heads or tails. Whether a coin lands heads or tails after spinning many times, however, is so sensitive to these initial conditions that is practically impossible to set the initial conditions accurately enough to predict the outcome. The fact that a coin lands heads about half the time is not simply because the coin is symmetric but because the uncertainty in selecting the initial conditions is such that about half the time we choose the initial conditions from the set of all possible initial conditions that leads to heads.

A similar argument may apply to radioactive decay of nuclei. We do not know enough about the states of nuclei to predict under what conditions a nucleus will radiate. If we observe enough decays, however, one can say something about the average rate. This average about the states of nuclei to predict under what conditions a nucleus will radiate. If we observe enough decays, however, one can say something about the average rate. This average rate is then considered as the effective macroscopic observable rather than the non-reproduce-able number of decays observed in a single experiment.

The mathematical formulation of this we obtain by introducing a probability distribu-tion of possible initial conditions. If $\{x_j(t)\}$ is the state of the system at time t, then $x_j(t)$ is

also a function $x_j(t, x_1(0), x_2(0), \cdots)$ of the initial state $x_1(0), x_2(0), \cdots$. If we assume that the initial states are chosen at random with a probability density $p(x_1(0), x_2(0), \cdots)$ and we are concerned with a function $f(x_1(t), x_2(t), \cdots)$, then the macroscopic quantity of interest is the expectation or average value of f over the probability distribution of initial states, i.e.,

$$\int \cdots \int dx_1(0)dx_2(0)\cdots dx_N(0)p(x_1(0),x_2(0),\cdots)f(x_1(t,x_1(0),x_2(0),\cdots), x_2(t,\cdots),\cdots) \tag{1.3}$$

where if $x_j(0)$ is a vector, $dx_j(0)$ represents a volume element in the space of $x_j(0)$ values. Fortunately the three types of averages described above, complement each other in that by using one of them we do not destroy the possibility of using another. Furthermore the result of successive averages is independent of the order in which they are done. For most macroscopic quantities one can in fact argue that what one observes involves all three averages i.e. a time average of a particle average of an average over some distribution of initial states. If, as one often hopes, the time average of some function f is independent of the initial state and independent of permutations of the particles or the particle average is independent of the initial state and time or the average over initial states is independent of time and permutations of the particles, then the use of more than one type of average is redundant because the second or third average becomes the average of a constant which always yields this same constant independent of the distribution over which one averages.

Each of the above arguments has an obvious application to traffic since we have always yields this same constant independent of the distribution over which one averages.

Each of the above arguments has an obvious application to traffic since we have nowhere used any properties of the state space or the equations of motion. We have assumed, however, that the equations are deterministic, i.e. $\{x_j(t)\}$ is uniquely determined from $\{x_j(0)\}$, but we could have assumed that for any initial state $\{x_j(0)\}$ there is a probability distribution

11

for the states $\{x_j(t)\}$. In the latter case one could perform still a fourth average over all possible motions of the system.

Unfortunately the above ideas which are so effective in describing mass phenomena in physics will never give more than a very crude description of highway traffic. In physics one is interested in time averages over times which are in effect very large but in traffic there is little interest for example in flow on a highway averaged over a year's time in which one averages out hourly variations, seasonal variations, etc. but to average only over a few minutes is usually not very effective in smoothing out random fluctuations. In physics a particle average is typically over enormously many particles but in traffic the number of cars one observes at any time is quite small. In physics the laws of nature are assumed to be valid for an indefinite length of time and one can repeat an experiment as many times as one wishes under what seem to be equivalent conditions. If the outcome of an experiment is random one can find the average behavior by repeated trials. In traffic one is never quite sure if one is repeating an experiment under nearly equivalent conditions, and furthermore one cannot repeat the experiment indefinitely because the traffic behavior is known to vary from year to year. The "laws" of traffic are not valid for all times in the future. In traffic one will always be confronted with the problem of not having as much data as one would like.

These inherent limitations on the accuracy of any theory of traffic flow are things that always be confronted with the problem of not having as much data as one would like.

These inherent limitations on the accuracy of any theory of traffic flow are things that one must learn to accept. Certainly a traffic engineer is not concerned about the detailed behavior of each and every driver in a city; he is only interested in the typical or the average. That the latter is only very crudely defined cannot be remedied by better theory or more accurate data. A theory that gives predictions to an accuracy of 25% or even a factor of 2 is

better than nothing and perhaps in some cases the ultimate accuracy of any prediction. That it does not give the 10% or the 1% accuracy that engineers are accustomed to expect in engineering applications of physics is not always a reflection of a poor theory but more likely a crudely stated but nevertheless relevant question.

Bibliographical Notes

There is no published literature in which statistical mechanics and highway traffic are compared although the papers listed below by Cohen, Newell, and Prigogine were obviously motivated by similarities in the two subjects and other works to be discussed later certainly mimic many of the techniques of statistical mechanics even if it was not intended. There is a very large number of books on statistical mechanics but most of the modern textbooks emphasize the methods of applying the conclusions to current physical problems. The references given below contain the most thorough discussion of foundations. The review paper by P. and T. Ehrenfest is a classic and gives a thorough history of the early developments of the kinetic theory of gases and a penetrating analysis of the controversies. Despite the early date, much of the discussion is as relevant today as then. The book by Tolman is also quite old in terms of the developments in physics but it is still one of the best books the early date, much of the discussion is as relevant today as then. The book by Tolman is also quite old in terms of the developments in physics but it is still one of the best books available. The book by Khinchin treats the subject from a rigorous mathematical point of view (to the extent that such is possible).

References

1. M. H. Cohen, "The relative distribution of households and places of work." Theory of Traffic Flow, R. Herman ed. Elsevier 1961 p. 79-84.

2. G. F. Newell, "Mathematical models of freely-flowing highway traffic." Operations Research 3 (1955), 176-186.

3. I. Prigogine, "A Boltzmann-like approach to the statistical theory of traffic flow." Theory of Traffic Flow, R. Herman ed. Elsevier 1961 p. 158-164.

4. P. and T. Ehrenfest, "The conceptual foundations of the statistical approach in mechanics" translated by M. J. Moravicsik, Cornell University Press 1959 from Vol. 4, Encyclopadie der Mathematischen Wissenschaften (1911).

5. R. C. Tolman, "The principles of statistical mechanics," Oxford 1938.

6. A. I. Khinchin, "Mathematical foundations of statistical mechanics," translated by G. Gamow, Dover 1949.

II. LOW DENSITY TRAFFIC (NO INTERACTION)

1. <u>Introduction</u>. Since any realistic mathematical description of traffic flow is certain to be so complicated that we would have great difficulty in analyzing it, the main object at the present state of development of traffic theory is to construct models which are as simple as possible but still contain some similarity with certain aspects of real traffic. We shall therefore begin by considering what seem to be the crudest possible models and then gradually add refinements to them.

The three basic ingredients of any theory are 1. a description of the population of drivers. 2. a description of the road network and 3. some equations of motion with an appropriate specification of an initial state. The simplest population of drivers is a population of identical drivers. The simplest road network is a homogeneous highway of length L say with one entrance at x = 0 and one exit at x = L. The simplest equations of motion are that each driver travels at a constant velocity v. We represent the highway as a one-dimensional line and a car by a point (the position of its center for example). If $x_j(t)$ is the position of the jth car at time t then the equations of motion are

$$\frac{dx_j(t)}{dt} = v \quad for\ 0 < x_j(t) < L \tag{1.1}$$

$$\frac{dx_j(t)}{dt} = v \quad for\ 0 < x_j(t) < L \tag{1.1}$$

the solution of which is

$$x_j(t) = x_j(0) + vt \quad if\ 0 < x_j(t) < L$$

or

$$x_j(t) = v(t - t_{jo}) \quad for\ 0 < t - t_{jo} < L/v$$

15

if t_{jo} is the time the car enters the highway at $x = 0$.

This system of equations is so simple that it seems hardly necessary to say more. To lay the groundwork for future refinements, however, it is convenient to use this model as a means of illustrating some of the dynamical and statistical concepts that will be necessary in the treatment of more complicated models and also to see how some of the ideas of chapter 1 relate to this model.

Unlike classical mechanics in which the equations of motion are second order differential equations and therefore require a specification of both initial positions and velocities, (1.1) is a system of first order equations. Even for more general models it seems reasonable to assume for traffic flow that the desired speed of a driver is a property of the individual driver (analogous to the mass of a particle), something which the driver retains for all time. For a system in which a driver deviates from his desired speed to stop for a traffic signal or to slow down as he overtakes another car, etc. it is still reasonable to assume that the driver's behavior at any time depends only (or at least mainly) upon his position on the highway and the positions of other cars. We might therefore postulate for a fairly general class of models that the state space of the system is the space of position coordinates only, but not also velocities (although the description of the population of drivers implies a specification of the properties of the jth driver including his desired speed).

but not also velocities (although the description of the population of drivers implies a specification of the properties of the jth driver including his desired speed).

The state space of the system under discussion thus consists of spatial coordinates x_j, $j = 1, 2..., N$ for which N represents the total of all cars which may use this highway during any specified time of observation. Those cars which are not on the road at any

particular time can be assigned any coordinates not in (0,L). They may be in parking lots or on other highways.

The equations of motion of our system consist of some rule whereby we can determine the position $x_j(t)$ at any time t if we know the corresponding positions at any earlier time. In the present situation we determine $x_j(t)$ of the cars in $(0, L)$ at time t from the positions at an earlier time for the cars in $(0,L)$ at the earlier time through (1.1), otherwise from the times at which they enter $(0, L)$. The entering times can be considered as properties of the population or as part of the equations of motion not covered by (1.1). The evolution of our system is then represented by a trajectory $(x_1(t), x_2(t), \cdot \cdot)$ in the space of coordinates $(x_1, x_2, \cdot \cdot)$. If cars not in $(0, L)$ are assigned the coordinates 0 or L, this trajectory is a piece-wise linear curve in the N dimensional space which changes direction each time a car either enters or leaves $(0, L)$.

One can also represent trajectories as a curve in an $N+1$-dimensional space $(t, x_1, x_2,..,x_N)$. The N-dimensional curve is then the projection of the N+1 dimensional curve onto the N-dimensional space with the parameter t represented as a parameter along the latter curve.

Because of the difficulty in visualizing trajectories in an N or $N+1$-dimensional space, it is more usual to draw two-dimensional projected (x_j,t) graphs. On the same graph of

position vs. time we draw each single car trajectory $x_j(t)$. This has the obvious advantage of simple geometric representation but, if there is an interaction between cars, it has the disadvantage that the trajectory of one car depends upon those of other cars whereas in N or N+1 dimensions there is just one trajectory which describes all cars simultaneously.

For the particular model proposed here, all cars travel at the same velocity. The single car trajectories are all straight lines of slope v as shown in figure II.2 rather than the more realistic type of picture such as in figure II.1. The spacing between the cars is fixed and there is no passing. Furthermore if the spacing between cars is larger than the range of influence between cars, the velocity v can be interpreted as some average free speed or desired speed of the drivers.

The most important feature of this model that makes the equations of motion manageable is that there are no interactions and so the trajectory of one car is functionally independent of any other car. Although we really have a system of N simultaneous equations, these equations are uncoupled and equivalent to the superposition of N independent one car systems.
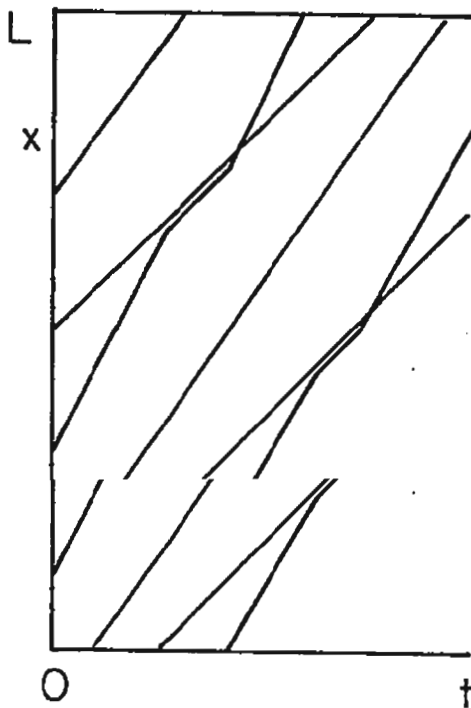

Fig. II 1
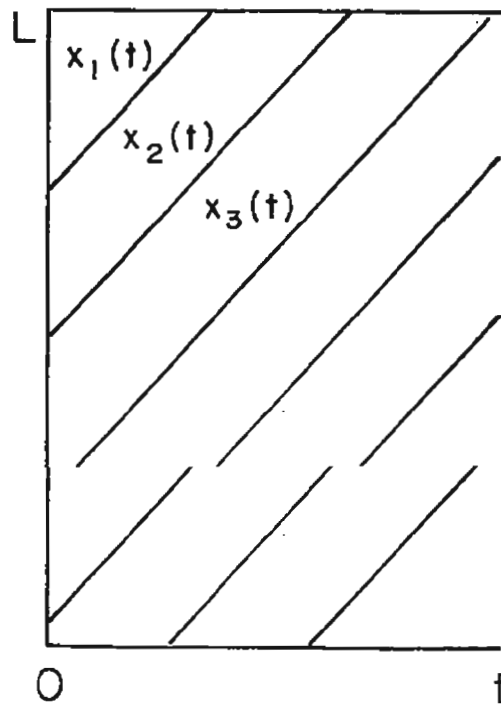Typical trajectories of cars

Fig. II 2
Trajectories for cars all traveling at the same velocity

2. <u>Macroscopic quantities</u>. Three quantities that experimentalists frequently measure are the density of cars on the highway, k, the mean velocity, v , and the flow, q. The flow, also called the volume of traffic, is, roughly speaking, the number of cars crossing a fixed point on the highway per unit time. In the present model in which all cars travel at the same velocity, the mean velocity must be v but the precise definitions of k and q are less obvious.

Density of cars, k, should represent the number of cars per unit length of highway. Unfortunately the number of cars in any section of highway must be integer and if we take an arbitrarily small interval about some point x, as would be usual in defining a density, it will contain either 0 or 1 cars. The density is either 0 or ∞ in the limit of zero length of highway. A similar problem arises, however, in fluid dynamics. The density of mass is defined as the mass per unit volume but since mass is atomic one cannot take an arbitrarily small volume to define density at a point.

There are two approaches to this problem. By analogy with the usual procedure in physics, one can imagine a length of highway sufficiently long as to contain many cars but sufficiently small that the gross aspects of the system can be considered constant over this section of highway. This scheme works well in physics because any element of volume that is small in terms of visible dimensions of length usually still contains an enormous number of section of highway. This scheme works well in physics because any element of volume that is small in terms of visible dimensions of length usually still contains an enormous number of particles. Also it is the nature of the physical laws that local concentrations of particles diffuse very rapidly. Even though there will be fluctuations in the number observed in volume elements of the same size, these fluctuations will be small compared with the number of particles observed. In traffic there are serious problems in observing a density because the

19

lengths of highway one would like to use in evaluating a density seldom contain very many cars but if one increases the length one is apt to exceed the range of highway over which traffic can be considered more or less homogeneous. Nevertheless this is the only practical way of evaluating a density unless one makes repeated observations and computes an average, but then one has the problem of ascertaining if the different observations were done under equivalent conditions.

The second approach to defining a density is mathematically more satisfactory but experimentally impractical. One first defines a joint probability distribution over the possible states of the system. Suppose $F(x_1, x_2, \cdots, x_N; t)$ is the probability that car 1 has coordinate less than x_1, car 2 less than x_2, etc. for $0 \leq x \leq L$ at time t. Any car not in (0,L) can be put at 0 or L. The marginal probability that car j has coordinate less than x_j is

$$F_{x_j}(x_j; t) = F(\infty, \infty, \ldots, x_j, \infty \ldots; t) . \tag{2.1}$$

If $F_{x_j}(x; t)$ is differentiable then there is a probability density for the coordinate $x_j(t)$ of car j

$$f_{x_j}(x; t) = \frac{\partial}{\partial x} F_{x_j}(x; t) \tag{2.2}$$

which is interpreted to mean that for small dx, $f_{x_j}(x; t)dx$ is the probability that $x_j(t)$ lies between x and x + dx. The density of cars at x is then

between x and x + dx. The density of cars at x is then

$$k(x, t) = \sum_{j=1}^{N} f_{x_j}(x; t) . \tag{2.3}$$

The connection between the two interpretations of density arises as follows. Let $x_j(t)$ be the random position of car j and for any small interval (a,b) in (0,L) let $\chi_{ab}(x)$ be the set characteristic function of (a,b) i.e.,

20

$$\chi_{ab}(x) = \begin{cases} 1 & \textit{if } a < x < b \\ 0 & \textit{otherwise} \end{cases} \tag{2.4}$$

$\chi_{ab}(x_j(t))$ is a random variable with value 1 if the jth car is in (a,b) at time t and zero otherwise.

$$\sum_{j=1}^{N} \chi_{ab}(x_j(t)) \tag{2.5}$$

is the number of cars in (a,b) at time t. If we divide this by (b-a) we obtain what is experimentally evaluated as the density.

We hope that the stochastic structure of the $\{x_j(t)\}$ is such that a law of large numbers applies, which in the present context would imply the value of (2.5) divided by the expectation of (2.5) would be nearly equal to one for almost all realizations of the random variables $x_j(t)$. For this to be true, however, we must choose (b-a) sufficiently large that many $\chi_{ab}(x_j(t))$ are non-zero. The second definition of density (2.3) is equivalent to

$$k(x, t) = \lim_{a \to b \to x} \frac{1}{(b-a)} E\{ \sum_{j=1}^{N} \chi_{ab}(x_j(t)) \} \tag{2.6}$$

in which $E\{x\}$ represents the expectation or average of x.

The interpretation of the flow q has similar problems, in fact the only difference between k and q is that k is the spatial density, the number per unit length of road, while
The interpretation of the flow q has similar problems, in fact the only difference between k and q is that k is the spatial density, the number per unit length of road, while q is the "time density." In the (x,t) space, k is the density of crossings of the x axis at fixed t and q is the density of crossings of the t axis for fixed x by the trajectories $x_j(t)$. Furthermore for the case of equal velocity v for all cars, the q and k are simply related by the equation

$$q = kv. \tag{2.7}$$

This one can see from the fact that if at any instant there are n = kx cars on an interval of highway (0,x), then these n cars traveling at velocity v will all pass the position x in a time x/v. Thus n = kx = q x/v.

Edie [1] has suggested another interpretation of q and k which is in practice somewhat easier to apply. Consider any area A (which for convenience we choose as a convex set) in the (x,t) plane (see figure II.3). Let Δx_j be the distance traveled in A by car j and Δt_j be the time traveled in A by car j. Now define the density of cars k and the flow of cars q in A by

$$k = \frac{1}{A} \sum_j \Delta t_j, \quad q = \frac{1}{A} \sum_j \Delta x_j \tag{2.8}$$

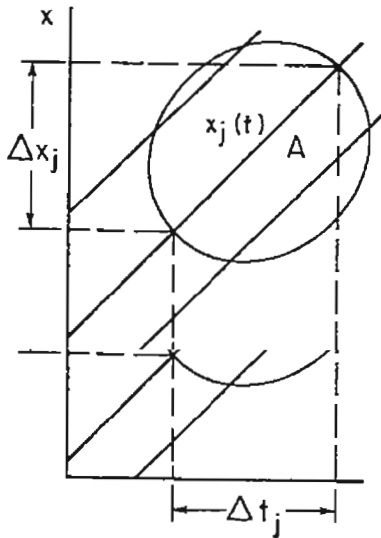in which A is the area of the region A (dimensions length time).

Fig. II 3
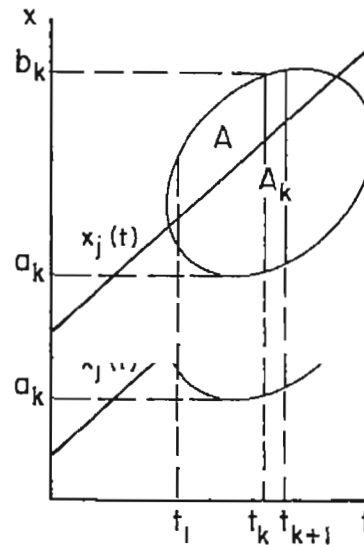Sections of trajectories enclosed by area

Fig. II 4
Equivalence of definitions

The k and q so defined are random variables and if A is too small we again have difficulties analogous to those described above because A may be empty or contain only one non-zero Δx_j. We can formally avoid this problem by defining k and q by

$$k(x, t) = \lim_{A \to 0} \frac{1}{A} \sum_j E\{\Delta t_j\} \;;\; q(x,t) = \lim_{A \to 0} \frac{1}{A} \sum_j E\{\Delta x_j\} \;. \qquad (2.9)$$

in which A->0 implies that the largest distance or time interval in A vanishes while A converges on the point (x,t).

If all relevant functions of x and t are continuous functions of x and t, then the definition of k(x,t) in (2.9) is equivalent to (2.3) or (2.6). To show that this is true, divide the region A into narrow vertical strips as shown in figure II 4, the kth strip A_k including times $t_k < t \le t_{k+1}$. The time Δt_j which the jth car spends in A can be written as the sum of the times spent in each of the A_k. Since the velocity v is finite (the trajectories are not vertical) most trajectories which enter A_k will spend a time $t_{k+1}-t_k$ in A_k. The only exceptions are those trajectories which enter or leave through the upper or lower ends of A_k near the positions b_k or a_k respectively. These latter, however, will contribute nothing in the limit of arbitrarily narrow strips if f(x,t) is continuous and the boundary of A is smooth.

It follows then that

$$E\{\Delta t_j\} \approx \sum_k E\{time\ spent\ in\ A_k\ by\ car\ j\}$$

It follows then that

$$E\{\Delta t_j\} \approx \sum_k E\{time\ spent\ in\ A_k\ by\ car\ j\}$$

$$\sim \sum_k (t_{k+1} - t_k)\, P\{j^{th}\ car\ is\ in\ A\ at\ time\ t_k\}$$

$$= \sum_k (t_{k+1} - t_k) \int_{a_k}^{b_k} f_{x_j}(x,\, t_k)dx \;.$$

In the limit $t_{k+1} - t_k \to 0$, this last sum is itself an integral over the time and we have

23

$$E\{\Delta\ t_j\} = \iint\limits_{A} dx\ dt\ f_{x_j}\ (x,\ t)\ .$$

This relation can be proved more rigorously and is in fact valid for any region A such that

the integral over the region A is mathematically well-defined.

If we now let A be an arbitrarily small region about some point (x, t) and assume

$f_{x_j}(x,\ t)$ is continuous and therefore essentially a constant over A, we have

$$\lim_{A\ \to\ 0} \frac{E\{\Delta t_j\}}{A} = \lim_{A\ \to\ 0} \frac{1}{A}\ \iint\limits_{A} dx\ dt\ f_{x_j}\ (x,\ t) = f_{x_j}(x,\ t)\ .$$

The terms of (2.9) and (2.3) are all equal.

Since in practice one usually infers the values of k(x,t) or q(x,t) from a single

observation, the usual measurement of k(x,t) as the actual number of cars in an interval of

highway about some position x at time t is essentially equivalent to (2.8) when A is

chosen to be an arbitrarily narrow vertical slit of the (x,t) plane, an A_k of figure II 4.

Similarly the usual direct observation of q(x,t) is equivalent to (2.8) when A is chosen to

be an arbitrarily narrow horizontal slit of the (x,t) plane. It is not obvious what shape or size

of A leads to the most accurate estimates of k(x,t) or q(x,t) for a single observation

particularly if k(x,t) and q(x,t) are not constant functions of x and t but (2.8) at least gives

one the option of choosing whatever shapes of A might seem most suitable in any situation.
particularly if k(x,t) and q(x,t) are not constant functions of x and t but (2.8) at least gives

one the option of choosing whatever shapes of A might seem most suitable in any situation.

As yet very little has been done to determine "best" statistical estimates of parameters

occurring in traffic models. Actually most models that exist now do not describe the statistical

aspects of traffic flow in sufficient detail as to admit the possibility of one even posing such

problems.

3. <u>Stochastic description</u>. So far we have avoided any detailed discussion of what sort of arrival pattern is reasonable in this problem of traffic flow on a homogeneous section of highway although in the definitions of macroscopic quantities it was necessary to have some fairly well behaved probability structure in order that these things have any meaning at all.

There is inherent in almost any traffic situation some uncertainty. One cannot predict exactly when or how many cars will enter a highway, for example. In principle, probability distributions can be determined experimentally by repeated observation but sometimes one can deduce certain features of them from plausible assumptions about uncertainties in human behavior.

One important human limitation, for example, is that people can not judge time with split-second precision, in fact most people cannot predict when they will depart on a journey to an accuracy better than one minute. If there is a probability density for a given driver's departure time, it should therefore be nearly constant over small times of the order of a minute perhaps. Furthermore the uncertainty of departure time for one driver should be statistically independent of the uncertainty for another driver.

To exploit this suppose that the section of highway under discussion is fed by a very large parking lot representing an idealized source of cars from which drivers start their journey. There is to be no interference between the cars in leaving the lot and there are to be large parking lot representing an idealized source of cars from which drivers start their journey. There is to be no interference between the cars in leaving the lot and there are to be no traffic signals or other external influences that would cause statistical correlations in the departure times of the cars. We will therefore assume that the departure times i.e., the times at which drivers enter the highway at $x = 0$ are subject to the statistical uncertainties described above.

25

Consider a short (perhaps a minute) interval of time between some arbitrarily selected time t and time $t+\tau$. Any j^{th} driver has a certain probability density associated with his random departure time and by hypothesis this density is essentially constant over the time $(t, t+\tau)$. If the density at time t is $\alpha_j(t)$ then the probability that the jth driver leaves in this time interval is approximately $\alpha_j(t)\tau$. It is to be implied also that this τ is sufficiently small that $\alpha_j(t)\tau$ is small compared with 1 for all j and t.

This $\alpha_j(t)$ is closely related to the probability density for the position of the jth car $x_j(t)$ at time t, the $f_{x_j}(x, t)$ of the last section. If the jth car crosses $x = 0$ in the time interval $(t, t + \tau)$ and travels with velocity v, it is certain to lie in the spatial interval $(0, v\tau)$ at time $t+\tau$. Thus $\alpha_j(t) = v\, f_{x_j}(0, t)$.

Suppose we now choose t as a time when some car has entered the highway and we ask what is the probability, $P_o(\tau, t)$, that no cars enter the highway during the time interval $(t, t + \tau)$? Since the departure times of different cars are assumed to be statistically independent,

$$P_o(\tau, t) = \prod_{j=1}^{M} [1 - \alpha_j(t)\tau],$$

if there are M cars still in the parking lot at time t and we label them with the index j. If

if there are M cars still in the parking lot at time t and we label them with the index j. If $\alpha_j(t)\tau$ is sufficiently small we can write $1-\alpha_j(t)\tau$ as $\exp[-\alpha_j(t)\tau]$ and obtain

$$P_o(\tau,t) \sim exp[-\alpha(t)\tau] \qquad (3.1)$$

with

26

$$\alpha(t) = \sum_{j=1}^{M} \alpha_j(t) \quad . \tag{3.2}$$

In this formula, the $\alpha_j(t)$ and consequently also the $\alpha(t)$ were to be nearly constant over the time interval t to $t+\tau$. Equation (3.1) admits the possibility of slow variations of $P_o(\tau,t)$ with t caused for example by hourly variations in demand but it is the more rapid exponential dependence upon τ which is of primary concern here. The quantity $1-P_o(\tau,t)$ considered as a function of τ represents the distribution function at time t for the random time interval between departures, the probability that this interdeparture time has a value less than τ.

To derive (3.1) as an exact limit distribution one would imagine that the $\alpha_j(t)$ depended upon M in such a way that for $M \to \infty$ we have $\alpha_j(t) = \alpha_j(t, M) \to 0$ for all j and t, while $\alpha(t, M) \to \alpha(t)$. In practice we do not have an arbitrarily large M and it is not natural to think of the $\alpha_j(t)$ being dependent upon the size of the parking lot. We can, however, treat (3.1) as a plausible conjecture which is subject to experimental confirmation. We can also define roughly the sort of experimental conditions under which one might expect (3.1) to be a reasonable approximation.

The conditions under which (3.1) is plausible are: (1) $1/\alpha(t)$, being a measure of the average time between departures, should be appreciably larger than about 3 seconds, three The conditions under which (3.1) is plausible are: (1) $1/\alpha(t)$, being a measure of the average time between departures, should be appreciably larger than about 3 seconds, three seconds being a typical time headway for congested traffic where the statistical independence assumptions are not expected to be valid, (2) τ must be small enough (a few minutes perhaps) so that we cannot single out any particular driver as having a significant probability of departing between time t and $t+\tau$; and (3) $\alpha(t)$ is nearly constant over a time τ. The last

condition is the least critical because on the one hand it is usually true (typically traffic demands or average headways should be nearly constant over about 10 to 20 minute intervals) and on the other hand if it is not true we can modify (3.1) accordingly. In the above derivation the $\alpha_j(t)\tau$ was used as an approximation to $\int_t^{t+\tau} \alpha_j(z)dz$. Similarly (3.1) is used as an approximation for a more general expression

$$P_o(\tau, t) \sim \exp\left[-\int_t^{t+\tau} \alpha(z) \ dz\right] \tag{3.3}$$

where the $\alpha(z)$ is still as given in (3.2).

We can now go one step further and ask: what is the probability that j cars enter the highway between time t and $t+\tau$. To simplify notation we will delete for now the dependence of $P_o(\tau,t)$, $\alpha(t)$, etc. upon t and take $P_o(\tau,t)=P_o(\tau)=e^{-\alpha\tau}$.

To evaluate the probability $P_1(\tau)$ that one car enters between time t and $t+\tau$, we note first that for any z, $0<z<\tau$, the probability that no car enters between t and $t+z$, that one enters between $t + z$ and $t + z + dz$ is $-P_o(z + dz) + P_o(z) = e^{-\alpha z}\alpha dz$.

The probability that there is one car in the above interval dz and no others in $(t, t + \tau)$ is the above probability multiplied by the conditional probability that there is also none in the interval $(t + z + dz, t + \tau)$ given that there is none in $(t, t + z)$ and one in the interval dz. If is the above probability multiplied by the conditional probability that there is also none in the interval $(t + z + dz, t + \tau)$ given that there is none in $(t, t + z)$ and one in the interval dz. If M is very large, the fact that one car has already been removed from the parking lot does not significantly alter the probability of no entrances in the final interval from $t + z + dz$ to $t + \tau$. We therefore conclude that the probability of one car in dz and no others in $(t, t + \tau)$ is

28

$$e^{-\alpha z}\alpha dz e^{-\alpha(\tau - z - dz)} \sim e^{-\alpha \tau}\alpha dz \quad .$$

and

$$P_1(\tau) = \int_o^\tau dz \alpha e^{-\alpha \tau} = \alpha \tau e^{-\alpha \tau} \quad .$$

Similarly one can argue that the probability of one car in (t, t+z), a second car in (t+z, t+z+dz) and no others is approximately

$$\alpha z e^{-\alpha z}\alpha dz e^{-\alpha(\tau-z)} = \alpha^2 z dz e^{-\alpha \tau}$$

and so the probability of two cars in (t, t+τ) is

$$P_2(\tau) = \int_o^\tau dz e^{-\alpha \tau}\alpha^2 z = \frac{1}{2}(\alpha \tau)^2 e^{-\alpha \tau} \quad .$$

More generally the probability of j cars in (t,t+τ) is

$$P_j(\tau) = \frac{(\alpha \tau)^j}{j!} e^{-\alpha \tau} \quad . \tag{3.4}$$

This is the famous Poisson distribution for the number of events j. It arises in a wide variety of applications. Typically the Poisson distribution arises in any situation where one asks for the probability that j events will occur (an event here is the entrance of some specified car onto the highway) when there is a very large number of independent events that could occur (here there are many cars in the lot that could enter the highway) but the specified car onto the highway) when there is a very large number of independent events that could occur (here there are many cars in the lot that could enter the highway) but the probability that any specified event occurs is very small (there is a small probability that any specified car will enter the highway during the time τ).

The fact that the Poisson distribution like the exponential distribution of (3.1) can be derived as an exact limit distribution is again somewhat academic because the appropriate

29

mathematical limit cannot be physically realized, but one can describe roughly the conditions under which (3.4) is plausible. One certainly would not use (3.4) for example to calculate $P_j(\tau)$ for j so large that j cars entering the highway during a time interval would necessarily imply some congestion.

If α is a slowly varying function of t, the corresponding form for $P_j(\tau,t)$ is obtained by the substitution of $\alpha(t)$ for α in (3.4). If α varies more rapidly (3.4) can be further generalized to the form

$$P_j(\tau,\ t)\ =\ [\int_t^{t+\tau} \alpha(z)dz]^j\ \exp[-\int_t^{t+\tau} \alpha(z)dz]\ /j!\ . \tag{3.5}$$

Since we have already defined q(0,t) as the expected number of cars that enter the highway per unit time we can now identify the $\alpha(t)$ in the above formulas as the observed flow since

$$q(0,\ t)\ =\ \lim_{\tau \to 0} \frac{1}{\tau} \sum_j j\ P_j(\tau,\ t)\ =\ \alpha(t)\ . \tag{3.6}$$

For the model considered here in which all cars travel at the same constant velocity v, the position of any car at time t is uniquely determined by the time at which it enters the highway at x = 0. A specification of the probability distribution for all entrance times, such as the ones described above, therefore implies a specification of the probability distributions for highway at x = 0. A specification of the probability distribution for all entrance times, such as the ones described above, therefore implies a specification of the probability distributions for any events whatsoever that may occur on the highway (insofar as the event in question is determined by the positions of cars at various times).

If j cars enter the highway at $x = 0$ between time t and t+τ, the same j cars must cross a position x during the time interval t+ x/v and $t + x/v +\tau$. Thus if $P_j(\tau,t,x)$ represents the probability of j cars crossing x during a time t to t + τ

$$P_j(\tau,\ t,\ x) = P_j(\tau,\ t - x/v,\ 0) \tag{3.7}$$

identically for all j,τ, t and x. It follows also that

$$q(x,\ t) = q(0,\ t-x/v) \tag{3.8}$$

and
$$k(x,\ t) = k(0,\ t - x/v).$$

If the traffic entering the highway has a Poisson distribution in time at $x = 0$, it must also have the same Poisson distribution of crossing the position x at a time x/v later.

The spatial distribution of cars at a fixed time will also be Poisson distributed because there can be j cars in an interval of highway between x and x+ξ at time t if and only if these j cars entered at $x = 0$ during the time interval t-x/v and t-(x+ξ)/v. If the latter has the Poisson distribution (3.4) with parameter qτ for a time interval τ, the probability of j cars in an interval of highway of width ξ is

$$P_j(\xi/v)\ =\ \frac{(k\xi)^j}{j!}\ e^{-k\xi}$$

a Poisson distribution with parameter qξ/v = kξ i.e., the Poisson distribution which has a mean spatial density of cars k. If the q and k vary with x and t one must of course use a Poisson distribution with parameter qξ/v = kξ i.e., the Poisson distribution which has a mean spatial density of cars k. If the q and k vary with x and t one must of course use the appropriate values of q and k as given by (2.18).

4. Poisson distributions and Poisson processes. The Poisson distribution arises in a wide variety of traffic situations besides the one mentioned above. For future reference, we review

31

here a few of its properties and origins. For a more complete discussion of some of the following, see the book by Feller [2] or the review by Gerlough [3].

In most text books on probability theory, the Poisson distribution is derived as a limit of a binomial distribution. Suppose an experiment results with probability p in the event E, sometimes called "success" in the context of Bernoulli trials but in the present context the event that a driver enters the highway during the time (t, t + τ). If we make M statistically independent trials of the <u>same</u> experiment, the probability of j occurrences of E in the M trials is

$$P_j = \binom{M}{j} p^j (1 - p)^{M-j} \tag{4.1}$$

where

$$\binom{M}{j} = \frac{M!}{j!(M - j)!}$$

If we keep λ = pM fixed while p → 0 and M → ∞ , then

$$p_j \rightarrow \frac{\lambda^j e^{-\lambda}}{j!} . \tag{4.2}$$

This method of derivation of the Poisson distribution is certainly simpler than the one described in the last section, but it is not quite as general. We could use this simple

This method of derivation of the Poisson distribution is certainly simpler than the one described in the last section, but it is not quite as general. We could use this simple derivation to show the following: If the parking lot contains M cars and each has a probability $\alpha_j(t)dt$,

$$\alpha_j(t) = \alpha(t)/M$$

of entering the highway between t and $t + dt$, independent of the entrance times of any other cars, then the probability of its entering between t and $t + \tau$ is

$$p = \int_t^{t+\tau} \alpha_j(t)dt = \frac{1}{M} \int_t^{t+\tau} \alpha(t)dt$$

and $P_j(\tau, t)$ is as given by (3.5). To derive (3.5) in the last section, we did not, however, need to assume that the $\alpha_j(t)$ were all equal. Different drivers could have different distributions for entrance times. The assumption that one driver's entrance time is statistically independent of any other is, however, basic to both derivations.

In deriving the Poisson distribution, we have already made use of the fact that the number of cars entering during one interval of time $(t, t+ \tau_1)$ is (for $M \to \infty$ and $\alpha_j(t) \to$ 0) statistically independent of the number entering during any non-overlapping interval $(t + \tau_1, t + \tau_1 + \tau_2)$ say. It follows that the probability, $P(j_1, j_2, ..., j_k)$, that j_1 cars enter during $(t, t + \tau_1)$, j_2 in $(t + \tau_1, t + \tau_1 + \tau_2)$, ... , and j_k in $(t + \tau_1 + \cdots \tau_{k-1}, t + \tau_1 + \cdots \tau_k)$ is

$$P(j_1, j_2, ... , j_k) = \frac{\lambda_1^{j_1} e^{-\lambda_1}}{j_1!} \times \frac{\lambda_2^{j_2} e^{-\lambda_2}}{j_2!} \times ... \times \frac{\lambda_k^{j_k} e^{-\lambda_k}}{j_k!} \qquad (4.3)$$

where

$$\lambda_1 = \int_t^{t+\tau_1} \alpha(t)dt \ , \ \lambda_2 = \int_{t+\tau_1}^{t+\tau_1+\tau_2} \alpha(t)dt \ , \ ... \quad .$$

$$\lambda_1 = \int_t^{t+\tau_1} \alpha(t)dt \ , \ \lambda_2 = \int_{t+\tau_1}^{t+\tau_1+\tau_2} \alpha(t)dt \ , \ ... \quad .$$

This is the multiple-Poisson distribution which in most textbooks is derived as a limit of the multinomial distribution in the special case of repeated trials of the same experiment. From this we can define a Poisson process.

33

In the general theory of random processes [4], a random process is a family of random variables X(t), parametrically dependent upon some parameter t which takes values in some arbitrary space. The parameter space will usually be the real numbers, $-\infty < t < +\infty$, having the physical interpretation of time. For any values of t; t_1, t_2, ..., t_n, $X(t_1)$, $X(t_2)$, ..., $X(t_n)$ are random variables. A point process is a special case of a random process for which X(t) assumes only the values 0 or 1, 1 if an event happens at time t, 0 otherwise. In the present context an event is the entrance of a car onto the highway, and the point process the process of entrances. In describing the probability structure of a point process, it is sometimes more convenient to specify the joint probabilities for the times t_1, t_2, ... at which events occur, rather than the joint probabilities associated with the X(t) at various arbitrary times.

The probability distributions of entrance times were originally defined here by specification of the probability densities $\alpha_j(t)$ for the entrance time of each jth car, along with the assumption that these entrance times were statistically independent. This description not only defines the probabilities for the times at which cars enter the highway (the process of entrance times) but also which car enters at any time. Equation (4. 3) represents first of all only a limit distribution or an approximate property of the entrance times but secondly does not include a description of which cars enter during any time intervals. A process which satisfies (4. 3) for all values of t, τ_1, τ_2, $\cdot \cdot \cdot$ is called a Poisson process with rate $\alpha(t)$.
description of which cars enter during any time intervals. A process which satisfies (4. 3) for all values of t, τ_1, τ_2, $\cdot \cdot \cdot$ is called a Poisson process with rate $\alpha(t)$.

Most elementary texts on probability theory treat only Poisson processes for which $\alpha(t) = \alpha$, a constant independent of t. The term "Poisson process" is also used frequently to imply that $\alpha(t)$ is constant. We shall use the term here usually in this more restricted sense, but in cases where there may be some confusion we shall use the more specific expression

34

"homogeneous Poisson process" to refer to this special case. The non-homogeneous processes will arise in some discussions of traffic flow. It is obvious that they are relevant to the study of flows in which q(x, t) varies with time because of rush hour traffic, etc.

Homogeneous Poisson processes have certain special stochastic properties not shared by non-homogeneous processes. For a homogeneous Poisson process, the probability of no events during $(t, t + \tau)$,

$$P_0(\tau, t) = e^{-\alpha\tau} , \tag{4.4}$$

is independent of t. In particular, this is true even if t is the time at which an event has occurred. If in this case, we let T be the time of the first event after t, given that there is one at t, then

$$P_0(\tau, t) = P\{T > \tau\} = 1 - F_T(\tau)$$

$$F_T(\tau) = 1 - e^{-\alpha\tau}.$$

Thus F_T is the distribution function for the time between the event at t and the next event.

If we let T_1, T_2, $\cdots$ be the time intervals between successive events, it follows also that T_j has a distribution function

$$P\{T_j > \tau\} = e^{-\alpha\tau} \tag{4.5}$$

independent of the values of any other T_k, $k \neq j$. The set of random variables $\{T_j\}$ are mutually independent. A homogeneous Poisson process is also <u>uniquely</u> defined by saying independent of the values of any other T_k, $k \neq j$. The set of random variables $\{T_j\}$ are

mutually independent. A homogeneous Poisson process is also <u>uniquely</u> defined by saying

that the times T_j are independent random variables with a distribution function (4.5), and

that the (marginal) probability of finding some event during a time interval $(t, t + dt)$ is αdt.

The process is also uniquely defined if we say that there is a probability $\alpha\,dt$ of an event

35

during $(t, t + dt)$ for any t, independent of the times of any other events outside this interval.

The latter interpretation has an obvious generalization to non-homogeneous processes. A non-homogeneous Poisson process is one for which there is probability $\alpha(t)dt$ for an event occurring during $(t, t + dt)$ for any t, independent of the times of any other events outside this interval. The former interpretation of a Poisson process in terms of the times T_j between events, does not however have a simple generalization. The times T_j for a non-homogeneous Poisson process are <u>not</u> statistically independent.

A homogeneous Poisson process is also a special case of another well-known type of point process known as a <u>renewal process</u>. A renewal process is one for which the times T_j between events are independent identically distributed random variables with some general distribution function, not necessarily exponential. As a special case of a renewal process, the most important property of the exponential distribution for the T_j is the following. If we know that $T_j > a$; i.e., a time a has already elapsed since the last event occurred, then

$$P\{T_j - a > t | T_j > a\} = \frac{P\{T_j > t + a\}}{P\{T_j > a\}} = \frac{e^{-\alpha(t + a)}}{e^{-\alpha a}} = e^{-\alpha t} . \qquad (4.6)$$

Thus the distribution of time until the next event, T_j - a, given that there has been none for a time a, is also $e^{-\alpha t}$, independent of a . The time until the next event is independent of Thus the distribution of time until the next event, T_j - a, given that there has been none for a time a, is also $e^{-\alpha t}$, independent of a . The time until the next event is independent of how long it has been since the last event.

A non-homogeneous Poisson process is not in general a renewal process. The non-homogeneous Poisson process is a different type of generalization of a homogeneous Poisson process. It is, in fact, a simpler type of process to treat than the renewal process. It is always

36

possible to map a non-homogenous process into a homogeneous Poisson process. If we define a new time t^* by

$$t^*(t) = \int_0^t \alpha(z)dz \tag{4.7}$$

then $t^*(t)$ is a monotone increasing function of t if $\alpha(t) > 0$. There is a one-to-one correspondence between t and t^*. As a function of t^*, however, the Poisson process is homogeneous.

If real traffic were known to conform with a homogeneous Poisson process of unknown flow q (i. e., α) the estimation of q from a single observation of the number of cars which enter the highway during a time τ is a classic problem in parameter estimation. It is discussed in most books on mathematical statistics as the problem of estimation of the mean for a Poisson distribution.

If j cars are observed in a time τ, the usual estimate of q is

$$\hat{q} = j/\tau$$

The $\hat{q}$ is a random variable. If the experiment were repeated under identical conditions (same q), $\hat{q}$ would have a probability distribution of possible values determined by the fact the number of cars observed in a time τ has a Poisson distribution with parameter $q\tau$.

number of cars observed in a time τ has a Poisson distribution with parameter $q\tau$.

$$P\{\hat{q} = k/\tau\} = \frac{(q\tau)^k}{k!} e^{-(q\tau)} , \quad k = 0, 1, \ldots .$$

The estimate of $\hat{q}$ is unbiased: i.e.,

$$E\{\hat{q}\} = q .$$

The standard deviation of $\hat{q}$ is

$$\left[E\{(\hat{q} - q)^2\}\right]^{1/2} = \left(\frac{q}{\tau}\right)^{1/2} \quad .$$

The standard deviation divided by the mean is a measure of fractional deviation one would expect between the estimate and the actual value of q. This ratio is $(q\tau)^{-1/2} \sim j^{-1/2}$.

Even if the traffic is not Poisson distributed, it is generally a "good rule of thumb" to say that if one is using some count of events as an estimate of some parameter, the fractional error in the estimate will be of the order of magnitude of (number of events)$^{-1/2}$.

Unfortunately in most traffic counts, the value of τ and consequently $q\tau$ is limited by the desire to have q nearly constant over the time τ. Even under the most favorable conditions, one cannot usually expect flows to remain constant for more than 15 minutes or perhaps an hour. On the other hand, if the hypothesis of a Poisson distribution is to apply, we should have average time headways, 1/q , of the order of 10 seconds or more (otherwise there will be some interactions between cars). Typically, the best one can do is $q\tau$ in the range of 10^2 to 10^3 , $(q\tau)^{-1/2}$ is then of the order 10^{-1} or 10%.

It is rather difficult to devise a meaningful statistical test for the hypothesis that a traffic stream is Poisson distributed. Cars have finite size and interact strongly at close headways. The Poisson process is certainly not an exact representation of traffic, and one can devise statistical tests (sensitive to short headways) which would almost always lead one to reject the

hypothesis that the traffic is a Poisson process. The conventional theory of statistical testing is not ideally suited to the rather poorly defined question: is a hypothesis (known to be false) approximately correct? It would be possible to devise some suitable tests, but the logic would become rather involved.

5. <u>Velocity distributions</u>. The model just discussed in which all cars travel at the same velocity gives only a rather crude description of the way cars actually behave. It is one which, however, is very useful in estimation of the proper synchronization of traffic signals or the description of flows on networks in which the network aspect of the problem is of main concern and further refinements of the detailed behavior of traffic on single highways adds little insight but considerable complication.

The most obvious fault of this model is that it does not permit passing nor the natural spreading of a platoon due to variations in the speed of the cars. To incorporate this into the model we will simply assign to each jth driver his own free speed v_j which may be different for different drivers. Since different single car trajectories will have different slope, some trajectories will intersect (the cars will pass each other). Depending upon the nature of the highway a passing may or may not produce a significant delay to the driver who wishes to pass. We will assume for the present that there is no interaction between cars. This will be valid if either the passings can be executed with negligible change in velocity or if passings are so rare that one can neglect their consequences (the density is very low).

Again the detailed dynamics is simple. The trajectories are given by

$$x_j(t) = x_j(0) + v_j t \tag{5.1}$$

but certain modifications must now be made in the definitions of the macroscopic variables

$$x_j(t) = x_j(0) + v_j t \tag{5.1}$$

but certain modifications must now be made in the definitions of the macroscopic variables and the stochastic description of the system.

A density k and a flow q can still be defined as the number of cars per unit length of highway and the number passing a fixed point per unit time respectively. Since there are

39

many different velocities, however, the simple relation q = kv no longer has an obvious meaning.

Suppose we have only finitely many possible velocities $v^{(1)}$, $v^{(2)}$, $\cdots$, $v^{(m)}$ and for each velocity $v^{(i)}$ there is a density $k^{(i)}$ of cars with this velocity and a flow $q^{(i)}$. For these cars alone, all with the same velocity, it is true that

$$q^{(i)} = v^{(i)} k^{(i)} \quad or \quad k^{(i)} = q^{(i)}/v^{(i)}. \tag{5.2}$$

Since cars do not interfere with one another, the complete set of trajectories is simply a superposition of trajectories corresponding to the different velocities. Therefore

$$q = \sum_i q^{(i)} \quad and \quad k = \sum_i k^{(i)} \tag{5.3}$$

which with (5.2) gives

$$q = k \sum_i k^{(i)} v^{(i)}/k \quad or \quad k = q \sum_i q^{(i)}/(v^{(i)} q). \tag{5.4}$$

To define an empirical probability distribution of velocities one might quite naturally proceed in either of two ways. If we take an aerial photograph and count the fraction of cars with velocity $v^{(i)}$, we might consider

$$\frac{k^{(i)}}{k} = \frac{q^{(i)}/v^{(i)}}{\sum_j q^{(j)}/v^{(j)}} \tag{5.5}$$

as the velocity distribution. But if we stood at a fixed point on the highway and counted the fraction of cars with velocity $v^{(i)}$ that passed this point, we might consider
as the velocity distribution. But if we stood at a fixed point on the highway and counted the fraction of cars with velocity $v^{(i)}$ that passed this point, we might consider

$$\frac{q^{(i)}}{q} = \frac{v^{(i)} k^{(i)}}{\sum_j v^{(j)} k^{(j)}} \tag{5.6}$$

as the velocity distribution. These two distributions are clearly not the same. The extreme example is when some $v^{(i)}$ is zero. These cars would be observed from an aerial photograph to

40

have density $k^{(i)}$ but since they have zero rate of passing a fixed point, a stationary observer would see none of them.

If we define a spatial mean velocity as

$$E_s\{v\} = \sum_{(i)} v^{(i)} k^{(i)} / k \qquad (5.7)$$

then (5.4) gives

$$q = k \, E_s\{v\}, \qquad (5.8)$$

the analogue of the equation $q = kv$ in the case of equal velocities. If one wishes to use the time distribution $q^{(i)}/q$ then

$$k = q \, E_t\{v^{-1}\} \qquad (5.9)$$

with

$$E_t\{v^{-1}\} = \sum_i (1/v^{(i)}) \; (q^{(i)}/q). \qquad (5.10)$$

In terms of the time distribution, one can define a harmonic mean velocity as

$$E_h\{v\} \;=\; \frac{1}{E_t\{v^{-1}\}}$$

which is the same as $E_s\{v\}$.

If one uses the technique of defining q and k from segments of trajectories contained in an area A of the x-t plane as in figure 3, one again obtains (2.8), which does not contain the velocities of the cars. It follows then that the spatial mean velocity is contained in an area A of the x-t plane as in figure 3, one again obtains (2.8), which does not contain the velocities of the cars. It follows then that the spatial mean velocity is

$$E_s(v) \;=\; \frac{q}{k} \;=\; \frac{\sum_j \Delta x_j}{\sum_j \Delta t_j} \quad .$$

This along with many other similar type relations are contained in [1]. The recognition that the velocity distribution of cars observed per unit time at a fixed position is different from

that observed at a fixed time per unit length of highway seems to have originated with Wardrop [5], although analogous situations were known in the kinetic theory of gases.

In the previous discussion of cars leaving a parking lot, we were only concerned with how many cars left during any time interval since all cars had identical behavior. We must now construct a model giving not only departure times but also speeds. In the absence of any specific knowledge indicating that high velocity cars are more likely to depart at one time than another or more likely to leave after a fast car than a slow car, etc., the only reasonable postulate we can make is the following. The parking lot contains so many cars that we can define a velocity distribution

$$F_v(v) = \textit{fraction of cars in the lot with velocity less than } v \qquad (5.11)$$

which can be approximated by some continuous distribution function. Any car which is due to depart from the parking lot is now sampled at random from this distribution independent of any previous departures.

There are a number of equivalent descriptions of the complete stochastic structure of the departing cars, for example:

1.   The time intervals between successive departures are independent identically distributed random variables with an exponential distribution

   $$P\{ \textit{ time interval } \leq \tau \} = 1 - e^{-q\tau}.$$
   ed random variables with an exponential distribution

   $$P\{ \textit{ time interval } \leq \tau \} = 1 - e^{-q\tau}.$$

The velocities of successive cars are independent random variables with a distribution function $F_V(v)$ independent of the departure times. (For a discrete velocity distribution the distribution function $F_V(v)$ corresponds to the frequency distribution $q^{(i)}/q$ of (5.6).)

42

2. The probability of j departures in a time interval of duration τ with the first car having a velocity less than or equal to v_1, the second car a velocity less than or equal to v_2, etc., is

$$\frac{(q\tau)^j}{j!} \, e^{-q\tau} \, F_v(v_1) \, F_v(v_2) \, .. \, F_v(v_j) \, . \tag{5.12}$$

This event is also statistically independent of the number or velocities of any departures during times outside the time interval τ.

3. The cars with velocity less than any given value v also define a Poisson process. The probability that j cars with velocity less than or equal to v depart during a time τ is

$$\frac{[q\tau F_v(v)]^j}{j!} \, \exp \, [-q\tau F_v(v)] \, . \tag{5.13}$$

This event is statistically independent of the number or the velocities of the cars having a velocity larger than v that depart during the time τ and is independent of the number or velocities of departures during times outside the interval τ.

We will leave it to the reader to prove the equivalence of the above or to formulate other descriptions.

One of the main objections to the above postulates about the stochastic properties is that it is very difficult to test experimentally and it is also difficult to justify on theoretical
grounds alone. Before we divided the population of drivers according to their velocities, we had some hope that the flow would remain constant long enough (half an hour perhaps) so that we would obtain large enough samples to test the Poisson nature of the traffic. Now with the same amount of data we want to separate drivers into various ranges of velocities and

then test not only that the number of cars in each velocity range has a Poisson distribution but that they are also statistically independent. This is more than one can do in a satisfactory way.

If we try to take larger times we must expect not only that the flow may change but also the distribution of velocities. Certainly the five o'clock rush of commuters is likely to have a different distribution of desired speeds than the midday shoppers and commercial traffic.

As regards the justification from postulates about uncertainties in departure times of individuals we have previously argued that uncertainties in the departure time of an individual should be a few minutes perhaps, during which time many drivers were likely to depart. If, however, we take a subpopulation of drivers in some narrow velocity range, it is no longer plausible that many such drivers are likely to depart during a time of the order of the uncertainty in the departure time of a single driver. Indeed if we take an arbitrarily small velocity range, the times between departures of cars with velocities in this range can be made as large as we please and any fluctuations in individual departure times will be negligible by comparison. The whole picture begins to look more like a deterministic one than a stochastic one.

one.

6. <u>Time dependent flow</u>. One important consequence of distributed velocities is that a concentration of cars on the highway tends to disperse. To investigate this and related problems we consider a joint density of both position and velocity. Let

$$\rho_s(x, v, t)dx\, dv = E\ \{number\ of\ cars\ between\ x\ and \tag{6.1}$$

x + dx with velocity between v and v + dv at time t}.

This is the two dimensional density in the space of position x and velocity v at time t.

Correspondingly we define a joint density of flow and velocity

$$\rho_t(x, v, t)dt\, dv = E\{ \text{ number of cars crossing the position } x \text{ between time t and}$$

$$t + dt \text{ with velocity between v and } v+dv\} . \qquad (6.2)$$

The $\rho_s(x,v,t)$ is the analogue for a continuous velocity distribution of the $k^{(i)}$ and $\rho_t(x,v,t)$

is the analogue of the $q^{(i)}$ of equations (5.2) and (5.3) except that we are now admitting the

possibility that these quantities may vary with position and time. If these functions are

continuous in x and t, they are related by the equations

$$\rho_t(x, v, t) \;\; = \;\; v\, \rho_s\, (x, v, t) \qquad (6.3)$$

analogous to the equation $q^{(i)} = v^{(i)}k^{(i)}$ of (5.2) .

The time dependence of ρ_s and ρ_t is dictated by the equations of motion for the

individual cars. Any car of velocity v that lies between x and x+dx at time t will lie

between $\dot{x} + v\tau$ and $x + v\tau + dx$ at time $t + \tau$. Thus

$$\rho_s(x, v, t) \;\; = \;\; \rho_s(x + v\tau, v, t + \tau) \text{ for all } \tau, \qquad (6.4)$$

and in particular

$$\rho_s(x,v,t) = \rho_s(0,v,t\text{-}x/v)$$

$$= \rho_s(x\text{-}vt,v,0)$$

$$\rho_s(x,v,t) = \rho_s(0,v,t\text{-}x/v)$$

$$= \rho_s(x\text{-}vt,v,0)$$

gives ρ_s at x,v,t in terms of its values at the origin x = 0 or at some initial time t = 0.

These equations are the generalizations of (3.8) for distributed velocities.

The density of cars k(x,t) is given by

$$k(x, t) = \int_0^\infty dv \, \rho_s(x, v, t) = \int_0^\infty dv \, \rho_s(x - vt, v, 0) \, , \qquad (6.5)$$

and the flow is

$$q(x, t) = \int_0^\infty dv \, \rho_t(x, v, t) = \int_0^\infty dv \, v \, \rho_s(x, v, t) \, . \qquad (6.6)$$

It is important to observe that one cannot evaluate the density at time t from only the density $k(x,0)$ at time 0. One must know the distribution of velocities also. The distribution of velocities, however, is not constant but is given by

$$f_s(x, v, t) = \frac{\rho_s(x, v, t)}{k(x, t)} \qquad (6.7)$$

and

$$f_t(x, v, t) = \frac{\rho_t(x, v, t)}{q(x, t)} \qquad (6.8)$$

respectively for the spatial and time distributions.

The above equations can be used to describe the diffusion of cars due to a distribution of velocities. Suppose, for example, that initially we know that the density of cars is non-zero only at $x = x_o$ but the expected number of cars at x_o is 1. We represent this by

$$\rho_s(x,v,0) = \delta(x-x_o) \, f_s(v)$$

in which $\delta(x-x_o)$ is the Dirac δ-function and $f_s(v)$ is the spatial density of velocities taken from the parking lot population for example i.e.,
in which $\delta(x-x_o)$ is the Dirac δ-function and $f_s(v)$ is the spatial density of velocities taken from the parking lot population for example i.e.,

$$f_s(v) = \frac{v^{-1} \, dF_V(v)/dv}{\int_0^\infty v^{-1} \, dF_V(v)} \cdot \qquad (6.9)$$

The Dirac δ-function is, in the modern mathematical literature, called a "generalized function." It is a mathematical notion by which a concentrated unit mass at a single point x_o

46

is represented as if it had a density. Loosely speaking $\delta(x-x_o)$ is a function that is zero everywhere except at x_o but at x_o is infinite in such a way that the integral of $\delta(x-x_o)$ is one. The key property of $\delta(x-x_o)$ is that for any function $\phi(x)$ continuous at x_o

$$\int_{-\infty}^{+\infty} \delta(x - x_o)\ \phi(x)\ dx = \phi(x_o)$$

The integral of a function $\phi(x)$ with $\delta(x-x_o)$ is thus a linear mapping of the function $\phi(x)$ into itself, the identity operator on (continuous) functions.

The density at any later time, is given by (6.5)

$$k(x,\ t) = \int_0^\infty dv\ \delta(x - vt - x_0)f_s(v)$$

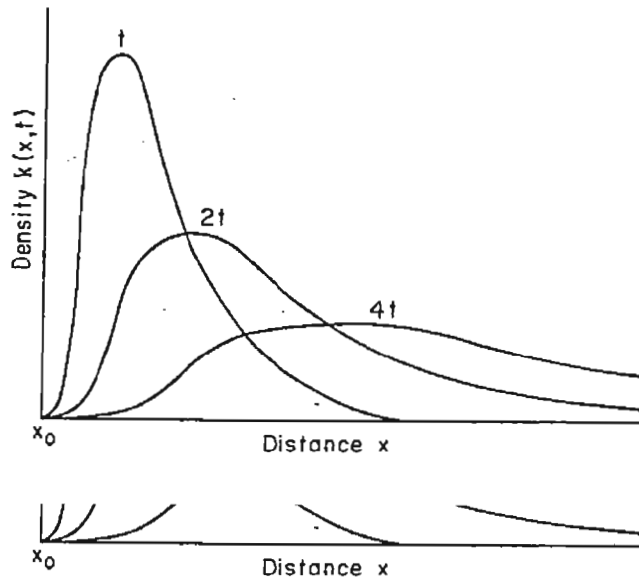$$= \frac{1}{t}\ f_s\!\left(\frac{x - x_0}{t}\right) . \tag{6.10}$$

Fig. II 5
Dispersion of traffic due to distributed velocities

The spatial density is for all times given by the same function f_s except for a rescaling of the spatial coordinates measured from x_o by a factor t and a rescaling of the density itself by a

47

factor t. This is shown in figure (II.5) for some typical shape for $f_s(v)$. Similarly the flow is given by

$$q(x, t) = \left(\frac{x - x_0}{t^2}\right) f_s\left(\frac{x - x_0}{t}\right) .$$ (6.11)

As another interesting example suppose that at $x = 0$ we have a fixed cycle traffic signal fed by a time homogeneous traffic source. The flow leaving the traffic signal will then be a periodic function of time, i.e.,

$$q(0, t) = q(0, t + T)$$ (6.12)

if T is the cycle time. If we choose the velocity distribution independent of t at $x = 0$, then

$$\rho_t(0, v, t) = f_v(v) \, q \, (0, t),$$ (6.13)

with

$$f_v(v) = dF_v(v)/dv,$$

is a product of the time independent probability density for the velocities and the periodic time-dependent flow.

The flow at some $x > 0$ is given by

$$q(x, t) = \int_0^\infty dv \, \rho_t(x, v, t) = \int_0^\infty dv \, \rho_t(0, v, t - x/v)$$
$$= \int_0^\infty dv \, f_v(v) \, q(0, t - x/v) .$$ (6.14)

The output from any cycle of the signal will diffuse and we expect that for sufficiently

$$= \int_0^\infty dv \, f_v(v) \, q(0, t - x/v) .$$ (6.14)

The output from any cycle of the signal will diffuse and we expect that for sufficiently large x the pulses of traffic created by the signal will overlap i.e., fast cars from one cycle will overtake slower cars from the previous cycles. Eventually so many pulses will overlap that the flow should become nearly constant. This effect can be deduced from (6.14).

If we let $u = 1/v$ in (6.14) then

48

$$q(x, t) = \int_0^\infty \frac{du}{u^2} f_v\left(\frac{1}{u}\right) q(0, t - ux) \quad . \tag{6.15}$$

We can decompose this integral into the contributions coming from successive cycles.
Formally

$$q(x, t) = \sum_{j=0}^\infty \int_{jT/x}^{(j+1)T/x} \frac{du}{u^2} f_v\left(\frac{1}{u}\right) q(0, t - ux)$$

in which the argument t-ux of q increases by T in each range of integration $(jT/x \, , \, (j+1)T/x)$.

