

Heterogeneity, Lane-Changing and Instability in Traffic: A Mathematical Approach

Jonathan A. Ward

Department of Engineering Mathematics

University of Bristol



A dissertation submitted to the University of Bristol in
accordance with the requirements of the degree of
Doctor of Philosophy in the Faculty of Engineering.

2009

Abstract

This thesis is concerned with the mathematical analysis of spatio-temporal pattern formation in highway traffic. The general approach is multiscale in that microscopic (individual driver) models and their simulation are studied using macroscopic (partial differential equation) techniques. Three problems are tackled:

1. **Spatial Heterogeneity.** The stationary patterns that result from simulations of a car-following model with spatial heterogeneity in the form of a simple capacity drop bottleneck are studied. These states are analysed at progressively smaller scales, from simple conservation principles and characteristics arguments to second order continuum models and ultimately the linearised microscopic model itself.
2. **Stability.** Two new results concerning the stability analysis of microscopic car-following models at the *linear level* are presented. Firstly, a simple formula for the linear stability of a general class of car following models is derived and then extended to include heterogeneous driver populations (for example mixes of car and trucks). Secondly, the convective propagation of disturbances on the open-road is analysed using group velocity and transform techniques.
3. **Lane-Changes.** Complex microscopic lane-changing models are analysed in a highly simplified setting, for which macroscopic flow rates between lanes are derived. A flexible simulation framework has been developed for the simulation of multilane car-following models, which has subsequently been used to verify our macroscopic theory.

Acknowledgements

First and foremost, Eddie; it's been a roller-coaster ride but you've always been totally upfront and honest, which has made you a great supervisor, collaborator and mate. I'd also like to thank all the academic and support staff in the Engineering Maths department as well as the post-docs and PhD students past and present — it has been a stimulating place to study with a great atmosphere, both for work and social life. In particular I'd like to thank Jo Mason and Tom Melvin; I can honestly say that the last three (and a bit) years have been the best of my life and you guys have been a massive part of that. I guess I've also got to thank Howell Jordan for the nick-name.

I'm very grateful to David Lloyd for his in-depth and extended technical discussions about fluid dynamics, travelling waves and every other aspect of surfing. I've learnt more than I had ever hoped for. I'm also eternally grateful to the whole Lloyd clan: Chris, Jan, Mike and Dotty; you're my second family. To all my other friends, in particular Phil and Sam, cheers for all the great times. To my brother Chris, who turned out to be my best mate and to Grandma, Mum and Dad; I can't express how much I love you guys and how appreciated you are. And finally, to the most important person in my life, Claire Haseler, simply put, I love you.

Author's Declaration

I declare that the work in this dissertation was carried out in accordance with the regulations of the University of Bristol. The work is original except where indicated by special reference in the text and no part of the dissertation has been submitted for any other degree.

Any views expressed in the dissertation are those of the author and in no way represent those of the University of Bristol.

The dissertation has not been presented to any other University for examination either in the United Kingdom or overseas.

Signed:

Dated:

Contents

1	Introduction	1
1.1	Motivation	1
1.2	Overview of Mathematical Modelling Methods	6
1.2.1	Microscopic Models	6
1.2.2	Coarse-graining Microscopic models	11
1.2.3	Continuum Models	12
1.3	Outline of Thesis	14
2	Multiscale Modelling of Bottleneck Pattern Formation	17
2.1	Motivation	17
2.2	Problem Set-Up and Simulation Results	19
2.3	Density and Flow Balances	23
2.4	Wave Selection via Characteristics	24
2.5	Further Characteristic Analysis	27
2.5.1	Cumulative Flux Method	31
2.5.2	Calculation of the Two/Three-Plateau Boundary	34
2.6	Second Order Modelling	37
2.7	Discrete Spectral Analysis	42
2.8	Summary	46
3	Instability in Car-Following Models	49
3.1	Motivation	49
3.2	General Linear Stability Analysis	50
3.3	Heterogeneous Driver Behaviour	53
3.4	Absolute and Convective Instability	60
3.5	Inductive Asymptotic Methods	66
3.6	Laplace Transform Development	73
3.7	Summary	76
4	Macro Analysis of Micro Lane-Changing	79
4.1	Motivation	79
4.2	Microscopic Multi-Lane Model	81

4.3	Simulation Framework	83
4.4	Longitudinal Quasi-Equilibrium	87
4.5	Partition of (h_0, h_1) Space	89
4.6	Worked Example	94
4.7	Taylor Approximation of λ	97
4.8	Comparison with Microscopic Simulation	100
4.9	Lane-Changing and Instability	103
4.10	Summary	110
5	Conclusions	111
5.1	Review of Thesis	111
5.2	Future Research Directions	113
5.2.1	Spatial Heterogeneity	113
5.2.2	Instability	113
5.2.3	Lane-Changing	114
5.3	Final Remarks	115
A	Derivation of Multiclass Stability Analysis Identities	117
B	Fourier Analysis of the Linearised General Car-Following Model	119
	References	121

List of Tables

- 4.1 Each element corresponds to the value of λ for the case indicated (see Figure 4.8). 90
- 4.2 Comparison of microscopic data with QEA data (a) and linearised QEA solutions (b). Data is for the OVRV model with four different initial headways (first two columns) using three lane-changing timescales $\mu = 0.1, 0.01$ and 0.001 . Model parameters are $\alpha = 2, \beta = 0.5$ and the lane-changing thresholds are $\sigma = 2$ and $\tau = 1.4$. The third and fourth columns in each table are the number of vehicles in each lane. The right-most three columns are the fractional density difference error $\varepsilon_{a,b}$ for (a) and (b) respectively. **Table (a):** For $\mu = 0.1$ the agreement is poor for most cases, however this improves significantly for $\mu \leq 0.01$. Note that there is very good agreement for all values in the second row. This is because there is very little change in the equilibrium density differences, i.e. few lane-changes take place. **Table (b):** In most cases there is very little quantitative difference between the results for the linearised QEA solutions and the macroscopic QEA data (Table (a)). The largest discrepancy between the approximation methods is for the initial headways $(h_0, h_1) = (2.5, 18)$. This initial condition is close to the I_1 boundary (see Figure 4.13), where λ is strongly non-linear. In this case, the linearised QEA solution under estimates the density difference. Note that in contrast to the other initial conditions, there agreement is better at *large* μ values. 102

List of Figures

- 1.1 Variable Message Signs (a) are a flexible means of informing drivers about current road conditions, expected travel times, road works etc. The Active Traffic Management scheme (b) introduced on the M42 around Birmingham combines dynamic control of speed limits and hard shoulder operation during peak periods to increase capacity, smooth traffic flow and reduce vehicle emissions [2]. The Highways Agency have recently deployed Traffic Officers (c) to provide support for motorists in difficulty and for rapid incident response. Images reproduced courtesy of the Highways Agency website <http://www.highways.gov.uk> [1, 3]. 2

- 1.2 Data taken from inductance loops over one minute averages. The relationship between velocity v and flow q (panel (a)), velocity v and density ρ (panel (b)) and flow q and density ρ (panel (c)) are illustrated. There is a uni-modal relationship between flow and velocity/density, and velocity decreases with density. 3

- 1.3 An example of spatio-temporal patterns in inductance loop data from the north-bound M42 Motorway in the UK. Distance in kilometres is plotted on the vertical axis against time in hours on the horizontal axis. Colour corresponds to speed in kilometres per hour. There are three junctions present in this figure, the position of which are labelled with J4–J6. Drivers travel up the page, in the direction of increasing distance and time. Vehicles are forced to slow down considerably around Junction J6, and numerous stop-and-go waves can be seen propagating from this location against the flow of traffic. 4

- 1.4 Spatio-temporal evolution of the Cellular Automata model proposed by Nagel and Schreckenberg [64] to emulate traffic flow. The road is composed of discrete lattice sites i that are either occupied by a vehicle or are unoccupied at any discrete instance in time j . Vehicles have integer values of velocity $v = 1, 2, \dots, v_{\max}$. Vehicular velocity is used to denote which lattice sites are occupied in the simulation depicted above for the case where $v_{\max} = 5$. A period is used to denote unoccupied sites. The lattice is updated from time j to $j + 1$ using the following algorithm: 1. *Increase* the velocity of vehicles with $v < v_{\max}$ by 1, $[v \rightarrow v + 1]$; 2. *Decrease* the velocity of vehicles with $h < v$ free lattice spaces ahead to $h - 1$, $[v \rightarrow h - 1]$; 3. *Decrease* the velocity of vehicles with $v > 0$ by 1 with probability p , $[v \rightarrow v - 1]$; 4. *Move* all vehicles forward by v lattice sites. These rules ensure collision-free dynamics. CA models require very little computational time during each update, hence systems with large numbers of vehicles can be modelled over long time intervals. These models also exhibit transitions from laminar flow to stop-and-go behaviour (an example of the latter is illustrated above). Figure reproduction courtesy of K. Nagel and M. Schreckenberg [64]. 7
- 1.5 Illustration of a simple car-following model set up. Vehicles occupy a single lane travelling in the increasing x -direction (in this case to the right). Associated with each vehicle is a position x_n and a velocity v_n . The spacing between cars is known as the headway h_n 8
- 1.6 Microscopic simulations of the OV model (1.5) with OV function (1.6) and unstable parameters: $\alpha = 1.9$, $h_* = 2$. The initial velocity of a single vehicle is perturbed by a small amount. This is amplified by subsequent vehicles, eventually forming stop-and-go waves which can be seen in the top-right-hand corner as a striped pattern in the vehicle trajectories due to a sharp change in velocity. Note that the flow is uniform in the bottom-left-hand corner. 9
- 1.7 A sketch of the Fundamental Diagram (compare with the empirical data in Figure 1.2(c)). Characteristic speeds are given by the gradient $Q'(\rho)$, and shock speeds correspond to the gradient of the chord connecting the up- and downstream densities, ρ_u and ρ_d respectively. In this case the shock velocity is negative and hence moves against the flow of vehicles. 13

List of Figures

- 1.8 Sketch in (t, x) of the characteristics for an initial pulse in density with sharp up- and downstream fronts (panel (b)), with density profiles at times t_0 and t_1 (panels (a) and (c) respectively). The pulse density is initially ρ_j , corresponding to negative characteristic velocity, and the background density is ρ_f , corresponding to positive characteristic velocity. The downstream front forms an expansion fan. The upstream front propagates as a shock against the flow of traffic until the expansion fan refracts its trajectory and eventually annihilates the pulse altogether. . . . 14

- 2.1 Stationary $t \rightarrow \infty$ coarse-grained density profiles $\rho(x)$. The portions of solution profiles within the bottleneck are indicated by shading. (a) Light traffic $h_* = 7.0$ ($\rho_* = 0.142857$), see Example 1; (b) Medium traffic $h_* = 2.5$ ($\rho_* = 0.4$), see Example 3; (c) Heavy traffic $h_* = 1.0$ ($\rho_* = 1.0$), see Example 2. 22

- 2.2 Solution structure of (2.5,2.6). Panels (a-c) correspond directly to panels (a-c) in Fig. 2.1. Top row (i) indicates solutions of (2.7) and bottom row (ii) of Eq. (2.8). The extra numbering in panels (c)(i,ii) allows the (ρ_1, ρ_B) solution pairs to be identified. 24

- 2.3 Characteristic analysis for the (ρ_1, ρ_B) solution pairs from Fig. 2.2(c)(i,ii). Panel (iii) shows the location of solution pairs (pair (1) is not fully contained in the figure), joined by horizontal lines representing flow balance, on the fundamental diagrams Q and $r_B Q$. Characteristic pictures for each of the three root pairs are shown in panels (iv), (v) and (vi): the bottleneck is denoted by shading. Panels (v) and (vi) cannot give stationary profiles since they predict an expansion fan at the up- and down-stream ends of the bottleneck respectively. Hence solution pair 2 from panel (iv) is selected. Note that in panel (iii), this solution pair is non-straddling in the sense that both ρ_1 and ρ_B are the same side of the fundamental diagram maximum. 25

- 2.4 Characteristic analysis for the three-plateau case: only the configuration shown with flow maximised in the bottleneck avoids expansion fans. Note that the characteristics inside the bottleneck have zero velocity and hence this structure is on the very boundary of becoming an expansion fan. The internal shock between ρ_1 and ρ_2 is classical since at it the characteristics converge. 27

- 2.5 Phase diagram derived from (2.12) for bottleneck length $\hat{L} = 0.25$. The points marked (a), (b), and (c) correspond to panels (a), (b) and (c) in Figure 2.1 and Figure 2.2. The line segments denoted A and B indicate to where the phase boundary would move for $\hat{L} = 0.5$ and $\hat{L} = 0.75$ respectively. 28
- 2.6 Illustration of the tangent chord function $\sigma = G(\rho)$ plotted on the Fundamental Diagram. The tangent to the flux-density relation $Q(\rho)$ at σ , intersects the density ρ_u . Non-classical shocks occur between upstream densities ρ_u and downstream densities $\rho_d > \sigma$ 29
- 2.7 Evolution of characteristics around a single bottleneck when the initial flux in free flow is *less* than the capacity of the bottleneck, i.e. $\rho_* < \rho_L$. The Fundamental Diagram is drawn in panel (a) where the initial density ρ_* and the changes in density that occur when characteristics pass into and out of the bottleneck are labelled ρ_b and ρ_f respectively. The development of characteristics are illustrated in Panel (b), the bottleneck region is shaded in grey. The expansion fan that results when characteristics cross the upstream bottleneck front and the shock that propagates away from the downstream bottleneck front are labelled. 30
- 2.8 Evolution of characteristics around a single bottleneck when the initial flux in free flow is *greater* than the capacity of the bottleneck, i.e. $\rho_* > \rho_L$. The Fundamental Diagram is drawn in panel (a), where the initial density ρ_* and the changes in density when characteristics pass into and out of the bottleneck are labelled. In the later case, the density drops to ρ_f . Characteristics which cross into the bottleneck can no longer conserve flux. Hence these are forced to connect to a much higher density, ρ_R . Characteristics are illustrated in panel (b). A shock can be seen propagating upstream from the bottleneck entrance, which corresponds to the development of a queue. Furthermore, the characteristics in the bottleneck are dominated by the expansion fan. The density in the bottleneck therefore approaches ρ_{\max} , and hence maximises flow. 31

List of Figures

- 2.9 Sketch of characteristic evolution on a ring-road when $\rho_L < \rho_* < \rho_{\max}$. The shocks emanating from the bottleneck collide, from which the resulting shock propagates forwards since $\rho_* < \rho_{\max}$. This is then refracted by the expansion fan at (x_c, t_c) . There are two subsequent outcomes. The characteristics upstream of the shock have increasing density, approaching ρ_L . Hence the velocity of the shock tends to zero in the free flow region. If the shock crosses into the bottleneck however, its velocity is *increased* by the downstream fan. 32
- 2.10 The Riemann initial value problem. The fundamental diagram is illustrated in panel (a). Piecewise constant initial densities (ρ_u upstream and ρ_d downstream) are initially separated by a discontinuity at $x = 0$ (illustrated in panel (b)). We choose these densities such that $Q'(\rho_d) < Q'(\rho_u)$, hence the discontinuity propagates as a shock. Characteristics are sketched in the (x, t) plane in panel (c). The shock path is labelled 'S' and the up- and downstream characteristics that intersect the x -axis at c_u and c_d respectively coincide on the shock at point I. 33
- 2.11 Numerical (ρ, ρ_x) phase plots corresponding to Figure 2.1(a-c). The discs mark the boundaries of the bottleneck. (a) Light traffic: heteroclinic cycle connecting $(\rho_1, 0)$ and $(\rho_B, 0)$. (b) Three plateau case: saddle at $(\rho_1, 0)$; saddle-node at $(\rho_{\max}, 0)$ explaining S-shaped structure; complex fixed point at $(\rho_2, 0)$. (c) Heavy traffic: complex fixed points at $(\rho_1, 0)$ and $(\rho_B, 0)$ 37
- 2.12 Light traffic phase-plane topologies for (2.51, 2.52), with parameters chosen so as to agree with Experiment 1. (a) Unconstrained regime (2.53) and (b) bottleneck (2.54). Flows in the top half-plane go from left to right and flows in the bottom half-plane go from right to left. Dark lines indicate the manifolds associated with the saddle points at $(\rho_{fl}, 0)$ and $(\rho_{bl}, 0)$. The unstable manifolds connect through the upper half-plane with stable nodes at $(\rho_{fr}, 0)$ and $(\rho_{br}, 0)$. (c) Overlay of the separatrices from (a) and (b). Grey markers denote the fixed points and black markers denote where the separatrices meet and hence where the switch between model (2.53) and model (2.54) occurs. Compare with (d) which is a reproduction of Figure 2.1(b). 40
- 2.13 (a) Reproduction of Figure 2.11(b) where the portion of the trajectory corresponding to the bottleneck is highlighted in bold. (b) Overlay of (2.51, 2.52, 2.54) phase-plane trajectories (grey lines) with microscopic simulation data for the bottleneck from (a). The saddle node $(\rho_{sn}, 0)$ is marked with a cross. 42

- 2.14 The markers in (a-c) denote solutions $\lambda = \psi + i\omega$ to (2.62). In panels (a-c), parameters are chosen such that the spectrum calculated represents respectively: the light traffic bottleneck plateau, as in Figure 2.1(a); the medium density plateau just upstream of the bottleneck, as in Figure 2.1(b); and the high density plateau outside of the bottleneck, as in Figure 2.1(c). The complex conjugate pairs marked in bold in panels (b,c) are those that explain oscillatory x -increasing escape from heavy traffic. 44
- 3.1 Growth rates, λ_{+R} , for the OVRV model. Parameter values are $\alpha = 1.6$ and $\beta = 0.4, 0.2, 0$ for curves (a), (b) and (c) respectively. The temporal mode λ_{-R} does not affect the stability in these cases. The values $\lambda = 0$, $\theta = 0$ are always solutions. Curve (a) bends down at the origin, hence the parameters are stable. Curve (b) is tangent to the θ axis at the origin, and is thus marginally stable. Curve (c) bends up at the origin and hence the parameters are unstable. Thus instability emerges at small wave-number, i.e. at $\theta = 0$, and hence long wave-length. 52
- 3.2 Marginal stability curves of the OVRV model with driver heterogeneity for the parameter values $\alpha = 1.4$ and $\beta = 0.2$. Grey scale contours of the fraction of cars η are plotted in the (γ, v_*) plane, where γ is the OV scaling coefficient (i.e. how much slower the slow vehicles are) and v_* is the equilibrium velocity. At the point labelled (v_s, γ_s) , the values are $v_s = 1$, $\gamma_s = 0.8$ and marginal stability occurs at $\eta \approx 0.76$ 60
- 3.3 Unstable (panel (a)) and stable (panel (b)) velocity time-series of a single vehicle from simulations of the OVRV model with heterogeneous driver behaviour (3.34). Model parameters are $\alpha = 1.4$, $\beta = 0.2$, $\gamma = 0.8$, $N = 100$ vehicles and the equilibrium velocity is $v_* = 1$. In panel (a) there are $N_C = 70$ cars (hence $N_T = 30$ trucks) and thus the system is unstable. In panel (b) there are $N_C = 80$ cars (hence $N_T = 20$ trucks) and the system is stable. Notice that in panel (a), the velocity disturbance initially decays and growth is only observed over long timescales. 61

List of Figures

- 3.4 Numerical examples of convective (Cases (1) and (3)) and absolute (Case (2)) instability using the OVRV model. The OVRV model is used with unstable parameter values $\alpha = 0.5$, $\beta = 0.2$ and $h_* = 2.8, 2, 1.3$ for Cases (1–3) respectively. A fixed time-step of $dt = 0.025$ is used and the number of vehicles is $N = 300$. Velocity v_n is plotted against position x_n at two instances in time and these points are joined with a line. Blue lines are earlier times ($t = 125, 100, 225$ for Cases (1–3) respectively) than green lines ($t = 500, 125, 325$ for Cases (1–3) respectively). Vehicles move in the increasing x -direction, i.e. left to right. The initial disturbance originates close to the origin. Cases (1) and (3) illustrate convective instability because the perturbation is carried down- and upstream respectively. Case (2) illustrates absolute instability since the disturbance grows both up- and downstream simultaneously. 62
- 3.5 Growth rates $\lambda_{+R}(\theta)$ (blue) and absolute group velocities (green) for Cases (1–3). As in Figure 3.4, Cases (1–3) correspond to the OVRV model with $\alpha = 0.5$, $\beta = 0.2$ and $h_* = 2.8, 2.0, 1.3$ respectively. The largest unstable wave-numbers are labelled θ_d and the zero absolute group velocities are labelled θ_g . There is no θ_g solution in Case (1), whereas in Case (3) $\theta_g > \theta_d$. Hence Cases (1) and (3) are *convectively* unstable. In Case (2), θ_g is within the range of unstable wave-numbers (i.e. $\theta_g < \theta_d$). Hence Case (2) is *absolutely* unstable. 64
- 3.6 Instability phase diagram for the OVRV model in (α, h_*) space with $\beta = 0.2$ using the standard OV function (1.6). Regions of stability (labelled S, white), convective instability (labelled CU travelling upstream and CD travelling downstream, both in light grey) and absolute instability (labelled A, dark grey) are illustrated. The transition boundary at $h_* = h_c$ is found by solving (3.56). The dashed line corresponds to the transition boundary predicted by the small wave number expansion (3.58). The parameter values for Cases (1–3) ($h_* = 1.3, 2.0, 2.8$ and $\alpha = 0.5$ respectively) are marked with circles and labelled (1–3) respectively; these classifications are in agreement with Figures 3.5 and 3.4. . . . 66

- 3.7 Temporal evolution of the headways h_n of three vehicles ($n = 15, 45$ and 90) using the OVRV model with the standard OV function (1.6). Microscopic simulations (blue) are compared with solutions of the linearised equations of motion (green), calculated using (3.73). Parameters are chosen such that uniform flows are unstable ($\alpha = 1.0$ and $\beta = 0.2$), the equilibrium spacing is $h_* = 2.0$ and the initial perturbation is $\delta = 0.48$. Hence nonlinear instability develops rapidly, yet there is good agreement between the linearisation and the microscopic simulation up to $n = 45$, after which the growth of the disturbance is slowed by nonlinearities in the microscopic simulation. 70
- 3.8 Solutions $\tilde{h}_n(\tau)$ (blue) plotted in n for three times, $\tau = 25, 37.5, 50$ (panels (a–c) respectively for the OVRV model with $\alpha = 1.5$, $\beta = 0.1$ and the standard OV function (1.6). The leading order term (3.78) (green) and Gaussian envelope of (3.79) (red) are also plotted. The leading order term proves to be a poor approximation, and actually gets worse as t increases. This appears to be due to dispersion in the leading order term that must be counter-acted by lower order terms. 71
- 3.9 Coefficients of vehicle $n = 20$ $a_m^{(20)}$ are plotted in panel (a) and in panel (b) these are multiplied by t^m for $t = 25$. In both panels, the real part is plotted in blue, the imaginary part in green and the absolute value in red. The leading order term ($m = 19$) is actually the *smallest* term in panel (a). Furthermore, the main contribution depends on the time t , illustrated in panel (b), where at $t = 25$, the dominant contributions come from the terms $m = 10$ and $m = 15$. Note however the large magnitude of the scale in panel (b); hence solutions trajectories must result from a complicated cancellation of terms. 72
- 4.1 Neighbours of vehicle n (grey circle) denoted n_f, n_b, n_{fo} and n_{bo} . Headways $h_n^{(f)}, h_n^{(b)}, h_n^{(fo)}$ and $h_n^{(bo)}$ are measured longitudinally and take positive values by convention. 82

List of Figures

- 4.2 Multilane simulation flow diagram. Initial data (including all positions, headways, velocities, pointers and parameters) is read as input, after which the main loop is executed. This consists of four stages. First the lane-changing criteria are checked. Second, the pointers and headways are updated to account for any lane-changes. Third, the vehicles are accelerated according to the updated headways and velocity differences. Finally, the pointers and headways are updated again to account for any vehicles passing their downstream neighbour in the adjacent lane. The new variable values are written to an output file and the loop continues for the specified number of iterations. 84
- 4.3 Lane-changing Case (A). Vehicle n (white circle) changes-lanes (grey circle). Its four neighbours are illustrated with black circles. The white and striped rectangles between the up- and downstream neighbours represent platoons of vehicles. In this case, the upstream platoon must change their front-other pointer from vehicle n to $p_f(n)$ after the manoeuvre. 85
- 4.4 Lane-changing Case (B). Again vehicle n (white circle) changes-lanes (grey circle), its four neighbours are illustrated with black circles and the white and striped rectangles between the up- and downstream neighbours represent platoons of vehicles. In this case, the upstream platoon must change their front-other pointer to vehicle n after the manoeuvre. 85
- 4.5 Pointer update algorithm for a single vehicle n changing lanes. When one or more lane-changes occur during a single time step, the algorithm is applied to each vehicle individually. The temporary scalar vehicle references (n_f , n_b , etc.) are fixed prior to the pointer/headway update. First the algorithm identifies whether vehicles are arranged according to Case (A) or (B). A while loop is then updates the pointers in succession through the platoon of vehicles upstream of, and including n_{bo} and n_b in Cases (A) and (B) respectively. Finally, the pointers of vehicle n and its neighbours are updated. 86
- 4.6 Pointer update algorithm for a passing manoeuvre. The temporary scalar neighbouring vehicle references (n_f , n_b , etc.) are fixed prior to the pointer/headway update. A loop through the vehicle identities determines if such a manoeuvre has occurred and the pointers and headways are updated accordingly. 88
- 4.7 Definition of phase variable θ which measures the normalised position of an arbitrary vehicle n between vehicles n_{fo} and n_{bo} in the adjacent lane. 89

- 4.8 The three generic cases for $I(\theta)$ and $S(\theta)$. The incentive and safety criteria are either satisfied for all θ (Case I_c and Case S_a resp.), over an interval (Case I_b and Case S_b resp.) or not at all (Case I_a and Case S_c resp.). In both Case I_b and Case S_b , the criteria have zeros θ_I and θ_S respectively. 91
- 4.9 Grey scale contours of the mass transfer coefficient λ in the (h_0, h_1) plane. Values of λ between 0 and 1 are illustrated with shades between white and black. The OVRV model is used with parameters $\alpha = 2$, $\beta = 0.5$; the incentive threshold set at $\sigma = 2$. The safety threshold τ is 1.4, 0.8 and 0.2 in panels (a–c) respectively. The solid blue line I_1 marks the $I(1; h_0, h_1) = 0$ boundary, the dashed blue line S_0 marks the $S(0; h_0, h_1) = 0$ boundary (only panels (b) and (c)) and the dashed green line S_1 marks the $S(1; h_0, h_1) = 0$ boundary. The edge of the region where lane-changing occurs $\theta_I = \theta_S$ is marked with a black dashed line (panels (a) and (b)). Note that the size of this region diminishes as τ decreases, in fact no lane-changing occurs in panel (c). There are also large regions where lane-changing does not occur in both panels (a) and (b), in particular when $h_0 > h_1$ 96
- 4.10 Grey scale contours of the mass transfer coefficient λ in the (h_0, h_1) plane for the IDM model with standard parameters [79]. The lane-changing thresholds suggested in [48] are used: $\sigma = 0.1$, $\tau = 4$. As previously, the solid blue line I_1 marks the $I(1; h_0, h_1) = 0$ boundary, the dashed blue line S_0 marks the $S(0; h_0, h_1) = 0$ boundary and the dashed green line S_1 marks the $S(1; h_0, h_1) = 0$ boundary. 98
- 4.11 Comparison of averaged microscopic density and macroscopic numerical QEA data for the OVRV model. Model parameters are $\alpha = 2$, $\beta = 0.5$, lane-changing thresholds are $\sigma = 2$, $\tau = 1.4$ and initial headways are $\bar{h}_0 = 1.5$ and $\bar{h}_1 = 12$. Panels (a–c) correspond to lane-changing rates $\mu = 0.1, 0.01, 0.001$ respectively. There is a good quantitative agreement for values $\mu < 0.01$ where lane-changes are infrequent. Note that the equilibrium density of the microscopic model changes with the lane-changing rate, a feature that is not reproduced by the approximation method. 101

List of Figures

- 4.12 Comparison of averaged microscopic density and the explicit linearised QEA solutions for the OVRV model. Parameters and initial conditions are the same as in Figure 4.11: $\alpha = 2$, $\beta = 0.5$, $\sigma = 2$, $\tau = 1.4$, $\bar{h}_0 = 1.5$, $\bar{h}_1 = 12$. Panels (a–c) correspond to lane-changing rates $\mu = 0.1, 0.01, 0.001$. Again, there is a good quantitative agreement for values $\mu < 0.01$ where lane-changes are infrequent and the equilibrium density is the same for different lane-changing rates μ 101
- 4.13 Mass exchange coefficient in the (ρ_0, ρ_1) plane for the OVRV model with $\mu = 0.01$ and the same parameters and lane-changing thresholds used in Tables 4.2(a) and (b), namely $\alpha = 2$, $\beta = 0.5$, $\sigma = 2$ and $\tau = 1.4$. Overlaid on this are the averaged microscopic data trajectories for the initial headways $(h_0, h_1) = (2.5, 18)$, $(1.5, 6)$ and $(1.5, 12)$ labelled (a–c) respectively. Conservation of mass across the two lanes means that these trajectories are linear. Note that λ is highly non-linear at the I_1 boundary, however this is not reflected in the microscopic data. 104
- 4.14 Evolution of density in lane 0 (panel (a)) and lane 1 (panel (b)) from initially homogeneous conditions. The set-up is identical to that used in Section 4.2. Position on the loop x is plotted against time t and the grey-scale corresponds to the coarse-grained density. Parameters are such that flows are linearly unstable, $\alpha = 0.5$, $\beta = 0.25$; the lane-changing threshold parameters are $\sigma = 0.01$ and $\tau = 2.0$; the initial conditions are $h_0 = 1.8$, $n_0 = 30$, $h_1 = 2.7$ and $n_1 = 20$ ($L = 54$); the lane-changing rate is $\mu = 0.01$. Crosses mark lane-changes from lane 0 to 1 and circles mark lane-changes from lane 1 to 0. A stop-and-go wave emerges and *synchronises* across the lanes. The series of lane-changes that lead up to this are marked by p_1 – p_4 106

- 4.15 Evolution of a small perturbation in density in lane 0 with unstable parameter values $\alpha = 0.5$, $\beta = 0.25$. Lane-changing rate panel (a) occurs with probability $\mu = 0.01$ in unit time, but is not permitted in panel (b) (i.e. $\mu = 0$). Position on the loop x is plotted against time t and the grey-scale corresponds to the coarse-grained densities. Lane-changing threshold parameters in panel (a) are $\sigma = 0.01$ and $\tau = 2.0$ and initial conditions are $h_0 = h_1 = 1.8$ and $n_0 = n_1 = 1000$. A small perturbation is introduced by setting the velocity of one vehicle in lane 0 to $1.1V(h_0)$. Crosses mark lane-changes from lane 0 to 1 and circles mark lane-changes from lane 1 to 0. The number of steps is chosen so that the leading and trailing edges of the resulting disturbance (f_d and f_u respectively) do not interact. Lane-changes appear to affect the long-term spatial development of individual jams (cf. Z_1 and Z_2). 108
- 4.16 Evolution of traffic with heterogeneous driver behaviour. Position on the loop x is plotted against time t and the colour-scale corresponds to coarse-grained density for each lane, labelled ρ_0 and ρ_1 . Crosses mark lane-changes from lane 0 to 1 and circles mark lane-changes from lane 1 to 0. The vehicle population comprises equal numbers of slow ($\gamma = 0.6$) and fast ($\gamma = 1.0$) cars, which are initially distributed evenly between lanes but whose order is chosen randomly. The initial spacing is constant $h_* = 2$, as is the number of vehicles starting in each lane $n_0 = n_1 = 400$. Model parameters are $\alpha = 0.6$, $\beta = 0.3$, lane-changing thresholds are $\sigma = 0.1$, $\tau = 0.5$ and lane-changes are deterministic (i.e. $\mu = 1/\Delta t$). In contrast to Figures 4.14 and 4.15, we observe that the density profiles become *asynchronised* (compare areas A and B). In addition to stop-and-go waves propagating upstream (J), density waves due to heterogeneity also propagate *downstream* (H). 109

Chapter 1

Introduction

“Euclid’s fifth proposition, perfect in conception (5)”, Hendra 98(7d) [31]

1.1 Motivation

More than 25 million people take to the roads everyday in the UK [20, 23]. Congestion is inevitable. Travel however is essential to the success of the economy, from the transport of goods to the daily commute. Delays cost money: the estimated cost to businesses in the UK is £20 billion annually [68]. Thus improvements to the nation’s highways, or ‘strategic network’, are of critical importance to the Government [22].

The solution to congestion is not simply to build new roads. Rather, it is “new technology and other smart ways of getting more out of existing capacity” [3] that is the primary focus of the Highways Agency (HA), who are responsible for operation and maintenance of the strategic road network in England. The field in which digital technology is integrated into physical roadside infrastructure is known as Intelligent Transport Systems (ITS), and there are three key areas (Figure 1.1) where ITS methodologies are applied:

1. **Information.** Dissemination of real-time traffic information using Variable Message Signs (VMS), websites, web-cams, radio stations and information points at service stations.
2. **Management.** Dynamic control of speed limits (such as ‘Controlled Motorways’ on the M25 and M1), traffic signals at junctions (known as ‘Ramp Metering’) and large-scale schemes to increase capacity (such as Active Traffic Management (ATM) on the M42).
3. **Safety.** Automatic queue protection, delivered by the MIDAS (Motorway Incident Detection and Automatic Signalling) system, alerts drivers approaching stationary traffic; control centres with extensive CCTV coverage coordinate Traffic Officers and Incident Support Units who can respond quickly to incidents and aid motorists at risk.



Figure 1.1. Variable Message Signs (a) are a flexible means of informing drivers about current road conditions, expected travel times, road works etc. The Active Traffic Management scheme (b) introduced on the M42 around Birmingham combines dynamic control of speed limits and hard shoulder operation during peak periods to increase capacity, smooth traffic flow and reduce vehicle emissions [2]. The Highways Agency have recently deployed Traffic Officers (c) to provide support for motorists in difficulty and for rapid incident response. Images reproduced courtesy of the Highways Agency website <http://www.highways.gov.uk> [1, 3].

Modern highways are currently equipped with CCTV and inductance loops to monitor traffic conditions and in the future hi-tech instruments such as microwave/radar, infrared, fibre optic, magnetic anomaly and acoustic detectors [1] may be deployed. Inductance loops are coils of wire buried in the road surface that detect passing vehicles. These are typically placed at 500m intervals in Western Europe, the majority of which are double loop detectors that measure *flow* (number of vehicles that pass per unit time), *occupancy* (the proportion of time a detector is covered) and vehicle *velocities* and *lengths* (velocities and lengths are estimated using the signal on and signal off times on each half of the loop, and thus may not be observed directly with single inductance loops). The data recorded by these detectors is usually averaged over one minute intervals before being sent to a control centre. One can also approximate the mean traffic *density* by dividing flow by velocity.

It was noted as early as the 1930s [30] that there is an implicit relationship between average velocity and flow. If we plot inductance loop data in this plane (see Figure 1.2(a)), we find that the flow of traffic is maximised at intermediate velocities (approximately 80km/h). To understand this, consider how the average velocity changes with density, see Figure 1.2(b). When the distance between vehicles is large, and hence the density is low, drivers are free to drive at their desired speed, thus the average velocity is approximately the speed limit. Conversely, when the distance between vehicles is small, and the density high, one must drive slowly for safety reasons.

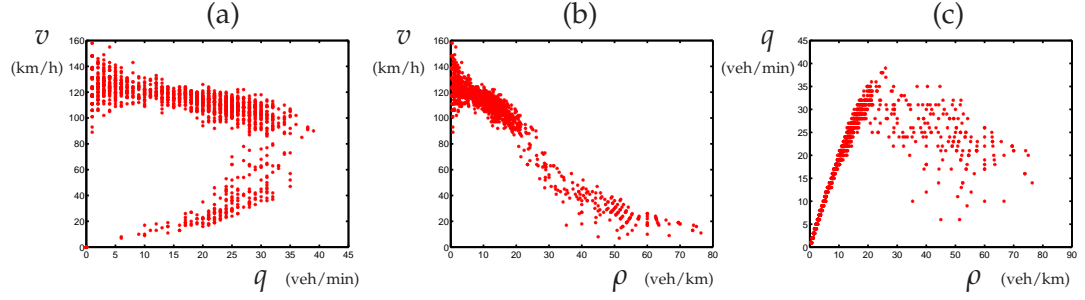


Figure 1.2. Data taken from inductance loops over one minute averages. The relationship between velocity v and flow q (panel (a)), velocity v and density ρ (panel (b)) and flow q and density ρ (panel (c)) are illustrated. There is a uni-modal relationship between flow and velocity/density, and velocity decreases with density.

Hence the average velocity decreases as the density increases. Since fast moving traffic is sparse, the number of vehicles that pass a point in unit time is low. Similarly, if vehicles are close together but moving slowly, the flow of vehicles is also low. Thus flow is maximised at intermediate velocities and densities.

Thus central to the study of traffic flow is the consequent relationship between flux and density, known as the *Fundamental Diagram* (see Figure 1.2(c)). Flow increases approximately linearly with density up to concentrations of 20-30 vehicles/km/lane. Over this range the average velocity of vehicles is approximately constant, i.e. the speed limit. At high densities vehicles must slow down, hence the flow decreases with density but interestingly the data points become widely scattered. This scatter has received considerable attention in the traffic science literature and has led to an intense study of the macroscopic (large length scales and long time scales) spatio-temporal patterns observed in traffic data (see Figure 1.3).

Several spatio-temporal phenomena are common to daily traffic patterns around the world. Typically very early in the morning (or late at night), traffic flows smoothly, a regime known as ‘free flow’. In this regime vehicles interact only very weakly with each other and each vehicle drives more or less steadily at its own chosen velocity, facilitated by lane-changes. Consequently one tends to observe that average velocities vary between lanes. Moreover, an observer standing by the side of the road recording the passing time of each vehicle will discover a Poisson distribution for inter-arrival times, emphasising that vehicles’ motions are largely independent of each other.

In contrast, during periods of congestion (typically the morning and evening commutes), inter-arrival times are no longer Poissonian, indicating that vehicles’ motions are correlated with each other. In addition congestion appears to organise itself into spatio-temporal patterns (see Figure 1.3), the cause and classification of which is still under intense debate [18, 44, 73, 83]. As the density of vehicles increases, a decrease in

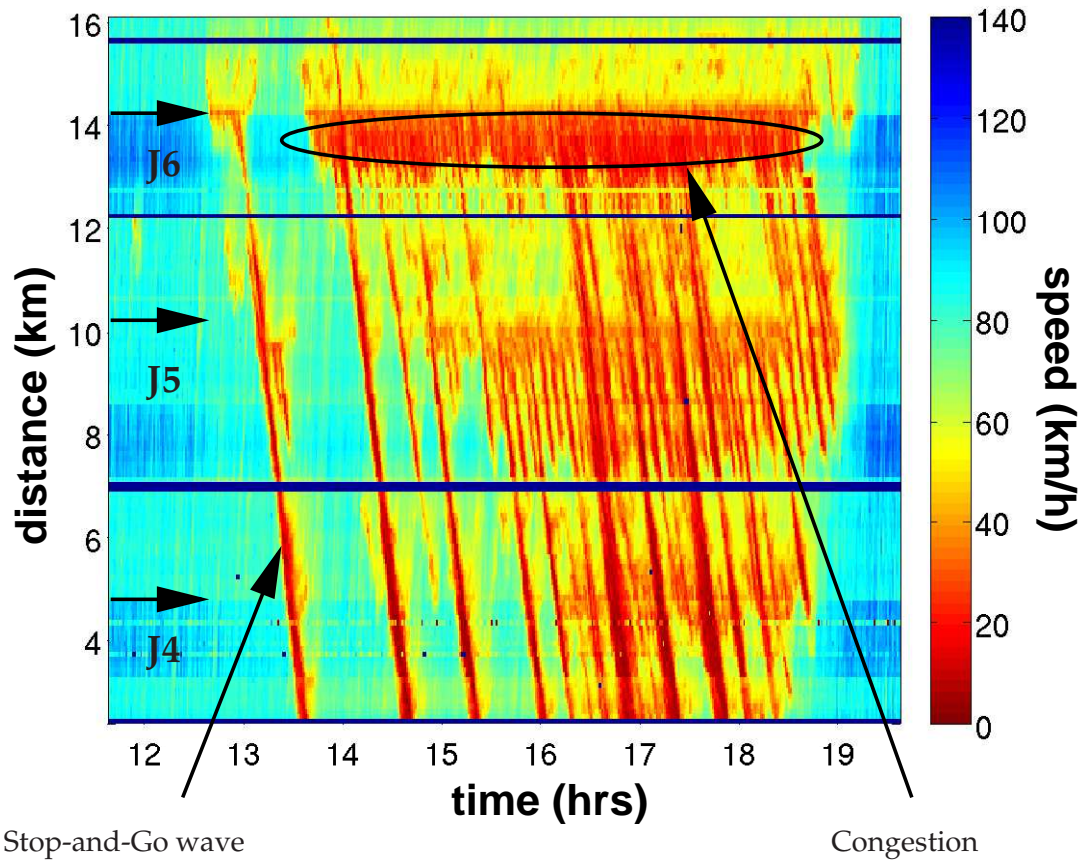


Figure 1.3. An example of spatio-temporal patterns in inductance loop data from the north-bound M42 Motorway in the UK. Distance in kilometres is plotted on the vertical axis against time in hours on the horizontal axis. Colour corresponds to speed in kilometres per hour. There are three junctions present in this figure, the position of which are labelled with J4–J6. Drivers travel up the page, in the direction of increasing distance and time. Vehicles are forced to slow down considerably around Junction J6, and numerous stop-and-go waves can be seen propagating from this location against the flow of traffic.

velocity variance is observed [47] and queues of vehicles form upstream of junctions with a large influx of vehicles. Of particular interest are so called ‘stop-and-go waves’ (also known as ‘phantom-jams’, or ‘wide-moving jams’) which are near-stationary columns of vehicles up to a kilometre long that propagate *upstream*, sometimes for tens of kilometres. A striking characteristic of these jams is the near universal velocity of their downstream front, approximately 15-20km/h [73]. Stop-and-go waves only dissipate when the incoming upstream vehicle flow becomes sufficiently low. It is generally thought that the creation of stop-and-go waves is due to some type of instability, and there are several theories that use combinations of linear and non-linear stability analysis to describe the growth of these traffic jams [6, 24, 38, 46, 52, 69, 82, 83].

Thus at the heart of transportation science is the identification of fundamental spatio-temporal pattern formation mechanisms. Why is this of any wider importance? The adoption of ITS strategies has introduced a wealth of areas where technological advances can potentially cut delays, reduce emissions and improve reliability [1, 21]. Thus, there is value in developing basic understanding of traffic patterns since this ultimately leads to the improvement of traffic engineering schemes.

In this thesis, we are concerned with underlying theory of spatio-temporal pattern, which to date has largely involved highly simplified settings such as single lanes, spatial homogeneous roads and identical drivers [6, 24, 46, 64, 69]. Our approach is to relax such assumptions and to generalise the existing analysis. Furthermore, our philosophy is multiscale in that we try to understand the output of microscopic (individual driver) models and their simulation using approximate macroscopic (partial differential equation) techniques that relate variations in density, velocity and flow. We make three separate, yet ultimately connected contributions:

1. **Spatial Heterogeneity.** Most existing analyses deal with traffic patterns in a homogeneous setting. But motorways are not spatially uniform: they have junctions where traffic exits and enters; the number of lanes can change; there may be road-works, or other physical attributes that change capacity such as hills or curves etc. Furthermore, spatial heterogeneity *itself* may generate spatio-temporal traffic patterns. We develop an analysis based on a highly simplified situation in which vehicles drive round a ring-road with a simple capacity-drop bottleneck. We find rich families of patterns even at parameters where classical stop-and-go waves are impossible.
2. **Instability.** The linear stability analysis of microscopic traffic models dates back to the 1950s [12, 26] and has been used to explain the formation of (macroscopic) stop-and-go waves. More recent work [24, 69] has concentrated on nonlinear

stability properties. However, we have two new contributions, at the *entirely linear level*. Firstly, we can show that there is a simple formula that determines the linear stability of a wide range of models, which may be generalised to consider heterogeneous driver populations (mixes of cars and trucks, for example). Secondly, it seems that the convective propagation of disturbances on the open-road is a relatively unstudied area, and we make new insights using group velocity computations and transform techniques.

3. **Lane-Changes.** Classical mathematical (e.g., stability) analyses of highway traffic suppose that the order of vehicles is preserved, and consequently they do not generalise easily to multi-lane situations. Based on microscopic dynamics, we have derived a macroscopic approximation that describes the lumped flow rate between lanes. The macroscopic model is then verified with a simulation framework that we have also developed. Moreover, our simulation experiments have uncovered a wealth of dynamical behaviour that has not been reported in the literature to date.

