STATE OF CALIFORNIA • DEPARTMENT OF TRANSPORTATION

TECHNICAL REPORT DOCUMENTATION PAGE

TR0003 (REV 10/98)

ADA Notice

For individuals with sensory disabilities, this document is available in alternate formats. For information call (916) 654-6410 or TDD (916) 654-3880 or write Records and Forms Management, 1120 N Street, MS-89, Sacramento, CA 95814.

1. REPORT NUMBER	2. GOVERNMENT ASSOCIATION NUMBER	3. RECIPIENT'S CATALOG NUMBER
CA17-2876		

4. TITLE AND SUBTITLE	5. REPORT DATE
A Unified Framework for Analyzing and Designing for Stationary Arterial Networks	May 17, 2017
	6. PERFORMING ORGANIZATION CODE

7. AUTHOR	8. PERFORMING ORGANIZATION REPORT NO.
Wenlong Jin and Shizhe Shen	N/A

9. PERFORMING ORGANIZATION NAME AND ADDRESS	10. WORK UNIT NUMBER
Institute of Transportation Studies Civil and Environmental Engineering University of California, Irvine Irvine, CA 92697-3600	11. CONTRACT OR GRANT NUMBER UTC Agreement 65A0528

12. SPONSORING AGENCY AND ADDRESS	13. TYPE OF REPORT AND PERIOD COVERED
Department of Transportation Division of DRISI-HQ MS #38 P. O. Box 942874 Sacramento, CA 94274-0001	Final Report 4/1/2015-1/31/2017 14. SPONSORING AGENCY CODE

15. SUPPLEMENTARY NOTES

16. ABSTRACT

This research aims to develop a unified theoretical and simulation framework for analyzing and designing signals for stationary arterial networks. Existing traffic flow models used in design and analysis of signal control strategies are either too simple to be realistic or too detailed to be efficient.

In this research we apply the link transmission model to formulate, analyze, and simulate traffic dynamics in a signalized arterial network. We first analytically derive approximate macroscopic fundamental diagrams for stationary traffic patterns with different network topologies, road conditions, driving behaviors, and signal settings. We then analyze congestion mitigation effects of different signal settings, including cycle lengths, green splits, and offsets. We further formulate and solve an optimization problem with the network flow-rate as performance measure to find optimal signal control parameters. We derived simple formulas for the optimal signal cycle length and offset under different traffic conditions to improve arterial network performance.

17. KEY WORDS	18. DISTRIBUTION STATEMENT
Arterial networks, MFD, cycle lengths, green splits, green ratio, offsets, queues.	No restrictions. This document is available to the public through the National Technical Information Service, Springfield, VA 22161

19. SECURITY CLASSIFICATION *(of this report)*	20. NUMBER OF PAGES	21. COST OF REPORT CHARGED
Unclassified	43	N/A

DISCLAIMER STATEMENT

This document is disseminated in the interest of information exchange. The contents of this report reflect the views of the authors who are responsible for the facts and accuracy of the data presented herein. The contents do not necessarily reflect the official views or policies of the State of California or the Federal Highway Administration. This publication does not constitute a standard, specification or regulation. This report does not constitute an endorsement by the Department of any product described herein.

For individuals with sensory disabilities, this document is available in alternate formats. For information, call (916) 654-8899, TTY 711, or write to California Department of Transportation, Division of Research, Innovation and System Information, MS-83, P.O. Box 942873, Sacramento, CA 94273-0001.

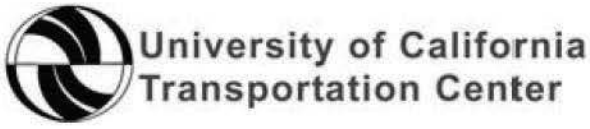
University of California Transportation Center

A Unified Framework for Analyzing and Designing for Stationary Arterial Networks

Wenlong Jin
Ph.D., Associate Professor
Civil and Environmental Engineering
Institute of Transportation Studies
University of California, Irvine

Shizhe Shen
Graduate Student
Civil and Environmental Engineering
Institute of Transportation Studies
University of California, Irvine

2017



ABSTRACT

This research aims to develop a unified theoretical and simulation framework for analyzing and designing signals for stationary arterial networks. Existing traffic flow models used in design and analysis of signal control strategies are either too simple to be realistic or too detailed to be efficient.

In this research we apply the link transmission model to formulate, analyze, and simulate traffic dynamics in a signalized arterial network. We first analytically derive approximate macroscopic fundamental diagrams for stationary traffic patterns with different network topologies, road conditions, driving behaviors, and signal settings. We then analyze congestion mitigation effects of different signal settings, including cycle lengths, green splits, and offsets. We further formulate and solve an optimization problem with the network flow-rate as performance measure to find optimal signal control parameters. We derived simple formulas for the optimal signal cycle length and offset under different traffic conditions to improve arterial network performance.

TABLE OF CONTENTS

LIST OF FIGURES

ACKNOWLEDGEMENTS

The work is partially supported by a University of California Transportation Center faculty grant.

Three research articles have been generated from this research fund:

1. Wen-Long Jin and Yifeng Yu, 2016. Performance analysis and signal design for a stationary signalized ring road.
    a. This paper has been submitted to Transportation Science for publication.
    b. This paper lays the foundation for performance analysis and signal design based on the stationary states in the link transmission model and approximate macroscopic fundamental diagrams.
2. Shizhe Shen and Wen-Long Jin, 2016. Analysis of a stationary two-signal ring road with offsets.
    a. This paper was submitted to the 2017 Transportation Research Board Annual Meeting. We are revising the paper for a future submission.
    b. This paper is an extension of the preceding paper to include the offsets in addition to the cycle lengths and green ratios.

Chapter 1

Literature Review

1.1 Background

Traffic signals have been widely deployed to resolve conflicts among various traffic streams and improve safety of drivers and pedestrians at busy urban intersections. But signalized intersections are also major network bottlenecks, inducing stop-and-go traffic patterns, travel delays, and vehicle emissions. Many efforts have been devoted to mitigating the congestion effects of isolated and coordinated intersections by optimally designing phase sequences, cycle lengths, green splits, offsets, and other parameters of traffic signals (Papageorgiou et al., 2005).

In existing signal analysis and design methods, performance measures include individual vehicles' delays or the level of service at signalized intersections (Webster, 1958; Lo, 1999, 2001; Li, 2010), the bandwidth of a set of coordinated signalized intersections (Roess et al., 2010), the whole traffic system's throughput (Li, 2010), a combination of delays and early arrival flows (He et al., 2010), or the mean of excess delays (Zhang et al., 2010). To evaluate these performances, underlying most of existing signal analysis and design methods are two types of traffic flow models: simple formulas for aggregate delay and bandwidth or traffic simulation models. The first type of methods are usually analytical, and the second type simulation-based. However, existing methods for traffic signal analysis and design are either too simplistic to be physically realistic or too complicated to be mathematically tractable, and there still lacks a systematic method (even) for mathematically analyzing and designing traffic signals for a large-scale arterial network, even for ''a one-way arterial" (Newell, 1989).

Since the introduction of the celebrated LWR model (Lighthill and Whitham, 1955; Richards, 1956), kinematic wave theory has been successfully applied to describe traffic dynamics on both freeways and arterial roads. It has been shown to be capable of capturing shock and rarefaction waves and the initiation, propagation, and dissipation of traffic queues, caused by various bottlenecks and interactions among vehicles. In particular, with the Cell Transmission Model (CTM) and other network kinematic wave theory (Daganzo, 1995; Lebacque, 1996), traffic dynamics in a road network can be systematically modeled. Compared with microscopic models, such network kinematic wave models are more suitable for studying traffic dynamics in large-scale arterial road networks. In addition, with the availability of various types of traffic data and the development of connected and automated vehicles, it is high time to develop effective and efficient methods for analyzing and designing signals for large-scale arterial networks.

1.2 Existing signal design and analysis methods

Traditionally performance measures used in the analysis and design of traffic signals include individual vehicles' delays or the level of service at signalized intersections (Webster, 1958; Lo, 1999, 2001; Li, 2010), the bandwidth of a set of coordinated signalized intersections (Roess et al., 2010), the whole traffic system's throughput (Li, 2010), a combination of delays and early arrival flows (He et al., 2010), or the mean of excess delays (Zhang et al., 2010). To evaluate these performances, underlying most of existing signal analysis and design methods are

two types of traffic flow models: simple formulas for aggregate delay and bandwidth or traffic simulation models. The first type of methods are usually analytical, and the second type simulation-based.

1.2.1 Analytical methods

For signal control on local arterials, it can be classified into two types according to the number of targeted intersections: for isolated intersections only and for coordinated intersections. In the literature, there have been a number of signal control strategies proposed for each category. In the following subsections, we provide a review of some prevailing strategies.

For isolated intersections

According to (Papageorgiou et al., 2003), fixed-time control strategies for a single intersection can be stage-based or phase-based. For stage-based strategies, the stage settings are fixed, and the proposed strategies are developed to find optimal splits and cycle lengths by minimizing the total delay or maximizing the total throughput at the intersection. To calculate vehicle's average delay, the delay formulation proposed by Webster (Webster, 1958) has been widely used in the literature, which can be formulated as follows:

$$d = \frac{\frac{1}{2}C\left(1-\frac{g}{C}\right)^2}{1-\frac{g}{C}X} + \frac{X^2}{2v(1-X)} - 0.65\left(\frac{C}{v^2}\right)^{\frac{1}{3}} X^{2+5\frac{g}{C}} \quad (1)$$

where

C =cycle length,

g =effective green time,

v =arrival flow-rate,

c =capacity of the intersection approach,

s =saturation flow-rate,

$X = \frac{v}{c} = \frac{vC}{sg}$, the degree of saturation.

The above delay calculation consists of three parts: the first part is the uniform delay; the second part is the random delay; and the third part is the empirical adjustment. In (Webster, 1958), optimal cycle lengths were obtained by minimizing the total delay at the intersection under given arrival flow-rates. In (Miller, 1963b), to obtain optimal settings of splits and cycle lengths, various arrival patterns were taken into account in the calculation of random delay. Since it is possible that an approach may have right of way in more than one stage within a cycle, SIGSET was proposed in (Allsop, 1971a; Allsop, 1971b) to take into account such a case, and Webster's delay formula was used in the delay estimation. By minimizing the total delay with the capacity, cycle lengths, and minimum green time constraints, optimal settings of cycle length and effective green time for each stage were obtained. Under similar constraints as those in

SIGSET, another program called SIGCAP was proposed in (Allsop, 1972; Allsop, 1976) to maximize the practical capacity at signalized intersections.

Different from stage-based control strategies, phase-based control strategies are developed to further consider optimal stage settings. One example can be found in (Improta and Cantarella, 1984), in which the constraint of fixed staging was released. Instead, incompatibility of traffic streams was introduced as a constraint in the optimization problem. By either minimizing the total delay or maximizing the intersection capacity, optimal settings of splits, cycle lengths, and stage settings can be obtained. In (Improta and Cantarella, 1984), the optimization problem was formulated as a binary-mixed-integer-linear- programming (BMILP) problem, and solutions were obtained using a branch-and-bound method.

Besides fixed-time control strategies, there also exist traffic-responsive control strategies that utilize the real-time loop detector data in the field. In (De la Breteque and Jezequel, 1979), examples such as the Vehicle Interval strategy, the Volume Density strategy, and Miller's algorithm were provided. In the Vehicle Interval strategy, each stage has a set of pre-specified minimum and maximum green times. If a vehicle is detected to cross the intersection, a critical interval (CI) will be used to extend the green time to allow that vehicle to pass. A similar control logit was used in the Volume Density strategy. But it further takes into account queue lengths and vehicles' waiting times during the red phases while deciding the switching time instants. In (Miller, 1963a), a computer program was used to determine whether to switch the signal immediately or to delay the switch for a user-defined time interval at every time step. Such a decision is made based on the evaluation of the time gain in postponing the switch. If the time gain is negative, the signal is switched immediately; otherwise, it remains unchanged for the next time step.

For coordinated intersections

If traffic signals in an arterial are close enough, the dissipation of vehicles is usually in platoons. Therefore, it is possible to synchronize the signals so as to allow vehicles travel along the arterial from the beginning to the end without stopping. In this case, bandwidth in one traffic direction is defined as the time difference between the first and the last vehicles that satisfy the above requirement. In the literature, there have been studies trying to maximize the bandwidths along the arterial. For example, with given cycle and speed ranges, MAXBAND was introduced in (Little, 1966) to obtain optimal offset settings so as to maximize the total bandwidths of a two-way arterial. The optimization problem was formulated as a mixed-integer-linear-programing (MILP) problem, and a branch-and-bound method was used to solve it. Later in (Gartner et al., 1991), MULTI-BAND was proposed to add new features such as determination of left-turn phases and different bandwidths among the links into the optimization problem. In (Robertson, 1969), TRANSYT (TRAffic Network StudY Tool) was proposed to obtain multi-directional green waves so as to minimize the total delay. Such a model consists of two parts: (i) with given network information such as road geometries, turning ratios at intersection, and demands, a platoon dispersion model is used to describe vehicle's progression inside a link; (ii) a "hill-climbing" method is used to solve the optimization problem. Performance Index (PI) is introduced to evaluate the improvements at each optimization step. The program will stop when a (local) minimum is found.

Due to the fact that demands and turning movements at intersections are changing as time elapses, traffic-responsive coordinated strategies have also been proposed in the literature. SCOOT (Split, Cycle and Offset Optimization Technique), which is a traffic-responsive version of TRANSYT, was proposed in (Hunt et al., 1982; Hunt et al., 1981). While keeping similar optimization structure as in TRANSYT, SCOOT works in a real-time fashion: it utilizes real-time measurements of flows and occupancies from vehicle loop detectors to predict delay and stops; the signal optimizer works in real time, and new signal settings are implemented directly on the street. Besides SCOOT, another algorithm called OPAC (Optimization Policies for Adaptive Control), which is a model-based traffic-responsive strategy, was proposed in (Gartner, 1983). In OPAC, splits, offsets, and cycles are not explicitly considered. A rolling horizon approach is used for real-time applications: at time t, the optimization method calculates an optimal switching scheme for the time interval $[t-h, t+H-h]$ $(H>h)$ based on the data in the time interval $[t-h, t]$ and applies it to the time interval $[t, t+h]$; then the optimization time horizon moves to the next step, $t+h$. Note that since OPAC employs complete enumeration in the optimization, it is not real-time feasible for multiple intersections (Papageorgiou et al., 2003).

