A mechanism to describe the formation and propagation of stop-and-go waves in congested freeway traffic

BY JORGE A. LAVAL1,* AND LUDOVIC LECLERCQ2

1School of Civil and Environmental Engineering, Georgia Institute of Technology, Atlanta, GA, USA
2Laboratoire Ingénierie Circulation Transport LICIT (INRETS/ENTPE), Université de Lyon, France

This paper introduces a parsimonious theory for congested freeway traffic that describes the spontaneous appearance of oscillations and their ensuing transformation into stop-and-go waves. Based upon the analysis of detailed vehicle-trajectory data, we conclude that timid and aggressive driver behaviours are the cause for this transformation. We find that stop-and-go waves arise independently of the details of these behaviours. Analytical and simulation results are presented.

Keywords: stop-and-go waves; microscopic traffic-flow models; kinematic-wave model

1. Introduction

Stop-and-go driving is a nuisance for motorists throughout the world. Not only does it increase fuel consumption and emissions, but it also imposes safety hazards. Unfortunately, our understanding of this type of oscillation in congested traffic is still limited. On the one hand, detailed vehicle-trajectory data are very scarce, and aggregated sensor data are often noisy and insufficient. On the other hand, few attempts have been made to validate the oscillations predicted by existing traffic-flow models, which are often a result of mathematical curiosities rather than driving behaviour.

A traffic oscillation has two components, formation and propagation. It is known that the formation can be caused by lane-changing activity (Laval 2005; Ahn & Cassidy 2006; Laval & Daganzo 2006) or, in general, any kind of moving bottleneck (Koshi et al. 1992; Laval 2005). The propagation component remains unclear. In particular, we still do not know the precise mechanism that makes oscillations grow, even in the absence of lane changes; almost invariably, a seemingly small perturbation turns into a full stop, which may or may not dissipate as it propagates upstream. This phenomenon was first observed in the famous Lincoln tunnel experiments (Edie 1961). Nor do we know why oscillations

*Author for correspondence (jlaval3@ce.gatech.edu).

One contribution of 10 to a Theme Issue ‘Traffic jams: dynamics and control’.

This journal is © 2010 The Royal Society
tend to exhibit a regular period somewhere between 2 and 15 min (Mauch & Cassidy 2002; Ahn et al. 2004; Ahn 2005; Laval et al. 2009); see figure 1. This paper will shed some light on these matters.

Earlier research has unveiled important insights. Car-following theories that explicitly incorporate driver reaction time (Kometani & Sasaki 1958; Newell 1961) were shown to explain oscillation growth, but with very small periods of the order of a few seconds. To explain the 5 min periods seen in the Lincoln tunnel, Newell (1962, 1965) conjectured that there must be two congested branches in the flow–density fundamental diagram: an upper branch containing traffic states when vehicles decelerate and a lower branch when vehicles accelerate. He did not specify, however, the driver-behaviour mechanisms, or equivalently, the ‘paths’ in the fundamental diagram for going from one branch to the other.

Lately, significant contributions have been made, and we now have several traffic-flow models that predict oscillation growth and dissipation. They may be roughly classified into ‘fully stochastic’ and ‘unstable’ car-following models. In the latter case, Wilson (2008) concluded that oscillations with a period of several car spacings may be expected to grow in a very general class of car-following models,
so long as basic ‘common-sense’ driving behaviours are observed. This indicates that stop-and-go is inherent to driving behaviour, regardless of the specific details (but within reason). This result has been reinforced by Igarashi et al. (2001) and Orosz et al. (2009), who studied nonlinear jam formation for well-known car-following models. They showed that when traffic becomes dense enough, the nonlinear effect may trigger congestion waves for sufficiently large perturbations caused by a single vehicle. Unfortunately, these works did not provide a detailed description of how perturbations occur and did not relate jam formation with experimental observations. Some physical insights into the formation and the propagation of stop-and-go waves are still lacking.

Fully stochastic models include random components to account for variable driver behaviour. A classic example is the seminal cellular automata of Nagel & Schreckenberg (1992). It produces oscillations owing to a braking probability that can be constant or a function of the speed (Barlovic et al. 1998, 2002), as do the second-order macroscopic models proposed by Khoshyaran & Lebacque (2007). Gas-kinetic models are inherently stochastic (Helbing & Treiber 1998; Shvetsov & Helbing 1999; Helbing et al. 2001; Ngoduy et al. 2006; Ngoduy 2008), but have decreased in popularity, probably because of the large number of non-physical parameters needed, and the complexity of their implementation. Del Castillo (2001) and Kim & Zhang (2008) took Newell’s conjectures, supposed that the paths between the two congested branches were random and showed how this produces oscillations that grow or dissipate. In particular, Del Castillo (2001) assumed that in the deceleration phase, headways are random variables, whereas in the acceleration phase, they are deterministic. Kim & Zhang (2008) postulated that the follower draws his/her reaction time from a probability distribution each time the leader changes speed. In Kerner (2004), drivers chose their speed and spacing somewhat randomly between two branches in the fundamental diagram based on the hypotheses that drivers continuously vary their spacing trying to look for lane-changing opportunities.

Unfortunately, with the exception of Kerner (2004), all the models cited previously lack an explicit connection with the mechanisms of driver behaviour responsible for these instabilities. Such models predict sudden stops (or severe slow down) that are only owing to the stochastic process and not correlated with any explicit and physical explanation. Note that Yeo & Skabardonis (2009) recently interpreted the I-80 Next Generation Simulation (NGSIM) trajectory data (NGSIM 2006) using Newell’s conjectures and pointed out that the cause for oscillations might be human error, i.e. anticipation and overreaction. They based their explanation on the existence of five different traffic phases. A model, however, is still lacking.

