K.B. DAVIDSON

THE THEORETICAL BASIS OF A FLOW - TRAVEL TIME RELATIONSHIP FOR USE IN TRANSPORTATION PLANNING

ABSTRACT

The flow-travel time relationship, first proposed in 1966, has been found to have all the required characteristics of such a relationship; however, beyond a loose association with queueing theory, no adequate theoretical basis had been found. This paper demonstrates that the relationship can be derived from either of two queueing models, both of which assume the road operates as a specific number of sequential queueing elements. In one model the delay parameter J is the proportion of elements which cause delay and in the other it has a close association with the Erlang number. This association is explored and comments on practical use of the relationship are made.

INTRODUCTION

The author proposed in Davidson (1966) a flow travel time relationship (*Fig. 1*) which has the form:

$$T = t\,\frac{1 - c(1 - J)}{1 - c} \qquad (1)$$

or

$$T = t\,\frac{S - Q(1 - J)}{S - Q} \qquad (2)$$

where

Q = vehicular flow rate
S = saturation flow rate
t = 'zero flow' travel time
T = travel time at flow rate Q
c = degree of saturation Q/S
J = a parameter.

The relationship was generated from concepts of queueing theory and where $J = 1$, $T/t = 1/(1-c)$ is the ratio of time in the system to service time under steady-state conditions in a single channel facility with a random arrival rate and exponentially distributed service rate.

For the simple system described above, the ratio of delay in the queue to service time is $c/(1-c)$ but the paper stated that

> traffic on a road is not truly a single continuous queueing situation. Rather, delay on a road is caused by a succession of queueing situations such that a varying amount of the total service time is subject to queueing delays.

The delay parameter, J, was introduced to represent this variation.

It has always been recognised that this was not a rigorous derivation, but the relationship had all the desirable shape characteristics of a flow-travel time relationship and had the added advantage that the delay parameter J allowed infinite variation to the shape of the curve to thus provide for the individual characteristics of different roads.

The basic problem with any attempt at a rigorous queueing derivation is that simple single-channel queueing systems require that only one unit be in service at any time. If a length of road is considered as a queueing system then, clearly, more than one unit is in service and the values of mean service time and the reciprocal of mean service rate are different (mean service time for a road is the time to travel over the section when there are no delays caused by other vehicles whilst mean service rate is the saturation flow rate), whereas in simple single channel queueing systems these values are identical.

TWO NON-RIGOROUS MODELS

Perhaps some progress can be made if a system is defined which equates these values. One way to do this is to define a series of service facilities along the road such that each facility has a mean service time equal to the headway at saturation flow rate (i.e. the reciprocal of mean service rate). If $1/u$ = mean service time, then $u = S$, the saturation flow rate. If the section of road requires time t to traverse in the absence of any delays caused by other vehicles, then there are, by definition, L service facilities in series, each with mean service time $1/u$ so that

$$t = \frac{L}{u} \qquad (3)$$

MODEL 1

The original derivation of the model can, in terms of this system, be more precisely stated if it is considered that, of the L service facilities, only a proportion J operate as such, and that they do so with a random service time distribution. The remainder are considered not to operate and contribute no queueing delay at all, $1/u$ being the time to traverse each one. The mean queueing delay at one facility, $\overline{Wq}$, when both arrival and service distributions are random is

$$\overline{W}q = \frac{1}{u}\,\frac{c}{1-c} \tag{4}$$

If it is further assumed that the series operation of these service elements does not disturb the randomness of the arrival pattern at any element, then total queueing delay

$$= JL\frac{1}{u}\,\frac{c}{1-c} \tag{5}$$

since there are L elements and a proportion, J, of them are operating. The total time to traverse the section of road, T, is then the sum of queueing delay and the service time in all of the elements:

$$T = J\frac{L}{u}\,\frac{c}{1-c} + \frac{L}{u} \tag{6}$$

$$= \frac{L}{u}\cdot\frac{1-c(1-J)}{1-c} \tag{7}$$

$$= t\,\frac{1-c(1-J)}{1-c} \tag{8}$$

which corresponds to eqn (1). This model can thus be used to generate the relationship.

