Keywords: Three phase traffic theory; Conservation law; Phase transition

Nomenclature

- \( t \): Time (s)
- \( u \): Vehicle speed (\( \text{ms}^{-1} \))
- \( a(\rho, u) \): Speed adaptation coefficient
- \( \mu \): Average acceleration (\( \text{ms}^{-2} \))
- \( \rho(x,t) \): Density of traffic in a continuum (vehicles/m)
- \( \tau \): Relaxation time (s)
- \( d \): Vehicle length (m)
- \( q=\rho u \): Flow rate (vehicles/hour)
- \( x \): Displacement (m)
- \( P \): Traffic pressure (N/m\(^2\))
- \( A(U) \): Jacobian matrix
- \( F \): Free traffic flow
- \( C \): Congested traffic
- \( S \): Synchronised flow
- \( J \): Wide moving jam

- \( F \rightarrow S \): Phase transition from free flow phase to synchronized flow phase
- \( S \rightarrow J \): Phase transition from synchronized flow phase to wide moving jam phase

\[
\frac{\partial \rho(x,t)}{\partial t} : \text{Rate of change of traffic density with time}
\]
\[
\frac{\partial q(x,t)}{\partial x} : \text{Rate of change of traffic flow with space}
\]

Introduction

Transport forms a key component of creating a competitive business environment as well as a means through which various economic, social and environmental objectives of a nation can be achieved. Congestion of vehicular traffic experienced within urban areas has adverse effects on people’s quality of life due to delays, accidents and environmental pollution. An efficient transport system that minimizes travel times as well as externalities should be put into consideration. One way of eliminating the problem of traffic congestion is to increase the capacity of existing roadways by developing transport corridors, surface streets and by-passes. However, this is not always practical due to limited financial resources, space, environmental effects and local politics.

For effective traffic management and control, proper understanding of traffic congestion is needed. These are the insights into what causes congestion, what determines the time and location of traffic breakdown and how the congestion propagates through the road network. In addition, spatiotemporal behavior of empirical traffic congested patterns should also be studied closely. This is because traffic congestion is observed to take place in space and time in form of spatiotemporal congested traffic patterns that propagate within roadways. For this reason, huge numbers of publications that are devoted to empirical studies of traffic congestion and associated traffic flow theories have been done.

The earliest well-articulated theories of traffic flow is known as the kinematic wave theory, LWR traffic flow congestion theory, and classical flow congestion by James Lighthill and Gerald Whitham.
The theory combines three elements: the fundamental identity of traffic flow theory, that flow is the product of density and velocity, the equation of continuity, which is the conservation of mass for a fluid and an assumed technological relationship between velocity and density.

This being a first order approximate model of traffic flow dynamics has a number of serious deficiencies; firstly, the model predicts infinite deceleration when a vehicle crosses a shock. This reveals that acceleration or deceleration of a vehicle stream is proportional to traffic concentration and concentration gradient. Secondly, the model assumes that the equilibrium speed-concentration also holds for non-Equilibrium traffic. In reality, traffic flow is hardly in equilibrium and its dynamics is a result of the retarded response of drivers to various frontal stimuli. Thirdly, the model lacks a mechanism for traffic to accelerate or decelerate at a finite speed when the concentration gradient is large.

Due to these deficiencies, Payne [2] derived an equation of motion from car-following theory, response=sensitivity $\times$ stimulus. The model is called PW-models.

The Payne’s model together with its computer implementation aroused considerable interest in higher-order continuum traffic flow models. However, the application of this model, reported mixed results.

Daganzo [3], points out that most of the problems were attributed to the fundamental flaw that the model produces a ‘wrong way travel’ that is negative travel speed.

Daganzo [3] pointed out that information in the PW model can travel faster than vehicles. There was an enhanced continuum traffic flow theory that removes certain deficiencies of the LWR theory without introducing new flaws. In deriving the new theory Zhang neglected higher order terms. This raised concern by Daganzo [3] as to whether one can neglect higher order terms when the concentration on the road is rapidly changing, as occurs near a shock path. However he showed that these higher order terms can be neglected if the concentration function $\rho(x,t)$ is well behaved and temporal spatial scales are properly treated.