The aim of this paper is to describe a parsimonious mechanism of driver behaviour based on empirical observation able to explain the formation and propagation of oscillations. Towards this end, §2 provides background information, and §3 analyses the US-101 and I-80 NGSIM trajectory datasets on a single freeway lane, away from lane changes. Based on these observations, §4 proposes a theory that explicitly models timid and aggressive driver behaviour. Simulations are provided in §5 in order to analyse the type of oscillations produced at upgraded freeway sections. Finally, a discussion is provided in §6.

Phil. Trans. R. Soc. A (2010)
2. Newell’s car-following model

The car-following model of Newell (2002) gives the exact solution of the kinematic-wave theory of Lighthill & Whitham (1955) and Richards (1956) with a triangular flow–density diagram Daganzo (2005); referred to here as the kinematic-wave triangular (KWT) model. The triangular fundamental diagram requires only three observable parameters: free-flow speed $u$, wave speed $-w$ and jam density $k$; see figure 2a. Additionally, it is the only fundamental diagram to produce kinematic-wave solutions where acceleration and deceleration waves travel upstream at a nearly constant speed and without rarefaction fans, as observed empirically (Foster 1962; Cassidy & Windover 1995; Kerner & Rehborn 1996; Muñoz & Daganzo 2000; Ahn et al. 2004; Chiabaut et al. 2009a, b). In the context of car-following models, it is convenient to express the decreasing branch of the triangular fundamental diagram in terms of vehicle speed $v$ and spacing $S(v)$, i.e.

$$S(v) = \delta + \tau v,$$

where $\tau = 1/(wu)$ is the wave trip time between two consecutive trajectories and $\delta = 1/k$ is the jam spacing; see figure 2b. Both $\delta$ and $\tau$ are assumed identical for all drivers and can be estimated using $w$ and $k$, which are readily observable.

The exact KWT solution in continuous time–space is

$$x_{i+1}(t) = \min \{x_{i+1}(t - \tau) + u\tau, x_i(t - \tau) - \delta\},$$

where $x_{i+1}(t)$ is the position of vehicle $i + 1$ at time $t$. Under free-flow conditions, $x_{i+1}(t)$ is given by the first term of equation (2.2), which implies infinite vehicle-acceleration capability; under congestion, the second term dominates. When these two terms are equal, it means that the vehicle reaches a shockwave, i.e. the back of a queue marking the transition between free flow and congestion; see points ‘1’ and ‘2’ in figure 2a, c, respectively. The second term implies that vehicle $(i + 1)$’s trajectory is identical to that of its leader $i$ but shifted $\tau$ time units forward in time and $\delta$ distance units upstream in space; see figure 2b.

Solution (2.2) is a consequence of having a single wave speed independent of the traffic state so long as it is confined to a single regime, free flow or congestion. Therefore, traffic information, such as speeds, flows, spacing, etc., propagate unchanged along characteristics travelling at either $u$ or $-w$. For example, differentiating the second term of KWT solution (2.2) with respect to time gives

$$v_{i+1}(t) = v_i(t - \tau),$$

which indicates that the speed of all vehicles along a wave is identical when traffic is congested. It is important to note that in non-stationary conditions, the follower’s spacing at time $t$, $s_{i+1}(t)$, should be interpreted as

$$s_{i+1}(t) = S(v_i(t - \tau)),$$

rather than $s_{i+1}(t) = x_i(t) - x_{i+1}(t)$, which coincides with equation (2.4) only under stationary conditions. This property can be observed in figure 2b, which also shows the graphical interpretation $S(v)$.

Phil. Trans. R. Soc. A (2010)
3. Data analysis

In this section, we analyse the NGSIM trajectory dataset collected on a 640 m segment on southbound US-101 in Los Angeles, CA on 15 June 2005 between 07.50 and 08.35. For generality, we have supplemented these measurements with NGSIM data collected on eastbound I-80 in the San Francisco Bay area in Emeryville, CA on 13 April 2005. We focus on vehicle trajectories when traversing oscillations at the time of their formation.

Figure 1 presents a time–space diagram of the median-lane trajectories. This is an uphill segment with percentage grades ranging from 2 to 4.6%. The maximum speed is around 60 km h$^{-1}$, which indicates that this segment is under the congestion of a downstream bottleneck. Note how the oscillations propagate upstream at a constant wave speed, $w$, of approximately 16 km h$^{-1}$. Although most oscillations originate downstream of the segment (there is a long 5.6% grade approximately 2 km downstream responsible for the congestion in this segment), there are several oscillations that arise within the segment. This is especially true during the first 15 min, where oscillations arise at a fixed location with a striking regular period of about 2 min.

Figure 3 shows a more detailed view of the first 15 min of the data. A different speed colour scale has been used to highlight the changes in speed. It can be seen how

R-1: speeds begin to decrease gradually well before the oscillation starts propagating; this is more evident in figure 3$b,c$, which show that this ‘precursor period’ lasts for about 1 min,
R-2: after the precursor period comes the oscillation, which propagates upstream at the wave speed, and

R-3: the minimum speed inside oscillations decreases gradually until reaching a complete stop after some 20 vehicles. This stop lasts about 30 s.

