A Concrete Approach to Mathematical Modelling

Mike Mesterton-Gibbons

Department of Mathematics
Florida State University, Tallahassee

A Wiley-Interscience Publication
JOHN WILEY & SONS, INC.
New York • Chichester • Brisbane • Toronto • Singapore

6.7 When Does a T-junction Require a Left-turn Lane?

Consider a T-junction on a narrow two-lane road.[2] As drawn in Fig. 6.5, the major road runs from north to south, with the minor road leaving to the west. There is no left-turn lane on the major road at the junction. Thus if southbound traffic is heavy, then cars turning left may have difficulty entering the minor road and may cause delays to northbound traffic. When is this problem serious enough to warrant building a left-turn lane? In this section, we will attempt to answer that question.

Let λ denote the mean traffic flow from north to south, expressed as number of cars per second. Let μ denote the corresponding flow from south to north. Let α denote the probability that one of these northward bound cars will turn left at the junction; hence $1-\alpha$ is the probability that it will continue forward. We will suppose that northbound drivers first cross the bar marked "Decision" in Fig. 6.5, then decide to turn left (with probability α) or continue forward (with probability $1-\alpha$). We will call this bar the *decision bar*. In practice, the decision bar might be 20 or 30 yards

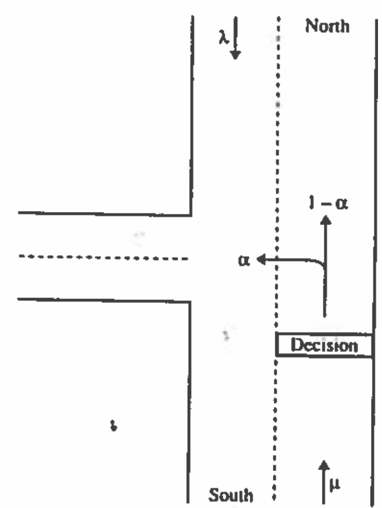


Fig. 6.5 Plan of T-junction.

[2]You may wish to review Section 5.4 before proceeding.

south of the junction; but it will be convenient to imagine that drivers make this decision at the last possible moment, so that the decision bar is about a car-length away from the turning point (this is, anyway, how almost everyone in Tallahassee seems to drive). Let T be the minimum time gap in the southbound traffic flow that will allow a north-bound car to turn left (assumed the same for all cars). Because T is a time, as opposed to a distance, it is immaterial whether a southbound car turns right or goes forward at the junction. We will assume throughout that drivers joining the main road at the junction (from the west) interrupt neither the northbound nor southbound flow, but rather enter the main stream of traffic whenever it is completely clear for them to do so. We will also assume that drivers act responsibly, so that southbound cars are not impeded by drivers turning left when such a move would be dangerous. It is then reasonable to regard the approach of southbound cars to the junction as unrestricted; and hence, as for the crossing control problem of Section 5.4, to model it as a Poisson process, with an average of λt arrivals in an interval of length t seconds. Let the random variable S denote the time gap in seconds between southbound cars. Then, from (5.52), the distribution of S is given by

$$\text{Prob}(S > t) = e^{-\lambda t}, \qquad 0 \le t < \infty. \tag{6.65}$$

Similarly, if the random variable N denotes the time gap in seconds between northbound cars whenever northbound traffic is moving freely, then it seems reasonable to assume that

$$\text{Prob}(N > t) = e^{-\mu t}, \qquad 0 \le t < \infty, \tag{6.66}$$

with associated probability density function

$$f_N(t) = \frac{d}{dt}\{\text{Prob}(N \le t)\} = \mu e^{-\mu t}, \qquad 0 \le t < \infty. \tag{6.67}$$

Let us define a *tailback* to be a group of one or more northbound cars that have been brought to a halt behind a car turning left at the junction. Because a tailback may form, northbound traffic need not move freely at the junction itself, but it will still move freely behind the tailback! Thus (6.66) governs times between northbound arrivals at the tailback, if there is a tailback; or at the junction, if not. From our comments above, it follows that the first car of any tailback will be waiting just behind the decision bar.

Now suppose that a northbound car has reached the T-junction and is waiting to turn left. In Fig. 6.5, it has just crossed the decision bar. Let the random variables $S_1, S_2, S_3, \ldots$ denote the sequence of time gaps in the southbound flow which confronts this driver; and let U_j be the event that she enters the minor road during the jth such gap. Then, because the gaps are independent, her predicament is identical to that of a crossing

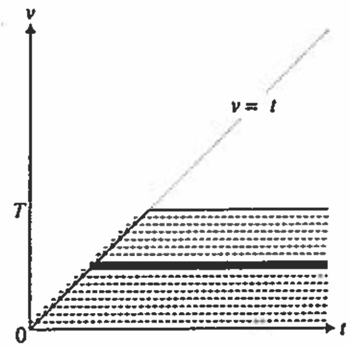


Fig. 6.6 The region in the t–v plane for which $v \le T$ and $t > v$. This infinite trapezoid is sliced into infinitesimally thin strips, parallel to the t-axis and extending from $t = v$ to $t \to \infty$. For fixed v, the inner integration with respect to t in (6.70) is performed along a strip, as indicated schematically. The outer integration with respect to v then sums over all such strips.

pedestrian, and we deduce immediately from (5.53) that

$$\begin{aligned}\text{Prob}(U_j) &= \text{Prob}(S_1 \le T, S_2 \le T, \ldots, S_{j-1} \le T, S_j > T)\\ &= (1 - e^{-\lambda T})^{j-1}\, e^{-\lambda T}.\end{aligned}$$

Let W denote the event that no tailback forms behind our waiting driver, and let q_i denote the probability that southbound car i arrives before the next northbound car, conditional upon southbound cars $1, 2, \ldots, i-1$ having already arrived; and all conditional upon time gaps not exceeding T in the southbound flow. Then, because a Poisson process is memoryless, $q_i = \text{Prob}(N > S \mid S \le T)$ is independent of i. Let A denote the hatched trapezoidal region in Figure 6.6, and note that S has probability density function

$$f_S(v) = \lambda e^{-\lambda v} \tag{6.69}$$

by analogy with (6.67). Then, because (A.11) implies that $\text{Prob}(N > S \mid S \le T)$ equals $\text{Prob}(N > S \textit{ and } S \le T)$ divided by $\text{Prob}(S \le T)$, we have

$$\begin{aligned} q_i &= \frac{1}{1 - e^{-\lambda T}} \int_A\!\!\int f_N(t)\cdot f_S(v)\, dv\, dt \\ &= \frac{1}{1 - e^{-\lambda T}} \int_0^T \lambda e^{-\lambda v} \int_v^\infty \mu e^{-\mu t}\, dt\, dv \\ &= \frac{\lambda}{\mu + \lambda}\left(\frac{1 - e^{-(\mu+\lambda)T}}{1 - e^{-\lambda T}}\right),\end{aligned} \tag{6.70}$$

on using the results in Exercise 6.10.[3]

[3] I am very grateful to Peter Watterson for pointing out that I neglected to condition on gaps no greater than T in earlier editions of this book, in effect assuming that $T \to \infty$ in (6.70)–(6.74).

Now, *given* that our driver enters the minor road during gap j, the probability that no tailback forms behind her is simply the probability that the $(j-1)$th southbound car passes her before the next northbound car arrives at the junction. In other words,

$$\text{Prob}(W \mid U_j) = q_{j-1}q_{j-2}\cdots q_2 q_1 = \left(\frac{\lambda\{1-e^{-(\mu+\lambda)T}\}}{\{\lambda+\mu\}\{1-e^{-\lambda T}\}}\right)^{j-1}, \tag{6.71}$$

on using (6.70). But the infinite sequence of events $\{U_1, U_2, U_3, \ldots\}$ exhausts all possibilities for entering the minor road. Hence from (A.13),

$$\text{Prob}(W) = \sum_{j=1}^{\infty} \text{Prob}(W \mid U_j)\cdot \text{Prob}(U_j) \tag{6.72}$$

$$= e^{-\lambda T}\sum_{j=1}^{\infty}\left\{\frac{\lambda\{1-e^{-(\mu+\lambda)T}\}}{\mu+\lambda}\right\}^{j-1}, \tag{6.73}$$

on using (6.68) and (6.71). Let θ denote the probability that a tailback does form behind a left-turning driver. Then (Exercise 6.10)

$$\theta = 1 - \text{Prob}(W) = 1 - \frac{(\mu+\lambda)e^{-\lambda T}}{\mu + \lambda e^{-(\mu+\lambda)T}}. \tag{6.74}$$

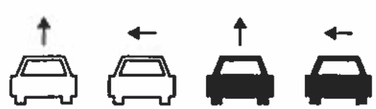
Note that because the formation of a tailback is governed by two memoryless Poisson processes, θ is also the probability that a tailback containing two or more cars will form behind a tailback containing precisely one.