MODEL 2

Alternatively, it is interesting to note that a single channel, infinite queue system with random arrivals and an Erlang service distribution gives, according to Blunden (1971), p. 63 or Drew (1968) p. 250:

$$\overline{W}q = \frac{K+1}{2K}\cdot\frac{c}{1-c}\cdot\frac{1}{u} \tag{9}$$

where K is the Erlang Number in the distribution (after Blunden 1971, p. 51)

$$F(t) = e^{-Kut}\sum_{n=0}^{K-1}\frac{(Kut)^n}{n!} \tag{10}$$

Mean time in the system, $\overline{W} = \overline{W}q$ + time in service,

$$\text{i.e.}\qquad \overline{W} = \overline{W}q + \frac{1}{u} \tag{11}$$

$$= \frac{1}{u}\cdot\frac{K+1}{2K}\,\frac{c}{1-c} + \frac{1}{u} \tag{12}$$

$$= \frac{1}{u}\,\frac{c\,\frac{K+1}{2K} + (1-c)}{1-c} \tag{13}$$

$$= \frac{1}{u}\,\frac{1-c\left(1-\frac{K+1}{2K}\right)}{1-c} \tag{14}$$

This corresponds to the flow-travel time relationship if

$$J = \frac{K+1}{2K} \tag{15}$$

$$T = L\overline{W} \tag{16}$$

$$t = L\frac{1}{u} \tag{3}$$

In this case $\overline{W}$ is interpreted as the average time in the system for each of the L elements, which, on average, behave as though they each had random arrivals and Erlang service. It cannot be said that each element actually behaves in this way since the output from each element, and hence the input to the next, would tend to be Erlang rather than random. However, as shown in Blunden (1971) p. 64, delay characteristics of a system with Erlang distributions in either the arrival pattern or the service pattern or both are very similar and could be at least well approximated to by a delay curve for a system with random arrivals and Erlang service where the Erlang number was properly selected. Hence, defining an Erlang number which represents the average service behaviour of each of the L elements when random arrivals are assumed is not likely to lead to an erroneous shape for the delay curve. Again, this does not constitute a strict derivation of the relationship but it is suggested as a reasonable model.

SUMMARY

The relationship can thus be derived from one of two alternative concepts. In one case it is assumed that a proportion J of L potential queueing-system elements in the section of road are operating as delay-producing elements and that they do so with random arrivals and service. In the other case, all L elements operate, and on average they behave as though they were single-channel infinite-queue systems each with random arrivals and Erlang service, with the Erlang number being such that $J = (K + 1)/2K$.

RELATIONSHIP BETWEEN DELAY PARAMETER AND ERLANG NUMBER

With the second interpretation, J is directly related to the Erlang number of the service distribution

$$J = \frac{K+1}{2K} \tag{15}$$

Thus, if $J = 1$ then $K = 1$ and this corresponds to a single-channel infinite-queueing case with random arrivals and exponential service. If $J = 0.5$, $K = \infty$, corresponding to regular service which results in exactly half the delay of the exponential (random) service case. (Note in *Fig. 1* that $T - t$ for $J = 0.5$ is half the value of $T - t$ when $J = 1$.)

