Diverge Nodes and Dynamic Traffic Equilibria

Stephen D. Boyles Christopher Melson Tarun Rambha Jennifer C. Duthie

March 28, 2013

Abstract

In most dynamic network loading models, oversaturation at a diverge node results in a queue forming on the upstream link, restricting flow to all downstream approaches. When combined with simplified flow models where travel speed is constant for all densities below the critical density, the resulting dynamic equilibrium problem may have infinitely many user equilibrium solutions, despite a unique system optimum solution. We demonstrate this with a simple diverge-merge network, which we also use to show that the price of anarchy in such systems may be unbounded. We feel that this issue is an artifact of modeling assumptions, rather than a description of a phenomenon in the field, and discuss piecewise-linear fundamental diagrams as one possible resolution.

Keywords: Diverge models, dynamic user equilibrium, multiple equilibria, price of anarchy

1 Introduction

This paper discusses the implications of two common assumptions in dynamic network loading models: (1) queues at diverge nodes obstruct traffic to all downstream links; and (2) vehicles travel at free-flow speed whenever vehicle density is subcritical. These assumptions underlie many of the large-scale dynamic traffic assignment (DTA) models used in practice today, greatly simplifying the calculations involved. However, we present a small example in which, under these two assumptions, literally every feasible route assignment is a user equilibrium, although these solutions correspond to vastly different flow conditions and the system optimum solution is unique. In passing, we show that the price of anarchy in such systems can be arbitrarily high.

The setting for our examples is a basic diverge-merge network (Figure 1), similar to the ones used by Daganzo (1998) and Nie (2010) to discuss other complications in dynamic network loading. Daganzo uses this network to show that these models can exhibit chaotic behavior due to queue spillback, and in particular that increasing the capacity on a "bottleneck" link may worsen conditions due to self-optimizing drivers choosing shortest routes. Nie demonstrates that the user equilibrium solutions in such a network are not unique, and proposes stability and efficiency criteria for distinguishing among these equilibria. In fact, the equilibria we study in this paper were briefly noted by (Nie, 2010, Equilibria IV in Section 3), but the focus of his paper was on the other, "efficient" equilibria.

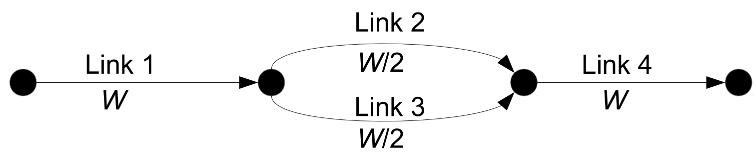

Figure 1: Network studied in this paper, with link capacities shown.

In this paper, we perform a deeper analysis of the "inefficient" equilibria. In contrast to both Daganzo (1998) and Nie (2010), which focused on the effects of congestion arising at a merge node, we consider congestion arising at the *diverge* node: in our examples, the only network congestion will occur upstream of the diverge node.

As discussed in more detail below, our concern is not that this phenomenon occurs in practice, but that it may occur during solution of DTA models, giving the illusion of much greater congestion than would exist in reality. Furthermore, this phenomenon may be difficult to detect in large networks. Modifying the fundamental diagram to include more than one linear piece for subcritical densities can reduce the impact of this phenomenon without significantly complicating the network loading. Our intent in this paper is primarily to call attention to this possibility, and to spur further research into its prevalence and potential resolutions.

The remainder of the paper is organized as follows: Section 2 describes the network loading model and assignment logic, presents the basic example, and remarks on the price of anarchy. Section 3 discusses the practical implications of our example, and how it may manifest in larger networks. Piecewise-linear fundamental diagrams are discussed in Section 4, and Section 5 briefly concludes and points to future research.