Daganzo [3] argued that second order models violate the principle that a car is anisotropic particle and responds to frontal stimuli. This motivated Rascle [4] to develop a model with which would rectify the above inconsistencies.

The Aw-Rascle model respects all the natural requirements (frontal stimuli) and the inconsistencies of ‘wrong way travel’ of second order models disappear.

Klar et al. [5] derived a macroscopic traffic flow model from a macroscopic follow the leader models. The model obtained by Klar was of Aw-Rascle type.

Krhithi et al. [6] developed a macroscopic traffic flow model based on Kerner [7] theory using Integral-Differential Equations of kinetic models. The model has the following equations.

$$\frac{\partial \rho}{\partial t} + \frac{\partial (\rho u)}{\partial x} = 0$$

$$\frac{\partial (\rho u)}{\partial t} + \frac{\partial (\rho u^2)}{\partial x} - A(\rho, u) \frac{\partial u}{\partial x} = pR(\rho, u)$$

$R(\rho, u)$ is the relaxation term which describes the tendency of traffic flow to relax to an equilibrium velocity.

$$A(\rho, u) = \frac{C\rho}{(1 - \rho)}$$

Where $A(\rho, u) = \frac{C\rho}{(1 - \rho)}$.

Kerner [7] in his work describes traffic in several major cities in the world in three phases; free flow, wide moving jam and synchronized flow. In our Kenyan roads the three phases exist and knowledge on how the phase transition takes place can be a great contribution in traffic management. The speed of the vehicles in the synchronized and wide moving jam is influenced by the aggression of the driver. The same driver aggression influences the phase transition from free to congested region.

In this study, we develop a macroscopic traffic flow model motivated by Kerner [7] traffic flow theory with considerations of driver aggressiveness as experienced on Kenyan roads such Thika superhighway.

We shall formulate the driver aggression as a function of velocity and develop a model to investigate how the driver aggression can influence the phase transitions from Free Flow to Synchronized Flow and then to Wide Moving Jam. The model will help us deduce the causes of congestion in Kenyan roads in order to design comfortable and safe roads, to solve road congestion problems and to design adequate traffic management measures.

Driver aggressiveness

A driving behavior is aggressive if it is deliberate, likely to increase the risk of collision and is motivated by impatience, annoyance, hostility and/or an attempt to save time. These driving actions that markedly exceed the norms of safe driving behavior, directly affect other road users by placing them in unnecessary danger.

All these behaviors are exacerbated by the stress and time pressures of modern life. Increasingly crowded and congested roads also lead to feelings of frustration and are responsible for cases of aggressive driving and lack of respect for other drivers such as illegal use of hard shoulder, changing lanes without indicating or preventing other vehicles from entering a traffic lane.

In our study we look at the driver aggressiveness as a result of lane-drop bottleneck from 3 lanes to 2 lanes. In normal circumstances drivers absolves themselves to the two proceeding lanes far upstream of the bottleneck. This does not cause much traffic jam in case of a lane-drop bottleneck. An aggressive driver on the other hand will take advantage of the free third lane and drive all the way to merge at the lane-drop bottleneck. This does not cause much traffic jam in case of a lane-drop bottleneck. An aggressive driver on the other hand will take advantage of the free third lane and drive all the way to merge at the lane-drop bottleneck. This behavior forces other drivers to emergency braking and as a result traffic jam forms from the lane-drop point and grows upstream of the highway.

Governing equations

The equations governing flow in traffic flow are based on the momentum and conservation of mass. The mathematical modeling of traffic flow often rests on a fluid flow analogy, treating the traffic stream as a two dimensional compressible fluid from which it can be deduced that traffic flow is conserved. The number of vehicles entering a certain region equals the number leaving the same region. Gas kinetic traffic flow models were first proposed by Prigogine [8] and are based upon the analogy between gas flows and traffic flows. Where in the former case the dynamics are governed by interacting gas particles the latter deals with interacting vehicles. In order to realistically describe traffic flows the specification of the vehicle interactions must obviously differ from that of gas particles, the main difference being that drivers do not behave according to physical laws.
Continuity equation

A continuity equation is a differential equation that describes the transport of some kind of conserved quantity, in particular-mass. The continuity equation expresses the idea that the total amount inside any region can only change by the amount that passes in or out of the region through the boundary. A conserved quantity cannot increase or decrease, it can only move from place to place.