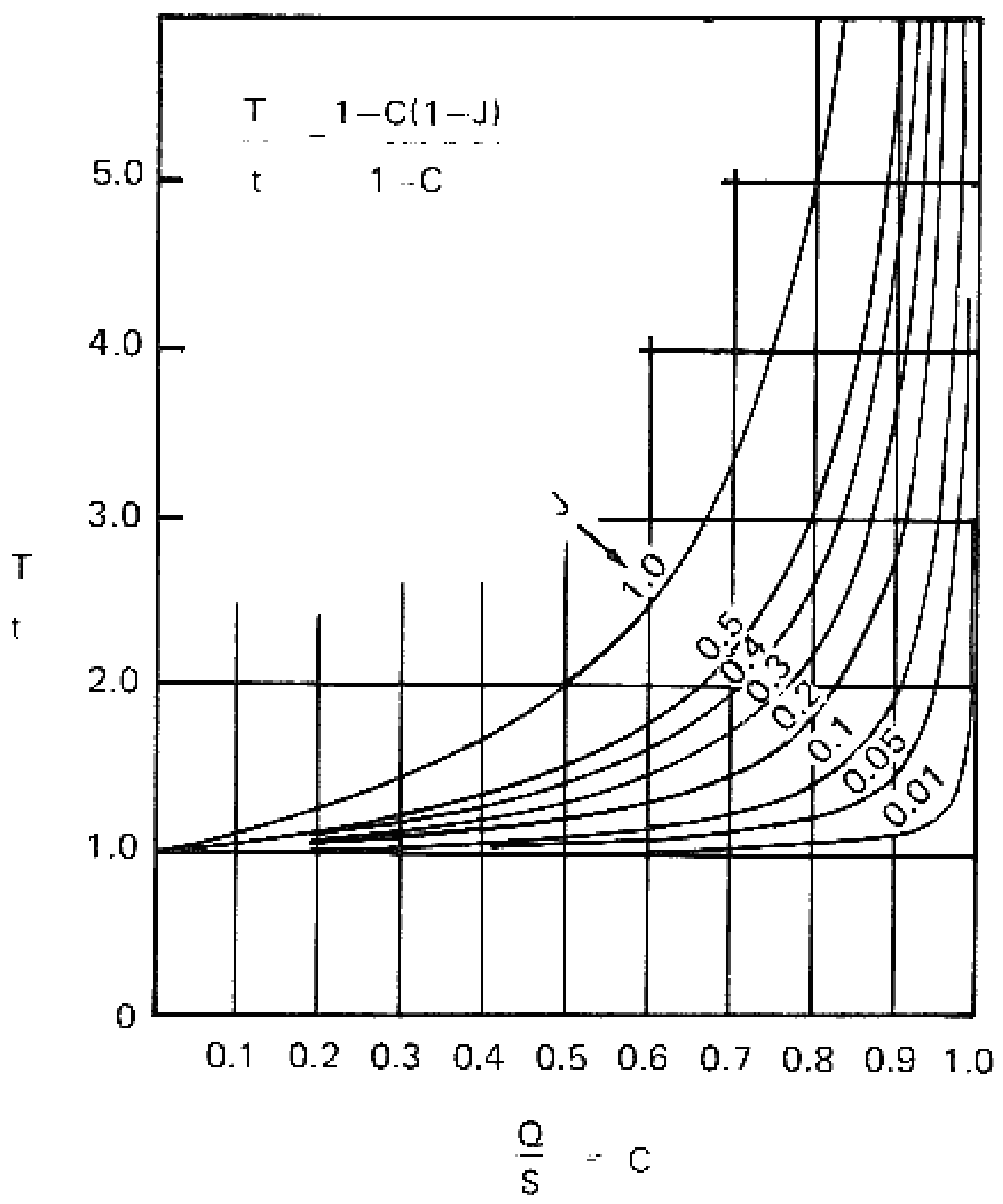

Fig. 1 — Proposed flow-travel time relationship showing a range of values of the delay parameter (J)

However, if the service distribution goes from exponential (random) to regular as J goes from 1 to 0.5 (K goes from 1 to ∞), what interpretation can be placed on values of J of less than 0.5? *Fig. 2* shows the relationship between J and K for the whole range of J values. At $J = 0.5$ there is a discontinuity with K going from ∞ to $-\infty$ and as $J \rightarrow 0$, $K \rightarrow -1$.

