IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received January 3, 2020, accepted January 18, 2020, date of publication February 4, 2020, date of current version February 13, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.2971661

Toward the Formalization of Macroscopic Models of Traffic Flow Using Higher-Order-Logic Theorem Proving

ADNAN RASHID[1], MUHAMMAD UMAIR[2], OSMAN HASAN[1], (Senior Member, IEEE), AND MOHAMED H. ZAKI[3], (Member, IEEE)

[1]School of Electrical Engineering and Computer Science (SEECS), National University of Sciences and Technology (NUST), Islamabad 44000, Pakistan
[2]Research Center for Modeling and Simulation (RCMS), National University of Sciences and Technology (NUST), Islamabad 44000, Pakistan
[3]Department of Civil, Environmental, and Construction Engineering, College of Engineering and Computer Science, University of Central Florida, Orlando, FL 32816, USA

Corresponding author: Adnan Rashid (adnan.rashid@seecs.nust.edu.pk)

ABSTRACT Next-generation transportation will be integrated, interconnected and highly autonomous. One key challenge in traffic management is ensuring safety while maintaining the required level of service quality. In such future mobility systems, rigorous formalization and validation will thus become critical to ensure that the transportation network operates as intended, traffic properties are reliably maintained, and services resume in a timely manner after potential disruptions. Formal methods have recently gained considerable attention as a modeling and verification paradigm capable of addressing many challenges associated with next-generation transportation systems. In this regard, we propose to use higher-order-logic theorem proving for formally analyzing transportation systems. As a first step towards this direction, we present a logical framework for the formal analysis of macroscopic models of traffic flow. Leveraging upon the high expressiveness of the underlying logic, we formally model the continuous components of macroscopic models while capturing their real behavior. In particular, we present a formalization of some foundation concepts of macroscopic models, namely density, flow rate, mean speed, relative occupancy, and shockwave using the higher-order-logic theorem prover HOL Light. This choice is primarily motivated by the fact that the macroscopic models deal with the traffic flow dynamics and thus play a vital role in planning strategies in allocating resources for implementing optimized and balanced transportation systems. For illustration, we perform the formal input-output and shockwave analysis of a German freeway. The case study demonstrated the practicability of this formal approach due to the high expressiveness of the underlying logic. The proposed research is first step towards formalizing the foundational mathematical theories and core concepts of traffic flow theory. This accomplishment will open new ways to plan and model various components of the transportation systems such as highway links, ramp metering, merging behavior and eventually address the problem of routing vehicles in a network of automated vehicles.

INDEX TERMS Input-output analysis, macroscopic model, shockwave analysis, theorem proving, transportation modeling, validation methods.

I. INTRODUCTION

Safety and mobility are contending characteristics while developing transportation networks, where a trade-off in design prioritizes one of those properties over the other. As transportation becomes integrated (connected through computer control and communications) and more autonomous, such trade-off is becoming less relevant, and thus the focus is shifting on how to eliminate safety issues while maintaining the desired level of service quality. Rigorous formalization and validation are thus necessary to ensure that the transportation network operates as intended, traffic properties are reliably maintained, and services function properly after potential disruptions. Formal verification of transportation systems is an emerging field [1], [2] that provides an indispensable tool for future development of

The associate editor coordinating the review of this manuscript and approving it for publication was Rashid Mehmood.

transportation networks. Formal verification has emerged as an established tool for safety-critical computer systems, and some demonstrative applications for traffic applications have been presented in [3]–[5]. Based on similar motivation, we provide a novel formal verification framework for the analysis of transportation networks with a focus on the fundamental concepts and properties of traffic flow.

Formal methods [6] are computer-based system analysis techniques that address shortcomings associated with conventional modeling and verification approaches. Theorem proving [7] is a widely used formal method that allows the verification of mathematical relations, including continuous variables, by leveraging upon the expressiveness of higher-order logic, and thus is appropriate for analyzing traffic flow problems. As a first step towards the formal analysis of traffic flow problems, we present a framework for the formal verification of macroscopic models in traffic flow. This choice is primarily motivated by the fact that the macroscopic models deal with traffic flow dynamics and thus play a vital role in planning strategies in allocating resources for implementing optimized and balanced transportation systems [8], [9]. We use the HOL Light theorem prover [10] for conducting the proposed formal analysis due to its extensive support for formally reasoning about multivariate calculus theories.

We present a higher-order-logic formalization of macroscopic model characteristics, namely relative occupancy [11], density [11], flow rate [12], mean speed [12] and shockwave [13]. Based on this formalization, we verify the properties depicting the relationship of relative occupancy and shockwave with the basic parameters of the traffic flow. In the macroscopic model, the continuous traffic flow under equilibrium and non-equilibrium conditions is modeled by the continuous-time partial differential equations, known as *conservation equations* or *continuity equations* [14], [15] and *random number generations* [16]. These equations can be solved to find a relation between density and flow rate of the traffic. These fundamental parameters are further used to calculate the queue size/number of vehicles. An abrupt change in this queue size, due to some obstruction, i.e., crashes, diversion, etc., results into the phenomenon of shockwave [17], which is a boundary between two regions having vehicles with different average values of density, flow rate, and speed. As time progresses, this shockwave moves in the direction of the traffic flow, by creating new shockwaves that replace the earlier shockwaves, depending on the average values of these parameters in the respective regions. The analysis based on these foundations, called shockwave analysis [13], provides the rate of formation or dissipation of the congestion [18] and thus the identification of the congested areas by calculating the queue size/number of vehicles. The work in this paper identifies the mathematical foundations of transportation systems that are required to conduct such analysis within the sound core of a higher-order-logic theorem prover. Moreover, it describes a step-wise procedure to develop a formal model of the given traffic flow problems in higher-order logic and reason about its corresponding properties using an interactive theorem prover.

To illustrate the practical effectiveness of our formalization, we present a formal analysis of a German freeway [19] by verifying its traffic flow properties, and input-output [13], [20] and shockwave analysis related expressions [13]. Our proposed framework, providing the formal verification of the foundations of the transportation systems, can equally be used to plan and model various components of the transportation systems, such as, highway links, diverges, merges and stations, and thus to address the problem of routing vehicles in the network of automated transportation [21].

The rest of the paper is organized as follows: We present an overview of the state-of-the-art of formal verification of safety-critical systems and transportation in Section II. We provide a brief overview of traffic flow theory and the HOL Light theorem prover in Section III. Section IV presents the proposed framework for the formalization of the macroscopic traffic flow models and their properties. In Section V, we provide the formalization of the traffic flow theory foundations, which include the density, flow rate, mean speed, relative occupancy, and shockwave. Moreover, we utilize our foundational formalization to verify some of the properties depicting the relationship of the relative occupancy and shockwave with the macroscopic model parameters, including flow rate and density. To demonstrate the practical utilization and effectiveness of the proposed formalization, we present a formal input-output and shockwave analysis of a German freeway in Section VI. Finally, Section VII concludes the paper by highlighting some future directions.

II. RELATED WORK

Traffic flow theory [11] is developed to describe the interactions between vehicles, the drivers and the transportation infrastructures with operation managed through as highway signals, markings and control devices. All these parameters, contributing towards the dynamics of the transportation systems, are mathematically modeled and analyzed to obtain an optimal and balanced traffic flow with minimal congestion [22]. Traffic flow theory mainly consists of two models, namely *microscopic* and *macroscopic*. Microscopic models [23] capture the dynamic behavior of the underlying transportation system based on the individual behaviors of the vehicles and drivers, and their mutual interaction [24]. On the other hand, the macroscopic model considers the behavior of multiple vehicles simultaneously and it is characterized by its fundamental parameters, such as flow rate, density, mean speed, relative occupancy and shockwave [25]. Thus, in other words, the macroscopic model captures the behavior of all of the vehicles in a certain cross-section as opposed to the microscopic model, which includes the analysis of an individual vehicle.

The continuous traffic flow models of a cross-section highway area are discretized in time and space to facilitate their analyses using computer arithmetic and numerical techniques. This kind of discretization compromises

the completeness of analysis and thus the accuracy of the results [14]. Just like the case of macroscopic models, the paper-and-pencil proof methods and simulation tools, like VISSIM (a microscopic traffic flow simulator) [26], [27] and MITSIMLab (microscopic traffic simulation laboratory) [28], used for analyzing microscopic models, also suffer from the accuracy limitations, described above. Computer algebra systems, such as Mathematica and Maple, have also been used to solve differential equations symbolically and to overcome the inaccuracies introduced by computer arithmetic based computations and numerical methods. However, the algorithms used by these systems are not rigorously verified and thus can produce error-prone results [29]. These flaws in the traditional techniques are tremendously undesirable in case of the high safety-critical domain of transportation, as ignoring some corner cases may lead to dire consequences, such as frequent traffic congestions, road accidents and loss of human lives in worst cases.

Formal Methods based Approach for Verifying Continuous Models: Theorem proving has been widely used for formally analyzing the Cyber-physical Systems (CPS), which are used in various domains including robotics, medicine, avionics and autonomous automobiles. Loos *et al.* [30] formally verified the safety and collision avoidance properties of the distributed car control system using KeYmaera [31], a theorem prover for analyzing hybrid systems. Similarly, Mitsch *et al.* [5] used KeYmaera to formally analyze a distributed intelligent speed adaptation system by incorporating the speed limit control and mainly verified its safety properties. However, both these works are based on the quantified differential dynamic logic, which is a first-order logic. Hasan *et al.* [32] used higher-order logic to formalize the block diagram representations of control systems. Moreover, these representations were also used for the formal analysis of steady-state errors in the feedback [32] and unity-feedback [33] control systems using the HOL Light theorem prover. However, both these works only present the frequency domain analysis of the feedback control systems. Recently, Rashid *et al.* [34], [35] presented the higher-order logic formalization of the linear control systems by modeling their dynamical behaviour using differential equations and formally analyzed the systems based on Laplace transform [36]. The authors also formalized the foundations of the linear control systems, which include the controllers, compensators, phase and gain margins using HOL Light.

Rashid *et al.* formally verified the transfer function of an Unmanned Free Swimming Submersible (UFSS) vehicle [34] and 4-π soft error crosstalk model [37], and frequency response of an Automobile Suspension System (ASS) [38] based on the higher-order logic formalization of Laplace [36], [39] and Fourier [38] transforms, respectively, using HOL Light. Similarly, Sanwal and Hasan [40] formalized the homogenous linear differential equations in the same theorem prover and used it for the formal analysis of the CPS. The authors used their proposed formalization for analyzing a heart pacemaker and a fluid-filled catheter, which are widely used in the domain of bio-medicine. Siddique *et al.* [41] presented the formal verification of the integrated photonic systems using HOL Light. In particular, the authors provided the formal specification of a photonic mirroring resonator and verified its various properties, such as spectral power and rejection ratio. Farooq *et al.* [42] developed the support for the formal kinematic analysis of two-link planar manipulator in HOL Light theorem prover and used it for analyzing the two-dimensional biped walking robot. Later, Affeldt *et al.* [43] extended this framework by formalizing the foundations for analyzing three-dimensional robot manipulator using the Coq theorem prover [44] and used it for the verification of the SCARA robot manipulator. Recently, Rashid and Hasan [45] provided the formal modeling and analysis of the 2-DOF robotic cell injection systems using HOL Light. However, to the best of our knowledge, higher-order-logic based theorem proving has never been used in the context of macroscopic models, which is the scope of the current paper.

III. PRELIMINARIES

In this section, we provide a brief introduction to the macroscopic model of traffic flow theory and the HOL Light theorem prover to facilitate the understanding of the paper.

A. TRAFFIC FLOW THEORY - MACROSCOPIC MODEL

The macroscopic model of traffic flow theory considers all of the vehicles in a cross-section of a road simultaneously [11], [12]. In order to understand the widely used notions of relative occupancy, flow rate, density and mean speed, consider Figure 1, which depicts two rectangular regions, namely S_1 and S_2, representing areas in the t-x space. The region S_1 corresponds to a measurement over a road section ΔX during an infinitely small time interval dT, whereas the region S_2 corresponds to an infinitely small road length dX at a fixed location over a time period ΔT. It is assumed that n and m vehicles move through regions S_1 and S_2, respectively.