In traffic flow, the number of vehicles is conserved hence we use the equation of continuity. In our study we are considering a multilane road with three lanes but one lane is varnishing at some point leaving two lanes proceeding.

\[ \frac{\partial \rho(x,t)}{\partial t} + \frac{\partial q(x,t)}{\partial x} = 0 \]

The equation expresses the relation between the rates of change of the density with respect to time and flow with respect to space.

Dynamic velocity equations

This equation is used to describe time varying and spatially varying average velocities \( u(x,t) \) such as those that occur in traffic jams or stop-jams, Helbing [9].

\[ \frac{\partial q(x,t)}{\partial t} = \beta \left( \rho \neu - a(\rho) \frac{\partial u}{\partial x} \right) \]

Where \( a(\rho) \) is the speed adaptation coefficient.

The equation by Kimathi et al. (2012) is as follows;

\[ \frac{\partial q(x,t)}{\partial t} = \beta \left( \rho \neu - \frac{\rho}{1-\rho} \frac{\partial u}{\partial x} \right) \]

The variables, \( \beta \) is a measure of driver aggressiveness and in our study we shall approximate it as a function of velocity as follows:

\[ \beta = Ae^u \]

Where \( A \) is a constant.

The range will be \( 0.3 < \beta < 1 \). As \( \beta \) tends to zero, the model produces the negative travel velocities. If \( \beta=0.3 \), the model becomes three phase traffic flow, if \( \beta=1 \), the model becomes Aw Rascle.

Substituting for \( \beta \) in the above equation into we get

\[ \frac{\partial \nu u}{\partial t} + \frac{\partial \nu u^2}{\partial x} - A \left( e^u \frac{\partial u}{\partial x} \right) = 0 \]

Features of the aggressive model

We can re-write the model equations in non-conservative form in terms of \( \tau \) and \( u \) as follows

\[ \frac{\partial \rho(x,t)}{\partial t} + \frac{\partial q(x,t)}{\partial x} = 0 \]

\[ \frac{\partial u}{\partial t} + \left( u - Ae^u \rho \right) \frac{\partial u}{\partial x} = 0 \]

Using the expression (1) while substituting C for A on (2) yields

\[ \frac{\partial u}{\partial t} + \left( u - e^u \rho \right) \frac{\partial u}{\partial x} = 0 \]

Comparing (4) with the Aw Rascle model, its non-conservative form is

\[ \frac{\partial \rho(x,t)}{\partial t} + \frac{\partial q(x,t)}{\partial x} = 0 \]

In order to express the equations in conservative form we multiply (2) with \( p' \) to obtain \( p' \frac{\partial \rho}{\partial t} - p \frac{\partial p}{\partial t} - p' \frac{\partial \rho}{\partial x} \). Substituting the result in (4) and multiplying by \( e^u \), we get

\[ e^u \frac{\partial u}{\partial t} + We \frac{\partial u^2}{\partial x} + p' \frac{\partial p}{\partial t} + p \frac{\partial e^p}{\partial t} = 0 \]

Grouping this expression gives

\[ \partial \left( \rho e^p \right) + \partial \left[ pu(e^p + p) \right] = 0 \]

Multiplying (1.1) by \( (e^p + p) \) we get

\[ (e^p + p) \partial \rho + (e^p + p) \partial u = 0 \]

Multiply (5) by \( \rho \) and add result to (6) to get the conservative form after regrouping the terms:

\[ \frac{\partial \rho}{\partial t} + \frac{\partial \nu u}{\partial x} = 0 \]

\[ \partial \left( \rho e^p \right) + \partial \left[ pu(e^p + p) \right] = 0 \]

Where the conservative variables are \( \rho \) and \( y = pe^\rho + pp \)