1.2 Overview of Mathematical Modelling Methods

Road transport modelling may be separated into two broad categories: 1. Traffic on networks of minor routes that make up towns and cities; and 2. Traffic on major arterial trunk roads (motorways or highways in the UK) that connect towns and cities. This thesis is solely concerned with the latter, where the complexity that arises from large numbers of driver interactions has led to families of mathematical models posed at different scales. Broadly speaking, models may be either microscopic, and deal with the trajectories of individual vehicles (see Section 1.2.1), or they may be macroscopic (see Section 1.2.3) and deal with average fluid-like quantities. See [13, 33, 63] for reviews. The multiscale approach in this thesis uses approximate macroscopic models to understand the emergent dynamics of microscopic models and their simulation, hence we briefly describe in Section 1.2.2 how these different scales may be connected.

1.2.1 Microscopic Models

Microscopic models attempt to describe the behaviour of individual drivers. There are two main types: 1. Car-following models, which use ordinary or delay differential equations that specify the trajectories of discrete vehicles in continuous space and time, whereas 2. Cellular Automata (CA) model vehicles as discrete entities moving in discrete space and time, see Figure 1.4. The remainder of the thesis is concerned

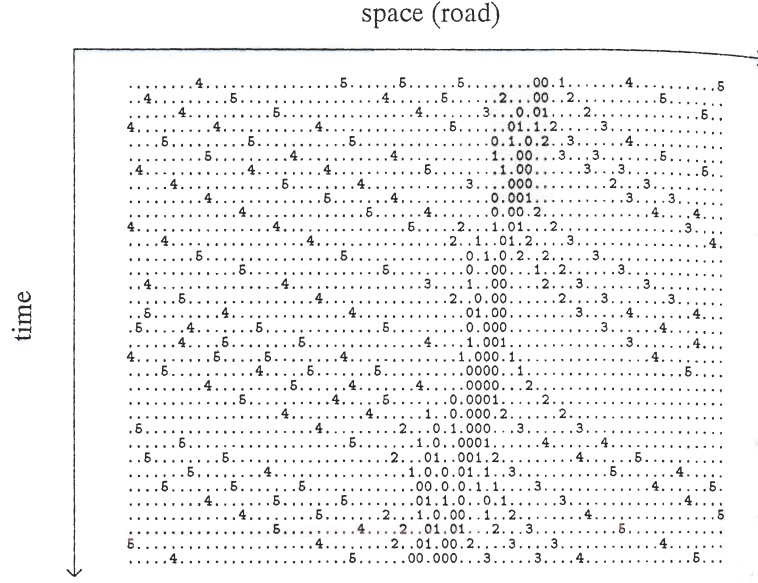


Figure 1.4. Spatio-temporal evolution of the Cellular Automata model proposed by Nagel and Schreckenberg [64] to emulate traffic flow. The road is composed of discrete lattice sites i that are either occupied by a vehicle or are unoccupied at any discrete instance in time j . Vehicles have integer values of velocity $v = 1, 2, \dots, v_{\max}$. Vehicular velocity is used to denote which lattice sites are occupied in the simulation depicted above for the case where $v_{\max} = 5$. A period is used to denote unoccupied sites. The lattice is updated from time j to $j + 1$ using the following algorithm: 1. *Increase* the velocity of vehicles with $v < v_{\max}$ by 1, $[v \rightarrow v + 1]$; 2. *Decrease* the velocity of vehicles with $h < v$ free lattice spaces ahead to $h - 1$, $[v \rightarrow h - 1]$; 3. *Decrease* the velocity of vehicles with $v > 0$ by 1 with probability p , $[v \rightarrow v - 1]$; 4. *Move* all vehicles forward by v lattice sites. These rules ensure collision-free dynamics. CA models require very little computational time during each update, hence systems with large numbers of vehicles can be modelled over long time intervals. These models also exhibit transitions from laminar flow to stop-and-go behaviour (an example of the latter is illustrated above). Figure reproduction courtesy of K. Nagel and M. Schreckenberg [64].

with car-following models only, which form the basis of our numerical simulation experiments. See [10] for a historical review.

We suppose each vehicle is a particle with a position x_n and velocity

$$\dot{x}_n = v_n, \quad (1.1)$$

where dot denotes differentiation with respect to time and the subscripts n reference vehicles (see Figure 1.5). Initially, we consider single lanes of traffic where overtaking is not allowed and the order of vehicles is preserved. We adopt the convention that vehicles are numbered sequentially *upstream*. Drivers respond to various stimuli; of

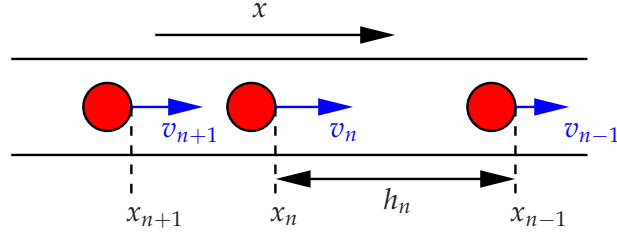


Figure 1.5. Illustration of a simple car-following model set up. Vehicles occupy a single lane travelling in the increasing x -direction (in this case to the right). Associated with each vehicle is a position x_n and a velocity v_n . The spacing between cars is known as the headway h_n .

particular importance is the distance to the car in front, known as *headway*,

$$h_n = x_{n-1} - x_n. \quad (1.2)$$

The velocity difference between vehicles is consequently a rate of headway,

$$\dot{h}_n = v_{n-1} - v_n. \quad (1.3)$$

Driver behaviour is mimicked by modelling acceleration response to driving conditions. This idea emerged in the 1950s, at which time models typically took the form

$$\dot{v}_n(t) = c v_n^m(t) \frac{\dot{h}_n(t - \tau)}{h_n^l(t - \tau)}, \quad (1.4)$$

where τ represents a reaction time [12, 26] and l and m are positive integer exponents. In this model, one tends to accelerate if the vehicle in front is driving away, or decelerate if one is closing on the vehicle ahead, and moreover, the response is accentuated the smaller the headway. These early studies included stability analysis and some comparison with experimental data.

A limitation of (1.4) is that there is no well-defined equilibrium velocity for a given headway. That is, any combination of v_n and h_n satisfy $\dot{v}_n = 0$ provided $\dot{h}_n = 0$, which is at odds with Figure 1.2(b). This observation leads to the *Optimal Velocity* (OV) model [6]

$$\dot{v}_n = \alpha \{V(h_n) - v_n\}, \quad (1.5)$$

where drivers adjust their velocity towards a safe optimum $V(h_n)$ based on their headway. Here $\alpha > 0$ is known as the sensitivity parameter. The OV function V is increasing to represent the fact that a safe driving velocity increases with headway up to the given speed limit. Equilibria of the OV model (1.5) are known as uniform flow solutions and are parameterised in terms of an equilibrium spacing h_* , such

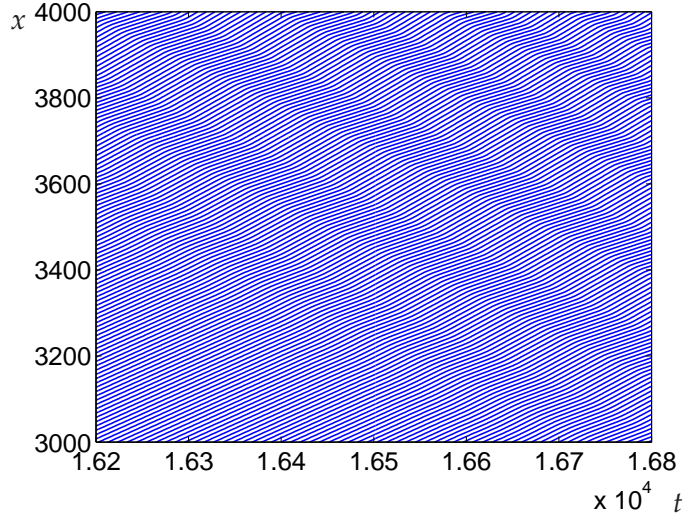


Figure 1.6. Microscopic simulations of the OV model (1.5) with OV function (1.6) and unstable parameters: $\alpha = 1.9$, $h_* = 2$. The initial velocity of a single vehicle is perturbed by a small amount. This is amplified by subsequent vehicles, eventually forming stop-and-go waves which can be seen in the top-right-hand corner as a striped pattern in the vehicle trajectories due to a sharp change in velocity. Note that the flow is uniform in the bottom-left-hand corner.

that vehicle positions are simply translated forwards at the equilibrium velocity, i.e. $x_n = -h_*n + V(h_*)t$ and $v_n = V(h_*)$.

When the OV function is S-shaped, for example the commonly used non-dimensional form

$$V(h) = \tanh(h - 2) + \tanh(2), \quad (1.6)$$

interesting dynamics can occur. It was shown by Bando *et al.* [6] that uniform flow solutions are linearly unstable if

$$\frac{\alpha}{2} < V'(h_*). \quad (1.7)$$

Using numerical simulations, Bando *et al.* then illustrated that nonlinear effects curb the growth of perturbations, giving rise to large amplitude solutions that resemble stop-and-go waves (see Figure 1.6). Later, in Chapter 3, we will present a substantial generalisation of Bando's linear stability analysis.

However the OV model is limited in that it does not include the effect that the velocity difference with the car in front has on driver acceleration. Hence we introduce a simple adaptation of the OV model to include *relative velocity*, namely the OVRV model

$$\dot{v}_n(t) = \alpha \{V(h_n) - v_n\} + \beta \dot{h}_n, \quad (1.8)$$

in which velocity difference is introduced linearly with the additional parameter $\beta >$

0. In more complex models the relative velocity term is scaled with headway, since the importance of the velocity of the vehicle in front diminishes as the distance to it increases. An example of such a model is the *Intelligent Driver Model* (IDM), developed by Treiber *et al.* [79], which takes the form

$$\dot{v}_n(t) = a \left[1 - \left(\frac{v_n}{v_{\max}} \right)^\delta - \left(\frac{s(v_n, \dot{h}_n)}{h_n} \right)^2 \right], \quad (1.9)$$

where

$$s(v_n, \dot{h}_n) = s_0 + s_1 \sqrt{\frac{v_n}{v_{\max}}} + T v_n - \frac{\dot{h}_n v_n}{2\sqrt{ab}}. \quad (1.10)$$

There are eight parameters in total, and standard values and their interpretations can be found in [79]. An equilibrium headway relation can only be determined numerically in the most general case, however setting $s_0 = s_1 = 0$ yields an explicit equilibrium velocity-headway relation.

Other features that have been introduced into car-following models include driver reaction time delay and multiple look-ahead (where the headway to several cars in front is considered). However, in this thesis we are concerned with models that fall into the general class described by

$$\dot{h}_n = v_{n-1} - v_n, \quad (1.11)$$

$$\dot{v}_n = f(h_n, \dot{h}_n, v_n), \quad (1.12)$$

from which one can identify an equilibrium velocity function $V(h)$ which satisfies

$$f(h_*, 0, V(h_*)) = 0 \quad \text{for all } h_* > 0, \quad (1.13)$$

with the properties

$$1. \quad V(0) = 0, \quad (1.14)$$

$$2. \quad V' \geq 0 \quad \text{and} \quad (1.15)$$

$$3. \quad V(h) \rightarrow V_{\max} \quad \text{as } h \rightarrow \infty. \quad (1.16)$$

For rational driver behaviour, the partial derivatives of f should satisfy

$$D_h f, \quad D_{\dot{h}} f > 0 \quad \text{and} \quad D_v f < 0. \quad (1.17)$$

This simply states that drivers increase their acceleration in response to an increase in headway or relative velocity but tend to decrease their acceleration as their own velocity increases.

The simplest method of analysing microscopic simulation results is to observe the motion of vehicles in space and time. However the explicit positions of vehicles are not present in the general form (1.11) and (1.12); hence to obtain them we must supplement (1.11) and (1.12) with the kinematic equation (1.1). Alternatively we could write the headway variables in terms of positions and velocities using (1.2) and (1.3) and swap the velocity difference (1.11) with the kinematic equation (1.1). This yields

$$\dot{x}_n = v_n, \quad (1.18)$$

$$\dot{v}_n = f_n(x_{n-1} - x_n, v_{n-1} - v_n, v_n). \quad (1.19)$$

In these variables, uniform flow solutions are relative, rather than absolute, equilibria.

1.2.2 Coarse-graining Microscopic models

The simplest way to relate microscopic and macroscopic variables is via $\rho(x_n, t) = 1/h_n(t)$, although it is well-known [8, 56] that this relationship holds exactly only for equilibrium situations. To model more general non-equilibrium dynamics, macroscopic continuum models have been derived from microscopic Car-following models via perturbation [8] and coarse-graining [56] methods. However, these approximations necessarily introduce errors due to the truncation of terms in series expansions. In this thesis we calculate coarse-grained densities [56] from the distributional point density of discrete vehicles by convolving with a normalised Gaussian test function. A macroscopic flow variable $q(x, t)$ may be obtained in a similar way by coarse-graining the discrete velocity v_n , and then a coarse-grained velocity is given by $v(x, t) := q(x, t)/\rho(x, t)$.

Following the method of Lee, Lee and Kim [56], the distribution point density and flux are given by

$$\hat{\rho}(x, t) = \sum_n \delta(x_n(t) - x), \quad (1.20)$$

$$\hat{v}(x, t) = \sum_n v_n(t) \delta(x_n(t) - x), \quad (1.21)$$

where x_n and v_n are microscopic vehicle positions and velocities respectively and $\delta(x)$ is the Dirac delta function. In general, these point-fields are then convolved with an envelope function $\phi(x, t)$. This function must be non-negative valued, normalised $\int \phi(x, t) dx dt = 1$ and peaked at $x, t = 0$. For concreteness, in this thesis we only coarse-grain in *either* space or time using a standard Gaussian test function

$$\phi(z) = \frac{1}{w\sqrt{2\pi}} \exp\left(-\frac{z^2}{2w^2}\right). \quad (1.22)$$

Thus, the coarse-grained density and flux are then

$$\rho(x, t) = \int \phi(x - x') \hat{\rho}(x', t) dx' \quad (1.23)$$

$$q(x, t) = \int \phi(x - x') \hat{v}(x', t) dx', \quad (1.24)$$

respectively when coarse-grained in space at time t .

The choice of the Gaussian test function width w affects the resulting coarse-grained profiles. If w is chosen too large, then fine details of the velocity and density may be lost. However, choosing a value of w that is too small leads to noisy results. Thus one must experiment with the width w in each case to capture the correct behaviour and check that the coarse-grained profile reflects the underlying microscopic dynamics correctly. As a guide, we typically use widths w of the order of the largest headway, however the exact value will be described in each instance.

1.2.3 Continuum Models

Macroscopic models associate continuous densities $\rho(x, t)$ and velocities $v(x, t)$ with the traffic state at positions x and times t ; hence they are commonly referred to as continuum models. On a region of road with no entrances or exits, the number of vehicles remains constant unless there is an influx from upstream or an out flow downstream. This property can be described mathematically with a conservation law,

$$\rho_t + (\rho v)_x = 0, \quad (1.25)$$

where the subscripts denote partial differentiation. To close (1.25), it was proposed by Lighthill, Whitham and Richards (LWR) [58, 70] that the velocity of traffic has a functional dependence on the density, $\mathcal{V}(\rho)$. This takes the form of a decreasing function to emulate empirical data (see Figure 1.2(b)).

Supplementing (1.25) with $v = \mathcal{V}(\rho)$, one finds that the density is constant along lines (or line segments) in the (t, x) plane with gradient (i.e. velocity) $Q'(\rho)$. These are known as characteristics. The spatio-temporal evolution of density is subsequently determined by following the progress of characteristics. If a fast region of low density is initially downstream of a slow region of high density, the density profile spreads out in what is known as an expansion fan. If the situation is reversed however, the density profile steepens and a discontinuous shock develops. The velocity of shock-waves can be obtained from the gradient of the chord on the fundamental diagram connecting the densities up- and downstream of the shock front (see Figure 1.7).

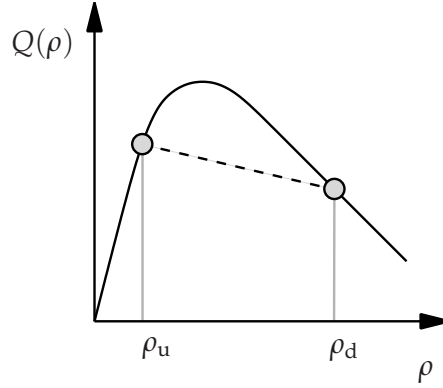


Figure 1.7. A sketch of the Fundamental Diagram (compare with the empirical data in Figure 1.2(c)). Characteristic speeds are given by the gradient $Q'(\rho)$, and shock speeds correspond to the gradient of the chord connecting the up- and downstream densities, ρ_u and ρ_d respectively. In this case the shock velocity is negative and hence moves against the flow of vehicles.

Sharp transitions in density are indeed seen in empirical traffic data, which typically travel against the flow of traffic (see Figure 1.3). Such behaviour is predicted by hyperbolic conservation laws and there are many other examples where they can be used effectively [82]. However LWR models cannot reproduce every feature of empirical traffic data, in particular they fail to capture stop-and-go waves. Such structures, with sharp up- and down-stream fronts, cannot persist in LWR type models since one of the fronts inevitably bleeds away as an expansion fan (see Figure 1.8). To remedy this, higher order terms have been introduced, for example

$$v_t + vv_x = \alpha [\mathcal{V}(\rho) - v] - \mu \frac{\rho_x}{\rho} + \nu \frac{v_{xx}}{\rho}. \quad (1.26)$$

The left-hand side of (1.26) is the total derivative of velocity v ; hence (1.26) is likened to momentum balance equations in fluid dynamics and (1.26) is often referred to as the ‘momentum equation’. The first term on the right-hand side models relaxation to a safe driving velocity given by the velocity function $\mathcal{V}(\rho)$, which can be related to the equilibrium velocity function of car-following models described in the previous section. The remaining terms are a so called *pressure term* proportional to ρ_x and a *diffusion term* proportional to v_{xx} . Examples of the form (1.26) include the Payne-Whitham model [36, 82], the Kühne model [52] and the Kerner-Konhäuser model [46]. Similar models have also been derived from microscopic models [8, 56], further details of which are presented in Chapter 2.

Instability can occur in higher order models from which stop-and-go waves can develop. However, the inclusion of a momentum equation introduces several unde-

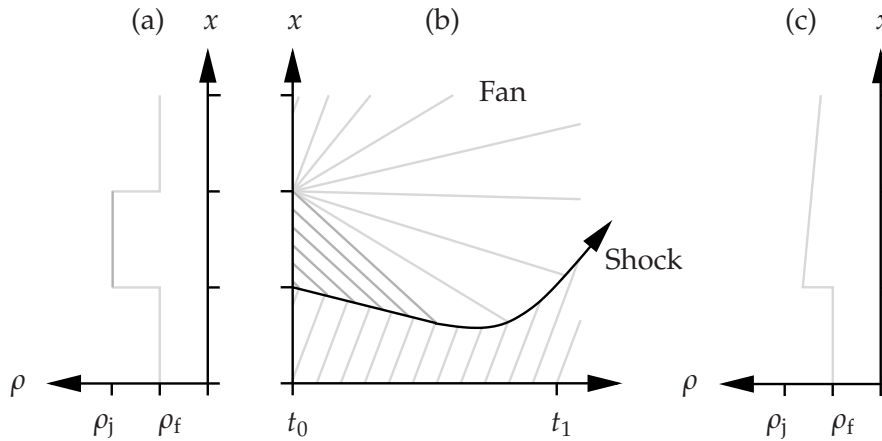


Figure 1.8. Sketch in (t, x) of the characteristics for an initial pulse in density with sharp up- and downstream fronts (panel (b)), with density profiles at times t_0 and t_1 (panels (a) and (c) respectively). The pulse density is initially ρ_j , corresponding to negative characteristic velocity, and the background density is ρ_f , corresponding to positive characteristic velocity. The downstream front forms an expansion fan. The upstream front propagates as a shock against the flow of traffic until the expansion fan refracts its trajectory and eventually annihilates the pulse altogether.

sirable features [16]; in particular, one of the characteristic speeds allows information to propagate down-stream faster than the traffic velocity. There have been several attempts to repair these problems [5, 29, 85], however in most cases it can be shown that the fix eliminates instability [83]. It is an open problem, considerably beyond the scope of this thesis, to identify macroscopic models which are provably well-posed yet have sufficiently rich dynamics to produce the full range of empirically observed traffic patterns.

1.3 Outline of Thesis

The three distinct problems tackled in this thesis are described in each of the three subsequent chapters. In Chapter 2, we illustrate the types of stationary pattern that emerge from microscopic simulations on a ring-road with a simple bottleneck. We use a numerical technique to coarse-grain the microscopic simulation data, from which we obtain macroscopic density profiles. We then analyse these states at progressively smaller scales. First we use simple conservation principles and the method of characteristics to obtain a phase diagram relating model parameters to each steady state. We then use second order continuum models to qualitatively reproduce the internal front structure. Finally, we return to the microscopic model to describe the features missing from the macroscopic analysis. In short, even this very simple spatially het-

erogeneous setting displays rich traffic patterns which require sophisticated tools for their explanation.

In Chapter 3, we turn our attention to the classical problem of the stability of single-lane traffic. We describe how the linear stability of a wide range of car-following models can be captured by a simple condition for both the ring-road and the open-road cases. Furthermore, we extend this analysis to include heterogeneous driver behaviour, where the acceleration rules of individual vehicles can be different. We then develop an analytical theory based on group velocity, which distinguishes between models with absolute and convective linear instability (since real data suggests that traffic is only convectively unstable in the upstream direction, measured relative to the frame of the road). In the open road setting, we outline a novel method of analysis for a localised disturbance growing/decaying progressively through a chain of vehicles. Finally, we illustrate how this technique might be used to obtain asymptotic formulae using transform methods.

In Chapter 4, we consider two-lane traffic where overtaking is allowed. Firstly we describe the multilane set-up and our numerical simulation framework. To date, the complexity of lane-changing models has left the subject area largely devoid of mathematical examination. We analyse the simple case where vehicles are identical and lane-changes are rare. Using features common to car-following models with rational acceleration behaviour, we obtain a series of conditions that ensure safe lane-changing for a generic lane-changing model. Furthermore, we derive a macroscopic approximation to the lane-changing rate between two lanes of traffic with different density and, via expansions, suggest how higher order terms can be incorporated into macroscopic lane-changing models. The analytical theories are shown to be in good agreement with numerical simulation. Finally, we present simulation experiments that both challenge the assumptions of our multilane analysis and uncover phenomena that have not been reported on in the traffic literature to date.

In Chapter 5, we summarise our results and discuss a number of unresolved issues and open problems. We emphasise the new developments, and how these may be used to tackle the fundamental issues at the heart of spatio-temporal pattern formation mechanisms in empirical traffic data.

Chapter 2

Multiscale Modelling of Bottleneck Pattern Formation

“Computer program reversing trend to Operations Research (6)”, Chifonie 12(5a) [31]

2.1 Motivation

Since the mid 1990s, there has been much interest in the spatio-temporal dynamics of highway traffic flow models (see [13, 33, 44]). In particular, Kerner and collaborators [46] understood that the classical gas dynamics models developed by Lighthill, Whitham and Richards [82] could not capture all of the qualitative features in empirical data collected from inductance loop systems. Consequently, a major goal has been to develop better models with one objective being that they reproduce the so-called *spontaneous flow breakdown* effect, where small disturbances in the traffic may amplify in some circumstances to give large amplitude *stop-and-go* waves.

Intensive studies of such complex dynamics in a spatially uniform setting can be found in the applied mathematics and theoretical physics literature. However, data collected from inductance loop systems suggests that spatio-temporal wave patterns are usually triggered by spatial heterogeneity [43, 79]. Practical examples include a change in the speed limit (e.g., restrictions due to roadworks), a change in the number of lanes, a merge with an input road, or a change in grade. Such changes that result in a drop in capacity are known as bottlenecks, and their flow patterns have been studied extensively from the empirical point of view [7, 9, 11]. However, the analysis of bottlenecks in the engineering literature has been restricted to the use of simple macroscopic conservation principles [65, 66], extending most recently to the analysis of moving bottlenecks [17, 55, 67] that model the interaction of fast traffic streams with individual slow-moving vehicles. Numerical simulations of microscopic models [45, 79] indicate that very rich dynamics may be induced by bottlenecks, and in our view there is presently a lack of an underlying theory that explains the mechanisms

for complex traffic pattern formation caused by road heterogeneities. The highly simplified situation that we analyse here should be seen as a first step in developing such a theory, rather than an attempt to model any one real-world traffic situation in an accurate way.

In this chapter, we analyse the wave structures of the microscopic Optimal Velocity (OV) model [6], in which vehicles are modelled as discrete entities moving in continuous time and space. The parameters are chosen so that uniform flows are stable. Consequently stop-and-go waves are impossible, yet we force patterns by introducing spatial heterogeneity in the form of a crude bottleneck model. In fact, we expect that the patterns which we discover will also be found in broader classes of car-following model provided the bottleneck is implemented in the same sort of way, i.e. by a (possibly nonlinear) scaling down of the equilibrium flow-density relationship.

In order to limit the variety of patterns that may form and to simplify the study, we analyse a ring road which forms a closed loop. This removes the need to prescribe an upstream boundary condition and also prohibits certain types of wave structure, e.g. expansion fans cannot persist as $t \rightarrow \infty$ since their up- and down-stream tails would eventually interact with each other. Our set-up is thus identical to the recent paper by Hanaura *et al.* [32], but our analysis goes further in that we extract a phase diagram from the resulting flow and density balances and we explain in detail the wave selection principles that were previously overlooked. Furthermore, the latter part of this chapter is concerned with a detailed analysis of wave structure.

In what follows, we make use of the idea that the modelling of traffic flow can be understood at two distinct levels: microscopically, whereby each vehicle is considered individually, and macroscopically whereby traffic is considered as a continuous fluid. Often, these two approaches are treated independently. In contrast, the key point of this chapter is that the simple discrete model that we consider here develops stable stationary patterns as $t \rightarrow \infty$, which can be understood by drawing parallels with continuum models. Hence our approach is fundamentally multiscale in that we use macro theory to describe micro phenomena.

This chapter is set out as follows. In Section 2.2, we describe the OV model set-up, including precise details of how the spatial heterogeneity is applied. Then we outline our numerical simulation and coarse-graining procedure, and we show results of three numerical experiments with qualitatively different solution structures as $t \rightarrow \infty$ (see Fig. 2.1). Sections 2.3 and 2.4 investigate these experiments using classical kinematic wave theory, firstly by analysing simple flow and density balances and then by using characteristic arguments to explain the wave selection principles. We then consider the evolution of characteristics from the initial value problem in Section 2.5, and use

a shock tracking technique to obtain the transition boundary between different solution structures. In Section 2.6, we investigate the structures of the various shock waves, using a travelling-wave phase-plane of a continuum (PDE) model that approximates the microscopic dynamics on a macroscopic scale [56]. In fact this method is insufficient in that it cannot explain all wave features and, therefore, a spectral analysis of the original microscopic model is carried out in Section 2.7. Finally in Section 2.8, we conclude and indicate the success and failures of higher order continuum models in explaining the fine details of the solution structure. This work has been published [81] with co-authors Dr R.E. Wilson and Dr P. Berg, and is based on material presented at Traffic and Granular Flow 2005 [80].

2.2 Problem Set-Up and Simulation Results

We consider the traffic patterns formed by a large number N of identical vehicles driving on a unidirectional single-lane ring-road of length L . Overtaking is not considered. Vehicles move in continuous space x and time t , and their displacements and velocities are labelled $x_n(t)$ and $v_n(t)$ respectively. We suppose that the direction of motion is in increasing x , and moreover that vehicles are labelled $n = 1, 2, \dots, N$ in the upstream direction. For the vehicles' equations of motion, we extend the well-known Optimal Velocity (OV) car-following model (see Section 1.2.1)

$$\dot{x}_n = v_n, \quad (2.1)$$

$$\dot{v}_n = \alpha \{V(h_n; x) - v_n\}. \quad (2.2)$$

Here, as usual, dot denotes differentiation with respect to time, and the rate constant $\alpha > 0$ is known as the sensitivity. The variable $h_n := x_{n-1} - x_n$ gives the headway, or gap to the vehicle in front, and loosely speaking the OV model describes the relaxation of traffic to a safe speed which is defined in terms of this gap. Note that under open boundary conditions one would need to prescribe the trajectory of the lead vehicle 1, but on the ring-road we assume merely that it follows vehicle N , so that $h_1 = L + x_N - x_1$.

The heterogeneous OV function that we use here takes the form

$$V(h_n; x) := \begin{cases} r_B V(h_n), & 0 \leq x \bmod L < \hat{L}L, \\ V(h_n), & \hat{L}L \leq x \bmod L < L, \end{cases} \quad (2.3)$$

and which is thus scaled down by a reduction factor $0 < r_B < 1$ for a proportion $0 < \hat{L} < 1$ of the ring-road under consideration. Note that for the sake of simplicity,

the vehicles' displacements $x_n(t)$ are set-up as monotone increasing and unbounded, although henceforth we interpret all results modulo L .

In (2.3), V with a single argument denotes a spatially independent OV function, and for concreteness, we adopt the standard S-shape (1.6). However, qualitatively similar results should be recovered by any V for which 1. $V(0) = 0$, 2. $V' \geq 0$ and 3. $V(h) \rightarrow V_{\max}$ as $h \rightarrow \infty$, and for which the flow is linearly stable. Therefore, throughout this chapter we choose $\alpha \geq 2 \max V'(h)$, thereby ensuring linear stability (1.7). Consequently, the patterns that we observe are forced only by spatial heterogeneity and not by spontaneous flow breakdown effects.

Next we supplement (2.1–2.3) with the uniformly spaced initial data

$$x_n = L - nh_* \quad \text{and} \quad v_n = V(h^*) \quad \text{for } n = 1, 2, \dots, N. \quad (2.4)$$

Here $h^* := L/N$ is the mean spacing. Note that for the limiting (no bottleneck) cases where either $r_B = 1$ or $\hat{L} = 0$, (2.4) gives a uniform flow solution of (2.1–2.3) in which $x_n = L - nh_* + tV(h^*)$. However, in general we should expect the bottleneck to redistribute traffic. In order to investigate the resulting patterns, we solve the initial value problem (2.1–2.4) numerically using a standard fixed step fourth-order Runge-Kutta solver.

After extensive experimentation with the solver, we conclude that the traffic always settles down to a stationary profile as $t \rightarrow \infty$, although the transient processes can sometimes be very long. Here stationarity means that suitably defined macroscopic density and velocity variables become steady, although they are non-trivially dependent on space x . Consequently vehicles' motions are in fact periodic as $t \rightarrow \infty$, since as they drive around the ring-road, they move repeatedly through the spatial pattern and experience traffic jams, free-flowing regimes etc. Note however that if we chose smaller values of sensitivity α than presented here, so as to force the linear instability of a range of uniform flows, then the macroscopic variables could also be non-trivially time-dependent as $t \rightarrow \infty$.

Taking into account the above discussion, the results that we display shortly show stationary macroscopic density profiles $\rho(x)$ rather than individual vehicle trajectories. As noted in Section 1.2.2, the simplest way to relate microscopic and macroscopic variables is via $\rho(x_n, t) = 1/h_n(t)$, however this relationship only holds exactly in equilibrium situations [8, 56]. Therefore we use a coarse-grained density [56] which is obtained from the distributional point density by convolving with a normalised Gaussian test function whose characteristic length scale is chosen large enough so as to smooth out individual vehicles but small enough so as to retain macroscopic features.

However, flows are stationary in our study. Hence it proves sufficient to coarse-grain the trajectory data of a single vehicle with a temporal kernel only, thus providing a substantial simplification of the general case presented in [56]. Moreover, since the time step in the initial value solver may be chosen as small as we wish, any macroscopic feature can be adequately resolved by this approach. The width parameter of the Gaussian envelope used in this chapter is $w = 0.05$ in each case.

We now give three examples of the eventual stationary profiles $\rho(x)$ which show how the structure changes as the mean headway h^* is varied. To simplify matters, all other parameters are held fixed as follows: $N = 100$ vehicles, bottleneck reduction factor $r_B = 0.6$, bottleneck nondimensionalised length $\hat{L} = 0.25$ and sensitivity $\alpha = 3.0$. We should emphasise that the resulting structures, shown in Figure 2.1, have also been described in previous studies [32, 80]. The novel feature here is that we analyse in detail how the qualitative solution structures change as functions of the three problem parameters $\rho_* := 1/h_*$, r_B and \hat{L} .

Example 1. We take $h_* = 7.0$ which corresponds to light traffic (large h_* , small ρ_*), shown in Figure 2.1(a). The $t \rightarrow \infty$ steady density profile $\rho(x)$ adopts a two-plateau form, with an almost constant density $\rho_B \approx 0.12$ attained in the bottleneck and a lower (almost constant) density $\rho_1 \approx 0.20$ on the remainder of the loop. At each end of the bottleneck, the two density plateaus are joined by sharp, almost shock-like fronts.

Example 2. We take $h_* = 1.0$ which corresponds to heavy traffic, exhibited in Figure 2.1(c). In a similar fashion to Example 1, $\rho(x)$ adopts a two-plateau form. However this time the bottleneck density $\rho_B \approx 0.71$ is less than the density $\rho_1 \approx 1.09$ on the unconstrained part of the loop. Like Example 1, there are also sharp, shock-like fronts at each end of the bottleneck, although here they have a more complicated oscillatory structure.

Example 3. We now take $h_* = 2.5$ which may be regarded as an intermediate case, presented in Figure 2.1(b). In contrast to the two previous examples, $\rho(x)$ now has a three-plateau form. The density as before adopts an almost constant (but slightly S-shaped) profile $\rho \simeq 0.36$ within the bottleneck, with fronts at each end. Downstream of the bottleneck is a low density $\rho_1 \approx 0.17$ region, whereas upstream is a high density $\rho_2 \approx 0.64$ region, which may be thought of as a queue waiting to enter the bottleneck. There is thus an extra internal shock-like front in the unconstrained part of the loop, where the fast traffic that emerges from the bottleneck rejoins the queue to enter it. Unlike the other fronts we have encountered so far, the wave joining ρ_1 and ρ_2 is not locked onto a discontinuity in the model; nevertheless, it is stationary.

Further simulation may be used to show how the Figure 2.1 profiles are related to each other. If one starts with the Figure 2.1(b) structure (Example 3) and decreases the

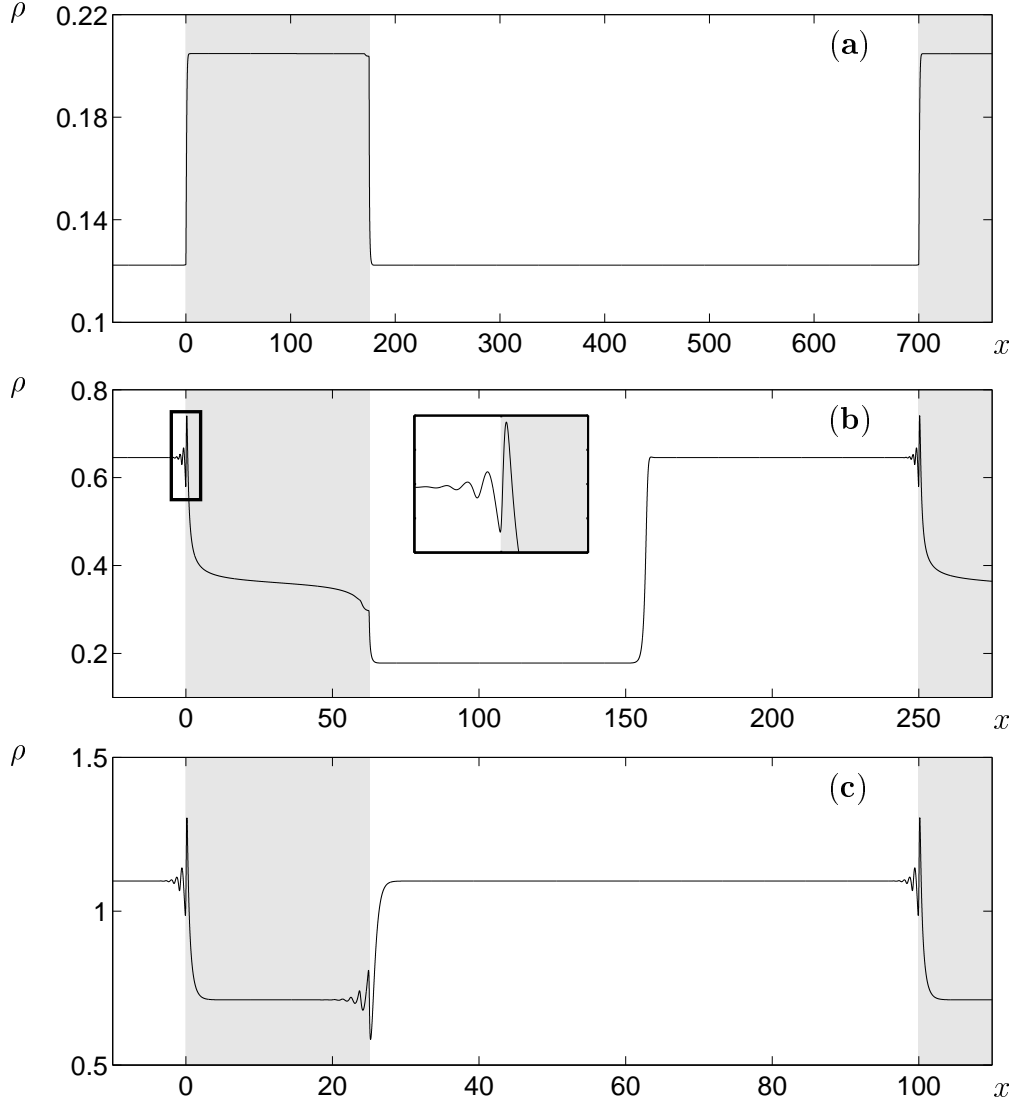


Figure 2.1. Stationary $t \rightarrow \infty$ coarse-grained density profiles $\rho(x)$. The portions of solution profiles within the bottleneck are indicated by shading. (a) Light traffic $h_* = 7.0$ ($\rho_* = 0.142857$), see Example 1; (b) Medium traffic $h_* = 2.5$ ($\rho_* = 0.4$), see Example 3; (c) Heavy traffic $h_* = 1.0$ ($\rho_* = 1.0$), see Example 2.

mean headway (increases the mean density), then the queue upstream of the bottleneck grows in length until it reaches the downstream boundary of the bottleneck, and swamps the entire unconstrained part of the loop. At this point, the internal shock vanishes and the Figure 2.1(c) structure is recovered. Conversely, if one starts with Figure 2.1(b) and decreases the mean density, the queue upstream of the bottleneck shortens until it vanishes altogether. At that point, the internal shock is absorbed into the upstream boundary of the bottleneck and the Fig. 2.1(a) structure is recovered.

2.3 Density and Flow Balances

We now begin an explanation of the structures seen in Section 2.2. Later we derive a phase diagram which predicts when each will occur. Since the observed structures resemble constant density plateaus separated by classical shocks, we initially attempt an explanation based on kinematic wave theory [82]. To this end, we recall the fundamental (flow) diagram $Q(\rho) = \rho\mathcal{V}(\rho)$ where $\mathcal{V}(\rho) = V(1/\rho)$ is the continuum counterpart to the discrete OV function V , (1.6). As is well-known, Q is usually a unimodal function. With choice (1.6), Q attains its maximum value $Q_{\max} \simeq 0.58$ at $\rho_{\max} \simeq 0.36$. In the bottleneck, the fundamental diagram Q is scaled by r_B .

Firstly we consider the two-plateau structures of (2.1)(a) and (c). Since the fronts are sharp, negligibly few vehicles are contained within them at any one time. We may therefore approximate the density $\rho(x)$ with a piecewise-constant profile consisting of ρ_B within the bottleneck and ρ_1 in the unconstrained part of the loop. It thus follows that

$$\hat{L}\rho_B + (1 - \hat{L})\rho_1 = \rho_*, \quad (2.5)$$

$$Q(\rho_1) = r_B Q(\rho_B), \quad (2.6)$$

which describe respectively the conservation of vehicles and a flow balance (the latter is necessary since the observed profiles are stationary). Equations (2.5, 2.6) are thus a pair of simultaneous equations to solve for ρ_1 and ρ_B , where the remaining parameters ρ_* , r_B and \hat{L} are prescribed.

We must therefore examine the (ρ_1, ρ_B) solution structure of Equations (2.5, 2.6) and this is achieved via Figure 2.2. To see this, note that ρ_B may be eliminated from Equations (2.5, 2.6) to give

$$Q(\rho_1) = r_B Q\left(\frac{\rho_* - (1 - \hat{L})\rho_1}{\hat{L}}\right), \quad (2.7)$$

and the left and right hand sides of this equation are plotted against ρ_1 in Figure 2.2(a-c)(i). Note alternatively that ρ_1 can be eliminated from Equations (2.5, 2.6) to give

$$r_B Q(\rho_B) = Q\left(\frac{\rho_* - \hat{L}\rho_B}{1 - \hat{L}}\right), \quad (2.8)$$

and as a cross-check, the left and right hand sides of this equation are plotted against ρ_B in Figure 2.2(a-c)(ii). Furthermore, parameters have been chosen so that the panels (a-c) correspond directly to panels (a-c) in Figure 2.1. Firstly, the light traffic diagrams

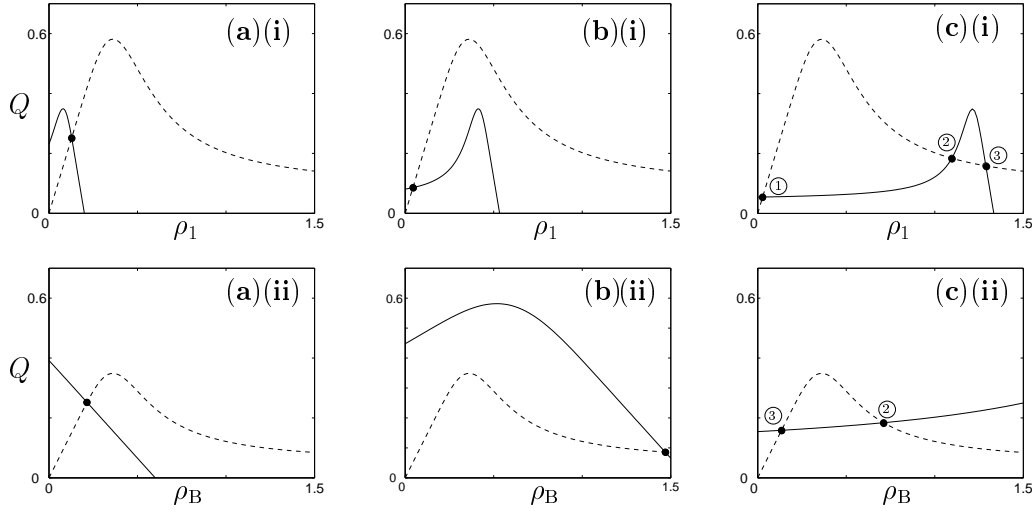


Figure 2.2. Solution structure of (2.5,2.6). Panels (a-c) correspond directly to panels (a-c) in Fig. 2.1. Top row (i) indicates solutions of (2.7) and bottom row (ii) of Eq. (2.8). The extra numbering in panels (c)(i,ii) allows the (ρ_1, ρ_B) solution pairs to be identified.

Figure 2.2(a)(i,ii) indicate a unique (ρ_1, ρ_B) solution pair and it may be shown that this corresponds to values obtained in Example 1.

However, in the heavy traffic diagrams Figure 2.2(c)(i,ii), there are clearly three (ρ_1, ρ_B) solution pairs: what determines which pair is selected in the corresponding Example 2? Finally, in the intermediate case of Figure 2.2(b)(i,ii), there is a unique (ρ_1, ρ_B) solution pair, however, the corresponding Example 3 selects instead a three-plateau structure. It now remains to identify extra principles which explain the solution selection in cases (b) (Example 3) and (c) (Example 2).

2.4 Wave Selection via Characteristics

We now use characteristic arguments from kinematic wave theory [82, Chap. 2] to explain the observed wave selection behaviour. We focus initially on Example 2 (heavy traffic), see Figure 2.1(c) and Figure 2.2(c)(i,ii), and then later we consider the three-plateau case.

We recall that in kinematic wave theory, characteristics are lines (or line segments) in the (x, t) plane on which density is conserved. Furthermore, it is well-known that the local velocity of a characteristic with density ρ is given by $Q'(\rho)$. Consequently, characteristics propagate downstream in light traffic and upstream in heavy traffic. When characteristics converge, one obtains a classical shock, whereas when they diverge, one obtains a (non-stationary) expansion fan.