2 Primary Results

2.1 Network model

The phenomenon raised in this paper arises under rather general assumptions, to be discussed in more detail in Section 3. However, for concreteness, we describe our example based on the LWR hydrodynamic theory (Lighthill and Whitham, 1955; Richards, 1956), which postulates (1) that the rate of flow q at any point and time is a function of density k alone — that is, $q(x,t) = Q_{x,t}(k(x,t))$ at any location x and time t; (2) that the conservation law $\partial q/\partial x + \partial k/\partial t = 0$ holds everywhere these derivatives exist; and (3) that the space-mean speed u is related to q and k by the *fundamental equation* $q = uk$. The function $Q_{x,t}$ is often called the *fundamental diagram.* Given appropriate boundary conditions, this network loading model can be formulated as the solution to a hyperbolic partial differential equation. The solution of this model is greatly simplified by assuming a piecewise-linear fundamental diagram; under this assumption, the solution is readily obtained by the method of characteristics (Newell, 1993a,b,c), a finite difference scheme (Daganzo, 1994, 1995), or through the solution of a shortest-path problem (Daganzo, 2005a,b).

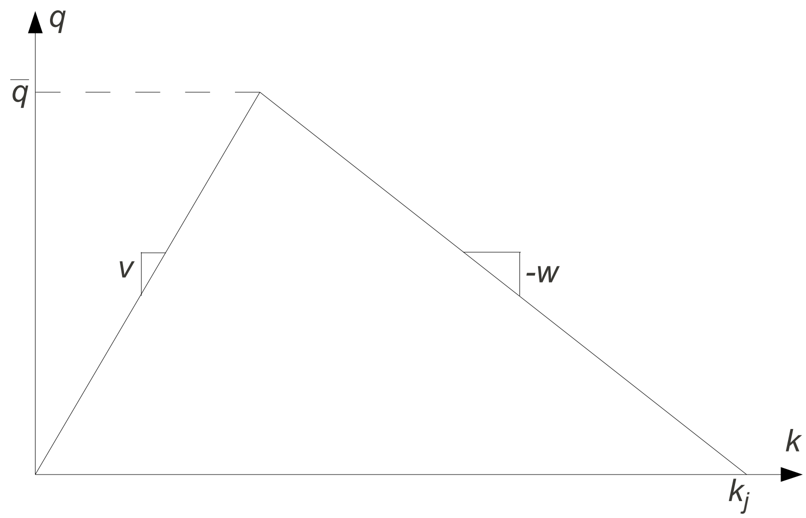

Figure 2: Triangular fundamental diagram specified by any three of the four parameters shown.

Omitting the location and time indices for brevity, the fundamental diagram $Q(k)$ is continuous, concave, and has two zeros: one at $k = 0$ and the other at the *jam density* $k = k_j$. We define the *capacity* $\overline{q} = \max_k Q(k)$, and the *critical density* $k_c = \min\{k : Q(k) = \overline{q}\}$. A density value k is *subcritical* if $k < k_c$. In the case of a triangular fundamental diagram (Figure 2), $Q(k)$ can be completely specified by any three of the following four parameters: the free-flow speed v, the backward wave speed $-w$, the capacity $\overline{q}$, and the jam density k_j. It is convenient to write $Q(k) = \min\{S(k), R(k)\}$ in terms of the *sending flow* $S(k) = \min\{kv, \overline{q}\}$ and *receiving flow* $R(k) = \min\{w(k_j - k), \overline{q}\}$ in this case.[1] Other piecewise linear fundamental diagrams have been proposed in the literature; for instance, the trapezoidal diagram in Daganzo (1994) is commonly used, with a horizontal piece corresponding to the capacity $\overline{q}$. For brevity we initially assume a triangular fundamental diagram of two pieces; however, in all examples in this paper density remains subcritical, so need only concern ourselves with the increasing portion of the fundamental diagram, and the remarks in this paper would apply equally to trapezoidal fundamental diagrams or any other with a single increasing linear piece.

