School of Physics and Astronomy

Group Project
Traffic Flow in Vehicular Transport

Baxter Claire, Clegg Charlie, Kennedy Ross, Mutny Mojmir, Skea Christiana

September 4, 2018

Abstract

Under proposal for a UK Research Institute into traffic flow modelling, an evaluation of funding of models in vehicular traffic is presented. After inspection of the literature, we conclude that research focusing on discrete models would have a large impact on traffic flow modelling if funding were to be increased.

Declaration

We declare that this project and report is my own work.

Supervisor: Prof. Martin Evans   Date: September 4, 2018

Contents

1 Introduction

Traffic flow is a ubiquitous part of modern life due to the rapid expansion of vehicular traffic in the last century. It seems natural to strive for a better understanding of traffic flow systems, with a view to improving road network infrastructure. Traffic flow can also be thought of, in more general terms, as the structured flow of particles or information and, as such, can be found in many fields outwith vehicular traffic. For example, vesicle transport in biology [1, 2]. However, in this report, the focus exclusively lies in vehicular traffic flow modelling. A number of methods have been devised to model the evolution of a traffic flow system. The main models are summarised and evaluated in terms of their validity and potential for applications.

Lighthill and Whitham first modelled vehicular traffic in 1955 [3], which formed the basis for modern continuum models. The concept of fluid dynamics is used to model a collection of vehicles as one continuous fluid, and predicts the evolution of the system as a whole, not individual particles. The system is described by differential equations in both space and time. These models use locally averaged values, such as flow and density, to describe the system. Including the intricacies of vehicular traffic creates a very complicated problem, which in most cases is insolvable. However, in 1971, Prigogine developed a more simplistic kinetic theory of traffic flow [4]. As before, the vehicles are seen collectively as one system but here they are treated as a weakly interacting gas. The state of the system is described by a probability distribution of position and velocity. Similar to the fluid dynamics model, the particles are not identified individually.

Modern technological advances in computational power have changed the landscape for traffic flow modelling. Models have been created such that they naturally lend themselves to computational simulation. In 1992 Nagel-Schreckenberg developed a discretised model [5], which is still used as a foundation for models in practice today. This model describes the system through each individual particle's position and velocity, and evolves the system in discrete timesteps.

1.1 Outline

In Section 2, motivations behind studying traffic flow are outlined. Subsequently, an overview of the approaches used in traffic flow modelling are presented in Sections 3, 4 and 5. An evaluation of each model will be presented with a viewpoint of practical applications and possible developments. In Section 6, potential future problems in traffic flow are presented and their relation to the reviewed traffic models discussed. Section 7 briefly introduces other important methods used in traffic flow optimisation. Finally, in Section 8 a judgement is made as to whether increased funding in these areas will be of financial and social benefit.

2 Motivations for Studying Traffic Flow

2.1 Economic Motivation

Whether it's the time spent on a journey or the money spent on petrol, tolls and congestion charges: reducing these costs is of clear economic benefit [6]. The more that is known about the behaviour of traffic flow, particularly in areas of congestion, the more can be done to reduce the cost of a journey.

In London there is a considerable congestion problem and many attempts have been made to try and streamline traffic. Traffic models are used to provide valuable data and insights on the possible effects of any proposed changes. For example, in 2007 GLA Economics commissioned a study to see if reducing the number of traffic lights could decrease congestion [6]. A theoretical junction was modelled and simulations run using cross roads with directional priority instead of traffic lights. This procedure highlights the benefits of using traffic models: data can be collected about new road configurations without the cost and disruption of implementing them in real road environments.

It can also be a cost to city councils and governments in general if a road network is designed that is not functional and so requires time consuming and expensive upgrades. The modelling of traffic flow can therefore be a cost effective way to develop new networking systems and help effectively improve existing ones.

The ability to predict the nature of traffic flow can be a huge advantage for city planning and development. For example, when road tolls were abolished in Scotland in 2007, due to a change in legislation, the Scottish government used traffic flow models to assess the impact of removing tolls from the Forth and Tay road bridges [7]. By modelling local traffic in the immediate area surrounding the bridges, changes in congestion patterns could be predicted and their effect managed before the legislation took effect, and tolls were eradicated.

The ability to model a potential road layout before construction has more than just economic benefits. By reducing the likelihood of queuing traffic, not only is the journey time and costs reduced but so is the effect on the environment. Pollutants emitted from vehicles are in highest concentrations when they are accelerating. If the traffic can be kept flowing steadily the environmental impact can be minimised.

2.2 Academic Motivation

The advent of automobiles and highways during the last 70 years caused an interesting field of science to emerge - traffic flow modelling. This research area aims to describe the rich dynamics that occurs inside and outside cities on road networks. Traffic flow research historically derives from statistical physics as these techniques were traditionally used to analyse the systems [8]. To this day traffic flow is still considered a part of statistical physics, and many renowned physics academic journals publish work regarding traffic flow modelling, such as Physica A [9].

A physicist's motivation for studying traffic flow system differs from that of an engineer's, although their ultimate goal tends to be the same - understanding the intricacies of traffic flow. Physicists usually focus on a minimalistic system that reproduces the important dynamics of traffic such as velocity profiles, collision frequency, and jam transitions. Of special interest to statistical physicists is the concept of phase transitions which occur often in traffic, namely the transition from free flowing state to a jammed state. The search for an elegant, realistic and reliable model that would replicate these transitions is the pinnacle of the current statistical physics research in traffic flow models.

Although the motivation and initial analysis of many models in traffic flow comes from vehicular traffic, many of the concepts found a role in other seemingly unrelated fields such as financial market modelling [10] and cell biology [11]. Therefore, as it is common in scientific disciplines, the concepts of traffic models transcend the original idea and the abstract concepts can be used elsewhere. Traffic flow remains a very interesting and prolific research area with many branches.

3 Kinetic Theory of Traffic Flow

A conceptually simple way of modelling traffic flow is to treat the traffic system like a weakly interacting gas. Analogous to the kinetic theory of gases, vehicles are modelled in a similar way to gas particles using a time varying function of position and momentum through the use of a Boltzmann-like equation, the classical form of which is shown in Equation (1). However, there are some distinct differences between the treatment of gas particles and vehicles. For example, particles do not have directional preferences, whereas vehicles have a tendency to react to the environment in front of them more than behind [12]. The classical Boltzmann equation describes the time evolution of a distribution function as:

$$\frac{\partial f}{\partial t} + v\frac{\partial f}{\partial x} = Q. \tag{1}$$

In this equation v is the velocity of particles and f is a distribution function dependent on position, velocity and time. The term of the right hand side, Q, is called the Boltzmann collision operator and describes the interaction between particles when they collide [13]. In the absence of collisions, Q is zero and Equation (1) reduces to a continuity equation on which the discussed models are based [14]. Modelling these continuous functions can be computationally expensive, even with the advent of faster and more powerful machines. This expense can put a real restraint on the resolution and practicality of kinetic models.

Many approaches in this field use the concept of a desired state: a state of flow that the system tends to in equilibrium. Apart from interactions with other vehicles, which are treated as instantaneous, deviations from the idealised state can include adverse weather conditions, poor road conditions, etc.

The remainder of this section will discuss two established models for the kinetic theory of traffic flow, the first by Prigogine et al. in 1971 [4] followed by a development by Paveri-Fontana in 1975 [14]. Both the models make the assumption that each point particle (car) is identical.

3.1 Prigogine Model

As mentioned above, a principle model in the kinetic theory was developed by Prigogine et al. [4]. This was one of the first attempts to model a Boltzmann-like theory of traffic and remains a significant reference for activity in the field receiving 18 citations in 2009, almost 35 years after its first publication [15]. This theory uses the idea of a desired distribution function $f_{des}(x, v)$, towards which the system will tend in the absence of interactions. It is assumed that this desired function does not change. To define the distribution of the vehicles in the model let $f(x, v, t)dxdv$ be the number of vehicles in the range $x \to x + dx$ with velocity $v \to v + dv$, much like the Boltzmann distribution in the theory of gases. Prigogine proposes an expression for the time evolution of the function $f(x, v, t)$ in a form similar to the Boltzmann equation shown above. It is of the form:

$$\frac{\partial f}{\partial t} + v\frac{\partial f}{\partial x} = Rel + Int. \tag{2}$$

Where Rel and Int stand for the relaxation term and interaction term respectively. In the following sections the assumptions made to gain explicit forms of these two terms are discussed.

3.1.1 Relaxation Term

The relaxation term models the temporal change in the distribution function towards its desired distribution f_{des} in the absence of vehicle - vehicle interactions. To create an explicit expression for this term Prigogine makes two main assumptions. Firstly that the relaxation process only depends on a single relaxation time τ_{rel}. This can be interpreted as the time taken for a vehicle to return to the desired state after an interaction. Therefore it can be written that:

$$Rel = -\frac{f - f_{des}}{\tau_{rel}}. \tag{3}$$

The second assumption made is that, since the desired distribution function does not change, the distribution of the desired speeds is independent of the concentration of vehicles, $c(x, t)$, where $c = \int f(x, v, t)dv$. As a consequence, Prigogine express this as:

$$f_{des}(x, v, t) = c(x, t)F_{des}(v). \tag{4}$$

Where, as shown, the desired distribution is a function of two independent terms: the concentration and the desired velocity (v_{des}) distribution. Combining these assumptions, Equation (3) can be re-expressed as:

$$Rel = -\frac{f(x, v, t) - c(x, t)F_{des}(v)}{\tau_{rel}}. \tag{5}$$

This is Prigogine's expression of the relaxation of the system towards the desired state. It shows that the further the current state of the system is from the desired distribution, the greater the rate of relaxation towards the desired distribution (state).

3.1.2 Interaction Term

Looking back at Equation (2), the second term on the right hand side describes the interaction between vehicles. When a faster moving car approaches a slower car from behind there are two possible scenarios:

1. The faster car can overtake (velocity unchanged), with probability p
2. The faster car will slow down, with probability $(1-p)$

In both situations the velocity of the slower car is unaffected. In case 2, the change in velocity of the approaching car is taken to be instantaneous [14]. Since, in case 1, neither car changes velocity only case 2 will contribute to the interaction term. As mentioned above, the model is limited as the interaction term only considers two car exchanges and does not facilitate multi-vehicle interactions. It also relies on the assumption that the two vehicles are not interrelated and therefore can be represented as two independent functions: $F(x', v', x, v, t) = f(x, v, t)f(x', v', t)$ [8]. If the faster car slows on approach, with probability $(1-p)$, its one vehicle distribution function is affected. At position $x' = x$ the change in velocity experienced is $(v' - v)$, where v' is the velocity of the faster car. The interaction term is then defined by integrating over all possible approach velocities within the system. This has form:

$$Int = f(x, v, t) \int_0^\infty dv'(1-p)(v'-v)f(x, v', t). \tag{6}$$

3.1.3 Limitations of Prigogine Model

There have been many questions raised about the validity of this model. The first assumption that, while overtaking, the vehicle's speed remains constant does not generally hold in real-life. For example, the overtaking vehicle may slow upon approach or accelerate while passing. The assumptions that both the length scales of the vehicles and the time taken to change speed are negligible may initially appear unreasonable. However, as long as they are significantly smaller than the characteristic time and length scales of the system these approximations are valid [14].

The main concern in this model lies in the assumption leading to Equation (4). Prigogine claims that the distribution of the desired velocities, $F(v_{des})$, is independent of the concentration due to its time invariance [14]. Paveri-Fontana showed this claim to be false by examining the situation in which all vehicles are travelling at their desired velocity [14] - the proof can be found in Appendix A. This failing calls into question the validity of the desired distribution function used, as it has been shown to be time dependent. As a result the vehicles will be tending towards a unrepresentative state and so the use of this model is questionable. Paveri-Fontana proposed an Improved Boltzmann Equation which he claims overcomes this issue [14].