For $x \to \infty$, the range of integration of each integral goes to zero and if $f(1/u)/u^2$ is continuous for $0 \le u \le \infty$ this factor is essentially constant over the range of integration. Thus

$$q(x, t) \to \sum_{j=0}^\infty \frac{1}{u_j^2} f_v\left(\frac{1}{u_j}\right) \int_{jT/x}^{(j+1)T/x} du \, q(0, t - ux)$$

with

$$u_j = jT/x \quad .$$

The remaining integrals are integrals of a periodic function,

$$x \int_{jT/x}^{(j+1)T/x} du \, q(0, t - ux) = \int_{jT}^{(j+1)T} dy \, q(0, t - y) = \int_0^T dy \, q(0, y) \quad ,$$

which is independent of both t and x. If the traffic signal is fed by a homogeneous stream

which is independent of both t and x. If the traffic signal is fed by a homogeneous stream

with flow q, then we can interpret this integral as the average number of cars leaving per

cycle or qT. Thus

49

$$q(x,\ t) \rightarrow \sum_{j=0}^{\infty} \frac{1}{u_j^{\ 2}}\ f_V\left(\frac{1}{u_j}\right)\frac{qT}{x}\quad .$$

But now for $x \rightarrow \infty$ the u_j become uniformly dense with spacing T/x so this sum becomes a Riemann integral

$$q(x,\ t) \rightarrow q \int_0^{\infty} \frac{du}{u^2}\ f_V\left(\frac{1}{u}\right) = q \int_0^{\infty} dv\ f_V(v) = q\quad . \tag{6.16}$$

The time dependent flow at $x = 0$ thus returns to the constant flow q that exists before the signal as $x \rightarrow \infty$.

This represents only one special case of a much wider class of possible flow patterns $\rho_t(0,v,t)$ which will give rise to a flow q(x,t) that converges to a constant q for $x{\rightarrow}\infty$. Similarly there is a wide class of possible initial spatial distributions $\rho_s(x,v,0)$ which give rise to a density k(x, t) that converges to some constant k for $t \rightarrow \infty$. Note that any flow pattern of the former type can be mapped into a flow pattern of the latter type if we simply inter-change the roles of x and t. Any set of straight line trajectories with x plotted vs t, remains a set of straight line trajectories if we plot t vs x.

For the flow to approach a constant q one must impose certain conditions on $\rho_t(0,v,t)$, however. If there is a discrete component of the velocity distribution, i.e., some single velocity v appears with non-zero probability, than any irregularities in the starting flow however. If there is a discrete component of the velocity distribution, i.e., some single velocity v appears with non-zero probability, than any irregularities in the starting flow pattern of this velocity will propagate with no dispersion and will exist for arbitrarily large x. We have avoided this in the above example by having a continuous density $f_V(v)$. If for very high velocities $v{\rightarrow}\infty$, the $f_V(v)$ does not go to zero fast enough some irregularities in the starting flow may reach x almost instantaneously. In practice this of course cannot happen

because velocities are bounded. Here we avoided this problem by choosing $f(1/u)/u^2$ continuous at $u = 0$. Finally, this dispersion only smooths local variations in flow. One must be able to define some long time average flow q over a time interval $(t,t+T^*)$ for some sufficiently large T.* One must also have some long range uniformity over t of the velocity distributions to avoid "focusing." These properties are satisfied in the above example by the periodic flow and time-independent velocity distribution but they would obviously also be true of a more general class of flows.

7. The Poisson tendency of traffic and reversibility.

In the last section we considered only the densities of position and velocity or expectations for the number of cars. We made use of the fact that these expectations at some position x and time t are related to similar quantities at other values of x and t. To study the probabilities for various events we must, however, consider the evolution of stochastic properties of the system other than just these expectations.

If in the present model we specify the entrance times at $x = 0$, or the positions of the cars at $t = 0$ and the velocities of all cars, the positions and velocities of the cars are uniquely determined for all time. Similarly if we specify the joint probability distributions of all entrance times at $x = 0$ or positions at $t = 0$ and all velocities, the complete probability uniquely determined for all time. Similarly if we specify the joint probability distributions of all entrance times at $x = 0$ or positions at $t = 0$ and all velocities, the complete probability structure of all trajectories is uniquely defined. In particular the joint distribution of positions and velocities of cars at time t and the joint distribution of the times at which cars cross some point x and the velocities are both uniquely determined.

If positions and velocities have a joint probability density and we number the cars in some way, let

$$\rho_s(x_1, x_2, ..; v_1, v_2, ..; t)dx_1dx_2..; dv_1dv_2..$$

be the probability that the jth car has a position in (x_j, x_j+dx_j) and a velocity in (v_j, v_j+dv_j) for all j at time t. The equations of motion then require that (if ρ_s is continuous in all its arguments)

$$\rho_s(x_1 + v_1\tau, x_2 + v_2\tau, ..; v_1, v_2, ..; t + \tau) \qquad (7.1)$$

$$= \rho_s(x_1, x_2, ..; v_1, v_2 ..; t)$$

for all τ since, if car j is at x_j at time t, it is certain to be at $x_j + v_j\tau$ at time t+τ, Similarly if

$$\rho_t(x; v_1, v_2, ..; t_1, t_2, ..) dv_1dv_2 .. dt_1dt_2..$$

is the probability that car j has a velocity in (v_j, v_j+dv_j) and crosses a position x during the time interval (t_j, t_j+dt_j), then

$$\rho_t(x+ \xi; v_1, v_2, ..; t_1+ \xi/v_1, t_1 + \xi/v_2, ..) \qquad (7.2)$$

$$= \rho_t(x; v_1, v_2, ..; t_1, t_2, ..) \text{ for all } \xi.$$

The Poisson distribution with statistically independent velocities described earlier as a reasonable model for the entrance time distribution plays a unique role in the present theory as does the Poisson distribution of position of cars with independent velocities. We shall see later that a wide class of possible distributions for velocities and starting times at x = 0
as does the Poisson distribution of position of cars with independent velocities. We shall see later that a wide class of possible distributions for velocities and starting times at x = 0
should for $x \rightarrow \infty$ converge to this distribution in some sense. Correspondingly a wide class of possible initial spatial distributions of cars will under the dynamical motion of the system approach a Poisson spatial distribution with independent velocities for t→∞.

To prove this and investigate the nature and rate of convergence is one of the key problems in the logical foundations of this theory. Although we have argued that this Poisson distribution is a reasonable model for entrance times from a highly idealized parking lot, the usefulness of this distribution in practical applications would be severely limited if one could not also show that it had some stability; that initial deviations from this distribution or any of a wide variety of disturbances caused by traffic signals, occasional interactions between cars, etc., would cause only transient effects and the distribution would eventually return to the Poisson distribution.

The problem under consideration is almost identical to one which occurs in statistical mechanics and is the origin of many famous paradoxes and controversies. The controversy in statistical mechanics centers around the question of how one can deduce the irreversibility of thermodynamic systems (the second law of thermodynamics says that the entropy of an isolated system never decreases with time) from the equations of dynamics which are invariant to changing time to the negative of the time i.e., for any motion of a dynamical system there is another one that will exactly reverse it. If one is not very careful in defining the entropy in statistical mechanics one will usually come to the conclusion that the entropy of a system cannot change at all with time.

In physics, as in the present model of traffic flow, we can define a state of the system entropy of a system cannot change at all with time.

In physics, as in the present model of traffic flow, we can define a state of the system by specifying the positions and velocities of all particles or cars. In both cases we have equations of motion which determine the state at time t from the state at time 0 and vice versa i.e., the state at any time uniquely determines not only the future but also the past. If we define entropy as

53

$$S(t) = -\int .. \int dx_1 .. dv_1 .. \rho_s(x_1, .. ; v_1, .. ; t) \log \rho_s(x_1, .. ; v_1, .. ; t) \qquad (7.3)$$

then here, as in physics, the equations of motion as expressed by (7.1) are such that S(t) is independent of t. To prove this, substitute (7.1) with $\tau = -t$ into (7.3):

$$S(t) = -\int .. \int dx_1 .. dv_1 .. \rho_s(x_1 - v_1 t, x_2 - v_2 t, .. ; v_1, v_2, .. ; 0) \times$$

$$\log \rho_s(x_1 - v_1 t_1, x_2 - v_2 t, ..; v_1, v_2, ..; 0) .$$

Now change the variables of integration. Let

$$x_j' = x_j - v_j t \qquad , \quad v_j' = v_j .$$

The Jacobian of this transformation is 1. Equivalently if we integrate with respect to x_j' or x_j, $j = 1,2,..$, for fixed values of the v_j, then integrate with respect to the v_j, the integral over all x_j, $-\infty < x_j < +\infty$, is the same as an integral over all x_j', $-\infty < x_j' < +\infty$, for any fixed v_j. Also $\partial x_j'/\partial x_j = 1$ for all fixed v_j.

Thus

$$S(t) = -\int .. \int dv_1 .. dx_1' .. \rho_s(x_1', x_2', .. ; v_1', v_2', .. ; 0) \times$$

$$\log \rho_s(x_1', x_2', .. ; v_1', v_2', .. ; 0)$$

$$= S(0) . \qquad\qquad for\ all\ t .$$

Entropy has been used to define a measure of the amount of information we have about the system [7]. That the entropy is constant can also be described by the statement that

since there is a one-to-one correspondence between the distribution at time t and time 0, the specification of this distribution at one time gives exactly the same information as at any other time.

Some attempts have been made to show that the Poisson distribution with independent velocities has some aspects of stability [8-14], but this is a rather delicate problem. One should always keep in mind the negative aspect of it, that for this model with deterministic equations of motion there can be one and only one spatial distribution at time 0 that can give exactly the Poisson distribution of cars with independent velocities at any finite time t. Indeed we shall see later, that this Poisson distribution is time invariant and consequently the only one that can reproduce this distribution.

Despite this discouraging note, there are some positive aspects to the problem. First of all, one can show, even for this deterministic model, that certain coarse features of the distributions do change with time and may converge to the corresponding features of the Poisson distribution. This happens because the existence of a probability density at time $t = 0$, although it implies the same for any finite t, does not guarantee that the probability density has a limit for $t \to \infty$. If in describing the behavior for $t \to \infty$, we interchange limiting processes, we might, for example, find that a probability density has no limit for $t \to \infty$ but a distribution function does have a limit and that the limiting distribution function has a density. This limiting distribution function may also define an entropy which is not the same as that for any finite t, i.e., the limit of the entropy for $t \to \infty$ is not the same as the entropy of the limit distribution.
for any finite t, i.e., the limit of the entropy for $t \to \infty$ is not the same as the entropy of the limit distribution.

Whether or not the above type of mathematical procedure is meaningful or not depends upon the physical context in which the mathematical problem was posed. We should first recognize that even if drivers could maintain a velocity which is exactly constant with infinite precision one could not experimentally estimate a probability density without

55

appealing to some smoothness on a scale of measurement having only a finite accuracy. Secondly, and probably much more important, is the fact that drivers cannot maintain a constant velocity exactly and any fluctuations in velocity are likely to give rise to increasing uncertainties in the future position of a car.

In physics, the equations of motion, the laws of classical dynamics, quantum mechanics or whatever, are not customarily considered as approximations to something inherently stochastic, and the reversibility vs. irreversibility paradox has never been completely resolved. Here, however, we have a natural way out of this trap in that the above deterministic equations of motion are only a mathematical idealization to something that is stochastic and inherently irreversible.

Some of the odd effects described here appear even in the simple examples of the last section. If there is only one car in the system, then the densities (6.1) and (6.2) have the interpretation of probability densities for the one car and are, therefore, special cases of (7.1) and (7.2).

For the initial δ-function spatial distribution of (6.9) we know at $t = 0$ exactly where a car is located but know only the probability distribution of velocity. At any later time we know less about where the car is located but if we observe where the car is we will know its velocity exactly. If the car is at x at time t and started at x_o at time 0, it must have velocity know less about where the car is located but if we observe where the car is we will know its velocity exactly. If the car is at x at time t and started at x_o at time 0, it must have velocity $(x-x_o)/t$. Information about position at time 0 is thus transformed into information about velocity at some later time.

If, in this example, we wish to consider the future distribution of cars with the distribution at time $t_o > 0$ taken as a new "initial state", we must not now claim ignorance of

56

the existing correlations between position and velocity by reassigning to each car a velocity at time t_0 taken at random from the distribution of velocities that existed at t = 0. Here lies one of the interesting qualitative differences between diffusion of cars and the diffusion of molecules in the kinetic theory of gases. In the diffusion of gases, molecules undergo collisions and, after a few collisions, a molecule more or less forgets what its velocity may have been originally. Here one can reasonably reassign the molecules new velocities chosen at random from some distribution of velocities. The result of this is that the mean dispersion in distance traveled by a molecule in time t increases proportional to $\sqrt{t}$ whereas that of cars increases proportional to t.

The other example of the traffic signal illustrates some other points. We started at x = 0 with a fairly general periodic flow q(0,t) and a time-independent velocity distribution but at any finite non-zero distance x from the signal both the velocity distribution and the flow will be periodic and non-constant. As x increases the flow q(x,t) becomes smoother but the velocity distribution becomes more irregular. For very large x, after many platoons have overlapped, the velocity distribution becomes highly oscillatory because if one observes cars crossing x at some time t, one will find certain velocities are absent entirely because cars with these velocities would have to have left x = 0 during a red time in order to reach x at the time t. Cars which leave x = 0 during the subsequent green time, however, with only with these velocities would have to have left x = 0 during a red time in order to reach x at the time t. Cars which leave x = 0 during the subsequent green time, however, with only slightly high velocities will appear. The velocity distribution at any finite but large x may have a density but it is one that oscillates with v between zero and some non-zero values. In the limit x → ∞, the velocity distribution has no limiting density but the distribution function

57

does have a limit and the limit distribution function has a derivative. In fact the derivative of the limit distribution function is the probability density for velocities at $x = 0$.

To see that this last statement must be true, suppose we extract from the flow at $x = 0$ only those cars with velocity in some range v to $v + \Delta v$ with $\Delta v > 0$. These cars will themselves define a new periodic flow at $x = 0$ with a time independent velocity distribution in the range Δv. This flow also satisfies the condition that $q(x,t) \to q$, (6.15), where the $q(x,t)$ is now only the flow in this velocity range and the q is now the original time-average flow q multiplied by $F_v(v + \Delta v) - F_v(v)$, the total fraction of cars in this velocity range. The smaller Δv, the longer it takes for platoons from different signal cycles to overlap, and the slower the rate of convergence to the constant flow. If, however, we let $x \to \infty$ and then let $\Delta v \to \infty$ we will conclude that the flow in each velocity range Δv becomes constant, the total flow is also constant, and the fraction of cars in $(v, v + \Delta v)$ is $F_v(v + \Delta v) - F_v(v)$ independent of t. This in   turn defines for $\Delta v \to 0$ the density $f_v(v)$.

The above arguments can be translated into mathematical theorems but this is perhaps academic. We shall see later that, despite the implications of such theorems, this model with constant velocities does lead to some unrealistic predictions relating to the motion of cars through synchronized traffic signals.

If it is meaningful to discuss the behavior of traffic for $x \to \infty$ even though highways through synchronized traffic signals.

If it is meaningful to discuss the behavior of traffic for $x \to \infty$ even though highways are always of finite length, one could also imagine that one places another traffic signal at x with a fixed synchronization relative to the signal at $x = 0$. Whether or not a car is stopped at x by a red signal phase in some j^{th} cycle is now very sensitive to the velocity of the car. If the velocity were only slightly different, it would pass in a neighboring green phase or in

some different cycle. The delay at the second signal is thus particularly sensitive to any rapid oscillations in the velocity distribution at x = 0 similar to those which we smoothed out by using the distribution function in the above arguments. To discuss such problems of synchronized signals one must use more realistic models with stochastic rather than deterministic equations of motion.

As regards the more general problem of showing when and how a rather general initial distribution of positions and velocities will approach the Poisson distribution with independent velocities, Weiss and Herman [8] and Miller [9] have shown that if, at t = 0, the cars are distributed along the highway so that the spacings $x_j(0)$-$x_{j-1}(0)$ are independent identically distributed random variables and the velocities v_j are also independent identically distributed random variables, independent also of the spacings, then for t → ∞ the spacings between adjacent cars become exponentially distributed. Breiman [10,11] and Thedéen [13] consider the same problem with a much more general initial distribution of cars. Any of these results can, of course, be translated into corresponding behavior of the distribution of velocities and crossing times for x →∞ given these distributions at x = 0.

Here again we will be content to give heuristic arguments and explanations rather than proofs.

Breiman assumes that for t = 0 the positions $x_k(0)$ are random but restricted only by proofs.

Breiman assumes that for t = 0 the positions $x_k(0)$ are random but restricted only by the conditions that (a) with probability one an average density exists in the sense that

$$\lim_{x \to \infty} [number\ of\ cars\ in\ (-x,\ 0)]/x = k$$

(b) the expected number of cars in any finite interval of highway is bounded by some number depending only upon the length of the interval. The conditions which he imposes on the

59

velocities are much more severe. He assumes that the $\{v_k\}$ are independent identically distributed random variables having a density $f_s(v)$ independent of the positions $x_k(0)$. The $f_s(v)$ must also be reasonably smooth (continuous almost everywhere and bounded in every finite velocity interval). The key assumption, however, is the statistical independence.

From these assumptions Breiman proves that in the limit $t \to \infty$, the number of cars in any interval of highway of length y has a Poisson distribution with mean ky. It follows also if we extract from these cars only those with velocity in some range v to $v + \Delta v$ with $\Delta v > 0$, these cars by themselves satisfy the same conditions and consequently the number of cars in y with velocity in Δv has a Poisson distribution with mean

$$ky[F_s(v+\Delta v)-F_s(v)], \qquad f_s(v) = dF_s(v)/dv.$$

As a heuristic argument one can say first that under the above conditions we should expect that if a density $k(x,t)$ of cars is defined in the sense of (2.3) at t = 0 then $k(x,t) \to k$ for $t \to \infty$. Here the problem is essentially the same as the traffic signal problem discussed above, which suggests that time-dependent flows at x = 0 approach a constant flow for x $\to$ ∞. The existence of the spatial density $k(x,t)$ is obviously not essential, however, if one only counts cars in non-zero length intervals of highway.

That the number of cars should have a Poisson distribution, however, is a consequence of the independence of the velocities. If we were to specify the positions of all cars at t = 0,

That the number of cars should have a Poisson distribution, however, is a consequence of the independence of the velocities. If we were to specify the positions of all cars at t = 0, then the condition that some j^{th} car appear in an interval y of the highway at time t is a condition imposed upon the velocity of that car. The velocities of cars, however, are assumed to be independent and consequently the probability that some j^{th} car appear in this interval at time t is statistically independent of whether or not any other cars may be there. As t

becomes large, the probability of any given j^{th} car being in this interval of highway is very small. Again we have the typical situation that leads to a Poisson distribution. We ask how many events occur (how many cars are in the y-interval at time t) when there are infinitely many possible events to choose from (infinitely many cars) but the probability of any one occurring is very small and they are independent. A rigorous proof of this is straightforward but requires some care.

Although Breiman showed only that the number of cars in any interval of highway of length y approaches a Poisson distribution with mean ky for t → ∞, Thedéen [13] pointed out that the numbers of cars in disjointed sections of highway are also asymptotically independent for t → ∞. Consequently the positions of cars define a Poisson process for t → ∞. The same is also true of only those cars having velocities in some range of velocities v to v + Δv. Furthermore for t→∞, the Poisson limit process for cars of velocity between v and v + Δv is statistically independent of the process of cars with velocities not in this range. It would seem then that for t→∞, we are obtaining a Poisson process with statistically independent velocities. One should note, however, that this result can be derived only if we let t → ∞ for y > 0 and Δv > 0. Then perhaps let y → 0 or Δv → 0. We do not let y → 0 and/or Δv→0 first and then let t → ∞ .

Actually it is possible to relax considerably the assumptions concerning the stochastic

Actually it is possible to relax considerably the assumptions concerning the stochastic structure of the velocities and still obtain the same result. Suppose for example that the velocities of cars are independent only if they are initially sufficiently far apart, but some restriction is imposed to prevent most neighboring cars on the highway from having almost equal velocities. It seems plausible perhaps that if cars are far enough apart they will be

independent. If now two or more cars appear in the y interval at time t it is almost certain that they were far apart at t = 0 and consequently independent. Otherwise these cars would not only have to be near each other at t = 0 but also have nearly identical velocities in order still to be within a distance y of each other at time t.

There are limits on how much one can relax the conditions on the initial distribution and still prove a convergence of some sort to a Poisson distribution. If one analyzes the distribution in too much detail one runs into contradictions from the reversibility arguments. The entropy does seem to be increasing and in fact this Poisson distribution with independent velocities is the distribution with the maximum entropy per unit length of highway among all distributions with the same average density of cars in every velocity range [15],[15a].

To see how the entropy or information changes, suppose that initially the $x_k(0)$ are fixed (not random) but satisfy (a) and (b) above. We then have complete information about the positions at t = 0. At any later time, one can find at any <u>point</u> x only cars with certain discrete velocities determined by the condition that the velocity must have the value [x-$x_k(0)$]/t for some k. In the above argument for the approach to a Poisson distribution we did not, of course, look at a point x; we looked at an interval of highway and in an interval we obtain a continuous range of velocities. Part of the information at t = 0 thus goes into rapid variations of the velocity distribution. In addition to this some other information which one obtain a continuous range of velocities. Part of the information at t = 0 thus goes into rapid variations of the velocity distribution. In addition to this some other information which one has for finite t is lost in the limit t → ∞.

If for any finite t we observe a car at x, we know that the same car cannot also be observed somewhere else at time t. This implies some statistical dependence between the number of cars observed in one interval of highway and the number observed in other

intervals. There will in fact be situations such that if some car is observed in one interval some other interval must have no cars. As t increases, however, this correlation becomes spread over longer and longer distances. The probability that a given car appears in any given interval of the highway becomes very small anyway. Once we observe a car in one interval, the small probability that it would have been seen in some other interval must be changed to zero. This information is again lost through a mathematical scheme where we evaluate in the limit t →∞ mathematical quantities which are not sensitive to these weak long range correlations. This does not mean, however, that these correlations might not be significant to the evaluation of some other quantities than those considered here.

Finally, we conclude the discussion of this model by showing that the Poisson distribution with independent velocities either for the crossings at x = 0 or the positions at t = 0 gives the same for the crossings at any x or the positions at any t. This follows almost immediately from the following more general property of these distributions: If at t = 0 we have this distribution for initial positions and velocities and we take any (measurable) set of points A_o in the two dimensional (x,v) space of positions and velocities, then the number of cars with position and velocity in A_o at time t = 0 will have a Poisson distribution.

If A_o is a rectangle then we are asking for the distribution of the number of cars at t = 0 in some interval of highway with velocities in some specified range. That this is Poisson

If A_o is a rectangle then we are asking for the distribution of the number of cars at t = 0 in some interval of highway with velocities in some specified range. That this is Poisson distributed and independent of the number of cars in any non-overlapping rectangle can be interpreted as the given hypothesis. The sum of independent random variables each with a Poisson distribution, however, is itself Poisson distributed. Consequently the statement is also true if A_o is the union of non-overlapping rectangles. If A_o is the limit of the union of

rectangles, the probabilities on A_0 will be defined as the corresponding limits of the probabilities on approximating unions of rectangles. This implies that the statement is also true for limits of sets generated from the rectangles, which include all sets of interest here.

Any car that has position $x(t)$ and velocity v in some set A_t at time t, i.e., $(x(t),v)\varepsilon A_t$, must have been in a set A_0 at time 0 where A_0 is the set of points $(x(0),v)$ such that $(x(0)+vt,v)\varepsilon A_t$. The probability that there are n cars in A_t at time t is therefore equal to the probability that these n cars are in A_0 at time 0. The latter, however, has a Poisson distribution, independent of the number of cars in any non-overlapping set. The same must therefore also be true for time t.

If we were to ask instead how many cars cross some point x with times and velocities in some point set B_x of the (t,v) space, we can also determine a unique set of initial conditions A_0 that these cars must satisfy at time 0 in order to cross x with coordinates in B_x. Since the number in any A_0 is Poisson distributed the number of crossings in B_x is also Poisson distributed. Actually one can take essentially any property at all for positions, velocities or crossing times of fixed or moving points and conclude that the number of cars with this property will be Poisson distributed and independent of the number of cars having any other exclusive property. For example the number of cars passing a given car in some time interval will be Poisson distributed [12].
exclusive property. For example the number of cars passing a given car in some time interval will be Poisson distributed [12].

8. Velocity fluctuations. We saw in the last section that any theory with completely deterministic equations of motion can lead to incorrect conclusions. If some of the controversies appear academic, some consequences to be described later are less so.

Very little is known about the stochastic properties of individual drivers but certain rather obvious qualitative features of human behavior do severely limit the range of acceptable models. We wish to retain the concept that different drivers have different behavior and in particular each driver is labeled with a desired speed v_j which, however, we now interpret as some type of average speed for this j^{th} driver when uninhibited by other cars, traffic signals, etc.

In addition to the probability distribution associated with the selection of different drivers with different v_j, which still conforms more or less with the pattern described previously, the actual velocity $v_j(t)$ of a driver with given desired speed shall be now interpreted as a random function of t. In the absence of any interaction between cars the time-dependence of one of the $v_j(t)$ should be statistically independent of the velocity fluctuations of any other driver. The stochastic nature of the $v_j(t)$ is assumed to arise from human or mechanical fluctuation, for example, chance fluctuations in the pressure a driver exerts on the accelerator pedal.

If the highway is spatially homogeneous and there are no external time-dependent influences, the velocity $v_j(t)$ should be a stationary process in the sense that for any set of time points $t_1, t_2, .., t_n$, the joint probability distribution of $v_j(t_1), .., v_j(t_n)$ should be invariant to simultaneous translations of $t_1, t_2, .. t_n$, i.e. the joint distribution is the same as that for $v_j(t_1+\tau)$, time points $t_1, t_2, .., t_n$, the joint probability distribution of $v_j(t_1), .., v_j(t_n)$ should be invariant to simultaneous translations of $t_1, t_2, .. t_n$, i.e. the joint distribution is the same as that for $v_j(t_1+\tau)$, .., $v_j(t_n+\tau)$ for any τ and any choice of n, $t_1, t_2, .. , t_n$.

There is a tremendous literature on the theory of stationary stochastic processes including several books [16-18]. Much of this theory deals with the analysis of the covariance $E\{v_j(t)v_j(t+\tau)\}$ and its Fourier representations much of which is potentially relevant to traffic

theory, and the extrapolation or prediction problems which are probably of less practical interest in relation to the subject discussed in this chapter.

We will assume that time averages can be evaluated over sufficiently long times that we may idealize to infinite time averages, and that the process $v_j(t)$ for a given driver is ergodic, which by definition means

$$E\{v_j(t)\} = \lim_{t \to \infty} \frac{1}{T} \int_0^T v_j(t)dt \qquad (8.1)$$

i.e. the long time average of $v_j(t)$ is (with probability one) equal to the expectation for any fixed t. For a stationary process $E\{v_j(t)\}$ must, of course, be independent of t and we shall now identify this as the desired speed for this model with random velocities

$$v_j = E\{v_j(t)\}. \qquad (8.2)$$

It is also convenient to consider the mean and the fluctuations separately. Let

$$\eta_j(t) = v_j(t) - v_j \qquad (8.3)$$

so that

$$E\{\eta_j(t)\} = 0$$

and

$$v_j(t) = v_j + \eta_j(t).$$

All of the expectations here are conditional expectations given the particular driver.

$$v_j(t) = v_j + \eta_j(t).$$

All of the expectations here are conditional expectations given the particular driver.

If a driver is chosen at random from some population, the velocity $v(t)$ of the randomly selected car will also define a stationary stochastic process but it will not necessarily be ergodic. The time average speed will be the desired speed v_j of whatever driver

66

happened to be selected but the selected driver does not necessarily have a desired speed equal to the mean $E\{v\}$ of all desired speeds from the population.

Ergodic processes play an important role in the theory of stationary stochastic processes and in this theory the value of v_j would be described as belonging to the "deterministic part" of the process $v(t)$. This terminology is motivated by the fact that in predicting the future behavior of the process $v(t)$, the value of v_j can be inferred from observation on the past behavior.

The stochastic properties of $\eta_j(t)$ will in general also be different for different drivers and perhaps different even for different drivers with the same v_j. There are, no doubt, other parts of $v(t)$ which in a realistic theory would be considered as deterministic. It is perhaps sufficient at this stage in the development of traffic theory to disregard this, however, and imagine that the stochastic properties of $\eta_j(t)$ are the same for all drivers or at least for all drivers with the same v_j.

If the velocity of a car is random, so is its position which is determined from the equation

$$x_j(t) = x_j(0) + v_j t + \int_0^t \eta_j(\tau) d\tau .$$ (8.4)

For the conditional expectation given j, we have

For the conditional expectation given j, we have

$$E\{x_j(t)\} = E\{x_j(0)\} + v_j t$$ (8.5)

thus the mean position is determined by the desired or mean speed v_j.

67

We can anticipate the probable existence of two qualitatively quite different types of behavior for $v_j(t)$ and $x_j(t)$. Mathematically one could have many others but the following seem most natural.

1. A bus driver is expected to maintain a fixed schedule and knows at what time he should arrive at various points along his route. If chance fluctuations cause him to get ahead or behind schedule, he will decrease or increase his speed accordingly. His goal is to keep the fluctuations in distance traveled within certain limits independent of the time. In this case it is reasonable to assume that $\int_0^t \eta_j(\tau)d\tau$ approaches a stationary process for large t, or perhaps, for all t, it is a process obtained from a stationary processes conditional that the processes have the value 0 at $t = 0$.

A similar example is furnished by an alert driver traveling through a sequence of synchronized traffic signals. If he gets a little ahead or behind the ideal speed he can perhaps correct for his errors early enough to keep in phase i.e. he will correct before the errors are so large that he is stopped by some signal. In effect, his observations of the signals give him a measure not of his instantaneous velocity but of his accumulated errors in velocity or his travel distance relative to what it should be to keep in phase.

2. In the absence of clocks or other means of measuring and correcting errors in distance traveled, a driver will observe only his velocity and make occasional corrections for deviation
2. In the absence of clocks or other means of measuring and correcting errors in distance traveled, a driver will observe only his velocity and make occasional corrections for deviation from the speed he wishes to maintain. With this type of control $\int_0^t \eta_j(\tau)d\tau$ is not necessarily bounded for $t \to \infty$; it is not a stationary process. In the terminology of the theory of random time series, it would be classed as a process of stationary increments, if $\eta_j(t)$ is a stationary process.

If a significant fraction of the traffic behaves in a manner analogous to (1) above, the traffic is not likely to be Poisson distributed even in any limiting sense. The population of buses on a given route, scheduled so they do not pass each other and perhaps do not even get very close, certainly is not Poisson distributed. The arguments of the last section about the approach to a Poisson distribution break down mainly because the probability distribution of desired speeds should not be considered as continuous. All bus drivers are constrained to maintain nearly the same long time average speed. In effect the distribution of speeds is nearly discrete and even fluctuations in velocity will not smear the ordered state of flow enough to create a completely random situation. Actually scheduled buses form a part of traffic with highly predictable behavior which violates almost any of the stochastic assumptions used for other types of traffic. If it forms a significant part of the total traffic flow, it must be considered separately.

Cars that can pass through long sequences of synchronized signals also have effectively a discrete velocity distribution. The spatial distribution or distribution of arrivals for these cars will not become Poisson either.

If the desired speeds did have a continuous probability distribution, then even fluctuations of type 1 would remove some of the reversibility paradoxes of the deterministic models and give a continuous increase in entropy approaching that for a Poisson distribution, fluctuations of type 1 would remove some of the reversibility paradoxes of the deterministic models and give a continuous increase in entropy approaching that for a Poisson distribution, but it is not obvious where such a model would be applicable.

For the second type of velocity fluctuations it is reasonable to assume that for sufficiently large τ, $\eta_j(t)$ and $\eta_j(t+\tau)$ are statistically independent. There should in fact be some characteristic time τ_o, say, that gives a measure of the relaxation time for fluctuations.

69

If $\tau \ll \tau_o$ then $\eta_j(t)$ and $\eta_j(t + \tau)$ should be nearly equal i.e. the time τ is too short for any driver to change his speed significantly. But if $\tau \gg \tau_o$ memory of the fluctuation $\eta_j(t)$ at time t has been lost by time $t + \tau$. We will not make any further speculations on the probability structure of the $\eta_j(t)$.

Over short time intervals ($\tau \ll \tau_o$), the velocity of a car is nearly constant and there is no real distinction between a car having desired speed v_j plus a fluctuation $\eta_j(t)$ at time t and some other car having a desired speed $v_j + \eta_j(t)$ at that time but no fluctuations. In effect we have a motion equivalent to that given by the model with constant velocities except that the distribution of speeds is that of $v_j + \eta_j(t)$.

The long time behavior is more interesting, however. The distance traveled is essentially a Brownian motion about the mean [19] and any model with the stochastic properties described above behaves in a manner very similar to the following rather artificial prototype. Suppose that at each integer time point measured say in units of τ_o a driver suddenly selects a new velocity fluctuation. At time $k\tau_o$ he chooses a velocity $v_j+\eta_j(k\tau_o)$ with the $\eta_j(k\tau_o), k=0,1, \cdots$ independent identically distributed random variables with $E\{\eta_j(k\tau_o)\}=0$. The distance traveled in time $\eta\tau_o\!\sim$ is then

$$x_j(n\tau_0) = x_j(0) + v_j n\tau_0 + \tau_0 \sum_{k=0}^{n-1} \eta_j\,(k\tau_0) \quad . \tag{8.6}$$

$$x_j(n\tau_0) = x_j(0) + v_j n\tau_0 + \tau_0 \sum_{k=0}^{n-1} \eta_j\,(k\tau_0) \quad . \tag{8.6}$$

The last term is the fluctuation. It is the sum of random variables, and according to the central limit theorem, the distribution for this sum will be asymptotically normal for large n with a mean given by (8.5) and variance

$$Var\{x_j(n\tau_o)\} = n\tau_0^2 \; Var\{\eta_j(0)\} \; .$$

70

The standard deviation of the $x_j(t)$ thus increases proportional to $t^{1/2} = (n\tau_o)^{1/2}$.

For any more realistic models with continuously varying $\eta_j(t)$ but statistical dependence that decays rapidly enough with time, $x_j(t)$ will be asymptotically normal with again the same mean as given by (8.5) but with a variance which is some constant (not necessarily 1) times $\tau_o t \, \text{Var}\{\eta_j(0)\}$.

This behavior relates to the motion of a given j^{th} car with specified desired speed v_j. Over a long period of time t, this car travels to a first approximation a distance $v_j t$ but then in the next approximation we must add to this the relatively much smaller normally distributed fluctuations with a standard deviation that increases only as $t^{1/2}$. This motion is exactly the same as the much studied one-dimensional Brownian motion of a particle except that it is superimposed upon a particle (car) traveling with some speed v_j. The probability density for $x_j(t)$ satisfies the usual diffusion or heat conduction equation having as its initial value Green's function the normal distribution described above.

If we were to choose a driver at random from some population, then the desired speed v is also a random variable. If, in addition, each car behaves according to (8.6) for example, we take $x_j(0)=0$, and assume that all cars have identically distributed $\eta_j(k\tau_o)$, then the distance traveled in time $n\tau_o$ has a variance $\text{Var}\{x(n\tau_o)\} = n^2\tau_o^2\text{Var } v + n\tau_o^2 \, \text{Var}\{\eta(0)\}$, the sum of the variance due to the distribution in desired speeds and that due to the Brownian motion. As noted previously the fluctuations in travel distance due to differences in desired speed increase as t (the variance as t^2) whereas that of the Brownian motion grows only as $t^{1/2}$. For large t, the former, of course, dominates the latter (provided Var v $\neq$ 0) but this does not necessarily mean that the latter can be disregarded.

71

For this model with velocity fluctuations of the second type, one can now prove a much stronger type of convergence to a Poisson process than that given in section 7. The heuristic argument for the case of no fluctuations was that if we pick an interval of highway of length y , the probability of finding in y any specified j^{th} car goes to zero for t→∞, and the probability of the jth car being in y is (asymptotically) independent of whether or not any other car is in y. These are the key stochastic properties that give rise to a Poisson distribution for the number in y. The probability structure in this case was derived in section 7 from the assumption that at time zero each car had a desired speed statistically independent of any other, and that the speed distribution had a continuous probability density. The argument fails, however, if, for example, all cars have exactly the same velocity (or any discrete velocity distribution) because cars with exactly the same velocity maintain for all time the spacing they had at time 0. There is no tendency for the spacing distribution, arbitrarily specified at time 0, to approach an exponential distribution.

Suppose now we add the velocity fluctuations and reconsider even this extreme case in which all cars have the same desired speed (now interpreted as the time average speed), plus specified positions at time 0. The position of a car at time t has a variance now (due to the fluctuations) that increases linearly with time. The distribution of the position of any j^{th} car at time t becomes spread over a distance of order $t^{1/2}$. This spread becomes infinite for t→∞ fluctuations) that increases linearly with time. The distribution of the position of any j^{th} car at time t becomes spread over a distance of order $t^{1/2}$. This spread becomes infinite for t→∞ and the probability of finding this car in any given finite interval of length y goes to zero like $t^{-1/2}$. Since the fluctuations in velocity of one car are assumed to be statistically independent of those of any other car, it follows also that the probability for one car to be in y is statistically independent of whether or not any car is in y. We have the same basic

72

properties that give rise to a Poisson distribution for the number in y, but they arise from a different mechanism. The convergence to the Poisson distribution is, in some sense, slower than before since the probabilities for a car being in y go to zero only as $t^{-1/2}$ whereas for a distribution of desired speeds these probabilities go to zero as t^{-1}.

The entropy now increases continuously with time due to the inherently irreversible features of the velocity fluctuations and there are no tricks such as interchange of limits. The fluctuations also destroy correlations between velocity and positions that appear in the deterministic model.

As a simple illustration one could reconsider the flow from a fixed cycle traffic signal as in (6.12) assuming now that all cars have exactly the same desired speed but independent velocity fluctuations. The flow $q(x,t)$ will again approach a constant for $x \to \infty$. The scale of distance on which this takes place is again the distance x which cars must travel before the uncertainty in arrival times at x are comparable with the cycle time T, so that cars from one cycle overtake those from other cycles. The mechanism and the rate of spreading are different, however. The present mechanism should be much slower on two accounts. First we would expect Var η to be much smaller than Var v, the errors of an individual in selecting his velocity should normally be much less than the differences in desired speeds of different drivers. Secondly the variance increases with time or distance traveled at a slower rate.
his velocity should normally be much less than the differences in desired speeds of different drivers. Secondly the variance increases with time or distance traveled at a slower rate.

9. <u>Further notes</u>. Most of the techniques and arguments used in this chapter originated in fields other than traffic theory and are in some cases more than 100 years old. Application of probability models to traffic problems dates from the 1930s. Some of the early contributions

were by Kinzer (1934)[20], Adams (1936)[21], and Garwood (1940)[22]. In particular the use of the Poisson distribution in highway traffic originates with Kinzer and Adams.

The use of the velocity distribution of cars to describe the dispersion of a pulse was apparently suggested first by Pacey [23] who also did some crude experiments to check the validity of the theory.

Grace and Potts [24-25] have evaluated the spatial density $k(x,t)$ from (6.5) explicitly for a number of special initial distributions $k(x,0)$ other than the δ-function distribution of (6.10) and for a normal distribution of velocities $f_s(v)$. They point out that if $f_s(v)$ is normal then $k(x,t)$ satisfies a partial differential equation of the type

$$\frac{\partial k(k,\ t)}{\partial (t^2)} = A\ \frac{\partial^2 k(x,\ t)}{\partial (x\ -\ mt)^2} \tag{9.1}$$

where A and m are constants depending upon the variance and mean of the distribution $f_s(v)$. This is the diffusion equation relative to the independent variables t^2 and x-mt (instead of the more usual variables of the diffusion equation t and x). This equation is valid, however, only if the initial velocity distribution is normal and also independent of x.

For a model with all cars traveling with the same desired speed v and with a Brownian motion for velocity fluctuations, one also obtains an equation like (9.1) in terms of the variables t and x-vt but for quite different reasons. That there is a t^2 in one equation Brownian motion for velocity fluctuations, one also obtains an equation like (9.1) in terms of the variables t and x-vt but for quite different reasons. That there is a t^2 in one equation where there is a t in the other is, of course, associated with the fact that in one model a pulse spreads linearly with t but as $t^{1/2}$ in the other.

Other experimental studies of platoon spreading have been done by B.J. Lewis [26], Graham and Chenu [27], and Herman, Potts, and Rothery [28]. These experiments all agree

74

qualitatively with the constant velocity model but none are sufficiently detailed to distinguish between various similar theories. No experiments have been done to investigate the distinction between time average speeds and fluctuations.

References
Chapter II

1. L.C. Edie, Discussion of traffic stream measurements and definitions. Proc. second international symposium on the theory of road traffic flow, London 1963, pp. 139-154, Organisation for Economic Cooperation and Development, Paris (1965).

2. W. Feller, An introduction to probability theory and its applications, Vol. 1, second ed. Wiley, New York (1957).

3. D.L. Gerlough, The use of the Poisson distribution in problems of street and highway traffic," Poisson and Traffic," The Eno Foundation, New Haven (1955).

4. M. Rosenblatt, Random Processes, Oxford University Press (1962).

5. J.G. Wardrop, Some theoretical aspects of road traffic research, Proc. Inst. Civil Engrs. II 1(2), 325-362 (1952).

6. M.J. Lighthill, An introduction to Fourier analysis and generalized functions, Cambridge University Press (1958).

7. C. Shannon and W. Weaver, The mathematical theory of communication, University of Illinois Press (1949).

8. G. Weiss and R. Herman, Statistical properties of low density traffic. Quart. Appl. Math. 20, 121-130 (1962).

9. A.J. Miller, On the theoretical justification of the random queues model for road traffic flow, unpublished (1962).

10. Leo Breiman, On some probability distributions occurring in traffic flow, Bulletin of the Institute of International Statistics 39, 155-161 (1962).

11. Leo Breiman, The Poisson tendency in traffic distribution, Ann. Math. Stat. 34, 308-311 (1963) Institute of International Statistics 39, 155-161 (1962).

11. Leo Breiman, The Poisson tendency in traffic distribution, Ann. Math. Stat. 34, 308-311 (1963).

12. A.Rényi, On two mathematical models of the traffic on a divided highway, J. Appl. Prob. 1, 311-320 (1964).

13. T. Thedéen, A note on the Poisson tendency in traffic distribution, Ann. Math. Statist. 35, 1823-1824 (1964).

14. T. Thedéen, Convergence and invariance questions for road traffic with free overtaking, Proc. third international symposium on theory of traffic flow, New York (1965).

15. A. Rényi, On an extremal property of the Poisson process, Ann. Inst. Statist. Math 16, 129-133 (1964).

15a. J.A. McFadden, The entropy of a point process, Soc. Industrial Appl. Math Jour. 13, 988-994 (1965).

16. U. Grenander and M. Rosenblatt, Statistical analysis of stationary time series, Wiley and Sons, New York (1957).

17. J.H. Laning, Jr. and R.H. Battin, Random processes in automatic control, McGraw-Hill, New York (1956).

18. A.M. Yaglom, An introduction to the theory of stationary random functions, translated by R.A. Silverman, Prentice-Hall (1962).

19. N. Wax, ed., Selected papers on noise and stochastic processes by J.L. Doob, S. Chandrasekhar, S.O. Rice, G.E. Uhlenbeck, M. Kac, L.S. Ornstein, Dover, New York (1954).

20. J.P. Kinzer, Application of the theory of probability to problems of highway traffic, abstract by L.F. Rader, Proc. Inst. Traffic Engrs. 5, 118-124 (1934).

21. W.F. Adams, Road traffic considered as a random series, J. Inst. Civil Engrs. 4, 121-130 (1936).

22. F. Garwood, An application of the theory of probability to the operation of vehicular controlled traffic signals, J. Royal Statist. Soc. (London) supp. 8, 7, 65-77 (1940).

23. G.M. Pacey, Progress of a bunch of vehicles released from a traffic signal, Road Research Laboratory, England, Report RN/2665, unpublished (1956).

24. M.J. Grace and R.B. Potts, Diffusion of traffic platoons, Proc. first conf. Australian Road Research Board, pp. 260-267 (1962).

24. M.J. Grace and R.B. Potts, Diffusion of traffic platoons, Proc. first conf. Australian Road Research Board, pp. 260-267 (1962).

25. M.J. Grace and R.B. Potts, A theory of the diffusion of traffic platoons, Opns. Res. 12, 255-275 (1964).

26. B.J. Lewis, Platoon movement of traffic from an isolated signalized intersection, Highway Research Board Bull., No. 178 (1958).

27. E.P. Graham and D.C. Chenu, A study of unrestricted platoon movement of traffic, Traffic Eng. 32, 11-13 (1962).

28. R. Herman, R.B. Potts, and R.W. Rothery, Behavior of traffic leaving a signalized intersection, Traffic Engr. and Control, 5, 529-533 (1964).

III. LOW DENSITY TRAFFIC (WEAK INTERACTIONS)

1. <u>Introduction</u>. In the theory described in chapter II, interactions between cars were neglected completely. The inclusion of interactions makes the theory drastically more complicated because one cannot follow the motion of one car without considering simultaneously the motion of other cars. Furthermore, if passing is allowed, a car will, over a long period of time, interact with many other cars. It appears that one must consider simultaneously the motion of virtually <u>all</u> cars.

The theory to be described in this chapter is essentially a "perturbation theory" in which we try to estimate the lowest order effects of weak interactions. To a large extent the theory mimics the treatment of transport phenomena (heat conduction, diffusion, etc.) in the kinetic theory of gases. In the theory of gases, a molecule is assumed to travel at constant velocity until it collides with another molecule. A collision occurs only if the molecules come sufficiently close together and for rarefied gases lasts only for a time which is short compared with the mean time between collisions. The analogue of a molecular collision is not interpreted here as a collision of two cars in the usual sense, but rather some short range interaction, the end result of which is that a faster car passes the slower one after suffering some delay. One important difference between cars and molecules, however, is that after two interaction, the end result of which is that a faster car passes the slower one after suffering some delay. One important difference between cars and molecules, however, is that after two cars interact, they each return to approximately the same speed they had before the interaction whereas two molecules will exchange energy during the collision and acquire new velocities. This exchange of energy during collision is crucial in the argument for the existence of such a thing as a diffusion or a heat conduction equation for gases. Although there are certain

similarities in the approach to the treatment of low density traffic and a nearly ideal gas, this preservation of desired speeds by drivers will soon cause a divergence in the type of development and lead to a macroscopic theory for cars quite different from that of a gas.

The theory of chapter II can be interpreted here either as an idealization for traffic situations in which cars do not interact very strongly with each other as for example on a multilane highway at low density, or it can be considered as an exact theory for the limit behavior on essentially any homogeneous highway for $k \to 0$, because for $k \to 0$ cars hardly ever meet each other anyway. The theory to be considered in this chapter is an extension of the previous theory in the sense of the latter interpretation. If the density of cars, k, is small but nonzero, the probability that a given car is within some finite distance d of some other car should be approximately of order kd. The probability that there are two cars within a distance d of a given car should be of order $(kd)^2$ for $k \to 0$.

If d is chosen to be a measure of the range of interaction within which one might expect some deviations from the constant velocity trajectories of chapter 2, then the theory of chapter 2 is correct "most of the time," except possibly during a fraction of the time of order kd. The theory to be considered now is directed toward an investigation of the first order effects of interactions, the effects of relative order kd. This includes the effects of interactions between pairs of cars. But we will systematically neglect all effects of order $(kd)^2$, which effects of interactions, the effects of relative order kd. This includes the effects of interactions between pairs of cars. But we will systematically neglect all effects of order $(kd)^2$, which will mean that we neglect the consequences of simultaneous interactions between three cars. The purpose of the theory, however, is two-fold. First it is an end in itself, a description of traffic at low densities, but secondly it is a source of suggestions regarding the properties of

traffic at moderate densities, the mathematical analysis of which is certain to be very complicated.

In section 3.2 we determine the effect of interactions on the long time average speed of a car with given desired speed. This is a very simple calculation, but it gives little insight into the stochastic structure of the traffic. In a subsequent section, we determine how the interactions distort the Poisson process of steady flow for non-interacting cars. This is followed by some discussion of time-dependent flows and flows on expressways.

2. Average speed. Consider a long homogeneous section of highway with a uniform flow of traffic which, in the absence of any interactions, would have a homogeneous Poisson distribution of cars of spatial density k and independently distributed velocities with a spatial velocity probability density $f_s(v)$. On a typical rural highway most drivers when overtaking a slower car will slow down and look for an opportunity to pass. Once they pass, they return to essentially the same speed as before. The net effect of this is that the faster driver has lost some time or equivalently his time average velocity is less than if he could pass without slowing down. If the density of cars is very low, the frequency of passing will be low and the loss in average speed will also be small. To estimate this loss, however, we must know the rate at which our reference car overtakes other cars. This, in turn, depends
be low and the loss in average speed will also be small. To estimate this loss, however, we must know the rate at which our reference car overtakes other cars. This, in turn, depends upon the velocity and spatial distribution of the other cars which have also been perturbed by interactions amongst themselves. The basic idea in the perturbation scheme is that to calculate the first order effects of interactions we will use the lowest order (no interaction) approxima-

tion to the distribution of positions and velocities; i.e., we assume a Poisson distribution with

independent velocities.

If we have a Poisson distribution of cars of density k, velocity distribution $f_s(v)$ and no

interaction between cars, then a car of velocity v passes cars of velocity between v' and $v' +$

dv' , $v' < v$, at a rate

$$(v - v')kf_s(v')dv'. \tag{2.1}$$

For each such passing, we assume that the car of velocity v suffers a loss either in travel time

or distance traveled. Suppose we let $d(v, v')$ be the average loss in distance traveled for each

passing of a car of velocity v' by a car of velocity v. Figure III 1 shows the space-time

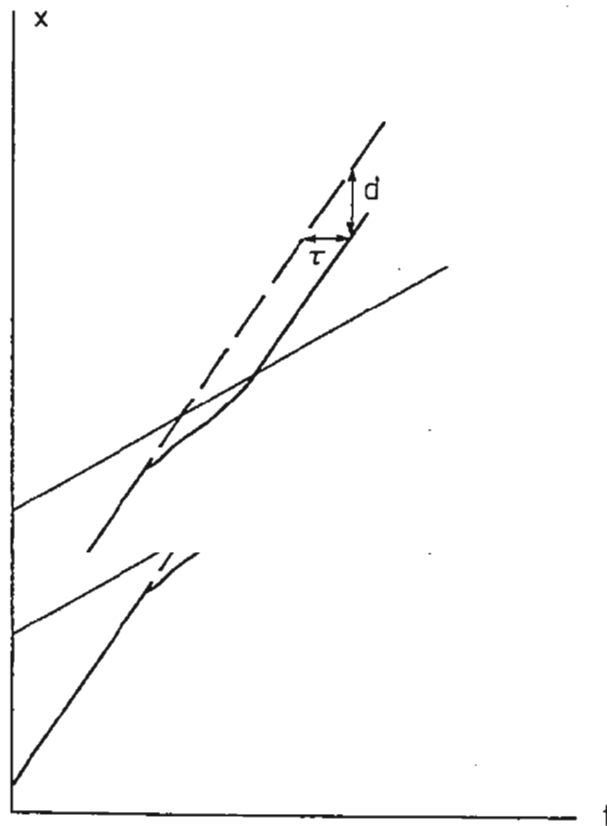trajectories of two cars. The faster car, velocity v, has the same velocity before and after

Fig. III 1
Space and time loss due to passing

82

passing. The distance $d(v, v')$ is the distance between the real trajectory and the corresponding trajectory (broken line) with no delay, or the extrapolation of the trajectory before passing.

Fortunately, the properties of the flow which are of practical interest are not very sensitive to the detailed form of this function $d(v, v')$. One can propose various mechanisms for the loss, none of which is very realistic in detail but any one of which will give approximately the same order of magnitude for the average delays in practical applications. A small value of $d(v, v')$ results if we assume that the loss is due only to the fact that the path of travel into the passing lane and back is longer than the direct path that one would have if cars could pass through each other. A larger value results if we assume that the fast car starts to decelerate when it is a "safe driving distance" behind the slower car and gradually decelerates until it is at the speed v'. The fast car then returns to the velocity v as it passes. Finally, we may consider that to pass on a typical rural road, a driver must find a gap in the traffic using the second lane. The fast car is forced to travel at the speed v' until it finds an acceptable gap. We do not attempt here to calculate $d(v, v')$; we assume it is "given."

The total average loss of distance traveled per unit time by a car of velocity v due to passing slower cars of any velocity $v' < v$ is

$$k \int_0^v dv' d(v, v')(v - v') f_s(v') \ .$$

The time average velocity of a car is the average distance traveled per unit time. If we call

$$k \int_0^v dv' d(v, v')(v - v') f_s(v') \ .$$

The time average velocity of a car is the average distance traveled per unit time. If we call this $u(v)$ then

$$u(v) = v - k \int_0^v dv' d(v, v')(v - v') f_s(v') + O(k^2) \ . \tag{2.2}$$

83

We have added a term $O(k^2)$, which will be justified later, as a reminder that we have made many approximations in the derivation of (2.2) and that the formula is valid only for small k.

This formula already displays the obvious facts that for small k the decrease in velocity is linear in k and is largest for the highest velocities v (provided one uses some reasonable form for d(v, v′)). From (2.2) we can also evaluate the average velocity for all cars.

$$E\{u\} = E\{v\} - k \int\limits_0^\infty dv f_s(v) \int\limits_0^v dv' d(v, v')(v - v') f_s(v') + O(k^2) \tag{2.3}$$

If we knew d(v, v′) and $f_s(v)$, we could evaluate the integral here but in any case we see that E{u} decreases linearly with k for small k. This prediction that a curve of average velocity E{u}vs. k has a negative slope at k = 0 is confirmed for most rural roads [1]. Not enough is known about d(v, v′), however, to make quantitative comparisons of theoretical and experimental values for the slope. Crude models for the passing maneuver do give values of d which are at least in qualitative agreement (within a factor of two perhaps) with experimental measurements of E{u}. For practical purposes it suffices at present to say that there is some effective average loss d for all passings, that the integral in (2.3) which contains the velocity difference (v - v′) as a factor in its integrand is also roughly proportional to the standard deviation σ of the velocity distribution, and therefore (2.3) has the form

difference (v - v′) as a factor in its integrand is also roughly proportional to the standard deviation σ of the velocity distribution, and therefore (2.3) has the form

$$E\{u\} = E\{v\} - kd \, \sigma C + O(k^2) \tag{2.4}$$

in which C is some dimensionless constant, the value of which depends upon the functional form of d(v, v′) and $f_s(v)$.

The method of computation for passing rates and the perturbation scheme used here are patterned after a similar method used by Maxwell for computing collision rates in the kinetic theory of gases a hundred years ago. These procedures, which will be analyzed more carefully in the next section, were first applied to the study of low density traffic by Newell [2], Bartlett [3], and Carleson [4]. The theory of Carleson, however, is somewhat more elaborate and has potential application to moderately dense traffic. It will be discussed in more detail in chapter 5 along with some other extensions of the theory by Miller [5].

Since this derivation of the first order effects of the density on the average velocity was so simple, one might be tempted to extend the perturbation scheme and evaluate the second order terms proportional to k^2. Although we have as yet only computed the first order changes in the average velocities, we could, by methods to be described in section 3, also evaluate the first order changes in the distribution of velocities, etc. From this one might expect that, following the usual iteration of perturbation techniques, we could evaluate the average velocity u to second order in k. Unfortunately this procedure suddenly becomes quite tedious due to a number of new phenomenon that enter into the second order theory, a few of which are listed below.

a. In the above derivation, we considered passings involving only two cars at a time. The density of interacting pairs of cars, however, should be proportional to k^2 and so the rate at
a. In the above derivation, we considered passings involving only two cars at a time. The density of interacting pairs of cars, however, should be proportional to k^2 and so the rate at which a car overtakes such pairs is also proportional to k^2. Three car interactions must, therefore, be considered in any second order theory. Very little is known about the form of $d(v, v')$ above. Even less is known about the queueing phenomenon that exists when two cars wish to pass a third car.

b. In the first order theory we used a Poisson distribution of cars (or at least certain properties of the Poisson distribution) in order to estimate passing rates. This is no longer valid for the next approximation. First of all, the delays give rise to a higher probability of finding cars close together particularly a fast car behind a slow one because the interactions cause cars to stick together for a while. Although one could estimate to first order in k the difference between the conditional density of cars given the position of one car and the corresponding expression for a Poisson distribution (for which the given position of one car is irrelevant), the necessary use of the distributions for pairs of cars in the second order theory adds complications. This, however, is not as annoying as the consequences of the fact that these interactions also give rise to correlations between velocities and positions. Suppose, for example, we know that three cars have interacted simultaneously. The two fastest cars pass the slowest one; but if the fastest of the three cars is the last to pass, then we know for certain that this fastest car will soon want to pass the car of intermediate speed. We can no longer say that the expected rate of passings in the future is independent of the past. Already in the second order theory we are beginning to see some of the complications inherent in a general exact theory.

c. In the first order theory it is assumed that the distribution of desired speeds on the highway $f_s(v)$ is the same as the unperturbed distribution. It is clear that the interactions cause

c. In the first order theory it is assumed that the distribution of desired speeds on the highway $f_s(v)$ is the same as the unperturbed distribution. It is clear that the interactions cause deviations between the desired speeds and the actual speeds but they can also cause a change in the observed distribution of desired speeds. Suppose, for example, we consider again the highway of length L which is fed by a parking lot in which the velocities have a distribution function $F_v(v)$. It is perhaps reasonable to assume that the manner in which cars are selected

from the parking lot does not depend upon the rate at which they leave; i.e., $f_i(v)$ is still the probability density of desired speeds observed per unit time at the entrance. The car of desired speed v, however, actually travels at a time average speed $u(v)$. The number of cars per unit length of highway with desired speed between v and $v + dv$ is $qf_i(v)dv/u(v)$ rather than $qf_i(v)dv/v$. Thus the probability density of desired speeds per unit length of highway is

$$f_s(v) = \frac{[u(v)]^{-1}dF_v(v)/dv}{\int [u(v)]^{-1}dF_v(v)} \tag{2.5}$$

and this depends upon k through the $u(v)$. The probability density of fast cars is increased due to the interactions because their time average speed is reduced most and so their time required to traverse the highway is increased. The longer cars of any velocity v stay on the highway, the larger is their density at any given time.

No one has tried to formulate a second order theory correctly, but it is clear from the above description of what has been neglected in the first order theory that the errors in (2.2) are of order k^2 for $k \to 0$ as already indicated. From a practical point of view, a second order theory would probably not be of great value since it will on the one hand contain so many unknown parameters that one could not evaluate very much; and on the other hand, there is probably only a rather narrow range of densities in which the second order theory could give much improvement over the first order theory before one would need also to there is probably only a rather narrow range of densities in which the second order theory could give much improvement over the first order theory before one would need also to include the third, fourth, etc., order terms as well. Various extensions of the theory to higher densities will be considered in chapter 5, but the object will then be to treat simplified models or obtain incomplete information without perturbation type arguments.

87

3. <u>Equilibrium distributions</u>. The derivations in the last section were quite insensitive to the detailed features of the model. Before we can study probability distributions of the traffic, it is desirable to define the model more carefully.

We will assume first that there is some finite zone of interaction around each car. Any driver who is not in the zone of interaction of any other car will travel at his desired speed just as in chapter II. We are not interested in following the detailed behavior of cars within the zone of interaction. It is convenient, therefore, to idealize all trajectories by piecewise linear curves. When a driver enters the zone of interaction of another car, we will imagine that this driver continues to travel at his desired speed until he actually coincides with the car he wishes to pass. He then instantaneously assumes the speed of the latter and keeps this speed for some non-zero length of time, after which he returns to his desired speed to complete the passing. The time spent in passing is to be so chosen that the idealized trajectory coincides with the true one outside of the zone of interaction. Thus in figure III 1 the correct trajectory is replaced by the dotted curve. This approximation will have no effect upon the calculation of car positions outside the zone of interaction. Cars which are interacting with each other will be identified by the coincidence of two cars.

Each ith car is permanently assigned a desired speed. He is also assigned a set of distances d_{ij}, the loss in distance traveled when the ith car passes the jth car. In the stochastic

Each ith car is permanently assigned a desired speed. He is also assigned a set of distances d_{ij}, the loss in distance traveled when the ith car passes the jth car. In the stochastic treatment of this model, we will, however, assume that the ith driver with velocity v does not identify other drivers by their number, but only by their velocity v'. For fixed v and v', the d_{ij} will be considered as identically distributed random variables. It is also reasonable to assume, at least for very light traffic, that all the d_{ij} are statistically independent. The justification for

this is that in the light traffic two passings are not expected often to occur close enough to each other in both space and time so as to influence one another. Furthermore, if one driver passes another, he is not likely ever to encounter the same driver again and thereby be influenced by his past associations.

Throughout this section low density traffic shall be interpreted to mean that the zones of interaction or the d_{ij} are small compared with the average distance between cars; i.e., $d_{ij} \ll 1/k$. Any calculation will include only first order effects, i.e., terms linear in the dimensionless parameter kd_{ij}.

If again we consider the finite section of highway having length L with entrance at x = 0 as discussed in chapter II and we specify (a) the starting time of each jth car, (b) the velocity of each jth car, (c) the values of the d_{ij} , and (d) if two cars coincide at x = 0, the remaining distance lost or remaining time needed before the completion of the passing that is in progress, then the future behavior of each car is uniquely defined. The actual computation or graphical construction of the trajectories is straightforward, although not nearly as simple as with no interaction. In figure III 2, for example, we show a series of trajectories starting at x = 0 with no coincidences (no passings already in progress).

Trajectories can be constructed by at least three types of iterative schemes.

I. Start at x = 0 and draw straight line trajectories with the assigned velocities. Observe
Trajectories can be constructed by at least three types of iterative schemes.