To obtain the conserved form of Aw-Rascle model, we multiply (4) by \( \rho \) and

\[ \partial \rho \left( u + p(\rho) \right) + u \partial \left( \rho u + p(\rho) \right) = 0 \]

By \( u+\rho(p) \) and add up the two equations to get

\[ \partial \rho \left( u + p(\rho) \right) + \partial \left( \rho u + p(\rho) \right) = 0 \]

Therefore the conservative variables of the model are \( \rho \) and \( y = pu + pp(\rho) \)

The properties of the system are largely dictated by the Eigen values of the Jacobian matrix \( A(U) \), and are determined by the characteristic polynomial \( \det(A - \lambda I) = 0 \).

Expanding the new model (2) and (4) in the form

\[ \frac{dU}{dt} + A(U) \frac{dU}{dx} = 0 \]

We have

\[ U = \begin{bmatrix} u \\ p \end{bmatrix} \]

\[ A(U) = \begin{bmatrix} 0 & \rho \\ 0 & u - e^\rho p(\rho) \end{bmatrix} \]

So that

\[ u - \lambda \begin{bmatrix} 0 & \rho \\ 0 & u - e^\rho p(\rho) - \lambda \end{bmatrix} = 0 \]

This yield

\[ \lambda_1(u) = u - e^\rho p(\rho) \]

Expanding the equation gives:

\[ \lambda_2(u) = (2u - e^\rho p(\rho) + e^\rho p(\rho)) \]

Applying the quadratic formula gives

\[ \lambda = \frac{- (2u - e^\rho p(\rho) + e^\rho p(\rho)) \pm \sqrt{(2u - e^\rho p(\rho) + e^\rho p(\rho))^2 - 4(0 - (2u - e^\rho p(\rho))} \]

leading to

\[ \lambda_1(u) = u - e^\rho p(\rho) \] and \( \lambda_2(u) = u + e^\rho p(\rho) \)
These Eigen values are the characteristic speeds that govern the propagation of information in the traffic system. The largest Eigen value is equal to the flow velocity. This now means no traffic information travels faster than the traffic and therefore the anisotropic character of vehicular traffic flow is preserved.

The Eigen values are real and distinct hence the system of equations is purely hyperbolic.

Expressing the Aw-Rascle model in the form (12)

We have \( \mathbf{U} = \begin{bmatrix} u_1 \\ u_2 \\ \rho \end{bmatrix} = \begin{bmatrix} \rho \\ u_1 - \rho p(\rho) \end{bmatrix} \) and

\[
A(\mathbf{U}) = \begin{bmatrix} 0 & \rho \\ -u_2 & 0 \end{bmatrix}
\]

The Eigen values become;

\[
\lambda_1 = 1 - \rho p'(\rho) \quad \lambda_2 = 1 - \rho p'(\rho)
\]

This shows that the system is purely hyperbolic and the anisotropic character of vehicular traffic is preserved.

We now turn to other hyperbolic features of the systems.

We refer to the waves associated with \( \lambda_1 \) as 1-wave and those associated with \( \lambda_2 \) as 2-waves. In order to determine those waves, we calculate the right eigenvectors \( R^{(i)} = (r_{1i}, r_{2i}) \) of the matrix \( A(\mathbf{u}) \) corresponding to Eigen values \( \lambda_i \). We now evaluate \( R^{(i)}(\mathbf{u}) \)

\[
\begin{bmatrix} u \\ 0 \\ u - e^{-\lambda_i} \rho p(\rho) \end{bmatrix} = \begin{bmatrix} r_1 \\ r_2 \\ r_3 \end{bmatrix}
\]

This means that

\[
\begin{bmatrix} u_1 \\ u_2 \\ \rho \end{bmatrix} = \begin{bmatrix} (u - e^{-\lambda_i} \rho p(\rho))r_1 \\ (u - e^{-\lambda_i} \rho p(\rho))r_2 \\ (u - e^{-\lambda_i} \rho p(\rho))r_3 \end{bmatrix}
\]

Leading to \( r_1 = r_2 = (e^{-\lambda_i} p(\rho)) \) and \( r_2 = 1 \) hence we have

\[
R^{(i)}(\mathbf{u}) = \begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \end{bmatrix}
\]

This shows that the system is purely hyperbolic and the anisotropic character of vehicular traffic is preserved.