Prigogine also claims that the two vehicle problem has separable solutions such that that $F(x', v', x, v, t) = f(x, v, t)f(x', v', t)$. However, Paveri-Fontana highlights the scenario in which vehicles are queuing. In this situation, a very high correlation between the vehicles positions and speeds is implied. For the separable assumption to be valid, the operation of this model (and indeed most models in kinetic theory) is only valid for dilute traffic conditions. This constraint heavily limits the use of these models in practice.

3.2 Paveri-Fontana Model

In order to try correct for some of the drawbacks of the Prigogine model, such as the correlation between vehicles while queuing, Fontana adapted the model to consider the vehicles as partially independent. It was argued that, instead of the system collectively tending towards a desired distribution, each individual vehicle has a desired velocity, v_{des}, towards which it relaxes in the absence of interactions. It is assumed that the desired velocities do not change with time i.e. that each driver does not change their plans on what speed they would like to be travelling at.

The improved model introduces v_{des} (the desired velocity) as a dependent variable in the one-vehicle distribution function: $g(x, v, t; v_{des})$. Fontana models the time evolution of the system in the same manner as Prigogine:

$$\frac{\partial g}{\partial t} + v\frac{\partial g}{\partial x} = Rel + Int. \tag{7}$$

Although this expression is of the same form as that proposed by Prigogine, each vehicle's desired velocity now effects its evolution in the system. This is a progression towards a more realistic representation as it is more reasonable that each individual driver will have a desired speed rather than the system as a whole tending to a velocity distribution.

3.2.1 Relaxation Term

Using similar assumptions to Prigogine, the relaxation process depends on a single relaxation time (τ), Fontana expresses the relaxation term as:

$$Rel = \left(\frac{\partial g}{\partial t}\right)_{rel} = \frac{\partial}{\partial v}\left(\frac{v - v_{des}}{\tau} g\right) \tag{8}$$

using $\frac{d}{dt} = \frac{d}{dv}(\frac{dv}{dt})$. It is assumed that all accelerations occur over a unique time constant τ [14]. In practice this will not necessarily be the case.

3.2.2 Interaction Term

To define the interaction term the same approach is taken as that under Prigogine's regime. However, due to the redefinition of the one vehicle distribution function, it results in a more complicated expression.

The form of the Prigogine and Paveri-Fontana models are remarkably similar with the main difference residing in the definition of $g(x, v, t; v_{des})$ and the mathematical complication resulting from this difference. However, importantly, Fontana's model does not make the assumption that $f_{rel}(x, v, t) = c(x, t)F_{rel}(v)$ and so is considered, in his own words, as the Improved Boltzmann Equation [14].

The Paveri-Fontana model overcomes some of the conceptual issues found in Prigogine's. However, real difficulty arises if analytical solutions are sought in situations where the interaction term is important and cannot be neglected. Fontana also assumes that no driver changes their desired velocity. In realistic situations drivers desired speeds can and will adapt to a number of environmental changes such as road layout and surface conditions. These adaptations to the desired velocity can be included in the kinetic model in principle but vastly increase its complexity.

3.2.3 Multi-lane Extension

The Paveri-Fontana model can be extended to include multi-lane traffic more explicitly. Although Fontana included overtaking in his model, the different lanes were not defined. In this explicit modification, the lanes can be precisely defined and enumerated. Similar to the above models, a continuity equation is constructed to describe the evolution of the system. In addition to the relaxation and interaction terms discussed above three extra terms are included: A lane changing term and two further terms to describe the traffic entering and leaving the system through slip roads.

Another interesting consideration of this model is the variance in the velocities dependent on road restrictions (such as speed limits). For example, if the speed limit on an area of a road was 30 mph, a reasonable range of velocities could be within 5 mph of this value. However, if the speed limit was 70 mph a much larger range of 15 mph is not unreasonable. This shows progression towards the modelling of more realistic traffic behaviour, however it does increase the complexity of the model and thus makes it more computationally expensive to analyse [16].

A complication in the overtaking process arises when dealing with high concentration, well defined, multi-lane traffic. Here it is no longer necessarily a two body interaction as vehicles in neighbouring lanes will affect the decision and ability to complete the manoeuvre. This is managed by instead defining a concentration dependent overtaking probability. There may be other influencing factors affecting the probability such as the velocity of approaching vehicles or environmental factors like visibility and reaction time. [16]

3.3 Applications

Kinetic theory has been used to model intelligent transportation systems such as those involving adaptive cruise control (ACC), where the ACC vehicle will slow down or speed up depending on the behaviour of the leading car [17]. The introduction of these vehicles is expected to smooth the distribution of velocities within the system. Kinetic models are able to model macroscopic quantities and so allow differences in the average velocity distribution to be observed, making it a suitable choice for this purpose. This is a recent area of development and could have great influences for the future of driving comfort and safety. However, it needs to be evaluated in advance in order to minimise any negative results. This is done using the kinetic model by Ngoduy, where he compares his results to real data [17].

A fundamental problem in trying to practically apply gas-like kinetic models of traffic is the shear number of variables. However, since these models largely require dilute concentrations, the density of vehicles can be extremely small. Computationally, when performing calculations with extremely small numbers, they are subject to considerable fluctuations [16]. This can make it difficult to compare these models directly with data. A possible way to reduce the number of variables and hence minimise the computational expense could be to parametrise certain variables instead of expressing them as continuous functions. Parameterising allows a complex or very detailed computation to be represented by a simplified process. A similar approach is used in climate modelling where, for example, variables are parameterised over a grid box, reducing the resolution and therefore computational expense of the model.

3.4 Development and Funding

Several approaches could be taken to try and develop kinetic models with the aim of more accurately representing traffic flow. One of the recent developments to the kinetic model is to include a new variable known as activity. Instead of treating the vehicles as classical particles they can now be called *active particles* [18]. An active particle allows the inclusion of effects due to intelligent behaviour such as driving skill and reaction times. In the models discussed above, only the mechanical nature of the vehicles was considered in combination with a programmed desired velocity. This new approach also considers the differing properties of the vehicles themselves, it is hoped that by accounting for the inhomogeneous nature of traffic systems kinetic models can become more applicable for practical modelling. Research using this new activity variable is ongoing and has been approached by applied mathematicians, as it has many possible practical motivations [19]. For example, applying this to crowd dynamics instead of vehicles could provide an opportunity to predict dangerous/panic inducing situations when leaving stadiums or crowded places [19]. Due to this, funding of kinetic models for this kind of progression may lead to some success. It is worth noting that if development is continued to be made in this manner, via the inclusion of more and more variables, computational expense will become a limiting factor and parametrisation may need to be considered.

3.5 Conclusion of Kinetic Models

Some small individualities of drivers within the models presented have been discussed, such as differing desired velocities and velocity variances due to speed limit restrictions. However, there are many more important individual differences these models do not include. Variances for different vehicle types, which will have an effect on the velocity and driving style have not been incorporated (consider the differing driving approach between a Smart car and a large lorry). Different *types* of drivers (such as aggressive or apprehensive) have also not been accounted for, in fact there is a long list of additional features that could be included to make these models more realistic. It poses some interesting questions: can these models ever completely represent real-life traffic flow? Is a complete representation even necessary for some practical applications?

Currently kinetic models are unable to accurately represent traffic flow in practice: the continuity equation on which they are based is not entirely valid since, in the dilute limit, the distances between vehicles cannot be neglected. In a continuous medium a continuity equation states that the incoming flow is equal to the outgoing flow in equilibrium. In kinetic models, the dilute condition means that the distances between the vehicles is so large that probability distribution can no longer be considered continuous and so the continuity restriction cannot be definitively applied.

The methods used also require a far greater number of vehicles than physically occur in road systems. Because of this, the assumption that the functions are continuous is not entirely valid. The vast number of potential variables, and the interactions between such variables, makes it extremely difficult to conceive an explicit representation for real traffic flow. However, for use in city planning and road development this level of accuracy is not necessarily required as even a simplified general model would allow for observations and comparisons between different road layouts. The main issue with using kinetic models for applications is their requirement for dilute traffic conditions. The kinetic models presented are unable to represent traffic jams which, for road planning and development in cities, make them of extremely limited use.

4 Fluid Dynamics

While modelling the flow of traffic from the viewpoint of kinetic theory can be done with relative success, there will always be an issue when considering more densely populated systems, as the models themselves break down in this limit. Hence using fluid models would appear to be a sensible choice, especially when traffic models are going to be of most practical use in areas of dense traffic. It is also possible to consider traffic flow along highways as analogous to fluid in a pipe. As in the kinetic models (Section 3), we assume identical particles (vehicles).

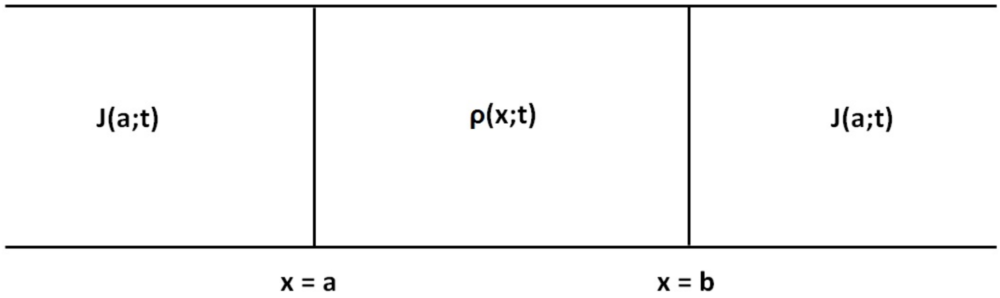

Figure 1: This diagram shows the flow, J, of a fluid with density ρ into a region. The flow is defined as the density of the material times the velocity; $J(x,t) = \rho(x,t)v(x,t)$. Original image.

4.1 Lighthill-Whitham (LW) Model

Much of the research in continuum modelling stems from the pioneering work by Lighthill and Whitham (and independently by Richards) in the 1950s. The basic premise of the model is conservation of the number of cars along a section of road. As shown in Figure 1, the number of cars in this portion of road is given by the number of cars entering at $x = a$ subtracting the cars leaving at $x = b$.

This can be algebraically expressed in terms of the density of cars [$\rho(x,t)$] and the traffic flow [$J(x,t)$]:

$$\frac{\partial}{\partial t}\left[\int_a^b \rho(x,t)dx\right] = J(a,t) - J(b,t), \tag{9}$$

where the rate of change of number of cars in the interval (a,b) on the left has been equated to the difference in flow between the two endpoints on the right.

Considering the limit where $(b-a)$ gets infinitesimally small this becomes the more familiar continuity equation:

$$\frac{\partial \rho(x,t)}{\partial t} + \frac{\partial J(x,t)}{\partial x} = 0. \tag{10}$$

This is the central equation of the Lighthill-Whitham Model.

The next step is to state the relationship between the flow and density of the traffic. This cannot be defined explicitly due to the intricacy of the system, but can be approximated. The LW approximation is the assumption that:

$$J(x,t) = j(\rho(x,t)). \tag{11}$$

In other words, the dependence of flow on position (x) and time (t) is only defined through the functional dependence of the density. Since the flow is defined as $J(x,t) =$

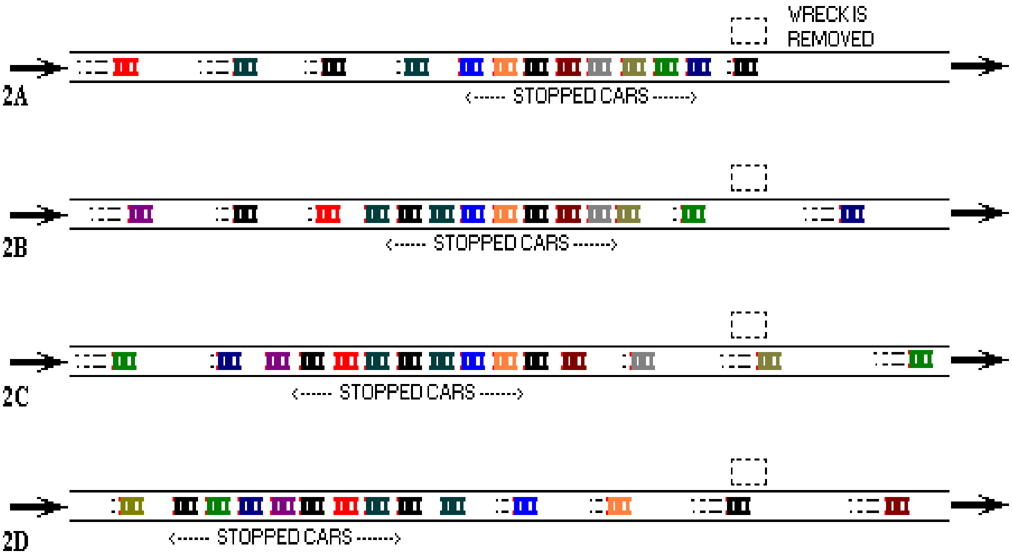

Figure 2: This figure demonstrates the backwards propagation of a density wave caused by a car crash. It is important to note that the density wave propagates backwards with respect to the traffic flow, it can still move forward with respect to the road. This figure has been taken from [20].