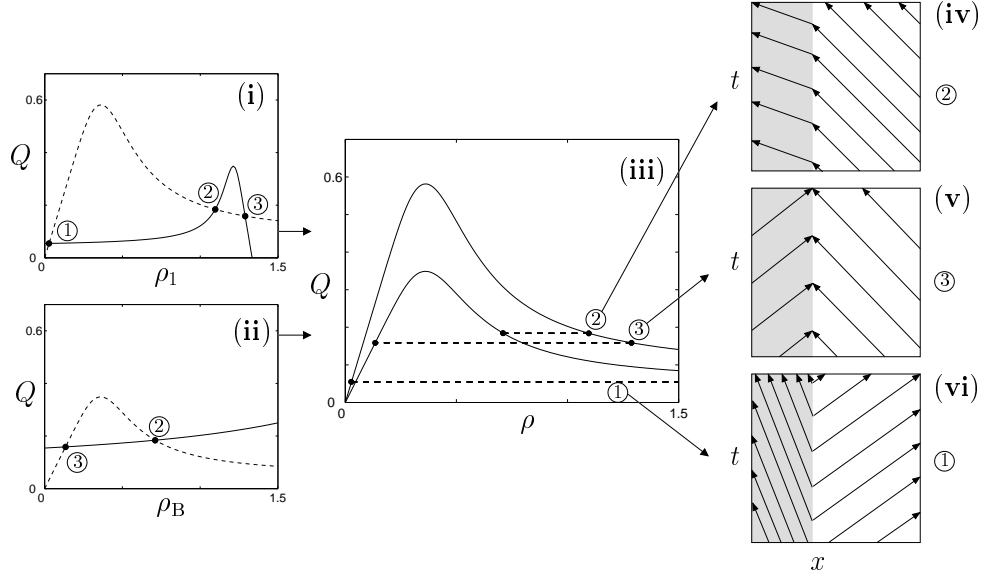


Figure 2.3. Characteristic analysis for the (ρ_1, ρ_B) solution pairs from Fig. 2.2(c)(i,ii). Panel (iii) shows the location of solution pairs (pair (1) is not fully contained in the figure), joined by horizontal lines representing flow balance, on the fundamental diagrams Q and $r_B Q$. Characteristic pictures for each of the three root pairs are shown in panels (iv), (v) and (vi): the bottleneck is denoted by shading. Panels (v) and (vi) cannot give stationary profiles since they predict an expansion fan at the up- and down-stream ends of the bottleneck respectively. Hence solution pair 2 from panel (iv) is selected. Note that in panel (iii), this solution pair is non-straddling in the sense that both ρ_1 and ρ_B are the same side of the fundamental diagram maximum.

Figure 2.3 develops a characteristic analysis of the (ρ_1, ρ_B) solution pairs found in Figure 2.2(c)(i,ii). The key point to note is that the solution pairs numbered 1 and 3 straddle $\rho = \rho_{\max}$ at which both the unconstrained $Q(\rho)$ and bottleneck $r_B Q(\rho)$ fundamental diagrams attain their maxima. These solution pairs can be disregarded, because the consequent density profiles would involve patterns of characteristics with both positive and negative slopes. This means that at either the upstream or downstream boundary of the bottleneck, there would necessarily be a non-stationary expansion fan which would not agree with the $t \rightarrow \infty$ stationary results.

In contrast, solution pair 2 is non-straddling and involves only characteristics with negative slopes, see Figure 2.3 panel (iv). In this sketch, neither the upstream or downstream boundary of the bottleneck has a classical (compressive) shock. Rather, at each boundary the characteristics cross through the shock which is forced solely by the model discontinuity at that point. It may be shown that this solution agrees with that found by discrete simulation in Section 2.2 and moreover that it is a proper solution of the Lighthill-Whitham-Richards model in that it may be reached via the solution of

the initial value problem [42].

We now turn our attention to the three-plateau case (Example 3, Equation (2.1)(b)), for which it may be shown that the analysis of Section 2.3 predicts a straddling, and hence invalid solution pair (ρ_1, ρ_B) . The resolution is to approximate the density $\rho(x)$ by a piecewise-constant profile with three components: ρ_B (density in bottleneck) and ρ_1, ρ_2 (densities in unconstrained part of loop). The density and flow balances thus yield respectively

$$\hat{L}\rho_B + \beta(1 - \hat{L})\rho_1 + (1 - \beta)(1 - \hat{L})\rho_2 = \rho_*, \quad (2.9)$$

$$Q(\rho_1) = Q(\rho_2) = r_B Q(\rho_B), \quad (2.10)$$

where $0 < \beta < 1$ parametrises the internal shock position separating ρ_1 and ρ_2 . We thus have three equations, but four unknowns, namely β, ρ_1, ρ_2 and ρ_B , and we require extra information to fix a unique solution.

By studying characteristic diagrams, it becomes clear that a solution without diverging characteristics (and hence non-stationary expansion fans) is only possible if

$$\rho_B = \rho_{\max}, \quad (2.11)$$

i.e., if the flow inside the bottleneck is maximised. Further, when supplemented by (2.11), system Equations (2.9, 2.10) can be solved uniquely for ρ_1, ρ_2 and β , and it may be shown that this solution agrees with the discrete simulations. The characteristic structure is shown in Figure 2.4. In particular, it involves non-standard waves at the up- and down-stream ends of the bottleneck. However it will be illustrated in Section 2.5 that these are admissible solutions of the Lighthill-Whitham-Richards model via the solution of the initial value problem.

We now turn our attention to the computation of a phase diagram. Since in the three-plateau case we have $\rho_B = \rho_{\max}$, the values of ρ_* where solutions change from two plateau solutions to three plateau solutions can be calculated. At the thresholds, β is either 0 or 1 and $\rho_B = \rho_{\max}$, thus eliminating ρ_1 or ρ_2 in (2.10) using (2.9), leaves only

$$r_B Q(\rho_{\max}) = Q\left(\frac{\rho_* - \hat{L}\rho_{\max}}{1 - \hat{L}}\right), \quad (2.12)$$

as a relation between the problem parameters that holds at the transition, see Figure 2.5. In particular, we may partition the (ρ_*, r_B) plane according to whether the three-plateau solution occurs, or according to which type of two-plateau solution occurs, and the boundary in this plane depends on \hat{L} in a manner that we can determine explicitly. In particular, increasing the length of the bottleneck shrinks the domain where the three-plateau solution occurs.

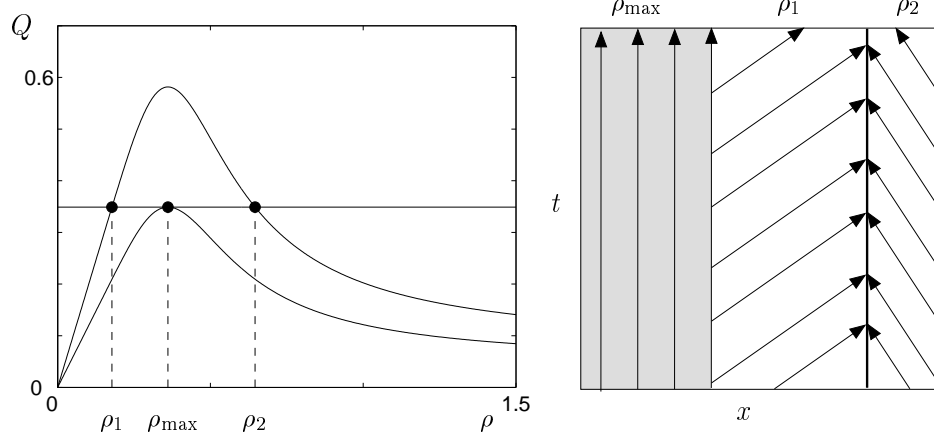


Figure 2.4. Characteristic analysis for the three-plateau case: only the configuration shown with flow maximised in the bottleneck avoids expansion fans. Note that the characteristics inside the bottleneck have zero velocity and hence this structure is on the very boundary of becoming an expansion fan. The internal shock between ρ_1 and ρ_2 is classical since at it the characteristics converge.

2.5 Further Characteristic Analysis

We now illustrate how the phase-diagram in the (ρ_*, r_B) plane might be obtained via solution of the initial value problem using the method of characteristics. To achieve this we introduce Newell's cumulative flux method [65] to track the progress of shocks, from which we obtain the transition boundary (2.12).

Let us first consider the inflection point in the flux-density relation Q , introduced by the OV function (1.6) at $\rho_I = 1/2$. It is well known that non-convex flux functions give rise to non-classical compound shocks [57]. To describe how these occur, we define the *tangent chord* function $G(\rho) = \sigma \neq \rho$, such that the line tangent to the point $(\sigma, Q(\sigma))$ intersects $(\rho, Q(\rho))$ (see Figure 2.6). Classical shocks satisfy what is known as the 'Lax Entropy Condition'. A geometric interpretation is that the chord on the Fundamental Diagram connecting the up- and downstream densities (ρ_u and ρ_d respectively) must lie below $Q(\rho)$. This condition is violated if $\rho_d > G(\rho_u)$. Thus we can identify situations in which non-classical shocks arise and subsequently choose parameters that prevent them from occurring at the transitions between two and three plateau solutions.

To simplify matters, we first consider the initial value problem for the case of a single bottleneck on an open road when the initial density is less than that which maximises flow, i.e. $\rho_* < \rho_{\max}$. A rigorous treatment can be found in [42]. Two qualitatively different situations occur when the initial flux in the free flow region is

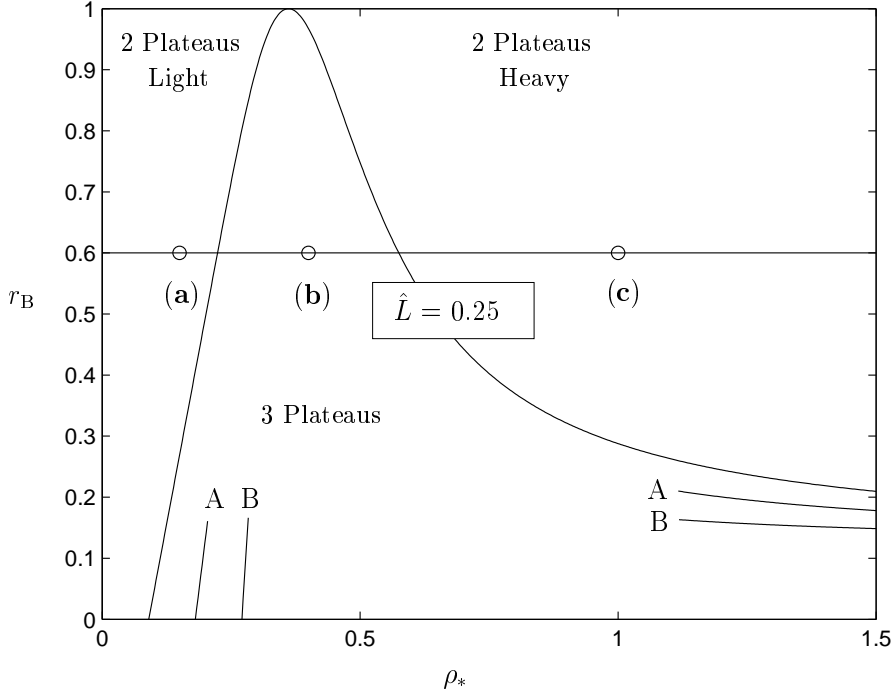


Figure 2.5. Phase diagram derived from (2.12) for bottleneck length $\hat{L} = 0.25$. The points marked (a), (b), and (c) correspond to panels (a), (b) and (c) in Figure 2.1 and Figure 2.2. The line segments denoted A and B indicate to where the phase boundary would move for $\hat{L} = 0.5$ and $\hat{L} = 0.75$ respectively.

above and below the maximum flux in the bottleneck. To distinguish between these we introduce the notation $\rho_{L,R}$ where $Q(\rho_{L,R}) = r_B Q(\rho_{\max})$ and $\rho_L < \rho_{\max} < \rho_R$.

The characteristic development for the case where $\rho_* < \rho_L$ is illustrated in Figure 2.7. Characteristics which cross the discontinuity at the bottleneck boundaries must conserve flow [42]. The initial characteristic velocity in the free flow region is greater than in the bottleneck since $Q'(\rho_*) > r_B Q'(\rho_*)$. When characteristics cross from the free flow region into the bottleneck their density increases in order to maintain a constant flux. We denote the new density $\rho_b > \rho_*$. Furthermore, their velocity *decreases* below that of characteristics that originate inside the bottleneck. This is because the flux function is convex since $\rho_* < \rho_{\max} < \rho_I$. This gives rise to an expansion fan that originates inside the bottleneck.

Conversely, characteristics which cross out of the bottleneck decrease in density and hence increase in speed. Those characteristics with density ρ_* are reduced to ρ_f and similarly ρ_b changes back to ρ_* . The speed of those that originate inside the bottleneck is therefore greater than those in free flow, hence a shock propagates downstream from the bottleneck (see Figure 2.7(b)).

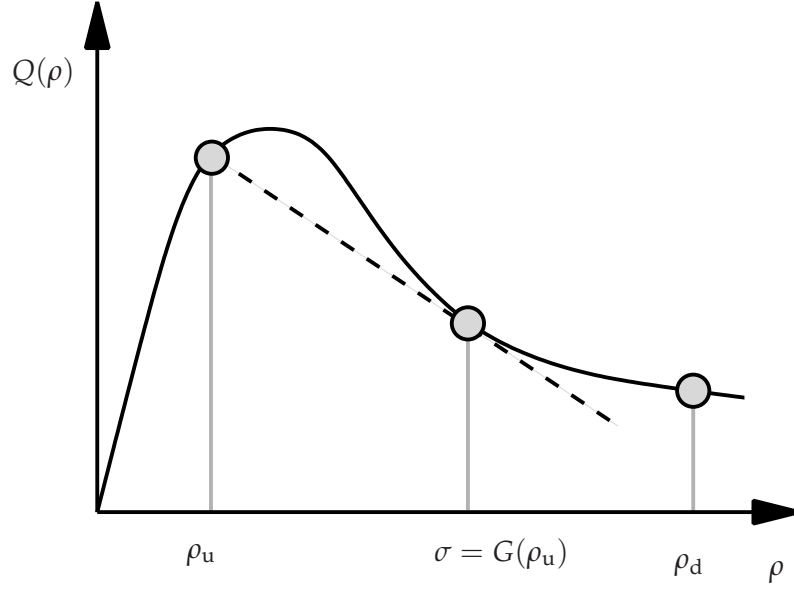


Figure 2.6. Illustration of the tangent chord function $\sigma = G(\rho)$ plotted on the Fundamental Diagram. The tangent to the flux-density relation $Q(\rho)$ at σ , intersects the density ρ_u . Non-classical shocks occur between upstream densities ρ_u and downstream densities $\rho_d > \sigma$.

The characteristics development for the case where $\rho_L < \rho_* < \rho_{\max}$ is illustrated in Figure 2.8. The behaviour at the downstream front is essentially the same as in the previous case illustrated in Figure 2.7. Characteristics at the upstream front however can no longer conserve flux if they cross into the bottleneck. They therefore first connect to a high density ρ_R on the *free flow* Fundamental Diagram, which preserves the conservation of flux at the boundary, thus creating a shock that propagates upstream, away from the bottleneck. Since characteristics do not cross into the bottleneck at all, the behaviour there is dominated by the expansion fan. The flux inside the bottleneck therefore approaches the bottleneck capacity asymptotically.

We now consider how periodic boundary conditions affect the characteristics described above. In both cases the initial development remains exactly the same. When $\rho_* < \rho_L$, the shock originating at the downstream bottleneck front eventually traverses the free flow region and reaches the bottleneck entrance. Its speed changes in the bottleneck, but continues propagating around the loop. This progress is hindered by the expansion fan which eventually corrodes the shock both up- and downstream. The expansion fan dominates the subsequent evolution and the flux approaches a constant value across the entire loop. The density in the bottleneck and the free flow regions therefore tend towards two *different* constant values. This process may take a very long time, particularly if the characteristic speeds $Q'(\rho_*)$ and $Q'(\rho_b)$ are similar. In summary, when $\rho_* < \rho_L$, only two plateau solutions emerge.

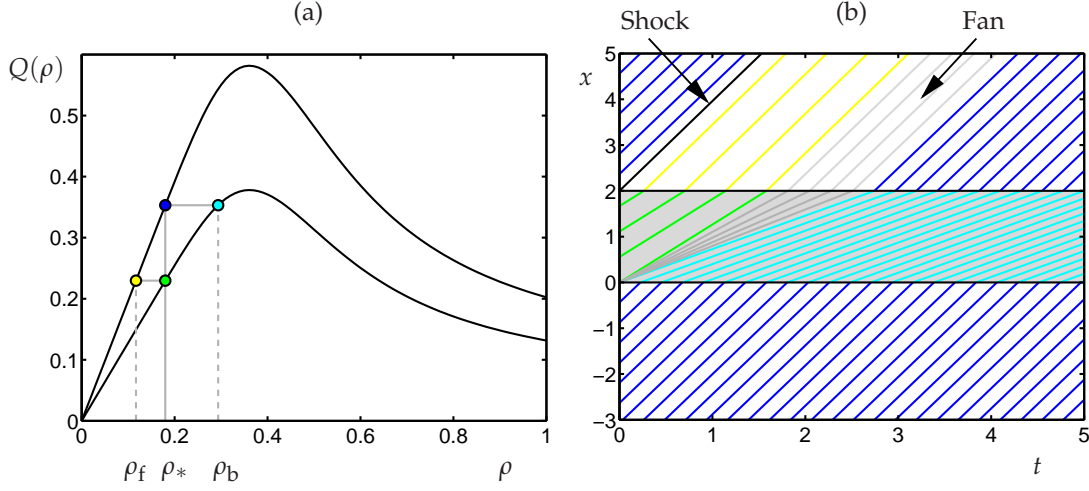


Figure 2.7. Evolution of characteristics around a single bottleneck when the initial flux in free flow is *less* than the capacity of the bottleneck, i.e. $\rho_* < \rho_L$. The Fundamental Diagram is drawn in panel (a) where the initial density ρ_* and the changes in density that occur when characteristics pass into and out of the bottleneck are labelled ρ_b and ρ_f respectively. The development of characteristics are illustrated in Panel (b), the bottleneck region is shaded in grey. The expansion fan that results when characteristics cross the upstream bottleneck front and the shock that propagates away from the downstream bottleneck front are labelled.

When $\rho_L < \rho_* < \rho_{\max}$, the shocks originating from both the bottleneck entrance and exit collide (see Figure 2.9). Note that the shock propagating from the bottleneck exit may be refracted by the expansion fan first. The velocity of the characteristics in the *fan* in free flow approach $Q'(\rho_L) > 0$. Hence the shocks always collide within the free flow region since their velocities are of opposing sign. The resultant shock velocity is positive since $Q(\rho_f) < Q(\rho_R)$ (see Figure 2.8). If the shock crosses into the bottleneck then the expansion fan in the bottleneck *increases* the shocks velocity. Note that after this occurs, the characteristic density downstream of the front is always less than ρ_{\max} , hence the final pattern is always a two plateau solution since connections that straddle the maximum have been eliminated. The progress of the shock through the free flow region may however be slowed by the expansion fan (this occurs at the point (x_c, t_c) in Figure 2.9). The density of the characteristics in the fan vary from ρ_f to ρ_L . The difference in flux up- and downstream of the shock therefore approaches zero, and hence the shock does not necessarily cross into the bottleneck. In such cases, three-plateau solutions are observed. In summary, the transition between two and three plateau solutions occurs when the refracted shock trajectory is asymptotic to the bottleneck entrance.

We have noted that in some cases, non-classical shocks may arise. In the description of the characteristics above, three shocks occurred; at the bottleneck entrance

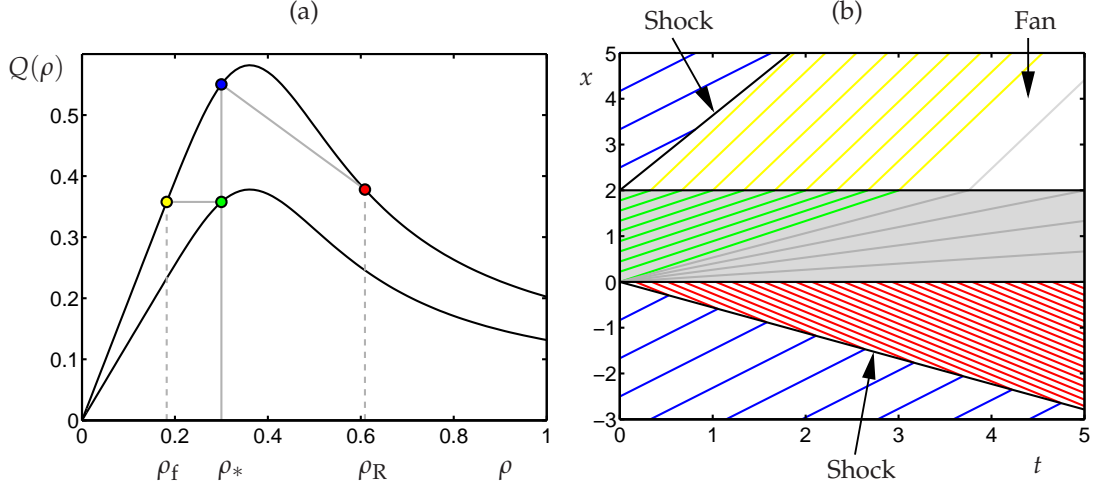


Figure 2.8. Evolution of characteristics around a single bottleneck when the initial flux in free flow is *greater* than the capacity of the bottleneck, i.e. $\rho_* > \rho_L$. The Fundamental Diagram is drawn in panel (a), where the initial density ρ_* and the changes in density when characteristics pass into and out of the bottleneck are labelled. In the later case, the density drops to ρ_f . Characteristics which cross into the bottleneck can no longer conserve flux. Hence these are forced to connect to a much higher density, ρ_R . Characteristics are illustrated in panel (b). A shock can be seen propagating upstream from the bottleneck entrance, which corresponds to the development of a queue. Furthermore, the characteristics in the bottleneck are dominated by the expansion fan. The density in the bottleneck therefore approaches ρ_{\max} , and hence maximises flow.

and exit, and when these shocks collided (and subsequently refracted). The upstream shock will be a classical shock while

$$\rho_R > G(\rho_*). \quad (2.13)$$

Hence r_B must be sufficiently large to prevent this condition from being violated. For initial densities $\rho_* < \rho_{\max}$, the downstream shock and the combined (and subsequently refracted) shocks are all classical since they all have positive velocity. There is a qualitative symmetry when $\rho_* > \rho_{\max}$, however non-classical shocks are inevitable between $(\rho_*, Q(\rho_*))$ and $(\rho_f, Q(\rho_f))$. We therefore only consider the low density transition between two and three plateaus, assuming inequality (2.13) holds.

2.5.1 Cumulative Flux Method

Our goal now is to track the refracted shock in Figure 2.8, and determine when its position is asymptotic to the bottleneck entrance. To achieve this we briefly describe Newell's cumulative flux method [65], and then apply it to our problem. The cumu-

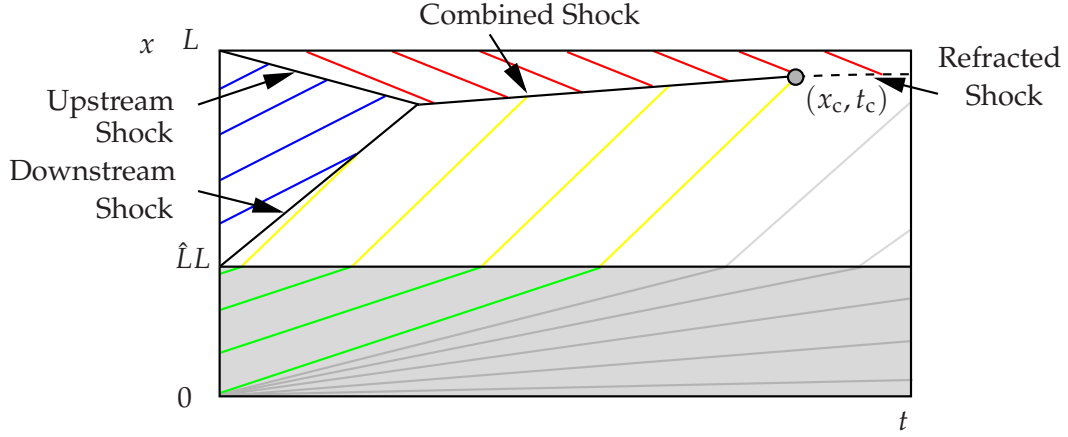


Figure 2.9. Sketch of characteristic evolution on a ring-road when $\rho_L < \rho_* < \rho_{\max}$. The shocks emanating from the bottleneck collide, from which the resulting shock propagates forwards since $\rho_* < \rho_{\max}$. This is then refracted by the expansion fan at (x_c, t_c) . There are two subsequent outcomes. The characteristics upstream of the shock have increasing density, approaching ρ_L . Hence the velocity of the shock tends to zero in the free flow region. If the shock crosses into the bottleneck however, its velocity is *increased* by the downstream fan.

lative flux $A(x, t)$ must be defined according to some reference point (x_0, t_0) where $A(x_0, t_0) = 0$. The flux and density are given by

$$q(x, t) = \frac{\partial A(x, t)}{\partial t}, \quad \text{and} \quad (2.14)$$

$$\rho(x, t) = -\frac{\partial A(x, t)}{\partial x}, \quad (2.15)$$

respectively. Conservation is a natural consequence and is found by equating the second derivatives of (2.14) and (2.15) with respect to x and t respectively. The change in cumulative flux along characteristics of constant density is

$$dA = \frac{\partial A}{\partial t} dt + \frac{\partial A}{\partial x} dx \quad (2.16)$$

$$= (Q(\rho) - \rho Q'(\rho)) dt. \quad (2.17)$$

The key advantage of this method comes from the following facts,

1. Discontinuities in $A(x, t)$ correspond to shock waves,
2. If $A(x, t)$ is multivalued, then the lower envelope is the solution.

We illustrate these points with a simple example. Consider the Riemann problem where at time $t = 0$, a jump in density is located at $x = 0$, which separates the densities

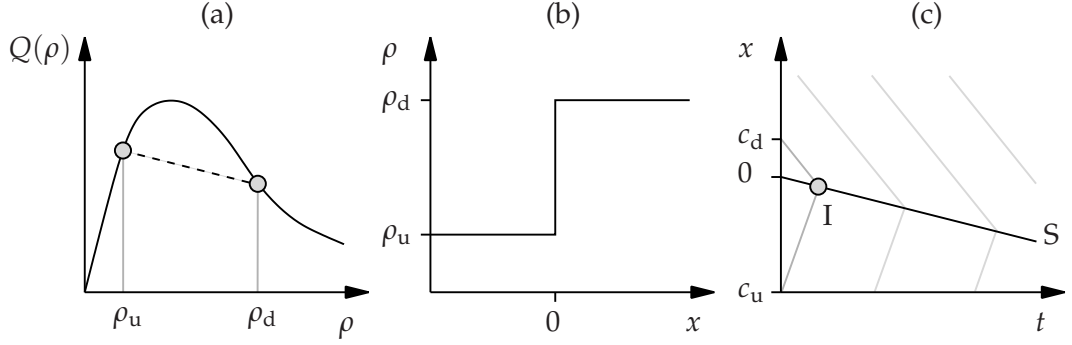


Figure 2.10. The Riemann initial value problem. The fundamental diagram is illustrated in panel (a). Piecewise constant initial densities (ρ_u upstream and ρ_d downstream) are initially separated by a discontinuity at $x = 0$ (illustrated in panel (b)). We choose these densities such that $Q'(\rho_d) < Q'(\rho_u)$, hence the discontinuity propagates as a shock. Characteristics are sketched in the (x, t) plane in panel (c). The shock path is labelled 'S' and the up- and downstream characteristics that intersect the x -axis at c_u and c_d respectively coincide on the shock at point I.

ρ_d downstream (i.e. $x > 0$), and ρ_u upstream (i.e. $x < 0$), where $Q'(\rho_d) < Q'(\rho_u)$ (see Figure 2.10). The equations of two arbitrary up- and downstream characteristics satisfy

$$x_u = Q'(\rho_u)t + c_u, \quad (2.18)$$

$$x_d = Q'(\rho_d)t + c_d. \quad (2.19)$$

We are interested in pairs of characteristics that converge on the shock such that

$$x(t) = x_u(t) = x_d(t). \quad (2.20)$$

The cumulative flux consists of two envelopes, corresponding to the up- and downstream densities, which are equal along the shock. Integrating (2.17) allows us to equate the cumulative flux along each characteristic, this yields

$$A(x(t), t) - A(c_u, 0) = -\rho_u^2 \mathcal{V}'(\rho_u) t, \quad (2.21)$$

$$A(x(t), t) - A(c_d, 0) = -\rho_d^2 \mathcal{V}'(\rho_d) t. \quad (2.22)$$

We are free to choose the reference point $A(c_u, 0) = 0$, thus to determine $A(c_d, 0)$ we must integrate (in space) the flux between c_u and c_d at $t = 0$, to give

$$A(c_d, 0) = c_d \rho_d - c_u \rho_u. \quad (2.23)$$

We then equate $A(x, t)$ in (2.21) and (2.22) with $A(c_u, 0) = 0$, which allows us to write

c_d in terms of c_u , we then substitute this expression for c_d into the characteristic equation (2.19). Then using (2.18) to eliminate c_u , we obtain the shock trajectory

$$x = \frac{Q(\rho_u) - Q(\rho_d)}{\rho_u - \rho_d} t := v_s t. \quad (2.24)$$

Thus as is well known, the shock velocity v_s is just the gradient of the chord connecting $(\rho_u, Q(\rho_u))$ and $(\rho_d, Q(\rho_d))$ on the fundamental diagram.

2.5.2 Calculation of the Two/Three-Plateau Boundary

We now turn our attention to the asymptotic position of the refracted shock in Figure 2.9. To simplify matters, we imagine that the characteristics at the up- and downstream bottleneck interfaces can be extended, hence we can effectively ignore the bottleneck. The density of characteristics in the fan in free flow decreases from ρ_f to ρ_L . We denote the density of an arbitrary characteristic in the expansion fan in free flow by ρ_F , such that $\rho_L < \rho_F < \rho_f$. Our method is to determine when the shock position coincides with the bottleneck entrance at $x = L$, in the limit $\rho_F \rightarrow \rho_L$.

The equation of an arbitrary characteristic downstream of the shock with density ρ_R is

$$x_d = Q'(\rho_R)t + c_d. \quad (2.25)$$

The equation of an arbitrary characteristic upstream of the shock in the expansion fan is

$$x_u = Q'(\rho_F)t + c_F, \quad (2.26)$$

where the intersection with the x -axis, c_F , depends on the characteristic density ρ_F . To obtain this relationship, consider the speed of characteristics in the fan *prior* to leaving the bottleneck. This can be parameterised in terms of the escape time from the bottleneck τ . The velocity is then $\hat{L}L/\tau$, and this is related to the density of characteristics in the expansion fan in the bottleneck ρ_B by

$$\frac{\hat{L}L}{\tau} = r_B Q'(\rho_B). \quad (2.27)$$

When the characteristics cross into the free flow region flux is conserved, hence

$$r_B Q(\rho_B) = Q(\rho_F). \quad (2.28)$$

The intersection c_F can be related to the escape time τ via

$$c_F = \hat{L}L - Q'(\rho_F)\tau. \quad (2.29)$$

Therefore combining (2.27–2.29) we find that the equation of an arbitrary characteristic in the expansion fan is

$$x_u = Q'(\rho_F)t + f(\rho_F)\hat{L}L, \quad (2.30)$$

where

$$f(\rho) = 1 - \frac{Q'(\rho)}{r_B Q' \left(Q^{-1} \left(\frac{Q(\rho)}{r_B} \right) \right)}. \quad (2.31)$$

The range of densities in the expansion fan ensure that the inverse flux function Q^{-1} is well defined.

We now consider the change of cumulative flux along the characteristics, given by

$$A(x(t), t) - A(c_F, 0) = -\rho_F^2 \mathcal{V}'(\rho_F) t, \quad (2.32)$$

$$A(x(t), t) - A(c_d, 0) = -\rho_R^2 \mathcal{V}'(\rho_R) t. \quad (2.33)$$

We choose the reference point $A(c_F, 0) = 0$; hence by integrating the contribution from each of the regions between c_F and c_d , we find

$$\begin{aligned} A(c_d, 0) &= \int_{f(\rho_F)\hat{L}L}^{f(\rho_f)\hat{L}L} \rho(x, 0) dx + \int_{f(\rho_f)\hat{L}L}^{\hat{L}L} \rho(x, 0) dx \\ &\quad + \int_{\hat{L}L}^L \rho(x, 0) dx + \int_L^{c_d} \rho(x, 0) dx, \end{aligned} \quad (2.34)$$

$$\begin{aligned} &= \int_{f(\rho_F)\hat{L}L}^{f(\rho_f)\hat{L}L} \rho(x, 0) dx + (1 - f(\rho_f))\hat{L}L\rho_f \\ &\quad + (1 - \hat{L})L\rho_* + (c_d - L)\rho_R. \end{aligned} \quad (2.35)$$

To simplify the integral in (2.35), we change variables from positions x to densities ρ using (2.30) with $t = 0$,

$$\int_{f(\rho_F)\hat{L}L}^{f(\rho_f)\hat{L}L} \rho(x, 0) dx = \hat{L}L \int_{\rho_F}^{\rho_f} \rho \frac{\partial f(\rho)}{\partial \rho} d\rho, \quad (2.36)$$

$$= (f(\rho_f)\rho_f - f(\rho_F)\rho_F)\hat{L}L + \hat{L}L \int_{\rho_f}^{\rho_F} f(\rho) d\rho. \quad (2.37)$$

We can eliminate c_d from the initial cumulative flux (2.35) by equating the characteristic positions (2.25) and (2.30),

$$c_d = [Q'(\rho_F) - Q'(\rho_R)] t + f(\rho_F)\hat{L}L. \quad (2.38)$$

Then equating the cumulative fluxes (2.32) and (2.33) with $A(c_F, 0) = 0$ and (2.35), and

substituting the identity $\rho^2 \mathcal{V}'(\rho) = \rho Q'(\rho) - Q(\rho)$, we find

$$\begin{aligned} [(\rho_F - \rho_R)Q'(\rho_F) - \{Q(\rho_F) - Q(\rho_R)\}] t &= \hat{L}L \int_{\rho_f}^{\rho_F} f(\rho) d\rho \\ &+ (\rho_f \hat{L} - \rho_R)L + (1 - \hat{L})L\rho_* - (\rho_F - \rho_R)f(\rho_F)\hat{L}L. \end{aligned} \quad (2.39)$$

This equation tells us the time at which two characteristics with densities ρ_R and ρ_F coincide on the refracted shock. Using (2.39) to eliminate t from (2.30), we thus find the corresponding position, which we set to $x = L$. This yields

$$\begin{aligned} L [(\rho_F - \rho_R)Q'(\rho_F) - \{Q(\rho_F) - Q(\rho_R)\}] &= -f(\rho_F)\hat{L}L \{Q(\rho_F) - Q(\rho_R)\} \\ &+ Q'(\rho_F) \left\{ \hat{L}L \int_{\rho_f}^{\rho_F} f(\rho) d\rho + (\rho_f \hat{L} - \rho_R)L + (1 - \hat{L})L\rho_* \right\}. \end{aligned} \quad (2.40)$$

Taking the limit $\rho_F \rightarrow \rho_L$ then gives

$$\rho_L = \hat{L} \int_{\rho_f}^{\rho_L} f(\rho) d\rho + \hat{L}\rho_f + (1 - \hat{L})\rho_*. \quad (2.41)$$

We can simplify the integral via two substitutions. First we use $s = Q(\rho)/r_B$,

$$\int_{\rho_f}^{\rho_L} f(\rho) d\rho = (\rho_L - \rho_f) - \int_{\rho_f}^{\rho_L} \frac{Q'(\rho)}{r_B Q' \left(Q^{-1} \left(\frac{Q(\rho)}{r_B} \right) \right)} d\rho, \quad (2.42)$$

$$= (\rho_L - \rho_f) - \int_{Q(\rho_*)}^{Q(\rho_{\max})} \frac{ds}{Q'(Q^{-1}(s))}. \quad (2.43)$$

Then we insert $\hat{\rho} = Q^{-1}(s)$, which gives

$$\int_{Q(\rho_*)}^{Q(\rho_{\max})} \frac{ds}{Q'(Q^{-1}(s))} = \int_{\rho_*}^{\rho_{\max}} d\hat{\rho} = (\rho_{\max} - \rho_*). \quad (2.44)$$

Equation (2.41) then reads

$$(1 - \hat{L})\rho_L = \rho_* - \hat{L}\rho_{\max}. \quad (2.45)$$

With $Q(\rho_L) = r_B Q(\rho_{\max})$ we obtain the three-plateau condition (2.12), namely

$$Q \left(\frac{\rho_* - \hat{L}\rho_{\max}}{1 - \hat{L}} \right) = r_B Q(\rho_{\max}). \quad (2.46)$$

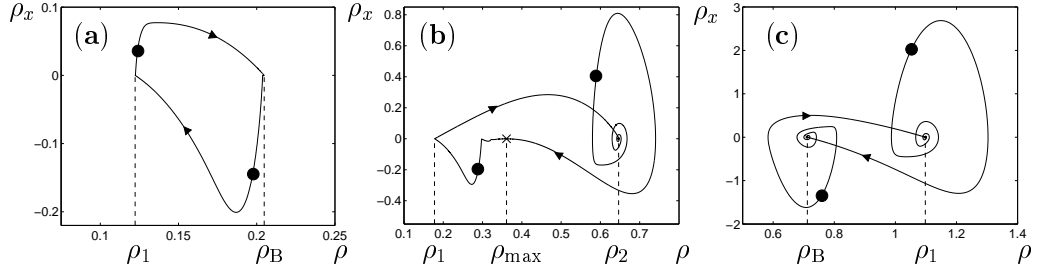


Figure 2.11. Numerical (ρ, ρ_x) phase plots corresponding to Figure 2.1(a-c). The discs mark the boundaries of the bottleneck. (a) Light traffic: heteroclinic cycle connecting $(\rho_1, 0)$ and $(\rho_B, 0)$. (b) Three plateau case: saddle at $(\rho_1, 0)$; saddle-node at $(\rho_{\max}, 0)$ explaining S-shaped structure; complex fixed point at $(\rho_2, 0)$. (c) Heavy traffic: complex fixed points at $(\rho_1, 0)$ and $(\rho_B, 0)$.

2.6 Second Order Modelling

We have shown in Sections 2.3–2.5 how a first order kinematic wave theory approximation explains the plateau-like features of Figure 2.1, in terms of piecewise constant profiles connected together by perfect discontinuities. However, the displacements and velocities arising from (2.1–2.3) are continuous, even at the bottleneck boundary, and moreover smoothing is involved in the coarse-graining process. Hence the fronts that connect plateaus are in fact continuous, and we now turn our attention to their structure. The main challenge is to examine whether a higher order macroscopic theory can explain the interface structure in the coarse-grained solution profiles, or whether it is necessary to analyse the microscopic model (2.1–2.3) directly.

In summary, our answer is that some features can be explained by second order PDE modelling (this section), whereas other features, that would not occur on a homogeneous road but which are forced directly by the bottleneck discontinuity, have such short wavelength that only an examination of the spectrum of (2.1, 2.2) can explain them (see Section 2.7).

To focus this discussion, we are interested specifically in whether fronts are monotone (as in Figure 2.1(a)), or whether they involve rapid oscillations (e.g., the upstream bottleneck boundary in Figure 2.1(b) and both bottleneck boundaries in Figure 2.1(c)). We would also like to explain the apparently S-shaped plateau in the bottleneck density profile of Figure 2.1(b).

To emphasise these features, we re-map the coarse-grained data used to produce Figure 2.1 into a (ρ, ρ_x) phase-plane, where ρ_x is obtained by numerical differentiation. The corresponding new figures are shown in Figure 2.11(a-c). These diagrams are sometimes referred to as Poincaré phase-planes. The advantage of these coordinates is

that almost-constant solution behaviour along a plateau is compressed into the neighbourhood of so-called *fixed points* $((\rho_1, 0)$ and $(\rho_B, 0)$ in Figure 2.11(a) and 2.11(c), and $(\rho_1, 0)$, $(\rho_{\max}, 0)$ and $(\rho_2, 0)$ in Figure 2.11(b)), whereas the sharp transitions between plateaus are stretched out. In particular, the rapid oscillations from Figure 2.1 now express themselves in the form of spiralling features. A disadvantage of these coordinates is that the point of discontinuity in the model is now hidden; consequently we add solid discs to Figure 2.11 to mark the points on the solution trajectories where the bottleneck boundaries are crossed.

Here we attempt to explain the structures in Figure 2.11 using a macroscopic theory developed by Kim, Lee and Lee [56] (see also Berg, Mason and Woods [8] for a similar theory based on perturbation expansions). Their analysis shows that, in the long-wavelength near-equilibrium limit, the coarse-grained density $\rho(x, t)$ and velocity $v(x, t)$ arising from the Optimal Velocity model are approximate solutions of the PDE system

$$\rho_t + (\rho v)_x = 0, \quad (2.47)$$

$$v_t + vv_x = \alpha \{ \mathcal{V}(\rho) - v \} + \alpha \left[\mathcal{V}'(\rho) \frac{\rho_x}{2\rho} + \frac{v_{xx}}{6\rho^2} \right], \quad (2.48)$$

which resembles the Kerner-Konhäuser model [46], in that it consists of a conservation equation and a momentum equation, the latter being composed of an Optimal Velocity relaxation term plus a balance of pressure and diffusion effects. Here (2.47, 2.48) must be generalised later to allow for spatial heterogeneity by making \mathcal{V} and \mathcal{V}' depend explicitly on x .

We now follow [84] by seeking steady solutions

$$v = W(x) \quad \text{and} \quad \rho = X(x), \quad (2.49)$$

to reduce (2.47, 2.48) to an ordinary differential equation problem. The continuity equation (2.47) thus becomes $(WX)' = 0$, where $'$ denotes differentiation with respect to x . Direct integration then gives

$$W = q/X, \quad (2.50)$$

where q is an integration constant that equals the (spatially uniform, time-independent) flow around the ring. For example, if we are interested in solutions involving a non-bottleneck plateau with $\rho = \rho_1$, then $q = Q(\rho_1) = \rho_1 \mathcal{V}(\rho_1)$. Note that in [84], this argument is worked through with the travelling wave ansatz $v = W(z)$, $\rho = X(z)$ and $z = x - ct$, when one obtains the standard Rankine-Hugoniot result; that is, wave speed c is given by the gradient of a chord cutting the so-called fundamental diagram

$Q(\rho)$, and furthermore that q is given by the intercept of this chord with the vertical Q axis.

By using (2.50), the velocity variable W can be eliminated from (2.48), resulting in a second order ODE in terms of the density variable X . This may be re-cast as the first order system

$$X' = Y, \quad (2.51)$$

$$Y' = \frac{6X^3}{q} \left[\frac{q^2 Y}{\alpha X^2} + \{Q(X; x) - q\} \right] + \left[\frac{2Y}{X} + \frac{3X^3 \mathcal{V}'(X; x)}{q} \right] Y. \quad (2.52)$$

The subtlety here is the spatial heterogeneity due to the bottleneck, which makes (2.52) depend explicitly on space x . Thus (2.51, 2.52) actually represent a pair of phase-plane problems, one (unconstrained flow) in which

$$Q(X; x) = Q(X) \quad \text{and} \quad \mathcal{V}'(X; x) = \mathcal{V}'(X), \quad (2.53)$$

and one (bottleneck case) in which

$$Q(X; x) = r_B Q(X) \quad \text{and} \quad \mathcal{V}'(X; x) = r_B \mathcal{V}'(X). \quad (2.54)$$

Note that $X', Y' = 0$ fixed points correspond to plateaus. For such we require $Y = 0$ from (2.51), with (2.52) imposing the natural constraint $Q(X; x) = q$. Therefore the density values of plateaus are found by horizontal sections through the $Q(\rho)$ curve and all is consistent with Section (2.3). Furthermore, each of the (2.51, 2.52, 2.53) and (2.51, 2.52, 2.54) phase-planes typically has two fixed points $(\rho_l, 0)$ and $(\rho_r, 0)$ where ρ_l and ρ_r are respectively to the left and right of the maximum of Q at $\rho = \rho_{\max}$.