Let us now define four traffic configurations, labelled $i = 0, 1, 2, 3$, as shown in Fig. 6.7:

Configuration	Last car to drive over decision bar	Status of northbound lane
$i = 0$	has continued forward	no tailback
$i = 1$	is waiting to turn left	no tailback
$i = 2$	has continued forward	tailback
$i = 3$	is waiting to turn left	tailback

(6.75)

Because transitions between these configurations are governed by two memoryless Poisson processes, we may regard them as the states of a Markov chain. Notice that each transition corresponds to a physical event. States 0 or 1 may be entered only when a car crosses the decision bar,



0 1 2 3

Fig. 6.7 Possible traffic configurations at a T-junction.

states 2 or 3 either when a car crosses the decision bar or when one joins the tailback. Thus time is measured nonlinearly. Therefore, let $P = [p_{ij}]$ be the transition matrix; i.e., let p_{ij} be the probability that j is the next configuration if the current one is i, $0 \le i, j \le 3$. Then, because state 2 cannot be entered from states 0 or 1, and because state 3 cannot be entered from state 0, we have $p_{02} = p_{03} = p_{12} = 0$.

Suppose that the junction is in state 0. Then northbound traffic is moving freely, the last car over the decision bar has gone forward, and the next state will be 0 or 1 according to whether the next car goes forward or turns left, after passing the decision bar. Hence $p_{00} = 1-\alpha$, $p_{01} = \alpha$.

Suppose that the junction is in state 1. Thus a car is waiting to turn, and traffic behind it is still moving freely. If it fails to turn before the arrival of the next northbound car, however, then a tailback will form and the junction will enter state 3. Hence $p_{13} = \theta$. Otherwise, the car succeeds in turning before the next arrival, no tailback forms and the next state will be 0 or 1 according to whether the next northbound car goes forward or turns. We have thus determined p_{1j}, $0 \le j \le 3$.

Next suppose that the junction is in either state 2 or state 3; let r denote the probability of exactly one car in the tailback (conditional on at least one car) just before the next transition; and let ρ denote the corresponding probability just after the previous transition. As noted above, the probability that the tailback does not lengthen during the interim is $1-\theta$, and so $r = (1-\theta)\rho$. But $\rho \le 1$. Hence

$$\rho \le 1 - \theta. \tag{6.76}$$

This inequality will yield an answer to the question we set out to address—in effect, a sufficient condition for a left-turn lane to be necessary!

Now suppose that the junction is in state 2 for sure. Then a car has just gone forward, and the junction will enter state 0 or state 1 with probability r, state 2 or state 3 with probability $1-r$. If it enters state 0 or state 1, then it will enter state 1 with probability α and state 0 with probability $1-\alpha$. Similarly for state 2 or state 3. This is sufficient to determine p_{2j}, $0 \le j \le 3$.

Suppose, finally, that the junction is in state 3. Then it will enter state 0 or state 1 only if there is precisely one car in the tailback and the waiting car succeeds in turning before the next arrival. Hence $p_{30} + p_{31} = r(1-\theta)$. If the junction enters state 0 or state 1, then it will enter 1 with probability α and 0 with probability $1-\alpha$. Otherwise it will enter state 2 or state 3. It will enter state 2 if the waiting car succeeds in turning before the next arrival, *and* the tailback contains more than one car, *and* the next car goes forward. Thus

$$p_{32} = (1-\theta)(1-r)(1-\alpha). \tag{6.77}$$

It will remain in state 3 if *either* the waiting car fails to turn before a new northbound car arrives, regardless of how many cars are in the tailback; *or* if the tailback contains more than one car, *and* the waiting car turns before an arrival, *and* the next car then waits to turn left. Thus

$$p_{33} = \theta + (1 - r)(1 - \theta)\alpha. \tag{6.78}$$

Combining our results, we find that the transition matrix for the Markov chain is

$$\mathbf{P} = \begin{bmatrix} 1-\alpha & \alpha & 0 & 0 \\ (1-\theta)(1-\alpha) & (1-\theta)\alpha & 0 & \theta \\ r(1-\alpha) & r\alpha & (1-r)(1-\alpha) & (1-r)\alpha \\ r(1-\theta)(1-\alpha) & r(1-\theta)\alpha & (1-r)(1-\theta)(1-\alpha) & \theta+(1-r)(1-\theta)\alpha \end{bmatrix}. \tag{6.79}$$

Inspection of (6.79) reveals that it is the transition matrix of an irreducible ergodic Markov chain; see the end of Section 6.1. Hence, irrespective of the value of $\pi(0)$, the T-junction's distribution vector $\pi(n)$ will converge to a unique stationary distribution $\pi(\infty)$. On using (6.10) this stationary distribution is found to be (Exercise 6.10)

$$\pi(\infty) = \frac{1}{\Gamma}\begin{bmatrix} r(1-\theta)(1-\alpha) \\ r(1-\theta)\alpha \\ \alpha\theta(1-\theta)(1-\alpha)(1-r) \\ \alpha\theta\{(1-\alpha)r+\alpha\} \end{bmatrix}, \tag{6.80}$$

where

$$\Gamma \equiv \{1 - \theta(1-\alpha)\}r(1-\alpha\theta) + \alpha\theta\}. \tag{6.81}$$

Note that the numerator of $\pi_2(\infty)$ is the product of five numbers with magnitude less than one; whereas for $j = 0, 1, 3$, the numerator of $\pi_j(\infty)$ is the product of only three numbers with magnitude less than one. Thus 2 is always the least probable configuration, as confirmed by numerical results.

From (6.80), we deduce that the stationary proportion of blocked configurations is $B = \pi_2(\infty) + \pi_3(\infty) = \alpha\theta\{1 - (1-\alpha)(1-r)\theta\}/\{(\alpha\theta + (1-\alpha\theta)r)(1 - (1-\alpha)\theta)\}$, which decreases with r and hence is least for $r = 1 - \theta$, by (6.76): if B is too large for $r = 1 - \theta$, then B is too large, and thus we arrive at our sufficient condition for requiring a left-turn lane. Let $\Delta(\lambda, \mu, \alpha, T)$ denote the lower bound on B. Then, on setting $r = 1 - \theta$ in B and writing $\theta(\lambda, \mu, T)$ in place of (6.74), we obtain

$$\Delta(\lambda, \mu, \alpha, T) = \cfrac{1}{1 + \cfrac{\{1-\theta(\lambda,\mu,T)\}^2}{\alpha\theta(\lambda,\mu,T)\{1-\{1-\alpha\}\theta(\lambda,\mu,T)^2\}}}. \tag{6.82}$$

Let δ_c denote the maximum proportion of traffic configurations in which a tailback is deemed acceptable. Of course, defining δ_c may require a value judgment on the part of a traffic engineer or community activist, depending on the purpose for which the model is being used, but we shall suppose that somehow δ_c is known. Then a left-turn lane is necessary if

$$\Delta(\lambda, \mu, \alpha, T) > \delta_c. \tag{6.83}$$

If $\Delta > \delta_c$ but there is no left-turn lane, then drivers may try to reduce Δ to δ_c by lowering their perception of the value of T, thus increasing the risk of collisions at the junction.

The junction of High Road with Hartsfield Road in Tallahassee was a T-junction like the one we have described above. High Road, the main road, runs north and south; Hartsfield Road, the minor road, runs east and west. High Road is not used excessively by large vehicles, and those that do use it almost never turn left into Hartsfield. Thus assuming T to be the same for all cars is a better approximation to reality than it might be at many other junctions. Unfortunately, the other major source of such variation, namely, the personality of the driver, cannot be so easily eliminated.

On April 10, 1986, a normal working day, the City of Tallahassee Traffic Engineering Department enumerated peak-hour traffic movements at this junction. In the morning, 437 vehicles traversed the junction in the southbound direction and 154 in the northbound direction, of which 41 turned left into Hartsfield. For the morning rush hour, it therefore seems reasonable to take $\lambda = 437/3600 = 0.121$, $\mu = 154/3600 = 0.0428$ and $\alpha = 41/154 = 0.266$. In the evening, 383 vehicles traversed the junction in the southbound direction and 337 in the northbound direction, of which 133 turned left into Hartsfield. For the evening rush hour, it therefore seems reasonable to take $\lambda = 383/3600 = 0.106$, $\mu = 337/3600 = 0.0936$ and $\alpha = 133/337 = 0.395$. For these morning and evening rush hours, Δ has been plotted in Fig. 6.8 as a function of T. For $T = 6$, for example, the model predicts that the northbound lane may have been brought to a halt behind cars turning left for only 3.2% of traffic configurations in the morning rush hour, but for more than 8.1% of traffic configurations in the evening rush hour. Unfortunately, data for T were not available, though it seems likely that T would be less than 10 seconds (which is why the graph has been truncated at that value).