This behaviour can be seen qualitatively in the real data shown in figure 3d,e. These figures show a detailed view of a handful of vehicle-trajectory data from figure 3b,c. Dashed lines are called ‘Newell’s trajectories’ and represent vehicle trajectories shifted according to the second term in equation (2.2). Therefore, the deviations between vehicle \((i + 1)\) trajectory (solid lines) and its leader’s Newell’s trajectory is an indication of the degree of \((i + 1)\)’s non-equilibrium. Note that this shift has been performed ‘manually’ for every trajectory pair in such a way that prior to the deceleration waves, both Newell’s and the follower’s trajectories appear superimposed.
Mechanism to describe stop-and-go waves

In particular, when the solid line is above the relevant dashed line, it means that \((i + 1)\) is accepting spacings that are shorter than the equilibrium spacing; conversely, when the solid line is below the dashed line, \((i + 1)\) maintains a spacing larger than the equilibrium spacing. These two types of driver behaviour are called here ‘aggressive’ and ‘timid’, and are labelled using asterisks and plus symbols in the figure. The important point here is that vehicle speeds drop to zero after the interaction of just a few timid and aggressive drivers along the oscillation; we call this oscillation growth.

To verify this observation, figure 4 shows additional platoons facing a deceleration wave and causing the oscillation to grow. The following remarks can be made.

R-4: before drivers are forced to decelerate owing to the driver in front, they follow Newell’s trajectory very closely; while decelerating, some drivers change this equilibrium behaviour to either timid or aggressive behaviour,

R-5: out of the 74 trajectories in figures 3 and 4, approximately 40 per cent exhibit timid behaviour, 20 per cent aggressive behaviour and the remaining 40 per cent are well described by Newell’s model,

R-6: most trajectories exhibit a single behaviour, but occasionally (approx. 5%) both timid and aggressive behaviour are observed for the same driver,

R-7: timid drivers seem to have a bigger impact on oscillation growth because ‘shying-away’ behaviour invariably leads to a lower speed; after his/her target spacing has been reached, the timid driver accelerates to catch up with his/her Newell’s trajectory later on; however, the speed decrease created earlier tends to propagate upstream, and

R-8: aggressive drivers can also create a speed decay. This aggressive behaviour can only last for a short period of time to avoid collision. When the aggressive driver reaches his/her target spacing and realizes that his/her leader is not accelerating, he/she is forced to return to his/her equilibrium spacing by driving slower than his/her leader.
To study timid and aggressive driver behaviour in more detail, figure 5 shows a few trajectory pairs and their corresponding flow–density evolution as they cross an oscillation. These have been computed using Edies’ generalized definitions (Edie 1961), which for an $n$-vehicle platoon inside an arbitrary time–space region $A$ are

$$k = n \sum_{i=1}^{n} \frac{T_i}{|A|}, \quad q = n \sum_{i=1}^{n} \frac{D_i}{|A|} \quad \text{and} \quad v = \frac{q}{k} = \frac{n \sum_{i=1}^{n} D_i}{\sum_{i=1}^{n} n T_i},$$

where the symbols $k$, $q$ and $v$ stand for density, flow and speed in $A$, respectively, $|A|$ is the area of $A$, and $T_i$ and $D_i$ are the $i$th vehicle travel time and distance travelled inside $A$, respectively.

Each small circle in the fundamental diagram in figure 5 corresponds to flow–density measurements inside a particular trapezoidal area in the corresponding time–space diagram. The number ‘1’ in each figure indicates the first such measurement; the rest of the measurements are consecutively added as the trajectory pair travels downstream. Figure 5a,b corresponds to a timid follower, while figure 5c,d to an aggressive follower. One can see that

R-9: timid drivers seem to decelerate along the equilibrium branch ($1 \to 5$ and $1 \to 8$ in figure 5a,b, respectively), and accelerate along a ‘lower’ branch to come back to the initial traffic state $1$,

R-10: aggressive drivers decelerate along an ‘upper’ branch ($1 \to 7$ and $1 \to 3$ in figure 5c,d, respectively), and accelerate along the equilibrium branch to come back to equilibrium, and

R-11: shortly after the appearance of these non-equilibrium behaviours, the speed of vehicles upstream rapidly drops to zero.
These remarks suggest that, unlike Newell’s conjectures, there appears to be multiple branches on the fundamental diagram in congestion. At least one for equilibrium states and two others for timid and aggressive driving behaviour.

The vehicle speed decay towards zero along the oscillation can be seen in figures 6 (US-101) and 7 (I-80). These figures present trajectories crossing four oscillations at the time of their formation, and shows the speed of vehicles along characteristics near deceleration waves. These characteristics are depicted as lines...
in each time–space diagram, and the speed of each vehicle was measured at the
time they cross this line. It can be seen that

R-12: vehicle speeds seem to decay from vehicle to vehicle according to a random
walk with a negative drift,
R-13: this negative drift has an average of approximately $\text{-1.13 km h}^{-1}$ between
two consecutive vehicles and a standard deviation of $0.28 \text{km h}^{-1}$, both
sites combined, $\text{-1.25 and 0.29 km h}^{-1}$ at US-101, and $\text{-1 and 0.19 km h}^{-1}$
at I-80; we can conclude that the drift at both sites appears to be drawn
from the same distribution, and
R-14: from figures 6a,d and 7a,b,d, there seems to be room for speculating
that the functions depicted there are convex; i.e. speed drops faster at
higher speeds.

Next, we propose a theory based on the above remarks. We will show that the
existence of non-equilibrium behaviour alone is not enough to produce the kind
of oscillations seen in the data.