In the network in Figure 1 there are four links with unit length; relative to each link, $x = 0$ denotes the upstream end and $x = 1$ the downstream end. We assume that the fundamental diagram is homogeneous on each link and constant with time, using $\overline{q}_i$ to refer to the capacity at each point on link i, k_j^i the jam density on link i, and so forth. At the diverge, we are interested in the inflow rates r_2 and r_3 from the downstream end of link 1 to the upstream ends of links 2 and 3, respectively. Defining the sending and receiving flows at location x on link i at time t as $S_i(x, t)$ and $R_i(x, t)$, r_2 and r_3 depend on $S_1(1, t)$, $R_2(0, t)$, and $R_3(0, t)$. Using $p_{1i}(t)$ to denote the fraction of the sending flow wishing to travel to link i ($i \in \{2, 3\}$), we have $r_i(t) = \phi(t)p_{1i}(t)S_1(1, t)$ where $\phi(t) = \min_{i\in\{2,3\}}\{1, R_i(0, t)/[p_{1i}(t)S_1(1, t)]\}$. Notice that if one of the downstream links i' is

[1] We describe the problem in continuous time; discrete equivalents of these formulas can be found in Daganzo (1995), Yperman (2007), and Nie et al. (2008).

oversaturated — that is, if $R_{i'}(0,t) < p_{1i'}(t)S_1(1,t)$ — the flow to *both* links 2 and 3 is restricted.

A variety of merge models exist (Daganzo, 1995; Lebacque, 1996; Jin and Zhang, 2003; Ukkusuri et al., 2012). However, in all of the examples in this paper, we assume $\overline{q}_4 \geq \overline{q}_2 + \overline{q}_3$, so there is no congestion at the merge node and we only need to specify the trivial case of no merge congestion: $q_{i4}(t) = S_i(1,t)$ for all t and $i \in \{2,3\}$.

We assume a total of D vehicles enter the upstream end of link 1 at a uniform rate W. (That is, the inflow rate is W for $t \in [0, D/W]$ and zero afterward.) Route choice is often expressed in terms of the rate at which vehicles depart on each path in the network. However, in our network, it is simpler to express route choice solely through the splitting proportions at the diverge. Since $p_{12}(t)$ and $p_{13}(t)$ are nonnegative and sum to one, we can use the single parameter $p(t) \in [0,1]$ to describe route choice at each point in time, with $p_{12}(t) = p(t)$ and $p_{13}(t) = 1 - p(t)$. Expressing route choice in this way simplifies the calculations and is equivalent to specifying path departure rates under mild regularity assumptions; for instance, it is sufficient to assume path departure rates and $p(t)$ are continuous almost everywhere, an assumption we adopt. We are interested in the functions $p(t)$ which create dynamic user equilibria according to Wardrop 's principle (1952): that no traveler may reduce his or her travel time by unilaterally switching routes. In a continuous-flow setting, this implies that for any departure time, the travel times on all used paths are equal and minimal.

In our example, we let $\overline{q}_2 = \overline{q}_3 = W/2$, $\overline{q}_1 = \overline{q}_4 = W$, $v^i = 3$ on all links (so the free-flow time through the network is 1), and $k^i_j = D$ on all links (so there will be no queue spillback). With this selection of parameters, the only place congestion could possibly occur is upstream of the diverge node. Since the route choice $p_{12}(t)$ is only made at the diverge node itself (the downstream end of any queue on link 1), it should be clear that any possible function $p(t)$ results in a user equilibrium, because all links downstream of the diverge node will be at free-flow. Different $p(t)$ functions will create different amounts of delay upstream of the queue, but this delay never affects the vehicle at the front of the queue, which is the only place where route choice can be exercised. Since the user equilibrium assumption is only concerned with a driver's own delay, and never that of other drivers in the system, it follows that any choice of $p(t)$ satisfies the user equilibrium condition. The following subsections explore this issue quantitatively, respectively deriving the same result by direct calculation of delays, modifying the network to illustrate an unbounded price of anarchy, and conducting a broader sensitivity analysis with less contrived parameters.