The T-junction model we developed in this section represents a considerable improvement upon Liu and Yao (1977); see Exercise 5.12. Our model allows arbitrary values of the four most relevant traffic parameters, namely, southbound traffic flow λ, northbound traffic flow μ, left turn probability α, and minimum crossing time T. By contrast, Liu and Yao's model allows arbitrary values of only two parameters, namely, the left turn probability α, and the probability that a southbound vehicle will turn right. The latter is one of the least relevant parameters. More fundamentally, however, Liu and Yao chose traffic configurations in which a northbound vehicle can turn left only after a southbound vehicle has turned right. This is clearly unrealistic, because left-turning vehicles could not take advantage of breaks

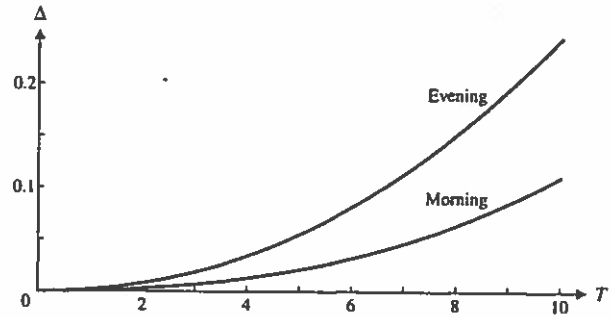


Fig. 6.8 Proportion of traffic configurations, $\Delta(\lambda, \mu, \alpha, T)$, as a function of minimum crossing time T, in which flow along a narrow two-lane major road develops a tailback in one direction behind cars turning left onto a minor road at a T-junction. The values of the parameters λ, μ, and α are fixed for the morning and evening rush hours at a T-junction in Tallahassee, Florida.

in the oncoming flow and would block the northbound lane permanently if all southbound traffic were continuing forward! (In the morning rush hour described above, only 53 of the 437 southbound cars, i.e., only 12%, turned right.)

I devised this model toward the end of 1986. The junction of High with Hartsfield in Tallahassee has since been redesigned.

Exercises

6.1 By solving (6.8) and (6.9), with 2 replacing 5 in (6.9) and S given by (5.76), verify directly that (6.20) is the unique stationary distribution of the linear time elevator model introduced in Section 5.6.

6.2 Verify directly that (5.92) is the unique stationary distribution of the bird chain considered in Section 5.7. Confirm that, when the distribution is stationary, the mean number of birds in the North end of the cage is 2, as suggested by (5.91).

6.3 Show that the modified bird chain in Exercise 5.8 has the same stationary distribution as in Exercise 6.2. This further illustrates that stationary distributions exist which cannot be found by the indirect method, at least as described at the beginning of this chapter. The relationship of this stationary distribution to the pattern discovered in Exercise 5.9 is explained, for example, by Hoel et. al. (1972, pp. 78–79). It is interesting mathematically but has little to do with the subject of this book.

6.4 Verify (6.31). *Hint*: First differentiate

$$\sum_{j=0}^{N} \rho^j = \frac{1-\rho^{N+1}}{1-\rho}$$

with respect to ρ.

6.5 Verify (6.35) and (6.36). *Hint*: Adapt Exercise 6.4.

6.6 Derive (6.40).

6.7 (i) Let the random variables A and I denote, respectively, the number of active servers and the number of idle servers in a stationary supermarket. Then

$$I = \begin{cases} 0, & \text{if } A \geq k; \\ k - A, & \text{if } A \leq k-1. \end{cases}$$

Show that the expected number of idle servers is $E[I] = k - \rho$.
(ii) Deduce that the expected number of active servers agrees with intuition. *Hint*: Assume N infinite and use the fact that $jC_j = \rho C_{j-1}$ if $j \leq k-1$, whereas $kC_j = \rho C_{j-1}$ if $j \geq k$.

***6.8** A single worker is responsible for the maintenance of N machines. The time she takes to repair a machine is exponentially distributed with mean s, and the time until failure of a machine is exponentially distributed with mean μ. Model this as a stationary birth and death process in which the birth rate b_j actually depends upon j. Show that the expected number of machines in need of repair is

$$L = N - (1 - \pi_0)\frac{\mu}{s},$$

where

$$\pi_0 = \frac{1}{\sum_{j=0}^{N} \frac{N!}{(N-j)!}\left(\frac{s}{\mu}\right)^j}.$$

Hints: It may be easier to think of the birth rate as being a function of $N - j$, the number of machines in working condition. You may find the mathematics easier if you observe that $C_j = s(N-j+1)C_{j-1}/\mu$. Note that the model would apply as well to a team of repair workers if every member were needed to fix each machine.

6.9 (i) Obtain (6.48b) by successive application of (5.40). *Hint*: Remember that (6.27) is valid only for $j \geq 1$; $d_0 = 0$.
(ii) Verify that the integrand in (6.48b) is the derivative of (6.48a).

an unsteady heat flow governed by (11.5). We must therefore solve (11.5) subject to the boundary conditions

$$\phi(0,t) = T_1, \tag{11.10}$$

$$\phi(D,t) = T_3, \tag{11.11}$$

and the initial condition

$$\phi(x,0) = T_s(x). \tag{11.12}$$

The solution of (11.5) subject to (11.10)–(11.12) is

$$\phi(x,t) = T_1 + (T_3 - T_1)\frac{x}{D} + \frac{2(T_3 - T_2)}{\pi}\sum_{n=1}^{\infty}\frac{(-1)^n}{n}e^{-\kappa n^2\pi^2 t/D^2}\sin\left(\frac{n\pi x}{D}\right). \tag{11.13}$$

You can easily see that (11.13) satisfies (11.10) and (11.11), and you can verify by differentiation that (11.13) satisfies (11.5) because every term in the infinite series is itself a solution. The fact that (11.13) also satisfies (11.12) depends upon a result from the theory of Fourier series, namely, that for $0 < x < D$,

$$\frac{\pi x}{2D} = \sum_{n=1}^{\infty}\frac{(-1)^{n+1}}{n}\sin\left(\frac{n\pi x}{D}\right). \tag{11.14}$$

Given this result, you can verify for yourself that the initial condition is satisfied. A proof that the series (11.13) converges will be found in any substantial text on Fourier analysis and does not belong in a text on mathematical modelling; see, for example, Chapter 7 of Carslaw (1930).

Taking the limit of (11.13) as $t \to \infty$ we see that the temperature approaches a new steady state:

$$T_s^{new}(x) = T_1 + (T_3 - T_1)\frac{x}{D}, \tag{11.15}$$

in perfect agreement with (2.50). Moreover, because the terms in the infinite series are dominated by the first term and decrease so rapidly, we see that the glass has virtually evolved toward its new equilibrium after time t_1, where t_1 is a low multiple of $D^2/\kappa\pi^2$.

The terms in (11.13) that fade away as $t \to \infty$ are known as *transients*. In this section, the steady state that then emerges is a static one, but if the external temperature varied cyclically, then we might expect the steady state to be a limit cycle. Clearly, when seeking such cycles directly, it is no longer permissible to set $\partial/\partial t = 0$ and solve (11.8); rather, we must seek an appropriate solution of the unsteady equation (11.5). An example of this emerges in Exercise 11.1, which you should now attempt.

11.2 How Does Traffic Move after the Train Has Gone By? A First Look

We have already seen two models of traffic flow in a single lane. One was an equilibrium model (Section 2.3) in which traffic density was independent of both time and distance along the road.[2] The other was a nonequilibrium model (Section 1.13) in which a driver responded only to the car in front of him. Now, the latter is probably perfectly adequate in a tunnel, because it is difficult to see the brake lights of the car in front of the car in front of you (or at least that was true in the old days, before manufacturers began to affix additional brake lights below the rear window or on the roof). On the open road, however, we would expect a driver to look ahead and react to changes in traffic conditions, such as variations of traffic density with distance along the road. But if traffic density depends upon distance, and in such a way that drivers react to conditions far ahead of them, then neither of our existing models will suffice. We are obliged to invent a new one.

An adaptation suggests itself immediately if we think of cars as pollutants and roads as rivers of traffic. As we said in Chapter 9, the lake core model of Section 9.4 is essentially a river model, and we have already adapted it once to describe population "flow." To adapt the same model for our present purpose, all we need do is to regard traffic density as a concentration of contaminant—but per unit length, rather than per unit volume. As shown in Section 9.6, this is accomplished simply by setting A, the cross-sectional area of the "river," equal to 1.

Accordingly, let $\phi(z,t)$ denote the density of cars, or the number of cars per unit roadlength, on a single lane of highway at station z at time t. Suppose that the road between $z = z_1$ and $z = z_2$ ($> z_1$) has no entrances or exits and that passing is not allowed. Then the only way cars can enter the region $z_1 \le z \le z_2$ is from "upstream," i.e., by crossing $z = z_1$; and the only way they can leave is by going "downstream," i.e., by crossing $z = z_2$. Accordingly, we can replace (9.40) by

Rate of change of traffic = Traffic inflow − Traffic outflow

$$\frac{\partial}{\partial t}\int_{z_1}^{z_2}\phi(z,t)dz \quad = \quad F(z_1,t) \quad - \quad F(z_2,t), \tag{11.16}$$

where $F(z,t)$ is the flux of traffic, or the number of cars per unit time passing station z at time t. This leads, in the same way as (9.40) yielded (9.44), to

$$\frac{\partial\phi}{\partial t} + \frac{\partial F}{\partial z} = 0. \tag{11.17}$$

We must now choose F to mimic the way in which drivers behave. Clearly, this will depend upon the kind of road and the traffic conditions.

[2] You may wish to review Sections 2.3, 3.5, and 9.4–9.6 before proceeding.

