Common-Lines and Passenger Assignment in Congested Transit Networks

Roberto Cominetti • José Correa
Universidad de Chile, Departamento de Ingeniería Matemática and
Centro de Modelamiento Matemático (UMR 2071 UCHILE-CNRS),
Casilla 170/3 Correo 3, Santiago, Chile

We analyze a Wardrop equilibrium model for passenger assignment in general transit networks, including the effects of congestion over the passengers' choices. The model is based on the common-line paradigm, which is applied to general networks using a dynamic programming approach. Congestion is treated by means of a simplified bulk queue model described in the appendix. We provide a complete characterization of the set of equilibria in the common-line setting, including the conditions for existence and uniqueness. This characterization reveals the existence of ranges of flow in which a Braess-like paradox appears, and in which a flow increase does not affect the system performance as measured by transit times. The congested common-line model is used to state an equilibrium model for general transit networks, and to establish the existence of a network equilibrium.

Introduction

Passenger assignment models aim to describe the way users of a public transportation system employ the available infrastructure for traveling between different origins and destinations in the network. Several models have been proposed, differing with respect to the assumptions on passenger behavior, network structure, and modeling of congestion. Here is a short overview of work in this area.

Earlier studies, such as those of Dial (1967), Fearnside and Draper (1971), and Le Clercq (1972), neglected congestion and assumed that passengers traveled along shortest paths on each origin-destination (OD) pair. The length of a path in this context corresponds to the total transit time including waiting as well as in-vehicle travel time. Later, considering a single corridor served by a set of bus lines, Chriqui and Robillard (1975) introduced the notion of common-lines suggesting that passengers could bundle together a subset of the available lines in order to reduce the waiting and hence the overall transit time. The assignment of passengers to bus lines was done proportionally to the nominal frequencies of each common-line. The extension of the common-line idea to general networks led Spiess (1984) and Spiess and Florian (1989) to introduce the notion of *strategy*, which was later expressed in graph-theoretic language by Nguyen and Pallottino (1988) under the denomination of *hyperpath*, namely, an acyclic subgraph connecting a given OD pair. In these models—which can handle simultaneously several OD pairs, overlapping bus lines, and transfers at intermediate nodes on each trip—passengers are assumed to travel along shortest hyperpaths. Despite this generality, the models did not consider explicitly the increase in waiting times induced by congestion, and the assignment of passengers to bus lines was done proportionally to the nominal frequencies. However, Nguyen and Pallottino consider flow-dependent travel times, modeling the on-board crowding of buses which may affect the passengers' choices.

Transportation Science © 2001 INFORMS
Vol. 35, No. 3, Summer 2001 pp. 250–267
0041-1655/01/3503/0250$05.00
1526-5447 electronic ISSN

The first attempt to incorporate the congestion effects on the passenger distribution and waiting times at bus stops seems to be Gendreau (1984). This study was based on a bulk queue model describing the waiting process, but its complexity prevented the analysis of the network equilibrium. To overcome this difficulty and the complexity of dealing with hyperpaths, De Cea and Fernández (1993) studied a network model in which passengers travel by following a sequence of intermediate transfer nodes. To move between successive transfer nodes, passengers are supposed to solve a common-line problem in which waiting times as well as passenger distribution depend not only on the flow boarding at the node but also on the flow already on the bus. While this work may be considered as the first network model to incorporate congestion, it has some important drawbacks. First, the functional form used to represent congestion is only justified heuristically, and practical computations show that the model may produce line loads beyond the line capacities. Second, the common-lines between transfer nodes are computed according to a heuristic method which does not guarantee the Wardrop equilibrium condition to be fulfilled. Third, and related to the previous point, all passengers traveling between two transfer nodes are assumed to use the same set of common-lines (we shall see that in general there may be no equilibrium of this form).

More recently Wu et al. (1994) studied a congested network assignment model in which passengers travel according to shortest hyperpaths. Travel times as well as waiting times are considered to be flow dependent, but the passenger assignment is based on the nominal frequencies of the lines. Finally, Bouzaïene-Ayari et al. (1995 a–c) extended this model to study existence, uniqueness, and computation of network equilibria for the case in which the assignment is also flow dependent: the flow distribution is done proportionally to the inverses of the waiting times of each line. This congestion model assumes that waiting times obey an inverse additive law of the form

$$\frac{1}{W_s(v)} = \sum_{i\in s} \frac{1}{W_{\{i\}}(v)}$$

where v represents the vector of line-flows, W_s denotes the waiting time corresponding to the set of lines s, and $W_{\{i\}}$ is the waiting time of line i taken individually. The model is based on evidence obtained from simulations, and the functions $W_{\{i\}}$ are calibrated empirically. To establish the existence of an equilibrium the authors assume these functions to have uniformly bounded gradients. This assumption precludes the use of congestion functions based on queuing theory, which exhibit a level of flow saturation at which queue lengths and waiting times explode to infinity. Additionally, the travel-time functions are assumed to be strongly monotone, which prevents the model from being used in the simplest case of constant travel times. Moreover, uniqueness of equilibrium is established only for the case in which this strongly monotone term dominates over the waiting time.

The goal of this paper is to study a network equilibrium model with congestion, in which travel times are not necessarily monotone and congestion affects both the waiting times and the flow distribution. Passengers are assumed to travel according to shortest hyperpaths, which are described in a simplified form in terms of local strategies. The model—an outgrowth of Correa (1999)—exploits the common-line idea in a dynamic programming approach and it can handle congestion functions obtained from queuing models as the ones proposed by Gendreau (1984). In Appendix A we provide a queue-theoretic support and discuss the limitations of our treatment of congestion, justifying the additive law for inverses of waiting times as well as the flow distribution proportionally to these inverses.

The plan of the paper is as follows. In §1 we analyze the common-line problem under congestion, proving the existence of equilibria in terms of *strategy-flows* and giving a complete characterization in terms of an equivalent optimization problem satisfied by the equilibrium *line-flows*. This characterization shows that the set of equilibrium line-flows forms a simple convex polytope, and readily gives the conditions for uniqueness. Moreover, it reveals some unexpected features such as the existence of ranges of flow in which no single set of common lines can yield an equilibrium, and also the fact that a flow increase need not induce larger transit times. More interestingly, in those ranges the equilibrium is not the social optimum since by appropriately restricting the user's choices everyone can strictly improve his transit time. In other words, the equilibrium need not be Pareto optimal. This can be seen as an analog for Braess' paradox in the setting of congested public transportation networks.

It is worth pointing out that the congested common-line equilibrium model could be stated as a variational inequality which is however nonmonotone as well as nonsymmetric. This fact prevents the use of the existing machinery for monotone variational inequalities. In particular, the set of equilibrium strategy-flows is in general a nonconvex subset of a smooth (nonlinear) manifold. The existence of an optimization equivalent for the corresponding equilibrium line-flows is therefore a very fortunate fact, which is extremely convenient for stating an equilibrium model for general transit networks. The latter is done in §2, where the common-line model is used as a building block exploiting a dynamic programming approach. The resulting network model consists of a set of common-line problems (one for each OD pair) linked together by flow conservation constraints. The philosophy behind the model is closely related to the strategy/hyperpath approach. However, using the common-line framework we avoid handling explicitly the flows on strategies and we write a concise model directly in terms of line-flows. The existence of a network equilibrium is obtained by using fixed point arguments.

1. The Common-Line Problem Under Congestion

Consider the simplest network consisting of an origin O connected to a destination D by n bus lines l_i, $i \in A = \{1, \dots, n\}$, each one characterized by an in-vehicle travel time t_i and a frequency f_i (Figure 1). According to Chriqui and Robillard (1975), for the purpose of traveling from O to D, passengers select a subset of *common lines* $s \subset A$ boarding the first incoming bus from this set. The chosen strategy s should minimize the expected transit time T_s, including the waiting time $1/\sum_{i\in s} f_i$ and the expected in-vehicle travel time $\sum_{i\in s} t_i \pi_i^s$, where $\pi_i^s = f_i/\sum_{j\in s} f_j$ is the probability of boarding line l_i, i.e.,

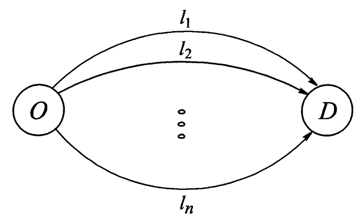

Figure 1 The Common-Line Problem

$$T_s = \frac{1 + \sum_{i\in s} t_i f_i}{\sum_{i\in s} f_i}.$$

Notice that some models adopt a waiting time of the form $\alpha/\sum_{i\in s} f_i$ with $\alpha \in (0, 1)$, which can be easily reduced to the present setting by changing the frequencies to f_i/α. Now, intuitively, a passenger consenting to board l_i should also consider the faster lines l_j with $t_j < t_i$ (see Lemma 1.2 below), leading to a linear-time algorithm for computing an optimal s^* (see Chriqui and Robillard 1975): initialize $s^* = \emptyset$ and repeat the update $s^* \leftarrow s^* \cup \{i \notin s^* : t_i = \min_{j\notin s^*} t_j\}$ until $\min_{j\notin s^*} t_j \geq T_{s^*}$.

This model neglects congestion and assumes that passengers can effectively board the first incoming bus. However, large flows and limited bus capacity may prevent (from time to time) a passenger from boarding the bus that comes first, increasing his waiting time. A simple approach to model this situation is to suppose that the frequency of line l_i is a decreasing function of the flow v_i on this line. In Appendix A we provide a queue-theoretic support for such a model and discuss its limitations.

Let us then assume that each line l_i is characterized by a differentiable *effective frequency* function $f_i : [0, \bar{v}_i) \to (0, \infty)$, with $f_i'(\cdot) < 0$ and $f_i(v_i) \to 0$ as $v_i \to \bar{v}_i$. The constant $\bar{v}_i > 0$ (eventually $\bar{v}_i = +\infty$) is called the *saturation flow* of the line. In this setting, the expected transit time of strategy s becomes a function of the line-flows

$$T_s(v) := \frac{1 + \sum_{i\in s} t_i f_i(v_i)}{\sum_{i\in s} f_i(v_i)} \qquad (T)$$

so that the optimal decision of each passenger is affected by the choices of the others, and then one must think of the common-line problem in terms of equilibrium. Specifically, consider a total flow $x > 0$ of passengers going from O to D and denote by $\mathcal{S}$ the set of nonempty subsets $s \subset A$ (the possible strategies). The flow x splits into flows $y_s \geq 0$ along the strategies so that $x = \sum_{s\in\mathcal{S}} y_s$. Assuming that a passenger from strategy s boards each line $i \in s$ with probability $\pi_i^s = f_i(v_i)/\sum_{j\in s} f_j(v_j)$ (see Appendix A), and denoting by $\mathcal{S}_i := \{s \in \mathcal{S} : i \in s\}$ the set of strategies containing line l_i, the vector of strategy-flows $y = (y_s)_{s\in\mathcal{S}}$ determines a unique vector of line-flows $v = v(y)$ through the system of equations (see §1.1)

$$v_i = \sum_{s\in\mathcal{S}_i} y_s \frac{f_i(v_i)}{\sum_{j\in s} f_j(v_j)} \quad \text{for } i = 1, \dots, n. \qquad (E)$$

Using Wardrop's principle, we say that the strategy-flow vector y^* is an *equilibrium* if and only if all strategies carrying flow are of minimal time, that is,

$$y_s^* > 0 \Rightarrow T_s(v(y^*)) = \widehat{T}(v(y^*)), \qquad (W)$$

where $\widehat{T}(v) := \min_{s\in\mathcal{S}} T_s(v)$.

