Approximate Network Loading and Dual-Time-Scale Dynamic User Equilibrium

Terry L. Friesz[a*] Taeil Kim[a†] Changhyun Kwon[b‡]
Matthew A. Rigdon[a§]
[a]*Department of Industrial and Manufacturing Engineering, Pennsylvania State University, PA 16803, USA*
[b]*Department of Industrial and Systems Engineering, State University of New York, Buffalo, NY 14260, USA*

Abstract

In this paper we present a dual-time-scale formulation of dynamic user equilibrium (DUE) with demand evolution. Our formulation belongs to the problem class that Pang and Stewart (2008) refer to as differential variational inequalities. It combines the within-day time scale for which route and departure time choices fluctuate in continuous time with the day-to-day time scale for which demand evolves in discrete time steps. Our formulation is consistent with the often told story that drivers adjust their travel demands at the end of every day based on their congestion experience during one or more previous days. We show that analysis of the within-day assignment model is tremendously simplified by expressing dynamic user equilibrium as a differential variational inequality. We also show there is a class of day-to-day demand growth models that allow the dual-time-scale formulation to be decomposed by time-stepping to yield a sequence of continuous time, single-day, dynamic user equilibrium problems. To solve the single-day DUE problems arising during time-stepping, it is necessary to repeatedly solve a dynamic network loading problem. We observe that the network loading phase of DUE computation generally constitutes a differential algebraic equation (DAE) system, and we show that the DAE system for network loading based on the link delay model (LDM) of Friesz et al. (1993) may be approximated by a system of ordinary differential equations (ODEs). That system of ODEs, as we demonstrate, may be efficiently solved using traditional numerical methods for such problems. To compute an actual dynamic user equilibrium, we introduce a continuous time fixed-point algorithm and prove its convergence for effective path delay operators that allow a limited type of nonmonotone path delay. We show that our DUE algorithm is compatible with network loading based on the LDM and the cell transmission model (CTM) due to Daganzo (1995).We provide a numerical example based on the much studied Sioux Falls network.

Keywords: *dynamic user equilibrium; differential variational inequalities; differential algebraic equations; dual-time-scale; fixed-point algorithm in Hilbert space*

1 Introductory Remarks

Dynamic traffic assignment (DTA) is usually viewed as the positive (descriptive) modeling of time varying flows on vehicular networks consistent with established traffic flow. This paper is concerned with a specific type of dynamic traffic assignment known as *continuous time dynamic user equilibrium* (DUE) for which unit travel cost, including early and late arrival penalties, is identical for those route and departure time choices selected by travelers between a given origin-destination pair. We study the special case for which travel demand is constant within each given day of interest, although it evolves from day to day. This special case

*Corresponding author. Tel. +1-814-865-2445; fax: +1-814-863-4745; e-mail: tlf13@psu.edu
†Tel. +1-814-865-7654; fax: +1-814-863-4745; e-mail: tzk115@psu.edu
‡Tel. +1-716-645-4705; fax: +1-716-645-3302; e-mail: chkwon@buffalo.edu
§Tel. +1-814-865-7654; fax: +1-814-863-4745; e-mail: mar409@psu.edu

is the simplest plausible circumstance under which a discrete-time, day-to-day model of demand learning may be coupled to a continuous-time, within-day DUE model. In presenting such a dual-time-scale theory, we employ demand evolution dynamics motivated by evolutionary game theory. No doubt more complicated dual-time-scale models can and will be proposed; nonetheless, it is fitting that, as the first such model reported in the dynamic traffic assignment literature, our formulation is particularly easy to understand and to solve by a combination of decomposition via time-stepping and fixed point iterations in a function space[1].

In Sections 2 and 3, we quickly review some already well known material concerning the history of dynamic user equilibrium modeling and computation. In Section 4, we present our dual-time-scale formulation. In the same section, we also show that any continuous time within-day variational inequality formulation of DUE is naturally re-expressible as a differential variational inequality. We go on to show that such a differentiable variational inequality is easily analyzed using elementary optimal control theory. That analysis, elegant in its simplicity and conciseness, is important from a pedagogical point of view and should make the theory of continuous time DUE accessible to a wider body of scholars. Section 6 present a fixed point algorithm for within-day DUE and its convergence for a limited class of non–monotone delay operators. The same section also presents an algorithm for calculating the day-to-day evolution of travel demand consistent with intra-day dynamic user equilibria. Section 7 discusses dynamic network loading from the point of view of differential algebraic equations and gives a detailed procedure for DNL when the link delay model of Friesz et al. (1993) is employed. Section 8 presents the results of several numerical examples.

2 Current State of DUE Scholarly Inquiry

Numerous scholarly teams working independently around the globe have slowly made advances in modeling and computing dynamic use equilibria. In fact, DUE modeling and computation have now reached a point where substantial agreement exists regarding the general content of a mathematical model of dynamic user equilibrium, the desired standards of performance for algorithms that compute DUE flow patterns, and critical unanswered research questions. To document this emerging consensus, in Section 3 we review some of the most significant analytical dynamic user equilibrium models that have been proposed, along with associated algorithms for their solution. That review stresses DUE modeling perspectives that are widely held and widely employed for DUE computational research. Those perspectives must be assessed in the light of available and emerging mathematical and algorithmic tools; when that is done, the following observations may be made:

Observation I Among analytical DUE models, there are no dual-time-scale models recognizing tactical routing and departure time decisions are made in continuous time (the within-day time scale) while demand evolves in discrete time (the day-to-day time scale) and that the two time scales are coupled, although there is considerable agreement that this dichotomy of time scales is apropos.

Observation II We may distinguish two essential aspects of modeling dynamic user equilibrium for the within-day time scale: (i) dynamic network loading and (ii) simultaneous route and departure time equilibria, where dynamic network loading (DNL) subsumes the modeling of delay, flow evolution (arc dynamics) and flow propagation (enforcement of traffic laws during flow evolution).

Observation III Simultaneous route and departure time choice are integral to the definition of a dynamic user equilibrium and have to date been mainly expressed as variational inequalities, quasi-variational inequalities or complementarity problems, either in discrete time or continuous time. However, the emerging literature on abstract *differential variational inequalities* has not been well exploited for either modeling or computing simultaneous route and departure time equilibria.

Obsevation IV Little agreement exists regarding an appropriate mathematical formulation of network loading. Furthermore, the emerging literature on *differential algebraic equations*, despite its focus on problem structures like those encountered in network loading, has not been exploited.

[1]Evolutionary game theory is an appealing foundation for modeling day-to-day demand evolution, as it is principally concerned with dynamic learning processes. Excellent presentations of evolutionary game theory are provided by Hofbauer and Sigmund (1998) and Samuelson (1998).

Observation V Fully general path delay operators may fail to be monotonic and/or differentiable. Rigorously convergent algorithms for determining path departure rates that constitute a user equilibrium for such general path delay operators have not been available.

In this paper we make contributions related to each of the above observations. Specifically, we accomplish the following:

1. the formulation of a dual-time-scale model of dynamic user equilibrium that endogenizes day-to-day evolution of travel demand;
2. the expression of a version of simultaneous path and departure time equilibrium as a *differential* variational inequality that subsumes several of the key models we review, simplifies the analysis of equilibrium conditions, and provides direct access to the growing literature on differential variational inequalities;
3. a *continuous time algorithm* for our differential variational inequality formulation that accommodates a range of DUE models, including those with path delays that are non-differentiable, and/or determined by a separate simulation model;
4. a proof of convergence for the continuous time algorithm that admits a limited class of non-monotonic effective delay operators; and
5. a simplified network loading procedure, in the form of a system of ordinary differential equations, that is applicable when the link delay model of Friesz et al. (1993) is employed to describe arc delay and that allows the efficient calculation of path delays for each vector of departure rates encountered during the computation of a dynamic user equilibrium.

Our dual-time-scale formulation of dynamic user equilibrium with demand growth belongs to the problem class known as *differential variational inequalities*, according to terminology introduced by Pang and Stewart (2008). Our formulation combines the within-day time scale for which route and departure time choices fluctuate in continuous time with the day-to-day time scale for which demand evolves in discrete time steps. For our differential variational inequality formulation, we present and establish convergence of an algorithm that solves day-to-day subproblems using a time-stepping approach and within-day subproblems using a continuous time fixed point scheme. We report numerical tests indicating our DUE algorithm and our DNL procedure, separately and in tandem, are scalable and practical for real applications in a deliberate planning context.

3 Details of Prior DUE Models and Algorithms

Dynamic user equilibrium models, as noted by Peeta and Ziliaskopoulos (2001), tend to be comprised of five essential submodels:

1. a model of path delay;
2. flow dynamics;
3. flow propagation constraints;
4. a route and departure-time choice model; and
5. a model of demand growth.

Dynamic user equilibrium models from the early 1990s forward have been largely concerned with the solution of submodels 1 through 4 above, concentrating on the so-called within-day time scale for which drivers make tactical routing and departure decisions. The notion of "day" here is quite arbitrary and could be any portion of an actual day for which there is a significant, discernible fluctuation in travel demand. It has become common place to use the appellation *dynamic network loading* (DNL) to refer to the determination of arc-specific volumes, arc-specific exit rates and experienced path delay when departure rates are known for each path. As such, dynamic network loading is typically represented by submodels 1, 2 and 3 above, while submodel 4 (simultaneous route and departure choice) mathematically articulates the notion of dynamic user equilibrium in computable form; it is the actual traffic assignment aspect of dynamic user equilibrium modeling. Submodel 5, demand growth, occurs on the day-to-day time scale and allows travel demand to be updated. It is not meaningful to speak of a sequential solution of submodels 1 through 4 without feedback among them. However, we will see ultimately that submodel 5 can be treated sequentially when certain assumptions are made.

The analytical DUE models developed in the early 1990s were influenced greatly by the dynamic system optimal models of Merchant and Nemhauser (1978a,b) who proposed an especially simple type of arc dynamics that still influences present day DUE models. In particular, if one posits that it is possible to specify, or to mathematically derive from some plausible theory, functions that describe the rate at which traffic exits a given network arc for any given volume of traffic present on that arc, one is lead to some deceptively simple arc dynamics. To express this supposition symbolically, let $x_a(t)$ denote the volume of traffic on arc a at time t and let $g_a[x_a(t)]$ be an exit function that gives the rate at which traffic exits from link a. Also let the rate at which traffic enters arc a be denoted by $u_a(t)$. Note that both $g_a[x_a(t)]$ and $u_a(t)$ are rates; that is, they have the units of volume per unit time, so it is appropriate to refer to them as exit flow and entrance flow, respectively. A natural flow balance equation can now be written for each link:

$$\frac{dx_a}{dt} = u_a(t) - g_a[x_a(t)] \qquad \forall a \in \mathcal{A} \tag{1}$$

where every arc of the network of interest is directed and $\mathcal{A}$ denotes the set of all arcs. In (1) each $u_a(t)$ is treated as a control variable. Although (1) is a fairly obvious identity, it is an approximation that only becomes exact when arcs are of infinitesimal length and traffic interactions with other links do not occur. The same dynamics were employed by Friesz et al. (1989) and Wie et al. (1995) in an effort to develop a model of dynamic user equilibrium. However, their model relied on difficult to interpret dynamic shadow prices and does not yield true dynamic user equilibria relative to route and departure time decisions. Moreover, the exit flow functions used in (1) have been widely criticized as difficult to specify and measure. Exit flow functions allow for the potential violation of the first-in-first-out (FIFO) queue discipline as illustrated and discussed by Carey (1986, 1987, 1992, 1995). However, Carey and McCartney (2004) shows that the violation of FIFO is largely dependent upon time and space discretization schemes. Additionally, Carey (2004a) and Carey (2004b) show that exit flow functions satisfy FIFO but not causality. Another problem with exit flow funcions is that an inflow at any time t instant affects the outflows at the same time instant t. These difficulties have caused almost all researchers studying dynamic network flow problems to abandon dynamics based on exit flow functions.

We point out that the Merchant-Nemhauser model employed flow conservation constraints of the following form for the case of a single origin-destination pair:

$$S_k(t) = \sum_{a \in \mathcal{A}(k)} u_a(t) - \sum_{b \in \mathcal{B}(k)} g_a[x_a(t)] \qquad \forall k \in \mathcal{N} \tag{2}$$

where $\mathcal{A}(k)$ is the set of arcs with tail node k, $\mathcal{B}(k)$ is the set of arcs with head node k, and $\mathcal{N}$ is the set of all network nodes.

An intriguing modification of the Merchant-Nemhauser arc dynamics was proposed by Ran et al. (1993). Their idea was to employ dynamics that treat both arc entrance and exit flows as control variables; that is

$$\frac{dx_a^{ij}}{dt} = u_a^{ij}(t) - v_a^{ij}(t) \qquad \forall a \in \mathcal{A}, (i,j) \in \mathcal{W} \tag{3}$$

where $\mathcal{W}$ is the set of all origin-destination pairs, $x_a^{ij}(t)$ is the flow on arc a traveling between origin-destination pair $(i,j) \in \mathcal{W}$, while $u_a^{ij}(t)$ and $v_a^{ij}(t)$ denote the rates at which traffic also traveling between (i,j) enters and exits arc a, respectively. Two types of flow propagation constraints for preventing instantaneous flow propagation and ensuring the FIFO queue discipline have been suggested by Ran et al. (1993) and Ran and Boyce (1996) for dynamics (3). The first is stated as:

$$U_a^p(t) = V_a^p[t + \Delta_a(t)] \qquad \forall a \in A, p \in \mathcal{P}$$

where $U_a^p(\cdot)$ and $V_a^p(\cdot)$ are the cumulative numbers of vehicles associated with path p which are entering and leaving link a respectively, while $\Delta_a(t)$ denotes the time needed to traverse link a at time t and $\mathcal{P}$ is the set of all paths. The meaning of these constraints is fairly intuitive: vehicles entering an arc at a given moment in time must exit at a later time consistent with the arc traversal time. The second Ran et al. type of flow propagation constraint is much more notationally complex and is omitted here for the sake of brevity. Suffice it to say that constraints of this second type are articulated in terms of path-specific arc volumes and are meant to express the idea that path-specific traffic on an arc must ultimately visit a downstream arc or exit the network at the destination node of the path in question. Ran et al. argue that by enforcing this consideration they rule out FIFO violations and instantaneous flow propagation anomalies. Ran et al. (1993) employ a Beckmann-type objective function to create an optimal control model with dynamic user equilibria as solutions. Subsequently Ran and Boyce (1996) use the same general flow dynamics and constraints with variational inequalities to model dynamic user equilibrium.

Friesz et al. (1993) introduced the notion of exit time functions together with a variational inequality to describe dynamic user equilibrium; that model is consistent with FIFO for appropriate arc delay functions, even though explicit flow propagation constraints are not employed. In particular, they introduce a function $\xi_{a_i}^p(t)$ that expresses the time of exit from arc a_i of every path

$$p = \left\{a_1, a_2, ..., a_{i-1}, a_i, a_{i+1}, ..., a_{m(p)}\right\} \in \mathcal{P} \ ,$$

where $\mathcal{P}$ is the set of all network paths. The exit time functions obey the recursive relationships

$$\xi_{a_1}^p = t + D_{a_1}\left[x_{a_1}(t)\right] \qquad \forall\, p \in \mathcal{P} \tag{4}$$

$$\xi_{a_i}^p = \xi_{a_{i-1}}^p(t) + D_{a_i}\left[x_{a_i}\left(\xi_{a_{i-1}}^p(t)\right)\right] \qquad \forall\, p \in \mathcal{P},\ i \in [2, m(p)] \tag{5}$$

where $D_{a_i}\left[x_{a_i}(t)\right]$ is the time to traverse arc a_i; it is a function of the number of vehicles x_{a_i} in front of the entering vehicle at the time of entry. This model of arc delay is herein called the *link delay model* (LDM). Friesz et al. (1993) also give a continuous time articulation of flow conservation based on a fixed within-day trip matrix:

$$\sum_{p \in \mathcal{P}_{ij}} \int_0^T h_p(t)\, dt = Q_{ij} \qquad \forall\, (i,j) \in \mathcal{W} \tag{6}$$

where $\mathcal{W}$ is the set of all origin-destination pairs, $\mathcal{P}_{ij}$ is the set of paths connecting $(i,j) \in \mathcal{W}$ and $h_p(t)$ is the departure rate from the origin of path $p \in \mathcal{P}_{ij}$, while Q_{ij} is the travel demand between $(i,j) \in \mathcal{W}$ and $[0,T] \in \Re_+^1$ is the continuous time interval representing a single day or commuting period of interest. They used (4) and (5) together with dynamics expressed as integral equations involving inverse exit time functions to define an effective path delay operator. That operator, in turn, was used with (6) and non-negativity restrictions to construct an infinite dimensional variational inequality whose solutions are dynamic user equilibria; their formulation is the first expression of dynamic user equilibrium as a variational inequality. Subsequently, Wu et al. (1998) and Xu et al. (1999) developed algorithms for the Friesz et al. (1993) model. In particular they studied the use of the projected gradient method and solved some modest size test problems, but did not provide useful convergence results. Zhu and Marcotte (2000) prove the existence of solutions to the Friesz et al. (1993) model when departure rates are stipulated to be bounded from above. More recently, Bliemer and Bovy (2003) have extended the Friesz et al. (1993) formulation by introducing multiple user classes, thereby creating a quasi-variational inequality.

Friesz et al. (2001) employed path delays computed from (4) and (5) with dynamics

$$\frac{dx_{a_1}^p(t)}{dt} = h_p(t) - g_{a_1}^p(t) \qquad \forall p \in \mathcal{P} \tag{7}$$

$$\frac{dx_{a_i}^p(t)}{dt} = g_{a_{i-1}}^p(t) - g_{a_i}^p(t) \qquad \forall\, p \in \mathcal{P}, i \in [2, m(p)] \ , \tag{8}$$

where $x_{a_i}^p(t)$ is the volume of traffic on arc a_i of path p for $i \in [1, m(p)]$ and $g_{a_i}^p(t)$ denotes the flow exiting that same arc, to formulate the dynamic user equilibrium problem as a differential variational inequality that is completely equivalent to the Friesz et al. (1993) infinite dimensional variational inequality formulation. Friesz et al. (2001) included in their formulation the flow propagation constraints

$$g_{a_1}\left(t+D_{a_1}\left[x_{a_1}(t)\right]\right)\left(1+D'_{a_1}\left[x_{a_1}(t)\right]\dot{x}_{a_1}\right)=h_p(t) \tag{9}$$

$$g^p_{a_i}\left(t+D_{a_i}\left[x_{a_i}(t)\right]\right)\left(1+D'_{a_i}\left[x_{a_i}(t)\right]\dot{x}_{a_i}(t)\right)=g^p_{a_{i-1}}(t)$$
$$\forall\ p\in P,\ i\in[2,m(p)] \tag{10}$$

which are identical to those proposed by Astarita (1995) and which include consideration of expanding/contracting platoons of vehicles. Friesz and Mookherjee (2006) propose and test a fixed point algorithm implemented in continuous time to solve the differential variational inequality formulation of Friesz et al. (2001); that algorithm requires monotonic path delay operators to assure convergence and, hence, is a heuristic in practice.