Suppose that the road has a single lane in each direction, one going north and one going south, and is traversed from west to east, at $z = 0$, by a railroad. We'll consider only the northbound traffic, in which direction z increases. Now imagine that a long, slow train has just passed by. There is therefore a long line of cars (bumper to bumper) in $z < 0$, and the northbound lane in $z > 0$ is completely clear. Let the gates finally open at time $t = 0$. If we are interested solely in the vicinity of the crossing, then it is legitimate to pretend that the road extends to infinity on either side of $z = 0$. Hence the initial conditions are

$$\phi(z,0) = \begin{cases} 0, & \text{if } z > 0; \\ \phi_{max}, & \text{if } z < 0, \end{cases} \tag{11.18}$$

where $\phi_{max} = 1/\{\text{CAR LENGTH}\}$. Moreover, cars cannot pass after the gates have opened because of the queue in the opposite direction; and we will simply assume that there are no entrances or exits near the crossing. The assumptions embodied in (11.16) are thus satisfied.

The initial density profile (11.18) is sketched in Fig. 11.2 as the dashed curve, together with some solid curves that suggest the shape that the profile should take at later times, as the traffic begins to flow. Doesn't this diagram remind you of Fig. 9.7? It is the mirror image of the diagram that we would have obtained in Section 9.5 if the lake pollution had been allowed to diffuse into $y < 0$, and it suggests the hypothesis

$$F = -\Gamma \frac{\partial \phi}{\partial z}, \tag{11.19}$$

where Γ (> 0) is a constant. From (11.17) and (11.19) we then obtain the partial differential equation

$$\frac{\partial \phi}{\partial t} = \Gamma \frac{\partial^2 \phi}{\partial z^2}. \tag{11.20}$$

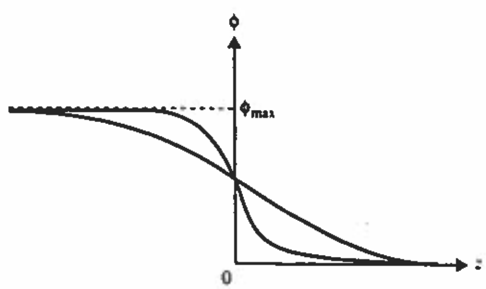


Fig. 11.2 Evolution of (11.18) according to (11.20). The dashed profile is (11.18). The solid curves are for two later times.

You can readily show (Exercise 11.2) that (11.18) and (11.20) are both satisfied by

$$\phi(z,t) = \frac{\phi_{max}}{\sqrt{\pi}} \int_{z/2\sqrt{\Gamma t}}^{\infty} e^{-\xi^2} d\xi, \qquad -\infty < z < \infty, 0 < t < \infty. \tag{11.21}$$

With the help of Exercise 11.2, notice that $\phi(z,t) \to \phi_{max}/2$ as $t \to \infty$. Moreover, by repeating the argument that led to (9.69), you can show that the road at station z would be within 5% of this limiting density $\phi_{max}/2$ after time $127z^2/\Gamma$—*if* the model were correct.

Is the model correct? For a short time after the gates have opened, in the immediate vicinity of the crossing (11.21) may well lead to density profiles with roughly the right shape (see Fig. 11.4). Indeed, were this not so, we would never have adopted (11.19) in the first place. The model also predicts, however, that

$$F(z,t) = -\Gamma \frac{\partial \phi}{\partial z} = \frac{\Gamma \phi_{max}}{2\sqrt{\pi \Gamma t}} e^{-z^2/4\Gamma t} \tag{11.22}$$

tends to zero as $t \to \infty$, and on replacing x by ϕ in Exercise 2.5, we have in general that

Number of cars passing per unit of time	=	Number of cars per unit length of road	×	Length of road covered per unit of time	(11.23)
$F(z,t)$	=	$\phi(z,t)$	×	$v(z,t)$,	

where v is the velocity of traffic flow. Note, from (9.37) with $A = 1$, that this is just the traffic flow analogue of (9.38). Thus $v(z,\infty) = F(z,\infty)/\phi(z,\infty) = 2F(z,\infty)/\phi_{max} = 0$; i.e., the model predicts that traffic slows to a halt for $\phi < \phi_{max}$. We know that this is not correct. Our model must therefore be rejected.

Knowing this in advance, of course, I could have omitted this section; but I wanted to illustrate yet again the importance of trial and error in constructing mathematical models. Yet the moral of this section is not, as you might think, that models must sometimes be rejected. Rather, it's that successful models are often the results of failures, as Section 11.4 will demonstrate. Meanwhile, try Exercise 11.3.

11.3 How Does Traffic Move after the Train Has Gone By? A Second Look

In this section, we will take a different approach to modelling the effect of a railroad crossing on traffic flow. Recall that in our road tunnel model of Chapter 2 we found an equilibrium velocity v for any traffic density $\phi(z,t)$ in the range $0 \le \phi \le \phi_{max}$, where t denotes time and z distance. Because the flow was in equilibrium, $\phi(z,t) = \phi$ was a constant. But the value of that constant was arbitrary (within the range $0 \le \phi \le \phi_{max}$). For each such value, the velocity was given by a relationship of the form

$$v = v(\phi) \tag{11.24}$$

and the traffic flow by

$$F = F(\phi) = \phi v(\phi). \tag{11.25}$$

Now, let's *extend* these ideas by assuming that relationships (11.24) and (11.25) are valid even when the traffic flow is not in equilibrium, i.e., by assuming that

$$v(z,t) = v(\phi(z,t)), \qquad F(z,t) = F(\phi(z,t)). \tag{11.26}$$

The rationale behind this assumption is that if density changes are not too violent, then traffic conditions will adjust continually, as though passing quickly through a series of stable equilibria, in such a way as to maintain "dynamic equilibrium" between F and $\phi v(\phi)$. Be sure you appreciate that (11.26) is an assumption and represents a considerable extension of (11.24) and (11.25). As usual, the quality of this assumption will be determined by how well the model's predictions accord with reality.

Let us now combine the nondifferential dynamic model (11.26) with the model (11.17) for traffic flow in a single lane, which we developed in the previous section. We obtain

$$\frac{\partial \phi}{\partial t} + \frac{\partial F}{\partial z} = \frac{\partial \phi}{\partial t} + F'(\phi)\frac{\partial \phi}{\partial z} = 0, \tag{11.27}$$

where, as usual, a prime denotes differentiation with respect to argument. Hence, if (11.26) is a good assumption, then the solution of (11.27), subject to the initial conditions (11.18), will provide an adequate description of how traffic flows in the immediate aftermath of a long, slow train. In principle, $v(\phi)$ can have any form that gives $F(\phi)$ the overall shape of the fundamental diagram of Exercise 2.5, but we shall use the form that gave such good agreement with data in Section 4.6. On rearranging (4.26), therefore, we have

$$v(\phi) = \begin{cases} v_{max}, & \text{if } 0 \le \dfrac{\phi}{\phi_{max}} \le e^{-v_{max}/\lambda}; \\[2ex] \lambda \ln\left(\dfrac{\phi_{max}}{\phi}\right), & \text{if } e^{-v_{max}/\lambda} \le \dfrac{\phi}{\phi_{max}} \le 1; \end{cases} \tag{11.28}$$

where λ is the velocity that maximizes the traffic flux. From (11.25) and (11.28) we easily deduce that

$$F'(\phi) = \begin{cases} v_{max}, & \text{if } 0 \le \dfrac{\phi}{\phi_{max}} \le e^{-v_{max}/\lambda}; & (11.29\text{a}) \\[2ex] \lambda\left\{\ln\left(\dfrac{\phi_{max}}{\phi}\right) - 1\right\}, & \text{if } e^{-v_{max}/\lambda} \le \dfrac{\phi}{\phi_{max}} \le 1. & (11.29\text{b}) \end{cases}$$

The solution of (11.27) subject to (11.18) and (11.29) is (Exercise 11.4)

$$\begin{aligned} \frac{\phi(z,t)}{\phi_{max}} &= 1, && \text{if } -\infty < z \le -\lambda t; \\ &= e^{-(z/\lambda t + 1)}, && \text{if } -\lambda t \le z \le (v_{max} - \lambda)t; \\ &= e^{-v_{max}/\lambda}, && \text{if } (v_{max} - \lambda)t \le z \le v_{max} t; \\ &= 0, && \text{if } v_{max} t < z < \infty. \end{aligned} \tag{11.30}$$

The velocities corresponding to each of the four regions in the z-t plane defined by (11.30) are given in Fig. 11.3. Note that both $\phi(z,t)$ and $v(z,t)$ are continuous except along the line $z = v_{max}t$, which corresponds to the motion of the leading vehicle into the clear road ahead of the crossing. Notice also that cars cross the rails with velocity λ, corresponding to maximum flux. Thus one easy way to measure the road's capacity would be to stand at the crossing for a minute or two after a train goes by and count the number of vehicles.