$\rho(x,t)v(x,t)$ it is implied that velocity, too, has its spatial and time dependence defined through the density i.e. $v = v(\rho(x,t))$. Which, when substituted into Equation (10), using the chain rule gives:

$$\frac{\partial \rho(x,t)}{\partial t} + \left[v(\rho) + \rho(x,t)\frac{dv(\rho)}{d\rho}\right]\frac{\partial \rho(x,t)}{\partial x} = \frac{\partial \rho(x,t)}{\partial t} + v_g \frac{\partial \rho(x,t)}{\partial x} = 0, \tag{12}$$

where the group velocity is defined as $v_g = \frac{dJ}{d\rho} = v(\rho) + \rho\frac{dv}{d\rho}$.

This equation is non-linear in general. However, if v_g is taken to be constant ($v_g = v_0$) it is possible to write a general solution of the form:

$$\rho(x,t) = f(x - v_0 t). \tag{13}$$

This represents density wave motion - an initial density profile which travels at speed v_0 without changing shape. It is important here to note that velocity will be a strictly decreasing function of density - the denser the traffic, the slower the traffic will travel. Hence $\frac{dv(\rho)}{d\rho} < 0$, implying that the density wave travels more slowly than the fluid itself. Therefore, the density wave travels backwards in relation to the motion of the fluid. The propagation of the density wave through traffic after a car crash is demonstrated in Figure 2.

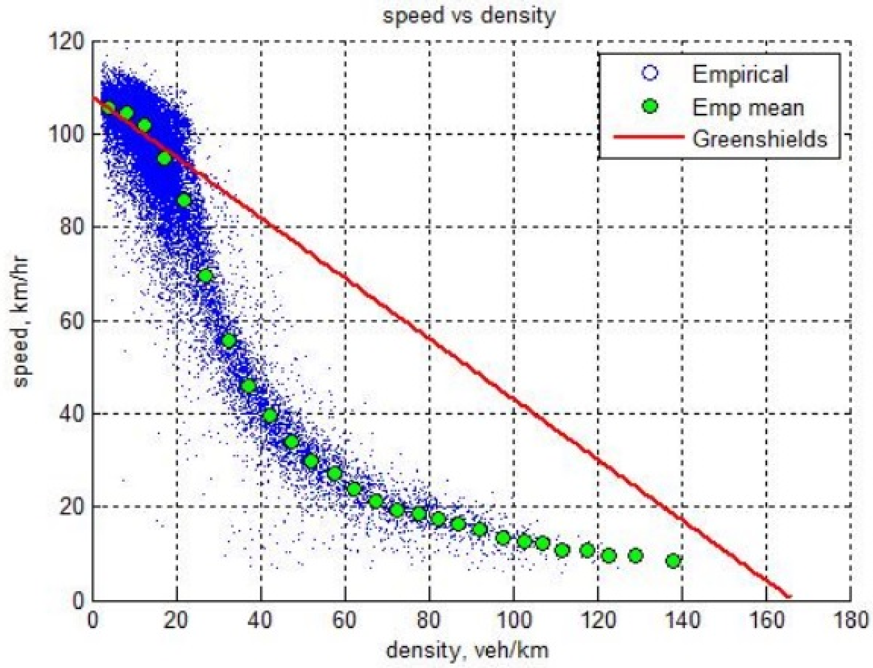

Figure 3: Here the speed is plotted against density, with the real data in blue and the Greenshields model prediction in red. The blue 'cloud' of data points represents one years worth of data with each point representative of the traffic condition at a fixed position taken at five minute intervals. This picture has been taken from [21].

To progress further it is necessary to state the functional form of the velocity. Many different variations have been suggested, but the majority follow the same basic principle - defining a maximum density, ρ_{max}, at which flow grinds to a halt - and many take the general form:

$$v(\rho) = v_{max}\left[1 - \left(\frac{\rho}{\rho_{max}}\right)^{\alpha}\right]^{\beta}. \tag{14}$$

Where the maximum 'jam' density has been defined such that $v(\rho_{max}) = 0$ and is analogous to a phase transition from liquid to solid. The maximum velocity (v_{max}) as well as α and β are positive constants that are defined by the particular conditions of the model and can be calibrated to match physical data. For example, the simplest version of the above equation is given by the Greenshields model ($\alpha = \beta = 1$) [21]. The predictions of the Greenshields model have been compared to data taken from a 4-lane highway in Georgia, USA and are shown in the Figure 3.

Clearly the Greenshields model is too simplistic as it deviates significantly from the data. However, there are other forms of velocity that fit the data more closely as shown in Figure 4.

One model that appears to fit the data well is the Underwood Model (turquoise) in Figure

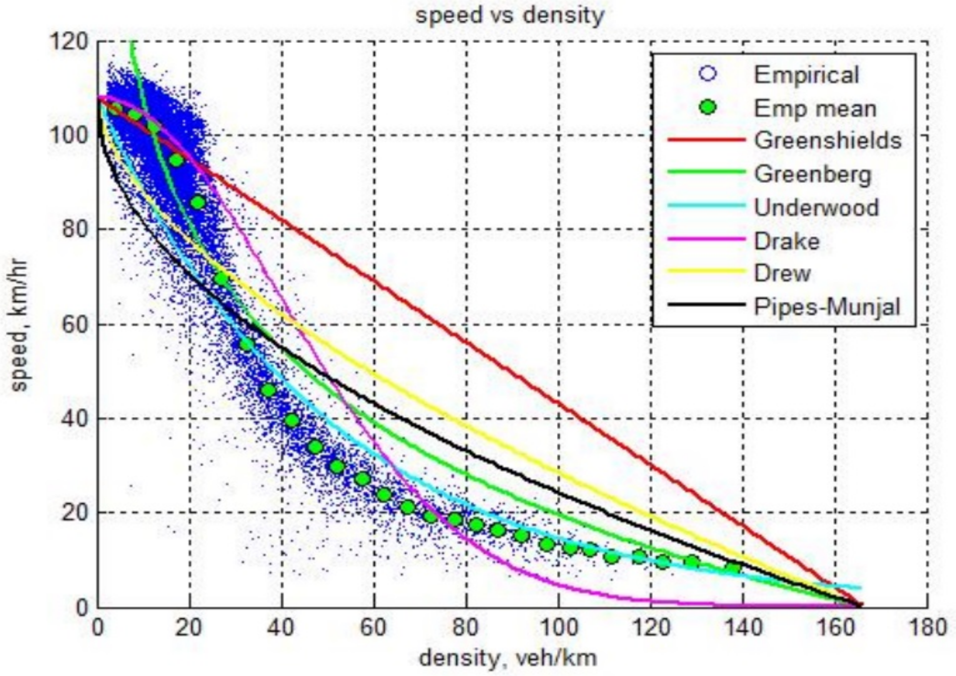

Figure 4: Here a number of different forms of the density dependent velocity function have been plotted against real data for comparison. Figure taken from [21].

4, which models velocity by a Boltzmann distribution

$$v = v_{max} \exp\left(-\frac{\rho}{\rho_{max}}\right), \tag{15}$$

which is also used in many kinetic theory models.

4.2 Extensions of the LW Model

Whilst the basic Lighthill-Whitham model is useful in modelling simple, single lane flow, for it to be more applicable to real-life situations other factors need to be taken into consideration such as a human drivers, the varying size of vehicles on the road and interweaving multi-lane traffic. Here, some improvements to make the model more realistic have been explored.

4.2.1 Diffusion Term

In order to address some of the limitations of the above model, LW proposed some improvements yielding a second order model in space. Firstly, assume that the flow is not only dependent on the density of the fluid, but the gradient of the density as

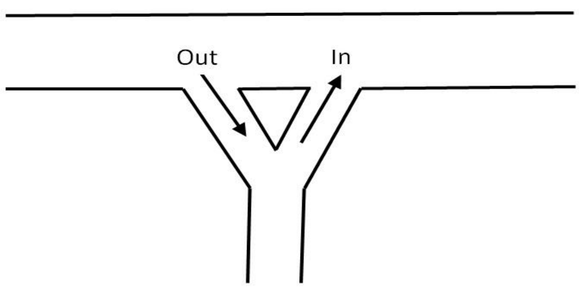

Figure 5: A simple schematic of entrance and exit terms for a one-lane system. An original image.

well:

$$J(\rho) = j(\rho) - D\frac{\partial v}{\partial \rho}, \tag{16}$$

where D is the diffusion coefficient, a positive constant.

This diffusion term can be interpreted as the drivers ability to adjust their speed based on what is in front of them. When they approach to an area of higher density, they reduce their speed and vice versa. A new expression can be constructed based on the redefinition of $J(\rho)$,

$$\frac{\partial \rho}{\partial t} + v_g\frac{\partial \rho}{\partial x} - D\frac{\partial^2 \rho}{\partial x^2} = 0, \tag{17}$$

where, again, the group velocity is defined by $v_g = \frac{dJ}{d\rho}$.

4.2.2 Entrances and Exits

It is possible to introduce incoming and outgoing vehicles to the first order form of the LW model when still assuming single lane as so:

$$\frac{\partial \rho(x,t)}{\partial t} + \frac{\partial J(x,t)}{\partial x} = In - Out, \tag{18}$$

where In is the flow of cars onto the road and Out the flow off the road, as shown in Figure 5.

Using the same mathematical formalism, it is possible to consider a two lane system. In this case there would be an equation of the same form as Equation (18) for each lane and the entrance/exit terms of each can be seen as cars coming from (going into) the other lane. This results in a simplistic multi-lane system. By including more lanes and entrance/exit ramps, motorway-like situations can be modelled.

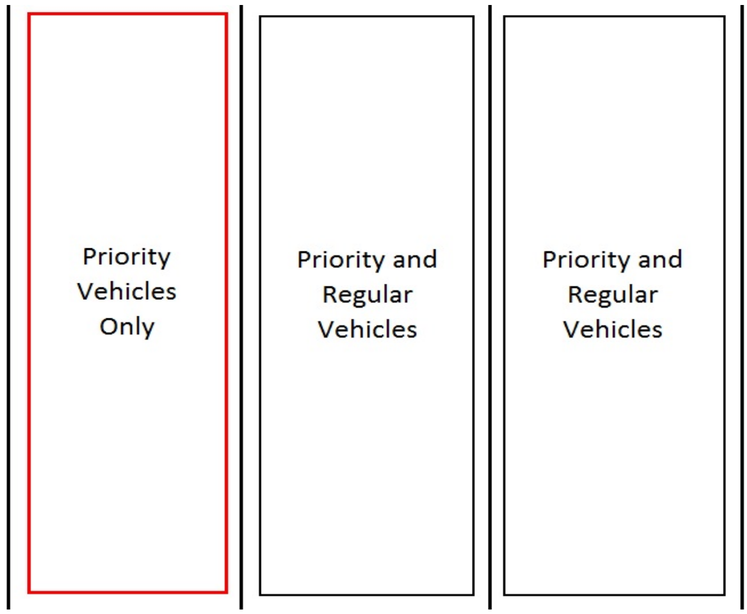

Figure 6: A simple example of a three lane system with one priority vehicle lane and two regular vehicle lanes. An original image.