The paper by Li et al. (2000) is one of several that uses the Friesz et al. (1993) recursive equations (4) and (5) that are based on exit time functions along with the flow propagation constraints (9) and (10) to express path delay and assure physically meaningful flow. They express the DUE conditions in discrete time and show it is equivalent to a finite dimensional variational inequality. They offer an ad hoc algorithm without discussing convergence.

The recursive equations (4) and (5) and flow propagation constraints (9) and (10) also have much in common with the path delay submodel employed by Tong and Wong (2000) in their study of dynamic user equilibrium with bounds on queue length. They too employ a finite dimensional variational inequality formulation to represent a discrete time version of DUE. However, they also propose and successfully test a time-dependent shortest path algorithm in conjunction with incremental loading to compute DUE flows. Huang and Lam (2002) determine path delay using a nested delay function reminiscent of that arising from the recursive relationships (4) and (5); they like several authors before them also use a finite dimensional variational inequality representation of discrete time DUE. Their choice of algorithm is a heuristic route swapping technique that becomes exact only when monotonicity is present, a result that is common in the DUE literature.

An alternative to dynamics like (1), (3), (7) and (8) is provided by the *cell transmission model* (CTM), which is the name given by Daganzo (1994) to discrete time dynamics that approximate the Lighthill-Whitman-Richards (LWR) kinematic wave model of traffic flow (Lighthill and Whitham, 1955; Richards, 1956) and have the form:

$$x_j(t+1)-x_j(t) = y_j(t)-y_{j+1}(t) \tag{11}$$
$$y_j(t) = \min\left\{x_{j-1}(t), Q_j(t), \alpha\left[N_j(t)-x_j(t)\right]\right\} \tag{12}$$

where t is now a discrete time index and a unit time step is employed. In the above, the subscript $j\in C$ refers to a spatially discrete physical "cell" of the highway segment of interest while $(j-1)\in C$ refers to the cell downstream; C is of course the set of cells needed to describe the highway segment of interest. Similarly to before, $x_j(t)$ refers to the traffic volume of cell j. Furthermore, $y_j(t)$ is the actual inflow to cell j, Q_j is the maximal rate of discharge from cell j, N_j is the holding capacity of cell j, and α is a parameter. Daganzo (1995) shows how (11) and (12) can be extended to deal with network structures through straightforward bookkeeping. Note that (12) is a constraint on the variables $x_j(t)$ and $y_j(t)$. The language introduced previously is readily applicable to the cell transmission model; in particular (11) are arc (cell) dynamics (although now several dummy arcs can make up a real physical arc) and (12) are flow propagation constraints. The cell transmission model also includes an implicit notion of arc delay. That notion, however, is somewhat subtle: namely delay is that which occurs from traffic flowing in accordance with the fundamental diagram of road traffic. This is because (12), as explained by Daganzo (1994), is really a piecewise linear approximation of the fundamental diagram of road traffic.

Lo and Szeto (2002) employed the CTM to create link dynamics; they also present a route travel time extraction procedure that allows the CTM to subsume the role of the effective path delay operator described by Friesz et al. (1993). The Lo and Szeto (2002) model is expressed in discrete time and is, hence, finite dimensional. They use a discrete version of (6) to express flow conservation and describe the dynamic user equilibrium itself as a finite dimensional variational inequality that is the discrete time equivalent of the infinite dimensional variational inequality formulation of Friesz et al. (1993). They propose and test an alternating direction method for solving their variational inequality formulation. Convergence is proven based on the assumption that the relevant operators are co-coercive. Co-coercive mappings are monotone and Lipschitz continuous over their domain of definition, although the converse does not hold in general. As a consequence the Lo and Szeto (2002) convergence result does not consider the potentially non-monotonic delay operators we address in subsequent sections of this paper. Szeto and Lo (2004) extend the Lo and Szeto (2002) modeling framework to consider dynamic user equilibrium with elastic travel demand.

Nie and Zhang (2010) use the Friesz et al. (1993) notion of an exit time function to develop recursive formulae for travel time that are equivalent to (4) and (5). They also use (6) to express flow conservation. To these model features, they add arc dynamics and nodal dynamics, the latter to enable the direct consideration of queues. Three different models of arc dynamics are considered: the link delay model, the spatial queue model, and a kinematic wave model based on the LWR theory of traffic flow. Nie and Zhang (2008a) also present a rather general model of traffic flow through nodes, based on notions of virtual demand and virtual supply. The dynamic origin-destination demand estimation problem itself is modeled as a finite dimensional variational inequality that is the discrete time equivalent of the infinite dimensional variational inequality formulation of Friesz et al. (1993). Furthermore, Nie and Zhang (2008a) propose to solve their variational inequality formulation using an equivalent mathematical program based on the idea of a gap function; they test feasible direction methods and the method of successive averages for solving that program. They too do not provide regularity conditions that assure convergence of the algorithms tested.

Ban et al. (2008) employ the arc dynamics (7) and (8) and flow propagation constraints that are a modified version of (9) and (10). They also express flow conservation at the nodal level using constraints similar to (2); as a consequence their definition of dynamic user equilibrium identifies least travel cost as the difference of nodal dual variables. Moreover, they formulate their notion of dynamic user equilibrium as a mixed complementarity problem (MiCP) for which they provide a discrete time approximation. They propose and test a heuristic algorithm based on a partial relaxation of the MiCP.

Important contributions to the literature on arc and path delay dynamics were made in a series of papers by Perakis (2000), Kachani and Perakis (2001), Kachani and Perakis (2002) and Kachani and Perakis (2010) based on the LWR theory but without creating cells. Using the physical length of a path, they derive a constrained differential equation for each path whose solution gives path delay in continuous time. However, the interactions of paths traversing the same arc are not considered and FIFO is not assured when this model is applied to a general network. Perakis and Roels (2006) continue the aforementioned exploration of the LWR model. They are able to derive delay functions for specific flow, spatial and time regimes associated with the propagation of kinematic waves along a roadway arc. They do not consider traffic interactions among arcs incident to the same node, although their delay functions do include exogenous time-dependent functions associated with arc entry and exit that can in principle be adjusted to account for such interactions. They are able to show that their derived delay functions are monotonic, continuous and preserve the FIFO queue discipline provided the rate of evolution of density obeys certain inequalities. They mathematically articulate a dynamic user equilibrium without departure time choice as a path-based infinite dimensional variational inequality whose principal operator is a pure function of time. They propose a version of the Frank-Wolfe algorithm for application to a discrete time approximation and present numerical results for a two arc example problem. Their procedure for computing a dynamic user equilibrium depends on a flow propagation mechanism that does not have a complete mathematical articulation; in particular, Perakis and Roels (2006) remark that "Further research is necessary to develop a mathematical formulation of the flow propagation (mechanism) ...".

4 The Within-Day Differential Variational Inequality Formulation

We will, for the time being, assume the time interval of analysis is a single commuting period

$$[t_0, t_f] \subset \Re^1_+$$

where $t_f > t_0$. The most crucial ingredient of a dynamic user equilibrium model is the path delay operator, which provides the delay on any path p per unit of flow departing from the origin of that path; it is denoted by

$$D_p(t,h) \quad \forall p \in \mathcal{P}$$

where $\mathcal{P}$ is the set of all paths employed by travelers, t denotes departure time, and h is a vector of departure rates. From these we construct effective unit path delay operators $\Psi_p(t,h)$ by adding the so-called schedule delay $F\left[t + D_p(t,h) - T_A\right]$; that is

$$\Psi_p(t,h) = D_p(t,h) + F\left[t + D_p(t,h) - T_A\right] \quad \forall p \in P$$

where T_A is the desired arrival time and $T_A < t_f$. The function $F(\cdot)$ assesses a penalty whenever

$$t + D_p(t,h) \neq T_A \tag{13}$$

since $t + D_p(t,h)$ is the clock time at which departing traffic arrives at the destination of path $p \in \mathcal{P}$. As we have noted in Section 3, the path delay operators may be obtained from an embedded delay model, data combined with response surface methodology, or data combined with inverse modeling. Unfortunately, regardless of how derived, realistic path delay operators do not possess the desirable property of monotonicity; they may also be non-differentiable. We will have more to say about path delays when we discuss dynamic network loading in Section 7.4.

For the time being, there will be a fixed trip matrix

$$Q = (Q_{ij} : (i,j) \in \mathcal{W})$$

where each $Q_{ij} \in \Re^1_{++}$ is the fixed travel demand, expressed as a volume, between origin-destination pair $(i,j) \in \mathcal{W}$ and $\mathcal{W}$ is the set of all origin-destination pairs. Additionally, we will define the set $\mathcal{P}_{ij}$ to be the subset of paths that connect origin-destination pair $(i,j) \in \mathcal{W}$. We denote the space of square integrable functions for the real interval $[t_0, t_f]$ by $L^2[t_0, t_f]$. We stipulate that

$$h \in \left(L^2_+[t_0, t_f]\right)^{|\mathcal{P}|}$$

We write the flow conservation constraints as

$$\sum_{p \in \mathcal{P}_{ij}} \int_{t_0}^{t_f} h_p(t)\, dt = Q_{ij} \quad \forall (i,j) \in \mathcal{W} \tag{14}$$

where (14) is comprised of Lebesgue integrals. We define the set of feasible flows by

$$\Lambda_0 = \left\{ h \geq 0 : \sum_{p \in \mathcal{P}_{ij}} \int_{t_0}^{t_f} h_p(t)\, dt = Q_{ij} \quad \forall (i,j) \in \mathcal{W} \right\} \subseteq \left(L^2_+[t_0, t_f]\right)^{|\mathcal{P}|}$$

Let us also define the infimum of effective travel delays

$$v_{ij} = ess\inf\left[\Psi_p(t,h) : p \in \mathcal{P}_{ij}\right] \quad \forall (i,j) \in \mathcal{W}$$

We now offer the following definition of dynamic user equilibrium first articulated by Friesz et al. (1993):

Definition 1 *Dynamic user equilibrium. A vector of departure rates (path flows) $h^* \in \Lambda_0$ is a dynamic user equilibrium if*

$$h^*_p(t) > 0, p \in P_{ij} \Longrightarrow \Psi_p\left[t, h^*(t)\right] = v_{ij}$$

We denote this equilibrium by $DUE(\Psi, \Lambda_0, [t_0, t_f])$.

The meaning of Definition 1 is clear: positive departure rates at a particular time along a particular path must coincide with least effective travel delay. An implication of the definition is that

$$\Psi_p(t,h^*) > v_{ij}, p \in P_{ij} \Longrightarrow h^*_p = 0$$

Using measure theoretic arguments, Friesz et al. (1993) established that a dynamic user equilibrium is equivalent to the following variational inequality under suitable regularity conditions:

$$\left.\begin{array}{c} \text{find } h^* \in \Lambda_0 \text{ such that} \\ \sum_{p\in\mathcal{P}} \displaystyle\int_{t_0}^{t_f} \Psi_p(t,h^*)(h_p - h_p^*)dt \geq 0 \\ \forall h \in \Lambda_0 \end{array}\right\} \tag{15}$$

It has not been previously noted, however, that (15) is equivalent to a differential variational inequality. This is most easily seen by noting that the flow conservation constraints may be re-stated as

$$\left.\begin{array}{c} \dfrac{dy_{ij}}{dt} = \displaystyle\sum_{p\in\mathcal{P}_{ij}} h_p(t) \quad \forall (i,j) \in \mathcal{W} \\ y_{ij}(t_0) = 0 \quad \forall (i,j) \in \mathcal{W} \\ y_{ij}(t_f) = Q_{ij} \quad \forall (i,j) \in \mathcal{W} \end{array}\right\}$$

which is recognized as a two point boundary value problem. As a consequence (15) may be expressed as the following differential variational inequality (DVI):

$$\left.\begin{array}{c} \text{find } h^* \in \Lambda \text{ such that} \\ \sum_{p\in\mathcal{P}} \displaystyle\int_{t_0}^{t_f} \Psi_p(t,h^*)(h_p - h_p^*)dt \geq 0 \\ \forall h \in \Lambda \end{array}\right\} DVI(\Psi,\Lambda,[t_0,t_f]) \tag{16}$$

where

$$\Lambda = \left\{ h \geq 0 : \frac{dy_{ij}}{dt} = \sum_{p\in P_{ij}} h_p(t), \ y_{ij}(t_0) = 0, \ y_{ij}(t_f) = Q_{ij} \quad \forall (i,j) \in \mathcal{W} \right\}$$

Theorem 1 *Dynamic user equilibrium equivalent to a differential variational inequality. Assume* $\Psi_p(\cdot,h) : [t_0,t_f] \longrightarrow \Re^1_+$ *is measurable and strictly positive for all* $p \in \mathcal{P}$ *and all* $h \in \Lambda$. *A vector of departure rates (path flows)* $h^* \in \Lambda$ *is a dynamic user equilibrium if and only if* h^* *solves* $DVI(\Psi,\Lambda,[t_0,t_f])$*, as defined by (16).*

Proof: This proof may be most easily carried out using the necessary conditions of optimal control theory. The measurability assumption assures that the integrals used in articulating $DVI(\Psi,\Lambda,[t_0,t_f])$ are well defined. The proof is in two parts.

(i) $[DVI(\Psi,\Lambda,[t_0,t_f]) \Longrightarrow DUE(\Psi,\Lambda,[t_0,t_f])]$ The differential variational variational inequality of interest may be written as

$$\sum_{(i,j)\in\mathcal{W}} \sum_{p\in\mathcal{P}_{ij}} \int_{t_0}^{t_f} \Psi_p(t,h^*)h_p dt \geq \sum_{(i,j)\in\mathcal{W}} \sum_{p\in\mathcal{P}_{ij}} \int_{t_0}^{t_f} \Psi_p(t,h^*)h_p^* dt \quad \forall h \in \Lambda$$

which means that the solution $h^* \in \Lambda$ satisfies the optimal control problem

$$\min J_0 = \sum_{(i,j)\in\mathcal{W}} v_{ij}\left[Q_{ij} - y_{ij}(t_f)\right] + \sum_{(i,j)\in\mathcal{W}} \sum_{p\in\mathcal{P}_{ij}} \int_{t_0}^{t_f} \Psi_p(t,h^*)h_p dt \tag{17}$$

subject to

$$\frac{dy_{ij}}{dt} = \sum_{p\in\mathcal{P}_{ij}} h_p(t) \quad \forall (i,j) \in \mathcal{W} \tag{18}$$

$$y_{ij}(t_0) = 0 \quad \forall (i,j) \in \mathcal{W} \tag{19}$$

$$h \geq 0 \tag{20}$$

where the v_{ij} are presently dual variables for the terminal conditions on the state variables. The Hamiltonian for problem (17), (18), (19) and (20) is

$$\begin{aligned} H &= \sum_{(i,j)\in\mathcal{W}} \sum_{p\in\mathcal{P}_{ij}} \Psi_p(t,h^*)h_p + \sum_{(i,j)\in\mathcal{W}} \lambda_{ij} \sum_{p\in\mathcal{P}_{ij}} h_p \\ &= \sum_{(i,j)\in\mathcal{W}} \left\{ \sum_{p\in\mathcal{P}_{ij}} \left[\Psi_p(t,h^*)+\lambda_{ij}\right] h_p \right\} \end{aligned}$$

where the adjoint equations are

$$\frac{d\lambda_{ij}}{dt} = -\frac{\partial H}{\partial y_{ij}} = 0 \quad \forall (i,j)\in\mathcal{W}, p\in\mathcal{P}_{ij}, t\in[t_0,t_f] \tag{21}$$

with transversality conditions

$$\begin{aligned} \lambda_{ij}(T) &= \frac{\partial \sum_{(i,j)\in\mathcal{W}} v_{ij}\left[Q_{ij}-y_{ij}(T)\right]}{\partial y_{ij}(T)} = -v_{ij} = \text{constant} \\ & \qquad \forall (i,j)\in\mathcal{W}, p\in\mathcal{P}_{ij}, t\in[t_0,t_f] \end{aligned} \tag{22}$$

The implication of (21) and (22) is of course that

$$\lambda_{ij}(t) = -v_{ij} \quad \forall (i,j)\in\mathcal{W}, t\in[t_0,t_f]$$

The controls must obey the minimum principle in $\Re^{|\mathcal{P}|}$ for each instant of time; that is, they must solve

$$\min H \quad \text{s.t.} \quad -h\leq 0$$

for which the Kuhn-Tucker conditions are

$$\Psi_p(t,h^*) - v_{ij} = \rho_p \geq 0 \quad \forall (i,j)\in\mathcal{W}, p\in\mathcal{P}_{ij}, t\in[t_0,t_f] \tag{23}$$

where the ρ_p are dual variables satisfying the complementary slackness conditions

$$\rho_p h_p = 0 \quad \forall (i,j)\in\mathcal{W}, t\in[t_0,t_f] \tag{24}$$

From (23) and (24) we have immediately the conditions of a dynamic user equilibrium, namely

$$\begin{aligned} h_p^* &> 0, p\in\mathcal{P}_{ij} \Longrightarrow \Psi_p(t,h^*) = v_{ij} \\ \Psi_p(t,h^*) &> v_{ij}, p\in\mathcal{P}_{ij} \Longrightarrow h_p^* = 0 \end{aligned}$$

with the obvious interpretation that each dual variable v_{ij} is the essential infimum of the effective unit path delay $\Psi_p(t,h^*)$. Thus, we are assured that any solution of our differential variational inequality is a dynamic user equilibrium relative to path and departure time choice.