2.2 Delay calculations

The previous subsection argued that any choice of $p(t)$ produces a user equilibrium solution, due to the selfish nature of equilibrium routing and the network structure. This subsection verifies this logic by writing expressions for the travel time on each path given $p(t)$, which will also serve as the basis for additional analysis.

As stated above, the only possible congestion in the network occurs at the downstream end of link 1. The diverge model ensures that the inflow rates to links 2 and 3 never exceed their capacity, and since the capacity of link 4 is the sum of the capacities of links 2 and 3, no congestion will occur at the merge. Until all D vehicles exit link 1, its outflow rate will be $W\phi(t)$. Therefore, the cumulative number of vehicles which have exited link 1 by time t is given by

$$N_1^{\downarrow}(t) = \int_0^t S_1(t')\phi(t')\ dt' \tag{1}$$

while the cumulative number of vehicles that have entered link 1 by time t is $N_1^{\uparrow}(t) = Wt$. Let T be the time at which the last vehicle exits link 1, that is, $N_1^{\downarrow}(T) = D$. Substituting the parameters of the network model into the diverge equation, we have

$$\phi(t) = \min\{1/[2p(t)], 1/[2(1-p(t))]\} \tag{2}$$

for $t \in (1/3, T)$ since there is no spillback from links 2 and 3, implying $R_2(0,t) = R_3(0,t) = W/2$.

Furthermore, $S_1(t) > 0$ and $\phi(t) > 0$ for $t \in (1/3, T)$ imply that $N_1^{\uparrow}(t)$ and $N_1^{\downarrow}(t)$ are strictly increasing and thus have well-defined inverse functions $t_1^{\uparrow}(n)$ and $t_1^{\downarrow}(n)$, respectively denoting the time at which the n-th vehicle enters and exits link 1. Thus, for $n \in (0, D)$, the travel time of the n-th vehicle on link 1 is given by

$$\tau_1(n) = t_1^{\downarrow}(n) - t_1^{\uparrow}(n) \tag{3}$$

Since both paths in the network consist of three links, two of which are never congested and have free-flow time 1/3, the total travel time of the n-th vehicle is

$$\tau(n) = t_1^{\downarrow}(n) - t_1^{\uparrow}(n) + 2/3 \tag{4}$$

In particular, $\tau(n)$ *does not depend on the path chosen* — for all vehicles, the travel time is the same on either path available to them, so *all* functions $p(t)$ produce solutions satisfying the Wardrop user equilibrium conditions. However, as seen in the next section, these equilibria are *not* equivalent in terms of experienced travel time.

2.3 Price of anarchy

It is common to distinguish among multiple equilibria in terms of the total travel time of all vehicles in the system; in particular, following Koutsoupias and Papadimitriou (1999), one can define the "price of anarchy" ρ of a system to be the ratio between the *worst* user equilibrium (in terms of total travel time) and the system-optimal solution which minimizes total travel time (which need not be a user equilibrium). In this subsection, we show that by varying the parameters in our example, the price of anarchy can be arbitrarily high. This is in contrast to many other network equilibrium problems, in which the price of anarchy can be bounded under general assumptions: Roughgarden (2002) shows that $\rho \leq 4/3$ for a static equilibrium problem with affine cost functions, and that $\rho = \Theta(p/\log p)$ for polynomial cost functions of degree p. Explicit bounds have been found for other equilibrium variants, including static equilibrium with congestion pricing (Han and Yang, 2008), static stochastic user equilibrium with logit disturbance terms (Guo et al., 2010; Huang et al., 2011), and dynamic single-bottleneck models (Doan and Ukkusuri, 2012). Anshelevich and Ukkusuri (2009) showed that the price of anarchy can be arbitrarily large in dynamic networks;

however, in their example, the price of anarchy only grew with the network size. Below, we show that the price of anarchy can be made large even with a four-link network.