4.2.3 Special Lane Traffic

A useful progression of the LW model that has been researched are the dynamics of Special Lane traffic [22]. Special Lane traffic is defined when there are two classes of vehicle. One that can only travel along a given number of regular lanes and the second priority class that can enter all of these lanes, plus other lanes that the first vehicle cannot. A simple example is a road with bus lanes. However, the same basic framework can be applied to a stretch of motorway in the lead up to an exit. In this case, the cars wishing to take the exit ramp are considered the regular vehicles - confined to the lane nearest the exit - and those staying on the road can be in any lane.

As before, the continuity equation is used to model the evolution of the traffic flow, but it now has vector form:

$$\frac{\partial \boldsymbol{\rho}}{\partial t} + \frac{\partial \mathbf{J}}{\partial x} = 0, \tag{19}$$

where $\mathbf{J} = \binom{J}{j}$ etc. Upper case letters represent priority vehicles and lower case represent the regular vehicles.

With this configuration it is possible to consider two regimes in which the system will reside. A two pipe regime where each class of vehicle stays in their own lanes, which essentially implies existence of two separate pipes with two separate fluids, $V > v$. If the speed of flow in the priority lanes is slower than that of the regular lanes, we assume that the priority vehicles will move across to the quicker, regular lanes until the advantage is negated. This results in being able to consider all of the lanes in a one pipe regime with one coalesced fluid and one speed $V = v$. For simplicity, it is assumed that the transition between these two regimes is instantaneous [22].

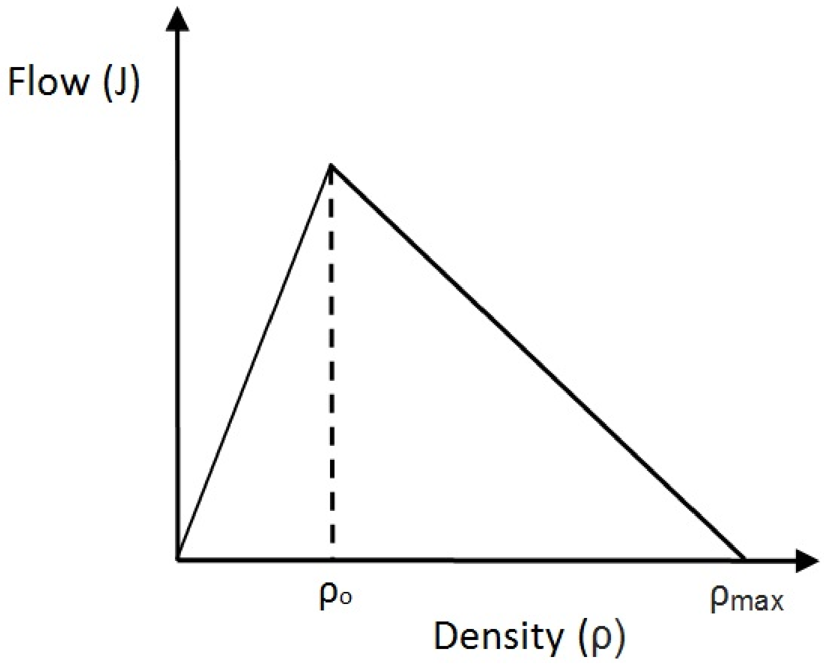

Figure 7: Original Image. The figure demonstrates a triangular flow-density relation where ρ_0 is the optimal density (the density at which flow is at its highest) and ρ_{max} is the maximum density, where flow stops.

The transition between the two regimes is primarily based on an a change of density of the system and it is possible to establish the point at which this regime change occurs. Consider a general velocity function, u, (which decreases with density) that describes the velocity of flow in any regime [22]. Once the form of this has been determined, comparing the velocity condition for the transition from the two pipe regime to the single pipe regime ($V = v$) it is possible to deduce the density condition for said transition. The one pipe regime combines the density of all lanes

$$V = v = u(P + \rho), \tag{20}$$

where, for 2 pipes,

$$v = u(\rho/y) \qquad V = u(P/Y). \tag{21}$$

In Equation (21), Y and y give the fraction of priority and regular lanes ($y + Y = 1$). Hence $\frac{P}{Y} > \frac{\rho}{y}$ implies the one pipe regime and vice versa. From here it is possible to predict the evolution of a system with LW theory when assuming a simplified triangular flow-density relationship, as shown in Figure 7.

While it is possible, depending on the boundary conditions, to solve this system analytically, in many cases only a numerical solution currently exists. In addition, this method has not been tested against real data and therefore, though a good model in principle, it is difficult to judge the merit of such an approach [22].

4.3 Limitations

Although considered by many as one of the most applicable models of traffic flow, LW model has recently received a lot criticism. In particular, Kerner identified three crucial problems with traffic flow modelling using conventional approaches, such as the LW model.[23] These are:

Spatiotemporal behaviour: Empirical results point to the fact that the emergence of a jam at a bottleneck is more localised than the theory predicts.

Time Scale: The transition of free flow to jam is accompanied by a significant velocity drop. In reality this velocity drop occurs very quickly, but the models overestimate the decreasing interval by a factor of at least 10.

Space Scale: In reality jam transition points (or bottlenecks) are localised to around 200-300 m while the simulation results give distances bigger by a factor of at least 10.

Due to the above reasons, Kerner [23] argues that fluid models have, and had, very limited impact on modelling traffic flow in practice.

4.4 Applications

There are a number of methods through which researchers have been able to model traffic on a single lane road whilst taking into consideration concepts such as density gradient dependent velocities which are, qualitatively, applicable in practice. It is possible to see the application of the different forms of the LW model in Figure 4 as the density dependence of velocity on an American highway is shown to match that of a model [21]. Another use for LW models is in areas of urban traffic. In this instance, the traffic flow dynamics are dominated by the numerous traffic signals as opposed to the intrinsic dynamics of the flow and the limitations of the LW model are not as detrimental in this case [24].

4.5 Development and Funding

With the addition of other factors (multi-lane traffic [25], junctions [26] etc.) it becomes significantly more difficult to find analytic solutions. In many cases, an analytic solution is simply not possible so numerical solutions, if obtainable, have to be used. There is a case where a fluid model was used in the Los Angeles area to predict the occupancy of a highway based on inflow and outflow data from ramp meters. They show that the speed density relationship can be fitted to the model by monitoring the stream of vehicles on entrance and exit ramps of a freeway [27]. However, in this particular case the simulation splits the time and position into sections, hinting towards the discretisation to be discussed in the following models.

Thus, it appears that, for the application purposes, the next step in the development is to convert fluid models into a more discrete form, in order for them to be processed computationally. As a result, it would not appear that an increase of funding into fluid dynamics would be of major benefit. For the complex systems that cannot be calculated analytically, there is no guarantee that investment will result in any advancement of the field. In order to computationally solve fluid models, they have be artificially discretised while there is a possible formulation of a naturally discrete system. Lastly, it has been already pointed that fluid models do not reproduce the crucial empirical facts and thus their correspondence to reality ranges from only approximate to non-existent [23].

5 Discrete Models

The models which have been mentioned so far fall into the category of continuous models. This means that states of the system are represented by continuous variables. In contrast, for discrete models, time varies as integer multiples of a given timestep.

The analysis of discrete models is very difficult without the aid of computer simulations, hence, the fluid models were more popular before the era of cheap computation. However, with invention and increased availability of computers the interest of scientists and engineers shifted to discrete models. Discrete models are currently used in practical analysis of road networks and are a popular research topic in academia as well.

5.1 Theoretical Background

Discrete models of vehicular flow suppose that states of the system can be enumerated with integers. In other words, they are countable. Systems with a finite number of states that obey transitional rules are generally labelled by the common name automata. Therefore, discrete models are often called automata models of traffic flow, and these terms will be used interchangeably [8].

automata models of traffic flow usually model states which are fully specified by position, speed, acceleration and time. For example, a single lane road can be represented as a one dimensional lattice whose cells can be occupied by a car. Transitions between two states going from state t to state $t+1$ is given by set of rules which govern the dynamics of the system. Different rules have been proposed, from Cremer and Ludwig [28] to the most widely known Nagel-Schreckenberg model [5].

The theory of automata tells us that the rules illustrated in Figure 8 can sometimes produce unpredictable or surprising outcomes despite looking very simple. It is common that a system with very simple automata obeying straightforward rules can lead to chaotic or cyclic behaviour. For example, Rule 30 according to Wolfram notation, which has simple rules for subsequent row update such as those shown in Figure 8, produces chaotic behaviour, as can be seen in Figure 9.

This unpredictability means that dynamics of models based on automata are very hard to explore without computer simulation, hence why discrete models have been introduced relatively late to the field of traffic modelling. To explore the system, or to make

current pattern	111	110	101	100	011	010	001	000
next pattern	0	0	0	1	1	1	1	0

Figure 8: Rules of automaton 30. To compute the next line in automaton we look at the previous line and see the numbers in the cells directly above the computed cell. We assign values to the next row according to the table. A visual description where white - 0 and grey - 1 can be found in Figure 9.

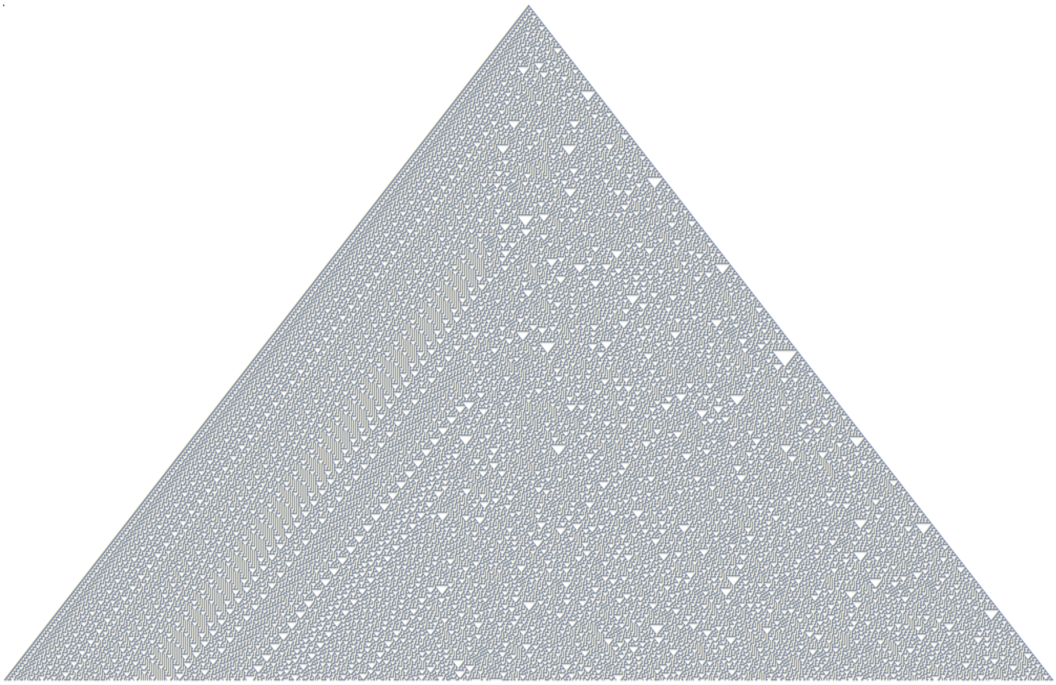

Figure 9: automaton with rule 30. One can notice that on the right a chaotic behaviour occurs. It fulfils even rigorous criteria of chaos as the final result is very sensitive to changes in the initial conditions [29]. The picture has been taken from [30].

predictions, one has to simulate the automaton in order to obtain results. This introduces pragmatic constraints on the accuracy and length of the simulation due to limited computational power.

In the following sections the most common models and approaches to discretised vehicular traffic flow are briefly described. The initial focus will be on single-lane roads then progressing towards more complex traffic flow models including junctions. The section ends with a discussion of a few practical systems in use today.