We now turn to other hyperbolic features of the systems.

We refer to the waves associated with \( \lambda_1 \) as 1-wave and those associated with \( \lambda_2 \) as 2-waves. In order to determine those waves, we calculate the right eigenvectors \( R^{(i)} = (r_{1i}, r_{2i}) \) of the matrix \( A(\mathbf{u}) \) corresponding to Eigen values \( \lambda_i \). We now evaluate \( R^{(i)}(\mathbf{u}) \)

\[
\begin{bmatrix} u \\ 0 \\ u - e^{-\lambda_i} \rho p(\rho) \end{bmatrix} = \begin{bmatrix} r_1 \\ r_2 \\ r_3 \end{bmatrix}
\]

This means that

\[
\begin{bmatrix} u_1 \\ u_2 \\ \rho \end{bmatrix} = \begin{bmatrix} (u - e^{-\lambda_i} \rho p(\rho))r_1 \\ (u - e^{-\lambda_i} \rho p(\rho))r_2 \\ (u - e^{-\lambda_i} \rho p(\rho))r_3 \end{bmatrix}
\]

Leading to \( r_1 = r_2 = (e^{-\lambda_i} p(\rho)) \) and \( r_2 = 1 \) hence we have

\[
R^{(i)}(\mathbf{u}) = \begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \end{bmatrix}
\]

Similarly for \( R^{(i)}(\mathbf{u}) \)

\[
\begin{bmatrix} u \\ 0 \\ u - e^{-\lambda_i} \rho p(\rho) \end{bmatrix} = \begin{bmatrix} r_1 \\ r_2 \\ r_3 \end{bmatrix}
\]

Yielding to \( r_2 = 0 \)

The Eigen vectors for the Aw-Rascle model becomes

\[
R^{(1)}(\mathbf{u}) = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}
\]

\[
R^{(2)}(\mathbf{u}) = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix}
\]

We now determine the kind of waves associated to each Eigen value \( \lambda_i = 1,2 \) by checking whether the dot product \( \nabla \lambda_i(\mathbf{u}); R^{(i)}(\mathbf{u}) \) vanishes or not.

For \( \lambda_1 \) we have

\[
\begin{bmatrix} \frac{\partial_c \lambda_1} {\partial \mathbf{U}} \\ \frac{\partial_c \lambda_2} {\partial \mathbf{U}} \end{bmatrix} = \begin{bmatrix} 1 \\ -e^{-\lambda_i} p(\rho) \end{bmatrix}
\]

\[
\begin{bmatrix} 1 \\ -e^{-\lambda_i} p(\rho) \end{bmatrix} = \begin{bmatrix} 1 \\ -e^{-\lambda_i} p(\rho) \end{bmatrix} + (1 - e^{-\lambda_i} p(\rho))(-e^{-\lambda_i} p'\rho) = 0
\]

This implies that the 1st characteristic is genuinely non-linear.

Evaluating \( \lambda_2(\mathbf{u}) \), we have

\[
\begin{bmatrix} \frac{\partial_c \lambda_1} {\partial \mathbf{U}} \\ \frac{\partial_c \lambda_2} {\partial \mathbf{U}} \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}
\]

Hence the 2nd characteristic field is linearly degenerate.

Comparing this model with the Aw Rascle model, we have \( \nabla \lambda_{1}(\mathbf{u}) \cdot R^{(i)}(\mathbf{u}) = 0 \). This shows that the 1st characteristic field is genuinely non-linear.

\( \nabla \lambda_{2}(\mathbf{u}) \cdot R^{(i)}(\mathbf{u}) = 0 \) shows that the 2nd characteristic field is linearly degenerate.

Now we compute the Riemann Invariants for the Aggressive model across each wave as follows;

Across the \( \lambda_1 \) -wave, we have \( \frac{d \rho}{d u} = \frac{du}{-p'(\rho)} \) which on re-arrangement \( p'(\rho) = -e^{-\rho} \rho p' \) and on integration yields \( p(\rho) + e = C \)

Across the \( \lambda_2 \) -wave, we have \( \frac{d \rho}{d u} = \frac{du}{-p'(\rho)} \) which on integration leads to \( u = C \).