It is suggested that this has a fairly obvious physical interpretation. In the analysis, it is assumed that the arrival distribution is random; the delay is halved if the service rate goes from random to regular. If the arrival distribution remains random, the only way that delay can be further reduced is if the service distribution increasingly mirrors the arrival distribution. If the service distribution was exactly the same as the arrival distribution then there should be no delay at all until saturation flow was reached (as, for example, shown by Blunden (1971) on p. 65 for constant arrivals and constant service). This situation is reached when $J = 0$ ($K = -1$) and in this case the service distribution must exactly match the arrival distribution, that is, it must be random when looked at in isolation but exactly matching the arrival distribution when both distributions are seen together. $K = -1$ may therefore be thought of as representing a 'matched' random distribution and as the negative values of K approach $-\infty$ the 'matched' randomness approaches regularity in exactly the same way as the corresponding positive values of K go from normal randomness to regularity.

This form of operation is quite feasible in a road traffic situation. Its only application in a real queueing situation would occur if a server was truly Parkinsonian ('the work expands to fill the time available' — a proposition in 'Parkinson's Law'), i.e. if his rate of service depended upon the pressure of customers.

APPLICATIONS

The principal advantage of the relationship is that the delay parameter J may be regarded as a characteristic of the road or type of road so that each road type can have a different relationship between flow and delay at the same degree of saturation.

It must be remembered, however, that the relationship is derived from simple queueing theory and that a basic assumption is therefore that steady-state conditions exist. This is plainly not true in most traffic situations and the result is that delays are somewhat less than they would be if a steady state existed. Furthermore, in seeking to calibrate the parameters, measurement of the whole of the delay associated with the facility must be attempted — it may be quite easy otherwise to be measuring only service time.

Because of the above effects, a calibration using real data will tend to yield a saturation flow rate which is higher than known absolute capacity. The higher the value of J and the shorter the duration of steady-state conditions (e.g. the shorter the peak period) the more this will be so. It is suggested therefore that the function be arbitrarily cut off at a flow rate somewhat less than S, depending, as indicated above, on the value of J and the duration of the peak. To do so would also eliminate the potential modelling embarrassment of infinite travel time.