1.2.2 Simulation-based methods

In the second type of methods, various traffic flow models are used to simulate traffic dynamics, and optimal control problems are formulated to find best signal settings simultaneously subject to given demand patterns (Gazis and Potts, 1963; Gazis, 1964; D'ans and Gazis, 1976; Improta and Cantarella, 1984; Papageorgiou, 1995; Park et al., 1999; Chang and Lin, 2000; Chang and Sun, 2004). For example, the Cell Transmission Model (CTM) (Daganzo, 1995; Lebacque, 1996) has been incorporated into numerous formulations of the traffic signal problem (Lo, 1999; Li, 2010). Such models can are more realistic in traffic dynamics, but are too detailed to be amenable for mathematical analysis. In addition, these formulations usually end up as mixed-integer linear programing (MILP) problems, which can be effectively solved only for small networks and become numerically formidable for large-scale arterial networks.

1.3 Network traffic flow models

Since the introduction of the celebrated LWR model (Lighthill and Whitham, 1955; Richards, 1956), kinematic wave theory has been successfully applied to describe traffic dynamics on both freeways and arterial roads. It has been shown to be capable of capturing shock and rarefaction waves and the initiation, propagation, and dissipation of traffic queues, caused by various bottlenecks and interactions among vehicles. In particular, with the Cell Transmission Model (CTM) and other network kinematic wave theory (Daganzo, 1995; Lebacque,1996), traffic dynamics in a road network can be systematically modeled.

1.3.1 Cell transmission model

In kinematic wave theories, traffic flow is considered as a continuous media. Three location-and-time dependent variables, speed $v(x,t)$, density $k(x,t)$, and flow-rate $q(x,t)$, are used to describe the traffic flow characteristics at point x and time t. For a road section without any entrances and exits, flow conservation is hold, which can be written as

$$\frac{\partial k(x,t)}{\partial t}+\frac{\partial q(x,t)}{\partial x}=0 \tag{2}$$

In traffic flow, it is well known that there exists a fundamental relation between flow-rate (or speed) and density (Greenshields, 1935), i.e., $q(x,t) = Q(k(x,t))$ or $v(x,t) = V(k(x,t))$. Such a relation is known as the traffic flow fundamental diagram and can be validated using the vehicle loop detector data from freeways. Generally speaking, $Q(k)$ is a concave function and attains its capacity C at $k = k_c$, where k_c is the critical density. Introducing the fundamental diagram into Equation (2), the Lighthill-Whitham-Riahcrds (LWR) model (Lighthill and Whitham, 1955; Richards, 1956) is obtained.

$$\frac{\partial k(x,t)}{\partial t} + \frac{\partial Q(k(x,t))}{\partial x} = 0 \tag{3}$$

Equation (3) is a hyperbolic conservation law and is difficult to solve analytically under general initial and boundary conditions. Therefore, in (Daganzo, 1994; Daganzo, 1995), the cell transmission model (CTM) was introduced to numerically solve Equation (3). According to the Godunov method (Godunov, 1959), a link is equally divided into N cells with a length of Δx, and the whole time interval is partitioned into J time steps with an interval of Δt. In Figure 1.1, cell representation inside a regular link is provided. Then the discrete version of Equation (3) can be written as

$$\frac{k_i^{j+1} - k_i^j}{\Delta t} - \frac{f_{i-1}^j - f_i^j}{\Delta x} = 0 \tag{4}$$

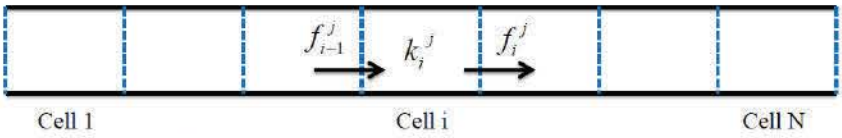


Figure 1.1 Cell representation inside a regular link

where k_i^j and k_i^{j+1} are the densities of cell i at time steps j and j+1, respectively, and f_{i-1}^j and f_i^j are the upstream and downstream boundary fluxes of cell i at time step j, respectively. Here, the choice of $\frac{\Delta t}{\Delta x}$ should follow the CFL condition (Courant et al., 1928), which requires a vehicle cannot travel across one cell at one time step. That is, $\frac{v_f \Delta t}{\Delta x} \leq 1$, where v_f is the free-flow speed of that link. Given densities and fluxes at time step j, the density at time step j+1 can be updated using the following equation:

$$k_i^{j+1} = k_i^j + \frac{\Delta t}{\Delta x}(f_{i-1}^j - f_i^j) \tag{2}$$

To obtain the fluxes crossing the cell boundaries, the definitions of demand D and supply S (Daganzo, 1995; Lebacque, 1996) are introduced and can be calculated as

$$D = Q(\min\{k, k_c\}) \tag{3}$$

$$S = Q(\max\{k, k_c\}) \tag{4}$$

Therefore, the flux through a cell boundary can be calculated by taking the minimum of the upstream cell's demand and the downstream cell's supply, which is

$$f_{i-1}^j = \min\{D_{i-1}^j, S_i^j\} \tag{5}$$

where D_{i-1}^j is the demand of cell i-1, and S_i^j is the supply of cell i at time step j. For freeway networks, network junction models such as those in (Daganzo, 1995; Lebacque, 1996; Jin et al.,

2009; Jin, 2010, Jin, 2012a; Jin, 2014a) are needed to model the traffic dynamics at various types of junctions. For urban networks, besides the network junction models, signal control should be considered in order to manage the conflicting traffic movements at the intersections.

1.3.2 Link transmission model

Traditionally, at a point (a, x_a) inside link a, the density $k_a(x_a, t)$, the speed $v_a(x_a, t)$, the flow-rate $q_a(x_a, t)$ are used as variables to describe the evolution of traffic flow. The flow conservation, $\frac{\partial k_a}{\partial t} + \frac{\partial q_a}{\partial x} = 0$, together with the traffic flow fundamental diagram, $q_a = Q_a(k_a)$, forms the LWR model. However, we also can use another type of state variable, which is the cumulative flow, $A_a(x_a, t)$, and is known as the Moskovitz function (Moskowitz, 1965). Since we have $k_a = -\frac{\partial A_a}{\partial x}$, and $q_a = \frac{\partial A_a}{\partial t}$, the flow conservation is automatically satisfied if we have $\frac{\partial^2 A_a}{\partial x \partial t} = \frac{\partial^2 A_a}{\partial t \partial x}$. Therefore, to solve the LWR model in Equation 3 is equivalent to solve the following Hamilton-Jacobi equation

$$\frac{\partial A_a}{\partial t} - Q_a\left(-\frac{\partial A_a}{\partial x}\right) = 0 \tag{6}$$

with the Hamiltonian $H\left(\frac{\partial A_a}{\partial x}\right) = -Q_a\left(-\frac{\partial A_a}{\partial x}\right)$. Besides CTM, another new solution to the LWR model, which is called the Link Transmission Model (LTM), was proposed in recent studies. The discrete version can be found in (Yperman, 2007), while its continuous version can be referred to (Jin, 2014b).

Here, the triangular traffic flow fundamental diagram (Haberman, 1977), $q = Q(k) = \min\{v_f k, w(k_j - k)\}$, is used. The initial cumulative flow at $x_a \in [0, L_a]$ is denoted as $N_a(x_a)$. The cumulative in-flow and the in-flux at the upstream boundary are denoted as $F_a(t)$ and $f_a(t)$, respectively. The cumulative out-flow and the out-flux at the downstream boundary are denoted as $G_a(t)$ and $g_a(t)$, respectively. To describe the congestion pattern inside a link, two variables, the link queue size $\alpha_a(t)$ and the link vacancy size $\beta_a(t)$ are used and can be calculated as follows:

$$\alpha_a(t) = \begin{cases} N_a\left(L_a - v_{a,f}t\right) - G_a(t) & t \le \dfrac{L_a}{v_{a,f}} \\ F_a\left(t - \dfrac{L_a}{v_{a,f}}\right) - G_a(t) & t > \dfrac{L_a}{v_{a,f}} \end{cases}$$

$$\beta_a(t) = \begin{cases} N_a(w_a t) + k_{a,j} w_a t - F_a(t) & t \le \dfrac{L_a}{w_a} \\ G_a\left(t - \dfrac{L_a}{w_a}\right) + k_{a,j} L_a - F_a(t) & t > \dfrac{L_a}{w_a} \end{cases} \tag{7}$$

Initially, we have $\alpha(0) = 0$ and $\beta(0) = 0$. In the LTM, either cumulative flows or link queue and vacancy sizes can be used as stable variables to describe the evolution of traffic dynamics. If the cumulative flow, i.e., $F_a(t)$ and $G_a(t)$, are used, we have the following evolution equations:

$$\frac{d}{dt}F_a(t) = f_a(t) \tag{8}$$

$$\frac{d}{dt}G_a(t) = g_a(t)$$

If the link queue and vacancy sizes, i.e., $\alpha_a(t)$ and $\beta_a(t)$, are used, we have the following evolution equations:

$$\begin{aligned}\frac{d\alpha_a(t)}{dt} &= \begin{cases} k_a\left(L_a - v_{a,f}t, 0\right)v_{a,f} - g_a(t) & t \le \dfrac{L_a}{v_{a,f}} \\ f_a\left(t - \dfrac{L_a}{v_{a,f}}\right) - g_a(t) & t > \dfrac{L_a}{v_{a,f}} \end{cases} \\ \frac{d\beta_a(t)}{dt} &= \begin{cases} -k_a(w_a t, 0)w_a + k_{a,j}w_a - f_a(t) & t \le \dfrac{L_a}{w_a} \\ g_a\left(t - \dfrac{L_a}{w_a}\right) - f_a(t) & t > \dfrac{L_a}{w_a} \end{cases} \end{aligned} \tag{9}$$

To update the evolution functions in Equations 11 and 12, the in-fluxes and out-fluxes are needed to be calculated/updated first. Here we define an indicator function $H(y)$ for $y \ge 0$, which is formulated as follows:

$$H(y) = \lim_{\Delta t \to 0^+} \frac{y}{\Delta t} = \begin{cases} 0 & y = 0 \\ +\infty & y > 0 \end{cases} \tag{10}$$

Then the link demand $d_a(t)$ and link supply $s_a(t)$ are defined as

$$\begin{aligned} d_a(t) &= \begin{cases} \min\{k_a\left(L_a - v_{a,f}t, 0\right)v_{a,f} + H\left(\alpha_a(t)\right), C_a\} & t \le \dfrac{L_a}{v_{a,f}} \\ \min\{f_a\left(t - \dfrac{L_a}{v_{a,f}}\right) + H\left(\alpha_a(t)\right), C_a\} & t > \dfrac{L_a}{v_{a,f}} \end{cases} \\ s_a(t) &= \begin{cases} \min\{k_{a,j}w_a - k_a(w_a t, 0)w_a + H\left(\beta_a(t)\right), C_a\} & t \le \dfrac{L_a}{w_a} \\ \min\{g_a\left(t - \dfrac{L_a}{w_a}\right) + H(\beta_a(t)), C_a\} & t > \dfrac{L_a}{w_a} \end{cases} \end{aligned} \tag{11}$$

At a junction j, macroscopic junction models are used to determine the in-fluxes and out-fluxes from the upstream link demands, downstream link supplies, and turning proportions, which in general can be written as follows:

$$\left(\boldsymbol{g}_j(t), \boldsymbol{f}_j(t)\right) = \boldsymbol{F}(\boldsymbol{d}_j(t), \boldsymbol{s}_j(t), \boldsymbol{\xi}_j(t))) \tag{12}$$

Here, $\boldsymbol{g}_j(t)$ is the set of in-fluxes, while $\boldsymbol{f}_j(t)$ is the set of out-fluxes. $\boldsymbol{d}_j(t)$ is the set of upstream link demands, while $\boldsymbol{s}_j(t)$ is the set of downstream link supplies. $\boldsymbol{\xi}_j(t)$ is a matrix that contains turning proportions from the upstream links to the downstream ones. As shown in (Jin, 2014b), non-invariant junction models cannot be used in the LTM, which may yield no solution to the traffic statics problem under certain traffic conditions. A set of invariant junction models can be found in (Jin et al., 2009; Jin, 2010, Jin, 2012a; Jin, 2014a).

With Equations 14 and 15, the in-fluxes and out-fluxes can be calculated and then be introduced into Equations 11 or 12 to update the state variables. But note that, as shown in Equation 14, link demands and supplies depend on the historical data, and therefore, Equations

11 and 12 are systems of ordinary differential equations (ODEs) with delays. Once the cumulative in-flows $F_a(t)$ and the cumulative out-flows $G_a(t)$ are obtained, traffic states inside link a can be obtained. More details can be referred to (Jin, 2014b).