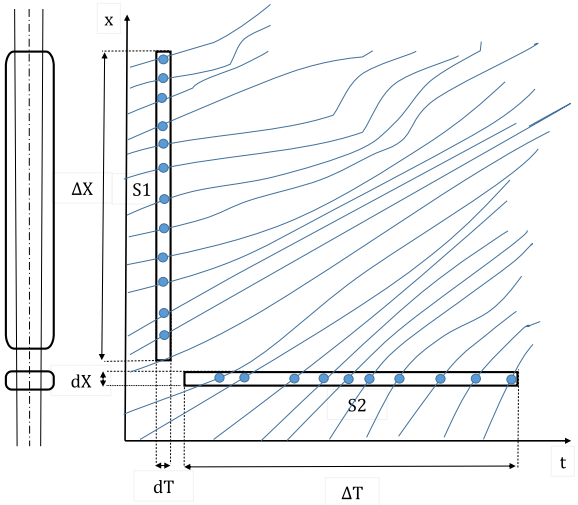

FIGURE 1. Time-space Diagram [11].

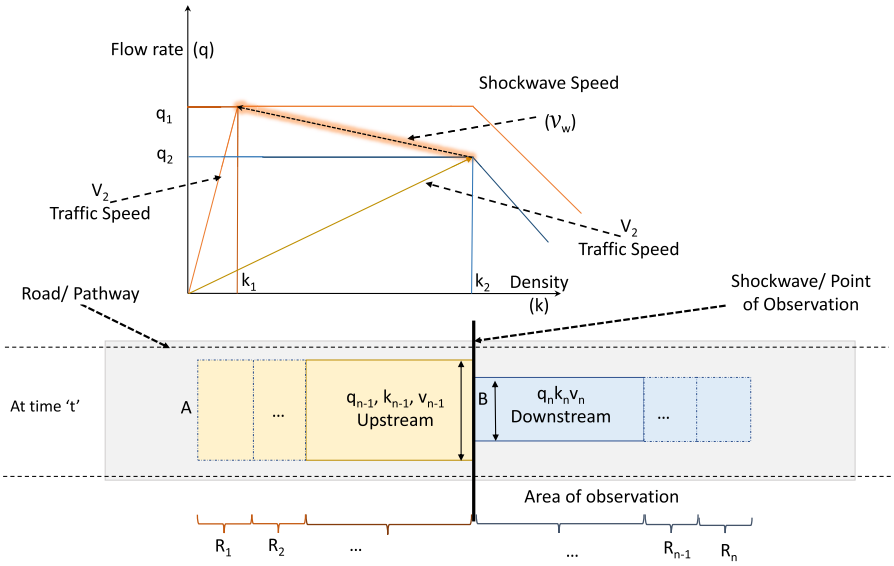

FIGURE 2. Flow-density Diagram.

Moreover, the area of the region S_1 is ΔXdT, whereas the area of the region S_2 is ΔTdX. Now, the relative occupancy is defined as the measurement of the fraction of time, for which the measurement location is occupied by the vehicles. In the region S_2, it is given by the following formula [12]:

$$b_{S_2} = \frac{1}{\Delta T}\sum_{i=1}^{n} O_i = \frac{1}{\Delta T}\sum_{i=1}^{n} \frac{L_i}{V_i} \tag{1}$$

where ΔT is the length of the region S_2. O_i is the occupancy of the i^{th} vehicle and is equal to the ratio of the length L_i and speed V_i of the vehicle, whereas n is the total number of vehicles in S_1.

Similarly, the occupancy of the vehicles in the region S_1 can be mathematically expressed as [12]:

$$b_{S_1} = \frac{1}{\Delta X}\sum_{i=1}^{m} L_i \tag{2}$$

where ΔX represents the length of the region S_1 and m is the total number of vehicles in the region S_1.

The flow rate of the traffic can be defined as the number of vehicles in a certain cross-section per unit time or, alternatively, as the ratio of the total distance covered by all vehicles in a region and the area of the region. In the region S_1, it is given by the following formula [12]:

$$q_{S_1} = \frac{1}{\Delta X}\sum_{i=1}^{m} V_i \tag{3}$$

where V_i is the velocity of the i^{th} vehicle in the region S_1. Similarly, the flow rate in the region S_2 is given by the following formula [12]:

$$q_{S_2} = \frac{n.dX}{\Delta T.dX} = \frac{n}{\Delta T} \tag{4}$$

where dX is the width of the region S_2.

The density of the traffic represents the number of vehicles in a certain cross-section or, alternatively, as the ratio of the total time spent by all vehicles in a region and the area of the region. In the region S_1, it is represented by the following mathematical expression [12]:

$$k_{S_1} = \frac{m.dT}{\Delta X.dT} = \frac{m}{\Delta X} \tag{5}$$

where dT is the width of the region S_1. The following formula represents the density in the region S_2 [12]:

$$k_{S_2} = \frac{1}{\Delta T}\sum_{i=1}^{n} \frac{1}{V_i} \tag{6}$$

where V_i is the velocity of the i^{th} vehicle.

The mean speed can now be defined as the ratio of the flow rate (q) and the density (k) of the traffic flow in each of the S_1 and S_2 regions. It is the time mean speed when calculated for the region S_2 and space mean speed when calculated for the region S_1.

$$u = \frac{q}{k} \tag{7}$$

In general, the time mean speed is the arithmetic mean of speeds observed at some point in a specific time interval and it is generally easier to measure. Whereas, the space mean speed used in the traffic models is calculated as the arithmetic mean of speeds in different time intervals at a region S_1 and is generally harder to measure [46], [47].

Now, to understand the phenomenon of shockwave, consider Figure 2, which mainly depicts the flow-density diagram [48]. Consider an area of observation, in which the traffic flows with some density, flow rate and speed where a sudden obstruction of the traffic flow, due to some accident or closed road or some diversion, splits this area in two regions, namely R_1 and R_2. This obstruction results

into a phenomenon of shockwave, which basically defines a boundary between the Regions R_1 and R_2 and each of these regions contain vehicles having different values of average density, flow rate and speed, i.e., q_1, k_1 and v_1 in R_1 and q_2, k_2 and v_2 in R_2, respectively at some time instant as depicted in Figure 2. With the passage of time, this shockwave moves along the flow of the traffic with some speed v_w by creating new shockwaves and thus regions and canceling the earlier shockwave and the corresponding old regions. The shockwave speed v_w thus plays a vital role in the identification of the congested area by capturing the rate of formation and dissolution of the congestions and finding out the number of vehicles in the respective regions. The shockwave speed, for two adjacent regions R_{n-1} and R_n is:

$$v_{w_n} = \frac{q_n - q_{n-1}}{k_n - k_{n-1}} \tag{8}$$

where q_n and q_{n-1} are the flow rates in Regions R_n and R_{n-1}, respectively. Similarly, k_n and k_{n-1} are the densities in Regions R_n and R_{n-1}, respectively. These densities and flow rates are related by the following mathematical expressions:

$$q_{n-1} = k_{n-1} v_{n-1} \tag{9}$$

$$q_n = k_n v_n \tag{10}$$

where v_{n-1} and v_n represent the average space mean speeds of the vehicles in Regions R_{n-1} and R_n, respectively. The relative speed of a vehicle to an observer is defined as the space mean speed relative to the shockwave speed. In Region R_1, it is mathematically represented as:

$$v_{R_1} = v_1 - v_w \tag{11}$$

Similarly, the relative speed in Region R_2 is given by:

$$v_{R_2} = v_2 - v_w \tag{12}$$

Figure 3 represents the time-space diagram for the macroscopic model depicting the shockwave speeds in three different regions. The queue size based on a shockwave analysis considering Regions R_1 and R_2 is mathematically expressed as follows [12], [13], [20]:

$$\text{Queue Size} = -v_{w_1} \Delta t \Delta k = -\frac{q_2 - q_1}{k_2 - k_1} \Delta t \Delta k \tag{13}$$

where Δk and Δt are the density range and time length, respectively for the shockwave speed v_{w_1}. Similarly q_1, k_1, q_2 and k_2 are the flow rates and densities in Regions R_1 and R_2 of traffic flow, respectively. Whereas $v_{w_1} = (q_2 - q_1)/(k_2 - k_1)$ in Figure 3. It is important to note that the outgoing flow rate is taken as positive and the ingoing flow rate as negative unlike the input-output model, as the queue size for input-output analysis is mathematically represented as [13], [20]:

$$\text{Queue Size} = (q_1 - q_2)\Delta t \tag{14}$$

The behavior of multiple shockwaves for three regions is depicted in the time-space domain [13] in Figure 3. Where

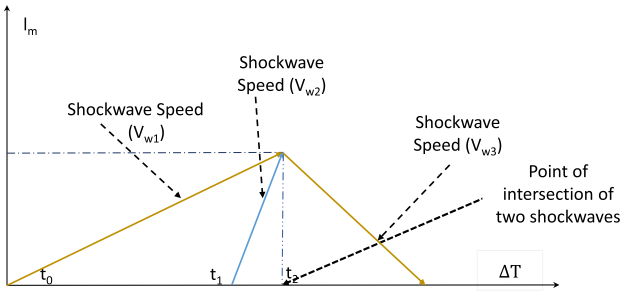

FIGURE 3. Multiple Shockwaves in Time-space Diagram [13].

two shockwaves v_{w1} and v_{w2} are overlapping in time-space graph from t_1 to t_2 [13] and intersect at time point t_2 where their effect disappears and consequently a new shockwave v_{w3} emerges at that point. We use the following generic mathematical expression to analyze the queue size N_{sw} for n regions that can be mathematically expressed as:

$$N_{sw} = \sum_{j=1}^{m} -(v_{w_j} \Delta t_j - \sum_{i=0}^{n} v_{w_i} \Delta t_i) \Delta k_j \tag{15}$$

where v_{w_j} has the longest duration with respect to time as compared to the rest of short ranged shockwaves v_{w_i}, which simultaneously exist in time domain with v_{w_j}. The negative sign is used in the above equation because in shockwave analysis, the outgoing flow rate is taken as positive and ingoing flow rate as a negative real number.

The Input-output analysis models the queue size for n number of regions as follows:

$$N_{io} = \sum_{i=1}^{n} (q_i - q_{i+1}) \Delta t_i \tag{16}$$

We consider the boundary space between two regions, i.e., between Regions R_1 and R_2, as a separate Region R_w and the average speed of the vehicles in this shockwave region is considered as v_w. This way, the average speed shift between the two regions is $v_1 - v_w$. The density range of this shockwave region is considered as k_1 to accommodate all the incoming vehicles from Region R_1. Similarly consider the time required for the whole queue size of R_1 to exit from the ending point of Region R_1 and to enter the Region R_w as Δt. Hence, Δt and Δk should be the same based on the universal law of conservation. Thus, the number of vehicles crossing the boundary of Region R_1 to R_2 can be expressed as:

$$N_1 = v_{R1} k_1 \Delta t = (v_1 - v_w) k_1 \Delta t = (\frac{q_1}{k_1} - v_w) k_1 \Delta t \tag{17}$$

Similarly, the incoming number of vehicles from the rear boundary in Region R_2 is given as:

$$N_2 = v_{R2} k_2 \Delta t = (v_2 - v_w) k_2 \Delta t = (\frac{q_2}{k_2} - v_w) k_2 \Delta t \tag{18}$$

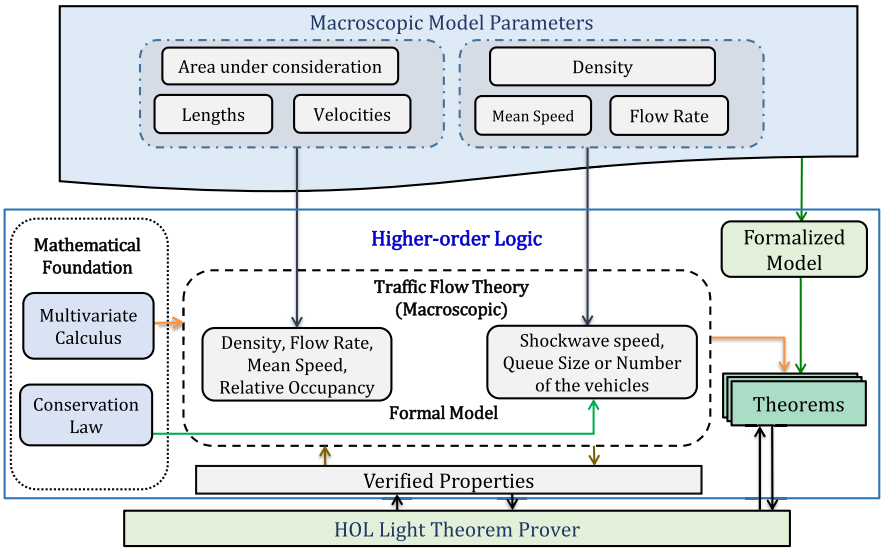

FIGURE 4. Proposed Framework.