(ii) $[DUE(\Psi,\Lambda,[t_0,t_f]) \Longrightarrow DVI(\Psi,\Lambda,[t_0,t_f])]$ For a dynamic user equilibrium we have

$$\Psi_p(t,h^*) \geq v_{ij} \quad \forall (i,j)\in\mathcal{W}, p\in\mathcal{P}_{ij}, t\in[t_0,t_f] \tag{25}$$

Therefore if $\left(h_p - h_p^*\right)\geq 0$ we have

$$\Psi_p(t,h^*)\left(h_p - h_p^*\right) \geq v_{ij}\left(h_p - h_p^*\right) \quad \forall (i,j)\in\mathcal{W}, p\in\mathcal{P}_{ij}, t\in[t_0,t_f] \tag{26}$$

However

$$\left(h_p - h_p^*\right) < 0 \Longrightarrow h_p^* > h_p \geq 0 \Longrightarrow h_p^* > 0$$

which requires (25) to hold as an equality, thereby assuring (26) is valid for any $h_p, h_p^* \in \Omega$. As a consequence, we may sum and integrate both sides of (26) to obtain

$$\begin{aligned}
\sum_{p\in\mathcal{P}} \int_{t_0}^{t_f} \Psi_p(t,h^*) \left(h_p - h_p^*\right) dt &\geq \sum_{p\in\mathcal{P}} \int_{t_0}^{t_f} v_{ij} \left(h_p - h_p^*\right) \\
&= \sum_{(i,j)\in\mathcal{W}} v_{ij} \sum_{p\in\mathcal{P}_{ij}} \int_{t_0}^{t_f} \left(h_p - h_p^*\right) \\
&= \sum_{(i,j)\in\mathcal{W}} v_{ij}(Q_{ij} - Q_{ij}) = 0
\end{aligned}$$

This completes the proof. ■

5 A Simple Dual-Time-Scale Model

We wish to give a mathematical statement of the following story. Tomorrow's travel demands are determined tonight after travelers come to rest and reflect on their past congestion experience; this may include consideration of the entire historical congestion record up to and including today. Tomorrow morning every traveler begins to fulfill his/her travel needs expressed as aggregate origin-destination travel demands, making tactical departure time and route choice decisions in continuous time. Tomorrow ends and reflection on the historical congestion experience (which is now one day longer) leads to new travel demands, to be implemented the next morning. Let

$$\tau \in \Upsilon \equiv \{1, 2, ..., N\}$$

be one typical discrete day and take the length of each day to be Δ, while the continuous clock time within each day reads $t \in [(\tau-1)\Delta, \tau\Delta]$ for all $\tau \in \{1, 2, ..., N\}$. The entire planning horizon spans N consecutive days. Let us suppose we have a demand growth model of the abstract form

$$\begin{aligned}
Q_{ij}^{\tau+1} &= \mathcal{F}_{ij}\left(Q_{ij}^{\tau}, h^1, h^2, ..., h^{\tau}; \theta\right) \\
&\qquad \forall (i,j) \in \mathcal{W},\ \tau \in \{1, 2, ...N-1\}
\end{aligned}$$

$$\begin{aligned}
Q_{ij}^{\tau} &\geq 0 \quad \forall (i,j) \in \mathcal{W},\ \tau \in \{2, ..., N\} \\
Q_{ij}^{1} &= K_{ij} \in \Re_+^1 \quad \forall (i,j) \in \mathcal{W}
\end{aligned}$$

where

$$\begin{aligned}
Q_{ij}^{\tau} &= \text{origin-destination travel demand } \forall (i,j) \in \mathcal{W},\ \tau \in \{1, 2, ..., N\} \\
Q^{\tau} &= \left(Q_{ij}^{\tau} : (i,j) \in \mathcal{W}\right) \\
Q &= \left(Q^{\tau} : \tau \in \Upsilon\right) \\
K_{ij} &= \text{a known, non-negative constant } \forall (i,j) \in \mathcal{W} \\
h^{\tau} &= \left(h_p^{\tau} : p \in \mathcal{P}\right) \\
h &= \left(h^{\tau} : \tau \in \Upsilon\right) \\
\theta &= \text{a vector of model parameters}
\end{aligned}$$

Note the change in notation: now h is a tuple of daily flow vectors h^{τ} rather than merely a vector of flows for one representative day. Also we define

$$\Lambda_{\tau}^{0}(Q^{\tau}) = \left\{ h^{\tau} \geq 0 : \sum_{p\in\mathcal{P}_{ij}} \int_{(\tau-1)\cdot\Delta}^{\tau\cdot\Delta} h_p^{\tau}(t)\, dt = Q_{ij}^{\tau} \quad \forall (i,j) \in \mathcal{W} \right\}$$

which is of course equivalent to

$$\Lambda_{\tau}(Q^{\tau}) = \left\{ h^{\tau} \geq 0 : \frac{dy_{ij}}{dt} = \sum_{p\in\mathcal{P}_{ij}} h_p(t),\ y_{ij}\left[(\tau-1)\cdot\Delta\right] = 0,\ y_{ij}(\tau\cdot\Delta) = Q_{ij}^{\tau} \ \ \forall (i,j) \in \mathcal{W} \right\}$$

We note that

$$\Lambda_\tau(Q^\tau) \in \left(L^2\left[\tau \cdot \Delta, (\tau-1)\cdot\Delta\right]\right)^{|\mathcal{P}|}$$

and also define

$$\Lambda(Q) = \prod_{\tau=1}^{N} \Lambda_\tau(Q^\tau)$$

A dual-time-scale model of dynamic user equilibrium with endogenous demand growth is

$$\left.\begin{array}{c}
\text{find } Q^* \geq 0 \text{ and } h^* \in \Lambda(Q^*) \text{ such that} \\[2ex]
\sum_{p\in\mathcal{P}} \displaystyle\int_{(\tau-1)\cdot\Delta}^{\tau\cdot\Delta} \Psi_p(t, h^{\tau*})\left(h_p^\tau - h_p^{\tau*}\right) dt \geq 0 \quad \forall \tau \in \Upsilon, h^\tau \in \Lambda_\tau\left(Q^{\tau*}\right) \\[2ex]
Q_{ij}^{\tau+1,*} = \mathcal{F}_{ij}\left(Q_{ij}^{\tau*}, h^{\tau*}, h^{\tau-1,*}, ..., h^{1*}; \theta\right) \qquad \forall (i,j) \in \mathcal{W}, \tau \in \{1, 2, ..., N-1\} \\[2ex]
Q_{ij}^{1*} = K_{ij} \qquad \forall (i,j) \in \mathcal{W}
\end{array}\right\} \tag{27}$$

This is the most basic model one can articulate about demand growth.

The dual-time-scale model (27) belongs to a class of differential variational inequalities that Pang and Stewart (2008) observe may be solved by time stepping, so that exactly one continuous time variational inequality is faced for each day τ. To understand why time-stepping works for (27), note that when $\tau = 1$ we know each $Q_{ij}^{1*} = K_{ij}$ so that we also know

$$Q^{1*} = (K_{ij} : (i,j) \in \mathcal{W})$$

Thus, we face the well-defined problem of finding $h^{1*} \in \Lambda_1\left(Q^{1*}\right)$ such that

$$\sum_{p\in\mathcal{P}} \int_0^\Delta \Psi_p(t, h^{1*})\left(h_p^1 - h_p^{1*}\right) dt \geq 0 \quad \forall h^1 \in \Lambda_1\left(Q^{1*}\right) \tag{28}$$

The solution of (28) allows us, using the day-to-day demand dynamics, to compute

$$Q_{ij}^{2*} = \mathcal{F}_{ij}\left(Q_{ij}^{1*}, h^{1*}; \theta\right) \quad \forall (i,j) \in \mathcal{W}$$

and thereby determine the vector Q^{2*}, setting the stage for solving the next within-day differential variational inequality to find h^{2*}. This process, known as time-stepping, leads us eventually to a complete solution of (27). It also focuses attention on the need for an algorithm to solve the continuous time differential variational inequality faced for each value of τ.

As an example of dynamics governing the evolution of travel demand, we postulate that, for each day τ, the travel demands Q_{ij}^τ between each given origin-destination pair are determined by the following system of difference equations:

$$Q_{ij}^{\tau+1} = \left[Q_{ij}^\tau - s_{ij}^\tau\left\{\frac{\displaystyle\sum_{p\in\mathcal{P}_{ij}}\sum_{j=1}^{\tau}\int_{(j-1)\cdot\Delta}^{j\cdot\Delta}\Psi_p\left(t, h^{j*}\right) dt}{|\mathcal{P}_{ij}|\cdot\tau\cdot\Delta} - \chi_{ij}\right\}\right]^+$$

$$\forall (i,j) \in \mathcal{W}, \tau \in \{1, 2, ..., L-1\} \tag{29}$$

with boundary conditions

$$Q_{ij}^1 = K_{ij}$$

where $K_{ij} \in \Re_+^1$ is the fixed travel demand for the OD pair $(i,j) \in \mathcal{W}$ for the first day and χ_{ij} is the so-called fitness level. The operator $[x]^+$ is shorthand for $\max[0, x]$. The parameter s_{ij}^τ is related to the rate of change of inter-day travel demand. The above system of difference equations assumes that the moving average of effective travel delay is the principal signal that influences demand adjustments. Such a model is explicitly non-Markovian for it makes today's travel demand dependent on the full history of travel demands. Clearly many other alternative models of demand evolution for the day-to-day time scale may be proposed.

There is a natural time stepping method for solving the dual-time-scale model (27). The main objective of the time stepping method is to decompose the day-to-day dynamics into a sequence of one period within-day DUE problems, so that exactly one DUE problem is faced for each day. Recall the day-to-day demand growth model of our interest is

$$Q_{ij}^{\tau+1} = \mathcal{F}_{ij}\left(Q_{ij}^{\tau}, h^1, h^2, ..., h^{\tau}; \theta\right) \quad \forall (i,j) \in \mathcal{W} \quad \tau \in \{1,2,...,N-1\}$$

$$\begin{aligned} Q_{ij}^{\tau} &\geq 0 \quad \forall (i,j) \in \mathcal{W}, \ \tau \in \{2,...,N\} \\ Q_{ij}^{1} &= K_{ij} \quad \forall (i,j) \in \mathcal{W} \end{aligned}$$

where $\theta \in \Re^n_{++}$ is a vector of parameters. The algorithm itself has the form given below:

Algorithm for the Dual-Time-Scale Model (27)

Step 0. Initialization. Given $Q_{ij}^{1,*} = K_{ij}$ for all $(i,j) \in \mathcal{W}$, choose the vector of model parameters $\theta \in \Re^n_{++}$. Set $\tau = 1$.

Step 1. Solving the Within-Day Model. Solve $DUE(\Psi, \Lambda_{\tau}(Q^{\tau *}), \Delta)$ for day τ by the fixed point algorithm in Section 6.1. Call the solution $h^{\tau,*}$.

Step 2. Update Demand. Using the current dynamic user equilibrium solution, compute the travel demand for the next day according to

$$Q_{ij}^{\tau+1,*} = \mathcal{F}_{ij}\left(Q_{ij}^{\tau,*}, h^{1,*}, h^{2,*}, ..., h^{\tau,*}; \theta\right) \quad \forall (i,j) \in \mathcal{W}$$

Step 3. Time Stepping. If $\tau = N$, stop. Otherwise set $\tau = \tau + 1$ and go to Step 1.

The above algorithm is nothing more than an observation that a decomposition into single-day models that are solved sequentially is possible; this is so because demand depends only on the demands and flows realized over previous days. In the event a total number of trips R_0 is known or stipulated for the N-day horizon, the dual-time-scale model takes a more complicated form, namely the following:

$$\left.\begin{array}{c} \text{find } Q^* \geq 0 \text{ and } h^* \in \Lambda(Q^*) \text{ such that} \\ \sum_{p\in\mathcal{P}} \displaystyle\int_{(\tau-1)\cdot\Delta}^{\tau\cdot\Delta} \Psi_p(t, h^{\tau *})\left(h_p^{\tau} - h_p^{\tau *}\right) dt \geq 0 \quad \forall \tau \in \Upsilon, h^{\tau} \in \Lambda_{\tau}(Q^{\tau *}) \\ Q_{ij}^{\tau+1,*} = \mathcal{F}_{ij}\left(Q_{ij}^{\tau *}, h^{\tau *}, h^{\tau-1,*}, ..., h^{1*}; \theta\right) \quad \forall (i,j) \in \mathcal{W}, \tau \in \{1,2,...,N-1\} \\ \sum_{\tau=1}^{N} Q_{ij}^{\tau} = R_0 \quad \forall (i,j) \in \mathcal{W} \\ Q_{ij}^{1*} = K_{ij} \quad \forall (i,j) \in \mathcal{W} \end{array}\right\} \tag{30}$$

Simple time stepping does not work for the model (30) and a different computational strategy is needed. The solution of (30) is a topic we intend to address in a separate paper at a future date.

6 Within-Day Computation

We have already commented that an algorithm for the within-day differential variational inequality is needed if the dual-time-scale model is to be solved. Solution of the within-day differential variational inequality, as we have also mentioned, is complicated by the fact that the effective delay operator

$$\Psi(t,h) = (\Psi_p(t,h) : p \in \mathcal{W})$$

is typically neither monotonic nor differentiable. Consequently, we must select an algorithm that places minimal restrictions on $\Psi(t,h)$. One such category of algorithms is that of iterative methods in Hilbert space for a fixed point equivalent of the within-day differential variational inequality

$$\left.\begin{array}{c} \text{find } h^{\tau*} \in \Lambda_\tau(Q^\tau) \text{ such that} \\ \sum_{p\in\mathcal{P}} \int_{(\tau-1)\cdot\Delta}^{\tau\cdot\Delta} \Psi_p(t,h^{\tau*})\left(h_p^\tau - h_p^{\tau*}\right)dt \geq 0 \\ \forall h^\tau \in \Lambda_\tau(Q^\tau) \end{array}\right\} DVI(\Psi,\Lambda_\tau,\Delta) \tag{31}$$

for every $\tau \in \Upsilon$. We will use the notation $DUE(\Psi,\Lambda_\tau,\Delta)$ for a within-day dynamic user equilibrium that solves $DVI(\Psi,\Lambda_\tau,\Delta)$ defined in (31) above[2]. With the preceding background, we are in a position to state and prove a result that permits the solution of the differential variational inequality (31) to be obtained by solving an appropriate fixed point problem. That result is:

Theorem 2 *Fixed point re-statement. Assume, for each $\tau \in \Upsilon$, that $\Psi_p(\cdot,h^\tau) : [(\tau-1)\cdot\Delta, \tau\cdot\Delta] \longrightarrow \Re^1_+$ is measurable for all $p \in \mathcal{P}$, $h^\tau \in \Lambda_\tau(Q^\tau)$. Then, for each $\tau \in \Upsilon$, the fixed point problem*

$$h^\tau = P_{\Lambda_\tau(Q^\tau)}\left[h^\tau - \alpha\Psi(t,h^\tau)\right], \tag{32}$$

is equivalent to $DVI(\Psi,\Lambda_\tau,\Delta)$ where $P_{\Lambda_\tau(Q^\tau)}[\cdot]$ is the minimum norm projection onto $\Lambda_\tau(Q^\tau)$ and $\alpha \in \Re^1_{++}$.

Proof: The fixed point problem considered requires that

$$h^\tau = \arg\min_h \left\{\frac{1}{2}\left\|h^\tau - \alpha\Psi(t,h^\tau) - h\right\|^2 : u \in \Lambda_\tau(Q^\tau)\right\} \tag{33}$$

That is, we seek the solution of the optimal control problem

$$\min_h J_\tau(h) = \sum_{(i,j)\in\mathcal{W}} \omega_{ij}\left[Q_{ij}^\tau - y_{ij}(\tau\cdot\Delta)\right] + \int_{(\tau-1)\cdot\Delta}^{\tau\cdot\Delta} \frac{1}{2}\sum_{(i,j)\in\mathcal{W}}\sum_{p\in\mathcal{P}_{ij}}\left[h^\tau - \alpha\Psi(t,h^\tau) - h\right]^2 dt \tag{34}$$

subject to

$$\frac{dy_{ij}}{dt} = \sum_{p\in\mathcal{P}_{ij}} h_p(t) \quad \forall (i,j)\in\mathcal{W} \tag{35}$$

$$y_{ij}[(\tau-1)\cdot\Delta] = 0 \quad \forall (i,j)\in\mathcal{W} \tag{36}$$

$$h \geq 0 \tag{37}$$

This problem has the Hamiltonian

$$H_\tau = \frac{1}{2}\sum_{(i,j)\in\mathcal{W}}\sum_{p\in\mathcal{P}_{ij}}\left[h_p^\tau - \alpha\Psi_p(t,h^\tau) - h_p\right]^2 + \sum_{(i,j)\in\mathcal{W}}\lambda_{ij}\sum_{p\in\mathcal{P}_{ij}} h_p \tag{38}$$

which is convex in its controls h and has no state dependence, so the minimum principle and supporting optimality conditions are both necessary and sufficient. Hence, it will be enough to show that any solution of (34), (35), (36) and (37) satisfies $DVI(\Psi,\Lambda_\tau,\Delta)$. The minimum principle for (38) has the Kuhn-Tucker conditions

$$\left[h_p^\tau - \alpha\Psi_p(t,h^\tau) - h_p^*\right](-1) + \lambda_{ij} = \rho_p \quad \forall (i,j)\in\mathcal{W}, p\in\mathcal{P}_{ij} \tag{39}$$

$$\rho_p h_p^* = 0 \quad \forall (i,j)\in\mathcal{W}, p\in\mathcal{P}_{ij} \tag{40}$$

$$\rho_p \geq 0 \quad \forall (i,j)\in\mathcal{W}, p\in\mathcal{P}_{ij}$$

[2] See Friesz (2010) for additional mathematical details about differential variational inequalities.

where h^* denotes the optimal solution of

$$\min H_\tau \quad \text{s.t.} \quad h \geq 0$$

By virtue of (33), $h_p^* = h_p^\tau$; therefore, we may re-state (39) as

$$\alpha \Psi_p \left(t, h^\tau\right) + \lambda_{ij} = \rho_p \qquad \forall \left(i,j\right) \in \mathcal{W}, p \in \mathcal{P}_{ij}$$

Note also that the adjoint equations and associated transversality conditions are

$$\begin{aligned} \frac{d\lambda_{ij}}{dt} &= \left(-1\right) \frac{\partial H_\tau}{\partial y_{ij}} = 0 \qquad \forall \left(i,j\right) \in \mathcal{W} \\ \lambda_{ij}\left(\tau \cdot \Delta\right) &= \frac{\partial \omega_{ij}\left[Q_{ij} - y_{ij}\left(\tau \cdot \Delta\right)\right]}{\partial y_{ij}\left(\tau \cdot \Delta\right)} = -\omega_{ij} \qquad \forall \left(i,j\right) \in \mathcal{W} \end{aligned}$$

Consequently

$$\lambda_{ij}(t) = -\omega_{ij} \qquad \forall \left(i,j\right) \in \mathcal{W}, t \in \left[(\tau - 1) \cdot \Delta, \tau \cdot \Delta\right]$$

Because $h_p^* = h_p^\tau$, we have by (40) that $\rho_p = 0$ when $h_p^\tau > 0$; hence

$$\begin{aligned} h_p^\tau &> 0,\ p \in \mathcal{P}_{ij} \Longrightarrow \alpha \Psi_p \left(t, h^\tau\right) = -\lambda_{ij} = \omega_{ij} \\ &\Longrightarrow \Psi_p \left(t, h^\tau\right) = \frac{\omega_{ij}}{\alpha} \equiv v_{ij} \end{aligned}$$

which is recognized as the essential feature of $DUE(\Psi, \Lambda_\tau, \Delta)$. ■

6.1 Convergence of the Within-Day Fixed Point Algorithm

Naturally Theorem 2 suggests the algorithm

$$h^{\tau,k+1} = P_{\Lambda_\tau(Q^\tau)} \left[h^{\tau,k} - \alpha \Psi\left(t, h^{\tau,k}\right)\right]$$

which is clearly a particular instance of the abstract algorithm

$$x^{k+1} = M\left(x^k\right) \tag{41}$$

for solving the fixed point problem

$$x = M(x)$$

where $x \in V$, a Hilbert space. The algorithm (41) does not enjoy a theorem assuring strong convergence of $\left\{x^k\right\}$, even when $M : V \longrightarrow V$ is non-expansive. However, there are variants of (41) that are provably strongly convergent (Xu, 2003) under fairly mild regularity conditions. One such is the fixed point algorithm addressed by the following theorem:

Theorem 3 *Strong convergence of modified fixed point algorithm. Consider $M : V \longrightarrow V$ where V is a Hilbert space. The sequence generated by*

$$x^{k+1} = \beta_k z^0 + \left(1 - \beta_k\right) M\left(x^k\right) \tag{42}$$

converges strongly to a fixed point x^* of M, that is

$$\lim_{k \longrightarrow \infty} \left\{x^k\right\} \underset{strongly}{\longrightarrow} x^*$$

$$x^* = M(x^*)\ ,$$

when M is non-expansive, $z^0 \in V$ is an arbitrary point and

(i) $\quad \beta_k \in [0,1]$

(ii) $\quad \lim_{k \longrightarrow \infty} \beta_k = 0$

(iii) $\quad \sum_{k=0}^{\infty} \beta_k = \infty$

(iv) $\quad \lim_{k \longrightarrow \infty} \left(\beta_k - \beta_{k-1}\right) \left(\beta_k\right)^{-1} = 0$

Proof: This result is generally attributed to Halpern (1967), although he used vastly different notation. A generalization whose proof is much more readable is due to Bauschke (1996). Although concerned with a different problem class in Banach spaces, the paper by Xu (2003) gives an informative summary of Theorem 3 and related results stemming from Halpern (1967). ■

Note that condition (*iii*) of Theorem 3 is equivalent to

$$\prod_{k=0}^{\infty}(1-\beta_k)=0$$

and that z^0 may be taken to be x^0, the starting solution. Note also that, because algorithm (42) converges strongly, we know it converges weakly.