1.4 Stationary states and macroscopic fundamental diagram

1.4.1 Stationary states

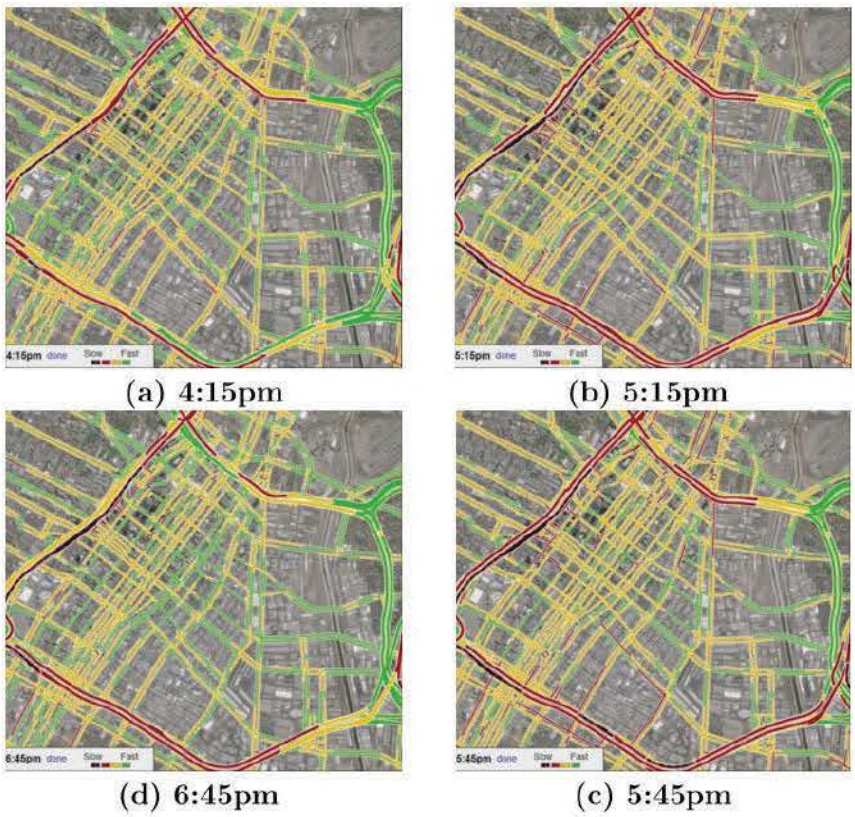


Figure 1.2 Typical traffic conditions in downtown Los Angeles during Monday's afternoon peak periods (Source: Google Maps Traffic)

Figure 1.2 demonstrates typical traffic conditions in downtown Los Angeles during Monday's afternoon peak periods. Comparing the left figures, (a) and (d), and the right figures, (b) and (c), we can see that traffic conditions are relatively stationary on the arterial network from 5:15 to 5:45; that is, congested links, queue lengths, and bottleneck locations remain almost the same during the period. Such stationary patterns have also been observed in freeway networks (Jin, 2015b), and we believe that stationary traffic patterns in arterial networks are also caused by that ''the traffic demand and origin-destination desires are relatively constant over the time period" (Wattleworth, 1967). The existence of such stationary traffic patterns is also consistent with our daily driving experience: we experience almost the same level of congestion on the same link at the same time every day.

In many studies on analysis, control, management, planning, and design of road networks during peak periods, traffic patterns have been assumed to be stationary (Merchant and

Nemhauser, 1978b; Yang and Yagar, 1995; Yang and Lam, 1996): in (Beckmann et al., 1956), the static traffic assignment problem was formulated to determine the aggregate route choice behaviors of vehicles; in (Godfrey, 1969), it was postulated that a network-wide macroscopic fundamental diagram (MFD) exists in such stationary, or steady, states, and this has been verified by observations (Geroliminis and Daganzo, 2008); in (Wattleworth, 1967), the local and global control problem of a freeway system was solved with linear programming methods; in (Potts and Oliver, 1972), network flow conservation problems are solved; and in (Payne and Thompson, 1974), the integrated traffic assignment and ramp metering problem was solved for stationary traffic patterns.

1.4.2 Macroscopic fundamental diagrams

In stationary urban road networks, it was postulated that there exists a relation between network-average flow and density in (Godfrey, 1969). Such a relation is called the macroscopic fundamental diagram (MFD) and has been shown to be unique in homogeneous networks, but not in non-homogeneous ones with simulations and observations (Ardekani and Herman, 1987; Mahmassani et al., 1987; Olszewski et al., 1995; Geroliminis and Daganzo, 2008; Buisson and Ladier, 2009; Cassidy et al., 2011; Geroliminis and Boyaci, 2012). In (Daganzo, 2007; Geroliminis et al., 2013), regional demand control strategies were developed based on MFD.

As a system-wide characteristic, MFD emerges from network traffic flow patterns, which are determined by network topology, signal and other control measures, and drivers' choices in destinations, modes, departure times, routes, and speeds. Some efforts have been devoted to deriving MFD in simple signalized networks from various traffic flow models. In (Gartner and Wagner, 2004), with cellular automaton simulations for traffic on a ring road, which has multiple identical signals, the relationship between flow-rate, density, and offset was obtained in relatively stationary states after a long time (2000 seconds), and it was found that offsets can have drastic impacts on the overall throughputs and, therefore, travel times on an arterial road. In (Daganzo and Geroliminis, 2008), a variational method was proposed to compute the approximate MFD for the relationship between long-time average flow-rates and densities in a ring road with multiple signals, but the study did not provide definitions of stationary states, or explicit formulas for the MFD, or optimal signal settings. In (Daganzo et al., 2011), the MFD in a double-ring network with turning movements was studied with heuristic double-bin approximations and cellular automaton simulations. In (Jin et al., 2013), steady or stationary states in a signalized double-ring network were defined as asymptotically periodic traffic states within the framework of a network kinematic wave theory, and impacts of signal settings and turning movements on MFD in stationary states were simulated with the Cell Transmission Model (CTM) (Daganzo, 1995). In (Muralidharan et al., 2015), the existence and stability of periodic solutions in a signalized road network with fixed-time control were studied with the point queue model, but queue spillbacks were not captured in this study. In (Gan et al., 2015), stationary states in a double-ring network with turning movements, which is equivalent to a homogeneous, symmetric road network with turning movements, were defined and analyzed with a Poincare map: it was shown that stationary states exist, but may be unstable, and there exist multiple average flow-rates at one density. The study was enabled by the link queue model, which is an ordinary differential equation approximation of the kinematic wave model, uses the link fundamental diagram, and captures the queue spillback phenomenon, but allows for instantaneous vehicle and wave propagation.

1.5 Summary

In this chapter, we provided the literature review for arterial signal design and analysis methods as well as network traffic flow theory. First, we reviewed existing signal design and analysis methods, including both analytical and simulation-based methods. Second, we reviewed network kinematic wave theory, including the cell transmission model (CTM), and the link transmission model (LTM). Third, we reviewed studies related to stationary states and macroscopic fundamental diagram.

In the following we discuss limitations and improvements of the existing models and methods. (i) Existing methods for traffic signal analysis and design are either too simplistic to be physically realistic or too complicated to be mathematically tractable, and there still lacks a systematic method (even) for mathematically analyzing and designing traffic signals for a large-scale arterial network, even for ‘‘a one-way arterial”; (ii) Traffic dynamics in a signalized network are still too complicated to be mathematically tractable; (iii) there exists no systematic studies on stationary states in signalized networks, and there exists no explicit closed-form formula of MFD; in particular, signal settings have not been included as parameters in MFD, even for simple networks.

In this project, based on network kinematic wave theory, we will first rigorously define and discuss properties of stationary states in signalized networks, which are more complicated than those in freeway networks without signals. Then the MFD is formally defined as the relationship between network flow-rate and density, and we will derive closed-form formulas for macroscopic fundamental diagrams with signal settings as explicit parameters for different networks. Further we will provide analytical formulas for optimal signal settings for different densities in different road networks. Finally we will develop numerical methods for finding stationary states, MFD, and optimal signal settings in large-scale arterial road networks.

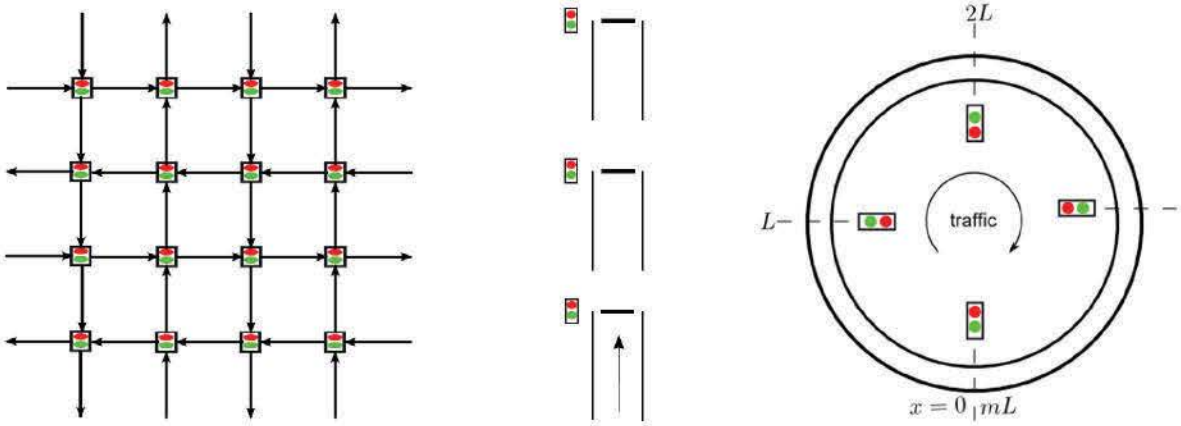
This chapter provides researchers and engineers a systematic view on the state-of-the-art traffic signal design and analysis as well as dynamic and stationary network traffic flow models, and can help them to appreciate the need and potential for developing more efficient and effective control and management schemes in urban networks in the future.

Chapter 2

Problem formulation

2.1 Network and signal settings

We consider a road network, as illustrated in Figure 2.1(a), in which turning movements can be ignored. Thus we effectively consider an infinite street as illustrated in Figure 2.2(b). We label the roads consecutively by integer numbers: $\cdots, -2, -1, 0, 1, 2, \cdots$. For road j, its length is L_j. Here we assume that the roads are periodic with a period of m, which is a natural number. That is, we assume that $L_{m+j} = L_j$. We denote the location on the street by x, which increases in the traffic condition; for simplicity, we assume that road 1 starts at $x = 0$ and ends at $x = L_1$.



(a) (b) (c)

Figure 2.1 Illustration of (a) A road network; (b) An infinite street with periodic roads; (c) A signalized ring road

For a fixed, periodic traffic signal at the downstream boundary of road j, we denote its cycle length by T_j, green ratio by π_j, and (cumulative) offset by $\delta_j = \Delta_j - \Delta_{j-1}$, where Δ_j is the cumulative offset at signal j. Without loss of generality, $\Delta_0 = 0$.i.e., signal 0's green interval starts at $t = 0$.

The binary signal can be represented by

$$\psi_j(t) = H\big(\pi_j T_j - (t - \Delta_j) \bmod T_j\big) = \begin{cases} 1, & t \in iT_j + \Delta_j + [0, \pi_j T_j) \\ 0, & t \in iT_j + \Delta_j + [\pi_j T_j, T_j) \end{cases} \tag{2.1}$$

where $a\ mod\ b$ returns the remainder of a divided by b, and $H(\cdot)$ is the Heaviside function. Here a signal cycle is divided into green and red intervals, and we omit the yellow intervals. It is reasonable to assume that the signals are also m-periodical; i.e., $T_{j+m} = T_j, \pi_{j+m} = \pi_j, \delta_{j+m} = \delta_j$.

Here we consider homogeneous roads: $L_j = L$ with $T_j = T$, $\pi_j = \pi$, $\delta_j = \frac{n}{m}T$, where n and m are co-prime numbers. Hence $\Delta_j = j\frac{n}{m}T$. There are four parameters for the signal settings: the cycle length, T, the green ratio, π, and the two numbers for the offset, (m, n). In this case, the infinite street is equivalent to a ring road with m signals, as illustrated in Figure 2.1(c). Therefore the signal function can be written as

$$\psi_j(t) = H\left(\pi T - (t - j\frac{n}{m}T)\ mod\ T\right) = \begin{cases} 1, & t \in iT + j\frac{n}{m}T + [0, \pi T)\ ; \\ 0, & t \in iT + j\frac{n}{m}T + [\pi T, T). \end{cases}$$

2.2 The link transmission model

We denote the cumulative flow at signal j by $G_j(t)$, with $G_0(0) = 0$. Initially each link's density is constant at k_0. It is reasonable to assume that traffic dynamics are also m-periodical; i.e., $g_{j+m}(t) = g_j(t)$, which is the flow-rate. Thus $G_{j+m}(t) = G_j(t) - mk_0$.

We assume that all links have the same triangular fundamental diagram:

$$q = Q(k) = \min\{uk, w(\kappa - k)\},$$

where u is the free-flow speed, -w the information propagation speed in congested traffic, and κ the jam density. In addition, we denote the critical density by κ_c, and the capacity by C:

$$\kappa_c = \frac{w}{u + w}\kappa, \quad C = u\kappa_c.$$

Then from the link transmission model, we have the following m equations ($j = 1, \cdots, m$) after a long time t:

$$G_j(t) = \begin{cases} \min\left\{G_{j-1}\left(t - \frac{L}{u}\right), G_{j+1}\left(t - \frac{L}{w}\right) + \kappa L, G_j\left(iT + \Delta_j\right) + \left(t - iT - \Delta_j\right)C\right\}, & t \in iT + \Delta_j + (0, \pi T] \\ G_j\left(iT + \Delta_j + \pi T\right), & t \in iT + \Delta_j + (\pi T, T] \end{cases} \quad (2.2)$$

Here $G_0(t) = G_m(t) + mk_0, G_{m+1}(t) = G_1(t) - mk_0$. Equation (2.2) shows that the cumulative flows inside the green intervals are determined by three waves, and those inside the red intervals are constant, since no vehicles can cross the intersection.

In Equation (2.2), $\frac{L}{u}$ equals vehicles' traverse time on a link, and $\frac{L}{w}$ equals the congestion building up time. We denote them by

$$\frac{L}{u} = \theta_1 T = (l_1 + \alpha_1)T, \tag{2.3}$$

$$\frac{L}{w} = \theta_2 T = (l_2 + \alpha_2)T, \tag{2.4}$$

where l_1 and l_2 are integers, and $\alpha_1, \alpha_2 \in [0,1)$.