Figure 11.3 can be interpreted in either of two ways: by fixing z and letting t vary, or by fixing t and letting z vary. First, therefore, consider an observer standing at $z = z_0$. What he sees is represented in the diagram by a vertical column of dashes. The first car passes the observer at time $t = z_0/v_{max}$. From then until time $t = z_0/(v_{max} - \lambda)$, he sees a short platoon of cars, density $\phi_{max}\exp(-v_{max}/\lambda)$, moving at maximum speed v_{max}. Thereafter, the density of cars is always increasing but never exceeds

$$\phi^* \equiv \frac{\phi_{max}}{e}, \tag{11.31}$$

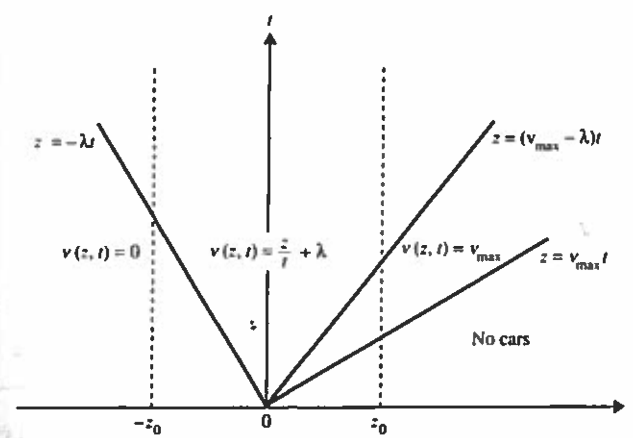


Fig. 11.3 Representation in the z-t plane of the traffic flow immediately after the crossing gates have opened.

where $e = \exp(1)$, and the speed of cars is always decreasing, but never falls below the optimal value λ. Therefore, in terms of the fundamental diagram introduced in Exercise 2.5, traffic is light ahead of the crossing. Now consider a second observer at $z = -z_0$. The car at this station is unable to move for time z_0/λ. Thus cars take a finite time to begin moving again after the train has gone by. This is in perfect agreement with reality. Thereafter, the density decreases but never falls below ϕ^*, and the velocity increases but never exceeds λ. In terms of the fundamental diagram, traffic is heavy behind the railroad crossing.

Second, we can examine the density profile at a series of fixed times as though we were photographing the entire road in the vicinity of the crossing. Three such profiles are sketched in Fig. 11.4. You can see that they are not totally unlike the profiles of the previous section, though the focal point has shifted down from 0.5 to $1/e$. But there is a crucial difference: there are now corners, where ϕ/ϕ_{max} takes the values $\exp(-v_{max}/\lambda)$ and 1. Thus our new model predicts that information—specifically the fact that the train has passed—cannot propagate more than a finite distance, either ahead of or behind the crossing, in any finite time (imagine that the observer at $z = -z_0$ does not realize that the gates are open until the car

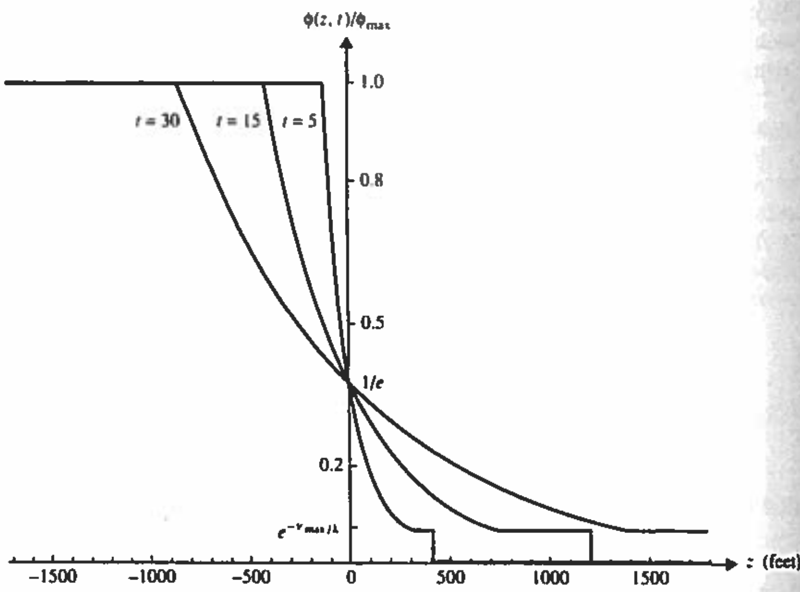


Fig. 11.4 Density profiles for 1800 feet of road on either side of a railroad crossing, within half a minute of the gates opening. Times are quoted in seconds; v_{max} is 55 m.p.h. and λ is 20 m.p.h.

at his station begins to move, and that the observer at $z = z_0$ realizes the gates are open only when the first car passes).

Notice, on using the Chain Rule and (11.27), that

$$\frac{d\phi}{dt} = 0 \tag{11.32}$$

whenever

$$\frac{dz}{dt} = F'(\phi). \tag{11.33}$$

Thus ϕ is constant along curves in the z-t plane defined by (11.33), which must therefore be straight lines. But the z-t plane is a mathematical fiction. What happens in reality is that the point on the road where a given density is found changes with time *as though* it moved (either forward or backward) with velocity $F'(\phi)$. Thus the information that the gates have opened moves back through the queue with speed λ because $F'(\phi_{max}) = -\lambda$, while the same information propagates forward into the open road with velocity v_{max} because $F'(0) = v_{max}$. The curves in the z-t plane along which a given value of ϕ is propagated are known as *characteristics* and are represented schematically in Fig. 11.5. But remember that characteristics are just convenient mathematical fictions. Don't go down to a railroad crossing in the hope of meeting a characteristic. You'll be disappointed.

It's important, in particular, to distinguish between characteristics, along which information propagates, and *trajectories*, along which cars actually move (although the two coincide for the first car in the queue, as you can easily verify). A trajectory is the solution of the ordinary differential

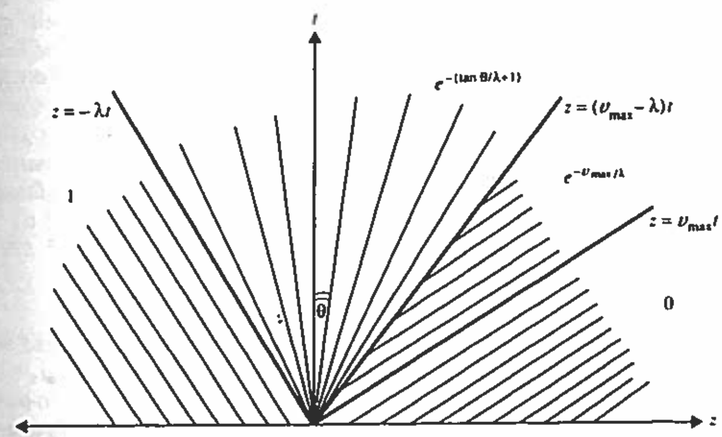


Fig. 11.5 Characteristics in the z-t plane for (11.30). Values of $\phi(z, t)/\phi_{max}$ in each of four regions are also given; $\theta = \arctan(z/t)$.

equation

$$\frac{dz}{dt} = v(z, t), \tag{11.34}$$

subject to the appropriate initial condition. But we have already found that the car at $z = -z_0$ takes time z_0/λ to begin moving. Hence, from Fig. 11.3, we must solve

$$\frac{dz}{dt} = \frac{z}{t} + \lambda, \qquad z\left(\frac{z_0}{\lambda}\right) = -z_0. \tag{11.35}$$

The solution (Exercise 11.5) is

$$z(t) = \begin{cases} \lambda t\left\{\ln\left(\dfrac{\lambda t}{z_0}\right) - 1\right\}, & \text{if } t \geq \dfrac{z_0}{\lambda}; \\ -z_0, & \text{if } t \leq \dfrac{z_0}{\lambda}. \end{cases} \tag{11.36}$$

A typical trajectory in the z-t plane is sketched in Fig. 11.6. Notice how it crosses the characteristics (dashed). Notice also that the car that queued a distance z_0 behind the crossing will ultimately drive over it at time

$$\frac{ez_0}{\lambda} \tag{11.37}$$

after the gates have opened.

How good is this model—does it fit the facts? Our remarks have indicated that there is certainly qualitative agreement with reality. But (11.37) also provides a quantitative test of our model. According to (11.37), after

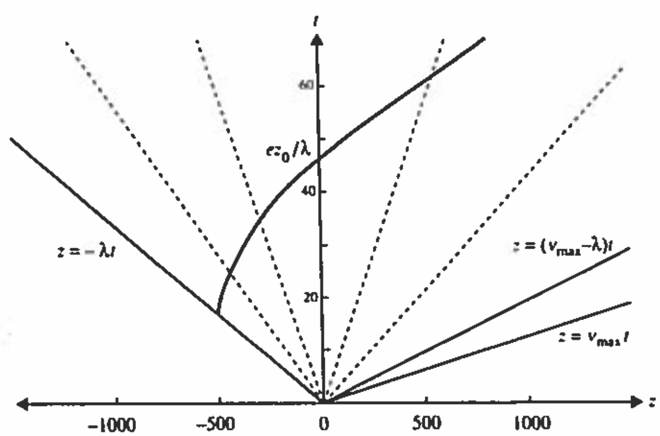


Fig. 11.6 Trajectory in the z-t plane of a car that waited in line at 500 feet from the crossing; $v_{max} = 55$ m.p.h., $\lambda = 20$ m.p.h. Distance is measured horizontally in feet; time is measured vertically in seconds.

the gates have opened, a car will take about 2.7 times as long to reach the crossing as it has to wait to get moving. Is this what actually happens? Perhaps you would like to perform the experiment! Of course, it isn't necessary to perform the experiment at a railroad crossing. You could use a stop light (provided no cars may turn off), a school bus unloading, or indeed any temporary obstruction of sufficient duration in the absence of entrances or exits.