4. The continuous-time model

The model presented in this section is the continuous-time version of the
model in Laval & Leclercq (2008) extended to account for the remarks in the
previous section. In the spirit of the kinematic-wave model: (i) dynamics will be
studied along characteristics travelling at $-w$, which is assumed constant, and
(ii) vehicle positions are obtained as the minimum of a free-flow term and a
congestion term.

Our main assumption is that in congestion, deceleration waves can trigger some
drivers initially in equilibrium to switch to two non-equilibrium ‘modes’, either
timid or aggressive (R-4).

(a) The free-flow term

Following Laval & Daganzo (2006), we shall remove infinite accelerations in
Newell’s model by incorporating the vehicle-kinematics model

$$a(v) = a_m \left(1 - \frac{v}{u}\right) - gG,$$

(4.1)

with $g = 9.81 \text{m s}^{-2}$. Equation (4.1) can be interpreted as the desired vehicle
acceleration when travelling at a speed $v$ on a freeway segment with a 100G\%
grade. The parameter $a_m$ gives the maximum acceleration, which is attained at
$v = 0$; also, $u/a_m$ can be interpreted as a relaxation time.

Here, we note that equation (4.1) can be solved analytically to obtain vehicle
displacements, $\tilde{x}_{i+1}(t)$, e.g. between $t - \tau$ and $t$,

$$\tilde{x}_{i+1}(t) = v_c \tau - (1 - e^{-br})(v_c - v_{i+1}(t - \tau)),$$

(4.2)
Mechanism to describe stop-and-go waves

Mechanism to describe stop-and-go waves

where $v_c = u(1 - gG/a_m)$ is the crawl speed and $b = a_m/u$. Then, the free-flow term becomes

$$x_{i+1}(t) = x_{i+1}(t - \tau) + \min\{u\tau, \ddot{x}_{i+1}(t)\} \quad \text{(in free flow).} \quad (4.3)$$

(b) The congestion term

In order to model non-equilibrium spacings in congestion, we use the dimensionless variable $\eta_{i+1}(t)$ to denote the ratio between vehicle $(i + 1)$’s spacing, $s_{i+1}(t)$, and its equilibrium spacing, i.e.

$$s_{i+1}(t) = \eta_{i+1}(t)S(v_i(t^*_i)),$$  
where $t^*_i$ is the time when the characteristic reaching the follower at time $t$ is emanated by the leader $i$ (see figure 8a), i.e.

$$t^*_i = t - \eta_{i+1}(t)\tau. \quad (4.5)$$

This is true because $S(v)$ is linear in both $\tau$ and $\delta$, and therefore the non-equilibrium fundamental diagram $\eta_{i+1}(t)S(v_i(t^*_i))$ can be interpreted as having wave trip time $\eta_{i+1}(t)\tau$ and jam spacing $\eta_{i+1}(t)\delta$. Therefore, vehicle trajectories in the proposed model can be obtained with the following generalization of equation (2.2):

$$x_{i+1}(t) = x_i(t^*_i) - \eta_{i+1}(t)\delta = x_i(t - \eta_{i+1}(t)\tau) - \eta_{i+1}(t)\delta. \quad (4.6)$$

As an illustration, consider figure 8a where the follower is in equilibrium at time $t_0$ and starts an aggressive behaviour by driving faster than Newell’s trajectory. As in Newell’s model, the trajectory of vehicle $i$ can also be obtained in terms of the lead-vehicle trajectory, i.e.

$$x_{i+1}(t) = x_0 \left( t - \tau \sum_{j=1}^{i+1} \eta_j(t^*_j) \right) - \delta \sum_{j=1}^{i+1} \eta_j(t^*_j). \quad (4.7)$$

Figure 8. (a) Time–space diagram of a pair of trajectories in the proposed model and (b) vehicle speed decays toward zero along the oscillation predicted by equation (4.10) using $\varepsilon = 100\,\text{vehicles h}^{-1}$.
Figure 9. Time–space diagrams of a leader introducing a speed perturbation and the reaction of: (a,c) a timid and (b,d) an aggressive driver proposed here. Drivers complete their non-equilibrium cycle (a,b) before and (c,d) after the perturbation has passed.

(c) The behavioural model

According to R-7 and R-8, we define two non-equilibrium ‘modes’:

— contraction mode: the driver seeks to shorten its spacing relative to the equilibrium spacing; i.e. the time derivative of $\eta$ is negative, $\eta'_{i+1}(t) < 0$, and

— expansion mode: the driver seeks to widen his/her spacing relative to the equilibrium spacing, i.e. $\eta'_{i+1}(t) > 0$.

It follows that, within this framework, a timid behaviour is composed of an expansion followed by a contraction, whereas an aggressive behaviour is a contraction followed by an expansion. This is shown in figure 9, which depicts a leader introducing a speed perturbation and the reaction of a timid (figure 9a,c) and an aggressive (figure 9b,d) driver as hypothesized here. Figure 9a,b depicts a follower that completes his/her non-equilibrium cycle before the perturbation is removed, and figure 9b,c depicts a follower that completes his/her non-equilibrium cycle after the perturbation is removed.