Hereafter we normalize time with respect to T, and Equation (2.2) is written as

$$G_j(t) = \begin{cases} \min\left\{G_{j-1}(t-\theta_1), G_{j+1}(t-\theta_2) + \kappa L, G_j\left(i + j\frac{n}{m}\right) + \left(t - i - j\frac{n}{m}\right)C\right\}, & t \in i + j\frac{n}{m} + (0, \pi] \\ G_j\left(i + j\frac{n}{m} + \pi\right), & t \in i + j\frac{n}{m} + (\pi, 1] \end{cases} \tag{2.5}$$

Chapter 3

Approximate macroscopic fundamental diagrams

3.1 Stationary states

We consider the simple periodic states with a period of T as stationary states in the signalized ring road, where $g(t+T) = g(t)$. Thus in stationary states we have

$$G(t+T) = G(t) + q_0 T, \tag{3.1}$$

where $q_0 \in [0, \pi C]$ is the average flow-rate during a cycle. According to Equation (2.5), the cumulative flow at the end of the green interval equals that at the start plus $q_0 T$; i.e.,

$$G_j\left(i + j\frac{n}{m} + \pi\right) = G_j\left(i + j\frac{n}{m}\right) + q_0 T = \min\left\{G_{j-1}\left(i + j\frac{n}{m} + \pi - \theta_1\right), G_{j+1}\left(i + j\frac{n}{m} + \pi - \theta_2\right) + \kappa L, G_j\left(i + j\frac{n}{m}\right) + \pi TC\right\}. \tag{3.2}$$

Clearly $q_0 \leq \pi C$, which is achieved when the boundary flow-rate equals the capacity during the whole green interval.

A macroscopic fundamental diagram gives the following relationship:

$$q_0 = Q(k_0; L, T, \pi, m, n), \tag{3.3}$$

which is a function of the density as well as the road and signal settings. We will derive approximate macroscopic fundamental diagrams from equation (3.2).

3.2 Approximate macroscopic fundamental diagrams for $(m, n) = (1, 0)$

When $(m, n) = (1,0)$, $\Delta_j = 0$, and all signals are synchronized in a simultaneous progression. Denote $G_1(t) = G(t)$. Then $G_j(t) = G(t) - (j-1)k_0 T$. As illustrated in Figure 3.1, Equation (3.2) can be written as

$$\begin{aligned} G(i+\pi) &= G(i) + q_0\, T \\ &= \min\{G(i+\pi-\theta_1) + k_0 L, G(i+\pi-\theta_2) + (\kappa - k_0)L, G(i) + \pi TC\}. \end{aligned} \tag{3.4}$$

We observe that $q_0 = \pi C$ if and only if $g(t) = C \cdot H(\pi - \phi(t))$ and

$$G(t) = G(0) + \Psi(t) TC,$$

where $\Psi(t) = \Phi(t)\,\pi + \min\{\phi(t), \pi\}$, $\Phi(t) = \lfloor t \rfloor$ is the floor function, and $\phi(t) = t - \Phi(t) = t \bmod 1$ is the remainder. Note that $\Psi(t)$ is the integral of the homogeneous signal function

$\beta(t)$. We have that $\pi t \le \Psi(t) \le \pi(t+1-\pi)$, where the left equal sign holds when $\phi(t) = 0$ and the right equal sign holds when $\phi(t) = \pi$.

Thus Equation (3.4) can be written as

$$\pi CT = \min\{\Psi(\pi-\theta_1)TC + k_0 L, \Psi(\pi-\theta_2)\,TC + (\kappa - k_0)L, \pi TC\},$$

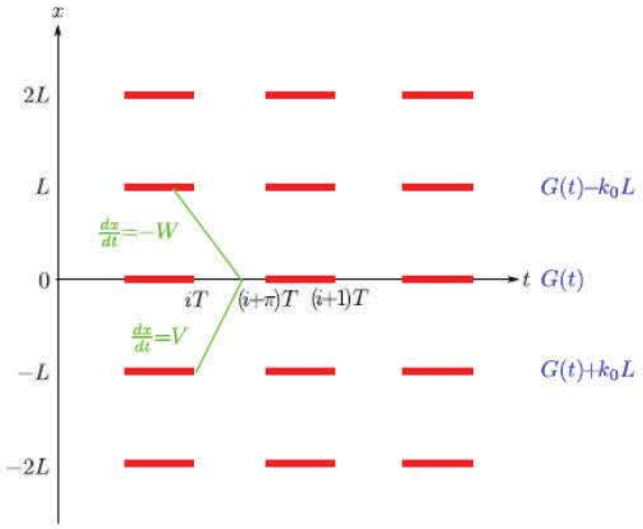


Figure 3.1 Illustration of the signal settings in the case $(m,n) = (1,0)$

which leads to $\Psi(\pi-\theta_1)TC + k_0 L \ge \pi TC$ and $\Psi(\pi-\theta_2)TC + (\kappa-k_0)L \ge \pi TC$. Therefore, $q_0 = \pi C$ if and only if

$$k_1 \le k_0 \le k_2,$$

where

$$k_1 = \frac{TC}{L}[\pi - \Psi(\pi-\theta_1)] = \frac{l_1 + \min\left\{\frac{\alpha_1}{\pi}, 1\right\}}{\theta_1}\pi\kappa_c,$$

$$k_2 = \kappa - \frac{TC}{L}[\pi - \Psi(\pi-\theta_2)] = \kappa - \frac{l_2 + \min\left\{\frac{\alpha_2}{\pi}, 1\right\}}{\theta_2}\pi\frac{C}{w}.$$

Further we have the following lemma.

Lemma 3.1 k_1 and k_2 satisfy $\pi\kappa_c \le k_1 \le \kappa_c \le k_2 \le \kappa - \pi\frac{C}{w}$.

Then we can approximate the macroscopic fundamental diagram by a trapezoidal function.

Theorem 3.2 The macroscopic fundamental diagram can be approximated by the following trapezoidal function

$$q_0 = Q(k_0; L, T, \pi, 1, 1) \approx \min\left\{\frac{k_0}{k_1}, 1, \frac{\kappa - k_0}{\kappa - k_2}\right\} \pi C. \tag{3.5}$$

Note that the macroscopic fundamental diagram is accurate when $k_1 \le k_0 \le k_2$. The approximate macroscopic fundamental diagram is shown in Figure 3.2.

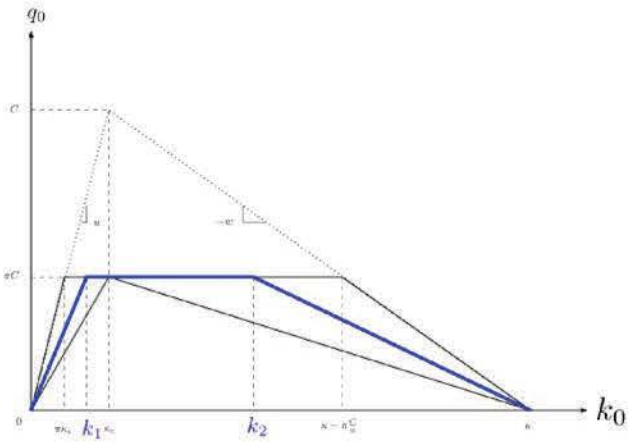


Figure 3.2 An approximate trapezoidal macroscopic fundamental diagram for $(m, n) = (1,0)$

3.3 Approximate macroscopic fundamental diagrams for $(m, n) = (2, 1)$

When $(m, n) = (2,1)$, $\Delta_j = j\frac{1}{2}T$. $G_0(t) = G_2(t) + 2k_0L, G_3(t) = G_1(t) - 2k_0L$, and Equation (3.2) can be written as

$$G_1\left(i + \tfrac{1}{2} + \pi\right) = G_1\left(i + \tfrac{1}{2}\right) + q_0T = \min\left\{G_2\left(i + \tfrac{1}{2} + \pi - \theta_1\right) + 2k_0L, G_2\left(i + \tfrac{1}{2} + \pi - \theta_2\right) + \kappa L, G_1\left(i + \tfrac{1}{2}\right) + \pi TC\right\}. \tag{3.6}$$

$$G_2(i + 1 + \pi) = G_2(i + 1) + q_0T = \min\{G_1(i + 1 + \pi - \theta_1), G_1(i + 1 + \pi - \theta_2) + (\kappa - 2k_0)L, G_2(i + 1) + \pi TC\}. \tag{3.7}$$

We observe that $q_0 = \pi C$ if and only if

$$G_2\left(i + \tfrac{1}{2} + \pi - \theta_1\right) + 2k_0L \ge G_1\left(i + \tfrac{1}{2}\right) + \pi TC, \tag{3.8}$$

$$G_1(i + 1 + \pi - \theta_1) \ge G_2(i + 1) + \pi TC, \tag{3.9}$$

$$G_2\left(i + \tfrac{1}{2} + \pi - \theta_2\right) + \kappa L \ge G_1\left(i + \tfrac{1}{2}\right) + \pi TC, \tag{3.10}$$

$$G_1(i + 1 + \pi - \theta_2) + (\kappa - 2k_0)L \ge G_2(i + 1) + \pi TC. \tag{3.11}$$

Also note that $G_2(i+1+t) = G_2(i+1) + \Psi(t)TC$, and $G_1\left(i+\frac{1}{2}+t\right) = G_1\left(i+\frac{1}{2}\right) + \Psi(t)TC$. Thus combining Equations (3.8) and (3.9), we obtain

$$k_0 \geq k_1 \equiv \frac{TC}{L}\left[\frac{\pi}{2} - \Psi\left(\pi - \frac{1}{2} - \theta_1\right)\right]. \tag{3.12}$$

Similarly, combining Equations (3.10) and (3.11), we obtain

$$k_0 \leq k_2 \equiv \kappa - \frac{TC}{L}\left[\frac{\pi}{2} - \Psi\left(\pi - \frac{1}{2} - \theta_2\right)\right]. \tag{3.13}$$

When $k_1 \leq k_2$, the approximate macroscopic fundamental diagram can be written as in Equation (3.5):

$$q_0 = Q(k_0; L, T, \pi, 2, 1) \approx \min\left\{\frac{k_0}{k_1}, 1, \frac{\kappa - k_0}{\kappa - k_2}\right\} \pi C. \tag{3.14}$$

Note that, however, k_1 may not be smaller than k_2. When $k_1 > k_2$, the maximum flow-rate of $q_0 < \pi C$. More discussions are provided in Chapter 4. Also note that we cannot assume that the number of vehicles on each street is always k_0. Further note that conditions in (3.12) and (3.13) are both necessary and sufficient conditions, but the proof is omitted here. Intuitively this is due to the translational symmetry and unique information propagation direction.

3.4 Approximate macroscopic fundamental diagrams for general (m, n)

For more general networks with different (m, n), if the flow-rate equals the capacity during the green intervals at all intersections. Then we have

$$G_{j-1}\left(i + (j-1)\frac{n}{m} + \pi + \frac{n}{m} - \theta_1\right) = G_{j-1}\left(i + (j-1)\frac{n}{m}\right) + \Psi\left(\pi + \frac{n}{m} - \theta_1\right)TC,$$
$$G_{j+1}\left(i + (j+1)\frac{n}{m} + \pi - \frac{n}{m} - \theta_2\right) = G_{j+1}\left(i + (j+1)\frac{n}{m}\right) + \Psi\left(\pi - \frac{n}{m} - \theta_2\right)TC.$$

Thus Equation (3.2) can be written as $(q_0 = \pi C)$

$$G_j\left(i + j\frac{n}{m}\right) + \pi CT$$
$$= \min\left\{G_{j-1}\left(i + (j-1)\frac{n}{m}\right) + \Psi\left(\pi + \frac{n}{m} - \theta_1\right)TC, G_{j+1}\left(i + (j+1)\frac{n}{m}\right) + \kappa L\right.$$
$$\left. + \Psi\left(\pi - \frac{n}{m} - \theta_2\right)TC, G_j\left(i + j\frac{n}{m}\right) + \pi TC\right\}.$$

Thus we have the following $2m$ inequalities: $(j = 1, \cdots, m)$

$$G_{j-1}\left(i + (j-1)\frac{n}{m}\right) + \Psi\left(\pi + \frac{n}{m} - \theta_1\right)TC \geq G_j\left(i + j\frac{n}{m}\right) + \pi TC, \tag{3.15}$$
$$G_{j+1}\left(i + (j+1)\frac{n}{m}\right) + \kappa L + \Psi\left(\pi - \frac{n}{m} - \theta_2\right)TC \geq G_j\left(i + j\frac{n}{m}\right) + \pi TC. \tag{3.16}$$

Combining the m inequalities in (3.15), we have

$$G_0(i) + m\Psi\left(\pi + \frac{n}{m} - \theta_1\right)TC \geq G_m(i+n) + m\pi TC.$$

Since $G_m(i+n) = G_m(i) + n\pi TC = G_0(i) - mk_0L$, we have

$$k_0 \geq k_1 \equiv \frac{TC}{L}\left[\left(1 + \frac{n}{m}\right)\pi - \Psi\left(\pi + \frac{n}{m} - \theta_1\right)\right] = \frac{C}{u}\frac{\left(1+\frac{n}{m}\right)\pi - \Psi\left(\pi + \frac{n}{m} - \theta_1\right)}{\theta_1}. \tag{3.17}$$

Combining the m inequalities in (3.16), we have

$$G_{m+1}\left(i+(m+1)\frac{n}{m}\right)+m\kappa L+m\Psi\left(\pi-\frac{n}{m}-\theta_2\right)TC \geq G_1\left(i+\frac{n}{m}\right)+m\pi TC.$$

Since $G_{m+1}\left(i+(m+1)\frac{n}{m}\right)=G_{m+1}\left(i+\frac{n}{m}\right)+n\pi TC=G_1\left(i+\frac{n}{m}\right)-mk_0L+n\pi TC$, we have

$$k_0 \leq k_2 \equiv \kappa-\frac{TC}{L}\left[\left(1-\frac{n}{m}\right)\pi-\Psi\left(\pi-\frac{n}{m}-\theta_2\right)\right]=\kappa-\frac{C}{w}\frac{\left(1-\frac{n}{m}\right)\pi-\Psi\left(\pi-\frac{n}{m}-\theta_2\right)}{\theta_2}. \quad (3.18)$$

Then when $k_1 \leq k_2$, which is equivalent to

$$\theta_1+\Psi\left(\pi+\frac{n}{m}-\theta_1\right)+\theta_2+\Psi\left(\pi-\frac{n}{m}-\theta_2\right) \geq 2\pi,$$

the approximate macroscopic fundamental diagram can be written as:

$$q_0=Q(k_0;L,T,\pi,m,n) \approx \min\left\{\frac{k_0}{k_1},1,\frac{\kappa-k_0}{\kappa-k_2}\right\}\pi C. \quad (3.19)$$

Note that (3.5) and (3.14) are special cases of (3.19). Also note that such an approximate macroscopic fundamental diagram can be derived for periodical road with varying length L_j, free-flow speed u_j, congestion wave speed w_j, and offset σ_j.