I. Start at x = 0 and draw straight line trajectories with the assigned velocities. Observe the first position where any two trajectories meet (point a of figure III 2). Put a jog in the trajectory of the faster of the two cars with the appropriate d_{ij}. Now look for the next smallest x at which two trajectories meet (point b) and put another jog in the trajectory of the faster of these two cars. Continue this procedure. If per chance three cars should interact
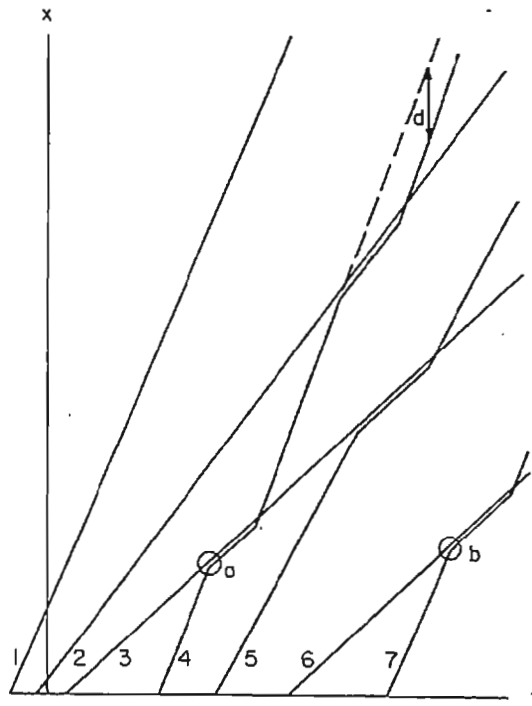
89

Fig. III 2
Construction of trajectories

simultaneously, one can introduce some more or less arbitrary finite delays. We assume that

these will occur so rarely as to be of little consequence in any quantitative calculation. To

keep the mathematics as simple as possible, however, it is convenient to imagine that the

delays suffered in a three car interaction are the same as those that would have occurred had

each passing taken place without interference with a third car; i.e., the total loss in distance

traveled by car is the sum of the d_{ij} for all cars j that have been passed.

each passing taken place without interference with a third car; i.e., the total loss in distance

traveled by car is the sum of the d_{ij} for all cars j that have been passed.

II. Draw the trajectory of the slowest of all the cars. This car will never pass any other

cars, and its trajectory will therefore be a straight line (car 3 of figure III 2). Next draw the

trajectory of the second slowest car. If this trajectory intersects that of the slowest car, insert

the appropriate jog. Continue to add trajectories of the faster cars. At each stage the trajectory

90

that is added will depend only upon the trajectories of the slower cars which have already been determined in previous steps.

III. If at $x = 0$ there is a first car to leave, number the cars so that $t_1 \leq t_2 \leq t_3$ -- at $x = 0$. (If we have an initial value problem with positions x_j specified at $t = 0$ and there is a largest x, number the cars so that $x_1 \geq x_2 \geq x_3$ --). Car 1 will never pass any other car (although others may pass it) and we can draw its straight line trajectory. Car 2 can only pass car 1 and so we can now draw its trajectory knowing that of the first car. At each stage the jth car can pass only cars $k < j$ whose trajectories are determined in previous steps.

Some of the definitions and relations discussed in Chapter II are still valid for this more general model. The definitions of density and flow $k(x, t)$ and $q(x, t)$ equations (II 2.3) to (II 2.9), are still meaningful. Although the existence of coincidences implies that the joint probability distribution for positions and velocities will not have a joint probability density (the conditional distribution of x_j given x_k will have a discrete component at $x_j = x_k$), it is still reasonable to assume that the marginal distributions for a single car have densities.

The relations between $k(x, t)$ and $q(x, t)$ at the same x and t as in (II 5.2) to (II 5.10), or between $\rho_t(x, v, t)$ and $\rho_s(x, v, t)$, equation (II 6.3), are also valid provided we interpret all velocities to be the actual velocities of cars at any instant rather than the desired 5.10), or between $\rho_t(x, v, t)$ and $\rho_s(x, v, t)$, equation (II 6.3), are also valid provided we interpret all velocities to be the actual velocities of cars at any instant rather than the desired speeds and all velocity distributions to be the distributions of the actual velocities at the point (x, t). It might even be reasonable to postulate that cars leaving an ideal parking lot at $x = 0$ have a Poisson distribution (no coincidences) with independent desired (or actual) speeds.

This could be true if there were no interference between cars in leaving the lot; i.e., the lot is like a highway with free passing.

Any equations describing the evolution of $k(x, t)$, $q(x, t)$, $\rho_s(x, v, t)$, etc. will be changed, however, because the equations of motion are now different. The Poisson distribution, for example, is no longer invariant since for any $x > 0$ there will be coincidences between single cars. The equations describing the evolution are structurally much more complicated. The $\rho_s(x, v, t)$ is no longer determined by its values at $t = 0$ in (II 6.4). To follow the motion of even a single car for a time t one must know at least the number of passings in time t which in turn is not uniquely determined by the average densities for single cars, but must be determined from the joint distribution for positions of all slower vehicles which the one in question must pass.

If a system contains only a few cars, we can follow in detail the motion of each car. If the system contains a very large (infinite) number of cars and we observe it over very long periods of time, we can also apply some asymptotic approximations to obtain equilibrium distributions. An exact analysis of any large but finite system, however, is very tedious. The evaluation of the equilibrium distributions for an infinite system [6] is easier than the analysis of distributions for small numbers of cars (even two), so we will consider the equilibrium distributions first.

As a preliminary exercise to illustrate the type of arguments to be used, we consider first the following very artificial situation. Along a highway at stations with positions y_1, y_2, ... , we place some stationary cars. A second set of cars, all with the same velocity v, now

enter the highway at $x = 0$ and at times t_1, t_2, ... (see figure III 3). When the ith moving car reaches the stationary car at the jth station, it is delayed a time τ_{ij}.
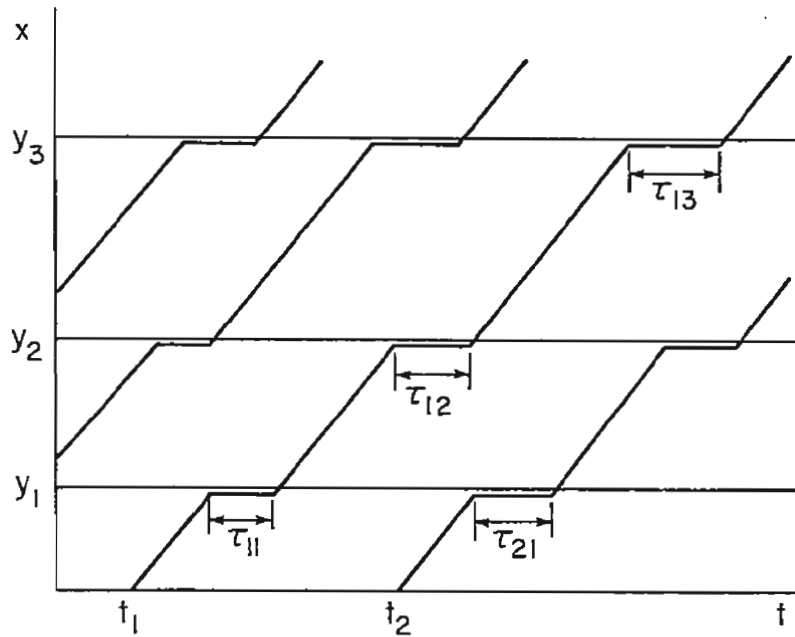

Fig. III 3
Fast cars pass stationary cars

We assume that the average interdeparture times $E\{t_{j+1} - t_j\}$ is large compared with the average delay $E\{\tau_{ij}\}$. Occasionally two or more cars will be at the jth station simultaneously but such events will produce no first order effects as long as the delays resulting from such interactions are finite. It may, however, be convenient to imagine that when such

from such interactions are finite. It may, however, be convenient to imagine that when such multiple interactions do occur, each of the cars is delayed at the jth station as if the others were not there. (In the terminology of queueing theory, each station acts like an ∞-channel server for the moving cars).

For the present model, the delays suffered by car j are statistically independent of those by car k, k $\neq$j. The arrival time of car i at station y_j is

93

$$t_{ij} = t_i + (y_j/v) + \sum_{k=1}^{j-1} \tau_{ik} \quad ,$$

its departure time from x = 0, plus the travel time to the jth station, plus the sum of all

delays at stations 1 to j - 1. For large j, the distribution of t_{ij} for fixed y_j and t_i should,

according to the central limit theorem, be approximately normal with mean

$$E\{t_{ij}\} = t_i + (y_j/v) + (j - 1)E\{\tau_{ij}\}$$

and variance

$$Var\ t_{ij} = (j - 1)\ Var\ \{\tau_{ij}\}.$$

This model is very similar to that discussed in sections II.8. There is little difference

between the random velocity fluctuation discussed there and the random delays due to

passing. The joint distribution of arrival times t_{ij} at y_j , given the t_i , is asymptotically joint

normal for j → ∞ with independent t_{ij}. For a fairly general class of distributions for the

initial times t_i such as discussed in sections II.7 and II.8, the process of arrival times at station

j will for j → ∞ approach that of a homogeneous Poisson process. The heuristic argument is

that in any finite interval of time τ, there is a small probability for any given i^{th} car to

arrive at station j during the time τ (the standard deviation of t_{ij} will be large compared

with τ) but the probability of one car arriving in τ is essentially independent of any others

because of their independent Brownian motions generated by the τ_{ij} .

with τ) but the probability of one car arriving in τ is essentially independent of any others

because of their independent Brownian motions generated by the τ_{ij} .

Although the equilibrium process of arrivals at station j or in fact any position x will

be a homogeneous Poisson process, the spatial distribution of cars is not exactly a homoge-

neous Poisson process because they tend to cluster at the points y_j. The number of cars in

any spatial interval (x, x + ξ) which does <u>not</u> contain any of the y_j will have a Poisson

94

distribution according to the same argument used near the end of section 11.7. If the mean flow of cars is q, the mean number in (x, x + ξ) will be qξ/v. The number of cars <u>at</u> y_j will also have a Poisson distribution. This one can prove directly by actually computing from the Poisson distribution of arrivals and the independent delay times, the probability that k cars arrives at y_j at such times as to still be at y_j at some time t. The simplest argument, however, is to appeal again to the origin of the Poisson distribution. If we have a large (∞) number of cars in the system, the probability of any specified cars being at station y_j at time t is small (zero) and independent of whether there are any others there. These are just the circumstances under which a Poisson distribution arises for the number of such events. The spatial process of all cars is a non-homogeneous Poisson process.

Each station y_j acts like an ∞ - channel service facility with Poisson arrivals. That the equilibrium queue length for such a system has a Poisson distribution has been derived before in many different ways [7 - 10] in many different contexts. The mean of the Poisson distribution is most easily calculated from the observation that over a long period of time the average total delay per unit time to all cars at the jth station is given by

total delay per unit time = mean number of cars at the jth station

$$= \textit{average delay per car} \times \textit{number of arrivals per unit time} \qquad (3.1)$$

$$= qE\{\tau_{ij}\}.$$

$$= \textit{average delay per car} \times \textit{number of arrivals per unit time} \qquad (3.1)$$

$$= qE\{\tau_{ij}\}.$$

This relation is also valid for much more general queueing systems [11].

Our interest in this queueing problem is, however, confined mainly to the case $qE\{\tau_{ij}\}$ << 1 because the model of an ∞- channel server at each station is realistic only if the

probability for more than one car to be at a station is negligibly small. The Poisson distribution of cars at y_j is, to first order in q, a distribution on just 0 or 1.

$$P\{no\ car\ at\ y_j\} \ = \ 1 \ - \ qE\{\tau_{ij}\} + O(q^2)$$

$$P\{one\ car\ at\ y_j\} \ = \ qE\{\tau_{ij}\} + O(q^2) \qquad\qquad (3.2)$$

$$P\ \{more\ than\ one\ car\ at\ y_j\} \ = \ O(q^2)$$

for $q \rightarrow 0$.

It also follows from the Poisson limit theorem of section II.4 that the number of cars at stations y_1, y_2, ... , for given values of the y_j , are independent of each other with a multiple Poisson distribution and independent of the Poisson process of cars not at points y_j . These are all conditional distributions, however, for given y_j and these conditional distributions depend upon the y_j through the fact that there is a nonzero probability of a fast car being at points y_j. If the y_j are themselves random, the complete distribution of all cars is obtained by multiplying these conditional distributions with the distribution for the y_j.

Of particular interest is the case in which the points y_j themselves define a Poisson process of density k_0. The equilibrium spatial distribution of all the cars with desired speed v will at any time t consist of the superposition of two statistically independent processes (1) a Poisson process of free cars with spatial density q/v, and (2) a Poisson process of points y_j at each of which there is a random Poisson distributed number of cars trying to pass the Poisson process of free cars with spatial density q/v, and (2) a Poisson process of points y_j at each of which there is a random Poisson distributed number of cars trying to pass the stationary ones, the number at each y_j being statistically independent of the numbers elsewhere. The superposition of these two processes is not a Poisson process because of the possible multiple occurrences of cars at the y_j points (the sum of a Poisson distributed number of independent Poisson distributed random variables is not Poisson distributed). If,

96

however, q is small enough that we can neglect queues of two or more cars at points y_j and treat the points y_j as points where 0 or 1 of the fast cars can be found, then the second process above will also be a Poisson process (the proof is left as an exercise). The combined process will be a Poisson process with spatial density $(q/v) + k_0 qE\{\tau_{ij}\}$.

If now we look at the combined distribution for all cars, those of desired speed v and the stationary ones, we see that to this approximation of 0 or 1 length queues both the cars of desired speed 0 or v define Poisson processes separately but the processes are not independent. The combined process can be interpreted as the superposition of three statistically independent processes, (1) single v-cars with spatial density q/v, (2) single stationary cars with spatial density $k_0 - k_0 qE\{\tau_{ij}\}$, and (3) coincident pairs of cars with a density of pairs $k_0 qE\{\tau_{ij}\}$. Any count of cars in (3) will have a distribution on only the even integers.

The total number of cars of both types observed in any interval of highway is <u>not</u> Poisson distributed. It is, however, the sum of a Poisson process for single cars and another independent Poisson process for coincident pairs, i.e., the sum of a Poisson process on all the integers and another process defined only on the even integers. Distributions of this type occur frequently in applications of probability. They are special cases of <u>compound Poisson distributions</u>, also of <u>infinitely divisible distributions</u> [12]. The process of cars is also a special case of a <u>compound Poisson process</u>, a process of points at each of which there is a random number of items. The births of babies in a hospital is a process of the type considered here, mostly single events, occasional pairs but rarely anything else.

It is also of interest to notice that the marginal distribution for the spacing between consecutive cars at any time t consists of a discrete component at spacing zero plus an

97

exponentially distributed tail. In this model, the actual trajectories of cars were idealized so as to be accurately represented when two cars were outside their range of interaction. The distribution of spacing for cars in a non-idealized model can differ from the above only in that the coincident pairs should be separated by some positive distance depending upon the detailed equations of motion but not by more than the range of the interaction. The distribution outside this range is unaffected by this idealization of the trajectories and will remain, to first order in q, the exponential distribution of the idealized model.

The results derived above for the special case in which there are two types of cars, some with velocity v, some with velocity 0, can be generalized first to the case of two velocities $v^{(1)}$ and $v^{(2)}$, neither of which is zero, then to three or more velocities $v^{(j)}$, and finally to a continuous distribution of desired speeds. Formally one can transform the velocities 0 and v into nonzero velocities $v^{(1)}$ and $v^{(2)}$ by going to a moving coordinate system but a few consequences of this need to be examined. Suppose the slower cars have density $k^{(1)}$, flow $q^{(1)}$ and desired speed $v^{(1)}$; the faster cars density $k^{(2)}$, flow $q^{(2)}$, and desired speed $v^{(2)}$. The trajectories are now as shown in figure III 4.

If the slow cars travel at a perfectly controlled velocity $v^{(1)}$, then any distribution of headways between the slow cars assigned at one time will be preserved for all times. There is no tendency for these cars to acquire an exponential headway distribution. In reality, however, headways between the slow cars assigned at one time will be preserved for all times. There is no tendency for these cars to acquire an exponential headway distribution. In reality, however, it is impossible for drivers to maintain a velocity $v^{(1)}$ with perfect precision or for two cars to keep exactly the same velocity. At best the velocity of any driver is some average $v^{(1)}$ plus a random time series. Any reasonable postulate one might make about the nature of this random part would be sufficient to guarantee that after a long period of time, the positions of the cars
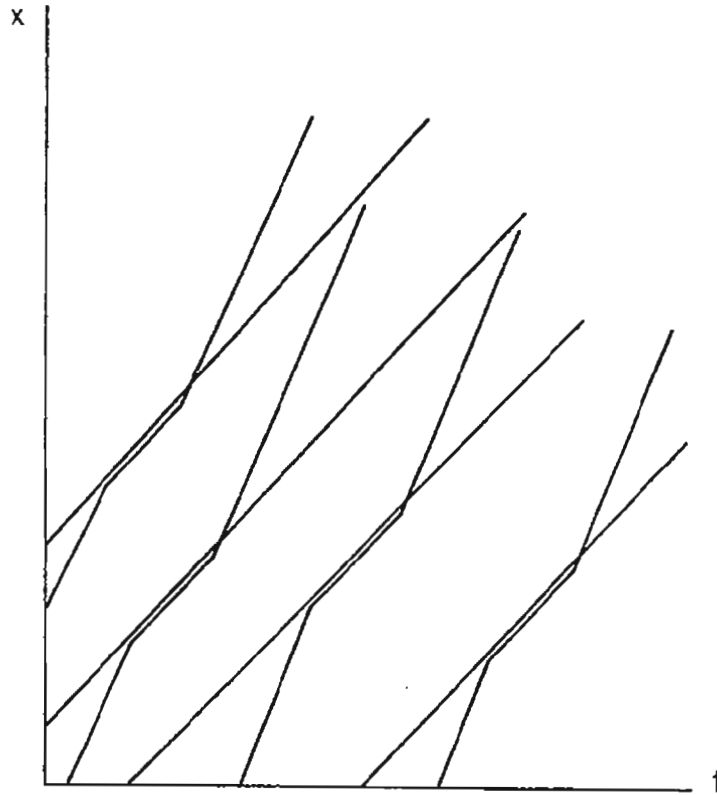
Fig. III 4
Fast cars passing slow cars

of velocity $v^{(1)}$ do become approximately a Poisson process. It is, therefore, reasonable to assign a Poisson process to the $v^{(1)}$ cars initially.

For $v^{(1)} = 0$ we argued that arrival times of the fast cars at any y_j would be a Poisson process, but the spatial distribution of fast cars would be non-Poisson and dependent upon the For $v^{(1)} = 0$ we argued that arrival times of the fast cars at any y_j would be a Poisson process, but the spatial distribution of fast cars would be non-Poisson and dependent upon the y_j. In the present case, $v^{(1)} \neq 0$, neither the spatial distribution nor the arrival times at a fixed point x will be a Poisson process. The exact analogue of the argument for $v^{(1)} = 0$ is to say that if one has some reasonable distribution of times at which the fast cars pass slow car number 1 say, which has a trajectory extending from $t = -\infty$ (thus for $x = -\infty$ to $+\infty$), then

99

the times at which these fast cars pass the jth slow car should for j $\rightarrow \infty$ approach a Poisson process. If all trajectories start at x = 0, we can still reach a similar conclusion provided the slow cars also have some reasonable (Poisson, for example) distribution of starting times so as to guarantee that fast cars that start later and later after the slow car number 1 and which must therefore pass more and more cars before reaching this slow car, will have some more or less stationary distribution of times for passing car 1.

We could also argue that if for sufficiently large x we ask for the <u>conditional</u> probability that j fast cars will cross x in some time τ or be in coincidences with some given slow car, or satisfy some other such condition, given the trajectories of the slow cars, these should all satisfy an appropriate Poisson distribution provided the events in question are rare events for any specified fast car. Again we appeal to the argument that fast cars which suffer many statistically independent passing delays before reaching x will have essentially independent probabilities for satisfying some condition at x.

The equilibrium process of times at which fast cars pass a given slow car should therefore be a Poisson process; i.e., the time intervals between passings are independent and exponentially distributed. The equilibrium number of fast cars crossing any fixed point x in any time τ or the number of fast cars in any interval of highway of length ξ at time t, given the number of slow cars in τ or ξ have a Poisson distribution. The crossing times or any time τ or the number of fast cars in any interval of highway of length ξ at time t, given the number of slow cars in τ or ξ have a Poisson distribution. The crossing times or the positions of cars, however, do not define homogeneous Poisson processes, the mean number in τ or ξ are not exactly proportional to τ or ξ but depend upon the number of slow cars in τ or ξ.

If we neglect the events that two or more fast cars be in coincidence with the same slow car, we can again say that for a Poisson distribution of slow cars the crossing time distribution or the spatial distribution of cars is approximately (to first order in k) a superposition of three independent Poisson processes (1) a distribution of single fast cars, (2) a distribution of single slow cars, and (3) a distribution of coincident pairs of one fast and one slow car.

To complete the analysis, we need only compute the means of the various Poisson distributions.

The slow cars travel freely so their distributions are as described in Chapter II. In particular, we have $q^{(1)} = v^{(1)}k^{(1)}$. Over a long period of time T the fast car will travel a distance $u^{(2)}T$ if $u^{(2)}$ is the time average velocity. The number of fast cars in the length of highway $u^{(2)}T$ will be the number that entered in time T, i.e., $q^{(2)}T$. Therefore

$$q^{(2)} = k^{(2)}u^{(2)}. \tag{3.3}$$

The average velocity $u^{(2)}$ is obtained by observing that over a long time T, the distance traveled $u^{(2)}T$ is the distance the car would travel if there were no passings, $v^{(2)}T$, less the average distance loss due to passings. The fast cars gain on the slower ones at an average relative velocity $u^{(2)} - v^{(1)}$. The average number of passings is therefore $(u^{(2)} - v^{(1)})k^{(1)}T$. Thus

$$u^{(1)} = v^{(2)} - (u^{(2)} - v^{(1)})k^{(1)}E\{d_{21}\},$$

relative velocity $u^{(2)} - v^{(1)}$. The average number of passings is therefore $(u^{(2)} - v^{(1)})k^{(1)}T$. Thus

$$u^{(1)} = v^{(2)} - (u^{(2)} - v^{(1)})k^{(1)}E\{d_{21}\},$$

or

101

$$u^{(2)} = v^{(2)} - \frac{(v^{(2)} - v^{(1)}k^{(1)})E\{d_{21}\}}{1 + k^{(1)}E\{d_{21}\}} .$$

In the present approximations we wish to keep terms only to first order in k so it is proper that we discard the denominator term above and let

$$u^{(2)} = v^{(2)} - (v^{(2)} - v^{(1)})k^{(1)}E\{d_{21}\} + O(k^2) \tag{3.4}$$

which is a special case (for discrete velocities) of (2.2).

Fast cars will overtake a given slow car at a rate of $k^{(2)}(u^{(2)} - v^{(1)})$ and each fast car holds the velocity $v^{(1)}$ until it has lost a distance d_{21}, thus an average time $E\{d_{21}\}/(v^{(2)} - v^{(1)})$. The average number of fast cars in coincidence with a slow car is therefore

$$\frac{k^{(2)}(u^{(2)} - v^{(1)})E\{d_{21}\}}{v^{(2)} - v^{(1)}} \sim k^{(2)}E\{d_{21}\} .$$

For low flows (to order k) we have a Poisson process of pairs having spatial density $k^{(1)}k^{(2)}E\{d_{21}\}$ and flow (of pairs)

$$q^{(1)}k^{(2)}E\{d_{21}\} = q^{(1)}E\{d_{21}/v_2\} = q^{(1)}q^{(2)}E\{\tau_{21}\}. \tag{3.5}$$

The free fast cars have a spatial density $k^{(2)} - k^{(1)}k^{(2)}E\{d_{21}\}$ and a flow

$$v^{(2)}k^{(2)} - v^{(2)}k^{(1)}k^{(2)}E\{d_{21}\}.$$

The single slow cars have a spatial density $k^{(1)} - k^{(1)}k^{(2)}E\{d_{21}\}$ and a flow

$$v^{(1)}v^{(1)} - v^{(1)}v^{(1)}v^{(2)}E\{d\}$$

The single slow cars have a spatial density $k^{(1)} - k^{(1)}k^{(2)}E\{d_{21}\}$ and a flow

$$v^{(1)}k^{(1)} - v^{(1)}k^{(1)}k^{(2)}E\{d_{21}\}.$$

The marginal distribution of spacing between consecutive cars still has a discrete component at spacing zero plus an exponential distribution. The probability that a car is the first car of a coincident pair and therefore has spacing zero is

102

$$p = \frac{k^{(1)}k^{(2)}E\{d_{21}\}}{k^{(1)} + k^{(2)}} \quad .$$

The distribution function for the spacing between consecutive cars is

$$P\{spacing < y\} = p + (1 - p)\{1 - exp[\ -k(1 - p)y]\}\ ,\ y > 0. \tag{3.6}$$

The presence of an interaction between cars, $p > 0$, causes the formation of pairs which

depletes the stream of free cars. As p increases, the amplitude of the exponential part of the

distribution decreases; but so does the rate of decay. For sufficiently large spacing, the

number of large spacings also increases. Figure III 5 shows a comparison of the distribution
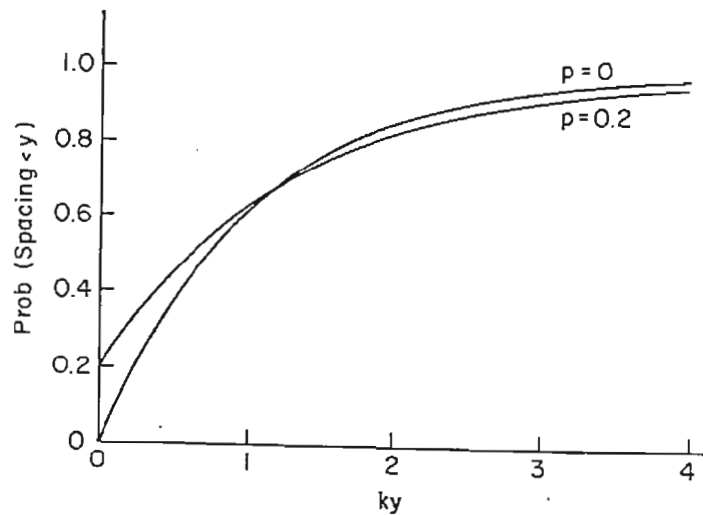
functions for $p = 0$ and $p > 0$.

Fig. III 5
ky

Fig. III 5
Distribution function for spacing

We can now generalize the above to the case of three or more velocities. Suppose we

now add some cars with density $k^{(3)}$, flow $q^{(3)}$ and velocity $v^{(3)} > v^{(2)}$. These cars will not

disturb the motion of the slower cars provided we neglect simultaneous interactions involving

three or more cars (in which a $v^{(3)}$-car while passing a $v^{(1)}$-car could delay a $v^{(2)}$-car). Also, if we neglect the three-car interactions, the delays to the $v^{(3)}$-car is the sum of the independent delays due to passing the $v^{(1)}$-cars and the $v^{(2)}$-cars.

By essentially the same arguments as above, we conclude that the equilibrium conditional probabilities that j of the $v^{(3)}$-cars will be in some interval of highway or cross in some interval of time, etc., given the trajectories of the $v^{(1)}$ and $v^{(2)}$ cars, will all be Poisson distributed with appropriate means depending perhaps upon the positions of the slower cars. If the slower cars are assigned their equilibrium probability distributions as described above, we also conclude that to first order in the interaction, the combined traffic stream can be represented as a superposition of six independent Poisson processes, processes for the single cars of velocities $v^{(1)}$, $v^{(2)}$, or $v^{(3)}$ plus processes for coincident pairs having desired speeds $v^{(1)}$ and $v^{(2)}$, $v^{(1)}$ and $v^{(3)}$, or $v^{(2)}$ and $v^{(3)}$.

To first order in the interaction, the mean spatial densities for the various pairs will be the same as they would have been in the absence of the cars having velocity different from those of the paired cars in question; i.e., the pairs will have spatial densities $k^{(1)}k^{(2)}E\{d_{21}\}$, $k^{(1)}k^{(3)}E\{d_{31}\}$, and $k^{(2)}k^{(3)}E\{d_{32}\}$ independent of $k^{(3)}$, $k^{(2)}$, and $k^{(1)}$ respectively. The densities of free cars are obtained by simply subtracting away the densities of cars caught in pairs, thus the density of free $v^{(3)}$-cars is
free cars are obtained by simply subtracting away the densities of cars caught in pairs, thus the density of free $v^{(3)}$-cars is

$$k^{(3)} - k^{(1)}k^{(3)}E\{d_{31}\} - k^{(2)}k^{(3)}E\{d_{32}\}.$$

The time-average velocity of the $v^{(3)}$-cars will be equal to the spatial mean velocity, the analogue of the $E_s\{v\}$ in chapter II; i.e., the average of $v^{(1)}$, $v^{(2)}$ and $v^{(3)}$ weighted according to the densities of $v^{(3)}$-cars that are in coincidence with $v^{(1)}$-cars, $v^{(2)}$-cars or free. Thus

104

$$u^{(3)} = v^{(3)} - (v^{(3)} - v^{(2)})k^{(2)}E\{d_{32}\} - (v^{(3)} - v^{(1)})k^{(1)}E\{d_{31}\}$$

which again is consistent with (2.2).

The extension of this to any finite number of velocities $v^{(i)}$ is obvious. The extension to arbitrary discrete or continuous velocity distribution can then be obtained since any continuous distribution can be approximated by a discrete distribution.

If we have a total density k and a continuous distribution of desired speeds with probability density $f_s(v)$, the density of cars with velocity between v and $v + dv$ will be $kf_s(v)dv$ (the analogue of the $k^{(i)}$ above). The spatial density of pairs of cars, one of which has a desired speed between v and $v + dv$, the other a desired speed between v' and $v' + dv'$, $v' < v$, is

$$k^2 d(v, v')f_s(v)f_s(v')dvdv' \tag{3.7}$$

(the analogue of the $k^{(i)}k^{(j)}E\{d_{ji}\}$ above) where $d(v, v')$ is the expectation of the distance loss for cars of velocity v passing cars of velocity v' as in equation (2.2).

These pairs will be randomly and independently distributed. The totality of all pairs with the slower car of any pair having a velocity v' in some velocity set A and the faster car a velocity v in some velocity set B will define a Poisson process with spatial density

$$\int_{v \in B, \, v' \in A, \, v' < v} dv \int dv' k^2 d(v, v')f_s(v)f_2(v') . \tag{3.8}$$

$$\int_{v \in B, \, v' \in A, \, v' < v} dv \int dv' k^2 d(v, v')f_s(v)f_2(v') . \tag{3.8}$$

A car of velocity v has a probability $kd(v, v')f_s(v')dv'$ of being in coincidence with a car of velocity v', $v' < v$ and therefore having an actual velocity v' instead of v. The average velocity $u(v)$ is, therefore, as already derived in (2.2).

$$p' = \int_0^\infty dv \int_0^v dv' q\tau(v,v') f_t(v) f_t(v') \qquad (3.12)$$

that the next arrival at x is a pair. This is the analogue of (3.9). In general $p \neq p'$.

$$P\{headway < t\} = p' + (1 - p')\{1 - exp[-q (1 - p')t]\}. \qquad (3.13)$$

It should be pointed out again that in the model considered here, we have represented interacting cars by coincidences and evaluated all distributions to only first order in the interaction. The form of the headway distribution (3.13) is, therefore, expected to be an accurate description (to first order) of the actual distribution of headways larger than the duration of an interaction between two cars. It is not correct for short times except in the sense that any excess probability of the true distribution from an exponential has been reassigned to headway zero. It is, of course, not possible to derive the correct distribution at short headways unless we have a model which includes a detailed description of how cars interact during and just before the passing maneuver.

It is possible to extend the present theory to the non-idealized trajectories and thereby obtain the headway distribution (to first order in k) for short headways as well as long ones. The formulas, however, would contain such things as the mean actual velocity of a car with desired speed v when it is a distance x from a car with speed v' which it is about to pass (in the idealized model, this velocity is v for x > 0 and v' for x = 0). It is not obvious that desired speed v when it is a distance x from a car with speed v' which it is about to pass (in the idealized model, this velocity is v for x > 0 and v' for x = 0). It is not obvious that such formulas would be of much practical value, however, because it would be easier to measure headway distributions than to measure properties of the trajectories from which these distributions could be calculated.

There is an abundant supply of experimental data on headway distributions. People have been collecting such data at least since the 1940s, if not earlier. Equipment exists which will automatically record both velocities and headways of cars passing some point on the highway. Numerous conjectures have been made regarding the form of these distributions, and several people have tried to fit various formulas to these data, Γ-distributions, translated exponential, combinations of normal and exponential, linear combinations of exponentials, etc. Invariably the formulas used have exponential tails because this seemed plausible even if it could not be derived. That the tail of the distribution is indeed exponential has been verified in essentially all cases, often with exceptional accuracy, well beyond that which could be derived from the present theory.

Most experimental data relating to headway distributions is for traffic densities which we would interpret here as moderate densities and is, therefore, relevant both to the present chapter and to chapter V. Several papers give fairly extensive reviews of the literature [13 - 15]. We will comment here only on a few points that are relevant to the present theory.

Schuhl [16] proposed that the probability density of the headway distribution should be approximately a linear combination of a relatively long range exponential plus a short range displaced exponential. Some further analysis of such a model has also been made by Petigny [17]. Schuhl's formula differs from (3.13) in that the first term p′ representing the range displaced exponential. Some further analysis of such a model has also been made by Petigny [17]. Schuhl's formula differs from (3.13) in that the first term p′ representing the distribution of paired cars at zero spacing, is replaced by a term of the form

$$p' \ exp \ [- \beta(t - t_0)] \ for \ \ t > t_0$$

$$0 \qquad\qquad\qquad t < t_0$$

for appropriate values of β and t_0 (the translated exponential). This was not derived from any mathematical model, but was intended to represent the distribution of headways for cars that are paired. The fit to experimental data was exceptionally good. Furthermore, the headway distribution is almost equivalent to (3.13) for headways larger than $t_0 + O(1/\beta)$.

Several people have suggested the use of a single translated exponential for the whole range of headways, on the grounds that for long spacing the headways are nearly exponential, but very short headways should be rare because cars cannot be on top of each other. This gives less satisfactory results and clearly is not quite the correct form. Equation (3.3) implies that if we plot

$$log\ P\{Headway > t\}\ =\ log\ (1 - p') - q(1 - p')t$$

vs. t , one should get a straight line, at least outside the range of interaction. The slope should be less than q and the $t = 0$ intercept $log\ (1 - p')$ should be negative because $p' > 0$. If one does the same thing with a translated exponential distribution, however, one obtains a slope greater than q and a positive intercept. In the latter distribution there is a deficiency of cars within the entire range of interaction, whereas there should be an excess as compared with a pure exponential. The extrapolation of the tail distribution back to $t = 0$ is probably as good a measure as any of the amount of pairing. The more customary method of identifying paired cars is to count all cars with headways less than some specified headway.
measure as any of the amount of pairing. The more customary method of identifying paired cars is to count all cars with headways less than some specified headway.

We have shown in this section, that, for small k, the traffic stream can be represented as a superposition of statistically independent processes, one for pairs and one for single cars. An alternative description of the traffic is to say that the spacings or headways are independent identically distributed random variables with a distribution function of the form

109

(3.10) or (3.13); i.e., the spacing is either zero or exponentially distributed. These two representations are not <u>exactly</u> equivalent but they are equivalent to within the order of approximation considered here. The latter description, for example, admits the possibility of two consecutive zero spacings (thus a triple of interacting cars) with probability p^2 or p'^2. We, however, neglect probabilities of this order in the former description.

In the literature on traffic theory and queueing theory there are many papers in which a traffic stream of moderate density is represented mathematically as a Poisson process of traveling queues [18 - 20], a renewal process (independent headways), or a sequence of alternating blocks and gaps [21]. The motivation for such postulates, however, was not that it was necessarily realistic but that it was mathematically tractable for the particular application under consideration. Although we have confirmed here that these are justifiable postulates for low density traffic, it is not possible to prove, in general, that any of these are necessarily correct for moderate density, including multiple car interactions. What happens in the next approximation depends much more critically upon the stochastic properties of the passing mechanism and is likely to give rise to a very complicated stochastic structure for headways. These other possible stochastic structures for traffic will be discussed further in chapter V.

The theory of this section determines not only the stochastic properties of the headways, it gives the complete stochastic properties of the trajectories including velocities as
The theory of this section determines not only the stochastic properties of the headways, it gives the complete stochastic properties of the trajectories including velocities as well as headways. Although most random variables or processes which we discussed involved statistical independence of almost everything, one important exception is the desired speeds of two cars of a pair. Equation (3.7) or (3.11), in effect, give the joint probability densities for the two velocities and it is in general not a product of marginal densities. Since the function

110

d(v , v') is not determined by the theory, it is not possible to make very meaningful comparisons with experimental data on joint distribution of velocities for consecutive cars. Yet this dependence is an important consideration in the comparison of alternative models of stochastic structure.

4. <u>Velocity fluctuations</u>. The last section dealt with only models for which cars maintain constant velocities except while passing. A more realistic model would result if we assumed that cars behave as described in section II.8 between passings. Their velocities while free would then be represented by a stationary time series

$$v_j(t) = v_j + \eta_j(t)$$

as in II 8.3 with $x_j(t)$ perhaps having the properties of a Brownian motion as discussed in II 8.6.

For models with passing delays the existence of random fluctuations in velocity was not crucial in the arguments relating to the existence of an equilibrium distribution, as they were in Chapter II, because the passing delay was chosen to be random and gave a mechanism for entropy increase.

The addition of fluctuations, however, introduces the possibility that the same pairs of cars may pass each other more than once. Two cars with exactly the same desired speed may
The addition of fluctuations, however, introduces the possibility that the same pairs of cars may pass each other more than once. Two cars with exactly the same desired speed may wish to pass if one car has a larger $\eta_j(t)$ than the other. At some later time the fluctuations might be reversed so that the car which had passed the other is overtaken by the one it passed. In a traffic stream where most drivers try to travel at the speed limit and therefore have a small variance of desired speeds, this phenomenon is likely to occur quite frequently.

Suppose that the relaxation time τ_0 for the fluctuations is large compared with the times required to perform the passing so that once a driver has started to pass another car, he does not change his mind in the middle of the process. Although the actual velocity of a car will be changing during the passing maneuver itself, we assume that the velocity (with fluctuation) of the passing car returns after the passing to the same value it had before. Over times of the order of the passing time, cars therefore behave as if their desired speeds were the speeds $v_j + \eta_j(t)$ at the time in question. In calculating the delays, we shall also imagine that the time losses τ_{ij} or distance losses d_{ij} depend only upon the velocities $v_j(t)$ at the time of passing rather than the desired speeds v_j. (This avoids problems of notation that might arise when a car of desired speed v_i may actually try to pass one of desired speed $v_j > v_i$ because the fluctuations cause $v_j + \eta_j(t)$ to be less than $v_i + \eta_i(t)$.)

We again take a density k sufficiently low that passings are rare and, to a lowest approximation, the cars have nearly the Poisson distribution they would have without the interactions.

Despite the obvious differences in the stochastic structure for the passing processes now as compared with that of the previous sections, average passing rates and average delays can be evaluated by essentially the same techniques used in section III 2, except that we should use velocity distributions for the actual speeds (with fluctuation) rather than the can be evaluated by essentially the same techniques used in section III 2, except that we should use velocity distributions for the actual speeds (with fluctuation) rather than the desired speeds. As the analogue of (2.1) we can say that the rate at which cars of actual speed v overtake of those with speed between v' and $v' + dv'$ with $v' < v$ is, to first order in k,

$$(v - v')kf^*(v')dv'$$

in which $f^*(v)$ is the probability density of the speeds $v(t) = v + \eta(t)$.

112

It is the f*(v) which is usually evaluated in experiments on velocity distributions. To find $f_s(v)$, in the presence of velocity fluctuations, it would be necessary to observe the long time average velocities of individual drivers, rather than their instantaneous values. Since $E\{\eta(t)\} = 0$ for cars of each desired speed v, the variance of v(t) is

$$Var\{v+\eta(t)\} = E\{(v+\eta(t) - E\{v\})^2\} = E\{(v-E\{v\})^2\} + E\{\eta^2(t)\} = Var\ v + E\{Var(\eta(t)|v)\}$$

in which the last terms is the variance of $\eta(t)$ for given v, averaged over all v. Thus the variance of v(t) is larger than for v. Since passing rates depend upon a spread or variance of speed, we expect the total passing rates with the fluctuations to exceed those without the fluctuations.

Equation (2.2) gave the time average speed u(v) of a car with desired speed v. The direct analogue of this is an expression for the average distance traveled per unit time of a car whose actual speed (except during passing) is v(t), but a time average over times small compared with τ_0 so that the random velocity v(t) is essentially constant over the time of observation. Thus

$$u(v(t)) = v(t) - k \int_0^{v(t)} dv' d(v(t), v')(v(t) - v') f_s^*(v') \tag{4.1}$$

is the average velocity of a car whose actual speed between passings is v(t) at time t. To obtain the average speed of a car with desired speed v, we must take the expectation of this is the average velocity of a car whose actual speed between passings is v(t) at time t. To obtain the average speed of a car with desired speed v, we must take the expectation of this over the distribution of $\eta(t)$ for fixed v. This is also the long time (over times large compared with τ_0) average speed of a single car with desired speed v. The expectation of (4.1) over $\eta(t)$ does not have a simple form.

The average speed of all cars, however, can be obtained directly from (4.1) by taking the expectation over the distribution of velocities v(t). Since $E\{\eta(t)\} = 0$

$$E\{u\} = E\{v\} - k\int_0^\infty dv f_s^*(v) \int_0^v dv' d(v, v')(v - v') f_s^*(v') + O(k^2) \qquad (4.2)$$

which differs from (2.3) only in that the f_s is replaced by f_s^*.

The equilibrium distributions can still be represented, as in the last section, by Poisson distributions of free cars plus Poisson distributions of pairs. Any subdivision of the population of cars according to their free speeds or their actual speeds or both will have independent Poisson distributions. The densities of the Poisson distributions, in particular the densities of the pairs, however, is determined by the distributions f* of actual speeds but is otherwise unchanged from that described in the last section.

The time dependence of flows, etc., under nonequilibrium conditions is hopelessly complicated to follow in detail but there are some curious effects that can be discussed at least qualitatively.

Suppose two cars with identical desired speeds are in coincidence at time 0 with one car having a velocity fluctuation $\eta_1(0)$, the other a velocity fluctuation $\eta_2(0) \neq \eta_1(0)$. Consider the idealized Brownian motion as discussed in (II 8.6) in which drivers retain their velocities until time τ_0 at which time they pick new η's independent of their past values. Consider the idealized Brownian motion as discussed in (II 8.6) in which drivers retain their velocities until time τ_0 at which time they pick new η's independent of their past values. They again pick new velocities at each time $j\tau_0$, $j = 1, 2, \ldots$. The relative motion is in essence a Brownian motion with relative velocities $\eta_1(t) - \eta_2(t)$ chosen independently at times $j\tau_0$. The relative positions of the cars at time $k\tau_0$ is the sum of these independent relative velocities multiplied by τ_0, i.e.,

114

$$x_1(k\tau_0) - x_2(k\tau_0) = \tau_0 \sum_0^{k-1} [\eta_1(j\tau_0) - \eta_2(j\tau_0)].$$ (4.3)

A second passing between these two cars will occur if the Brownian motion of relative positions crosses zero at some time t > 0. There is an extensive literature on the times between zero-crossing of Brownian motion. This is a problem closely related to, or a special case of, problems variously described as the problem of gambler's ruin, first passage time, absorbing barriers, zero-crossings, etc., which are discussed in most books on probability theory; see for example Feller [12]. There are even some books almost entirely devoted to problems of this type [22, 23].

If, in particular, the random variables $\eta_1(j\tau_0)$ and $\eta_2(j\tau_0)$ are statistically independent and each assumes only two values, say $\pm \eta$, with probability 1/2 each, then each term of (4.3) can be considered as the winnings in a single play of a game in which each player flips a coin and player one wins η times the number of heads (or tails) on the two coins. The entire expression (4.3) can be interpreted as $\tau_0 \eta$ times the cumulative winnings in a series of k plays or in a series of 2k tosses of a single coin. The times between successive passings of the same two cars is the analogue of the number of plays in a game between times when the opponents are even. In this particular model, the players might be even after each of several plays, but once one player has gone ahead, the problem is identical to the "gambler's ruin" problem, to calculate the number of plays until the player has lost his capital.
plays, but once one player has gone ahead, the problem is identical to the "gambler's ruin" problem, to calculate the number of plays until the player has lost his capital.

From the literature on such problems, we conclude the following:

1. With probability one, a second passing will occur within a finite time. In an infinite time the two cars will pass each other infinitely many times. This is true for essentially any distribution of the η_1 and η_2; i.e., the time between passings is a

115

proper random variable. The distribution of this random variable, however, depends upon the detailed properties of the η_1 and η_2.

2. There is a high probability that the second passing will occur within a few multiples of the time τ_0, in fact it is quite likely to occur within a time $2\tau_0$, i.e., the cars will pass again before they have a chance to drift very far apart.

3. The mean time between passings of the same two cars is infinite because the probability for very long passing times does not go to zero fast enough. In this case of equal desired speeds the probability that no passing occurs within a time t is of order $t^{-1/2}$ for $t \to \infty$. If T represents the time between passings, it has a distribution function and

$$1 - F_T(t) = 0(t^{1/2}) \to 0 \text{ for } t \to \infty$$

but

$$E\{T\} = \int^{\infty} t dF_T(t) = 0 \left[\int^{\infty} t \frac{dt}{t^{3/2}} \right] = \infty \ .$$

The reason this happens is that if the two cars have not met again after some fairly long time, they are likely to have drifted far apart. Although they are certain to meet again eventually, it may take a very long additional time.

The analogue of this property for the special case of coin tossing is discussed in most meet again eventually, it may take a very long additional time.

The analogue of this property for the special case of coin tossing is discussed in most clementary books on probability theory. It can be derived from the fact that over long times, $k \gg 1$, the cumulative sum in (4.3) is approximately normally distributed (the central limit theorem). That these sums are approximately normal is, however, true for essentially any

116

distribution of the η's (provided the first and second moments of the η's are finite, which is certainly true for velocities of cars).

If, instead of having just two cars, we have a low density spatial distribution of cars all with the same desired speed, then if a passing pair of cars fails to meet again within some reasonable length of time, they are likely to have drifted far apart and will encounter other cars. The time until a given car passes <u>some</u> other car is a random variable with finite expectation. For a uniform density of cars we have already concluded that the mean rate at which a given car passes other cars is proportional to the density k. This itself guarantees that the time between passings of cars has a finite expectation proportional to $1/k$. It becomes infinite for $k \to 0$ which implies that it must also be infinite for a single pair of cars as in 2 above.

Although everything appears to be consistent, the fact that the mean time between passings for a car in a stream of cars with identical desired speeds is proportional to $1/k$ comes about in a peculiar way. If car 1, say, has just passed a car 2 then the mean time until it meets car 2 again is infinite. Therefore this cannot by itself explain the mean time of order $1/k$. If we label as car 3 the car that car 1 is aiming for next after passing 2, cars 1 and 3 are initially a distance of order $1/k$ apart. The mean velocities of cars 1 and 3 are identical, however. If they are to meet, it will be because of the fluctuations which cause the relative initially a distance of order $1/k$ apart. The mean velocities of cars 1 and 3 are identical, however. If they are to meet, it will be because of the fluctuations which cause the relative position to drift with an uncertainty proportional to $t^{1/2}$. It is likely, therefore, to take car 1 a time of order k^{-2} to meet 3, which for $k \to 0$ is also too long a time to explain a passing rate proportional to $1/k$.

117

What happens is that after car 1 passes car 2, it is quite likely to pass 2 again within a short time (of order 1 relative to k). The distribution function $F_T(t)$ therefore rises quite rapidly (independent of k), but then starts to form the slowly decaying tail which would, if continued, cause $E\{T\}$ to be infinite. The presence of other cars, however, in effect cuts off the tail after t has become of order $1/k^2$. Car 1 is likely to pass 2 very many times before it drifts away. If or when it encounters some other car, it is likely to pass it many times, etc.

If cars have different desired speeds, the relative position between two cars is a Brownian motion with a net drift (like a gambler's winnings in a biased game). If, perchance, a car of desired speed v_1, should pass another with a desired speed $v_2 > v_1$ (but with $v_2 + \eta_2(t) < v_1 + \eta_1(t)$), then these two cars are certain to meet again. The time until they meet also has a finite expectation and is likely to occur soon after the time τ_0. If, however, $v_2 < v_1$, the two cars might meet again within a short time but with positive probability they will drift apart and never meet. With probability one, they will meet only finitely many times before the car with the higher v_j pulls away for good.

Although most of the properties described above were discussed in terms of an artificial discrete time model, they also hold for quite general types of random processes $\eta(t)$ as discussed in section II for 8 fluctuations of the second type; i.e., processes for which $\eta(t)$ is stationary but the position or integral of η has a variance that increases linearly with t. as discussed in section II for 8 fluctuations of the second type; i.e., processes for which $\eta(t)$ is stationary but the position or integral of η has a variance that increases linearly with t. The key property of the sum (4. 3) or the integral of a continuously varying $\eta_1(t) - \eta_2(t)$ is that over long periods of time, the sum or integral should satisfy a central limit theorem. This should be true for the real life processes.

5. <u>Time dependent flows</u>. To study time dependent flows particularly with finitely many cars, we return to the model of section (3.3), without the velocity fluctuations. If we know the entrance times of cars on the highway and their velocities, or if we know a probability distribution for such state variables, then the equations of motion determine the same quantities at any other position $x > 0$. We wish here to investigate how the probability distribution of state variables changes with x. To do this we must follow the trajectories of cars. We can do this by any one of the three schemes I, II, or III of section (III.3).

To follow trajectories according to scheme I (for each x look for the next passing between any two cars) seems awkward because any car could potentially pass any other. At each stage we must inspect all cars that could pass and see which ones really do. This may be convenient to describe the behavior over short times because only cars which are initially close together are capable of passing within a short time. One could perhaps derive some differential equations to describe the x - dependence of the distribution. (This might be useful to analyze equilibrium solutions for infinite systems.)

To obtain the solution over long times (with many passings) either scheme II or III seems more suitable. The procedure would be first to determine the evolution for the distribution of car 1 defined to be the slowest car in scheme II or the first car to depart in scheme III. One then determines the conditional distributions for car 2 given the motion of distribution of car 1 defined to be the slowest car in scheme II or the first car to depart in scheme III. One then determines the conditional distributions for car 2 given the motion of car 1, etc. These sets of conditional distributions then determine the joint distributions.

In either of these two schemes the first step should be to renumber the cars. We will assume that drivers have no identifying properties other than their positions and desired

119

speeds. Two states that differ only in the numbering of the cars will be considered equivalent. If we renumber the cars according to scheme II, it is convenient to define a new probability density ρ_t^*,

$$\rho_t^*(0; v_1, v_2, \ldots , v_n; t_1 t_2, \ldots , t_n) \qquad (5.1)$$

$$= \sum_p \rho_t (0; v_{i_1}, v_{i_2}, \ldots , v_{i_n} ; t_{i_1}, t_{i_2}, \ldots t_{i_n})$$
$$\text{for } v_1 < v_2 < \ldots < v_n$$
$$= 0 \text{ otherwise}$$

where P is any permutation $(i_1, i_2, \text{--}, i_n)$ of the numbers $(1, 2, \text{--} n)$. The quantity ρ_t^* dv_1 -- dv_n dt_1 -- dt_n is now the probability that among the n cars there is one car having a velocity between v_1 and $v_1 + dv_1$ (identified as car 1 in the new numbering), another with velocity between v_2 and $v_2 + dv_2$ etc. and the car with velocity in dv_j departs during the time interval $(t_j, t_j + dt_j)$. The density ρ_t^* vanishes on all but one of the n! parts of the original state space corresponding to various orderings of the velocities.

If we wish to apply the third scheme of numbering it is desirable to use a similar procedure applied to the times t_j rather than the velocities. We can define a probability density $\rho_t^\dagger$, so that $\rho_t^\dagger dv_1$ -- dv_n dt_1 -- dt_n is the probability that among the n cars there is one car with departure time between t_1 and $t_1 + dt_1$, a second with departure time between t_2

~~and t_2 + dt, etc. and the car that leaves in dt has a velocity in~~

one car with departure time between t_1 and $t_1 + dt_1$, a second with departure time between t_2 and $t_2 + dt_2$ etc., and the car that leaves in dt_j has a velocity in dv_j,

$$\rho_t^\dagger (0; v_1, v_2, \ldots, v_n; t_1, t_2, \ldots, t_n)$$

$$= \sum_p \rho_t (0; v_{i_1}, v_{i_2}, \ldots v_{i_n}; t_{i_1}, t_{i_2}, \ldots, t_{i_n}) \qquad (5.2)$$

$$\text{for} \quad t_1 < t_2 < \ldots t_n$$

$$= 0 \quad \textit{otherwise}$$

For any x > 0, the time order of the crossings at x may be different from those at x = 0. For sufficiently large x, the time ordering will, of course, become the reverse of the velocity ordering (the fastest car will reach x first for some sufficiently large x regardless of when it left x = 0). In the 2n dimensional space of velocities and times, the probability density eventually shifts onto one of n! further subdivisions of the space representing the various permutations of the times relative to the velocities.

Both schemes II and III are tedious to apply beyond the first few cars in their respective numberings. Scheme II leads to a quicker analysis of what happens at the rear of a platoon because, at least for large x, the rear of the platoon is determined mainly by the behavior of the slowest cars rather than the ones which started late. These are the low numbered cars in this scheme. Scheme III on the other hand leads to a quicker analysis of what happens at the front of a platoon because the expected first arrival at x must not only be numbered cars in this scheme. Scheme III on the other hand leads to a quicker analysis of what happens at the front of a platoon because the expected first arrival at x must not only be a fast car but must not have been delayed by too many passings. It is therefore likely to have started near the front of the platoon initially and have a low number in this third numbering scheme. We shall pursue only methods II here, however.

121

Starting from a given distribution ρ^*, we can define the marginal distribution for car 1 (the slowest). Since some of the distributions for arrival times will have discrete components, it is convenient to work with distribution functions relative to the time variables. Let

$$F^*_1(x, v_1, t_1)dv_1 = P\{v_1 < \text{velocity of car 1} < v_1 + dv_1,$$

$$\text{and the crossing time at } x \leq t_1\}. \tag{5.3}$$

For $x = 0$ this is determined from ρ_t^*

$$F_1^*(0, v_1, t_1) = \int_{-\infty}^{t_1} dt_1 \int_{-\infty}^{+\infty} \cdots \int dt_2 \cdots dt_n \int_{v_1}^{\infty} dv_2 \int \cdots \int_{v_{n-2}}^{\infty} dv_{n-1} \int_{v_{n-1}}^{\infty} dv_n \, \rho_t^*(0 \, ; \, \ldots) \quad . \tag{5.4}$$

Car I never passes any other. Consequently if it has velocity v_1, it can cross x before time t_1 if and only if it crosses 0 before time $t_1 - x/v_1$. Thus $F_1(x, v_1, t_1)$ is given by

$$F_1^*(x, v_1, t_1) = F_1^*(0, v_1, t_1 - x/v_1) \tag{5.5}$$

and this describes the behavior of car 1.

We consider next the conditional distribution of car 2 given that of car 1. Let

$$F^*_2(x, v_2, t_2 | v_1, t_1)dv_2 = P\{v_2 < \text{velocity of car 2} < v_2 + dv_2,$$

$$\text{the crossing time at } x \text{ of car } 2 < t_2, \text{ given the velocity } v_1 \text{ and} \tag{5.6}$$

$$\text{crossing time at } x, t_1, \text{ of car 1}\}.$$

Figure III 6 shows a possible trajectory for car 1 by the solid line and a series of possible

~~trajectories for car 2 by broken lines. Starting from the left hand side of figure III 6 we see~~

Figure III 6 shows a possible trajectory for car 1 by the solid line and a series of possible trajectories for car 2 by broken lines. Starting from the left-hand side of figure III 6, we see that if car 2 crosses $x = 0$ before car 1, it will not pass car 1. If it leaves just before car 1, it crosses x at the time $t_1 - x/v_1 + x/v_2$ (point a of Figure III 6), but if it leaves just after car I, it must pass car I and will arrive at x a time d_{21}/v_2 later (point b of figure III 6), or at time t_1 if x is so small that $x(1/v_1 - 1/v_2) < d_{21}/v_2$, i.e., the passing is not completed by time t_1. There

are no trajectories in the shaded area on the left-hand side of figure III 6. If car 2 has a trajectory anywhere in the shaded area on the right-hand side of figure III 6, it catches car 1 just before it reaches x and crosses x at the time t_1. For any trajectory to the right of this, there will be no passing before it reaches x.
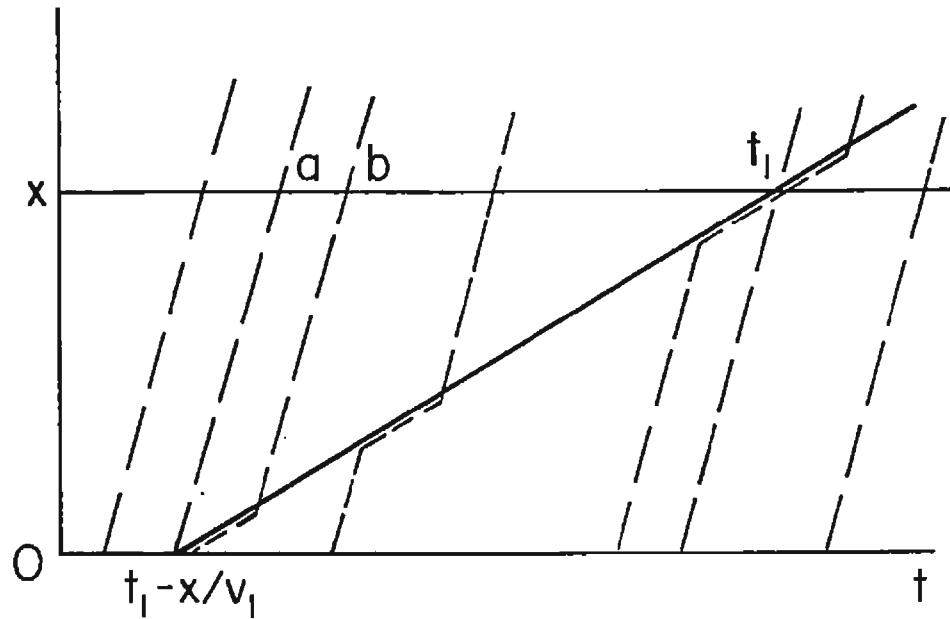

Fig. III 6
Possible trajectories of fast cars for a given slow car

If we choose a time $t_2 < a$, then car 2 will cross x before time t_2 with a velocity v_2 if and only if it crossed 0 before time t_2-x/v_2. The same is true if $t_2 \geq t_1$. Thus we have

If we choose a time $t_2 < a$, then car 2 will cross x before time t_2 with a velocity v_2 if and only if it crossed 0 before time t_2-x/v_2. The same is true if $t_2 \geq t_1$. Thus we have

$$F_2^*(x,v_2,t_2/v_1,t_1) = F_2^*(0,v_2,t_2\text{-}x/v_2/v_1,t_1\text{-}x/v_1)$$

$$\text{if } t_2 < t_1 - x/v_1 + x/v_2 \text{ or } t_2 > t_1. \qquad (5.7a)$$

123

As we have defined F_2^*, the conditions on car 1 are to be specified at the value of x in question. Values v_1, t_1 for car 1 at x are equivalent to values v_1, $t_1 - x/v_1$ at $x = 0$, either of which, of course, uniquely defines the entire trajectory for car 1.

Car 2 cannot cross x between points a and $\min\{b,t_1\}$ of figure III 6, therefore for t_2 in this range F_2^* has the same value as at point a, i.e.,

$$F_2^*(x, \ v_2, \ t_2/v_1, \ t_1) = F_2^*(0, \ v_2, \ t_1 - x/v_1/v_1, \ t_1 - x/v_1) \tag{5.7b}$$

$$if \ t_1 - x/v_1 + x/v_2 < t_2 < \min \ \{t_1 - x/v_1 + (x + d_{21})/v_2 \ , \ t_1\} \ .$$

If d_{21} is considered to be a random variable, then this is actually the conditional distribution given d_{21}. The complete distributions would be obtained from this by multiplying these conditional distributions by the distributions for the d_{21}.

Finally if t_2 lies between points b and t_1, car 2 crosses before time t_2 if and only if it crosses 0 before time $t_2-(x+d_{21})/v_2$. Thus

$$F_2^*(x, \ v_2, \ t_2/v_1, \ t_1) = F_2^*(0, \ v_2, \ t_2 - (x + d_{21})/v_2/v_1, \ t_1 - x/v_1)$$

$$if \ t_1 - x/v_1 + (x+d_{21})/v_2 < t_2 < t_1 \ . \tag{5.7c}$$

This conditional distribution (5.7 a,b,c) at x along with the marginal distribution for car 1 at x, (5.5), determines the joint distributions for cars 1 and 2 at x. For n=2, this is the complete description. The form of F_2^* is shown in figure III 7. It has a discontinuity at $t_2=t_1$.

The distributions for car 3 already become rather complicated because car 3 might complete description. The form of F_2^* is shown in figure III 7. It has a discontinuity at $t_2=t_1$.

The distributions for car 3 already become rather complicated because car 3 might pass either car 2 or car 1 or both. If we define a distribution $F_3^*(x, v_3, t_3|v_1, t_1, v_2, t_2)$ in the obvious way as the conditional distribution given both trajectories for cars 1 and 2, we can at least evaluate this for large or small t_3. Certainly car 3 will not pass either car 1 or 2 if it leaves $x = 0$ ahead of the others or if it leaves so late that it is not able to catch either car
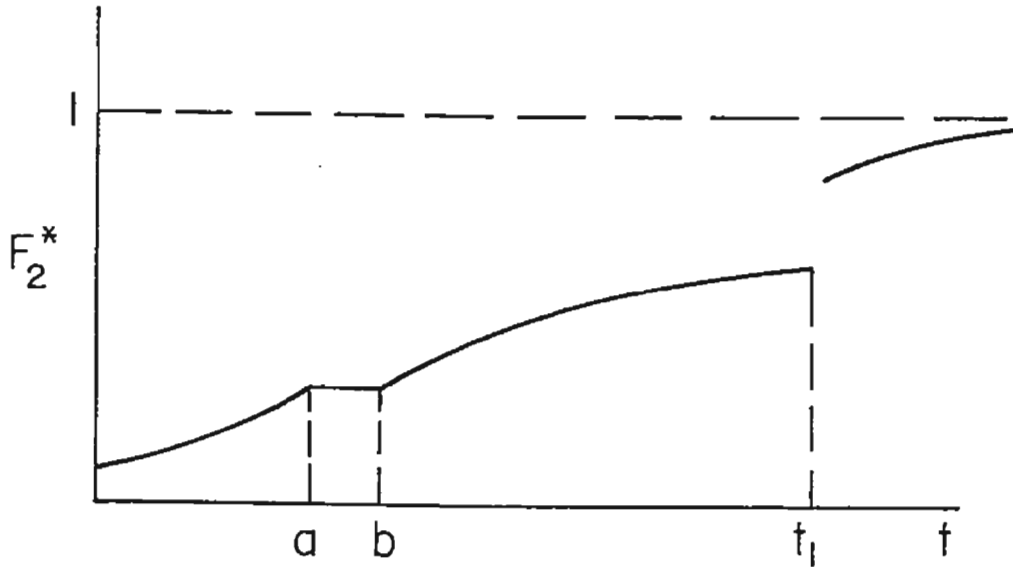
Fig. III 7
Distribution function for passing time

before they reach x. The former is true if and only if t_3-x/v_3 is less than both t_1-x/v_1 and

t_2-x/v_2 (regardless of whether or not cars 1 and 2 interact). The latter is true if $t_3>t_1$ and t_2. In

either of these two cases

$$F_3{}^*(x, v_3, t_3/v_1, t_1, v_2, t_2) = F_3{}^*(0, v_3, t_3{-}x/v_3/v_1, t_1{}', v_2, t_2{}') \qquad (5.8)$$

where $t_1{}'$ and $t_2{}'$ are the times cars 1 and 2 must leave $x = 0$ so as to cross x at time t_1 and t_2

(their values depend upon whether or not cars 1 and 2 pass). For intermediate values of t_3,

F.*   depends upon the configuration of the trajectories for cars 1 and 2.