We refer to the model above as the *TEW* equilibrium model. In §1.2 we prove the existence of equilibria y^* and we characterize the corresponding line flows $v(y^*)$ as the optimal solutions of an equivalent optimization problem. This characterization reveals some remarkable features of the equilibrium, including the conditions for having uniqueness. For instance, a priori one could expect that the only effect of congestion would be a change in the optimal strategy, which could then be computed by using Chriqui and Robillard's algorithm with suitable frequencies $f_i(v_i)$ (eventually by an iterative process adjusting the flows v_i). This view, which assumes that all passengers use the same strategy, is adopted for example by De Cea and Fernández (1993). However we will see that in some ranges of flow there cannot be an equilibrium where all passengers adopt the same strategy, and x must necessarily split among two or more strategies. Moreover, in such ranges the equilibrium transit time remains constant, i.e., an increase in the flow x does not deteriorate the performance of the system.

1.1. Relation Between Strategy-Flows and Line-Flows

For the *TEW* equilibrium model to be well defined we must check that (E) defines an implicit function $v = v(y)$. To this end we consider the change of variables $z_i := \ln f_i(v_i)$ defined for $v_i < \bar{v}_i$ (we set $f_i(v_i) = f_i(0) - v_i$ for $v_i < 0$). In these new variables system (E) becomes

$$0 = -f_i^{-1}(e^{z_i}) + \sum_{s\in\mathcal{S}_i} y_s \frac{e^{z_i}}{\sum_{j\in s} e^{z_j}} \quad \text{for all } i = 1, \dots, n, \qquad (1)$$

which amounts to $\nabla q(z) = 0$, where $q : \mathbb{R}^n \to \mathbb{R}$ is defined as

$$q(z) := \sum_{i=1}^{n} h_i(e^{z_i}) + \sum_{s\in\mathcal{S}} y_s \ln\left(\sum_{j\in s} e^{z_j}\right)$$

with $h_i : \mathbb{R}_+ \to \mathbb{R}$ given by

$$h_i(u) := -\int_1^u \frac{f_i^{-1}(\tau)}{\tau}\, d\tau.$$

It is easy to see that $h_i(u)$ is strictly convex for $i = 1, \dots, n$, hence $q(z)$ is also strictly convex and the solution of (1) is the unique minimizer of q, provided it exists. The next result establishes conditions for the existence of this minimum.

Theorem 1.1. *For $y \geq 0$, the following are equivalent*:

(a) $\sum_{s\in\mathcal{S}} y_s \min_{j\in s} x_j < \sum_{i=1}^n \bar{v}_i x_i$ *for all* $x \in \Delta := \{x \in \mathbb{R}^n : x \geq 0, \sum_{i=1}^n x_i = 1\}$.

(b) *There exist* $v_{si} \geq 0$ *such that* $y_s = \sum_{i\in s} v_{si}$ *for all* $s \in \mathcal{S}$ *and* $\sum_{s\in\mathcal{S}_i} v_{si} < \bar{v}_i$ *for* $i = 1, \dots, n$.

(c) $\sum_{s'\subset s} y_{s'} < \sum_{i\in s} \bar{v}_i$ *for all* $s \in \mathcal{S}$.

These conditions are necessary and sufficient for the existence and uniqueness of a minimum of q, and hence of a solution $v = v(y)$ of system (E).

Proof. Since q is convex, its minimum is attained iff $L(d) := \lim_{t\to\infty} \langle \nabla q(td), d\rangle > 0$ for all $d \neq 0$. A straightforward computation gives

$$L(d) = \begin{cases} -\sum_{i=1}^n \bar{v}_i d_i + \sum_{s\in S} y_s \max_{j\in s} d_j & \text{if } d \leq 0, \\ +\infty & \text{otherwise,} \end{cases}$$

so that q attains its minimum if and only if

$$-\sum_{i=1}^{n} \bar{v}_i d_i + \sum_{s\in S} y_s \max_{j\in s} d_j > 0 \quad \forall\, d \le 0,\, d \ne 0.$$

Normalizing so that $\sum_{i=1}^{n} d_i = -1$ and multiplying by -1, this is precisely condition (a).

(a) $\iff$ (b): Notice that (a) can be equivalently stated as $L > 0$ with

$$L := \min_{x,\lambda} \left\{ \sum_{i=1}^{n} \bar{v}_i x_i - \sum_{s\in\mathcal{S}} y_s \lambda_s : x \in \Delta;\ \lambda_s \le x_i \right.$$
$$\left. \text{for all } s \in \mathcal{S},\, i \in s \right\}.$$

This is a bounded linear program whose dual is

$$L = \max_{\substack{\gamma\in\mathbb{R}\\ (v_{si})\ge 0}} \left\{ \gamma : \sum_{i\in s} v_{si} = y_s \text{ for all } s \in \mathcal{S}; \right.$$
$$\left. \sum_{s\in\mathcal{S}_i} v_{si} \le \bar{v}_i - \gamma \text{ for } i = 1, \ldots, n \right\}$$

so that $L > 0$ is also equivalent to (b).

(a) $\Longrightarrow$ (c): For $s \in \mathcal{S}$ let $x_i = 1/|s|$ if $i \in s$ and $x_i = 0$ otherwise. Then (a) gives $\sum_{s'\subset s} y_{s'} < \sum_{i\in s} \bar{v}_i$.

(c) $\Longrightarrow$ (b): Consider an auxiliary graph (Figure 2) with node sets $\mathcal{S}$ and A, and two extra nodes f, p. Connect these nodes by the arcs $E_1 = \{(f, s) : s \in \mathcal{S}\}$, $E_2 = \{(i, p) : i \in A\}$, and $E_3 = \{(s, i) : s \in \mathcal{S},\, i \in s\}$ assigning to each arc e a capacity

$$c(e) = \begin{cases} y_s & \text{if } e = (f, s) \in E_1, \\ \bar{v}_i - \epsilon & \text{if } e = (i, p) \in E_2, \\ \infty & \text{if } e \in E_3, \end{cases}$$

with $\epsilon > 0$ chosen so that

$$\sum_{s'\subset s} y_{s'} \le \sum_{i\in s} \bar{v}_i - \epsilon, \quad \forall\, s \in \mathcal{S}. \tag{2}$$

Clearly, proving (b) amounts to showing that the maximum flow from f to p is $\sum_{s\in\mathcal{S}} y_s$. By the maxflow-mincut theorem this is equivalent to proving that every (f, p)-cut (W, W^c) has capacity $c(W) \ge \sum_{s\in\mathcal{S}} y_s$. Observe that it suffices to consider those cuts containing only arcs from E_1 and E_2, since otherwise $c(W) = \infty$. In this case, letting $U = W \cap A$ we get

$$c(W) = \sum_{i\in U} (\bar{v}_i - \epsilon) + \left(\sum_{s\in\mathcal{S}} y_s - \sum_{s'\subset U} y_{s'} \right)$$

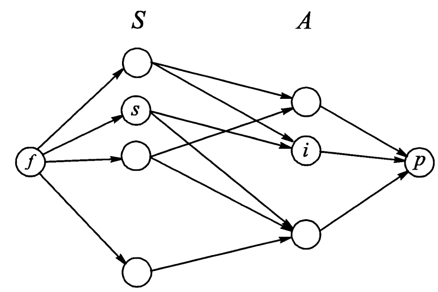

Figure 2 **The Auxiliary Graph**

so that using (2) we conclude

$$c(W) - \sum_{s\in\mathcal{S}} y_s = \sum_{i\in U} (\bar{v}_i - \epsilon) - \sum_{s'\subset U} y_{s'} \ge 0. \quad \square$$

Remark 1.1. Condition (b) amounts to saying that the strategy-flow vector $y = (y_s)_{s\in\mathcal{S}}$ may be decomposed into line-flows without saturating any line. To interpret condition (c) observe that the quantity $\sum_{s'\subset s} y_{s'}$ must flow through the set of lines s, so that the condition requires that each subset of lines has enough capacity to carry the flow required from them.

1.2. Existence and Characterization of Equilibria

Let $w(\alpha) = (w_i(\alpha))_{i=1}^{n}$ with $w_i : [0, \infty) \to [0, \bar{v}_i)$ the inverse of the differentiable and strictly increasing function $v_i \mapsto v_i / f_i(v_i)$, and let $h : [0, \infty) \to [0, \infty)$ be defined as

$$h(\alpha) := \sum_{i=1}^{n} \int_0^{\alpha} [\widehat{T}(w(\beta)) - t_i]_+ w_i'(\beta)\, d\beta, \tag{3}$$

where $[x]_+$ denotes the positive part of x (equal to x if $x \ge 0$ and 0 otherwise). We shall prove that the line-flows $v = v(y^*)$ corresponding to solutions y^* of *TEW* coincide with the optimal solutions of

$$\min_{v\ge 0} \left\{ \sum_{i=1}^{n} t_i v_i + h\left(\max_{i=1,\ldots,n} \frac{v_i}{f_i(v_i)} \right) : \sum_{i=1}^{n} v_i = x \right\}. \tag{P_x}$$

This equivalence will give the existence as well as a complete characterization of equilibria, including the conditions for uniqueness. To proceed we establish some elementary preliminary lemmas. In the sequel we call a *strict convex combination* (scc) any convex combination with strictly positive coefficients.

Lemma 1.1. *If $s \subset s'$, then $T_{s'}(v)$ is a scc of $T_s(v)$ and $\{t_j : j \in s' \backslash s\}$.*

Proof. It suffices to remark that

$$T_{s'}(v) = \left(\frac{\sum_{i\in s} f_i(v_i)}{\sum_{i\in s'} f_i(v_i)} \right) T_s(v) + \sum_{j \in s' \backslash s} \left(\frac{f_j(v_j)}{\sum_{i \in s'} f_i(v_i)} \right) t_j. \quad \square$$

Lemma 1.2. *A strategy s is optimal (i.e., $T_s(v) = \widehat{T}(v)$) iff $t_i \le T_s(v) \le t_j$ for all $i \in s$, $j \notin s$.*

Proof. Let s be optimal. For each $i \in s$ the time $T_s(v)$ is a scc of $T_{s\backslash\{i\}}(v)$ and t_i. Since $T_{s\backslash\{i\}}(v) \ge T_s(v)$, we must have $t_i \le T_s(v)$. Similarly, for $j \notin s$ we have that $T_{s\cup\{j\}}(v)$ is a scc of $T_s(v)$ and t_j with $T_{s\cup\{j\}}(v) \ge T_s(v)$, and then $t_j \ge T_s(v)$.

Conversely, suppose that $t_i \le T_s(v) \le t_j$ for all $i \in s$, $j \notin s$. Let s' be an optimal strategy and set $u = s \cap s'$ so that $T_{s'}(v)$ is a scc of $T_u(v)$ and $\{t_j\}_{j \in s' \backslash s}$. The latter times are larger than $T_s(v)$, and also $T_u(v) \ge T_s(v)$ since $T_s(v)$ is a scc of $T_u(v)$ and the times $\{t_i\}_{i \in s\backslash u}$ which are smaller than $T_s(v)$. Therefore $T_{s'}(v) \ge T_s(v)$, proving that s is also optimal. $\square$

Corollary 1.1. *Let $\hat{s} := \{i : t_i < \widehat{T}(v)\}$ and $\check{s} := \{i : t_i \le \widehat{T}(v)\}$. Then s is optimal iff $\hat{s} \subset s \subset \check{s}$.*

Proof. Lemma 1.2 implies that $\hat{s} \subset s \subset \check{s}$ for each optimal s. In particular $\widehat{T}(v) = T_s(v)$ is a scc of $T_{\hat{s}}(v)$ and $\{t_j\}_{j \in s\backslash\hat{s}} = \{\widehat{T}(v)\}$, so that $T_{\hat{s}}(v) = \widehat{T}(v)$. But then, for any $\tilde{s}$ with $\hat{s} \subset \tilde{s} \subset \check{s}$ the time $T_{\tilde{s}}(v)$ is a scc of $T_{\hat{s}}(v) = \widehat{T}(v)$ and $\{t_j\}_{j \in \tilde{s}\backslash\hat{s}} = \{\widehat{T}(v)\}$, so that $T_{\tilde{s}}(v) = \widehat{T}(v)$ and $\tilde{s}$ must be optimal. $\square$

Lemma 1.3. *$\alpha \mapsto \widehat{T}(w(\alpha))$ is continuous and strictly increasing, with $\widehat{T}(w(\alpha)) \to \infty$ as $\alpha \to \infty$.*