The left and right Riemann invariants now become;

\[
\begin{align*}
I_1(p_1, p_2) &= p(\rho) + e \\
I_2(p_1, p_2) &= u
\end{align*}
\]

Riemann Invariants for the Aw-Rascle model are;

Across the \( \lambda_1 \) -wave, we have \( \frac{d \rho}{d u} = \frac{du}{-p'(\rho)} \)\)

Therefore

\[
\begin{align*}
I_1(p_1, p_2) &= p(\rho) + u \\
I_2(p_1, p_2) &= u
\end{align*}
\]

And across the \( \lambda_2 \) -wave we have \( \frac{d \rho}{d u} = \frac{du}{-p'(\rho)} \) which on integration yields

\[
I_1(p_1, p_2) = u
\]

From the above facts, we can conclude that in the Aggressive model the 1-wave will either be a rarefaction or shock waves and the 2-waves will be contact discontinuities [10].

**Method of Solution**

The numerical solution to the conservative system (9-10) will be achieved by use of numerical scheme based on the Godunov method.

To numerically solve the homogenous system

\[
\partial \mathbf{U} + \partial \mathbf{F}(\mathbf{U}) = 0
\]

Suppose that at time \( t = t^n \), the initial data for (15) is given as \( \mathbf{U}(x, t^n) \).

Then the first step of Godunov scheme is the evolution of the solution to a time \( t^{n+1} = t^n + \Delta t \). That is achieved through considering the cell averages.

\[
\mathbf{U}^{n+1}_{i+1/2} = \frac{1}{\Delta x} \int_{x_{i+1/2}}^{x_{i+3/2}} \mathbf{U}(x, t^n) \, dx
\]

Which produces a piecewise constant approximation of the solution \( \mathbf{U}(x, t^n) \) as

\[
\mathbf{U}(x, t^n) = \mathbf{U}^{n+1}_{i+1/2}
\]

For all \( x \in C_k, i = 0,1, \ldots, m \).

The second step is obtaining the solution for the local Riemann problem \( \mathbf{R}(\mathbf{U}^{n+1}_{i+1/2}, \mathbf{U}^{n+1}_{i+3/2}) \) at the cell interface \( x_{i+1/2} \) with data \( \mathbf{U}^{n+1}_{i+1/2} \) and \( \mathbf{U}^{n+1}_{i+3/2} \) respectively on the left side and right side of position \( x_{i+1/2} \).
The solutions to this Riemann problem are self-similar solutions

\[ U(x,t) = \begin{cases} x/t, & t \geq x-x, \ x \in [x,x], t \in [t,t] \end{cases} \]

that is, they are functions of Riemann Problem local co-ordinates \( x/t \) and are constituted by the 1- and 2-wave. Now for a sufficiently small time step \( \Delta t \), such that there are no eave interactions, we obtain the global solution \( U(x,t) \) in the entire spatial domain for \( t \in [0,\Delta t] \) by gluing together the solution of the local Riemann Problem at each interface of the cell as below:

\[ U(x,t) = U_i, \quad \text{for all } (x,t) \in [x_i,x_{i+1}] \times [0,\Delta t] \]

Having obtained the solution \( U(x,t) \) the final step of the Godunov scheme entail the evolution of the solution to a time \( t^{n+1} = t^n + \Delta t \) by defining a new set \( \{ U_i^{n+1} \} \) of average values as follows;

\[ U_i^{n+1} = \frac{1}{\Delta t} \int_{x_i^{-1}}^{x_i} U_i(x,t^n) \, dx \]

Within \( C_i = [x_i^{-1}, x_i] \).