The transition between the two modes is modelled here using a single parameter, $\alpha$, to represent the notion of a target deviation from equilibrium; i.e. a target value of $\eta$, $\eta_T$; see figure 9. Until this target is not achieved, the
Mechanism to describe stop-and-go waves

<table>
<thead>
<tr>
<th></th>
<th>timid expansion</th>
<th>contraction</th>
<th>aggressive contraction</th>
<th>expansion</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\eta(0))</td>
<td>1</td>
<td>1 + (\alpha)</td>
<td>1</td>
<td>1 - (\alpha)</td>
</tr>
<tr>
<td>(\eta_T)</td>
<td>1 + (\alpha)</td>
<td>1</td>
<td>1 - (\alpha)</td>
<td>1</td>
</tr>
</tbody>
</table>

Table 1. Dependence of the initial and target \(\eta\) values on the type of behaviour and mode.

driver remains in the current mode. For simplicity in exposition, from now on, \(t = 0\) denotes the time when a driver switches from one mode to another. Both the initial value, \(\eta(0)\), at the beginning of a mode and the target \(\eta_T\) depend on the type of behaviour and mode, as shown in table 1.

It remains to establish how a driver goes from \(\eta(0)\) to \(\eta_T\), i.e. the shape of \(\eta'(t)\). Taking derivatives in equation (4.6), using \(S(v_i(t^*_i)) = tv_i(t^*_i) + \delta\) and rearranging, one finds

\[
\eta'_{i+1}(t) = \frac{v_i(t^*_i) - v_{i+1}(t)}{S(v_i(t^*_i))}.
\]

(4.8)

It follows that reasonable models for \(\eta'(t)\) should come from reasonable assumptions regarding the speed difference between the two vehicles along characteristics \(v_i(t^*_i) - v_{i+1}(t)\). Therefore, the proposed framework allows us to study a family of models that generalize equation (2.3) by postulating

\[
v_{i+1}(t) = f(v_i(t^*_i)),
\]

(4.9)

where \(f(\cdot)\) is an arbitrary function of the leader speed (and possibly some other arguments) that describes the follower’s behaviour; see figure 8a. A potential practical limitation of this approach is that for most specifications of \(f(\cdot)\), the ODEs (4.8) and (4.9) are nonlinear, and it may not be possible to solve them explicitly, or their solution might be non-algebraic. Fortunately, the specification

\[
f(v) = v - \varepsilon S(v),
\]

(4.10)

where the parameter \(\varepsilon\) has units of flow, accords well with remarks R-12–R-14. To see this, figure 8b presents the vehicle speeds predicted by equation (4.10) along a characteristic (thick line), together with three random walks around it. Note the slight convexity of the prediction line. Notably, equation (4.10) yields a linear solution for equations (4.8) and (4.9),

\[
\eta_{i+1}(t) = \eta_{i+1}(0) + \varepsilon t,
\]

(4.11)

where \(t = 0\) denotes when the driver switches modes. Depending on the sign of \(\varepsilon\), one can model expansions (\(\varepsilon > 0\)) or contractions (\(\varepsilon < 0\)).

\(d\) Some properties for the lead-vehicle problem

We shall analyse here the analytical solution of the proposed model for a lead vehicle that decelerates instantaneously to a speed \(v_0\) followed by a platoon in congestion. This platoon is assumed homogeneous (all drivers timid or aggressive)
Figure 10. Exact solution for a platoon of seven (a) timid and (b) aggressive drivers. Following table 1 and figure 9, the expansion for timid drivers is implemented such that at $t = 0$ (when the leader reduces its speed to $v_0$), $\eta > 0$ is set until $\eta = \eta_T = 1 + \alpha$ is reached; for the contraction of aggressive drivers, at $t = 0$, $\eta < 0$ is set until $\eta = \eta_T = 1 - \alpha$ is reached. It was assumed that $\alpha = 0.6$ and $|\eta| = 200$ vehicles h$^{-1}$.

and in a stationary initial (non-)equilibrium state characterized by $\eta(0) = \eta_0$; once the leader $i = 0$ decelerates at $(t, x) = (0, 0)$, all drivers seek a common target $\eta_T$ and stay at this target when it is reached.

Figure 10a shows the (exact) solution for a platoon of seven timid vehicles. According to table 1 and figure 9a, the expansion mode is implemented such that at $t = 0$ (when the leader reduces its speed to $v_0$), $\eta > 0$ is set until $\eta = \eta_T = 1 + \alpha$ is reached. The case of an aggressive platoon is shown in figure 10b. In this case, the contraction mode implies that at $t = 0$, $\eta < 0$ is set until $\eta = \eta_T = 1 - \alpha$ is reached.

(i) **Speeds in contraction and expansion modes**

One can obtain the speeds $v_1, v_2, \ldots, v_7$ in figure 10 by noting that they are time independent (since the speed of the first vehicle is constant) and ‘unfolding’ the iterative relationship (4.9) and (4.10): $v_{i+1} = (1 - \varepsilon \tau)v_i - \varepsilon \delta$. We find that for $t \geq 0$, the speed of vehicle $i$ is given by

$$v_i = (v_0 + w)(1 - \varepsilon \tau)^i - w.$$  \hspace{1cm} (4.12)

Of course, we also require $0 \leq v_i \leq u$. Noting that $|\varepsilon| S(v) \approx 1$ km h$^{-1}$ from figures 6, 7 and 8b, it follows that $|\varepsilon| \tau$ should be of the order of $10^{-2}$. Therefore, for fixed $|\varepsilon|$, the magnitude of the speed change along the platoon are almost identical for expansions ($\varepsilon > 0$) and contractions ($\varepsilon < 0$). Equation (4.12) can be solved for $i$ to obtain the number of vehicles upstream of the leader that it takes to reach a given speed $v$,

$$i(v) = \frac{\log(v + w)(v_0 + w)}{\log(1 - \varepsilon \tau)},$$ \hspace{1cm} (4.13)

which should be comparable with $(v_0 - v)/(v_0 + w)/\varepsilon \tau$ if $|v_0 - v|$ is not too big.
(ii) Iso-η contours