Proof. Clearly $w_i(\cdot)$ is differentiable with $w_i'(\cdot) > 0$ and $w_i(\alpha) \to \bar{v}_i$ when $\alpha \to \infty$. It follows that each function $T_s(w(\cdot))$ is continuous with $T_s(w(\alpha)) \to \infty$ when $\alpha \to \infty$, and therefore the same properties hold for their minimum $\widehat{T}(w(\cdot))$. To prove that $\widehat{T}(w(\cdot))$ is strictly increasing it suffices to show that $\frac{d}{d\alpha} T_s(w(\alpha)) > 0$ for all s such that $T_s(w(\alpha)) = \widehat{T}(w(\alpha))$. A straightforward computation gives

$$\frac{d}{d\alpha} T_s(w(\alpha)) = \frac{1}{\sum_{j \in s} f_j(w_j(\alpha))} \sum_{i \in s} [t_i - T_s(w(\alpha))] \times f_i'(w(\alpha)) w_i'(\alpha).$$

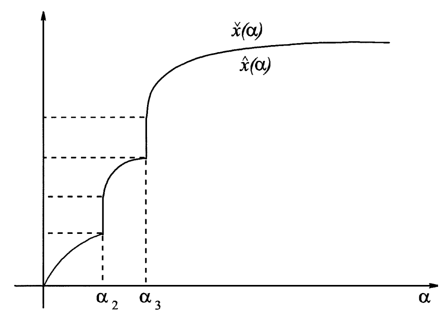

Figure 3 **The Functions $\hat{x}(\alpha)$ and $\check{x}(\alpha)$**

Lemma 1.2 implies $t_i \le T_s(w(\alpha))$ for all $i \in s$, while in general $T_s(w(\alpha)) > t_{i_0}$ for the fastest line $i_0 \in s$. Since $f_i'(w_i(\alpha)) < 0$ and $w_i'(\alpha) > 0$, the conclusion follows. $\square$

We are ready to state our main result. Consider the strictly increasing functions

$$\hat{x}(\alpha) := \sum_i \{w_i(\alpha) : t_i < \widehat{T}(w(\alpha))\},$$

$$\check{x}(\alpha) := \sum_i \{w_i(\alpha) : t_i \le \widehat{T}(w(\alpha))\}.$$

These functions are continuous and equal, except at the values α_k such that $\widehat{T}(w(\alpha_k)) = t_k$ where we have $\check{x}(\alpha_k) > \hat{x}(\alpha_k)$ (Figure 3). Moreover, $\hat{x}(0) = \check{x}(0) = 0$ and $\hat{x}(\alpha) = \check{x}(\alpha) \to \sum_{i=1}^n \bar{v}_i$ when $\alpha \to \infty$.

Theorem 1.2. *Let $x \in (0, \sum_{i=1}^n \bar{v}_i)$. Set $\widehat{T}_x := \widehat{T}(w(\alpha_x))$ with α_x the unique solution of $x \in [\hat{x}(\alpha), \check{x}(\alpha)]$. Then the optimal set V_x of (P_x) is the nonempty bounded polytope of all v's such that $\sum_{i=1}^n v_i = x$ and $0 \le v_i \le w_i(\alpha_x)$, with $v_i = w_i(\alpha_x)$ if $t_i < \widehat{T}_x$ and $v_i = 0$ if $t_i > \widehat{T}_x$. Moreover, every TEW-equilibrium y^* has equilibrium time $\widehat{T}(v(y^*)) = \widehat{T}_x$ and we have $V_x = \{v(y^*) : y^* \text{ solves TEW}\}$.*

Proof. Since $h(\cdot)$ is strictly increasing, problem (P_x) is equivalent to

$$\min_{v,\alpha} \left\{ \sum_{i=1}^n t_i v_i + h(\alpha) : \sum_{i=1}^n v_i = x; 0 \le v_i \le w_i(\alpha) \right\}. \quad (\widetilde{P}_x)$$

It is easy to check that every optimal solution (v, α) is such that $\alpha = \max_i v_i / f_i(v_i) > 0$ and satisfies the Mangasarian-Fromovitz constraint qualification, so that the Karush-Kuhn-Tucker (KKT) conditions hold: there exist multipliers $\lambda \in \mathbb{R}$, $\mu_i^1 \ge 0$, $\mu_i^2 \ge 0$ with

(a) $\mu_i^1 v_i = 0$ and $\mu_i^2 (v_i - w_i(\alpha)) = 0$ for $i = 1, \ldots, n$,

(b) $\lambda - t_i = \mu_i^2 - \mu_i^1$ for $i = 1, \ldots, n$,
(c) $h'(\alpha) = \sum_{i=1}^n \mu_i^2 w_i'(\alpha)$.

Since $\alpha > 0$ we have $w_i(\alpha) > 0$ for all i's, and then (a) implies that μ_i^1 and μ_i^2 cannot both be positive. This, combined with (b), yields $\mu_i^2 = [\lambda - t_i]_+$ and $\mu_i^1 = [\lambda - t_i]_-$, and then equation (c) becomes

$$\sum_{i=1}^n [\widehat{T}(w(\alpha)) - t_i]_+ w_i'(\alpha) = \sum_{i=1}^n [\lambda - t_i]_+ w_i'(\alpha),$$

which gives $\lambda = \widehat{T}(w(\alpha))$. These relations, together with (a), imply that $v_i = 0$ if $t_i > \widehat{T}(w(\alpha))$ and $v_i = w_i(\alpha)$ if $t_i < \widehat{T}(w(\alpha))$, while for the remaining lines we have $0 \le v_i \le w_i(\alpha)$. Hence

$$\begin{aligned}\sum\{w_i(\alpha) : t_i < \widehat{T}(w(\alpha))\} &\le \sum_{i=1}^n v_i \\ &\le \sum\{w_i(\alpha) : t_i \le \widehat{T}(w(\alpha))\}\end{aligned}$$

so that $\hat{x}(\alpha) \le x \le \check{x}(\alpha)$, and therefore $\alpha = \alpha_x$. The only variables which are not completely specified by KKT are the flows v_i for the lines l_i such that $t_i = \widehat{T}_x$. However, whatever value $v_i \in [0, w_i(\alpha_x)]$ is assigned to these variables will give the same objective value, and therefore we obtain the stated characterization for V_x. We have proved in fact that every feasible point (v, α) satisfying the KKT conditions is optimal for $(\widetilde{P}_x)$.

Let us show that $\widehat{T}_x$ is the equilibrium time for the *TEW* model and $V_x = \{v(y^*) : y^* \text{ solves } TEW\}$. Consider first an optimal $v \in V_x$ and let y^* be constructed by the following algorithm.

Initialize: $y_s^* \leftarrow 0$ for all s; $\alpha_i \leftarrow v_i/f_i(v_i)$ for all i
while $s := \{i : \alpha_i > 0\}$ is nonempty
  set $y_s^* := \delta \sum_{j\in s} f_j(v_j)$ with $\delta := \min\{\alpha_i : \alpha_i > 0\}$
  update $\alpha_i \leftarrow \alpha_i - \delta = \alpha_i - y_s^*/\sum_{j\in s} f_j(v_j)$
    for all $i \in s$
end

The update stage reduces the value of the positive α_i's, setting to zero the smallest among them. Hence the algorithm is finite and the generated y^* and v satisfy Equation (E) so that $v = v(y^*)$. Notice also that the strategies s found along the iterations are strictly decreasing, with all of them contained in $\check{s} := \{i : t_i \le \widehat{T}_x\}$ (since $\alpha_i = 0$ if $t_i > \widehat{T}_x$) and containing $\hat{s} := \{i : t_i < \widehat{T}_x\}$ (since these lines have the largest value of α_i, initially set to α_x). Thus, the times $T_s(v)$ are a scc of $T_{\hat{s}}(v)$ and $\{t_j\}_{j\in s\setminus\hat{s}} = \{\widehat{T}_x\}$. Now, Corollary 1.1 implies $T_{\hat{s}}(v) = T_{\hat{s}}(w(\alpha_x)) = \widehat{T}_x$ so that $T_s(v) = \widehat{T}_x$, and then $t_i \le T_s(v) \le t_j$ for all $i \in s$, $j \notin s$. Lemma 1.2 implies that the strategies generated by the algorithm are optimal, and therefore y^* is an equilibrium.

Conversely, let y^* be an equilibrium and $v = v(y^*)$. Defining $\alpha = \max_i v_i/f_i(v_i)$ we have $\sum_{i=1}^n v_i = x$ and $0 \le v_i \le w_i(\alpha)$ so that (v, α) is feasible for $(\widetilde{P}_x)$. Also, defining $\lambda = \widehat{T}(v)$, $\mu_i^2 = [\lambda - t_i]_+$ and $\mu_i^1 = [\lambda - t_i]_-$, it is evident that KKT conditions (b) and (c) hold. Since y^* is an equilibrium, Corollary 1.1 implies $\hat{s} \subset s \subset \check{s}$ for all s with $y_s^* > 0$, and using Equation (E) it follows that $v_i/f_i(v_i)$ is maximal (equal to α) for $i \in \hat{s}$ and is 0 for $j \notin \check{s}$. Hence $v_i = w_i(\alpha)$ for $i \in \hat{s}$ and $v_j = 0$ for $j \notin \check{s}$, proving that condition KKT (a) is also satisfied. As noticed above, this implies that (v, α) is optimal for $(\widetilde{P}_x)$, and consequently $v \in V_x$ as was to be proved. A fortiori $\alpha = \alpha_x$ and $\widehat{T}(v) = \widehat{T}_x$ so that every equilibrium y^* has the same equilibrium time $\widehat{T}_x$, completing the proof. □

Remark 1.2. Theorem 1.2 implies that for each $x \in (0, \sum_{i=1}^n \bar{v}_i)$ there exists at least one equilibrium. This equilibrium will be unique unless there are two or more lines l_i with $t_i = \widehat{T}_x$. Of course this may not occur if all travel times t_i are different, in which case we always have uniqueness.

Remark 1.3. For each $t_k > \widehat{T}_0$ there is a unique α_k with $\widehat{T}(w(\alpha_k)) = t_k$. If x does not belong to any of the intervals $[\hat{x}(\alpha_k), \check{x}(\alpha_k)]$, the unique equilibrium y^* is such that all the flow x is assigned to the strategy $s^* = \{i : t_i \le \widehat{T}_x\} = \{i : t_i < \widehat{T}_x\}$. When $x \in [\hat{x}(\alpha_k), \check{x}(\alpha_k)]$ we have $\alpha_x = \alpha_k$ and $\widehat{T}_x = t_k$ (notice that a flow increment in that range will not increase the equilibrium time). In this situation it may happen that every $v \in V_x$ has $0 < v_i/f_i(v_i) < \alpha_x$ for some line l_i, so that at equilibrium the flow x must necessarily split among two or more strategies (this follows from Equation (E)).

Remark 1.4. We notice that for $x \in (\hat{x}(\alpha_k), \check{x}(\alpha_k))$ the equilibrium is not the social optimum: if we force the flow x to use only the strategy $\check{s} = \{i : t_i \le \widehat{T}_x\}$, the time $T_{\check{s}}$ will be smaller than the equilibrium time $\widehat{T}_x$. Thus, by restricting the passenger's choices the system would reach a state where each and every

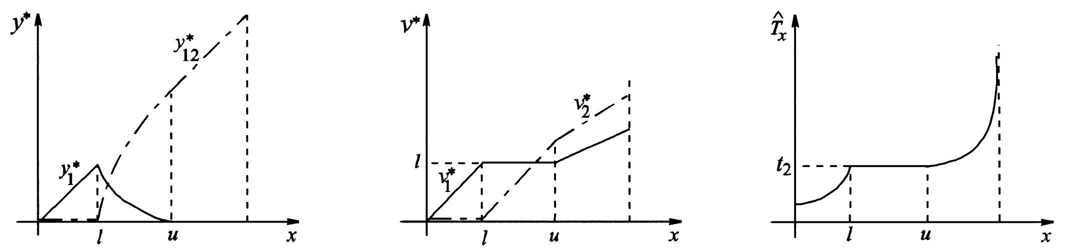

Figure 4 **Equilibrium Strategy-Flows, Line-Flows, and Transit Time**

passenger would be better off than in the equilibrium situation. This can be seen as an analog for Braess' paradox in the case of public transportation networks. This restricted state is not an equilibrium though, since the strategy $\hat{s} = \{i : t_i < \widehat{T}_x\}$ has a still smaller time $T_{\hat{s}} < T_{\check{s}} < \widehat{T}_x$, inducing a transfer of flow from $\check{s}$ to $\hat{s}$. This flow transfer increases the time of both strategies, and the equilibrium is reestablished when they reach the level $\widehat{T}_x$.