6.2 When the Effective Delay is Strongly Monotonic

Our interest now is in constructing a convergent algorithm for the fixed point formulation (32) of within-day dynamic user equilibrium. To that end we propose the algorithm

$$h^{\tau,k+1}=\beta_k h^{\tau,0}+(1-\beta_k)\,P_{\Lambda_\tau(Q^\tau)}\left[h^{\tau,k}-\alpha\Psi\left(t,h^{\tau,k}\right)\right] \tag{43}$$

In order to prove the convergence of (43) using Theorem 3, we will continue to assume that individual path delay operators are measurable and strictly positive; we will also employ some additional regularity conditions. Note that the feasible solutions considered in the regularity conditions below are named $h^{\tau,k}$ and $h^{\tau,k+1}$ for convenience, although they need not be consecutive iterates of the fixed point algorithm to study the non-expansive nature of the operator M_τ; rather they need only be distinct feasible solutions.

We will first consider strongly monotone delay operators; in that case, the regularity conditions for each day $\tau\in\Upsilon$ are:

R1. $\Lambda_\tau(Q^\tau)\subset\left(L_+^2[t_0,t_f]\right)^{|\mathcal{P}|}$; note in particular that $h^\tau\geq 0$.

R2. The unit path delay operator $\Psi(\cdot,h^\tau)$ is measurable and strictly positive on $\Lambda_\tau(Q^\tau)$.

R3. The unit path delay operator $\Psi(\cdot,h^\tau)$ is continuous on $\Lambda_\tau(Q^\tau)$.

R4. The unit path delay operator obeys the Lipschitz condition

$$\left\|\Psi^{\tau,k+1}-\Psi^{\tau,k}\right\|\leq\sqrt{K_0}\left\|h^{\tau,k+1}-h^{\tau,k}\right\|$$

on $\Lambda_\tau(Q^\tau)$, where

$$\Psi^{\tau,k}\equiv\Psi\left(\cdot,h^{\tau,k}\right)$$

and $K_0\in\Re^1_{++}$.

R5. The unit path delay operator $\Psi(\cdot,h^\tau)$ is strongly monotone on $\Lambda_\tau(Q^\tau)$ with constant $K_1\in\Re^1_{++}$; that is

$$\langle\Psi^{\tau,k+1}-\Psi^{\tau,k},h^{\tau,k+1}-h^{\tau,k}\rangle\geq\frac{K_1}{2}\left\|h^{\tau,k+1}-h^{\tau,k}\right\|^2$$

We note that, in the event the effective path delay operator is strongly monotone, an unsurprising version of a standard convergence result is obtained for (43):

Theorem 4 *Strong convergence of modified fixed point algorithm for $DVI(\Psi,\Lambda_\tau,\Delta)$ with strongly monotone delay operators. We assume that regularity conditions R1, R2, R3, R4 and R5 introduced above are in force. Then, provided conditions (i), (ii), (iii) and (iv) of Theorem 3 hold, the sequence $\left\{h^{\tau,k}\right\}$ generated by algorithm (43) converges strongly to $h^{\tau,*}$, a dynamic user equilibrium $DUE(\Psi,\Lambda_\tau,\Delta)$, for all $\tau\in\Upsilon$, provided the fixed point parameter $\alpha\in\Re^1_{++}$ is suitably small.*

Proof: Although quite simple, the proof is instructive as preparation for demonstrating our main result which follows as Theorem 5. In particular, we make note that the fixed point problem underlying (43) is equivalent to $DVI(\Psi, \Lambda_\tau, \Delta)$ and, hence, to $DUE(\Psi, \Lambda_\tau, \Delta)$. As such convergence of (43) is all that needs to be shown. To apply Theorem 3 to that end, we need to establish that the operator

$$M_\tau\left(h^{\tau,k}\right) \equiv P_{\Lambda_\tau(Q^\tau)}\left[h^{\tau,k} - \alpha\Psi\left(t, h^{\tau,k}\right)\right]$$

is non-expansive. It is well known that the minimum norm projection operator is non-expansive in Hilbert space; therefore we have

$$\left\|M_\tau\left(h^{\tau,k+1}\right) - M_\tau\left(h^{\tau,k}\right)\right\| \leq \left\|\left[h^{\tau,k+1} - \alpha\Psi\left(t, h^{\tau,k+1}\right)\right] - \left[h^{\tau,k} - \alpha\Psi\left(t, h^{\tau,k}\right)\right]\right\|$$

or

$$\left\|M_\tau\left(h^{\tau,k+1}\right) - M_\tau\left(h^{\tau,k}\right)\right\| \leq \left\|\left(h^{\tau,k+1} - h^{\tau,k}\right) - \alpha\left(\Psi^{\tau,k+1} - \Psi^{\tau,k}\right)\right\|$$

where the obvious notation

$$\Psi^{\tau,k} \equiv \Psi\left(t, h^{\tau,k}\right)$$

is employed. Thus

$$\left\|M_\tau\left(h^{\tau,k+1}\right) - M_\tau\left(h^{\tau,k}\right)\right\|^2 \leq \left\|\left(h^{\tau,k+1} - h^{\tau,k}\right) - \alpha\left(\Psi^{\tau,k+1} - \Psi^{\tau,k}\right)\right\|^2 \equiv \Phi^{\tau,k}$$

so that

$$\Phi^{\tau,k} = \left\|h^{\tau,k+1} - h^{\tau,k}\right\|^2 - 2\alpha\langle\Psi^{\tau,k+1} - \Psi^{\tau,k}, h^{\tau,k+1} - h^{\tau,k}\rangle + \alpha^2\left\|\Psi^{\tau,k+1} - \Psi^{\tau,k}\right\|^2$$

Lipschitz continuity requires

$$\left\|\Psi^{\tau,k+1} - \Psi^{\tau,k}\right\| \leq \sqrt{K_0}\left\|h^{\tau,k+1} - h^{\tau,k}\right\|$$

for some $K_0 \in \Re^1_{++}$. Thus we have

$$\begin{aligned}\Phi^{\tau,k} &\leq \left\|h^{\tau,k+1} - h^{\tau,k}\right\|^2 - 2\alpha\langle\Psi^{\tau,k+1} - \Psi^{\tau,k}, h^{\tau,k+1} - h^{\tau,k}\rangle + \alpha^2 K_0\left\|h^{\tau,k+1} - h^{\tau,k}\right\|^2 \\ &\leq \left(1 - 2\alpha\frac{K_1}{2} + \alpha^2 K_0\right)\left\|h^{\tau,k+1} - h^{\tau,k}\right\|^2\end{aligned}$$

We require for non-expansiveness of $M_\tau(\cdot)$ the following

$$1 - 2\alpha\frac{K_1}{2} + \alpha^2 K_0 \leq 1$$

or

$$\alpha\left(-K_1 + \alpha K_0\right) \leq 0 \Longrightarrow \alpha < \frac{K_0}{K_1}$$

■

6.3 When the Effective Delay is Component-Wise Strongly Pseudomonotone

Although a literature exists advocating the use of monotone operators for DUE problems, there is evidence that fully general DUE problems involve operators that are pervasively if not always non-monotone Nie (2010). Consequently, Theorem 4 is not entirely satisfactory. Instead, we seek convergence of (43) for non-monotone operators. Accordingly, we now turn our attention to the notion of strongly pseudomonotone operators. We begin our discussion of such operators with the following definition:

Definition 2 *Strongly pseudomonotone operator (Farouq, 2001). The mapping $F(v)$ is strongly pseudomonotone on V with constant $K_1 \in \Re^1_{++}$ if*

$$\langle F\left(v^1\right), v^2 - v^1\rangle \geq 0 \Longrightarrow \langle F\left(v^2\right), v^2 - v^1\rangle \geq \frac{K_1}{2}\left\|v^2 - v^1\right\|^2$$

or equivalently

$$\langle F\left(v^2\right), v^2 - v^1\rangle < \frac{K_1}{2} \Longrightarrow \langle F\left(v^1\right), v^2 - v^1\rangle < 0$$

for all $v^1, v^2 \in V$.

We will now introduce the notion of a component-wise pseudomontone operator; such an operator has components that are pseudomonotone with respect to their own-arguments. In particular, the following definition applies:

Definition 3 *Component-wise strongly pseudomonotone operator. The mapping $F(v)$ is component-wise strongly pseudomonotone on $V \subseteq (L_2[t_0, t_f])^n$ with constant $K_1 \in \Re^1_{++}$ if*

$$\int_{t_0}^{t_f} F_i\left(v^1\right)\left(v_i^2 - v_i^1\right) dt \geq 0 \Longrightarrow \int_{t_0}^{t_f} F_i\left(v^2\right)\left(v_i^2 - v_i^1\right) dt \geq \frac{K_1}{2}\left\|v^2 - v^1\right\|^2 \qquad i = 1, \ldots, n$$

for all $v^1, v^2 \in V$.

We note that

1. a merely monotone function is not strongly pseudomonotone;
2. a merely strongly pseudomonotone function is not strongly monotone;
3. a component-wise strongly pseudomonotone function may fail to be monotone or strongly pseudomonotone.

The first two of the above properties of pseudomonotone functions are discussed in Farouq (2001); the third is easily demonstrated by appeal to the definitions given above. However, the above properties notwithstanding, it may at first seem that no strongly pseudomonotone operator cannot be adequately distinct from a strongly monotone operator to reflect the kind of non-monotonic behavior needed to describe DUE effective path delay. However, that viewpoint is superficial and incorrect, as we illustrate by an example in Appendix A. In fact, that example verifies that functions one might loosely refer to as "seemingly non-monotone" may be component-wise strongly pseudomonotone. Further note that if

$$\sum_{p\in\mathcal{P}} \int_{t_0}^{t_f} \Psi_p^{\tau,k+1}\left(h_p^{\tau,k+1} - h_p^{\tau,k}\right) dt - \sum_{p\in\mathcal{P}} \int_{t_0}^{t_f} \Psi_p^{\tau,k}\left(h_p^{\tau,k+1} - h_p^{\tau,k}\right) dt \geq (-1)\frac{K_2}{2}\left\|h^{\tau,k+1} - h^{\tau,k}\right\|^2 < 0 \quad (44)$$

holds for arbitrary $h^{\tau,k+1}, h^{\tau,k} \in \Lambda_\tau(Q^\tau)$ and some $K_2 \in \Re^1_{++}$, we say that the delay operator is weakly monotone Farouq (2001).

In what follows, we will employ some of the prior regularity conditions as well as some conditions not previously invoked. Note that the feasible solutions considered in the regularity conditions are named $h^{\tau,k}$ and $h^{\tau,k+1}$ for convenience, although they need not be consecutive iterates of the fixed point algorithm to study the non-expansive nature of the operator M_τ; rather they need only be distinct feasible solutions. The regularity conditions for each day $\tau \in \Upsilon$ are discussed below. Note that, in Ŕ4 and Ŕ5, $h_p^{\tau,k}$ and $h_p^{\tau,k+1}$ need only be distinct points in $\Lambda_\tau(Q^\tau)$ and are not necessarily consecutive iterates of the fixed point algorithm:

Ŕ1. $\Lambda_\tau(Q^\tau) \subset \left(L_+^2[t_0, t_f]\right)^{|\mathcal{P}|}$ is compact; note in particular that $h^\tau \geq 0$.

Ŕ2. The unit path delay operator $\Psi(\cdot, h^\tau)$ is measurable and strictly positive on $\Lambda_\tau(Q^\tau)$.

Ŕ3. The unit path delay operator $\Psi(\cdot, h^\tau)$ is continuous on $\Lambda_\tau(Q^\tau)$.

Ŕ4. The unit path delay operator obeys the Lipschitz condition

$$\left\|\Psi^{\tau,k+1} - \Psi^{\tau,k}\right\| \leq \sqrt{K_0}\left\|h^{\tau,k+1} - h^{\tau,k}\right\|$$

on $\Lambda_\tau(Q^\tau)$, where

$$\Psi^{\tau,k} \equiv \Psi\left(\cdot, h^{\tau,k}\right)$$

and $K_0 \in \Re^1_{++}$.

Ŕ5. The unit path delay operator $\Psi(\cdot, h^{\tau})$ is component-wise strongly pseudomonotone (CWSPM) on $\Lambda_{\tau}(Q^{\tau})$ with suitably large constant $K_1 \in \Re^1_{++}$; that is

$$\int_{t_0}^{t_f} \Psi_p^{\tau,k}\left(h_p^{\tau,k+1} - h_p^{\tau,k}\right) dt \geq 0 \Longrightarrow \int_{t_0}^{t_f} \Psi_p^{\tau,k+1}\left(h_p^{\tau,k+1} - h_p^{\tau,k}\right) dt \geq \frac{K_1}{2}\left\|h^{\tau,k+1} - h^{\tau,k}\right\|^2 \tag{45}$$

or equivalently

$$\int_{t_0}^{t_f} \Psi_p^{\tau,k+1}\left(h_p^{\tau,k+1} - h_p^{\tau,k}\right) dt < \frac{K_1}{2}\left\|h^{\tau,k+1} - h^{\tau,k}\right\|^2 \Longrightarrow \int_{t_0}^{t_f} \Psi_p^{\tau,k}\left(h_p^{\tau,k+1} - h_p^{\tau,k}\right) dt < 0 \tag{46}$$

for all $p \in \mathcal{P}$ and all feasible vectors $h^{\tau,k}, h^{\tau,k+1} \in \Lambda_{\tau}(Q^{\tau})$.

Ŕ6. We assume that the effective delay operator is component-wise weakly monotone[3] on $\Lambda_{\tau}(Q^{\tau})$ with suitably small constant $K_2 \in \Re^1_{++}$; that is

$$\int_{t_0}^{t_f} \left(\Psi_p^{\tau,k+1} - \Psi_p^{\tau,k}\right)\left(h_p^{\tau,k+1} - h_p^{\tau,k}\right) dt \geq (-1)\frac{K_2}{2}\left\|h^{\tau,k+1} - h^{\tau,k}\right\|^2 \tag{47}$$

or equivalently

$$\int_{t_0}^{t_f} \left(\Psi_p^{\tau,k} - \Psi_p^{\tau,k+1}\right)\left(h_p^{\tau,k+1} - h_p^{\tau,k}\right) dt \leq \frac{K_2}{2}\left\|h^{\tau,k+1} - h^{\tau,k}\right\|^2$$

for all $p \in \mathcal{P}$ and all feasible vectors $h^{\tau,k}, h^{\tau,k+1} \in \Lambda_{\tau}(Q^{\tau})$.