From Equation (3.19) we can see that the maximum flow-rate, referred to as the network capacity, equals πC when $k_1 \leq k_2$. However, when $k_1 > k_2$, which is equivalent to

$$\theta_1+\Psi\left(\pi+\frac{n}{m}-\theta_1\right)+\theta_2+\Psi\left(\pi-\frac{n}{m}-\theta_2\right)<2\pi, \quad (3.20)$$

the network capacity is smaller than πC. In this case, from Equation the capacity is reached when

$$G_j\left(i+j\frac{n}{m}+\pi\right)=G_j\left(i+j\frac{n}{m}\right)+q_0T=G_{j-1}\left(i+j\frac{n}{m}+\pi-\theta_1\right)=G_{j+1}\left(i+j\frac{n}{m}+\pi-\theta_2\right)+\kappa L. \quad (3.21)$$

This equation will be solved in Chapter 4.

Chapter 4

Analysis of the impacts of signal settings

In this chapter, we analyze the impacts of signal settings on the network flow-rate q_0.

4.1 Impacts of the cycle length

For $(m,n) = (1,1)$ we analyze the impacts of the cycle length on the network flow-rate. We denote $Q_1(k_0;T,\pi) = \frac{k_0}{k_1}\pi C$, and $Q_2(k_0;T,\pi) = \frac{\kappa-k_0}{\kappa-k_1}\pi C$. Then we have the following properties for them:

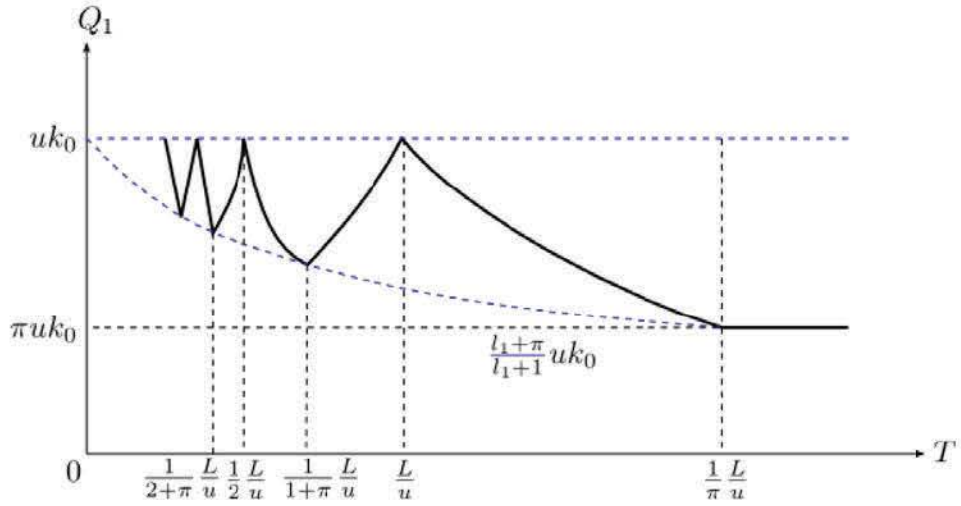
1. When $l_1 = 0$ and $0 < \alpha_1 < 1$, $Q_1(k_0;T,\pi) = \begin{cases} \pi u k_0, 0 < \alpha_1 \le \pi, \\ \alpha_1 u k_0, \pi < \alpha_1 < 1. \end{cases}$ When $l_1 \ge 1$ and $0 \le \alpha_1 < 1$, $Q_1(k_0;T,\pi) = \begin{cases} \frac{l_1+\alpha_1}{\pi l_1+\alpha_1}\pi u k_0, 0 \le \alpha_1 \le \pi, \\ \frac{l_1+\alpha}{l_1+1} u k_0, \pi < \alpha_1 < 1. \end{cases}$ Thus $Q_1(k_0;T,\pi)$ is continuous in θ_1; it retains the global minimum $\pi u k_0$ when $0 < \theta_1 = \alpha_1 \le \pi$, reaches the global maximum $u k_0$ when $\theta_1 = l_1$, and reaches local minima $\frac{l_1+\pi}{l_1+1} u k_0$ when $\theta_1 = l_1 + \pi$.
2. When $l_2 = 0$ and $0 < \alpha_2 < 1$, $Q_2(k_0;T,\pi) = \begin{cases} \pi(\kappa - k_0)w, \ 0 < \alpha_2 \le \pi, \\ \alpha_2(\kappa - k_0)w, \pi < \alpha_2 < 1. \end{cases}$ When $l_2 \ge 1$ and $0 \le \alpha_2 < 1$, $Q_2(k_0;T,\pi) = \begin{cases} \frac{l_2+\alpha_2}{\pi l_2+\alpha_2}\pi(\kappa - k_0)w, \ 0 < \alpha_2 \le \pi, \\ \frac{l_2+\alpha_2}{l_2+1}(\kappa - k_0)w, \pi < \alpha_2 < 1. \end{cases}$ Thus $Q_2(k_0;T,\pi)$ is continuous in θ_2; it retains the global minimum $\pi w(\kappa - k_0)$ when $0 < \theta_2 = \alpha_2 \le \pi$, reaches the global maximum $u(\kappa - k_0)$ when $\theta_2 = l_2$, and reaches local minima $\frac{l_2+\pi}{l_2+1} u(\kappa - k_0)$ when $\theta_2 = l_2 + \pi$.

Since $\theta_1 = \frac{L}{uT}$ and $\theta_2 = \frac{L}{wT}$, we can have the following properties for $Q_1(k_0;T,\pi)$ and $Q_2(k_0;T,\pi)$.

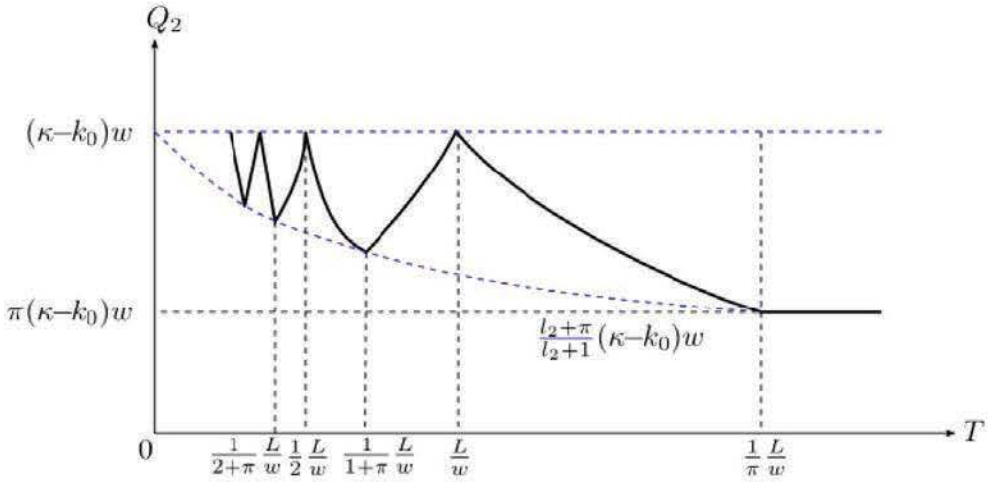
Lemma 4.1 $Q_1(k_0;T,\pi)$ and $Q_2(k_0;T,\pi)$ are functions of T, as shown in Figure 4.1.

1. $Q_1(k_0;T,\pi)$ is continuous in T; it retains the global minimum $\pi u k_0$ when $T \ge \frac{1}{\pi}\frac{L}{u}$, reaches the global maximum $u k_0$ when $T = \frac{1}{l_1}\frac{L}{u}$, and reaches local minima $\frac{l_1+\pi}{l_1+1} u k_0$ when $T = \frac{1}{l_1+\pi}\frac{L}{u}$. In particular, when $\frac{L}{u} \le T \le \frac{1}{\pi}\frac{L}{u}$, $Q_1(k_0;T,\pi)$ decreases in T: $Q_1(k_0;T,\pi) = \frac{k_0 L}{T}$.
2. $Q_2(k_0;T,\pi)$ is continuous in T; it retains the global minimum $\pi w(\kappa - k_0)$ when $T \ge \frac{1}{\pi}\frac{L}{W}$, reaches the global maximum $u(\kappa - k_0)$ when $T = \frac{1}{l_2}\frac{L}{w}$, and

reaches local minima $\frac{l_2+\pi}{l_2+1}u(\kappa - k_0)$ when $T = \frac{1}{l_2+\pi}\frac{L}{w}$. In particular, when $\frac{L}{w} \le T \le \frac{1}{\pi}\frac{L}{w}$, $Q_2(k_0;T,\pi)$ decreases in T: $Q_2(k_0;T,\pi) = \frac{(\kappa - k_0)L}{T}$.



(a)



(b)

Figure 4.1 Flow-cycle length relation: (a) $Q_1(k_0;T,\pi)$ (b) $Q_2(k_0;T,\pi)$

From (3.5) we can see that the flow-rate $Q(k_0;T,\pi)$ decreases in k_1 and increases in k_2. In particular, we have the following Lemma.

Lemma 4.2 In the following five region for k_0, $Q(k_0;T,\pi)$ varies with $k_1 \in [\pi\kappa_c, \kappa_c]$ and $k_2 \in \left[\kappa_c, \kappa - \pi\frac{C}{w}\right]$ as follows:

1. When $k_0 \in [0, \pi\kappa_c]$; i.e., when traffic is very sparse, $Q(k_0;T,\pi) = Q_1(k_0;T,\pi)$ for $k_1 \in [\pi\kappa_c, \kappa_c]$, which decreases in k_1 and is independent of k_2. In this case, the global maximum flow-rate is uk_0 when $T = \frac{1}{l_1}\frac{L}{u}$, and the global minimum flow-rate is $\pi u k_0$ when $T \geq \frac{1}{\pi}\frac{L}{u}$.
2. When $k_0 \in [\pi\kappa_c, \kappa_c]$; i.e., when traffic is sparse, $Q(k_0;T,\pi) = \min\{Q_1(k_0;T,\pi), \pi C\}$, which is first constant for $k_1 \in [\pi\kappa_c, k_0]$, then decreases for $k_1 \in (k_0, \kappa_c]$, and is independent of k_2. In this case, the global maximum flow-rate is πC when $T = \frac{1}{l_1}\frac{L}{u}$, and the global minimum flow-rate is $\pi u k_0$ when $T \geq \frac{1}{\pi}\frac{L}{u}$.
3. When $k_0 = \kappa_c$; i.e., when traffic is critical, $Q(k_0;T,\pi) = \pi C$, which is constant for any k_1 and k_2.
4. When $k_0 \in \left(\kappa_c, \kappa - \pi\frac{C}{w}\right]$; i.e., when traffic is dense, $Q(k_0;T,\pi) = \min\{Q_2(k_0;T,\pi), \pi C\}$, which is first increasing for $k_2 \in [\kappa_c, k_0]$, then constant for $k_2 \in (k_0, \kappa - \pi\frac{C}{w}]$, and is independent of k_1. In this case, the global maximum flow-rate is πC when $T = \frac{1}{l_2}\frac{L}{w}$, and the global minimum flow-rate is $\pi w(\kappa - k_0)$ when $T \geq \frac{1}{\pi}\frac{L}{w}$.
5. When $k_0 \in \left(\kappa - \pi\frac{C}{w}, \kappa\right]$; i.e., when traffic is very dense, $Q(k_0;T,\pi) = Q_2(k_0;T,\pi)$, which is first increasing in k_2, and is independent of k_1. In this case, the global maximum flow-rate is πC when $T = \frac{1}{l_2}\frac{L}{w}$, and the global minimum flow-rate is $\pi w(\kappa - k_0)$ when $T \geq \frac{1}{\pi}\frac{L}{w}$.

From Lemmas 4.1 and 4.2, we can then determine $Q(k_0;T,\pi)$ for any cycle length T and the effective green ratio π, and a density k_0.

4.2 Impacts of the offset

For general (m,n), from the property of $\Psi(t)$ we have that

$$\begin{aligned}(1-\pi)(\theta_1+\theta_2) + 2\pi &\geq \theta_1 + \Psi\left(\pi + \frac{n}{m} - \theta_1\right) + \theta_2 + \Psi\left(\pi - \frac{n}{m} - \theta_2\right) \\ &\geq (1-\pi)(\theta_1+\theta_2) + 2\pi^2.\end{aligned}$$

Thus, if $\theta_1 + \theta_2 \geq 2\pi$, or if $T \leq \frac{L\kappa}{2\pi C}$, then the approximate macroscopic fundamental diagram in (3.19) applies. Thus the necessary condition for the network capacity to be smaller than πC is

$$\theta_1 + \theta_2 < 2\pi.$$

From Figure 4.2 for the properties of $\Psi(\theta)$ and $\theta + \Psi(\theta)$, we can see that both are increasing functions in θ. Thus we have the following lemma.

Lemma 4.3. The sufficient and necessary condition for $\theta_1 + \Psi\left(\pi + \frac{n}{m} - \theta_1\right) + \theta_2 + \Psi\left(\pi - \frac{n}{m} - \theta_2\right) < 2\pi$ is $\theta_1 < \frac{n}{m}$, $\theta_2 < 1 - \frac{n}{m}$, and $\theta_1 + \theta_2 < \pi$. That is, if either $\theta_1 \geq \frac{n}{m}$ or $\theta_2 \geq 1 - \frac{n}{m}$ or $\theta_1 + \theta_2 \geq \pi$, then $\theta_1 + \Psi\left(\pi + \frac{n}{m} - \theta_1\right) + \theta_2 + \Psi\left(\pi - \frac{n}{m} - \theta_2\right) \geq 2\pi$, and the approximate macroscopic fundamental diagram in (3.19) applies.