Define the total cost of travel to be

$$\Sigma = \int_0^D \tau(n) dn \tag{5}$$

Clearly $\tau(n) \geq 1$ for all n (since that is the free-flow travel time). This lower bound can in fact be attained for all vehicles: $t_1^{\uparrow}(n)$ is simply n/W, and if $p(t) = 1/2$ for all t, $\phi(t) = 1$, $S_1(t) = W$ for $t \in [1/3, T]$, so

$$N_1^{\downarrow}(t) = \int_0^t S_1(t')\phi(t')\ dt' = \begin{cases} 0 & \text{if } t < 1/3 \\ W(t - 1/3) & \text{if } t \in [1/3, T] \\ D & \text{if } t > T \end{cases} \tag{6}$$

$N_1^{\downarrow}(t)$ must be continuous, so we have $T = 1/3 + D/W$, and can calculate the inverse function $t_1^{\downarrow}(n) = 1/3 + n/W$ for $n \in (0, D)$; hence $\tau_1(n) = (1/3 + n/W) - n/W = 1/3$ and $\tau(n) = 1$, so $\Sigma = D$ in this system-optimal solution.

To find an upper bound on Σ, we need to choose $p(t)$ to maximize $\tau_1^{\downarrow}(n)$ for all n. While in general such a maximization problem would require techniques from the calculus of variations, this example can be solved through a simpler argument: maximizing $\tau_1^{\downarrow}(n)$ is equivalent to maximizing $t_1^{\downarrow}(n)$ for all n (since $t_1^{\uparrow}(n)$ is given), which is equivalent to minimizing $N_1^{\downarrow}(t)$ for all t, which is equivalent to minimizing $\phi(t)$ for all t, which occurs if $p(t) = 1$ or 0 for all $t \in (0, T)$. Performing the calculations in this case, we have

$$T = 1/3 + 2D/W \tag{7}$$

$$N_1^{\downarrow}(t) = W(t - 1/3)/2 \text{ for } t \in [1/3, T] \tag{8}$$

$$t_1^{\downarrow}(n) = 1/3 + 2n/W \tag{9}$$

so finally

$$\tau_1(n) = 1/3 + n/W \tag{10}$$

$$\tau(n) = 1 + n/W \tag{11}$$

and

$$\Sigma = D + D^2/2W \tag{12}$$

Since all solutions are user equilibria, we may define the price of anarchy ρ to be the ratio between these upper and lower bounds on Σ found above:

$$\rho = \frac{D + D^2/2W}{D} = 1 + \frac{D}{2W} \tag{13}$$

which can clearly be made arbitrary large by increasing the ratio D/W.

2.4 Sensitivity analysis

This section discusses how the above analysis would change under different choices of parameters. The most important parameters in our example are $\overline{q}_2$, $\overline{q}_3$, and the inflow rate W to the network. The choice of $\overline{q}_4$ and $\overline{q}_1$ are less significant as long as $\overline{q}_4 \geq \overline{q}_2+\overline{q}_3$ and $\overline{q}_1 \geq W$, assumptions we retain as essential features of our demonstration. In particular, we allow $\overline{q}_2 \neq \overline{q}_3$, but enforce $\overline{q}_2 \leq \overline{q}_3$ by symmetry. Allowing W, $\overline{q}_2$, and $\overline{q}_3$ to take generic positive values within these assumptions, we divide the analysis into three major cases:

Case I: $W < \overline{q}_2$. In this case, no congestion can arise at the diverge node, since $\phi(t) = 1$ regardless of $p(t)$, and all solutions have $\Sigma = D$. Therefore all solutions are both user equilibria and system optima, and $\rho = 1$.