5.2 Nagel-Schreckenberg (NS) model

5.2.1 Basic Formulation

The most famous automata model was developed by Nagel and Schreckenberg [31]. Although it is not the first discrete model to be successful in theory and applications [28], it was the first model which significantly sparked interest in discrete modelling. The NS model is of interest due to the simplicity of its rules. Before describing the rules of the model, assumptions must be made about the variables that describe the state of the system. Let x_n, and v_n denote the position and the velocity of the n^{th} car repectively. As the cars cannot overtake in basic NS model, these quantities define the system completely. Next, suppose that $v \in \{0, \dots, v_{max}\}$, and $x_n \in \{0, \dots, L-1\}$ where L denotes number of cells on the single-lane road. For convenience, define the variable, $d_n = x_{n+1} - x_n$, the separation between the car labelled $n+1$ and the car labelled n. Also, suppose that the total number of cars in the system is $N \leq L$.

The set of rules that tell us how one can propagate the system (forwards or backwards) in time are called dynamical rules. In the NS model, there are four rules that are applied for each car in parallel [5]. In other words, the recipe is followed for each car with input variables of the state at the previous time. The recipe for updates goes as follows:

- Velocity update:
    1. $v_{n+1} = \min(v_n + a_n, v_{max})$, where a_n corresponds to the acceleration of the vehicle (which might differ for each vehicle). In the basic NS model, a_n is set to one for all cars.
    2. $v_{n+1} = \min(v_{n+1}, d_n - 1)$
    3. if $v_{n+1} > 0$ then with probability p, $v_{n+1} = max(v_{n+1} - 1, 0)$
- Position update
    1. $x_{n+1} = x_n + v_{n+1}$

The rules are not deterministic as there is a random element which makes our automaton model stochastic. This turns out to be a crucial feature of the system and will be further discussed.

The NS model, as described here, is a minimal model according to Chowdhury [8]. These rules are the minimum requirements to reproduce basic traffic flow behaviour, such as formation of jams, slowing of traffic with increasing density, and variation in driving styles. For example, step 1 in the velocity update reflects the desire of drivers to travel as fast as possible. Step 2 ensures that the driver brakes so that they do not collide with the vehicle in front of it. Step 3 is stochastic and tries to emulate different driving styles. For example, some drivers brake too much when encountering an obstacle or slower vehicle, these slight variations turn out to be crucial in the forming of jams [8].

It has been shown that relaxing any of the above conditions does not reproduce jamming properties of traffic flow [5]. However, this does not mean that this system and its rules are the only possible ways to reproduce congestion on roads i.e. there are multiple minimal models. For example, a deterministic model with a different set of transition rules, lacking the stochastic rule, was shown to replicate congestion reliably [32]. These models allow us, given a density of cars in the system, to estimate the point at which transitions occur for free traffic, stop and go, and congested traffic [5].

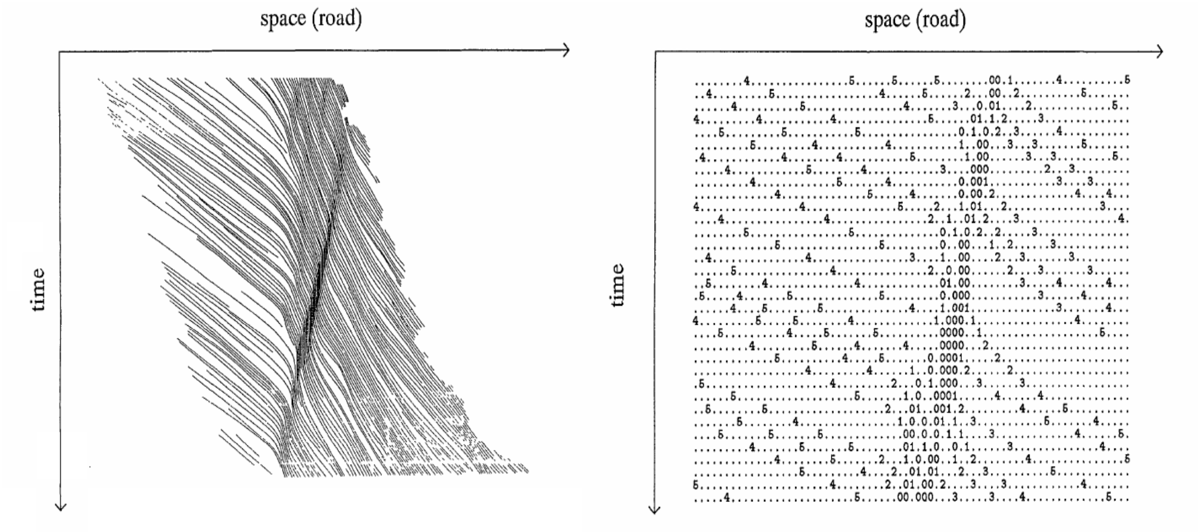

Figure 10: This picture contains simulation of NS model. The diagram on the right shows the evolution of time, where the number at each position in the lattice represents the speed of the car. The track of each individual car in time can also be seen. This picture represents a jam, as can be seen from the 0 - velocity cluster in the left and slowed progression of cars on right [5].

5.2.2 Multi-lane extension

The NS model can be easily extended to multiple lanes with the addition of another set of rules for lane changing [33, 8]. Specifically, at the beginning of the update a car may decide, based on the lane changing rules, whether it changes a lane and then proceeds with the basic NS model rules.

The rules of a two lane road may be symmetric or antisymmetric. There is a non-obvious choice to this, as symmetry is not often observed in practice but makes the dynamic rules less complex. For example, in some European states legislation dictates that the right lane should be preferred if unoccupied and that overtaking can be done only from the left. While building a multi-lane model it is important to take into account the reasons for lane swapping. For example, the other lane might allow higher speed or the driver might need to make a turn in the future which the current lane would not permit. Additionally, lane changing rules have to implement safety rules such that a collision does not occur. An example of a simple symmetric two-lane system has been proposed by Rickert and others [33]. If the following four criteria are satisfied then lane changing occurs:

1. The gap between one vehicle and the vehicle in front (in the current lane) is smaller than desired. The example of desired distance is usually taken as $\max(v_{max}, v+1)$.
2. If the situation in the alternate lane is better than in the current lane. In particular if the distance is bigger than $\max(v_{max}, v+1)$.
3. If the distance to the car behind in the other lane is bigger than required distance, then there there will be no collisions.
4. Lastly, Rickert included a stochastic aspect to the lane change rules. In other words if all previous criteria are satisfied the change only happens with certain probability

p_{change}. This ensures that there is no "ping-pong motion" [8] between lanes.

A unified review of the two-lane models can be found in the seminal work of Nagel et al. [34]. The theoretical analysis of such models often becomes very cumbersome, however simulation techniques are very successful in predicting results using these models, despite additional complexity. Most models used in practice are conceptually based on the NS model, with extensions to ensure that they are more applicable.

5.3 Biham-Middleton-Levine (BML) model

5.3.1 Basic Overview

The Automata models presented so far were concerned with traffic flow confined to one dimension. Another avenue of development is that of multi-directional flow, a simple example of which will be the topic of discussion of this next section - the Biham-Middleton-Levine model. Complications arise when considering multi-directional flow due to the inherent appearances of junctions, which play an important role when modelling traffic in urban areas. The model BML model, which was created in 1992 [35], models vehicles travelling in two perpendicular directions. In the examples provided these are chosen to the north and to the east.

As the BML model is a automaton model, it is also described by set of state variables. In this case, this includes the two dimensional position of car n, (x_n, y_n), and time, t. The dynamic rules are remarkably simple and are executed in parallel as follows [36]:

1. If t is odd - north travelling cars move $\uparrow$
   If the space above the vehicle is free, the car moves up a cell.
2. If t is even - east travelling cars move $\rightarrow$
   If the space to the right of the vehicle is free, the car moves a cell to the right.

This problem was originally explored based on the initial conditions. On an NxN grid a car was inserted with a probability p on each lattice cell and, as a consequence, p also represents density of the occupation on the grid. Hence, the terms probability and density will be used interchangeably in this section. The next step in the initialisation process is to choose a random half of the cars to be travelling north and the other half to be travelling east.

At low values of p free flow is observed, as in Figure 11. During free flow the average velocity is measured to be $\langle v \rangle = 1$ (where velocity is in moves per timestep), and the traffic does not experience jamming. Note here that the average velocity refers to averaging over all vehicles in the system and the maximum velocity is achieved when a car is able to move at every time step. However, due to inherent randomness in the system, this velocity is also averaged over a number of runs of the simulation at the same probability. A recent study showed that the system suffers a transition to an ordered flow, an intermediate state between the free flow and jamming, as shown in Figure 11. It has been shown that the average velocity of this transient phase is close to the value free flow, $\langle v \rangle = 1$ [36].

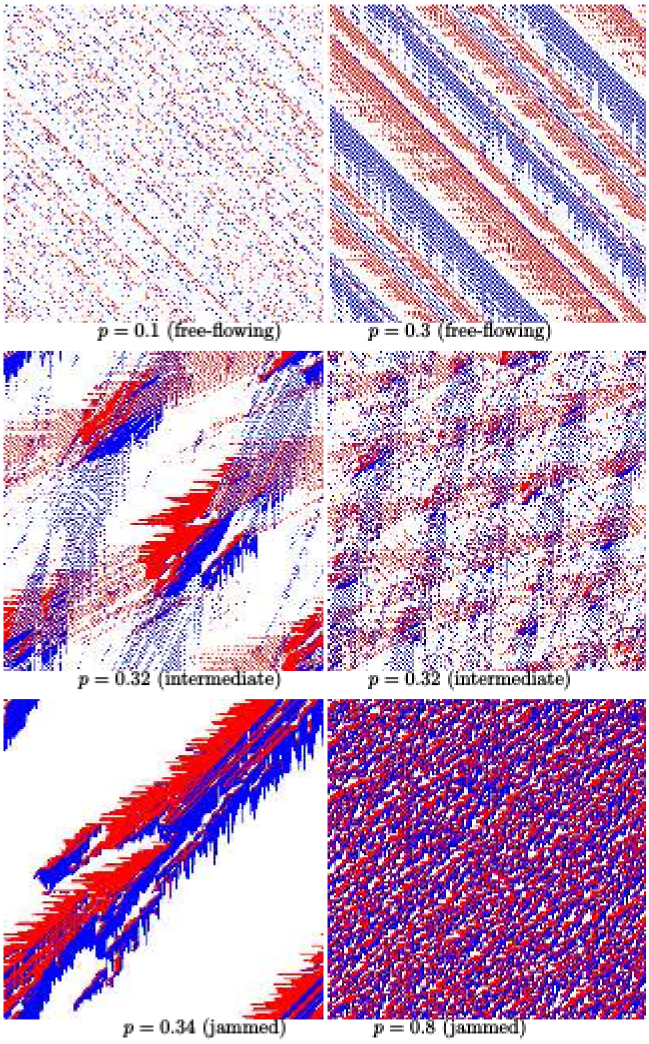

Figure 11: Examples of flows with varying densities of cars can be seen in this figure. East-facing and north-facing cars are shown in red and blue respectively. The picture is taken after 20,000 steps on a 200-by-200 grid. The figure has been adapted from [36].

When the average velocity is $\langle v \rangle = 0$ a jamming state occurs and has been experimentally shown to occur for probabilities $p > 0.35$, as shown graphically in Figure 11.

The BML model served as the initial inspiration for physicists and analysts that modelled city traffic [37]. Despite its elegance and remarkable properties, the BML model is known to be mathematically very intractable [38]. For this reason, researchers often resort to simulation techniques to explore the system, as is the case in the previously mentioned NS model (Section 5). Detailed analysis of the BML model is beyond the scope of this report, however this is not needed to appreciate the successes of the theory, such as modelling four-directional traffic [39]. Different modifications regarding the dynamics, such as random updates instead of odd/even distinction, can give rise to more complex phenomena for which practical interpretations have been considered [36].