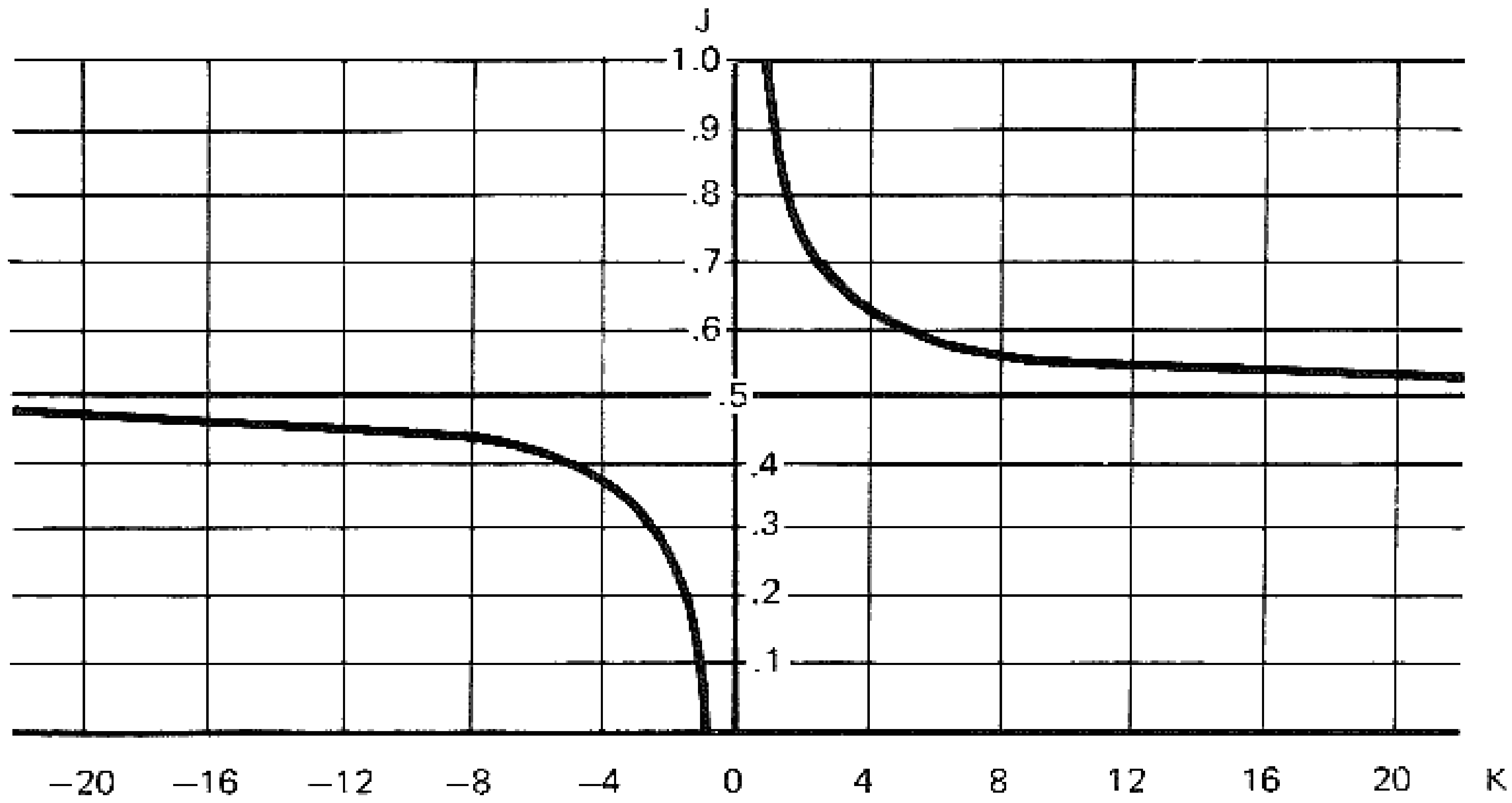

Fig. 2 — Relationship between delay parameter J and the Erlang number K of the service distribution

CONCLUSION

The flow-travel time relationship developed in Davidson (1966) is firmly based on queueing theory, and the delay parameter *J* is closely related to the Erlang number of the service distribution. Useful values of *J* correspond to negative Erlang numbers, but by developing the idea of a service facility which has a distribution which, to varying degrees, is 'matched' to the arrival distribution, positive and negative Erlang numbers of the same absolute value can be shown to reflect the same degree of randomness. Such a concept is clearly realistic in traffic modelling and it may have applications in other queueing systems.

REFERENCES

BLUNDEN, W.R. (1971). *The Land Use Transport System.* (Pergamon: Oxford.)
DAVIDSON, K.B. (1966). A flow travel time relationship for use in transportation planning. Proc. 3rd ARRB Conf. 3(1), pp. 183-94.
DREW, D.R. (1968). *Traffic Flow Theory and Control.* (McGraw Hill: New York.)

K.B. DAVIDSON,
B.E., M.E., F.C.I.T.

Ken Davidson joined the Queensland Main Roads Department in 1962 after graduating in Civil Engineering at Queensland University. He later gained an M.E. in the School of Traffic Engineering at the University of New South Wales. In 1968, he was appointed Lecturer in Civil Engineering at Queensland University with the task of establishing courses in transportation. He was promoted to Senior Lecturer in 1970. From 1974 to 1976 he was Assistant Secretary, Transportation, in the then federal Department of Urban and Regional Development and since then has been with the National Capital Development Commission, first as Director of Transportation Planning and now as Director of Engineering Planning, a role which encompasses transportation, hydraulics and environment. His major research interest is in the use of accessibility models as a basis for estimating the land use effects of transport changes and the efficiency of the transport-land use system. He has recently submitted a Ph.D. thesis on that topic.