Ŕ7. We require that

$$\mathcal{P}_6 \neq \emptyset$$

where

$$\mathcal{P}_6 = \left\{ p \in \mathcal{P}_2 : \int_{t_0}^{t_f} \Psi_p^{\tau,k+1}\left(h_p^{\tau,k+1} - h_p^{\tau,k}\right) dt \geq \frac{K_1}{2}\left\|h^{\tau,k+1} - h^{\tau,k}\right\|^2 \quad \forall h_p^{\tau,k}, h_p^{\tau,k+1} \in \Lambda_{\tau}(Q^{\tau}) \right\}$$

and

$$\mathcal{P}_2 = \left\{ p \in \mathcal{P} : \int_{t_0}^{t_f} \Psi_p^{\tau,k}\left(h_p^{\tau,k+1} - h_p^{\tau,k}\right) dt < 0 \quad \forall h_p^{\tau,k}, h_p^{\tau,k+1} \in \Lambda_{\tau}(Q^{\tau}) \right\}$$

We are now ready to state and prove the following result:

Theorem 5 *Strong convergence of modified fixed point algorithm for $DVI(\Psi, \Lambda_{\tau}, \Delta)$ with component-wise strongly pseudomonotone delay operators. We assume that regularity conditions Ŕ1, Ŕ2, Ŕ3, Ŕ4, Ŕ5, Ŕ6 and Ŕ7 introduced above are in force. Then, provided conditions (i), (ii), (iii) and (iv) of Theorem 3 hold, the sequence $\left\{h^{\tau,k}\right\}$ generated by algorithm (43) converges strongly to $h^{\tau,*}$, a dynamic user equilibrium $DUE(\Psi, \Lambda_{\tau}, \Delta)$, for all $\tau \in \Upsilon$, provided there is a fixed point parameter $\alpha \in \Re^1_{++}$ that is suitably small.*

Proof: We continue to employ the notation

$$\Psi^{\tau,k} \equiv \Psi\left(t, h^{\tau,k}\right)$$

The non-expansive nature of the projection operator, as we have already observed, assures that

$$\left\|M_{\tau}\left(h^{\tau,k+1}\right) - M_{\tau}\left(h^{\tau,k}\right)\right\|^2 \leq \left\|\left(h^{\tau,k+1} - h^{\tau,k}\right) - \alpha\left(\Psi^{\tau,k+1} - \Psi^{\tau,k}\right)\right\|^2 \equiv \Phi^{\tau,k}$$

We now observe that

$$\begin{aligned} \Phi^{\tau,k} \quad = \quad & \left\|h^{\tau,k+1} - h^{\tau,k}\right\|^2 + 2\alpha \sum_{i=3}^{6} \sum_{p \in \mathcal{P}_i} \int_{t_0}^{t_f} \left(\Psi_p^{\tau,k} - \Psi_p^{\tau,k+1}\right)\left(h_p^{\tau,k+1} - h_p^{\tau,k}\right) dt \\ & + \alpha^2 \left\|\Psi^{\tau,k+1} - \Psi^{\tau,k}\right\|^2 \end{aligned} \tag{48}$$

[3]Note that weak monotonicity is a very mild condition; it is possible for an operator to be both component-wise strongly pseudomonotone and component-wise weakly monotone, as illustrated in the Appendix.

where

$$\mathcal{P}_1 = \left\{ p \in \mathcal{P} : \int_{t_0}^{t_f} \Psi_p^{\tau,k} \left(h_p^{\tau,k+1} - h_p^{\tau,k} \right) dt \geq 0 \quad \forall h_p^{\tau,k}, h_p^{\tau,k+1} \in \Lambda_\tau \left(Q^\tau \right) \right\} \tag{49}$$

$$\mathcal{P}_2 = \left\{ p \in \mathcal{P} : \int_{t_0}^{t_f} \Psi_p^{\tau,k} \left(h_p^{\tau,k+1} - h_p^{\tau,k} \right) dt < 0 \quad \forall h_p^{\tau,k}, h_p^{\tau,k+1} \in \Lambda_\tau \left(Q^\tau \right) \right\} \tag{50}$$

$$\mathcal{P}_3 = \left\{ p \in \mathcal{P}_1 : \int_{t_0}^{t_f} \Psi_p^{\tau,k+1} \left(h_p^{\tau,k+1} - h_p^{\tau,k} \right) dt < \frac{K_1}{2} \left\| h^{\tau,k+1} - h^{\tau,k} \right\|^2 \quad \forall h_p^{\tau,k}, h_p^{\tau,k+1} \in \Lambda_\tau \left(Q^\tau \right) \right\} \tag{51}$$

$$\mathcal{P}_4 = \left\{ p \in \mathcal{P}_1 : \int_{t_0}^{t_f} \Psi_p^{\tau,k+1} \left(h_p^{\tau,k+1} - h_p^{\tau,k} \right) dt \geq \frac{K_1}{2} \left\| h^{\tau,k+1} - h^{\tau,k} \right\|^2 \quad \forall h_p^{\tau,k}, h_p^{\tau,k+1} \in \Lambda_\tau \left(Q^\tau \right) \right\} \tag{52}$$

$$\mathcal{P}_5 = \left\{ p \in \mathcal{P}_2 : \int_{t_0}^{t_f} \Psi_p^{\tau,k+1} \left(h_p^{\tau,k+1} - h_p^{\tau,k} \right) dt < \frac{K_1}{2} \left\| h^{\tau,k+1} - h^{\tau,k} \right\|^2 \quad \forall h_p^{\tau,k}, h_p^{\tau,k+1} \in \Lambda_\tau \left(Q^\tau \right) \right\} \tag{53}$$

$$\mathcal{P}_6 = \left\{ p \in \mathcal{P}_2 : \int_{t_0}^{t_f} \Psi_p^{\tau,k+1} \left(h_p^{\tau,k+1} - h_p^{\tau,k} \right) dt \geq \frac{K_1}{2} \left\| h^{\tau,k+1} - h^{\tau,k} \right\|^2 \quad \forall h_p^{\tau,k}, h_p^{\tau,k+1} \in \Lambda_\tau \left(Q^\tau \right) \right\} \tag{54}$$

Note that $\mathcal{P}_3 = \emptyset$ by virtue of the definition of component-wise strong pseudomonotonicity; furthermore

$$\begin{aligned} \mathcal{P} &= \mathcal{P}_1 \cup \mathcal{P}_2 \\ \mathcal{P} &= \mathcal{P}_3 \cup \mathcal{P}_4 \cup \mathcal{P}_5 \cup \mathcal{P}_6 \\ \mathcal{P}_1 \cap \mathcal{P}_2 &= \emptyset \\ \mathcal{P}_i \cap \mathcal{P}_j &= \emptyset \quad \forall i, j \in [3, 4, 5, 6] \quad \text{and} \quad i \neq j \end{aligned}$$

Next note that for any path $p \in \mathcal{P}_6$, we have

$$\begin{aligned} \int_{t_0}^{t_f} \Psi_p^{\tau,k} \left(h_p^{\tau,k+1} - h_p^{\tau,k} \right) dt &< 0 \\ \int_{t_0}^{t_f} \Psi_p^{\tau,k+1} \left(h_p^{\tau,k+1} - h_p^{\tau,k} \right) dt &\geq \frac{K_1}{2} \left\| h^{\tau,k+1} - h^{\tau,k} \right\|^2 \\ -\int_{t_0}^{t_f} \Psi_p^{\tau,k+1} \left(h_p^{\tau,k+1} - h_p^{\tau,k} \right) dt &\leq (-1) \frac{K_1}{2} \left\| h^{\tau,k+1} - h^{\tau,k} \right\|^2 \end{aligned}$$

Therefore

$$\int_{t_0}^{t_f} \Psi_p^{\tau,k} \left(h_p^{\tau,k+1} - h_p^{\tau,k} \right) dt - \int_{t_0}^{t_f} \Psi_p^{\tau,k+1} \left(h_p^{\tau,k+1} - h_p^{\tau,k} \right) dt < (-1) \frac{K_1}{2} \left\| h^{\tau,k+1} - h^{\tau,k} \right\|^2 \qquad p \in \mathcal{P}_6 \tag{55}$$

Consequently

$$2\alpha \sum_{p \in \mathcal{P}_6} \int_{t_0}^{t_f} \left(\Psi_p^{\tau,k} - \Psi_p^{\tau,k+1} \right) \left(h_p^{\tau,k+1} - h_p^{\tau,k} \right) dt < (-1) \alpha K_1 \cdot |\mathcal{P}_6| \cdot \left\| h^{\tau,k+1} - h^{\tau,k} \right\|^2$$

Furthermore, weak monotonicity assures

$$2\alpha \sum_{i=3}^{5} \sum_{p \in \mathcal{P}_i} \int_{t_0}^{t_f} \left(\Psi_p^{\tau,k} - \Psi_p^{\tau,k+1} \right) \left(h_p^{\tau,k+1} - h_p^{\tau,k} \right) dt \leq \alpha K_2 \cdot \Big(|\mathcal{P}_4| + |\mathcal{P}_5| \Big) \cdot \left\| h^{\tau,k+1} - h^{\tau,k} \right\|^2 \tag{56}$$

It follows from (55) and (56) that

$$\begin{aligned} E &\equiv 2\alpha \sum_{i=3}^{6} \sum_{p \in \mathcal{P}_i} \int_{t_0}^{t_f} \left(\Psi_p^{\tau,k} - \Psi_p^{\tau,k+1} \right) \left(h_p^{\tau,k+1} - h_p^{\tau,k} \right) dt & (57) \\ &< \alpha \Big[K_2 \cdot \Big(|\mathcal{P}_4| + |\mathcal{P}_5| \Big) - K_1 \cdot |\mathcal{P}_6| \Big] \cdot \left\| h^{\tau,k+1} - h^{\tau,k} \right\|^2 & (58) \end{aligned}$$

Therefore, by the Lipschitz assumption, we have

$$\begin{aligned}\Phi^{\tau,k} &\leq \left\|h^{\tau,k+1}-h^{\tau,k}\right\|^2+E+\alpha^2K_0\left\|h^{\tau,k+1}-h^{\tau,k}\right\|^2\\ &< \left(1+\alpha\left[K_2\cdot\Big(|\mathcal{P}_4|+|\mathcal{P}_5|\Big)-K_1\cdot|\mathcal{P}_6|\right]+\alpha^2K_0\right)\cdot\left\|h^{\tau,k+1}-h^{\tau,k}\right\|^2\end{aligned}$$

We seek $\Phi^{\tau,k}<1$, which is assured by

$$1+\alpha\left[K_2\cdot\Big(|\mathcal{P}_4|+|\mathcal{P}_5|\Big)-K_1\cdot|\mathcal{P}_6|\right]+\alpha^2K_0\ <1$$

It follows that we must have

$$K_2\cdot\Big(|\mathcal{P}_4|+|\mathcal{P}_5|\Big)-K_1\cdot|\mathcal{P}_6|+\alpha K_0<0\Longrightarrow\alpha<\frac{K_1\cdot|\mathcal{P}_6|-K_2\cdot\Big(|\mathcal{P}_4|+|\mathcal{P}_5|\Big)}{K_0}\equiv\bar{\alpha}$$

for K_1 and K_2 such that

$$K_1>\frac{K_2\cdot|\mathcal{P}_4|\cdot|\mathcal{P}_5|}{|\mathcal{P}_6|}$$

Consequently

$$\begin{aligned}\left\|M_\tau\left(h^{\tau,k+1}\right)-M_\tau\left(h^{\tau,k}\right)\right\| &\leq \sqrt{\Phi^{\tau,k}}<\sqrt{1+\alpha\left(K_2\cdot\Big(|\mathcal{P}_4|+|\mathcal{P}_5|\Big)-K_1\cdot|\mathcal{P}_6|+\alpha K_0\right)}\left\|h^{\tau,k+1}-h^{\tau,k}\right\|\\ &< \left\|h^{\tau,k+1}-h^{\tau,k}\right\|\end{aligned}$$

for

$$0<\alpha<\bar{\alpha}\in\Re^1_{++}$$

In other words $M_\tau(\cdot)$ is non-expansive. Also, conditions (i), (ii), (iii) and (iv) of Theorem 3 regarding the sequence $\{\beta_k\}$ are met by the given. Hence, for each $\tau\in\Upsilon$ and suitably small fixed point parameter α, we know by Theorem 3 that the sequence $\left\{h^{\tau,k}\right\}$ converges strongly to $h^{\tau,*}$, a within-day dynamic user equilibrium flow pattern. ■

6.4 Details of the Fixed Point Algorithm

The question of what rules to use in selecting the coefficients β_k in (43) that satisfy conditions (i), (ii) (iii) and (iv) of Theorem 3 naturally arises; one choice is

$$\beta_k=\left(\frac{1}{k+1}\right)^q$$

where q is a parameter to be designated and which must satisfy

$$0<q\ <1$$

The algorithm itself has the form given below:

Fixed Point Algorithm for $DUE(\Psi,\Lambda_\tau,\Delta)$

Step 0. Initialization. Select $h^{\tau,0}$ and the rule for generating the sequence $\{\beta_k\}$. Also select a stopping tolerance $\epsilon\in\Re^1_{++}$. Set $k=0$.

Step 1. Major iteration. Compute

$$h^{\tau,k+1}=\beta_kh^{\tau,0}+(1-\beta_k)\,P_{\Lambda_\tau(Q^\tau)}\left[h^{\tau,k}-\alpha\Psi\left(t,h^{\tau,k}\right)\right]$$

Step 2. Stopping test. If

$$\left\|h^{\tau,k+1}-h^{\tau,k}\right\|\leq\epsilon$$

stop and declare

$$h^{\tau,*}\approx h^{\tau,k+1}$$

Otherwise set $k=k+1$ and go to Step 1.

6.5 The Sub-Problems of the Within-Day Fixed Point Algorithm

Note that the major iterations (Step 1) of the fixed point algorithm presented above require that the following optimal control problem be solved:

$$\begin{aligned}\min_h J_\tau^k(h) &= \sum_{(i,j)\in\mathcal{W}} v_{ij}\left[Q_{ij}^\tau - y_{ij}\left(\tau\cdot\Delta\right)\right] \\ &\quad + \int_{(\tau-1)\cdot\Delta}^{\tau\cdot\Delta} \frac{1}{2}\sum_{(i,j)\in\mathcal{W}}\sum_{p\in\mathcal{P}_{ij}}\left[h^{\tau,k} - \alpha\Psi\left(t,h^{\tau,k}\right) - h\right]^2 dt \qquad (59)\end{aligned}$$

subject to

$$\frac{dy_{ij}}{dt} = \sum_{p\in\mathcal{P}_{ij}} h_p\left(t\right) \quad \forall\left(i,j\right)\in\mathcal{W} \qquad (60)$$

$$y_{ij}\left[\left(\tau-1\right)\cdot\Delta\right] = 0 \quad \forall\left(i,j\right)\in\mathcal{W} \qquad (61)$$

$$h \geq 0 \qquad (62)$$

Note that in order to solve this optimal control problem we must know the dual variables v_{ij} for all $(i,j)\in\mathcal{W}$; finding them is an essential task if we are to compute a within-day dynamic user equilibrium.

6.6 Finding the Subproblem Dual Variables v_{ij}

Finding the dual variables associated with the terminal time demand constraints turns out to be relatively easy. Note that the relevant Hamiltonian for (59), (60), (61) and (62) is

$$H_\tau^k = \frac{1}{2}\sum_{(i,j)\in\mathcal{W}}\sum_{p\in\mathcal{P}_{ij}}\left[h_p^{\tau,k} - \alpha\Psi_p\left(t,h^{\tau,k}\right) - h_p\right]^2 + \sum_{(i,j)\in\mathcal{W}}\lambda_{ij}\sum_{p\in P_{ij}} h_p$$

where each λ_{ij} is an adjoint variable obeying

$$\begin{aligned}\frac{d\lambda_{ij}}{dt} &= (-1)\frac{\partial H_\tau^k}{\partial y_{ij}} = 0 \qquad \forall\left(i,j\right)\in\mathcal{W} \\ \lambda_{ij}(\tau\cdot\Delta) &= \frac{\partial}{\partial y_{ij}(\tau\cdot\Delta)} v_{ij}\left[Q_{ij} - y_{ij}\left(\tau\cdot\Delta\right)\right] = -v_{ij} \qquad \forall\left(i,j\right)\in\mathcal{W}\end{aligned}$$

from which we determine that

$$\lambda_{ij}(t) = -v_{ij} = \text{constant} \qquad \forall\left(i,j\right)\in\mathcal{W}$$

The minimum principle is simply

$$\begin{aligned}h_p^{\tau,k+1} &= \arg\left\{\frac{\partial H_1}{\partial h_p} = 0\right\} \qquad p\in\mathcal{P} \\ &= \arg\left\{\left[\sum_{p\in\mathcal{P}_{ij}}\left(h_p^{\tau,k} - \alpha\Psi_p\left(t,h^{\tau,k}\right) - h_p\right)\right](-1) - v_{ij} = 0\right\} \qquad p\in\mathcal{P}\end{aligned}$$

from which we obtain

$$\sum_{p\in\mathcal{P}_{ij}}\left(-h_p^{\tau,k} + \alpha\Psi_p\left(t,h^{\tau,k}\right) + h_p^{\tau,k+1}\right) - v_{ij} = 0 \quad \forall\left(i,j\right)\in\mathcal{W}, p\in\mathcal{P}_{ij}$$

or

$$\sum_{p\in\mathcal{P}_{ij}} h_p^{\tau,k+1} = \left[\sum_{p\in\mathcal{P}_{ij}}\left(h_p^{\tau,k} - \alpha\Psi_p\left(t,h^{\tau,k}\right)\right) + v_{ij}\right]_+ \qquad \forall\left(i,j\right)\in\mathcal{W}, p\in\mathcal{P}_{ij} \qquad (63)$$

where $[\cdot]_+$, the elementary projection operator defined by

$$[u]_+ = \begin{cases} u & if \quad u \geq 0 \\ 0 & if \quad u < 0 \end{cases},$$

is employed to assure non-negativity of path flows. By flow conservation we have

$$\int_{(\tau-1)\cdot\Delta}^{\tau\cdot\Delta} \sum_{p\in\mathcal{P}_{ij}} h_p^{\tau,k+1}(t)\,dt = Q_{ij}^{\tau} \qquad \forall (i,j) \in \mathcal{W}$$

Consequently the dual variable v_{ij} must satisfy

$$\int_{(\tau-1)\cdot\Delta}^{\tau\cdot\Delta} \sum_{p\in\mathcal{P}_{ij}} \left[h_p^{\tau,k}(t) - \alpha\Psi_p\left(t, h^{\tau,k}\right) + v_{ij}\right]_+ dt = Q_{ij}^{\tau} \qquad \forall (i,j) \in \mathcal{W} \tag{64}$$

Recalling that each v_{ij} is time invariant and noting that the equations of (64) are uncoupled, we see that simple line searches will find the values of v_{ij} satisfying the above conditions. Once the v_{ij} satisfying (64) have been found, the new path flow vector $h^{\tau,k+1}$ may be computed from (63).

7 Dynamic Network Loading

The problem of finding link activity when travel demand and departure rates (path flows) are known is commonly referred to as the dynamic network loading problem. Effective path delays are constructed from arc delays that, directly or indirectly, depend on arc activity; moreover, activity on a given arc is influenced by the delays on paths that utilize that arc. Thus, dynamic network loading is intertwined with the determination of path delays. Recall that, in our formulation of and computational scheme for the dual-time-scale model presented above, we require the effective path delay operators to be are measurable, positive and strongly pseudomonotone. Since positive, strongly pseudomonotone delay operators may reflect the undulating character of the example given in Section 6.3, our formulation and solution method can accommodate path delay operators derived from a range of dynamic network loading models.

In the balance of this section we present a computable version of a dynamic network loading model proposed first by Friesz et al. (1993) and studied numerically by Wu et al. (1998). We introduce this dynamic network loading model to illustrate a simple scheme for approximate network loading that relies on off-the-shelf software.

7.1 The Structure of Network Loading Models

A dynamic network loading model captures the following three critical features of dynamic traffic assignment modeling:

1. path delay;
2. flow dynamics; and
3. flow propagation constraints.

As we saw in Section 3 flow propagation constraints must be imposed on the state dynamics, thereby creating a system of constrained ordinary differential equations (ODEs). The flow propagation constraints will most generally be in the form of a differential algebraic equation (DAE). Since an ordinary differential equation is a special case of a differential algebraic equation, the dynamic network loading model may itself be viewed as a system of differential algebraic equations.