We now overlay the phase-planes for each of the cases (2.51, 2.52, 2.53) and (2.51, 2.52, 2.54), and seek paths which continuously combine segments of trajectories so as to connect up the required fixed points in accord with the selection principles established in Section 2.4. Figure 2.12 illustrates how the procedure works. The reduction factor $r_B = 0.6$ has been chosen to match Example 1 and q has been set equal to the resulting flow value $q = Q(\rho_1) = r_B Q(\rho_B)$. See also Figure 2.1(a) and Figure 2.11(a). Figure 2.12(a,b) plots numerical solution trajectories for systems (2.51, 2.52, 2.53) and (2.51, 2.52, 2.54) respectively, whose left- and right-hand fixed points are denoted $(\rho_{fl}, 0)$, $(\rho_{fr}, 0)$ and $(\rho_{bl}, 0)$, $(\rho_{br}, 0)$. The selection principles of Section 2.4 require plateaus whose densities do not straddle the maximum point of Q . Here this means that we seek a connection from $(\rho_{fl}, 0)$ in the uncongested phase-plane to $(\rho_{bl}, 0)$ in the bottleneck phase-plane. Thus Figure 2.12(c) overlays Figure 2.12(a,b) and magnifies the region of interest.

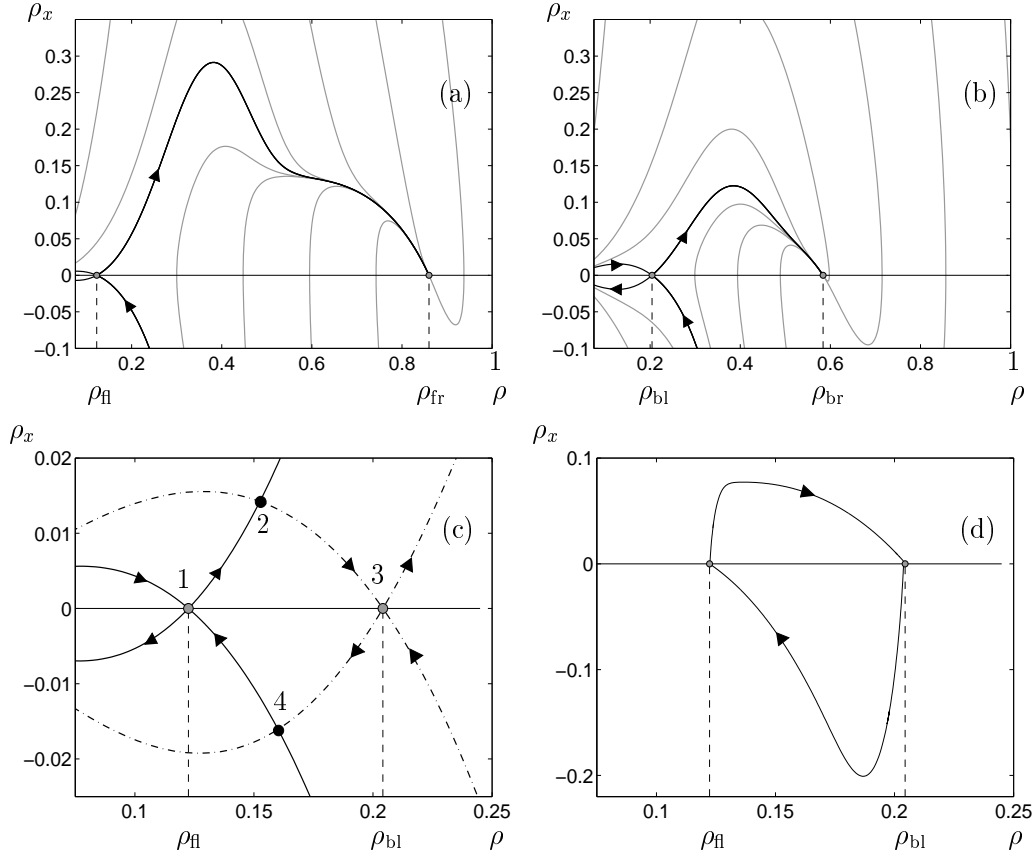


Figure 2.12. Light traffic phase-plane topologies for (2.51, 2.52), with parameters chosen so as to agree with Experiment 1. (a) Unconstrained regime (2.53) and (b) bottleneck (2.54). Flows in the top half-plane go from left to right and flows in the bottom half-plane go from right to left. Dark lines indicate the manifolds associated with the saddle points at $(\rho_{fl}, 0)$ and $(\rho_{bl}, 0)$. The unstable manifolds connect through the upper half-plane with stable nodes at $(\rho_{fr}, 0)$ and $(\rho_{br}, 0)$. (c) Overlay of the separatrices from (a) and (b). Grey markers denote the fixed points and black markers denote where the separatrices meet and hence where the switch between model (2.53) and model (2.54) occurs. Compare with (d) which is a reproduction of Figure 2.1(b).

The goal of the overlay process is to use Figure 2.12(c) to explain the structure of the discrete simulation result Figure 2.11(a) which we reproduce in Figure 2.12(d) in order to aid the comparison. At the qualitative level, there is excellent agreement and thus we may interpret the overall structure of Figure 2.12(d) as a heteroclinic cycle. Furthermore, the simple saddle-like nature of the states that are connected (which we explain shortly using linear theory), explains the monotone front structure of the density profile in Figure 2.1(a). However, at the quantitative level, the comparison fails, as may be seen by comparing the vertical scale of Figure 2.12(c,d). In practice, the coarse-grained density gradient ρ_x achieves much larger values than the solutions

of the PDE model (2.47, 2.48). This should not come as a surprise, since the shock-like structures which we describe break the small amplitude, long wavelength premise on which the PDE derivation is based.

As we have indicated, further insight may be obtained via the linearisation about fixed points of (2.51, 2.52), see [84], where the Jacobian takes the form

$$J = \begin{pmatrix} 0 & 1 \\ g_x & g_y \end{pmatrix}, \quad (2.55)$$

and

$$g_x := \left. \frac{\partial Y'}{\partial X} \right|_{(X,0)} = \frac{6X^2}{\mathcal{V}(X)} Q'(X), \quad (2.56)$$

$$g_y := \left. \frac{\partial Y'}{\partial Y} \right|_{(X,0)} = \frac{3X^2}{\alpha \mathcal{V}(X)} (2\mathcal{V}(X)^2 + \alpha V'(X)). \quad (2.57)$$

The classification of fixed points is then governed by the eigenvalues of J , and in particular by the trace g_y and the determinant $-g_x$.

Firstly note for light traffic fixed points where $X < \rho_{\max}$, we have $Q'(X) > 0$ and consequently $g_x > 0$. Since the determinant is negative, such fixed points are saddles (e.g., the points $(\rho_{\text{fl}}, 0)$ and $(\rho_{\text{bl}}, 0)$ in Figure 2.12).

For heavy traffic fixed points where $X > \rho_{\max}$, we have $Q'(X) < 0$ and consequently $g_x < 0$. Since the determinant is positive, the sign of the trace affects the classification. However, recall that here we operate solely in the stable regime where $\alpha > 2V'(h_*)$ and for such it may be shown that $g_y < 0$. The upshot is that such fixed points are either stable nodes or stable spirals in the (2.51, 2.52) phase-plane (e.g., $(\rho_{\text{fr}}, 0)$ and $(\rho_{\text{br}}, 0)$ in Figure 2.12) depending on how exactly other parameters are chosen.

We now use the PDE-based phase-plane argument to explain the bottleneck solution profile for the three-plateau structure of Example 2 (see Figure 2.1(b) and Figure 2.11(b)). Fig. 2.13(a) reproduces the coarse-grained simulation data of Fig. 2.11(b) with the bottleneck section highlighted in bold, which is then overlaid with numerical solution trajectories of the bottleneck phase-plane (2.51, 2.52, 2.54) in Figure 2.13(b). Since there is excellent agreement in the neighbourhood of the fixed point, we may use the linearisation to explain the structure there. We have $\rho_{\text{sn}} = \rho_{\max}$, and thus the critical condition $Q'(\rho_{\text{sn}}) = 0$ holds. Consequently, $(\rho_{\text{sn}}, 0)$ is a degenerate fixed point of (2.51, 2.52, 2.54) which corresponds to the coalescence of fixed points of saddle and stable node type (see [53, Chap. 3.3]). In the neighbourhood of such a point, the variation in terms of the removed variable x is polynomial rather than exponential.

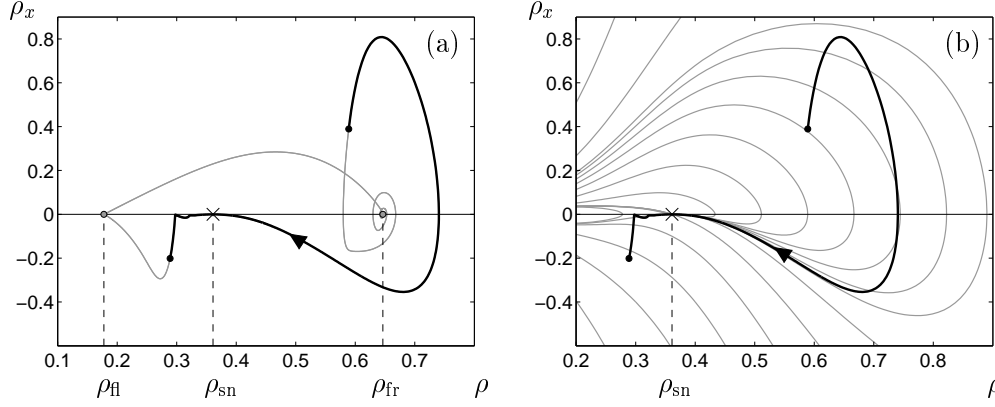


Figure 2.13. (a) Reproduction of Figure 2.11(b) where the portion of the trajectory corresponding to the bottleneck is highlighted in bold. (b) Overlay of (2.51,2.52,2.54) phase-plane trajectories (grey lines) with microscopic simulation data for the bottleneck from (a). The saddle node $(\rho_{sn}, 0)$ is marked with a cross.

Hence equilibration on to the plateau is slow, explaining the pronounced S-shape in the bottleneck profile in Figure 2.1(b).

Finally, we consider the qualitative limitations of the phase-plane argument which relate to self-intersections in Figure 2.11(b,c). For example, the trajectory in Figure 2.11(c) has self-intersections in the neighbourhood of the fixed points $(\rho_1, 0)$ and $(\rho_B, 0)$, and these are not consistent with the second order dynamics, (2.51,2.52,2.53) and (2.51,2.52,2.54) respectively, that should hold at these points. The inescapable conclusion is that the PDE system (2.47, 2.48) is unable to describe these effects, and a higher order model, perhaps a higher order PDE, is required.

The shortcomings of the second order theory are immediate when one considers the linearisation of heavy traffic fixed points $(X, 0)$ for which $X > \rho_{max}$. As shown above, these are either stable nodes or spirals (and hence have no unstable directions). Consequently, solution profiles may relax on to such plateaus in the increasing x direction, but have no subsequent means of escape. However in our set-up, escape is forced by the boundary conditions, and hence second order theory must fail.

2.7 Discrete Spectral Analysis

The goal now is to explain how the discrete model phase-plane trajectories escape from the fixed points in heavy traffic, since we have shown that this behaviour cannot be replicated by (2.51) and (2.52). In Figure 2.11, one can see that the heavy traffic fixed points $(\rho_2, 0)$ in Figure 2.11(b) and $(\rho_B, 0)$ and $(\rho_1, 0)$ in Figure 2.11(c) have

a complicated Shilnikov-type structure, which when projected into two dimensions results in self-intersections. It may well be possible to explain the structure of these fixed points by deriving PDE models which exploit higher order terms in the Taylor expansion of the discrete model, and hence include more dimensions of phase-space. However, the approach here is to work at the linear level and analyse the spectrum of the discrete problem (2.1, 2.2) directly.

The first step is to use $\dot{h}_n = v_{n-1} - v_n$, and consequently $\ddot{h}_n = \dot{v}_{n-1} - \dot{v}_n$ to eliminate velocity and write (2.1, 2.2) in terms of coupled second order ODEs

$$\frac{1}{\alpha} \ddot{h}_n + \dot{h}_n + \{V(h_n) - V(h_{n-1})\} = 0, \quad (2.58)$$

for the headway variables alone. In a travelling wave solution, each vehicle has the same motion as the one in front, but with a time delay of τ , which is given by the reciprocal of the flow rate in the stationary case that we consider here. Thus $h_{n-1}(t) = h_n(t + \tau)$, with

$$\tau = h_*/V(h_*) > 0, \quad (2.59)$$

reduces system (2.58) to the single advance equation

$$\frac{1}{\alpha} \ddot{h}(t) + \dot{h}(t) + \{V(h(t)) - V(h(t + \tau))\} = 0. \quad (2.60)$$

Our interest is in the near-equilibrium behaviour of (2.60), and so we write $h = h_* + \tilde{h}(t)$, $\tilde{h}(t)$ small, and linearise to find

$$\frac{1}{\alpha} \ddot{\tilde{h}} + \dot{\tilde{h}} + V'(h_*) \{\tilde{h} - \tilde{h}(t + \tau)\} = 0, \quad (2.61)$$

which we solve with the exponential ansatz $\tilde{h} = \text{Re}(A \exp(\lambda t))$, where A is real and λ is generally complex. This yields the characteristic equation

$$\frac{1}{\alpha} \lambda^2 + \lambda + V'(h_*) \{1 - e^{\tau \lambda}\} = 0. \quad (2.62)$$

This type of mixed polynomial / exponential equation has received much attention in the past, see e.g. [15]. In contrast to the second order dynamics of (2.51, 2.52), the characteristic equation (2.62) has infinitely many roots, and Figure 2.14 displays numerical solutions for representative parameter values. In each case the spectrum consists of a pair of eigenvalues which correspond to those of the PDE analysis, zero, and an open ‘fan’ of complex conjugate pairs in the right-hand complex plane. Thus in Figure 2.14(b,c), the complex conjugate pairs which are highlighted in bold describe the oscillatory escape from heavy traffic plateaus shown in Figure 2.1(b,c).

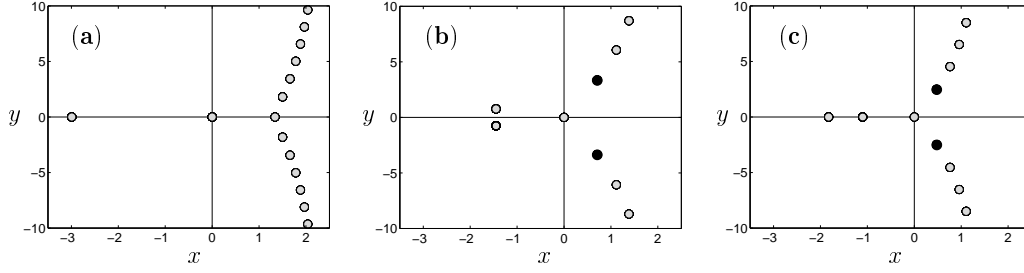


Figure 2.14. The markers in (a-c) denote solutions $\lambda = \psi + i\omega$ to (2.62). In panels (a-c), parameters are chosen such that the spectrum calculated represents respectively: the light traffic bottleneck plateau, as in Figure 2.1(a); the medium density plateau just upstream of the bottleneck, as in Figure 2.1(b); and the high density plateau outside of the bottleneck, as in Figure 2.1(c). The complex conjugate pairs marked in bold in panels (b,c) are those that explain oscillatory x -increasing escape from heavy traffic.

Since Figure 2.14 is based on numerical solution techniques, the remaining challenge is to prove that it is representative of all possible scenarios and that no other balance of eigenvalues between the left-hand and right-hand half-plane is possible. This analysis has three stages:

1. We show that for $\alpha > 2V'(h_*)$, purely imaginary solutions $\lambda = i\omega$, with $\omega \neq 0$, are impossible. Since solutions λ are continuous in problem parameters, spectrum may only cross from left to right half-planes or vice-versa by passing through the origin.
2. Consequently, we find conditions for small solutions λ by using Taylor expansions, and we show that spectrum may only cross the origin when we move between light and heavy traffic cases, i.e. when h^* passes through ρ_{\max} . Thus the number of solutions in the left-hand half-plane may be completely characterised by considering a single light traffic and a single heavy traffic case.
3. We thus focus on the special case $\alpha > 4V'(h_*)$ for which the spectrum may be shown to lie on loci of especially simple form. This procedure allows us to bound left-hand half-plane spectrum and to characterise the open ‘fan’ of eigenvalues in the right-hand half-plane. Since the spectrum does not cross the imaginary axis as we change α provided that we maintain $\alpha > 2V'(h_*)$, this argument establishes that Figure 2.14(a-c) is representative.

Part 1. We substitute $\lambda = i\omega$, ω real, in (2.62) and separate real and imaginary parts. Algebraic terms in ω may be eliminated and we may divide through by a factor of non-zero ω to obtain

$$\frac{\alpha}{V'(h_*)} = \frac{\sin^2 \omega \tau}{1 - \cos \omega \tau}. \quad (2.63)$$

We then expand the trigonometric terms using half-angle formulae to show that (2.63) has no solutions for $\alpha > 2\max V'(h_*)$. Note that (2.63) is equivalent to the neutral stability boundary derived by Bando et al [6].

Part 2. We consider the spectrum λ in the neighbourhood of the origin by expanding the characteristic equation (2.62) to $O(\lambda^2)$ to obtain

$$\lambda \left[\left(\frac{1}{\alpha} - \frac{\tau^2 V'(h_*)}{2} \right) \lambda + (1 - \tau V'(h_*)) \right] = 0, \quad (2.64)$$

which yields

$$\lambda = -2\alpha \left(\frac{1 - \tau V'(h_*)}{2 - \alpha \tau^2 V'(h_*)} \right), \quad (2.65)$$

for the small non-zero root which passes through the origin at

$$\tau V'(h_*) = \frac{h_* V'(h_*)}{V(h_*)} = 1. \quad (2.66)$$

By using $\mathcal{V}(\rho) := V(1/\rho)$ and $Q(\rho) := \rho \mathcal{V}(\rho)$, one may show that (2.66) corresponds to $Q'(\rho) = 0$, i.e., the boundary between light and heavy traffic. Moreover (2.65) is in asymptotic agreement with the small root of the PDE model spectrum, since a small eigenvalue $\hat{\lambda}$ expansion of the Jacobian (2.55) may be mapped directly onto (2.65) via the transformation $h_* = 1/X_*$, observing that $\lambda = \mathcal{V}(X_*)\hat{\lambda}$.

Part 3. We now sketch an analysis in the full complex plane, by writing (2.62) in the form

$$h(\lambda) = f(\lambda) - g(\lambda) = 0, \quad (2.67)$$

where

$$f(\lambda) = \frac{1}{\alpha} \lambda^2 + \lambda + V'(h_*) \quad \text{and} \quad g(\lambda) = V'(h_*) e^{\tau \lambda}. \quad (2.68)$$

Solutions must necessarily lie on the loci of $|f(\lambda)| = |g(\lambda)|$. Hence by writing $\lambda = \psi + i\omega$, where ψ and ω are real, we find

$$\omega^2 = -p(\psi) \pm \sqrt{p(\psi)^2 + \alpha^2 (g(\psi)^2 - f(\psi)^2)}, \quad (2.69)$$

where

$$p(\psi) = \left(\psi + \frac{\alpha}{2} \right)^2 + \alpha \left(\frac{\alpha}{4} - V'(h_*) \right). \quad (2.70)$$

Thus for $\alpha > 4V'(h_*)$, we have $p(\psi) > 0$ for all ψ . Hence provided $f(\psi) < g(\psi)$, (2.69) has a single positive root for ω^2 , and otherwise none. The characterisation of the loci $\omega = \omega(\psi)$ of $|f(\lambda)| = |g(\lambda)|$ is thus achieved via a bifurcation analysis of the real equation

$$f(\psi) = g(\psi), \quad (2.71)$$

whose solutions also yield real roots $\lambda = \psi$ of (2.62).

Firstly, note that $f(\psi) - g(\psi)$ is monotone increasing for $\psi < -\alpha/2$, and that $f(-\alpha) - g(-\alpha) = V'(h_*) - V'(h_*) \exp(-\alpha\tau) > 0$ whereas $f(-\alpha/2) - g(-\alpha/2) = V'(h_*) - \alpha/4 - V'(h_*)e^{\psi\alpha/2} < 0$. Hence there is a unique root $\psi_1 \in (-\alpha, -\alpha/2)$ to (2.71) and none further to the left. Next note that $\psi = 0$ is always a solution to (2.71). The remaining analysis breaks down into a consideration of the heavy traffic case, for which $f'(0) < g'(0)$, hence $f(0-) > g(0-)$, so that there is a consequently an additional root $\psi_2 \in (-\alpha/2, 0)$, and the light traffic case for which the additional root ψ_2 is positive.

An analysis of the pattern of sign changes of $f - g$ then yields the structure of the $|f| = |g|$ loci, which consists of a closed bounded curve in the left-hand half-plane, and an open curve in the right-hand half-plane which is attached to the positive real root in the light traffic case and to the origin in the heavy traffic case. Since f dominates g on any sufficiently large contour to the left of the open section of the locus, Rouché's theorem [71] proves that the solution count of (2.62) is the same as that of $f = 0$ to the left of this curve (i.e., two roots). Hence we establish that Figure 2.14 has captured all the spectrum.

2.8 Summary

In this chapter we have found and analysed interesting macroscopic wave structures that emerge in a simple microscopic traffic model with spatial heterogeneity. Because we posed the problem with periodic (i.e., ring-road) boundary conditions, there is a natural flow balance which leads to stationary density profiles consisting of sets of plateaus separated by sharp interfaces. In the infinite line (open road) setting, one would expect a more complicated outcome in which wave structures can propagate to infinity and this warrants further study, as do more general settings such as merges, lane drops, and moving bottlenecks.

Our overall approach has been multiscale, in that we have tried to understand the macroscopic solution profiles of an underlying microscopic model at four distinct levels of increasing sophistication but also of decreasing tractability:

1. Basic flow and density balances (essentially the zeroth order model): This gives rise to algebraic equations with isolated but non-unique solutions for plateau values.

2. First order conservation law (Lighthill-Whitham model): Characteristic methods impose uniqueness for plateau values but fronts are represented unrealistically by discontinuous shocks.
3. Second order traffic models (of similar type to that of Kerner-Konhäuser [46], but derived [8, 56] directly from the original microscopic model): Global dynamical systems arguments (phase-planes) explain qualitatively most interface structure but cannot explain the (increasing x) escape of trajectories from heavy traffic plateaus.
4. Original microscopic traffic model: This analysis is only tractable at the linearised level but we show how the extra spectra can explain the oscillatory escape from heavy traffic plateaus.

As a general remark, we believe that in a spatially homogeneous setting, all solution features can be adequately described by second order continuum theory. Hence, the principal novelty here is that the model discontinuity forces structures with such short length scales that the assumptions of the second order approximation break down, and consequently these features can only be analysed via the original microscopic model.

We should emphasise that the wave structures which we have found in Section 2.2 will also occur in most other types of car-following model, provided some simple conditions are met: (i) at equilibrium there is a single-valued velocity-headway relationship, that gives rise to (ii) a unimodal flow-density curve; (iii) the bottleneck is implemented by a (possibly nonlinear) down-scaling of this flow-density curve; and (iv) that all uniform flow equilibria are linearly stable. In this setting, the macroscale results in Sections 2.3–2.6 should carry across in a qualitatively similar way, although the details of the discrete spectral analysis in Section 2.7 might turn out to be quite different.

A relatively unexplored area that has emerged from this study is the analysis of global connections in the original (and exact) microscopic model, which is in essence a lattice differential equation. (We emphasise again that in this chapter, the microscopic model is analysed only at the linear level.) Hence there is future scope for a formal bifurcation analysis of the advance equation (2.60).

Chapter 3

Instability in Car-Following Models

“Inspect ruminants appearing in a range of colours (8)”, Mercury 6(17d) [31]

3.1 Motivation

The rich spatio-temporal patterns observed in empirical data are the primary focus of scientific studies of traffic flow [33]. Wave patterns arise frequently during periods of high congestion, and we illustrated some striking examples in Figure 1.3. Despite a long history of theoretical and empirical analysis, the fundamental mechanism that triggers such spatio-temporal patterns is still under debate. In a recent empirical study, Schönhof and Helbing [73] suggested that stop-and-go waves may be attributed to specific events (lane-changes for example) that trigger sizable localised disruptions to traffic flow, from which large-scale jams develop. Another rationale is that uniform traffic is inherently unstable. This philosophy is championed by Sugiyama *et al.* [75], who recently conducted traffic *experiments* in an artificial setting in which 22 participants drove around a 230m long circular road. Vehicles were tuned to initially drive around the loop at about 30km/h with near uniform spacing. Once the flow conditions were suitably smooth, drivers took control of their vehicles and were instructed to attempt to maintain their cruising velocity whilst driving safely. This initially uniform flow was observed to be unstable and quickly (less than a minute) developed into a stop-and-go wave, where vehicles oscillated between a stationary jam phase and a free flowing phase with speeds of approximately 40km/h. Movies of this experiment are available online [75]. However, this work has suffered criticism from those who believe that the drivers were influenced by the experimental setting, and that it therefore does not reflect the dynamics of real traffic.

The occurrence of wave phenomena triggered by congestion in empirical highway traffic data has captured the attention of mathematical modellers since the 1950s [12, 58, 70]. In fact, linear stability analysis of car-following models is well established in the traffic literature [6, 12, 34, 38, 83] and recent studies have also applied dynamical

systems techniques to analyse instability at the non-linear level [24, 69]. While stability and instability in traffic models has been well studied, there are still open problems, and in this chapter we present new insights developed from a mathematical perspective. In particular, these are general to car-following models of the form (1.12), and one need only know that the partial derivatives satisfy the mild constraints (1.17). In Section 3.2 we describe how linear instability of such microscopic traffic models can be captured in a simple condition and in Section 3.3, we extend this method to include models with heterogeneous driver behaviour. A key issue that has received very little attention in the traffic literature is the distinction between so called ‘absolute’ and ‘convective’ instability, which relates to the directional growth of disturbances. In fact, fully developed stop-and-go waves captured in empirical traffic data are only seen to propagate upstream. We use two methods to analyse the directional growth of perturbations: 1. In Section 3.4 we compare group velocities and growth rates (surprisingly only previously considered for the OV model [61]) and 2. In Sections 3.5 and 3.6 we explicitly solve the linearised equations of motion progressively upstream through the column of vehicles. The later parts of this chapter describes recent work that is still in progress. The ultimate goal in developing this area in a general setting is to identify the elements of modern traffic models that faithfully reproduce phenomena observed in empirical data, and hence benchmark future developments in car-following models.

3.2 General Linear Stability Analysis

To qualitatively reflect empirical traffic data, car-following models typically possess a range of parameter values at which uniform flows are linearly unstable and stop-and-go wave phenomena are observed. Thus linear stability analysis techniques have been used to study car-following models since their first appearance in the late 1950s [12, 34], in fact this procedure has become almost routine with each new development (for example the OV model [6]). In this section, we present a generalisation of linear stability analysis to the broad class of car-following models (1.12), for which the conditions (1.17) ensure rational driver behaviour (this method was recently introduced by Wilson [83]). Furthermore, we describe both ring-road and long straight road set ups and highlight the similarities and differences in their analysis.

We consider N vehicles driving on a unidirectional single-lane road with no entrances or exits, where overtaking is not allowed. Vehicles are labelled with subscripts $n = 0, 1, 2, \dots, N - 1$ in the upstream direction and have initial displacements $x_n = L - nh_*$, where $L = Nh_*$ is the length of the loop on a ring-road and $L = 0$ on a

long straight road. The initial spacings are therefore h_* and we set the initial velocities of all vehicles (*except* $n = 1$) to be the equilibrium velocity $V(h_*)$. We perturb the initial velocity of vehicle $n = 1$ by a small amount $\delta \ll V(h_*)$ such that $v_1(0) = V(h_*) + \delta$. On a ring-road the first vehicle follows the last, closing the system. Modelling a long straight road however presents more options. One possibility is to focus on a fixed region in space and specify the up- and downstream boundary conditions. We choose an alternative approach in which we fix the velocity of the lead vehicle $n = 0$ at the equilibrium value $V(h_*)$ for all time and keep track of all N vehicles in the system as they evolve from their initial conditions.

We consider small perturbations about the equilibrium state, $h_n = h_* + \tilde{h}_n$ and $v_n = V(h_*) + \tilde{v}_n$. Linearisation of the equations of motion (1.11) and (1.12) yields

$$\dot{\tilde{h}}_n = \tilde{v}_{n-1} - \tilde{v}_n, \quad (3.1)$$

$$\dot{\tilde{v}}_n = D_h f \tilde{h}_n + D_{h'} f \dot{\tilde{h}}_n + D_v f \tilde{v}_n. \quad (3.2)$$

Differencing the accelerations (3.2) of neighbouring vehicles and substituting in the velocity difference (3.1), we obtain the second-order inhomogeneous equation

$$\ddot{\tilde{h}}_n + (D_{h'} f - D_v f) \dot{\tilde{h}}_n + (D_h f) \tilde{h}_n = (D_{h'} f) \dot{\tilde{h}}_{n-1} + (D_h f) \tilde{h}_{n-1}. \quad (3.3)$$

Note that (3.3) applies to all vehicles on a ring-road but only vehicles $n > 0$ on the long-straight road. To analyse the growth of small perturbations, we introduce the ansatz $\tilde{h}(n, t) = \text{Re}(A \exp(i\theta n + \lambda t))$, and look for solutions with $\text{Re}(\lambda) > 0$. On the ring-road, periodic boundary conditions require that the length of the loop L is divisible by the wave-length of solutions, hence the wave-number θ is discrete. In contrast, no such constraints exist on the long-straight road and θ is continuous. For either boundary conditions, substitution of the ansatz yields

$$\lambda^2 + [D_{h'} f (1 - e^{-i\theta}) - D_v f] \lambda + D_h f (1 - e^{-i\theta}) = 0. \quad (3.4)$$

Solving for λ , we find the two temporal modes

$$\lambda_{\pm}(\theta) = \frac{1}{2} \left\{ - [D_{h'} f (1 - e^{-i\theta}) - D_v f] \pm \sqrt{[D_{h'} f (1 - e^{-i\theta}) - D_v f]^2 - 4D_h f (1 - e^{-i\theta})} \right\}. \quad (3.5)$$

Note that for solutions to (3.4) can be written in the form

$$\lambda(\theta) = \lambda_R(\theta) + i\lambda_I(\theta), \quad (3.6)$$

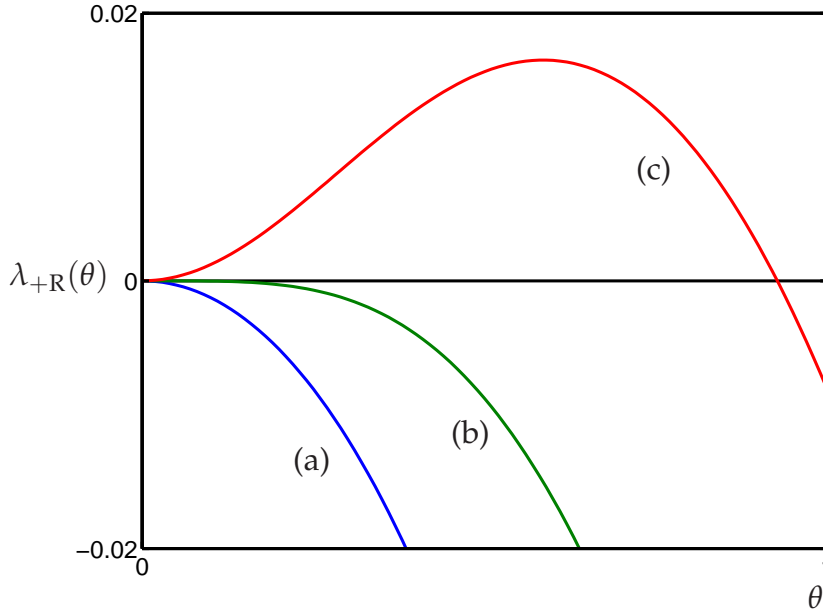


Figure 3.1. Growth rates, λ_{+R} , for the OVRV model. Parameter values are $\alpha = 1.6$ and $\beta = 0.4, 0.2, 0$ for curves (a), (b) and (c) respectively. The temporal mode λ_{-R} does not affect the stability in these cases. The values $\lambda = 0, \theta = 0$ are always solutions. Curve (a) bends down at the origin, hence the parameters are stable. Curve (b) is tangent to the θ axis at the origin, and is thus marginally stable. Curve (c) bends up at the origin and hence the parameters are unstable. Thus instability emerges at small wave-number, i.e. at $\theta = 0$, and hence long wave-length.

which, by substitution into (3.4), can be shown to have the symmetry $\lambda(-\theta) = \overline{\lambda(\theta)}$. Thus $\lambda_R(\theta)$ is an even function and $\lambda_I(\theta)$ is an odd function. Furthermore, $-\lambda_I(\theta)$ is the dispersion relation and $\lambda_R(\theta)$ is the growth rate.

It is shown in [83] that instability emerges at long wave-length. We illustrate this in Figure 3.1 by plotting the growth rates $\lambda_R(\theta)$ for the OVRV model. Three parameter values are chosen depicting stability, marginal stability and instability. Note that $\lambda = 0, \theta = 0$ solve (3.4) and thus the system only becomes unstable when the growth rate $\lambda_R(\theta)$ bends upwards at $\theta = 0$. To establish when this occurs we introduce the perturbation expansion

$$\lambda = i\lambda_1\theta + \lambda_2\theta^2 + O(\theta^3), \quad (3.7)$$

and determine the behaviour at $\theta = 0$. The first two coefficients are then

$$\lambda_1 = \frac{D_h f}{D_v f}, \quad \text{and} \quad \lambda_2 = \frac{D_h f}{D_v f^3} \left(\frac{D_v f^2}{2} - D_{hf} D_v f - D_h f \right). \quad (3.8)$$

The symmetry property of λ ensures that the stability is governed by the even coefficients. Hence the system is unstable when $\lambda_2 > 0$, which, with the sign constraints

(1.17), occurs when

$$\frac{1}{2}(D_v f)^2 - (D_h f) - (D_h f)(D_v f) < 0. \quad (3.9)$$

This is a sufficient condition for instability on the long straight road, however a large number of vehicles may be necessary to observe weak instability. Note that on the ring-road, the maximum possible excitable wave-length is the length of the loop. Thus satisfying the inequality (3.9) does not guarantee instability on a periodic domain.

3.3 Heterogeneous Driver Behaviour

We have developed a condition for instability in a general class of car-following models for the simple case where vehicles are identical. The traffic recorded by motorway detectors however is composed of a wide variety of vehicle types with different physical attributes, whose drivers have a completely unique behaviour. We are interested in how incorporating such distributions of attributes changes the properties of the system. In this section we extend the stability analysis of the general car-following model presented in Section 3.2 to include heterogeneous driver behaviour. Hence we consider the stability of mixed populations of vehicles with different characteristics, for example cars and trucks. Intuitively one would expect that the marginal stability boundary would lie between that of the extreme cases, what is not clear however is how exactly each population affects the stability of the system and whether the order of vehicles is important. Our goal in this section is to quantify how these factors impact system-wide stability on a ring-road via a generalisation of the criterion (3.9). Our technique combines the perturbation method described in Section 3.2 with the approach adopted by Mason and Woods [60], in which the entire system is described via a matrix equation.

There are subtle differences between the model set up for homo- and heterogeneous drivers. Once again we consider N cars travelling on a single-lane ring-road where overtaking is not allowed. We consider models of the form

$$\dot{h}_n = v_{n-1} - v_n, \quad (3.10)$$

$$\dot{v}_n = f_n(h_n, v_{n-1} - v_n, v_n). \quad (3.11)$$

Note that the subscript on the acceleration function f_n generalises (1.12) to include driver heterogeneity. The usual constraints on the partial derivatives apply, but the precise values differ between vehicle populations. We assume that one can identify an equilibrium velocity function with each driver V_n . Thus a given spacing may result in a range of equilibrium velocities. Equilibria of the system require that all vehicles have

the same velocity, hence we must specify the equilibrium velocity v_* and determine individual equilibrium headways such that

$$V_n(h_{*,n}) = v_*, \quad (3.12)$$

where the length of the loop is $\sum_n h_{*,n} = L$. It follows that vehicles are not necessarily distributed evenly around the ring when in equilibrium.

In a manner analogous to Section 3.2, we consider small perturbations to the equilibrium $h_n = h_{*,n} + \tilde{h}_n$ and $v_n = V(h_*) + \tilde{v}_n$. Thus, linearising the equations of motion (3.10) and (3.11) about the equilibria and differencing the acceleration of two adjacent vehicles yields the second-order inhomogeneous equation

$$\ddot{\tilde{h}}_n + (D_h f - D_v f_n) \dot{\tilde{h}}_n + (D_h f_n) \tilde{h}_n = (D_h f_{n-1}) \dot{\tilde{h}}_{n-1} + (D_h f_{n-1}) \tilde{h}_{n-1}. \quad (3.13)$$

We are interested in the spatio-temporal development of these small disturbances, hence we consider the ansatz

$$\tilde{h}_n = \text{Re} (A_n \exp (i\theta n + \lambda t)). \quad (3.14)$$

Substitution into (3.13) yields

$$\lambda^2 A_n + (D_h f - D_v f_n) \lambda A_n + (D_h f_n) A_n = [(D_h f_{n-1}) \lambda A_{n-1} + (D_h f_{n-1}) A_{n-1}] e^{-i\theta}. \quad (3.15)$$

We can write this as a system of equations in the form

$$\lambda^2 \begin{bmatrix} A_1 \\ \vdots \\ A_N \end{bmatrix} = \mathbf{M} \begin{bmatrix} A_1 \\ \vdots \\ A_N \end{bmatrix} \quad (3.16)$$

where the matrix \mathbf{M} is

$$\mathbf{M} = \begin{bmatrix} -(D_h f_1 - D_v f_1) \lambda - D_h f_1 & 0 & \dots & (D_h f_N \lambda + D_h f_N) e^{-i\theta} \\ (D_h f_1 \lambda + D_h f_1) e^{-i\theta} & -(D_h f_2 - D_v f_2) \lambda - D_h f_2 & \dots & 0 \\ 0 & (D_h f_2 \lambda + D_h f_2) e^{-i\theta} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & -(D_h f_N - D_v f_N) \lambda - D_h f_N \end{bmatrix}. \quad (3.17)$$

Thus solutions to (3.16) must satisfy

$$\begin{vmatrix} \lambda^2 + (D_h f_1 - D_v f_1)\lambda + D_h f_1 & 0 & \dots & -(D_h f_N \lambda + D_h f_N)e^{-i\theta} \\ -(D_h f_1 \lambda + D_h f_1)e^{-i\theta} & \lambda^2 + (D_h f_2 - D_v f_2)\lambda + D_h f_2 & \dots & 0 \\ 0 & -(D_h f_2 \lambda + D_h f_2)e^{-i\theta} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda^2 + (D_h f_N - D_v f_N)\lambda + D_h f_N \end{vmatrix} = 0, \quad (3.18)$$

where $|\cdot|$ denotes the determinant. The method of Mason and Woods produces a similar equation for the OV model, however it does not capture the *spatio*-temporal development.

The form of the matrix \mathbf{M} allows us to write (3.18) in the form

$$\prod_{n=1}^N (\lambda^2 + (D_h f_n - D_v f_n)\lambda + D_h f_n) - e^{iN\theta} \prod_{n=1}^N (\lambda D_h f_n + D_h f_n) = 0, \quad (3.19)$$

(see Appendix A for a derivation). The commutability of the products in (3.19) means that the stability of the system is independent of the order of vehicles. Note that as in the single-class case, $\lambda = 0$ and $\theta = 0$ solve (3.19). In the manner of [83], we seek small λ solutions via the perturbation expansion (3.7), namely $\lambda = i\lambda_1\theta + \lambda_2\theta^2 + \dots$, and attempt to determine the nature of the characteristic equation (3.19) at the origin. This analysis determines a sufficient condition for instability in this case. Substituting the perturbation expansion (3.7) into (3.19) yields

$$\begin{aligned} & \prod_{n=1}^N \left([i\lambda_1\theta + \dots]^2 + (D_h f_n - D_v f_n) [i\lambda_1\theta + \lambda_2\theta^2 + \dots] + D_h f_n \right) \\ & - \left(1 + iN\theta - \frac{N^2}{2}\theta^2 + \dots \right) \prod_{n=1}^N ([i\lambda_1\theta + \lambda_2\theta^2 + \dots] D_h f_n + D_h f_n) = 0. \end{aligned} \quad (3.20)$$

We refer to the components of this equation as the left-product (top-row of (3.20)), the exponential (bottom line left of (3.20)) and the right-product (bottom line right of (3.20)). The order θ terms are composed of the following: 1. The left-product contributes $(D_h f_n - D_v f_n)i\lambda_1$ for each n multiplied by the constant terms $D_h f_{m \neq n}$ of all other vehicles; 2. The right-product and $O(\theta^0)$ terms in the exponential contribute $-D_v f_n i\lambda_1$ for each n multiplied by the constant terms $D_h f_{m \neq n}$ of all other vehicles; 3. The $O(\theta^1)$ term in the exponential multiplied by the product of all constant terms $D_h f_n$

in the right product. Collecting all the order θ terms, we find

$$\begin{aligned} O(\theta^1) : \quad & i\lambda_1 \sum_n (D_h f_n - D_v f_n) \prod_{m \neq n} D_h f_m \\ & - i\lambda_1 \sum_n D_h f_n \prod_{m \neq n} D_h f_m + iN \prod_n D_h f_n = 0, \end{aligned} \quad (3.21)$$

and solving for λ_1 yields

$$\lambda_1 = \frac{N \prod_n D_h f_n}{\sum_n D_v f_n \prod_{m \neq n} D_h f_m}. \quad (3.22)$$

Similarly, the order θ^2 terms are composed of the following: 1. The $O(\theta^2)$ terms in the left-product contribute $-\lambda_1^2 + (D_h f_n - D_v f_n) \lambda_2$ for each n multiplied by the constant terms $D_h f_{m \neq n}$ of all other vehicles; 2. The left-product also contributes the cross-terms $-(D_h f_i - D_v f_i) (D_h f_j - D_v f_j) \lambda_1^2$ of all *combinations* of i and j multiplied by the constant terms $D_h f_{m \neq i,j}$ of all other vehicles; 3. The right-product and $O(\theta^0)$ terms in the exponential contribute $-D_v f_n \lambda_2$ for each n multiplied by the constant terms $D_h f_{m \neq n}$ of all other vehicles from the terms of $O(\theta^2)$; 4. The right-product and $O(\theta^0)$ terms in the exponential also contribute the cross-terms $-D_h f_i D_h f_j \lambda_1^2$ of all *combinations* of i and j multiplied by the constant terms $D_h f_{m \neq i,j}$ of all other vehicles; 5. The $O(\theta^1)$ terms in the exponential multiplied by the $O(\theta^1)$ terms in the right-product; 6. The $O(\theta^2)$ terms in the exponential multiplied by the product of all constant terms $D_h f_n$ in the right product. Collecting all order θ^2 terms, we find

$$\begin{aligned} O(\theta^2) : \quad & \sum_n [-\lambda_1^2 + (D_h f_n - D_v f_n) \lambda_2] \prod_{m \neq n} D_h f_m \\ & - \lambda_1^2 \sum_{i \neq j, j > i} (D_h f_i - D_v f_i) (D_h f_j - D_v f_j) \prod_{m \neq i,j} D_h f_m \\ & + \lambda_2 \sum_n D_h f_n \prod_{m \neq n} D_h f_m + \lambda_1^2 \sum_{i \neq j, j > i} D_h f_i D_h f_j \prod_{m \neq i,j} D_h f_m \\ & - N\lambda_1 \sum_n D_h f_n \prod_{m \neq n} D_h f_m + \frac{N^2}{2} \prod_n D_h f_n = 0. \end{aligned} \quad (3.23)$$

The next step is to collect the λ_2 terms on the right-hand side of (3.23). Furthermore, we expand the $(D_h f_i - D_v f_i) (D_h f_j - D_v f_j)$ term, from which we can cancel the $D_h f_i D_h f_j$ terms. This leaves us with all *permutations* of $D_h f_i D_v f_j$ and all *combinations* of

$D_v f_i D_v f_j$. These changes yield

$$\begin{aligned} \left[\sum_n D_v f_n \prod_{m \neq n} D_h f_m \right] \lambda_2 = & -\lambda_1^2 \sum_n \prod_{m \neq n} D_h f_m + \lambda_1^2 \sum_{i \neq j} D_h f_i D_v f_j \prod_{m \neq i, j} D_h f_m \\ & - \lambda_1^2 \sum_{i \neq j, j > i} D_v f_i D_v f_j \prod_{m \neq i, j} D_h f_m - N \lambda_1 \sum_n D_h f_n \prod_{m \neq n} D_h f_m + \frac{N^2}{2} \prod_n D_h f_n. \end{aligned} \quad (3.24)$$

We now introduce one of three identities required in this derivation (for further details see Appendix A). The first term on the right-hand side of (3.24) obeys the identity

$$\sum_n \prod_{m \neq n} D_h f_m = \frac{\sum_n D_h f_n \left[\prod_{m \neq n} D_h f_m \right]^2}{\prod_n D_h f_n}. \quad (3.25)$$

We then factor $\lambda_1^2 / \left[\prod_n D_h f_n \right]$ from the right-hand side of (3.24) to get

$$\begin{aligned} \left[\sum_n D_v f_n \prod_{m \neq n} D_h f_m \right] \lambda_2 = & \frac{\lambda_1^2}{\prod_n D_h f_n} \left\{ - \sum_n D_h f_n \left[\prod_{m \neq n} D_h f_m \right]^2 \right. \\ & + \left[\sum_{i \neq j} D_h f_i D_v f_j \prod_{m \neq i, j} D_h f_m \right] \left[\prod_n D_h f_n \right] - \left[\sum_{i \neq j, j > i} D_v f_i D_v f_j \prod_{m \neq i, j} D_h f_m \right] \left[\prod_n D_h f_n \right] \\ & \left. - \left[\sum_n D_h f_n \prod_{m \neq n} D_h f_m \right] \left[\sum_n D_v f_n \prod_{m \neq n} D_h f_m \right] + \frac{1}{2} \left[\sum_n D_v f_n \prod_{m \neq n} D_h f_m \right]^2 \right\}. \end{aligned} \quad (3.26)$$

There are five terms in braces on the right-hand side of (3.26), which we refer to as terms 1–5 respectively. The first of the two remaining identities relates to the permutation term (term 2 in (3.26))

$$\begin{aligned} \sum_{i \neq j} D_h f_i D_v f_j \prod_{m \neq i, j} D_h f_m = & \frac{1}{\prod_n D_h f_n} \left\{ \left[\sum_n D_h f_n \prod_{m \neq n} D_h f_m \right] \left[\sum_n D_v f_n \prod_{m \neq n} D_h f_m \right] \right. \\ & \left. - \sum_n D_h f_n D_v f_n \left[\prod_{m \neq n} D_h f_m \right]^2 \right\}, \end{aligned} \quad (3.27)$$

which on substitution into (3.26) allows term 4 in (3.26) to be cancelled. The final

identity relates to the combination term (term 3 in (3.26))

$$\sum_{i \neq j, j > i} D_v f_i D_v f_j \prod_{m \neq i, j} D_h f_m = \frac{1}{2 \prod_n D_h f_n} \left\{ \left[\sum_n D_v f \prod_{m \neq n} D_h f_m \right]^2 - \sum_n D_v f^2 \left[\prod_{m \neq n} D_h f_m \right]^2 \right\}, \quad (3.28)$$

which on substitution into (3.26) allows term 5 in (3.26) to be cancelled. Combining (3.26–3.28), we find

$$\lambda_2 = \frac{N^2 \prod_n D_h f_n}{\sum_n D_v f_n^2 \prod_{m \neq n} D_h f_m} \left\{ \sum_n \left[\frac{D_v f_n}{2} - D_h f_n D_v f_n - D_h f_n \right] \left[\prod_{m \neq n} D_h f_m \right]^2 \right\}. \quad (3.29)$$

With the sign constraints (1.17), for a system of traffic with heterogeneous driver behaviour on a ring-road, instability occurs if

$$\sum_n \left[\frac{D_v f_n^2}{2} - D_h f_n D_v f_n - D_h f_n \right] \left[\prod_{m \neq n} D_h f_m \right]^2 < 0. \quad (3.30)$$

Note that if there is only one population of vehicles then the instability condition (3.9) is recovered. As a simple example, we consider just two populations of vehicles. Of the total population N , we suppose that there are N_C cars and N_T trucks (where subscripts C and T refer to cars and trucks respectively). The instability condition (3.30) is then

$$\eta D_h f_T^2 \left[\frac{D_v f_C^2}{2} - D_h f_C D_v f_C - D_h f_C \right] \quad (3.31)$$

$$+ (1 - \eta) D_h f_C^2 \left[\frac{D_v f_T^2}{2} - D_h f_T D_v f_T - D_h f_T \right] < 0, \quad (3.32)$$

where $\eta = N_C / N$ is the fraction of cars (and hence $1 - \eta = N_T / N$), and (3.32) is thus independent of the total number of vehicles N .