11.4 Avoiding a Crash at the Other End. A Combination

In this section, we will attempt to model the other end of the queue of cars that forms at a railroad crossing. As usual, z will measure distance along the road, but $z = 0$ will now denote the rear of the queue instead of the head.

Suppose that the crossing is on a very lonely road, where the density is ordinarily very small. Nevertheless, the trains in this region are so long and so slow that a long line of cars always accumulates when the crossing gates are closed—so long, in fact, that if we are interested only in the rear of the queue then it is legitimate to pretend that

$$\phi(z, 0) = \phi_{max}, \qquad z > 0. \tag{11.38}$$

Now, along you come at time $t = 0$, approaching the queue at speed v_{max} but miles ahead of the following car. It is therefore legitimate to pretend that

$$\phi(z, 0) = 0, \qquad z < 0. \tag{11.39}$$

Suddenly, ahead of you, you see the queue for the crossing. What do you do? According to the model of Section 11.3, you either brake to a halt in zero distance or crash into the end of the queue!

The problem with the model of the previous section is this. It says that traffic velocity is determined by the local traffic density, according to $v = v(\phi)$, and takes no account of (spatial) *variations* of density; so that, in $z < 0$, you would drive at $v(0) = v_{max}$ right up until you hit the last car in the queue. In reality, of course, you look ahead, observe the sudden increase in traffic density, and react accordingly. We must find a way to model this.

In Section 11.2 we attempted to model traffic flow by using a flux law of the form

$$F = -\Gamma\frac{\partial\phi}{\partial z}. \tag{11.40}$$

This has the obvious advantage that it takes account of the density gradient, which is exactly what you would do if you were approaching a stationary queue of cars. On the other hand, it says that the effect of congestion ahead is to cause approaching traffic to flow backward, when, in reality, F could never be less than zero. Must we therefore abandon (11.40), or can we modify it to suit our present needs?

The answer lies in the observation that congestion ahead will decrease the flow, and dispersion ahead will increase the flow, not relative to zero but *relative to the flow that exists already*. Let us therefore combine the models of the two previous sections to obtain the hybrid flux law

$$F = \phi v(\phi) - \Gamma\frac{\partial\phi}{\partial z}, \tag{11.41}$$

which couples the effects of convection with those of diffusion. Then, from (11.17), we obtain the partial differential equation

$$\frac{\partial\phi}{\partial t} + \frac{\partial F}{\partial z} = \frac{\partial\phi}{\partial t} + \frac{\partial}{\partial z}\left(\phi v(\phi)\right) - \Gamma\frac{\partial^2\phi}{\partial z^2} = 0. \tag{11.42}$$

We can ensure that F in (11.41) is never negative by making Γ sufficiently small, which is anyhow what we would expect for a diffusivity.

You may be concerned that (11.23) and (11.41) are incompatible if $\Gamma \neq 0$. The correct interpretation of the apparent discrepancy is that (11.23) *defines* velocity; whereas (11.41) is a legitimate possibility for the flux function, in which $v(\phi)$ is an arbitrary function but not the velocity. It will be so close to the velocity if Γ is very small, however, that we prescribe it to be (11.28) in the subsequent development.

Equation (11.42) is difficult to solve in general, but just as the solution of (11.5) subject to (11.10)–(11.12) led to the steady state (11.15), so we might expect that the solution of (11.42) subject to (11.38) and (11.39) would also tend to a steady state. Moreover, because the density begins as steady except at $z = 0$, we might expect that the solution of (11.42) would differ from (11.38) and (11.39) only in a narrow transition layer near $z = 0$. As far as this layer is concerned, the regions $\phi = 0$ and $\phi = \phi_{max}$ will appear to be at infinity. Therefore, on setting $\partial/\partial t = 0$ in (11.42), it is sufficient to seek the solution $\phi = \phi(z)$ of the ordinary differential equation

$$\frac{d}{dz}(\phi v) - \Gamma\frac{d^2\phi}{dz^2} = 0 \tag{11.43}$$

subject to

$$\phi(\infty) = \phi_{max}, \qquad \phi'(\infty) = 0, \tag{11.44}$$

and

$$\phi(-\infty) = 0, \qquad \phi'(-\infty) = 0. \tag{11.45}$$

If this makes you feel uncomfortable, imagine that z is measured in miles in (11.38) and (11.39), but in inches in (11.43)–(11.45). On integrating (11.43), we obtain

$$F(z) = \phi v - \Gamma\frac{d\phi}{dz} = \text{constant}. \tag{11.46}$$

In view of (11.44) and (11.45), the constant must be zero. Thus the effect of the transition layer is to shift the road from a state in which the flux is zero by virtue of no cars to one in which the flux is zero by virtue of no movement. On using the form of $v(\phi)$ given by (11.28), we readily obtain the solution of the ordinary differential equation (11.46) in the implicit form

$$\frac{\phi(z)}{\phi_{max}} = \begin{cases} \alpha_1 e^{v_{max}z/\Gamma}, & \text{if } 0 \le \dfrac{\phi}{\phi_{max}} \le e^{-v_{max}/\lambda}, \quad (11.47a) \\ e^{-\alpha_2 e^{-\lambda z/\Gamma}}, & \text{if } e^{-v_{max}/\lambda} \le \dfrac{\phi}{\phi_{max}} \le 1, \quad (11.47b) \end{cases}$$

where α_1 and α_2 are constants (Exercise 11.7).

The constants α_1 and α_2 are determined from the continuity of ϕ and its derivative at the point where $\phi/\phi_{max} = \exp(-v_{max}/\lambda)$, i.e., at the point where $\phi(z)$ switches from (11.47a) to (11.47b). In deriving (11.47), however, we have not yet identified this point; denote it, therefore, by $z = \xi$. Because (11.47a) must satisfy (11.45), it describes the end of an open road. Because (11.47b) must satisfy (11.44), it describes the beginning of a congested road. Hence $\phi_{max}\exp(-v_{max}/\lambda)$ is a critical density, at which drivers must begin to adjust their speed in response to congestion ahead; and ξ is the value of z for which this critical density obtains. Requiring both $\phi(z)$ and $\phi'(z)$ to be continuous at $z = \xi$ yields:

$$\begin{aligned} \alpha_1 e^{v_{max}\xi/\Gamma} &= e^{-\alpha_2 e^{-\lambda\xi/\Gamma}} \\ \frac{\alpha_1 v_{max}}{\Gamma} e^{v_{max}\xi/\Gamma} &= \frac{\alpha_2\lambda}{\Gamma} e^{-\lambda\xi/\Gamma} e^{-\alpha_2 e^{-\lambda\xi/\Gamma}}. \end{aligned} \tag{11.48}$$

Division yields

$$v_{max} = \alpha_2\lambda e^{-\lambda\xi/\Gamma}, \tag{11.49}$$

whence

$$\phi(\xi) = e^{-v_{max}/\lambda} \tag{11.50}$$

for any value of ξ. The value of ξ being thus purely arbitrary, we choose $\xi = 0$ for convenience. Thus, from (11.48) we have

$$\alpha_1 = e^{-v_{max}/\lambda}, \qquad \alpha_2 = \frac{v_{max}}{\lambda}; \tag{11.51}$$

whence, from (11.47) and (11.28), we obtain the explicit expressions

$$\frac{\phi(z)}{\phi_{max}} = \begin{cases} e^{v_{max}\left(\frac{z}{\Gamma} - \frac{1}{\lambda}\right)}, & \text{if } z < 0; \\ e^{-\frac{v_{max}}{\lambda}e^{-\lambda z/\Gamma}}, & \text{if } z \ge 0; \end{cases} \tag{11.52}$$

$$\frac{v(z)}{v_{max}} = \begin{cases} 1, & \text{if } z < 0; \\ e^{-\lambda z/\Gamma}, & \text{if } z \ge 0. \end{cases} \tag{11.53}$$

Notice that, according to (11.52), the beginning of the queue, where ϕ reaches ϕ_{max}, is reached only in the limit as $z \to \infty$; but this is quite acceptable because what is short according to the long scale may be extremely long according to the short scale. Naturally, having stipulated that drivers begin to react at $z = 0$, we cannot also have the beginning of the queue at $z = 0$, for drivers require a finite distance in which to brake.