Remark 1.5. Theorem 1.2 provides a direct characterization of the equilibrium line-flows v, without using explicitly the strategy-flows y (although these can be recovered if necessary). This fact will be useful for stating a simple equilibrium model for general transit networks, by working directly in terms of arc-flows and avoiding dealing explicitly with the notion of strategy/hyperpath.

Example 1. Consider two bus line with times $t_1 < t_2$, Poisson arrivals of intensities μ_1 and μ_2, and equal capacity K. According to (9) and (8) in Appendix A we have $w_i(\alpha) = \mu_i \alpha(1 - [\alpha/(1+\alpha)]^K)$ and $f_i(v_i) = v_i[1/\rho_i(v_i) - 1]$ with $\bar{v}_i = K\mu_i$ and $\rho_i(v_i)$ the unique solution $\rho \in [0, 1)$ of the equation

$$\mu_i(\rho + \rho^2 + \cdots + \rho^K) = v_i.$$

For every $x \in (0, \bar{v}_1 + \bar{v}_2)$ there exists a unique equilibrium, which may only use the strategies $s = \{1\}$ and/or $s = \{1, 2\}$. To compute it explicitly, set $\lambda = \mu_1/(\mu_1 + \mu_2)$ and let α_2 be the solution of $\widehat{T}(w(\alpha)) = t_2$, or equivalently $f_1(w_1(\alpha)) = 1/(t_2 - t_1)$. If $\mu_1(t_2 - t_1) < 1$, there is no such α_2, and the equilibrium is $y^*_{\{1,2\}} = x$ with corresponding line-flows $v^*_1 = \lambda x$ and $v^*_2 = (1-\lambda)x$. Otherwise we have $\alpha_2 = \delta/(1-\delta)$ with $\delta := [1 - 1/\mu_1(t_2 - t_1)]^{1/K}$, and setting $l := \hat{x}(\alpha_2)$, $u := \check{x}(\alpha_2)$, and $t^\lambda = \lambda t_1 + (1-\lambda)t_2$ we get

$$y^*_{\{1\}} = \begin{cases} x & \text{if } x \le l, \\ l - \frac{\rho_2(x-l)}{[t_2-t_1][1-\rho_2(x-l)]} & \text{if } l < x < u, \\ 0 & \text{if } x \ge u; \end{cases} \quad y^*_{\{1,2\}} = x - y^*_{\{1\}};$$

$$v^*_1 = \begin{cases} x & \text{if } x \le l, \\ l & \text{if } l < x < u, \\ \lambda x & \text{if } x \ge u; \end{cases} \qquad v^*_2 = x - v^*_1;$$

$$\widehat{T}_x = \begin{cases} t_1 + \frac{\rho_1(x)}{x[1-\rho_1(x)]} & \text{if } x \le l, \\ t_2 & \text{if } l < x < u, \\ t^\lambda + \frac{\rho_1(\lambda x)}{x[1-\rho_1(\lambda x)]} & \text{if } x \ge u. \end{cases}$$

Note that for all $x \in (l, u)$ both $s^* = \{1\}$ and $s^* = \{1, 2\}$ carry a positive flow, so that the existence of an equilibrium using two strategies may not be considered as an isolated degenerate situation (Figure 4).

1.3. Dealing with Infinite Frequencies

When modeling general transit networks (see §2) one usually considers different types of arcs: boarding, alight, on-board, and walk arcs. The waiting processes affect mainly the boarding arcs, while the other arcs correspond to services which are always available so that their waiting times are either zero or negligible. This can be modeled by attaching an infinite frequency $f_i(\cdot) \equiv \infty$ to those arcs, with $\bar{v}_i = \infty$. The formulas and results in the previous section must be revised accordingly.

Denote by A^F and A^I respectively the finite and infinite frequency lines in A. To define the transit time T_s and boarding probabilities π^s_i for a strategy s containing one or more lines from A^I, we replace the infinite frequencies by a constant f_∞ and consider the limit of

T_s and π_i^s as $f_\infty \to \infty$. Hence, T_s is the average of the travel times of the lines in $s \cap A^I$ and y_s is distributed uniformly among these lines. Then, the flow on each line $i \in A^I$ can be computed directly, and system (E) must only be solved for $(v_i)_{i \in A^F}$. Defining $w_i(\cdot) \equiv \infty$ for all $i \in A^I$, Theorem 1.2 remains valid in this extended setting. The proof follows the same lines as before and is left to the reader. However, a few points must be kept in mind, namely:

(a) Lemmas 1.1 and 1.2, as well as Corollary 1.1, apply only to finite frequency strategies $s \subset A^F$.

(b) In Lemma 1.3, $\alpha \mapsto \widehat{T}(w(\alpha))$ is strictly increasing until a point α_∞ where it reaches the value $t_\infty := \min\{t_i : i \in A^I\}$. Beyond that point it remains constant and equal to t_∞.

(c) Since $\widehat{T}(w(\beta)) \le t_\infty$, only the lines in A^F contribute to the definition of $h(\cdot)$ in (3).

2. Passenger Assignment and Equilibrium in General Networks

In this section we study an equilibrium model for transit networks, supporting multiple origins and destinations, overlapping bus lines, as well as transfers at intermediate nodes on any given trip. The model exploits the common-line scheme using a dynamic programming approach. The idea is as follows. Consider a passenger heading towards destination k and reaching an intermediate node i in his trip. To exit from i he can use the arcs $a \in A_i^+$ to reach the next node $j(a)$. Denoting by t_a the travel time of arc a and by $\tau_{j(a)}^k$ the transit time from $j(a)$ to k, the decision problem faced at node i is a common-line problem with travel times $t_a + \tau_{j(a)}^k$ and effective frequencies corresponding to the services attached to the arcs $a \in A_i^+$. The solution of this problem determines the transit time from i to k, which can be used recursively to solve the common-line problems for the upstream nodes. The variables τ_i^k and the equilibrium flows must be determined simultaneously for each node i and every destination k. Thus, the transit network model consists of a family of *TEW* equilibrium models (one for each pair i, k) linked together by flow conservation equations. A detailed description follows.

Network. We state the model for a general directed graph $G = (N, A)$. However, to make the model more concrete, the reader may think of a graph G built as follows (Spiess and Florian 1989). Let S be a set of nodes representing the bus stops in the network. A bus line l is defined by a set of line-nodes B_l representing the sequence of stop-nodes visited by this line. Each line-node in B_l connects to the corresponding stop-node in S through boarding and alight arcs, as well as to the next line-node in the sequence through an on-board arc. Eventually, one may consider walk arcs connecting directly a pair of nodes in S. In the sequel we denote by $i(a)$ and $j(a)$ respectively the tail and head nodes of $a \in A$, and we let $A_i^+ = \{a : i(a) = i\}$ and $A_i^- = \{a : j(a) = i\}$ be the sets of arcs leaving and entering node $i \in N$.

Demand. $K \subset N$ denotes the set of destinations and $\{d^k\}_{k \in K} \subset \mathbb{R}^N$ the corresponding demands: $d_i^k \ge 0$ is the demand flow rate from node $i \ne k$ to destination k, and $d_k^k = -\sum_{i \ne k} d_i^k$.

Flows. $V := [0, \infty)^A$ is the space of arc-flows, and $\mathcal{V} := V^K$ is the space of destination-arc-flows. A vector $v = (v^k)_{k \in K} \in \mathcal{V}$ is said to be feasible iff $v_a^k = 0$ for all $a \in A_k^+$ (no flow with destination k leaving node k) with each v^k satisfying flow conservation

$$d_i^k + \sum_{a \in A_i^-} v_a^k = \sum_{a \in A_i^+} v_a^k \quad \text{for all } i \in N.$$

The quantity on the left is the flow entering node i with destination k and will be denoted x_i^{kv}.

Time-to-Destination. Given a frequency vector $f = (f_a)_{a \in A}$ with $f_a \in (0, \infty]$ and a travel time vector $t = (t_a)_{a \in A}$ with $t_a \in [0, \infty)$, the time-to-destination corresponding to a given $k \in K$ is the unique solution $\tau^k(f, t) = (\tau_i^k)_{i \in N}$ of the generalized Bellman equations (see Nguyen and Pallotino 1988, Spiess and Florian 1989, and §3)

$$\begin{cases} \tau_k^k = 0, \\ \tau_i^k = \min_{s \subset A_i^+} \dfrac{1 + \sum_{a \in s}(t_a + \tau_{j(a)}^k) f_a}{\sum_{a \in s} f_a} \quad \text{for all } i \ne k. \end{cases} \tag{B}$$

When s includes infinite frequency arcs the expression inside the minimum is interpreted as the average of the times $t_a + \tau_{j(a)}^k$ corresponding to these arcs, while for $s = \emptyset$ this quantity is taken as ∞.

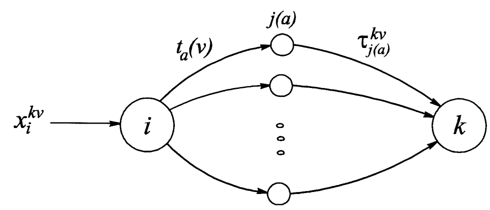

Figure 5 The Local *TEW* Equilibrium Problem

Travel Times and Effective Frequencies. Each arc $a \in A$ is characterized by a continuous travel time function $t_a : \mathcal{V} \to [0, \bar{t}_a)$, where $\bar{t}_a$ is a finite upper bound, and an effective frequency function $f_a : \mathcal{D}_a \subset \mathcal{V} \to (0, \infty]$ which is either identically ∞ with $\mathcal{D}_a = \mathcal{V}$, or everywhere finite and differentiable on $\operatorname{int}(\mathcal{D}_a)$ with $[\partial f_a / \partial v_a^k] < 0$ for all $k \in K$. Let $t(v) = (t_a(v))_{a\in A}$ and $f(v) = (f_a(v))_{a\in A}$, and denote by A^F and A^I the sets of finite and infinite frequency arcs respectively. The dependence of $f_a(\cdot)$ on the vector v reflects the fact that flow on other arcs may use part of the capacity of the service (think of the on-board passengers on a bus arriving at a given stop). Similarly, the dependence of $t_a(\cdot)$ on v can be used to model the passenger discomfort produced by the on-board crowding of buses. We make no assumption (except continuity) on the functions $t_a(\cdot)$, which may eventually be taken as constant.

Given $v \in \mathcal{D}_a$ and $k \in K$, we denote by $f_a^{kv}(\cdot)$ the mapping $f_a(\cdot)$ taken as a function of the flow v_a^k alone, the remaining variables being fixed at the values specified by v. For $a \in A^F$ the domain of $f_a^{kv}(\cdot)$ is assumed to be of the form $[0, \bar{v}_a^{kv})$ with $f_a^{kv}(z) \to 0$ as $z \to \bar{v}_a^{kv}$. Moreover, $f_a^{kv}(\cdot)$ is supposed to be maximal in the fully uncongested situation, namely $f_a^{k0}(\cdot) \geq f_a^{kv}(\cdot)$ for all $v \in \mathcal{V}$. Finally, we denote by $w_a^{kv}(\cdot) : [0, \infty) \to [0, \bar{v}_a^{kv})$ the inverse of $z \mapsto z/f_a^{kv}(z)$ for $a \in A^F$ and $w_a^{kv}(\cdot) \equiv \infty$ for $a \in A^I$.