B. HOL Light THEOREM PROVER

HOL Light [49] is a higher-order-logic proof assistant that ensures secure theorem proving using the Objective CAML (OCaml) language [10], which is a variant of the strongly-typed functional programming language ML. HOL Light users can interactively verify theorems by applying the available proof tactics and proof procedures. A HOL Light theory consists of types, constants, definitions and theorems. HOL Light theories are built in a hierarchical fashion and new theories can inherit the definitions and theorems of their parent theories. HOL Light consists of a rich set of formalized theories, including sets, natural numbers and the multi variable calculus theories. i.e., real analysis and vector calculus theories. The availability of these theories was the main motivation for choosing HOL Light for the proposed formalization as these foundations are required for reasoning about continuous (real-valued) variables and partial differential equations. Table 1 provides the mathematical interpretations of some of the HOL Light symbols and functions used in this paper.

TABLE 1. HOL Light Symbols and Functions.

HOL Symbol	Standard Symbol	Meaning
/\	and	Logical *and*
\/	or	Logical *or*
~	not	Logical *negation*
T	true	Logical true value
F	false	Logical false value
==>	$\longrightarrow$	Implication
<=>	$=$	Equality in Boolean domain
!x. t	$\forall x.t$	for all $x : t$
λx.t	$\lambda x.t$	Function that maps x to $t(x)$
num	$\{0, 1, 2, \ldots\}$	Natural number data type
real	All Real numbers	Real data type
SUC n	$(n + 1)$	Successor of natural number
&a	$\mathbb{N} \rightarrow \mathbb{R}$	Typecasting from Natural Numbers to Reals
HD L	$head$	Head element of list L
TL L	$tail$	Tail of list L
EL n L	$element$	n^{th} element of list L
CONS	::	Adds a new element to the top of the list
LENGTH L	$length$	Length of list L
(a,b)	a x b	A pair of two elements
FST	fst (a, b) = a	First component of a pair
SND	snd (a, b) = b	Second component of a pair

IV. PROPOSED FRAMEWORK

The proposed framework, shown in Figure 4, outlines the proposed approach for the formal analysis of macroscopic traffic flow models based on higher-order-logic theorem proving, which includes the formalization of their fundamentals, i.e., density, flow rate, means speed, relative occupancy and shockwave. The inputs to the framework are the macroscopic model parameters. For example, to find out the density, flow rate, mean speed and relative occupancy, these input parameters are the lengths and velocities of vehicles and the starting and ending points of the regions S_1 and S_2 (Figure 1). Similarly, to find out the shockwave speed and queue size (number of vehicles), flow rate, density and mean speed are used as the input parameters for our proposed framework.

The first step in conducting formal analysis is the construction of the higher-order-logic based formal model of the given system based on the given macroscopic model parameters. The higher-order-logic formalization, required for developing this model, can be broadly decomposed into two parts, which are depicted by the dotted rectangles in Figure 4. The first part is the core mathematical foundations of macroscopic model of traffic flow theory and the second part is composed of the definitions and theorems of traffic flow theory required for the analysis of macroscopic models. These mathematical foundations include Multivariate calculus theory and the conservation law. The traffic flow theory part builds upon the mathematical foundations and the formalization of the basic concepts of lengths and widths of the rectangular regions, density and flow rates, and the dependencies between them are shown in the Figures 2 and 4. We propose to formalize

the commonly used macroscopic characteristics i.e., density, flow rate, mean speed, relative occupancy, number of vehicles and shockwave speed for capturing the dynamics of the given transportation system. Furthermore, by using these definitions, we propose to verify the corresponding theorems that capture the characteristics of the macroscopic traffic flow model, e.g., the properties that depict the relationship of the relative occupancy and shockwave speed with the density, flow rate and number of vehicles of a transportation system.

Once the formal model, corresponding to the given macroscopic model parameters, is constructed then the next step is to verify its properties as higher-order-logic theorems. The proof goals can be expressed in higher-order logic and can be discharged by interacting with the proof assistant of the HOL Light theorem prover. The reasoning process, involved in this interactive verification, would be mainly based on the properties of the above-mentioned formalized notions of macroscopic model of traffic flow theory.

V. FORMALIZATION OF MACROSCOPIC MODELS

In this section, to the best of our knowledge, we present the first higher-order-logic formalization of the fundamentals of the macroscopic models of the traffic flow theory.

A. FORMALIZATION OF RELATIVE OCCUPANCY

We first present a higher-order-logic formalization of relative occupancy, which is one of the foremost elements of the macroscopic model of the traffic flow theory, as depicted in Figure 4. This formalization builds upon the formalizations of multi variable calculus and the notions of length and widths of the rectangular regions and velocities, density, flow rate and mean speed from the traffic flow theory part.

A macroscopic model of the traffic flow theory consists of two rectangular regions, namely region S_1 and region S_2 (Figure 1), and the lengths and speeds of the vehicles. We model the length and width of both of the regions in terms of their starting and ending points as a pair of real numbers $(\mathbb{R}, \mathbb{R})$, where the first element represents the starting point and the second element represents the ending point of the length and width. For example, taking a measurement of the traffic flow between 2km and 5km on a highway in the region S_1, the starting point of the length of this region is 2 and the ending point is 5 and it can be represented as a pair (2, 5). We formally describe the macroscopic model datatype for relative occupancy as:

Definition 1: Macroscopic Model Datatype for Relative Occupancy

```
new_type_abbrev "ts_macro_traffic_flow",
  :(((real × real) × (real × real) × (real × real) ×
     (real × real)) × ((real × real)list × (real × real)list))
```

In the above definition, a time-space model is a pair with the first element as a 4-tuple. The first element of the 4-tuple is a pair (real × real), which represents the starting and ending points of the length of the region S_1 (vertical rectangle). The second element of the 4-tuple is also a pair (real × real), representing the starting and ending points of the width of the region S_1. Similarly, the third and fourth elements of the 4-tuple are also pairs (real × real) representing the starting and ending points of the length and width of the region S_2 (horizontal rectangle). The second element of the macroscopic model pair is itself a pair. The first element (real × real)list of this pair is a list of pairs in which the first element represents the length of a vehicle and the second element is its corresponding speed in the region S_1. Likewise, the second element (real × real)list of this pair is also a list of pairs, where each pair represents the lengths and speed of the vehicles in the region S_2.

In order to obtain the characteristics of the macroscopic model of the traffic flow, i.e., relative occupancy, density, flow rate and mean speed, we need to find out the lengths, widths and the occupancy of the S_1 and S_2 regions. For this purpose, we use the following function that allows us to find the length and width of a rectangle:

Definition 2: Length/Width of the Rectangles

```
⊢def ∀x_t_rec. differ x_t_rec = SND x_t_rec - FST x_t_rec
```

The function differ accepts the starting and ending point of the length/width of the rectangle in the form of a pair and returns its length/width by taking the arithmetic difference between the elements of the given pair.

The list containing the occupancies of all of the vehicles can be obtained as follows:

Definition 3: List Containing the Occupancies of a Collection of Vehicles

```
⊢def occ_list [ ] = [ ] ∧
  occ_list (CONS h t) = CONS (FST h / SND h) (occ_list t)
```

The function occ_list:$((real \times real)\ list \rightarrow real\ list)$ accepts a list of pairs, where each pair represents the length and speed of a vehicle, and returns the list of their corresponding occupancies.

In order to obtain the relative occupancy in the region S_2, we need the sum of the occupancies of all of the vehicles:

Definition 4: Sum of the Occupancies of all the Vehicles

```
⊢def ∀L. occ_sum L =
  sum (1..LENGTH (occ_list L)) (λi. EL (i - 1) (occ_list L))
```

The function occ_sum:$((real \times real)\ list \rightarrow real)$ accepts a list of pairs, where each pair represents the length and speed of a vehicle, and returns a real number that is the sum of the occupancies of all of the vehicles. This definition uses the HOL Light function sum in order to take sum of a function over a range of values.

Now, we can obtain relative occupancy in the region S_2 (Equation (1)) by using Definitions 2 and 4 as follows:

Definition 5: Relative Occupancy in Region S_2

```
⊢def ∀xv tv xh lng_spd_v lng_spd_h th.
  rel_occ_s2 ((xv,tv,xh,th),lng_spd_v,lng_spd_h) =
        occ_sum lng_spd_h / differ th
```

The function rel_occ_s2:$(ts\_macro\_traffic\_flow \rightarrow real)$ accepts an element of data type *ts_macro_traffic_flow* and returns the corresponding relative occupancy of the vehicles in the region S_2.

To obtain the relative occupancy in the region S_1, we need to find the summation of the lengths of all of the vehicles.

Definition 6: Summation of the lengths of all of the vehicles in Region S_1

```
⊢def ∀L. sum_l_list L =
       sum (1..LENGTH (l_list L)) (λi. EL (i - 1) (l_list L))
```

The function **sum_l_list**:((*real* × *real*) *list* → *real*) accepts a list of pairs, where each pair represents the length and speed of a vehicle, and returns the sum of the lengths of all of the vehicles in the given list. The function **l_list** used in the above definition, takes the list of pairs containing lengths and speeds of the vehicles and returns a list containing their lengths only.

Now, the relative occupancy in the region S_1 (Equation (2)) is formalized as follows:

Definition 7: Relative Occupancy in Region S_1

```
⊢def ∀tv xh th lng_spd_h lng_spd_v xv.
       rel_occ_s1 ((xv,tv,xh,th),lng_spd_v,lng_spd_h) =
                  sum_l_list lng_spd_v / differ xv
```

The function **rel_occ_s1** accepts an element of data type (*ts_macro_traffic_flow*) and returns the relative occupancy of the vehicles in the region S_1.

Our next step is to formalize the notion of traffic density in the S_1 (Equation (5)) and S_2 regions (Equation (6)):

Definition 8: Density in Region S_1

```
⊢def ∀xh th lng_spd_h lng_spd_v tv xv.
       density_s1 ((xv,tv,xh,th),lng_spd_v,lng_spd_h) =
     (&(no_veh lng_spd_v) ∗ differ tv) / (differ xv ∗ differ tv)
```

The function **density_s1**:((*ts_macro_traffic_flow*) → *real*) accepts an element of data type (*ts_macro_traffic_flow*) and returns the corresponding traffic density in the S_1 region. The density in the region S_2 (Equation (6)) can be formalized as follows:

Definition 9: Density in Region S_2

```
⊢def ∀xv tv lng_spd_v lng_spd_h th xh.
       density_s2 ((xv,tv,xh,th),lng_spd_v,lng_spd_h) =
                  sum_v_inv lng_spd_h / (differ th)
```

The function **density_s2** takes an element of data type (*ts_macro_traffic_flow*) and returns the density in the region S_2. The function **sum_v_inv** in the above definition accepts a list of pairs, where each pair represents the length and speed of a vehicle, and returns the summation of the inverse of their speeds.

We formally define the traffic flow rate in the region S_1 (Equation (3)) as follows:

Definition 10: Flow Rate in Region S_1

```
⊢def ∀xh th lng_spd_h lng_spd_v tv xv.
       flow_rate_s1 ((xv,tv,xh,th),lng_spd_v,lng_spd_h) =
                    (sum_v_list lng_spd_v / differ xv)
```

The function **flow_rate_s1** takes a time-space macroscopic model and returns the flow rate in the region S_1. In this definition, the function **sum_v_list** accepts a list of pairs, where each pair represents the length and speed of a vehicle, in the region S_1 and returns the sum of their speeds.