5.3.2 Four-Directional Extension and Gridlock

Expanding the model to four directions is a necessary complication to enable the modelling of real-life junctions. Four-directional traffic simulations have been proposed in the article by Huang in 2006 [39]. In this setup, vehicles are defined to travel in one of the four directions. The initialisation step is very similar to the original BML model. The only difference is that east and west moving vehicles move on alternate rows and the north and south travelling cars travel on alternate columns. Similar to the original BML model, the horizontally travelling vehicles are updated on odd steps and the vertically travelling ones are updated on even steps.

The above complication introduces a new phenomenon where vehicles travelling in all four directions can simultaneously block themselves in a formation called gridlock, as can be seen in Figure 12. In this case, four cars trying to travel in their separate directions are blocked. The occurrence of these gridlocks increases with the density of cars, and at high enough densities a state at which no car can move is called a total gridlock [39]. Additionally, gridlocks can form islands where vehicles have been blocked, as in Figure 13. The nature of total gridlocks change based on the initial conditions e.g. at $p = 0.05$ they have the form shown in Figure 14 and are said to be in global gridlock. When the density increases island gridlocks start to emerge e.g. at $p = 0.1$ they look as shown in Figure 13.

Surprisingly, the occurrence of gridlock is found to be determined not, as is the case in the original BML model, by density, but by total number of vehicles. The minimum density value for gridlock is $p_{grid} = \frac{4}{L}$, where L represents the side of a square grid. Experimentally it has been determined that at $p = p_{grid}$ global gridlock occurs, and at $p \geq \frac{6}{L}$ island gridlocks do [39]. On a large lattice total gridlock is nearly inevitable at any finite density due to p_{grid} becoming small.

The average velocity $\langle v \rangle$ is an important variable in this situation and sharply drops with increasing p at any given grid size. In the case of global gridlock the velocity decreases linearly with time, while for island gridlocks ($p \geq \frac{6}{L}$) the decrease of average velocity is exponential. Having been mostly concerned with gridlocks in this chapter it is important to note that free flow is rarely found in four-directional traffic. For instance, on an 80x80 lattice, with density as low as $p = 0.025$ only 3 out of 1000 simulations achieved a free flowing state [39].

5.3.3 Traffic Light Extension

Another class of advances of the basic BML model keeps the two directional traffic setting, but extends the dynamic update rules. Gu et al. introduce a distinction between some of the lattice fields, in other words, the grid is divided into two classes of lattice cells [41]. At lattice cells A, of which there is concentration c_A, a car has to wait at least time τ_A to advance. Similarly, at B (concentration c_B) a car has to wait time τ_B. This is in contrast with the original model where all the lattice cells were identical with waiting times $1 = \tau_A = \tau_B$. Other update and initialisation rules are kept the same, as in the previous model.

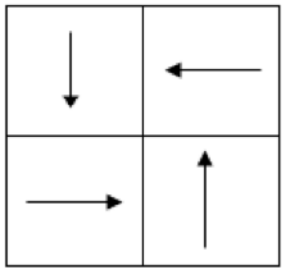

Figure 12: This figure shows a basic unit that is source of a gridlock. One can observe that none of the vehicles are able to move in any of the time steps. Since the update rule disables the possibility of two cars being on the same cell, these four vehicles have to remain in this configuration. This is an original image.

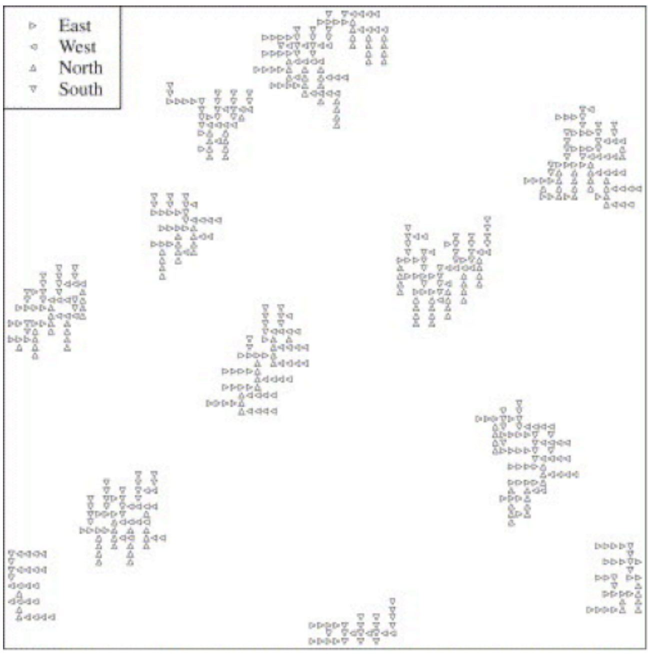

Figure 13: In this figure, island gridlocks are shown. These are formed at higher values of p, namely $p \geq \frac{6}{L}$. In this case the vehicles cannot move further, however they are concentrated in local regions rather than globally, in contrast to Figure 14 This picture is taken from [40].

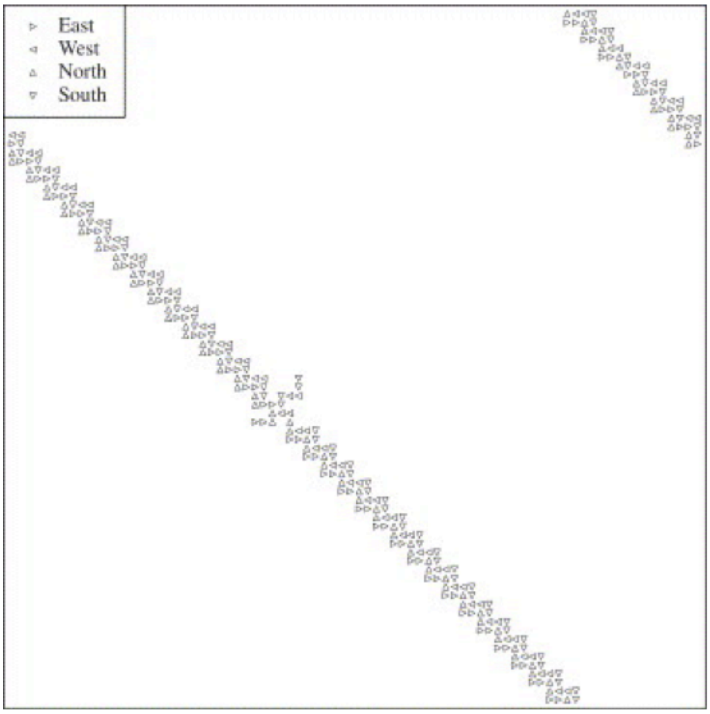

Figure 14: In this figure, the formation of gridlock is shown when the initialisation probability (or density, if one wishes) is p_{grid}. In this case, a so called global gridlock occurs where all the vehicles form a connected structure that cannot evolve further. This picture is taken from [40].

A realistic interpretation is offered for these different waiting times at some lattice cells in the form of non-synchronised traffic lights. Another important feature of this model is its ability to describe blocked roads. In this case one the cell waiting time τ_A is set to ∞, ensuring that the cars stay at the same spot. For this case fewer update steps are required to reach a jammed state. Additionally, the sharp transition from free flow to jamming that occurs in the original BML model is not present [41]. Through simulation it can be seen that even a small number of accidents will cause considerable jams in a city [41]. Relaxing the condition on τ_A to $1 \leq \tau_A \leq \infty$, keeping $\tau_B = 1$, reduces the average velocity of the system, but raises the critical density required for the system to jam. It has been shown for a system with denity $p \geq 0.26$ that adding more B sites reduces the jamming [41]. Hence it can be interpreted that, in higher densities, controlling the traffic lights in cities will keep the traffic moving by reducing the average speed.

However, in real-life, where there are restrictions, there are drivers who will violate these restrictions. Ding et al. introduced the concept of a driver who breaks the rules of the traffic [40]. The new kind of drivers are called violators and they conform to the following strategy: if the space in front is free it will move there in the next step, whether it is an even or odd step. Violators are randomly chosen at each step update, and can be travelling in either direction. The presence of violators complicates the situation and, therefore, an additional rule is needed to ensure there are no collisions. A normal driver, or a violator with a green light, gives way to a violator with a red light. The presence of the violators increase the average velocity of the system at low density. However, it decreases the critical density where jamming occurs. This is a result of the violators breaking the self organisation pattern that would otherwise occur at low densities [40].

Finally, the original BML model is considered with the inclusion of an overpass cell, allowing two vehicles to occupy the same site. The fraction of overpass cells included, f, is set before the simulation is begun. At $f = 1$ the system decouples into two one dimensional problems where the overpass sites allow north travelling and east travelling vehicles to coexist on the same site. The more overpasses, the higher the critical density required for jamming and the quicker the average velocity of the system. A conclusion that could be drawn from this model is that overpasses are the best device to reduce traffic jams. However, in most cases it is not a cost effective one [41].

5.4 Limitations

Discrete models are popular for realistic modelling of traffic, as can be seen from the previous section. However, these models are vastly complicated and often deviate significantly from the underlying simple models of NS or BML. These simulations are discretised but their update steps and set of dynamics are often more complex with constantly updating real-time conditions. Therefore, it is difficult to speak about applied traffic modelling with respect to the simple models addressed in this paper. They should, instead, be considered as inspirational paradigms for engineers progressing towards practical application. Kerner shows that there are fundamental empirical features that a traffic model should reproduce such as the possibility of different jam states [23] but it is difficult to show that a particular model conforms to states similar to the empirical results. However, as already said, discrete models appear to be the best candidates to achieve this.

5.5 Simulation in Practice and Funding

Before proceeding with a discussion of the particular limitations or successes of simulation, it should be noted that there is a significant difference between the approach of scientists and engineers to the modelling of traffic flow. As argued in Section 2.2, traffic flow scientists are often motivated by reasons that transcend particular applications in vehicular traffic flow. Therefore, it is possible that a particular model can be researched for reasons other than its applicability in the real world.

For some scientists, many of the models that are currently used in engineering are too complex to be approached analytically or even computationally. Therefore, many focus on simplistic models that usually are too simple to be practically used to optimise traffic flow, such as the NS model. However, the NS model provides a solid foundation for the understanding of the concepts that drive the most applicable models; as most discrete models are based on NS, it is clear that this theoretical model greatly influenced the evolution of practical traffic flow modelling. Previously, the most successful models were based on fluid dynamical approaches such as the Lighthill-Whitham model but these have been gradually replaced by discrete models [5]. This shows that scientific approach to traffic flow modelling, although not be initially motivated by practical problems, may have a great impact on it in long run.

Additionally, automaton simulations can be done very cost effectively as the models are relatively simple to simulate on small scales. They can provide useful insights into the effects of traffic jams and how accidents, traffic lights etc. can impact upon city traffic. These can be theoretically simulated at any density and do not pose mathematical limitations in this sense, in contrast to kinetic or fluid models. The models presented here, being fairly simple in nature, allow for the possibility of expansion and further research. The simulations created from these models could prove of further use in city traffic modelling as well as for a wide range of other phenomena. Because of this, it is concluded that funding in this area would most likely be beneficial.

5.6 Practical Applications and Successes

As far as back in 1992, Nagel states that discrete models have been used to improve traffic conditions in Germany [5]. The potential of traffic flow simulations has been acknowledged by the United States government in 1995 [42] with the first large-scale project, TRANSSIM, created at Los Alamos Laboratory. This suite of programs and tools are used to simulate and forecast regional traffic flow as well as the environmental impact of the traffic [43].

A recent project from 2012 [44] based on VANETs - vehicular ad hoc network - uses aspects of discrete models. VANET models use cars as nodes that report information about the state of traffic. The aforementioned model uses information from these VANET nodes to warn drivers of a congestion or a problematic situation on a given road. To show the effectiveness of these methods, authors have resorted to discrete model computer simulations for comparison.