(their values depend upon whether or not cars 1 and 2 pass). For intermediate values of t_3,

$F_3{}^*$ depends upon the configuration of the trajectories for cars 1 and 2.

Similarly car j passes no other cars if it leaves $x = 0$ ahead of all slower cars

1,2,--,j-1 or if it leaves so long after them that it cannot catch any of the slower cars before it

reaches x. The former is true if

125

$$t_j - x/v_j < \min_{k<j} (t_k - x/v_k) \ . \tag{5.9}$$

The latter is true if

$$t_j > \max_{k<j} t_k \ . \tag{5.10}$$

The quantities on the right hand side of (5.9) are not necessarily all departure times but the minimum of these must be the departure time of the first car among them to leave x = 0 because the first car to leave (among k < j) will not pass any others.

It is straightforward but tedious to write the complete distributions F_j^* for all j. For n cars, there are all together n! possible passing configurations (the initial spatial ordering can be any permutation of the final ordering or vice versa). For moderately large n, this can be come very awkward to disentangle. For very large n, however, one can make some approximations based upon what is typical rather than what is possible.

To illustrate the methods above and the effects of passing delays on diffusion, we consider the simplest examples with only two cars. Suppose that at x = 0 we have (1) the departure times t_1 and t_2 are independent identically distributed random variables with a probability density q(t) and distribution function Q(t), (2) velocities are independent identically distributed random variables with a probability density f(v) and distribution function F(v) independent of the t_j, and (3) the d_{21} is not random for given v_1 and v_2.
ly distributed random variables with a probability density f(v) and distribution function F(v) independent of the t_j, and (3) the d_{21} is not random for given v_1 and v_2.

The joint probability density for entrance times of cars 1 and 2 are

$$p_t(0; \ v_1, \ v_2; \ t_1, \ t_2) = f(v_1)f(v_2)q(t_1)q(t_2). \tag{5.11}$$

The density p_t^* of (5.1) for the cars renumbered according to their velocities is

126

$$\rho_I^*(0 \; ; \; v_1 \; , \; v_2 \; ; \; t_1 \; , \; t_2) \; = \; \begin{cases} 2f(v_1)f(v_2)q(t_1)q(t_2) \\ for \; 0 < v_1 < v_2 \\ 0 \quad otherwise \end{cases} \qquad (5.12)$$

At $x = 0$ the distribution functions F_1^* and F_2^* are

$$F_1^*(0,v_1, \; t_1) \; = \; 2Q(t_1)f(v_1)[1\text{-}F(v_1)]$$

$$F_2^*(0, \; v_2 \; , \; t_2 \; | \; v_1 \; , \; t_1) \; = \; \begin{cases} Q(t_2)f(v_2)/[1 - F(v_1)] \quad for \; v_1 < v_2 \\ 0 \quad for \; v_1 > v_2 \; . \end{cases} \qquad (5.13)$$

The marginal distribution for the second car is obtained by integration of this conditional distribution over the distribution of car one, i.e.,

$$F_2^*(x \; , \; v_2 \; , \; t_2) \; = \; \int\limits_0^{v_2} dv_1 \int\limits_{t_1=-\infty}^{+\infty} F_2^*(x \; , \; v_2 \; , \; t_2 \; | \; v_1 \; , \; t_1)dF_1^*(x \; , \; v_1 \; , \; t_1) \qquad (5.14)$$

where the differential dF_1^* is with respect to the time variable t_1. For $x=0$ this has the value

$$F_2^*(0,v_2,t_2) \; = \; 2Q(t_2)f(v_2)F(v_2). \qquad (5.15)$$

The flow density $\rho_t(x,v,t)$ is,

$$\rho_t(x \; , \; v \; , \; t) \; = \; \frac{d}{dt} \; [F_1^*(x \; , \; v \; , \; t) + F_2^*(x \; , \; v \; , \; t)] \; , \qquad (5.16)$$

the sum of the densities for each of the two cars. At $x=0$ substitution of (5.13) and (5.15) into

ut

the sum of the densities for each of the two cars. At $x=0$ substitution of (5.13) and (5.15) into

(5.16) gives

$$\rho_t(0,v,t) \; = \; 2q(t)f(v)\{[1\text{-}F(v)] + F(v)\} \qquad (5.17)$$

$$= \; 2q(t)f(v).$$

This merely reproduces a result that is self-evident from (5.11). The total flow density is just twice that of a single car (before the cars were renumbered according to their velocities). We have reformulated this problem in this seemingly awkward way (adding the flow for the slowest, second slowest, etc. cars) because for x>0 and d≠0 we cannot represent the flow as the superposition of statistically independent flows. This awkward formulation was chosen so that we could follow the evolution conveniently.

For x>0, we have from (5.5) and (5.13)

$$F_1{}^*(x,v_1,t_1) = 2Q(t_1\text{-}x/v_1)f(v_1)[1\text{-}F(v_1)] \tag{5.18}$$

a result which does not depend upon the interactions. It simply describes the free motion of car 1 which has a velocity probability density $2f(v)[1\text{-}F(v)]$. From (5.7) and (5.13) we have for $v_1 < v_2$

$$F_2{}^*(x,v_2,t_2|v_1,t_1) = 2Q(t_2\text{-}x/v_2)f(v_2)/[1\text{-}F(v_1)] \tag{5.19}$$

$$\text{if } t_2 < t_1 - x/v_1 + x/v_2 \text{ or } t_2 > t_1$$

$$= 2Q(t_1\text{-}x/v_1)f(v_2)/[1\text{-}F(v_1)]$$

$$\text{if } t_1\text{-}x/v_1+x/v_2 < t_2 < min\{t_1\text{-}x/v_1+(x+d_{21})/v_2,t_1\}$$

$$= 2Q(t_2\text{-}(x+d_{21})/v_2)f(v_2)/[1\text{-}F(v_1)]$$

$$\text{if } t_1\text{-}x/v_1+(x+d_{21})/v_2 < t_2 < t_1.$$

Substitution of (5.18) and (5.19) into (5.14) now gives (after some integrations and
$$\text{if } t_1\text{-}x/v_1+(x+d_{21})/v_2 < t_2 < t_1.$$

Substitution of (5.18) and (5.19) into (5.14) now gives (after some integrations and rearrangement) for the marginal distribution

$$F_2{}^*(x,v_2,t_2) = 2Q(t_2\text{-}x/v_2)f(v_2)F(v_2)$$

$$-f(v_2) \int_0^{v_2} dv_1 f(v_1)[Q(t_2 - x/v_2) - Q(t_2 - x'/v_2)] [Q(t_2 - x'/v_2)$$

128

$$+Q(t_2 - x/v_2) - 2Q(t_2 - x/v_1)] \qquad (5.20)$$

where $x^*/v_2 = \min\{(x+d_{21})/v_2, x/v_1\}$.

The term on the second line describes the effect of $d_{21} \neq 0$. It vanishes for $d_{21}=0$ because $x^*=x$ for all v_1. It is negative for $d_{21}>0$ and $x>0$ because all factors in the integrand are positive (this must be so because the passings cause a shift to later arrival times). It also vanishes for $x=0$ to agree with (5.15).

Even in this simple situation with only two cars we are obtaining some unpleasant integrals, particularly since d_{21} will, in general, be a function of v_1 and v_2. To evaluate $p_t(x,v,t)$ from (5.16), we must still differentiate (5.20) with respect to t_2 and to find $q(x,t)$ we must integrate this with respect to v_2.

In order to obtain some fairly simple _explicit_ results, we now make some further restrictions. Let

1. $d_{21} = d$ be independent of v_2 and v_1

2. the departure times uniformly distributed over the time interval $(0,1)$ in some arbitrary time units, i.e.,

$$Q(t) = \begin{cases} 1 \; for & 1 < t \\ t \; \; for & 0 < t < 1 \\ 0 \; for & t < 0 \end{cases} \qquad (5.21)$$

$$Q(t) = \begin{cases} 1 \; for & 1 < t \\ t \; \; for & 0 < t < 1 \\ 0 \; for & t < 0 \end{cases} \qquad (5.21)$$

and 3. the reciprocal velocities $1/v$ (the time required to travel unit distance) be uniformly distributed over some interval $(a, a+b)$, i.e.,

$$\frac{1}{u^2} f\left(\frac{1}{u}\right) = \begin{cases} \dfrac{1}{b} & for \quad a < \dfrac{1}{u} < a + b \\ 0 & otherwise . \end{cases} \tag{5.22}$$

First for $d = 0$ we have

$\rho_i(x,v,t) = 2q(t\text{-}x/v)f(v)$

$$q(x , t) = 2\int\limits_0^\infty dv \; f(v)q(t - x/v)$$

$= 2/b \times$ *length of the intersection of the intervals*

(a,a+b) and ((t-l)/x, t/x) $\tag{5.23}$

This function is illustrated in figure III 8. For each of several values of x, we have superimposed on the x,t plane, graphs of q(x,t) vs t. For x=0, q(x,t) is the rectangular graph q(0,t)=1 for 0<t<1. The fastest car travels with velocity 1/a thus the pulse of traffic vanishes to the left of the line t=ax of figure III 8. The slowest car travels at velocity 1/(a + b); no car can be to the right of where a slowest car would be if it started at t=1, i.e., to the right of t = (a + b)x + 1. These two lines are shown as the broken lines of figure III 8.

For small x, specifically for x<1/b, the distribution of velocities causes both the front and rear of the pulse to disperse. For ax < t < (a + b)x + 1, the flow will lack some slow cars. For ax + 1 < t < (a + b)x + 1, the flow will lack some fast cars. The middle range, and rear of the pulse to disperse. For ax < t < (a + b)x + 1, the flow will lack some slow cars. For ax + 1 < t < (a + b)x + 1, the flow will lack some fast cars. The middle range, however, retains q(x, t)=1.
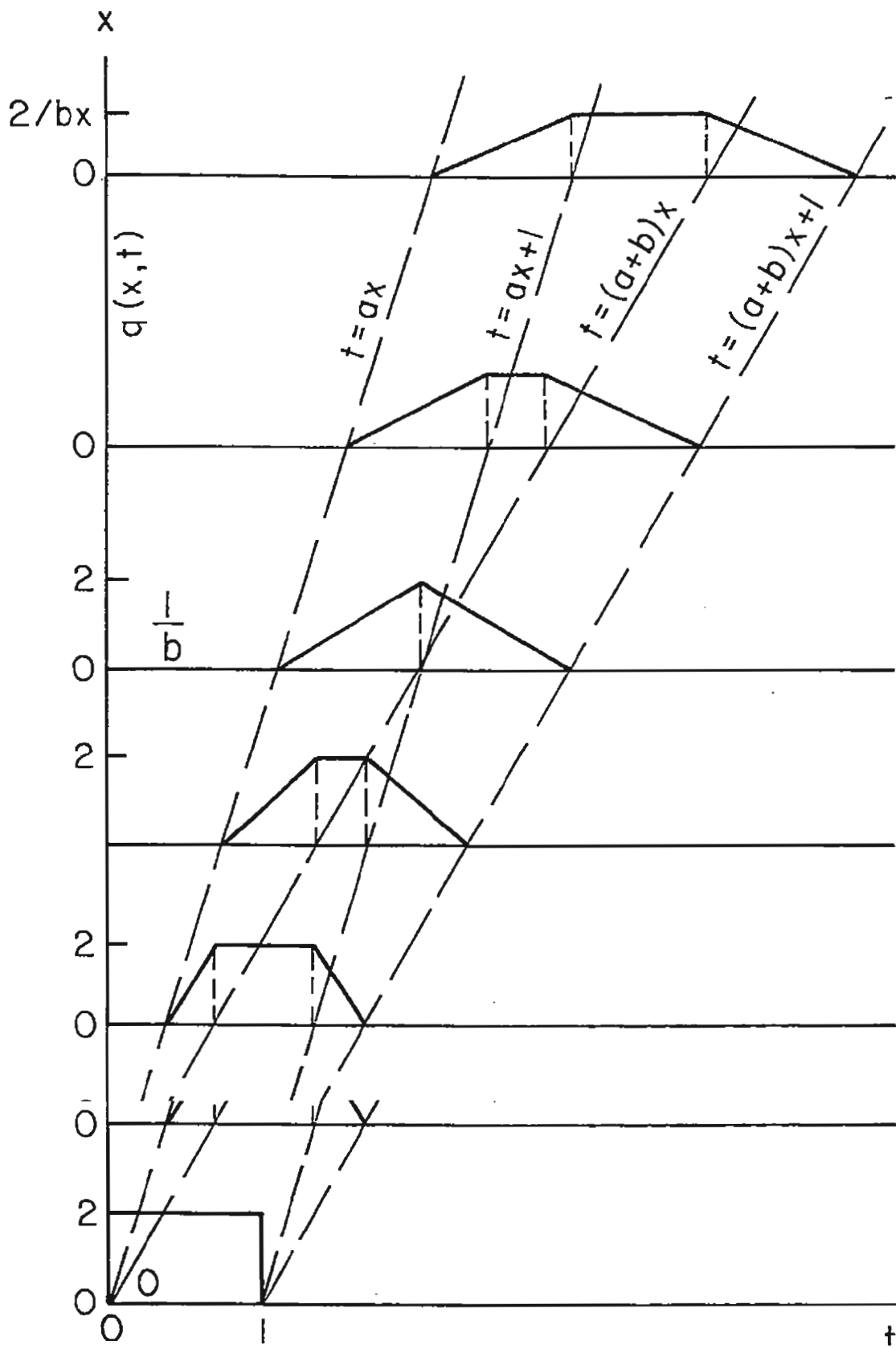
Fig. III 8
Platoon spreading with no interaction

131

When x=1/b, a fastest car from the rear of the pulse (t=1) can overtake a slowest car starting at the front. For any x>1/b, we will find, at any time t, only a limited range of possible velocities. The curve

of q(x,t) is again flat between ax+1 and (a+b)x but now it is so because we took a uniform velocity distribution, not because we started with a uniform q. In this middle range we have cars which might have started from any t, 0<t<1 but outside this range we can have only cars which started in some subset of (0,1).

The correction to the above formulas and graphs of q(x, t) when we take d>0, is the contribution to q(x,t) from the second line of (5.20). The evaluation of the integrals for the special case with Q(t) and f(v) given by (5.21) and (5.22) is elementary but somewhat tedious because this correction has different forms in various regions of the (x,t) plane. We shall describe the result below but not the derivation.

Figure III 9 shows the various regions of the (x,t) plane. For x<ad/b, no car is capable of completing a passing before it reaches x. Even a fast car with velocity 1/a which is ready at x=0 to begin passing a slowest car, velocity 1/(a + b), will arrive at x within a time (a+b)x traveling at the slowest speed but it will not complete the passing until x is so large that this time is equal to the time it would have arrived at x if it had started a distance d behind and traveled with the speed a, i.e., time a(x+d). Figure III 9 shows the trajectory of such a car time is equal to the time it would have arrived at x if it had started a distance d behind and traveled with the speed a, i.e., time a(x+d). Figure III 9 shows the trajectory of such a car starting at t = 0 (also from t=1). It is the line t = (a + b)x for x<ad/b and t=a(x+d) for x>ad/b. In the range x<ad/b, (we assume here that ad<1) we find for the correction Δq to the value for d=0, equation (5.23), the following:
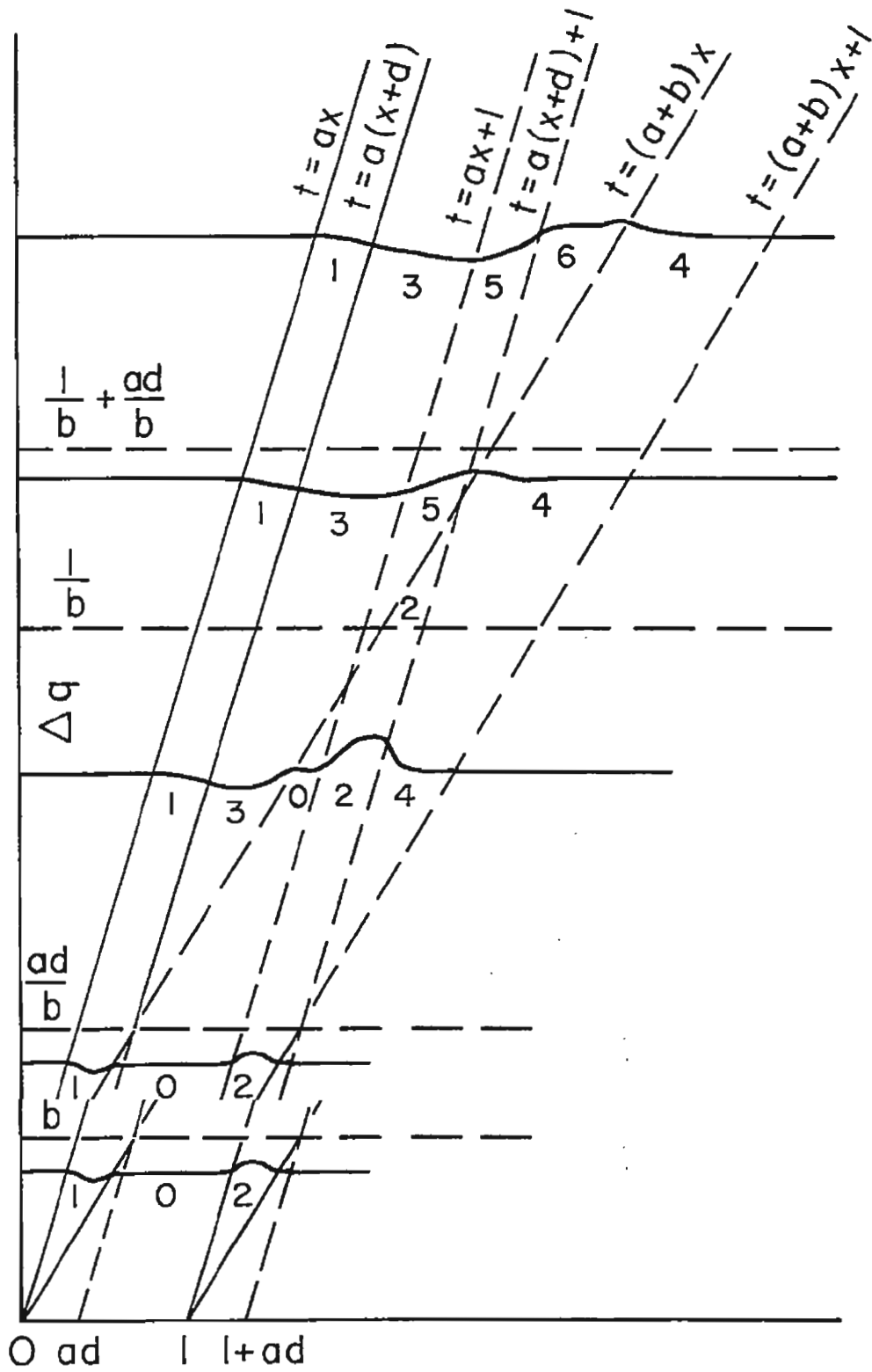
Fig. III 9
Correction to flow due to interactions

133

$$\Delta q = \begin{cases} -\dfrac{x}{b^2}\left[\dfrac{t}{x} - a\right]^2\left[a + b - \dfrac{t}{x}\right] \textit{for } a < \dfrac{t}{x} < a + b \quad (region1) \\[4mm] 0 \quad \textit{for } (a + b)x < t < ax + 1 (region0) \\[4mm] \dfrac{x}{b^2}\left[a + b - \dfrac{(t-1)}{x}\right]^2\left[\dfrac{t-1}{x} - a\right]\textit{for } a < \dfrac{t-1}{x} < a + b \; (region\ 2) \\[4mm] 0 \quad \textit{for} \dfrac{t}{x} < a \textit{ or } \dfrac{t-1}{x} > a + b \; . \end{cases} \qquad (5.24)$$

A typical curve of Δq vs t for some fixed x, x<ad/b is shown in figure III 9. This Δq

vanishes everywhere except in the range of x where q itself is either increasing or decreasing

as shown in figure III 8. In the range a≤t/x≤a+b, region 1 of figure III 9, we see from (5.24)

that Δq is a polynomial in t which has a second order zero at one end, t=ax, and a first order

zero at the other end. It is non-positive in this range, i.e., the passing delays tend to decrease

the flow at the front of the pulse as compared with d=0. The form of Δq for a<(t-1)/x<a+b is

just the mirror image of the above. It has a second order zero at the end t=(a+b)x+1 and is

non-negative. Delays tend to displace some of the flow into the rear of the pulse. The area

under the entire curve of Δq vs t must be zero since a change in d produces no change in the

total number of cars (two) crossing any x. The cars that are pushed out of region 1 of figure

III 9 appear later in region 2

total number of cars (two) crossing any x. The cars that are pushed out of region 1 of figure

III 9 appear later in region 2.

The formula (5.24) for region (1) also applies throughout the other regions labeled (1)

in figure III 9, i.e., for x>ad/b and ax<t<a(x+d). A fast car which has been delayed but still

arrives in one of these regions will not have completed its passing either since the earliest any

car can reach x after completion of a pass is at time a(x+d), the time a fastest car of velocity

1/a would arrive starting at t=0. In region 3 of figure III 9, equation (5.24) must be modified however because some cars will have completed a passing. Here Δq is less negative than given by (5.24), which is in effect the formula for d = ∞. Similarly in region 4, some cars which for d = ∞ would still be passing will, for finite d, have completed the pass. In region 2, the formula is still unchanged from (5.24).

For ad/b < x < 1/b, we find

$$
\Delta q = \begin{cases}
-\dfrac{x}{b^2}\left[\dfrac{t}{x}-a\right]^2\left[a+b-\dfrac{t}{x}\right] & \text{for } ax < t < a(x+d) \quad (\text{region } 1) \\[2ex]
-\dfrac{d}{b^2}\left[\dfrac{t^2}{x(x+d)}-a^2\right]\left[a+b-\dfrac{t}{x}\right] & \text{for } a(x+d) < t < (a+b)x \ (\text{region } 3) \\[2ex]
0 & \text{for } (a+b)x < t < ax+1 \ (\text{region } 0) \\[2ex]
+\dfrac{x}{b^2}\left[a+b-\dfrac{(t-1)}{x}\right]^2\left[\dfrac{t-1}{x}-a\right] & \text{for } ax+1 < t < a(x+d)+1 \ (\text{region } 2) \\[2ex]
+\dfrac{x}{b^2}\left[a+b-\dfrac{(t-1)}{x}\right]^2\left[\dfrac{(t-1)}{x(x+d)}\right] & \text{for } a(x+d) < 1 < t < (a+b)x+1 \ (\text{region } 4)
\end{cases}
\tag{5.25}
$$

The curve of Δq vs t in this range of x still has zeros at the same points as for (5.24) but the curve no longer has the front to rear mirror symmetry. The amplitude of (5.24) has been reduced in regions 3 and 4.

curve no longer has the front to rear mirror symmetry. The amplitude of (5.24) has been reduced in regions 3 and 4.

For $\dfrac{1}{b} < x < \dfrac{1}{b} + \dfrac{ad}{b}$, the formulas for Δq , (5.24) and (5.25), remains as before in

regions 1, 2, 3 and 4 but the 0 region has now disappeared and is replaced by a new region 5 where

$$\Delta q = -\frac{1}{b^2 x^2}[ax + 1 + (a+b)x - 2t] + \frac{1}{b^2 x(x+d)}[t - a(x+d)]^2[(a+b)x - t] \ . \tag{5.26}$$

(region 5)

Finally for 1/b + ad/b < x , the formulas for regions 1, 3, 4, and 5 remain as above but region 2 has disappeared and is replaced by region 6 where

$$\Delta q = -\frac{[1 + (a+b)x - 2t]d}{b^2 x^2 (x+d)} \quad (region \ 6) \ . \tag{5.27}$$

For x > > 1/b, we can make some simplifying approximations in which we keep only the largest terms of expansions in powers of x. Thus we find that

$$\Delta q \sim \begin{cases} -(t-ax)^2/xb & in \ region \ 1 \\ -ad[2(t-ax) - ad]/bx & in \ region \ 3 \\ +[(t-ax-ad)^2 - 1]/bx & in \ region \ 5 \\ -d(a+b-2t/x)/b^2 x^2 & in \ region \ 6 \\ [(a+b)x + 1 - t]^2 (a+b)d/b^2 x^2 & in \ region \ 4 \ . \end{cases} \tag{5.28}$$

In region 1, Δq vs x decreases quadratically from 0 at t = ax to $-a^2 d^2/xb$ at t = ax+ad. The latter is $-a^2 d^2/2$ times the maximum value $2(xb)^{-1}$ of q itself. From here Δq decreases almost linearly with t until t=ax+1 where Δq has the value -ad(2-ad)/bx. Since ad is not necessarily very small, this means that in regions 1 and 3, the passings can cause a significant

~~change in q(for ad=1 a change by a factor of 2)~~

necessarily very small, this means that in regions 1 and 3, the passings can cause a significant change in q(for ad=1 a change by a factor of 2).

In region 5, Δq starts to increase and for t=a (x+d)+1, Δq has returned to zero according to (5.28) but actually it reaches only some quantity of order $1/x^2$. These regions 1, 3, and 5 extend over a finite range of t even for x →∞. Region 6, however, has a width of order bx which becomes infinite for x →∞. In this region Δq increases linearly with t, but it

136

is uniformly of order $1/x^2$, smaller by $O(1/x)$ than in regions 1, 3, and 5. The area of Δq in region 6, however, must be essentially equal to the negative of that in regions 1, 3, and 5 (it has a small amplitude but a wide range). In region 4, Δq is also of order $1/x^2$ but this region has width 1, so it contributes nothing to the area for $x \to \infty$.

At least in this example, the passings cause for large x relatively large negative changes in q very near the front of the pulse but very small positive changes in q over most of the remaining range.

Although these examples illustrate some interesting physical phenomena and give some qualitative indication of how passing delays effect the dispersion of a pulse, they also suggest that even rather simple physical situations can lead to horrible computations. This problem is of considerable practical interest (one should know something about the spreading of a pulse in order to describe methods for the optimal synchronization of traffic signals, for example) but we shall not try to pursue it further here.

6. <u>Uniform flow on freeways</u>. One difference between freeways and the older type of rural road is that on the multilane freeway a car can pass another with essentially no change in velocity. Since at low density a passing car usually changes lanes gradually, it also suffers very little increase in distance traveled. Some drivers will even increase their speed during passing so as to produce a negative loss of time in terms of the model described above. The very little increase in distance traveled. Some drivers will even increase their speed during passing so as to produce a negative loss of time in terms of the model described above. The experimental fact [24, 25] is that the effective value of $d(v, v')$ in (2.3) or d in (2.4) is so small for freeways as to have no observable influence at all. Some experiments suggest that $u(v)$ may even increase slightly as the density increases for low densities so as to give in effect a negative value of d and a maximum value for $u(v)$ at some density greater than zero.

If we assume that two car interactions have a negligible effect, then we must look to the three car interactions as the lowest order interaction. If a highway has only two lanes for traffic moving in one direction then while one car passes another both lanes are temporarily blocked and any third car which might overtake this pair while they are passing will be forced to wait until the passing lane is free. Similarly for a three lane highway, one car will occasionally be passed simultaneously by two other cars thereby causing a temporary block of all three lanes to any fourth car that may wish to pass. Despite the fact that for low density these multiple passings should occur only rarely, the delays caused by any single occurrence may be sufficiently large that the total effect is a significant one.

For a highway two lanes wide (for the same direction of traffic), we assume that drivers are disciplined to keep in the right-hand lane except while passing. When one car wishes to pass a second car, the first car attempts to move into the passing lane when it has come to within a distance D behind the second car (measured center to center) and it stays in the passing lane until it has reached a point at a distance D ahead of the second car. We consider D to be the same for all cars. It should be approximately a "safe driving distance" between cars traveling at the prevailing average velocity. We also assume that for sufficiently low density of traffic, passings involving more than three cars are too infrequent to be of any importance.
low density of traffic, passings involving more than three cars are too infrequent to be of any importance.

We focus our attention on a reference car 0 that has desired speed v_0. This car retains the speed v_0 at all times except during some passing maneuver involving three cars. For such passings there are three types of situations to consider.

I. At the moment when car 0 has reached a point D behind a car 1 with desired speed

$v_1 < v_o$, the driver looks to see if there is a second car (car 2) in the passing lane already in the process of passing car 0 (and also car 1). We will consider this to be true if car 2 is within a distance D behind or ahead of car 0 and has a velocity $v_2 > v_o$. In this situation, we assume that cars 1 and 2 retain their respective desired speeds but that car 0 takes the speed v_1 and keeps this speed until car 2 has reached a point D in front of it, after which it regains its original speed v_o and passes car 1. See figure III 10.
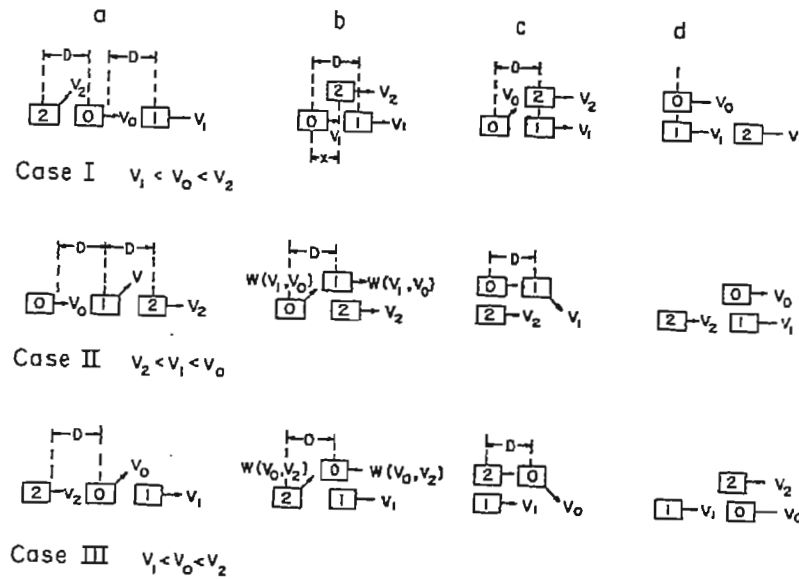

Fig. III 10
The types of passing maneuvers for three cars

II. At the moment when car 0 has reached a point D behind a car 1 with desired speed $v_1 < v_o$, car 1 is already in the process of passing a car 2 with speed $v_2 < v_1$. This occurs if car 1 lies within a distance D behind or ahead of car 2 at this instant. Our first impulse would be to

139

assume here that cars 1 and 2 retain their respective desired speeds and car 0 assumes the speed v_1 until car 1 has passed car 2, after which car 0 regains the speed v_o. Unfortunately, as will be seen below, this leads to the unrealistic conclusion that the traffic piles up behind cars for which v_1 and v_2 are nearly equal. To avoid this, it is necessary to make some assumption that will keep the time of passing finite. We will therefore assume that in the presence of a car 0 behind car 1, car 1 increases its speed to some value $w(v_1,v_o)$ with $v_1 \leq w(v_1,v_o) \leq v_o$ and that car 0 also travels at the speed $w(v_1,v_o)$ until car 1 has passed car 2.

III. As a result of the assumptions in II, we must also consider situations in which car 0 is in the position analogous to that car 1 in II. If car 0 passes car 1 and while doing so is overtaken by a car 2, then car 0 increases its speed to $w(v_o,v_2)$ for the remainder of the passing operation.

We will now evaluate the time average velocity $u(v_o)$ of a car with desired speed v_o following the same type of perturbation arguments used in sec. III.1, i.e., we will calculate passing rates and delays using an unperturbed Poisson distribution of cars.

To evaluate the delay due to maneuver I we observe first that car 0 with desired speed v_o overtakes cars with velocity between v_1 and $v_1 + dv_1$ at a rate of approximately

$$(v_o\text{-}v_1)k\,f_s(v_1)dv_1 \tag{6.1}$$

for $v_o > v_1$ as in (2.1). The probability that at the instant when car 0 overtakes car 1, there is

$$(v_o\text{-}v_1)k\,f_s(v_1)dv_1 \tag{6.1}$$

for $v_o > v_1$ as in (2.1). The probability that at the instant when car 0 overtakes car 1, there is also a car 2 with velocity between v_2 and $v_2 + dv_2$, $v_2 > v_o$, located at a position between x and x + dx relative to car 0 is

$$k\,f_s(v_2)dv_2dx. \tag{6.2}$$

140

(Here we have used the fact that for Poisson traffic, the probability density for car 2 is independent of the position or velocities of cars 0 and 1). In accordance with I above, car 0 must travel with velocity v_1 for a time

$$(D\text{-}x)/(v_2\text{-}v_1)$$

whenever this happens and $-D < x < D$. As a result car 0 loses a distance

$$(v_o\text{-}v_1)(D\text{-}x)/(v_2\text{-}v_1) \tag{6.3}$$

compared with what it would have traveled if it had not been delayed in passing. The total loss in distance traveled per unit time by car 0 due to all possible positions and velocities of cars 1 and 2 satisfying the conditions given in I is the product of (6.1), (6.2), and (6.3) integrated over all v_1, v_2 and x with $v_2 > v_o > v_1$ and $-D < x < D$, i.e.

$$2k^2D^2 \int_o^{v_o} dv_1 \int_{v_o}^{\infty} dv_2 \frac{(v_o - v_1)^2}{(v_2 - v_1)} f_s(v_1) f_s(v_2) \ . \tag{6.4}$$

In a similar way we find that the loss in distance traveled per unit time by car 0 due to interactions of type II is

$$2k^2D^2 \int_o^{v_o} dv_1 \int_o^{v_1} dv_2 \frac{(v_o - v_1)[v_o - w(v_1, v_o)]}{[w(v_1, v_o) - v_2]} f_s(v_1) f_s(v_2) \ , \tag{6.5}$$

while the loss (which is really a gain) due to conditions described in III is

while the loss (which is really a gain) due to conditions described in III is $\int_o^{v_o} \int_{v_o}^{\infty} \frac{(v_2 - v_o)[w(v_o, v_2) - v_o] f_s(v_1) f_s(v_2)}{}$ (6.6)

$$-2k^2D^2 \int_o^{v_o} dv_1 \int_{v_o}^{\infty} dv_2 \frac{(v_2 - v_o)[w(v_o, v_2) - v_o] f_s(v_1) f_s(v_2)}{[w(v_o, v_2) - v_1]} \ . \tag{6.6}$$

The total distance traveled per unit time by car 0, $u(v_o)$, is v_o minus the three expressions (6.4), (6.5), and (6.6).

The average velocity of all cars is

$$E\{u\} = E\{v_o\} - 2k^2D^2 \times$$

$$\left\{ \int_0^\infty dv_o \int_0^{v_o} dv_1 \int_{v_o}^\infty dv_2 \, f_s(v_o) \, f_s(v_1) \, f_s(v_2) \, (v_o - v_1)^2/(v_2 - v_1) \right.$$

$$+ \int_0^\infty dv_o \int_0^{v_o} dv_1 \int_0^{v_1} dv_2 \, f_s(v_o) \, f_s(v_1) \, f_s(v_2) \, (v_o - v_1) \frac{[v_o - w(v_1, v_o)]}{[w(v_1, v_o) - v_2]} \qquad (6.7)$$

$$\left. - \int_0^\infty dv_o \int_0^{v_o} dv_1 \int_{v_o}^\infty dv_2 \, f_s(v_o) \, f_s(v_1) \, f_s(v_2) \, (v_2 - v_o) \frac{[w(v_o, v_2) - v_o]}{[w(v_o, v_2) - v_1]} \right\}.$$

The form of $w(v_1, v_o)$ has not been specified, but, if we were to take $w(v_1, v_o) = v_1$ as suggested in II above, the integral in (6.5) diverges. The source of difficulty arises in this case because when car 1 passes car 2, it blocks the passing lane for a time of order $D/(v_1 - v_2)$. For $v_1 \to v_2$, this time becomes infinite, however, so rapidly, in fact, that we would be forced to conclude that a car spends most of its time in long queues even for arbitrarily small values of k. We must, therefore, either abandon the suggestion that $w(v_1, v_o) = v_1$ or revise the method of computing delays, because the assumption of negligible correlations in positions of cars would no longer be a valid approximation. Although the same singularity occurs in the integrand of (6.4), the numerator vanishes faster than the denominator for $v_1 \to v_2$ by virtue cars would no longer be a valid approximation. Although the same singularity occurs in the integrand of (6.4), the numerator vanishes faster than the denominator for $v_1 \to v_2$ by virtue of the restriction $v_2 \geq v_o \geq v_1$ in the range of integration.

There is no doubt that traffic on freeways even at low flows has a tendency to form occasional clusters and perhaps an effect such as this is one of the causes, but it does not seem realistic that the consequences of this are as violent as the above postulates might

suggest. If we modify this simple model, however, by taking some other form for $w(v_1,v_o)$ this can be done only at the expense of introducing at least one more parameter into the model to describe this function. At the present time there does not seem to be any experimental data that would be of much help in determining this.

One can easily deduce from this model with $v_l<w(v_1,v_o) \leq v_o$ that as v_o approaches the smallest allowed value then $u(v_o)$ - v_o goes to zero because all delays to a car of free speed v_o depend upon this car passing still slower cars and as v_o decreases so does the rate of passing slower cars. For sufficiently small v_o, $u(v_o)$ - v_o will be positive. The dominant effect for slow cars is that when they do pass still slower cars they will block the highway and will increase their speed to get out of the way, i.e. type III effects dominate those of type I and II. For high values of v_o, the passings give delays (type I and II losses overpower the gains from type III passings) and for most reasonable assumptions about w and the distribution $f_s(v)$, one will find that the average effect to all cars is a decrease in time average velocity. Since the low speed cars on the average increase their speed in passing and the high speed cars decrease their speed, the dispersion in time average velocity is reduced because of passing.

The integrals in (6.4), (6.5), and (6.6) all represent averages of functions depending only upon velocity differences. The integrals also have the dimension of velocity. For any reasonable distribution $f_s(v)$, these integrals will, therefore, be of the order of magnitude of only upon velocity differences. The integrals also have the dimension of velocity. For any reasonable distribution $f_s(v)$, these integrals will, therefore, be of the order of magnitude of the standard deviation of the free speed velocity distribution, σ, and $u(v_o)$ can be written in the form

$$u(v_o) = v_o - k^2 D^2 \sigma\, G\left(\frac{v_o - E\{v_o\}}{\sigma}\right) \tag{6.8}$$

143

where G is some non-dimensional function of the dimensionless velocity $(v_o - E\{v_o\})/\sigma$. The form of G will depend on the distribution $f_s(v)$ and the function w. The average of $u(v_o)$ for all cars has the form

$$E\{u\} = \int u(v_o) \, f_s(v_o) dv_o = E\{v\} - k^2 D^2 \sigma G \qquad (6.9)$$

for some constant G.

In contrast with the previous models that deal only with passings involving two cars and predict a linear decrease of u with k, the present model predicts that u decreases proportional to k^2. This is a direct consequence of the postulate that delays arise only from passings involving three cars. For dimensional reasons alone, k^2 must be multiplied by the square of some length and a velocity to give something with the dimension of a velocity. The velocity must be a measure of the dispersion since the number of passings depends only upon velocity differences and the length must be some length of highway in which a passing takes place. Thus, aside from the value of the number G in (6.9), the form of this equation is an obvious consequence of the basic postulates of the model and independent of the detailed mechanism of passing. Minor variations on this model can be incorporated in a new value for G.

We shall not try to construct a realistic model for three lane traffic here but only indicate what an obvious extension of the above two lane model will give. Real three lane traffic is more complicated than two lane traffic because drivers do not generally adhere to the rule that they should drive in the right-hand lane except to pass and even if they did there would be some question about delay to a driver who uses the second passing lane because the first two are blocked.

144

The analogue of the above two lane model would be that passings involving either two or three cars produce no delay because two cars can pass a third car simultaneously using all three lanes. Occasionally, however, a fourth car will overtake a trio of cars that are temporarily blocking all three lanes and be delayed until the lanes are unblocked. The only point we wish to make here is that even though the possibility exists for infinite delays in a model that is the analogue of the two lane model with $w(v_1, v_o) = v_1$, large delays do not occur often enough to cause a divergence in the low density expansion as occurred for the two lane highway.

For the two lane highway the divergence difficulty arose with $w(v_1, v_o) = v_1$ because the time consumed in passing was proportional to $D/(v_1-v_2)$ and one was led to integrals such as (6.5) which were of the form

$$\int_{x>0} dx/x$$

near $x = 0$. For the three lane highway, a block will last only until one or the other of the two passing cars has completed the passing. The time the highway is blocked by cars with velocities $v_3 > v_2 \geq v_1$ will be proportional to

$$\min \left\{ \frac{L_1}{v_3 - v_2} , \frac{L_2}{v_2 - v_1} \right\}$$

for some lengths L_1 and L_2. Although this delay is still infinite for $v_3 = v_2 = v_1$, the integrals one must evaluate to find the total delay will have the form

145

$$\int\limits_{x,\,y>0}\int dx\,dy\,\min\left\{\frac{1}{x},\frac{1}{y}\right\} = \int\limits_{x>0} dx\int\limits_{0}^{x}\frac{dy}{x} + \int\limits_{y>0} dy\int\limits_{0}^{y} dx\frac{1}{y}$$

$$= \int\limits_{x>0} dx + \int\limits_{y>0} dy$$

for small x and y. These integrals are finite.

For low densities, this model will lead to a formula analogous to (6.8)

$$u(v_o) = v_o - k^3 D^3 \sigma G((v_o - E\{v_o\})/\sigma)$$

in which G is finite. The delay proportional to k^3 results from the inclusion only of passings involving four cars.

That $u(v_o) - v_o$ should be proportional to k^3 and that passings involving three cars can be neglected is probably not realistic. Since the two lane highway model gave possible divergences, however, and the three lane highway does not, the three lane highway is probably less sensitive to passings than one might otherwise expect.

References
Chapter III

1. O.K. Norman and W.P. Walker, "Highway Capacity Manual," Bureau of Public Roads, U.S. Government Printing Office (1950).

2. G.F. Newell, "Mathematical models of freely flowing highway traffic," J. Opns. Res. Soc. America 3, 176-188 (1955).

3. M.S. Bartlett, "Some problems associated with random velocities," Publications de l'Institut de Statistique de l'Universite de Paris 6, 261-270 (1957).

4. L. Carleson, "A mathematical model for highway traffic," Nordisk Mat. Tidskr. 5, 176-180, 213 (1957).

5. A.J. Miller, "Traffic flow treated as a stochastic process," Proc. symposium on the theory of traffic flow, Detroit 1959, R. Herman, ed., Elsevier (1961), pp. 165-174.

6. G.F. Newell, Equilibrium probability distributions for low density highway traffic, J. Applied Probability 3, 247-260 (1966).

7. J. Riordan, Telephone traffic time averages, Bell System Tech. J. 30, 1129-1144 (1951).

8. V.E. Benês, Fluctuations of telephone traffic, Bell System Tech. J. 36, 965-973 (1957).

9. N.M. Mirasol, The output of an $M/G/\infty$ queueing system is Poisson, Opns. Res. 11, 282-284 (1963).

10. G.F. Newell, The $M/G/\infty$ queue, SIAM Journal 4, 86-88 (1966).

11. J.D.C. Little, A proof for the queueing formula $L = \lambda W$, Operations Res. 9, 383-387 (1961).

12. W. Feller, An introduction to probability theory and its applications, Vol. 1, 2nd ed., Wiley, New York (1957). (1961).

12. W. Feller, An introduction to probability theory and its applications, Vol. 1, 2nd ed., Wiley, New York (1957).

13. D.J. Buckley, Road traffic headway distributions, Proc. first conference Australian Road Research Board, pp. 153-187 (1962).

14. R.M. Oliver and B. Thibault, A high-flow traffic-counting distribution, Highway Res. Bd. Bull. 356, 15-27 (1962).

15. A.D. May, Jr., Gap availability studies, Highway Res. Record #72, 101-136 (1965).

16. André Schuhl, "The probability theory applied to distributions of vehicles on two-lane highways," Poisson and Traffic, pp. 59-75, The Eno Foundation for Highway Traffic Control, Saugatuck, Connecticut (1955).

17. B. Petigny, "Le calcul des probabilités et la circulation des véhicules sur les chaussées à deux ou trois voies," Ann. Ponts Chaussées, 131, 145-223 (1961).

18. A.J. Miller, A queueing model for road traffic flow, J. Royal Statist. Soc. B 23, 63-75 (1961).

19. R.M. Oliver, A note on a traffic counting distribution, Operational Res. Q. 13, 71-178 (1962).

20. A.J. Miller, Analysis of bunching in rural two-lane traffic, Opns. Res. 11, 236-247 (1963).

21. D.P. Gaver, Jr., Accommodation of second-class traffic, Opns. Res. 11, 72-87 (1963).

22. J.H.B. Kemperman, The passage problem for a stationary Marker chain, University of Chicago Press (1961).

23. F. Spitzer, Principles of random walk, Van Nostrand, Princeton (1964).

24. T.W. Forbes, "Speed, headway and volume relationships on a freeway, Proc. Inst. Traffic Engrs, pp. 103-126 (1951).

25. Highway Research Board Bulletin 235, Traffic behavior on freeways (1960).

IV. HIGH DENSITY TRAFFIC

1. <u>Introduction</u>. From the consideration of traffic flow at low densities for which cars behave

nearly independently, we now switch to the opposite extreme of very high density where the

dependence between drivers is so strong that they produce a highly ordered state of flow. It is

the consequences of passing that are difficult to analyze in any theory for moderate dense

traffic. By going to very high densities, we can minimize the effects of this because at high

densities cars have little room to make such a maneuver. We idealize the situation by

eliminating it completely. We also rule out any changing of lanes by cars so that a multilane

highway becomes, in effect, a collection of independent one-lane highways. The theory

described here, however, is not only a potential model for very dense traffic but is also

applicable to traffic of any density provided one can justify the no-passing condition on other

grounds. One of the main sources of experimental data for these studies is the traffic in the

tunnels connecting New York City and New Jersey under the Hudson River. These tunnels

are two-lane highways, but it is forbidden by law for cars to change lanes anywhere over a

two-mile stretch of road. Passing is also artificially eliminated occasionally when one lane of

a two-lane highway is blocked over long distances while it is being repaved.

The key feature of traffic flow with no passing that makes the mathematical analysis

The key feature of traffic flow with no passing that makes the mathematical analysis

tractable is that one can number the cars consecutively once and for all. Since the range of

view of a driver is finite, his motion depends only upon that of finitely many neighbors (the

same for all time). Furthermore, a driver is influenced mainly by the cars ahead of him and

only slightly, if at all, by those behind. Whatever detailed form the equations of motion may

take, the possibility exists that we can solve them iteratively, i.e., if we know the trajectory of

the jth, $(j-1)^{th}$, etc., cars, we can, in principle, determine the trajectory of the $(j+1)^{th}$, from which we can find that of the $(j+2)^{th}$, etc. This iterative scheme, however, cannot be readily applied if cars can pass each other because passing either alters the ordering of the cars, if we number the cars, or it alters the identity of the cars if we number their positions. The existence of both forward and backward interactions can also lead to some mathematical difficulties.

Unlike the theory of low density traffic which has evolved in a rather unsystematic way, the theory of traffic flow with no passing has developed through fairly well-defined stages. To a large extent, we shall follow here the historical order of evolution which begins with fairly simple models and mathematical techniques and progressively introduces new refinements.

The first significant work on traffic flow at high density was by Herrey and Herrey (1945)[1]. The mathematical analysis in this paper is rather crude (all trajectories are approximated by parabolas) but many features included in their model were eliminated by subsequent authors to simplify the detailed mathematics and were not reintroduced until 10 or 15 years later. Their paper seems unfortunately to have had rather little influence on later developments because many of the ideas introduced by Herrey and Herrey were forgotten and independently rediscovered later.

developments because many of the ideas introduced by Herrey and Herrey were forgotten and independently rediscovered later.

We consider first the simple case of a steady flow in which all cars travel at a constant velocity v. We assume that for any given velocity, a jth driver wishes, for reasons of safety, to maintain a certain minimum spacing $D_j(v)$ which depends upon the velocity v. Under congested conditions we assume that each does, in fact, travel at this minimum

150

spacing. It is perhaps difficult to ascertain (and irrelevant) whether or not drivers travel at what they consider a safe spacing or not, however, because in the final analysis the $D_j(v)$ will be what is experimentally observed and not what some safety expert may think is safe.

Much of what was previously described in Chapter 2 regarding models of traffic flow in which all cars travel at the same velocity still applies. In particular, one still defines a density k and a flow q in the same way. The main differences are that the model in which cars all travel at the same velocity makes more sense in the present context because cars cannot pass and therefore cannot maintain a difference in velocity very long. Also the specification that each driver has a spacing $D_j(v)$ implies that the velocity and the density are not functionally independent.

As regards the stochastic properties, we had previously postulated that for low densities the cars have a Poisson distribution, thus an exponential distribution for spacing with parameter k. Herrey and Herrey considered only the case in which the $D_j(v)$ were the same for all cars but one can easily extend the model by assuming that drivers are sampled from some population of drivers with different ideas about what is a safe distance. It is reasonable to assume, therefore, that for any fixed v the $D_j(v)$ are independent identically distributed random variables with a distribution function

$$F_D(d,v) = P(D_j(v) < d). \tag{1.1}$$

random variables with a distribution function

$$F_D(d,v) = P(D_j(v) < d). \tag{1.1}$$

The only difference from the low density case is that $F_D(d,v)$ is no longer the exponential distribution, $1 - e^{-kd}$.

In the language of probability theory, the number of cars in any interval of highway for the above model defines a "renewal process." There is an extensive literature on such

processes including a recent book by Cox [2]. Intuitively one expects that over any sufficiently long length L of highway there should be approximately kL cars. The distance L must be the sum of distances between cars, however, so

$$\sum_{j=1}^{kL} D_j(v) \sim L .$$

The $D_j(v)$ are independent random variables and by the law of large numbers, the sum of the $D_j(v)$ should be approximately kL times the expectation of the $D_j(v)$; thus

$$kL \; E\{D_j(v)\} = L$$

or

$$k = k(v) = 1/E\{D(v)\} . \tag{1.2}$$

A rigorous proof of this is one of the basic theorems in renewal theory which will not be reconstructed here.

The equation (II 2.7) q = kv, is still valid so that we also have that

$$q = q(v) = vk(v) = v/E\{D(v)\} . \tag{1.3}$$

A similar argument can be used to describe time headways. The time headway $t_j(v)$ is defined as the time between the crossings of a fixed position by the $(j-1)^{th}$ and j^{th} cars. Since the cars travel at velocity v, it follows that

$$t_j(v) = D_j(v)/v \tag{1.4}$$

the cars travel at velocity v, it follows that

$$t_j(v) = D_j(v)/v \tag{1.4}$$

and since v is fixed (not random)

$$E\{t(v)\} = \frac{1}{v} \; E\{D(v)\} \tag{1.5}$$

and

152

$$q(v) = \frac{1}{E\{t(v)\}} . \qquad (1.6)$$

What functional form one assumes for $D_j(v)$ or $E\{D(v)\}$ depends upon whether one postulates some model for safe driving or uses for $D(v)$ what in fact happens. Herrey and Herrey proposed that $D(v)$ was a quadratic function

$$D(v) = a + bv + cv^2 . \qquad (1.7)$$

The term a is the minimum spacing for velocity zero and is somewhat more than the length of a car. The constant b has dimensions of time and represents a reaction time; bv is the distance a car travels before the driver can start to brake. The last term represents the distance traveled if the driver decelerates at a constant rate after applying the brakes until he has stopped.

Equation (1.7) agrees qualitatively with what drivers actually do at moderately close spacing but for large spacing the velocity should have a finite limit. In later applications we will be more concerned with the inverse of this relation; i.e., the relation $V(d)$ giving the velocity as a function of the spacing. The qualitative shape for $V(d)$ is shown in Figure IV.1. For spacings less than $D(0)$, if possible, $V(d)$ is zero. For large spacings $V(d)$ tends to a finite limit $V(\infty)$ which one naturally interprets as the driver's free speed. Everywhere the curve has positive but monotone decreasing slope.
limit $V(\infty)$ which one naturally interprets as the driver's free speed. Everywhere the curve has positive but monotone decreasing slope.

One important consequence of the above shape for $V(d)$ is that the curve for the flow q as a function of density k has the shape shown in Figure IV.2. For low k,

$$q = vk \sim kV(\infty)$$

has slope $V(\infty)$ and is zero for $k = 0$. For large k, $k \to 1/D(0)$, the velocity goes to zero and
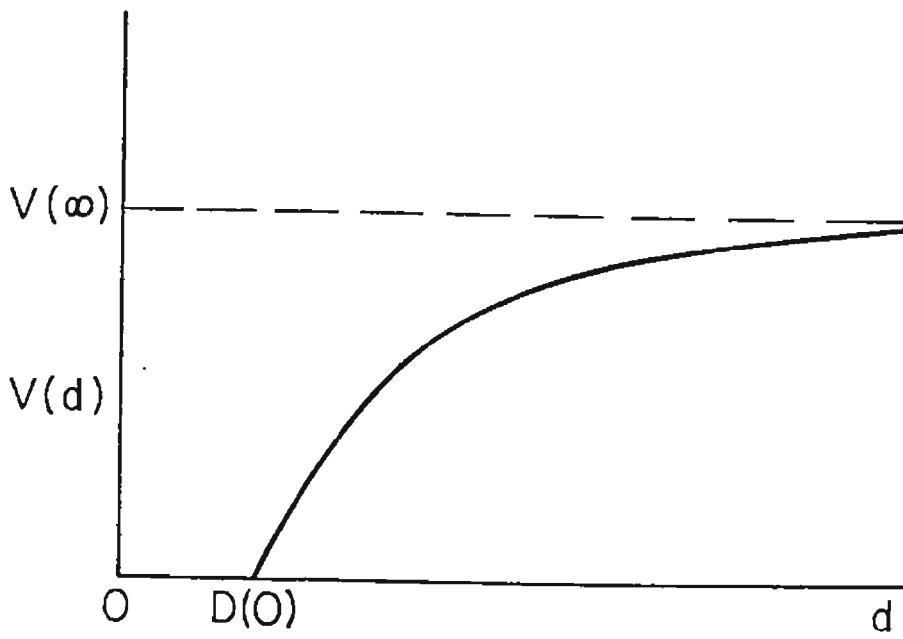
153

Fig. IV 1
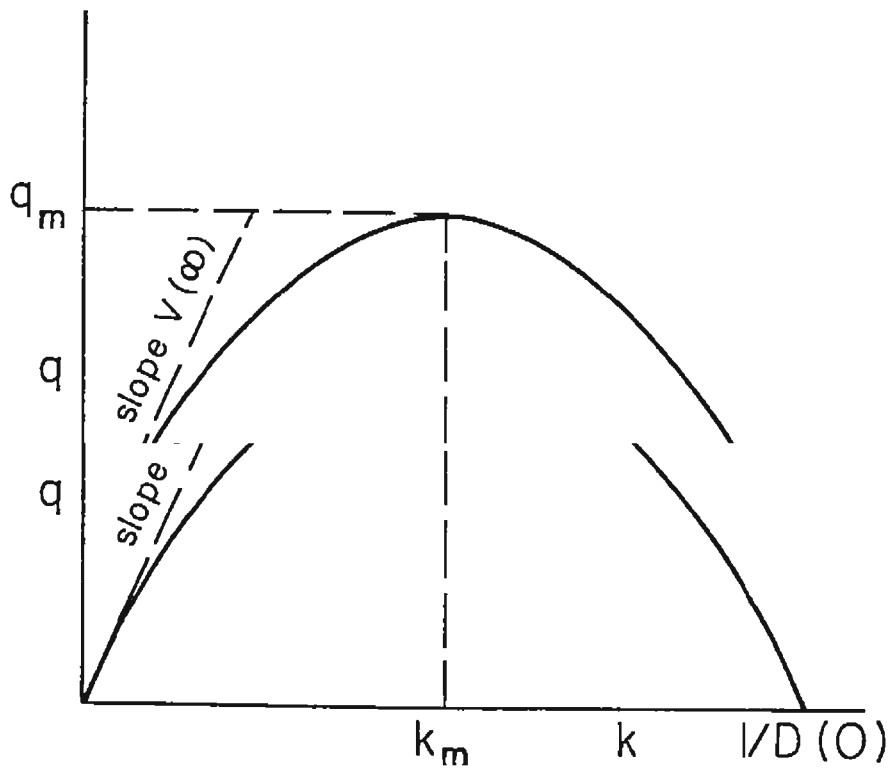A typical relation between velocity and spacing

Fig. IV 2
A typical relation between flow and density

154

$$q \sim v(k)/D(0)$$

again goes to zero. The curve of q vs k must have a maximum for some intermediate k and consequently there exists a density k_m and a corresponding velocity v_m which produces the maximum flow q_m. The value of q_m is usually called the capacity. The surprising experimental fact about these curves is that v_m is quite small for typical urban roads, usually less than 20 m/h.

Although considerable effort has been devoted to experimental measurements of these curves, various methods of averages, etc., the main emphasis in the theoretical work has been directed toward study of time-dependent flows.

The simplest assumption one can make about time-dependent flow is that even when the velocity varies with time, a driver follows the car ahead of him in such a way that his velocity and spacing are related by V(d). To simplify the mathematics somewhat, we will assume now that this relation is the same for all drivers. This is not quite what Herrey and Herrey did, but they did propose an algorithm for constructing trajectories based upon a safe driving distance and their results were almost equivalent to the above assumption. This specification that the velocity is related to the spacing implies a procedure for constructing trajectories. If $x_j(t)$ is the position of the jth car at time t, then the spacing is $x_{j-1}(t) - x_j(t)$. The velocity $v_j(t)$ of the jth car is $dx_j(t)/dt$ and so the assumption that velocity and spacing are

velocity $v_j(t)$ of the jth car is $dx_j(t)/dt$ and so the assumption that velocity and spacing are related through the function V implies that

$$\frac{dx_j(t)}{dt} = V[x_{j-1}(t) - x_j(t)] \ . \tag{1.8}$$

This is a system of first order (in general nonlinear) equations for $x_j(t)$ which can, in principle, be solved iteratively. If $x_0(t)$ is known, the equation for $j=1$ is an equation for $x_1(t)$ which can be solved (numerically or graphically if necessary) in terms of $x_0(t)$ and some initial condition $x_1(0)$. Knowing $x_1(t)$ and $x_2(0)$ one can then solve the equation with $j=2$ for $x_2(t)$, etc.
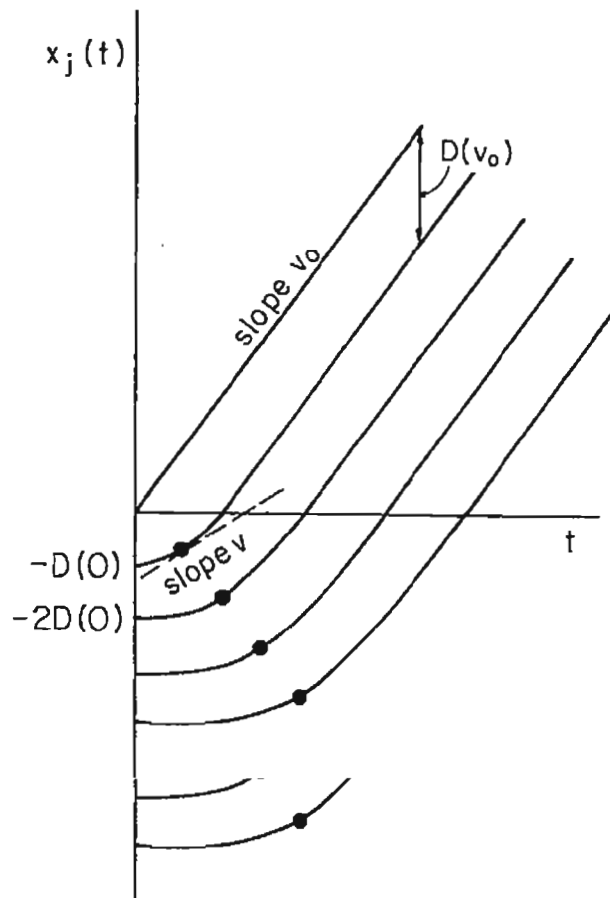

Fig. IV 3
Trajectories for a sequence of vehicles

Suppose that we have somehow constructed a series of trajectories. For example Figure IV 3 shows a hypothetical series of trajectories for cars which are initially at rest with spacing $D(0)$, but, at time 0, the lead car suddenly accelerates to velocity $V(\infty)$, as might occur for example when a traffic signal turns green. Before any car can increase its speed, it must wait for the spacing to increase to the safe driving distance for the higher velocity. If one marks on each trajectory the points where that car first reaches some velocity v, as represented in Figure IV 3 by the circles, and connects these points by a broken line, one obtains what Herrey and Herrey called the iso-velocity lines of velocity v. In the current terminology these are called the "waves" of velocity v. These waves are of considerable interest in the description of time-dependent flows because, as we shall see later, they also trace the propagation of any small disturbance at the velocity v. It should be noticed that this propagation of the starting wave back into the line of cars with a finite velocity is a result of the fact that drivers must increase their spacing as the velocity increases. We have said nothing here about "delayed reactions" of drivers. Herrey and Herrey also observed that the wave corresponding to the optimal speed v_m (for which is q_m is a maximum) is stationary. From this they concluded that the flow passing through a traffic intersection would almost immediately adjust to the value q_m and stay at that value provided the lead car has a velocity larger than v_m.

immediately adjust to the value q_m and stay at that value provided the lead car has a velocity larger than v_m.

Most of these ideas which were first introduced by Herrey and Herrey by means of rather crude analytical and graphical methods of approximation will be reconsidered later in a more sophisticated way but in the historical development the emphasis now changed to explicit analytic solution of simpler models. This turn to more exact mathematical analysis

has its drawbacks, however, and may even have delayed progress in the field because in the final analysis one will be forced to deal with realistic models by means of approximate techniques and one will want to know results only as accurately as can be read from a graph. Exact formulas are frequently so complicated that they only obscure relatively simple qualitative results.

2. Linear theories. The main obstacle to a simple analytic solution of (1.8) is that the equations are, in general, nonlinear. Most of the emphasis for the next 10 or 15 years after Herrey and Herrey's work was directed toward solution of various linear models. The first linear model was a straightforward linearization of (1.8); the function V(d) was taken to be linear in d,

$$\frac{dx_j(t)}{dt} = \alpha[x_{j-1}(t) - x_j(t)] - \beta \tag{2.1}$$

with α and β constants independent of $x_j(t)$, j, or t. This was proposed independently by Reuschel (1950) [3,4] and by Pipes (1953) [5].