More formally we note that a fairly general abstract differential algebraic equation is

$$f(x, \dot{x}, y, t) = 0 \tag{65}$$

where for any given instant of time t we have

$$\begin{array}{rcl} x, \dot{x} & \in & \Re^n \\ y & \in & \Re^m \\ f & : & G \subseteq \Re^n \times \Re^n \times \Re^m \times \Re^1 \longrightarrow \Re^{n+m} \end{array}$$

A natural approach to solving system (65) is to differentiate it as needed to create a system of ordinary differential equations; however, upon doing so, one is faced with finding a consistent set of initial conditions $(x_0, \dot{x}_0, y_0, t_0)$ such that

$$f(x_0, \dot{x}_0, y_0, t_0) = 0$$

Simple examples may be given that illustrate finding a consistent set of initial conditions can be problematic. In fact Pantelides (1988) provides such examples as well as a criterion for determining whether differentiation of a subset of the equations of a nonlinear DAE system provides additional constraints satisfied by the available initial conditions; he also provides a graph-theoretic algorithm for finding such subsets. See also Brenan et al. (1996), Ascher and Petzold (1998) and Pang and Stewart (2008) for more details regarding DAEs.

We observe in this section that the flow propagation constraints (9) and (10) derived from the link delay model, together with a model of arc dynamics, constitute a DAE system that describes network loading. We show that such a DAE system for network loading may be accurately approximated by a system of first order ordinary differential equations and efficiently solved by off-the-shelf software, thereby providing a much more computationally appealing method for network loading based on the link delay model than has previously been available.

7.2 The Link Delay Model

We of course describe arc volume as the sum of volumes associated with individual paths using an arc; that is

$$x_a(t) = \sum_{p \in \mathcal{P}} \delta_{ap} x_a^p(t) \qquad \forall a \in \mathcal{A}$$

where $x_a^p(t)$ denotes the volume on arc a associated with path p and

$$\delta_{ap} = \begin{cases} 1 \text{ if arc } a \text{ belongs to path } p \\ 0 \text{ otherwise} \end{cases}$$

We employ the simple deterministic link delay model suggested by Friesz et al. (1993). To articulate this delay model, let the time to traverse arc a_i for drivers who arrive at its tail node at time t be denoted by $D_{a_i}[x_{a_i}(t)]$. That is, the time to traverse arc a_i depends only on the number of vehicles in front of the entering vehicle at its time of entry. The recursive relationships (4) and (5) follow immediately once the link delay model is articulated .

7.3 The DAE System

It is well known that application of the chain rule to (4) and (5) allows one to derive the flow propagation constraints (9) and (10) from (4) and (5); see for example Friesz et al. (2001). Knowledge of the flow propagation constraints allows us to articulate the following differential algebraic equation (DAE) system describing dynamic network loading:

$$\frac{dx_{a_i}^p(t)}{dt} = g_{a_{i-1}}^p(t) - g_{a_i}^p(t) \qquad \forall\, p \in \mathcal{P}, i \in [1, m(p)] \tag{66}$$

$$x_{a_i}^p(0) = x_{a_i}^{p,0} \in \Re_+^1 \qquad \forall\, p \in \mathcal{P}, i \in [1, m(p)] \tag{67}$$

$$h_p^{\tau,k}(t) = g_{a_1}(t + D_{a_1}[x_{a_1}(t)]) \left(1 + D'_{a_1}[x_{a_1}(t)]\, \dot{x}_{a_1}\right) \tag{68}$$

$$g^p_{a_{i-1}}(t) = g^p_{a_i}(t + D_{a_i}[x_{a_i}(t)]) \left(1 + D'_{a_i}[x_{a_i}(t)]\,\dot{x}_{a_i}(t)\right) \qquad \forall\ p \in \mathcal{P},\ i \in [2, m(p)] \tag{69}$$

where $g^p_{a_i}(t)$ is the flow along path p that exits arc a_i at time t; by convention

$$g^{p,\tau,k}_{a_0} = h^{\tau,k}_p \qquad \forall\ p \in \mathcal{P}$$

is the departure rate from the origin of path $p \in \mathcal{P}$.

7.4 A Simplified Network Loading Procedure

The main difficulty in computing solutions to the DAE system (66), (67), (68) and (69) arises from the time shifts appearing in the flow propagation constraints. In preparation for dealing with this feature of the DAE system and creating an approximate network loading scheme, we introduce the following definition:

Definition 4 *Regularity of the dynamic network loading problem based on the link delay model. The DAE system (66), (67), (68) and (69) is regular if the exit flow functions $g^p_{a_i}(t)$ for all $p \in \mathcal{P}$ and $i \in [1, m(p)]$ are everywhere twice continuously differentiable with respect to their own arguments.*

Invoking regularity in the sense of Definition 4, we note that a second order Taylor series approximation of the time shifted term of the flow propagation constraints yields

$$g^p_{a_1}(t + D_{a_1}[x_{a_1}(t)]) \approx g^p_{a_1}(t) + \frac{dg^p_{a_1}(t)}{dt} D_{a_1}[x_{a_1}(t)] + \frac{d^2 g^p_{a_1}(t)}{dt^2} \frac{(D_{a_1}[x_{a_1}(t)])^2}{2} \qquad \forall\ p \in \mathcal{P}$$

$$g^p_{a_i}(t + D_{a_i}[x_{a_i}(t)]) \approx g^p_{a_i}(t) + \frac{dg^p_{a_i}(t)}{dt} D_{a_i}[x_{a_i}(t)] + \frac{d^2 g^p_{a_i}(t)}{dt^2} \frac{(D_{a_i}[x_{a_i}(t)])^2}{2} \qquad \forall\ p \in \mathcal{P}, i \in [2, m(p)]$$

Furthermore, if we let

$$\frac{dg^p_{a_i}(t)}{dt} \equiv r^p_{a_i}(t) \qquad \forall\ p \in \mathcal{P}, i \in [1, m(p)]$$

so that

$$\frac{dr^p_{a_i}(t)}{dt} = \frac{d^2 g^p_{a_i}(t)}{dt^2} \qquad \forall\ p \in \mathcal{P}, i \in [1, m(p)]\ ,$$

then the second order approximations may be written as

$$g^p_{a_1}(t + D_{a_1}[x_{a_1}(t)]) \approx g^p_{a_1}(t) + r^p_{a_1}(t)\, D_{a_1}[x_{a_1}(t)] + \frac{dr^p_{a_1}(t)}{dt} \frac{(D_{a_1}[x_{a_1}(t)])^2}{2} \qquad \forall p \in \mathcal{P}$$

$$g^p_{a_i}(t + D_{a_i}[x_{a_i}(t)]) \approx g^p_{a_i}(t) + r^p_{a_i}(t)\, D_{a_i}[x_{a_i}(t)] + \frac{dr^p_{a_i}(t)}{dt} \frac{(D_{a_i}[x_{a_i}(t)])^2}{2} \qquad \forall\ p \in \mathcal{P}, i \in [2, m(p)]$$

Based on the above, the flow propagation constraints themselves become

$$\left\{ g^p_{a_1}(t) + r^p_{a_1}(t)\, D_{a_1}[x_{a_1}(t)] + \frac{dr^p_{a_1}(t)}{dt} \frac{(D_{a_1}[x_{a_1}(t)])^2}{2} \right\} \cdot \left(1 + D'_{a_1}[x_{a_1}(t)]\,\dot{x}_{a_1}\right) - h^{\tau,k}_p(t) = 0 \qquad \forall p \in \mathcal{P}$$

$$\left\{ g^p_{a_i}(t) + r^p_{a_i}(t)\, D_{a_i}\left[x_{a_i}(t)\right] + \frac{dr^p_{a_i}(t)}{dt}\frac{\left(D_{a_i}\left[x_{a_i}(t)\right]\right)^2}{2}\right\} \cdot \left(1 + D'_{a_i}\left[x_{a_i}(t)\right]\dot{x}_{a_i}(t)\right) - g^p_{a_{i-1}}(t) = 0 \quad \forall\ p \in \mathcal{P}, i \in [2, m(p)]$$

From the above we obtain

$$\begin{aligned}
\frac{dr^p_{a_1}(t)}{dt} &= R^p_{a_1}\left(x, g, r, h^{\tau,k}\right) \qquad \forall\ p \in \mathcal{P} \\
\frac{dr^p_{a_i}(t)}{dt} &= R^p_{a_i}(x, g, r) \qquad \forall\ p \in \mathcal{P}, i \in [2, m(p)] \\
\frac{dg^p_{a_i}(t)}{dt} &= r^p_{a_i}(t) \qquad \forall\ p \in \mathcal{P}, i \in [1, m(p)]
\end{aligned}$$

where for all $p \in \mathcal{P}$ and $i \in [1, m(p)]$ we use the definitions

$$\begin{aligned}
R^p_{a_1}\left(x, g, r, h^{\tau,k}\right) &\equiv \frac{2h^{\tau,k}_p(t)}{\left(D_{a_1}\left[x_{a_1}(t)\right]\right)^2\left(1 + D'_{a_1}\left[x_{a_1}(t)\right]\dot{x}_{a_1}\right)} - \frac{2\left(g^p_{a_1}(t) + r^p_{a_1}(t)\, D_{a_1}\left[x_{a_1}(t)\right]\right)}{\left(D_{a_1}\left[x_{a_1}(t)\right]\right)^2} \\
R^p_{a_i}(x, g, r) &\equiv \frac{2g^p_{a_{i-1}}(t)}{\left(D_{a_i}\left[x_{a_i}(t)\right]\right)^2\left(1 + D'_{a_i}\left[x_{a_i}(t)\right]\dot{x}_{a_i}\right)} - \frac{2\left(g^p_{a_i}(t) + r^p_{a_i}(t)\, D_{a_i}\left[x_{a_i}(t)\right]\right)}{\left(D_{a_i}\left[x_{a_i}(t)\right]\right)^2}
\end{aligned}$$

The remaining time derivatives $\dot{x}_{a_i}$ for $i \in [1, m(p)]$ may be replaced by expressions involving arc exit flows; we suppress those details for the sake of brevity. Note also that there are accompanying initial conditions

$$g^p_{a_i}((\tau - 1) \cdot \Delta) = 0 \qquad \forall\ p \in \mathcal{P}, i \in [1, m(p)]$$

If there is no flow entering the paths at the first instant of the time horizon, then we impose the following additional initial conditions

$$r^p_{a_i}((\tau - 1) \cdot \Delta) = 0 \qquad \forall\ p \in \mathcal{P}, i \in [1, m(p)]$$

Thus, to find the arc volumes and arc exit flows for a given vector of departure rates $h^{\tau,k}$, one may solve the following system of ordinary differential equations with initial conditions:

$$\begin{aligned}
\frac{dx^p_{a_1}(t)}{dt} &= h^{\tau,k}_p(t) - g^p_{a_1}(t) \qquad \forall p \in \mathcal{P} && (70) \\
\frac{dx^p_{a_i}(t)}{dt} &= g^p_{a_{i-1}}(t) - g^p_{a_i}(t) \qquad \forall\ p \in \mathcal{P}, i \in [2, m(p)] && (71) \\
\frac{dg^p_{a_i}(t)}{dt} &= r^p_{a_i}(t) \qquad \forall p \in \mathcal{P}, i \in [1, m(p)] && (72) \\
\frac{dr^p_{a_1}(t)}{dt} &= R^p_{a_1}\left(x, g, r, h^{\tau,k}\right) \qquad \forall\ p \in \mathcal{P} && (73) \\
\frac{dr^p_{a_i}(t)}{dt} &= R^p_{a_i}(x, g, r) \qquad \forall\ p \in \mathcal{P}, i \in [2, m(p)] && (74) \\
x^p_{a_i}((\tau - 1) \cdot \Delta) &= x^{p,0}_{a_i} \qquad \forall\ p \in \mathcal{P}, i \in [1, m(p)] && (75) \\
g^p_{a_i}((\tau - 1) \cdot \Delta) &= 0 \qquad \forall\ p \in \mathcal{P}, i \in [1, m(p)] && (76) \\
r^p_{a_i}((\tau - 1) \cdot \Delta) &= 0 \qquad \forall\ p \in \mathcal{P}, i \in [1, m(p)] && (77)
\end{aligned}$$

Definition 5 *Regularity of the simplified dynamic network loading procedure. The simplified dynamic network loading problem expressed by the initial value problem (70) through (77) is regular if, for all* $p \in \mathcal{P}, \ i \in [1, m(p)]$ *and* $t \in [(\tau-1)\cdot\Delta, \tau\cdot\Delta]$, *the following conditions are satisfied:*

1. *the arc exit time functions* $\xi^p_{a_i}$ *are strictly monotonic (and thus invertible);*
2. *the arc delay functions* $D_{a_i}[x_{a_i}(t)]$ *are bounded for bounded arguments;*
3. $D_{a_i}[x_{a_i}(t)] > 0$ *;*
4. $D''_{a_i}[x_{a_i}(t)]$ *exists and is continuous; and*
5. $h^{\tau,k}_p(t)$ *is continuous*

We will need the following results:

Lemma 1 *For a domain* $\mathcal{D} \subset \Re^{n+1}$, *consider a vector function* $f : \mathcal{D} \longrightarrow \Re^n$. *If* f *is continuous in* $\mathcal{D}$ *and all components of the Jacobian*

$$\frac{\partial f}{\partial x} = \left(\frac{\partial f_i}{\partial x_j}\right)_{i,j=1,\dots,n}$$

are continuous in $\mathcal{D}$, *then the vector function* f *satisfies a local Lipschitz condition.*

Proof. See Walter (1988). ■

Theorem 6 *Consider a vector* $x \in \Re^n$ *and a vector function* $f : \mathcal{D} \longrightarrow \Re^n$ *for a domain* $\mathcal{D} \subset \Re^{n+1}$ *and the system*

$$\begin{aligned} \frac{dx_1}{dt} &= f_1(t, x_1, ..., x_n) \\ &\vdots \\ \frac{dx_n}{dt} &= f_n(t, x_1, ..., x_n) \end{aligned}$$

with initial conditions

$$x_i(t_0) = x_i^0 \qquad i \in [1, n]$$

If the vector function f *satisfies a local Lipschitz condition with respect to the vector* x, *the system has exactly one solution.*

Proof. See Walter (1988). ■

Next we establish existence and uniqueness for the simplified dynamic network loading initial value problem:

Theorem 7 *Existence and uniqueness of the simplified dynamic network loading. If the simplified dynamic network loading procedure expressed by the initial value problem (70) through (77) is regular in the sense of Definition 5, it has an unique solution.*

Proof. If the right hand sides of (70) through (77) are continuous and continuously differentiable with respect to $x_{a_i}, g^p_{a_i}$ and $r^p_{a_i}$, then we have existence and uniqueness of by Lemma 1 and Theorem 6. In fact closed formed expressions for the derivatives of $R^p_{a_1}$ may be given:

$$\begin{aligned} \frac{\partial R^p_{a_1}}{\partial x_{a_1}} &= X^{k,p,\tau}_{a_1} - \frac{2r^p_{a_1}(t)\,D'_{a_1}[x_{a_1}(t)]}{\left(D_{a_1}[x_{a_1}(t)]\right)^2} \\ &\quad + \frac{4\left(g^p_{a_1}(t) + r^p_{a_1}(t)\,D_{a_1}[x_{a_1}(t)]\right)D_{a_1}[x_{a_1}(t)]\,D'_{a_1}[x_{a_1}(t)]}{\left(D_{a_1}[x_{a_1}(t)]\right)^4} \end{aligned}$$

$$\frac{\partial R_{a_1}^p}{\partial g_{a_1}^p} = \frac{2h_p^{\tau,k}(t)}{(D_{a_1}[x_{a_1}(t)])^2} \frac{D'_{a_1}[x_{a_1}(t)]}{\left(1 + D'_{a_1}[x_{a_1}(t)]\left(h_p^{\tau,k}(t) - g_{a_1}^p(t)\right)\right)^2} - \frac{2g_{a_1}^p(t)}{(D_{a_1}[x_{a_1}(t)])^2}$$

$$\frac{\partial R_{a_1}^p}{\partial r_{a_1}^p} = -\frac{2}{D_{a_1}[x_{a_1}(t)]}$$

where

$$X_{a_1}^{k,p,\tau} = \frac{-2h_p^{\tau,k}(t)\left(Y_{a_1}^{k,p,\tau} + Z_{a_1}^{k,p,\tau}\right)}{(D_{a_1}[x_{a_1}(t)])^4\left(1 + D'_{a_1}[x_{a_1}(t)]\left(h_p^{\tau,k}(t) - g_{a_1}^p(t)\right)\right)^2}$$

$$Y_{a_1}^{k,p,\tau} = 2D_{a_1}[x_{a_1}(t)]\,D'_{a_1}[x_{a_1}(t)]\left(1 + D'_{a_1}[x_{a_1}(t)]\left(h_p^{\tau,k}(t) - g_{a_1}^p(t)\right)\right)$$

$$Z_{a_1}^{k,p,\tau} = (D_{a_1}[x_{a_1}(t)])^2\left(D''_{a_1}[x_{a_1}(t)]\left(h_p^{\tau,k}(t) - g_{a_1}^p(t)\right)\right)$$

By inspection, these are continuous under regularity in the sense of Definition 5. Similar results may be established for $i \in [2, m(p)]$ where $p \in \mathcal{P}$ is an arbitrary path. Thus, the desired result is proven. ■

7.5 Constructing the Path Delay for a Given $h^{\tau,k}$

An immediate consequence of the recursive relationships (4) and (5) is that the total traversal time for path p can be articulated in terms of the final exit time function and the departure time:

$$D_p = \sum_{i=1}^{m(p)}\left[\xi_{a_i}^p(t) - \xi_{a_{i-1}}^p(t)\right] = \xi_{a_{m(p)}}^p(t) - t \qquad \forall p \in \mathcal{P}$$

where $\xi_{a_i}^p(t)$ is the time of exit from arc $i \in [1, m(p)]$ for path $p \in \mathcal{P}$ given departure from the origin at time t. We have already introduced the potentially asymmetric arrival penalty operator $F(\cdot)$ with property (13) so that the effective delay operator is

$$\Psi_p = D_p + F[t + D_p - T_A^\tau] \qquad \forall p \in \mathcal{P}$$

where T_A^τ is the desired arrival time for day τ. If the path flows $h^{\tau,k}$ are known, it is possible to find the arc exit flows, volumes and delays by solving the system of ordinary differential equations (70) through (77) over the time horizon of interest. Let us denote the traffic volumes from solution of that system by

$$x^{k,\tau} = \left(x_{a_i}^{p,k,\tau} : p \in \mathcal{P},\ i \in [1, m(p)]\right)$$

and define

$$x_a^{k,\tau}(t) = \sum_{p\in\mathcal{P}} \delta_{ap} x_a^{p,k,\tau}(t) \qquad \forall a \in \mathcal{A}$$

The arc exit time functions may be computed for a path by first noting that

$$\xi_{a_1}^{p,k,\tau}(t) = t + D_{a_1}\left[x_{a_1}^{k,\tau}(t)\right]$$

where t is the departure time. Once the arc exit time function of the first arc has been computed, the arc exit time function for the next arc in the path may be computed as

$$\xi_{a_2}^{p,k,\tau} = \xi_{a_1}^{p,k,\tau}(t) + D_{a_2}\left[x_{a_2}\left(\xi_{a_1}^{p,k,\tau}(t)\right)\right]$$

and so forth until the arc exit times of all arcs have been computed. This procedure is carried out for each path $p \in \mathcal{P}$. When the arc exit time functions $\xi_{a_i}^{p,k,\tau}$ are known for all $p \in \mathcal{P}$ and $i \in [1, m(p)]$, the effective path delays may be computed as pure functions of time following

$$\Psi_p^{k,\tau}(t) = \xi_{a_{m(p)}}^{p,k,\tau,}(t) - t + F\left[\xi_{a_{m(p)}}^{p,k,\tau} - T_A^\tau\right]$$

where it will be recalled that $F[.]$ is the arrival penalty and T_A^τ is the desired time of arrival.