Local Equilibrium. Given a destination-arc-flow $v \in \mathcal{V}$ we let $E(v) := \prod_{k\in K; i\in N} E_i^k(v) \subset \mathcal{V}$, where $E_i^k(v)$ is the solution set of a *TEW* model (Fig. 5) defined by the flow x_i^{kv} entering node i with destination k, and by the arcs $a \in A_i^+$ with travel times $t_a^{kv} := t_a(v) + \tau_{j(a)}^k(f(v), t(v))$ and effective frequencies $f_a^{kv}(\cdot)$. This applies for $i \neq k$, while for $i = k$ we set $E_k^k(v) = \{0\}$.

According to Theorem 1.2, for $i \neq k$ and provided that $x_i^{kv} < \sum_{a\in A_i^+} \bar{v}_a^{kv}$, we have

$$E_i^k(v) = \left\{ (\tilde{v}_a^k)_{a\in A_i^+} : \begin{array}{ll} \multicolumn{2}{c}{\sum_{a\in A_i^+} \tilde{v}_a^k = x_i^{kv},\ 0 \leq \tilde{v}_a^k \leq w_a^{kv}(\alpha_i^{kv}),} \\ \tilde{v}_a^k = 0 & \text{if } t_a^{kv} > \widehat{T}_i^{kv}(\alpha_i^{kv}), \\ \tilde{v}_a^k = w_a^{kv}(\alpha_i^{kv}) & \text{if } t_a^{kv} < \widehat{T}_i^{kv}(\alpha_i^{kv}), \end{array} \right\} \quad (4)$$

where

$$\widehat{T}_i^{kv}(\alpha) := \min_{s\subset A_i^+} \frac{1 + \sum_{a\in s} t_a^{kv} f_a^{kv}(w_a^{kv}(\alpha))}{\sum_{a\in s} f_a^{kv}(w_a^{kv}(\alpha))}$$

and α_i^{kv} denotes the unique $\alpha \geq 0$ such that

$$\begin{aligned} \sum \{ w_a^{kv}(\alpha) : t_a^{kv} < \widehat{T}_i^{kv}(\alpha) \} &\leq x_i^{kv} \\ &\leq \sum \{ w_a^{kv}(\alpha) : t_a^{kv} \leq \widehat{T}_i^{kv}(\alpha) \}. \end{aligned} \quad (5)$$

Thus, the set-valued map $E : \mathcal{V} \rightrightarrows \mathcal{V}$ has compact convex values with domain

$$\operatorname{dom}(E) = \{ v \in \mathcal{V} : x_i^{kv} < \sum_{a\in A_i^+} \bar{v}_a^{kv} \text{ for all } k \in K,\ i \neq k \}.$$

Observe that a vector $\tilde{v} \in E(v)$ may fail to satisfy flow conservation, but we have

$$\sum_{a\in A_i^+} \tilde{v}_a^k = x_i^{kv} = d_i^k + \sum_{a\in A_i^-} v_a^k \quad \text{for all } i \neq k. \quad (6)$$

Global Equilibrium. Given the graph $G = (N, A)$, the demands $\{d^k : k \in K\}$, the arc travel time functions $t_a(\cdot)$, and the effective frequency functions $f_a(\cdot)$, we say that the destination-arc-flow $v \in \mathcal{V}$ is an equilibrium iff $v \in E(v)$. Using (6), this amounts to saying that v satisfies flow conservation with no flow towards k leaving node k (conservation at k results by adding (6)) and with $(v_a^k)_{a\in A_i^+}$ an equilibrium assignment at each node i and for each destination k.

Remark 2.1. Although hyperpaths and strategy-flows are underlying the model, we formulated it directly in terms of arc-flows. However, let us point out a difference with respect to previous models based on strategies/hyperpaths (Bouzaïene et al. 1995c, Nguyen and Pallotino 1988, Spiess and Florian 1989, Wu et al. 1994) namely, in the model above all passengers reaching an intermediate node i and heading towards the same destination k are mixed in x_i^{kv}, and we do not distinguish them in terms of their origin. This reduction allows for a simpler description of the system since we do not have to keep track of the different origins.

2.1. Existence of a Transit Network Equilibrium

To prove the existence of an equilibrium, we use Kakutani's fixed-point theorem (Berge 1997, Chapter VIII, §2). We shall restrict ourselves to the case where each destination $k \in K$ can be reached from every $i \neq k$ by following a path of infinite frequency arcs. This assumption implies $\operatorname{dom}(E) = \mathcal{V}$ and establishes a uniform bound on the time-to-destination $\tau_i^k(f(v), t(v)) \leq \tau_\infty$ for all $v \in \mathcal{V}$. This provides in turn a priori bounds on the arc-flows, needed for the fixed point argument. The condition may look stringent but it suffices for practical purposes: a passenger can always walk to his destination even if it takes very long. This is not to say that at equilibrium such "walking-paths" will be used, which is rather unlikely given the optimal character of equilibrium strategies. Let us establish some preliminary lemmas.

Lemma 2.1. *Let $\bar{\alpha}_i^k$ be such that $\sum_{a \in A_i^F} f_a^{k0}(w_a^{k0}(\bar{\alpha}_i^k)) \leq 1/\tau_\infty$ with $A_i^F = A^F \cap A_i^+$. Then $\alpha_i^{kv} \leq \bar{\alpha}_i^k$.*

Proof. The assumption $f_a^{k0}(\cdot) \geq f_a^{kv}(\cdot)$ implies $w_a^{kv}(\cdot) \leq w_a^{k0}(\cdot)$ so that

$$\frac{1}{\sum_{a \in A_i^F} f_a^{k0}(w_a^{k0}(\alpha))} = \frac{\alpha}{\sum_{a \in A_i^F} w_a^{k0}(\alpha)} \leq \frac{\alpha}{\sum_{a \in A_i^F} w_a^{kv}(\alpha)} = \frac{1}{\sum_{a \in A_i^F} f_a^{kv}(w_a^{kv}(\alpha))},$$

and then

$$\tau_\infty \leq \frac{1}{\sum_{a \in A_i^F} f_a^{kv}(w_a^{kv}(\bar{\alpha}_i^k))} \leq \min_{s \subset A_i^F} \frac{1 + \sum_{a \in s} t_a^{kv} f_a^{kv}(w_a^{kv}(\bar{\alpha}_i^k))}{\sum_{a \in s} f_a^{kv}(w_a^{kv}(\bar{\alpha}_i^k))}.$$

Using remark (b) in §1.3 we get $\alpha_i^{kv} \leq \bar{\alpha}_i^k$ as claimed. □

Lemma 2.2. *$E(\cdot)$ is upper-semicontinuous (usc) and the following functions are continuous:*
- (a) $v \mapsto x_i^{kv}$.
- (b) $(v, \alpha) \mapsto w_a^{kv}(\alpha)$.
- (c) $(f, t) \mapsto \tau^k(f, t)$.
- (d) $v \mapsto t_a^{kv}$.
- (e) $(v, \alpha) \mapsto \widehat{T}_i^{kv}(\alpha)$.
- (f) $v \mapsto \alpha_i^{kv}$.

Proof. Upper semicontinuity of $E(\cdot)$ follows easily from (4) and the continuity of (a)–(f). Now, (a) and (b) are straightforward while (c) will be proved in Corollary 3.1. Also (d) follows immediately from (c) while (e) is an easy consequence of (d) and (b), so we must only prove (f).

Let $v^n \to v$ and set $\alpha^n := \alpha_i^{kv_n}$, $x^n := x_i^{kv_n}$, and $w_a^n := w_a^{kv_n}(\alpha_i^{kv_n})$. Consider an accumulation point $\alpha^\infty = \lim \alpha^{n_j}$ with $n_j \to \infty$. If $t_a^{kv} < \widehat{T}_i^{kv}(\alpha^\infty)$ or $t_a^{kv} > \widehat{T}_i^{kv}(\alpha^\infty)$, using (d) and (e) we see that the same inequalities hold for j large enough with v and α^∞ replaced by v^{n_j} and α^{n_j} respectively, and then (5) implies

$$\sum \{ w_a^{n_j} : t_a^{kv} < \widehat{T}_i^{kv}(\alpha_\infty) \} \leq x^{n_j} \leq \sum \{ w_a^{n_j} : t_a^{kv} \leq \widehat{T}_i^{kv}(\alpha_\infty) \}.$$

Continuity of (a) and (b) gives $x^{n_j} \to x_i^{kv}$ and $w_a^{n_j} \to w_a^{kv}(\alpha^\infty)$, so that letting $j \to \infty$ we get $\alpha^\infty = \alpha_i^{kv}$. Since Lemma 2.1 implies that α^n is bounded, we conclude $\alpha^n \to \alpha_i^{kv}$ which completes the proof. □

Since $E(\cdot)$ is usc with compact convex values, proving the existence of a transit network equilibrium reduces to finding a compact convex set $\mathcal{K} \subset \mathcal{V}$ with $E(\mathcal{K}) \subset \mathcal{K}$. The difficulty with such a priori bounds comes from the fact that a local equilibrium $\tilde{v} \in E(v)$ need not satisfy flow conservation. To overcome this problem we analyze the general equilibrium problem by reduction to the following special case.

Lemma 2.3. *Suppose that for each $i, j \in N$ there is a unique infinite frequency arc from i to j, and that all these arcs have the same travel time $t_\infty > 0$. Then there exists an equilibrium.*

Proof. Let $\bar{\alpha}_i^k$ be taken as in Lemma 2.1 with $\tau_\infty = t_\infty$. Setting $\bar{v}_a^k := w_a^{k0}(\bar{\alpha}_i^k)$, for each $v \in \mathcal{V}$ and $\tilde{v} \in E(v)$ we have $\tilde{v}_a^k \leq w_a^{kv}(\alpha_i^{kv}) \leq w_a^{k0}(\alpha_i^{kv}) \leq \bar{v}_a^k$ for $a \in A^F$. Now, since all the infinite frequency arcs have the same travel time, the only arcs $a \in A^I$ which may carry a positive flow $\tilde{v}_a^k > 0$ towards k are the ones connecting directly to this destination. Thus, for $a \in A^I$ we have $\tilde{v}_a^k \leq \bar{v}_a^k$ with $\bar{v}_a^k$ equal to the sum of the bounds $\bar{v}_{a'}^k$, on the finite frequency arcs a' entering the tail node of a. Hence the compact convex set $\mathcal{K} = \{v \in \mathcal{V} : v_a^k \leq \bar{v}_a^k\}$ satisfies $E(\mathcal{K}) \subset \mathcal{K}$, and Kakutani's fixed point theorem implies the existence of $v \in \mathcal{V}$ such that $v \in E(v)$. □

We may now prove the announced result on the existence of transit network equilibria.

THEOREM 2.1. *Let $G = (N, A)$ be a graph with arc travel time functions $t_a(\cdot)$ and effective frequency functions $f_a(\cdot)$ as before, and let $\{d^k : k \in K\}$ be a family of demand vectors. Suppose that for each $k \in K$ and every $i \neq k$ there is a path from i to k in the infinite frequency subgraph $G^I = (N, A^I)$. Then there exists a transit network equilibrium $v \in \mathcal{V}$, i.e., a solution of $v \in E(v)$.*

PROOF. The existence of infinite frequency paths joining each i and k implies $\mathrm{dom}(E) = \mathcal{V}$ as well as $\tau_i^k(f(v), t(v)) \leq \bar{\tau}_i^k$ for each $v \in \mathcal{V}$, where $\bar{\tau}_i^k$ denotes the time of a shortest path from i to k in G^I computed with the arc travel time bounds $\bar{t}_a$.

Using Lemma 2.3, for each $r > 0$ we may find an equilibrium $v(r)$ for an auxiliary network G^r constructed from G by assigning arc travel time functions $t_a^r(v) := t_a(v) + 1/r$ to each $a \in A$, finite frequency functions $f_a^r(v) := r/(1 + \sum_{k \in K} v_a^k)$ to the arcs $a \in A^I$, and adding artificial arcs from each $i \in N$ to every $j \in N$ with infinite frequency and travel time $t_\infty = r$.