To verify (3.30) with a numerical example, we focus on the OVRV model (1.8), which we adapt in the following way to include driver heterogeneity

$$\ddot{h}_n = \alpha \{V_n(h_n) - v\} + \beta \dot{h}_n. \quad (3.33)$$

For two vehicle populations we choose

$$V_n(h) = \begin{cases} V(h) & : \text{Car} \\ \gamma V(h) & : \text{Truck} \end{cases}, \quad (3.34)$$

where $V(h)$ is the standard OV function (1.6) and $0 < \gamma < 1$. If all vehicles are trucks then the stability criterion is simply $\alpha/2 > \gamma V'(h_*)$, as could be obtained from (3.9). Similarly, if all vehicles are cars then the stability criterion is $\alpha/2 > V'(h_*) > \gamma V'(h_*)$; hence trucks *stabilise* the system. Note that one might also choose the sensitivity α_T of trucks to be lower than that of cars α_C . In this case, cars (larger sensitivity values) act to stabilise whereas trucks (smaller sensitivity values) act to destabilise.

Uniform flows are parameterised in terms of the equilibrium velocity v_* . Since the equilibrium headways differ between populations of vehicles, we introduce the equilibrium headway function $h = H(v)$. The equilibrium spacing for cars is therefore $h_{*,C} = H(v_*)$ and for trucks is $h_{*,T} = H(v_*/\gamma)$. The explicit form of the equilibrium headway for the standard OV function (1.6) is

$$H(v) = \operatorname{arctanh}(v - \tanh(2)) + 2, \quad (3.35)$$

for velocities in the range $v \in [0, V_{\max})$, where $V_{\max} = 1 + \tanh(2)$. It follows that the equilibrium velocity must satisfy $v_*/V_{\max} < \gamma$. The partial derivatives of the acceleration function are

$$D_h f_C = \alpha V'(H(v_*)), \quad (3.36)$$

$$D_h f_T = \gamma \alpha V'(H(v_*/\gamma)), \quad (3.37)$$

$$D_h f_C = D_h f_T = \beta \quad \text{and} \quad (3.38)$$

$$D_v f_C = D_v f_T = -\alpha, \quad (3.39)$$

which on substitution into the stability criterion (3.32) yields

$$\eta [\gamma V'(H(v_*/\gamma))]^2 \left(\frac{\alpha}{2} - \beta - V'(H(v_*)) \right) \quad (3.40)$$

$$+ (1 - \eta) [V'(H(v_*))]^2 \left(\frac{\alpha}{2} - \beta - \gamma V'(H(v_*/\gamma)) \right) < 0. \quad (3.41)$$

There are therefore five free parameters that control the stability of the system: the fraction of cars η , the equilibrium velocity v_* and the model parameters α , β and γ .

In Figure 3.2 we illustrate marginal stability curves (3.41) for different values of η (grey scale separated by solid lines) in the (γ, v_*) plane, with model parameter values $\alpha = 1.4$, $\beta = 0.2$. The grey scale has been split up into ten regions, the edges of which are marked with a contour labelled according to the value of η . We have marked the

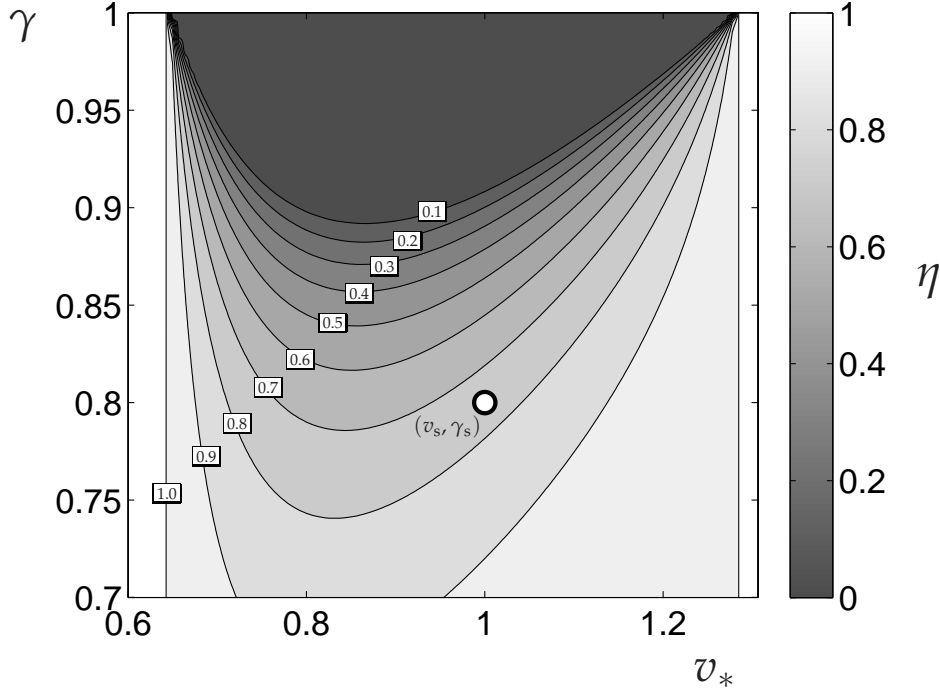


Figure 3.2. Marginal stability curves of the OVRV model with driver heterogeneity for the parameter values $\alpha = 1.4$ and $\beta = 0.2$. Grey scale contours of the fraction of cars η are plotted in the (γ, v_*) plane, where γ is the OV scaling coefficient (i.e. how much slower the slow vehicles are) and v_* is the equilibrium velocity. At the point labelled (v_s, γ_s) , the values are $v_s = 1$, $\gamma_s = 0.8$ and marginal stability occurs at $\eta \approx 0.76$.

point $(v_s, \gamma_s) = (1, 0.8)$, which corresponds to $\eta \approx 0.76$, i.e. between the contours $\eta = 0.7$ and $\eta = 0.8$. For example, if $N = 100$ vehicles, then the system is stable if there are $N_C = 70$ cars (hence $N_T = 30$ trucks) and unstable if there are $N_C = 80$ cars (hence $N_T = 20$ trucks). This is confirmed in Figure 3.3, where the velocity time series of a single vehicle in each case is compared. Notice that in the unstable case, Figure 3.3(a), there is initially decay and growth is only observed over a long enough period of time.

3.4 Absolute and Convective Instability

We now turn our attention to the convective nature of stop-and-go waves. Jams which propagate along highways develop from some initial perturbation (be it small or large), which in some traffic models may be carried upstream, downstream or even in both directions simultaneously depending on the velocity of the traffic. Such instabilities are known as ‘convective’ when the growth is uni-directional and ‘absolute’

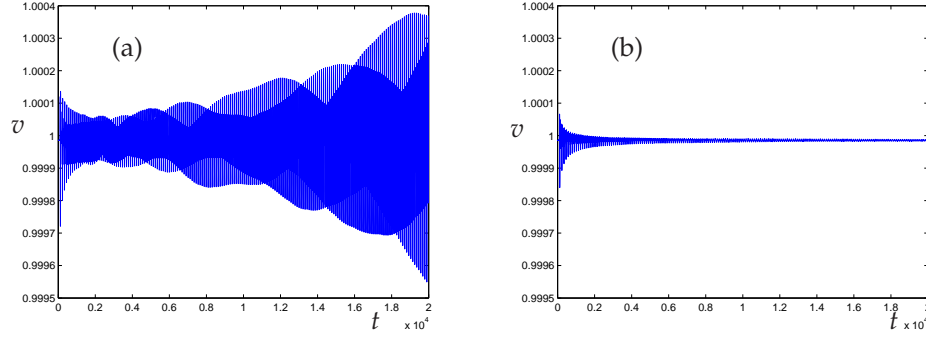


Figure 3.3. Unstable (panel (a)) and stable (panel (b)) velocity time-series of a single vehicle from simulations of the OVRV model with heterogeneous driver behaviour (3.34). Model parameters are $\alpha = 1.4$, $\beta = 0.2$, $\gamma = 0.8$, $N = 100$ vehicles and the equilibrium velocity is $v_* = 1$. In panel (a) there are $N_C = 70$ cars (hence $N_T = 30$ trucks) and thus the system is unstable. In panel (b) there are $N_C = 80$ cars (hence $N_T = 20$ trucks) and the system is stable. Notice that in panel (a), the velocity disturbance initially decays and growth is only observed over long timescales.

when the growth is bi-directional. Very little attention has been paid to this distinction in the traffic literature (to the best of our knowledge, only the OV model has been considered in [61]), despite the fact that convective and absolute instability have been well studied in the hydrodynamics community [41, 72]. Moreover, in recent studies of empirical data, Schönhof and Helbing [73] suggest that the underlying mechanism that generates ‘large moving clusters’ of traffic (essentially stop-and-go waves) is what they call the ‘boomerang effect’. From data collected in Germany, they observe that perturbations in density initially grow and propagate *downstream* before being ‘reflected’ back upstream by congested regions around a bottlenecks. Thus identifying transitions in the directional propagation of disturbances is of fundamental importance if traffic models are to capture such complex behaviour.

The fundamental difference between absolute and convective instability is whether at a *fixed location* (i.e. the frame of reference of the road, referred to as the laboratory frame in physics) one observes sustained growth from equilibrium or eventual decay. This is only applicable on the long straight road since vehicles on a ring-road periodically encounter a disturbance as it propagates around the loop. Thus we consider a long straight road set up in which the velocity of the lead (or zeroth) vehicle in a long chain of cars is held constant. Vehicles are identical and initially in equilibrium with spacing h_* and velocity $V(h_*)$. However, we perturb the velocity of the first vehicle in the chain at $t = 0$, and the system is subsequently left to evolve (as described in Section 3.2). When the system is linearly unstable, this perturbation grows as it propagates through the line of cars. The directional propagation of such a disturbance is

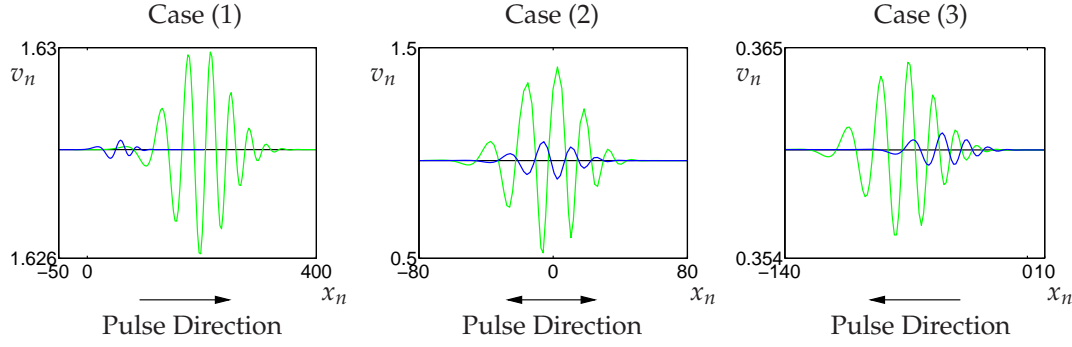


Figure 3.4. Numerical examples of convective (Cases (1) and (3)) and absolute (Case (2)) instability using the OVRV model. The OVRV model is used with unstable parameter values $\alpha = 0.5$, $\beta = 0.2$ and $h_* = 2.8, 2, 1.3$ for Cases (1–3) respectively. A fixed time-step of $dt = 0.025$ is used and the number of vehicles is $N = 300$. Velocity v_n is plotted against position x_n at two instances in time and these points are joined with a line. Blue lines are earlier times ($t = 125, 100, 225$ for Cases (1–3) respectively) than green lines ($t = 500, 125, 325$ for Cases (1–3) respectively). Vehicles move in the increasing x -direction, i.e. left to right. The initial disturbance originates close to the origin. Cases (1) and (3) illustrate convective instability because the perturbation is carried down- and upstream respectively. Case (2) illustrates absolute instability since the disturbance grows both up- and downstream simultaneously.

illustrated in Figure 3.4 with simulations of the OVRV model using the standard OV function (1.6). In Case (1) the velocity of traffic is high ($V(h_*) \approx 1.62$) and the perturbation (initially near the origin) grows downstream in the positive x -direction. In Case (3) the velocity of traffic is low ($V(h_*) \approx 0.35$) and the same perturbation is seen to grow upstream. Both Cases (1) and (3) are illustrations of convective instability. In Case (2) however, the traffic velocity is an intermediate value ($V(h_*) \approx 0.96$), and the disturbance grows both up- and downstream simultaneously; thus Case (2) is an example of absolute instability.

We now consider the linearisation of the general car-following model (3.3), which is posed in the ‘lattice frame’, i.e. moving with the vehicles. The lattice corresponds to the indices n , which by our convention increase against the flow of traffic. As a first step towards characterising absolute and convective instability, we introduce the *lattice group velocity*,

$$c_n(\theta) := -\lambda'_{\pm I}(\theta) = \text{Im} \left(-\frac{1}{2} \left\{ -D_h f \pm \frac{D_h f [D_h f (1 - e^{-i\theta}) - D_v f] - 2D_h f}{\sqrt{[D_h f (1 - e^{-i\theta}) - D_v f]^2 - 4D_h f (1 - e^{-i\theta})}} \right\} i e^{-i\theta} \right). \quad (3.42)$$

Note that (3.42) matches the standard definition $\omega'(k)$ for pure harmonic ansatz of the form $\tilde{h}_n(t) = \text{Re}(A \exp(i\theta n - i\omega t))$. The lattice group velocity is the speed at which a

particular amplitude of the *envelope* of the disturbance propagates through the lattice. However, we have noted that the salient reference frame with regards to convective and absolute instability in traffic is the reference frame of the road. A fixed position relative to the road moves with velocity $n/t = V(h_*)/h_*$ through the lattice. Hence we define the *absolute* (i.e. fixed position) group velocity

$$c_x(\theta) := c_n(\theta) - \frac{V(h_*)}{h_*}. \quad (3.43)$$

The absolute group velocity is thus the speed at which a particular amplitude of the *envelope* of the disturbance propagates with respect to the reference frame of the *road*.

A fixed position has zero absolute group velocity, hence we are interested in real wave-number θ_g solutions to

$$c_x(\theta_g) = 0. \quad (3.44)$$

Absolute instability occurs when the wave-number θ_g is within the range of unstable wave-numbers, since the amplitude of the *envelope* of the growing disturbance is constant at its origin and hence the perturbation is growing both up- and downstream simultaneously. Conversely, when the system is unstable but no such wave-number θ_g exists or lies outside of the range of unstable values, then the flow is convectively unstable.

We can determine whether the wave-number θ_g lies within the range of unstable wave-numbers if we can identify the largest unstable (real) wave-number θ_d , i.e. solutions to

$$\lambda_R(\theta_d) = 0, \quad (3.45)$$

for $\theta_d > \phi$ where $\lambda_{\pm R}(\phi) > 0$. Hence the system is absolutely unstable when

$$\theta_d > \theta_g. \quad (3.46)$$

In Figure 3.5, we illustrate how the growth rates and absolute group velocity can be used to distinguish between absolute and convective instability for the Cases (1–3), illustrated in Figure 3.4. The growth rates $\lambda_{+R}(\theta)$ are plotted in blue and the absolute group velocities (3.43) are plotted in green. In Case (1) there is no solution θ_g , whereas in Case (3) the solution lies outside of the unstable wave numbers, (i.e. $\theta_g > \theta_d$). In Case (2), the solution θ_g lies within the range of unstable wave numbers (i.e. $\theta_g < \theta_d$). Thus this method correctly predicts that Cases (1) and (3) are convectively unstable, whereas Case (2) is absolutely unstable.

Although we can apply this technique on a case by case basis, we would like to identify the regions in parameter space where the system is convectively/absolutely

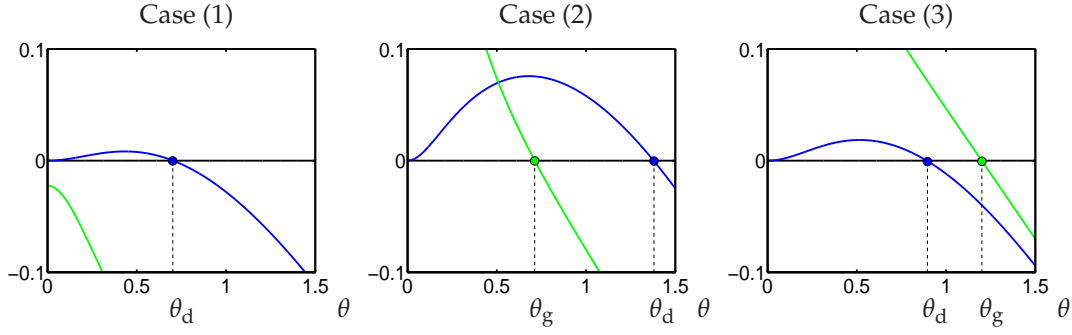


Figure 3.5. Growth rates $\lambda_{+R}(\theta)$ (blue) and absolute group velocities (green) for Cases (1–3). As in Figure 3.4, Cases (1–3) correspond to the OVRV model with $\alpha = 0.5$, $\beta = 0.2$ and $h_* = 2.8, 2.0, 1.3$ respectively. The largest unstable wave-numbers are labelled θ_d and the zero absolute group velocities are labelled θ_g . There is no θ_g solution in Case (1), whereas in Case (3) $\theta_g > \theta_d$. Hence Cases (1) and (3) are *convectively* unstable. In Case (2), θ_g is within the range of unstable wave-numbers (i.e. $\theta_g < \theta_d$). Hence Case (2) is *absolutely* unstable.

unstable. To achieve this, we extend the perturbation expansion (3.7) to fourth order to determine an approximation for the absolute group velocity and higher order terms in the growth rate $\lambda_{+R}(\theta)$, from which we can estimate the largest unstable wave-number θ_d . The first two coefficients of the fourth order perturbation expansion,

$$\lambda = i\lambda_1\theta + \lambda_2\theta^2 + i\lambda_3\theta^3 + \lambda_4\theta^4 + O(\theta^5), \quad (3.47)$$

are given by (3.8). The third and fourth order coefficients are more complicated,

$$\lambda_3 = -\frac{D_h f}{D_v f^5} \left\{ \frac{D_v f^4}{6} - D_h f D_v f^3 + D_h f^2 D_v f^2 - 2D_h f \left[\frac{D_v f^2}{2} - \frac{3}{2}D_h f D_v f - D_h f \right] \right\}, \quad (3.48)$$

and

$$\lambda_4 = -\frac{D_h f}{24D_v f^7} \left(D_v f^6 - 14D_h f D_v f^5 + 36D_h f^2 D_v f^4 - 144D_h f D_h f^2 D_v f^2 - 24D_h f^3 D_v f^3 + 108D_h f D_h f D_v f^3 - 240D_h f^2 D_h f D_v f - 14D_h f D_v f^4 + 72D_h f^2 D_v f^2 - 120D_h f^3 \right), \quad (3.49)$$

respectively. The growth rate can therefore be approximated by the second and fourth order coefficients

$$\lambda_R(\theta) \approx \hat{\lambda}_R(\theta) := \lambda_2\theta^2 + \lambda_4\theta^4. \quad (3.50)$$

For small wave-numbers, the largest unstable wave-number $\hat{\theta}_d > 0$ which solves

$\hat{\lambda}_R(\theta_d) = 0$ is

$$\hat{\theta}_d = +\sqrt{-\frac{\lambda_2}{\lambda_4}}. \quad (3.51)$$

The absolute group velocity can be approximated by

$$c_x(\theta) \approx \hat{c}_x(\theta) := -\lambda_1 - 3\lambda_3\theta^2 - \frac{V(h_*)}{h_*}. \quad (3.52)$$

Note that

$$\lambda_1 = \frac{D_h f}{D_v f} < 0, \quad (3.53)$$

and for unstable parameter values, the sign constraints (1.17) ensure that $\lambda_3 > 0$, since the term in square brackets in (3.48) is negative,

$$0 > \frac{D_v f^2}{2} - D_h f D_v f - D_h f > \frac{D_v f^2}{2} - \frac{3}{2} D_h f D_v f - D_h f. \quad (3.54)$$

It follows that for small wave numbers, the gradient of the group velocity is negative since

$$\hat{c}'_x(\theta) = -6\lambda_3\theta < 0, \quad (3.55)$$

and hence $\theta = 0$ is a maximum. Thus there is a transition between convective and absolute instability when the maximum of the absolute group velocity at $\theta = 0$ changes sign,

$$c_x(0) = \hat{c}_x(0) = -\lambda_1 - \frac{V(h_*)}{h_*} = 0. \quad (3.56)$$

In fact, this condition holds throughout the region of instability since $\theta_d > 0$. A further transition occurs when $\theta_d = \theta_g > 0$. We can approximate this by $\hat{\theta}_d = \hat{\theta}_g$, where $\hat{\theta}_g > 0$ is a solution to $\hat{c}_x(\hat{\theta}_g) = 0$,

$$\theta_g \approx +\sqrt{-\frac{(h_*\lambda_1 + V(h_*))}{3h_*\lambda_3}}. \quad (3.57)$$

Thus there is also a transition between convective and absolute instability when the model parameters satisfy

$$\frac{\lambda_2}{\lambda_4} = \frac{(h_*\lambda_1 + V(h_*))}{3h_*\lambda_3}. \quad (3.58)$$

In Figure 3.6, we illustrate a phase diagram for the OVRV model in (α, h_*) space depicting regions of stability (white area marked S), convective instability (light grey areas marked C) and absolute instability (dark grey area marked A) for the case $\beta = 0.2$ using the standard OV function (1.6). The stability boundary, $\lambda_2 > 0$, is given by $\alpha = 2[V'(h_*) - \beta]$ in parameter space. The absolute/convective transition boundary given

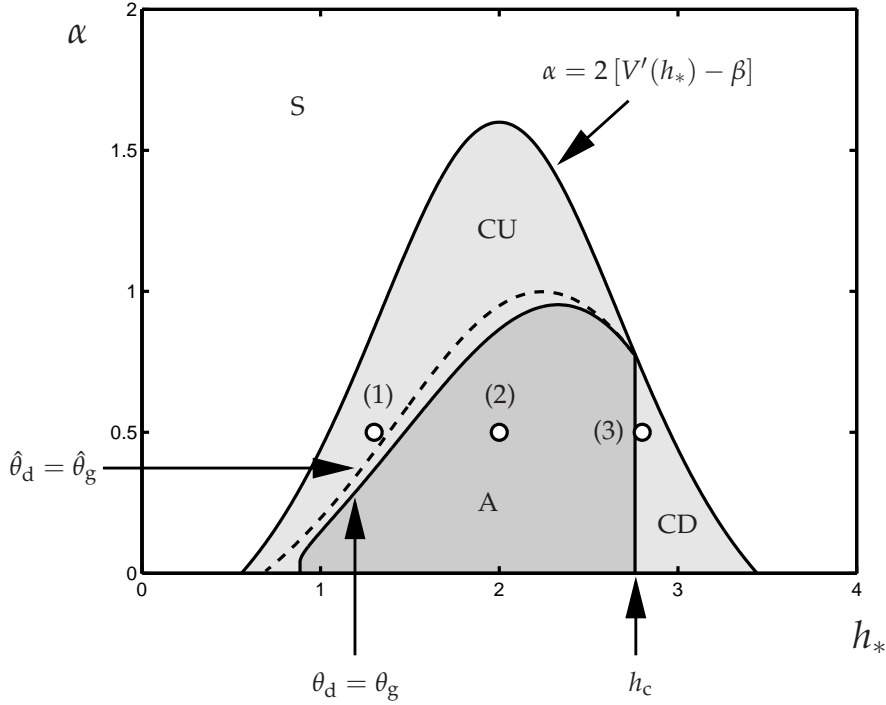


Figure 3.6. Instability phase diagram for the OVRV model in (α, h_*) space with $\beta = 0.2$ using the standard OV function (1.6). Regions of stability (labelled S, white), convective instability (labelled CU travelling upstream and CD travelling downstream, both in light grey) and absolute instability (labelled A, dark grey) are illustrated. The transition boundary at $h_* = h_c$ is found by solving (3.56). The dashed line corresponds to the transition boundary predicted by the small wave number expansion (3.58). The parameter values for Cases (1–3) ($h_* = 1.3, 2.0, 2.8$ and $\alpha = 0.5$ respectively) are marked with circles and labelled (1–3) respectively; these classifications are in agreement with Figures 3.5 and 3.4.

by (3.56), occurs at h_c where $V'(h_c) = V(h_c)/h_c$. We calculate the absolute/convective transition boundary numerically for both the full group velocity/dispersion relation comparison $\theta_d = \theta_g$ (solid line) and the small wave-number approximation $\hat{\theta}_d = \hat{\theta}_g$ (dashed line) using a continuation procedure. We observe that the small wave-number approximation over estimates the size of the region of absolute instability, however the agreement is surprisingly good. It is possible that this may be attributed to a cancellation of errors when we equate $\hat{\theta}_d$ and $\hat{\theta}_g$.

3.5 Inductive Asymptotic Methods

We now present an alternative approach to analyse the growth of a perturbations in the long straight road set up described in Sections 3.2 and 3.4. The basis of this technique is to solve the linearised equation of motion (3.3) progressively along the chain

of vehicles. This means that at each stage the non-local terms are known, and subsequent solutions can be obtained by solving a resonant forced oscillator equation. However, this proves to be difficult both analytically and numerically for large n and t , but the advantage of using this method is that we gain insight into the asymptotic nature of solutions. Furthermore, we build on this technique in the following section, where Laplace transforms are used to illustrate how a formula for the motion of the n th vehicle may be obtained.

In our long straight road set up, the velocity of the zeroth vehicle (i.e. $n = 0$) is held fixed for all time. We perturb the velocity of the first vehicle ($n = 1$, etc.) from equilibrium by a small amount $v_1(0) = V(h_*) + \delta$. The linearised equation of motion (3.3) with $n = 1$ is simplified by the fact that the velocity of the lead vehicle is held fixed, hence all non-local terms are zero,

$$\ddot{\tilde{h}}_1 + (D_h f - D_v f)\dot{\tilde{h}}_1 + (D_h f)\tilde{h}_1 = 0. \quad (3.59)$$

The sign constraints (1.17) ensure that solutions to the characteristic equation,

$$\lambda_{\pm} = \frac{1}{2} \left[-(D_h f - D_v f) \pm \sqrt{(D_h f - D_v f)^2 - 4(D_h f)} \right], \quad (3.60)$$

always have negative real part, and subsequently vehicle $n = 1$ returns to equilibrium. Note that the growth rates λ_{\pm} differ from the dispersion relation introduced in Section 3.2. In the following analysis we assume that λ_{\pm} are not equal. Solving (3.59) explicitly with the boundary conditions $\tilde{h}_1(0) = 0$ and $\dot{\tilde{h}}_1(0) = -\delta$ yields

$$\tilde{h}_1 = a_0^{(1)} e^{\lambda_+ t} + b_0^{(1)} e^{\lambda_- t}, \quad (3.61)$$

where

$$a_0^{(1)} = -\frac{\delta}{\lambda_+ - \lambda_-}, \quad b_0^{(1)} = -a_0^{(1)}, \quad (3.62)$$

and λ_{\pm} are given by (3.60).

Having determined the behaviour of the first vehicle, we move onto the second vehicle in the chain, vehicle $n = 2$. Substituting (3.61) into the linearised equation of motion (3.3) with $n = 2$ yields

$$\ddot{\tilde{h}}_2 + (D_h f - D_v f)\dot{\tilde{h}}_2 + (D_h f)\tilde{h}_2 = (D_h f\lambda_+ + D_h f)a_0^{(1)} e^{\lambda_+ t} + (D_h f\lambda_- + D_h f)b_0^{(1)} e^{\lambda_- t} \quad (3.63)$$

This is a resonant forced oscillator equation that can be solved using the method of undetermined coefficients. The forcing frequencies are the same as the eigen-frequencies, hence for each mode λ_{\pm} , the solution is the product of a first order polynomial in t

(whose coefficients must be calculated) and the exponential $e^{\lambda_{\pm}t}$. The coefficients of the polynomial can be determined by matching like powers of t . In fact, after completing this process for the second vehicle, we can continue in this manner and iteratively determine the solutions of the third, fourth, fifth, etc. vehicles. At each step the particular solution adds a term that is one order higher in t than the last solution. The general solution for the n th vehicle is therefore

$$\tilde{h}_n = p_n(t)e^{\lambda_+t} + q_n(t)e^{\lambda_-t}, \quad (3.64)$$

where $p_n(t)$ and $q_n(t)$ are polynomials of order $n - 1$. Note that $\text{Re}(\lambda_{\pm}) < 0$ ensures that the behaviour at large t is dominated by exponential decay. Hence all vehicles eventually return to equilibrium.

The goal now is to solve for the coefficients of the polynomials $p_n(t)$ and $q_n(t)$ which, due to linearity, are independent; hence we focus on solving for the coefficients of $p_n(t)$ since we can determine $q_n(t)$ using the same procedure. Substituting $p_n(t)e^{\lambda_+t}$ into (3.3) and cancelling the resonant terms we find

$$\ddot{p}_n + (2\lambda_+ + D_h f - D_v f) \dot{p}_n = (D_h f) \dot{p}_{n-1} + (\lambda_+ D_h f + D_h f) p_{n-1}. \quad (3.65)$$

We wish to expand the polynomial $p_n(t)$ and match like powers of t in order to determine the unknown coefficients. Hence, to associate the polynomial coefficients with particular vehicles we use the notation

$$p_n(t) = \sum_{m=0}^{n-1} a_m^{(n)} t^m, \quad (3.66)$$

where the subscript m refers to the m th order term, and the superscript (n) denotes the polynomial coefficients of the n th vehicle. Subsequent analysis does not contain powers of coefficients so our superscript notation only refers to vehicle allocation. To equate like powers of t in (3.65) using (3.66), we require the derivatives of the $p_n(t)$ terms,

$$\dot{p}_n(t) = \sum_{m=1}^{n-1} m a_m^{(n)} t^{m-1} = \sum_{m=0}^{n-2} (m+1) a_{m+1}^{(n)} t^m, \quad (3.67)$$

$$\ddot{p}_n(t) = \sum_{m=2}^{n-1} m(m-1) a_m^{(n)} t^{m-2} = \sum_{m=0}^{n-3} (m+2)(m+1) a_{m+2}^{(n)} t^m, \quad (3.68)$$

and similarly the derivatives of the $p_{n-1}(t)$ terms,

$$p_{n-1}(t) = \sum_{m=0}^{n-2} a_m^{(n-1)} t^m, \quad (3.69)$$

$$\dot{p}_{n-1}(t) = \sum_{m=1}^{n-2} m a_m^{(n-1)} t^{m-1} = \sum_{m=0}^{n-3} (m+1) a_{m+1}^{(n-1)} t^m. \quad (3.70)$$

Furthermore, we define

$$\mu_n = \lambda_+ D_h f + D_h f, \quad (3.71)$$

$$\mu_d = 2\lambda_+ + D_h f - D_v f, \quad (3.72)$$

and substituting (3.67–3.72) into (3.65) we find

$$\begin{aligned} & \sum_{m=0}^{n-3} (m+2)(m+1) a_{m+2}^{(n)} t^m + \mu_d \sum_{m=0}^{n-2} (m+1) a_{m+1}^{(n)} t^m \\ &= (D_h f) \sum_{m=0}^{n-3} (m+1) a_{m+1}^{(n-1)} t^m + \mu_n \sum_{m=0}^{n-2} a_m^{(n-1)} t^m. \end{aligned} \quad (3.73)$$

From this we can calculate the term $a_{m+1}^{(n)}$, providing we know the term one order higher $a_{m+2}^{(n)}$ for vehicle n , as well as the terms $a_{m+1}^{(n-1)}$ and $a_m^{(n-1)}$ of the preceding vehicle $n-1$. Thus, one must work progressively through the chain of vehicles, from $n=1$ upwards, and calculate the coefficients from *highest to lowest* order. The zeroth order coefficients must also be calculated using the initial conditions,

$$a_0^{(n)} = -\frac{a_1^{(n)} + b_1^{(n)} + \tilde{h}_n(0)\lambda_- - \dot{\tilde{h}}_n(0)}{\lambda_+ - \lambda_-}, \quad (3.74)$$

$$b_0^{(n)} = \frac{a_1^{(n)} + b_1^{(n)} + \tilde{h}_n(0)\lambda_+ - \dot{\tilde{h}}_n(0)}{\lambda_+ - \lambda_-}. \quad (3.75)$$

In summary, to calculate the coefficients of a given vehicle $n \geq 2$ (and hence the solution $\tilde{h}_n(t)$), one must first calculate *all* the coefficients of *every* preceding vehicle.

In Figure 3.7, we have used (3.73) to calculate the polynomial coefficients and solutions trajectories of three vehicles $n = 15, 45$ and 90 , which we compare with simulations of the OVRV model. From experiments with small initial disturbances and weak instability (i.e. close to marginal stability), we observe extremely good agreement. In fact, we have chosen highly unstable parameter values in Figure 3.7 ($\alpha = 1.0$ and $\beta = 0.2$), with a large initial disturbance ($V(h_*) = 0.96$ and $\delta = 0.48$), to illustrate that even with such extreme parameter values, there is still very good agreement between the solutions of the linearised equations (plotted in green) and the microscopic simu-

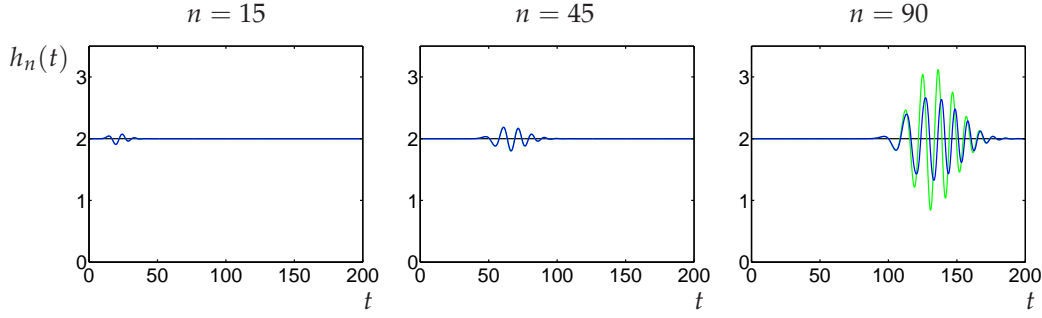


Figure 3.7. Temporal evolution of the headways h_n of three vehicles ($n = 15, 45$ and 90) using the OVRV model with the standard OV function (1.6). Microscopic simulations (blue) are compared with solutions of the linearised equations of motion (green), calculated using (3.73). Parameters are chosen such that uniform flows are unstable ($\alpha = 1.0$ and $\beta = 0.2$), the equilibrium spacing is $h_* = 2.0$ and the initial perturbation is $\delta = 0.48$. Hence nonlinear instability develops rapidly, yet there is good agreement between the linearisation and the microscopic simulation up to $n = 45$, after which the growth of the disturbance is slowed by nonlinearities in the microscopic simulation.

lations (plotted in blue), in this case up to vehicle $n = 45$. For vehicle $n = 90$ however, non-linearities have slowed the growth of the perturbation in the microscopic simulation and hence there is a discrepancy between this and the solutions to the linearised equation.

The number of calculations required to obtain the n th vehicle's coefficients is order $(n - 1)^2$. Furthermore, calculations become increasingly difficult as n increases due to repeated multiplication. Hence we would like to determine whether (3.73) can be approximated asymptotically. Equating the coefficients of the *largest* powers of t , i.e. the $m = n - 2$ terms, from (3.73) yields

$$a_{n-1}^{(n)} = \frac{1}{(n-1)} \mu a_{n-2}^{(n-1)}, \quad (3.76)$$

where $\mu = \mu_n / \mu_d$. Successive applications of (3.76) leads inductively to

$$a_{n-1}^{(n)} = \frac{1}{(n-1)!} \mu^{n-1} a_0^{(1)}, \quad (3.77)$$

hence the leading order term of *every* vehicle depends only on the initial conditions of the first vehicle $n = 1$. Thus for large t , we may consider approximating the solutions

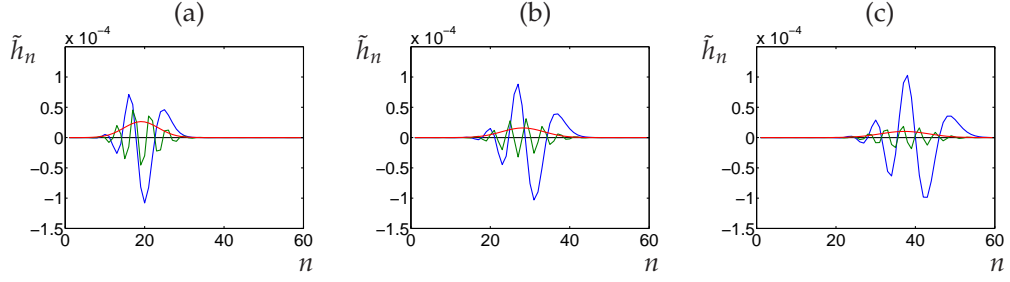


Figure 3.8. Solutions $\tilde{h}_n(\tau)$ (blue) plotted in n for three times, $\tau = 25, 37.5, 50$ (panels (a–c) respectively for the OVRV model with $\alpha = 1.5, \beta = 0.1$ and the standard OV function (1.6). The leading order term (3.78) (green) and Gaussian envelope of (3.79) (red) are also plotted. The leading order term proves to be a poor approximation, and actually gets worse as t increases. This appears to be due to dispersion in the leading order term that must be counter-acted by lower order terms.

by

$$\tilde{h}_n(t) \approx a_0^{(1)} \exp(\lambda_{+I}t) \frac{(\mu t)^{n-1}}{(n-1)!} \exp\{-|\lambda_{+R}|t\}, \quad (3.78)$$

$$\approx a_0^{(1)} \exp(\lambda_{+I}t) \left(\frac{\mu}{|\lambda_{+R}|} \right)^{n-1} \frac{1}{\sqrt{2\pi|\lambda_{+R}|t}} \exp\left\{-\frac{[(n-1) - |\lambda_{+R}|t]^2}{2|\lambda_{+R}|t}\right\}, \quad (3.79)$$

using the normal distribution approximation to the Poisson distribution. We gain insight into the asymptotic nature of solutions from this approximation since at large values of t , solutions in n are simply modulated Gaussians. However, we illustrate in Figure 3.8 that the leading order term is *not* a good approximation and seems to diffuse, rather than grow.

To understand why this is the case, we begin by equating the t^{n-3} terms,

$$(n-1)(n-2)a_{n-1}^{(n)} + \mu_d(n-2)a_{n-2}^{(n)} = D_h f(n-2)a_{n-2}^{(n-1)} + \mu_n a_{n-3}^{(n-1)}. \quad (3.80)$$

Thus using $\mu = \mu_n/\mu_d$, we can write

$$a_{n-1}^{(n)} = \frac{1}{(n-1)!} \mu^{n-1} a_0^{(1)} \quad \text{and} \quad (3.81)$$

$$a_{n-2}^{(n-1)} = \frac{1}{(n-2)!} \mu^{n-2} a_0^{(1)}. \quad (3.82)$$

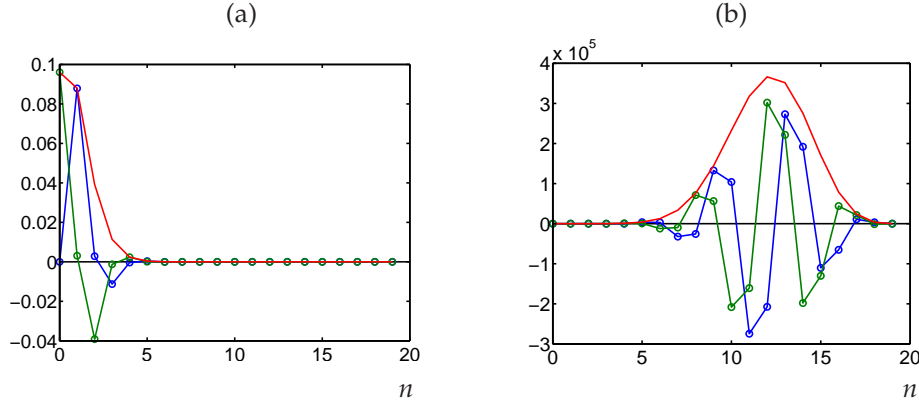


Figure 3.9. Coefficients of vehicle $n = 20$ $a_m^{(20)}$ are plotted in panel (a) and in panel (b) these are multiplied by t^n for $t = 25$. In both panels, the real part is plotted in blue, the imaginary part in green and the absolute value in red. The leading order term ($m = 19$) is actually the *smallest* term in panel (a). Furthermore, the main contribution depends on the time t , illustrated in panel (b), where at $t = 25$, the dominant contributions come from the terms $m = 10$ and $m = 15$. Note however the large magnitude of the scale in panel (b); hence solutions trajectories must result from a complicated cancellation of terms.