The velocity and density profiles (11.52) and (11.53) are sketched in Fig. 11.7(a). Distance is measured on the horizontal scale in units of Γ/λ. Observe that the effective width of the transition layer is about $5\Gamma/\lambda$, which is small because Γ is small. The suddenness of this transition is sufficient justification, when solving (11.43), for imagining the regions of constant density to be at infinity; the mathematics is simplified but still produces a density profile with effectively constant regions that are but a short distance from $z = 0$. If you view the road from afar (the locality of infinity, as it were) then the profiles predicted by (11.43)–(11.45) will look like those in Fig. 11.7(b); the transition layer will appear as a discontinuity. Because $\phi(z)$ is a steady state, which may have persisted since $t = 0$, we can view the transition layer as being hidden in the discontinuity of (11.38) and (11.39), which are the initial conditions for a distant observer. To this observer, when you approach the queue, you appear to brake from v_{max} to

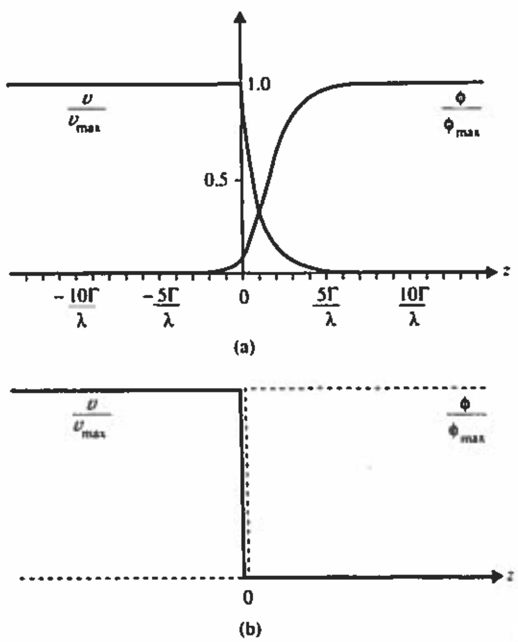


Fig. 11.7 Double take of the steady state density and velocity profiles; $v_{max}/\lambda = 55/20 = 2.75$. (a) Close-up of transition layer; (b) view from afar.

zero in zero distance, but a local observer would know from Fig. 11.7(a) that you brake to a halt over a distance $5\Gamma/\lambda$ and thus avoid crashing without the use of magic. Similarly, in the model of the previous section, the discontinuity of (11.18) hides the transition layer between ϕ_{max} and 0; it becomes the "expansion fan" of Fig. 11.5, in which characteristics diverge from the origin.

The concept of a narrow transition layer is a fruitful one in mathematical modelling (particularly in fluid dynamics, where it is usually dubbed a shock or boundary layer). It may remind you of the metered models of Chapters 1 and 5. There, when we had two different time scales, we were able to use one model for the small time scale and another for the large one. Here, when we have two different length scales, we are able to use one model, Fig. 11.7(a), for the small length scale and another, Fig. 11.7(b), for the large length scale. If you still feel uncomfortable about all this, please rest assured that a formidable arsenal of mathematical artillery exists for formalizing the use of several competing length and time scales. But because our brief is modelling, not methods, we restrict our attention to a heuristic approach. If you wish to pursue this matter more rigorously, then you might begin with Nayfeh (1981, 1973).

We conclude this section by observing that we have used the traffic diffusivity Γ without knowing its value (except that it's small), but the model of Fig. 11.7(a) can be used to obtain a crude estimate. It probably takes a car about 150 feet, or 5/176 mile, to brake to a halt from 55 m.p.h. Taking $\lambda = 20$ m.p.h., and equating this braking distance and $5\Gamma/\lambda$, we obtain the crude estimate

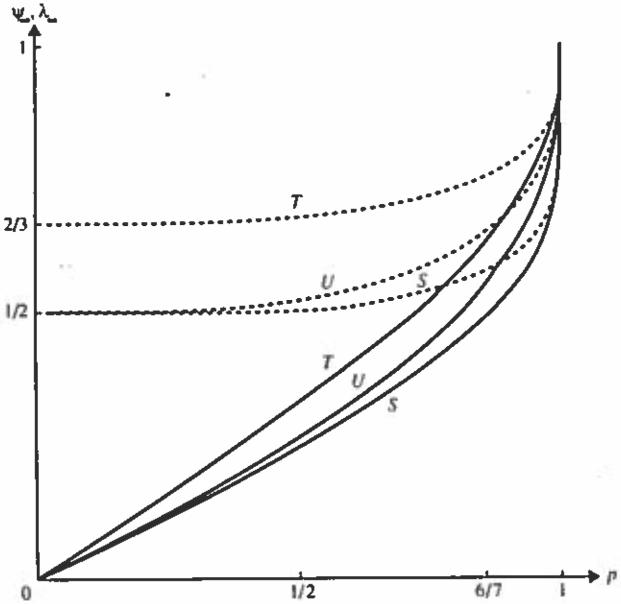
$$\Gamma = \frac{5}{44} \approx 0.1\mathrm{m}^2/\mathrm{h}. \tag{11.54}$$

11.5 Spreading Canal Pollution. An Adaptation

Suppose that a finite amount of pollution, P_0, is instantly injected into a canal, or indeed any other body of stagnant water.[3] How do we describe the subsequent diffusion? Now, we have already developed a model for diffusion of contaminant through stagnant water—that of Section 9.5. But we assumed in doing so that, initially, the contaminant was dispersed uniformly throughout the body of water. We must therefore adapt our model for circumstances in which, initially, the pollution is concentrated.

Let y denote distance along our canal and $y = 0$ the place of injection. Let t denote time, $t = 0$ the instant of injection, and $\phi(y, t)$ the concentration of contaminant per unit length of canal, at station y at time t. Because diffusion is so slow, let's assume that the canal extends to infinity in both

[3] You may wish to review Section 9.5 before proceeding.



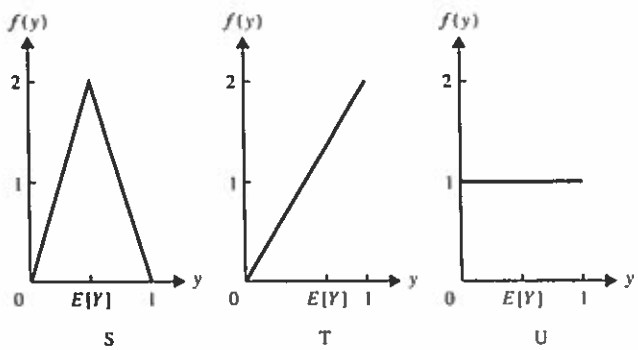


Fig. 12.1 (Top) Graphs of $\psi = \psi_{\infty}(p)$ (the dashed curves) and $\lambda = \lambda_{\infty}(p)$ (the solid curves) for the three merit distributions sketched in Fig. 12.2.
Fig. 12.2 (Bottom) Probability density functions of three possible distributions of merit between 0 and 1.

would be 10 candidates, then the maximum expected value of the score of the appointed candidate would, from Table 12.1, be 0.861. There is therefore a high price to be paid for any uncertainty in the length of the field of candidates.

Like the stationary buying policy of Section 12.1, the stationary interviewing policy of this section has the attractive feature that it is particularly easy to implement, the critical score being the same for each interview. It also has the attractive feature of flexibility, as discussed in Section 7.8. Curves like those in Fig. 12.1 could easily be drawn by computer for any distribution of merit, and so the model could be used by each of two college deans who have similar personal preferences (appoint the best candidate) but differ in their opinions of the quality of the field. Moreover, the graphs in Fig. 12.1 afford them the luxury of knowing the optimal policy at a glance, for an arbitrary value of p. On the other hand, it has the unattractive feature that the expected value of the quality of the candidate appointed will be significantly lower than if the number of candidates were known for certain. Thus an interviewer should strive to apply the policy of the previous section whenever possible. Now try Exercise 12.11.

12.4 The Motorist's Dilemma. Choosing the Optimal Parking Space

A motorist is approaching the central crossroads of an old-fashioned town, one that underwent ribbon development in the nineteenth century and hasn't changed much since (Fig. 12.3). She would like to park as close as possible to the central crossroads O, but she knows that the central parking spaces may not be free when she reaches them. She must therefore be prepared to park away from the center. How far from the center should she accept a vacant parking space? In this section, we will attempt to answer that question.

For the sake of definiteness, we will assume that the motorist approaches the crossroads from the west. Because traffic is extremely heavy, she decides that it will be virtually impossible to cross the road for a parking space. She therefore looks for a parking space on only her side of the road. The motorist has undertaken this journey many times before and knows how far each parking space is from the center. Her policy is to take the first available parking space if it is sufficiently close to the center. If she reaches the crossroads without finding a space, however, then she takes the next available space. If we assume that parking spaces are equally likely to be free along any of the four approach roads, then it makes no difference whether the motorist turns right (or left, though the heavy traffic will strongly discourage this) or carries on, if she reaches the crossroads. For the sake of definiteness, however, let's assume that she turns right—south in Fig. 12.3. Thus the qualitative form of the optimal policy is known. Our problem is to determine the optimal interpretation of "sufficiently close."