Note that in our model, the value of η is not constant along characteristics. To see this, note that in our example, $\eta_i(t) = \eta_0 + \epsilon(t - i\eta_0\tau)$ and $x_i(t) = -i\delta\eta_0 + (t - i\eta_0\tau)v_i$, where $v_i$ is given by equation (4.12). Eliminating $i$ in these two equations yields the iso-η contours $\eta = c$, i.e.

$$x(t) = \tilde{c}(v_0 + w)(1 - \epsilon\tau)^{(1-\tilde{c})/(\eta_0\tau)} - tw \quad \text{(iso-η contours),}$$

where $\tilde{c} = (c - \eta_0)/\epsilon$. This indicates that for expansions ($\epsilon > 0$), the first term on the right-hand side tends to 0 as $\epsilon \to \infty$, so that all iso-η contours tend to the line $x = -wt$ from above; in contraction, this term tends to $\infty$ as $\epsilon \to \infty$ and all iso-η contours tend to be vertical as time passes. This can be verified in figure 10. In fact, it can be seen in figure 10a that the iso-η contours for timid drivers always travel faster than $-w$, with the the iso-η contour being the fastest (as shown by the bold curve in the figure). The first vehicle to reach $\eta_T$ is $i = 1$ at point ‘1’ in the figure, where a wave is emanated. Once vehicles cross the iso-η contour, their trajectory is given by Newell’s model in the sense that vehicle speeds along the characteristics are constant. This produces the narrow bands shown in the figure, which result in a piece-wise linear acceleration fan.

Figure 10b shows that in the case of aggressive drivers, the iso-η contours travel at a speed lower than $-w$, which gradually decreases with $c$. The first vehicle to reach $\eta_T$ is $i = 1$ at point ‘2’ in the figure. From this point on, its speed changes to $v_0$ and a wave is emanated upstream travelling at $-w$. When this wave crosses each vehicle in the platoon, they decelerate gradually since they are still in contraction mode. They reach their target $\eta_T$ at the bold line in the figure, which happens to be flat in this particular example, but can even travel downstream, bearing a striking resemblance with the precursor region described in figure 3. This region can be described as a deceleration fan.

(e) Flow–density paths

The path in the flow–density diagram for a given trajectory pair, where the leader’s speed is $v(t)$, can be obtained in parametric form by $\{k(t) = 1/(\eta(t)S(v(t)))\}$, $q(t) = v(t)/k(t)$, which is valid only in $1 - \alpha \leq \eta(t) \leq 1 + \alpha$. In the case of a leader advancing at constant acceleration, $v(t) = v_0 + at$, eliminating $t$ in the above parametric equations gives

$$q(k) = \sqrt{\left(\frac{a}{\epsilon}(\kappa w - \eta_0(v_0 + w)k) + \frac{1}{4}\left(\frac{a}{\epsilon} + v_0 + w\right)^2k\right) k - \frac{1}{2}\left(\frac{a}{\epsilon} - \eta_0 - v_0 + w\right) k},$$

(4.15)

which is now valid only in $w(\kappa/(1 + \alpha) - k) \leq q(k) \leq w(\kappa/(1 - \alpha) - k)$. Note that $q(k) = v_0k$ when $a = 0$; i.e. contractions and expansions are performed along lines of constant speed when the leader does not accelerate. This can be seen in the data of figure 5a,b.

Interestingly, the leader’s acceleration and the model parameter $\epsilon$ always appear in equation (4.15) as a ratio, $a/\epsilon$. This implies that the situations $\epsilon \geq 0$, $a \geq 0$ and $\epsilon \leq 0$, $a \leq 0$ are described by the same family of
Figure 11. Families of flow–density trajectories for different values of $a/\epsilon$; equation (4.15) is evaluated for $v_0 = \{0, 10, 20, 30, 40, 60, 90\}$. Note how timid behaviour in figure 5a can be obtained by combining paths $1 \to 2$ in (c), $2 \to 3$ in (f) and $3 \to 1$ in (e); the aggressive behaviour of figure 5c can be described within (e).

flow–density trajectories; so are $\epsilon \leq 0$, $a \geq 0$ and $\epsilon \geq 0$, $a \leq 0$. Six such families are shown in figure 11. In each flow–density diagram in the figure, the value of $a/\epsilon$ is fixed and equation (4.15) is evaluated for six different values of $v_0$; this process is shown in figure 11a,d for a single value of $v_0$.

One can obtain flow–density trajectories similar to the ones in figure 5 by carefully combining different branches from different parts in figure 11. For example, the timid behaviour in figure 5a can be seen (at least qualitatively) by combining paths $1 \to 2$ in figure 11c, $2 \to 3$ in figure 11f and $3 \to 1$ in figure 11c. The aggressive behaviour of figure 5c can be described within a single family, as shown by the path depicted in figure 11e. Note that for maximum analogy with figure 5c, for this path we have assumed that the jam density cannot be surpassed. However, this restriction does not change the macroscopic behaviour of the model significantly.

It is worth noting that the proposed model does not specify a behavioural model when the leader accelerates ($a > 0$), provided that the leader has decelerated previously and thus triggered non-equilibrium behaviour upstream.

Next, we resort to the numerical resolution of the proposed theory to show that uphill grades may produce perturbations that trigger the formation of traffic oscillations with regular periods.