One can think of this linear relation between velocity and spacing either as a crude approximation to V(d) over a wide range of spacings d or as an accurate representation over a small range of d. In the latter case one, in effect, replaces the nonlinear V(d) by its tangent approximation to V(d) over a wide range of spacings d or as an accurate representation over a small range of d. In the latter case one, in effect, replaces the nonlinear V(d) by its tangent line at some fixed value of d about which small changes are to be considered. The values of α and β/α are the slope and intercept respectively of the tangent line.

More recent models also include a reaction time of the driver. These were first proposed independently by Chandler, Herman and Montroll (1958) [6] and by Kometani and

Sasaki (1958) [7] after which a series of quite a few papers appeared giving further details and refinements [8-12].

Kometani and Sasaki proposed that a driver chooses his velocity at time t+T according to the spacing he observes at an earlier time t. The time T is interpreted as a reaction time, the time required for the driver to respond to any changes in the spacing. They thus proposed a set of equations

$$\frac{dx_j(t + T)}{dt} = \alpha[x_{j-1}(t) - x_j(t)] - \beta .$$

(2.2)

Chandler, Herman, and Montroll argued that a driver does not control his velocity directly but only the accelerator pedal or brakes of his car. They therefore proposed that the driver chooses his acceleration at time t+T on the basis of whatever information he may have about his state at time t. They suggested a number of models in which the acceleration was linearly dependent upon deviations in the spacing from some given value, differences in velocities between the jth and $(j-1)^{th}$ cars, $v_{j-1}(t) - v_j(t)$, and even changes involving the next nearest neighbor car. Experiments designed to estimate the parameters in the model suggested, however, that the acceleration depended mainly upon the velocity difference. Most of this work and subsequent work by others centered around an equation of the form

work and subsequent work by others centered around an equation of the form   (2.3)
$$\frac{d^2x_j(t + T)}{}$$

$$\frac{d^2x_j(t + T)}{dt^2} = \alpha[v_{j-1}(t) - v_j(t)] .$$

(2.3)

This equation is just the time derivative of (2.2). Even though the two models (2.2) and (2.3) arise from slightly different motivation, the end result is the same. Since the first

159

step in solving (2.3) would normally be to integrate the equation into the form (2.2) and since (2.2) includes (2.1) as a special case (for T=0), it suffices to consider only (2.2).

There are a number of analytic methods for solving equations such as (2.2). For any fixed j, (2.2) would be called a linear differential-difference equation with constant coefficients because it contains both derivatives in t and displaced times t and t+T but none of the coefficients α or β depends upon t. We would consider the jth equation of (2.2) as an equation for $x_j(t)$ with $x_{j-1}(t)$ as an inhomogeneous term. The solution would give $x_j(t)$ in terms of $x_{j-1}(t)$. There is an extensive literature on differential-difference equations including a book by Bellman and Cooke [13]. Considered as an equation describing the dependence of $x_j(t)$ on j, (2.2) would also be considered a difference equation with constant coefficients provided α and β were the same for all drivers.

To establish the existence of solutions of (2.2) and also to evaluate the solutions explicitly for moderate values of t/T and j, we notice that if one is given $x_o(t)$ for $0<t< \infty$, $x_1(t)$ for $0<t\leq,T$, $x_2(t)$ for $T<t<2T$, etc., $x_j(t)$ for $(j-1)T<t<jT$, then one can determine $x_j(t)$ from (2.2) for all t, jT<t, by iterative methods. For example, the equation for j=1 gives $v_1(t)$ for $T<t\leq 2T$ in terms of $x_1(t)$ and $x_o(t)$ with $0<t\leq T$, both of which are given. If we just integrate this equation directly we obtain

this equation directly we obtain

160

$$x_1(t) = x_1(T) + \int_0^{t-T} d\tau[\alpha[x_o(\tau) - x_1(\tau)] - \beta]$$

$$for \ T < t < 2T$$

in which everything on the right-hand side is given. Having evaluated $x_1(t)$ in (T,2T) one can next obtain it in (2T,3T) by the same procedure. Generally, if one already knows $x_j(t)$ and $x_{j-1}(t)$ for (k-1)T<t≤kT, one can evaluate $x_j(t)$ for kT<t≤(k+1)T from the equation

$$x_j(t) = x_j(kT) + \int_{(k-1)T}^{t-T} d\tau\{\alpha[x_{j-1}(\tau) - x_j(\tau)] - \beta\} \tag{2.4}$$

$$for \ kT < t \le (k + 1)T \ .$$

This suffices to determine $x_j(t)$ for all jT<t.

From a computational point of view, this scheme becomes quite tedious if one must iterate very many times. One can obtain the exact formulas this way, however, and also evaluate the asymptotic behavior for large j and/or t, but there are simpler ways of finding the asymptotic behavior. The main advantage of this iterative method is that it would apply even if (2.2) were nonlinear in the $x_j(t)$ and even if the coefficients α and β depended upon j and t although in this case the possibility of evaluating the integrals in any reasonable form is very remote.

and t although in this case the possibility of evaluating the integrals in any reasonable form is

very remote.

That the coefficients α and β in (2.2) do not depend upon either j or t and the equations are linear suggests immediately that you could advantageously use Laplace or Fourier transforms in time and/or generating functions, Laplace or Fourier-Stieltjes transforms

in j. The main point is that (2.2) is invariant to translations in t or integer translations in j; one car is the same as the next and one time is equivalent to any other time.

The inhomogeneous term β in (2.2) can be removed if we let

$$z_j(t) = x_j(t) + j\beta/\alpha .\tag{2.5}$$

In effect we remove from the highway the minimum spacing β/α between each pair of cars and move all cars forward to take up the gaps. There are j gaps between $x_o(t)$ and $x_j(t)$ and so the jth car is moved forward by $j\beta/\alpha$. The equation for $z_j(t)$ becomes homogeneous

$$\frac{dz_j(t + T)}{dt} = \alpha \ [z_{j-1}(t) - z_j(t)] .\tag{2.6}$$

Before becoming involved in detailed analytic solution by rather formal methods using transforms, let us anticipate the nature of the solution using some rather crude mathematics. Suppose we conjecture that the solutions $z_j(t)$ should be, in some appropriate sense, slowly varying in both t and j. We think of j as a continuous parameter, $z_j(t)$ as a function of two real variables, and d/dt in (2.6) as $\partial/\partial t$, the partial derivative for fixed j. We make the approximations

$$\frac{dz_j(t + T)}{dt} \sim \frac{\partial}{\partial t} \ z_j(t)$$

$$z_{j-1}(t) - z_j(t) \sim -\frac{\partial}{\partial t} \ z_j(t)$$

$$z_{j-1}(t) - z_j(t) \sim -\frac{\partial}{\partial j} \ z_j(t)$$

so that (2.6) becomes

162

$$\frac{\partial z_j(t)}{\partial t} + \alpha \, \frac{\partial z_j(t)}{\partial j} \sim 0 \, . \tag{2.7}$$

This is the simplest of all partial differential equations, a linear partial differential equation of first order in two independent variables. It says, in effect, that $z_j(t)$ is constant in the direction $(1,\alpha)$ in the (t,j) plane. Equivalently, if we make a change in variables

$$t' = t - j/\alpha \tag{2.8}$$

i.e., we consider a new time coordinate t' which depends upon the car number j and which for each jth car is displaced by an amount $1/\alpha$ from that of its predecessor $(j-1)$, and we let

$$z_j(t) = z_j'(t') = z_j'(t-j/\alpha)$$

i.e., $z_j'(t')$ is the corresponding function to $z_j(t)$ in the new coordinate system (t',j), then (2.7) becomes

$$\frac{\partial z_j'(t')}{\partial j} \sim 0 \tag{2.9}$$

where the partial derivative with respect to j now means the derivative with t' fixed.

The general solution of (2.7) or (2.9) is

$$z_j'(t') = z_j(t) \sim h(t') = h(t - j/\alpha) \tag{2.10}$$

in which h is an arbitrary function. If we are given the trajectory of the lead car $x_o(t)$, then

h(t) must be $x_o(t)$ and so
in which h is an arbitrary function. If we are given the trajectory of the lead car $x_o(t)$, then

h(t) must be $x_o(t)$ and so

$$z_j(t) \sim z_o(t-j/\alpha)$$

or

$$x_j(t) \sim x_o(t-j/\alpha) - j\beta/\alpha \, . \tag{2.11}$$

The physical interpretation of this is that every car approximately reproduces the
motion of the lead car except that it is measured from a position translated by a distance $j\beta/\alpha$
and a time translated by j/α. A hypothetical set of trajectories corresponding to (2.11) is
shown in Figure IV 4. The circles mark corresponding points on the trajectories $x_j(t)$ and if
one connects these points by a line as was done in Figure IV 3 to indicate the propagation of
a wave, the line will always have slope $-\beta$. We say that the wave propagates with the
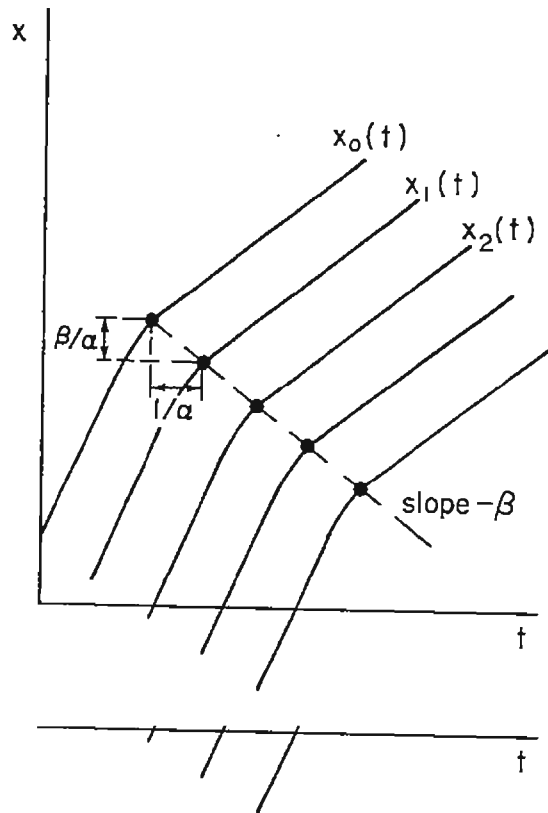velocity $-\beta$. In this linear model, all waves have the same velocity.

Fig. IV 4
Approximate trajectories

For certain idealized situations, it is a simple exercise to compute the wave velocity $-\beta$
directly from Figure IV 4. Suppose we assume that, before some disturbance, cars were

164

traveling at a steady speed v, and spacing $D(v_1)$. After the disturbance cars traveled at a steady speed v_2 and spacing $D(v_2)$. Regardless of whether or not each car mimics exactly the motion of its predecessor, if the final motion is one of constant velocity we can extrapolate the straight line trajectories of the initial and final motion until they intersect. We can then interpret the path of the intersections, the locus of the discontinuities of slope, as the path of wave propagation. Knowing the spacings and slopes of the two families of straight lines, it is a simple exercise to evaluate the slope of the intersection. In essence this is, in fact, the geometrical interpolation of the above mathematical approximation.

The approximations which led to (2.11) are rather crude and neglect some other important effects. It is, therefore, worthwhile to extend the approximation one step further. We now make the substitutions

$$z_j(t + T) = z_j'(t' + T) = z_j'(t') + T \frac{\partial}{\partial t'} z_j(t') + ..$$

$$z_{j-1}(t) = z_{j-1}'(t - (j-1)/\alpha) = z_{j-1}'(t' + 1/\alpha) =$$

$$= z_j'(t') + \frac{1}{\alpha} \frac{\partial}{\partial t'} z_j'(t) - \frac{\partial}{\partial j} z_j'(t) + \frac{1}{2\alpha^2} \frac{\partial^2}{\partial t'^2} z_j'(t') + ..$$

into (2.6). If we neglect all third or higher derivative terms and all derivatives of $\partial z(t')/\partial j$

which according to (2.9) is already small we obtain

into (2.6). If we neglect all third or higher derivative terms and all derivatives of $\partial z(t')/\partial j$ which according to (2.9) is already small we obtain

$$\frac{\partial z_j'(t')}{\partial j} \sim \frac{(1-2\alpha T)}{2\alpha^2} \frac{\partial^2 z_j'(t')}{\partial t'^2} . \tag{2.12}$$

This equation is also well-known. It is the heat conduction or diffusion equation in which j now takes the role normally taken by the time, t' the role normally taken by the space

coordinate, $z_j'(t')$ the role of temperature or concentration and $(1-2\alpha T)/(2\alpha^2)$ the conduction or diffusion coefficient. The coefficient of heat conduction or diffusion is always positive. If it were negative this would be equivalent to reversing the direction of time or in (2.12) reversing the direction of increasing j. For heat conduction or diffusion one knows that a concentration of heat or of mass will in the course of time disperse and approach a uniform distribution in space. The reverse processes with time running backwards must take a nearly uniform distribution into a highly irregular one. By analogy we expect that (2.12) with $1-2\alpha T<0$ would give rise to a highly irregular motion of the jth car even if $x_o(t)$ were slowly varying. Such a behavior would violate the assumptions used in deriving (2.12) so that we cannot justify the use of (2.12) for $1-2\alpha T<0$ but we shall see later by means of a more exact treatment of (2.6) that this condition does in fact lead to an amplification of small disturbances as (2.12) suggests. Actually, the solution of (2.12) with a negative coefficient is not always mathematically well defined.

On the other hand if $1-2\alpha T>0$, we know that even an irregular motion for $x_o(t)$, corresponding to an irregular initial concentration in diffusion, will lead to a smooth $x_j(t)$ for large j. Thus we expect that the larger j, the better one can justify the assumptions used in deriving (2.12) from (2.6). This one can readily verify by actually constructing the solution of (2.12) and showing that the terms which have been neglected in (2.12) are in fact small. deriving (2.12) from (2.6). This one can readily verify by actually constructing the solution of (2.12) and showing that the terms which have been neglected in (2.12) are in fact small.

The solution of (2.12) is well-known. The fundamental solution or initial value Green's function is

$$G(t', \tau) = \frac{\alpha}{[2\pi j(1-2T\alpha)]^{1/2}} \exp\left[\frac{-\alpha^2(t'-\tau)^2}{2j(1-2T\alpha)}\right] = G(t'-\tau) \qquad (2.13)$$

in terms of which the solution for arbitrary $z_0'(t')$ is

166

$$z_j'(t') = \int_{-\infty}^{+\infty} d\tau \, G(t',\tau) z_0'(\tau) = \int_{-\infty}^{+\infty} d\tau \, G(t'-\tau) z_o'(\tau)$$

(2.14)

$$= \int_{-\infty}^{+\infty} d\tau \, G(\tau) \, z_o(t'-\tau) \ .$$

The trajectories themselves are therefore given by

$$x_j(t) = -j\beta/\alpha + \int_{-\infty}^{+\infty} d\tau \, G(t - j/\alpha , \tau) x_o(\tau)$$

(2.15)

$$= -j\beta/\alpha + \int_{-\infty}^{+\infty} d\tau \, G(\tau) x_o(t - \tau - j/\alpha) \ .$$

This approximation reduces to (2.11) if we neglect the diffusion effects or equivalently if we set $1-2\alpha T=0$ in (2.12). If we let $1-2\alpha T\to 0$ in (2.13), $G(t',\tau)$ becomes $\delta(t'-\tau)$ and (2.15) goes over to (2.11). In a certain sense the first order effect is still that a disturbance propagates from the zeroth car to the jth with a propagation time of j/α. The second order effect is that any sudden change of the lead car becomes smeared out. Whereas, the propagation time grows linearly with j, the spreading of the disturbance grows only as $j^{\frac{1}{2}}$. Since our approximation of j by a continuous variable implies that j is large, it is in this sense that the first effect is large compared with the second.

The time derivative of (2.15) gives for the velocities
that the first effect is large compared with the second.

The time derivative of (2.15) gives for the velocities

$$v_j(t) = \int_{-\infty}^{+\infty} d\tau \, G(\tau) v_o(t - \tau - j/\alpha) \ .$$

Further differentiation of this gives similar equations for the accelerations, their derivatives, etc.

The simplest functional form for $v_o(t)$ which illustrates the propagation and diffusion of a disturbance is

$$v_o(t) = (2\pi\gamma)^{-1/2} exp(-t^2/2\gamma).$$

This represents a motion in which the lead car starts at $t = -\infty$ with velocity 0, accelerates gradually to a velocity $(2\pi\gamma)^{-1/2}$ at $t=0$ and then decelerates in a similar way back to zero velocity. It travels unit distance in the process. The jth car has a velocity

$$v_j(t) = \alpha \int_{-\infty}^{+\infty} d\tau \frac{exp\left[\dfrac{-\alpha^2\tau^2}{2j(1 - 2T\alpha)} - \dfrac{(t - \tau - j/\alpha)^2}{2\gamma}\right]}{2\pi\gamma^{1/2} [j(1 - 2T\alpha)]^{1/2}}$$

$$= exp\frac{\left[-\dfrac{(t - j/\alpha)^2}{2[\gamma + j(1 - 2T\alpha)/\alpha^2]}\right]}{(2\pi)^{1/2} [\gamma + j(1 - 2T\alpha)/\alpha^2]^{1/2}} .$$

Whereas car 0 has a Gaussian shaped velocity pulse centered at $t = 0$ with a duration measured by $\gamma^{1/2}$ and a peak velocity $[2\pi\gamma]^{-1/2}$, car j has a Gaussian pulse centered at time j/α, a duration $[\gamma+j(1-2T\alpha)/\alpha^2]^{1/2}$ and a maximum velocity proportional to the reciprocal of this. The pulse propagates with a time lag j/α but spreads and loses amplitude. For large j j/α, a duration $[\gamma+j(1-2T\alpha)/\alpha^2]^{1/2}$ and a maximum velocity proportional to the reciprocal of this. The pulse propagates with a time lag j/α but spreads and loses amplitude. For large j the spread increases proportional to $j^{1/2}$. The total distance traveled by the jth car is 1, however, for all j; the spacing between cars returns to the same value at $t = +\infty$ with which it started at $t = -\infty$. The curves of velocity vs time for this motion are shown in Figure IV 5.
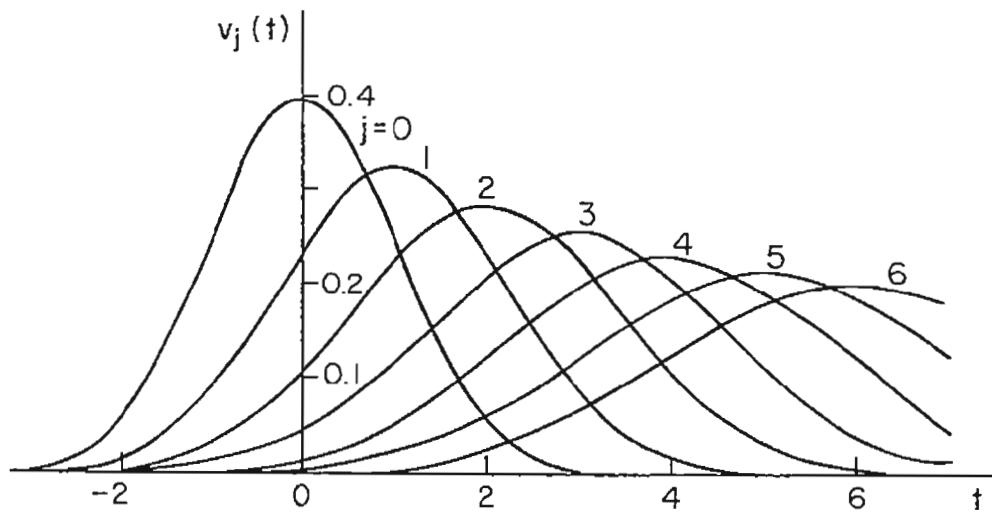
168

Fig. IV 5
Propagation of Gaussian shaped velocity pulse

Equations similar to (2.15) were first obtained by Chandler, Herman, and Montroll [6] from exact solutions of (2.3). The heuristic derivation given here mimics one used by Newell [14] in discussing some nonlinear models.

3. Solution by Laplace transform. We shall now reconsider (2.6) and obtain exact solutions in terms of Laplace transforms. To find the general solution with arbitrary starting conditions for all cars is quite cumbersome and not very illuminating. We are mainly concerned with the propagation of disturbances originating from car 0 at some time $t > 0$ and not with a succession of disturbances originating from each jth car because it may start initially with some velocity different from that of its predecessor.

169

In this section we shall specifically define $z_j(t)$ as the solution of (2.6) that satisfies the conditions

$$z_o(t) \text{ is given for } t > 0$$

and   (3.1)

$$z_j(t) = 0 \text{ for all } t \le jT.$$

This means that all cars are initially at rest and subsequently following a prescribed motion of the zero[th] car. Since (2.6) is linear in $z_j(t)$, however, the sum of this $z_j(t)$ and any other solution of (2.6) is also a solution. In particular

$$z_j(t) + (t-j/\alpha)v$$

is a solution of (2.6) with values $(t-j/\alpha)v$ for $t \le jT$ and thus represents the motion of a series of cars starting from a steady flow with velocity v.

If

$$f^*(s) = \int_0^\infty dt e^{-st} f(t) \qquad (3.2)$$

denotes the Laplace transform of f(t), then the Laplace transform of (2.6) with the conditions (3.1) gives

$$se^{sT}z_j^*(s) = \alpha[z_{j-1}^*(s) - z_j^*(s)]$$

or

$$se^{sT}z_j^*(s) = \alpha[z_{j-1}^*(s) - z_j^*(s)]$$

or

$$z_j^*(s) = z_{j-1}^*(s)[1 + se^{sT}/\alpha]^{-1} = z_o^*(s)[1 + se^{sT}/\alpha]^{-j} . \qquad (3.3)$$

The purpose of the Laplace transform was to obtain this simple difference equation for the $z_j^*(s)$. The method is also useful if different cars have different values of α, say, α_k. One then obtains

170

$$z_j^*(s) = z_o^*(s) \prod_{k=1}^{j} [1 + se^{sT}/\alpha_k]^{-1} ,$$

and can study statistical problems associated with the propagation of disturbances through cars with randomly selected properties. We shall not pursue this here, however.

We obtain the trajectories themselves from (3.3) by inverting the Laplace transform using the formula

$$z_j(t) = \frac{1}{2\pi i} \int_{\gamma-i\infty}^{\gamma+i\infty} ds \, e^{st} z_j^*(s) = \frac{1}{2\pi i} \int_{\gamma-i\infty}^{\gamma+i\infty} ds \, e^{st} \frac{z_o^*(s)}{[1 + se^{sT}/\alpha]^j} \qquad (3.4)$$

in which the integral is a line integral in the complex s-plane along some vertical path from $\gamma -i\infty$ to $\gamma +i\infty$ with γ so chosen that the path lies to the right of singularities of the integrand.

If we assume that $z_0(t)$ is finite (as it certainly must be) and that after some finite time Γ is a motion with constant velocity v_0.

$$z_0(t) = v_0 t + a_0 \quad for \quad t > \Gamma$$

then

$$z_0^*(t) = \int_0^\Gamma dt e^{-st} [z_0(t) - v_0 t - a_0] + \int_0^\infty dt e^{-st}(v_0 t + a_0)$$

$$= \frac{v_0}{s^2} + \frac{a_0}{s} + \text{an entire function}^1 \text{ of s.}$$

$$= \frac{v}{s^2} + \frac{v}{s} + \text{an entire function}^1 \text{ of s.}$$

Thus the only singularity of $z_0^*(s)$ with $|s|$ finite, for this fairly general class of motions, is at $s = 0$. Furthermore, the entire function is no larger in amplitude than some constant times $|e^{-s\Gamma}|$ in the left half-plane.

---

[1] An "entire function" is a function which is analytic (has a Taylor series expansion) at every point in the finite complex s-plane.

It is not possible in general to evaluate (3.4) in closed form but one can deduce from it the behavior of $z_j(t)$ for large t and/or large j which was particularly awkward to obtain from (2.4). We consider first the behavior for large t and fixed j.

For $t > \Gamma$, the integral (3.4) can be evaluated if we close the contour from $\gamma - i\infty$ to $\gamma + i\infty$ by a large (infinite) semi-circle in the left half-plane. The numerator of the integrand $e^{st}z_o^*$ vanishes for $s \to \infty$ at least as fast as $|\exp[s(t-\Gamma)]|$ in the left half-plane. The denominator has an infinite number of zeros as described below which extend to ∞ in the left half-plane, but one can always find a semi-circle of arbitrarily large radius which avoids these zeros and along which this factor of the integrand is bounded. The integral along the semi-circle vanishes and so (3.4) can be expressed as the sum of the residues of the integrand at each of its poles.

The second order pole at $s = 0$ of the factor $z_o^*(s)$ contributes a term

$$[t-j/\alpha]v_o+a_o$$

which we recognize as the equilibrium solution for motion with a constant velocity v_o. The other singularities of the integrand occur at points s_k which are roots of the equation.

$$a+s_k \exp (s_k T) = 0. \tag{3.5}$$

At each such root the integrand has a pole order of j and a residue of the form

$$P_{j-1}(s_k,t) \exp (s_k t)$$

At each such root the integrand has a pole order of j and a residue of the form

$$P_{j-1}(s_k,t) \exp (s_k t)$$

in which $P_{j-1}(s_k,t)$ is a polynomial in t of degree j-1. The integral thus has the form

$$z_j(t) = [t - j/\alpha]v_o + a_o + \sum_k P_{j-1}(s_k , t) \exp (s_k t) . \tag{3.6}$$

If we number the roots according to the values of Real s_k so that Real $s_1 \geq$ Real $s_2 \geq$ ..., then for large t the dominant terms of the series will be the ones with the largest values

172

for Real s_k. If Real $s_1 > 0$ this means that the motion is unstable because $z_j(t)$ grows exponentially with t. It is also unstable if Real $s_1 = 0$. In such cases each following car is so sensitive to any changes in the motion of the lead car that it is unable to control its motion. If s_1 is complex but $Re\ s_1 < 0$ the motion is stable in the sense that the transient terms decay but the transient will be oscillatory with exponentially decreasing amplitude. Since s_k and $\bar{s}_k$ both[2] satisfy (3.5), it follows that, if s_1 is complex, $\bar{s}_1$ is also a root and the two terms of (3.6) combine to give a real valued contribution to $z_j(t)$. Finally if s_1 is real and negative, the transient is a non-oscillatory exponential decay.

To locate the roots of (3.5) we note that this equation implies that the s_k are solutions of the two real equations.

$$Arg\ S + Im\ S = (2n+1)\pi \tag{3.7}$$

$$\alpha T = |S|\ exp\ (Re\ S) \tag{3.8}$$

in which $S = sT$, Arg S is the argument or complex phase of S, Im S is the imaginary part of S, Re S the real part of S, and n is any integer. Equation (3.7) does not contain αT, and so all roots s_k (for any αT) must lie on the curves in the S-plane which satisfy (3.7). These curves are shown in Figure IV 6 by the solid lines. One branch of (3.7) is the negative real axis. Another starts at (-1,0), moves upward and to the right having the line $i\pi$ as an asymptote. It also moves downward and to the right with $-i\pi$ as an asymptote. All other real axis. Another starts at (-1,0), moves upward and to the right having the line $i\pi$ as an asymptote. It also moves downward and to the right with $-i\pi$ as an asymptote. All other branches vary from an asymptote $2n\pi i$ at the left to $(2n+1)\pi i$ at the right if n = 1,2, -- or $(2n-1)\pi i$ at the right if n = -1, -2, --.

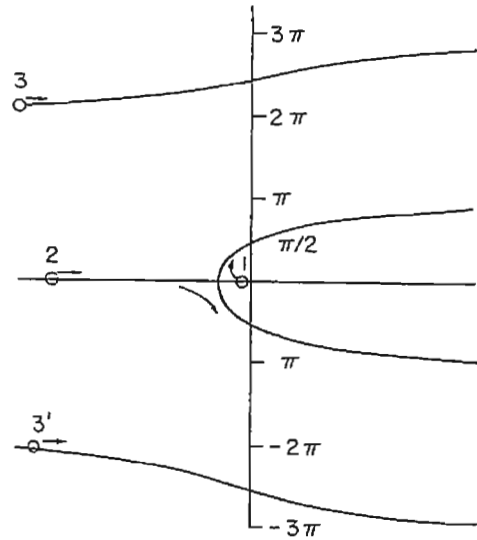[2] $\bar{s}_k$ is the complex conjugate of s_k

Fig. IV 6
Locus of roots in the complex plane

For $\alpha T \ll 1$ in (3.8), we see that either $|S| \ll 1$ and, because of (3.7) or Figure IV 6, S is negative real, or $-Re\,S \gg 1$ and because of (3.7) or Figure IV 6, $Im\,S \sim 2\pi n$ for some integer n. Thus the root with largest $Re\,S$ is

$$S = -\alpha T \exp(-ReS) = -\alpha T[1+\alpha T + \dots]. \qquad (3.9)$$

The other roots are at

$$ReS = -|\log \alpha T| - \log |S|$$

$$ReS = -|\log \alpha T| - \log |S|$$

or

$$S \sim 2\pi ni - |\log \alpha T| - \log |2\pi ni + \log \alpha T|. \qquad (3.10)$$

These roots are shown in Figure IV 6 by the circles numbered in order of decreasing $Re\,S$. The arrows show the direction in which the roots move as αT increases. The roots 1 and 2 are moving toward each other, and one can readily verify that when $\alpha T = 1/e$ the two roots

174

coalesce at $S = -1$. For $\alpha T > 1/e$ there are no negative real solutions of (3.8). The two roots which combined at $S = -1$ now move off the axis along the other curve of Figure IV 6 that crosses $S = -1$ as shown by the arrows. As αT increases further these roots will cross the imaginary axis and acquire positive real components at $S = i\pi/2$ when $\alpha T = \pi/2$. Throughout this variation of αT the roots 1 and 2 remain the ones with the largest real part.

We thus conclude that for $0 < \alpha T \leq 1/e$, the longest surviving transient is stable and non-oscillatory, for $1/e < \alpha T < \pi/2$ the transient is oscillatory but stable, and for $\pi/2 \leq \alpha T$ the motion is unstable.

Some numerical evaluations of the complete trajectories have been made for various initial conditions by Kometani and Sasaki [7] and by Herman, Montroll, Potts, and Rothery [9].

To estimate $z_j(t)$ when j is large but $\alpha T < \pi/2$, we choose a path of integration in (3.4) to the right of the origin but otherwise up the imaginary axis, $\gamma = 0$. Along this path with $s = i|s|$,

$$|1 + se^{sT}/\alpha|^j = |1 + \frac{|s|^2}{\alpha^2} - 2\frac{|s|}{\alpha}\sin(|s|T)|^{j/2} .$$
(3.11)

If $\alpha T < 1/2$ this function has a minimum at $s = 0$. An expansion in powers of $|s|$ gives

$|-2|

gives

$$|1 + se^{sT})/\alpha|^j = |1 + \frac{|s^2|}{\alpha^2}(1 - 2\alpha T) + .. |^{j/2}$$

which for large j will be a very sharp one. Except near $s = 0$ the integrand of (3.4) will, therefore, be relatively small and it suffices to estimate the integrand only near $s = 0$. Here

$$\left(1 + \frac{s e^{sT}}{\alpha}\right)^j \sim \left(1 + \frac{s}{\alpha} + \frac{s^2 T}{\alpha}\right)^j \sim \exp\left[j\frac{s}{\alpha} - \frac{js^2}{2\alpha^2}(1 - 2\alpha T) + ..\right]$$

and so

$$z_j(t) \sim \int_{-i\infty}^{+i\infty} ds \; e^{s(t-j/\alpha)} \; e^{\frac{js^2}{2\alpha^2}(1 - 2\alpha T)} \; z_o^*(s) \;. \tag{3.12}$$

For large j this integrand will be highly oscillatory along the imaginary axis unless t - j/α is sufficiently small. An oscillation causes cancellation of positive and negative terms and so this integral will be relatively small, except for contributions from any poles of $z_0^*(s)$ at the origin, unless the first factor of the integrand varies with s at a rate comparable with the second for large j; i.e., for

$$t - j/\alpha = O\left[j\frac{(1-2\alpha T)}{2\alpha^2}\right]^{1/2} \;.$$

If we apply the convolution theorem to (3.12) we obtain for $z_j(t)$ the expression derived earlier in (2.14). Although this derivation is also somewhat heuristic, it is possible to estimate error terms for (3.12) and so obtain a more convincing argument.

If $1/2 < \alpha T < \alpha/2$, the asymptotic form of (3.4) is more complicated. The factor (3.11) no longer has a minimum at $|s| = 0$ but a local maximum. There are minima, however, at points $s = s_0$ along the imaginary axis where the derivative of (3.11) vanishes namely where

points $s = s_0$ along the imaginary axis where the derivative of (3.11) vanishes namely where

$$\frac{1}{\alpha T} = \frac{\sin|s_o T|}{|s_o T|} + \cos|s_o T|$$

or if αT is close to 1/2 at

176

$$\frac{2}{3} |s_o T|^2 \sim 2 - \frac{1}{\alpha T} \, .$$

The absolute value of the integrand of (3.4) will have a sharp maximum at $s = s_o$ for large j. We can expand the integrand in the vicinity of s_o in much the same manner as in (3.12) and so obtain the asymptotic form for $z_j(t)$. The details are too awkward to describe here. It suffices to note that the dominant factor of the integral will be the value of the integrand at $s = s_o$, particularly $|1 + (s_o/\alpha) \exp(s_o T)|^{-j}$, which is the jth power of a number of absolute value greater than 1. This already points to an exponential growth in $z_j(t)$ with j. For any large fixed j, the time dependence will be oscillatory with a wave number approximately $|s_o|$ coming from the factor $\exp(s_o t)$ of the integrand at s_o, but with a time varying amplitude that is nearly Gaussian shaped with a peak near $t = j/\alpha$ coming from the integral itself.

The main conclusion is that, for $1/2 < \alpha T$, a disturbance propagates with increasing amplitude from one car to the next so that even a small perturbation of the trajectory $z_o(t)$ will eventually (for large j) grow into one of arbitrarily large size. Some actual graphs of such trajectories are given in [9].

Other types of linear theories have also been proposed in which velocities or accelerations of a jth car are linearly dependent upon positions, velocities, etc., of the $(j-1)^{th}$ and/or the $(j-2)^{th}$ cars but the main weakness of these models seems to be the assumption of accelerations of a jth car are linearly dependent upon positions, velocities, etc., of the $(j-1)^{th}$ and/or the $(j-2)^{th}$ cars but the main weakness of these models seems to be the assumption of linearity and not the detailed form of the linear equations. We turn next to some nonlinear models.

4. Continuum theories. Even before the linear theories described above had developed very far, a different type of theory was proposed independently by Lighthill and Whitham (1955) [15] and by Richard (1956) [16]. This was a nonlinear theory but one in which discrete cars were replaced by a continuum much as, in fluid dynamics, an atomic gas is treated as a continuous fluid. Since the theory describes in relatively simple form most of the important large scale effects associated with a nonlinear discrete theory such as (1.8), it is desirable to consider this theory before returning to the nonlinear discrete case.

The basic assumption in the continuum theory is that the relations between q and k which were described in sec IV 1 for a steady flow also apply to time-dependent flow. Thus, if k = k(x,t) and q = q(x,t) are functions of position and time, then these are related by the equation

$$q(x,t) = Q(k(x,t)) \tag{4.1}$$

in which Q(k) is the function determined experimentally from the steady state relation q = Q(k) when k is independent of x and t.

In order for k(x,t) and q(x,t) to be experimentally well defined, it is necessary that they vary so slowly that they are nearly constant over distances containing many cars, otherwise the meaning of the continuum approximation becomes questionable. Even with this, however, the assumption (4.1) is not as innocent as it may at first seem. It certainly does not otherwise the meaning of the continuum approximation becomes questionable. Even with this, however, the assumption (4.1) is not as innocent as it may at first seem. It certainly does not hold for very low densities where, as we have seen in chapter II, the flow q(x,t) depends upon not only the density but also the velocity distribution, which is itself time and space dependent. The continuum theory, in fact, does not contain the velocity distribution at all. We shall see in the next section, however, that the theory does describe a well defined limit for a

178

theory with no passing such as (1.8), but to what extent it is also a reasonable approximation for moderately dense traffic in which passings do occur has not yet been determined.

Consider any interval of highway of length Δx at x, which we will think of as being very small on a scale of variation of $k(x,t)$ but which must still contain many cars. The net rate at which cars enter Δx at time t is

$$q(x,t)- \ q(x + \Delta x,t) \sim -\Delta x \partial q(x,t)/\partial x.$$

Since cars are "conserved," and cannot vanish, this must be the rate of change of the number in Δx, $\Delta x \partial \ k/\partial x$. Thus

$$\frac{\partial k(x \ , \ t)}{\partial t} + \frac{\partial}{\partial x} q(x \ , \ t) = 0 \ . \tag{4.2}$$

This is a type of equation familiar in many branches of physics for which there is a conservation of mass, momentum, or energy, etc. In a stochastic model of traffic, this equation also applies if k and q are interpreted as expectations in the sense of chapter II.

The combination of (4.1) and (4.2) with a specified function $Q(k)$ gives a self-contained theory because we can eliminate either q or k from (4.1) and (4.2) to obtain a single equation for k or q. For example

$$\frac{\partial k(x \ , \ t)}{\partial t} + Q'(k(x \ , \ t)) \frac{\partial k(x \ , \ t)}{\partial x} = 0 \tag{4.3}$$

in which

$$\frac{}{\partial t} \quad \approx \ (k(x \ , \ t)) \ \frac{}{\partial x} \quad \vee$$

in which

$$Q'(k(x \ , \ t)) \equiv \frac{dQ(k)}{dk} \ | \ k = k(x \ , \ t) \ . \tag{4.4}$$

The solution of (4.3) for $k(x,t)$ will then also determine $q(x,t)$ through (4.1).

This is again a first order partial differential equation such as (2.7), but this equation is nonlinear since Q' depends upon k. We can interpret (4.3) as stating that the derivative of $k(x,t)$ at the point (x,t) vanishes along the direction $(Q',1)$ in the (x,t) plane. If $k(x,t)$ does not vary in this direction neither can the direction $(Q',1)$ since it depends only upon k.

If we know $k(x,0)$ and the function $Q(k)$, we can construct the solution of (4.3) as follows. In the (x,t) plane we draw through each point $(x,0)$ a straight line of slope $Q'(k(x,0))$ as in figure IV 7. This is known, since, for each x, we know $k(x,0)$ and for each such value of k, we can determine $Q'(k(x,0))$ the slope of the q vs k curve for that value of k. Everywhere along this line the derivative of k vanishes and so $k(x,t)$ is constant. Figure IV 7 thus represents the "contour" map of $k(x,t)$.

The formal solution of (4.3) is

$$k(x + tQ'(k(x,0)),t) = k(x,0) \quad \textit{for all t.} \tag{4.5}$$

The lines of Figure IV 7 would be called "characteristics" in the theory of partial differential
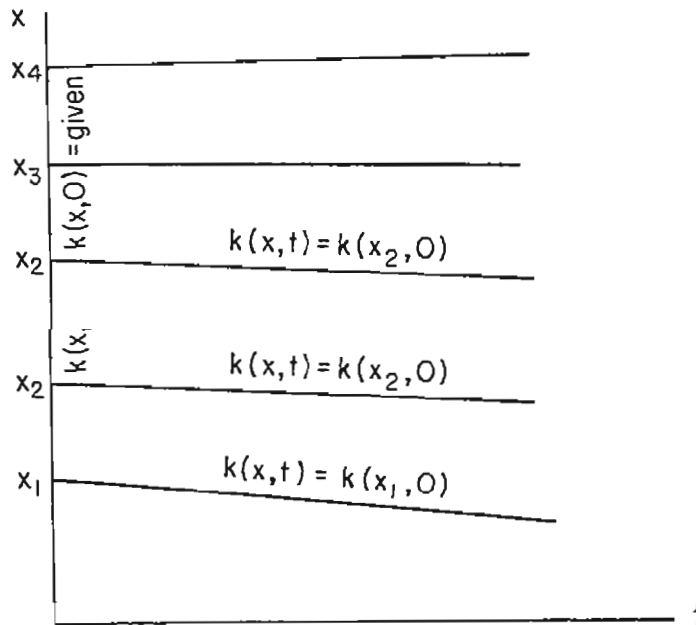

Fig. IV 7
Lines of constant density

equations. Here we call them waves of constant density. Since the value of k determines q and, therefore, also the average velocity v = q/k, these waves must represent the same waves described previously in section 1 of Chapter IV. The wave velocity is Q' (k). Because of the nonlinearity of the equations, the wave velocity is now a function of k, however.

In the continuum theory and in most macroscopic studies of traffic, it is common practice to postulate a relation between q and k as done here. In discrete theories, it has been customary to consider relations between the average velocity v and the average spacing D. The macroscopic quantities are related, however, by

$$q = kv \text{ and } d = 1/k$$

so that any relations between q and k or between v and d implies a relation between any other pair of these four quantities. If spacings differ from one car to the next, these macroscopic quantities can, of course, be identified by suitable expectations as in (1.2) and (1.3).

For the q vs k curve shown in Figure IV 8, the wave velocity at any value of k,k_o, is
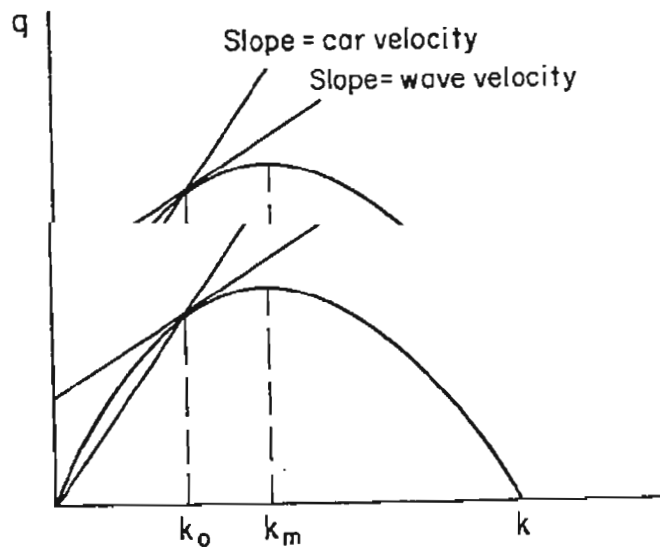

Fig. IV 8
Interpretation of the wave velocity

181

the slope $Q'(k_o)$ of the tangent line at k_o. This is positive (a disturbance propagates forward) if $k_o < k_m$, it is zero (the wave of maximum q is stationary) if $k = k_m$, and is negative (a wave propagates backwards) if $k_o > k_m$. The average car velocity, $v = q/k$, is the slope of the line from the origin to (q,k).

For the v vs. d curve of Figure IV 9, the wave velocity is[3]

$$\frac{dq}{dk} = \frac{d(v/d)}{d(1/d)} = v - d\frac{d(v)}{d(d)} \ . \tag{4.6}$$

At any spacing d_o, this is the velocity intercept of the tangent line to the v vs. d curve at d_o. The flow $q = v/d$ at d_o is the slope of the line from the origin to (v_o, d_o).

In section 2, we postulated a linear relation between v and d, equation (2.1) which, for small variations in d, we were prepared to interpret as the tangent to a nonlinear v vs. d curve. This line was assigned a slope α and velocity intercept $(-\beta)$. We later found that the wave velocity for this discrete linear theory, Figure IV 4, was also $(-\beta)$ which agrees with the value deduced here from the continuum theory.

One of our objectives in considering the continuum theory was to investigate the effects of the nonlinearity. Since waves of different density travel with different wave velocities as shown in Figure IV 7, these waves may either fan out as shown or converge and intersect. If $k(x,0)$ is a decreasing function of x, or equivalently if the velocity is an velocities as shown in Figure IV 7, these waves may either fan out as shown or converge and intersect. If $k(x,0)$ is a decreasing function of x, or equivalently if the velocity is an increasing function of x, so that cars are accelerating, then, from Figures IV 8 or IV 9, we notice that Q', assumed to be a decreasing function of k, will be an increasing function of x.

---

[3] The notation in which d represents the spacing is quite awkward here, because d is also used to denote a differential. The reader will need to interpret the meaning of d from the way the symbol appears.
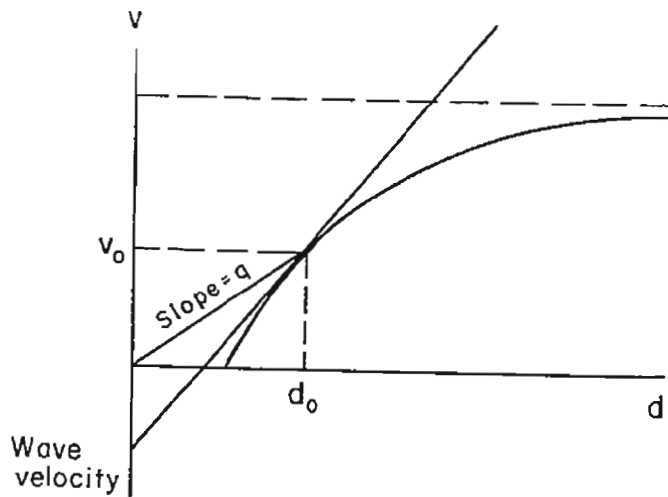
Fig. IV 9
Evaluation of the wave velocity
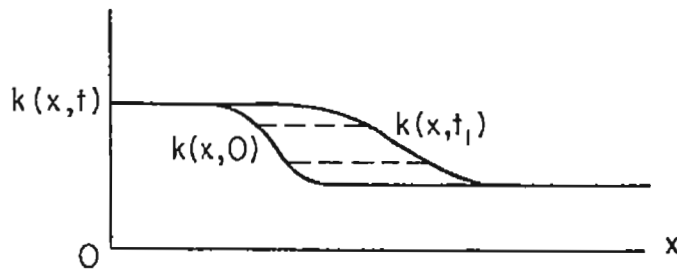
Fig. IV 10
Dispersion of an acceleration wave

This corresponds to the situation illustrated in Figure IV 7 in which the higher velocity waves are initially ahead of the slower velocity waves and so the region of acceleration tends to expand linearly with time as the waves spread further and further apart. Figure IV 10 shows a decreasing density $k(x,0)$ at time zero by a solid line. A short time later the high density

183

region has moved forward perhaps as indicated by the short arrow but not as much as the low density region. A sharp initial change in density thus tends to disperse.

Although the linear discrete theory also predicted a dispersion of an acceleration wave, it was of a quite different type. In the linear theory the dispersion increased as $t^{1/2}$, like a diffusion of particles, but here the dispersion is due to a velocity difference for propagation and increases as t; it behaves like the dispersion due to a distribution of velocities as described for low density traffic.

If k(x,0) is an increasing function of x, i.e., cars are decelerating, then the higher velocity waves at low density are initially behind the slower velocity wave for the high density. A gradually increasing k(x,0) tends to become steeper as shown in Figure IV 11 and eventually the profile of k(x,t) will have a vertical tangent. This occurs as soon as any two waves of neighboring values of k in Figure IV 7, which are now converging, actually intersect. At this time the solution (4.5) breaks down and, if it were continued beyond this time, would assign to some (x,t) point more than one value of k, because more than one wave passes through this point. The existence of a vertical tangent, however, means that ∂k/∂x becomes infinite and the differential equation from which (4.5) was derived is no longer valid.
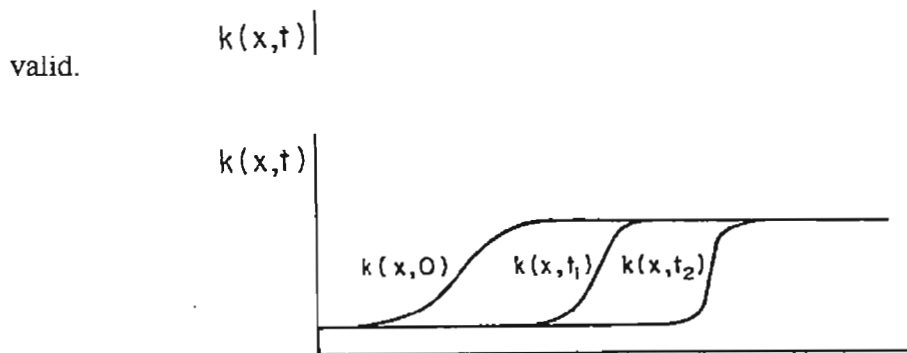
valid.

$k(x,t)|$

$k(x,t)|$

Fig. IV 11
Focusing of deceleration waves

184

A phenomena such as this is familiar in fluid dynamics. The failure of the differential equation is corrected by allowing discontinuities in k(x,t). In fluid dynamics such discontinuities are called shocks and this term has been carried over into traffic theory. Although the shock is interpreted here as a mathematical discontinuity in the equations for a continuous fluid-like motion of cars, it should be remembered that the continuum equations were meant to describe only density changes which are small over distances comparable with the distance between cars. If one did have a large change in density over a distance comparable with the distance between cars, then one could only represent this as a discontinuity in the continuum approximation.

The differential equation (4.3) for this continuum theory was derived from a conservation principle which itself does not require k to be differentiable. Suppose we take any section of highway between points a(t) and b(t) in which a(t) ≤ b(t) are differentiable functions of t and such that k(x,t) is continuous at a(t) and b(t). The rate at which cars enter the interval (a(t), b(t)) at point a(t) is the relative velocity between the cars and the point a(t) times the density, i.e.,

$$\left[v(a(t) , t) - \frac{da(t)}{dt} \right] k(a(t) , t) .$$

Similarly the rate at which they enter at b(t) is

$$\llcorner \ulcorner \qquad\qquad \lrcorner \urcorner$$

Similarly the rate at which they enter at b(t) is

$$- \left[v(b(t) , t) - \frac{db(t)}{dt} \right] k(b(t) , t) .$$

The conservation principle implies that this must be the rate of change of the number in (a(t), b(t)); i.e.,

185

$$\frac{d}{dt} \int_{a(t)}^{b(t)} k(x , t)dx = \left[v(a(t) , t) - \frac{da(t)}{dt} \right] k(a(t) , t)$$

(4.7)

$$- \left[v(b(t) , t) - \frac{db(t)}{dt} \right] k(b(t) , t) .$$

If $\partial k(x,t)/\partial x$ and $\partial k(x,t)/\partial t$ exist, let $a(t) \to b(t)$ uniformly in t; for example let $b(t) - a(t) = \epsilon$ independent of t and let $\epsilon \to 0$. Then the left-hand side of (4.7) becomes approximately

$$\epsilon \frac{d}{dt} k(a(t) , t) = \epsilon \left[\frac{\partial}{\partial t} k(x , t) + \frac{da(t)}{dt} \frac{\partial k(x , t)}{\partial x} \right] x = a(t) ,$$

and the right-hand side becomes approximately

$$- \epsilon \left\{ \frac{\partial}{\partial x} \left[v(x , t) k(x , t) \right] - \frac{da(t)}{dt} \frac{\partial k(x , t)}{\partial x} \right\} x = a(t) .$$

Equating these, we recover the partial differential equation (4.3), as we should.

If, however, there is a path of discontinuity between a(t) and b(t) and we let a(t) and b(t) converge on the two sides of this discontinuity we obtain for $a(t) \to b(t)$ that the left-hand side of (4.7) must go to zero, thus

$$\left[v(a(t) , t) - \frac{da(t)}{dt} \right] k(a(t) , t) = \left[v(b(t) , t) - \frac{da(t)}{dt} \right] k(b(t) , t)$$

(4.8)

$$\left[v(a(t) , t) - \frac{da(t)}{dt} \right] k(a(t) , t) = \left[v(b(t) , t) - \frac{da(t)}{dt} \right] k(b(t) , t)$$

(4.8)

in which k(a(t),t) and k(b(t),t) are the values of k on either side of the curve a(t) = b(t). This equation simply implies that the flow into the discontinuity must be the same as the flow out. This equation is called the shock equation and da(t)/dt is called the shock velocity.

186

Since k, q, v, and d are all related to each other through the q - k or v - d relations, we can also write (4.8) in the form

$$\frac{da(t)}{dt} = \frac{q_b - q_a}{k_b - k_a} = v_a - \frac{(v_b - v_a)}{(d_b - d_a)} d_a \qquad (4.9)$$

in which q_a, q_b, etc. represent q(a(t),t), q(b(t),t), etc., the values of these quantities on the two sides of the shock discontinuity.

In terms of the q vs k curve, (4.9) implies that the shock velocity is the slope of the line joining (q_b, k_b) with (q_a, k_a) as shown in Figure IV 12. In the limit of very weak shocks, $k_b \rightarrow k_a$, this shock line converges to the tangent line of the q - k curve and the shock velocity becomes the wave velocity. Similarly (4.9) shows that the shock velocity is the velocity intercept of the line joining (v_b, d_b) with (v_a, d_a), as shown in Figure IV 13 which, for weak shocks, converges to the tangent line of the v - d curve that determines the wave velocity.
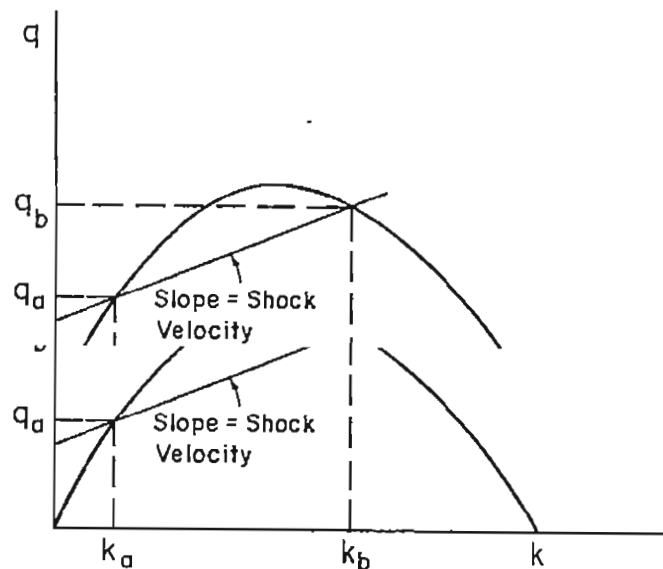

Fig. IV 12
Evaluation of the shock velocity from the q-k curve
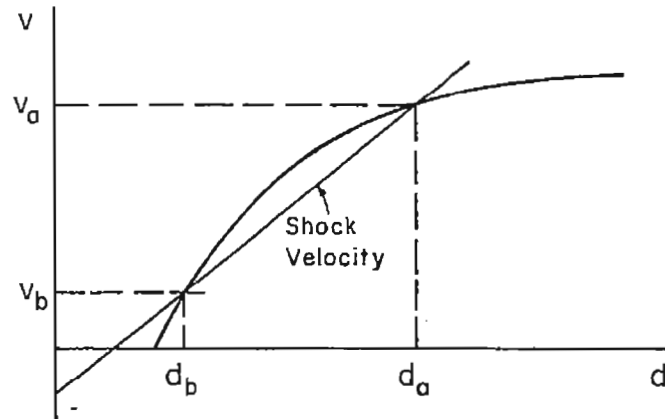
Fig. IV 13
Evaluation of shock velocity from the v-d curve

To evaluate $k(x,t)$, (4.8) or (4.9) is used mainly to determine the path of the shock. Figure IV 14 illustrates how in a typical problem one can construct the solution $k(x,t)$ graphically. If one is given $k(x,0)$ then one can draw the waves of constant k, or use the solution (4.5), to determine $k(x,t)$ at least until such time t_0 when two waves first intersect at a point (x_0, t_0) of Figure IV 14. At this moment a shock starts to form.

Except for very exceptional functions $k(x,0)$ for which there is a perfect focusing of the waves such that waves with a non-zero range of k values converge on the single point O, the shock starts as an "infinitesimal shock"; the intersecting waves have nearly the same velocity. The shock also will start with a velocity equal to those of the two intersecting the shock starts as an "infinitesimal shock"; the intersecting waves have nearly the same velocity. The shock also will start with a velocity equal to those of the two intersecting waves. Furthermore it is clear, from figures IV 12 or IV 13, that the shock velocity for a jump in density from k_a to k_b actually has a value between the values for the wave velocities of densities k_a and k_b, i.e., the slope of the shock line in Figure IV 12 must be less than the slope of the tangent at one of the densities k_a or k_b and larger than the slope of the tangent at

the other. For weak shocks, the shock velocity is approximately the average of wave velocities on either side.

Another method to determine x_o and t_o is based upon the following arguments. From Figure IV 14 we see that two waves starting at points x and $x + a$ will intersect at time t_o and position x_o provided

$$x_o = x+c(x)t_o = x+a+c(x+a)t_o \qquad (4.10)$$
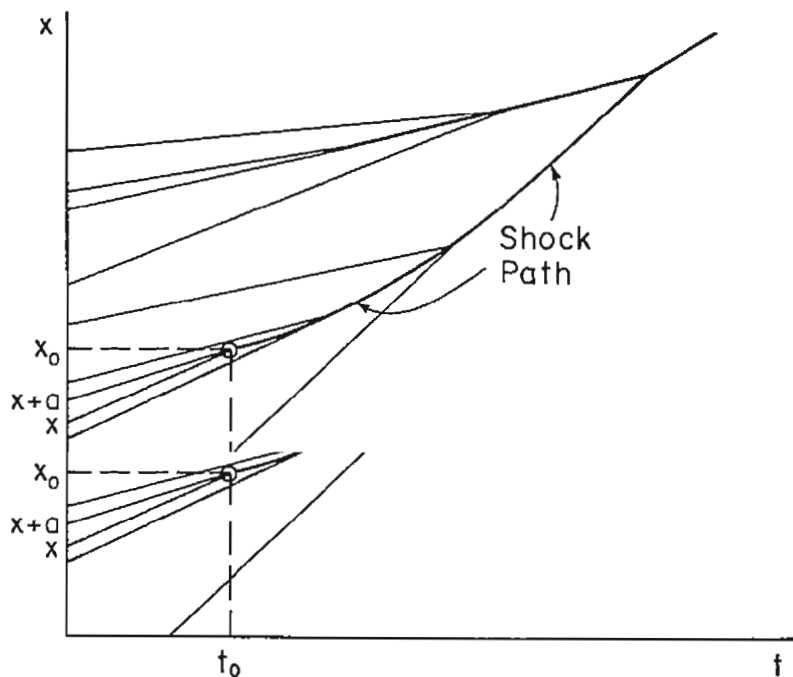
where

$$c(x) = Q'(k(x,0)) \qquad (4.11)$$

Fig. IV 14
Construction of the shock path

189

is the initial wave velocity at x. Since t_o is the first time any two waves intersect, it follows that

$$\frac{1}{t_o} = \max_{x,a} \frac{c(x) - c(x + a)}{a} . \qquad (4.12)$$

If c(x) has a continuous derivative, it follows also from the mean value theorem that - [c(x)-c(x+a)]/a is equal to the derivative of c at some point between x and x+a. Therefore

$$\frac{1}{t_o} = \max_x (-) \, dc(x)/dx . \qquad (4.13)$$

The values of x for which the maximum is realized also give the values of x in (4.10) that determine the position x_o where the shock originates.

We now know where the shock starts and its initial velocity, thus its position a short time later, at time $t_o + \Delta t$. If no new shocks form between time t_o and $t_o + \Delta t$, the values of k on either side of the shock at time $t_o + \Delta t$ are still determined by the waves starting at $t = 0$. The density on one side of the shock is determined by the wave which intersects the shock at time $t_o + \Delta t$ and approaches the shock from the same side. Another wave will approach from the other side and determine the density on that side. We now know the densities on either side of the shock at time $t_o + \Delta t$ and we can reevaluate the shock velocity at time $t_o + \Delta t$.

We can then estimate the shock position at a still later time $t_o + 2\Delta t$ perhaps, find the side of the shock at time $t_o + \Delta t$ and we can reevaluate the shock velocity at time $t_o + \Delta t$.

We can then estimate the shock position at a still later time $t_o + 2\Delta t$ perhaps, find the densities at this time etc. In essence we are performing a graphical integration of (4.9) for the shock path a(t).

If some other shocks should form elsewhere we proceed similarly to find their paths. If two shocks should intersect, we simply combine them into a single shock as shown in

Figure IV 14. The shock velocity for the single shock is determined by the densities in regions adjacent to the shock not including the region annihilated in the collision.

As a practical example of how the continuum theory can be applied, we consider what happens, according to this theory, when a steady flow of traffic is suddenly stopped, at a road intersection perhaps, and then released again, as would occur at a traffic signal. The trajectories of cars which the theory should predict are shown in Figure IV 15. The stopping point is at $x = 0$ and the first car to be stopped is designated by $x_1(t)$. The continuum theory does not describe in detail the trajectory of car 1. If we are given the density k_i, the flow q_i, or the velocity v_i of the initial approaching traffic stream, all other quantities are also determined by the q vs. k relations. Thus, in particular, the approach velocity of car 1 is specified. The theory also predicts the final velocity v_f of the lead car. If the car is stopped long enough, this car will have a nearly empty road ahead, i.e., $k_f = 0$ and the corresponding final velocity for $k_f = 0$ is also specified by the q vs. k relations. The continuum theory does not describe the details of the deceleration and acceleration of car 1. However, on a scale of time and distance in which the continuum theory is meant to apply, we expect the deceleration and acceleration of car 1 to take a negligible length of time.

The deceleration of the first car, being nearly instantaneous, creates a shock wave immediately. It is a shock from the initial state q_i, k_i on the q vs. k curve to the state $q = 0$,

The deceleration of the first car, being nearly instantaneous, creates a shock wave immediately. It is a shock from the initial state q_i, k_i on the q vs. k curve to the state $q = 0$, $k = 1/D(0)$ as shown in Figure IV 16. The slope of the shock line of Figure IV 16 between states determines the shock velocity. In Figure IV 15, the shock starts at $x = 0$, $t = 0$ and travels backwards with constant speed as shown by the broken line. This shock line represents, in effect, the rear of the queue caused by the traffic signal and the continuum

191

approximation implies that this deceleration occurs practically instantaneously (on a scale of distance in which the distance between cars is also essentially zero). The shock relations are only relations describing the conservation of cars and, in Figure IV 15, if we specify the spacing between the approaching cars and the spacing between these cars when they are stopped, the shock line occurs at the only place it can which will guarantee the continuity of these trajectories; namely, at the intersections of the trajectories for the approaching cars and their trajectories when stopped.