A more recent project, being run by the University of Duisburg, is forecasting the traffic conditions of North Rhine-Westphalia (NRW) highways in Germany. This project is supported by the German government and consists of a simulation with a website interface (`http://www.autobahn.nrw.de/`) that predicts traffic situations for the next 30 or 60 minutes (longer predictions are available) on all German highways in NRW. An example of predictions with colour indications can be see in Figure 15. This model is much more complicated than a simple NS model, with a multitude of constantly updating simulations, however it is based on the same core understanding of traffic using discrete models. Another scheme receiving investment from the German government is trying to estimate, using simulations, the outcome of the possible imposition of payment for the use of German Highways [45] i.e. introducing toll roads.

6 Developing the Traffic Models for Future

In an ideal situation a traffic model could completely describe the dynamics of a traffic system, allowing the forecasting of significant events such as jams. For this to be physically viable, the density of the vehicles must be known. This could be achieved by strategically placed cameras or pressure pads. Knowledge of the road network and an

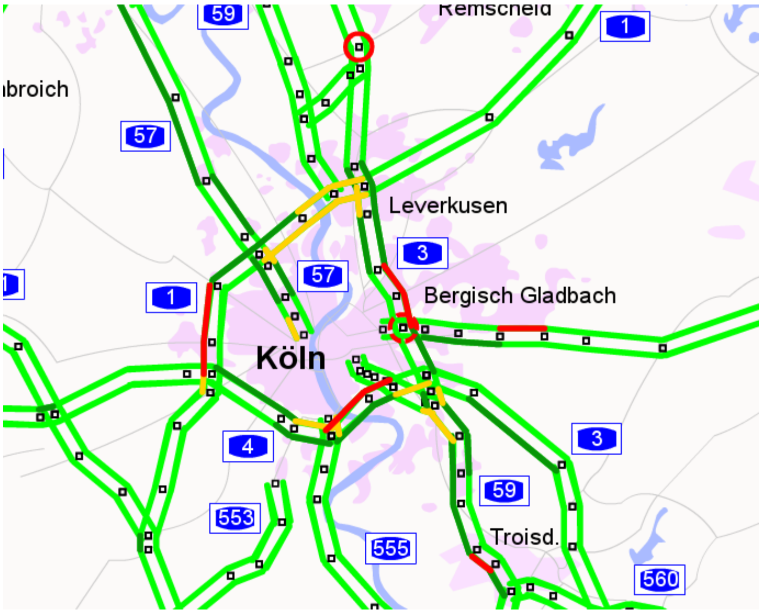

Figure 15: This screenshot from the Internet application shows 30 and 60 minute predictions for traffic flow around Cologne. The colours of the roads represent the density of traffic with green being the least occupied, and red the most occupied. This pictures is taken from `http://www.autobahn.nrw.de/`.

ability to model flow across a variety of junctions would also be necessary. Mapping of road systems could be readily provided with current route planning resources, such as satellite navigation. Multiple extensions of basic traffic models could be used to determine behaviour at certain junctions. For example, a 4-directional cross-roads as an extension of the BML model [39], or a fluid-dynamic approach to roundabouts [26]. Although, many other junction configurations exist and have yet to be modelled successfully.

It could be argued that it will never be possible to effectively predict how traffic will evolve over an extended period of time as vehicular traffic is an intrinsically human problem. Stress, driving skill and experience are all factors that cause differing driving behaviour, the combination of which can result in significant differences in traffic flow. A simple example to show the impact of human drivers is the caterpillar effect. This is when the drivers, whilst trying to stay at a constant speed, inevitably deviate slightly and cause random traffic jams. The creation of these 'phantom' traffic jams highlights the unpredictability that human drivers introduce and calls into question the validity of any models which do not take this into account.

Due to the vast spectrum of human characteristics, factoring these into equations for models is complicated. Nevertheless, attempts have been made to determine the effect of different driving techniques. For example, in an extension of the BML model the behaviour of violators (vehicles not stopping at red traffic lights) was shown to increase the overall speed of traffic, but cause jams to occur more often than usual [40]. An attempt was also made to adjust for these discrepancies in kinetic models, where they introduced an "activity variable" [18].

New technology could provide a way of eliminating this human error, for example driverless car technology, such as Google's variant [46]. These cars would eliminate the random aspect plaguing the accuracy of discrete models. This would allow discrete models to be adjusted so that they completely and simplistically model the behaviour of the traffic. If all cars have the same database, the routes of all vehicles could be known. This would enable careful planning to minimise the occurance of jams, making driving a much more predictable and time efficient mode of transport. Additionally these cars could drastically reduce traffic accidents and make city traffic flow much more efficient, which is vital in an age where limited resources highlight the need for economical fuel consumption. There could, however, be a complicated transition between now, where there are mainly human-driven cars, to an age when the roads are dominated by driverless vehicles. Despite this, we believe this is genuinely a viable solution rather than a science fiction ambition.

7 Other Approaches to Solving Traffic Flow Problems

In the following two sections a general description of mechanisms that can be used in traffic flow optimisation, other than modelling, are presented.

7.1 Economic Policies

As previously explained, traffic congestion can be seen as a human problem and - by extension - could have a human solution. Congested roads can be seen as an example of the 'tragedy of commons' [47]. As roads are usually free to use, there is no incentive not to use them up to the point of a traffic jam. It can be argued that, by changing the attitudes of road users, jam occurrences can be reduced by limiting the density of cars on the road - hence stopping the problem at the source. This may seem an unreasonable proposition at first but, according to the 2009 US Census, 76% of American commuters drive to work alone which, if reduced, could have a significant impact on the amount of traffic on roads and the subsequent jams [48].

One way to reduce the number of cars on the roads is through the introduction of high occupancy ('carpool') lanes. These are designated lanes for cars with more than a given number of passengers. This has been widely applied in Canada, with particular success in Calgary where average occupancy of cars has gone from 1.32 to 1.39, corresponding to a 32% reduction in the number of cars on the road [49]. It has also been applied in Australia in the opening of the South East Busway in Brisbane. This scheme reportedly saved the taxpayer $62 million (Australian) and increased property values in the area by 20% as a result of the improved transport links [49]. However, one of the drawbacks of these high occupancy lanes is the need to enforce restrictions which, in most cases, takes the form of fines.

Another method for reducing the volume of cars on the road is the use of toll roads. The basic concept is that vehicles have to pay a small charge for the use of these roads. Whilst the introduction of these charges has been somewhat controversial - an example being the establishment of the National Alliance Against Tolls [50] - there have been studies that show the benefits of toll roads in practice. In particular, data taken from a newly installed toll road in Seattle reveals a 3.5% reduction in travel time and an estimated 12% reduction in greenhouse gas emissions [51]. A similar approach is the use of congestion charges in cities, such as in London [52].

7.2 Mathematical Optimisation

Related to the economic approaches mentioned above, mathematical optimisation is a method that has been used extensively to study and forecast the optimality of traffic flow [23]; it has also been used in the study of congestion tolls [53]. In this approach, the cost of travelling on a particular segment of connected road is represented as a function. These cost functions can then be combined to form a total cost function. It is assumed that the drivers behave rationally, and want to choose the best possible scenario for themselves in order to optimise their costs - money, time and other expenditures.

To demonstrate the approach qualitatively, consider that each road requires investment from the government for maintenance. This expense is included in the roads individual cost function. Using economic assumptions [53], a cost function for each, called the Link Travel Cost is created. Finally, the cost function for each segment of road is combined under the further assumption that traffic flow variables within each segment are constant. This gives a Total Travel Cost function. The form of the Total Travel Cost function can be chosen such that it has a unique, and easily obtainable, minimum value. Hence the choice of available routes reduces to a simple minimisation problem.

8 Discussion of Traffic Flow Modelling

Throughout this report, three approaches to traffic modelling, which are popular among traffic scientists and engineers, have been presented. Each of the approaches have different advantages and disadvantages. In our opinion, the applicability of a model, and appeal to scientists or engineers, can be considered from three different perspectives. The ability to fundamentally predict behavior at different scales; numerical difficulty and impact of their predictions.

8.1 Scales of Modelling

Systems of considerable size and complexity can be described by macroscopic or microscopic variables. Macroscopic descriptions use a small number of averaged values in order to build a broad representation of the system. Conversely, a microscopic perspective provides a complete description of the system's constituents. Kinetic and fluid models are characterised by macroscopic quantities such as average velocities and flows presented in Sections 3 and 4. Although these general properties are more analytically manageable, they often contain limited information for use in applications. The inability to access localised information, such as regional densities, makes the use of these models in areas such city planning and road development inviable. The microscopic description of discrete models (Section 5.2.1) means that the exact state of each vehicle is known in terms of its position and velocity. In practice, this approach allows more informative conclusions to be drawn from the simulation. However, this precise description of individual vehicles could be computationally expensive to model.

8.2 Mathematical and Computational Difficulty

Historically, the first models introduced were fluid models as noted in Section 4. These models are based on continuity equations and handle traffic with a certain predictability. Whilst being useful for setting up problems, and behaving (to some degree) like real traffic, these problems have proved to be mathematically very complicated. Their theoretical evaluation gets difficult extremely quickly, and analytical solutions are often unobtainable. Additionally, computational solutions tend to be one of the most difficult, numerical PDEs in particular, as was described in Section 4.1.

It is natural to ask whether the practicality of the fluid models is hindered due to their insolvable nature. Is it better to change the paradigm of the field completely? Instead of formulating the fluid model, the problem could be broken down into a number of simpler ones. As the solutions to many differential equations are too difficult to obtain, as shown in Section 4.3, we believe that this field should not be the principle recipient of funding. However, we believe innovative research that tries to formulate these problems into more tractable settings should be supported in some form.

A similar change in paradigm, that was eluded to in the above section, is behind the creation of kinetic models (Section 3). Kinetic models are based on continuity equations, similar to fluid models. However they assume a dilute limit, which significantly simplifies their formulation, hence many problems can be solved analytically. This type of traffic modelling still remains an active area of research, as mentioned in Section 3.1. However, with the ability to achieve quick analytical solutions comes other problems, such as lower predictive power.

On the other hand, discrete models, which are inherently computational, do not principally aim to have analytic solutions. As a result, they perform much better in terms of mathematical tractability. They are usually automaton simulations, most of which can be run for a long time on a small scale. Large scale simulation requires significant computing power. However, due to rapidly increasing computational power, we believe that this is likely to be a weak constraint in the future. Therefore, as in Section 5.6, we conclude that discrete models are not only of academic interest to scientists but have significant potential for further development in practice.

8.3 Predictability Impact

As eluded to in Sections 3.3 and 8.2, the use of kinetic models in traffic engineering or city planning is extremely limited. Their inability to model dense traffic, and the nature of the assumptions made in their formulation, create heavy restrictions on their practical applications , for example, predicting traffic jams. On the other hand, fluid models were the most popular methods to simulate traffic flow before the advent of cheap computation. They have been used in various cases, particularly the revolutionary Lighthill-Whitham model. However more recently, criticism has emerged (such as that by Kerner [23]) where it is shown that the LW model does not satisfy the empirical assumptions that were listed in Section 4.3.

Currently the most researched, and useful, topics in traffic modelling is the discrete models, which use automaton as a theoretical basis (Section 5). These include models which are based on seminal papers by Nagel and Schrekenberg [5] and Biham et.al. [35]. They are remarkably successful in optimising traffic, for example, by including addtional traffic lights jamming denisities were shown to decrease (Section 5.3.3). These theoretical findings could be used directly in the planning or designing of traffic flow solutions. Although the engineering models such as TRANSSIM [42] or VANETs [44], that are discussed in Section 5.6, deviate significantly from basic NS model, they follow the same underlying principles. Due to the simplicity of the models there appears to be considerable room for expansion. With respect to city planning, and real-life vehicular traffic, we feel that funding should focus on discrete models rather than continuum, as there seems more potential for progress. For example, the modelling of new technology such as the aforementioned VANETs or even self driving vehicles as discussed in Section 6.