5. The discrete model

In order to perform numerical simulations, we resort to the discrete model proposed in Laval & Leclercq (2008), which turns out to be a discretization of a special case of the continuous-time framework proposed here.

*Phil. Trans. R. Soc. A* (2010)
Let the superscript \( j \) denote the value of a variable at discrete time \( t_j = j\tau, j = 0, 1, 2, \ldots \), e.g. \( x_j^i = x_i(t_j) \). Here, we use

\[
x_{i+1}^{j+1} = \min\{x_j^i + \min\{u\tau, \tilde{x}_{i+1}^{j+1}\}, x_j^i + \tau v_{i+1}^{j+1} - \eta_{i+1}^{j+1} S(v_{i+1}^{j+1})\}. \tag{5.1}
\]

The free-flow term is exact as it corresponds to equation (4.3). The congestion term corresponds to an ‘Euler-type’ discretization in that it gives the exact solution of our continuous-time model assuming that during \( t_j < t \leq t_{j+1} \), both \( v_i(t) \) and \( h_i(t) \) are constant and equal to \( v_{i+1}^{j+1} \) and \( h_{i+1}^{j+1} \), respectively; see Laval & Leclercq (2008). Note how model (5.1) gives Newell’s model (2.2) when \( a = \infty \) and \( \eta_{i+1}^{j+1} = 1 \), as expected. The update of \( \eta \) is now

\[
\eta_{i+1}^{j+1} = \eta_{i+1}^{j} + \varepsilon \tau \quad \text{(deterministic)}, \tag{5.2}
\]

as suggested by equation (4.11). Two other models will be tested here,

\[
\eta_{i+1}^{j+1} = \eta_{i+1}^{j} + \varepsilon \tau + N(0, \sigma^2 \tau) \quad \text{(random walk)} \tag{5.3}
\]

and

\[
\eta_{i+1}^{j+1} = \eta_{i+1}^{j} + N(0, \sigma^2 \tau) \quad \text{(fully stochastic)}, \tag{5.4}
\]

where \( N(0, \sigma^2 \tau) \) denotes a normal random variable with zero mean and variance \( \sigma^2 \tau \). In all models, we impose \( 1 - \alpha \leq \eta_{i+1}^{j+1} \leq 1 + \alpha \).

Model (5.3) was suggested by R-12, as it implies that \( \eta_{i+1}^{j+1} \) is a Brownian motion (or a random walk with normally distributed steps). This specification allows us to ‘generalize’ (4.10), which was chosen for mathematical convenience—although justified by figures 6 and 7—rather than driving behaviour considerations. Model (5.4) is ‘fully stochastic’ as described in the introduction; i.e. at each time step, drivers randomly select to increase or decrease their speed independently of the logic proposed in this paper. Recalling equation (4.8), speed differences \( v_i(t_i^*) - v_{i+1}(t) \) in this model will be approximately \( N(0, (S(v_i(t_i^*)) \sigma)^2 / \tau) \).

6. Periodic oscillations at uphill segments

In this section, we study a level-1-lane road with an uphill segment of length \( L = 300 \text{ m} \) and a \( 100\% \) grade. The desired acceleration of all vehicles is given by equation (4.1) and the proportion of timid and aggressive drivers is \( r_T \) and \( r_A \), respectively; the remaining proportion \( 1 - r_T - r_A \) obeys Newell’s model. The system is empty at \( t = 0 \), when vehicles enter the upstream boundary of the segment in equilibrium \( (\eta_i = 1) \) and at capacity (traffic state ‘1’ in figure 2a). Vehicles exit the system at the downstream boundary located 500 m downstream of the top of the uphill.
The three models in the previous section were run for all combinations of variables $G$, $\alpha_m$, $r_A$ and $r_T$ and parameters $\epsilon$, $\alpha$ and $\sigma$, each one varying within a range observed to produce notable qualitative changes. The length of the segment is 3 km, the upgrade is located at $x = 2$ km and its length is fixed at $L = 300$ m. The simulation time is 1 h in all cases, and the fundamental diagram parameters are: $u = 120$ km h$^{-1}$, $w = 20$ km h$^{-1}$ and $\kappa = 150$ vehicles km$^{-1}$.

Figure 12 shows typical speed maps from the deterministic model; figure 12$a,c,e$ only considers timid drivers, whereas figure 12$b,d,f$ only considers aggressive. It can be seen how different oscillation periods can be obtained by varying the model parameters. Note the precursor regions in figure 12$a$, similar to the ones observed in figures 1 and 3. It was verified that the experiments with only aggressive drivers produce outputs similar to figure 12$a,c,e$, but also the more complex patterns in figure 12$b,d,f$. In particular, note in figure 12$b$ how certain waves seem to merge with adjacent ones or simply dissipate, while only a few ‘survive’ and propagate upstream unchanged.

Figure 13 shows typical speed maps obtained with the random walk in figure 13$a,b$ and fully stochastic models in figure 13$c,d$. Figure 13$a,b$ can be compared with figure 12$a,b$, which shows, at least qualitatively, that these two
models exhibit the same properties; this was verified for a wide range of parameter values. Figure 13c, d show that the fully stochastic model is able to produce periodic oscillations (figure 13c), but also exhibits rather questionable patterns (figure 13d).

To quantify the main differences of the models, the oscillation period and amplitude of the oscillations were analysed over a 1 h speed time series at location $x = 1$ km for all runs (approximately 2000 per model). We used the method proposed in Li et al. (2010), which uses Fourier spectrum analysis to minimize the bias in estimating amplitudes and periods of traffic time-series data. Figure 14 shows the oscillation period and amplitude against the dimensionless ratio $gG/a_m$, for all three models. The different data points were obtained with different values of the parameter $\alpha$ and averaged across all remaining parameters. This parameter is common to all models and is a measure of driver-behaviour variability.