To guarantee that the interface of the \( i \)-waves, \( i=1,2 \) is entirely contained within cell \( C \), we impose the following CFL condition.

\[ \Delta t \leq \frac{CFL \cdot \Delta t}{\max_i \{ \int_{x_i} \mid U_i \mid \, dx \}} \]

\( CFL \) is called the Courant number and is usually set to 1. The CFL condition together with the integral form of the conservation law allows us to alternatively express \( U_i^{n+1} \) in the following form;

\[ U_i^{n+1} = U_i^n + \Delta t \left( F_i^n - F_i^{-1} \right) \]

With the inter cell numerical flux given by

\[ F_i^{-1} = F \left( U_i \left( 0; U_i^n, U_i^{n+1} \right) \right) \]

Results

Phase transitions

Figure 1 shows that as more vehicles perform the mandatory lane changes from lane 3 to lanes 1 and 2 with high flow rate, there is a decrease in velocity within deterministic disturbance.

A deterministic disturbance in free flow at the lane-drop bottleneck happens when vehicles merging from lane 3 onto the other two lanes compel the vehicles on these two lanes to decelerate in the vicinity of the lane-drop zone. This decrease in velocity causes an abrupt increase in density leading to F→S transition. As a result of the upstream propagation of the downstream front of synchronized flow that was initially fixed at the bottleneck there will be an increase in velocity within the deterministic disturbance. This velocity increase has some limit upon which the velocity of vehicles in synchronized flow increases and density decrease drastically, see Figure 2a, thus leading to a return to free flow at the lane-drop bottleneck. Hence S→F phase transition occurs, completing the hysteresis loop where the upper part is the deceleration branch associated with F→S transition while the lower part of the loop is the acceleration branch associated with S→F transition. However, this free flow traffic state exists for only a short period of time and then another F→S transition occurs spontaneously at the bottleneck as it can be seen in the Aw Rascle and Aggressive curves in Figure 2a. For Aw-Rascle model at the same location \( x=0 \), the traffic hysteresis loop is more pronounced unlike in the Aggressive model. This is as a result of more space gaps in the Aw-Rascle model as vehicles from lane 3 merge into the proceeding lanes 1 and 2 far upstream of the bottleneck and allows free movement in the lane drop merging zone. In the Aggressive model vehicles overlap at the bottleneck leaving little space to the approaching vehicles hence creating a dense traffic at the merging zone. This lowers the flow rate as compared to the Aw-Rascle model as Figure 2a shows.

Figure 1: Lane-drop bottleneck.

Figure 2: Flow-density relation for the Aw-Rasclle and Aggressive model in (a) and (b).
Figure 2b also shows the flow-density relations at a location of the considered highway, i.e. upstream of the lane-drop, at x=-20. The growing narrow moving jams at x=0 in Figure 2a, merge later on to transform into wide moving jams at x=-20. In the course of formation of these narrow moving jams, the downstream front starts to propagate upstream from the lane-drop bottleneck. As a result, the density within the deterministic disturbance decreases past the limit necessary for sustenance of synchronized flow leading to a return to free flow. However this free flow state is short-lived and so another traffic breakdown occurs at the bottleneck and in turn another upward moving narrow jam emerges. This process is repeated resulting to a sequence of narrow moving jams of low velocity within them.

Spatiotemporal congested traffic patterns

In the previous section we have discussed the nature of traffic breakdown at a highway bottleneck (lane-drop). In this section we look at the variations of the two models in the features of spatiotemporal congested patterns which results in a highway once traffic breakdown has occurred at a lane-drop bottleneck (Figure 2).