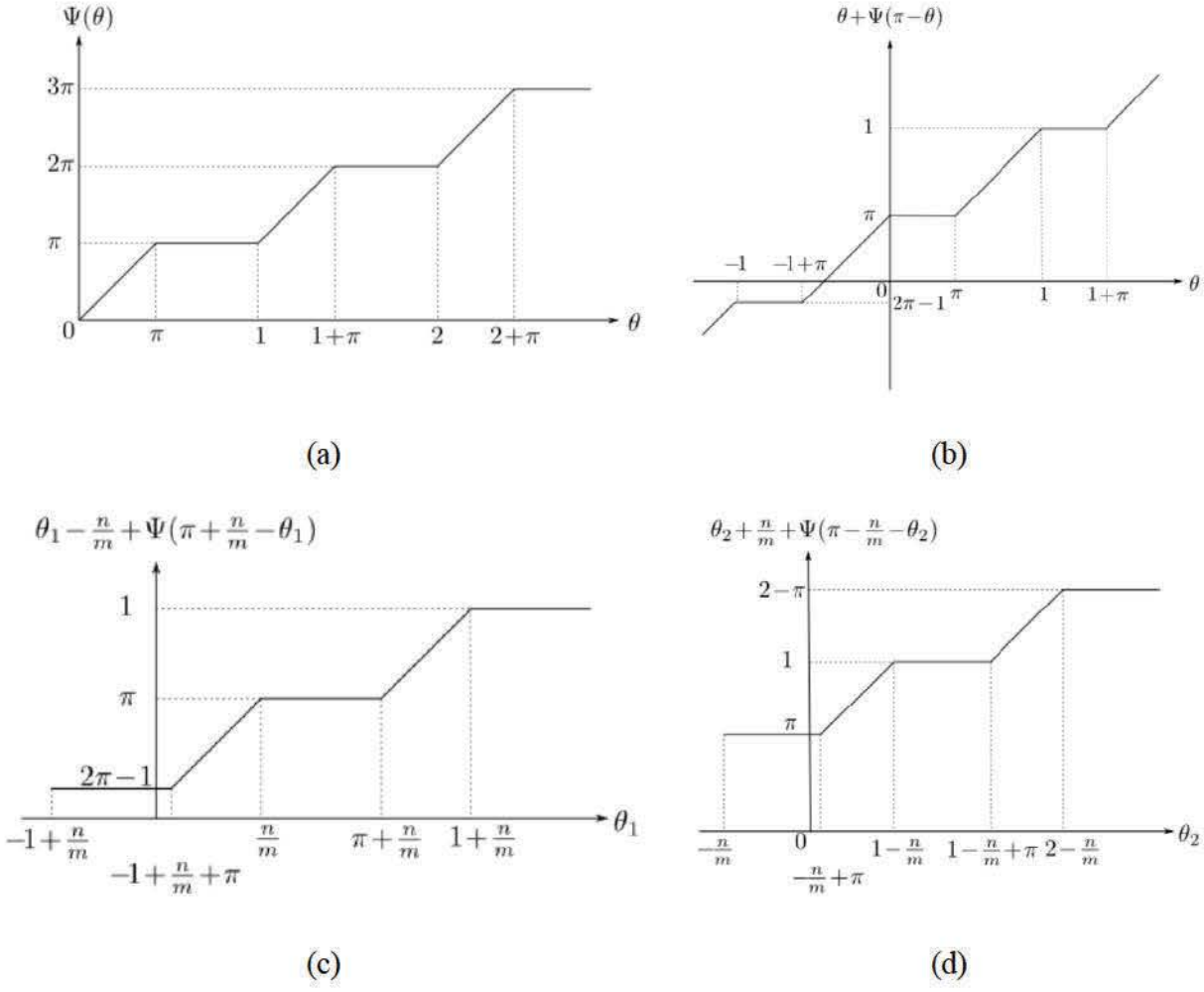


Figure 4.2 Illustrations of $\Psi(\theta)$, $\theta + \Psi(\pi - \theta)$, $\theta_1 - \frac{n}{m} + \Psi\left(\pi + \frac{n}{m} - \theta_1\right)$, and $\theta_2 + \frac{n}{m} + \Psi(\pi - \frac{n}{m} - \theta_2)$

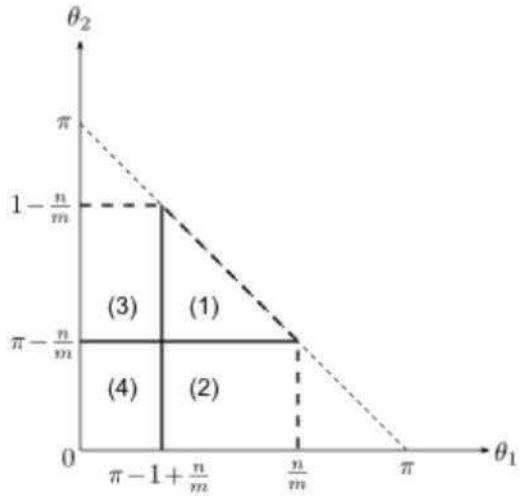


Figure 4.3 Four regions when the network capacity is smaller than πC

When the network capacity is smaller than πC, or equivalently, $\theta_1 + \Psi\left(\pi + \frac{n}{m} - \theta_1\right) + \theta_2 + \Psi\left(\pi - \frac{n}{m} - \theta_2\right) < 2\pi$, Equation (3.21) can be written as

$$G_j\left(i + j\frac{n}{m} + \pi\right) = G_{j-1}\left(i + j\frac{n}{m} + \pi - \theta_1\right) = G_{j+1}\left(i + j\frac{n}{m} + \pi - \theta_2\right) + \kappa L, \tag{4.1}$$

which is equivalent to

$$G_j\left(i + j\frac{n}{m} + \pi\right) = G_{j-1}\left(i + (j-1)\frac{n}{m} + \pi + \frac{n}{m} - \theta_1\right) = G_{j+1}\left(i + (j+1)\frac{n}{m} + \pi - \frac{n}{m} - \theta_2\right) + \kappa L. \tag{4.2}$$

According to Lemma 4.3, when the network capacity is smaller than πC, there are four regions of (θ_1, θ_2), as illustrated in Figure 4.3. In the following we find the network capacity q_* and the corresponding density k_* for the four regions:

1. In region 1, $\pi - 1 + \frac{n}{m} \le \theta_1 < \frac{n}{m}$ and $\pi - \frac{n}{m} \le \theta_2 < 1 - \frac{n}{m}$. Thus $\pi < \pi + \frac{n}{m} - \theta_1 \le 1$ and $-1 + \pi < \pi - \frac{n}{m} - \theta_2 \le 0$, and Equation (4.2) can be written as

   $$G_j\left(i + j\frac{n}{m} + \pi\right) = G_{j-1}\left(i + (j-1)\frac{n}{m} + \pi\right) = G_{j+1}\left(i + (j+1)\frac{n}{m}\right) + \kappa L. \tag{4.3}$$

   Hence $G_m(i + n + \pi) = G_0(i + \pi)$, which leads to $G_m(i + \pi) + nq_*T = G_m(i + \pi) + mk_*L$, or equivalently

   $$nq_*T = mk_*L. \tag{4.4}$$

   In addition, $G_{j+1}\left(i + (j+1)\frac{n}{m} + \pi\right) = G_j\left(i + j\frac{n}{m} + \pi\right) = G_{j+1}\left(i + (j+1)\frac{n}{m}\right) + \kappa L$, which leads to $q_*T = \kappa L$, or equivalently

$$q_* = \frac{\kappa L}{T} = (\theta_1 + \theta_2)C. \tag{4.5}$$

From Equation (4.4) we further have

$$k_* = \frac{n}{m}\kappa. \tag{4.6}$$

2. In region 2, $\pi - 1 + \frac{n}{m} \le \theta_1 < \frac{n}{m}$ and $0 \le \theta_2 < \pi - \frac{n}{m}$. Thus $\pi < \pi + \frac{n}{m} - \theta_1 \le 1$ and $0 < \pi - \frac{n}{m} - \theta_2 \le \pi - \frac{n}{m}$, and Equation (4.2) can be written as

$$G_j\left(i + j\frac{n}{m} + \pi\right) = G_{j-1}\left(i + (j-1)\frac{n}{m} + \pi\right) = G_{j+1}\left(i + (j+1)\frac{n}{m}\right) + (\pi - \frac{n}{m} - \theta_2)CT + \kappa L. \tag{4.7}$$

Hence $G_m(i + n + \pi) = G_0(i + \pi)$, which leads to $G_m(i + \pi) + nq_*T = G_m(i + \pi) + mk_*L$, or equivalently

$$nq_*T = mk_*L. \tag{4.8}$$

In addition, $G_{j+1}\left(i + (j+1)\frac{n}{m} + \pi\right) = G_j\left(i + j\frac{n}{m} + \pi\right) = G_{j+1}\left(i + (j+1)\frac{n}{m}\right) + (\pi - \frac{n}{m} - \theta_2)CT + \kappa L$, which leads to $q_*T = \left(\pi - \frac{n}{m} - \theta_2\right)CT + \kappa L$, or equivalently

$$q_* = \frac{\kappa L}{T} + (\pi - \frac{n}{m} - \theta_2)C = \left(\pi - \frac{n}{m} + \theta_1\right)C. \tag{4.9}$$

From Equation (4.8) we further have

$$k_* = \frac{n}{m}\frac{\left(\pi - \frac{n}{m} + \theta_1\right)}{\theta_1 + \theta_2}\kappa. \tag{4.10}$$

3. In region 3, $0 \le \theta_1 < \pi - 1 + \frac{n}{m}$ and $\pi - \frac{n}{m} \le \theta_2 < 1 - \frac{n}{m}$. Thus $1 < \pi + \frac{n}{m} - \theta_1 \le \pi + \frac{n}{m}$ and $-1 + \pi < \pi - \frac{n}{m} - \theta_2 \le 0$, and Equation (4.2) can be written as

$$G_j\left(i + j\frac{n}{m} + 1\right) = G_{j-1}\left(i + (j-1)\frac{n}{m} + 1\right) + (\pi + \frac{n}{m} - 1 - \theta_1)CT = G_{j+1}\left(i + (j+1)\frac{n}{m}\right) + \kappa L. \tag{4.11}$$

Hence $G_m(i + n + \pi) = G_0(i + \pi) + m(\pi + \frac{n}{m} - 1 - \theta_1)CT$, which leads to $G_m(i + \pi) + nq_*T = G_m(i + \pi) + mk_*L + m(\pi + \frac{n}{m} - 1 - \theta_1)CT$, or equivalently

$$nq_*T = mk_*L + m(\pi + \frac{n}{m} - 1 - \theta_1)CT. \tag{4.12}$$

In addition, $G_{j+1}\left(i + (j+1)\frac{n}{m} + 1\right) = G_j\left(i + j\frac{n}{m} + 1\right) + (\pi + \frac{n}{m} - 1 - \theta_1)CT = G_{j+1}\left(i + (j+1)\frac{n}{m}\right) + \kappa L + (\pi + \frac{n}{m} - 1 - \theta_1)CT$, which leads to $q_*T = \kappa L + (\pi + \frac{n}{m} - 1 - \theta_1)CT$, or equivalently

$$q_* = \frac{\kappa L}{T} + (\pi + \frac{n}{m} - 1 - \theta_1)C = \left(\pi + \frac{n}{m} - 1 + \theta_2\right)C. \quad (4.13)$$

From Equation (4.12) we further have

$$k_* = \frac{\frac{n}{m}\left(\pi+\frac{n}{m}-1+\theta_2\right)-\left(\pi+\frac{n}{m}-1-\theta_1\right)}{\theta_1+\theta_2}\kappa. \quad (4.14)$$

4. In region 4, $0 \le \theta_1 < \pi - 1 + \frac{n}{m}$ and $0 \le \theta_2 < \pi - \frac{n}{m}$. Thus $1 < \pi + \frac{n}{m} - \theta_1 \le \pi + \frac{n}{m}$ and $0 < \pi - \frac{n}{m} - \theta_2 \le \pi - \frac{n}{m}$, and Equation (4.2) can be written as

$$G_j\left(i + j\frac{n}{m} + 1\right) = G_{j-1}\left(i + (j-1)\frac{n}{m} + 1\right) + (\pi + \frac{n}{m} - 1 - \theta_1)CT = G_{j+1}\left(i + (j+1)\frac{n}{m}\right) + (\pi - \frac{n}{m} - \theta_2)CT + \kappa L. \quad (4.15)$$

Hence $G_m(i + n + \pi) = G_0(i + \pi) + m(\pi + \frac{n}{m} - 1 - \theta_1)CT$, which leads to $G_m(i + \pi) + nq_*T = G_m(i + \pi) + mk_*L + m(\pi + \frac{n}{m} - 1 - \theta_1)CT$, or equivalently

$$nq_*T = mk_*L + m(\pi + \frac{n}{m} - 1 - \theta_1)CT. \quad (4.16)$$

In addition, $G_{j+1}\left(i + (j+1)\frac{n}{m} + 1\right) = G_j\left(i + j\,\frac{n}{m} + 1\right) + (\pi + \frac{n}{m} - 1 - \theta_1)CT = G_{j+1}\left(i + (j+1)\frac{n}{m}\right) + (\pi + \frac{n}{m} - 1 - \theta_1)CT + (\pi - \frac{n}{m} - \theta_2)CT + \kappa L$, which leads to $q_*T = (2\pi - 1)CT$, or equivalently

$$q_* = (2\pi - 1)C. \quad (4.17)$$

From Equation (4.16) we further have

$$k_* = \frac{\theta_1 + (1-\pi)(1-2\frac{n}{m})}{\theta_1+\theta_2}\kappa. \quad (4.18)$$

Note that regions (2) and (4) are removed when $\pi - \frac{n}{m} \le 0$, and regions (3) and (4) are removed when $\pi - 1 + \frac{n}{m} \le 0$. In addition, in the four regions, the approximate macroscopic fundamental diagram can be written as

$$q_0 = Q(k_0; L, T, \pi, m, n) \approx \min\left\{\frac{k_0}{k_*}, \frac{\kappa - k_0}{\kappa - k_*}\right\} q_*. \quad (4.19)$$

Chapter 5

Design of traffic signals

From Equation (3.19), when $\theta_1 + \Psi\left(\pi + \frac{n}{m} - \theta_1\right) + \theta_2 + \Psi\left(\pi - \frac{n}{m} - \theta_2\right) \geq 2\pi$, the network flow-rate q_0 is approximately a trapezoidal function of the density k_0. According to Little's law (Little, 1961), the average travel time of each vehicle equals

$$\frac{k_0 L}{q_0} \approx \frac{k_0 L}{\min\left\{\frac{k_0}{k_1}, 1, \frac{\kappa - k_0}{\kappa - k_2}\right\}\pi C}, \tag{5.1}$$

which is a function of both density and signal settings. In the design of traffic signals, the objective is to minimize the average travel time at a given density, or equivalently to maximize the average flow-rate:

$$\underset{T,\pi,m,n}{\operatorname{argmin}} \frac{k_0 L}{q_0} = \underset{T,\pi,m,n}{\operatorname{argmax}} q_0 \approx \underset{T,\pi,m,n}{\operatorname{argmax}} \min\left\{\frac{k_0}{k_1}, 1, \frac{\kappa - k_0}{\kappa - k_2}\right\}\pi C. \tag{5.2}$$

5.1 Design of green ratio

From Equation (5.2), we can see that the larger green ratio, π, the better. However, the green ratio of one movement conflicts with that of its competing movement. In general, π has to be determined by considering the demands of different movements. This is beyond the topic of this research. Here we assume that $\pi \in (0,1)$ is given and attempt to find optimal cycle length and offset.