Similarly, the flow rate in the region S_2 (Equation (4)) can be defined as follows:

Definition 11: Flow Rate in Region S_2

```
⊢def ∀xv tv lng_spd_v lng_spd_h th xh.
       flow_rate_s2 ((xv,tv,xh,th),lng_spd_v,lng_spd_h)
     = (&(no_veh lng_spd_h) ∗ differ xh)/(differ th ∗ differ xh)
```

The function **flow_rate_s2**:((*ts_macro_traffic_flow*) → *real*) accepts an element of data type (*ts_macro_traffic _flow*) and returns the flow rate in the region S_2. In this function, the function **no_veh** takes a list of pairs containing the lengths and speeds of the vehicles in region S_2 and returns the number of vehicles in the region. This function uses the **HOL Light** function **LENGTH**, which accepts a list of any data type and returns its length as a positive integer.

We next formalize the mean speed in both regions. The mean speed in the region S_1 (Equation (7)) is defined as:

Definition 12: Mean Speed in Region S_1

```
⊢def ∀xv tv xh th lng_spd_v lng_spd_h.
mean_speed_s1 ((xv,tv,xh,th),lng_spd_v,lng_spd_h) =
     flow_rate_s1 ((xv,tv,xh,th),lng_spd_v,lng_spd_h)
             / density_s1 ((xv,tv,xh,th),lng_spd_v,lng_spd_h)
```

The function **mean_speed_s1** takes an element of data type *ts_macro_traffic_flow* and returns the mean speed in the region S_1. The mean speed in the region S_2 is given as follows:

Definition 13: Mean Speed in Region S_2

```
⊢def ∀xv tv xh th lng_spd_v lng_spd_h.
mean_speed_s2 ((xv,tv,xh,th),lng_spd_v,lng_spd_h) =
     flow_rate_s2 ((xv,tv,xh,th),lng_spd_v,lng_spd_h)
             / density_s2 ((xv,tv,xh,th),lng_spd_v,lng_spd_h)
```

In order to ensure the correctness and soundness of our definitions, we use them to verify a couple of properties representing some important characteristics of the macroscopic model. The first property deals with the case when length of all of the vehicles is the same then the relative occupancy in the region S_1 is equal to the length times the density of vehicles in the region. The second property captures the same characteristic under the same assumption for the vehicles in the region S_2.

We verify the first property as the following theorem:

Theorem 1: Relationship of Relative Occupancy and Density in Region S_1

```
   ⊢thm ∀xv tv xh th lng_spd_v lng_spd_h c.
[A1:] ~(NULL lng_spd_v) ∧
[A2:] &0 < (SND xv - FST xv) ∧
[A3:] &0 < (SND tv - FST tv) ∧
[A4:] (∀ i. EL i (l_list lng_spd_v) = c)
  ⟹ rel_occ_s1 ((xv,tv,xh,th),lng_spd_v,lng_spd_h) =
       c ∗ density_s1 ((xv,tv,xh,th),lng_spd_v,lng_spd_h)
```

The variable **lng_spd_v** represents the list of pairs having lengths and velocities of the vehicles, whereas, **xv** and **tv** represent the starting and ending points of the length (ΔX) and width (dT) of the region S_1, respectively. The assumption **A1** ensures that the list **lng_spd_v** is not empty. The assumptions **A2-A3** guarantee that each of the length and width of the region are always positive, as these are the distance and time. The assumption **A4** represents the condition that the lengths of all the vehicles is same. Finally, the conclusion of the theorem describes the relationship of the relative occupancy to the density of the vehicle.

The reasoning process of Theorem 1 is primarily based on the definitions of the functions **rel_occ_s1** and **density_s1**, and a lemma that says if all the elements of a list are same, i.e., equal to some constant c, then the summation of this list is equal to c times the length of the list.

Lemma 1: Summation of a List having Same Element c is c Times Length of List

```
⊢thm ∀c L. [A1:] ∼(NULL L) ∧
           [A2:] (∀i. EL i L = c)
       ⟹ sum (1..LENGTH L) (λi. EL (i - 1) L) =
                &(LENGTH L) ∗ c
```

Similarly, the second property depicting the relationship of the relative occupancy with the density in the region S_2 is given by the following theorem:

Theorem 2: Relationship of Relative Occupancy and Density in the Region S_2

```
⊢thm ∀xv tv xh th lng_spd_v lng_spd_h c.
[A1:] ∼(NULL lng_spd_v) ∧
[A2:] &0 < (SND th - FST th) ∧
[A3:] &0 < (SND xh - FST xh) ∧
[A4:] (∀i. FST (EL i lng_spd_h) = c)
   ⟹ rel_occ_s2 ((xv,tv,xh,th),lng_spd_v,lng_spd_h) =
      c ∗ density_s2((xv,tv,xh,th), lng_spd_v, lng_spd_h)
```

The variable lng_spd_h represents the list of pairs having lengths and velocities of the vehicles, whereas, xh and th represent the starting and ending points of the length (ΔT) and width (dX) of the region S_2, respectively. All the assumptions of this theorem are same as that of Theorem 1, but in the context of the region S_2. The conclusion of the Theorem 2 describes the relationship of relative occupancy with the density of the vehicles. The verification process of this theorem is similar to the one of Theorem 1 and more details can be found in the source code of the formalization [50].

B. FORMALIZATION OF SHOCKWAVE

For the shockwave analysis, we have modeled a single region as a pair $((q, k), v)$, where the first element itself represents a pair i.e., q and k represent the flow rate and density and the second element depicts the shockwave speed, respectively. All these parameters are real-valued, i.e., $q, k, v \in \mathbb{R}$. For example, traffic flow of 2000 veh/hr, density of 80 veh/km and shockwave speed of 1 km/hr on a highway is represented as ((2000, 80), 1). Consequently all the regions generating the shockwaves (Figure 2) can be individually modeled using a list of regions. Then the multiple shockwaves, shown in Figure 3, modeled from the pairs of regions, are added for all the regions shown in Figure 2. To simplify the reasoning process about shockwave phenomenon, we encode the above information using the three type abbreviations in HOL Light, namely, ptrgn, sw and sw_d as follows:

Definition 14: Macroscopic Model Datatype for Shockwave Analysis

```
new_type_abbrev "ptrgn",:(real × real) × real
new_type_abbrev "sw",: ((ptrgn)list × (real × real)) ×
(num × num)
new_type_abbrev "sw_d",:((sw)list × (real × real))
```

where the first element of sw is itself a pair, in which the first element represents the list of the regions. The second pair of sw is a pair (real × real) representing *time* points corresponding to start and end of a shockwave. Similarly, the second element of sw is also a pair (num × num), representing the indices of the two adjacent regions as shown in Figure 2.

Similarly a shockwave regional model sw_d is a pair, which models the dynamic behavior of all the shockwaves in an entire area, as shown in Figure 3. The first element of sw_d is a list of sw elements and the second element having data type real × real, shows the initial and final density points of the area. Consequently the following data type would be able to model the dynamic behavior for the area of observation on a road or highway as a composition of shockwave's elements, i.e., flow, density, region's index and density range with multiple regions depicted in Figure 2.

Definition 15: Point List in the Regions

```
⊢def ∀p. pt_list p = FST(FST p)
```

The function pt_list:$(sw \rightarrow (ptrgn)list)$ accepts a variable of data type sw and returns the first element of the first pair of variable, i.e., (ptrgn)list. The returned element is itself a list of points and each point in the list represents the flow rate, density and shockwave speed in one region.

Definition 16: Time Range

```
⊢def ∀p. time p = SND(SND(FST p)) - FST(SND(FST p))
```

The function time:$(sw \rightarrow real)$ accepts a variable of data type sw and returns the difference of the second element with the first part of the pair (data type: sw), which represents the considered time length for a shockwave in the area of consideration.

Definition 17: First Index of Point from Points/Regions List

```
⊢def ∀p. ind_m p = FST(SND p)
```

The function ind_m:$(sw \rightarrow num)$ accepts a variable of data type sw and returns the index of the first region for the shockwave considered in the area of consideration.

Definition 18: Second Index of Point from Points/Regions List

```
⊢def ∀p. ind_n p = SND(SND p)
```

The function ind_n:$(sw \rightarrow num)$ accepts a variable of data type sw and returns the index of the second region for the shockwave considered in the area of consideration.

Definition 19: Average Flow Rate of One Point/Region

```
⊢def ∀t. flow_rate t = FST(FST t)
```

The function flow_rate:$(ptrgn \rightarrow real)$ accepts an element of data type ptrgn and returns the first element of the first pair of a variable, i.e., the average flow rate of the vehicles.

Definition 20: Average Density of One Point/Region

```
⊢def ∀t. density t = SND(FST t)
```

The function density:$(ptrgn \rightarrow real)$ accepts an element of data type ptrgn and returns the second element of the first pair of a variable, i.e., the average density of the vehicles.

Definition 21: Shockwave Speed Associated with one Point/Region

```
⊢def ∀t. shock_wv t = SND t
```

The function shock_wv:$(ptrgn \rightarrow real)$ accepts a variable of data type ptrgn and returns the second element of the pair, i.e., the shockwave speed corresponding to underlying region.

Definition 22: Number of Vehicles Crossing Line (n) from Region R_{n-1} (During Some Time Period Δt)

```
⊢def ∀p n. n_crossing p n =
((flow_rate (EL n (pt_list p)) / density (EL n (pt_list p)) -
  shock_wv (EL n (pt_list p))) ∗
        density (EL n (pt_list p))) ∗ time p
```

The function n_crossing:$((sw \rightarrow num) \rightarrow real)$ accepts two variables of data types sw and num, respectively, and

returns the number of vehicles crossed to or from the boundary of the region with index n (according to Equations (17) and (18)).

Definition 23: List Containing the Shockwaves of All the Areas

```
⊢def sw_list [ ] = [ ] ∧ sw_list (CONS h t) =
  CONS (shock_wv (EL (ind_n h) (pt_list h))
                 * time h) (sw_list t)
```

The function sw_list:$((sw)list \rightarrow (real)list)$ accepts an element of data type (sw)list and multiplies each element of the list with time (the required time length of the shockwave to be considered for the analysis) and then returns the new list of real values. This represents the space regions in the time-space diagram or a shockwave over a time length according to the shockwave analysis [13].

Definition 24: Summation of the Shockwaves

```
⊢def ∀L. sum_sw L =
       sum (1..LENGTH L) (λi. EL (i - 1) (sw_list L))
```

The function sum_sw:$((real)list \rightarrow real)$ accepts a list of real numbers and returns a real number, i.e., the shockwave sums in the required time lengths according to shockwave analysis [13].

Next, we use the above-mentioned formalization for the verification of shockwave equation [20], [51], [52], which elaborates the average speed shift or speed change between two adjacent regions in terms of average flow rates and densities in those regions.

Theorem 3: Shockwave Speed Verification in two Regions R_m and R_n

```
⊢thm ∀p.
[A1:] n_crossing p (ind_n p) =
          n_crossing p (ind_m p) ∧
[A2:] density (EL (ind_n p) (pt_list p)) ≠
          density (EL (ind_m p) (pt_list p)) ∧
[A3:] density (EL (ind_n p) (pt_list p)) ≠ &0 ∧
[A4:] density (EL (ind_m p) (pt_list p)) ≠ &0 ∧
[A5:] shock_wv (EL (ind_n p) (pt_list p)) =
          shock_wv (EL (ind_m p) (pt_list p)) ∧
[A6:] &0 < time p
   ⟹ shock_wv (EL (ind_n p) (pt_list p)) =
       (flow_rate (EL (ind_n p) (pt_list p)) -
        flow_rate (EL (ind_m p) (pt_list p))) /
            (density (EL (ind_n p) (pt_list p)) -
             density (EL (ind_m p) (pt_list p)))
```

The variable p is a pair having data type sw, representing a list of points in regions, i.e., indices for the list of points in regions and the time interval, to compute a single shockwave, respectively. The assumption A1 ensures that the number of vehicles crossing (Definition 22) from one region (Equation (17)) would be the same as to the other region's incoming number of vehicles crossing (Equation (18)) from the rear adjacent boundary, according to the universal law of conservation. The assumption A2 describes that the densities in both regions are not the same and thus making them separate regions. The assumptions A3-A4 model the conditions that the densities of vehicles in both of the considered regions are non zero. The assumption A5 models the expression for the shockwave speed in both of the considered regions. The assumption A6 ensures that the considered time interval is not negative. Finally, the conclusion of this theorem describes the relationship of the shockwave with flow rates and densities in any two regions.