Let's begin by converting the decision rule to the form (12.24), in which action is taken if a certain variable exceeds a certain threshold. The motorist's potential parking spaces are as labelled in Fig. 12.3. For the

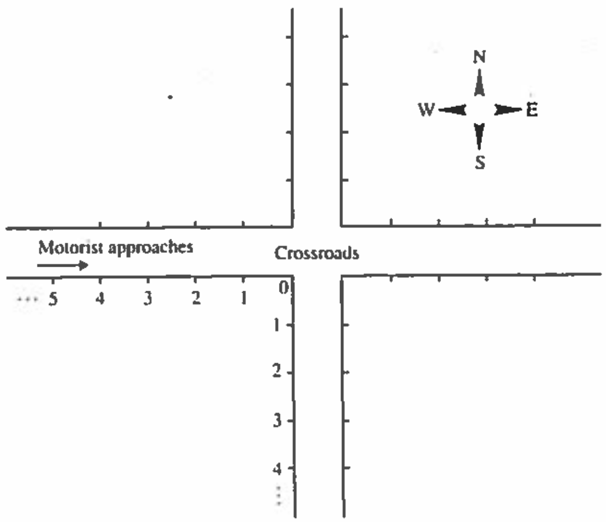


Fig. 12.3 Plan of central crossroads in an old-fashioned town.

sake of simplicity, we will assume that they are equidistant, so that the kth parking space from the center on the western approach road can be assumed to lie a distance k from the center. If they were not equidistant, however, then you could measure distance nonlinearly, as we measured time in Chapters 5 and 10, so that our model could still be applied. Let us agree that parking space k on the western approach road is "sufficiently close" to the center if $k \le k_c$, where k_c is an arbitrary integer. Then the decision rule can be cast in the form (12.24) by writing

$$\begin{aligned} &\text{If } k > k_c \text{ then DO NOT park (whether or not the space is empty);} \\ &\text{if } k \le k_c \text{ then DO park (if the space is empty).} \end{aligned} \tag{12.49}$$

We will attempt to determine the optimal value of k_c by adapting the model of the previous section.

Let us assume that parking spaces are occupied or unoccupied independently of one another. Let p denote the probability that a parking space is empty; let $q = 1-p$. We will determine the optimal value of k_c as follows. Let us pretend that the motorist has already set k_c, arbitrarily. Then, for each $k \le k_c$ along the western approach, she "interviews" parking space k, accepting it if it is empty and rejecting it if it is occupied, as implied by (12.49). Let the random variable V_k denote distance from the center at which the motorist will park, when k spaces remain before O. Thus the sample space of V_k is the set of integers $\{1, 2, \ldots, k\}$, $k \le k_c$. The motorist wishes to park as close to the center as possible, and so her utility, the quantity she would like to maximize, is $-V_k$. But this is a random variable, which cannot itself be maximized; so we will assume that our motorist agrees instead to maximize its expected value, $E[-V_k]$.[4] This is the same as minimizing $-E[-V_k] \equiv E[V_k]$, so we will refer to V_k as *disutility*. Let $L(k)$ denote $E[V_k]$. As usual, this expected value is computed with respect to possibilities arising at the time of "interview." Because the motorist wishes to minimize it, we will refer to $L(k)$ as her *disreward*, when $k (\le k_c)$ candidates remain and her strategy is to grab the next available parking space. We will calculate $L(k)$ by recursion. If there is any value of k between 0 and k_c for which $L(k) < L(k_c)$, then k_c has been set too high and should be lowered, because the motorist would then accept parking space k_c even though a nearer one to the center corresponds to a smaller disreward. If there is no value of k between k and k_c for which $L(k) < L(k_c)$, then k_c may have been set too low and may need to be raised, because a further parking space from O might be associated with a lower disreward if the policy of taking the next available space were applied earlier (on the other hand, k_c could have been guessed just exactly right). In practice, the pretense we have just made is not necessary because if we calculate the sequence $\{L(k) : k = 0, 1, \ldots\}$ on the assumption that k_c is at least as great as the current value of k, then $L(k)$ will decrease from $L(0)$ to a minimum $L(k^*)$ at $k = k^*$ and then increase again (see below). Thus k^* is the optimal value of k_c.

To begin our recursion, we must first find $L(0)$. With this in mind, suppose that the motorist reaches O without success, and let the (integer-valued) random variable D denote her parking distance from the center. Then $\text{Prob}(D = j)$ is the probability that the first $j-1$ spaces are filled and the jth is empty—$q^{j-1}p$—because parking spaces are available or occupied independently of one another. Thus, if she passes O, then her disreward is

$$E[D] = \sum_{j=1}^{\infty} j \cdot \text{Prob}(D = j) = \sum_{j=1}^{\infty} j \cdot q^{j-1} p = \frac{1}{p}, \tag{12.50}$$

by the same calculation as in Exercise 12.8. We deduce immediately that

$$L(0) = \frac{1}{p}. \tag{12.51}$$

Now suppose that the motorist is "interviewing" parking space k, where $0 \le k \le k_c$. There are two possibilities. Either the space is free, in which case she takes it and her disutility is $V_k = k$, or else the space is occupied in which case her disutility is the perceived disreward from the remaining $k-1$ parking spaces, i.e., $L(k-1)$. Hence her disutility is

[4] Because she is a frequent visitor to town, this is quite rational; see Section 8.8.

the random variable

$$V_k = \begin{cases} k, & \text{if the space is empty;} \\ L(k-1), & \text{if the space is occupied.} \end{cases} \tag{12.52}$$

This is a discrete random variable whose expected value over the two possibilities that now arise can be calculated from (A.31). Hence

$$L(k) = E[V_k] = p \cdot k + q \cdot L(k-1). \tag{12.53}$$

Thus the sequence $\{L(k)\}$ can be calculated from the recursion

$$L(0) = \frac{1}{p}, \qquad L(k) = pk + qL(k-1), \qquad k = 1, 2, \ldots. \tag{12.54}$$

The motorist would like to know the value, k^*, for which the integer-valued function $L(k)$ achieves its minimum.

The values of $L(k)$ for values of k between 0 and 5 are presented in Table 12.2. The table suggests that

$$\begin{aligned} L(k) &= p\sum_{i=1}^{k} iq^{k-i} + \frac{q^k}{p} = pq^k \sum_{i=1}^{k} i\left(\frac{1}{q}\right)^i + \frac{q^k}{p} \\ &= k - \frac{q}{p} + \frac{q^k(1+q)}{p}, \end{aligned} \tag{12.55}$$

by essentially the same calculation as in Exercise 6.4. The truth of (12.55) can be established by mathematical induction (see Exercise 12.12). Because

$$L(k+1) - L(k) = 1 - (1+q)q^k, \tag{12.56}$$

we see that L attains its minimum at $k = k^*(p)$, where $k^*(p)$ is the smallest integer exceeding

$$\frac{\ln(2-p)}{-\ln(1-p)} = \frac{\ln(1+q)}{\ln\left(\frac{1}{q}\right)}. \tag{12.57}$$

The staircase-like function $k = k^*(p)$ is plotted in Fig. 12.4 as the solid curve, with (12.57) the dashed one. Notice that the motorist should hold

Table 12.2 The first few recursions of (12.54).

k	$L(k)$
0	$1/p$
1	$p + q/p$
2	$p(2+q) + q^2/p$
3	$p(3+2q+q^2) + q^3/p$
4	$p(4+3q+2q^2+q^3) + q^4/p$
5	$p(5+4q+3q^2+2q^3+q^4) + q^5/p$

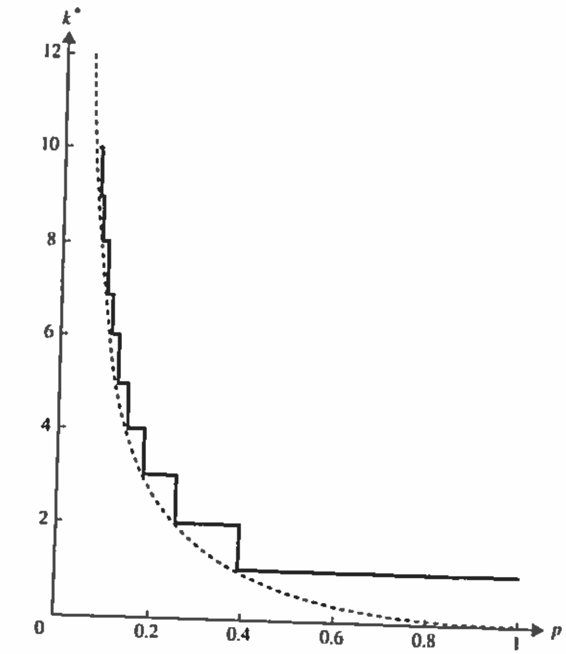


Fig. 12.4 Graphs of $k = k^*(p)$ (solid curve) and (12.57) (dashed curve). A motorist who wishes to park as close as possible to the center of town should not accept an empty parking space if it is more than k^* spaces from the center.

out for the final parking space $k = 1$ unless it is less than about 40% likely to be vacant (the precise percentage is $50\{3 - \sqrt{5}\}$). This is because, in such circumstances, the motorist has a high probability of finding a good consolation parking space on the road going south if the one she really wants is filled. If p is very small, however, then k^* is off the diagram in Fig. 12.4, and the motorist must be prepared to park a long way from town. Naturally, the value of the probability p must be supplied (most likely, subjectively) before the optimal policy can be implemented, but the fact that p can be retained as a parameter is an attractive feature of our model. Variations of it are considered in Exercise 12.13.

12.5 How Should a Bird Select W[redacted]

Suppose [redacted] worms. What type of worm should she be looking [redacted] thin, skinny one? To be su[redacted] worm contains more energy (calo[redacted]