Case II: $W > \overline{q}_2$ but $W \leq \overline{q}_2 + \overline{q}_3$. In this case, it is possible to avoid congestion by suitably dividing the outflow from link 1 between links 2 and 3, so in the system-optimum solution all vehicles experience unit delay and $\Sigma = D$. However, congestion can be caused if p is too large: if $p(t) > \overline{q}_2/W$, $\phi(t) < 1$, but such solutions remain user equilibria because the only source of delay is upstream of the diverge node. Repeating the analysis from the previous subsection, we find that the worst-case delay is obtained with $p(t) = 1$ for all t, for which

$$\Sigma = D + \frac{D^2}{2}\left(\frac{1}{\overline{q}_2} - \frac{1}{W}\right) \tag{14}$$

and

$$\rho = 1 + \frac{D}{2}\left(\frac{1}{\overline{q}_2} - \frac{1}{W}\right) \tag{15}$$

and again the price of anarchy can be made arbitrarily large by increasing D. (The previous subsection's analysis is a special case.)

Case III: $W > \overline{q}_2 + \overline{q}_3$. In this case, some congestion is inevitable, since links 2 and 3 do not have enough capacity to handle the outflow from link 1. Delay is minimized by maximizing $\phi(t)$, which occurs if $\overline{q}_2/p(t) = \overline{q}_3/(1-p(t))$ or $p(t) = \overline{q}_2/(\overline{q}_2+\overline{q}_3)$. In this case, the resulting total delay is

$$\Sigma = D + \frac{D^2}{2}\left(\frac{1}{\overline{q}_2+\overline{q}_3} - \frac{1}{W}\right) \tag{16}$$

as the reader may verify. As with Case II, delay is maximized with $p(t) = 1$ for all t, resulting in

$$\Sigma = D + \frac{D^2}{2}\left(\frac{1}{\overline{q}_2} - \frac{1}{W}\right) \tag{17}$$

As before, all solutions are user equilibria, but the system optimum is unique ($p(t) = \overline{q}_2/(\overline{q}_2 + \overline{q}_3)$). Therefore, the price of anarchy is

$$\rho = \frac{1 + D(1/\overline{q}_2 - 1/W)/2}{1 + D(1/(\overline{q}_2+\overline{q}_3) - 1/W)/2} \tag{18}$$

which is bounded for any fixed values of $\overline{q}_2$ and $\overline{q}_3$ even as $D \to \infty$, but which increases without bound if $\overline{q}_2 + \overline{q}_3 \to W$ at the same time as $D \to \infty$, again yielding the results of the previous subsection in the limit.

To summarize, regardless of the capacity values chosen, all feasible solutions are user equilibria if there is no downstream congestion at the merge; the implications in terms of the delay depend on the relative values of the upstream link outflow rate and the downstream links' capacities. If the inflow rate is sufficiently small, no congestion will occur and all feasible solutions are system optima; if the inflow rate is sufficiently large, a certain level of congestion is inevitable, but the system optimal solution is unique and the price of anarchy is strictly greater than one. For inflow rates between these extremes, some user equilibria produce congestion while others do not, and the price of anarchy can thus be made arbitrarily high by manipulating the total demand.

3 Discussion

The previous section demonstrated that simple dynamic diverge-merge networks exist where every feasible path assignment is a user equilibrium, despite major differences in total travel time as path assignments vary. In this section, we discuss the implications of these findings, addressing three questions in turn: whether this phenomenon is likely to exist in reality, or simply a modeling artifact; the implications on DTA modeling and algorithm termination; and whether this phenomenon would be localized, or if impacts could be felt throughout a larger network.

Regarding the first question, we believe this phenomenon is unlikely to occur in the field for several reasons. First, from entropy considerations, if both paths have equal travel times, it is improbable that all travelers would only select one path. Further, the previous results rely heavily on the modeling assumption that travel times will continue to be equal on both paths despite unequal usage, which follows from the assumption of the triangular fundamental diagram. In practice, due to heterogeneity in driver behavior, average speed will drop even at lower densities if overtaking is prohibited, and a travel time-minimizing driver would probably choose the link with lesser density to maximize his or her probability of driving at the preferred speed. If this is the case, the unique equilibrium occurs with $p(t) = 1/2$ for all t. In Section 4, we return to the fundamental diagram as one possible resolution of the issue.