The reasoning process of Theorem 3 is primarily based on the definitions of the functions n_crossing, flow_rate, density, pt_list, ind_m and ind_n and a lemma that ensures that the inverse of all non-zero real numbers would also be a non-zero quantity and another lemma that describes the cross multiplication of four real numbers [50].

We formalize the queue size/number of vehicles via an input-output analysis (Equation (14)) as follows:

Definition 25: Queue Size (Number of Vehicles) via Input-output Analysis

```
⊢def ∀p. n_io p =
       (flow_rate (EL ind_m (pt_list p)) -
        flow_rate (EL ind_n (pt_list p))) * time p
```

The function n_io:$(sw \rightarrow real)$ accepts an element of data type sw and returns a real number, which is the number of vehicles in a region.

In order to obtain the number of vehicles in n regions, we write the following HOL Light function:

Definition 26: Queue Size for n Regions via Input-output Analysis

```
⊢def io_list [ ] = [ ] ∧
io_list (CONS h t) = CONS ((n_io h) (io_list t))
```

The function io_list:$((sw)list \rightarrow real\ list)$ accepts a list of the elements of data type sw, and a list with each of its element as a real number, and its each element is the number of vehicles in any region.

Definition 27: Summation of the Number of Vehicles via Input-output Analysis

```
⊢def ∀L. sum_io L =
          sum (1..LENGTH L) (λi. EL (i - 1) (io_list L))
```

The above function sum_io:$((sw)list \rightarrow real)$ accepts sw list and returns the total number of vehicles in all the regions, i.e., the sum of the number of vehicles via input-output analysis according to (Equation (16)).

Next, in order to formalize (Equation (15)), we first model a single term, i.e., for $n = 1$, as follows:

Definition 28: Number of Vehicles in a Single Region

```
⊢def ∀r. sw_rgn r =
  −(shock_wv (EL (ind_n (HD(FST r)))
        (pt_list (HD(FST r))) * time (HD(FST r)) -
   sum_sw (TL(FST r))) * (SND(SND r) - FST(SND r))
```

The function sw_rgn:$(sw\_d \rightarrow real)$ accepts an element of data type sw_d and returns the number of vehicles in a single region via shockwave analysis.

In order to obtain the number of vehicles in n regions, we write the following HOL Light function:

Definition 29: Queue Size (Number of Vehicles for n Regions) via Shockwave Analysis

```
⊢def sw_rgn_list [ ] = [ ] ∧ sw_rgn_list (CONS h t) =
          CONS (sw_rgn h) (sw_rgn_list t)
```

The function sw_rgn_list:$((sw\_d)list \rightarrow (real)list)$ accepts an (sw_d)list and returns a list containing the number of vehicles in n regions. It uses sw_rgn (Definition 28) to obtain the number of vehicles in a single region.

In order to obtain the summation of shockwaves in an entire area, we need to sum up the individual accumulative shockwaves effect in the individual regions.

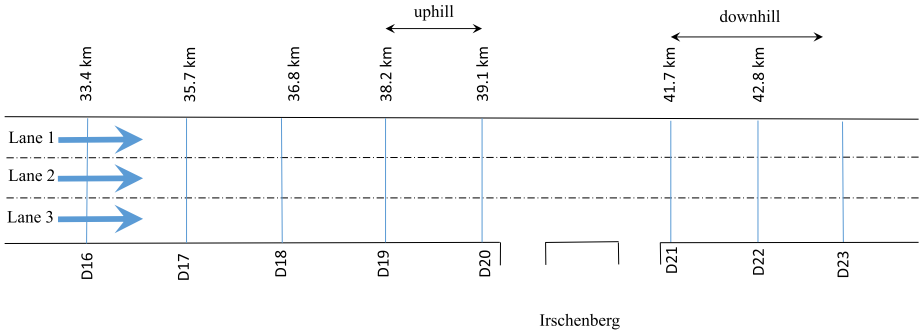

FIGURE 5. German Freeway [19].

Definition 30: Sum of the Regional List

$\vdash_{\text{def}}$ ∀L. sum_sw_rgn L =
sum (1..LENGTH L) (λi. EL (i - 1) (sw_rgn_list L))

This function sum_sw_rgn:$((sw\_d)list \rightarrow real)$ accepts an sw_d list and returns the total number of vehicles in all the regions, i.e., total number of vehicles via shockwave analysis.

The formalization presented in this section took about 500 lines-of-code and 50 man-hours. All the verified theorems are of generic nature as all the variables are universally quantified and can be specialized to obtain the formal analysis of any transportation system.

VI. CASE STUDIES

In order to illustrate the utilization and effectiveness of our proposed framework, we formally analyze the German freeway by verifying its foremost property depicting the average vehicle flow in different lanes [19]. We also present the formal shockwave and input-output analysis and consistency between both of these analyses [13].

A. GERMAN FREEWAY

We utilize our formalization, presented in Section V-A, to formally model and verify some vital properties of a macroscopic model of the German A8-East from Munich to Salzburg freeway [19], as shown in Figure 5.

There are three traffic lanes on this freeway as shown in Figure 5. In this case study, we consider two, three and four vehicles traveling on lanes 1, 2 and 3 of the freeway, respectively. Based on these parameter the macroscopic traffic flow model for the first lane is given by the following definition:

Definition 31: Macroscopic Model of Lane 1

$\vdash_{\text{def}}$ ∀xv tv xh th L11v V11v L12v V12v L11h V11h L12h V12h. german_freeway_lane_1 xv tv xh th
L11v V11v L12v V12v L11h V11h L12h V12h
= (xv,tv,xh,th), [L11v,V11v; L12v,V12v],
[L11h,V11h; L12h,V12h]

where xv, tv, xh and th are the pairs containing the starting and ending points of the lengths and widths of the regions S_1 and S_2, respectively. Similarly, Lijv and Lijh represent the length of the j^{th} vehicle in the i^{th} lane in the regions S_1 and S_2, respectively, whereas Vijv and Vijh represent the speed of the j^{th} vehicle in the i^{th} lane in the regions S_1 and S_2, respectively.

The function german_freeway_lane_1 accepts all of these parameters and returns the time-space macroscopic model of the traffic flow on Lane 1 of the German freeway.

Similarly, the following definitions provide the macroscopic models of traffic flow for Lanes 2 and 3, respectively.

Definition 32: Macroscopic Model of Lane 2

$\vdash_{\text{def}}$ ∀xv tv xh th L21v V21v L22v V22v L23v V23v L21h V21h L22h V22h L23h V23h.
german_freeway_lane_2 xv tv xh th L21v V21v L22v V22v L23v V23v L21h V21h L22h V22h L23h V23h =
(xv,tv,xh,th), [L21v,V21v; L22v,V22v; L23v,V23v],
[L21h,V21h; L22h,V22h; L23h,V23h]

Definition 33: Macroscopic Model of Lane 3

$\vdash_{\text{def}}$ ∀xv tv xh th L31v V31v L32v V32v L33v V33v L34v V34v L31h V31h L32h V32h L33h V33h L34h V34h.
german_freeway_lane_3 xv tv xh th L31v V31v L32v V32v L33v V33v L34v V34v L31h V31h L32h V32h L33h V33h L34h V34h = (xv,tv,xh,th),
[L31v,V31v; L32v,V32v; L33v,V33v; L34v,V34v],
[L31h,V31h; L32h,V32h; L33h,V33h; L34h,V34h]

Our next step is to formally verify the lane-averaged vehicle flow of the considered German freeway:

Theorem 4: Lane-Averaged Mean Velocity of the German Freeway

$\vdash_{\text{thm}}$ ∀xv tv xh th L11v V11v L12v V12v L11h V11h L12h V12h L21v V21v L22v V22v L23v V23v L21h V21h L22h V22h L23h V23h L31v V31v L32v.
[A1:] &0 < (SND th - FST th) ∧
[A2:] &0 < (SND xh - FST xh) ∧
[A3:] L = 3
⟹ sum (1..L) (λi. EL (i - 1) [
flow_rate_s2 (german_freeway_lane_1 xv tv xh th
L11v V11v L12v V12v L11h V11h L12h V12h) / &L;
flow_rate_s2 (german_freeway_lane_2 xv tv xh th
L21v V21v L22v V22v L23v V23v L21h V21h
L22h V22h L23h V23h) / &L;
flow_rate_s2 (german_freeway_lane_3 xv tv xh th
L31v V31v L32v V32v L33v V33v L34v V34v L31h
V31h L32h V32h L33h V33h L34h V34h) / &L]) =
&9 / (&L ∗ (SND th - FST th))

where the assumptions A1-A2 ensure that the length and width of the region S_2 are positive, as they represent time ΔT and distance dX, respectively. The assumption A3 represents the number of lanes in the freeway. The conclusion of Theorem 4 represents the lane-averaged mean velocity of the freeway. The proof process starts by rewriting with the definitions of the functions german_freeway_lane_1,

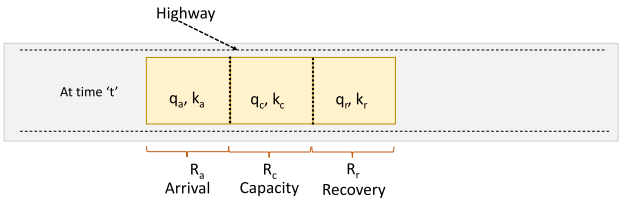

FIGURE 6. Highway Regions.

german_freeway_lane_2, german_freeway_lane_3 and density_s2. Next, the goal is verified using some properties from the list theory and the sum function and some arithmetic reasoning [50].

B. FORMAL INPUT-OUTPUT AND SHOCKWAVE ANALYSES AND THEIR CONSISTENCY

We use our formalization of shockwave, presented in Section V-B, to formally verify the queue size/number of vehicles based on both the input-output and shockwave analysis [13]. The input-output model (also called cumulative arrival and departure model) is commonly used to describe traffic congestions on highways. Conventionally, the queue size at any time can be measured by the difference between the cumulative arrival and the departure curves (shown in the Figure 7). In the same way, shockwave analysis keeps track of the queue propagation, discharging and dissipation. The queue size is measured by the product of the queue length and density at any time via Shockwave analysis (Figure 7). The difference between these two analyses is that the input-output analysis keeps track of queue length and also travel time by considering the time dimensions only, unlike the shockwave analysis, which considers both dimensions, i.e., time and density.

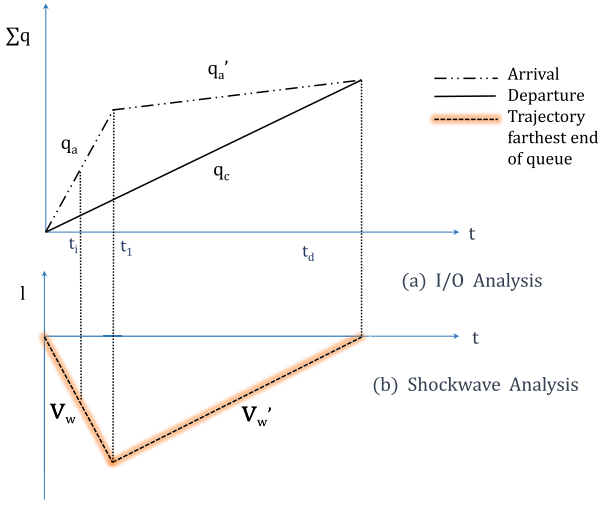

FIGURE 7. Queuing Dynamics with a Change in Arriving Flow Rate [13].