Substituting (3.81) and (3.82) into (3.80), we find

$$a_{n-2}^{(n)} = \frac{1}{(n-2)} \mu a_{n-3}^{(n-1)} + \frac{1}{(n-2)!} \left[\frac{D_h f - \mu}{\mu_d} \right] \mu^{n-2} a_0^{(1)} \quad (3.83)$$

$$= \frac{1}{(n-2)} \mu \left\{ \frac{1}{(n-3)} \mu a_{n-4}^{(n-2)} + \frac{1}{(n-3)!} \left[\frac{D_h f - \mu}{\mu_d} \right] \mu^{n-3} a_0^{(1)} \right\} \quad (3.84)$$

$$+ \frac{1}{(n-2)!} \left[\frac{D_h f - \mu}{\mu_d} \right] \mu^{n-2} a_0^{(1)} \quad (3.85)$$

$$= \frac{1}{(n-2)!} \left\{ a_0^{(2)} + (n-2) \left[\frac{D_h f - \mu}{\mu_d} \right] a_0^{(1)} \right\} \mu^{n-2}. \quad (3.86)$$

Hence we cannot say by a simple comparison with (3.77) that either coefficient is necessarily larger, thus the leading order term may not be the dominant contribution. In fact in Figure 3.9 we plot the absolute value of the coefficients of vehicle $n = 20$ (which is representative of the coefficients of other vehicles). We observe that higher order coefficients are actually *smaller* in magnitude than lower order terms. Furthermore, we also plot the product $a_m^{(20)} t^{20-1}$ for $t = 25$. We choose this value because we observe that vehicle $n = 20$ is at the centre of the pulse at this time in Figure 3.8(a). This illustrates that the largest contribution comes from intermediate values of m and moreover, we have observed that these change with different values of time t . However, note that the scale is much larger than the magnitude of the pulse. Hence the complete solution emerges from a complicated cancellation of terms. Thus determining the asymptotic nature of the pulse remains an open (and difficult) problem.

3.6 Laplace Transform Development

In the previous section we gained insight into the asymptotic nature of solutions to the linearised equation of motion (3.3) by solving the differential equations progressively along the chain of vehicles. In this section, we develop this analysis and describe how Laplace transforms can be used to obtain a formula for the trajectory of the n th vehicle. This represents on-going work, and we hope that in the future this method will yield an asymptotic relation for the solution trajectory of the n th vehicle.

Defining the Laplace Transform of the perturbation,

$$\mathcal{L} \{ \tilde{h}_n(t) \} := F_n, \quad (3.87)$$

and transforming the linearised equation of motion (3.3) of the first vehicle $n = 1$ we find

$$(z^2 + (D_h f - D_v f)z + D_h f) F_1 = z\tilde{h}_1(0) + \dot{\tilde{h}}_1(0) + \tilde{h}_1(0)(D_h f - D_v f). \quad (3.88)$$

The initial conditions of the first vehicle are $\tilde{h}_1(0) = 0$ and $\dot{\tilde{h}}_1(0) = -\delta$, hence solving for F_1 yields

$$F_1 = -\frac{\delta}{(z - \lambda_+)(z - \lambda_-)}. \quad (3.89)$$

Note that by considering (3.60), λ_{\pm} are complex conjugates when $(D_h f)^2 < 2D_h f$. When this is true, the following analysis simplifies to some extent; however we describe the general case here. The initial conditions of the second vehicle are $\tilde{h}_2(0) = 0$ and $\dot{\tilde{h}}_2(0) = +\delta$, and transforming the equation of motion (3.3) with $n = 2$, we find

$$(z^2 + (D_h f - D_v f)z + D_h f) F_2 = D_h f z F_1 + D_h f F_1 + \delta, \quad (3.90)$$

then solving for F_2 yields

$$F_2 = \frac{(D_h f z + D_h f) F_1}{(z - \lambda_+)(z - \lambda_-)} + \frac{\delta}{(z - \lambda_+)(z - \lambda_-)} \quad (3.91)$$

$$= -\frac{\delta (D_h f z + D_h f)}{(z - \lambda_+)^2 (z - \lambda_-)^2} + \frac{\delta}{(z - \lambda_+)(z - \lambda_-)}. \quad (3.92)$$

The initial conditions of vehicles $n \geq 3$ are $\tilde{h}_n(0) = 0$ and $\dot{\tilde{h}}_n(0) = 0$. The transform of the third vehicle is

$$(z^2 + (D_h f - D_v f)z + D_h f) F_3 = D_h f z F_2 + D_h f F_2, \quad (3.93)$$

solving for F_3 yields

$$F_3 = \frac{(D_h f z + D_h f) F_2}{(z - \lambda_+)(z - \lambda_-)} \quad (3.94)$$

$$= -\frac{\delta (D_h f z + D_h f)^2}{(z - \lambda_+)^3 (z - \lambda_-)^3} + \frac{\delta (D_h f z + D_h f)}{(z - \lambda_+)^2 (z - \lambda_-)^2} \quad (3.95)$$

Thus for vehicles $n \geq 3$, we find

$$\begin{aligned} F_n = & -\frac{\delta}{(z - \lambda_+)(z - \lambda_-)} \left[\frac{D_h f z + D_h f}{(z - \lambda_+)(z - \lambda_-)} \right]^{n-1} \\ & + \frac{\delta}{(z - \lambda_+)(z - \lambda_-)} \left[\frac{D_h f z + D_h f}{(z - \lambda_+)(z - \lambda_-)} \right]^{n-2} \end{aligned} \quad (3.96)$$

To obtain the solution trajectory of a given vehicle n we must apply the inverse transform to (3.96). We describe the salient features of this calculation in the following analysis. We can use partial fractions to write terms on the right-hand-side of (3.96) in the form

$$F := \frac{1}{(z - a)(z - b)} \left[\frac{A}{z - a} + \frac{B}{z - b} \right]^n = \sum_{k=0}^n \binom{n}{k} \frac{A^k B^{n-k}}{(z - a)^{k+1} (z - b)^{n-k+1}}. \quad (3.97)$$

Hence the inverse transform is

$$\begin{aligned} \mathcal{L}^{-1}\{F\} = & \text{Res} \left\{ \sum_{k=0}^n \binom{n}{k} A^k B^{n-k} (z - a)^{-(k+1)} (z - b)^{-(n-k+1)} e^{zt}; a \right\} \\ & + \text{Res} \left\{ \sum_{k=0}^n \binom{n}{k} A^{n-k} B^k (z - b)^{-(k+1)} (z - a)^{-(n-k+1)} e^{zt}; b \right\}, \end{aligned} \quad (3.98)$$

where $\text{Res}\{f(z); z_0\}$ denotes the residue of $f(z)$ at z_0 . Note that we have used the symmetry of the binomial coefficients to run the sum in the second term of the right-hand side of (3.98) backwards. The purpose of this is to introduce common factors in the residue at a and b .

Thus we are interested in solutions to

$$\text{Res} \left\{ (z - a)^{-p} (z - b)^{-q} e^{zt}; a \right\} = \frac{1}{(p-1)!} g^{(p-1)}(a), \quad (3.99)$$

where

$$g(z) = g_1(z)g_2(z) = e^{zt}(z - b)^{-q}. \quad (3.100)$$

Recalling Leibniz's product rule,

$$g^{(p-1)}(z) = \sum_{l=0}^{p-1} \binom{p-1}{l} g_1^{(l)}(z) g_2^{(p-1-l)}(z), \quad (3.101)$$

we must therefore determine the n th derivatives of $g_1(z)$ and $g_2(z)$. In the first case, this is simply

$$g_1^{(l)}(z) = t^l e^{zt}, \quad (3.102)$$

and in the second case we find

$$g_2^{(p-1-l)}(z) = (-1)^{p-1-l} \frac{(p-1-l+q-1)!}{(q-1)!} (z-b)^{-(p-1-l+q)}. \quad (3.103)$$

Combining (3.99–3.103) yields

$$\begin{aligned} \text{Res} \{ (z-a)^{-p} (z-b)^{-q} e^{zt}; a \} = \\ \frac{e^{at}}{(p-1)!} \sum_{l=0}^{p-1} \binom{p-1}{l} (-1)^{p-1-l} \frac{(p-l+q-2)!}{(q-1)!} (a-b)^{-(p-1-l+q)} t^l. \end{aligned} \quad (3.104)$$

It follows that the inverse transform (3.98) is

$$\begin{aligned} \mathcal{L}^{-1} \{F\} = & \sum_{k=0}^n \binom{n}{k} A^k B^{n-k} \frac{1}{k!} \sum_{l=0}^{k-1} \binom{k}{l} (-1)^{k-l} \frac{(n-l)!}{(n-k)!} (a-b)^{-(n-l+1)} t^l e^{at} \\ & + \sum_{k=1}^n \binom{n}{k} A^{n-k} B^k \frac{1}{k!} \sum_{l=0}^{k-1} \binom{k}{l} (-1)^{k-n-1} \frac{(n-l)!}{(n-k)!} (a-b)^{-(n-l+1)} t^l e^{bt}, \end{aligned} \quad (3.105)$$

which, by switching the order of summation and collecting binomial terms, we rewrite as

$$\begin{aligned} \mathcal{L}^{-1} \{F\} = & \sum_{l=0}^n (a-b)^{-(n-l+1)} \frac{t^l}{l!} \\ & \times \sum_{k=l}^n \binom{n}{k} \binom{n-l}{n-k} \left[(-1)^{k-l} A^k B^{n-k} e^{at} + (-1)^{k-n-1} A^{n-k} B^k e^{bt} \right]. \end{aligned} \quad (3.106)$$

This equation consists of two order $n-1$ polynomials in t multiplied by exponential in t , hence it is exactly the kind of solution that we were seeking in Section 3.5, however we have an explicit relation from which we can calculate the coefficients of the polynomial. Furthermore, we observe that the lower order terms sum the binomial coefficients of all *higher* order terms. By comparing the coefficients calculated numerically using the inductive method in Section 3.5 and the Laplace method in this section,

we find that both methods produce the same values. The future development of this area will be to consider the asymptotic nature of this sum and relate this to the stability and convective growth described in Sections 3.2 and 3.4.

3.7 Summary

In this chapter we have described new developments relating to the *linear* stability analysis of a general class of car-following models. In Section 3.2 we reviewed the method proposed by Wilson [83], and then extended this to include heterogeneous driver behaviour in Section 3.3. In Section 3.4, we considered the convective growth of small perturbations on a long straight road and, by comparing growth rates and group velocities, derived criteria that distinguish between so called absolute and convective instability. Finally, in Sections 3.5 and 3.6 we described a novel method in which the solutions of the linearised equations of motion were calculated progressively along a chain of vehicles. This method helped to develop deeper insight into the asymptotic nature of solutions which may be improved further by analysing the explicit solutions that result from our Laplace transform method.

In parallel with the work presented in this chapter are two further areas that link microscopic and macroscopic traffic models, and hence compliment our multiscale theme. Lattice equations can be homogenised via the (Taylor expansion) identity

$$h_{n-1}(t) = \exp(-\partial_n)h(n, t) = \exp(h_*\partial_z)\varepsilon(z, t), \quad (3.107)$$

(see [49]) where $z = -h_*n$. Applying this technique to the general car-following model (3.3) results in a (complicated) macroscopic PDE. One approach is to consider more general time varying perturbations as source terms and use Fourier analysis techniques to analyse the solutions. In fact this procedure runs parallel to the inductive methods described in Sections 3.5 and 3.6. The result however is an integral equation, which may prove easier to analyse asymptotically than the binomial sum derived in Section 3.6 (see Appendix B for a further discussion). The second approach is to simplify the resulting PDE using the approximation of Padé [49],

$$\exp(h_*^2\partial_n) - 1 \approx \frac{1}{2} \frac{h_*^2\partial_z^2}{1 - \frac{h_*^2}{12}\partial_z^2} + \frac{h_*\partial_z}{\left(1 - \frac{h_*^2}{12}\partial_z^2\right)^2}. \quad (3.108)$$

This method results in a linear macroscopic model, although non-standard terms with higher order time-derivatives and mixed space/time derivatives do appear. However, an interesting feature of the resulting equation is that there is no diffusion term (i.e a

term proportional to ρ_{xxx}) in the linearisation. There is an active debate in the traffic literature about the form of macroscopic models (see [5, 16, 29, 83]) and this has identified problems with some of the continuum models that have been derived from microscopic car-following rules [8, 56]. Hence it would be interesting to study the approach described above to determine whether any of the previous problems that result from homogenisation can be amended.

Linear stability analysis is commonplace in the traffic literature, yet surprisingly we have still been able to identify open problems, whose solutions are beneficial to the wider traffic community. The ultimate goal of the work described here is to develop a deeper awareness of the fundamental features present in faithful traffic models, and hence bench-mark future model developments. It has emerged from this work that a review and classification of the large number of microscopic traffic models is needed, with the aim of uniting the various elements that faithfully reproduce phenomena observed in empirical traffic data.

Chapter 4

Macro Analysis of Micro Lane-Changing

“A digital recording? (11)”, Rufus 77(9d) [31]

4.1 Motivation

It is becoming increasingly clear from data that lane-changes are a critical factor in the development of large-amplitude stop-and-go waves [4, 73]. However, there has been little progress in the analysis of multilane traffic models. This is largely due to 1. The complexity of vehicle interactions and 2. The lack of ‘high fidelity’ models calibrated against data. This second point results from the fact that time-averaged empirical data collected from induction loops is typically too coarse to quantify the cause and effect of individual lane changes [73, 78]. This situation is set to change with the recent emergence of high resolution Individual Vehicle Data (IVD) [59]. Re-identification algorithms [14, 59] have been developed that can track the progress of large numbers of vehicles over long distances. Such data sets will allow lane-changing models to be refined through quantitative testing against empirical data, and thus transform the development of multilane models.

Empirical traffic data is typically recorded on multilane highways where lane-changes occur frequently. In contrast, the study of spatio-temporal phenomena has focused on the analysis of single-lane traffic models [33], where overtaking is not considered. Those theoretical studies that have examined multilane highways have primarily used ad hoc models of macroscopic lane changing rates. One of the earliest approaches was by Gazis [25] in the 1960s, who analysed lane-changing in a simplified spatially independent setting using a linear model based on the density difference between lanes with time delay. This was extended by Munjal and Pipes in the 1970s [62] and later Holland and Woods [37] to include spatial effects. Further development of macroscopic models has been via gas kinetic models [39, 50] and semi-discretisations, or ‘hybrid’ models [54].

Numerical simulations with microscopic multilane models suggest that lane-changing gives rise to a rich variety of dynamics [19, 40, 51]. However, we believe that there is currently a lack of basic theory that describes how these complex spatio-temporal patterns emerge from lane-changes. It is our goal in this chapter to present a first step towards developing such a theory by considering a highly simplified setting for a general class of car-following models. We focus on the analysis of spatially independent *macro* dynamics present in *micro* simulations. Hence the novelty of this work is that the microscopic multilane driving rules are linked directly to the subsequent macroscopic (non-linear) lane-changing rates.

Lane-changing models typically break the decision making process into several distinct steps [27, 48]. At a strategic level, drivers may have to perform mandatory lane-changes due to the road infrastructure, such as at motorway merges or on/off ramps [77]. There is also a tactical level, where drivers may manoeuvre themselves to take advantage of a particular situation or even cooperate with other vehicles [35]. Finally, vehicles must decide if a lane-change is both desirable and safe [48]. Lane-changing models typically focus on this later stage and often fall into the class of ‘gap-acceptance’ rules [19, 28, 35, 40, 51, 76]. As suggested, these models consider the distance to the car behind in the adjacent lane, referred to as the gap (as well as neighbouring vehicle velocities) to determine whether changing lanes is both advantageous and safe.

The work described in this chapter has resulted in two principal advances. We have developed a flexible simulation framework, which has been used to test a number of lane-changing and longitudinal (within-lane) acceleration rules. This prompted a further analysis of the so called *Minimising Overall Breaking In Lane-changing* (MOBIL) framework developed by Kesting *et al.* [48]. In this framework, lane-changing is modelled via criteria that measure each driver’s incentive to change lanes and the safety of the manoeuvre. Longitudinal acceleration and lane-changing are separate driving elements in the MOBIL framework, hence one is free to choose any driver behaviour model. The lane-changing criteria however are based on longitudinal accelerations, thus the characteristics of the longitudinal model are directly incorporated into the decision making process. The criteria compare current accelerations with putative accelerations if the lane-change were to take place for both the vehicle considering the manoeuvre and the trailing vehicles upstream, hence incorporating both the incentive of the driver wishing to change lanes and the safety of its followers. If these differences exceed a given threshold then the criteria are satisfied.

We analyse these rules based on the general car-following model (1.12) in a highly simplified setting in which parameters are stable, vehicles identical and lane changes are relatively rare. Specifically, we assume that the longitudinal relaxation timescale

of traffic is shorter than the lane-changing timescale and this subsequently allows us to simplify and analyse the MOBIL criteria. Furthermore, we only consider the case where the lane-changing criteria are the same in both lanes. This is similar to driving rules used in North America where vehicles can pass in any lane. In the UK and Europe however, this is not the case and vehicles must move closer to the central reservation to overtake. The MOBIL criteria can replicate this feature if the lane-changing thresholds are biased towards particular lanes, however for simplicity we choose not to include this feature. In this simple set-up, we can identify links between complex microscopic interactions and simple macroscopic lane-changing rates. We stress that our goal is to initiate the development of a multiscale theory of lane-changing, rather than model any particular situation in an accurate way. To achieve this we are forced to make a number of simplifying assumptions. However, we contrast this reductionist approach with a selection of experiments illustrating spatio-temporal patterns that emerge when both instability and lane-changing are present, which cannot be accounted for via our simplifications.

This chapter is organised as follows. In Section 4.2 we present the MOBIL lane-changing criteria and our model set-up. We have developed a flexible simulation framework and experimented with a range of lane-changing and acceleration models, which we describe in Section 4.3. In Section 4.4 we explain our ‘quasi-equilibrium’ assumption and the simplifications that it introduces. We follow this with an analysis of the lane-changing criteria and determine necessary conditions for safe lane changes in Section 4.5, which we then illustrate with examples in Section 4.6. In Section 4.7 we develop an approximation for the macroscopic lane-changing rate based on the microscopic model. We then compare microscopic simulation data with results based on the quasi-equilibrium assumption in Section 4.8. We contrast this in Section 4.9 with results from three experiments where instability and driver heterogeneity force spatial pattern that are not accounted for in our simplified theory. Finally, we conclude with a discussion of our approach in Section 4.10.

4.2 Microscopic Multi-Lane Model

We consider a large number N of identical vehicles in discrete lanes numbered $l \in \{0, 1\}$ travelling in continuous space x and time t on a ring road of length L . Each vehicle n has a position $x_n(t)$ and velocity $v_n(t)$ and travels in the increasing x direction. Note that in contrast to the traditional set-up for single-lane models, the subscripts $n = 1, 2, \dots, N$ do not correspond to the order of vehicles because overtaking is allowed.

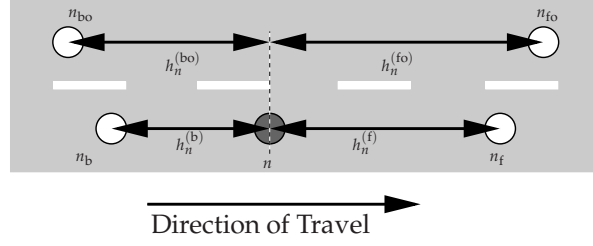


Figure 4.1. Neighbours of vehicle n (grey circle) denoted n_f , n_b , n_{fo} and n_{bo} . Headways $h_n^{(f)}$, $h_n^{(b)}$, $h_n^{(fo)}$ and $h_n^{(bo)}$ are measured longitudinally and take positive values by convention.

In order to prescribe rules for longitudinal motion and for lane-changing, we introduce subscript notation to reference the neighbours of vehicle n , see Fig. 4.1. We let n_f and n_b denote respectively the vehicles which are immediately in front and behind in the same lane, and we let n_{fo} and n_{bo} denote the vehicles which are immediately in front and behind in the adjacent (i.e., other) lane. In particular, the front-to-front distances to neighbouring vehicles are measured longitudinally and are given by

$$h_n = h_n^{(f)} = x_{n_f} - x_n, \quad (4.1)$$

$$h_n^{(b)} = x_n - x_{n_b}, \quad (4.2)$$

$$h_n^{(fo)} = x_{n_{fo}} - x_n \quad \text{and} \quad (4.3)$$

$$h_n^{(bo)} = x_n - x_{n_{bo}}. \quad (4.4)$$

Observe carefully the role that the bracketed superscript plays: for example $h_n^{(f)} = h_{n_f}^{(b)}$ or $h_n^{(bo)} = h_{n_{bo}}^{(f)} - h_n^{(fo)}$.

The model definition then comes in two parts. Firstly, we suppose that the longitudinal motion of each vehicle is modelled by the general form (1.12) with the additional constraints on the partial derivatives (1.17). We also note that the equilibrium velocity function V associated with (1.12) is a monotonically increasing function (1.15). Secondly, we prescribe the rules for lane-changing which will occasionally shuffle the order of the traffic streams to which the longitudinal laws apply. Complex lane-changing models from the engineering literature are often inextricably linked to the particular longitudinal model in force (see [27]). However, our approach is to use the MOBIL framework [48] in which general longitudinal laws are incorporated in the simplest possible way, namely via a comparison of current vehicle accelerations with putative accelerations that the longitudinal law would provide if a lane change were to take place. Thus to change lane, vehicle n tests for positivity of an *incentive function*

$$I_n := [\tilde{a}_n - a_n] + p [\tilde{a}_{n_b} - a_{n_b}] + p [\tilde{a}_{n_{bo}} - a_{n_{bo}}] - \sigma, \quad (4.5)$$

where a_n is its present acceleration and \tilde{a}_n is the acceleration that would result if it were to change lane. Here $\sigma \geq 0$ is a threshold parameter and the weight $0 \leq p \leq 1$ measures altruism. Thus a vehicle is inclined to change lane in return for an increase in its acceleration value (and thus for subsequently higher speed), but if $p > 0$, then the adverse effect of the lane-change on the drivers immediately behind is also considered.

With $p = 1$, a lane-change which severely disadvantages the ‘over-the-shoulder’ vehicle n_{bo} is unlikely to occur. However, the MOBIL framework includes a separate gap-acceptance criterion where vehicle n also tests for positivity of a *safety function*

$$S_n := \tilde{a}_{n_{bo}} + \tau. \quad (4.6)$$

Here $\tau > 0$ is a threshold parameter. In this thesis we are mostly concerned with the simple ‘selfish’ case $p = 0$, where a safety criterion with sufficiently large τ becomes essential in preventing lane-changes which cause collisions.

In the standard MOBIL framework, if $I_n > 0$ and $S_n > 0$ at any given instant, then vehicle n changes lane, provided it has not already done so within the last $T > 0$ units of time. This latency is necessary to prevent combinations of manoeuvres in which individual vehicles repeatedly change lane with high frequency. In this respect we differ from the standard rules by incorporating latency in a stochastic way. Specifically, if $I_n > 0$ and $S_n > 0$ at any given instant, we implement a lane change according to a Poisson process with rate $\mu > 0$, i.e., at the current epoch with probability $\mu\Delta t$, where the model framework’s time-step is Δt . The consequent latency has time-scale $1/\mu$. This approach has the computational advantage that the history of vehicle trajectories need not be stored and moreover, solutions will not lock into regular patterns as a result of phase effects.

4.3 Simulation Framework

We now describe the numerical implementation of the microscopic multilane model posed in Section 4.2. The key computational challenge is maintaining references to neighbouring vehicles. These are simply lists of the identities of relevant vehicles, known as pointers. Here, $p_f(n)$ is the pointer which links car n_f to car n . We omit the argument if we are referring to the entire list, i.e. p_f , p_b and p_{fo} ; in fact these are all that are required since $p_{bo} = p_b(p_{fo})$. Similarly, only the headways $h_n^{(f)}$ and $h_n^{(fo)}$ are needed since $h_n^{(b)}$ and $h_n^{(bo)}$ can be determined using the appropriate combination of pointers, for example

$$h_n^{(bo)} = h_{p_b(n_{fo})}^{(f)} - h_n^{(fo)}. \quad (4.7)$$

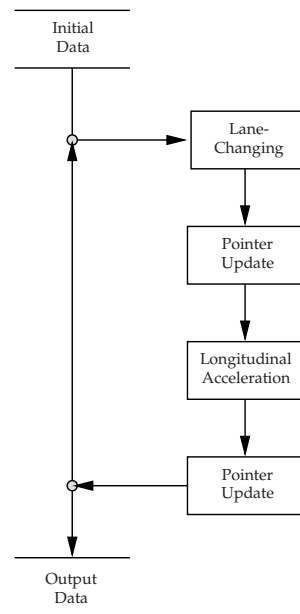


Figure 4.2. Multilane simulation flow diagram. Initial data (including all positions, headways, velocities, pointers and parameters) is read as input, after which the main loop is executed. This consists of four stages. First the lane-changing criteria are checked. Second, the pointers and headways are updated to account for any lane-changes. Third, the vehicles are accelerated according to the updated headways and velocity differences. Finally, the pointers and headways are updated again to account for any vehicles passing their downstream neighbour in the adjacent lane. The new variable values are written to an output file and the loop continues for the specified number of iterations.

A schematic overview of the program flow is illustrated in Figure 4.2. After the initial data has been read as input, the simulation enters its main loop and first determines which (if any) vehicles satisfy the lane-changing criteria. The identity of positive matches is stored and passed into a routine that updates the pointers and headways after a lane-change has occurred. The updated variables are then used by the acceleration law to advance each vehicle via a standard fixed time-step fourth order Runge-Kutta method. The vehicle acceleration phase may result in a number of cars passing their ‘front-other’ neighbour. If this occurs, the pointers are updated after which one iteration of the loop is complete. The lists of variables are consistently updated as they pass through each component of the simulation framework. This modularity and standard interfacing means that we can ‘plug-and-play’ a variety of lane-changing and longitudinal models with no changes to the underlying program core. We have used this feature to perform extensive tests with a number of trial lane-changing criteria as well as a variety of longitudinal models.

The post lane-change pointer update algorithm must distinguish between two sit-

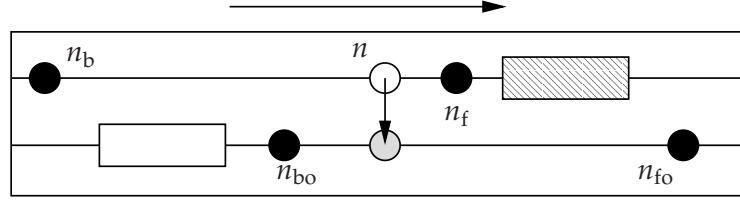


Figure 4.3. Lane-changing Case (A). Vehicle n (white circle) changes-lanes (grey circle). Its four neighbours are illustrated with black circles. The white and striped rectangles between the up- and downstream neighbours represent platoons of vehicles. In this case, the upstream platoon must change their front-other pointer from vehicle n to $p_f(n)$ after the manoeuvre.

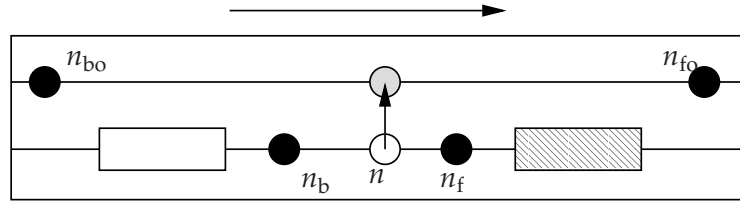


Figure 4.4. Lane-changing Case (B). Again vehicle n (white circle) changes-lanes (grey circle), its four neighbours are illustrated with black circles and the white and striped rectangles between the up- and downstream neighbours represent platoons of vehicles. In this case, the upstream platoon must change their front-other pointer to vehicle n after the manoeuvre.

uations, illustrated in Figures 4.3 and 4.4. We refer to these as Case (A) and (B) respectively. In both examples, vehicle n wishes to change lanes and its current (desired) position is illustrated with a white (grey) circle. The in-lane and out-of-lane nearest neighbours are illustrated with black circles and these may be separated by platoons of an arbitrary number of vehicles (illustrated with white and striped rectangles in Figures 4.3 and 4.4). These are immediately upstream of n_{bo} and n_b in Cases (A) and (B) respectively.

A flow diagram of the pointer update algorithm is illustrated in Figure 4.5. The left column corresponds to Case (A) and the right to Case (B). The upstream platoon of vehicles regard $p_{fo}(n)$ as their front-other prior to the lane-change in Case (A). These pointers are switched to n before any others are changed because they do not affect subsequent pointer shuffles. Similarly in Case (B), vehicles in the upstream platoon between n_b and n_{bo} see vehicle n as their front-other prior to the lane-change. These pointers are switched to n_{fo} first, since once again they do not affect subsequent shuffles.

Further changes are common to both Cases (A) and (B). These are illustrated in the bottom row of Figure 4.5. First the pointers of the nearest neighbours are updated. This involves changing the backward pointers of the downstream neighbours and the

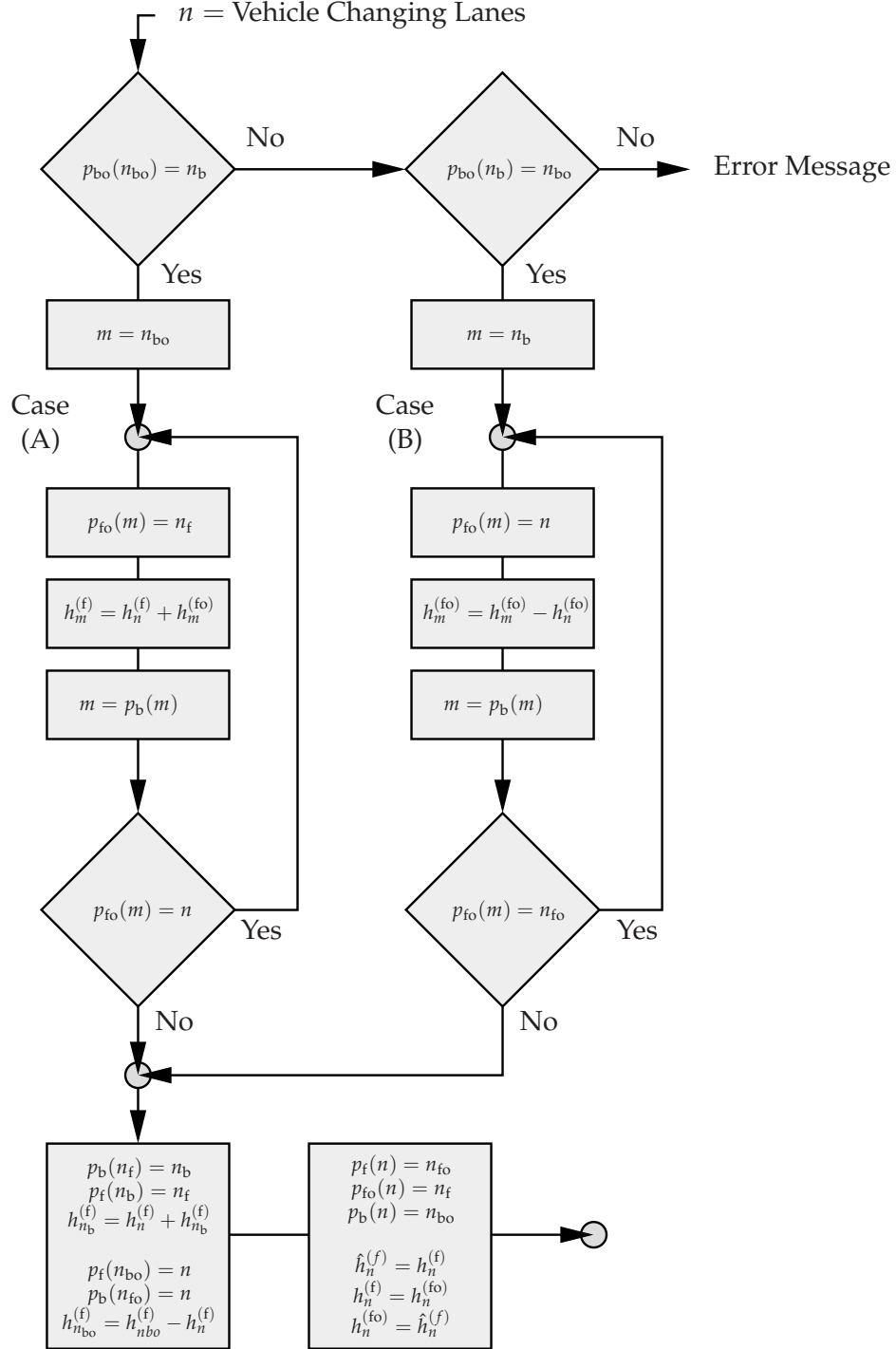


Figure 4.5. Pointer update algorithm for a single vehicle n changing lanes. When one or more lane-changes occur during a single time step, the algorithm is applied to each vehicle individually. The temporary scalar vehicle references (n_f , n_b , etc.) are fixed prior to the pointer/headway update. First the algorithm identifies whether vehicles are arranged according to Case (A) or (B). A while loop is then updates the pointers in succession through the platoon of vehicles upstream of, and including n_{bo} and n_b in Cases (A) and (B) respectively. Finally, the pointers of vehicle n and its neighbours are updated.

forward pointers of the upstream neighbours. Finally, the front $p_f(n)$ and front-other $p_{fo}(n)$ pointers of n are swapped and the back pointer $p_b(n)$ is changed to point to the back-other vehicle. Note that our pointers require that there is always at least one vehicle in a given lane at any one time. Strangely, if there is only one vehicle in either lane on the loop, it sees itself as the vehicle in front. This situation is highly unlikely in our set up where there are typically a large number of vehicles and the lane-changing criteria are the same in both lanes. However, when the lane-changing criteria are biased (as is the case when trying to model driving rules in the UK and Europe for example where there are specific passing lanes) and the spacing between vehicles is large there may not be any vehicles in one of the lanes. We note that this situation introduces problems with the front-other p_{fo} pointers which must be addressed when considering asymmetric lane-changing rules, however such a situation is very unlikely to occur in our set up.

The algorithm for updating pointers after a vehicle passes another is illustrated in Figure 4.6. This comprises of a loop over the vehicle indices which tests for negative front-other headways. If this is the case then the vehicle in question has passed what was previously its front-other neighbour and the front-other pointers of each vehicle are updated accordingly. This method assumes that one vehicle in a given lane cannot be passed by two vehicles in the adjacent lane during one time-step.

4.4 Longitudinal Quasi-Equilibrium

We now return to mathematical analysis of the MOBIL framework. Let us suppose that the lane-changing time-scale $1/\mu$ is much larger than the relaxation time-scale of the longitudinal model. In this setting, lane-changes are relatively rare, so that although each lane-change causes vehicles locally to accelerate/decelerate, in a short time, fluctuations have corrected themselves and the traffic is smooth again. In this limit, we expect flow to be longitudinally quasi-static, and the pairwise evolution of the traffic streams may thus be analysed entirely via consideration of the mass exchange between uniform flows.

The situation that we thus consider sees lane 1 occupied by a uniform flow of headway h_1 and velocity $V(h_1)$, sliding past a slower uniform flow of headway $h_0 < h_1$ and velocity $V(h_0) < V(h_1)$ in lane 0. The traffic in lane 0 thus has good reason to try and join lane 1. Here our goal is to analyse what proportion of time λ each vehicle n in lane 0 satisfies the incentive and safety criteria ($I_n > 0$ and $S_n > 0$) simultaneously. There follows a consequent mass transfer of traffic from lane 0 to lane 1 at rate $\lambda\mu$, so that

$$\dot{\rho}_0 = -\lambda\mu\rho_0 \quad \text{and} \quad \dot{\rho}_1 = +\lambda\mu\rho_0, \quad (4.8)$$

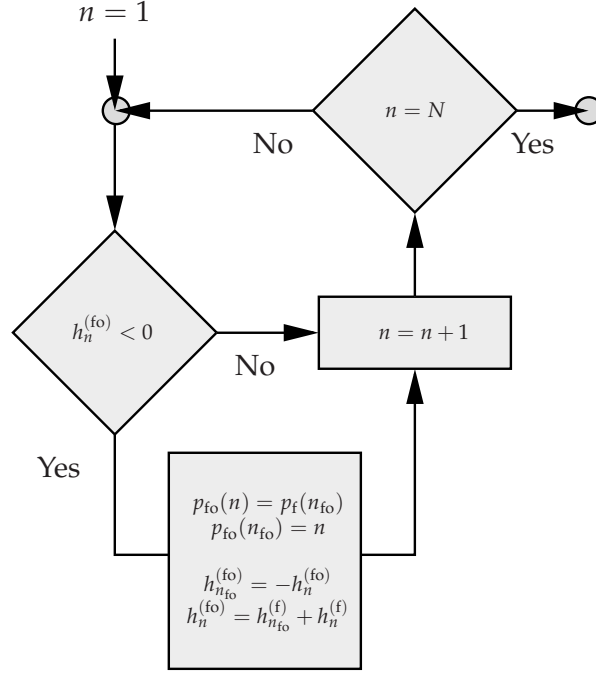


Figure 4.6. Pointer update algorithm for a passing manoeuvre. The temporary scalar neighbouring vehicle references (n_f , n_b , etc.) are fixed prior to the pointer/headway update. A loop through the vehicle identities determines if such a manoeuvre has occurred and the pointers and headways are updated accordingly.

where $\rho_0 = 1/h_0$ and $\rho_1 = 1/h_1$ give the density of traffic in lanes 0 and 1 respectively, and the rate λ depends on ρ_0 and ρ_1 through h_0 and h_1 . Note that there is no left-right bias in our model, hence all the arguments we present work through with the roles of the lanes reversed.

See Fig. 4.7 for the set-up. We introduce a phase variable $\theta \in [0, 1)$ which parameterises the position of vehicle n uniformly in space and time relative to the vehicles immediately in front and behind in the adjacent lane. At this point the incentive and safety functions may be written in the form $I(\theta; h_0, h_1)$ and $S(\theta; h_0, h_1)$, i.e, principally in terms of the phase θ , but also in terms of the quasi-static headways h_0 and h_1 (and of course the MOBIL parameters σ , τ , and p , and any other parameters from the longitudinal model). For the non-altruistic $p = 0$ case on which we focus, we have

$$I(\theta; h_0, h_1) = f(\theta h_1, V(h_1) - V(h_0), V(h_0)) - \sigma, \quad (4.9)$$

and

$$S(\theta; h_0, h_1) = f((1 - \theta)h_1, V(h_0) - V(h_1), V(h_1)) + \tau, \quad (4.10)$$

where f is the acceleration function of the general car-following model (1.12).

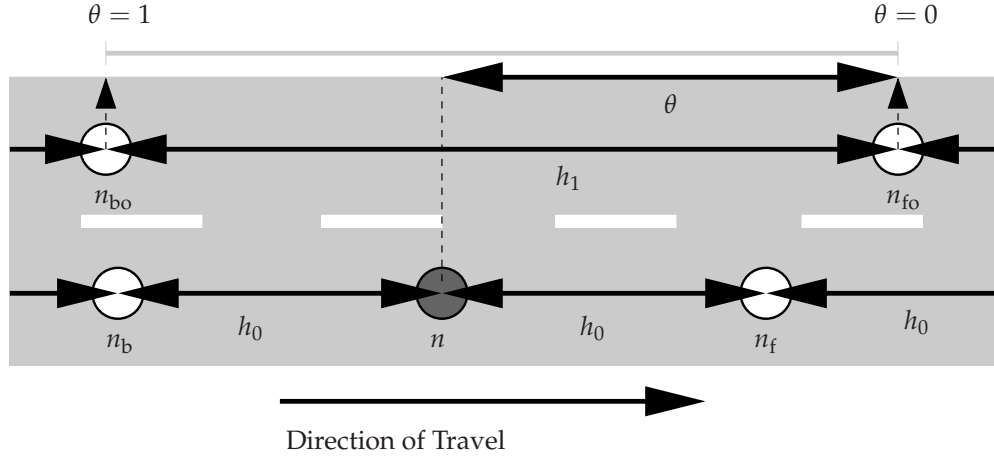


Figure 4.7. Definition of phase variable θ which measures the normalised position of an arbitrary vehicle n between vehicles n_{fo} and n_{bo} in the adjacent lane.

The mass transfer coefficient that we seek is then given by

$$\lambda = \int_0^1 H(I(\theta; h_0, h_1)) H(S(\theta; h_0, h_1)) d\theta, \quad (4.11)$$

where H is the Heaviside function. In averaging over the phase θ in this manner, we make an ergodic assumption which may be satisfied by supposing that either the phases are well distributed at any instant or that the lane-changing rate is sufficiently slow for the phases to be mixed over the characteristic lane-changing time-scale.

4.5 Partition of (h_0, h_1) Space

To calculate λ from (4.11), we need a detailed picture concerning where $I(\theta; h_0, h_1)$ and $S(\theta; h_0, h_1)$ change sign. If for a moment we suppose that $h_1 < h_0$, we may apply the sign constraints (1.17) and the monotonicity condition (1.15) to equation (4.9) and, by comparison with $f(h_0, 0, V(h_0)) = 0$, deduce that $I < 0$ for all θ . Hence in this quasi-static framework, no vehicle changes from the faster lane to the slower lane and we may revert to consideration of the opposite case without loss of generality.

We now consider the dependence of the incentive function on θ . For $p = 0$ we have

$$\partial I / \partial \theta = (D_1 f) h_1, \quad > 0, \quad (4.12)$$

using (4.9) and (1.17). Hence I increases with θ and from the point of view of the incentive function, lane changes are most desirable when $\theta = 1$, i.e., when vehicle n pulls out immediately in front of n_{bo} , since then it has the most space in front of it to

	Case I_a	Case I_b	Case I_c
Case S_a	0	0	0
Case S_b	0	$\max(0, \theta_S - \theta_I)$	θ_S
Case S_c	0	$1 - \theta_I$	1

Table 4.1. Each element corresponds to the value of λ for the case indicated (see Figure 4.8).

accelerate into. Clearly this conclusion is unacceptable from the gap acceptance point of view. One fix is to introduce altruism $p > 0$, to give

$$I(\theta; h_0, h_1) = f(\theta h_1, V(h_1) - V(h_0), V(h_0)) - \sigma \\ + pf(2h_0, 0, V(h_0)) + pf((1 - \theta)h_1, V(h_0) - V(h_1), V(h_1)), \quad (4.13)$$

and thus $\partial I / \partial \theta = [(D_1 f)^* - p(D_1 f)^{**}] h_1$, whose sign is not determined by (1.17). Here $(D_1 f)^*$ and $(D_1 f)^{**}$ denote values of the first partial derivative of f at different arguments. Since I need no longer be monotone increasing, it is possible for a lane-change to be maximally incentivised when vehicle n is roughly mid-way between n_{bo} and n_{fo} .

However, our approach is to set $p = 0$ and enforce gap acceptance via the safety criterion function given by (4.10). Note that

$$\partial S / \partial \theta = -(D_1 f) h_1, \quad < 0, \quad (4.14)$$

using (4.10) and (1.17). Results (4.12) and (4.14) are thus in opposition, since the former favours lane changes most when $\theta = 1 -$ and the latter favours lane changes most when $\theta = 0 +$ (since then vehicle n_{bo} need react the least to the lane-change).