Since each arc in G^r has strictly positive travel time, a flow $v_a^k(r) > 0$ implies that the transit time from $j(a)$ to k is strictly smaller than the time from $i(a)$ to k. It follows that $v(r)$ cannot send flow along a directed cycle for any destination $k \in K$, which combined with flow conservation implies that $v(r)$ is uniformly bounded. This implies in turn that when $r \to \infty$ the time of a shortest path from i to k in G^I computed with times $t_a^r(v(r)) + 1/f_a^r(v(r))$ will eventually become smaller than $\bar{\tau}_i^k$ and a fortiori smaller than $t_\infty = r$, so that the artificial arcs cannot carry a positive flow.

Redefine the vector $v(r)$ by removing the zero components corresponding to the artificial arcs. Clearly, this restricted $v(r)$ is an equilibrium in the network G, namely

$$\begin{cases} \sum_{a \in A_i^+} v_a^k(r) = x_i^{kv(r)}, & 0 \leq v_a^k(r) \leq w_a^{kv(r)}(\alpha_i^{kv(r)}), \\ v_a^k(r) = 0 & \text{if } t_a^{kv(r)} > \widehat{T}_i^{kv(r)}(\alpha_i^{kv(r)}), \\ v_a^k(r) = w_a^{kv(r)}(\alpha_i^{kv(r)}) & \text{if } t_a^{kv(r)} < \widehat{T}_i^{kv(r)}(\alpha_i^{kv(r)}), \end{cases} \qquad (S^r)$$

with all the involved functions computed according to the modified frequencies $f_a^r(\cdot)$ and times $t_a^r(\cdot)$.

Let $v = \lim v(r_n)$ be an accumulation point of $v(r)$ with $r_n \to \infty$. We claim that v satisfies the equilibrium conditions (S^r) for the original network G with travel time functions $t_a(\cdot)$ and infinite frequencies on the arcs $a \in A^I$. To prove this it suffices to let $n \to \infty$ in the conditions (S^{r_n}) satisfied by $v(r_n)$, provided that all the involved quantities converge to the appropriate limits.

To check the latter set $v_n = v(r_n)$ and $t^n = t^{r_n}(v_n)$, and let f^n be the vector with components $f_a(v_n)$ for $a \in A^F$ and $f_a^{r_n}(v_n)$ for $a \in A^I$. Clearly $t^n \to t(v)$ and $f^n \to f(v)$ so that Corollary 3.1 implies

$$\begin{aligned} t_a^{kv_n} &:= t_a^n + \tau_{j(a)}^k(f^n, t^n) \longrightarrow t_a^{kv} \\ &:= t_a(v) + \tau_{j(a)}^k(f(v), t(v)). \end{aligned}$$

It is also clear that $x_i^{kv_n} \to x_i^{kv}$. Now, a slight modification of the proof of (f) in Lemma 2.2 to take into account the variation of frequencies $f^n \to f(v)$ and times $t^n \to t(v)$, implies that the equilibrium values $\alpha_i^{kv_n}$ converge towards α_i^{kv}. From this it follows easily that $w_a^{kv_n}(\alpha_i^{kv_n}) \to w_a^{kv}(\alpha_i^{kv})$ and $\widehat{T}_i^{kv_n}(\alpha_i^{kv_n}) \to \widehat{T}_i^{kv}(\alpha_i^{kv})$. All these convergences allow us to pass to the limit in (S^{r_n}). $\square$

3. Time-to-Destination and Shortest Hyperpaths

Let $G = (N, A)$ be a directed graph with each arc $a \in A$ having an associated frequency $f_a \in (0, \infty)$ and a travel time $t_a \in [0, \infty)$. Set $f = (f_a)_{a \in A}$ and $t = (t_a)_{a \in A}$, and denote $\mathcal{F} = (0, \infty)^A$ and $\mathcal{T} = [0, \infty)^A$. Consider a fixed $k \in N$ and suppose that for each $i \neq k$ there is a directed path from i to k in G. The time-to-destination for destination k is the solution of the generalized Bellman equations (see Nguyen and Pallotino 1988, Spiess and Florian 1989)

$$\begin{cases} \tau_k = 0, \\ \tau_i = \min\limits_{s \subset A_i^+} \dfrac{1 + \sum_{a \in s}(t_a + \tau_{j(a)}) f_a}{\sum_{a \in s} f_a} \qquad \text{for all } i \neq k. \end{cases} \qquad (B)$$

PROPOSITION 1. *For each $(f, t) \in \mathcal{F} \times \mathcal{T}$ there is a unique solution $\tau^k(f, t) = (\tau_i)_{i \in N}$ of (B) which, moreover, depends continuously on (f, t).*

Proof. Let $M : [0,\infty]^N \to [0,\infty]^N$ with $M_i(\tau) = \min_{s\subset A_i^+}[1+\sum_{a\in s}(t_a+\tau_{j(a)})f_a]/\sum_{a\in s} f_a$ for $i \neq k$ and $M_k(\tau)=0$. Then, (B) corresponds to the fixed point equation $\tau = M(\tau)$. Notice that M is componentwise monotone: $\tau^1 \le \tau^2 \Rightarrow M(\tau^1) \le M(\tau^2)$.

Existence. Consider the iteration $\tau^{n+1} = M(\tau^n)$ with τ^0 given by $\tau_i^0 = \infty$ for all i. Clearly $\tau^1 \le \tau^0$ and inductively we get $\tau^{n+1} \le \tau^n$. Moreover, since every node i can be connected to k by a path with at most $|N|-1$ arcs, it follows that for $n \ge |N|$ the vector τ^n is finite. Then τ^n monotonically converges towards a certain $\tau \ge 0$, and continuity of M implies $\tau = M(\tau)$.

Uniqueness. Let $\tau^1 \neq \tau^2$ be two solutions and set $d = \max_{i\in N} \tau_i^1 - \tau_i^2$. Exchanging τ^1 and τ^2 we may suppose $d > 0$. Let i_0 be such that $\tau_{i_0}^1 - \tau_{i_0}^2 = d$, and take $s_0 \subset A_{i_0}^+$ attaining the minimum for $M_{i_0}(\tau^2)$. Then

$$d \le \frac{1+\sum_{a\in s_0}(t_a+\tau^1_{j(a)})f_a}{\sum_{a\in s_0} f_a} - \frac{1+\sum_{a\in s_0}(t_a+\tau^2_{j(a)})f_a}{\sum_{a\in s_0} f_a} = \frac{\sum_{a\in s_0} f_a(\tau^1_{j(a)}-\tau^2_{j(a)})}{\sum_{a\in s_0} f_a}.$$

The right-hand side is a strict convex combination of the quantities $\tau^1_{j(a)} - \tau^2_{j(a)}$, which are bounded from above by d so they must all be equal to d. Take $i_1 = j(a)$ with $a \in s_0$ having minimal $t_a + \tau^2_{j(a)}$. Proceeding inductively from this i_1, we find a sequence of nodes with $\tau_i^1 - \tau_i^2 = d$ for $i = i_0, i_1, i_2, \ldots$. Since the quantities τ_i^2 are strictly decreasing along this sequence, the nodes are all different so that at some point we get $i_j = k$, and then $\tau_k^1 - \tau_k^2 = d > 0$, which gives a contradiction.

Continuity. Let $(f^n, t^n) \to (f, t)$ and set $\tau^n := \tau^k(f^n, t^n)$. Consider the mapping M as a function of (τ, f, t), namely, $(\tau, f, t) \mapsto M(\tau, f, t)$ which is clearly continuous on $[0,\infty)^N \times \mathcal{F} \times \mathcal{T}$. Let $\tau = \lim \tau^{n_j}$ be an accumulation point of τ^n with $n_j \to \infty$. Letting $j \to \infty$ in the equality $\tau^{n_j} = M(\tau^{n_j}, f^{n_j}, t^{n_j})$ we obtain $\tau = M(\tau, f, t)$, so that by uniqueness we get $\tau = \tau^k(f, t)$. Since the sequence τ^n is clearly bounded, we deduce $\tau^n \to \tau^k(f, t)$ and therefore $\tau^k(\cdot,\cdot)$ is continuous. □

It is easy to see that $\tau^k(\cdot,\cdot)$ is nondecreasing with respect to each time variable t_a separately, and also nonincreasing with each frequency f_a. As a matter of fact, if $t_a + \tau^k_{j(a)}(f, t) < \tau^k_{i(a)}(f, t)$, then an increment in t_a or f_a will induce respectively an increase or a decrease of τ^k at the tail node $i(a)$, and inductively at any other node whose optimal strategy leads to $i(a)$. Otherwise, an increase in t_a or f_a will not affect τ^k. This monotonicity allows us to define $\tau^k(f, t)$ when some components of f are infinite. Namely, for $t \in \mathcal{T}$ and $f \in \mathcal{F}^\infty := (0,\infty]^A$ we let f^r be the vector with components $f_a^r = \min\{f_a, r\}$ and define $\tau^k(f, t)$ as the monotonically decreasing limit

$$\tau^k(f, t) = \lim_{r\to\infty} \tau^k(f^r, t).$$

Corollary 3.1. *The extended mapping $\tau^k : \mathcal{F}^\infty \times \mathcal{T} \to [0,\infty)^N$ is continuous.*

3.1. Hyperpath-Dijkstra

Time-to-destination and shortest hyperpaths can be computed using the algorithm given in Spiess (1984) and Spiess and Florian (1989). The existence proof in Proposition 3.1 suggests the alternative method $\tau^{n+1} = M(\tau^n)$ which can be proved to be finitely convergent, but not very efficient in terms of computational complexity. We propose a third method which is an adaptation of Dijkstra's shortest path algorithm. The method iteratively updates a time vector τ and a set of *solved nodes* S—containing those j's for which the current τ_j already gives the time-to-destination—adjusting the times τ_i to give the optimal transit time from i to k by using a strategy s^i that jumps directly from i to S. An auxiliary variable f^i accumulates the frequencies of the arcs in s^i.

Hyperpath-Dijkstra
Initialize: $\tau_i = \infty$, $s^i = \emptyset$, and $f^i = 0$ for $i \neq k$;
    $\tau_k = 0$; $S = \emptyset$;
while $(S \neq N)$ **do**
{find $j \notin S$ with smallest τ_j and update $S \leftarrow S \cup \{j\}$
  for $(a \in A$ with $j(a) = j$ and $i(a) \notin S)$ **do**
  {set $i = i(a)$ and $\tilde{t}_a = t_a + \tau_j$
    if $(\tilde{t}_a < \tau_i)$ **then** add(a)
    while $(\exists\, b \in s^i$ with $\tilde{t}_b \ge \tau_i)$ **do** remove(b)
  }
}

The **for** loop updates τ_i and s^i for those i's connected to the newly incorporated node $j \in S$, in order to keep τ_i equal to the optimal time from i to k using a strategy s^i jumping directly into S. The add(a) phase incorporates the arc a to the strategy s^i (if this helps reducing τ_i) as follows:

$$\{s^i \leftarrow s^i \cup \{a\};\ \tau_i \leftarrow (f^i\tau_i + f_a\tilde{t}_a)/(f^i + f_a);\ f^i \leftarrow f^i + f_a\}$$

with the convention $f^i\tau_i = 1$ for the initial case $f^i = 0, \tau_i = \infty$. As a consequence of the reduction of τ_i, some arcs b added to s^i on previous iterations may no longer be advantageous and are eliminated in the REMOVE(b) phase by doing

$$\{s^i \leftarrow s^i \backslash \{b\};\ \tau_i \leftarrow (f^i\tau_i - f_b\tilde{t}_b)/(f^i - f_b);\ f^i \leftarrow f^i - f_b\}.$$

Notice that REMOVE(b) further reduces the time τ_i affecting the test in the inner **while** loop, implying eventually the removal of some additional arcs. The justification for the ADD and REMOVE steps is given by Lemma 1.1.