The traffic flow patterns for the considered scenarios are shown in Figures 7 and 8, indicating the changes in arrival and discharging flow rates, where a queue is considered as

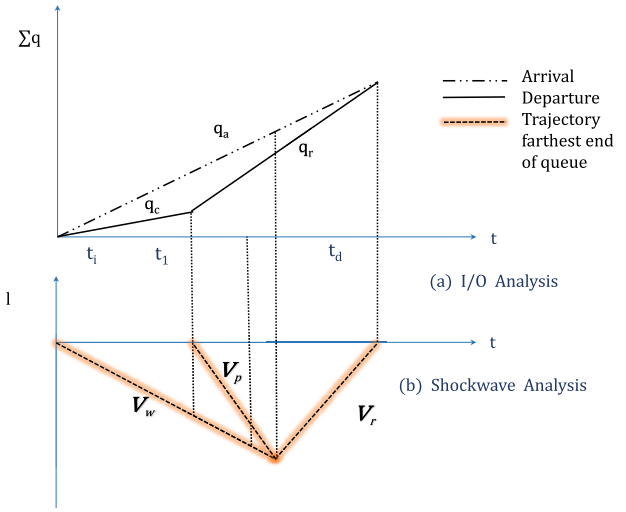

FIGURE 8. Queuing Dynamics with a Change in Discharging Flow Rate [13].

a lane or sequence of vehicles that are waiting for their turn to be attended. A region where the vehicles experience bottleneck or some obstructions on some highway are called highly dense regions. The upstream region considered along the direction of traffic flow is an area before the point of observation and the downstream region is formed after the point of observation. Considered the scenario depicted in Figure 6, having an upstream density and flow rate of q_a and k_a and a downstream flow and density of q_c and k_c at one time instant. After some time, i.e., at time t_1 of Figure 7, congestion is observed. As a result, the density and flow rate in the arrival region gets disturbed due to congestion in the upcoming region (capacity) and thus a new region is formed having flow $q_{a'}$ and density $k_{a'}$.

Then, at the next time instant, when the queue starts dissolving then the discharging (congestion distortion) rate is introduced in the process. Figure 8 shows the comparative sketches for the change of queue dissolving rates for the input-output model and the shockwave analysis. Whereas, at time t_1, another boundary or releasing wave (shockwave), i.e., v_p (due to partial removal of incident) occurs for time $t_1 \rightarrow t_2$. After time instant t_2, when v_p reaches the farthest end of the the queue, another wave v_r is formed (Figure 8), due to the introduction of a new Region R_r (Figure 6) in the process, and terminates until the complete removal of incident/congestion till time t_d as shown in Figure 8.

In order to illustrate the effectiveness of our proposed formalization (Section V-B), we present the formal input-output and shockwave analysis of a highway considering three regions as shown in Figure 6. Moreover, we prove the consistency of both these analyses by verifying that the number of vehicles in both cases, i.e., via input-output and shockwave analysis are same for each of the region.

For time interval $t_0 \rightarrow t_1$ (Figure 7), congestion propagates upstream and the queue size can be described by the input-output and shockwave analysis. Changes in flow rate in

the upstream region is called a change in the arriving demand. Then at t_1, the arriving flow rate is reduced to $q_{a'}$, (or q_{a_2}) due to the accumulation of vehicles in the arrival region.

The queue size for the change in arriving demand for two time intervals can be represented, using Equation (16), as:

$$N_{io} = \sum_{i=1}^{2} (q_{m_i} - q_{n_i}) \Delta t_i \tag{19}$$

$$N_{io} = (q_a - q_c)t_1 + (q_{a_2} - q_c)(t - t_1) \tag{20}$$

We formally verify the queue size for the case of change in arriving demand via input-output analysis as follows:

Theorem 5: Queue Size for Change in Arriving Demand via Input-output Analysis

```
⊢thm ∀vw vw2 qa qa2 qc ka ka2 kc t1 t.
sum_io [([(qa, ka), vw; (qc, kc), vw], &0, t1), 0, 1;
        ([(qa2, ka2), vw2; (qc, kc), vw2], t1, t), 0, 1] =
                    (qa - qc) * t1 + (qa2 - qc) * (t - t1)
```

where vw and vw2 represent the shockwave speeds before and after the change in arrival flow, respectively. Similarly, qa, qc, qa2 and ka, kc, ka2 represent flow rates and densities in approaching/arrival, capacity and changed approaching state, respectively. The reasoning process of Theorem 5 is based on rewriting with Definition 27, properties of HOL Light function sum along with some arithmetic reasoning.

In the same way, Equation (15) for the arrival/approaching change, as shown in Figure 7, for shockwave analysis is mathematically expressed as:

$$N_{sw} = \sum_{j=1}^{2} -(v_{w_j} \Delta t_j - \sum_{i=0}^{0} v_{w_i} \Delta t_i) \Delta k_j \tag{21}$$

$$N_{sw} = -(v_{w_1}(t_1 - t_0)(k_c - k_a) - 0) - (v_{w_2}(t - t_1)(k_c - k_{a_2}) - 0) \tag{22}$$

Using the values of v_{w_1} and v_{w_2} in Equation (22) along with some arithmetic simplification results into the following equation [13].

$$N_{sw} = (q_a - q_c)t_1 + (q_{a_2} - q_c)(t - t_1) \tag{23}$$

Similarly, we verify the queue size as described in the above equations for the case of change in arriving time via shockwave analysis as the following HOL Light theorem:

Theorem 6: Queue Size for Change in Arriving Demand via Shockwave Analysis

```
⊢thm ∀vw vw2 qc qa qa2 ka kc ka2 t1 t.
[A1:] ka ≠ kc ∧    [A2:] ka2 ≠ kc ∧
[A3:] ka ≠ &0 ∧    [A4:] ka2 ≠ &0 ∧
[A5:] kc ≠ &0 ∧
[A6:] (∀p. n_crossing p (ind_n p) =
                n_crossing p (ind_m p)) ∧
[A7:] (∀p. shock_wv (EL (ind_n p) (pt_list p)) =
                shock_wv (EL (ind_m p) (pt_list p))) ∧
[A8:] (∀p. &0 < time p)
⟹ sum_sw_rgn
  [[([(qa, ka), vw; (qc, kc), vw], &0, t1), 0, 1], ka,kc;
   [([(qa2, ka2), vw2; (qc, kc), vw2], t1, t), 0, 1], ka2, kc] =
            (qa - qc) * t1 + (qa2 - qc) * (t - t1)
```

where vw and vw2 represent the shockwave speeds before and after the change in arriving flow (Figure 7), respectively. Similarly, qa, qc, qa2 and ka, kc, ka2 represent flow rates and densities in arrival, capacity and changed arrival regions, respectively. The assumptions A1-A2 of Theorem 6 ensure that the densities in both regions are not same. The assumptions A3-A5 model the conditions that the densities of vehicles in the considered three regions are non-zero. The assumption A6 says that the number of vehicles crossing (Definition 22) from one region's front boundary (Equation (17)) would be the same as to the other adjacent region's number of vehicles crossing from the rear adjacent boundary and is thus according to the universal law of conservation. The assumption A7 models the condition that the shockwave speed in both regions are same. The assumption A8 models the non-negativity condition for the time interval. Finally, the conclusion of Theorem 6 presents the number of vehicles (queue size) for change in arrival demand via shockwave analysis. The reasoning process for Theorem 6 is based on Definition 30, properties of HOL Light's function sum along with some real arithmetic reasoning. More details about its proof can be found at [50].

Then at the next time instant, the arriving flow rate reaches its maximum value and the discharging flow rate in the downstream region starts changing. Thus leads to two cases, i.e., first for the queue propagation when the queue size increases and the second for the queue dissipation when the queue size decreases. This phenomenon is shown in Figure 8.

Changes in queue size, such as the introduction of discharging rate in the process, can be represented as follows:

$$N_{io} = \sum_{i=1}^{2} (q_{m_i} - q_{n_i}) \Delta t_i \tag{24}$$

$$N_{io} = (q_a - q_c)t_1 + (q_a - q_r)(t - t_1) \tag{25}$$

The queue size (number of vehicles) for the case of queue propagation via input-output analysis is formalized as follows:

Theorem 7: Queue Size for Case of Queue Propagation via Input-output Analysis

```
⊢thm ∀vw vp qa qc ka kc kr t1 t.
 sum_io [([(qa, ka), vw; (qc, kc), vw], &0, t1), 0, 1;
   ([(qa, ka), vp; (qr, kr), vp], t1, t), 0, 1] =
          (qa - qc) * t1 + (qa - qr) * (t - t1)
```

where all the input variables are the same as in Theorem 5 except kc and kr, which represent the densities in the capacity and the recovery regions (Figure 8), respectively. Similarly vp is the shockwave speed. The proof of this theorem is similar to the one for Theorem 5, and its details can be found in [50].

The queue size for the case of queue propagation via shockwave analysis, is mathematically expressed as:

$$N_{sw} = \sum_{j=1}^{2} -(v_{w_j} \Delta t_j - \sum_{i=0}^{1} v_{w_i} \Delta t_i) \Delta k_j \tag{26}$$

$$N_{sw} = -\{v_w(t - t_0) - v_p(t - t_1)\}(k_c - k_a) + -\{v_p(t - t_1) - 0\}(k_r - k_a) \tag{27}$$

Using the values of v_w and v_p in Equation (27) along with some arithmetic simplification results into the following

equation [13].

$$N_{sw} = (q_a - q_c)t_1 + (q_a - q_r)(t - t_1) \tag{28}$$

Similarly, we verify the queue size as described in the above equation for the case of change in propagation via shockwave analysis as the following HOL Light theorem:

Theorem 8: Queue Size for the case of Queue Propagation via Shockwave Analysis

```
⊢thm ∀vp vw qa qc qr ka kc kr t1 t.
[A1:] ka ≠ kc ∧   [A2:] kc ≠ kr ∧
[A3:] ka ≠ &0 ∧   [A4:] kc ≠ &0 ∧
[A5:] kr ≠ &0 ∧
[A6:] (∀p. n_crossing p (ind_n p) =
           n_crossing p (ind_m p)) ∧
[A7:] (∀p. shock_wv (EL (ind_n p) (pt_list p)) =
           shock_wv (EL (ind_m p) (pt_list p))) ∧
[A8:] (∀p. &0 < time p)
  ⟹ sum_sw_rgn
   [[([qc, kc), vp; (qr, kr), vp], t1, t), 0, 1], ka, kr;
    [([qa, ka), vw; (qc, kc), vw], &0, t), 0, 1;
     ([qc, kc), vp; (qr, kr), vp], t1, t), 0, 1], ka, kc] =
      (qa - qc) * t1 + (qa - qr) * (t - t1)
```

where vw and vp represent the shockwave speeds before the propagation starting in arrival region and after the propagation introduced at time t_1 (Figure 8), respectively. vw is the shockwave between arrival and capacity regions and vp is between capacity and recovery regions (Figure 6). Similarly, qa, qc, qr and ka, kc, kr are representing flow rates and densities in approaching/arrival, capacity and recovery regions, respectively. All the assumptions for Theorem 8 are the same as for Theorem 6 and also describing the same conditions. Similarly, the verification process for the above theorem is the same as that of Theorem 6.

The conclusions of both Theorems 7 and 8 show that the queue size in the case of input-output and shockwave analyses are the same, which means both analyses are consistent.

The releasing wave vp reaches the maximum possible end of the queue at time t_2. This is where the recovery of the normal operation of the highway begins i.e., the congestion of the queue starts dissipating (this phenomenon is called queue dissipation) and another wave v_r starts to grow as a result of discharging flow q_r (Figure 8). This emerging wave v_r moves downstream until the complete removal of the queue congestion at t, and the road way section returns to its normal operating condition.

The queue size via Equation (16) for the case of queue dissipation via input-output analysis is represented as follows:

$$N_{io} = \sum_{i=1}^{3}(q_{m_i} - q_{n_i})\Delta t_i \tag{29}$$

$$N_{io} = (q_a - q_c)t_1 + (q_a - q_r)(t_2 - t_1) + (q_a - q_r)(t - t_2) \tag{30}$$

Theorem 9: Queue Size for the Case of Queue Dissipation via Input-output Analysis

```
⊢thm ∀vw vp vr qa qc qr ka kc kr t1 t.
sum_io [([(qa, ka), vw; (qc, kc), vw], &0, t1), 0, 1;
  ([(qa, ka), vr; (qr, kr), vr], t1, t2), 0, 1;
    ([(qa, ka), vr; (qr, kr), vr], t2, t), 0, 1] =
      (qa - qc) * t1 + (qa - qr) * (t2 - t1) + (qa - qr) *(t - t2)
```

The verification process for the above theorem is very similar to that of Theorems 5 and 7.