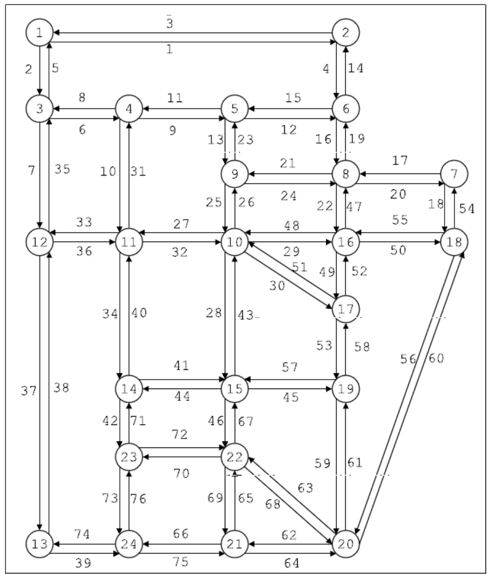


Figure 1: Sioux Falls Network

8 Numerical Example

In this section, we consider a numerical example based on the 76-arc, 20-node Sioux Falls network with 10 origin-destination (OD) pairs and 200 paths, depicted in Figure 1. We consider network loading based on two link performance models:

1. link delay model (LDM) approximated as a system of ordinary differential equations; and
2. the cell-transmission model (CTM) as implemented by Lo and Szeto (2002).

A common early/late arrival penalty function is used for all examples and is assumed to be of the form

$$0.5\left(t+D_p\left(t\right)-T_A\right)^2$$

where T_A is the desired arrival time for all users. For the above arrival penalty, both the single period within-day model and the dual-time-scale model are solved; in the latter case, dynamic user equilibria extending over a multiday period when demand evolves day-to-day according to the dynamics (29) are calculated. As already noted, for LDM-based network loading, we employed an ODE approximation of a DAE system. For the CTM, we employed the so-called route travel time extraction procedure presented in Lo and Szeto (2002), which is not a DAE system. Coding was done using MATLAB, and every attempt was made to avoid the use of numerical tricks that could not be applied equally to all models. For the numerical results presented below, DUE solutions obtained with the LDM and CTM link performance models display qualitative similarity: the solutions are always moving parabolic wave forms, showing the progression of flow through the network with departure rate peaks that correspond to minimum effective path delays. Zhang and Nie (2005) suggest that such similarity of solutions across loading models is due to the lack of spillback phenomena; numerical experiments supporting their observation were not conducted by us.

For the Sioux Falls network, we considered the following set of ten specific origin-destination pairs:

$$\mathcal{W}=\left\{(1,20)\,,\ \ (2,20)\,,\ \ (4,20)\,,\ \ (5,20)\,,\ \ (7,20)\,,\ \ (9,20)\,,\ \ (11,20)\,,\ \ (13,20)\,,\ \ (15,20)\,,\ \text{and}\ \ (16,20)\right\}$$

each having a fixed travel demand of 400 users; that is

$$Q_{ij}=400\quad\forall\left(i,j\right)\in\mathcal{W}$$

Linear arc delay functions of the form

$$D_a = A_a + B_a x_a$$

were employed; parameters for the delay functions are given in Table 1, where (k, ℓ) denotes an arc directed from node k to node ℓ. Furthermore

$$B_a = 0.007$$

for all arcs.

(k,ℓ)	A_a	(k,ℓ)	A_a	(k,ℓ)	A_a	(k,ℓ)	A_a
(1,3)	2	(3,1)	2	(2,6)	2	(6,2)	2
(3,4)	2	(4,3)	2	(4,5)	2	(5,4)	2
(5,6)	10	(6,5)	10	(3,12)	3	(12,3)	3
(4,11)	3	(11,4)	3	(5,9)	2	(9,5)	2
(6,8)	2	(8,6)	2	(9,10)	2	(10,9)	2
(1,2)	6	(2,1)	6	(9,8)	4	(8,9)	4
(8,7)	2	(7,8)	2	(8,16)	2	(16,8)	2
(7,18)	4	(18,7)	4	(12,11)	4	(11,12)	4
(11,10)	3	(10,11)	3	(10,16)	2	(16,10)	2
(16,18)	2	(18,16)	2	(12,13)	14	(13,12)	14
(11,14)	3	(14,11)	3	(10,15)	4	(15,10)	4
(10,17)	6	(17,10)	6	(16,17)	2	(17,16)	2
(17,19)	6	(19,17)	6	(18,20)	16	(20,18)	16
(14,15)	4	(15,14)	4	(15,19)	2	(19,15)	2
(14,23)	2	(23,14)	2	(15,22)	2	(22,15)	2
(23,22)	2	(22,23)	2	(23,24)	2	(24,23)	2
(22,21)	8	(21,22)	8	(22,20)	6	(20,22)	6
(19,20)	4	(20,19)	4	(13,24)	4	(24,13)	4
(24,21)	4	(21,24)	4	(21,20)	4	(20,21)	4

Table 1: Delay Function Parameters

The DUE solution obtained by the fixed point algorithm is illustrated in Figures 2 and 3, where we observe a moving parabolic wave. Two representative DUE path flows (departure rates) are displayed in Figure 2. The volumes of six representative arcs are displayed in Figure 3.

The dual-time-scale model is solved using the following parameters for the day-to-day demand dynamics:

$$\begin{array}{ll} \chi_{1,20} = 750 & \chi_{9,20} = 750 \\ \chi_{2,20} = 450 & \chi_{11,20} = 1050 \\ \chi_{4,20} = 650 & \chi_{13,20} = 1100 \\ \chi_{5,20} = 750 & \chi_{15,20} = 1150 \\ \chi_{7,20} = 800 & \chi_{16,20} = 1200 \end{array}$$

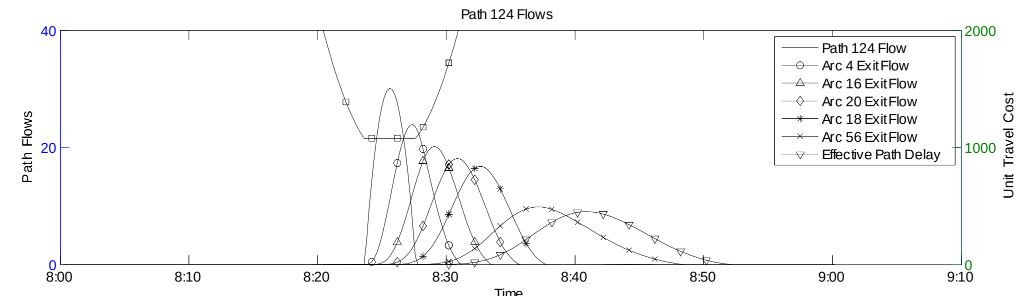
$$Q_{ij}^1 = 400 \qquad \forall\,(i,j) \in \mathcal{W}$$

$$s_{ij}^\tau = 0.1 \qquad \forall\,(i,j) \in \mathcal{W} \qquad \tau \in \Upsilon \equiv \{1, 2, ..., N\}$$

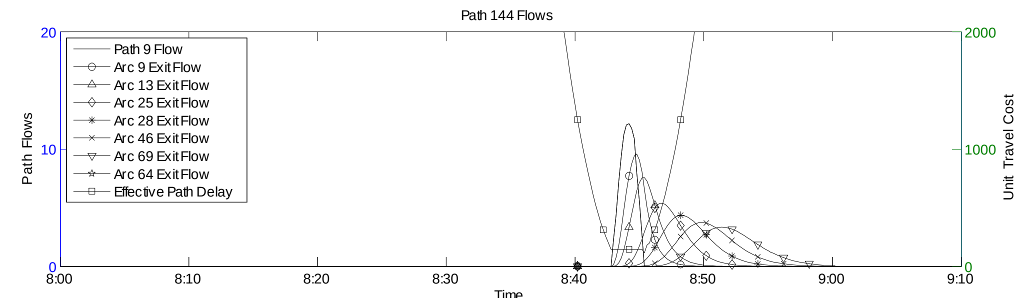
Using the same desired arrival times as used for the single day simulation of this network, results for the seven day simulation based on the dual-time-scale model are presented in Figure 4.

8.1 Performance of the Fixed Point Algorithm

The numerical examples were solved by the continuous time DUE fixed point algorithm, supported by the simplified DNL procedure developed in Section 7. The combined DUE/LDM-DNL solution approach was coded in MATLAB 7 and solved on a standard desktop computer with the following attributes: Windows Vista with an Intel Core2 Duo @ 2.20GHz and 1.5GB RAM. In these calculations, the day-to-day time

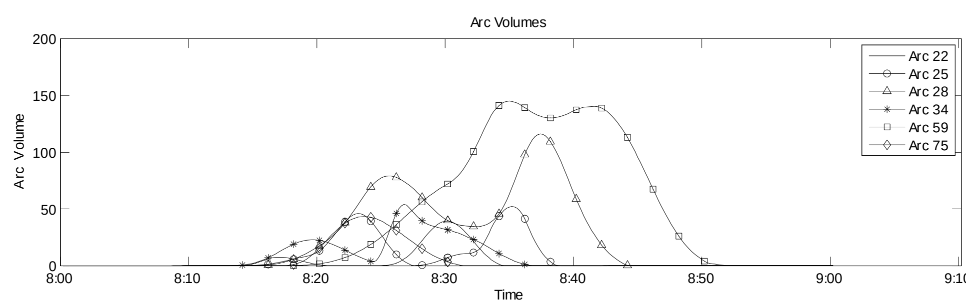


LDM

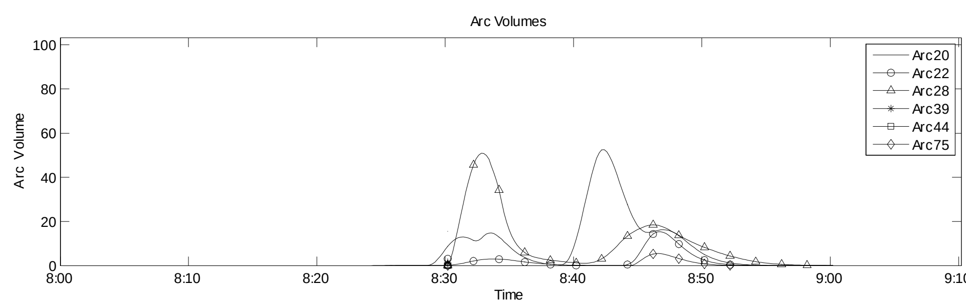


CTM

Figure 2: DUE Flows and Effective Delays for LDM and CTM Based Network Loading (Sioux Falls network)



LDM



CTM

Figure 3: DUE Arc Volumes for LDM and CTM Based Network Loading (Sioux Falls network)

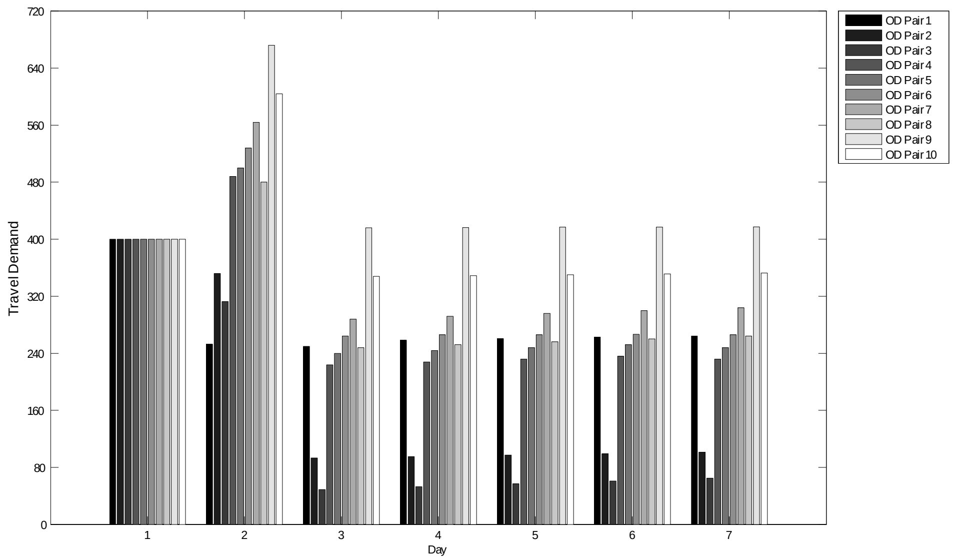


Figure 4: 7 Day Travel Demand Evolution for 10 OD pairs (Large network)

stepping to update demand is extremely fast. Moreover, the within-day dynamic user equilibrium calculations for the second and subsequent days are also extremely fast since a warm-start protocol is employed whereby a trial solution based on the previous within-day flow pattern is used. Thus, the most important measure of computational performance is the number major fixed point iterations for an initial (cold-start) within-day user equilibrium calculation. We report the number of such iterations in Table 2.

Sioux Falls Network	**LDM**	**CTM**
major iterations	14	13
cpu seconds	1975	3136

Table 2: DUE Fixed-Point Major Iterations

9 Concluding Remarks

In this paper we have explored a differential variational inequality (DVI) formulation of within-day dynamic user equilibrium and its integration with a day-to-day model of travel demand evolution. The following have been demonstrated/accomplished:

1. the differential variational inequality may be very easily analyzed to establish its equivalence to a dynamic user equilibrium using the minimum principle from optimal control theory;

2. its dual variables for terminal time constraints are easily computed and equal the minimum effective travel delay;

3. the proposed fixed point algorithm is shown to converge for merely component-wise strongly pseudomonotone delay operators that are also component-wise weakly monotone; such operators form a class of non-monotone operators;

4. the rapidly growing literature on differential variational inequalities offers a rich collection of ideas for improving the proposed fixed point algorithm and for innovating alternative algorithms;[4]

5. a dual-time-scale model of dynamic network traffic flows may be constructed that integrates a basic day-to-day demand evolution model with a differential variational inequality formulation of within-day dynamic user equilibrium;

6. simple time stepping in combination with the within-day fixed point algorithm has been shown to be an effective means of solving the dual-time-scale model;

7. although LDM-based network loading has been emphasized, through numerical examples, the proposed fixed point algorithm has been shown to also work effectively with CTM-based network loading; and

8. an approximation method for LDM-based network loading has been presented and shown to be at least as computationally efficient as CTM-based network loading for the computation of dynamic user equilibria with the proposed fixed point algorithm.

Several extensions of the ideas put forward in this paper are possible. These include;

1. explicit consideration of spillbacks and oscillatory traffic behavior, as noted in Nie and Zhang (2008b) and Nie (2010); and

2. establishing error bounds associated with approximate network loading.

References

Ascher, U., Petzold, L., 1998. Computer Methods for Ordinary Differential Equations and Differential-Algebraic Equations. SIAM Publications, Philadelphia.

Astarita, V., 1995. Flow propagation description in dynamic network loading models. Y.J Stephanedes, F. Filippi, (Eds). Proceedings of IV international conference of Applications of Advanced Technology in Transportation(AATT), 599–603.

Ban, X., Liu, H., Ferris, M., Ran, B., 2008. A link-node complementarity model and solution algorithm for dynamic user equilibria with exact flow propagations. Transportation Research Part B 42 (9), 823–842.

Bauschke, H., 1996. The approximation of fixed-points of compositions of nonexpansive mappings in hilbert space. Journal of Mathematical Analysis and Applications 202 (1), 150–159.

Bliemer, M., Bovy, P., 2003. Quasi-variational inequality formulation of the multiclass dynamic traffic assignment problem. Transportation Research Part B 37 (6), 501–519.

Brenan, K., Campbell, S., Petzold, L., 1996. Numerical Solution of Initial-Value Problems in Differential Algebraic Equations. Vol. 14 of Classics in Applied Mathematics. SIAM Publications, Philadelphia.

Carey, M., 1986. A constraint qualification for a dynamic traffic assignment problem. Transportation Science 20 (1), 55–58.

Carey, M., 1987. Optimal time-varying flows on congested networks. Operations Research 35 (1), 58–69.

Carey, M., 1992. Nonconvexity of the dynamic traffic assignment problem. Transportation Research Part B 26 (2), 127–132.

Carey, M., 1995. Dynamic congestion pricing and the price of fifo. In: Gartner, N. H., Improta, G. (Eds.), Urban Traffic Networks. Springer-Verlag, New York, pp. 335–350.

Carey, M., 2004a. Link travel times i: desirable properties. Networks and Spatial Economics 4 (3), 257–268.

[4]See Pang and Stewart (2008) for a review of differential variational inequality algorithms.

Carey, M., 2004b. Link travel times ii: properties derived from traffic-flow models. Networks and Spatial Economics 4 (4), 379–402.

Carey, M., McCartney, M., 2004. An exit-flow model used in dynamic traffic assignment. Computers and Operations Research 31 (10), 1583–1602.

Daganzo, C., 1994. The cell transmission model. Part I: A simple dynamic representation of highway traffic. Transportation Research Part B 28 (4), 269–287.

Daganzo, C., 1995. The cell transmission model. Part II: Network traffic. Transportation Research Part B 29 (2), 79–93.

Farouq, N. E., 2001. Pseudomonotone variational inequalities: convergence of proximal methods. Journal of Optimization Theory and Applications 109(2), 311 – 326.

Friesz, T., 2010. Dynamic Optimization and Differential Games. Springer.

Friesz, T., Bernstein, D., Suo, Z., Tobin, R., 2001. Dynamic network user equilibrium with state-dependent time lags. Networks and Spatial Economics 1(3/4), 319–347.