8.4 General Discussion

It has been argued in Section 2.1, from an economical standpoint, that improving the condition of traffic flow is economically beneficial for the individual as well as the collective. Because of this, we are of the opinion that a UK Research Institute into traffic flow modelling would be valuable. The above three sections summarised the critical evaluation of the three types of traffic flow modelling on which we have focused in our review: kinetic, fluid and discrete. The review revealed that the most applicable models, and indeed those that could have the greatest potential impact, are the discrete models. Therefore, it is felt that discrete models would be of most use due to their simplicity, versatility of prediction and possible development for the future. It should be noted that although kinetic and fluid models are limited, either due to their restrictive assumptions or their inability to be solved analytically, they still pose interesting and active fields of academic research, as has been pointed out in Section 8.2.

Finally, as previously stated, traffic flow modelling is not purely confined to vehicular traffic. The strong underlying basis in statistical physics means these models can be translated to many applied fields. For example, automata models have been used to model the growth and characteristics of cancerous tumours [54].

9 Conclusion

In this review paper, we have presented a theoretical outline of three main approaches used to model vehicular traffic flow: kinetic, fluid and discrete. These models are, to various degrees, applicable in real traffic flow modelling and have been thoroughly reviewed in this work. The models were compared on their predictive ability and computational complexity. The impact of these models was discussed in relation to road engineering and consequently society. A proposal for future development has been made at the end of the report, and an evaluation of funding efficiency has been presented for each approach to vehicular traffic modelling. After a careful literature review, it is concluded that, under the proposal of a UK Research Institute into traffic flow modelling, funding should be principally directed into the field of discrete models due to their predictability and impact on traffic flow modelling in practice.

References

[1] Eshel Ben-Jacob. From snowflake formation to growth of bacterial colonies ii: Cooperative formation of complex colonial patterns. *Contemporary Physics*, 38(3):205–241, 1997.

[2] Eshel Ben-Jacob. From snowflake formation to growth of bacterial colonies: Part i. diffusive patterning in azoic systems. *Contemporary Physics*, 34(5):247–273, 1993.

[3] LM Gramani Cumin. On kinematic waves i flood movement in long rivers. ii a theory of trac ow on long crowded roads. *International Journal of Non Linear Mechanics*, 44:263–268, 2009.

[4] Ilya Prigogine and Robert Herman. Kinetic theory of vehicular traffic. Technical report, 1971.

[5] Kai Nagel and Michael Schreckenberg. A cellular automaton model for freeway traffic. *Journal de physique I*, 2(12):2221–2229, 1992.

[6] Ashish Chandra John Siraut. Economic impact of traffic signals.

[7] Steer Davis Gleave. Toll impact study - final report.

[8] Debashish Chowdhury, Ludger Santen, and Andreas Schadschneider. Statistical physics of vehicular traffic and some related systems. *Physics Reports*, 329(4):199–329, 2000.

[9] Physica a: Description of journal contents, Accessed Jan 2015. `http://www.journals.elsevier.com/physica-a-statistical-mechanics-and-its-applications/`.

[10] Philip W Anderson, Kenneth Joseph Arrow, and David Pines. *The economy as an evolving complex system*. Addison Wesley Publishing Company, 1988.

[11] Stuart Kauffman. *At home in the universe: The search for the laws of self-organization and complexity*. Oxford university press, 1995.

[12] V Coscia, M Delitala, and Paolo Frasca. On the mathematical theory of vehicular traffic flow ii: Discrete velocity kinetic models. *International Journal of Non-Linear Mechanics*, 42(3):411–421, 2007.

[13] Ma.utexas.edu. Boltzmann equation - mwiki, Acessed Jan 2015. `https://www.ma.utexas.edu/mediawiki/index.php/Boltzmann_equation`.

[14] SL Paveri-Fontana. On boltzmann-like treatments for traffic flow: a critical review of the basic model and an alternative proposal for dilute traffic analysis. *Transportation Research*, 9(4):225–235, 1975.

[15] Libra.msra.cn. Kinetic theory of vehicular traffic, Acessed Jan 2015. `http://libra.msra.cn/Publication/2010081/kinetic-theory-of-vehicular-traffic`.

[16] Dirk Helbing and Andreas Greiner. Modeling and simulation of multilane traffic flow. *Physical Review E*, 55(5):5498, 1997.

[17] D Ngoduy. Application of gas-kinetic theory to modelling mixed traffic of manual and acc vehicles. *Transportmetrica*, 8(1):43–60, 2012.

[18] LM Gramani Cumin. On the modeling of granular traffic flow by the kinetic theory for active particles trend to equilibrium and macroscopic behavior. *International Journal of Non Linear Mechanics*, 44:263–268, 2009.

[19] Nicola Bellomo. *Modeling complex living systems: a kinetic theory and stochastic game approach*. Springer, 2008.

[20] Trafficwaves.org. Science hobbyist: Traffic waves, physics for bored commuters, Accessed Jan 2015.

[21] Daiheng Ni. *Equilibrium Traffic Flow Models*. University of Massachusetts, 1 edition.

[22] Carlos F Daganzo. A continuum theory of traffic dynamics for freeways with special lanes. *Transportation Research Part B: Methodological*, 31(2):83–102, 1997.

[23] BS Kerner. Criticism of generally accepted fundamentals and methodologies of traffic and transportation theory. In *Sixth International Conference on Nonlinear Mechanics (ICNM-6)*, page 21. DEStech Publications, Inc, 2013.

[24] Markos Papageorgiou. Some remarks on macroscopic traffic flow modelling. *Transportation Research Part A: Policy and Practice*, 32(5):323–329, 1998.

[25] Carlos F Daganzo. A behavioral theory of multi-lane traffic flow. part i: Long homogeneous freeway sections. *Transportation Research Part B: Methodological*, 36(2):131–158, 2002.

[26] Fluid-dynamic model at the junctions, Accessed Jan 2015. `http://www.iac.rm.cnr.it/~bretti/junctions.pdf`.

[27] Harold J Payne, Warren A Thompson, and Leif Isaksen. Design of a traffic-responsive control system for a los angeles freeway. *Systems, Man and Cybernetics, IEEE Transactions on*, (3):213–224, 1973.

[28] M Cremer and J Ludwig. A fast simulation model for traffic flow on the basis of boolean operations. *Mathematics and Computers in Simulation*, 28(4):297–303, 1986.

[29] Gianpiero Cattaneo, Michele Finelli, and Luciano Margara. Investigating topological chaos by elementary cellular automata dynamics. *Theoretical computer science*, 244(1):219–241, 2000.

[30] Robert Jackson. Rule 30: More hacking..., Accessed Jan 2015. `http://robertjackson.info/index/2011/12/rule-30-more-hacking-ooo-and-all-that/`.

[31] Kai Nagel and Hans J Herrmann. Deterministic models for traffic jams. *Physica A: Statistical Mechanics and its Applications*, 199(2):254–269, 1993.

[32] Kai Nagel and Hans J Herrmann. Deterministic models for traffic jams. *Univ. Cologne Appl. Math. WP 93.133*, 2008.

[33] Marcus Rickert, Kai Nagel, Michael Schreckenberg, and Andreas Latour. Two lane traffic simulations using cellular automata. *Physica A: Statistical Mechanics and its Applications*, 231(4):534–550, 1996.

[34] Kai Nagel, Dietrich E Wolf, Peter Wagner, and Patrice Simon. Two-lane traffic rules for cellular automata: A systematic approach. *Physical Review E*, 58(2):1425, 1998.

[35] Ofer Biham, A Alan Middleton, and Dov Levine. Self-organization and a dynamical transition in traffic-flow models. *Physical Review A*, 46(10):R6124, 1992.

[36] Omer Angel, Alexander E Holroyd, and James B Martin. The jammed phase of the biham-middleton-levine traffic model. *Electronic Communications in Probability*, 10(17):167–178, 2005.

[37] Zhong-Jun Ding, Rui Jiang, and Bing-Hong Wang. Traffic flow in the biham-middleton-levine model with random update rule. *Physical Review E*, 83(4):047101, 2011.

[38] Raissa M D'Souza. Geometric structure of coexisting phases found in the biham-middleton-levine traffic model. *arXiv preprint cond-mat/0402351*, 2004.

[39] Ding-wei Huang and Wei-neng Huang. Biham–middleton–levine model with four-directional traffic. *Physica A: Statistical Mechanics and its Applications*, 370(2):747–755, 2006.

[40] Zhong-Jun Ding, Xiao-Yan Sun, and Bing-Hong Wang. Violating traffic light behavior in the biham-middleton-levine traffic flow model. *Procedia engineering*, 31:1072–1076, 2012.

[41] GQ Gu, KH Chung, and PM Hui. Two-dimensional traffic flow problems in inhomogeneous lattices. *Physica A: Statistical Mechanics and its Applications*, 217(3):339–347, 1995.

[42] Kai Nagel, Paula Stretz, Martin Pieck, Rick Donnelly, and Christopher L Barrett. Transims traffic flow characteristics. *arXiv preprint adap-org/9710003*, 1997.

[43] CL Barrett, RJ Beckman, KP Berkbigler, KR Bisset, BW Bush, S Eubank, JM Hurford, G Konjevod, DA Kubicek, MV Marathe, et al. Transims (transportation analysis simulation system). volume 0: Overview. *Los Alamos National Lab., Los Alamos, NM, Rep. LA-UR-99-1658*, 1999.

[44] Florian Knorr, Daniel Baselt, Michael Schreckenberg, and Martin Mauve. Reducing traffic jams via vanets. *Vehicular Technology, IEEE Transactions on*, 61(8):3490–3498, 2012.

[45] University of Duisburg. Experimental research and simulation of the impact of traffic toll on highways in germany, Accessed Jan 2015. `https://www.uni-due.de/ptt/projekte/esam_start.shtml`.

[46] Lawrence D Burns. Sustainable mobility: a vision of our transport future. *Nature*, 497(7448):181–182, 2013.

[47] Garrett Hardin. The tragedy of the commons. *science*, 162(3859):1243–1248, 1968.

[48] Brian McKenzie and Melanie Rapino. *Commuting in the United States: 2009.* US Department of Commerce, Economics and Statistics Administration, US Census Bureau, 2011.

[49] C Dixon and K Alexander. Literature review of hov lane schemes, 2005.

[50] Notolls.org.uk. National alliance against tolls - home page, Accessed Jan 2015. `http://www.notolls.org.uk/index.htm`.

[51] Steven Danna, Keibun Mori, Jake Vela, and Michelle Ward. A benefit-cost analysis of road pricing in downtown seattle. *Evans School Review*, 2:26–46, 2012.

[52] Rémy Prud'homme and Juan Pablo Bocarejo. The london congestion charge: a tentative economic appraisal. *Transport Policy*, 12(3):279–287, 2005.

[53] Stella Dafermos and Frederick T Sparrow. Optimal resource allocation and toll patterns in user-optimised transport networks. *Journal of Transport Economics and Policy*, pages 184–200, 1971.

[54] Werner Duchting. Tumor growth simulation. *Chaos and Fractals: A Computer Graphical Journey*, page 15, 1998.

A Failure in Prigogine Model Proof

Assume that all vehicles are travelling at v_{des}. At time $t = 0$ let the function be of Gaussian form.

$$f(x, v, 0) = NF_{des}(v)\frac{1}{\sigma\sqrt{\pi}}\exp\left(\frac{x^2}{\sigma^2}\right) \tag{22}$$

Since all vehicles are travelling at their desired speeds which do not change with time:

$$f(x, v, t) = NF_{des}(v)\frac{1}{\sigma\sqrt{\pi}}\exp\left(\frac{(x - vt)^2}{\sigma^2}\right) \tag{23}$$

Prigogine claims that $f_{des} = F_{des}(v)c(x, t)$ where $c(x, t) = \int_0^\infty dv f(x, vt)$

Using the expressions above for $f(x, v, t)$:

$$\frac{f_{des}(x, v, t)}{c(x, t)} = F_{des}(v)\frac{\exp\left(\frac{(x-vt)^2}{\sigma^2}\right)}{\int_0^\infty f_{des}(v)\exp\left(\frac{(x-vt)^2}{\sigma^2}\right)} \tag{24}$$

Which $\neq F_{des}(v)$. Therefor a non varying F_{des} does not imply $f_{des} = F_{des}(v)c(x, t)$.

Proof Adapted from [14]