We verify the queue size for the case of queue dissipation via shockwave analysis, which is mathematically expressed as:

$$N_{sw} = \sum_{j=1}^{2} -(v_{w_j}\Delta t_j - \sum_{i=0}^{0} v_{w_i}\Delta t_i)\Delta k_j \tag{31}$$

$$N_{sw} = \{-(v_w - 0)(t_2 - t_0) + -(v_r - 0)(t - t_2)\}(k_r - k_a) \tag{32}$$

Using the values of v_w and v_r in Equation(32) along with some arithmetic simplification results into the following equation [13].

$$N_{sw} = (q_a - q_c)t_1 + (q_a - q_r)(t_2 - t_1) + (q_a - q_r)(t - t_2) \tag{33}$$

Theorem 10: Queue Size for the Case of Queue Dissipation via Shockwave Analysis

```
⊢thm ∀vp vw vr qa qc qr ka kc kr t1 t2 t.
[A1:] ka ≠ kc ∧  [A2:] kc ≠ kr ∧  [A3:] ka ≠ kr ∧
[A4:] ka ≠ &0 ∧ [A5:] kc ≠ &0 ∧ [A6:] kr ≠ &0 ∧
[A7:] (∀p. n_crossing p (ind_n p) =
           n_crossing p (ind_m p)) ∧
[A8:] (∀p. shock_wv (EL (ind_n p) (pt_list p)) =
           shock_wv (EL (ind_m p) (pt_list p))) ∧
[A9:] (∀p. &0 < time p) ∧
[A10:] sw_rgn ([([(qa, ka), vw; (qc, kc), vw],
          (&0, t2)), 0, 1], ka, kr) = sum_sw_rgn [
        [([(qc,kc), vp; (qr, kr), vp], t1, t2), 0, 1], ka, kr;
        [([((qa,ka), vw); (qc, kc), vw], (&0, t2)), (0, 1);
        ([((qc,kc), vp); (qr, kr), vp], t1, t2), 0, 1], ka, kc]
⟹ sum_sw_rgn
   [[([(qa,ka), vw; (qc, kc), vw], (&0, t2)), (0, 1)], ka, kr;
     [([(qa,ka), vr; (qr, kr), vr], (t2, t)), (0, 1)], ka, kr] =
       (qa - qc) * t1 + (qa - qr) * (t2 - t1) + (qa - qr) * (t - t2)
```

where the assumptions A1-A9 of the above theorem are the same as that of Theorems 6 and 8. While the assumption A10 describes that the queue size is already evaluated before time t_2 (in the last time interval) in Theorem 8. Finally, the conclusion of the above theorem represents the queue size in the case of queue dissipation.

The proof of this theorem is similar to Theorems 6 and 8. More details about whole of the formalization, presented in this section, can be found at [50]. Again, the conclusions of both theorems, i.e., Theorems 9 and 8, show that the queue size in the case of input-output and shockwave analyses are the same, which means both analyses are consistent.

The formal analysis presented, in Sections VI and V, took about 1200 lines-of-code and 110 man-hours. Moreover, the straightforward proof scripts for the properties, verified in this section, clearly indicate the usefulness of our foundational formalization presented in Section V of this paper. The effort involved in the verification of the individual theorem in the form of proof lines and the man-hours is presented in Table 2. The verification of Theorem 4 took only 95 lines of HOL Light code and 5 man-hours, which clearly illustrates the benefit of Theorems 1 and 2, and Lemma 1 regarding formalization of the macroscopic models of traffic flow. It is

TABLE 2. Verification Details for each Theorem.

Formalized Theorems	Proof Lines	Man-hours
Theorem 1 (Relationship of Relative Occupancy and Density in Region S_1)	90	7
Lemma 1 (Summation of a List having Same Element c is c Times Length of List)	105	8
Theorem 2 (Relationship of Relative Occupancy and Density in Region S_2)	225	17
Theorem 3 (Shockwave Speed Verification in two Regions R_m and R_n)	165	13
Theorem 4 (Lane-Averaged Mean Velocity of the German Freeway)	95	5
Theorem 5 (Queue Size for Change in Arriving Demand via Input-output Analysis)	20	3
Theorem 6 (Queue Size for Change in Arriving Demand via Shockwave Analysis)	145	17
Theorem 7 (Queue Size for the Case of Queue Propagation via Input-output Analysis)	20	4
Theorem 8 (Queue Size for the case of Queue Propagation via Shockwave Analysis)	170	18
Theorem 9 (Queue Size for the Case of Queue Dissipation via Input-output Analysis)	25	5
Theorem 10 (Queue Size for the Case of Queue Dissipation via Shockwave Analysis)	140	13

important to note that the man-hours are based on the number of lines of code as well as the complexity of the proof. So the number of lines of the proof script do not have a direct relationship with the man-hours. For example, the man-hours for the verification of Theorems 4 and 9 are same, whereas the proof lines for the verification of the former are greater than that for the later.

Our formalization can be utilized to formally reason about many other macroscopic model related properties and the results would be guaranteed to be correct due to the inherent soundness of theorem proving. Moreover, our theorems are generic in nature, i.e., all the variables in these theorems are universally quantified and thus can be specialized to any values based on a particular scenario. Thus, our formally verified relative occupancy and shockwave alongside other properties of the macroscopic model can be used to ensure an uninterrupted flow of traffic and its resumption in the case of any interruptions in the flow, resulting in a more reliable traffic flow. To the best of our knowledge, no other computer-based analysis technique can provide such benefits.

VII. CONCLUSIONS

With the increasing complexity of the transportation network through connectivity and automation, it is becoming more important to rely on more rigorous methods to ensure that the designed network follows the intended design. This task is performed at different levels of abstraction. For example, at the microscopic level, verification and proof of correct functionality can be performed to ensure protocols are followed by autonomous vehicles during safety-critical events when vehicles are in a collision course. At the macro-level, a guarantee of the particular level of service is maintained while satisfying all the network constraints and limitations, such as maximum density, for example. Simulation methods tend to address such problems by sampling the solution space and providing confidence in the performance outcome, like in the case of Monte Carlo simulation. Covering all possible traffic scenarios (the whole possible state of the network) is not possible using simulation-based methods due to the high computational overhead. Moreover, as the transportation network becomes more technologically dependent, it becomes more prone to computer vulnerability from bugs to security deficiencies and sample-based methods can be short of guaranteeing the resilience or even the reliability of the network. Therefore, there is an urgency to develop new validation paradigms for transportation applications.

In this paper, we propose to use higher-order-logic theorem proving for analyzing macroscopic models of traffic flow. Due to the high expressiveness of the underlying logic, we can formally model the continuous components of macroscopic models while capturing their actual behavior, and the soundness of theorem proving guarantees correctness of results. We formally model the necessary parameters of a transportation system, which include density, flow rate, speed, relative occupancy, and shockwave and used our formalization to formally analyze a German freeway and a commonly used highway, by performing the input-output and shockwave analyses. The primary challenge in the proposed approach is the enormous amount of user intervention required due to the undecidable nature of the logic. We propose to overcome this limitation by formalizing the foundational mathematical theories and core concepts of traffic flow so that these available results can be built upon to minimize user interaction. The case studies demonstrated the practicability of this approach. Our proposed approach can equally be used to plan and model various components of the transportation systems, such as, highway links, diverges, merges and stations, and thus to address the problem of routing vehicles in a network of automated vehicles [21].

We plan to expand on the formalization of basic concepts in macro-modeling theory, including queuing models, driver behavior and level of service. We also plan to formally verify the dynamic user-equilibrium, which involves algorithms minimizing the travel cost function for each of the travelers, departing from the same origin to same destination at the same time, which should be equal and minimal for all routes. Another future direction is to develop formal reasoning support for the microscopic models of traffic flow theory. Modeling of the equations, capturing the dynamics of the microscopic model, would include the formal modeling of interactions between vehicles and their characteristics [11], [24]. At the micro level, conflict-based indicators are essential measures of traffic safety that complement collision based data. Different indicators represent competing views for safety hazardous situations based on temporal, spatial predictions of potential collisions. While

different transportation agencies widely adopt methods, such as time to collision, there is no formal procedure for estimating their severity. A formal approach for modeling and verifying those indicators based on deductive methods would fill a significant research gap in modern traffic safety; as those indicators are considered vital for communicating vehicles [53], [54]. The next stage of the research is to consider modeling and verifying connectivity, computer, and autonomous elements in intelligent transportation systems, with focus on functionality soundness of platooning dynamics, control strategies, and safety requirements [55]–[57].