When the lead car accelerates again, it sends out a fan of acceleration waves for all car velocities between 0 and v_f. The slowest wave travels backwards with the wave velocity associated with car velocity 0. This wave velocity is given by the slope of q vs. k curve at
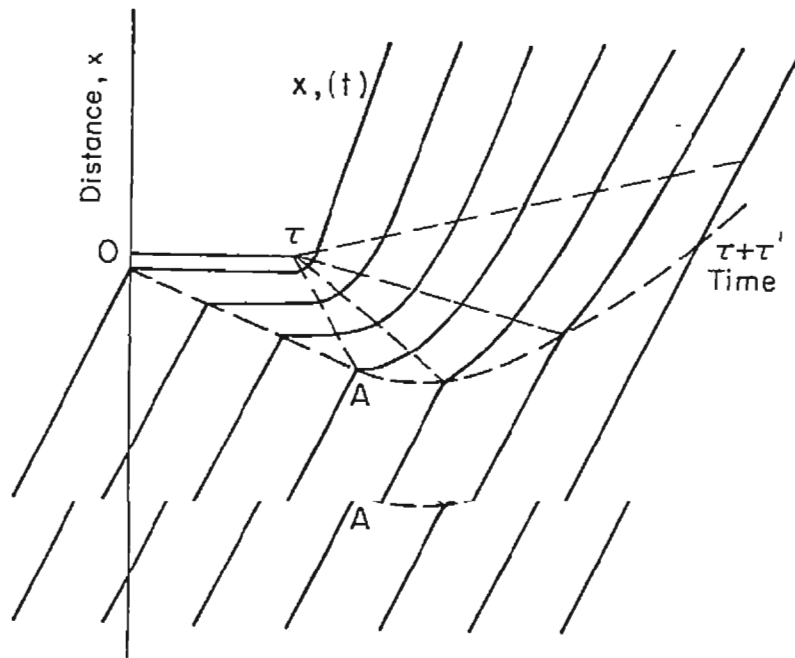
Fig. IV 15
Cars stopped at a traffic signal

$k = 1/D(0)$ as shown in Figure IV 16. When this wave starting at $x = 0$, $t = \tau$ as shown by the broken line of Figure IV 15 intersects the shock line at point A, this signals the car at the

192
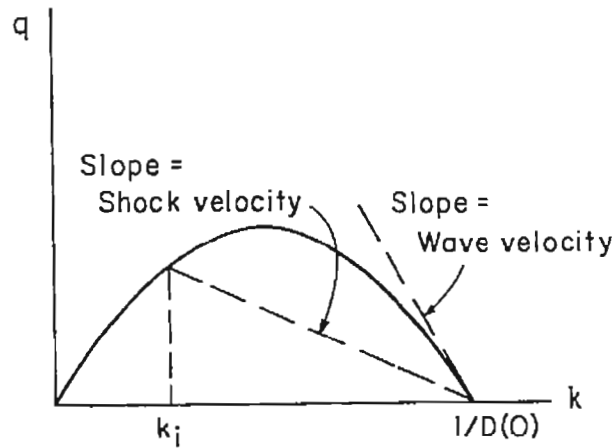
Fig. IV 16

Evaluation of flows from a traffic signal

tail of the queue to move. The shock does not disappear. The waves that intersect the shock only assign new values for the density or car velocity on the front side of the shock and these in turn assign to the shock a new velocity as given by (4.7) or (4.8). As the car velocity at the front of the shock increases, the shock gains forward speed. It eventually moves with a positive velocity and crosses the intersection or coordinate x = 0.

The time at which the shock crosses the intersection can be found very easily. We know that the wave with wave velocity zero is the wave corresponding to q_m since the tangent to the q vs. k curve is horizontal at $q = q_m$. The number of cars that cross the intersection before the shock arrives is therefore q_m multiplied by the time τ' for the shock to arrive. This must. however. also be equal to the total number of cars that have arrived by this time, before the shock arrives is therefore q_m multiplied by the time τ' for the shock to arrive. This must, however, also be equal to the total number of cars that have arrived by this time, i.e.,

$$q_i(\tau + \tau') = q_m \tau' \tag{4.14}$$

in which τ is the length of time the traffic is stopped; τ, q_i, and q_m are known.

193

The shock, as it moves forward, becomes weaker and weaker. The car velocity at the rear of the shock remains at v_i but the car velocity at the front keeps increasing. Eventually, the shock will overtake all waves for car velocities less than v_i and the shock will degenerate into a wave or a shock of zero jump. The waves of car velocity larger than v_i, however, move forward faster than the shock and the shock never reaches them.

There are many equivalent ways of constructing the trajectories and waves of Figure IV 15, but, in any case, they are uniquely determined by the q vs. k curve. In this simple example where all waves originate from the same point one can give a fairly simple explicit form of the solution.

Finally, we have already noted that the continuum theory and the linear discrete theory agree to "first order"; i.e., they give essentially the same velocity for propagation of a wave, but the "second order" effects, such as the spreading of a disturbance as it propagates, are quite different in the two theories. In the linear theory disturbances typically spread as $t^{1/2}$ whereas in the nonlinear theory an acceleration wave spreads proportional to t and a deceleration wave compresses to form a shock. These differences will, however, be resolved in the next section where we consider nonlinear discrete theories.

As regards stability, the continuum theory is always stable for small disturbances, at least in the sense that if some car should accelerate slightly and then decelerate back to its

original velocity (or vice versa), the waves generated by this disturbance will annihilate each other. The argument is that the deceleration waves will form a shock and the accelerations a fan of expanding waves analogous to those shown in Figure IV 16 for a large disturbance. The shock velocity, however, will always have a value within the range of velocities covered

194

by the acceleration waves. Regardless of whether the acceleration occurs before or after the deceleration, some of the acceleration waves will be either faster or slower than the shock. They will eventually overtake the shock and start systematically to annihilate it.

5. <u>Nonlinear discrete theories</u>. Although the linear discrete theory and the nonlinear continuum theory both describe the same first order effects of wave propagation (also in a way which is in fairly good agreement with experimental observations) they give quite different conclusions about the second order effects (neither of which is obviously correct). The first order effects are almost a direct consequence of the conservation of cars which is necessary in any theory and is implied by each of the two theories, but the second order effects are the consequences of the more detailed structure of the theory which has not yet been firmly established. Before investigating various other possible theories, however, it is worthwhile at least to show that the two theories described so far are compatible with each other and represent only limiting cases of a more general nonlinear discrete theory such as described by (1.8).

In the linear theory we neglect the effects of variation in wave velocities in comparison with the effects of the discrete nature of the traffic, an approximation which can, at best, be valid only for sufficiently short times and for disturbances of small amplitude. It is comparison with the effects of the discrete nature of the traffic, an approximation which can, at best, be valid only for sufficiently short times and for disturbances of small amplitude. It is not enough, however, that disturbances just be of small amplitude because two waves with even the slightest difference in velocity will eventually either diverge or converge at a rate (proportional to t) which will eventually swamp the diffusion type spreading of the disturbance predicted by the linear theory (which grows only as $t^{\frac{1}{2}}$). The continuum theory, on

195

the other hand, neglects the discrete nature of traffic, an approximation which is at best valid only if the variations in k,q, etc., are small over distances comparable with the spacing between cars, an approximation which also fails when a shock forms.

To investigate the solution of (1.8) or the obvious generalization of it which includes a reaction time T; namely

$$\frac{dx_j(t + T)}{dt} = V[x_{j-1}(t) - x_j(t)] , \tag{5.1}$$

we can either try to extend the approximation scheme described in section 2 to sufficiently high order as to include both the nonlinear effects and the dispersion effects or we can look for other special (nonlinear) forms of the function V for which exact solutions can be found.

In the first procedure we simply mimic that used in Section IV.2. Equation (2.7) is, in essence, equivalent to the continuum approximation provided we interpret α as the velocity dependent slope of the velocity-spacing curve. The solution (2.11) is, in turn, equivalent to the solution (4.5) of the continuum equation in the absence of shocks. The shocks, however, must also appear in the solution of the nonlinear form of (2.7). In the next approximation, one has the analogue of (2.12), except that the function $z_j(t')$ is the position of the jth car viewed from a moving coordinate system and the coordinate system $t' = t - j/\alpha$ moves with a time lag $1/\alpha$ per car which itself depends upon the velocity of the jth car. The "solution" (2.15) is now from a moving coordinate system and the coordinate system $t' = t - j/\alpha$ moves with a time lag $1/\alpha$ per car which itself depends upon the velocity of the jth car. The "solution" (2.15) is now only an integral equation for $x_j(t)$ because the α in (2.15) itself depends upon the solution. One can see, however, from this equation the qualitative consequences of the competition between the nonlinear and the diffusion effects. The phenomena that is taking place here also has analogues in fluid dynamics where in a first approximation to the fluid equations, without

196

viscosity and heat conduction, one has a first order partial differential equation which gives shocks, but in the next approximation a second derivative term appears which prevents the discontinuity from developing and also describes the "shock structure".

For an acceleration wave of small amplitude and short duration the wave velocity is nearly the same for all car velocities in the disturbance. The first effect is that the diffusion of the linear theory spreads the disturbance over a longer time and smoothes out any irregularities. By the time the disturbance has spread over a distance large compared with the spacing between cars, the continuum approximation can be used to describe the long time behavior. This predicts that eventually the disturbance will spread proportional to t times the difference in the wave velocities over the disturbance. Since the two dispersive effects compliment each other, the total spread of the disturbance at any time is at least as large as either of them alone.

For a deceleration wave the two effects compete with each other; the nonlinearity tries to produce a shock and the diffusion tries to smooth it out. If one starts initially with a very gradual deceleration, the continuum theory applies during the early stages until two waves try to intersect. Before this can happen, however, the diffusion term in the differential equation becomes large and prevents the discontinuity from developing. The disturbance eventually achieves an equilibrium shape in which the two effects balance. The "shock" develops into a becomes large and prevents the discontinuity from developing. The disturbance eventually achieves an equilibrium shape in which the two effects balance. The "shock" develops into a region of rapid but not discontinuous transition which travels with the shock velocity of the continuum theory but acquires a width and a shape that eventually propagates with no change. Furthermore, the equilibrium shape of the shock depends only upon the car velocities on either side of the shock and is independent of the initial rate of deceleration. If, on the other

197

hand, the disturbance was initially very rapid, so rapid that the diffusion effect is too large to balance the nonlinearity, the diffusion will cause the disturbance to spread until the two effects are again in balance as above.

To prove rigorously that these conclusions are correct might be difficult, although it is obvious that, during the initial stages of the wave propagation the solution is progressing in the direction described above. The only point that may be questionable is whether or not a deceleration wave does eventually acquire an equilibrium shock structure. For this to be true one must show that there exists an equilibrium shock structure to which the deceleration wave can converge.

For an equilibrium shock structure to exist there must be solutions of (5.1) of the form

$$x_j(t) = x_{j-1}(t - T^+) + D^+ \tag{5.2}$$

for some appropriate values of T^+ and D^+ which are independent of j and t; i.e., the trajectory of the jth car must be exactly the same as that of the $(j-1)^{th}$ car except that the motion of the $(j-1)^{th}$ car is mimicked by the jth car at a time T^+ later and at a position D^+ ahead (if $D^+ > 0$) as shown in Figure IV 17. From Figure IV 17, the shock velocity must be interpreted to be
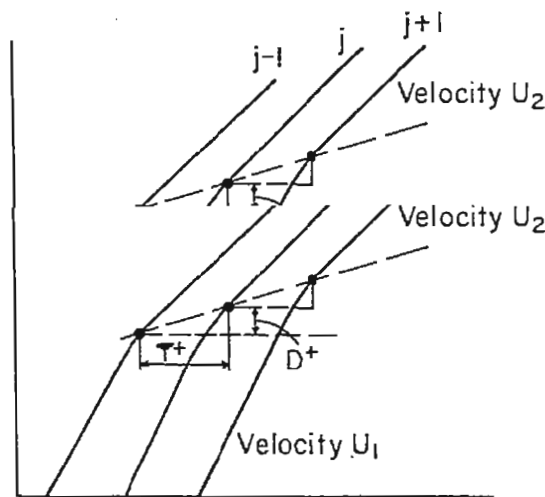
Fig. IV 17
An equilibrium shock

198

D^+/T^+.

Substitution of (5.2) into (5.1) gives the single differential difference equation

$$\frac{dx_j(t + T)}{dt} = V[x_j(t + T^+) - D^+ - x_j(t)] .$$ (5.3)

We further specify the boundary conditions that, for $t \to -\infty$, the cars approach the shock at a given constant speed U_1 and for $t \to +\infty$ they decelerate to a speed U_2 with $U_2 < U_1$. Thus, for $t \to -\infty$

$$x_j(t+T^+) - x_j(t) \to T^+U_1$$

and from (5.3) we have the condition

$$U_1 = V[T^+U_1 - D^+].$$ (5.4)

Similarly for $t \to +\infty$

$$x_j(t+T^+) - x_j(t) \to T^+U_2.$$

and so

$$U_2 = V[T^+U_2 - D^+].$$ (5.5)

If U_1 and U_2 are specified, (5.4) and (5.5) represent two simultaneous equations for the quantities T^+ and D^+ of (5.2).

The solution of these two equations can be found graphically from the velocity-spacing curve. Thus, in Figure IV 18, if we observe the two velocities U_1 and U_2 and the The solution of these two equations can be found graphically from the velocity-spacing curve. Thus, in Figure IV 18, if we observe the two velocities U_1 and U_2 and the corresponding spacings D_1 and D_2, then (5.4) and (5.5) imply that

$$T^+U_1 - D^+ = D_1 \text{ and } T^+U_2 - D^+ = D_2.$$ (5.6)

These two linear equations for T^+ and D^+ show that

$$T^+ = (D_1 - D_2)/(U_1 - U_2)$$
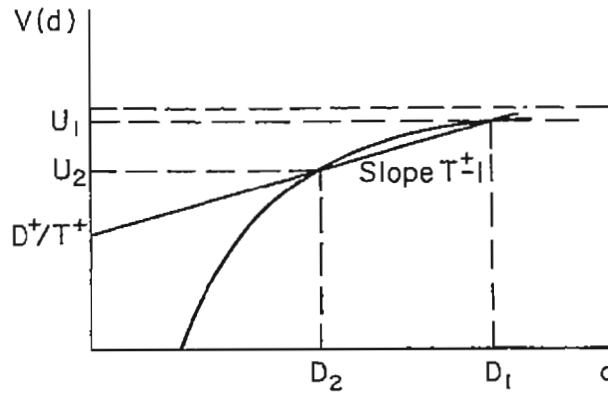
Fig. IV 18
Evaluation of the shock parameters

is the reciprocal slope of the line joining the points (U_1, D_1), and (U_2, D_2) on the V vs. d curve. The velocity intercept of this line is D^+/T^+. Thus the shock velocity of this discrete non-linear theory agrees with that derived earlier, as illustrated in Figure IV 13, from the continuum approximation, as indeed it should, since the shock velocity is determined from a conservation principle.

The existence and nature of the solutions of (5.3) have not as yet been established in general, but it is possible to find approximate solutions for "weak" shocks [14] in which $x_j(t)$ is expected to be slowly varying. If $V^{-1}(.)$ denotes the inverse of the function V, then (5.3) implies

$$V^{-1}(v_j(t)) = x_j(t+T^+-T) - x_j(t-T) - D^+.$$

implies

$$V^{-1}(v_j(t)) = x_j(t+T^+-T) - x_j(t-T) - D^+.$$

We expand the right hand side in a Taylor series in the time displacements to obtain

$$V^{-1}(v_j(t)) = T^+ v_j(t) - D + \frac{T^+(T^+ - 2T)}{2} \frac{dv_j(t)}{dt}$$

$$+ T^+ \frac{(T^{+2} - 3T^+ + 3T^2)}{3} \frac{d^2v_j(t)}{dt^2} \tag{5.7}$$

and in the lowest approximation neglect the second derivative term so that

$$-\frac{1}{2} T^+(T^+ - 2T) \frac{dv_j(t)}{dt} \sim T^+ v_j(t) - D^+ - V^{-1}(v_j(t)) . \tag{5.8}$$

The right-hand side of (5.8) has a simple geometrical interpretation as the horizontal distance in Figure IV 18 between the velocity-spacing curve and the displayed shock line, at the velocity $v_j(t)$. If the velocity-spacing curve is concave this is positive for $U_2 < v_j(t) < U_1$ and vanishes for $v_j(t) = U_1$ or U_2. Equation (5.8), therefore, potentially has a solution if $U_2 < U_1$ and $T^+ > 2T$, since it will give a negative value for $dv_j(t)/dt$; i.e., a deceleration. Since T^+ is the reciprocal slope of the shock line which, for weak shocks, is also essentially the reciprocal slope of the tangent to the velocity-spacing curve, it corresponds to α^{-1} in the linear theory. The condition $T^+ > 2T$, in turn, is equivalent to the stability condition $\alpha T < 1/2$ of the linear theory.

Integration of (5.8) gives

$$\frac{2t}{T^+(T^+ - 2T)} = \int^{v_j(t)} \frac{dv}{T^+ v - D^+ - V^{-1}(v)} . \tag{5.9}$$

$$\frac{2t}{T^+(T^+ - 2T)} = \int^{v_j(t)} \frac{dv}{T^+ v - D^+ - V^{-1}(v)} . \tag{5.9}$$

The origin of time has not been specified and this is reflected in (5.9) by the fact that it contains an arbitrary integration constant or an unspecified lower limit of integration. Aside from this, however, (5.9) determines t as a function of $v_j(t)$ or conversely $v_j(t)$ as a function of t and so gives the history of the car's velocity as it travels through the shock.

To see in more detail what this relation implies, we might further assume that, for weak shocks, we can approximate the function $V^{-1}(v)$ in the range $U_2 \leq v \leq U_1$ by a quadratic function. We write the right-hand side of (5.8) in the form

$$T^+v\text{-}D^+\text{-}V^{-1}(v) = D'' (v\text{-}U_1)(U_2\text{-}v)/2, \tag{5.10}$$

since the left-hand side certainly vanishes at $v = U_1$ or $v = U_2$. The constant D'' is the second derivative of $V^{-1}(v)$ evaluated at some value of v between U_1 and U_2.

The integral in (5.9) can now be evaluated explicitly and gives (except for an unspecified origin of t)

$$2v_j(t) = U_1 + U_2 - (U_1 - U_2)tanh(\gamma t) \tag{5.11}$$

with

$$\gamma = \frac{D''(U_1 - U_2)}{T^+(T^+ - 2T)} . \tag{5.12}$$

This solution for $v_j(t)$ shows that for $t \rightarrow -\infty$, $v_j(t) \rightarrow U_1$ and for $t \rightarrow +\infty$, $v_j(t) \rightarrow U_2$ as it should. The important new feature is that it gives the "shock thickness." The time required for a car to traverse the shock is measured by the time constant γ^{-1}.

These formulas were derived under the assumption that the shocks were weak enough that we could (1) approximate the curve $V^{-1}(v)$ by a parabola over the range of velocities in question and (2) expand the velocities as in (5.7) for times of the order T^+. The latter that we could (1) approximate the curve $V^{-1}(v)$ by a parabola over the range of velocities in question and (2) expand the velocities as in (5.7) for times of the order T^+. The latter condition is valid provided $\gamma^{-1} \gg T^+$, i.e., the time to traverse the shock is large compared with the time lag for propagation of a wave from one car to the next. This also implies that the spatial width of the shock is so large as to cover a large number of cars simultaneously.

From (5.12) we see that the time constant is inversely proportional to the curvature D'' so that $\gamma^{-1} \to \infty$ in the limit of the linear theory. In the linear theory, any disturbance would disperse, consequently there can be no finite shock widths. The time constant is also inversely proportional to U_1-U_2; thus weak shocks have a large thickness. It is also proportional to T^+ - $2T$ which suggests that as one approaches the limit of stability $2T \to T^+$, the shock widths become small and, therefore, give large decelerations. The approximations used here, of course, break down in the limit of small shock thickness but this at least suggests a trend and a potential danger that this theory might give rise to decelerations within the shock front which exceed the braking rate of cars if $2T$ is too close to T^+.

6. <u>Some exact solutions</u>. The approximation methods discussed in the last section can be used to describe the formation and propagation of weak shocks or the propagation of acceleration waves, for a fairly general type of velocity-spacing relation. It is not, however, possible to obtain simple exact solutions of (5.1) except for very special functions $V(d)$. So far the only exact time-dependent solutions that have been found are those which result when $T = 0$ and $V_j(d)$ has the special form

$$V_j(d) = V_j - V_j \exp [-(\alpha_j d - \beta_j)/V_j] \tag{6.1}$$

in which V_j, α_j, and β_j are constants that may perhaps depend upon the car number j .

$$V_j(d) = V_j - V_j \exp [-(\alpha_j d - \beta_j)/V_j] \tag{6.1}$$

in which V_j, α_j, and β_j are constants that may perhaps depend upon the car number j .

Theories of this type were proposed independently, but presented in somewhat different ways, by Newell [14] and Franklin [17-19].

Relation (6.1) is shown in Figure IV 19. It gives zero velocity at spacing $d = \beta_j/\alpha_j$.
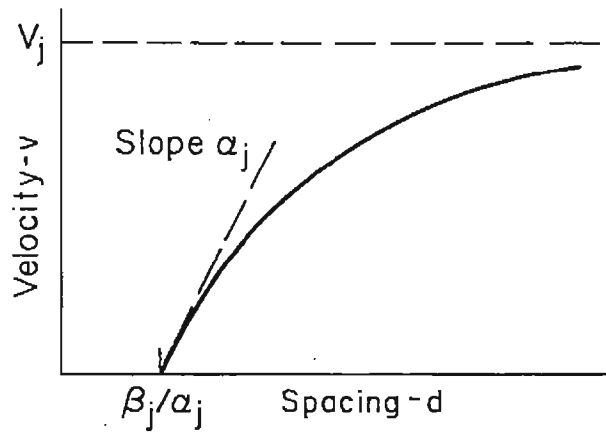
Fig. IV 19
An exponential velocity-spacing relation

For small velocities we can expand the exponential and obtain

$$V_j(d) \sim \alpha_j d - \beta_j \qquad (6.2)$$

which is the same form as the linear equation (2.1). The notation has been chosen so that the α_j and β_j in this non-linear model correspond to the α and β of (2.1), if we consider linear wave propagation at small velocities. If we approximate (6.1) by a tangent line at some velocity other than zero, we will, of course, have different slopes and intercepts. For $d \to \infty$, $V_j(d) \to V_j$, so V_j can be interpreted as the free speed of the jth car.

Fortunately the shape of (6.1) is approximately that observed for the steady state relation between velocity and spacing. In fact, by suitable choice of parameters, one can fit

relation between velocity and spacing. In fact, by suitable choice of parameters, one can fit

such a formula to any presently available data with as much accuracy as the data justify.

It is not possible to obtain simple solutions of (5.1) and (6.1) with a non-zero time lag T. The equations are solvable by the iterative method used in (2.4), but the form of the solution is not very illuminating. We will, however, solve these equations with T = 0, if only

204

to establish that the approximations of the last section are valid and that this theory gives rise to strong shocks as well as the weak shocks predicted in the last section.

We can first eliminate the parameters β_j from (6.1) by letting

$$z_j(t) = x_j(t) + \sum_{k=1}^{j} \beta_k/\alpha_k \tag{6.3}$$

This is the same trick as used in (2.5) for the linear theory. In effect, we are cutting out of the highway the minimum spacing β_j/α_j ahead of each jth car. Equation (5.1) now takes the form

$$v_j(t) = \frac{dz_j(t)}{dt} = V_j - V_j \exp\{-\alpha_j[z_{j-1}(t) - z_j(t)]/V_j\} \tag{6.4}$$

If we now make a non-linear substitution

$$Z_j(t) = \exp[-\alpha_j z_j(t)/V_j] \tag{6.5}$$

so that

$$v_j(t) = \frac{dz_j(t)}{dt} = \frac{-V_j}{\alpha_j} \frac{d \log Z_j(t)}{dt} \tag{6.6}$$

we obtain from (6.4) the linear differential equation for $Z_j(t)$,

$$\frac{1}{\alpha_j} \frac{d Z_j(t)}{dt} + Z_j(t) = [Z_{j-1}(t)]^{\mu(j,j-1)} \tag{6.7}$$

$$\frac{1}{\alpha_j} \frac{d Z_j(t)}{dt} + Z_j(t) = [Z_{j-1}(t)]^{\mu(j,j-1)} \tag{6.7}$$

$$\mu(j, j-1) = \alpha_j V_{j-1}/(\alpha_{j-1} V_j) . \tag{6.8}$$

That certain non-linear equations can be transformed into linear equations by non-linear substitutions has been exploited in various ways for at least two centuries. Franklin [19] arrives at this in a slightly different way. He started from the postulate that the acceleration of a driver is given (in terms of the notation here) by

$$\frac{dv_j(t)}{dt} = \alpha_j \left[1 - \frac{v_j(t)}{V_j} \right] \left[v_{j-1}(t) - v_j(t) \right].$$

This is, in essence, the derivative of (6.4). As an equation for $v_j(t)$ with $v_{j-1}(t)$ given, this is a form of the Riccati equation (the right-hand side of the equation is a quadratic in $v_j(t)$). The transformation (6.5) is the well-known change of dependent variables which converts the Riccati equation into a linear equation [20].

We can easily integrate equation (6.7) and obtain $Z_j(t)$ in terms of $Z_{j-1}(t)$. But for $\mu(j,j-1) \neq 1$, $Z_j(t)$ is a non-linear functional of $Z_{j-1}(t)$ which is difficult to iterate with respect to j so as to find $Z_j(t)$ in terms of $Z_o(t)$. Even if $\mu(j,j-1) = 1$ but α_j depends upon j, the iteration of the solution is a bit awkward. This is still worth pursuing, however, because there are some interesting effects associated with j-dependent free speeds V_j. If the free speed of a car is less than the actual speed of the car ahead, then this car will fall further and further behind. We can still treat a rather artificial version of this problem if we allow both α_j and V_j to depend upon j in such a way that $\mu(j,j-1) = 1$ or equivalently

$$\frac{\alpha_j}{V_j} = \frac{\alpha_{j-1}}{V_{j-1}} = \epsilon \qquad (6.9)$$

independent of j. This is reasonable at least to the extent that the more aggressive drivers might be expected to have both a large α_j (a steeper V vs. d curve) and a larger V_j. The independent of j. This is reasonable at least to the extent that the more aggressive drivers might be expected to have both a large α_j (a steeper V vs. d curve) and a larger V_j. The minimum spacing, which depends upon β_j, is still arbitrary but has been driven out of the equations.

Hereafter we will consider only the special cases for which (6.9) is true. Equations (6.5) and (6.7) now simplify to

$$Z_j(t) = exp \left[- \varepsilon \ z_j(t) \right] , \qquad (6.5a)$$

$$v_j(t) = - \frac{d \ log \ Z_j(t)}{\varepsilon \ dt} , \qquad (6.6a)$$

and

$$\frac{1}{\alpha_j} \frac{d \ Z_j(t)}{dt} + Z_j(t) = Z_{j-1}(t) . \qquad (6.7a)$$

Equation (6.7a) is similar to the equations that appear in the linear theories such as (2.6) except that the $z_j(t)$ now has a different interpretation. In fact for $\alpha_j = \alpha$, independent of j, the equations are exactly the same as (2.6) with $T = 0$.

To study an initial value problem for given $x_j(0)$, we can again use Laplace transform methods. If cars are moving forward, then $x_j(t)$ and $z_j(t)$ are increasing with t, and $Z_j(t)$ is decreasing. The Laplace transform of $Z_j(t)$, therefore, exists for all $Re \ s > 0$. If we take the Laplace transform (6.7a) we have

$$Z_j^*(s) = \frac{\alpha_j \ Z_{j-1}^*(s) + Z_j(0)}{s + \alpha_j} \qquad (6.10)$$

a linear finite difference equation for the $Z_j(s)$ with solution

$$Z_j^*(s) = \frac{1}{s + \alpha.} \left\{ Z_j(0) + \frac{\alpha_j \ Z_{j-1}(0)}{s + \alpha.,} + \frac{\alpha_j \ \alpha_{j-1} \ Z_{j-2}(0)}{(s + \alpha.,)(s + \alpha._,)} \right.$$

$$Z_j^*(s) = \frac{1}{s + \alpha_j} \left\{ Z_j(0) + \frac{\alpha_j \ Z_{j-1}(0)}{s + \alpha_{j-1}} + \frac{\alpha_j \ \alpha_{j-1} \ Z_{j-2}(0)}{(s + \alpha_{j-1})(s + \alpha_{j-2})} \right.$$

$$(6.11)$$

$$+ \ ... \ + \frac{\alpha_j \ \alpha_{j-1} \ ... \ \alpha_2 \ Z_1(0)}{(s + \alpha_{j-1}) \ ... \ (s + \alpha_1)} + \left. \frac{\alpha_j \ \alpha_{j-1} \ ... \ \alpha_1 \ Z_o^*(s)}{(s + \alpha_{j-1}) \ ... \ (s + \alpha_1)} \right\} .$$

If car 0 is chosen to be at x = 0 at time 0 and we assume that a steady state had existed prior to time 0, i.e., all cars were traveling at some velocity u, independent of j, then, according to (6.4), the initial values of $Z_j(t)$ must satisfy the relations

$$\frac{Z_{j-1}(0)}{Z_j(0)} = 1 - \frac{u}{V_j} \quad , \quad Z_o(0) = 1$$

or

$$Z_j(0) = \Pi_{k=1}^{j} (1 - u/V_k)^{-1} \tag{6.12}$$

provided that u is less than all the free speeds V_k.

We can obtain a more convenient form of the solution (6.11) with these initial conditions if we first observe that the steady state solution of (6.7a) is

$$Z_{j,o}(t) = Z_j(0)e^{-\varepsilon u t} \tag{6.13}$$

having transform

$$Z_{j,o}*(s) = Z_j(0)/(s + \varepsilon u) \,.$$

Since (6.7a) is a linear difference equation, sums or differences of solutions are also solutions. In particular $Z_j(t) - Z_{j,o}(t)$ is a solution satisfying the simpler initial conditions that it vanish at t = 0 for all j. We have now the same type of transform equation as (6.10) but with no inhomogeneous term

no inhomogeneous term

$$[Z_j^*(s) - Z_{j,o}^*(s)] = \frac{\alpha_j[Z_{j-1}(s) - Z_{j-1,o}(s)]}{s + \alpha_j} \,. \tag{6.10a}$$

Thus

$$Z_j^*(s) = \frac{1}{(s + \varepsilon u)\Pi_{k=1}^j (1 - u/V_k)} + \frac{Z_o^*(s) - (s + \varepsilon u)^{-1}}{\Pi_{k=1}^j(1 + s/\alpha_k)} . \qquad (6.11a)$$

The inversion of this transform is straightforward; it involves only rational functions of s or rational functions times $Z_0^*(s)$, but to see some of the consequences of the theory it is convenient to study some special cases first.

If all cars are identical $\alpha_j = \alpha$, $V_j = V$ for all j, then

$$Z_j^*(s) = \frac{1}{(s + \varepsilon u)} \left\{ \frac{1}{(1 - u/V)^j} - \frac{1}{(1 + s/\alpha)^j} \right\} + \frac{Z_o^*(s)}{(1 + s/\alpha)^j} \qquad (6.14)$$

or, from the convolution theorem, we have

$$Z_j(t) = \frac{e^{-\varepsilon ut}}{(1 - u/V)^j} \left\{ 1 - \int_0^{(1-u/v)\alpha t} d\tau \frac{e^{-\tau} \tau^{j-1}}{(j - 1)!} \right\} + \int_0^t d\tau \frac{e^{-\alpha\tau} (\alpha\tau)^{j-1}}{(j - 1)!} Z_o (t - \tau) . \qquad (6.15)$$

The first term above can be written as an incomplete gamma function. We define

$$\Gamma(j , t) = \int_t^\infty d\tau e^{-\tau} \tau^{j-1} , \gamma (j , t) = \int_0^t d\tau e^{-\tau} \tau^{j-1} \qquad (6.16)$$

$$\Gamma(j) = (j -1)! = \Gamma(j , t) + \gamma(j , t) = \Gamma(j , o) = \gamma(j , \infty).$$

To specialize still further, suppose that at t = 0, the lead car suddenly acquires a velocity v ≠ u and travels at velocity v for all t > 0, i.e.,

$$Z (t) = exp (-\varepsilon vt) \ for \ t > 0$$

velocity v ≠ u and travels at velocity v for all t > 0, i.e.,

$$Z_o(t) = exp (-\varepsilon vt) \ for \ t > 0.$$

The second term of (6.15) is now also an incomplete gamma function and

$$Z_j(t) = \frac{e^{-\varepsilon ut} \Gamma(j , (1 - u/V)\alpha t)}{\Gamma(j) (1 - u/V)^j} + \frac{e^{-\varepsilon vt}\gamma(j , (1 - v/V)\alpha t)}{\Gamma(j) (1 - v/V)^j} . \qquad (6.17)$$

From this we can now evaluate $z_j(t)$, $x_j(t)$, or $v_j(t)$ from (6.5a), (6.3), and (6.6a). In particular, the velocity is given by

$$v_j(t) = u\kappa + v(1 - \kappa) \tag{6.18a}$$

with

$$\kappa^{-1} = 1 + \frac{\mu^{-j} e^{\mu} \gamma(j, \mu)}{\upsilon^{-j} e^{\upsilon} \Gamma(j, \upsilon)} \tag{6.18b}$$

$$\mu = (1 - v/V)\alpha t \quad, \upsilon = (1 - \mu/V)\alpha t . \tag{6.18c}$$

This formula describes the velocity of the jth car after an impulse acceleration or deceleration of the lead car at time 0. If the approximation methods of the previous sections are correct, it must, in appropriate limiting cases, describe (a) the diffusion effects and wave propagation of the linear theory if u - v,j, and t are sufficiently small, (b) the formation of expanding waves if v > u and t is large enough, and (c) the formation of shocks if v < u and t sufficiently large. These three types of behavior do emerge from (6.18) and will be considered in order.

(a). If in (6.18) we let (v - u) → 0 with j and t fixed, and we evaluate $v_j(t)$ only to terms linear in v - u, then in (6.18a), which is equivalent to

$$v_j(t) = u + (v-u)(1-\kappa), \tag{6.19}$$

it suffices to evaluate κ only for v = u. But for v = u we have

$$v_j(t) = u + (v-u)(1-\kappa), \tag{6.19}$$

it suffices to evaluate κ only for v = u. But for v = u we have

$$1 - \kappa = 1 - \frac{1}{1 + \gamma(j, \mu)/\Gamma(j, \mu)} = \frac{\gamma(j, \mu)}{\Gamma(j)} . \tag{6.20}$$

This gives the exact solution of the linearized equations of motion that would result if we approximated V(d) by its tangent line at the velocity v. Solutions of this type were obtained originally by Reuschel [4] and Pipes [5].

210

The acceleration of the jth car is

$$\frac{dv_j(t)}{dt} = -(v - \mu)\frac{d\kappa}{dt} = \frac{\alpha(v - \mu)(1 - v/V)}{\Gamma(j)}e^{-\mu}\mu^{j-1}.$$

The time dependence of the acceleration is proportional to $\mu^{j-1}e^{-\mu}$ and this has a maximum at

$\mu = (1-v/V)\alpha t = j-1$. Curves of acceleration are shown in Figure IV 20. Thus the point of
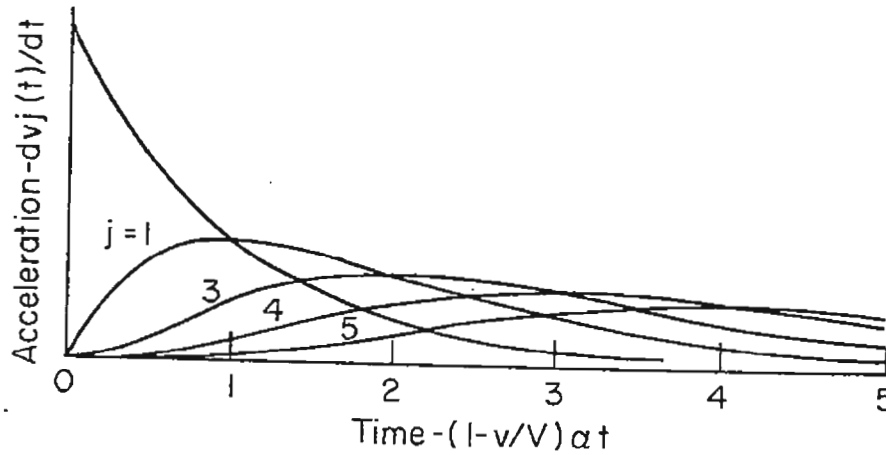
Fig. IV 20
Acceleration of *jth* car following an impulse acceleration of car O

maximum acceleration propagates with a time lag of $(1-v/V)^{-1}\alpha^{-1}$ per car, consequently with a

constant velocity, even for small j values. The shape of the acceleration wave changes rapidly

for the first few j values but for large j approaches a Gaussian shape, specifically if for large j

for the first few j values but for large j approaches a Gaussian shape, specifically if for large j

$$\mu - (j-1) = O(j^{1/2})$$

then

$$\frac{dv_j(t)}{dt} \sim \frac{\alpha(v - \mu)(1 - v/V)}{[2\pi(j - 1)]^{1/2}}\exp\left[\frac{-(\mu - j + 1)^2}{2(j - 1)}\right]. \qquad (6.21)$$

This follows from the well-known normal approximations to the Poisson distribution.

211

Except for a slight difference in notation (the $\alpha(1-v/V)$ in the present theory corresponds to the α of the linear theory with $T = 0$), this result simply reconfirms those of the heuristic analysis for the linear theory following (2.12); the pulse travels with a constant velocity and spreads proportional to $j^{1/2}$.

To obtain these results, we first let $v - u \to 0$ and then let j and/or t become large. For any non-zero $(v - u)$, the approximation of (6.18b,c) by (6.20) is valid, however, for only a finite range of t and j. By evaluating κ for $v = u$, we have made the approximations

$$(\mu/\upsilon)^j e^{(\mu-\upsilon)} \sim 1 \quad and \quad \Gamma(j, \upsilon) \sim \Gamma(j, \mu) .$$

For $(v/u)/V << 1$ and $j >> 1$, however, we have

$$\left(\frac{u}{\upsilon}\right)^{-j} e^{u-\upsilon} = \exp\left\{ j \, \log \frac{(1 - u/V)}{(1 - v/V)} - (v - u) \frac{\alpha t}{V} \right\}$$

$$\sim \exp\left\{ j \left[\frac{(v - u)}{V(1 - u/V)} - \frac{(v - u)^2}{2V^2(1 - u/V)^2} + ... \right] - \frac{(v - u) \, \alpha t}{V} \right\} . \tag{6.22}$$

The linear theory is, therefore, valid only if $v-u$ is so small that

$$\frac{(v - u)}{V} \left| \frac{j}{1 - u/v} - \alpha t \right| << 1 \tag{6.23a}$$

and

$$\frac{v}{} \quad \left| 1 - u/v \right| \quad \left|$$

and

$$j \frac{(v - u)^2}{(V - u)^2} << 1 . \tag{6.23b}$$

The approximations for the Γ-functions are valid under similar conditions.

For $(v-u)/V <<1$, and $j >> 1$, the disturbance for the jth car is, according to (6.22), confined mainly to the time range where

$$\left| \frac{j}{1 - u/V} - \alpha t \right| = \frac{O(j^{\frac{1}{2}})}{1 - u/V} . \tag{6.24}$$

In this range of time, however, (6.23b) implies (6.23a), so the important restriction on j is (6.23b).

The restriction (6.23b), of course, must arise physically from the fact that a small disturbance at car velocity u will reach the jth car at a time $j[\alpha(1-u/V)]^{-1}$, approximately, whereas a disturbance at car velocity v reaches the jth car at a time $j[\alpha(1-v/V)]^{-1}$. The sudden change in velocity of the lead car from velocity u to velocity v can, therefore, be considered as a single small disturbance propagating at a fixed velocity only so long as the duration of the disturbance as represented by (6.24) is large compared with the time interval between the arrivals of the waves of velocities u and v at the jth car, i.e.,

$$\frac{j^{\frac{1}{2}}}{\alpha(1 - u/V)} >> \frac{j(v - u)}{\alpha(1 - u/V)^2 \, V}$$

which is indeed equivalent to (6.23b).

(b). To investigate the formation of expanding waves and/or shocks we wish to evaluate κ, (6.18b), for large j, but with more or less arbitrary values for u and v. For this we need the following asymptotic formulas for the Γ-functions [21].

$$\gamma(j , \mu) \sim \Gamma(j) \left\{ \frac{1}{2} + \pi^{-\frac{1}{2}} \, Erf \, [(\mu - j + 1)/(2j - 2)^{\frac{1}{2}}] \right. \tag{6.25a}$$

$$\gamma(j , \mu) \sim \Gamma(j) \left\{ \frac{1}{2} + \pi^{-\frac{1}{2}} \, Erf \, [(\mu - j + 1)/(2j - 2)^{\frac{1}{2}}] \right. \tag{6.25a}$$

$$if \; (\mu - j + 1) = O[(j - 1)^{\frac{1}{2}}]$$

$$\gamma(j , \mu) \sim \frac{e^{-\mu}\mu^j}{(j - 1 - \mu)} \left\{ 1 - \frac{(j - 1)}{(j - 1 - \mu)^2} + ... \right\} \tag{6.25b}$$

$$if \; \mu - j + 1 < O[(j - 1)^{\frac{1}{2}}]$$

$$\gamma(j, \mu) \sim \Gamma(j) \quad if \quad \mu - j + 1 > 0[(j - 1)^{\frac{1}{2}}] \tag{6.25c}$$

$$\Gamma(j, \upsilon) \sim \Gamma(j) \left\{ \frac{1}{2} - \pi^{-\frac{1}{2}} \, Erf \left[(\upsilon - j + 1)/(2j - 2)^{\frac{1}{2}} \right] \right\} \tag{6.26a}$$

$$if \quad (\upsilon - j + 1) = 0 \left[(j - 1)^{\frac{1}{2}} \right]$$

$$\Gamma(j, \upsilon) \sim \Gamma(j) \quad if \; (\upsilon - j + 1) < 0[(j - 1)^{\frac{1}{2}}] \tag{6.26b}$$

$$\Gamma(j, \upsilon) \sim \frac{e^{-\upsilon}\upsilon^{j}}{(\upsilon - j + 1)} \left\{ 1 - \frac{(j - 1)}{(\upsilon - j + 1)^2} + ... \right\} \tag{6.26c}$$

$$if \quad (\upsilon - j + 1) > 0[j - 1)^{\frac{1}{2}}] \; .$$

If conditions (6.23a,b) hold, then we revert to the situation described above. If $j^{1/2}(v-u)/(V-u) = 0(1)$, then the factor (6.22) is no longer 1. Also in the range of t where (6.25a) applies, (6.26a) also applies, but the arguments of both Erf functions are of order 1 and significantly different from each other. This is a range of j values where the non-linear effects start to compete with the dispersion effects. Nothing cancels in the formulas for κ.

For

$$\frac{j^{\frac{1}{2}} \mid v - u \mid}{(V - u)} >> 1 \tag{6.27}$$

the behavior of κ depends critically on the sign of (v - u), as is expected. For an acceleration, v > u and u < v. The ith car will accelerate from velocity u to v, so κ must obviously decrease

the behavior of κ depends critically on the sign of (v - u), as is expected. For an acceleration, v > u and μ < v. The jth car will accelerate from velocity u to v, so κ must obviously decrease with t from 1 at t = 0 to 0 at t →∞. Our first task is to locate the range of t where most of this happens. The earliest range of t indicated in (6.25a) to (6.25c) is that where (6.25b) and (6.26b) both apply, but one can easily check that κ is still close to 1 here. Before we get out of the range where (6.25b) applies, we will pass completely through the range where (6.25a)

applies, but still κ is close to 1. The range where things do occur is where (6.25b) and (6.26c) hold. Here we have

$$\kappa^{-1} = 1 + \frac{(\upsilon - j + 1)}{(j - 1 - \mu)} \left\{ 1 + \frac{(j - 1)}{(\upsilon - j + 1)^2} - \frac{(j - 1)}{(j - 1 - \mu)^2} + ... \right\}$$

with

$$\mu - j + 1 < 0(j^{\frac{1}{2}}) < \upsilon - j + 1 .$$

Substitution of this into (6.18a) gives

$$v_j(t) \sim V \left[1 - \frac{(j - 1)}{\alpha t} + ... \right]$$

for

$$(1 - v/V)[1 + 0(j^{-\frac{1}{2}})] < \frac{j - 1}{\alpha t} < (1 - u/V) [1 + 0 (j^{-\frac{1}{2}})] .$$

In the limit $j \to \infty$, $t \to \infty$ with $j/(\alpha t)$ finite we find

$$v_j(t) \to \begin{cases} u & 1 - u/V < j/(\alpha t) \\ V[1 - j/(\alpha t)] & 1 - v/V < j/(\alpha t) < 1 - u/V \\ v & 1 - v/V > j/(\alpha t) \end{cases} \tag{6.28}$$

This represents the solution we would obtain from the continuum theory. All waves start from $j = 0$ at $t = 0$. A wave of velocity v', $u < v' < v$, reaches the jth car at approximately a time $j/(1 - v'/V)$ later. If we substitute this time into (6.28) we do indeed obtain $v_j(t) = v'$.

~~Figure IV 21 shows the velocity of a sequence of cars when the lead car suddenly~~ j/(1-v'/V) later. If we substitute this time into (6.28) we do indeed obtain $v_j(t) = v'$.

Figure IV 21 shows the velocity of a sequence of cars when the lead car suddenly accelerates from $u = 0$ to $v = 1/2V$ at $t = 0$. The velocity is plotted vs. $\alpha t/j$ to show the convergence of this velocity profile to the wave solution (4.28) for $j \to \infty$. The convergence is rather slow, particularly near the wave edges $\alpha t/j = 1$ and 2, because the diffusion effects depend upon $j^{1/2}$.
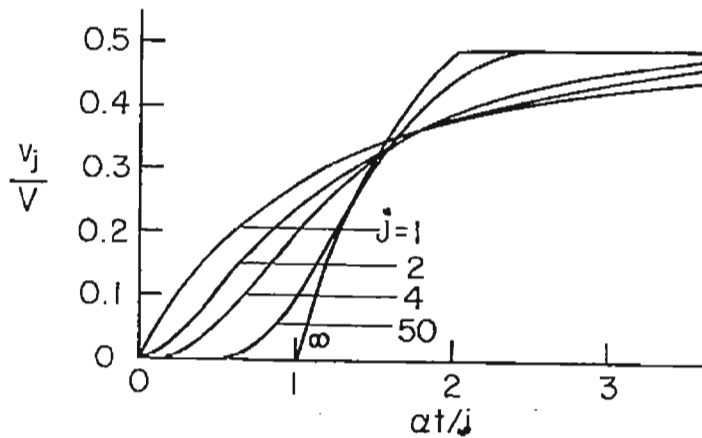
Fig. IV 21
Convergence of a sudden acceleration to the wave solution

(c). If $v < u$ and $\mu > v$, we again look for the range of t where most of the variation in velocity occurs. This time we find, by essentially the same arguments as above, that it occurs where (6.25c) and (6.26b) apply. The Γ-functions cancel in (6.18b) and we have from (6.21) that

$$\kappa^{-1} = 1 + \exp\left\{ j \log \frac{(1 - u/V)}{(1 - v/V)} - (v - u) \frac{\alpha t}{V} \right\}$$

and

$$v_j(t) = \frac{(u + v)}{2} + \frac{(v - u)}{2} \tanh\left\{ \frac{j}{2} \log \frac{(1 - u/V)}{(1 - v/V)} - \frac{(v - u)}{2V} \alpha t \right\} \tag{6.29}$$

$$v_j(t) = \frac{(u + v)}{2} + \frac{(v - u)}{2} \tanh\left\{ \frac{j}{2} \log \frac{(1 - u/V)}{(1 - v/V)} - \frac{(v - u)}{2V} \alpha t \right\}. \tag{6.29}$$

This represents the shock profile. We notice that each jth car has exactly the same motion as any other except for a displacement in time. The displacement in time from one car to the next is

$$\frac{1}{\alpha'} = \frac{V\{\log(1 - u/V) - \log(1 - v/V)\}}{(v - u)\alpha}.$$ (6.30)

If we mark the two points on the velocity-spacing curve corresponding to velocities u and v as in Figure IV 18, we can recognize immediately from (6.1) that this time lag is the difference in spacing divided by the difference in velocity for these two points. This checks with the properties of shock propagation as described in sections IV 4 and IV 5. The shape of $v_j(t)$, (6.29), also agrees with that derived in (5.11) for weak shocks (except that (5.11) does not include the j-dependent time lags.)

Figure IV 22 shows the velocities of a sequence of cars when the lead car suddenly decelerates from 1/2 V to 0. The velocities are plotted vs $\alpha(t-j/\alpha')$ for several values of j. The time coordinate of the jth car is displaced by the propagation time for the shock to show the convergence for $j \to \infty$ to the equilibrium shock profile given by (6.29).
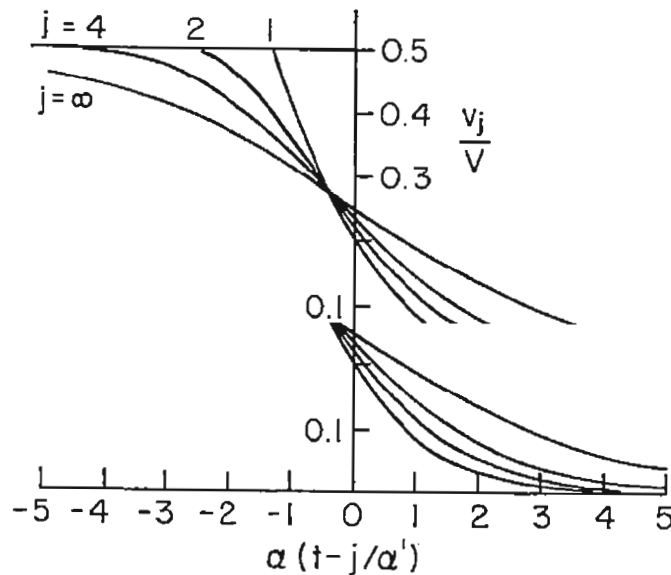
Fig. IV 22
Convergence of a sudden deceleration to the shock profile

217

One can easily show that if the lead car decelerates from velocity u to velocity v in any manner whatsoever, the resulting disturbance will eventually propagate as a shock. The shape of the equilibrium shock will depend upon the initial and final velocities u and v, but will otherwise be independent of the shape of the initial disturbance. If the lead car decelerates very slowly, the decelerations of subsequent cars must increase until they reach the values associated with the equilibrium shock trajectories. If the shock trajectories involve high decelerations, it is possible for a lead driver who is cautious and decelerates very slowly to initiate a disturbance that might eventually become strong enough to cause some accidents.

By differentiating (6.29) we see that the maximum deceleration in the shock is

$$\frac{(v - u)^2}{4V} \, \alpha \leq \frac{V\alpha}{4} \ .$$

The latter is for a deceleration between u = V and v = 0 and typically has a value of about 10 ft/sec^2. This is a moderately hard rate of deceleration, but not outside the range of braking power for most cars.

By comparing (6.29) with (5.12), however, we see that a non-zero reaction time T is likely to cause still larger decelerations. It does so for weak shocks at least. This phenomenon quite likely contributes to the causes of chain collisions.

Before going on to other aspects of car-following theories we will simply note that all

Before going on to other aspects of car-following theories we will simply note that all the theories described so far are consistent with each other. Both the linear car following theories and the continuum theories are limiting cases of a more general non-linear discrete theory. Most large-scale phenomena are most easily described in terms of the continuum theory or a continuum theory modified by inclusion of the diffusion effects as described in section IV 5.

7. Non-identical cars.

a. Linear theories. One of the logical deficiencies of the car-following theories described so far is that they imply the existence of some lead car whose motion must be explicitly specified. One might naturally ask: who is the lead driver and why doesn't he satisfy the same equations as everyone else?

In some situations there is an obvious way of identifying a lead driver. He may be the first car stopped by a red signal and when the signal turns green he is not consciously following anyone. His motion is not described by the equations above because they apply to a time-independent homogeneous highway. The traffic signal is a time inhomogeneity but we might imagine that the effect of the signal can be approximately represented by specifying its effect on car 0 say, and then postulating that subsequent cars move as if they were simply following car 0 but were otherwise not influenced by the signal.

In other instances car 0 may be a slow car that no one can pass. This driver is different from other drivers in that his desired speed is less than anyone else. His motion is not described by the above theory because we postulated that all drivers are essentially the same.

Any complete theory of traffic flow should be such that if a line of cars is obviously following some particular car, the theory should itself predict which car is leading, how it is

Any complete theory of traffic flow should be such that if a line of cars is obviously following some particular car, the theory should itself predict which car is leading, how it is leading, and why. As yet we do not have such a theory, but some significant generalizations in this direction will result if we include the possibilities that not all drivers are identical and that highway characteristics (the relations between flow, density, velocity, etc.) may be explicitly dependent upon position along the highway and the time. Many potential

219

applications of the car-following theory, however, apply to the motion of traveling queues; there is some limited passing but not enough to prevent long queues from forming behind slow cars. In such cases the lead car, the queue lengths, etc. are determined by the passing conditions. We do not have very good models for describing such things.

In this section we consider briefly some consequences of random selection of drivers. We will assume that each driver has his own velocity-spacing relation which he follows exactly with no fluctuations, but that different drivers have different relations and the drivers are selected at random from some population of drivers.

The discussion of linear models in section IV 3 can easily be generalized to non-identical drivers. If we take equation (2.2) with α, β, and T replaced by α_j, β_j, T_j for each jth driver, we can subtract away the minimum spacings β_j/α_j as done in (6.3). The generalization of (3.4) is, therefore,

$$z_j(t) = \frac{1}{2\pi i} \int_{\gamma-i\infty}^{\gamma+i\infty} ds \; e^{st} \frac{z_o^*(s)}{\Pi_{k=1}^j\left[1 + s \; \alpha_k^{-1} \exp(sT_k)\right]} \; .$$

For any given j, this integral can again be expanded in terms of residues. The transient motion of the jth car will, for sufficiently large t, be dominated by the residue from the zero of the denominator having the largest real part. This residue might come from any one of the motion of the jth car will, for sufficiently large t, be dominated by the residue from the zero of the denominator having the largest real part. This residue might come from any one of the factors of the integrand. There is no simple rule to determine which one because the roots of the kth factor depend upon both $\alpha_k T_k$ and T_k, the latter being the scale factor for S in Figure IV 6.

We can say that the dominant transient will be a pure exponential if $\alpha_k T_k < 1/e$ for all $k \leq j$ and it will be unstable if $\alpha_k T_k > \pi/2$ for any $k \leq j$. It will also a damped oscillation if $1/e < \alpha_k T_k < \pi/2$ for all $k \leq j$; but if some $\alpha_k T_k$ lie in this range, and others are less than $1/e$ the transient might be either a pure or oscillatory exponential. In any case the transient for the jth car persists at least as long as that of any car ahead of it.

The asymptotic properties of $z_j(t)$ for $j \to \infty$ can again be found by expansion of the integrand near $s = 0$ as in (3.11) and (3.12). We write

$$\Pi_{k=1}^j \left[1 + s \, \alpha_k^{-1} \, \exp(sT_k) \right] \sim \Pi_{k=1}^j \left[1 + s\alpha_k^{-1} + s^2 \, T_k \, \alpha_k^{-1} \right]$$

$$\sim \exp \left[s\left(\sum_1^j \alpha_k^{-1}\right) - \tfrac{1}{2} s^2 \left(\sum_1^j (1 - 2\alpha_k \, T_k)/\alpha_k^2\right) + .. \right] .$$

The analogue of (3.12) is

$$z_j(t) = \int_{-i\infty}^{i\infty} ds \, \exp \left[s\left(t - \sum_1^j \alpha_k^{-1}\right) \right] \exp \left[\frac{s^2}{2} \sum_1^j \frac{(1 - 2\alpha_k \, T_k)}{\alpha_k^2} \right] z_o^*(s) .$$

This has exactly the same form as (3.12) except that the signal reaches the jth car at a time approximately $\sum_1^j \alpha_k^{-1}$, the sum of the times for adjacent cars, instead of j/α; and the new "diffusion constant" is now

$$\sum_1^j (1 - 2\alpha_k \, T_k) \, \alpha_k^{-2} .$$

If drivers are independently sampled from some population the sums over k are sums

$$\sum_1^j (1 - 2\alpha_k \, T_k) \, \alpha_k^{-2} .$$

If drivers are independently sampled from some population, the sums over k are sums of independent identically distributed random variables. A wave does not travel through a series of cars with exactly a constant wave velocity as in Figure IV 4 but, for all practical purposes, travels as if the signal moved a distance $(-\beta_k/\alpha_k)$ in a time $1/\alpha_k$ in going from the

$(k-1)^{th}$ to the kth car. It follows sort of a Brownian path. For sufficiently large j, however, this random path of propagation will have a time average velocity defined by

$$wave\ velocity\ =\ \lim_{j \to \infty} \frac{-\sum_{k=1}^{j} \beta_k/\alpha_k}{\sum_1^j 1/\alpha_k} = \frac{-E\{\beta/\alpha\}}{E\{1/\alpha\}}\ .$$

The last step follows from the law of large numbers. The expectations are expectations over the population of all drivers. This wave velocity is not necessarily equal to $-E(\beta)$.

If some $\alpha_j=0$, the equations of motion imply that the velocity of this car is independent of spacing which in turn means that this driver pays no attention to what the car ahead does. He will stop the propagation of any waves. We will encounter a more realistic version of what happens if $E\{1/\alpha\} = \infty$ in the non-linear theory.

To guarantee stability, we want a positive diffusion constant. This will certainly be the case if $\alpha_k T_k < 1/2$ for all cars. If there is a non-zero probability that $\alpha T > 1/2$, then there is also a non-zero probability that the sum in question is negative for any finite j. A disturbance may, therefore, be amplified at least in the early stages of its propagation, but, if

$$E\{(1 - 2\alpha T)/\alpha^2\} > 0\ ,$$

the disturbance will, with probability 1, eventually decay. The safe drivers can undo the bad effects of the unsafe ones, in the present context.

More interesting consequences of non-identical drivers arises in the non-linear effects of the unsafe ones, in the present context.

More interesting consequences of non-identical drivers arises in the non-linear theories. If the maximum velocities (desired speeds) of individual drivers are not all equal, it is possible for a line of (j-1) fast cars to pull away from a slow car if their speed exceeds the desired speed of the latter. It is also possible for fast cars to form queues behind a slow car. It is not too difficult to solve (at least approximately) the equations of motion in any

particular situation. The real problem is that the solution is sensitive to the ordering of the cars; so many different things can happen that it is difficult to formulate what might be considered a typical problem. On an infinitely long highway with infinitely many cars, all cars behind the slowest eventually queue behind the slowest. Consequently there is not even an equilibrium distribution. For any finite length highway or a highway with entrances and exits, one must keep a careful account of the entrance or exit of any slow cars.

We consider here a few miscellaneous features of this problem but we will raise more questions than we will answer.

b. <u>Exact solutions</u>. The non-linear theory discussed in section IV 6 can be used to study some of the analytic aspects of traffic theory with non-identical cars. We will consider only the special case of this theory in which the α_j are proportional to the V_j as in (6.9).

In equation (6.10a), we have already the transform $Z_j^*(x)$ for a line of cars initially moving with velocity u (u less than any of the V_k). This equation has been shown to give the wave propagation, diffusion, expanding waves, and shocks, but now we want it also to give a separation of flow into platoons traveling behind slow cars.

Suppose again that the lead car suddenly changes velocity from u to v so that

$$Z_j^*(s) = 1/(s + \varepsilon v) .$$

then

$$Z_j^*(s) = 1/(s + \varepsilon v) .$$

then

$$Z_j^*(s) = \frac{1}{(s + \varepsilon u)\Pi_{k=1}^{j} (1 - u/V_k)} - \frac{\varepsilon (v - u)}{(s + \varepsilon u) (s + \varepsilon v) \Pi_{k=1}^{j} (1 + s/\alpha_k)} \qquad (7.1)$$

This is a rational function of s. If the α_j are all different, the inverse transform is

223

$$Z_j(t) = \frac{e^{-\varepsilon v t}}{\Pi_{k=1}^{j} (1 - v/V_k)} - \sum_{m=1}^{j} \frac{e^{-\alpha_m t} (v - u)/V_m}{(1 - u/V_m)(1 - v/V_m) \Pi_{\substack{k=1 \\ k \neq m}}^{j} (1 - \alpha_m/\alpha_k)} \,. \qquad (7.2)$$

If $v < V_m = \alpha_m/\varepsilon$ for all m, $m \leq j$, so that the lead car travels slower than the free speed of the jth car or any car ahead of it, then for sufficiently large t, the first term of (7.2) will be large compared with those of the sum. This term alone, however, gives the steady state motion of the jth car at velocity v with the appropriate spacing. All cars with $V_j > v$ which also follow cars with free speeds $V_m > v$ (for all $m \leq j$) will, therefore, eventually follow the lead car. If, however, εv is larger than some α_m, $m \leq j$, the dominant term of (7.2) will be the one with the smallest α_m, thus the smallest V_m. The jth car will eventually find itself traveling with the velocity min V_m, $m \leq j$. The final velocity of the jth car is monotone non-increasing with j. Each car with a free speed less than any predecessor will travel slower than those ahead and also prevent any car behind from traveling any faster.

These qualitative properties obviously had to come out of the theory somehow. There are some other properties, however, that require a more detailed analysis. Suppose cars start with $u = 0$ at a traffic signal. We may wish to know how the fluctuations in driver behavior effect the average flow across the intersection. Does the separation into queues behind slow cars occur soon enough to cause a drop in flow at the intersection? It may also be of interest effect the average flow across the intersection. Does the separation into queues behind slow cars occur soon enough to cause a drop in flow at the intersection? It may also be of interest to investigate the shock profile to see how the cooperative behavior of many cars forces the behavior of different individual drivers.

Equation (7.2) is not a very convenient representation of the solution for large j. To find asymptotic solutions, it is better to go back to (7.1) and apply some asymptotic methods to the inversion formula for the Laplace transform. We write $Z_j(t)$ in the form

$$Z_j(t) = \frac{1}{\Pi_1^j (1 - u/V_k)} \int_{\gamma-i\infty}^{\gamma+i\infty} \frac{ds\, e^{st}}{(s + \varepsilon u)} \left(1 - \frac{\varepsilon(v - u)\, \Pi_1^j (1 - u/V_k)}{(s + \varepsilon v)\, \Pi_1^j (1 + s/\alpha_k)} \right) . \qquad (7.3)$$

The possible singular points of the integrand are at $s = -\varepsilon u$, $-\varepsilon v$, $-\alpha_1$, $--$, $-\alpha_j$. We have assumed already that $u < V_m$, thus $-\varepsilon u > -\alpha_j$. Of these points, the one furthest to the right in the s-plane is, therefore, either $-\varepsilon u$ or $-\varepsilon v$ depending upon whether the motion is an acceleration or a deceleration. The two cases we already know must behave quite differently. We consider the acceleration first, $-\varepsilon u$ is to the right of $-\varepsilon v$.

The quantity in the bracket of (7.3) vanishes at $s = -\varepsilon u$, so this is actually not a singular point of the integrand. For large j, the second term of the bracket is, however, rapidly varying with s. It has the value 1 at $s = -\varepsilon u$ but along the real axis it drops off rapidly for $s > -\varepsilon u$ and becomes very large for $s < -\varepsilon u$. It behaves almost like an exponential.