Regarding the period we observe: (i) for all models, there is a well-defined decreasing relationship with $gG/a_m$, (ii) in the range $0.4 \leq gG/a_m \leq 0.8$, all models produce a very similar period in the range 3–10 min, (iii) in the range $gG/a_m \leq 0.4$, the deterministic and random-walk model predict significantly higher oscillation periods than the fully stochastic model, and (iv) aggressive driver behaviour systematically implies longer periods. Similarly, for the amplitude we note that: (v) the deterministic and random-walk model predict a well-defined decreasing relationship with $gG/a_m$, whereas the fully stochastic model predicts a concave relationship, (vi) for all values of $gG/a_m$, the fully stochastic model results in significantly lower amplitude values, and (vii) aggressive driver behaviour systematically implies larger amplitudes.

Figure 13. Typical speed maps obtained with the (a, b) random walk and (c, d) fully stochastic models. (c) $\sigma = 20$ h$^{-1}$; $a = 0.35$; $a_m = 1$ m s$^{-2}$; $r = 0.05$; $G = 3\%$ and (d) $\sigma = 20$ h$^{-1}$; $a = 0.35$; $a_m = 0.6$ m s$^{-2}$; $r = 0.2$; $G = 3\%$. 

Phil. Trans. R. Soc. A (2010)
Figure 14. Oscillation period against the dimensionless ratio $gG/a_m$ for: (a) deterministic, (b) random-walk and (c) fully stochastic models.

7. Discussion

We have introduced a framework based on empirical observation, which predicts the speed decay inside traffic oscillations produced by small perturbations on a single lane with no lane changes. This initial perturbation may have many causes, and this paper showed that upgrade sections is one of them. Moreover, the proposed theory seems to capture the mechanism that triggers periodic oscillations in this type of bottleneck.

The empirical evidence presented here seems to contradict existing conjectures. We found no evidence that oscillation growth is caused by drivers seeking lane-changing opportunities, as suggested in Kerner (2004). Nor did we
Mechanism to describe stop-and-go waves

find evidence supporting Newell’s acceleration and deceleration branches. We believe it seems more appropriate to talk about aggressive and timid branches, which act more like reflecting bounds for the flow–density paths described here.

The iso-\(\eta\) contours \(\eta = \eta_T\) are significant because in real data they would appear to be waves travelling at nearly constant speeds, but this constant speed can be very different in each realization; see figure 10. This may be an explanation for the shape of iso-speed contours in the precursor periods of figure 3b,c.

Our simulation results indicate that fully stochastic models exhibit questionable behaviour. To verify this assertion, we have analysed the loop-detector data located just upstream of the bottom of the 5.6 per cent uphill segment in US-101 alluded to in §3. This steep grade should correspond to \(gG/a_m\) values of about 0.7. The speed time series during the morning rush exhibits a period of about 7 min with an amplitude of about 15 km h\(^{-1}\). The reader can verify that this data point accords very well with all models, except the fully stochastic one.

The simulation experiment also revealed that the deterministic and random-walk model are essentially identical, at least with respect to the type of oscillations they produce. Therefore, there may be numerous variations of the basic deterministic model presented here that would produce similar results. We conclude that oscillations are a consequence of driver heterogeneous reactions to deceleration waves as postulated here, but independently of the details of this behaviour. Note the analogy to the finding of Wilson (2008) mentioned in the introduction.

The physical explanation of the correlation between the model parameters and variables and oscillation characteristics is not included here for brevity, but is the subject of current research by the authors. Suffice it to say that in the case of figure 12a,e, the oscillation period, or the time until the next oscillation, is a direct consequence of the discharge flow from the previous oscillation and of the speed of those vehicles when approaching the bottom of the uphill. This is true because a deceleration wave is felt whenever the speed of incoming vehicles is greater than the speed on the upgrade (which is likely to be close to the crawl speed \(u(1 - gG/a_m)\)). And the probability of an oscillation to form increases with the incoming flow, which in turn increases as the oscillation propagates upstream; see the ‘lower flow’ and ‘higher flow’ zones in figure 12c. Consideration shows that these two factors may explain the high correlation between the period and the dimensionless ratio \(gG/a_m\): the steeper the upgrade, the lower the crawl speed, the faster vehicles inside the oscillation will reach a full stop, and the faster these vehicles will be able to accelerate back to the crawl speed.

Finally, the proposed mechanism can also explain the hysteresis phenomenon (Treiterer & Myers 1974). Accordingly, hysteresis happens whenever a driver reaches equilibrium after the disturbance has passed, as in figure 9c,d. In both cases, flow–density measurements before and after the deceleration wave would appear in different branches of the fundamental diagram. However, these branches are not related to vehicle accelerations or decelerations as previously thought (Treiterer & Myers 1974), but rather to aggressive and timid behaviours. In fact, as shown in figure 11c,f, the acceleration path of an aggressive driver is above its deceleration path, which contradicts Treiterer & Myers (1974).
This research was partially supported by the Georgia Department of Transportation project 2006N78, and the National Science Foundation award number 0856218. The authors would like to thank two anonymous referees and D. Chen, B. Chihukuri, H. Zhengbing, X. Li and Y. Ouyang, whose comments and suggestions greatly improved the quality of this paper.

References


Mechanism to describe stop-and-go waves