The monotonicity of the incentive and safety criteria with respect to θ , (4.12) and (4.14) respectively, means that there are three generic cases to consider for each. These are illustrated in the top row of Figure 4.8 (Cases $I_a - I_c$) for the incentive criterion and in the bottom row of Figure 4.8 (Cases $S_a - S_c$) for the safety criterion. Note that the horizontal axes runs backwards to indicate how θ varies in spatial coordinates when vehicles drive to the right. In the leftmost column (Cases I_a and S_a), neither criteria are satisfied for any values of θ . In the centre column (Cases I_b and S_b), the incentive criterion is satisfied over the range $0 < \theta < \theta_I$ and the safety criterion is satisfied over the range $\theta_S < \theta < 1$. In the rightmost column (Cases I_c and S_c), both criteria are satisfied over the full range of θ . There are therefore nine possible combinations of the cases illustrated in Figure 4.8. For each of these combinations we print the mass transfer coefficient λ in Table 4.1. Note that column Case I_c and row Case S_c allow lane-changes that are not safe at $\theta = 1$ and $\theta = 0$. The central element (Case I_b

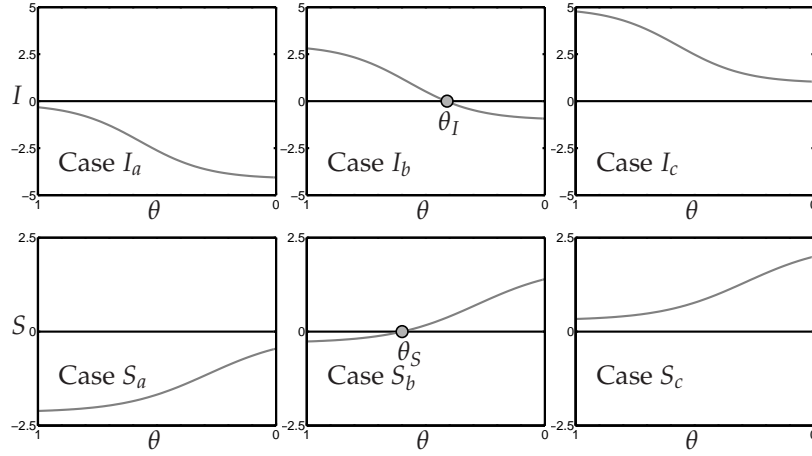


Figure 4.8. The three generic cases for $I(\theta)$ and $S(\theta)$. The incentive and safety criteria are either satisfied for all θ (Case I_c and Case S_a resp.), over an interval (Case I_b and Case S_b resp.) or not at all (Case I_a and Case S_c resp.). In both Case I_b and Case S_b , the criteria have zeros θ_I and θ_S respectively.

and S_b) is the ideal situation where lane-changing can occur at intermediate values of θ . However, lane-changing does not occur if $\theta_I > \theta_S$, since the incentive and safety criteria are not positive at the same values.

We would like to ensure that Cases I_c and S_c do not occur, by setting the thresholds σ and τ appropriately. However, it is also possible to set thresholds so that at certain headway pairs, no lane-changes occur at all (i.e., for any θ as in Cases I_a and S_a). To see this, consider either

$$I(1; h_0, h_1) = f(h_1, V(h_1) - V(h_0), V(h_0)) - \sigma, \quad (4.15)$$

or

$$S(0; h_0, h_1) = f(h_1, V(h_0) - V(h_1), V(h_1)) + \tau, \quad (4.16)$$

and apply monotonicity arguments using (1.17) and (1.15).

In the first case (4.15), $f(h_1, V(h_1) - V(h_0), V(h_0)) > 0$ by a comparison with either $f(h_0, 0, V(h_0)) = 0$ or $f(h_1, 0, V(h_1)) = 0$ and using the sign conditions (1.17) and monotonicity (1.15). Hence the overall sign in (4.15) cannot be determined a priori. In particular if $I(1; h_0, h_1) < 0$, then (4.12) rules out lane-changes for all θ . From (1.17) and (1.15) it is also true that $f(h_1, V(h_1) - V(h_0), V(h_0)) < f(h_1, V(h_1), 0)$. Therefore it is necessary that

$$\lim_{h_1 \rightarrow \infty} I(1, 0, h_1) > 0, \quad (4.17)$$

if the incentive criterion is to be satisfied at all.

In the second case (4.16), $f(h_1, V(h_0) - V(h_1), V(h_1)) < 0$ by comparison with $f(h_1, 0, V(h_1)) = 0$ and using the sign conditions (1.17) and monotonicity (1.15). Similarly, the overall sign in (4.16) cannot be determined a priori. If $S(0; h_0, h_1) < 0$, then (4.14) rules out lane-changes for all θ . Since $f(h_1, V(h_0) - V(h_1), V(h_1)) < 0$, it is necessary that

$$\tau > 0, \quad (4.18)$$

if the safety criteria is to be satisfied at all.

We have identified two necessary conditions for lane-changing, (4.17) and (4.18). We now turn our attention towards preventing dangerous manoeuvres. Similar considerations give

$$I(0; h_0, h_1) = f(0, V(h_1) - V(h_0), V(h_0)) - \sigma, \quad (4.19)$$

but the overall sign is unclear since $f(0, V(h_1) - V(h_0), V(h_0)) > 0$. If we wish to prevent lane-changing into the immediate wake of vehicle n_{fo} , we must set σ sufficiently large. From the sign constraints (1.17) and monotonicity (1.15), it is true that

$$f(0, V(h_1) - V(h_0), V(h_0)) < f(0, V(h_1), 0), \quad (4.20)$$

where $f(0, V(h_1), 0) > 0$ by comparison with $f(0, 0, 0) = 0$. Therefore σ should be chosen such that

$$\lim_{h_1 \rightarrow \infty} I(0; 0, h_1) < 0; \quad (4.21)$$

hence $I(0; h_0, h_1)$ is *never* satisfied and vehicles cannot move into the immediate wake of another.

The safety threshold τ should prevent lane-changes immediately downstream of the new follower, vehicle n_{bo} . To achieve this consider

$$S(1; h_0, h_1) = f(0, V(h_0) - V(h_1), V(h_1)) + \tau. \quad (4.22)$$

Here $f(0, V(h_0) - V(h_1), V(h_1)) < 0$ by comparison with either $f(h_1, 0, V(h_1)) = 0$ or $f(0, 0, 0) = 0$, and using the sign conditions (1.17) and monotonicity (1.15). Hence the overall sign of (4.22) is undetermined. It is also true that $f(0, V(h_0) - V(h_1), V(h_1)) > f(0, -V(h_1), V(h_1))$. Hence τ should be chosen such that

$$\lim_{h_1 \rightarrow \infty} S(1; 0, h_1) < 0, \quad (4.23)$$

to prevent the safety criterion from being satisfied at *all* values of (θ, h_0, h_1) . However this does not exclude the possibility of lane-changes occurring directly in front of the

new follower. Such manoeuvres always result in a collision since $V(h_1) > V(h_0)$.

One might consider choosing τ small enough that $S(1; h_0, h_1) < 0$ for all (h_0, h_1) pairs. This requires that $f(0, V(h_0) - V(h_1), V(h_1)) < -\tau$, however by comparison with $f(0, 0, 0) = 0$, we can always choose (h_0, h_1) such that $f(0, V(h_0) - V(h_1), V(h_1))$ is arbitrarily close to zero. Hence, to ensure that $S(1; h_0, h_1) < 0$ for all (h_0, h_1) pairs we must choose $\tau = 0$, however this contradicts (4.18) since it prevents the safety criteria from being satisfied at all. Thus we must choose τ such that $S(1; h_0, h_1) < 0$ when the incentive criteria are satisfied, i.e. when $I(1; h_0, h_1) > 0$.

To rectify this situation, let us consider how the incentive and safety criteria depend on the headways h_0 and h_1 . Firstly we have

$$\frac{\partial I}{\partial h_1} = (D_1 f)\theta + (D_2 f)V'(h_1), \quad > 0, \quad (4.24)$$

$$\frac{\partial I}{\partial h_0} = [(D_3 f) - (D_2 f)]V'(h_0), \quad < 0. \quad (4.25)$$

Note that the partial derivatives are evaluated at $(\theta h_1, V(h_1) - V(h_0), V(h_0))$. Hence, as we might expect, the incentive to change lane increases as the headway (and speed) in the adjacent lane increase, but the incentive decreases as ones present headway (and speed) increase. The trends of the safety criterion are more complicated. We have

$$\frac{\partial S}{\partial h_1} = (1 - \theta)(D_1 f) + [(D_3 f) - (D_2 f)]V'(h_1), \quad (4.26)$$

$$\frac{\partial S}{\partial h_0} = (D_2 f)V'(h_0), \quad > 0. \quad (4.27)$$

Note that the partial derivatives are evaluated at $((1 - \theta)h_1, V(h_0) - V(h_1), V(h_1))$. The sign of $\partial S / \partial h_1$ cannot be determined a priori, however $D_3 S(1; h_0, h_1) < 0$. Hence the safety criterion decreases monotonically in both h_0 and h_1 at $\theta = 1$. Thus, at $\theta = 1$ the sign of the partial derivatives of the safety criterion S are opposite to the incentive criterion I . Furthermore, both criteria are monotonic at $\theta = 1$. Thus, the region where $S(1; h_0, h_1) < 0$ (and hence safe lane-changing manoeuvres) occurs at lower h_0 values and higher h_1 values than the curve $S(1; h_0, h_1) = 0$. Similarly, the region where $I(1; h_0, h_1) > 0$ (and hence the incentive criteria are satisfied) also occurs at lower h_0 values and higher h_1 values than the curve $I(1; h_0, h_1) = 0$. Subsequently, we would like the curve $S(1; h_0, h_1) = 0$ to occur at larger h_0 and smaller h_1 values than the curve $I(1; h_0, h_1) = 0$, to ensure safe lane-changing manoeuvres. In cases where

$$|D_2 I| \geq |D_2 S| \quad \text{and} \quad |D_3 I| \geq |D_3 S|, \quad (4.28)$$

we can prevent dangerous lane-changes by choosing parameters such that

$$h_I \geq h_S, \quad (4.29)$$

where

$$I(1;0,h_I) = 0 \quad \text{and} \quad S(1;0,h_S) = 0, \quad (4.30)$$

given that σ and τ satisfy (4.17) and (4.23) respectively.

4.6 Worked Example

We now consider the OVRV model (1.8) as an example. To begin with, we fix the model parameters at $\alpha = 2$ and $\beta = 0.5$, and choose the standard OV function (1.6). The incentive criterion is

$$I(\theta;h_0,h_1) = \alpha \{V(\theta h_1) - V(h_0)\} + \beta \{V(h_1) - V(h_0)\} - \sigma, \quad (4.31)$$

and the safety criterion is

$$S(\theta;h_0,h_1) = \alpha \{V((1-\theta)h_1) - V(h_1)\} - \beta \{V(h_1) - V(h_0)\} + \tau. \quad (4.32)$$

To justify choosing the OVRV model, let us for a moment consider the incentive criterion of the standard OV model,

$$I(\theta;h_0,h_1) = \alpha \{V(\theta h_1) - V(h_0)\} - \sigma. \quad (4.33)$$

In this case vehicles will only change lanes if they can increase their headway by some threshold amount. We believe this to be too restrictive and not representative of real lane-changing manoeuvres. In contrast, vehicles changing lanes will accept a smaller headway in exchange for an increase in relative velocity with the OVRV model (4.31).

In Section 4.5, we derived a number of conditions on model parameters such that the lane-changing criteria could be satisfied and would not compromise driver safety. In the first case we found that parameters must satisfy (4.17) if the incentive criterion is to be satisfied at any (h_0, h_1) pair, for the OVRV model this means that

$$\sigma < \{\alpha + \beta\} V_{\max} \approx 4.91, \quad (4.34)$$

where $V_{\max} = 1 + \tanh(2)$. This requires that the incentive threshold is less than the maximum acceleration. The corresponding condition for the safety criterion to be satisfied (4.18) is simply $\tau > 0$. Condition (4.21) is necessary in order to prevent

lane-change into the immediate wake of vehicles in front in the neighbouring lane n_{fo} , thus

$$\sigma > \beta V_{\max} \approx 0.98. \quad (4.35)$$

To prevent the safety criterion from being satisfied at every point (θ, h_0, h_1) , it is necessary that condition (4.23) is satisfied,

$$\tau < (\alpha + \beta) V_{\max} \approx 4.91. \quad (4.36)$$

Similarly to (4.34), this requires that the safety threshold is less than the maximum deceleration. Because of the symmetry between acceleration and deceleration in the OVRV model, the upper bounds on σ and τ are subsequently the same.

The partial derivatives at $\theta = 1$ are

$$D_2 I = -(\alpha + \beta) V'(h_0), \quad D_3 I = (\alpha + \beta) V'(h_1), \quad (4.37)$$

$$D_2 S = \beta V'(h_0), \quad D_3 S = -(\alpha + \beta) V'(h_1); \quad (4.38)$$

hence (4.28) is satisfied and we can prevent lane-changes immediately upstream of vehicle n_{fo} via a proper choice of parameters. The intersections $I(1; 0, h_I) = 0$ and $S(1; 0, h_S) = 0$ are given by

$$(\alpha + \beta) V(h_I) - \sigma = 0 \quad \text{and} \quad -(\alpha + \beta) V(h_S) + \tau = 0, \quad (4.39)$$

respectively. Hence monotonicity of the OV function (1.15) means that we should choose

$$\sigma > \tau \quad (4.40)$$

to prevent dangerous lane-changes.

In Figure 4.9(a–c) we plot the mass transfer coefficient λ in the (h_0, h_1) plane for $\sigma = 2$ and $\tau = 1.4, 0.8, 0.2$ respectively. Lighter shades correspond to lower values of λ . The lines I_1 , S_1 and S_0 are zero level sets of (4.15), (4.22) and (4.16) respectively. The qualitative differences between panels (a–c) are due to changes in the position of the line S_0 . The boundary $\theta_I = \theta_S$ is also plotted with a black dashed line, which we calculate with a crude numerical algorithm. This encloses the region where $\lambda > 0$ and hence where lane-changing occurs.

The S_0 curve is not present in panel (a). For the OVRV model, the safety criterion is satisfied at $\theta = 0$ when

$$-\beta \{V(h_1) - V(h_0)\} + \tau > 0, \quad (4.41)$$

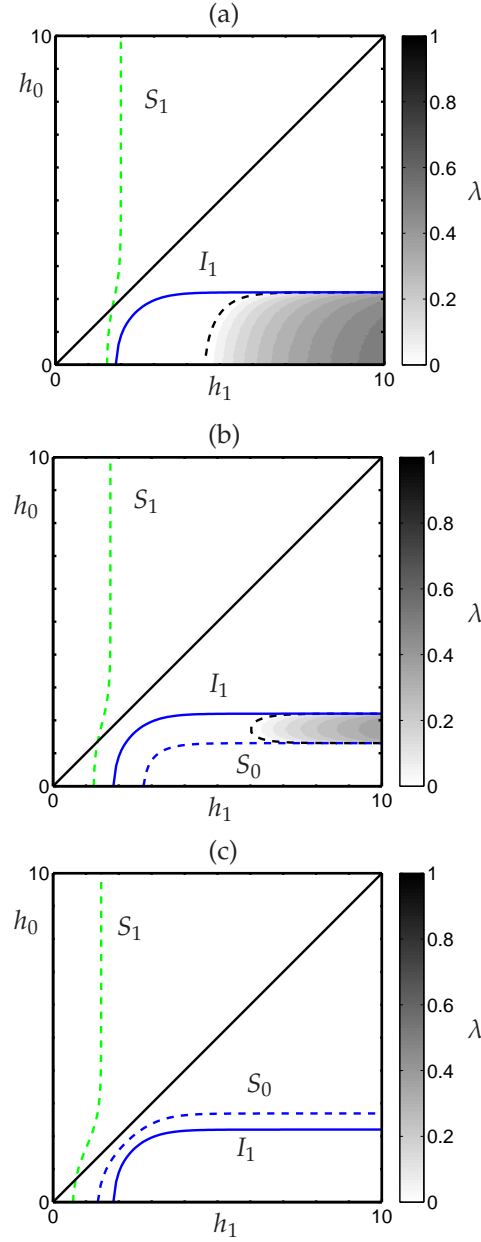


Figure 4.9. Grey scale contours of the mass transfer coefficient λ in the (h_0, h_1) plane. Values of λ between 0 and 1 are illustrated with shades between white and black. The OVRV model is used with parameters $\alpha = 2$, $\beta = 0.5$; the incentive threshold set at $\sigma = 2$. The safety threshold τ is 1.4, 0.8 and 0.2 in panels (a–c) respectively. The solid blue line I_1 marks the $I(1; h_0, h_1) = 0$ boundary, the dashed blue line S_0 marks the $S(0; h_0, h_1) = 0$ boundary (only panels (b) and (c)) and the dashed green line S_1 marks the $S(1; h_0, h_1) = 0$ boundary. The edge of the region where lane-changing occurs $\theta_I = \theta_S$ is marked with a black dashed line (panels (a) and (b)). Note that the size of this region diminishes as τ decreases, in fact no lane-changing occurs in panel (c). There are also large regions where lane-changing does not occur in both panels (a) and (b), in particular when $h_0 > h_1$.

hence the only negative contribution is $-\beta V(h_1)$. Thus when $\tau > \beta V_{\max}$, the safety criterion S_0 is satisfied at all (h_0, h_1) pairs. This is the case for panel (a) since $\tau = 1.4$ and $\beta V_{\max} \approx 0.98$. This condition is a result of monotonicity of the OVRV safety criteria at $\theta = 0$, which is a specific property of the OVRV model. A similar condition will apply to models where this is not the case, however one may have to resort to a numerical approximation rather than an explicit relation. The region $\lambda > 0$ where lane-changing can occur is bounded by I_1 and the $h_0 = 0$ axis. This implies that stationary vehicles can change into moving streams of traffic. Surprisingly, vehicles in the target lane must be travelling *above* a minimum speed before lane-changes become desirable/safe. This provides vehicles changing lanes with enough space in which to accelerate and hence prevents the new follower from braking severely.

In panel (b), $\tau = 0.8$; hence $S(0; h_0, h_1)$ can be negative and the region $\lambda > 0$ is bounded by I_1 and S_0 . This greatly reduces the number of situations in which the lane-changing criteria are satisfied. This is taken to the extreme in panel (c) where the relative position of the S_0 and I_0 curves switch and no lane-changing occurs. For the OVRV model, the incentive criterion is satisfied at $\theta = 1$ if

$$(\alpha + \beta) \{V(h_1) - V(h_0)\} - \sigma > 0, \quad (4.42)$$

hence combining (4.41) and (4.42) yields

$$\tau > \frac{\beta}{\alpha + \beta} \sigma. \quad (4.43)$$

This prescribes another necessary condition for lane-changing which is again due to monotonicity of the safety criteria at $\theta = 0$ and is hence specific to the OVRV model.

In Figure 4.10 we plot the mass transfer coefficient λ in the (h_0, h_1) plane for the IDM model using the standard parameters suggested in [79]. We also use the thresholds $\sigma = 1$ and $\tau = 4$ as in [48]. In contrast to the OVRV model, the S_0 curve is not monotonic, hence the region where the safety criterion is not satisfied is finite. The mass exchange coefficient is otherwise qualitatively the same as those in Figure 4.9 for the OVRV model.

4.7 Taylor Approximation of λ

Our aim is to connect microscopic lane-changing with macroscopic density exchange rates between lanes of traffic with different density. So far, we have used the quasi-equilibrium assumption to link the microscopic MOBIL lane-changing criteria with

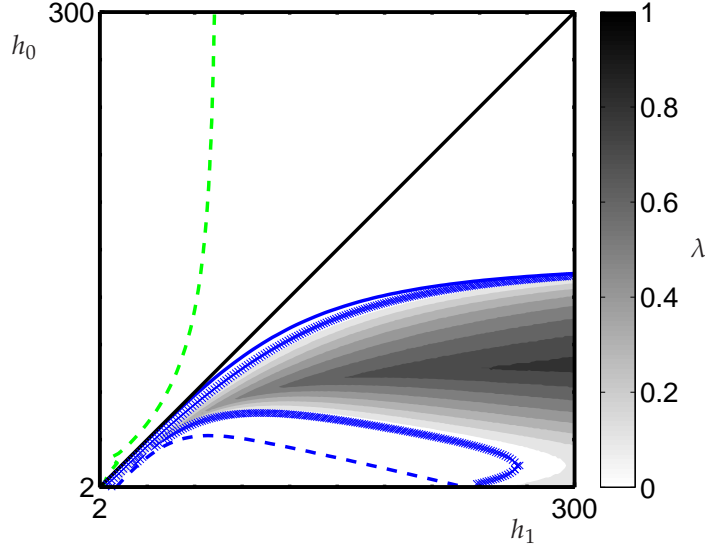


Figure 4.10. Grey scale contours of the mass transfer coefficient λ in the (h_0, h_1) plane for the IDM model with standard parameters [79]. The lane-changing thresholds suggested in [48] are used: $\sigma = 0.1$, $\tau = 4$. As previously, the solid blue line I_1 marks the $I(1; h_0, h_1) = 0$ boundary, the dashed blue line S_0 marks the $S(0; h_0, h_1) = 0$ boundary and the dashed green line S_1 marks the $S(1; h_0, h_1) = 0$ boundary.

the mass exchange coefficient λ . However, we have had to resort to numerical methods to visualise this in the (h_0, h_1) plane. In this section we use a *linear* Taylor expansion about the initial conditions to obtain an approximation of λ . Surprisingly, we can use this to solve the macroscopic density exchange rate equations (4.8) explicitly.

We assume that the initial densities $(\bar{\rho}_0, \bar{\rho}_1)$ lie within a region where lane-changing occurs. Thus, with the conditions on model parameters described in Section 4.5, the mass transfer coefficient is

$$\lambda = \theta_S - \theta_I. \quad (4.44)$$

We also assume that the zeros of the incentive and safety functions can be calculated a priori at $t = 0$, which we denote $\bar{\theta}_I$ and $\bar{\theta}_S$ respectively. Since θ_I and θ_S change with the densities $\rho_0 = 1/h_0$ and $\rho_1 = 1/h_1$, we choose to write them implicitly, such that

$$I(\theta_I(\rho_0, \rho_1), 1/\rho_0, 1/\rho_1) = 0, \quad (4.45)$$

$$S(\theta_S(\rho_0, \rho_1), 1/\rho_0, 1/\rho_1) = 0. \quad (4.46)$$

Expanding $\theta_I(\rho_0, \rho_1)$ and $\theta_S(\rho_0, \rho_1)$ to first order about the initial densities $(\bar{\rho}_0, \bar{\rho}_1)$, we obtain

$$\lambda \approx \lambda_0 + [D_1\theta_S - D_1\theta_I] (\rho_0 - \bar{\rho}_0) + [D_2\theta_S - D_2\theta_I] (\rho_1 - \bar{\rho}_1), \quad (4.47)$$

where D_1 (D_2) corresponds to the partial derivative with respect to the first (second) variable and $\lambda_0 = \bar{\theta}_S - \bar{\theta}_I$.

Differentiating the incentive function (4.45) with respect to the density ρ_0 whilst holding ρ_1 fixed, and vice-versa for ρ_1 , we obtain

$$D_1\theta_I = \frac{D_2I}{\bar{\rho}_0^2 D_1I} \quad \text{and} \quad D_2\theta_I = \frac{D_3I}{\bar{\rho}_1^2 D_1I}, \quad (4.48)$$

where the partial derivatives are evaluated at $(\bar{\theta}_I, 1/\bar{\rho}_0, 1/\bar{\rho}_1)$. Note that the incentive function and its derivatives are the same as those used in Sections 4.4 and 4.5, since we have accounted for the mapping between headway and density variables via the factor $-1/\bar{\rho}_{0,1}^2$. Similarly for the safety function (4.46), we find

$$D_1\theta_S = \frac{D_2S}{\bar{\rho}_0^2 D_1S} \quad \text{and} \quad D_2\theta_S = \frac{D_3S}{\bar{\rho}_1^2 D_1S}. \quad (4.49)$$

where the partial derivatives are evaluated at $(\bar{\theta}_S, 1/\bar{\rho}_0, 1/\bar{\rho}_1)$.

Combining (4.8) and (4.47–4.49), we find

$$\dot{\rho}_0 = -\dot{\rho}_1 = -\mu [\lambda_0 + d_0(\rho_0 - \bar{\rho}_0) + d_1(\rho_1 - \bar{\rho}_1)] \rho_0 \quad (4.50)$$

where the constants d_0 and d_1 are given by

$$d_0 = \frac{1}{\bar{\rho}_0^2} \left(\frac{D_2S}{D_1S} - \frac{D_2I}{D_1I} \right) \quad \text{and} \quad (4.51)$$

$$d_1 = \frac{1}{\bar{\rho}_1^2} \left(\frac{D_3S}{D_1S} - \frac{D_3I}{D_1I} \right), \quad (4.52)$$

evaluated at their respective initial values. We first note, as in [25], that $\dot{\rho}_0 + \dot{\rho}_1 = 0$, which we can integrate directly to obtain $\rho_0 + \rho_1 = \bar{\rho}_0 + \bar{\rho}_1$. Thus we can simplify the system by eliminating ρ_1 ; this yields

$$\dot{\rho}_0 = -\mu [\lambda_0 + (d_0 - d_1)(\rho_0 - \bar{\rho}_0)] \rho_0 \quad (4.53)$$

The equilibrium density ρ_0^* is therefore

$$\rho_0^* = \bar{\rho}_0 - \frac{\lambda_0}{d_0 - d_1}, \quad (4.54)$$

note that this is independent of the rate μ .

Note that (4.53) is in the form of a Bernoulli equation. Hence dividing through by ρ_0^2 and making the change of variables $\phi = 1/\rho_0$ yields the first order differential

equation

$$\dot{\phi} = \mu [\lambda_0 - (d_0 - d_1)\bar{\rho}_0] \phi + \mu(d_0 - d_1), \quad (4.55)$$

from which we obtain the solution

$$\rho_0(t) = \bar{\rho}_0 \left\{ \frac{\lambda_0 - (d_0 - d_1)\bar{\rho}_0}{\lambda_0 \exp \{ \mu [\lambda_0 - (d_0 - d_1)\bar{\rho}_0] t \} - (d_0 - d_1)\bar{\rho}_0} \right\}. \quad (4.56)$$

Furthermore, $\rho_0^* > 0$ ensures stability of the equilibria.

4.8 Comparison with Microscopic Simulation

We now test the longitudinal Quasi-Equilibrium Assumption (QEA) of Section 4.4 and the macroscopic approximation of Section 4.7 against microscopic simulation data. As in Section 4.6, we use the OVRV model with $\alpha = 2$ and $\beta = 0.5$. We also choose the lane-changing thresholds $\sigma = 2$ and $\tau = 1.4$. We obtain qualitatively similar results for other parameter values. Densities are calculated from the microscopic simulations by simply taking the reciprocal of the mean headway at each output time. We therefore use a large number of vehicles (typically of the order of 10^3) to obtain smooth macroscopic averages. We compare these with 1. Densities obtained from numerical solutions of the macroscopic differential equation (4.8) where λ is calculated using (4.11) (which we refer to as numerical QEA data); and 2. Explicit solutions to the linearised quasi-equilibrium rate equation (4.56) (which we refer to as explicit linearised QEA solutions).

In Figure 4.11(a–c) we compare microscopic data (white and black circles) with the numerical QEA data (solid lines) for three different lane-changing rates $\mu = 0.10$, 0.01 and 0.001 respectively with the initial spacings $\bar{h}_0 = 1.5$ ($\bar{\rho}_0 \approx 0.67$) and $\bar{h}_1 = 12$ ($\bar{\rho}_1 \approx 0.08$). Lower values of μ correspond to longer lane-changing timescales, hence we expect that our quasi-equilibrium approximation to become increasingly accurate as μ is decreased. We see in Figure 4.11 that this is indeed the case and we have a good agreement at $\mu = 0.01$. Similarly, in Figure 4.12 we compare the same microscopic data with the explicit linearised QEA solutions using the same parameter values and initial conditions as illustrated in Figure 4.11. Again the agreement improves as μ becomes smaller and there is almost no apparent difference between Figures 4.12 and 4.11. Note that the equilibrium density of the microscopic data changes with the lane-changing rate μ , a feature that is not present in either of the approximation methods.

We would like to judge the performance of our quasi-equilibrium approach; hence we need to know values of μ that give qualitatively correct macroscopic lane-changing

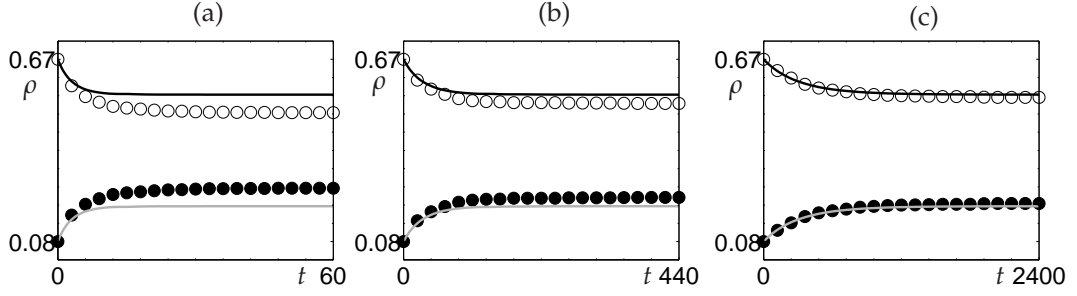


Figure 4.11. Comparison of averaged microscopic density and macroscopic numerical QEA data for the OVRV model. Model parameters are $\alpha = 2$, $\beta = 0.5$, lane-changing thresholds are $\sigma = 2$, $\tau = 1.4$ and initial headways are $\bar{h}_0 = 1.5$ and $\bar{h}_1 = 12$. Panels (a–c) correspond to lane-changing rates $\mu = 0.1, 0.01, 0.001$ respectively. There is a good quantitative agreement for values $\mu < 0.01$ where lane-changes are infrequent. Note that the equilibrium density of the microscopic model changes with the lane-changing rate, a feature that is not reproduced by the approximation method.

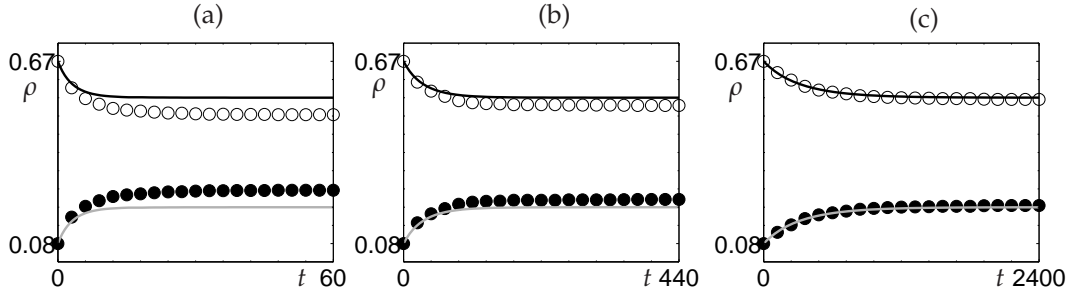


Figure 4.12. Comparison of averaged microscopic density and the explicit linearised QEA solutions for the OVRV model. Parameters and initial conditions are the same as in Figure 4.11: $\alpha = 2$, $\beta = 0.5$, $\sigma = 2$, $\tau = 1.4$, $\bar{h}_0 = 1.5$, $\bar{h}_1 = 12$. Panels (a–c) correspond to lane-changing rates $\mu = 0.1, 0.01, 0.001$. Again, there is a good quantitative agreement for values $\mu < 0.01$ where lane-changes are infrequent and the equilibrium density is the same for different lane-changing rates μ .

rates. Empirical data indicates that the mean lane-changing rate on highways is a uni-model function of density, with the maximum being of the order of hundreds of lane-changes per lane per kilometre per hour [48, 74]. Note that in our highly simplified experimental setting, the lane-changing rate tends to zero. The constant lane-changing rate observed in data is due to heterogeneous driver behaviour, hence it is difficult to relate this directly to our experimental setting. However, preliminary attempts using the dimensional IDM model to recreate simulation results obtained in [48] with distributed driver behaviour indicate that realistic macroscopic lane-changing rates are obtained with values in the range $\mu=0.01$ – 0.1 . To compare this with those used in Figures 4.11 and 4.12 we must first estimate how the timescale of the OVRV model scales with dimensional seconds using the common ‘2-second rule’. This states that a safe

<i>Density Difference (a)</i>						
h_0	h_1	n_0	n_1	$\mu = 0.100$	$\mu = 0.010$	$\mu = 0.001$
2.5	18	1080	150	0.5112	0.1592	0.0340
0.1	15	1500	10	0.0471	0.0118	0.0031
1.5	6	1000	250	0.4725	0.1617	0.0322
1.5	12	1000	125	0.4756	0.1880	0.0531

<i>Density Difference (b)</i>						
h_0	h_1	n_0	n_1	$\mu = 0.100$	$\mu = 0.010$	$\mu = 0.001$
2.5	18	1080	150	-0.1798	-0.3698	-0.4193
0.1	15	1500	10	0.0563	0.0181	0.0082
1.5	6	1000	250	0.4734	0.1618	0.0324
1.5	12	1000	125	0.4494	0.1606	0.0344

Table 4.2. Comparison of microscopic data with QEA data (a) and linearised QEA solutions (b). Data is for the OVRV model with four different initial headways (first two columns) using three lane-changing timescales $\mu = 0.1, 0.01$ and 0.001 . Model parameters are $\alpha = 2, \beta = 0.5$ and the lane-changing thresholds are $\sigma = 2$ and $\tau = 1.4$. The third and fourth columns in each table are the number of vehicles in each lane. The rightmost three columns are the fractional density difference error $\varepsilon_{a,b}$ for (a) and (b) respectively. **Table (a):** For $\mu = 0.1$ the agreement is poor for most cases, however this improves significantly for $\mu \leq 0.01$. Note that there is very good agreement for all values in the second row. This is because there is very little change in the equilibrium density differences, i.e. few lane-changes take place. **Table (b):** In most cases there is very little quantitative difference between the results for the linearised QEA solutions and the macroscopic QEA data (Table (a)). The largest discrepancy between the approximation methods is for the initial headways $(h_0, h_1) = (2.5, 18)$. This initial condition is close to the I_1 boundary (see Figure 4.13), where λ is strongly non-linear. In this case, the linearised QEA solution under estimates the density difference. Note that in contrast to the other initial conditions, there agreement is better at *large* μ values.

headway should allow at least two seconds to pass before reaching the position of the vehicle in front at the time counting commenced. This is equivalent to saying that the maximum gradient of the equilibrium velocity function is $1/2$. Since the maximum gradient of the OV function is 1, the timescale of the OVRV model is of the order of seconds. We therefore conclude that our approximation has worked well considering the highly simplified set-up, however further experiments with heterogeneous drivers are necessary.

To make a quantitative comparison between simulation data and our approximations, we first difference the equilibrium densities between lanes for the microscopic

simulation data $\Delta\rho_s$, the QEA data $\Delta\rho_a$ and the linearised QEA solutions $\Delta\rho_b$. We then calculate the fractional error of the density differences $\varepsilon_{a,b}$ for the QEA data and linearised QEA solutions respectively,

$$\varepsilon_{a,b} = \frac{\Delta\rho_s - \Delta\rho_{a,b}}{\Delta\rho_s}. \quad (4.57)$$

We present results in Tables 4.2(a) and (b) for a selection of initial headway pairs (h_0, h_1) (first two columns, corresponding number of vehicles are in the third and fourth columns). Data values are for the OVRV model with three lane-changing rates $\mu = 0.1, 0.01$ and 0.001 , for the model parameters $\alpha = 2$, $\beta = 0.5$ and lane-changing thresholds $\sigma = 2$, $\tau = 1.4$. We run the simulations until the density is approximately constant (i.e. once lane-changing stops and the lanes decouple) to quantify how close the equilibria of the QEA methods are to the microscopic simulations. The rightmost three columns of Tables 4.2(a) and (b) display the fractional errors $\varepsilon_{a,b}$ respectively. There is very little quantitative difference between the QEA approximation methods in Tables 4.2(a) and (b). The only considerable deviation is in the first row. This initial condition lies very close to the I_1 boundary where λ is highly nonlinear (see Figure 4.11). In contrast to the other initial conditions, the linear QEA approximates the final density in this particular case *better* at *smaller* lane-changing timescales (hence large μ values).

We plot the mass exchange coefficient λ in the (ρ_0, ρ_1) plane in Figure 4.13. On this we overlay the microscopic trajectory data for $\mu = 0.01$ and the initial conditions corresponding in the first, third and fourth rows of Tables 4.2(a) and (b) ($(h_0, h_1) = (2.5, 18)$), $(1.5, 6)$ and $(1.5, 12)$ respectively). The trajectories in the (ρ_0, ρ_1) plane are linear because cars move directly from one lane to the other, thus the total density remains constant. This figure also illustrates why there is a difference in equilibrium density between the QEA approximation methods for initial conditions close to the I_1 boundary. The mass exchange coefficient is highly non-linear in this region, hence our linear approximation is not valid. However, it appears that this non-linearity is *not* a good approximation of the microscopic data, since we observe that the microscopic data trajectory overshoots this boundary.

4.9 Lane-Changing and Instability

So far, the analysis and results presented in this chapter have been in stable parameter regimes, where we assumed that the mass exchange between lanes is spatially independent. In this section we use three numerical experiments to illustrate how lane-changing can force spatio-temporal pattern when parameters are unstable and/or

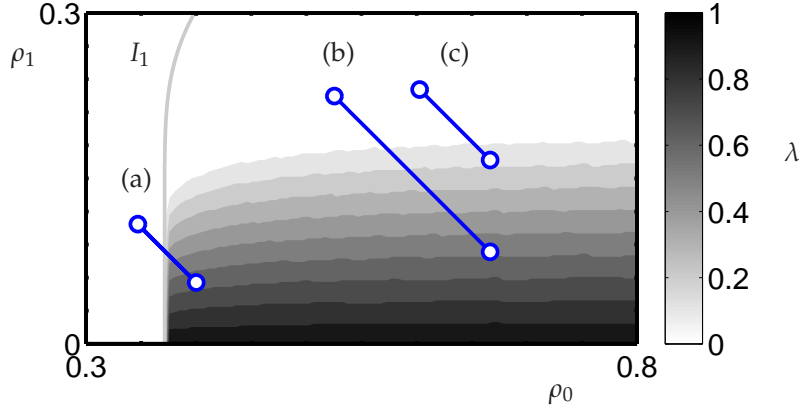


Figure 4.13. Mass exchange coefficient in the (ρ_0, ρ_1) plane for the OVRV model with $\mu = 0.01$ and the same parameters and lane-changing thresholds used in Tables 4.2(a) and (b), namely $\alpha = 2$, $\beta = 0.5$, $\sigma = 2$ and $\tau = 1.4$. Overlaid on this are the averaged microscopic data trajectories for the initial headways $(h_0, h_1) = (2.5, 18)$, $(1.5, 6)$ and $(1.5, 12)$ labelled (a–c) respectively. Conservation of mass across the two lanes means that these trajectories are linear. Note that λ is highly non-linear at the I_1 boundary, however this is not reflected in the microscopic data.

driver behaviour is heterogeneous. To achieve this we must relax the conditions on model parameters, (4.17), (4.18), (4.21) and (4.23), outlined in Section 4.5. Permitting lane-changes that seriously disadvantage the new follower can only be tolerated if they can react appropriately. Increasing the breaking capacity of vehicles (i.e. increasing α and β) resolves this to some extent, however it eliminates the desired instability. We therefore decrease the lane-changing timescale (and hence increase μ), so that vehicles can respond quickly to inhibitive manoeuvres (for example by moving into the space left by the vehicle changing lanes).

When parameters are unstable, small perturbations develop into large-scale stop-and-go waves. Lane-changes act to perturb *both* the source and target lanes, since traffic in the source lane can accelerate into the previously occupied gap and traffic in the target lane is forced to brake because of the additional vehicle. To illustrate the subsequent traffic patterns, we use the coarse-graining procedure described in Chapter 2 to produce numerical spatio-temporal densities for individual lanes. The coarse-grained density for a given time step is calculated by convolving a normalised Gaussian test function on the point distribution of vehicles, the characteristic length

scale of which is large enough to smooth out individual vehicles and small enough to preserve spatial structure.

We illustrate results using the OVRV model with the standard OV function (1.6). The three experimental set-ups are as follows: 1. Identical drivers initially in equilibrium with different equilibria in each lane; 2. Identical drivers initially in the same equilibrium in both lanes with a small perturbation to a single driver in one lane; 3. Heterogeneous drivers initially far from equilibrium. Driver heterogeneity is introduced via a linear scaling of the OV function, i.e. $V_n(h) = \gamma_n V(h)$ where $\gamma \in (0, 1]$ is the scaling factor. Rather than attempt a full mathematical analysis, our goal is simply to describe and compare the patterns observed in simulation.

Experiment 1. We duplicate the initial set-up of Section 4.2. All vehicles are identical and initially both lanes are in equilibrium, however the equilibria in each lane differ; the initial spacing in lane 0 is $h_0 = 1.8$ and in lane 1 is $h_1 = 2.7$. Model parameters $\alpha = 0.5$, $\beta = 0.25$, thus the system is acutely unstable to force large amplitude stop-and-go waves. The lane-changing thresholds are $\sigma = 0.01$ and $\tau = 2.0$. A small system of vehicles is used (initially $n_0 = 30$ and $n_1 = 20$), since we observed that the same length-scale of jams in larger systems. Coarse-grained spatio-temporal densities are plotted in Figure 4.14, where lane-changes from lane 0 to 1 are marked with a cross and from lane 1 to 0 with a circle.

The first striking feature is that despite different initial conditions, the density profiles of each lane *synchronise*. This phenomena results from a series of catalytic events, which we mark with labels p_1 – p_4 in Figure 4.14. At the marker p_1 , a vehicle switches from lane 0 to lane 1 where the density is lower. This triggers a further lane-change from lane 1 to lane 0 at a slightly later time *downstream*, and the subsequent disturbance propagates upstream in lane 0. At approximately the same time further upstream at marker p_2 , two lane-changes from lane 0 perturb the density in lane 1. A burst of manoeuvres at p_3 subsequently synchronise the two wave profiles. Further lane-changes are concentrated at the up- and downstream fronts (see marker p_4), which gradually remove slight phase differences between the jams. Note however that lane-changes cease once the density profiles fully synchronise and hence the lanes become *decoupled*.

Experiment 2. All vehicles are identical and initially both lanes have the *same* equilibrium spacings $h_0 = h_1 = 1.8$ and hence velocities, however we perturb the velocity of a single vehicle in lane 0. Model parameters are $\alpha = 0.5$, $\beta = 0.25$ and the lane-changing threshold parameters are $\sigma = 0.01$ and $\tau = 2.0$. We use a large number of vehicles and present data prior to the interaction of waves travelling up- and downstream around the loop. Coarse-grained density results are presented in Figure 4.15;

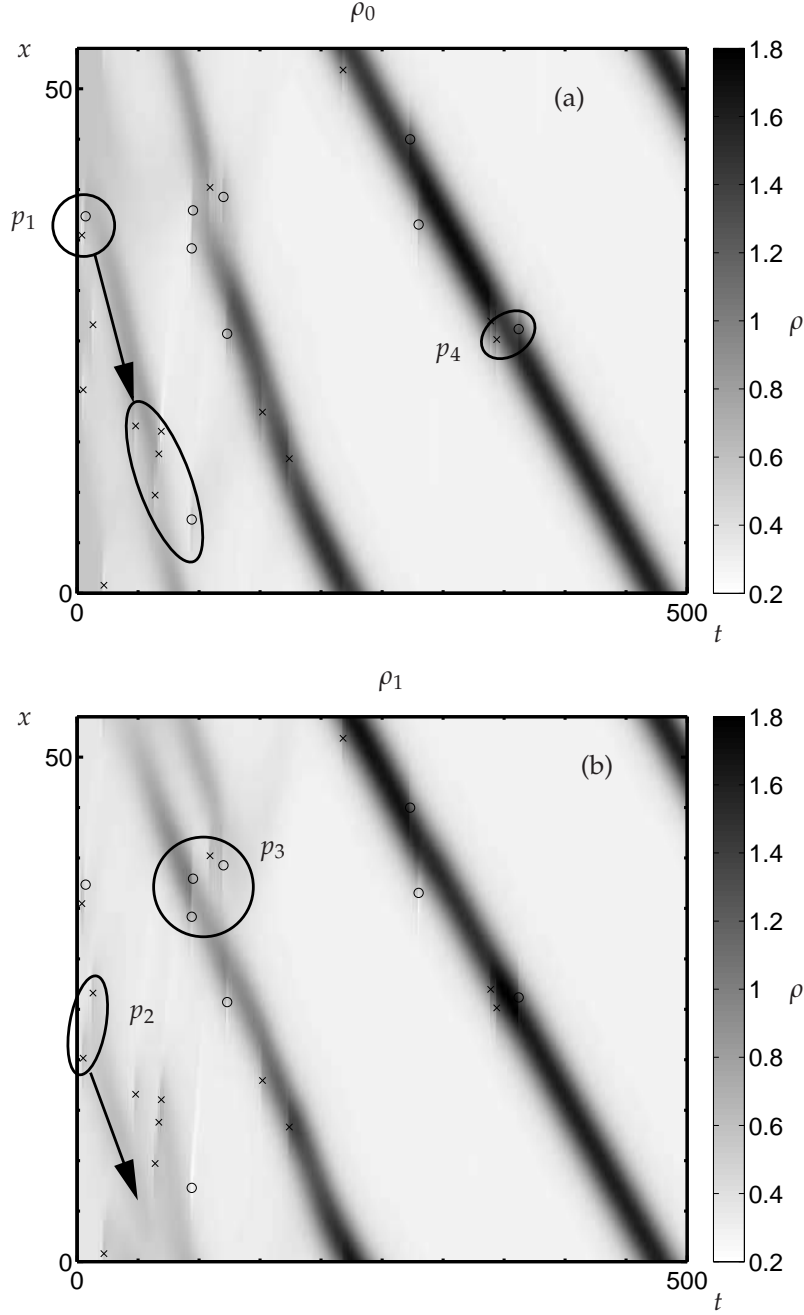


Figure 4.14. Evolution of density in lane 0 (panel (a)) and lane 1 (panel (b)) from initially homogeneous conditions. The set-up is identical to that used in Section 4.2. Position on the loop x is plotted against time t and the grey-scale corresponds to the coarse-grained density. Parameters are such that flows are linearly unstable, $\alpha = 0.5$, $\beta = 0.25$; the lane-changing threshold parameters are $\sigma = 0.01$ and $\tau = 2.0$; the initial conditions are $h_0 = 1.8$, $n_0 = 30$, $h_1 = 2.7$ and $n_1 = 20$ ($L = 54$); the lane-changing rate is $\mu = 0.01$. Crosses mark lane-changes from lane 0 to 1 and circles mark lane-changes from lane 1 to 0. A stop-and-go wave emerges and *synchronises* across the lanes. The series of lane-changes that lead up to this are marked by p_1 – p_4 .

again lane-changes from lane 0 to 1 are marked with a cross and from lane 1 to 0 with a circle. Panel (a) of Figure 4.15 is the density profile of lane 0 when lane-changing is permitted ($\mu = 0.01$) and panel (b) is the density profile of lane 0 when lane-changing is *not* permitted ($\mu = 0$). Lane-changing quickly synchronises the density of lanes 0 and 1, hence we do not plot the density profile of lane 1.