The selection of $j \notin S$ with minimal τ_j can be done efficiently by keeping S^c as an ordered list. Using a heap-list (where insertion has logarithmic computational cost; Cormen et al. 1991) the overall effort is $O(|N|\ln|N|)$ comparisons. Similarly, for the inner **while** loop we may keep the s^i's as ordered heap-lists, involving a total work of $O(\sum_{i\in N}|A_i^+|\ln|A_i^+|)$ comparisons, which can be bounded from above by $O(|A|\ln\delta^+)$, where $\delta^+ = \max\{|A_i^+| : i \in N\}$ denotes the maximal out-degree. Notice that $\delta^+ \leq |N|$.

PROPOSITION 2. *Hyperpath-Dijkstra computes the time-to-destination $\tau^k(f,t)$ and corresponding optimal strategies in at most $O(|N|\ln|N| + |A|\ln\delta^+)$ comparisons and $O(|A|)$ arithmetic operations.*

PROOF. Let us prove inductively that on each stage we have $\tau_j = \tau_j^k(f,t)$ for all $j \in S$. This is obvious for the first iteration in which node k is added to S with $\tau_k = 0$. Suppose by induction that at a given stage all nodes in S have the correct time-to-destination and let $j \notin S$ with minimal τ_j. Consider for each $i \in N$ the optimal strategy $\hat{s}^i$ given by Corollary 1.1, so that $\tau_{j(a)}^k(f,t) \leq t_a + \tau_{j(a)}^k(f,t) < \tau_i^k(f,t)$ for all $a \in \hat{s}^i$. These strict inequalities imply that starting from j and following the strategies $\hat{s}^i$, we must reach a node $i \notin S$ having all its successors by $\hat{s}^i$ in S. For this i we clearly have $\tau_i = \tau_i^k(f,t)$, and then $\tau_j^k(f,t) \geq \tau_i^k(f,t) = \tau_i \geq \tau_j$. Since obviously $\tau_j \geq \tau_j^k(f,t)$, we get $\tau_j = \tau_j^k(f,t)$ and j can be added to S, completing the induction step.

To estimate the worst-case complexity of the method we notice that each arc a is processed at most twice: once in the ADD phase and possibly a second time in a REMOVE step. The overall work involved in these operations is $O(|A|)$ (in fact, at most $11|A|$ arithmetic operations). Handling S^c and the s^i's as heap-lists dominates the number of comparisons with complexity $O(|N|\ln|N| + |A|\ln\delta^+)$. □

The algorithm can be easily adapted to handle infinite frequencies as follows:

HYPERPATH-DIJKSTRA
Initialize: $\tau_i = \tau_i^\infty = \infty$, $s^i = s_\infty^i = \emptyset$, and $f^i = 0$
  for $i \neq k$; $\tau_k = \tau_k^\infty = 0$; $S = \emptyset$;
while ($S \neq N$) **do**
{ find $j \notin S$ with smallest $\bar{\tau}_j = \min\{\tau_j, \tau_j^\infty\}$
  and update $S \leftarrow S \cup \{j\}$
 for ($a \in A$ with $j(a) = j$ and $i(a) \notin S$) **do**
 {set $i = i(a)$ and $\tilde{t}_a = t_a + \bar{\tau}_j$
  if ($f_a = \infty$) and ($\tilde{t}_a < \tau_i^\infty$) **then** $\{s_\infty^i \leftarrow \{a\};$
  $\tau_i^\infty \leftarrow \tilde{t}_a\}$
  if ($f_a < \infty$) and ($\tilde{t}_a < \tau_i$) **then** ADD(a)
  while ($\exists\ b \in s^i$ with $\tilde{t}_b \geq \tau_i$) **do** REMOVE(b)
 }
}

REMARK 3.1. The algorithm proposed by Spiess and Florian (1984, 1989) is based on a different idea: after initializing as in Hyperpath-Dijkstra, each loop of the method selects an arc $a = (i,j)$ with minimal $t_a + \tau_j$ among the arcs not previously processed. If $t_a + \tau_j$ is smaller than the current τ_i, then ADD(a) is used to incorporate a to the strategy s^i and to update τ_i and f^i. Otherwise a is simply discarded.

In contrast with our method, this algorithm does not require a REMOVE(b) phase and the optimal s^i and τ_i are only obtained upon completion of the algorithm while Hyperpath-Dijkstra solves one node per loop (just like the classical shortest path Dijkstra). Concerning the worst-case computational complexity, a heap-list implementation runs in $O(|A|\ln|A|)$ $=O(|A|\ln|N|)$ comparisons and $O(|A|\delta^-)$ arithmetic operations, with $\delta^- = \max\{|A_i^-| : i \in N\}$ the maximal in-degree. Thus Hyperpath-Dijkstra attains a better complexity bound, at the price of a slightly more complex programming. In any case, both methods attain low computational complexities, proving in fact that *shortest hyperpath is not harder than shortest path*. From a practical perspective, the apparently useless work involved in Hyperpath-Dijkstra when adding an arc which may eventually be removed later, balances against the fact that not all arcs will be processed (after j is added to S the arcs in A_j^+ are no longer considered). Thus, it is probably the case that each method outperforms the other depending on the particular instance being solved. Such an empirical comparison has not been done.

Acknowledgments

This work was supported by the FONDAP grant in Applied Mathematics (FONDECYT-Chile). The paper benefited greatly from discussions with Professor Jaime San Martín on the queue-theoretic treatment of congestion (Appendix A) and the optimization equivalent for the *TEW* equilibrium model. Constructive criticism by three anonymous referees also contributed significantly to enlarge the scope of the paper and to improve its presentation.

4. Appendix A: Waiting Times and Effective Frequencies

In this section we provide a queue-theoretic support for the congestion model adopted in §1, justifying the formulas

$$W_s = \frac{1}{\sum_{i\in s} f_i(v_i)},$$

$$\pi_i^s = \frac{f_i(v_i)}{\sum_{j\in s} f_j(v_j)}$$

and giving analytic expressions for the effective frequency functions $f_i(v_i)$ as well as for the inverse $w_i(\alpha)$ of $v_i \mapsto v_i/f_i(v_i)$.

Most of this material is a slight variation of standard models which can be found e.g. in Serfozo (1999) and Tijms (1994). The analysis concerns only the case of Poisson queues, so that further investigation is required to validate the congestion model in other situations. Furthermore, we point out a limitation of the *TEW* equilibrium model coming from the fact that, in the case of several overlapping strategies, the effective frequency functions $f_i(\cdot)$ may depend on the strategy-flow variables y_s and not only on the line-flows v_i.

4.1. Waiting Time and Effective Frequency for a Bus Line

Consider a bus stop with passengers arriving according to a Poisson process of rate v (Figure 6). The bus stop is served by a bus line with independent Poisson arrivals of rate μ and variable capacity C with $\mathbb{P}(C=j)=q_j$ for $j=0,\ldots,K$.

Figure 6 Queuing for a Single-Line Service

If the available capacity of an arriving bus is larger than the queue length, the latter reduces to zero. Hence, the queue length is a continuous time Markov chain with transition rates

$$\begin{cases} p_{k,k+1} = v & \text{for } k\geq 0, \\ p_{k+j,k} = \mu q_j & \text{for } k\geq 1 \text{ and } 0\leq j\leq K, \\ p_{k,0} = \mu\sum_{j=k}^K q_j & \text{for } 1\leq k\leq K, \end{cases}$$

and therefore its stationary distribution $\{\pi_k\}_{k\geq 0}$ is characterized by the balance equations

$$\begin{cases} (v+\mu)\pi_k = v\pi_{k-1} + \mu\sum_{j=0}^K q_j\pi_{k+j} & \text{for } k\geq 1, \\ (v+\mu)\pi_0 = \mu\sum_{j=0}^K q_j\left[\sum_{k=0}^j \pi_k\right], \end{cases}$$

whose solution is $\{\pi_k = (1-\rho)\rho^k\}_{k\geq 0}$ with $\rho=\rho(v)\in[0,1)$ such that

$$\mu\sum_{j=0}^K q_j(\rho+\cdots+\rho^j) = v. \tag{7}$$

The expression on the right-hand side is a convex increasing function of ρ, so that (7) has a unique solution in $[0,1)$ provided $v\in[0,\bar{v})$, where $\bar{v} := \mu\sum_{j=0}^\infty jq_j$ is the expected flow capacity of the line. The expected queue length may be computed explicitly as

$$\mathbb{E}(L) = \sum_{k=0}^\infty k\pi_k = (1-\rho(v))\sum_{k=1}^\infty k\rho(v)^k = \frac{\rho(v)}{1-\rho(v)}$$

which explodes as $v\to\bar{v}$, so that $\bar{v}$ can be interpreted as the saturation flow of the line. Also, using Little's formula, the expected waiting time turns out to be

$$W(v) = \frac{1}{v}\mathbb{E}(L) = \frac{1}{v}\frac{\rho(v)}{1-\rho(v)}$$

which also tends to ∞ as $v\to\bar{v}$. This function is differentiable with $W(v)>0$ and $W'(v)>0$ for $v\in[0,\bar{v})$, so that we may define the effective frequency function $f:[0,\bar{v})\to(0,\infty)$ as

$$f(v) := \frac{1}{W(v)} = v\left[\frac{1}{\rho(v)}-1\right] \tag{8}$$

which is also differentiable with $f'(v)<0$ and $f(v)\to 0^+$ when $v\to\bar{v}$.

To compute the inverse $w(\alpha)$ of $v\mapsto v/f(v)$ we observe that $v/f(v)=\alpha$ iff $\rho(v)=\alpha/(1+\alpha)$, whose solution is easily obtained by replacing $\rho=\alpha/(1+\alpha)$ in (7), yielding the concave increasing function

$$w(\alpha) = \mu\alpha\sum_{j=0}^K q_j\left[1-\left(\frac{\alpha}{1+\alpha}\right)^j\right]. \tag{9}$$

4.2. Waiting Time and Boarding Probabilities for a Strategy

Consider now a bus stop served by a set of lines $\{l_i : i \in s\}$ with independent Poisson arrivals of rate μ_i and variable capacities C_i with $\mathbb{P}(C_i = j) = q_j^i$ for $j = 0, \ldots, K_i$. Passengers arrive according to an independent Poisson process of rate y (Figure 7).