5.2 Design of cycle length

Here we solve the optimal cycle length for the special case with $(m, n) = (1,0)$. In this case, the optimization problem (5.2) becomes

$$\max_{\theta_1,\theta_2} \min\left\{\frac{\theta_1}{\pi - \Psi(\pi - \theta_1)} \frac{k_0}{\kappa_c}, 1, \frac{\theta_2}{\pi - \Psi(\pi - \theta_2)} \frac{(\kappa - k_0)w}{C}\right\}\pi C. \tag{5.3}$$

Thus we have the following results:

1. For very sparse traffic when $k_0 \leq \pi\kappa_c$, the optimal θ_1^* satisfies $\pi\theta_1^* + \Psi(\pi - \theta_1^*) = \pi$; i.e., $\phi(\theta_1^*) = 0$. Correspondingly the optimal cycle length is $T^* = \frac{1}{\Phi(\theta_1^*)} \frac{L}{u}$.

2. For sparse traffic when $\pi\kappa_c < k_0 < \kappa_c$, the optimal θ_1^* satisfies $k_1 \le k_0$; i.e., $\pi - \Psi(\pi - \theta_1^*) \le \frac{k_0}{\kappa_c}\theta_1^*$, which has multiple solutions. In particular, $\phi(\theta_1^*) = 0$ is an optimal solution.
3. For critical traffic when $k_0 = \kappa_c$, the optimal θ_1^* and θ_2^* satisfies $\theta_1^* + \Psi(\pi - \theta_1^*) \ge \pi$ and $\theta_2^* + \Psi(\pi - \theta_2^*) \ge \pi$, which are true for any θ_1^* and θ_2^*. Therefore, any T will lead to the optimal solution.
4. For dense traffic when $\kappa_c < k_0 < \kappa - \pi\frac{C}{w}$, the optimal θ_2^* satisfies $k_2 \ge k_0$; i.e., $\pi - \Psi(\pi - \theta_2^*) \le \frac{(\kappa - k_0)w}{C}\theta_2^*$, which has multiple solutions. In particular, $\phi(\theta_2^*)$ is an optimal solution.
5. For very dense traffic when $k_0 \ge \kappa - \pi\frac{C}{w}$, the optimal θ_2^* satisfies $\pi\theta_2^* + \Psi(\pi - \theta_2^*) = \pi$; i.e., $\phi(\theta_2^*) = 0$. Correspondingly the optimal cycle length is $T^* = \frac{1}{\Phi(\theta_2^*)}\frac{L}{w}$.

However, in reality due to limited reaction times and bounded acceleration rates of drivers and vehicles, there exists a start-up lost time, δ. The total effective green time for a cycle with two phases is only $(T - 2\delta)$. We assume that the effective green ratio is π_0, which allocates the total effective green time to the studied road. Then the effective green time is $\pi T = (T - 2\delta)\pi_0$. Therefore we have the following effective green ratio

$$\pi = \left(1 - \frac{2\delta}{T}\right)\pi_0, \tag{5.4}$$

and (5.3) can be written as

$$\max_{\theta_1,\theta_2} \min\left\{\frac{\theta_1}{\pi - \Psi(\pi - \theta_1)}\frac{k_0}{\kappa_c}, 1, \frac{\theta_2}{\pi - \Psi(\pi - \theta_2)}\frac{(\kappa - k_0)w}{C}\right\}\left(1 - \frac{2\delta}{T}\right)\pi_0 C. \tag{5.5}$$

Thus we can see that when T is very small for a stop sign, the network flow-rate becomes very low, and we should avoid very small cycle length.

When $\delta \ll 1$, $\pi \approx \pi_0$, and (5.5) can be approximated by

$$\max_{\theta_1,\theta_2} \min\left\{\frac{\theta_1}{\pi_0 - \Psi(\pi_0 - \theta_1)}\frac{k_0}{\kappa_c}, \left(1 - \frac{2\delta}{T}\right), \frac{\theta_2}{\pi_0 - \Psi(\pi_0 - \theta_2)}\frac{(\kappa - k_0)w}{C}\right\}\pi_0 C. \tag{5.5}$$

Then we have the following theorem.

Theorem 5.1 The optimal cycle length considering the start-up lost time is given in the following:

1. For very sparse traffic, $T = \frac{L}{u\theta_1}$, and the maximum flow-rate is $q^* \approx u\,k_0 \le \left(1 - \frac{2\delta}{T^*}\right)\pi_0 C$, for which there exist multiple optimal cycle lengths: $T^* = \frac{1}{\Phi(\theta_1^*)}\frac{L}{u}$ with $\phi(\theta_1^*) = 0$.

2. For sparse traffic, $T = \frac{L}{u\theta_1}$, and the optimal θ_1^* is determined by $\frac{\theta_1^*}{\pi_0 - \Psi(\pi_0 - \theta_1^*)}\frac{k_0}{\kappa_c} = 1 - \frac{2\delta u}{L}\theta_1^*$, which leads to $T^* = \frac{k_0 L}{\pi_0 C} + 2\delta$.
3. For critical traffic, the maximum flow-rate $q^* = \max\left(1 - \frac{2\delta}{T^*}\right)\pi_0 C$. Thus $T^* = \infty$.
4. For dense traffic, $T = \frac{L}{w\theta_2}$, and the optimal θ_2^* is determined by $\frac{\theta_2^*}{\pi_0 - \Psi(\pi_0 - \theta_2^*)}\frac{(\kappa - k_0)w}{C} = 1 - \frac{2\delta w}{L}\theta_2^*$, which leads to $T^* = \frac{(\kappa - k_0)L}{\pi_0 C} + 2\delta$.
5. For very dense traffic, the maximum flow-rate is $q^* \approx w(\kappa - k_0) \leq \left(1 - \frac{2\delta}{T^*}\right)\pi_0 C$, for which there exist multiple optimal cycle lengths: $T^* = \frac{1}{\Phi(\theta_2^*)}\frac{L}{w}$ with $\phi(\theta_2^*) = 0$.

If we denote the congestion level by

$$\mathrm{X} = \frac{\min\{uk_0, C\}}{\min\{C, (\kappa - k_0)w\}}, \tag{5.6}$$

which is the ratio of stationary demand over supply, then we have the following corollary from Theorem 5.1.

Corollary 5.2 The optimal cycle length at different congestion level considering the start-up lost time is given in the following:

1. For very sparse traffic with $\mathrm{X} \in [0, \pi_0)$, $T = \frac{L}{u\theta_1}$, and the maximum flow-rate is $q^* \approx u\, k_0 \leq \left(1 - \frac{2\delta}{T^*}\right)\pi_0 C$, for which there exist multiple optimal cycle lengths: $T^* = \frac{1}{\Phi(\theta_1^*)}\frac{L}{u}$ with $\phi(\theta_1^*) = 0$.
2. For sparse traffic with $\mathrm{X} \in [\pi_0, 1)$, $T = \frac{L}{u\theta_1}$, and the optimal θ_1^* is determined by $\frac{\theta_1^*}{\pi_0 - \Psi(\pi_0 - \theta_1^*)}\frac{k_0}{\kappa_c} = 1 - \frac{2\delta u}{L}\theta_1^*$, which leads to $T^* = \frac{\mathrm{X}L}{\pi_0 u} + 2\delta$.
3. For critical traffic with $\mathrm{X} = 1$, the maximum flow-rate $q^* = max\left(1 - \frac{2\delta}{T^*}\right)\pi_0 C$. Thus $T^* = \infty$.
4. For dense traffic with $\mathrm{X} \in (1, \frac{1}{\pi_0}]$, $T = \frac{L}{w\theta_2}$, and the optimal θ_2^* is determined by $\frac{\theta_2^*}{\pi_0 - \Psi(\pi_0 - \theta_2^*)}\frac{(\kappa - k_0)w}{C} = 1 - \frac{2\delta w}{L}\theta_2^*$, which leads to $T^* = \frac{\mathrm{X}L}{\pi_0 w} + 2\delta$.
5. For very dense traffic with $\mathrm{X} \in (\frac{1}{\pi_0}, \infty)$, the maximum flow-rate is $q^* \approx w(\kappa - k_0) \leq \left(1 - \frac{2\delta}{T^*}\right)\pi_0 C$, for which there exist multiple optimal cycle lengths: $T^* = \frac{1}{\Phi(\theta_2^*)}\frac{L}{w}$ with $\phi(\theta_2^*) = 0$.

When $\mathrm{X} < 1$, traditionally Webster's formula has been used to find the optimal cycle length: even though the above formula is substantially different from Webster's optimal cycle length formula, it is consistent in principle with the latter, as it increases in both the congestion level and the lost time. But here we also obtain a simple formula for dense and very dense traffic with

queue spillback when X > 1, and the optimal cycle length still increases in the lost time but decreases in the congestion level.

5.3 Design of offset

Here we assume that $\theta_1 + \Psi\left(\pi + \frac{n}{m} - \theta_1\right) + \theta_2 + \Psi\left(\pi - \frac{n}{m} - \theta_2\right) \geq 2\pi$. With offset, Equation (5.5) can be written as

$$\max_{\theta_1,\theta_2} \min \left\{ \frac{\theta_1}{\left(1+\frac{n}{m}\right)\pi_0 - \Psi\left(\pi_0 + \frac{n}{m} - \theta_1\right)} \frac{k_0}{\kappa_c}, \left(1 - \frac{2\delta}{T}\right), \frac{\theta_2}{\left(1-\frac{n}{m}\right)\pi_0 - \Psi\left(\pi_0 - \frac{n}{m} - \theta_2\right)} \frac{(\kappa - k_0)w}{C} \right\} \pi_0 C. \quad (5.7)$$

Theorem 5.3 The optimal cycle length, T^*, and offset, $\frac{n^*}{m^*}$, considering the start-up lost time are given in the following:

1. For very sparse traffic, $T = \frac{L}{u\theta_1}$, and the maximum flow-rate is $q^* \approx u\, k_0 \leq \left(1 - \frac{2\delta}{T^*}\right)\pi_0 C$, for which there exist multiple optimal cycle lengths and offsets: $T^* = \frac{1}{\Phi(\theta_1^*) + \frac{n^*}{m^*}} \frac{L}{u}$ with $\phi(\theta_1^*) = \frac{n^*}{m^*}$.
2. For sparse traffic, $T = \frac{L}{u\theta_1}$, and the optimal θ_1^* and offset $\frac{n^*}{m^*}$ are determined by $\frac{\theta_1^*}{\left(1+\frac{n^*}{m^*}\right)\pi_0 - \Psi\left(\pi_0 + \frac{n^*}{m^*} - \theta_1^*\right)} \frac{k_0}{\kappa_c} = 1 - \frac{2\delta u}{L}\theta_1^*$, which leads to $T^* = \frac{k_0 L}{\left(1+\frac{n^*}{m^*}\right)\pi_0 C} + 2\delta$.
3. For critical traffic, the maximum flow-rate $q^* = \max\left(1 - \frac{2\delta}{T^*}\right)\pi_0 C$. Thus $T^* = \infty$.
4. For dense traffic, $T = \frac{L}{w\theta_2}$, and the optimal θ_2^* and offset $\frac{n^*}{m^*}$ are determined by $\frac{\theta_2^*}{\left(1-\frac{n^*}{m^*}\right)\pi_0 - \Psi\left(\pi_0 - \frac{n^*}{m^*} - \theta_2^*\right)} \frac{(\kappa - k_0)w}{C} = 1 - \frac{2\delta w}{L}\theta_2^*$, which leads to $T^* = \frac{(\kappa - k_0)L}{\left(1-\frac{n^*}{m^*}\right)\pi_0 C} + 2\delta$.
5. For very dense traffic, the maximum flow-rate is $q^* \approx w(\kappa - k_0) \leq \left(1 - \frac{2\delta}{T^*}\right)\pi_0 C$, for which there exist multiple optimal cycle lengths and offsets: $T^* = \frac{1}{\Phi(\theta_2^*) + 1 - \frac{n^*}{m^*}} \frac{L}{w}$ with $\phi(\theta_2^*) = 1 - \frac{n^*}{m^*}$.

For a homogeneous road, there is a certain level of equivalence between the cycle length and the offset: given an offset, we can find the optimal cycle length; and given a cycle length in a range, we can find the optimal cycle length. However, for an inhomogeneous road, one can choose different offsets for different signals, so that all signals can have the same cycle length.

Chapter 6

Conclusions and Future Research

In this report, we applied the link transmission model to formulate and analyze traffic dynamics in a signalized arterial network. In particular we (1) analytically derived macroscopic fundamental diagrams for stationary traffic patterns with different network topologies, road conditions, driving behaviors, and signal settings; (2) quantified congestion mitigation effects of different signal settings, including cycle lengths, green splits, and offsets; (3) formulated an optimization problem with the network flow-rate as performance measure to find optimal signal control parameters under certain demand levels, and (4) developed a set of simple decision-support tools for arterial network improvement.