Friesz, T., Luque, J., Tobin, R., Wie, B., 1989. Dynamic network traffic assignment considered as a continuous-time optimal control problem. Operations Research 37 (6), 893–901.

Friesz, T., Mookherjee, R., 2006. Solving the dynamic network user equilibrium problem with state-dependent time shifts. Transportation Research Part B 40 (3), 207–229.

Friesz, T. L., Bernstein, D., Smith, T., Tobin, R., Wie, B., 1993. A variational inequality formulation of the dynamic network user equilibrium problem. Operations Research 41 (1), 80–91.

Halpern, B., 1967. Fixed points of nonexpanding maps. Bulletin of the American Mathematical Society 73 (6), 957–961.

Hofbauer, J., Sigmund, K., 1998. Evolutionary Games and Replicator Dynamics. Cambridge University Press, Cambridge.

Huang, H., Lam, W., 2002. Modeling and solving the dynamic user equilibrium route and departure time choice problem in network with queues. Transportation Research Part B 36 (3), 253–273.

Kachani, S., Perakis, G., 2001. Second-order fluid dynamics models for travel times in dynamic transportation networks. In: Proceedings of the 4th International IEEE Conference on Intelligent Transportation Systems, Oakland, California.

Kachani, S., Perakis, G., 2002. *Fluid dynamics models and their application in transportation and pricing.* Tech. rep., Sloan School of Management, MIT.

Kachani, S., Perakis, G., 2010. A dynamic travel time model for spillback. Networks and Spatial Economics 9 (4), 595–618.

Li, J., Fujiwara, O., Kawakami, S., 2000. A reactive dynamic user equilibrium model in network with queues. Transportation Research Part B 34 (8), 605–624.

Lighthill, M., Whitham, G., 1955. On kinematic waves. ii. a theory of traffic flow on long crowded roads. Proceedings of the Royal Society of London. Series A 229 (1178), 317–345.

Lo, H., Szeto, W., 2002. A cell-based variational inequality formulation of the dynamic user optimal assignment problem. Transportation Research Part B 36 (5), 421–443.

Merchant, D., Nemhauser, G., 1978a. A model and an algorithm for the dynamic traffic assignment problems. Transportation Science 12 (3), 183–199.

Merchant, D., Nemhauser, G., 1978b. Optimality conditions for a dynamic traffic assignment model. Transportation Science 12 (3), 200–207.

Nie, Y., 2010. Equilibrium analysis of macroscopic traffic oscillations. Transportation Research Part B 44 (1), 62–72.

Nie, Y., Zhang, H., 2008a. A variational inequality formulation for inferring dynamic origin-destination travel demands. Transportation Research Part B 42 (7-8), 635–662.

Nie, Y., Zhang, H. M., 2008b. Oscillatory traffic flow patterns induced by queue spillback in a simple road network. Transportation Science 44 (2), 236 – 248.

Nie, Y., Zhang, H. M., 2010. Solving the dynamic user optimal assignment problem considering queue spillback. Networks and Spatial Economics 10 (2), 1 – 23.

Pang, J.-S., Stewart, D., 2008. Differential variational inequalities. Mathematical Programming 113 (2), 345–424.

Pantelides, C., 1988. The consistent initialization of differential-algebraic systems. SIAM Journal on Scientific and Statistical Computing 9 (2), 213–231.

Peeta, S., Ziliaskopoulos, A., 2001. Foundations of dynamic traffic assignment: the past, the present and the future. Networks and Spatial Economics 1 (3), 233–265.

Perakis, G., 2000. *The dynamic user equilibrium problem through hydrodynamic theory.* Sloan School of Management, MIT, preprint.

Perakis, G., Roels, G., 2006. An analytical model for traffic delays and the dynamic user equilibrium problem. Operations Research 54 (6), 1151–1171.

Ran, B., Boyce, D., 1996. Modeling Dynamic Transportation Networks: An Intelligent Transportation System Oriented Approach. Springer-Verlag, New York.

Ran, B., Boyce, D., LeBlanc, L., 1993. A new class of instantaneous dynamic user optimal traffic assignment models. Operations Research 41 (1), 192–202.

Richards, P. I., 1956. Shock waves on the highway. Operations Research 4 (1), 42–51.

Samuelson, L., 1998. Evolutionary Games and Equilibrium Selection. MIT Press.

Szeto, W., Lo, H., 2004. A cell-based simultaneous route and departure time choice model with elastic demand. Transportation Research Part B 38 (7), 593–612.

Tong, C., Wong, S., 2000. A predictive dynamic traffic assignment model in congested capacity-constrained road networks. Transportation Research Part B 34 (8), 625–644.

Walter, W., 1988. Ordinary Differential Equations. Springer.

Wie, B., Tobin, R., Friesz, T., Bernstein, D., 1995. A discrete-time, nested cost operator approach to the dynamic network user equilibrium problem. Transportation Science 29 (1), 79–92.

Wu, J., Chen, Y., Florian, M., 1998. The continuous dynamic network loading problem: a mathematical formulation and solution method. Transportation Research Part B 32 (3), 173–187.

Xu, H.-K., 2003. Iterative algorithms for nonlinear operators. Journal of the London Mathematical Society 66 (1), 240–256.

Xu, Y., Wu, J., Florian, M., Marcotte, P., Zhu, D., 1999. Advances in the continuous dynamic network loading problem. Transportation Science 33(4), 341–353.

Zhang, H. M., Nie, Y., 2005. Modelling network flow with and without link interactions: Properties and implications. In: 84th annual meeting of the Transportation Research Board, Washington. p. 21.

Zhu, D. L., Marcotte, P., 2000. On the existence of solutions to the dynamic user equilibrium problem. Transportation Science 34 (4), 402–414.

A Component-Wise Strongly Pseudomonotone Functions That Are Also Component-Wise Weakly Monotone

Condition $\acute{R}5$ introduced in Section 6.3, the requirement that the path delay operator be component-wise strongly pseudomonotone, requires some elaboration since pseudomonotonicity has not been used heretofore in studying dynamic user equilibrium. To better understand $\acute{R}5$, it is helpful to consider a related problem without an explicit time dependence. We also need to discuss the component-wise weak monotonicity condition $\acute{R}6$ and show that both properties may coexist.

To these ends, let us consider the following abstract scalar function of a single variable:

$$F(x) = \frac{38}{100} + \frac{1}{2}(x - .2)^2 - \frac{1}{3}(x - .05)^3 \tag{A.1}$$

whose plot is given in Figure A.1. The function (A.1) is clearly not montonone on

$$X = \{x : 0 \leq x \leq 1.1\} \subset \Re^1_+ \tag{A.2}$$

However, as we will show, (A.1) is strongly pseudomonotone on X for some K_1, as well as weakly monotone

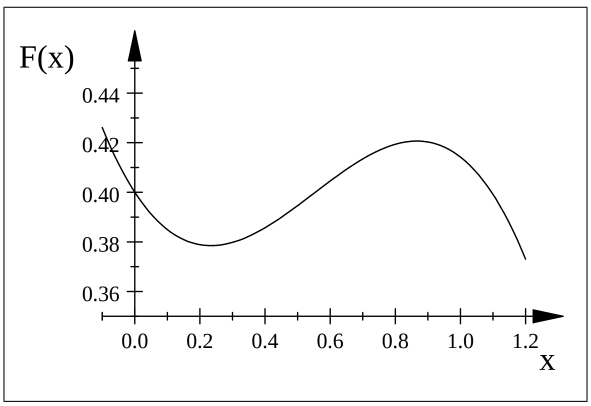


Figure A.1: Plot of the Example Function (A.1)

on X for some K_2.

A.1 Strong Pseudomonotonicity

For $F(x)$ to be strongly pseudomonotone, the constant $K_1 > 0$ must be such that

$$F(x_1)(x_2 - x_1) \geq 0 \Longrightarrow F(x_2)(x_2 - x_1) \geq \frac{K_1}{2}(x_2 - x_1)^2 \tag{A.3}$$

or equivalently

$$F(x_2)(x_2 - x_1) < \frac{K_1}{2}(x_2 - x_1)^2 \Longrightarrow F(x_1)(x_2 - x_1) < 0 \tag{A.4}$$

where

$$x_1, x_2 \in X \subset \Re^1_+$$

Note that because our function $F(x)$ is strictly positive in the first quadrant, (A.3) may only be employed when $x_2 \geq x_1$. For the same reason, (A.4) may only be employed when $x_2 < x_1$.

To illustrate that $F(x)$ is pseudomonotone on X, we consider the following sets:

$$S_1 = \left\{(x_1, x_2) : \left(\frac{38}{100} + \frac{1}{2}(x_1 - .2)^2 - \frac{1}{3}(x_1 - .05)^3\right) \cdot (x_2 - x_1) \geq 0\right\}$$

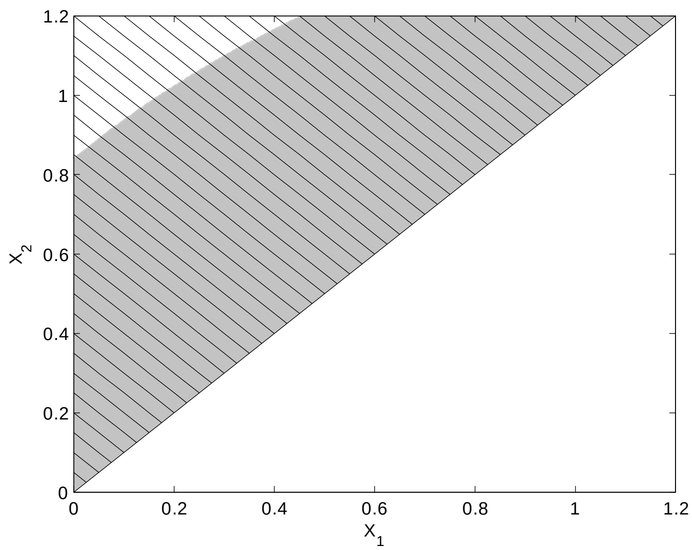


Figure A.2: $(S_2 \cap X) \subset (S_1 \cap X)$ for $K_1 = 1.000 \Longrightarrow SPM$ on X not demonstrated

$$S_2 = \left\{(x_1, x_2) : \left(\frac{38}{100} + \frac{1}{2}(x_2 - .2)^2 - \frac{1}{3}(x_2 - .05)^3\right) \cdot (x_2 - x_1) \geq \frac{K_1}{2}(x_2 - x_1)^2\right\}$$

$$S_3 = \left\{(x_1, x_2) : \left(\frac{38}{100} + \frac{1}{2}(x_2 - .2)^2 - \frac{1}{3}(x_2 - .05)^3\right) \cdot (x_2 - x_1) < \frac{K_1}{2}(x_2 - x_1)^2\right\}$$

$$S_4 = \left\{(x_1, x_2) : \left(\frac{38}{100} + \frac{1}{2}(x_1 - .2)^2 - \frac{1}{3}(x_1 - .05)^3\right) \cdot (x_2 - x_1) < 0\right\}$$

for specific values of $K_1 > 0$. The defining properties (A.3) and (A.4) are equivalent to

$$S_1 \subseteq S_2 \text{ for } x_2 \geq x_1$$
$$S_3 \subseteq S_4 \text{ for } x_2 < x_1$$

In the following plot for $K_1 = 1.000$, the black crosshatched region above the 45 degree line corresponds to S_1, and the grey region corresponds to S_2 in Figure A.2. As such, it is clear from the definition (A.2) of X that

$$(S_2 \cap Y) \subset (S_1 \cap Y)$$

where

$$Y = \{(x_1, x_2) \in X : x_2 \geq x_1\}$$

Hence, function (A.1) has not been demonstrated to be strongly pseudomonotone (SPM) on X by the choice of $K_1 = 1.0$. However, if we choose $K_1 = 0.730$, the sets S_1 (black crosshatch) and S_2 (solid grey) have a different depiction in Figure A.3, where we see that

$$(S_2 \cap Y) = (S_1 \cap Y)$$

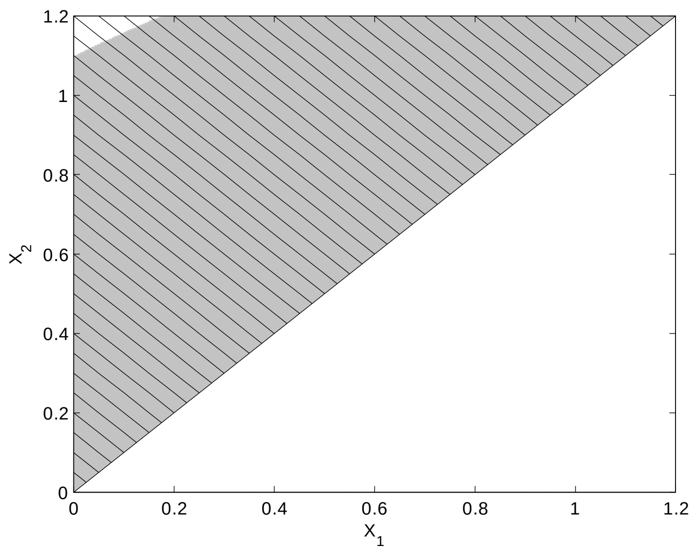


Figure A.3: $(S_2 \cap Y) = (S_1 \cap Y)$ for $K_1 = 0.730$

Moreover, for the same choice $K_1 = 0.730$, we find that the plots of S_3 and S_4 yield Figure A.4, where the black cross hatched area corresponds to S_3 and the grey area corresponds to S_4. Thus, we see that

$$(S_3 \cap Y^c) = (S_4 \cap Y^c)$$

where

$$Y^c = \text{the complement of } Y = \{x_1, x_2 \in X : x_2 < x_1\}$$

Hence, $F(x)$ is in fact strongly pseudomonotone on X, even though it is not monotone; at a superficial level, this result seems surprising, since $F(x)$ gives the visual impression of being "extremely non-monotone". The strictly positive nature of (A.1) on X is critical to it being strongly pseudomonotone. Furthermore, a close inspection of (A.3) reveals that strong pseudomonotonicity of a strictly positive scalar function may be informally described as a requirement that it does not change "too rapidly".

A.2 Weak Monotonicity

To show $F(x)$ is weakly monotone on X, we need to show that

$$\Big[F(x_1) - F(x_2)\Big]\,(x_1 - x_2) \geq (-1)\frac{K_2}{2}\,(x_1 - x_2)^2 \quad \forall x_1, x_2 \in X \tag{A.5}$$

where X is defined by (A.2). The relevant set is

$$S_5 = \left\{(x_1, x_2) : \Big[F(x_1) - F(x_2)\Big]\,(x_1 - x_2) \geq (-1)\frac{K_2}{2}\,(x_1 - x_2)^2 \quad \forall x_1, x_2 \in X\right\} \tag{A.6}$$

Based on (A.1), expression (A.5) is easily restated as

$$\left[-\frac{1}{3}\,(x_1)^3 + 0.55\,(x_1)^2 - 0.202\,5x_1 + \frac{1}{3}\,(x_2)^3 - 0.55\,(x_2)^2 + 0.202\,5x_2\right](x_1 - x_2) \geq (-1)\frac{K_2}{2}\,(x_1 - x_2)^2 \tag{A.7}$$

Using (A.6) together with (A.7), the set S_5 and its boundaries are depicted below in Figure A.5 for the value $K_2 = 0.500 < K_1$, from which it is apparent that $F(x)$ is weakly monotone (as well as strongly pseudomonotone) for all $x \in X$. Note that the 45-degree line is part of set of points that satisfy (A.7). By inspection we see that $X \subset S_5$; therefore, $F(x)$ is weakly monotone on X.

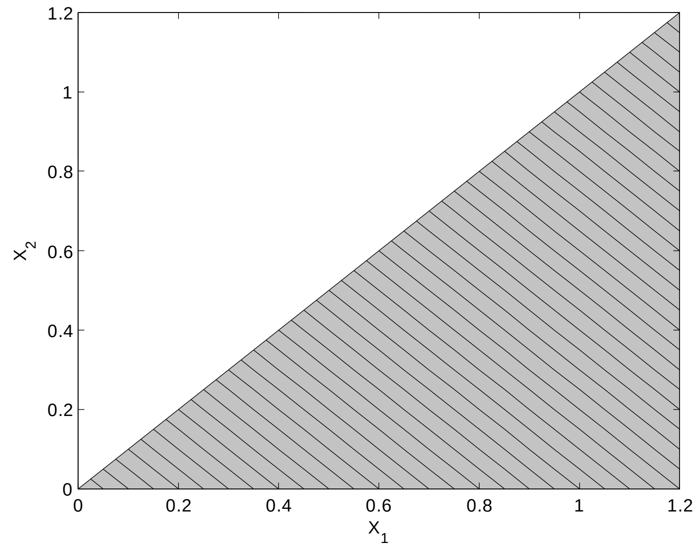


Figure A.4: $(S_3 \cap Y^c) = (S_4 \cap Y^c)$ for $K_1 = 0.730 \Longrightarrow F(x)$ is SPM on X

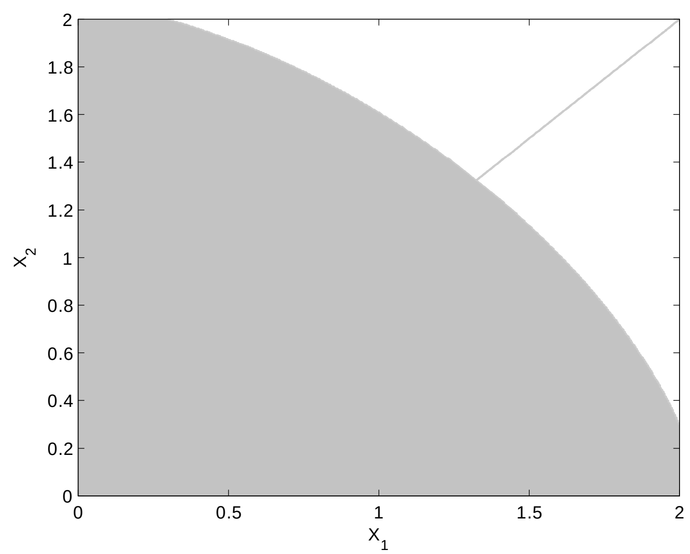


Figure A.5: Plot of set S_5 illustrating weak monotonicity of $F(x)$ on X for $K_2 = 0.500$

A.3 Remark

The shape of DUE effective delay operators that correspond to specific dynamic network loading models, including relevant flow propagation and non-negativity constraints, are not known. We can say that the gently rolling nature of Figure A.1 seems to be plausible and that, therefore, component-wise strongly monotone operators that are also component-wise weakly monotone may be a class of operators consistent with models of network loading, although that is not presently known and beyond the reach of the analysis tools employed herein.