The balance equations for the stationary distribution of the queue length are in this case

$$\begin{cases} (y + \sum_{i\in s} \mu_i)\pi_k = y\pi_{k-1} + \sum_{i\in s} \mu_i \sum_{j=0}^{K_i} q_j^i \pi_{k+j} & \text{for } k \geq 1, \\ (y + \sum_{i\in s} \mu_i)\pi_0 = \sum_{i\in s} \mu_i \sum_{j=0}^{K_i} q_j^i [\sum_{k=0}^{j} \pi_k], \end{cases}$$

whose solution is again of the form $\{\pi_k = (1-\rho)\rho^k\}_{k\geq 0}$ with $\rho = \rho(y) \in [0, 1)$ such that

$$\sum_{i\in s} \mu_i \sum_{j=0}^{K_i} q_j^i (\rho + \cdots + \rho^j) = y.$$

As before, using Little's formula we obtain the expected waiting time

$$W(y) = \frac{1}{y}\mathbb{E}(L) = \frac{1}{y}\frac{\rho(y)}{1-\rho(y)}.$$

Now, the probability of boarding line i is $\pi_i^s = v_i/y$ with v_i the expected flow on line i. The latter is equal to $\mu_i z_i$ with z_i the average number of passengers boarding on each arrival of line i so that, conditioning on the available capacity of the bus and the queue length, a straightforward computation yields

$$\begin{aligned} v_i = \mu_i z_i &= \mu_i \sum_{j=0}^{K_i} q_j^i \left[\sum_{k=0}^{j} k\pi_k + j\sum_{k>j} \pi_k\right] \\ &= \mu_i \sum_{j=0}^{K_i} q_j^i [\rho(y) + \cdots + \rho(y)^j]. \end{aligned}$$

Thus, denoting by $\rho_i(\cdot)$ and $f_i(\cdot)$ the root and effective frequency functions corresponding to line l_i as defined by (7) and (8), we have $\rho_i(v_i) = \rho(y)$ and therefore

$$\frac{v_i}{f_i(v_i)} = \frac{\rho_i(v_i)}{1-\rho_i(v_i)} = \frac{\rho(y)}{1-\rho(y)} = yW(y) \quad \text{for all } i \in s.$$

since $\sum_{i\in s} v_i = y$, we deduce

$$W(y) = \frac{1}{\sum_{j\in s} f_j(v_j)}$$

as well as

$$\pi_i^s = \frac{f_i(v_i)}{\sum_{j\in s} f_j(v_j)}$$

with the expected flow $v = (v_i)_{i\in s}$ being the unique solution of the system

$$v_i = y\frac{f_i(v_i)}{\sum_{j\in s} f_j(v_j)} \quad \text{for all } i \in s. \qquad (E)$$

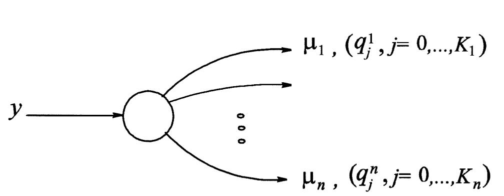

Figure 7 Queuing for a Multiple-Line Service

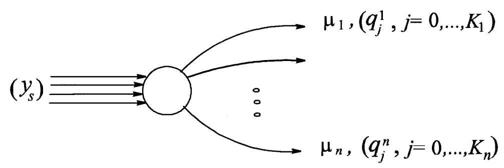

Figure 8 Queuing for Multiple Overlapping Strategies

4.3. Waiting Times for Multiple Overlapping Strategies

Consider next a bus stop served by a set of lines $\{l_1, \ldots, l_n\}$ as in §4.2, but where arriving passengers may have different strategies $s \subset A = \{1, \ldots, n\}$, each one corresponding to an independent Poisson process $\{N^s(t)\}_{t\geq 0}$ of rate y_s (Figure 8). In this situation one has a different queue for each strategy s, and these queues interfere with one another as soon as they share one or more lines. The explicit computation of the stationary distribution for the queue lengths $L^s(t)$ becomes rather intricate, so we adopt a different approach.

Let $\{B_i(t)\}_{t\geq 0}$ denote the Poisson process corresponding to the number of arrivals of line l_i up to time t, and let $\{B_s(t)\}_{t\geq 0}$ be the merging of $\{B_i(t) : i \in s\}$, which is also Poisson with intensity $\mu_s := \sum_{i\in s} \mu_i$. Consider a specific passenger joining the queue L^s, and let p_s be his probability of boarding the first incoming bus from s. Let Y_s be the random variable representing the time between the arrival of this passenger and the next arrival of a bus from strategy s (the excess life), and let $X_s^1, X_s^2, \ldots$ be the subsequent interarrival times for the process $B_s(t)$. Then, the waiting time of this passenger will be Y_s with probability p_s, $Y_s + X_s^1$ with probability $(1-p_s)p_s$, $Y_s + X_s^1 + X_s^2$ with probability $(1-p_s)^2 p_s$, and so on. Since all the involved processes are Poisson and independent, the variables Y_s and X_s^k are exponentially distributed with $\mathbb{E}(Y_s) = \mathbb{E}(X_s^k) = 1/\mu_s$, and then the expected waiting time for strategy s may be computed as

$$W_s = \frac{1}{\mu_s}\sum_{k=0}^{\infty}(k+1)(1-p_s)^k p_s = \frac{1}{\mu_s p_s}. \qquad (10)$$

The probability $p_s = p_s(y)$ is a function of the rate vector $y = (y_s)_s$ (as well as the rates and capacities of the services, which are however considered as fixed). In particular, the expression of W_s for strategy $s = \{i\}$ suggests defining $f_i(y) := 1/W_{\{i\}} = \mu_i p_{\{i\}}$ as the *effective frequency function* of line i. With this we obviously have $W_{\{i\}} = 1/f_i(y)$ but also, more generally, for every strategy s we get

$$W_s = \frac{1}{\sum_{i\in s} f_i(y)}. \qquad (11)$$

Indeed, consider a passenger in queue L^s. The conditional probability that a bus from strategy s arriving at the bus stop belongs to line $i \in s$ is μ_i/μ_s. Once such a bus arrives, the boarding probability of this passenger is the same as the passengers in queue $L^{\{i\}}$, and therefore Bayes' formula gives

$$p_s = \sum_{i\in s} \frac{\mu_i}{\mu_s} p_{\{i\}} = \frac{1}{\mu_s} \sum_{i\in s} f_i(y)$$

which plugged into (10) yields (11).

Let us compute next the probability π_i^s that a passenger of strategy s boards line $i \in s$. Conditioning on the first arrival of a bus from strategy s we have

$$\pi_i^s = \sum_{j \in s} \mathbb{P}[\text{boarding } l_i/l_j \text{ comes first}] \frac{\mu_j}{\mu_s}.$$

The probability $\mathbb{P}[\text{boarding } l_i/l_j \text{ comes first}]$ is equal to $(1-p_{\{j\}})\pi_i^s$ for $j \neq i$ (not boarding the present bus from line l_j and boarding line l_i later on), and equal to $p_{\{i\}} + (1-p_{\{i\}})\pi_i^s$ for $j = i$ (boarding the present bus from line l_i *or* not boarding it and boarding line l_i later on). Therefore,

$$\pi_i^s = \frac{p_{\{i\}}\mu_i + \pi_i^s \sum_{j\in s}(1-p_{\{j\}})\mu_j}{\mu_s}$$

from which we get

$$\pi_i^s = \frac{\mu_i p_{\{i\}}}{\sum_{j\in s} \mu_j p_{\{j\}}} = \frac{f_i(y)}{\sum_{j\in s} f_j(y)}. \tag{12}$$

According to these probabilities the expected value of the flow on line l_i is

$$v_i = \sum_{s\in\mathcal{S}_i} y_s \pi_i^s = \sum_{s\in\mathcal{S}_i} y_s \frac{f_i(y)}{\sum_{j\in s} f_j(y)}$$

which corresponds to (E) but with frequency functions $f_i(\cdot)$ depending on y and not just on v_i.

Comments. We have not been able to obtain analytic expressions nor closed form solutions for the functions $f_i(y)$ appearing in the expressions for W_s and π_i^s. Using simulation we have confirmed all the previous results including the fact that, when there is flow on two or more strategies, the effective frequency function of line i cannot be expressed as a function of v_i alone as in the previous subsections, nor even as a function of the whole vector of line flows v. It appears that $f_i(y)$ depends in an essential and subtle manner on the whole rate vector y, reflecting the complex interactions among the different queues L^s.

As a consequence, the model presented in §1 may only be considered as a first approximation. However, in view of the results in §4.2, the equilibrium predicted by the *TEW* equilibrium model is fully justified when there is a unique strategy s^* carrying flow. As seen from Theorem 1.2 and the remark following it (see also Example 1.1), this occurs when the total flow x does not belong to the intervals $(\hat{x}(\alpha_k), \check{x}(\alpha_k))$. In contrast, when x belongs to such intervals the equilibrium carries flow on more than one strategy and the congestion model is not mathematically correct. Nevertheless, in the latter situation, our simulation experiences show that the deviation of the line flows predicted by the model with respect to the "*true*" equilibrium flows is relatively small. In any case, more investigation is required to obtain precise estimates of this error or, even better, to analyze an equilibrium model with the effective frequency functions depending on y and not just on v. Notice that several of the preliminary results in §1 as well as some of the structural properties of the equilibrium do not depend on the particular functional form of $f_i(\cdot)$. For instance, the fact that more than one strategy may have to be used at equilibrium as well as the existence of ranges for the flow x where the equilibrium time will remain constant, are some of the features that one must expect from a more general model where $f_i(\cdot)$ is a function of y.

Let us also point out that the previous inaccuracy of the model may be negligible compared to other sources of error. Indeed, the analysis of congestion in this section concerns exclusively the case of independent Poisson queues. This assumption may be reasonable when the processes are independent, since the merging of many independent processes with small rates converges to a Poisson process. However, when the independence assumption does not hold as in the case of scheduled services, the analysis becomes much harder and the model adopted in the paper may be far from the real situation. Hence, the specific form of the effective frequency functions to be considered in the model may have to be adjusted and empirically calibrated. Notice however that the model and results derived in the paper support very general forms for the effective frequency functions $f_a(\cdot)$ providing a relative freedom in the modeling process.

Finally let us mention that, in the case of general networks, the waiting time at each stop node depends on the on-board loads of the incoming buses and, indirectly, on the queuing processes occurring throughout the network. In our model, these dependencies can be handled through the functional dependence of the effective frequency functions $f_a(\cdot)$ on the complete destination-arc-flow vector v, and not only on the corresponding component v_a. For instance, in the queuing model above these effects may be incorporated by assuming that the capacity distribution of the lines (q_j^i) depend on v. However, a formal treatment of these issues requires the study of complex stochastic networks which seem out of reach for the present state-of-the-art network queuing models (see Serfozo 1999, Tijms 1994).

References

Berge, C. 1997. *Topological Spaces*. Dover Publications Inc., Mineola, NY.

Bouzaïene-Ayari, B., M. Gendreau, S. Nguyen. 1995a. On the modeling of bus stops in transit networks, Part I: a survey of literature and new formulations. Technical Report CRT-95-55, Centre de Recherche sur les Transports, Univ. de Montréal.

——, ——, ——. (1995b). On the modeling of bus stops in transit networks, Part II: a numerical comparison of different models. Technical Report CRT-95-56, Centre de Recherche sur les Transports, Univ. de Montréal.

——, ——, ——. (1995c). An equilibrium-fixed point model for passenger assignment in congested transit networks. Technical Report CRT-95-57, Centre de Recherche sur les Transports, Univ. de Montréal.

Chriqui, C., P. Robillard. 1975. Common bus lines. *Transportation Sci.* **9** 115–121.

Cormen, T., Ch. Leiserson, R. Rivest. 1991. *Introduction to Algorithms.* McGraw-Hill, MIT Press, Boston, MA.

Correa, J. 1999. Asignación de flujos de pasajeros en redes de transporte público congestionadas, M.A. thesis, Univ. de Chile, Santiago, Chile.

De Cea, J., E. Fernández. 1993. Transit assignment for congested public transport systems: An equilibrium model. *Transportation Sci.* **27**(2) 133–147.

Dial, R. B. 1967. Transit pathfinder algorithms. *Highway Res. Record* **205** 67–85.

Fearnside, K., D. P. Draper. 1971. Public transport assignment—A new approach. *Traffic Engrg. Control* 298–299.

Gendreau, M. 1984. Etude approfondie d'un modèle d'équilibre pour 1'affectation de passagers dans les réseaux de transports en commun. Publication 384, Centre de Recherche sur les Transports, Univ. de Montréal.

Le Clercq, F. 1972. A public transport assignment model. *Traffic Engrg. Control* 91–96.

Nguyen, S., S. Pallotino. 1988. Equilibrium traffic assignment in large scale transit networks. *Eur. J. Oper. Res.* **37**(2) 176–186.

Serfozo, R. 1999. *Stochastic Networks.* Springer-Verlag, Berlin, Germany.

Spiess, H. 1984. Contribution à la théorie et aux outils de planification des réseaux de transport urbains. Publication 382, Centre de Recherche sur les Transports, Univ. de Montréal.

——, M. Florian. 1989. Optimal strategies: A new assignment model for transit networks. *Transportation Res.* **23B**(2) 83–102.

Tijms, H. 1994. *Stochastic Models: An Algorithmic Approach.* John Wiley & Sons, New York.

Wu, J., M. Florian, P. Marcotte. 1994. Transit equilibrium assignment: A model and solution algorithms. *Transportation Sci.* **28**(3) 193–203.

Received: January 1999; revisions received: September 2000; accepted: January 2001.