The second term of the bracket multiplied by e^{st}, with $t > 0$, will have a minimum along the real axis. For small enough s it is decreasing, but the exponential will eventually, for $s \to \infty$ overpower the polynomial decay and give a saddle point (a minimum along the real along the real axis. For small enough s it is decreasing, but the exponential will eventually, for $s \to \infty$ overpower the polynomial decay and give a saddle point (a minimum along the real axis but a maximum along the imaginary direction through this point). For small positive t, this saddle point occurs for some large value of s. If we evaluate $Z_j(t)$ by taking γ so that the path of integration crosses the real axis at the saddle point, the second term of the bracket

will be small compared with the first everywhere on this path. Its contribution can be neglected and so we obtain for $Z_j(t)$

$$Z_j(t) \sim \frac{e^{-\varepsilon ut}}{\Pi_1^j(1 - u/V_k)} .$$

For small t, the motion has the constant velocity u that is expected to persist for some non-zero time.

We will not obtain any significant deviation from this behavior until t becomes so large that the saddle point moves close to $s = -\varepsilon u$ where the two terms of the bracket are comparable.

Let s_o denote the position of this saddle point. Then s_o occurs where the logarithm of e^{st} times the second term of the bracket also has a saddle point and a zero derivative, i.e.,

$$\frac{d}{ds} \left\{ st - \log (s + \varepsilon v) - \sum_1^j \log \left(1 + s/\alpha_k\right) \right\}_{s = s_o} = 0 .$$

or

$$t = \frac{1}{\varepsilon v + s_o} + \sum_1^j \frac{1}{\alpha_k + s_o} . \tag{7.4}$$

The right-hand side of (7.4) is monotone decreasing in s_o to the right of any

$$\varepsilon v + s_o \qquad \overline{1} \qquad \alpha_k + s_o$$

The right-hand side of (7.4) is monotone decreasing in s_o to the right of any singularities. This equation therefore defines a unique one-to-one relation between s_o and t. It can be used either to determine the saddle point for given t or the time t when the saddle point occurs at a specified s_o. The time when the saddle point occurs at $-\varepsilon u$ is

226

$$t_o = \frac{1}{\varepsilon(v-u)} + \sum_1^j \frac{1}{\alpha_k - \varepsilon u} = \frac{1}{\varepsilon(v-u)} + \sum_1^j \frac{1}{\alpha_k(1-u/V_k)} . \qquad (7.5)$$

The meaning of this time is clear. Each term of the sum, $[\alpha_k(1-u/V_k)]^{-1}$ is the reciprocal slope of the velocity spacing curve for the kth car at the velocity u. From the discussion of wave propagation in the linear theory, we have seen that this is the effective propagation time of the wave of velocity u from the $(k-1)^{th}$ to the kth car. The sum is, therefore, the total propagation time for the disturbance at velocity u to reach the jth car from car 0. For large j, the sum dominates the extra term $\{\varepsilon(v-u)\}^{-1}$. This extra term can be identified with the diffusion effects associated with the impulse acceleration of the lead car. It is not particularly relevant here. As is expected, the jth car does not start to deviate from velocity u until a wave for velocity u can reach it at time t_o.

For t near t_o, we can approximate the integrand in the vicinity of $t \sim t_o$ and $s \sim s_o$. Specifically one expands the logarithm of the second term of the bracket in powers of $(s-s_o)$ through order $(s-s_o)^2$ so that this term is represented as an exponential of a quadratic form in s. This approximate integral can then be evaluated in terms of error functions, etc. What happens is the analogue of the wave edge phenomenon shown in Figure IV 21. For large j, the transition takes place in a time proportional to $j^{1/2}$ as compared with a duration of the entire disturbance which will be proportional to j. We shall not show the details of this but the transition takes place in a time proportional to $j^{...}$ as compared with a duration of the entire disturbance which will be proportional to j. We shall not show the details of this but asymptotic forms of $Z_j(t)$ can be found for this time range in a straightforward way.

After t has passed over this range, the saddle point s_o has shifted to the left of -εu. We again send the path of integration through the saddle point; there is no difficulty about moving the contour to the left of -εu because this is not a pole of the integrand. Along this

227

path, the second term of (7.3) is now much larger than the first and we can discard the 1 from the bracket. Until t is so large that the saddle point moves close to one of the poles at $(-\epsilon v)$ or one of the α_k, the logarithm of the integrand can again be expanded in powers of $(s-s_o)$ near the saddle point.

From (6.6a) we see that

$$v_j(t) = -\frac{1}{\epsilon} \frac{d\,Z_j/dt}{Z_j}\,.$$

The differentiation of Z_j with respect to t simply adds a factor s to the integrand of (7.3). For large j, the integrand is rapidly varying with s and the factor s is nearly constant, equal to s_o, over the range of s from which most of the contribution to the integration comes. Therefore

$$\frac{dZ_j(t)}{dt} \sim s_o\, Z_j(t)$$

and

$$v_j(t) \sim -\frac{s_o}{\epsilon} \tag{7.6}$$

This along with the equation (7.4) for s_o gives

$$t = \frac{1}{\epsilon(v - v_j(t))} + \sum_1^j \frac{1}{\alpha_k(1 - v_j(t)/V_k)}\,. \tag{7.7}$$

$$t = \frac{1}{\epsilon(v - v_j(t))} + \sum_1^j \frac{1}{\alpha_k(1 - v_j(t)/V_k)}\,. \tag{7.7}$$

As with equation (7.5) we can interpret this to mean that the velocity $v_j(t)$ of car j and the time t are related so that $v_j(t)$ is that velocity that can reach the jth car at time t if it propagates from the lead car with a time lag $\alpha_k^{-1}(1-v_j(t)/V_k)^{-1}$ from the $(k-1)^{th}$ to the kth car. Again one can carry out the asymptotic estimates to whatever accuracy one wishes.

228

For sufficiently large t, the value of s_0 given by (7.4) will move toward one of the poles of the right-hand side of (7.4). There are several possible situations to consider.

1. Suppose $v < \alpha_k/\varepsilon = V_k$ for all k<j and furthermore even for j→∞, $v < \inf_k V_k$, so that the poles at $s = -\alpha_k$ are bounded away from the one at $s = -\varepsilon v$. Then for large t, s_0 as given by (7.4) will lie close to $-\varepsilon v$. The method of expanding the logarithm of the integrand near s_0 will not necessarily give very accurate results if the behavior of the integrand near the saddle point is dominated by the contribution from a single pole. The derivation of (7.6) is still valid as a first approximation but to obtain subsequent approximations it is better to use a procedure similar to that used to get past the point $-\varepsilon u$ in (7.3). We can write the second factor in the bracket of (7.3) as

$$\frac{\varepsilon(v - u) \, \Pi_1^j \, (1 - u/V_k)}{(s + \varepsilon v) \, \Pi_1^j \, (1 + s/\alpha_k)} = \frac{\varepsilon(v - u) \, \Pi_1^j \, (1 - u/V_k)}{(s + \varepsilon v) \, \Pi_1^j \, (1 - v/V_k)} \tag{7.8}$$

$$+ \frac{\varepsilon(v - u) \, \Pi_1^j \, (1 - u/V_k)}{(s + \varepsilon v) \, \Pi_1^j \, (1 - v/V_k)} \left[- 1 + \frac{\Pi_1^j (1 - v/V_k)}{\Pi_1^j (1 + s/\alpha_k)} \right] .$$

We have simply added and subtracted the simple pole at $s = -\varepsilon v$. The first term of (7.8) contributes to $Z_j(t)$ a term

$$\frac{e^{-\varepsilon ut}}{\Pi_1^j (1 - u/V_k)}$$

$$\frac{e^{-\varepsilon ut}}{\Pi_1^j (1 - u/V_k)}$$

which alone would correspond to a final motion at the constant velocity v. The second term of (7.8), the transient part, has no pole at $s = -\varepsilon v$. Its contribution to $Z_j(t)$ can be handled in the same way that (7.3) was evaluated when the saddle point crossed over $s = -\varepsilon u$. The saddle point of the second term in the bracket of (7.8) satisfies the same type of equation as (7.4)

229

but with the term $(\varepsilon v+s_o)^{-1}$ missing. Once this newly defined s_o moves past $-\varepsilon v$, the contribution of the second term of (7.8) to $Z_j(t)$ dies very rapidly. As for s_o near $-\varepsilon u$, there will be a transition region as the wave of final velocity v passes over the jth car.

2. If $v = \inf_k V_k$, i.e., the poles at $s = -\alpha_k$ crowd close to the one at s_o, then the procedure used in 1 is probably not necessary until t is so large for some finite j that the single term of (7.4), $(\varepsilon v+s_o)^{-1}$, makes a contribution to t comparable with that of all the other terms combined. The phenomenon that is at issue here is that if some of the free speeds V_k are very close to the final velocity v, the wave propagation is retarded significantly by cars which will approach the velocity v only after the spacing has become very large. The velocity of the jth car, which in all cases is rather close to $-s_o/\varepsilon$ of (7.4) approaches v only at very large t.

3. If $v > \alpha_k/\varepsilon = V_k$ for some $k < j$, then the roles of the εv and the largest $(-\alpha_k)$ are reversed. The saddle point s_o cannot get to the left of this value of $(-\alpha_k)$ and the velocity v_j approaches the lowest velocity V_k, $k \le j$. Again we expect as in cases 1 vs. 2 above that the behavior of $Z_j(t)$ will depend upon whether there is a single slow car dominating the behavior or many slow cars approach their maximum speeds. In the latter case the wave propagation is slower because each slow car wishes to have a very large spacing.

The detailed evaluation of asymptotic formulas for all these cases is straightforward slower because each slow car wishes to have a very large spacing.

The detailed evaluation of asymptotic formulas for all these cases is straightforward but somewhat tedious. The important thing is that waves do propagate in this non-linear theory as one would expect from the analysis of the linear theory, provided the car velocity associated with the wave does not exceed or come too close to the free speeds of any cars.

For a deceleration $v < u < \max V_k$, the asymptotic evaluation of (7.3) is a little easier. The singularity furthest to the right in the s-plane is at $s = -\varepsilon v$, the next is at $s = -\varepsilon u$. If we shift the contour of integration so that $-\varepsilon u < \alpha < -\varepsilon v$ then, for large j, the second factor in the bracket will be small compared with the first everywhere along the path. In shifting the contour we have passed over the pole at $s = -\varepsilon v$ and so we must pick up the residue there. This may or may not be significant but in any case we have

$$Z_j(t) \sim \frac{e^{-\varepsilon u t}}{\Pi_1^j(1 - u/V_k)} + \frac{e^{-\varepsilon v t}}{\Pi_1^j(1 - v/V_k)} \ . \tag{7.9}$$

The term which we have neglected can still be estimated by a saddle point integration. It describes the approach to a shock structure as j increases but (7.9) alone describes the final shock for $j \rightarrow \infty$.

From (6.6a) and (7.9) we obtain

$$v_j(t) = \frac{(u + v)}{2} + \frac{(v - u)}{2} \tanh \left\{ \frac{(u - v)\, \varepsilon(t - \tau_j)}{2} \right\} \tag{7.10}$$

with

$$\tau_j = \frac{\sum_{k=1}^j \log(1 - v/V_k) - \sum_{k=1}^j \log(1 - u/V_k)}{(u - v)\varepsilon} \tag{7.11}$$

These reduce to (6.29) and (6.30) if $V_k = V$ independent of k.

$$\tau_j = \frac{}{(u - v)\varepsilon}$$

These reduce to (6.29) and (6.30) if $V_k = V$ independent of k.

Even for this case of non-identical drivers we see from (7.10) that every car has exactly the same shock trajectory as any other, except for a translation in time and space. That these trajectories are all similar is, no doubt, a consequence of the restriction that $\alpha_j/V_j = \varepsilon$ be independent of j. The translations are, however, subject to variation due to the

231

differences in the V_k. The propagation time τ_j is the sum of time lags for cars 1 to j. The shock wave, therefore, travels along a random path in much the same manner as the waves propagate.

8. Space and time-dependent highways. We have considered until now only highways that had time and space independent properties, i.e., we assumed that the q vs k relation was not explicitly dependent upon x or t. It is of some interest to know also what happens if a highway has curves, grades, etc., so that the q vs k relation varies with x, or what happens if it starts to snow so that the q vs k relation depends upon t. We will consider here only the modifications to the continuum theory since this describes in essence what happens also in the more complicated discrete theories. Some of the results described here were obtained by Lighthill and Whitham [15], De [22], and Robertello [23].

Suppose that at each position x and t, there is a specified flow-density relation so that

$$q = Q(k,x,t)$$

and that this relation is valid even if k is itself a function of x and t. The conservation equation

$$\frac{\partial k}{\partial t} + \frac{\partial q}{\partial x} = 0 \tag{8.1}$$

is still valid except that we must now interpret

$$\frac{\partial k}{\partial t} + \frac{\partial q}{\partial x} = 0 \tag{8.1}$$

is still valid except that we must now interpret

$$\frac{\partial q(k,x,t)}{\partial x} = Q_k(k,x,t)\,\frac{\partial k}{\partial x} + Q_x(k,x,t) \tag{8.2}$$

in which Q_k and Q_x represent the partial derivatives of Q(k,x,t) with respect to the first and second arguments, respectively, for fixed values of the remaining two.

This equation is still a first order partial differential equation for $k(x,t)$ but, with $Q_x \neq 0$, it is not generally possible to obtain simple explicit solutions. Solutions can, however, be constructed graphically or numerically and the existence of solutions can be established [24].

If at time $t = 0$ we are given $k(x,0)$, then we can also evaluate $Q_k(k(x,0),x,0)$ and $Q_x(k(x,0),x,0)$ from the given function of $Q(k,x,t)$. Through each point $(x,0)$ in the (x,t)-plane, we now draw a line (characteristic) of slope $Q_k(k(x,0),x,0)$ as illustrated in Figure IV 23. We interpret the quantity

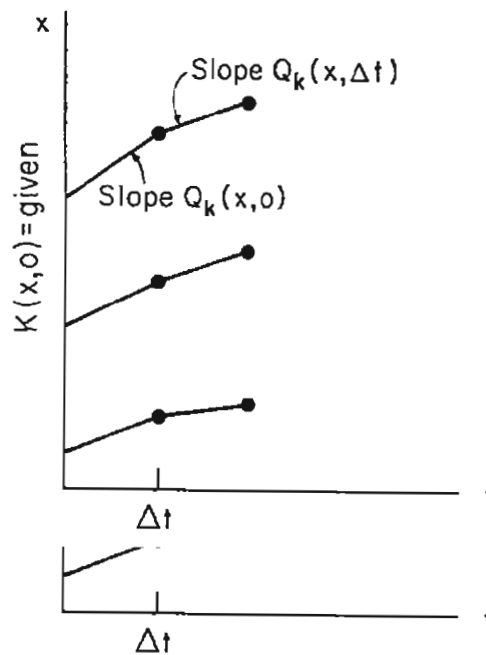$$\frac{\partial k}{\partial t} + Q_k \frac{\partial k}{\partial x} = \frac{dk}{dt}$$

Fig. IV 23
Iterative solution scheme

as the time derivative of k along the direction of the characteristic line on which

233

$$\frac{dx}{dt} = Q_k .$$

The conservation equation requires that in a small time Δt, the change in k along this characteristic is given by

$$\Delta k \sim \Delta t \; \frac{dk}{dt} \sim Q_x(k(x,0),x,0) \; \Delta t . \tag{8.3}$$

By moving along each of the characteristics we can now evaluate k(x,t) for all x at time t = Δt (provided no characteristics intersect). The same information is now available at time Δt as originally at time 0. The procedure can be continued with the k(x,Δt) as "initial data."

Since Q_k and Q_x depend upon k, as well as x and t, the slopes of the characteristics will be slightly different at time Δt than at t = 0. The iterative procedure generates a family of non-linear characteristic curves, curves which at every point (x,t) have a slope Q_k(k(x,t),x,t), and the value of k(x,t). If characteristics should intersect we must create shocks as in the homogeneous solutions. The shock equations are exactly the same as before except that the relation between q and k will be x and t-dependent.

Since the solutions of the differential equations are rather difficult to construct, it is perhaps of more interest to consider the simpler solutions for the two special cases in which q is a function of k and either x or t but not both.

If q = Q(k,t) so that Q_x = 0, then according to (8.3), Δk = 0 along the characteristic.
is a function of k and either x or t but not both.

If q = Q(k,t) so that Q_x = 0, then according to (8.3), Δk = 0 along the characteristic. Thus k is a constant along the characteristic. The equations for the characteristics themselves are also fairly simple now. Starting from some initial point (x_o, 0), we observe the value of k (x_o, 0) which will be the value of k everywhere along the characteristic passing through the point (x_o, 0). The equation for the characteristic is now the ordinary differential equation.

$$\frac{dx(t)}{dt} = Q_k(k(x_o,0),t), \quad x(0) = x_o$$

or

$$x(t) = x_o + \int_o^t Q_k(k(x_o, 0), \tau)\, d\tau.$$

The function of $Q(k,t)$, and consequently also $Q_k(x,t)$, are supposed to be given as are the values of $k(x_o,0)$. The integrand is, therefore, a known function, and the characteristics can be found directly from the initial data.

The differential equation (8.1) shows the duality of the interchange of q and t for k and x, respectively, which has been encountered many times before. It appears here because the conservation equation is equivalent to the requirement that car trajectories be continuous curves in the (x,t) plane with spatial density k. They also define continuous curves in the (t,x) plane with a "time density" q.

If $q = Q(k,t)$ gives solutions with k constant along characteristics as shown above, one might surmise from this duality that if $q = Q(k,x)$ in which the t-dependence is replaced by an x-dependence, that k=constant along a characteristic will be replaced by q=constant. It might, in this case, also be convenient to think of the q-k relation as defining k as a function of q instead of q as a function of k. If then $k = K(q,x)$, the conservation equation becomes an

~~equation for q~~

instead of q as a function of k. If then $k = K(q,x)$, the conservation equation becomes an equation for q

$$\frac{\partial K(q,x)}{\partial t} + \frac{\partial q}{\partial x} = 0$$

$$K_q(q,x)\frac{\partial q}{\partial t} + \frac{\partial q}{\partial x} = 0.$$

This equation does imply that q is constant along the characteristic curves defined by either

235

$$\frac{dx}{dt} = Q_k(k,x) \quad or \quad \frac{dt}{dx} = K_q(q,x) \ .$$

Of the various specific problems of the two types described above, the second type, q = Q(k,x) furnishes the more interesting examples partly because the function k = K(q,x) is double valued. Instead of considering an initial value problem, suppose we have a section of highway fed by some input at x = 0. The input flow q(0,t) is given. Suppose also that the highway has a bottleneck; Q(k,x), for any given k, is monotone decreasing in x until some point x_3 and then it increases again. A possible family of q-k curves for various values of x are shown in Figure IV 24. As x increases it goes from x_1 to x_2 to x_3 and then perhaps back to x_2 to x_1.
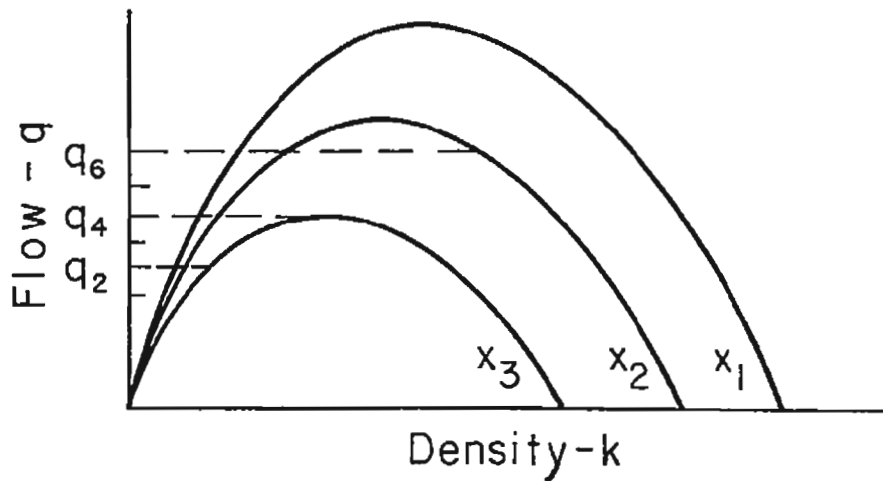
Fig. IV 24
Density-k

Fig. IV 24
A sequence of q-k curves

Prior to time 0 we suppose that a steady flow existed at a flow value q_1 and a density k(x) given by k(x)=K_(q_1,x), the smaller of the two values of k associated with any value of

236

q_1 and x. Now let q(0,t) increase from q_1, to q_2 to q_3, etc., until it exceeds the capacity of the bottleneck, i.e.,

$$q > \max_{k} Q(k,x_3) \ .$$

The various characteristics are shown in Figure IV 25. The initial increase in q
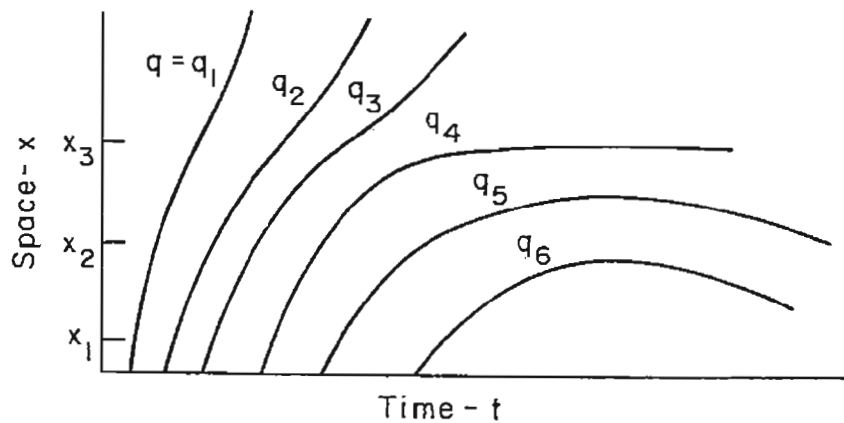
Fig. IV 25
Formation of a shock behind a bottleneck

starting at time t = 0 say propagates along the characteristics for flow q, whose equation is given by

$$t(x) \ = \ \int_{o}^{x} K_q(q_1,\xi)d\xi \ .$$

The values of Q or K at various positions x can be obtained directly from the q-k curves

$$t(x) \ = \ \int_{o}^{} K_q(q_1,\xi)d\xi \ .$$

The values of Q_k or K_q at various positions x can be obtained directly from the q-k curves. For flow q_1, they are the slopes or reciprocal slopes, respectively, of the q-k curves at $q = q_1$. The slopes are seen to be decreasing as x goes from x_1 to x_2 to x_3. These are also the slopes of the characteristics at x_1, x_2, x_3. The characteristic for $q = q_1$ is thus concave for $x < x_3$. Along this characteristic and to the left of it in Figure IV 25, the flow is everywhere q_1.

237

For flow q_2 the characteristics have smaller slopes at any given x. They are also decreasing with x for $x < x_3$. For flow q_3, just below the capacity of the bottleneck, the characteristic becomes almost horizontal as it passes the bottleneck. As q approaches the capacity of the bottleneck, the characteristics approach a limit characteristic represented by q_4 which has a horizontal asymptotic at the bottleneck.

For $q > q_4$, a characteristic at flow q_5, for example, will become horizontal at that location where the capacity of the road is q_5. Unlike the characteristic for q_4, however, which approaches the horizontal asymptotically in time, this characteristic will become horizontal at a finite time and subsequently turn back upstream as illustrated in Figure IV 25.

At location x_2, for example, there are two possible densities on the q-k curve corresponding to a flow q_5. At the lower density Q_k is positive, but at the higher density Q_k is negative. The former value of Q_k gives the slope of the characteristic as it goes forward past x_2. The latter value of Q_k gives the slope of the characteristic after it has turned back and passed x_2 again.

The details of how this evolves are not important. Certainly a flow q_5 cannot pass the bottleneck and any excess flow above q_4 will cause vehicles to pile-up behind the bottleneck, thus pushing the density over to the left-hand side of the q-k curve. That characteristics which have turned upstream will intersect other characteristics for higher flows moving downstream thus pushing the density over to the left-hand side of the q-k curve. That characteristics which have turned upstream will intersect other characteristics for higher flows moving downstream means that a shock must form somewhere upstream of the bottleneck.

9. Experimental evidence. The theories described in this chapter have been shown to be consistent with each other and have been described in sufficient detail that one can obtain a

238

clear picture of the qualitative predictions that the theory would produce when applied to a variety of real physical situations. Some of these predictions agree approximately with what is actually observed, enough that one cannot afford to scrap the theory, but unfortunately the theory still has some serious faults, not just quantitative inaccuracies which one would normally expect from such a crude model, but gross errors in the qualitative description of what really happens in certain situations. Before going on to more elaborate theories, it is therefore advantageous to review some of the successes and failures of these theories.

The key assumption made in the theory was that the relations between velocity and spacing or between flow and density determined under steady-state conditions are also true for time dependent flows. The main conclusions of the theory follow from this assumption alone which was made primarily as a mathematical convenience or plausible conjecture. There is no a-priori reason why it should be true. It implies that if the spacing between cars is known then the velocity is uniquely determined and is therefore independent of accelerations, past history of the motion, etc. We have considered the possibility that different drivers may have velocity-spacing relations, which accounts for some random effects, but we have not considered the consequences of the fact that a specified driver will experience random fluctuations in velocity and/or spacing about the mean velocity-spacing curve for this particular driver.

fluctuations in velocity and/or spacing about the mean velocity-spacing curve for this particular driver.

Experiments on traffic flow are very difficult to interpret. The experimental data almost always show violent fluctuations and it is very hard to decide whether some observed behavior is just some random deviation from a typical pattern or if it is itself the typical pattern. At the present time, what is "good" or "bad" agreement with experiment is often a

matter of personal judgement because seldom are enough data available to estimate the probability distributions for driver behavior or variations in the q vs k relations from one point to the next along a highway which one would need to know in order to make any real test of significance. We consider first some of the experiments which, for the most part, support the theory described here.

A fairly extensive series of experiments have been performed at General Motors Corporation, Detroit, by R. Herman along with various collaborators. This began with the work of Chandler, Herman and Montroll [6] in 1958 and has continued to the present time. Most of these experiments are "car-following" experiments in which cars equipped with various recording devices to make records of speed and spacing are driven by various people usually on a test track. One driver, a leader, is instructed to perform a series of maneuvers and a second (possibly also others) is instructed to "follow the lead car at what you consider to be a minimum safe distance at all times." This type of experiment has the obvious weakness that what drivers will do in an experiment while following certain instructions is not necessarily what they would do when driving under normal conditions on a highway, but it should at least give some hints.

As mentioned before, these experiments were first designed to establish best linear relations between the acceleration of a car and the velocity differences and/or spacing between that car and its leader. Chandler, Herman and Montroll first concluded from these experiments that the dependence of acceleration upon spacing was not large enough to improve significantly the correlation and so they concentrated on fitting an equation of the form (2.3) to the experimental data. From a large collection of data, values of the acceleration

at times t + T were compared with values of $v_{j-1}(t) - v_j(t)$. For any fixed T, a regression line was drawn giving the slope α in (2.3) that minimized the mean square deviation from (2.3). The correlation coefficient (which would have the value one if the two above quantities were exactly proportional) of these variables was then considered as a function of T. The values of T and α which maximized the correlation coefficients were then considered as the "observed values."

The success of the theory must be measured by the fact that the correlation coefficients (depending upon the driver) were mostly in the range 0.8 to 0.9. Whether this is "good" or not is difficult to say, but it does not seem unreasonable in consideration that the difference in behavior of eight drivers gave a range of T values from about 1 second to 2 seconds, a range of values of α between 0.17 sec^{-1} to 0.74 sec^{-1}, and a range of values for αT between 0.18 and 1.04. There are indeed tremendous variations in behavior from driver to driver and one is trying to fit a simple model for the average behavior in which the fluctuations in small samples nearly swamp the thing one is trying to observe.

An important feature of these experiments is that they correlate accelerations with velocity differences and not velocities with spacings. Most of the data is, therefore, taken during time dependent motion and not in the steady state. If there were a significant dependence of the velocity-spacing relations on the acceleration, for example, the value of α during time dependent motion and not in the steady state. If there were a significant dependence of the velocity-spacing relations on the acceleration, for example, the value of α one obtains by fitting (2.3) would not necessarily be the same as that obtained by fitting (2.2) to data taken in the steady state. The fact that one is forcing a best fit of the data to (2.3) and that (2.3) is the derivative of (2.2) is likely to bias the statistical estimates so as to

underestimate any differences that may exist, but, on the other hand, there is no a-priori reason why the theory should fit at all.

The value of αT, which is quite crucial to the question of macroscopic stability, was found to have an average value for the various drivers slightly above 1/2, the theoretical limit of stability, but certain drivers had values of αT considerably above 1/2. This represents a potential failure of the theory, however, that must be examined more carefully in conjunction with other observations. For if αT is larger than 1/2, or could be made greater than 1/2 by a suitable choice of drivers, the present theory would predict a type of instability that would seem to be more violent than is consistent with observations of macroscopic behavior. More recent experiments with buses [31], however, give values of $\alpha T < 1/2$.

In subsequent papers in this series [25-31] non-linear models were fitted to experimental data which were, in effect, the derivative of (5.1) for suitable functions V. The form of V deduced from dynamic experiments did agree quite well with those obtained under supposedly steady state conditions [25]. A review of most of the results of these experiments is contained in the papers by Herman and Potts [26], and Herman and Rothery [30]. In the latter it is shown that models with only interactions between adjacent cars are nearly equivalent experimentally to models containing interactions also between next nearest neighbor cars. It is also mentioned in both of these papers that the effective or average α in equivalent experimentally to models containing interactions also between next nearest neighbor cars. It is also mentioned in both of these papers that the effective or average α in (2.3) is somewhat larger for decelerations than for accelerations. Although this difference does not seem very large, it has potentially important consequences that are not consistent with the theories described here.

Kometani and Sasaki [7] in their early work on car-following also did some experiments. They instructed a lead driver to make periodic hard accelerations and decelerations. The following driver was again instructed to follow at what he considered a safe driving distance. They found values of αT closer to 1 than 1/2. The experiments were not done with long lines of cars so one could not check the theoretical prediction that the propagation of waves would be highly unstable for $\alpha T = 1$. The periodic motion of the leader car is also quite unnatural. Helly [32] also analyzed some other experimental data of Forbes et al [33] and found values of αT greater than 1/2.

All of these experiments were done with drivers who knew they were being watched and were also driving under instructions to follow at what they considered a safe distance which, no doubt, means that they were actually driving somewhat closer than they ordinarily would drive.

The agreement between theory and experiment is not very sensitive to the value of T; it is much more sensitive to α. It may be that the experiments are not accurate enough to give meaningful estimates of T or, if the theory does not have quite the correct form, the procedure of forcing the best fit to the theory of experimental data may force T to assume values which have little to do with reaction times. Certainly the experimental estimates of T from car-following experiments have raised more questions than they have answered.
values which have little to do with reaction times. Certainly the experimental estimates of T
from car-following experiments have raised more questions than they have answered.
Although there is considerable experimental evidence for instability of wave propagation, it is not as violent a type of instability as would be predicted by a theory with $\alpha T > 1/2$.

Further evidence (and perhaps the most convincing evidence) for the validity of the continuum or car-following theories has been obtained by Foster [35]. He observed the

acceleration of queues formed at a traffic signal. Crossing times of cars past six points near the intersection were recorded, from which approximate trajectory plots were drawn. Calculations were made of headway and spacing distributions at various speeds, average speed-density curves, and wave speeds. The queues were about 10 or 12 cars long for single observations but observations were made of many such queues.

The experiments showed that the flow-density relation (over a limited range of densities) had a maximum and a shape approximately as postulated in the theory described above. The acceleration waves showed a tendency to fan somewhat with a wave speed that increased with car speed. Statistical fluctuations were quite large. This is about as much as the experiment showed conclusively.

In this experiment all data were for cars undergoing acceleration. There was no way of checking if the flow-density relation obtained under this acceleration is the same as that which would exist in a steady state or under deceleration.

In contrast with the above experiments, all of which were done with platoons of about ten or fewer cars at a time, a long series of experiments, starting in the early 1950s, have been in progress to study the flow of traffic in the Holland and Lincoln tunnels of New York City. Most of these experiments were done by or in collaboration with Leslie Edie and Robert Foote [27-29, 32, 33, 35-42]. These tunnels are about two miles long and, although they are City. Most of these experiments were done by or in collaboration with Leslie Edie and Robert Foote [27-29, 32, 33, 35-42]. These tunnels are about two miles long and, although they are two lanes wide, it is against the law to change lanes. Under congested conditions, the tunnels will hold over a hundred cars any given time so this is an ideal experimental ground to test the large-scale effects of car-following. It is particularly these experiments that have shown

244

serious qualitative inconsistencies with the theory, which have still not been resolved in a satisfactory way.

Some of the unexplained observations are the following:

1. Flows averaged over periods of about an hour and involving a thousand or more cars still show fluctuations from hour to hour or from day to day of as much as 20 percent. These are not fluctuations in demand; there is a queue of cars waiting at the entrance at all times.

2. The average velocity of cars depends upon the position in the tunnel but the pattern of this variation is quite different in different tunnels. It also shows wide variation from day to day. [36].

3. The flow shows instability. There is a point particularly in the Holland Tunnel identified as a bottleneck. At various times a disturbance will originate at the bottleneck and start to move backwards toward the entrance. It grows in amplitude very quickly until it causes a complete stoppage which then continues to travel all the way back to the entrance. An acceleration wave is formed afterwards and follows the stoppage wave but, contrary to the theory described above, this acceleration wave does not overtake the deceleration wave. The deceleration wave does form what could be described as a shock; its shape seems to remain more or less fixed; and it does have a velocity in qualitative agreement with what the theory deceleration wave does form what could be described as a shock; its shape seems to remain more or less fixed; and it does have a velocity in qualitative agreement with what the theory would predict (about 10 m/h). The acceleration wave, however, does not seem to fan out as much as it should and is not obviously different in shape from the deceleration wave; it could also be called a shock. These disturbances occur at random time intervals with approximately an exponential distribution and a mean time between occurrences of about <u>four minutes.</u>

245

There can be more than one stoppage in the tunnel simultaneously and there is no evidence that this is the result of some oscillation caused perhaps by a wave traveling back and forth between the entrance and the bottleneck. The four minute "time constant" is about 100 times larger than any of the time constants in the car-following theory [37-40].

4. A variety of different averaging procedures have been used to produce graphs of average flow vs average density. Data of this type are recorded automatically and are available in almost unlimited amounts. Typically these graphs show a) for densities just above the value which produces the maximum observed average q, the fluctuations in q become relatively large; b) despite what should be congested flow behind the bottleneck, the densities of cars are typically at or less than that which gives the maximum flow instead of more, as the theory of the last section predicts; and c) as k increases past the point of maximum q, the average q seems to drop quite suddenly as if there were a discontinuity and then remains fairly constant as k increases still further. [41].

5. Traffic signals were installed at the entrance of the tunnel to limit the flow entering the tunnel. This has the effect of increasing the velocities (which for some reason also improves the stability) and also introduces gaps in the traffic steam that will absorb any shocks that might still be generated. Measurements of the q vs k curves under this controlled flow were significantly different from those of the uncontrolled flow. This means, of course, shocks that might still be generated. Measurements of the q vs k curves under this controlled flow were significantly different from those of the uncontrolled flow. This means, of course, that the relation between q and k is not a unique one as implied by the theory [40,42]. By limiting the number of cars which can enter the tunnel during any interval of time, these controls have actually given a slight increase in total average flow. They also reduce

accidents, stalled cars, ventilation problems, etc. Improvements have been made despite the limited understanding of what goes on.

6. To collect the type of data from which one can draw the actual trajectories of cars is rather tedious. One cannot make aerial observations of the tunnel traffic and so one must have either human or electronic observers stationed simultaneously at a large number of points in the tunnel. Experiments with about eight observations posts have been made. Trajectory plots show the large-scale propagation of stoppages but more detailed analysis of these has only produced a maze of contradictory conjectures. If on these trajectory plots one draws contours of constant velocity, constant density, and constant flow, they do not coincide; they do not even show necessarily the same qualitative shape. One thing does seem to appear quite consistently. If the flow approaching the bottleneck increases, the density at the bottleneck also increases and creates what one might describe as a super-saturated flow. This can be sustained for a certain time but it usually collapses after awhile giving a rather sharp drop in both the velocity and the flow (by as much as a factor of two). Once the flow has dropped, it is rather slow to recover. During this motion, the state of flow in the q-k plane follows a path that loops clockwise. A behavior of this type was noted in the paper of Lighthill and Whitham [15] even in 1955 based upon experiments in Amsterdam.

Instability that is perhaps closely related to that observed in the tunnels also appears in Lighthill and Whitham [15] even in 1955 based upon experiments in Amsterdam.

Instability that is perhaps closely related to that observed in the tunnels also appears in other forms. If one has a long queue, as for example at a toll gate on an expressway, the flow of traffic through the toll gate may be quite smooth but if one goes back a mile or so in the queue one finds that the traffic is "stop and go." Since there is no obvious strong fluctuation in flow at the toll gate one can only conclude that this stop and go driving is self-generated.

Apparently no experimental studies have been made of the behavior of long queues, but some experiments were performed on the Merritt Parkway in Connecticut [43] when a temporary bridge was in use after a flood. These experiments showed some similarities to those done in the Holland tunnel; they gave some odd-shaped average q - k curves and some oscillations when the flow exceeded the capacity of the bottleneck.

The art of aerial reconnaissance as applied to traffic studies is developing rapidly. Photographic studies of various traffic patterns are now being made from helicopters which can hover for a long time over a single spot. Even with the detailed data that is presently becoming available, however, it is a difficult and tedious procedure to try to separate fluctuations from what may be systematic trends. It does appear, however, from photographs that drivers do not always drive at their shortest safe distance even in congested traffic. Photographs taken of traffic on the George Washington Bridge by Dickens and Jordan of the Port of New York Authority show quite clearly that when a driver switches from one traffic lane to another, he does not usually produce any sudden reaction from the other drivers in either lane. A driver in one lane, however, suddenly finds that his spacing to the car in front has been reduced by perhaps a factor of two or more because a new car has moved into the middle of the gap. A driver in the other lane suddenly finds his spacing increased by a factor of about two. If drivers insist upon having some set spacing, the former driver would put on middle of the gap. A driver in the other lane suddenly finds his spacing increased by a factor of about two. If drivers insist upon having some set spacing, the former driver would put on his brakes and the latter would accelerate. What happens in fact is that neither driver does much of anything. Some readjustments usually occur within a half minute or so but it is not obvious that any drivers intentionally try to make these adjustments.

The car-following experiments were done with the instruction that the following driver was to maintain what he considered a safe driving distance. It is not obvious that this is what drivers typically are doing.

References
Chapter IV

1. E.M.J. Herrey and H. Herrey, "Principles of physics applied to traffic movement and road conditions," American Journal of Physics 13, 1-14 (1945).

2. D.R. Cox, "Renewal Theory," Methuen and John Wiley (1962).

3. A. Reuschel, "Fahzeugbewegungen in der Kolonne bei Gleichförmig Beschleunigtem oder vergögertem Leitfahrzeug." Zeits. Östereich. Ing. u. Arch. Vereines 95, 59-62, 73-77 (1950).

4. A. Reuschel, "Fahrzeugbewegungen in der Kolonne," Öst. Ing. Archiv 4, 193-215 (1950).

5. L.A. Pipes, "An operational analysis of traffic dynamics," J. Appl. Phys. 24, 274-81 (1953).

6. R.E. Chandler, R. Herman, and E.W. Montroll, "Traffic dynamics; studies in car following," Opns. Res. 6, 165-184 (1958).

7. E. Kometani and T. Sasaki, "On the stability of traffic flow," J. Opns. Res. Japan 2, 11-26 (1958).

8. T.S. Chow, "Operational analysis of a traffic-dynamics problem," Opns. Res. 6, 827-834 (1958).

9. R. Herman, E.W. Montroll, R.B. Potts, and R.W. Rothery, "Traffic dynamics: analysis of stability in car following," Opns. Res. 7, 86-106 (1959).

10. T. Sasaki, "On the stability of traffic flow, report II," J. Opns. Res. Soc. Japan 2, 60-79 (1959).

11. E. Kometani and T. Sasaki, "A safety index for traffic with linear spacings," Mem. Fac. Engrg. Kyoto University 21, 229-246 (1959); Opns. Res. 7, 704-720 (1959).

11. E. Kometani and T. Sasaki, "A safety index for traffic with linear spacings," Mem. Fac. Engrg. Kyoto University 21, 229-246 (1959); Opns. Res. 7, 704-720 (1959).

12. Takasi Kisi, "Traffic dynamics: analysis as sampled-data control systems," J. Opns. Res. Soc. Japan 2, 114-123 (1960).

13. R. Bellman and K.L. Cooke, "Differential-difference equations," Academic Press, New York-London (1963).

14. G.F. Newell, "Nonlinear effects in the dynamics of car following," Opns. Res. 9, 209-229 (1961).

15. M.J. Lighthill and G.B. Whitham, "On kinematic waves, II. A theory of traffic flow on long crowded roads." Proc. Royal Soc. London A 229, 317-345 (1955). Also reprinted in "An introduction to traffic flow theory," edited by D.L. Gerlough and D.G. Capelle, Highway Research Board special report 79, Washington (1964).

16. P.I. Richards, "Shock waves on the highway," Opns. Res. 4, 42-51 (1956).

17. R.E. Franklin, "Single lane traffic flow," Civil Eng. and Public Works Rev. 55, 1429- (1960).

18. R.E. Franklin, "The structure of a traffic shock wave," Civil Eng. and Public Works Rev. 56, 1186-1188 (1961).

19. R.E. Franklin, "The propagation of disturbances in a single lane of traffic," Civil Eng. and Public Works Rev. 57, 344-346, 487-489 (1962).

20. E.L. Ince, "Ordinary differential equations," Dover, New York (1944).

21. Bateman Manuscript Project, "Higher transcendental functions," Vol. 2, McGraw-Hill, New York (1953).

22. S.C. De, "Kinematic wave theory of bottlenecks of varying capacity," Proc. Cambridge Phil. Soc. 52, 564-72 (1956).

23. R. Robertello, Port of New York Authority, unpublished.

24. R. Courant and D. Hilbert, "Methods of Mathematical Physics," Vol. II, Interscience, New York (1962).

25. D.C. Gazis, R. Herman, and R.B. Potts, "Car following theory of steady state traffic flow," Opns. Res. 7, 499-505 (1959).

26. R. Herman and R.B. Potts, "Single lane traffic theory and experiment," Proc. symposium on the theory of traffic flow, Detroit (1959), R. Herman, ed., Elsevier, Amsterdam (1961), pp. 120-146.

26. R. Herman and R.B. Potts, "Single lane traffic theory and experiment," Proc. symposium on the theory of traffic flow, Detroit (1959), R. Herman, ed., Elsevier, Amsterdam (1961), pp. 120-146.

27. D.C. Gazis, R. Herman, and R.W. Rothery, "Non-linear follow-the-leader models," Opns. Res. 9, 546-567 (1961).

28. R. Herman and R.W. Rothery, "Microscopic and macroscopic aspects of single lane traffic flow," J. Opns. Res. Soc. Japan 5, 74-93 (1962).

29. L.C. Edie, R.S. Foote, R. Hermans, and R. Rothery, "Analysis of single lane traffic flow," Traffic Engr. 33, 21-27 (1963).

30. R. Herman and R. Rothery, "Car following and steady state flow," Proc. second symposium on the theory of road traffic flow, London (1963).

31. R. Rothery, R. Silver, R. Herman, and C. Torner, "Analysis of experiments on single lane bus flow," Opns. Res. 12, 913-933 (1964).

32. W. Helly, "Simulation of bottlenecks in single lane traffic flow," Proc. symposium on the theory of traffic flow, Detroit 1959, Elsevier, Amsterdam (1961), pp. 207-238.

33. T.W. Forbes, M.J. Zagorski, E.L. Holshouser, and W.A. Deterline," Measurement of driver reactions to tunnel conditions," Proc. Highway Research Board 37, 345-357 (1958).

34. J. Foster, "An investigation of the hydrodynamic model for traffic flow with particular reference to the effect of various speed-density relationships," Proc. first conference of the Australian Road Research Board 1962, Vol. 1, p. 229-257.

35. E.S. Olcott, "The influence of vehicular speed and spacing on tunnel capacity," Opns. Res. 3, 147-167 (1955).

36. R.I. Strickland, "Traffic operation of vehicle tunnels," Proc. Highway Research Board 33, 395-404 (1954).

37. L.C. Edie and R.S. Foote, "Traffic flow in tunnels", Proc. Highway Research Board 37, 334-344 (1958).

38. L.C. Edie and R.S. Foote, "Experiments on single-lane flow in tunnels," Proc. symposium on the theory of traffic flow, Detroit 1959, R. Herman, ed., Elsevier, Amsterdam (1961).

39. L.C. Edie and R.S. Foote, "Effect of shock waves on tunnel traffic flow," Proc. Highway Research Board 39, 492-505 (1960).

40. H. Greenberg and A. Daou, "The control of traffic flow to increase the flow," Opns. Res. 8, 524-532 (1960).

41. L.C. Edie, "Car-following and steady-state theory for non-congested traffic," Opns. Res. 9, 8, 524-532 (1960).

41. L.C. Edie, "Car-following and steady-state theory for non-congested traffic," Opns. Res. 9, 66-76 (1961).

42. R.S. Foote, "Single lane traffic flow control," Proc. 2nd international symposium on the theory of road traffic flow, London (1963).

43. M.R. Palmer, The development of traffic congestion, Bureau of Highway Traffic, Yale University, May 1959.

V. MODERATELY DENSE TRAFFIC

1. Introduction. Some of the difficulties associated with mathematical models for moderately dense traffic were discussed in chapter 3. Since most of the practical engineering problems of highway traffic deal with flows of moderate congestion where passings, merging, etc., occur, models of such traffic flow could be very useful. Indeed the future for research in traffic theory would be rather bleak if we had no hope of ever obtaining some reasonable models for such traffic. The mathematical problems that arise in the analysis of even the crudest models, however, are quite formidable. For the low density or high density traffic discussed in the previous chapters, one can question whether or not the theory is accurate enough yet to be of practical value but few people would question the basic framework of the theory. Greater accuracy will come from the addition of certain obvious refinements of the theory which will certainly make the mathematical analysis more tedious but not necessarily more difficult. For the models of moderately dense traffic, on the other hand, the difficulties are of a more basic nature. To model even crudely certain obvious features of such traffic, one is forced to set up equations for which no known methods of solution yet exist.

Since we are not interested in the detailed motion of a specified small collection of cars, we are committed to some type of stochastic model. Furthermore we are interested in the consequences of interactions between cars and this implies a rather complicated statistical dependence between the behaviors of the various cars. If the jth car has a trajectory $x_j(t)$, then we must consider the functions $\{x_j(t)\}$ as a set of statistically dependent random time series. From a practical point of view, this identification is of no obvious help, however, because there are no general mathematical techniques for concrete analysis of such things.

253

In this chapter we consider only the evolution and equilibrium distributions of cars on long homogeneous highways. There are essentially five types of models for moderately dense traffic, all of which are very crude. Each one is designed primarily to avoid rather than overcome mathematical difficulties of more realistic models.

The first types of models are simply attempts to represent the point processes of crossings of cars at a single fixed position, or of the positions of cars at a fixed time, as a stationary point process of rather simple stochastic structure. No attempt is made to deduce these processes from any theory of the dynamic motion of cars, nor are these models used in any way to predict the motion of cars. Indeed, in most cases, there is no hypothetical motion of cars for which the proposed processes would be invariant. Yet these models are useful as a means of approximate representation of experimental results and for the calculation of various quantities which are not expected to be very sensitive to the detailed stochastic structure, for example, the rate at which cars can merge or cross a traffic stream is determined primarily by the marginal probability distribution of headways.

The second theory to be discussed here is due to Carleson [1]. This theory is similar in many respects to the weak interaction theory discussed in chapter III, section 2. Some of the mathematical difficulties discussed in chapter 3 in extending such theories to higher densities are avoided, however, because Carleson considers only the question of how the time the mathematical difficulties discussed in chapter 3 in extending such theories to higher densities are avoided, however, because Carleson considers only the question of how the time average velocity of a car of given desired speed depends upon the density and free speed distribution of the other cars. To evaluate this he then uses a model specifically designed so that this average velocity does not depend upon the detailed stochastic structure of the traffic. Specifically he assumes that the passing delay to a given car depends only upon the number

of cars it passes but is independent of whether these cars are in queues or not. If one takes the delay per passing to be the average delay suffered in the absence of queues, this approximation has the attractive feature that it should never overestimate the consequences of passing because it does not include the additional delays that result when a driver must wait in a queue for other cars to pass some slower car.

The third type of theory is due mainly to Miller [2]. He introduces an artificial stochastic model for the passing discipline designed so as to guarantee that in statistical equilibrium the probability distributions for the positions of queues having any specified range of velocities define a Poisson process. Although this model is very crude it is mathematically tractable because of the special mathematical properties of Poisson traffic. This theory describes the evolution of the system, as well as its equilibrium properties.

The fourth type of theory originates with Tanner [3]. He considers a two-lane road in which all cars in the same lane travel with the same constant velocity and with a preassigned equilibrium probability distribution of spacings. He then introduces a single fast car into one lane and investigates the passing delays for this car. To pass a slower car, the fast driver must find a gap in the opposing traffic stream. Tanner takes a sensible model for the passing mechanism but an artificial model for the probability distribution of slow cars.

The fifth model is due to Prigogine [4]. It is a semi-macroscopic theory in which the mechanism but an artificial model for the probability distribution of slow cars.

The fifth model is due to Prigogine [4]. It is a semi-macroscopic theory in which the velocity distribution of cars is assumed to obey a certain differential equation. The main objection to this theory is that it is based upon postulates regarding the collective behavior of cars which are not obvious consequences of any well-defined notion of how drivers behave individually.

In each of these theories one is groping to obtain some crude results even if one cannot understand the relation of these models to what really happens. The real mathematical problem of how to handle statistically dependent trajectories is avoided in each of these theories. Carleson avoids it by asking only a simple question and by choosing a model in which the answer to the question is not sensitive to the stochastic structure. Miller avoids it by picking a model that is certain to yield only Poisson processes the properties of which are well understood. Tanner avoids it by having all cars but one travel at the same velocity so that the interactions exist only between the reference car and a stream of non-interacting cars. Prigogine avoids it by jumping directly to a macroscopic theory.

2. <u>Point processes</u>. In some applications to merging and crossing problems, it is necessary to have some stochastic description of headways or spacings in a traffic stream of moderate density. It is not necessary that we know the history of the process or that the description gives a completely correct representation of the statistical dependencies. It is important that the models be fairly simple so that subsequent uses of them will not become prohibitively complicated. Here we shall review a few properties of general point processes and some other properties of special processes which are convenient representations of highway traffic. We will discuss these in reference to events in time, namely crossing times of cars at a fixed properties of special processes which are convenient representations of highway traffic. We will discuss these in reference to events in time, namely crossing times of cars at a fixed highway position, although the same considerations also apply to events in space, the positions of cars at a fixed time.

a. <u>Stationary point processes</u>. Some heuristic discussions of point processes were given already in chapter 2 as a background for the description of Poisson processes. A more

rigorous definition of (stationary) point processes requires some care. If we let $--, T_{-2}, T_{-1}, T_o,$ $T_1, T_2, --$ be the random time points of a point process, we might be tempted to define probabilities through a joint distribution function of all the T_j, i.e., a function

$$F(--, t_{-2}, t_{-1}, t_o, t_1, --) = P\{ -- T_{-2} < t_{-2}, T_{-1} < t_{-1}, T_o < t_o, --\}.$$

It is not obvious, however, that such a function has any precise meaning, particularly in view of the heuristic interpretation of a stationary process: any specified T_o, say, is "equally likely to be anywhere" and has, therefore, probability zero of being in any finite interval. There are several other approaches that are more meaningful, however. For example,

(1) Suppose that for every collection of finitely many disjoint intervals, $I_1, I_2, -- I_m$ of the real line, we specify the joint probability distribution of the numbers, $N_1, N_2, -- N_m$ of T_j in $I_1, I_2,$ $--, I_m$ respectively. (Note that we do not try to identify <u>which</u> of the T_j are in I_1, I_2, etc.). This is a generalization of the definition of a Poisson process as one for which the N_j are all independent and Poisson distributed. Since these probabilities must be specified for every choice of m and every choice of the I_j, there is an enormous range of possibilities. These probabilities are restricted, however. If, for example, we choose m = 2, I_1 the interval (t, t $+\tau_1$) and I_2 the interval $(t + \tau_1 , t + \tau_1 + \tau_2)$, then choose m = 1 with I_1' the interval (t, t + τ_1 $+ \tau_2$), the random variable N_1' for I_1' must be the sum of the random variables N_1+N_2 of I_1 and I_2, i.e.,
$+ \tau_2$), the random variable N_1 for I_1 must be the sum of the random variables N_1+N_2 of I_1 and I_2, i.e.,

$$P\{N_1' = n_1'\} = \sum_{n_1 + n_2 = n_1'} P\{N_1 = n_1 , N_2 = n_2\} .$$

A stationary point process is one for which these probabilities are all invariant to a translation of the time axis. If I_1 is the interval (t_1, t_1'), I_2 the interval (t_2, t_2'), etc., then the joint

probability distribution of N_1, N_2, etc. is the same as the joint probability distribution for the numbers N_1', N_2'-- in the intervals $(t_1+\tau, t_1'+\tau)$, $(t_2+\tau, t_2'+\tau)$ etc. for every τ.

(2)   In view of this restriction that the number of events in the union of two sets I_j must be the sum of the events N_j in the sets, it may be more convenient to define a random function

$$N(t) = \begin{cases} + \textit{ number of events in } (0 , t) & \textit{for } t > 0 \\ - \textit{ number of events in } (+ t , 0) & \textit{for } t < 0 . \end{cases}$$

The probability structure of N(t) is defined if for arbitrary times $t_1 < t_2 < -- < t_m$ we give the joint probability distribution of the random variables $N(t_1)$, $N(t_2)$,.., $N(t_m)$. These $N(t_j)$ are restricted only by the condition $N(t_1) \leq N(t_2) \leq .. \leq N(t_m)$. If we know the probabilities as defined in (1), we can evaluate the probabilities for the $N(t_j)$ or vice versa under the rule that the number N_1 in an interval (t_1, t_2) is $N(t_2)-N(t_1)$.

The process N(t) is not itself stationary. It is called a process of stationary increments if the distributions of the random variables $N(t_1+\tau) - N(\tau)$, $N(t_2+\tau) - N(\tau)$, ... are independent of τ. Processes with stationary increments were encountered before. The position x(t) of a car undergoing a Brownian random motion, with the velocity v(t) defining a stationary process, is of this type.

(3)   Suppose we pick some reference point, at time τ. Let

$$T_o(\tau) \equiv \textit{smallest } t_j \textit{ with } T_j \geq \tau$$

(3)   Suppose we pick some reference point, at time τ. Let

$$T_o(\tau) \equiv \textit{smallest } t_j \textit{ with } T_j \geq \tau$$

and --, $T_{-2}(\tau)$, $T_{-1}(\tau)$, $T_o(\tau)$, $T_1(\tau)$ --be a renumbering of the T_j so that $T_j(\tau) \leq T_{j+1}(\tau)$. Now define the process by giving, for every n, the joint probability distribution of the $T_j(\tau)$, $-n \leq J \leq n$, or equivalently the joint distribution of $T_o(\tau)$ and the variables $T_j(\tau)-T_{j-1}(\tau)$. The process is stationary if these distributions are independent of τ.

(4)   From interpretation (1), we see that if $I_1, --, I_m$ is a partition of the interval (t, t+1) into m equal parts, then

$$E\{N_1 + N_2 + ... + N_m\} = \sum_{j=1}^{m} E\{N_j\} = m \ E\{N_1\}$$

because stationarity implies $E\{N_j\}=E\{N_1\}$ for all j. If we denote by q the mean number in the unit interval, the left-hand side above, then the mean number in an arbitrarily small interval of length dt= 1/m is q dt. If, however, events occur only one at a time and only finitely many occur in any finite time, then the number in an arbitrarily small interval of length dt must be either 0 or 1 with probabilities 1-qdt and qdt, respectively. Thus for a stationary process of single events there is a probability qdt of an event between t and t+dt, q being independent of t. If an event occurs in this interval, we label it as event 0 with T_o=t and proceed now to specify the joint probabilities of all finite collections of the times T_j-T_{j-1}. The stationarity is specified through the fact that these distributions do not depend upon t.

The first two ways of describing a point process are perhaps most directly suited to the study of traffic flow past a point, since it gives directly the number crossing in any time intervals. The last two are most convenient for describing headways and are also mathematically usually the most convenient because in some sense they involve the least amount of redundant information (particularly (4)).
mathematically usually the most convenient because in some sense they involve the least amount of redundant information (particularly (4)).

One of the peculiar features of general stationary processes is that the headways $T_j(\tau)$-$T_{j-1}(\tau)$ as specified in (3) are not the same as the headways T_j-T_{j-1} as described in (4). In (3) we started from an arbitrary time τ and called the next event T_o whereas in (4) we chose an arbitrary time t but identified a T_o only if an event occurs in a small interval (t, t+dt). In the

259

former case, the distributions of the $T_j(\tau)-T_{j-1}(\tau)$ depend upon the time interval $T_o-\tau$ and furthermore they are not necessarily independent of j. In the latter case we may not find an event in (t,t+dt) in which case we repeat the experiment until we do. When we find an interval containing an event, the times T_j-T_{j-1} will be identically distributed and independent of t. The distinction between these two will be described briefly here. For a more thorough investigation see McFadden and Weissblum [5,6].

If we adopt the fourth way of defining the point process, it can be specified through the value of q and, if

$$X_j = T_j-T_{j-1},$$

of

$$F(x) = P\{X_j < x\}$$

$$F_k(x,y) = P\{X_j < x,\ X_{j+k} < y\} \quad k = 1,2, --$$

$$F_{k,\ell}(x,y,z) = P\{X_j < x, X_{j+k} < y,\ X_{j+k+\ell} < z\}\ ,\ k,\ell = 1,2,--,\ etc.$$

These distributions are all independent of j because of the stationarity. Each distribution function of n variables implies those of n-1, n-2, etc., for example,

$$F_k(x,y) = F_{k,\ell}(x,y,\infty),$$

thus some of the above description is redundant. But we cannot define the distribution function for infinitely many variables so that it would imply the distributions of any finite thus some of the above description is redundant. But we cannot define the distribution function for infinitely many variables so that it would imply the distributions of any finite number.

Suppose now we derive the distribution functions of the $T_j(\tau)$ of the third description in terms of the above functions of the fourth representation. If we start at some arbitrary time τ, there is a probability qdt_o that an event will occur between t_o and t_o+dt_o (independent of τ).

If this happens we identify this event as the event T_o. There is a probability $1-F(t_o-\tau)$ that T_o- $T_{-1} > t_o-\tau$, so that the event at t_o is the next one after τ, and therefore a $T_o(\tau)$. The distribution function of the time to the first arrival $T_o(\tau)-\tau$ is therefore the product of these two probabilities, the probability that $T_o(\tau) = T_o$ and T_o is between t_o and t_o+dt_o, integrated over t_o, i.e.,

$$P\{T_o(\tau) - \tau < z\} = q \int_{\tau}^{z+\tau} [1 - F(t_o - \tau)] \, dt_o$$

(2.1)

$$= q \int_o^z [1 - F(z')] \, dz' \ .$$

Similarly the distribution function of the time since the last arrival is

$$P\{\tau - T_{-1}(\tau) < z\} = q \int_o^z [1 - F(z')] dz' \ .$$

(2.1a)

If this is to be a proper distribution function, it must go to 1 for $z \rightarrow \infty$, i.e.,

$$q \int_o^\infty [1 - F(z')] dz' = 1 \ .$$

We can integrate this by parts to obtain

We can integrate this by parts to obtain

$$q\{ z'[1 - F(z')]_o^\infty + \int_o^\infty z' \, dF(z') \} = 1$$

but $1-F(z') \rightarrow 0$ for $z' \rightarrow \infty$ and if $q \neq 0$, it is also true that $z'[1-F(z')] \rightarrow 0$ for $z' \rightarrow \infty$. Therefore, the end terms both vanish and

261

$$q \int_o^\infty [1 - F(z')] \, dz' = q \int_o^\infty z' \, dF(z') = 1 \ .$$

The last integral we recognize as the mean time between events. Thus

$$q \, E\{T_j - T_{j\text{-}1}\} = 1.$$

This relation is heuristically obvious. Over a long length of time L, the mean number of events in L is qL. But L must also be the number of events times the mean time between events (except possibly for some fractions of interevent times at the ends). Thus

$$q \, L \, E\{T_j - T_{j\text{-}1}\} = L.$$

Consider next the distribution of the time $T_o(\tau) - T_{\text{-}1}(\tau)$, the time between the events containing τ. The probability that T_o is between t_o and $t_o + dt_o$ and $T_{\text{-}1}$ lies between $t_{\text{-}1}$ and $t_{\text{-}1} + dt_{\text{-}1}$ is

$$q \, dt_o \, dF(t_o - t_{\text{-}1}).$$

The probability that $T_o(\tau) - T_{\text{-}1}(\tau)$ lies between z and z+dz is obtained by fixing $t_o\text{-}t_{\text{-}1} = z$ and integrating overall values of t_o with $t_{\text{-}1} = t_o\text{-}z < \tau$ and $\tau < t_o$. It is

$$dF(z) \, q \int_\tau^{\tau+z} dt_o = qz \, dF(z) \ .$$

The distribution function of $T_o(\tau) - T_{\text{-}1}(\tau)$ is, therefore,

The distribution function of $T_o(\tau) - T_{\text{-}1}(\tau)$ is, therefore,

$$F^*(z) = q \int_o^z z' \, dF(z') \ . \tag{2.2}$$

The first important difference between the $T_j - T_{j\text{-}1}$ and $T_j(\tau) - T_{j\text{-}1}(\tau)$ is that $T_o(\tau) - T_{\text{-}1}(\tau)$ has a probability density (if it exists) that is weighted by an additional factor of z

compared with that of T_o-T_{-1}. Another interpretation of this is that if we were to select a point at random on the infinite time axis, we are more likely to land in a long headway than a short one. In fact the probability of landing in a headway is proportional to its length.

We have seen that the marginal distribution of $T_o(\tau)$ - $T_{-1}(\tau)$ is not the same as that of T_o - T_{-1}. If the T_j - T_{j-1} are statistically independent, then the distribution of $T_j(\tau)$ - $T_{j-1}(\tau)$, $j\neq0$, are also independent of $T_o(\tau)$ - $T_{-1}(\tau)$. Furthermore, $T_j(\tau)$ - $T_{j-1}(\tau)$, $j\neq0$, have the same distributions as T_j - T_{j-1}. The headway one lands in at time τ is not typical, but all others are. If the T_j - T_{j-1} are statistically dependent, however, the $T_j(\tau)$ - $T_{j-1}(\tau)$ will in general not have the same distribution as the T_j-T_{j-1} for <u>any</u> j. The joint distributions of the $T_j(\tau)$ - $T_{j-1}(\tau)$ can be evaluated from those of the T_j- T_{j-1}, however. The procedure can be illustrated by the case of $T_1(\tau)$ - $T_o(\tau)$ and $T_o(\tau)$ - $T_{-1}(\tau)$.