For the homogeneous network without turning movements, this research successfully fills the gap between methods based on delay formulas and those based on traffic simulation by presenting a new method that is both physically realistic and mathematically tractable. There are three particular contributions in this study. First, we obtained a simple link transmission model for the boundary flows on a signalized ring road, which forms the foundation for solving and analyzing stationary states. Second, we derived an explicit approximate macroscopic fundamental diagram, in which the average flow-rate is a function of both traffic density and signal settings. Third, we presented formulas for optimal cycle lengths under five levels of congestion with a start-up lost time.

In the future, we will be interested in applying the method to study an inhomogenoeus road network, in which different roads have different lengths, speed limits, and offsets. We will also be interested in studying the congestion mitigation effect of speed limits and road lengths. In addition, we will also examine the impacts of different start-up lost time caused by different queues.

References

1. Ardekani, S. and Herman, R. (1987). Urban network-wide traffic variables and their relations. Transportation Science, 21(1):1--16.
2. Aboudolas, K., Papageorgiou, M., and Kosmatopoulos, E., 2007, July. Control and optimization methods for traffic signal control in large-scale congested urban road networks. In American Control Conference, 2007. ACC'07 (pp. 3132-3138). IEEE
3. Aboudolas, K., Papageorgiou, M., and Kosmatopoulos, E., 2009. Store-and-forward based methods for the signal control problem in large-scale congested urban road networks. Transportation Research Part C: Emerging Technologies, 17(2), 163-174.
4. Allsop, R. E., 1971a. Delay-minimizing settings for fixed-time traffic signals at a single road junction. IMA Journal of Applied Mathematics 8 (2), 164–185.
5. Allsop, R.E., 1971b. SIGSET: a computer program for calculating traffic signal settings. Traffic Engineering & Control.
6. Allsop, R.E., 1972. Estimating the traffic capacity of a signalized road junction. Transportation Research, 6(3):245-255.
7. Allsop, R.E., 1976. SIGCAP: A computer program for assessing the traffic capacity of signal controlled road junctions. Traffic Engineering & Control, 17(Analytic)
8. Almasri , Essam and Bernhard Friedrich, 2005. Online offset optimisation in urban networks based on cell transmission model. ITS HANNOVER.
9. Beckmann, M., McGuire, C. B., and Winsten, C. B. (1956). Studies in the economics of transportation. Yale University Press, New Haven, Connecticut. Also published as Rand-RM-1488-PR, Rand Corporation, Santa Monica, CA, May 12, 1955.
10. Buisson, C. and Ladier, C. (2009). Exploring the impact of homogeneity of traffic measurements on the existence of macroscopic fundamental diagrams. Transportation Research Record: Journal of the Transportation Research Board, 2124:127--136.
11. Cassidy, M., Jang, K., and Daganzo, C. F. (2011). Macroscopic fundamental diagrams for freeway networks. Transportation Research Record: Journal of the Transportation Research Board, 2260:8--15.
12. Chang, T.-H. and Lin, J.-T. (2000). Optimal signal timing for an oversaturated intersection. Transportation Research Part B: Methodological, 34(6):471--491.
13. Chang, T.-H. and Sun, G.-Y. (2004). Modeling and optimization of an oversaturated signalized network. Transportation Research Part B, 38(8):687--707.
14. Daganzo, C. F., 1994. The cell transmission model: A dynamic representation of highway traffic consistent with the hydrodynamic theory. Transportation Research Part B: Methodological, Vol. 28, No. 4, pp. 269–287.
15. Daganzo, C.F., 1995. The cell transmission model, part II: network traffic. Transportation Research Part B: Methodological, Vol. 29, No. 2, pp. 79–93.
16. Daganzo, C. F. (2005). A variational formulation of kinematic waves: basic theory and complex boundary conditions. Transportation Research Part B, 39(2):187--196.
17. Daganzo, C. F. (2007). Urban gridlock: Macroscopic modeling and mitigation approaches. Transportation Research Part B, 41(1):49--62.

18. Daganzo, C. F. and Geroliminis, N. (2008). An analytical approximation for the macroscopic fundamental diagram of urban traffic. Transportation Research Part B, 42(9):771--781.
19. D'ans, G., Gazis, D., 1976. Optimal control of oversaturated store-and-forward networks. Transportation Science 10 (1), 1–19.
20. De la Breteque, L., and R. Jezequel, 1979. Adaptive control at an isolated intersection-a comparative study of some algorithms. Traffic Engineering & Control.
21. Dion, F., Rakha, H., and Kang, Y. (2004). Comparison of delay estimates at under-saturated and over-saturated pre-timed signalized intersections. Transportation Research Part B: Methodological, 38(2):99--122.
22. Feldman, Olga, and Mike Maher, September, 2002. "The Application of the Cell Transmission Model to the Optimization of Signals on Signalized Roundabouts." PROCEEDINGS OF THE AET EUROPEAN TRANSPORT CONFERENCE, HOMERTON COLLEGE, CAMBRIDGE, UK-CD ROM.
23. Flötteröd, G. and Rohde, J. (2011). Operational macroscopic modeling of complex urban road intersections. Transportation Research Part B, 45(6):903--922.
24. Gartner, N. H., 1983. OPAC: A demand-responsive strategy for traffic signal control. Transportation Research Record, (906).
25. Gartner, N. H., Little, J. D., and Gabbay, H. (1975). Optimization of traffic signal settings by mixed-integer linear programming: Part i: The network coordination problem. Transportation Science, 9(4):321--343.
26. Gartner, N. H. and Wagner, P. (2004). Analysis of traffic flow characteristics on signalized arterials. Transportation Research Record: Journal of the Transportation Research Board, 1883:94--100.
27. Gazis, D. C., 1964. Optimum control of a system of oversaturated intersections. Operations Research 12 (6), 815–831.
28. Gazis, D. C., and R. B. Potts, 1963. The oversaturated intersection. Technical report.
29. Gazis, D. C. (1964). Optimum control of a system of oversaturated intersections. Operations Research, 12(6):815--831.
30. Geroliminis, N. and Boyaci, B. (2012). The effect of variability of urban systems characteristics in the network capacity. Transportation Research Part B, 46(10):1607--1623.
31. Geroliminis, N. and Daganzo, C. F. (2008). Existence of urban-scale macroscopic fundamental diagrams: Some experimental findings. Transportation Research Part B, 42(9):759--770.
32. Geroliminis, N., Haddad, J., and Ramezani, M. (2013). Optimal perimeter control for two urban regions with macroscopic fundamental diagrams: A model predictive approach. Intelligent Transportation Systems, IEEE Transactions on, 14(1):348--359.
33. Godfrey, J. (1969). The mechanism of a road network. Traffic Engineering and Control, 8(8):323--327.
34. Greenshields, B.D., 1935. A study of traffic capacity. Proc. Highway Research Board 14, 448–477.
35. Haberman, R., 1977. Mathematical models. SIAM.
36. He, Qing, Wei-Hua Lin, Hongchao Liu, and K. Larry Head, 2010. Heuristic algorithms to solve 0–1 mixed integer LP formulations for traffic signal control problems. In Service

Operations and Logistics and Informatics (SOLI), 2010 IEEE International Conference. pp. 118-124.
37. Hunt, P., D. Robertson, R. Bretherton, and M. Royle, 1982. The SCOOT on-line traffic signal optimisation technique. Traffic Engineering & Control, 23(4).
38. Hunt, P., D. Robertson, R. Bretherton, and R. Winton, 1981. SCOOT-a traffic responsive method of coordinating signals. Technical report.
39. Improta, G., and Cantarella, G., 1984. Control system design for an individual signalized junction. Transportation Research Part B: Methodological 18 (2147–167.
40. Jin, W.-L., 2010. Continuous kinematic wave models of merging traffic flow. Transportation research part B: methodological, 44(8):1084-1103.
41. Jin, W.-L., 2012a. A kinematic wave theory of multi-commodity network traffic ow. Transportation Research Part B: Methodological, 46(8):1000-1022.
42. Jin, W.-L., 2012b. A link queue model of network traffic flow. arXiv preprint arXiv:1209.2361.
43. Jin, W.-L. (2014). Analysis of kinematic waves arising in diverging traffic flow models. Transportation Science, 49(1):28--45.
44. Jin, W.-L. (2015a). Continuous formulations and analytical properties of the link transmission model. Transportation Research Part B, 74:88--103.
45. Jin, W.-L. (2015b). On the existence of stationary states in general road networks. Transportation Research Part B, 81:917--929.
46. Jin, W.-L., L. Chen, and E. G. Puckett, 2009. Supply-demand diagrams and a new framework for analyzing the inhomogeneous Lighthill-Whitham-Richards model. In Transportation and Traffic Theory 2009: Golden Jubilee, pages 603{635. Springer.
47. Jin, W.-L., Gan, Q.-J., and Gayah, V. V. (2013). A kinematic wave approach to traffic statics and dynamics in a double-ring network. Transportation Research Part B, 57:114--131.
48. Jin, W.-L. and Yu, Y. (2015). Asymptotic solution and effective hamiltonian of a hamilton-jacobi equation in the modeling of traffic flow on a homogeneous signalized road. Journal de Mathematiques Pures et Appliqu´ ees.
49. Lebacque, J., 1996. The Godunov scheme and what it means for first order traffic flow models. In Internaional symposium on transportation and traffic theory, pp. 647–677.
50. Li, Zichuan, 2010. Modeling arterial signal optimization with enhanced cell transmission formulations. Journal of Transportation Engineering 137, no. 7: 445-454.
51. Lighthill, M. and G. Whitham, 1955. On kinematic waves II: A theory of traffic flow on long crowded roads. Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences, Vol. 229, No. 1178, pp. 317–345.
52. Little, J.D., 1966. The synchronization of traffic signals by mixed-integer linear programming. Operations Research, 14(4):568-594.
53. Lo, H. K., 1999. A novel traffic signal control formulation. Transportation Research Part A: Policy and Practice, 33(6), 433-448.
54. Lo, H. K., 2001. A cell-based traffic control formulation: strategies and benefits of dynamic timing plans. Transportation Science, 35(2), 148-164.
55. Lo, H. K., Chang, E., & Chan, Y. C., 2001. Dynamic network traffic control. Transportation Research Part A: Policy and Practice, 35(8), 721-744.
56. Lo, H. K., & Chow, A. H., 2004. Control strategies for oversaturated traffic. Journal of Transportation Engineering, 130(4), 466-478.

57. Mahmassani, H., Williams, J., and Herman, R. (1987). Performance of urban traffic networks. Proceedings of the Tenth International Symposium on Transportation and Traffic Theory.
58. Merchant, D. and Nemhauser, G. (1978). Optimality conditions for a dynamic traffic assignment model. Transportation Science, 12(3):200--207.
59. Miller, A. J., 1963a. A computer control system for traffic networks.
60. Miller, A. J., 1963b. Settings for fixed-cycle traffic signals. OR, 373–386.
61. Moskowitz, K., 1965. Discussion of freeway level of service as influenced by volume and capacity characteristics by DR Drew and CJ Keese. Highway Research Record, 99:43-44.
62. Newell, G. F. (1989). Theory of highway traffic signals. Technical report.
63. Newell, G. F. (1993). A simplified theory of kinematic waves in highway traffic I: General theory. II: Queuing at freeway bottlenecks. III: Multi-destination flows. Transportation Research Part B, 27(4):281--313.
64. Olszewski, P., Fan, H., and Tan, Y. (1995). Area-wide traffic speed-flow model for the singapore cbd. Transportation Research Part A, 29(4):273--281.
65. Papageorgiou, M. (1995). An integrated control approach for traffic corridors. Transportation Research Part C: Emerging Technologies, 3(1):19--30.
66. Papageorgiou, M., Diakaki, C., Dinopoulou, V., Kotsialos, A., Wang, Y., 03. Review of road traffic control strategies. Proceedings of the IEEE 91 (12), 2043–2067.
67. Papageorgiou, M., C. Diakaki, V. Dinopoulou, A. Kotsialos, and Y. Wang. Review of road traffic control strategies. Proceedings of the IEEE, 91(12):2043--2067, 2005.
68. Payne, H. and Thompson, W. (1974). Allocation of freeway ramp metering volumes to optimize corridor performance. IEEE Transactions on Automatic Control, 19(3):177--186.
69. Potts, R. B. and Oliver, R. M. (1972). Flows in transportation networks. Academic Press.
70. Richards, P., 1956. Shock waves on the highway. Operations research, pp. 42–51.
71. Robertson, D. I., 1969. "TRANSYT" method for area traffic control. Traffic Engineering & Control, 11(6).
72. Roess, R., Prassas, E., and McShane, W. (2010). Traffic engineering. Prentice Hall.
73. Smith, M. and Ghali, M. (1990). The dynamics of traffic assignment and traffic control: A theoretical study. Transportation Research Part B, 24(6):409--422.
74. Su, Dongyan, Alex Kurzhanskiy, and Roberto Horowitz, 2013. Simulation of Arterial Traffic Using Cell Transmission Model. In 92nd Transportation Research Board Annual Meeting. No. 13-2387.
75. Tampere, C., Corthout, R., Cattrysse, D., and Immers, L. (2011). A generic class of first order node models for dynamic macroscopic simulation of traffic flows. Transportation Research Part B, 45(1):289--309.
76. Wattleworth, J. (1967). Peak period analysis and control of a freeway system/with discussion.
77. Highway Research Record, 157:1--21.
78. Webster, F. V., 1958. Traffic signal settings. Tech. rep.
79. Yang, H. and Lam, W. H. (1996). Optimal road tolls under conditions of queueing and congestion. Transportation Research Part A, 30(5):319--332.
80. Yang, H. and Yagar, S. (1995). Traffic assignment and signal control in saturated road networks. Transportation Research Part A, 29(2):125--139.

81. Yperman, I., 2007. The link transmission model for dynamic network loading. PhD dissertation.
82. Yperman, I., Logghe, S., Tampere, C., and Immers, B. (2006). The Multi-Commodity Link Transmission Model for Dynamic Network Loading. Proceedings of the TRB Annual Meeting.
83. Zhang, Lihui, Yafeng Yin, and Yingyan Lou, 2010. "Robust signal timing for arterials under day-to-day demand variations." Transportation Research Record: Journal of the Transportation Research Board, 156-166.