REFERENCES

[1] A. Fantechi, F. Flammini, and S. Gnesi, "Formal methods for intelligent transportation systems," in *Proc. Int. Symp. Leveraging Appl. Formal Methods, Verification Validation* LNCS, vol. 7610. Berlin, Germany: Springer, 2012, pp. 187–189.
[2] M. Kamali, L. A. Dennis, O. Mcaree, M. Fisher, and S. M. Veres, "Formal verification of autonomous vehicle platooning," *Sci. Comput. Program.*, vol. 148, pp. 88–106, Nov. 2017.
[3] A. Platzer, "Verification of cyber-physical transportation systems," *Intell. Syst.*, vol. 24, no. 4, pp. 10–13, 2009.
[4] A. Tiwari, "Formal verification of transportation cyber physical systems," in *Proc. Nat. Workshop Res. High-Confidence Transp. Cyber-Phys. Syst. Automot., Aviation, Rail Cyber-Phys. Syst.*, 2008, pp. 18–20.
[5] S. Mitsch, S. M. Loos, and A. Platzer, "Towards formal verification of freeway traffic control," in *Proc. IEEE/ACM Third Int. Conf. Cyber-Phys. Syst.*, Apr. 2012, pp. 171–180.
[6] O. Hasan and S. Tahar, "Formal verification methods," in *Encyclopedia of Information Science and Technology*. Hershey, PA, USA: IGI Global, 2015, pp. 7162–7170.
[7] J. Harrison, *Handbook of Practical Logic and Automated Reasoning*. Cambridge, U.K.: Cambridge Univ. Press, 2009.
[8] N. Gartner, C. J. Messer, and A. K. Rathi, "Traffic flow theory: A state-of-the-art report," Oak Ridge Nat. Lab., Transp. Res. Board US Dept. Transp., Tech. Rep. 165, 2001.
[9] S. K. Wu, M. P. Hunter, C. Lee, and M. O. Rodgers, "Evaluation of traffic signal control system using a system-wide performance measure under two-fluid model theory," *KSCE J. Civil Eng.*, vol. 15, no. 2, pp. 395–403, Feb. 2011.
[10] J. Harrison, "HOL light: An overview," in *Theorem Proving in Higher Order Logics*. Berlin, Germany: Springer, 2009, pp. 60–66.
[11] S. Maerivoet and B. De Moor, "Traffic flow theory," 2005, *arXiv:physics/0507126*. [Online]. Available: https://arxiv.org/abs/physics/0507126
[12] L. H. Immers and S. Logghe, "Traffic flow theory," Fac. Eng., Dept. Civil Eng., Sect. Traffic Infrastruct., Kasteelpark Arenberg, Belgium, vol. 40, 2002.
[13] P. Yi, Z. Tian, and Q. Zhao, "Consistency of input-output model and shockwave analysis in queue and delay estimations," *J. Transp. Syst. Eng. Inf. Technol.*, vol. 8, no. 6, pp. 146–152, Dec. 2008.
[14] J. Barcelo, *Fundamentals of Traffic Simulation*. Springer, 2010, vol. 145.
[15] H. M. Zhang, R. Kühne, and P. Michalopoulos, "Continuum flow models, traffic flow theory: A state-of-the-art report," in *Transportation Research Board, original text by Reinhart Kuhne and Panos Michalopoulos*, 2001, ch. 5.
[16] R. P. Roess, E. S. Prassas, and W. R. McShane, *Traffic Engineering*. Upper Saddle River, NJ, USA: Prentice-Hall, 2004.
[17] S. P. Hoogendoorn, H. Botma, and M. M. Minderhoud, *Traffic Flow Theory and Simulation*. Delft, The Netherlands: Faculty of Civil Engineering and Geosciences, 2007.
[18] N. Motamedidehkordi, M. Margreiter, and T. Benz, "Shockwave analysis on motorways and possibility of damping by C2X applications," *Mobil. TUM-Technol., Solutions Perspect. Intell. Transp. Syst.*, Jun./Jul. 2015.
[19] M. Treiber and D. Helbing, "Reconstructing the spatio-temporal traffic dynamics from stationary detector data," *Cooperat. Transp. Dyn.*, vol. 1, no. 3, pp. 1–3, 2002.
[20] H. Rakha and W. Zhang, "Consistency of shock-wave and queuing theory procedures for analysis of roadway bottlenecks," in *Proc. 84th Annu. Meeting Transp. Res. Board*, 2005, pp. 219–226.
[21] A. J. Pue, "Macroscopic traffic models for vehicle-follower automated transportation systems," *Transp. Res. B, Methodol.*, vol. 16, no. 2, pp. 125–142, Apr. 1982.
[22] H. Lieu, N. Gartner, C. Messer, and A. Rathi, "Traffic flow theory," *Public Roads*, vol. 62, pp. 45–47, 1999.
[23] B. S. Kerner, "The physics of traffic," *Phys. World*, vol. 12, no. 8, p. 25, 1999.
[24] S. Panwai and H. Dia, "Comparative evaluation of microscopic car-following behavior," *IEEE Trans. Intell. Transp. Syst.*, vol. 6, no. 3, pp. 314–325, Sep. 2005.
[25] B. D. Greenshields, W. Channing, and H. Miller, "A study of traffic capacity," in *Highway Research Board Proceedings*, vol. 1935. Washington, DC, USA: National Research Council, 1935.
[26] M. Fellendorf and P. Vortisch, "Microscopic traffic flow simulator VISSIM," in *Fundamentals of Traffic Simulation*. New York, NY, USA: Springer, 2010, pp. 63–93.
[27] J. Cao, M. Hadiuzzaman, T. Z. Qiu, and D. Hu, "Real-time queue estimation model development for uninterrupted freeway flow based on shockwave analysis," *Can. J. Civ. Eng.*, vol. 42, no. 3, pp. 153–163, Mar. 2015.
[28] M. Ben-Akiva, H. N. Koutsopoulos, T. Toledo, Q. Yang, C. F. Choudhury, C. Antoniou, and R. Balakrishna, "Traffic simulation with MITSIMLab," in *Fundamentals of Traffic Simulation*. New York, NY, USA: Springer, 2010, pp. 233–268.
[29] A. J. Durán, M. Pérez, and J. L. Varona, "The misfortunes of a trio of mathematicians using computer algebra systems. Can we trust in them?" *Notices Amer. Math. Soc.*, vol. 61, no. 10, p. 1249, Nov. 2014.
[30] S. M. Loos, A. Platzer, and L. Nistor, "Adaptive cruise control: Hybrid, distributed, and now formally verified," in *Formal Methods* (Lecture Notes in Computer Science), vol. 6664. Berlin, Germany: Springer, 2011, pp. 42–56.
[31] A. Platzer and J. D. Quesel, "Keymaera: A hybrid theorem prover for hybrid systems (system description)," in *Proc. Int. Joint Conf. Automated Reasoning* Lecture Notes in Computer Science, vol. 5195. Berlin, Germany: Springer, 2008, pp. 171–178.
[32] O. Hasan and M. Ahmad, "Formal analysis of steady state errors in feedback control systems using HOL light," in *Proc. Conf. Design, Autom. Test Eur.*, 2013, pp. 1423–1426.
[33] M. Ahmad and O. Hasan, "Formal verification of steady state errors in unity-feedback control systems," in *Proc. Int. Workshop Formal Methods Ind. Crit. Syst.* LNCS, vol. 8718. Cham, Switzerland: Springer, 2014, pp. 1–15.
[34] A. Rashid and O. Hasan, "Formal analysis of linear control systems using theorem proving," in *Int. Conf. Formal Eng. Methods* Lecture Notes in Computer Science, vol. 10610. Cham, Switzerland: Springer, 2017, pp. 345–361.
[35] A. Rashid, "Formalization of transform methods using higher-order-logic theorem proving," Ph.D. dissertation, School Elect. Eng. Comput. Sci., Nat. Univ. Sci. Technol., Islamabad, Pakistan, 2019.
[36] S. H. Taqdees and O. Hasan, "Formalization of laplace transform using the multivariable calculus theory of HOL light," in *Proc. Conf. Log. Program., Artif. Intell., Reasoning* LNCS, vol. 8312. Berlin, Germany: Springer, 2013, pp. 744–758.
[37] A. Rashid and O. Hasan, "Formalization of Lerch's theorem using HOL light," *J. Appl. Log. IFCoLog J. Log. Their Appl.*, vol. 5, no. 8, pp. 1623–1652, 2018.
[38] A. Rashid and O. Hasan, "On the formalization of Fourier transform in higher-order logic," in *Proc. Int. Conf. Interact. Theorem Proving* Lecture Notes in Computer Science, vol. 9807. Springer, 2016, pp. 483–490.
[39] A. Rashid and O. Hasan, "Formalization of transform methods using HOL light," in *Intelligent Computer Mathematics* (LNAI), vol. 10383. Cham, Switzerland: Springer, 2017, pp. 319–332.
[40] M. U. Sanwal and O. Hasan, "Formally analyzing continuous aspects of cyber-physical systems modeled by homogeneous linear differential equations," in *Proc. Int. Workshop Design, Modeling, Eval. Cyber Phys. Syst.* LNCS, vol. 9361. Cham, Switzerland: Springer, 2015, pp. 132–146.
[41] U. Siddique, O. Hasan, and S. Tahar, "Formal modeling and verification of integrated photonic systems," in *Proc. Annu. IEEE Syst. Conf. (SysCon)*, Apr. 2015, pp. 562–569.
[42] B. Farooq, O. Hasan, and S. Iqbal, "Formal kinematic analysis of the two-link planar manipulator," in *Proc. Int. Conf. Formal Eng. Methods* LNCS, vol. 8144. Berlin, Germany: Springer, 2013, pp. 347–362.
[43] R. Affeldt and C. Cohen, "Formal foundations of 3D geometry to model robot manipulators," in *Certified Programs and Proofs*. New York, NY, USA: ACM, 2017.

[44] B. Barras, S. Boutin, C. Cornes, J. Courant, J. C. Filliatre, E. Gimenez, H. Herbelin, G. Huet, C. Munoz, and C. Murthy, "The COQ proof assistant reference manual: Version 6.1," Ph.D. dissertation, Inria, Paris, France, 1997.
[45] A. Rashid and O. Hasan, "Formal analysis of robotic cell injection systems using theorem proving," in *Proc. Int. Workshop Design, Modeling, Eval. Cyber Phys. Syst.* LNCS, vol. 11267. Springer, 2017, pp. 127–141.
[46] F. L. Hall, "Traffic stream characteristics," in *Traffic Flow Theory*. Washington, DC, USA: Federal Highway Administration, 1996.
[47] V. Knoop, S. P. Hoogendoorn, and H. van Zuylen, "Empirical differences between time mean speed and sace mean speed," in *Traffic and Granular Flow*. Orsay, France: Springer, 2009, pp. 351–356.
[48] B. S. Kerner, *The Physics of Traffic: Empirical Freeway Pattern Features, Engineering Applications, and Theory*. Berlin, Germany: Springer, 2012.
[49] J. Harrison, "HOL light: A tutorial introduction," in *Formal Methods in Computer-Aided Design* LNCS, vol. 1166. Springer-Verlag, 1996, pp. 265–269.
[50] M. Umair. (2019). *Towards Formal Verification of Macroscopic Models in Traffic Flow Theory*. [Online]. Available: http://save.seecs.nust.edu.pk/projects/tftheory
[51] N. Saxena and N. Jain, "Evaluation of road traffic congestions by shockwave theory and reduction strategies," *Recent Innov. Trends Comput. Commun.*, vol. 5, no. 5, pp. 125–129, 2017.
[52] L. Elefteriadou, *An Introduction to Traffic Flow Theory*. Berlin, Germany: Springer, 2014.
[53] J. R. Ward, G. Agamennoni, S. Worrall, A. Bender, and E. Nebot, "Extending time to collision for probabilistic reasoning in general traffic scenarios," *Transp. Res. C, Emerg. Technol.*, vol. 51, pp. 66–82, Feb. 2015.
[54] C. Schwarz, "On computing time-to-collision for automation scenarios," *Transp. Res. F, Traffic Psychol. Behaviour*, vol. 27, pp. 283–294, Nov. 2014.
[55] D. Jia, K. Lu, and J. Wang, "On the network connectivity of platoon-based vehicular cyber-physical systems," *Transp. Res. C, Emerg. Technol.*, vol. 40, pp. 215–230, Mar. 2014.
[56] K. Li and P. Ioannou, "Modeling of traffic flow of automated vehicles," *IEEE Trans. Intell. Transp. Syst.*, vol. 5, no. 2, pp. 99–113, Jun. 2004.
[57] Y. Li, L. Zhang, H. Zheng, X. He, S. Peeta, T. Zheng, and Y. Li, "Nonlane-discipline-based car-following model for electric vehicles in transportation-cyber-physical systems," *IEEE Trans. Intell. Transp. Syst.*, vol. 19, no. 1, pp. 38–47, Jan. 2018.

ADNAN RASHID received the M.Sc. and M.Phil. degrees in electronics from the Department of Electronics, Quaid-i-Azam University (QAU), Islamabad, Pakistan, in 2008 and 2012, respectively, and the Ph.D. degree in information technology from the School of Electrical Engineering and Computer Science (SEECS), National University of Science and Technology (NUST), Islamabad, in 2019. He has also worked as a Visiting Researcher at the Hardware Verification Group (HVG), Concordia University, Canada, in 2018. He is currently working as a Research Associate with the System Analysis and Verification (SAVe) Laboratory, SEECS, NUST. He has a strong interest in formal methods, with their applications in control systems, analog circuits, biological systems, robotic systems, communication systems, and transportation systems. He has served as the Chair of the Doctoral program at the Conference on Intelligent Computer Mathematics, Edinburgh, U.K., in 2017.

MUHAMMAD UMAIR received the B.S. and M.S. degrees in electrical engineering and systems engineering from the National University of Science and Technology (NUST), Islamabad, Pakistan, in 2013 and 2017, respectively. He was a Research Assistant with the System Analysis and Verification (SAVe) Laboratory, School of Electrical Engineering and Computer Science (SEECS), NUST. He is currently working with the Forensic Science Laboratory, Lahore, Pakistan.

OSMAN HASAN (Senior Member, IEEE) received the B.Eng. (Hons.) degree from the University of Engineering and Technology, Peshawar, Pakistan, in 1997, and the M.Eng. and Ph.D. degrees from Concordia University, Montreal, Quebec, Canada, in 2001 and 2008, respectively. Before his Ph.D., he worked as an ASIC Design Engineer, LSI Logic, from 2001 to 2004. He worked as a Postdoctoral Fellow at the Hardware Verification Group (HVG), Concordia University, for one year until August 2009. He is currently an Associate Professor and the Head of the Department of Electrical Engineering, School of Electrical Engineering and Computer Science, National University of Science and Technology (NUST), Islamabad, Pakistan. He is the Founder and Director of the System Analysis and Verification (SAVe) Lab, NUST, which mainly focuses on the design and formal verification of energy, embedded, and e-health related systems. He has received several awards and distinctions, including the Pakistan's Higher Education Commission's Best University Teacher (2010) and Best Young Researcher Award (2011) and the President's Gold Medal for the Best Teacher of the University from NUST, in 2015. He is a member of the ACM, the Association for Automated Reasoning (AAR), and the Pakistan Engineering Council.

MOHAMED H. ZAKI (Member, IEEE) received the Ph.D. degree from the Hardware Verification Group, Concordia University, Montreal, in 2008. He was a Research Associate at the Bureau of Intelligent Transportation Systems and Freight Security, The University of British Columbia. He is currently an Assistant Professor of transportation engineering with the Civil, Environmental and Construction Engineering Department, University of Central Florida. His multidisciplinary research focuses on solving tomorrow's smart cities problems; from the computing and information to its facilities infrastructure. He studies road safety and road-users' behavior through the automated analysis of traffic data. He serves on the Transportation Research Board (TRB) ABJ70 Committee on Artificial Intelligence and Advanced Computing Applications.