The probability that $T_o(\tau)$ is between t_o and $t_o + dt_o$, $T_{-1}(\tau)$ is between t_{-1} and $t_{-1} + dt_{-1}$, is

$$q \, dt_o \, dF_j(t_o\text{-}T_{-j}, \, t_1 - t_o)$$

$$= q \, f_j(t_o - t_{-j}, \, t_1 - t_o) \, dt_o dt_{-1} \, dt_1$$

if f (.,.) represents the two dimensional density of X_j, X_{j+1}. The joint density of $T_o(\tau)$-$T_{-1}(\tau)$ and $T_1(\tau)$- $T_o(\tau)$ at x,y is obtained by integrating this overall t_o with t_o-t_1 and t_1-t_o fixed at x and y respectively. Thus
and $T_1(\tau)$- $T_o(\tau)$ at x,y is obtained by integrating this overall t_o with t_o-t_1 and t_1-t_o fixed at x and y respectively. Thus

$$f_1^*(x \, , \, y) = \int_{\tau}^{\tau+x} dt_o \, q \, f_1(x \, , \, y) = q \, x f_1 \, (x \, , \, y) \, .$$

The joint distribution function is

$$F_1^*(x, y) = \int_0^x dx' \int_0^y dy' \; q \; x' \; f_1(x', y') \; .$$

If we let $y \to \infty$ we obtain the marginal distribution function of $T_0(\tau) - T_{-1}(\tau)$

$$F^*(x) = \int_0^x dx' \; qx' \int_0^\infty f_1(x', y') = \int_0^x dx' \; qx'f(x') = \int_0^x qx'dF(x')$$

as derived earlier. If the $T_j - T_{j-1}$ are independent

$$f_1(x', y') = f(x')f(y')$$

and

$$F_1^*(x, y) = \left[\int_0^x dx' \; qx' \; f(x')\right]\left[\int_0^y dy' \; f(y')\right]$$

$$= F^*(x) \; F(y) \; .$$

Thus as described before $T_1(\tau) - T_0(\tau)$ is statistically independent of $T_0(\tau) - T_{-1}(\tau)$ although they do not have the same marginal distribution functions.

If, we let $x \to \infty$ we obtain the marginal distribution function of $T_1(\tau) - T_0(\tau)$

$$F_1^*(\infty, y) = \int_0^\infty dx' \int_0^y dy' \; qx' \; f_1(x', y')$$

$$F_1^*(\infty, y) = \int_0^\infty dx' \int_0^\cdot dy' \; qx' \; f_1(x', y')$$

which, in general, is not the same as $F(y)$.

One can construct all sorts of hypothetical processes with peculiar properties. For example, we could have a stationary process in which $T_j - T_{j-1}$ were non-random but alternate

between the values 1 and 2, i.e., if $T_j - T_{j-1} = 1$ then $T_{j+1} - T_j = 2$ and vice versa. For this process

$$P\{T_j - T_{j-1} = 1\} = 1/2 \ , \ P\{T_j - T_{j-1} = 2\} = 1/2$$

$$E\{T_j - T_{j-1}\} = 3/2 \quad , \ q = 2/3$$

The joint distributions are such that once $T_o - T_{-1}$ is specified the other $T_j - T_{j-1}$ are also specified. All probabilities are either 1/2 or 0. But $P\{T_o(\tau)-T_{-1}(\tau) = 1\} = 1/3$, $P\{T_o(\tau) - T_{-1}(\tau) = 2\} = 2/3$. Also the probabilities for any choices of $T_j(\tau) - T_{j-1}(\tau)$ will have only the values 1/3, 2/3, or 0. The $T_j(\tau) - T_{j-1}(\tau)$ and $T_j - T_{j-1}$ do not have the same marginal distributions for any j, not even for j→∞.

Although the stochastic properties of headways in highway traffic could be quite complicated, we expect that they should always have one property not shared by the above example, namely X_j and X_k should be statistically independent (or nearly so) for $|j-k|$ sufficiently large. This also implies that for sufficiently large j the joint distribution of the $T_j(\tau) - T_{j-1}(\tau)$ should become nearly identical with those of the $T_j - T_{j-1}$. In this regard we are still thinking of the traffic stream ideally as a stationary point process, which is valid only for finite times of the order of an hour at most. These statistical correlations should, however, decay in a much shorter time, probably less than one minute.

Although the concepts of joint distribution functions are mathematically well defined decay in a much shorter time, probably less than one minute.

Although the concepts of joint distribution functions are mathematically well defined and can, in principle, be evaluated experimentally, there are obvious limitations to what one can do in practice. Most experimental work has been concentrated on the evaluation of the marginal distributions of single headways, $F(x)$. As a practical matter again, these headway

distributions are usually inferred from measurements on successive headways starting from a randomly chosen time τ as in the third scheme above.

The "thought experiment" associated with the fourth scheme is to choose a short time interval (t, t + dt), see if there is an event in dt. If there is, one can then observe an arbitrary number of headways $T_j\text{-}T_{j\text{-}1}$, j = 1,2,-- limited only by the requirement that the process be stationary, all of which have the same distribution. If one fails to observe an event in dt, however (which will happen most of the time), then one must repeat the experiment under what are supposed to be identical and statistically independent conditions, a concept that is experimentally not too well defined. The difficulty with this procedure is that one cannot afford to give up the experimental trial every time one fails to observe an event during the prescribed interval of time dt.

If one can devise an experimental procedure based upon the third procedure, one is at least guaranteed of obtaining some data within a finite time. The difficulty, however, is that the first few headways are not typical. One probably would not ordinarily observe or try to use the time $T_o(\tau)\text{-}\tau$ in the estimation of F(x). Suppose we were to start with the headway $T_j(\tau)\text{-} T_{j\text{-}1}(\tau)$, j$\geq$1 and consider the quantity,

$$\tilde{F}(x) = \frac{1}{n} \times number\ of\ T_k(\tau)\ -\ T_{k\text{-}1}(\tau)\ less\ than\ x,\ k=j\ ,\ j+1\ ,\ ...\ j+n-1$$

$$\tilde{F}(x) = \frac{1}{n} \times number\ of\ T_k(\tau)\ -\ T_{k\text{-}1}(\tau)\ less\ than\ x,\ k=j\ ,\ j+1\ ,\ ...\ j+n-1$$

$$= \frac{1}{n} \sum_{k=j}^{j+n-1} Z_k(x)$$

$$
\text{with} \quad Z_k(x) = \begin{cases} 1 & \text{if } T_k(\tau) - T_{k-1}(\tau) < x \\ 0 & \text{if } T_k(\tau) - T_{k-1}(\tau) > x \end{cases}
$$

for the n headways starting with $T_j(\tau) - T_{j-1}(\tau)$. This $\tilde{F}(x)$ or the mean of several observations of $\tilde{F}(x)$ is what one would typically use as an estimate of $F(x)$. The accuracy of this estimate, however, depends upon the degree of statistical dependence.

If the $T_k(\tau) - T_{k-1}(\tau)$, $k > 1$ are statistically independent, they will all have the same distribution function $F(x)$. The quantity $n\tilde{F}(x)$ for fixed x is an integer valued random variable with a binomial distribution corresponding to a probability of "success or failure," $Z_k(x) = 1$ or 0, of $F(x)$ and $1-F(x)$, respectively. For $n \to \infty$, $\tilde{F}(x)$ will with probability 1 converge to $F(x)$ (law of large numbers) and furthermore $\tilde{F}(x)$ will be asymptotically normal (central limit theorem) with mean $F(x)$ and variance

$$
Var \, \tilde{F}(x) = \frac{2}{n} F(x)[1 - F(x)] .
$$

Most textbooks on mathematical statistics also describe more detailed properties of the stochastic properties of the dependence of $\tilde{F}(x)$ upon x. This is equivalent to a classic problem of estimation of a distribution function from independent repeated trials.

If the $T_k(\tau) - T_{k-1}(\tau)$, $k > 1$ are statistically dependent, one would probably still use the same estimate $\tilde{F}(x)$ but the estimate will be less accurate for a given finite value of n for two

If the $T_k(\tau) - T_{k-1}(\tau)$, $k > 1$ are statistically dependent, one would probably still use the same estimate $\tilde{F}(x)$ but the estimate will be less accurate for a given finite value of n, for two reasons. First of all, $\tilde{F}(x)$ might be a biased estimate of $F(x)$, i.e.,

$$
E\{\tilde{F}(x)\} \neq F(x)
$$

because the first few $T_k(\tau) - T_{k-1}(\tau)$ at least will not have a distribution function $F(x)$. Since, before one does any experiments, one does not know the joint distributions of the $T_k(\tau) -$

267

$T_{k-1}(\tau)$, one cannot very well correct for any bias. The safest thing to do is to disregard the first few headways, however many one thinks might be correlated with the biased selection of $T_o(\tau) - T_{-1}(\tau)$.

In addition, the variance of $\tilde{F}(x)$ is likely to be larger with statistical dependence. Suppose that $T_k(\tau) - T_{k-1}(\tau)$ was statistically independent of $T_\ell(\tau) - T_{\ell-1}(\tau)$ for $(k-\ell) > M$ for some number M. We could extract from the headways only every Mth observation. These $T_k(\tau) - T_{k-1}(\tau)$ would then be independent but from an experiment of given length n we would use only about n/M observations and therefore

$$Var \; \tilde{F}(x) \; = \; \frac{2M}{n} \; F(x)[1 \, - \, F(x)] \; .$$

Unless one knows something about the covariances of the $T_k(\tau) - T_{k-1}(\tau)$, one cannot evaluate Var $\tilde{F}(x)$ for the full number of n observations. One can show, however, that for n→∞, $\tilde{F}(x)$ → F(x) with probability 1, and that $[\tilde{F}(x) - F(x)]/n^{1/2}$ will be normally distributed with a variance of order $M^{1/2}$.

To determine experimentally an estimate of $F_1(x,y)$ one can proceed similarly. Let

$$\tilde{F}_1(x,y) \; = \; \frac{1}{n} \; \times \; number \; of \; k \; for \; which \; T_k(\tau) \, - \, T_{k-1}(\tau) \, < \, x \; and \; T_{k+1}(\tau) \, - \, T_k(\tau) \, < \, y \; .$$

Some references to experimental results were given in chapter III, reference III [13-
n

Some references to experimental results were given in chapter III, reference III [13-15]. Most of this literature deals with the function F(x) and various empirical formulas for its representation. One of the practical difficulties in evaluation of even F(x) is that F(x) depends upon the flow and upon such things as the geometry of the highway. It is not easy to decide

from a given collection of data taken with a slowly varying q(t), when a temporary rise in the apparent flow is just a fluctuation or part of a systematic pattern.

b. Renewal processes. A stationary point process as described above with statistically independent $T_j - T_{j-1}$ is, in the probability literature, called a renewal process. The name derives from its application to the theory of replacement or renewal of repair parts. For example, light bulbs may have statistically independent lifetimes. If a light bulb is instantaneously replaced by a new one every time it burns out, the times of renewal define a renewal process. There is a vast literature on renewal theory and its applications, including a small book by Cox [7]. Most of this theory deals with the evaluation of probability distributions for the number of events in some given time interval from the distributions of times between events. Although this has obvious relevance to highway traffic theory, we will not try to review here the theory of renewal processes.

It is difficult to trace the origin of the model that highway traffic be represented as a renewal process. Many people who measured or used headway distributions, and observed that they were not exponential, assumed that the headways were statistically independent without realizing that it could be otherwise. The literature on queueing theory also contains many papers that claim solutions for certain problems for an "arbitrary" distribution of arrivals or service times when they really mean an arbitrary renewal process of arrivals. many papers that claim solutions for certain problems for an "arbitrary" distribution of arrivals or service times when they really mean an arbitrary renewal process of arrivals.

There are no experiments that obviously point to the unsuitability of a renewal process as the process of arrivals. There is even some suggestion that it might be a reasonable approximation. For low density traffic as described in chapter III, section 3, we saw no

269

inconsistency with a renewal hypothesis although the traffic stream was described as being a superposition of Poisson processes for singlets and for pairs.

There are some experimental studies which describe, in a rather qualitative way, the tendency of traffic to form platoons [8,9]. The long headways have approximately an exponential distribution, the short headways some distribution about the mean spacing for "car-following." If we choose some, more or less arbitrary, time and consider all headways less than this to be platooned cars, we can then discuss the distribution of the number of consecutive platooned cars. If traffic is a renewal process, the probability of a headway being identified as a platoon headway, is statistically independent of whether or not the previous headway was short or long, or how many previous cars were considered to be in platoons. If p is the probability that a car will be in a platoon, then the number of cars in a platoon has a geometric distribution

$$P\{number\ in\ platoon\ =\ n\}\ =\ (1\ -\ p)pn$$

There are very few experiments on distribution of platoon lengths and most of these are very crude. Although one can probably find other distributions for the lengths that fit the experiments better than the geometric distribution, the geometric distribution has at least the correct qualitative shape and is not obviously incorrect.

As theoretical models become more refined and experimental data more plentiful, the correct qualitative shape and is not obviously incorrect.

As theoretical models become more refined and experimental data more plentiful, the assumption of a renewal process will undoubtedly be replaced by something better. It seems clearly inaccurate in some respects. For example, one would expect that a long headway is frequently followed (and caused by) a slow car, which in turn is frequently followed by a

platoon. Thus long headways are likely to be followed by a succession of small ones, contrary to the statistical independence postulate.

c. Random queues, block-gap processes. Miller [8] proposed a model of traffic as a compound Poisson process, i.e., a point process each point of which represents a platoon rather than a single car, the number of cars in the platoon being a random variable of unspecified distribution. A more realistic model can be obtained from this by assignment of a physical length to the queue and perhaps some internal structure of headways (but not necessarily a renewal process). Experiments do indicate that the marginal distribution of long headways is nearly exponential, so the assumption of Poisson process for the queues seems reasonable. Although it might be mathematically inconvenient to assume that platoon lengths are statistically independent of each other and of the long headways, this may not be physically desirable if very long headways are likely to be followed by long queues. It might be better to represent traffic as a Poisson process of platoons with the lengths of the platoons statistically dependent upon the headway proceeding the queue. Such a process is a Markov process. If one starts with a long headway one selects its value from an exponential distribution, then depending upon what was selected one next selects a platoon length from some other distribution. The following headway is then selected independent of the past again from the exponential distribution, etc. The starting points of successive platoons would in this some other distribution. The following headway is then selected independent of the past again from the exponential distribution, etc. The starting points of successive platoons would in this case define a renewal process, but the internal structure between these points would involve statistical dependencies.

Models of traffic of essentially the same types as the above have also evolved in a different context. One of the problems for which a knowledge of the stochastic properties of

271

the traffic stream is relevant, is the calculation of merging or crossing rates. In the analysis of merging or crossing, it is customary to think of traffic as an alternating sequence of "blocks and gaps." Blocks are periods of time during which there will be no headway larger than some preassigned value, so chosen as to guarantee that no crossing is possible within the block. Gaps are the remaining times during which a crossing may be possible. The headways used to define a blocked period may depend upon the crossing properties of drivers, the geometry of an intersection, etc., but otherwise there is no difference between the mathematical definition of a block and the definition of a platoon. In either case the maximum headway used to define a block or platoon is chosen in a rather arbitrary way.

Many papers have been written on distribution of block lengths, particularly as this relates to queueing for gaps. Calculation of block lengths for Poisson streams were made as early as 1936 by Garwood [1] and later by Tanner [11], and Raff [12]. These were extended to renewal processes by Mayne [13], Weiss and Maradudin [14] and Oliver [15]. More recently, however, the trend has been to postulate some arbitrary distribution for the block lengths to calculate queue lengths, etc. for crossing [16, 17].

3. Carleson theory. In the theory to be discussed in this section we imagine again a situation analogous to the homogeneous highway fed by an ideal parking lot as discussed in chapters II

and III. The cars in the lot have a probability distribution function for desired speeds $F_v(v)$ and they travel at their desired speeds except possibly when trying to pass other cars.

As in chapter III we wish to evaluate the time average velocity u(v) of a car with desired speed v. The present scheme parallels that used in chapter III but differs in the following respects.

1. The spatial distribution of velocities $f_s(v)$ in III (2.2) is taken to be the one described by III (2.5) corresponding to an assumption that the distribution of velocities of the cars selected from the lot does not depend upon the flow q. The average number of cars per unit length of highway with velocity between v and v + dv is thus taken as

$$kf_s(v)dv = \frac{qdF_v(v)}{u(v)} \qquad (3.1)$$

in which $F_v(v)$ is specified and independent of q.

2. The average rate at which cars of desired speed v overtake those of desired speed between v′ and v′ + dv′ is now chosen to be

$$[u(v) - u(v')] \, k \, f_s(v')dv' = q \, [u(v) - u(v')] \, d \, F_v(v')/u(v') \qquad (3.2)$$

as compared with III (2.1) where u(v) - u(v′) was replaced by v-v′.

3. We again define a function d(v,v′) as the average loss in the distance traveled that results when a car of desired speed v must pass one of desired speed v′. The function d(v,v′) is assumed to be known.

The time average distance traveled per unit time by a car with desired speed v is is assumed to be known.

The time average distance traveled per unit time by a car with desired speed v is given by an equation similar to III (2.2). It is the distance v less the average loss in distance traveled per unit time due to all passings evaluated from the above expressions (3.1) and (3.2), i.e.,

$$u(v) = v - q \int_o^v dF_v(v') \, d(v,v') \, [u(v) - u(v')]/u(v') \,. \tag{3.3}$$

To derive these relations one must make certain reasonable postulates about the stochastic properties of traffic. The implications of (3.1) have been described previously. For (3.2) we assume first that a car of desired speed v never passes one of desired speed $v' > v$ and that it passes one of speed $v' < v$ at most once, even if some queueing may occur. Over a long period of time T, a car of desired speed v will travel an average distance u(v)T. The actual distance traveled is a random variable but it is reasonable to assume that for sufficiently large T, the fluctuations in the distance traveled are small compared with the mean distance. Similarly $[u(v) - u(v')]$ is the average distance a car of desired speed v gains on one of desired speed v'. The average number of cars with velocity in dv' that must be passed is this latter distance times the average spatial density of cars with velocity in dv'. We are saying here that the average of a product, the product of a distance and a number of cars, is equal to the product of the averages. This can be justified, however, on the grounds that the distance in question is, for sufficiently large T, almost always very close to the average.

By an appropriate interpretation of $d(v,v')$, one could always guarantee that (3.3) is correct. The difficulty lies in the assumption that $d(v,v')$ is known. If there is queueing and other complicated interactions between cars, one must take for d(v v') the average of the loss

correct. The difficulty lies in the assumption that $d(v,v')$ is known. If there is queueing and other complicated interactions between cars, one must take for $d(v,v')$ the average of the loss in distance traveled averaged over all possible circumstances wherein a car of desired speed v passes one of desired speed v' including queueing times. If this is done, $d(v,v')$ becomes also a complicated functional of $F_v(\cdot)$ and a function of q. To avoid this problem we postulate

that there are no queueing times and that $d(v,v')$ is the loss that would exist if cars interact only two at a time.

With this interpretation of the loss $d(v,v')$, (3.3) and III (2.2) agree to first order in the density k or the flow q for small k or q. There is no reason why (3.3) should give quantitatively realistic values for $u(v)$ for higher flows, but the equation does describe the behavior for a hypothetical situation that has some qualitative similarities to real traffic (if we were to extrapolate III (2.2) to large k we would obtain negative velocities).

Equation (3.3), considered as an equation to be solved for $u(v)$ in terms of known functions $F_v(v)$ and $d(v,v')$ is, in its present form, a non-linear integral equation. If, however, we let

$$\phi(v) = \frac{1}{u(v)} \tag{3.4}$$

be the unknown we obtain for $\phi(v)$ the linear integral equation

$$\phi(v) = \frac{1}{v} + q \int_o^v d\, F_v(v') \, \frac{d(v,v')}{v} \, [\phi(v') - \phi(v)] \tag{3.5}$$

or

$$1 + q \int^v d\, F_v(v') \, d(v,v') \, \phi(v')$$

$$\phi(v) = \frac{1 + q \int_o^v d\, F_v(v') \, d(v,v') \, \phi(v')}{v + q \int_o^v dF_v(v') \, d(v,v')} \,. \tag{3.6}$$

Equation (3.6) is a Volterra type integral equation. If velocities are bounded away from 0, $v > v_m$ for some minimum velocity v_m, the solution of (3.6) is known to exist for any

275

$q > 0$ and can be constructed by successive substitution of the value of $\phi(v)$ given by the left side of (3.6) into the integral on the right. Even though the solution exists for some particular $d(v,v')$, it still is necessary to verify that the solution $u(v)$ is monotone increasing or equivalent $\phi(v)$ is monotone decreasing. If $d(v,v')$ were chosen to be a rapidly increasing function of v, there is the possibility that fast cars will be delayed so much as to have a slower time average velocity $u(v)$ than some cars of lower desired speed. If such is the case, however, equation (3.6) is no longer valid because it gives rise to an artificial negative rate of passing in (3.2) and a corresponding gain in speed.

There are a number of special cases of (3.6) that lead to fairly simple solutions. The method of successive substitutions as applied to (3.5) leads to an expansion in powers of q which for small q agrees with III (2.2) to order q but the complete expansion is quite awkward to use. If $F_V(v)$ is a discrete distribution let v have the value v_i with probability p_i, and order the v_i so that $v_1 < v_2 < --$. Then (3.6) becomes

$$\phi(v_j) = \frac{1 + q \sum_{i=1}^{j-1} p_i \, d(v_j,v_i) \, \phi(v_i)}{v_j + q \sum_{i=1}^{j-1} p_i \, d(v_j,v_i)} \tag{3.7}$$

$$\phi(v_1) = 1/v_1$$

This gives $\phi(v_j)$ in terms of the $\phi(v_i)$ for $i < j$. The solution is defined by successive

This gives $\phi(v_j)$ in terms of the $\phi(v_i)$ for $i < j$. The solution is defined by successive evaluation of $\phi(v_2)$, $\phi(v_3)$, etc. Thus

$$\phi(v_2) = \frac{1 + q p_1 \, d(v_2,v_1)/v_1}{v_2 + q \, p_1 \, d(v_2,v_1)} = \frac{1}{u(v_2)} \tag{3.8}$$

is the complete solution if there are only two velocities v_1 and v_2. Considered as a

function of q, other quantities being fixed, we see that $u(v_2)$ is a monotone decreasing function of q having the free speed value v_2 for $q = 0$ and the value v_1 for $q \to \infty$. There is nothing in this model that limits the density of cars or the flow so that the limit $q \to \infty$ is meaningful in this model even if it is physically unrealistic.

For three velocities (3.7) gives

$$\frac{1}{u(v_3)} = \phi(v_3) = \frac{1 + q \, p_2 \, d \, (v_3, v_2) \, \phi(v_2) + q \, p_1 \, d(v_3, v_1)/v_1}{v_3 + q \, p_2 \, d(v_3, v_2) + q \, p_1 \, d(v_3, v_1)} \tag{3.9}$$

in which the $\phi(v_2)$ is given by (3.8). The only point we wish to make here is that if q is large enough (3.9) gives

$$\phi(v_3) \sim \frac{p_2 \, d(v_3, v_2) \, \phi(v_2) + p_1 \, d(v_3, v_1)/v_1}{p_2 \, d(v_3, v_2) + p_1 \, d(v_3, v_1)} . \tag{3.10}$$

For $q \to \infty$, we have from (3.8) that $\phi(v_2) \to 1/v$, and therefore $\phi(v_3) \to 1/v$, also, but if $p_1 d \, (v_2, v_1)$ in (3.8) is small compared to the $p_2 d(v_3, v_2)$ and $p_1 d(v_3, v_1)$ in (3.9), i.e., the faster cars suffer much larger delays, we could find an intermediate range of q where (3.10) is valid but $\phi(v_2)$ is still nearly equal to $1/v_2$. In this case $\phi(v_3)$ lies between $1/v_2$ and $1/v_1$ and is, therefore, larger than $\phi(v_2) \sim 1/v_2$. This gives rise to the situation described above where the fast cars travel slower than the slow cars and the theory is not valid even though the equations have well-defined solutions.

travel slower than the slow cars and the theory is not valid even though the equations have well-defined solutions.

The behavior of $\phi(v)$ for a continuous velocity distribution is qualitatively similar to that of a discrete distribution. For the minimum velocity v_m we have that $\phi(v) = 1/v_m$ so if the monotone properties of $\phi(v)$ are preserved $\phi(v) \leq 1/v_m$ for all $v \geq v_m$. This means that, in (3.5), the term $\phi(v)$ and $1/v$ are bounded for all q and consequently, for $q \to \infty$, the integral in (3.5)

must go to zero. But this integral, having a non-negative integrand, can vanish only if $\phi(v')$ → $\phi(v)$ for all v and v′ where $dF_v(v')d(v,v') > 0$. This, in turn, implies that $\phi(v) \to 1/v_m$ for q→∞ provided $F_v(v)>0$ for any $v > 1/v_m$ (there are cars with velocity arbitrarily close to v_m).

There are also some special forms for $F_v(v)$ and $d(v,v')$ with continuous distributions that lead to relatively simple solutions for (3.5) or (3.6). Carleson considered in some detail the special case in which $d(v,v') = d\ v/v'$ for some constant d. The more general case in which $d(v,v')$ has a product form

$$d(v,v') = \alpha(v)\beta(v') \tag{3.11}$$

for some more or less arbitrary functions α and β is also convenient.

A formal solution of (3.5) can be obtained if we write (3.5) in the form

$$\frac{v\phi(v)}{\alpha(v)} = \frac{1}{\alpha(v)} + q \int_0^v d\ F_v(v')\beta(v')[\phi(v') - \phi(v)]$$

and then take the derivative (denoted by a prime) of both sides with respect to v. This gives

$$[1 - \frac{v\alpha'(v)}{\alpha(v)}]\ \phi(v) + [v + q \int_0^v d\ F_v(v')\ d(v,v')]\ \phi'(v) = -\frac{\alpha'(v)}{\alpha(v)} \tag{3.12}$$

which is a linear first order differential equation for $\phi(v)$. If all velocities are larger than v_m, we can everywhere replace lower limits of integration v′=0 by v′=v_m. For v=v_m we know that which is a linear first order differential equation for $\phi(v)$. If all velocities are larger than v_m, we can everywhere replace lower limits of integration v′=0 by v′=v_m. For v=v_m we know that $\phi(v_m)=1/v_m$ which gives us a boundary condition for the differential equation.

Equation (3.12) has an integrating factor

278

$$W(v) = \exp\left[\int_{v_m}^{v} \frac{dz\ [1 - z\ \alpha'(z)/\alpha(z)]}{z + q \int_{v_m}^{z} d\ F_v(x)\ d(z,x)}\right]$$ (3.13)

and a solution

$$\phi(v) = \frac{1}{v_m W(v)} - \frac{1}{W(v)} \int_{v_m}^{v} \frac{dz\ \alpha'(z)\ W(z)}{\alpha(z)\left[z + q \int_{v_m}^{z} d\ F_v(x)\ \alpha(z,x)\right]} .$$ (3.14)

This solution with arbitrary $\alpha(z)$ is a bit too complicated to be very instructive. One can, however, reconfirm that for $q \to \infty$, $W(v) \to 1$ and $\phi(v) \to 1/v_m$.

If $\phi(v)$ is to be monotone non-increasing in v, we must have $\phi'(v) \le 0$. From (3.12) we see that this implies

$$-\frac{\alpha'(v)}{\alpha(v)} \le \left[1 - \frac{v\ \alpha'(v)}{\alpha(v)}\right]\phi(v)$$

or

$$\frac{\alpha'(v)}{\alpha(v)} \le \frac{1}{v - u(v)} .$$ (3.15)

If this is to hold for arbitrarily large q, i.e., for $u(v) \to v_m$, then we must have

If this is to hold for arbitrarily large q, i.e., for $u(v) \to v_m$, then we must have

$$\frac{\alpha'(v)}{\alpha(v)} \le \frac{1}{v - v_m}$$

for all v. Carleson's choice of $\alpha(v) = v$ satisfies this condition but one cannot allow $\alpha(v)$ to increase as a power of v larger than one.

For the special case $\alpha(v) = v$, $W(v) = 1$ and (3.14) gives

$$\phi(v) = \frac{1}{v_m} - \int_{v_m}^{v} \frac{dz}{z^2 \left[1 + q \int_{v_m}^{z} d\, F_V(x)\, \beta(x) \right]}$$

as given by Carleson. If $\alpha(v) = 1$, (3.14) gives

$$\phi(v) = \frac{1}{v_m W(v)} = \frac{1}{v_m} \exp \left[- \int_{v_m}^{v} \frac{dz}{z + q \int_{v_m}^{z} d\, F_V(x)\, \beta(x)} \right].$$

The total number of passings per unit time and unit length of highway

$$q^2 \int_{o}^{\infty} \frac{d\, F_V(v)}{u(v)} \int_{o}^{v} \frac{d\, F_v(v')}{u(v')}\, [u(v) - u(v')]$$

$$= \frac{q^2}{2} \int_{o}^{\infty} d\, F_V(v) \int_{o}^{\infty} d\, F_V(v')\, |\, \phi(v) - \phi(v')\, |$$

For low flows this is proportional to q^2 but for high flows, $\phi(v) \rightarrow 1/v_m$ and one can

readily see from the above expressions for $\phi(v)$ that $\phi(v) - \phi(v')$ typically vanishes as q^{-1} for

$q \rightarrow \infty$. The rate of passing therefore increases only linearly with q for large q (provided the

delay per passing is independent of q). In fact the rate of passing probably decreases for very

$q \rightarrow \infty$. The rate of passing therefore increases only linearly with q for large q (provided the

delay per passing is independent of q). In fact the rate of passing probably decreases for very

large q because the delay per passing becomes large.

References
Chapter V

1. L. Carleson, "A mathematical model for highway traffic," Nordisk Mat. Tidskr. 5, 176-180, 213 (Swedish, English summary) (1957).

2. A.J. Miller, "Road traffic flow considered as a stochastic process," Proc. Cambridge Phil. Soc. 58 (1962), 312-325 corrigendum, ibid. 59, 508 (1963).

3. J.C. Tanner, "A simplified model for delays on a two-lane road," J. Roy. Statist. Soc. 20, 408-414 (1958).

4. I. Prigogine, "A Boltzmann-like approach to the statistical theory of traffic flow," Proc. first international symposium on the theory of traffic flow, Detroit 1959, pp. 158-164, Academic Press (1961).

5. J.A. McFadden, On the lengths of intervals in a stationary point process, J. Royal Statist. Soc. 24, 364-382 (1962).

6. J.A. McFadden and W. Weissblum, Higher-order properties of a stationary point process, J. Royal Statist. Soc. 25, 413-431 (1963).

7. D.R. Cox, Renewal Theory, Methuen, London (1962).

8. A.J. Miller, A queueing model for road traffic flow, J. Royal Statist Soc. B23, 63-75 (1961).

9. P. Athol, Headway groupings. Highway Research Record #72, Highway Research Board, pp. 137-155, 1965.

10. F. Garwood, An application of the theory of probability to the operation of vehicular-controlled traffic signals, J. Royal Statist. Soc. B7, 65-77 (1940).

11 . J.C. Tanner, The delay to pedestrians crossing a road, Biometrika 38, 383-392 (1951).
controlled traffic signals, J. Royal Statist. Soc. B7, 65-77 (1940).

11 . J.C. Tanner, The delay to pedestrians crossing a road, Biometrika 38, 383-392 (1951).

12. M.S. Raff, The distribution of blocks in an uncongested stream of automobile traffic, J. Amer. Statist. Assoc. 46, 114-123 (1951).

13. A.J. Mayne, Some further results in the theory of pedestrians and road traffic, Biometrika 41, 375-89 (1954).

14. G.H. Weiss and A. Maradudin, Some problems in traffic delay, Operations Res. 10, 74-104 (1962).

15. R.M. Oliver, Distribution of gaps and blocks in a traffic stream, Operations Res. 10, 197-217 (1962).

16. D.P. Gaver, Accommodation of second-class traffic, Operations Res. 11, 72-87 (1963).

17. G.F. Yeo and B. Weesakol, Delays to road traffic at an intersection, J. Applied Probability 1, 297-310 (1964).

VI. POSTSCRIPT

1. <u>Low density traffic</u>. Physicists and chemists who tried to apply the techniques of statistical mechanics to traffic flow theory translated much of the terminology and notation of physics into analogous features of traffic flow. This, however, was not always the most appropriate way to describe traffic.

In physics, the distribution of the velocities of molecules at some point in space and time depends on the local macroscopic properties (particularly the temperature). Positions and velocities are, of course, vectors in three dimensions, so, for a hypothetical one-dimensional system, it seemed natural to introduce a joint distribution of velocity and position, the ρ_s of II (6.1), and to consider the ρ_t of II (6.2), (6.3) as a "derived property" from the ρ_s.

In physics, all molecules (of the same chemical structure) are considered to be equivalent. Although the molecule may temporarily have a particular velocity, its velocity changes after it collides with other molecules and it then behaves the same as any other molecule with the same new velocity, independent of its past history. For highway traffic not all drivers are equivalent in the above sense. A driver may be aggressive or timid and he will take a trip from some specified origin to a specified destination regardless of how his motion may be affected by other drivers.

In dealing with flows over networks in which vehicles enter or leave some particular may be affected by other drivers.

In dealing with flows over networks in which vehicles enter or leave some particular road segment, it is generally necessary to keep account of where each driver will enter or leave the highway. Presumably, if one driver should pass another or otherwise interact with other vehicles, he will still remember where he wanted to go.

In "traffic flow theory" one does not typically stratify drivers by origin and destination. The implication is that the way a driver behaves in traveling over some section of road with no entrances or exits is independent of where his trip originated or where it will finally end. But even if there is some correlation between driver behavior and trip length, for example, there would still be some (marginal) distribution of driver types averaged over all trip lengths.

If drivers of any type (aggressive or timid) make a specified number of trips per unit time (hour or day), this would dictate the (average) <u>flow</u> past some point on a highway for vehicles of that type. Thus the relative proportions of aggressive or timid drivers passing some point on a highway should be a property of the origin and destination of trips, (nearly) independent of the interactions between cars. If, therefore, one wishes to describe a distribution of driver types in terms of a distribution of "desired speeds," it is more natural to describe it in terms of the $q^{(i)}$ of II (5.2) or the ρ_t rather than the $k^{(2)}$ or ρ_s. If, for some reason (passing delays or whatever), aggressive drivers cannot maintain their desired speeds, the <u>flow</u> of aggressive drivers should remain unchanged, but the spatial density of such drivers would increase.

Note that in three dimensions, velocity is a vector, and so would be the analogue of the flow q. It would be quite cumbersome in physics to introduce something analogous to the Note that in three dimensions, velocity is a vector, and so would be the analogue of the flow q. It would be quite cumbersome in physics to introduce something analogous to the ρ_t of one spatial dimension.

The issue raised here actually has little effect on the conclusions from chapters II and III. Even in chapter III we evaluated only the <u>relative</u> fractions of paired vehicles as compared with single vehicles and only to first order in k or q. The point here is more one of

style than substance, but in physics the natural question is: given the state of the system (the positions and velocities of all molecules) at time 0, what will be the state at some time t? In traffic theory the more natural question is: given the times and types of cars entering a highway at location x = 0, when will they pass some location x > 0.

Another difference between physics and traffic is that in physics it is natural to deal with the (vector) velocity or (vector) momentum, because there is a conservation of momentum as a result of any interaction between molecules. Also the kinetic energy depends on the square of the velocity and there is a conservation of energy. In traffic there is no meaningful analogue of momentum or energy; there is certainly no analogue of the conservation of momentum or energy. In traffic, trips will travel a specified distance dictated by the origin-destination of trips independent of how long it takes (a "conservation" of distance traveled).

Despite the fact that traffic engineers and the general public like to describe motion in terms of the velocity (or speed), the quantity which enters most naturally into any analysis is the reciprocal of the velocity, 1/v, the time to travel unit distance (the only common English word related to this seems to be "pace"). Even the measurement of velocity is typically made by observing the transit time between two detectors, i.e., the time to travel a specified distance.

The basic relation of q = kv, II (2.7) suggests that one would determine the flow q from the density k and the "space mean" velocity v, although one could write this as k = q/v. The 1/v, however, is also equal to the arithmetic mean of the "pace" of vehicles crossing a fixed location (the "time-mean" of the $1/v_j$'s), i.e., the sum of times for vehicles to pass

285

between detectors divided by the distance between detectors and the number of vehicles. The whole issue of space mean vs time mean, harmonic mean, etc. would be much less confusing if we could train ourselves to think in terms of the (average) pace and the flow q (which are the things most easily measured) and consider k as a derived quantity.

Also in relation to the material given in chapters I, II, and III, noticeably absent from the theory is any reference to the "Boltzmann-like" approach to traffic theory promoted by I. Prigogine and collaborators over a period of time starting in 1959 [1] and culminating in a book, "Kinetic Theory of Vehicular Traffic," by I. Prigogine and R. Herman [2] (P-H) in 1971. A discussion of this was promised in chapter V but never written. Since this theory was supposed to apply over a wide range of densities, it would logically form a part of the discussion for "moderate densities," but the theory is based mostly on arguments about the behavior of light traffic.

Reference [2], Chapter 3 on low-density traffic starts with an observation that the ρ_s of II (6.1) satisfies the differential equation

$$\frac{\partial \rho_s}{\partial t} + v \frac{\partial \rho_s}{\partial x} = 0 \tag{1.1}$$

the solution of which is

$$\rho_s(x,v,t) = \rho_s(x-vt,v,0) \tag{1.2}$$

the solution of which is

$$\rho_s(x,v,t) = \rho_s(x-vt,v,0) \tag{1.2}$$

as in II (6.4). The interpretation of (1.1) is that the derivative $d\rho_s/dt$ as seen by an observer traveling with the velocity v is zero. P-H then note that a non-uniform spatial distribution k(x,t) will tend to become more uniform with increasing time due to the spread of velocities, but they do not point out that the velocity distribution becomes <u>less</u> uniform at the same time.

286

P-H now ask what would happen if this distribution were "perturbed as a consequence of various factors, such as obstacles, weather, or interaction with other drivers." Presumably this perturbation lasts for only a finite time after which drivers are free to travel as they wish. If we were to define t = 0 as the time when this perturbation is removed, then it would seem that the subsequent motion should again satisfy (1.2) with the distribution ρ_s at t=0 assigned whatever was the result of perturbation.

P-H, however, now propose that there exists some "desired speed distribution" which, in keeping with the notation used here, we will label as $\rho_s^o(x,v,t)$. This is specifically chosen to have a form

$$\rho_s^o(x,v,t) = k(x,t)f_o(v). \tag{1.3}$$

The $f_o(v)$ is some prespecified time and space independent probability density of (desired) velocities. The $\rho_s(x,v,t)$ is now postulated to satisfy an equation of the form

$$\frac{\partial \rho_s}{\partial t} + v \frac{\partial \rho_s}{\partial x} = \frac{\rho_o - \rho_s}{T}, \tag{1.4}$$

for some "relaxation time" T (which is "a few seconds").

The effect of such a relaxation term on the right-hand side of (1.4) is <u>not</u> that individual drivers recover their previous desired speeds (before the perturbation), but that cars near some location x will collectively try to acquire a velocity distribution $f_o(v)$. Thus, if at individual drivers recover their previous desired speeds (before the perturbation), but that cars near some location x will collectively try to acquire a velocity distribution $f_o(v)$. Thus, if at time 0, one had a local concentration of aggressive drivers (as would exist downstream of a traffic signal), these drivers would quickly be replaced by drivers sampled from the distribution $f_o(v)$. There clearly is no rational justification for this.

287

To model the effects of interactions between cars (at low density), P-H add another term to the right-hand side of (1.4) which describes a mechanism by which fast cars will overtake slower cars and follow the latter, but they do not include a mechanism for the fast cars to return to their previous velocities after passing. The rate of overtaking is evaluated in a way similar to that described in chapter III. If one includes in (1.4) the interaction term but not the relaxation term, the average speed of cars would <u>always</u> be decreasing with time as faster cars overtake slower ones (unless all cars travel with the same velocity). Presumably it is the relaxation term which provides the mechanism for velocities to increase after passing, but the relaxation term would reassign to the fast car a velocity sampled from the distribution f_o. The fast car would not recover its own previous velocity.

Our conclusion is that this "Boltzmann-like" theory bears little relation to the real world.

2. <u>High density traffic</u>. Perhaps the most important development in transportation theory during the last 30 years is the increased use of "cumulative curves" for the description of traffic behavior. This does not really represent any new "theory"; it is simply a more convenient notation for describing it.

For any particular class of vehicles or for all vehicles, let
convenient notation for describing it.

For any particular class of vehicles or for all vehicles, let

$A(x,t) = $ *cumulative number of vehicles to pass x by time t,*

starting from some reference vehicle labeled as 0.

Equivalently, if someone were to attach numbers to the vehicles in order as they passed him, and vehicles either did not pass each other or interchanged numbers if they did pass, then $A(x,t)$ is the number of the last vehicle to pass x before time t.

Since $A(x,t)$ is integer valued, the function $A(x,t)$ is a step function in the three dimensional space (A,x,t). A curve in the x,t plane along which $A(x,t)$ jumps from j-1 to j is the trajectory $x_j(t)$ of the jth vehicle. If, however, one smoothes out the steps so that the smoothed $A(x,t)$ has derivatives, one can define flows and densities as

$$q(x,t) = \partial A(x,t)/\partial t \ , \ k(x,t) = -\partial A(x,t)/\partial x. \tag{2.1}$$

The immediate advantage of dealing with $A(x,t)$ rather than trying to describe the details of individual trajectories, or even the flows or densities, is that any smoothing of the $A(x,t)$ will wipe out much of the detailed properties of vehicles which are not likely to be reproducible anyway. The more one can smooth irregularities in the curves or surfaces, the better one can "see" the "macroscopic" properties of the traffic. Furthermore, if one draws curves $A(x,t)$ at two specified locations x_1 and x_2, $x_2 > x_1$, $A(x_1,t)$ and $A(x_2, t)$, the vertical distance $A(x_1,t) - A(x_2,t)$ between the curves at time t is the number of vehicles between x_1 and x_2 at time t , and the horizontal distance between the curves at height j is the trip time of the jth vehicle from x_1 to x_2. Most descriptive properties of traffic behavior have simple geometric interpretations in terms of the $A(x,t)$. If one does not care to follow all the details the jth vehicle from x_1 to x_2. Most descriptive properties of traffic behavior have simple geometric interpretations in terms of the $A(x,t)$. If one does not care to follow all the details of the motion of vehicles between various points, it would suffice to draw the curves $A(x_j,t)$ only at certain critical locations x_j (bottlenecks, junctions, etc.).

The existence of a function A(x,t) automatically guarantees that vehicles are conserved (if one numbers the vehicles, they will not disappear). The conservation equation IV (4.2) is equivalent to the mathematical identity

$$\frac{\partial^2 A(x,t)}{\partial x \partial t} = \frac{\partial^2 A(x,t)}{\partial t \partial x} \, . \tag{2.2}$$

This fact has recently been exploited [3] to obtain a much simpler method of analyzing the solution of the continuum theory of Chapter IV, sec. 4, particularly if there are shocks.

In the absence of shocks, the assumption IV (4.1) implies that A(x,t) is a solution of the differential equation

$$\frac{\partial A(x,t)}{\partial t} = Q\left(- \frac{\partial A(x,t)}{\partial x}\right) \tag{2.3}$$

(for a homogeneous highway section). A formal solution of (2.3) is obtained in the same way as described in Chapter IV, sec. 4, by observing that q(x,t) and k(x,t) are constant along straight line characteristic curves (as in IV (4.5)). This, in turn, means that A(x,t) must be linearly increasing in x and t along such curves.

$$dA(x,t) = \frac{\partial A(x,t)}{\partial x} \, dx + \frac{\partial A(x,t)}{\partial t} \, dt$$

$$= - k \, dx + q \, dt \, .$$

$$= - k \, dx + q \, dt \, .$$

Thus from appropriate initial or boundary conditions where A(x,t), q(x,t), or k(x,t) are specified, one can integrate A(x,t) along the characteristic curves and evaluate the A(x,t) everywhere.

If characteristic curves should intersect, this would imply that there is more than one formal solution to (2.3) at certain points in the x-t plane. One of these solutions must be the "correct" solution. The only issue is: which solution is the correct one? It can be shown that (for a concave function $Q(k)$) the correct value of $A(x,t)$ is the smallest of all such formal solutions. This solution will automatically have discontinuities in slopes (k and/or q) where the formal solutions become multiple-valued. One does not need to determine the shock path in order to evaluate the $A(x,t)$; the shock path will emerge <u>from</u> the $A(x,t)$. This method avoids the rather tedious numerical integration of the shock equations described in Chapter IV, sec. 4.

This is the only significant development during the last 30 years specifically related to the type of theory described in Chapter IV.

There are deficiencies in this class of models as described, in part, in Chapter IV, sec. 9. Although for dense traffic with negligible passing, each vehicle is "following" the vehicle ahead of it, a driver will not necessarily follow at what he considers to be the minimum safe driving separation. If a driver has a spacing substantially larger than the minimum, he might reach some point a few seconds later than if he drove more aggressively. He might reach his destination a few seconds late, but, if this were an issue, he could close the gap just before he left the road. An individual driver has little incentive to travel close to the vehicle ahead. On

left the road. An individual driver has little incentive to travel close to the vehicle ahead. On the contrary, if he leaves a substantial cushion, he can drive in a more relaxed way without responding quickly to any action of the vehicles ahead.

Although driving at a minimum safe driving distance would be beneficial to all the drivers <u>behind</u> the one in question, the only apparent incentive to the driver himself is to

291

prevent other drivers, from adjacent lanes or on-ramps, for example, from jumping into the gap. Indeed, one can observe relatively large flows at certain times or locations where drivers have some motivation to be alert and drive with a short but safe spacing.

Models have been proposed [4,5] in which drivers traveling at some specified velocity will have one (average) spacing if they are accelerating so as to catch-up with the vehicles ahead, but a smaller spacing if they are decelerating so as to avoid collision. If they have some spacing between these two limits, they are rather indifferent to minor variations in the motions of individual vehicles. Such models can potentially predict "stop-and-go" driving, but it is not obvious how drivers cooperatively alternate between traveling in a lazy fashion (maybe even be stopped) and in an aggressive fashion. Certainly the laziness or indifference of drivers is an aspect of driver behavior which affects the collective behavior but the details are not well understood.

That traffic displays some instabilities has been recognized at least since the late 1950s. This instability, however, is not of a type associated with an unstable reaction time T as described in Chapter IV. Presumably the T is of the order of a few seconds, but the period of alternating stop-and-go driving is of the order of a few minutes.

There have been numerous attempts during the last 30 years to modify the continuum theory of Chapter IV, sec. 4 with the introduction of certain "second-order effects," relaxation,
anticipation, viscosity, etc. in an attempt to make certain analogies to phenomena related to fluid mechanics. Much like the Boltzmann-like theories mentioned in the previous section, these theories have little theoretical or experimental justification [6]. It seems that very little progress has been made during the last 30 years to develop more realistic models of dense

traffic than those described in Chapter IV even though it is known that these models are not always consistent with observed behavior of traffic.

References
Chapter VI

1. I. Prigogine, A Boltzmann-like approach to the statistical theory of traffic flow, Proc. symposium on the theory of traffic flow, Detroit (1959), R. Herman, ed., Elsevier Amsterdam (1961), pp. 158-164.

2. I. Prigogine and R. Herman, Kinetic theory of vehicular traffic. American Elsevier, New York (1971).

3. G.F. Newell, A simplified theory of kinematic waves in highway traffic I General theory, II Queueing at freeway bottlenecks, III Multi-destination flows, Transpn. Res. B 27, 281-313 (1993).

4. G.F. Newell, Theories of instability in dense highway traffic, J. Operations Res. Soc. Japan 5, 9-54 (1962).

5. G.F. Newell, Instability in dense highway traffic, a review, Proc. Second International Symp. on the theory of traffic flow, London (1963), OECD, Paris (1965).

6. C.F. Daganzo, Requiem for second order fluid approximations of traffic flow, (in press), Transportation Research B.

APPENDIX

BIBLIOGRAPHY ON CAR-FOLLOWING AND CONTINUUM THEORIES

Code: T - Theoretical study
              E - Experimental data included
              L - Linearized theories
              N - Nonlinear car-following
              F - Fluid theories

1945

Herrey, E.M.J. and Herrey, H., "Principles of physics applied to traffic movement and road conditions," Amer. J. of Physics 13, 1-14 (1945); T, N

1950

Reuschel, A., "Fahrzeugbewegungen in der Kolonne bei gleich förmig bescheunigtem oder verzögertem Leitfahrzueg," Zeits. d. Osterreich Ing. v. Arch. Vereines, 95, 59-62, 73-77 (1950); T, L

Reuschel, A., "Fahrzeugbewegungen in der Kolonne," Ost. Ing. Archiv, 4, 193-215 (1950); T, L

1953

Pipes, L.A., "An operational analysis of traffic dynamics," J. Appl. Physics 24, 274-281 (1953); T, L

1955

Lighthill, M.J. and Whitham, G.B., "On kinematic waves II, a theory of traffic flow on long

1955

Lighthill, M.J. and Whitham, G.B., "On kinematic waves II, a theory of traffic flow on long crowded roads," Proc. Roy. Soc. (London) A 229, 317-345 (1955). Reprinted in "An introduction to traffic flow theory," HRB, report 79; T, F

1956

Richards, P.I., "Shock waves on the highway," Opns. Res. 4, 42-51 (1956); T, F

De, S.C., "Kinematic wave theory of bottlenecks of varying capacity," Proc. Cambridge Phil. Soc. 52, 564-572 (1956); T, F

Lighthill, M.J., "Dynamic instability of transport systems: the hydrodynamic analogy," Operational Res. Q, 8, 109-114 (1957), T, F

1958

Edie, L.C., and Foote, R.S., "Traffic flow in tunnels," Highway Res. Bd. Proc. 37, 334-344 (1958); E, F

Chandler, R.E., Herman, R., and Montroll, E.W., "Traffic dynamics: studies in car following," Opns. Res. 6, 165-184 (1958); T, E, L

Kometani, E., and Sasaki, T., "On the stability of traffic flow (Report I)," J. of Opns. Res., Japan 2, 11-26 (1958); T, L

Chow, Tse-Sun, "Operational analysis of a traffic dynamics problem," Opns. Res., 6, 827-834 (1958); T, L

1959

Sasaki, T., "On the stability of traffic flow (Report II)," J. Opns. Res. Japan, 2, 60-79 (1959), T, L

Helly, W., "Dynamics of single-lane vehicular traffic flow," Research Report No. 2, Center for Operations Research, Mass. Institute of Technology (1959), Ph.D. Thesis, M.I.T. (1959); T, L

Greenberg, H., "An analysis of traffic flow," Opns. Res., 7, 79-85 (1959); T, N

Herman, R., Montroll, E.W., Potts, R.B., and Rothery, R.W., "Traffic dynamics: analysis of stability in car following," Opns. Res., 7, 86-106 (1959); T, L

Gazis, D.C., Herman, R., and Potts, R.B., "Car following theory of steady state traffic flow," Opns. Res. 7, 499-505 (1959); E, N
stability in car following," Opns. Res., 7, 86-106 (1959); T, L

Gazis, D.C., Herman, R., and Potts, R.B., "Car following theory of steady state traffic flow," Opns. Res. 7, 499-505 (1959); E, N

Newell, G.F., "A theory of platoon formation in tunnel traffic," Opns. Res. 7, 589-598 (1959), T, N

Kometani, E., and Sasaki, T., "A safety index for traffic with linear spacing," Mem. Fac. Engrg., Kyoto University 21, 229-246 (1959), Opns. Res. 7 (1959), 704-720; T, L

Kometani, E., and Sasaki, T., "Dynamic behavior of traffic with a nonlinear spacing-speed relationship," Proc. 1st symp. (1959), Theory of traffic flow (1961), pp. 105-119. Elsevier: T, E, N

Herman, R., and Potts, R.B., "Single lane traffic theory and experiment," Proc. 1st symp. 1959, Theory of traffic flow (1961), pp. 120-146, Elsevier; T, E, N

Montroll, E.W., "Acceleration noise and clustering tendency of vehicular traffic," Proc. 1st symp. (1959), Theory of traffic flow (1961), pp. 147-157, Elsevier; T, L

Edie, L.C., and Foote, R.S., "Experiments on single-lane flow in tunnels," Proc. 1st symp. 1959, Theory of traffic flow (1961), pp. 175-192, Elsevier; E, F

Newell, G.F., "A theory of traffic flow in tunnels," Proc. 1st symp. (1959), Theory of traffic flow (1961), pp. 193-206, Elsevier; T, N

Helly, W., "Simulation of bottlenecks in single lane traffic flow," Proc. 1st symp. (1959), Theory of traffic flow (1961), pp. 207-238, Elsevier; T, E, N

1960

Edie, L.C., and Foote, R.S., "Effect of shock waves on tunnel traffic flow," Highway Res. Bd., Proc. 39, 492-505 (1960); E, F

Bick, J.H., and Newell, G.F., "A continuum model for two-directional traffic flow," Q. Applied Math., 18, 191-204 (1960); T, F

Franklin, R.E., "Single lane traffic flow," Civ. Eng. & Public Works Rev., 55, 1429 (1960); T, N

Kisi, T., "Traffic dynamics: analysis as sampled-data control system," J. Opns. Res. Soc., Japan, 2, 114-123 (1960); T, L

Greenberg, H., and Daou, A., "The control of traffic flow to increase the flow," Opns. Res., 8, 524-532 (1960); E, F

Greenberg, H., and Daou, A., "The control of traffic flow to increase the flow," Opns. Res., 8, 524-532 (1960); E, F

1961

Newell, G.F., "Nonlinear effects in the dynamics of car following," Opns. Res., 9, 209-229 (1961); T, N

Edie, L.C., "Car-following and steady-state theory for noncongested traffic," Opns. Res., 9, 66-76 (1961); E, T, N

Gazis, D.C., Herman, R., and Rothery, R.W., "Nonlinear follow-the-leader models," Opns. Res., 9, 546-567 (1961); T, E, N

Franklin, R.E., "The structure of a traffic shock wave," Civil Eng. and Public Works, Rev. 56, 1186-1188 (1961); T, N

1962

Franklin, R.E., "The propagation of disturbances in a single lane of traffic," Civil Eng. and Public Works Rev. 57, 344-346, 487-489 (1962); T, N

Gazis, D.C., Herman, R., and Weiss, G.H., "Density oscillations between lanes on a multilane highway," Opns. Res., 10, 658-667 (1962); T, F

Fortet, R., "Theorie du Trafic de Convoi," Recherche Operationnelle (1962), 337-360; T, L

Newell, G.F., "Theories of instability in dense highway traffic," J. Opns. Res. Soc., Japan, 5, 9-54 (1962); T, N

Herman, R., and Rothery, R., "Microscopic and macroscopic aspects of single lane traffic flow," J. Opns. Res. Soc., Japan, 5, 74-93 (1962); T, E, N

Foster, J., "An investigation of the hydrodynamic model for traffic flow with particular reference to the effect of various speed-density relationships," Proc. 1st Conference Australian Road Research Board (1962), pp. 229-257; E, F

1963

Montroll, E.W., "Theory and observations of the dynamics and statistics of traffic on an open road," Proc., 1st symp. on Engineering Applications of Random Function Theory," Bagdanoff and Kozin, ed., pp. 231-269 (1963); T, L

Edie, L.C., Foote, R.S., Herman, R., and Rothery, R., "Analysis of single lane traffic flow," Traffic Eng., 33, 21-27 (1963); T, E, N

Edie, L.C., Foote, R.S., Herman, R., and Rothery, R., "Analysis of single lane traffic flow," Traffic Eng., 33, 21-27 (1963); T, E, N

Herman, R., and Rothery, R.W., "Car following and steady state flow," Proc. 2nd Int. Symp. Theory of Traffic Flow, London 1963, OECD, Paris (1965), pp. 1-11; T, E, N

Edie, L.C., "Discussion on traffic stream measurements and definitions," Proc., 2nd Int. Symp. Theory of Traffic Flow, London 1963, OECD, Paris (1965), pp. 139-154; T, F

Foote, R.W., "Single lane traffic flow control," Proc. 2nd Int. Symp. Theory of Traffic Flow, London 1963, OECD, Paris (1965), pp. 84-103; E, N

Fortet, R., "La circulation des voitures sur une route ou dans les artères d'une ville, à partie des lois de l'hydraulique, The fluid mechanics approach to town and country traffic flow," Houille Blanche 18 (1963), 909-916; T, F

Franklin, R.E., "On the flow--concentration for traffic," Proc. 2nd Int. Symp. on Theory of Traffic Flow, London 1963, OECD, Paris (1965), pp. 120-128; T, N

Gazis, D.C., Herman, R., and Rothery, R.W., "Analytical methods in transportation: mathematical car-following theory of traffic flow," Proc., Amer. Soc. Civil Engrs., J. Engr. Mechanics Div., Proc. paper 3724, (1963), pp. 29-46; T, N

Michaels, R.M., "Perceptual factors in car following," Proc. 2nd Int. Symp. Theory of Traffic Flow, London 1963, OECD, Paris (1965), pp. 44-59; E, T, N

Newell, G.F., "Instability in dense highway traffic, a review," Proc. 2nd Int. Symp. Theory of Traffic Flow, London 1963, OECD, Paris (1965), pp. 73-83; T, N

Oliver, R.M., Catoire, B., and Skeates, R.S., "On a mathematical function of traffic flow theory," Proc. 2nd Int. Symp. Theory of Traffic Flow, London 1963, OECD, Paris (1965), pp. 12-26; T, L

Wardrop, J.G., "Experimental speed/flow relations in a single lane," Proc., 2nd Int. Symp. Theory of Traffic Flow, London 1963, OECD, Paris (1965), pp. 104-119; E, F

1964

Rothery, R., Silver, R., Herman, R., and Turner, C., "Analysis of experiments on single-lane bus flow," Opns. Res., 12, 913-933 (1964); E, T, N

Pipes, L.A., "Hydrodynamic approaches, chap. 1, an introduction to traffic flow theory," HRB Special Report 79 (1964); T, F

Montroll, E.W., and Potts, R.B., "Car-following and acceleration noise, Chap. 2, an introduction to traffic flow theory," HRB, Special Report 79 (1964); T, N

Montroll, E.W., and Potts, R.B., "Car-following and acceleration noise, Chap. 2, an introduction to traffic flow theory," HRB, Special Report 79 (1964); T, N

Oliver, R.M., "Travel times through shock waves," Operational Res., Q. 15, 139-144 (1964); T, N

1965

Battle, C.T., "Models of car-following," Ph.D. thesis, Harvard University, September 1965; T, N

Leutzbach, W., "Testing the applicability of the theory of continuity on traffic flow after bottlenecks," Proc. 3d symp. (1965), Vehicular Traffic Science, Amer. Elsevier (1967), pp. 1-13; T, E, F

Herman, R., and Rothery, R., "Propagation of disturbances in vehicular platoons," Proc. 3d. symp. (1965), Vehicular Traffic Science, Amer. Elsevier (1967), pp. 14-25; T, E, N

Edie, L.C., and Baverez, E., "The generation and propagation of stop-start traffic waves," Proc. 3d symp. (1965), Vehicular Traffic Science, Amer. Elsevier (1967); pp. 26-37; T, E, N

Franklin, R.E., "Single-lane traffic flow on circular and straight tracks," Proc. 3d symp. (1965), Vehicular Traffic Science, Amer. Elsevier (1967), pp. 42-55, T, E, N

Oliver, R.M., and Lam, T., "Statistical experiments with a two-lane flow model," Proc. 3d symp. (1965), Vehicular Traffic Science, Amer. Elsevier (1967), pp. 170-180; T, E, F

Pipes, L.A., "Wave theories of traffic flow," J. Franklin Inst., 280, 23-41 (1965); T, F

1966

Lee, G., "A generalization of linear car-following theory," Opns. Res., 14, 595-606 (1966); T, L

1968

Treiterer, J., "Some aspects of the stability of traffic flow," Proc. 4th symp. Karlsruhe 1968, Beiträge zur Theorie des Verkehrsflusses, Bonn (1969), pp. 8-13; T, N

Bennett, B.T., Gazis, D.C., Edie, L.C., and Foote, R.S., "Control of the Lincoln tunnel traffic by an on-line digital computer," Proc. 4th symp. Karlsruhe 1968, Beiträge zur Theory des Verkehrsflusses, Bonn (1969), pp. 48-56; T, E, F

Herman, R., and Rothery, R., "Frequency and amplitude dependence of disturbances in a traffic stream," Proc. 4th symp. Karlsruhe 1968, Beiträge zur Theorie des Verkehrsflusses, Bonn (1969), pp. 14-21; T, E, L

Herman, R., and Rothery, R., "Frequency and amplitude dependence of disturbances in a traffic stream," Proc. 4th symp. Karlsruhe 1968, Beiträge zur Theorie des Verkehrsflusses, Bonn (1969), pp. 14-21; T, E, L

1969

Mika, H.S., Kreer, J.B., and Yaun, L.S., "Dual mode behavior of freeway traffic," HRB Record #279, 1-12 (1969); E, N

1970

Herman, R., Lam, T., and Rothery, R., "Further studies on single-lane bus flow; transient characteristics," Trans. Sci., 4, 187-216 (1970); T, E, N

Lam, T., and Rothery, R., "The spectral analysis of speed fluctuations on a freeway," Trans. Sci., 4, 293-310 (1970); T, E, N

1971

Herman, R., Lam, T., and Rothery, R., "The starting characteristics of automobile platoons," Proc. 5th int. symposium theory of traffic flow and transportation, Berkeley 1971, Elsevier (1972), pp. 1-18; E, N

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

Munjal, P.K., and Pipes, L.A., "Propagation of on-ramp density waves on uniform unidirectional multilane freeways," Trans. Sci., 5, 390-402 (1971); T, F

Wright, C.C., "Some properties of the fundamental relations of traffic flow," Proc. 5th int. symp. theory of traffic flow and transportation, Berkeley 1971, Elsevier (1972), pp. 19-32; T, E, N

Darroch, J.N., and Rothery, R.W., "Car following and spectral analysis," Proc. 5th int. symp. theory of traffic flow and transportation, Berkeley 1971, Elsevier 1972, pp. 47-56; T, E, L

1972

Hida, K., "A new approach to kinematic waves in traffic flow," J. Physical Soc., Japan, 33, 817-827 (1972); T, F

Preparata, F.P., "Analysis of traffic flow on a signalized one-way artery," Trans. Sci., 6, 32-51 (1972); T, F

Preparata, F.P., "Analysis of traffic flow on a signalized one-way artery," Trans. Sci., 6, 32-51 (1972); T, F