After traffic breakdown is realized at the lane-drop bottleneck, various patterns of synchronized flow may result. These patterns are differentiated by the behavior of their upstream and downstream fronts that separate the synchronized flow within, from the free flow outside of the congested traffic. With increased flow rate at the bottleneck due to merging of vehicles, a wave of dense traffic appears and propagates upstream. Traffic breakdown (F→S transition) at the lane-drop bottleneck, which leads to the emergence of MSPs shown in Figure 3a, is explained by lane changing. Vehicles which move initially in lane 3 change to lane 1 and 2 upstream of the lane-drop bottleneck increasing the density in the two lanes and fluctuations upstream of the bottleneck. For this reason, when an increase in traffic demand upstream of the lane-drop bottleneck occurs, traffic breakdown is observed upstream of the bottleneck. Consequently, the downstream front of the resulting synchronized flow occurs at the bottleneck while the upstream front of synchronized flow propagates upstream. However, a return of phase transition from synchronized flow to free flow (S→F transition) occurs at the bottleneck. The increase in traffic demand upstream of the lane-drop bottleneck, which has caused traffic breakdown and resulting synchronized flow emergence at the bottleneck, last a short time interval only in the Aw-Rascle model, as shown in Figure 3b. Then the resulting S→F transition can lead to the restoration of free flow only in a neighborhood of the bottleneck whereas there is still a region of synchronized flow upstream of the effective location of the lane-drop bottleneck. This causes the upstream propagation of the downstream front of the synchronized flow region, i.e., MSP formation. These MSPs or rather narrow jams grow (i.e. velocity decrease and density increase within them) as they move upstream of the lane-drop bottleneck. The phenomenon of growing narrow jams is that if two or more growing narrow moving jams are relatively close to one another and one of these narrow jams grow into a wide moving jam, then the further growth of the narrow jams that are nearby is suppressed and/or they merge with the wide jam that has been formed. Hence a S→J phase transition is realized. Moreover, if growing narrow moving jams are far enough from the wide moving jam, then the process of either suppressing and/or merging of these narrow jams result into another wide jam formation.

Conclusion

In this study, macroscopic traffic flow models within the framework of Kerner’s 3-phase traffic flow theory with consideration of driver aggressiveness in the context of Kenyan roads have been presented. The 3-phase traffic theory is constituted in the macroscopic equations through the relaxation term. By construction of the solutions to the Riemann problem, set up using the conservative form of the model, the model features have been explored further. The numerical method for solving the macroscopic model in conservative form is discussed and tests are carried out to show the effectiveness of the numerical method used. Using this numerical method we go ahead to simulate traffic flow on a roadway with a lane-drop bottleneck. Through these simulations, we assess the ability of the derived macroscopic traffic flow model i.e. the aggressive model to reproduce the complex spatiotemporal features of traffic flow. Namely the first order F→S transition and the coexistence of free flow (F), synchronized flow (S) and wide moving jams (J) as observed in real traffic flow. The empirical investigation indicate that unless synchronized flow is hindered, moving jams do not emerge in free flow but rather emerge in the synchronized flow phase of traffic. That is their emergence is due to a sequence of two first order phase transitions: F→S and S→J. This is because the onset of congestion in an initial free flowing traffic is associated with F→S transition and later on at some location upstream of the bottleneck, S→J transition occurs depending on the traffic demand. However the Aw-Rascle is shown not to reproduce these features. It has equally been shown that the Aggressive model respects the frontal aspects of traffic and does not produce negative travel.

References

1. Lighthill MJ, Whitham GB (1955) On Kinematic Waves: II. A Theory of Traffic Flow on Long Crowded Roads. Proc Royal Society London 229: 317-345.
2. Payne HJ (1971) Models of freeway traffic and control. Simulation council.
3. Daganzo CF (1995) The nature of freeway gridlock and how to prevent it. Institute of Transportation Studies. Trans Sci 34: 629-646.
4. Rascle AW, Rascle M (2000) Resurrection of ‘second-order’ models of traffic flow. SIAM J Appl Math 60: 916-938.
5. Klar A, Rascle M, Materne T (2002) Derivation of Continuum Traffic Flow Models from Microscopic Follow-the-Leader Models. SIAM Jornal on Applied Mathematics 63: 259-278.
6. Kimathi ME (2012) Mathematical model for the 3-phase Traffic flow Theory.
7. Kerner BS (2002) Effect of driver behaviors on spatiotemporal congested traffic patterns at highway bottlenecks in the frame work of three phase traffic flow.
8. Prigogine I (1961) A Boltzmann-like Approach to the Statistical Theory of Traffic Flow.
9. Helbing D (1998) Gas-kinetic-based traffic model explaining observed hysteretic phase transition. Phys Rev Letters 81: 3042-3045.
10. Muñoz JC, Daganzo CF (2000) Experimental characterization of multi-lane freeway traffic upstream of an off-ramp bottleneck.