The initial growth of perturbations is qualitatively similar when lane-changing is both allowed and disallowed, despite a large number of lane-changes around the origin of the disturbance. Comparing zones Z_1 and Z_2 , we observe that the jam length scale is much smaller when lane-changing is present. However, lane-changing triggers further development of stop-and-go waves at the upstream front in Figure 4.15(a) which are not present in Figure 4.15(b).

Experiment 3. Two equal populations of vehicle type are present, initially distributed evenly between lanes but whose order is selected randomly. The scaling factor of the slow population is $\gamma = 0.6$, whereas the fast population are not scaled, i.e. $\gamma = 1$. All vehicles initially have the same spacing $h_* = 2$, which determines their velocity via the (distributed) OV function. Hence the system is *not* initially in equilibrium. Model parameters are $\alpha = 0.6$, $\beta = 0.3$ and the lane-changing thresholds are $\sigma = 0.1$ and $\tau = 0.5$. The number of vehicles starting in each lane is $n_0 = n_1 = 400$. Stochasticity is not present, hence $\mu = 1/\Delta t$ where Δt is the simulation time-step. The spatio-temporal density profiles are illustrated in Figure 4.16 where the left panel, ρ_0 , is the density in lane 0 and the right panel, ρ_1 , is the density in lane 1.

In stark contrast to the previous experiments with identical vehicles, the density profiles become *asynchronised* (compare the densities in the white circles labelled A and B in ρ_0 and ρ_1 respectively). This is however a transient effect; the length scale of the jams increases in ρ_0 and decreases in ρ_1 until the lanes become *decoupled* (i.e. no further lane-changes) with *different* density values. The stop-and-go waves propagate upstream (labelled J) as usual prior to decoupling, however density waves also travel downstream with the speed of traffic (labelled H), even after the lanes decouple. The reason for this is that for a given velocity, the equilibrium spacing is different for each vehicle population, hence the equilibrium density is not homogeneous. The final configuration sees 218 vehicles in lane 0 and 182 vehicles in lane 1, of which 78 are fast (and hence 140 are slow) in lane 0 and 122 are fast (and hence 60 are slow) in lane 1.

Summing up our experiments, by introducing instability and driver heterogeneity into our lane-changing model, we have observed a variety of spatio-temporal patterns that cannot be captured by our quasi-equilibrium assumption. When vehicles are identical, we observe that lane-changing causes the density (and thus velocity)

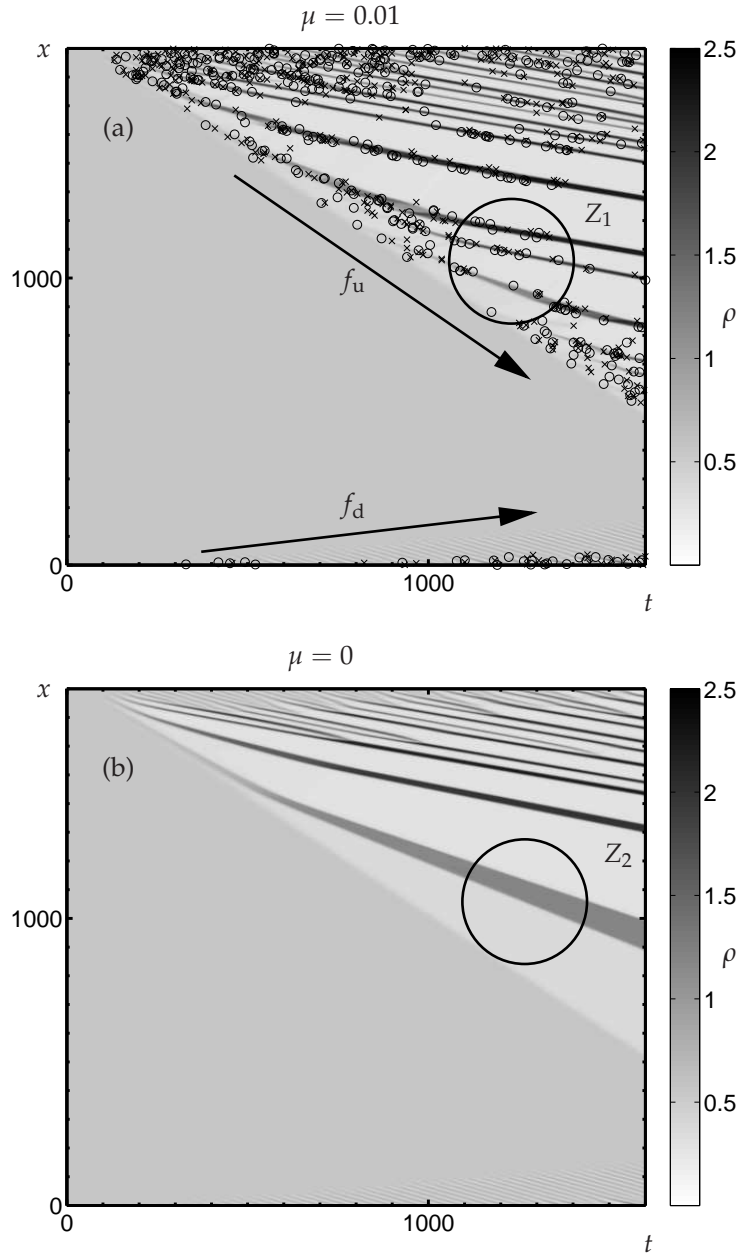


Figure 4.15. Evolution of a small perturbation in density in lane 0 with unstable parameter values $\alpha = 0.5, \beta = 0.25$. Lane-changing rate panel (a) occurs with probability $\mu = 0.01$ in unit time, but is not permitted in panel (b) (i.e $\mu = 0$). Position on the loop x is plotted against time t and the grey-scale corresponds to the coarse-grained densities. Lane-changing threshold parameters in panel (a) are $\sigma = 0.01$ and $\tau = 2.0$ and initial conditions are $h_0 = h_1 = 1.8$ and $n_0 = n_1 = 1000$. A small perturbation is introduced by setting the velocity of one vehicle in lane 0 to $1.1V(h_0)$. Crosses mark lane-changes from lane 0 to 1 and circles mark lane-changes from lane 1 to 0. The number of steps is chosen so that the leading and trailing edges of the resulting disturbance (f_d and f_u respectively) do not interact. Lane-changes appear to affect the long-term spatial development of individual jams (cf. Z_1 and Z_2).

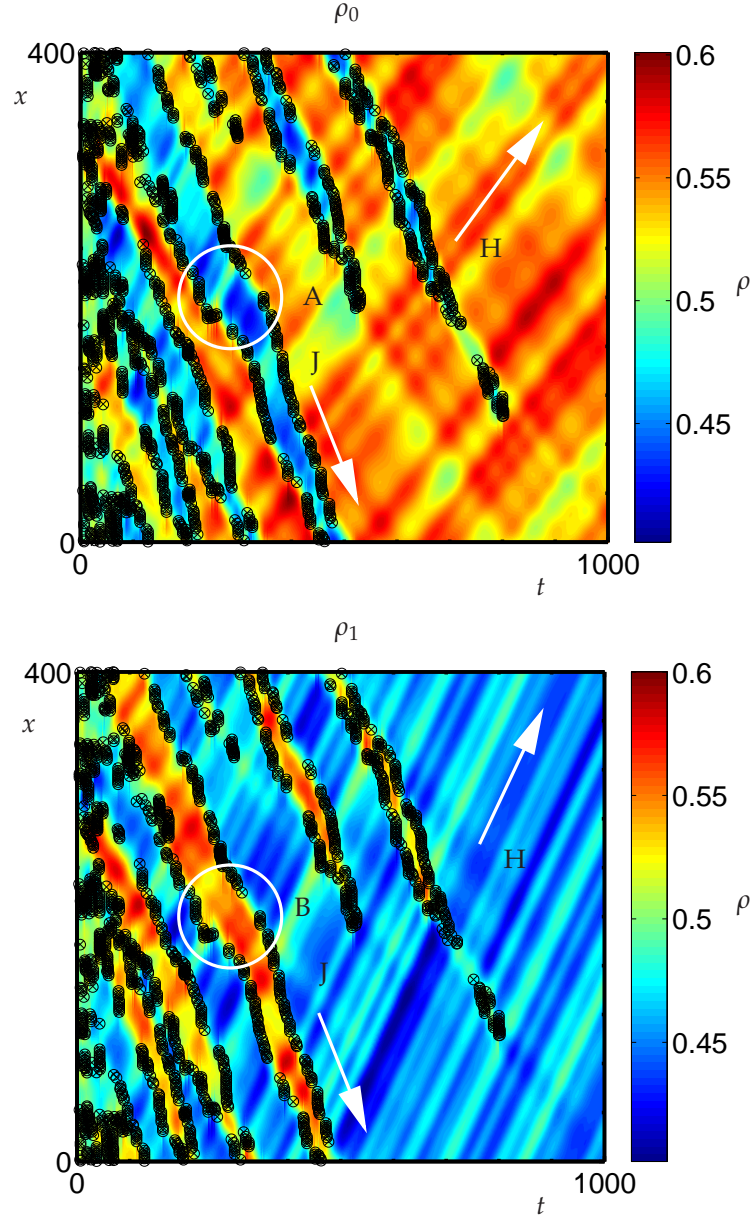


Figure 4.16. Evolution of traffic with heterogeneous driver behaviour. Position on the loop x is plotted against time t and the colour-scale corresponds to coarse-grained density for each lane, labelled ρ_0 and ρ_1 . Crosses mark lane-changes from lane 0 to 1 and circles mark lane-changes from lane 1 to 0. The vehicle population comprises equal numbers of slow ($\gamma = 0.6$) and fast ($\gamma = 1.0$) cars, which are initially distributed evenly between lanes but whose order is chosen randomly. The initial spacing is constant $h_* = 2$, as is the number of vehicles starting in each lane $n_0 = n_1 = 400$. Model parameters are $\alpha = 0.6$, $\beta = 0.3$, lane-changing thresholds are $\sigma = 0.1$, $\tau = 0.5$ and lane-changes are deterministic (i.e. $\mu = 1/\Delta t$). In contrast to Figures 4.14 and 4.15, we observe that the density profiles become *asynchronised* (compare areas A and B). In addition to stop-and-go waves propagating upstream (J), density waves due to heterogeneity also propagate *downstream* (H).

profiles to become spatially synchronised. Furthermore, lane-changing seems to trigger short length-scale jams, which do not grow in width. We also observe that driver heterogeneity can cause transient spatial asynchronisation, which leads to speed differential induced lane-decoupling.

4.10 Summary

In this chapter, we have presented a novel analysis of lane-changing in multilane car-following models. Our approach has been multiscale in that we used the microscopic MOBIL lane-changing criteria to derive a macroscopic mass exchange rate between two lanes of different density, based on the general car-following model (1.12) and the constraints (1.17). We considered a highly simplified setting, in which vehicles were identical and uniform flows were stable. Furthermore, we assumed that lane-changes occur when traffic is in quasi-equilibrium, i.e when the longitudinal relaxation timescale is shorter than the lane-changing timescale and hence lane-changes are rare. This assumption allowed us to determine a set of conditions on model parameters that ensure safe lane-changes. Moreover, we derived a spatially independent macroscopic lane-changing rate between two lanes of constant density from the microscopic lane-changing rules, which we found to be in good agreement with microscopic simulation. Finally, we illustrated how instability and driver heterogeneity introduces spatio-temporal pattern that is not captured by our quasi-equilibrium assumption.

We stress that our method is based on a highly simplified case, hence there is a need to extend our analysis to include features not captured, such as the variation of equilibrium density with lane-changing timescale. We believe that our multiscale approach is possibly the only way to develop a theory that can encapsulate the complex dynamics exhibited by multilane traffic models. Moreover, we imagine that the recent introduction of high resolution data [59] and cutting edge vehicle re-identification techniques [14] will prompt empirical calibration of lane-changing models and hence motivate a rapid theoretical development in the analysis of underpinning multilane models.

Chapter 5

Conclusions

“He used tax form that’s expired (9)”, Rufus 1(14d) [31]

5.1 Review of Thesis

In this chapter, we give a brief summary of the work presented in this thesis and describe a selection of open questions that have emerged during the course of study.

The topic of this thesis has been the study of spatio-temporal pattern formation mechanisms in highway traffic. Historically, the analysis of complex traffic patterns has focused on highly simplified settings such as spatially homogeneous single-lane roads with identical vehicles. It is this that has inspired our motivating philosophy: we have sought to relax such assumptions by considering the effect of lane-changing and heterogeneity in simple settings that have enabled us to develop new methods of analysis. Our approach has also been multiscale in that we have studied the dynamics of microscopic models and their simulation using macroscopic techniques that relate continuum variables such as traffic density, velocity and flow.

In Chapter 2, we analysed stationary wave structures found in simulations of the microscopic Optimal Velocity model with spatial heterogeneity. We considered only stable parameter regimes in which stop-and-go waves are impossible, yet we forced spatial pattern by introducing a crude capacity drop bottleneck. In fact we believe that the patterns observed are generic to this type of bottleneck. These consisted of sharp shock-like fronts separating constant density plateaus, which we observed in microscopic trajectory data using a numerical coarse-graining method. To analyse these stationary wave structures, we first considered simple conservation principles and characteristics arguments, which we used to map out transitions between two- and three-plateaus solutions in parameter space. The fact that such a simple approach works so well emphasises the wide range of models which will develop such patterns, since *all* single lane traffic models with no entrances or exits conserve the number of

vehicles. We then considered the internal front structure by appealing to second order continuum traffic models, derived from the OV model. This method gave qualitatively correct results, however we identified higher order effects that occur at the interfaces between high density plateaus. Thus we returned to the microscopic model and illustrated how additional eigenvalues in the spectrum can be used to explain the escape trajectories from high density plateaus. We stress that even in such a simple setting, introducing spatial heterogeneity resulted in rich traffic patterns that required sophisticated analysis techniques.

In Chapter 3, we described how the stability of a wide range of traffic models can be encapsulated via a simple condition on model parameters and moreover, how this can be extended to include heterogeneous driver behaviour. We also described absolute/convective instability, where the propagation of growing perturbations is respectively bi/uni-directional, and furthermore, how comparing the absolute group velocity (i.e. in the reference frame of the road) and the growth rates can be used to identify these states. We extended our perturbation analysis to fourth order to approximate the transition between absolute and convective instability in terms of the partial derivatives of the general car-following model and plotted a phase diagram in parameter space for the OVRV model. We then presented a novel technique in which we solved the linearised equations of motion progressively along a chain of vehicles and illustrated how transform methods may be used to shed light on the asymptotic nature of linear instabilities in microscopic traffic models.

In Chapter 4, we considered lane-changing in two-lanes of traffic. We described our versatile simulation framework which we used to test a number of lane-changing models. This led us to adopt the so called ‘MOBIL’ lane-changing framework [48], which uses accelerations computed from the longitudinal driver behaviour model to determine when changing lanes is both desirable and safe. We analysed the incentive and safety criteria of the MOBIL framework in a highly simplified setting in which vehicles are identical and lane-changes are rare. These assumptions allowed us to determine a set of criteria on model parameters that, when satisfied, result in safe lane-changing manoeuvres. Furthermore, in this setting we developed approximate *macroscopic* lane-changing rates and, using expansions, suggested how higher order terms might appear in continuum multilane models. We verified our approximations using our simulation software and found that they were in good agreement. Finally, we illustrated new phenomena, which to date have not been reported on in the traffic literature, that challenge our underlying assumptions.

5.2 Future Research Directions

We now summarise a selection of open problems for each of the chapters in this thesis.

5.2.1 Spatial Heterogeneity

In Chapter 2, we used spatial heterogeneity to induce spatio-temporal pattern. An interesting direction in which to develop this work would be to consider the macroscopic traffic patterns that emerge when a bottleneck is present and the system is unstable. For our simple capacity drop bottleneck, where the OV function is scaled linearly, there will be parameter ranges in which flows in the bottleneck are stable, yet flows in free-flow are unstable. It is possible that new stationary patterns may emerge which can be studied using the linear stability analysis presented in Chapter 2.

In fact, this leads to a much more serious evaluation of how the patterns observed in a range of microscopic bottleneck models fit with empirical data. This will ultimately help to develop a fundamental understanding of the underlying mechanisms behind many of the complex spatio-temporal patterns that have been observed and classified in data. In particular, features such as ‘Pinned Localised Clusters’, ‘Homogeneous Congested Traffic’ and ‘Oscillatory Congested Traffic’ identified by Schönhof and Helbing [73] seem to be triggered by combinations of bottlenecks and unstable high density traffic.

5.2.2 Instability

In the later parts of Chapter 3, we presented a selection of new analysis techniques which are still areas of current research. We expect that in the near future we will be able to identify the asymptotic behaviour of the solutions to the linearised equations of motion of the general car-following model (3.3), and thus conclude this area of research. A wider issue concerns the parameterisation of car-following models and confirming that the phenomena associated with stability and instability are consistent with empirical traffic patterns. In fact, we hope to classify a wide range of car-following models with the aim of highlighting the features that faithfully reproduce phenomena observed in empirical data and hence present a bench mark for car-following models developed in the future.

5.2.3 Lane-Changing

In Chapter 4 we presented a novel multiscale approach to analyse lane-changing in a highly simplified setting. This area of research is in its infancy, thus there are a wealth of interesting directions in which one might proceed. Our analysis was based on several simplifying assumptions, in particular, we only considered selfish lane changes where the parameter p was set to zero. Allowing drivers a degree of altruism not only affects our assumptions about the monotonicity of the incentive criteria, but leads to a subtle interplay between both the incentive and safety criteria. Multilane highways in the UK and Europe also have asymmetric passing rules and even in the US, where over- and undertaking is allowed, there is an inherent asymmetry introduced by the position of entrances and exits on the outermost lanes. Such asymmetries are known to cause the so called ‘lane inversion’ phenomena, in which the passing lanes have a higher density than the running lanes.

In contrast to the stable equilibrium solutions that emerge when vehicles are identical, simulation experiments with heterogeneous driver behaviour do not necessarily lead to such steady states and non-equilibrium phenomena may be observed. Furthermore, the application of periodic boundary conditions seems to play an important role in the traffic patterns observed in simulations over long periods of time. We have observed large scale jamming when vehicles with similar velocities pass one another. This effect needs further study in order to determine whether such phenomena are induced by periodic boundary conditions and how this affects the macroscopic properties of the system.

Another goal is an in-depth development of multiscale techniques in multilane problems with driver heterogeneity. In particular, macroscopic moving bottlenecks may be used to represent different classes of vehicles, which could be coupled with macroscopic lane-changing models. We would like then to tune such a macroscopic model to reflect the dynamics observed in microscopic simulation. Our goal would be to develop analysis techniques for the macroscopic model, which may be more tractable than attempting to directly describe the highly complex phenomena observed in microscopic multilane simulations with driver heterogeneity.

Finally, high resolution data is gradually becoming available for scientific use, from which large numbers of individual vehicle trajectories can be obtained. An important task in the future is a quantitative parametric evaluation exercise of lane-changing models using empirical data, for example in the neighbourhood of a merge where many lane changes are forced by the physical infrastructure. Ultimately this will lead to high-fidelity multilane traffic models and further theoretical developments.

5.3 Final Remarks

In summary, the work in this thesis has used a diverse array of advanced techniques at a range of scales to examine the dynamics of mathematical models of traffic with non-trivial features such as lane-changing and heterogeneity. These results carry serious weight in the mathematical traffic community since they are centred around areas of intense debate, which we intend to publish to a wider audience in the very near future. Further down the line, high resolution data will help to clarify many of the current issues, but there will still be a wealth of analysis problems arising from complex highway traffic models. We have illustrated this fact through each chapter in this thesis, and attempted to develop analytical techniques based on highly simplified settings. We believe that the multiscale philosophy presented in this thesis is the key to analysing the underlying spatio-temporal pattern formation mechanisms that arise from the vast number of unique individual interactions that occur perpetually on highways around the world.

The ultimate goal for mathematical traffic modellers is to describe the features present on real highways that challenge current modelling methodologies (such as lane-changing, spatial heterogeneity and distributed driver behaviour) via a fusion of cutting edge simulation and analysis techniques. The next generation of state of the art in traffic modelling will combine in-depth phenomenological knowledge of faithful models with new information/communication technology and ITS hardware, the benefits of which will be felt by both individual drivers and the economic prosperity of the nation as a whole.

Appendix A

Derivation of Multiclass Stability Analysis Identities

In the derivation of the stability criteria with heterogeneous driver behaviour, we use several formulae which we derive here.

Determinant Calculation. Consider a simple 3×3 matrix with the same form as (3.18),

$$\begin{aligned} \det(\mathbf{M}_3) \begin{vmatrix} x_{11} & 0 & x_{13} \\ x_{21} & x_{22} & 0 \\ 0 & x_{32} & x_{33} \end{vmatrix} &= x_{11} \begin{vmatrix} x_{22} & 0 \\ x_{32} & x_{33} \end{vmatrix} + x_{13} \begin{vmatrix} x_{21} & x_{22} \\ 0 & x_{32} \end{vmatrix}, \\ &= x_{11} \det(L_2) + x_{13} \det(U_2), \end{aligned} \quad (\text{A.1})$$

where L_2 is a 2×2 lower triangular matrix and U_2 is a 2×2 upper triangular matrix. For an $N \times N$ matrix \mathbf{M}_N of the same form, this becomes

$$\det(\mathbf{M}_N) = x_{11} \det(L_{N-1}) - (-1)^N x_{1(N-1)} \det(U_{N-1}). \quad (\text{A.2})$$

Thus the determinants of the triangular matrices L_{N-1} and U_{N-1} are simply the products of the main diagonal elements, hence we can write

$$\det(\mathbf{M}_N) = \prod_{n=1}^N x_{nn} - (-x_{1(N-1)}) \prod_{n=2}^N (-x_{n(n-1)}). \quad (\text{A.3})$$

Identities. There are three identities used. To obtain the first consider

$$\sum_n \mathcal{D}_h f_n \left[\prod_{m \neq n} \mathcal{D}_h f_m \right]^2 = \sum_n \prod_n \mathcal{D}_h f_n \prod_{m \neq n} \mathcal{D}_h f_m, \quad (\text{A.4})$$

$$\left[\prod_n \mathcal{D}_h f_n \right] \left[\sum_n \prod_{m \neq n} \mathcal{D}_h f_m \right], \quad (\text{A.5})$$

hence

$$\sum_n \prod_{m \neq n} D_h f_m = \frac{\sum_n D_h f_n \left[\prod_{m \neq n} D_h f_m \right]^2}{\prod_n D_h f_n}. \quad (\text{A.6})$$

To obtain the second and third identities consider

$$\begin{aligned} \left(\sum_n b_n \prod_{m \neq n} a_m \right) \left(\sum_n c_n \prod_{m \neq n} a_m \right) &= \sum_n b_n c_n \left(\prod_{m \neq n} a_m \right)^2 \\ &+ \sum_{i,j} b_i c_j \left(\prod_{m \neq i,j} a_m \right) \left(\prod_n a_n \right), \end{aligned} \quad (\text{A.7})$$

which in the case $b_n = c_n$ becomes

$$\begin{aligned} \left(\sum_n b_n \prod_{m \neq n} a_m \right)^2 &= \sum_n b_n^2 \left(\prod_{m \neq n} a_m \right)^2 \\ &+ 2 \sum_{i \neq j, j > i} b_i b_j \left(\prod_{m \neq i,j} a_m \right) \left(\prod_n a_n \right), \end{aligned} \quad (\text{A.8})$$

By substitution and rearranging (A.7) and (A.8), we obtain the second and third identities

$$\begin{aligned} \sum_{i \neq j} D_h f_i D_v f_j \prod_{m \neq i,j} D_h f_m &= \frac{1}{\prod_n D_h f_n} \left\{ \left(\sum_n D_h f_n \prod_{m \neq n} D_h f_m \right) \left(\sum_n D_v f \prod_{m \neq n} D_h f_m \right) \right. \\ &\quad \left. - \sum_n D_h f D_v f \left(\prod_{m \neq n} D_h f_m \right)^2 \right\}, \end{aligned} \quad (\text{A.9})$$

and

$$\begin{aligned} \sum_{i \neq j, j > i} D_v f_i D_v f_j \prod_{m \neq i,j} D_h f_m &= \frac{1}{2 \prod_n D_h f_n} \left\{ \left(\sum_n D_v f \prod_{m \neq n} D_h f_m \right)^2 \right. \\ &\quad \left. - \sum_n D_v f^2 \left(\prod_{m \neq n} D_h f_m \right)^2 \right\}, \end{aligned} \quad (\text{A.10})$$

respectively.

Appendix B

Fourier Analysis of the Linearised General Car-Following Model

We now briefly describe how one might use a multiscale approach to analyse the linearised general car-following model with temporal forcing,

$$\ddot{\tilde{h}}_n + (D_h f - D_v f) \dot{\tilde{h}}_n + (D_h f) \tilde{h}_n - (D_h f) \dot{\tilde{h}}_{n-1} - (D_h f) \tilde{h}_{n-1} = f_n(t). \quad (\text{B.1})$$

Homogenising via the identity $\tilde{h}_{n-1}(t) = \exp(-\partial_n) \tilde{h}(n, t)$, yields

$$\tilde{h}_{tt}(n, t) + [D_h f (1 - e^{-\partial_n}) - D_v f] \tilde{h}_t(n, t) + D_h f (1 - e^{-\partial_n}) \tilde{h}(n, t) = f(n, t), \quad (\text{B.2})$$

for which we are interested in the Green's function which satisfies

$$G_{tt}(n, t) + [D_h f (1 - e^{-\partial_n}) - D_v f] G_t(n, t) + D_h f (1 - e^{-\partial_n}) G(n, t) = \delta(n) \delta(t). \quad (\text{B.3})$$

The Green's function can be written as a Fourier superposition

$$G(n, t) = \frac{1}{(2\pi)^2} \int_{L_\lambda} \int_{F_\theta} G(\theta, \lambda) \exp(i\theta n + \lambda t) d\theta d\lambda. \quad (\text{B.4})$$

We now employ a series of standard Fourier analysis techniques, which are described in detail in [41].

We define the operator

$$D(-i\partial/\partial n, \partial/\partial t; R) = \partial_t^2 + [D_h f \exp(-\partial_n) - D_v f] \partial_t + D_h f [1 - \exp(-\partial_n)], \quad (\text{B.5})$$

where $R = \{D_h f, D_h f, D_v f\}$. Equation (B.2) then reads

$$D(-i\partial/\partial n, \partial/\partial t; R) G(n, t) = \delta(n) \delta(t). \quad (\text{B.6})$$

This translates in Fourier space to $D(\theta, \lambda) G(\theta, \lambda) = 1$. The formal solution to (B.3) is

then

$$G(n, t) = \frac{1}{(2\pi)^2} \int_{L_\lambda} \int_{F_\theta} \frac{1}{D(\theta, \lambda)} \exp(i\theta n + \lambda t) d\theta d\lambda. \quad (\text{B.7})$$

The inverse transform

$$G(n, \lambda) = \frac{1}{2\pi} \int_{L_\lambda} \frac{1}{D(\theta, \lambda)} \exp(\lambda t) d\lambda, \quad (\text{B.8})$$

can be determined by considering the residue at the poles associated with the zeros of the dispersion relation $D(\theta, \lambda) = 0$. The dispersion relation (3.4) is quadratic in λ and there are therefore two simple poles. The residues are $\mp 1/(\lambda_+ - \lambda_-)$, where λ_\pm are defined in (3.5). Thus, the Green's function (B.7) becomes

$$G(n, t) = -\frac{i}{2\pi} \sum_{\pm} \int_{F_\theta} \frac{\exp(i\theta n)}{\partial D / \partial \lambda(\theta, \lambda_\pm(\theta))} \exp(\lambda_\pm(\theta) t) d\theta. \quad (\text{B.9})$$

This form motivates the use of asymptotic methods. Note that the change of variables $z = 1 - \exp(-i\theta)$ transforms the F_θ contour into a unit circle centred at $z = 1$. We denote the new temporal modes

$$\lambda_\pm(k) = \hat{\lambda}_\pm(z) = \frac{1}{2} \left\{ -[D_h f z - D_v f] \pm \sqrt{[D_h f z - D_v f]^2 - 4D_h f z} \right\}, \quad (\text{B.10})$$

and then (B.9) becomes

$$G(n, t) = -\frac{i}{2\pi} \sum_{\pm} \int_{F_z} \frac{(1-z)^{-n}}{\pm \sqrt{(D_h f z - D_v f)^2 - 4D_h f z}} \exp(\hat{\lambda}_\pm(z) t) dz. \quad (\text{B.11})$$

Since the integration contour is now a closed curve, $G(n, t)$ can be determined exactly from the contributions at the poles of

$$f(z) = \frac{(1-z)^{-n} \exp(\hat{\lambda}_\pm(z) t)}{\pm \sqrt{(D_h f z - D_v f)^2 - 4D_h f z}}, \quad (\text{B.12})$$

This work is on-going, although we believe that the contour of integration must be deformed so that the branch cuts due to the square roots are not crossed. Care must be taken when doing this because the value of z differs on either side of the branch cut. Furthermore, calculating the residue at $z = 1$ proves difficult because we must repeatedly differentiate $f(z)(1-z)^n$ with respect to z .

References

- [1] Highways Agency. <http://www.highways.gov.uk/>.
- [2] Highways Agency. Atm Monitoring and Evaluation, 2008. <http://www.dft.gov.uk/pgr/roads/tpm/m42activetrafficmanagement/>.
- [3] Highways Agency. Highways Agency Business Plan, 2008. <http://www.highways.gov.uk/aboutus/18403.aspx>.
- [4] S. Ahn and M.J. Cassidy. Freeway traffic oscillations and vehicle lane-change maneuvers. In R.E. Allsop, M.G.H. Bell, and B.G. Heydecker, editors, *Transportation and traffic theory 2007*, pages 691–710. Elsevier Ltd., 2007.
- [5] A. Aw and M. Rascole. Resurrection of "second order" models of traffic flow. *Siam Journal of Applied Mathematics*, 60(3):916–938, 2000.
- [6] M. Bando, K. Hasebe, A. Nakayama, A. Shibata, and Y. Sugiyama. Dynamical model of traffic congestion and numerical simulation. *Phys. Rev. E*, 51(2):1035–1042, 1995.
- [7] J. H. Banks. Flow processes at a freeway bottleneck. *Transportation Research Record*, 1287:20–28, 1990.
- [8] P. Berg, A. Mason, and A. Woods. Continuum approach to car-following models. *Phys. Rev. E*, 61(2):1056–1066, 2000.
- [9] R. L. Bertini and M. J. Cassidy. Some observed queue discharge features at a freeway bottleneck downstream of a merge. *Transportation Research Part A*, 36:683–697, 2002.
- [10] M. Brackstone and M. McDonald. Car-following: a historical review. *Transportation Research Part F*, 2(4):181–196, 1999.
- [11] M. J. Cassidy and R. L. Bertini. Some traffic features at freeway bottlenecks. *Transportation Research Part B*, 33:25–42, 1999.
- [12] R.E. Chandler, R. Herman, and E.W. Montroll. Traffic dynamics: Studies in car following. *Operations Research*, 6(2):165–184, 1958.

- [13] D. Chowdhury, L. Santen, and Andreas Schadschneider. Statistical physics of vehicular traffic and some related systems. *Physics Reports*, 329:199–329, 2000.
- [14] B. Coifman and M. Cassidy. Vehicle reidentification and travel time measurement on congested freeways. *Transportation Research Part A*, 36(10):899–917, 2002.
- [15] K. L. Cooke and P. van den Driessche. On zeroes of some transcendental equations. *Funkcialaj Ekvacioj*, 29:77–90, 1986.
- [16] C. Daganzo. Requiem for second-order fluid approximations to traffic flow. *Transportation Research Part B*, 29(4):277–286, 1994.
- [17] C. F. Daganzo and J. A. Laval. On the numerical treatment of moving bottlenecks. *Transportation Research Part B*, 39:31–46, 2004.
- [18] C.F. Daganzo, M.J. Cassidy, and R.L. Bertini. Possible explanations of phase transitions in highway traffic. *Transportation Research A*, 33(5):365–379, 1998.
- [19] L. C. Davis. Multilane simulations of traffic phases. *Phys. Rev. E.*, 69(016108):1–5, 2004.
- [20] R. Eddington. The Eddington Transport Study, 2006.
- [21] Department for Transport. Understanding the benefits and costs of intelligent transport systems - a toolkit approach. <http://www.dft.gov.uk/itstoolkit/>.
- [22] Department for Transport. The future of transport, 2004. <http://www.dft.gov.uk/about/strategy/whitepapers/fot/>.
- [23] Department for Transport and National Statistics. Transport statistics great britain, 2007. <http://www.dft.gov.uk/pgr/statistics/>.
- [24] I. Gasser, G. Sirito, and Werner B. Bifurcation analysis of a class of ‘car following’ traffic models. *Physica D*, 197(3-4):277–296, 2004.
- [25] D. C. Gazis, R. Hermann, and G. H. Weiss. Density oscillations between lanes of a multilane highway. *Operations Research*, 10(5):658–667, 1962.
- [26] D.C. Gazis, R. Herman, and Rothery R.W. Nonlinear follow-the-leader models of traffic flow. *Operations Research*, 9(4):545–567, 1961.
- [27] P. G. Gipps. A model for the structure of lane-changing decisions. *Transportation Research Part B*, 20(5):403–414, 1985.

References

- [28] M. Goldbach, A. Eidmann, and A. Kittel. Simulation of multilane freeway traffic with detailed rules deduced from microscopic driving behavior. *Phys. Rev. E*, 61(2):1239–1246, 2000.
- [29] J.M. Greenberg. Extensions and amplifications of a traffic model of aw and rascle. *SIAM Journal of Applied Mathematics*, 62(3):729–745, 2002.
- [30] B.D. Greenshields. A study of traffic capacity. *Proceedings of the Highway Research Board*, 14:448–477, 1934.
- [31] The Guardian. Cryptic Crosswords Volume 1, 2003. See below for answers*.
- [32] H. Hanaura, T. Nagatani, and K. Tanaka. Jam formation in traffic flow on a highway with some slowdown sections. *Physica A*, 374:419–430, 2007.
- [33] D. Helbing. Traffic and related self-driven many-particle systems. *Rev. Mod. Phys.*, 73(4):1067–1141, 2001.
- [34] R. Herman, E.W. Montroll, R.B. Potts, and R.W. Rothery. Traffic dynamics: Analysis of stability in car following. *Operations Research*, 7(1):86–106, 1959.
- [35] P. Hidas. Modelling vehicle interactions in microscopic simulation of merging and weaving. *Transportation Research Part C*, 13(1):37–62, 2005.
- [36] Payne H.J. Freflo: A macroscopic simulation model for freeway traffic. *Transportation Research Record*, 722:68–77, 1979.
- [37] E. N. Holland and A. W. Woods. A continuum model for the dispersion of traffic on two-lane roads. *Transportation Research B*, 31:473–485, 1997.
- [38] E.N. Holland. A generalised stability criterion for motorway traffic. *Transportation Research Part B*, 32(2):141–154, 1997.
- [39] S. P. Hoogendoorn and P. H. L. Bovy. Platoon-based multiclass modeling of multilane traffic flow. *Networks and Spatial Economics*, 1(1):137–166, 2001.
- [40] D. W. Huang. Lane-changing behavior on highways. *Phys. Rev. E*, 66(026124):1–5, 2002.
- [41] P. Huerre and M. Rossi. *Hydrodynamics and Nonlinear Instabilities*, chapter Hydrodynamic instabilities in open flows. Cambridge University Press, 1998.
- [42] W. L. Jin and H. M. Zhang. The inhomogeneous kinematic wave traffic flow model as a resonant nonlinear system. *Trans. Sci.*, 37(3):294–311, 2003.

* Ideal, Editor, Spectrum, Fingerprint, Exhausted

- [43] B. S. Kerner. Empirical macroscopic features of spatial-temporal traffic patterns at highway bottlenecks. *Phys. Rev. E*, 65:1–30, 2002. Art. No. 046138.
- [44] B. S. Kerner. *The Physics of Traffic*. Springer-Verlag, 2004.
- [45] B. S. Kerner and S. L. Klenov. Microscopic theory of spatial-temporal congested traffic patterns at highway bottlenecks. *Phys. Rev. E*, 68:1–19, 2003. Art. No. 036130.
- [46] B. S. Kerner and P. Konhäuser. Cluster effect in initially homogeneous traffic flow. *Phys. Rev. E*, 48(4):R2335–R2338, 1993.
- [47] B.S. Kerner and H. Rehborn. Experimental properties of phase transitions in traffic flow. *Physical Review Letters*, 79(20):4030–4033, 1997.
- [48] A. Kesting, M. Treiber, and D. Helbing. General lane-changing model MOBIL for car-following models. *Transportation Research Record*, 1999:86–94, 2007.
- [49] P. G. Kevrekidis, I. G. Kevrekidis, A. R. Bishop, and E. S. Titi. Continuum approach to discreteness. *Phys. Rev. E*, 65(4):046613, Apr 2002.
- [50] A. Klar and R. Wegener. A hierarchy of models for multilane vehicular traffic i: Modeling. *SIAM J. Appl. Math.*, 59(3):983–1001, 1999.
- [51] S. Kurata and T. Nagatani. Spatio-temporal dynamics of jams in two-lane traffic flow with a blockage. *Physica A*, 318(3–4):537–550, 2003.
- [52] R.D. Kuühne. Macroscopic freeway model for dense traffic - stop-start waves and incident detection. In J. Volmuller and Hamerslag R., editors, *Proceedings of the ninth international symposium on transportation and traffic theory*, pages 21–42. VNU Science Press, Utrecht, Netherlands, 1984.
- [53] Y. A. Kuznetsov. *Elements of Applied Bifurcation Theory*. Springer-Verlag, 1995.
- [54] J. A. Laval and C. F. Daganzo. Lane-changing in traffic streams. *Transportation Research Part B*, 40(3):251–264, 2006.
- [55] L. Leclercq, S. Chanut, and J. B. Lesort. Moving bottlenecks in Lighthill-Whitham-Richards model: A unified theory. *Transportation Research Record*, 1883:3–13, 2004.
- [56] H. K. Lee, H. W. Lee, and D. Kim. Macroscopic traffic models from microscopic car-following models. *Phys. Rev. E*, 64(5):056126, 2001.
- [57] R.J. LeVeque. *Finite volume methods for hyprebolic problems*. Cambridge University Press, 2002.

References

- [58] M. Lighthill and G. B. Whitham. On kinematic waves ii. a theory of traffic on tong crowded roads. *Proc. Roy. Soc. London, Ser. A*, 229:317–345, 1955.
- [59] G. Lunt, R. E. Wilson, and M. Day. Enhancing motorway traffic data with novel vehicle re-identification algorithms. In *13th ITS World Congress*, 2006.
- [60] A. D. Mason and A. W. Woods. Car-following model of multispecies systems of road traffic. *Physical Review E*, 55(3):2203–2214, 1997.
- [61] N. Mitarai and H. Nakanishi. Convective instability and structure formation in traffic flow. *Journal of the Physical Society of Japan*, 69:3752–3761, 2000.
- [62] P. K. Munjal and L. A. Pipes. Propagation of on-ramp density perturbations on two- and three-lane freeways. *Transportation Research*, 5:241–255, 1971.
- [63] Takashi Nagatani. The physics of traffic jams. *Reports on Progress in Physics*, 65(9):1331–1386, 2002.
- [64] K. Nagel and M. Schreckenberg. A cellular automaton model for freeway traffic. *Journal de Physique*, 12(2):2221–2229, 1992.
- [65] G. F. Newell. A simplified theory of kinematic waves in highway traffic, part I: General theory. *Transportation Research Part B*, 27B:281–287, 1993.
- [66] G. F. Newell. A simplified theory of kinematic waves in highway traffic, part II: Queuing at freeway bottlenecks. *Transportation Research Part B*, 27B:289–303, 1993.
- [67] G. F. Newell. A moving bottleneck. *Transportation Research Part B*, 32:531–537, 1998.
- [68] Confederation of British Industry. The uk as a place to do business: Is transport holding us back? 2003.
- [69] G. Orosz, R. E. Wilson, and B. Krauskopf. Global bifurcation investigation of an optimal velocity traffic model with driver reaction time. *Physical Review E*, 70(2):026207, 2004.
- [70] P. Richards. Shockwaves on the highway. *Operations Research*, 4(1):42–51, 1956.
- [71] W. Rudin. *Real and Complex Analysis*. McGraw-Hill, 1960.
- [72] B. Sandstede and A. Scheel. Absolute and convective instabilities of waves on unbounded and large bounded domains. *Physica D*, 145:233–277, 2000.
- [73] M. Schönhof and D. Helbing. Empirical features of congested traffic states and their implications for traffic modeling. *Transportation Science*, 41(2):135–166, 2007.

- [74] V. Shvetsov and Helbing D. Macroscopic dynamics of multilane traffic. *Physical Review E*, 59(6):6328–6339, 1999.
- [75] Y. Sugiyama, M. Fukui, M. Kikuchi, K. Hasebe, A. Nakayama, K. Nishinari, S. Tadaki, and S. Yukawa. Traffic jams without bottlenecks—experimental evidence for the physical mechanism of the formation of a jam. *New Journal of Physics*, 10(033001), 2008. Movies: <http://www.iop.org/EJ/mmedia/1367-2630/10/3/033001/>.
- [76] T. Tang, H. Huang, and S. C. Wong. Lane changing analysis for two-lane traffic flow. *Acta Mech. Sin.*, 23(1):49–54, 2007.
- [77] T. Toledo, C.F. Chowdhury, and M.E. Ben-Akiva. Lane-changing model with explicit target lane choice. *Transportation Research Record: Journal of Transportation Research Board*, (1934):157–165, 2005.
- [78] M. Treiber and Helbing D. Reconstructing the spatio-temporal traffic dynamics from stationary detector data. *Cooperative Transportation Dynamics*, 1:3.1–3.24, 2002.
- [79] M. Treiber, A. Hennecke, and D. Helbing. Congested traffic states in empirical observations and microscopic simulations. *Phys. Rev. E*, 62(2):1805–1824, 2000.
- [80] J. Ward, R. E. Wilson, and Berg P. Wave selection problems in the presence of a bottleneck. In A. Schadschneider, T. Pöschel, R. Kühne, M. Schreckenberg, and D.E. Wolf, editors, *Traffic and Granular Flow '05*, pages 565–576. Springer-Verlag, 2007.
- [81] J. A. Ward, R. E. Wilson, and Berg P. Multiscale analysis of a spatially heterogeneous microscopic traffic model. *Physica D*, 236:1–12, 2007.
- [82] G.B. Whitham. *Linear and Nonlinear Waves*. John Wiley and Sons, New York, 1974.
- [83] R. E. Wilson. Mechanisms for spatiotemporal pattern formation in highway traffic models. *Philos. Transact. A Math. Phys. Eng. Sci.*, 366:2017–2032, 2008.
- [84] R. E. Wilson and P. Berg. Existence and classification of travelling wave solutions to second order highway traffic models. In M. Fukui, Y. Sugiyama, M. Schreckenberg, and D.E. Wolf, editors, *Traffic and Granular Flow '01*, pages 85–90. Springer-Verlag, 2003.
- [85] H.M. Zhang. A non-equilibrium traffic model devoid of gas-like behaviour. *Transportation Research Part B*, 36(3):275–290, 2002.