However, even if this phenomenon never occurs in the field, it poses great challenges from a modeling or algorithmic perspective. Given the example in the previous section, any DTA software using a gap criterion to terminate will stop after the first iteration, since the user equilibrium principle is satisfied exactly. If this first iteration is an all-or-nothing assignment, the reported level of congestion will be the highest possible value, even though the more likely field conditions would be the lowest possible value. Even continuing a DTA algorithm for another iteration after such "convergence" is obtained may not detect the second, unused, path: if ties are broken deterministically in the time-dependent shortest path computation, many simple path generation routines will return the first path again since the travel times remain at free-flow on both downstream segments. While it may be possible to check for the presence of this phenomenon in small networks, detecting this phenomenon in larger networks can be more problematic. Furthermore, this issue is also not confined to the LWR model and its variations; any flow model in which free-flow speeds are maintained even as density rises is subject to the same analysis, such as point queue models.

Still worse, this phenomenon can propagate through larger networks. In the example above, link

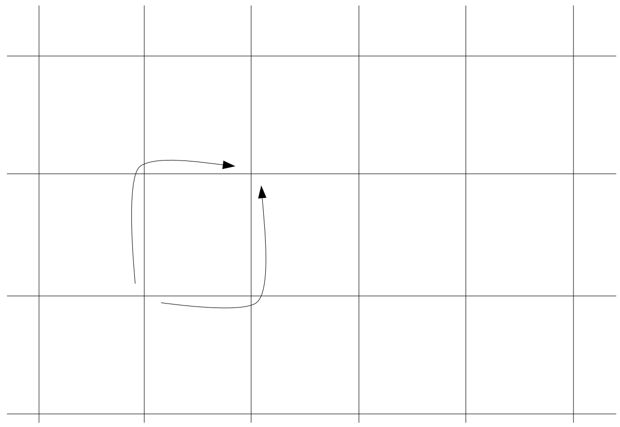

Figure 3: Grid networks have roughly similar topology to the diverge-merge network.

jam densities were chosen to avoid any spillback; however, if this phenomenon were to persist over time, spillback on the upstream link would likely occur, reducing capacity on additional upstream links. In a grid network (Figure 3), each block presents an alternative set of choices roughly similar to a diverge and merge. While not directly identical, due to the presence of intermediate nodes and the differences between diverge models and general intersections, congestion upstream of intersections combined with free-flow conditions downstream will still produce the same result. When modeling networks with high levels of congestion, even if this phenomenon occurs only at a small handful of blocks in a central business district, the resulting spillback effects could result in a huge increase in congestion.

One potential way to mitigate this modeling effect is to initially generate a working set of multiple paths per origin-destination pair, perhaps using a K-shortest path algorithm. This approach is often employed to avoid wasting early DTA iterations on highly congested solutions based on all-or-nothing assignments, and one may hope that this strategy may also avoid the worst of the equilibria discussed here. While studying this strategy is beyond the scope of this paper, we speculate that its impact may be more limited than one would hope. In grid networks, the number of equal-length paths between two points is quite large, and missing even one such path may trigger spillback and obstruction of other links and origin-destination pairs.

The root causes of the phenomenon described in this paper are queuing located upstream of diverges, and free-flow conditions downstream (even with unequal link utilizations). This suggests two approaches to resolving the issue and the difficulties it imposes on DTA models in determining convergence. One alternative is to modify the diverge model, perhaps providing separate queues for travelers heading for different links, as with independent turn lanes. However, this only pushes the problem further upstream: it does not make sense for the capacity for each of these separate queues to exceed the capacity for the downstream links, and effectively the diverge point has been moved to the point at which the turn lanes split, rather than the point at which the links themselves split.

However, the second cause admits a more promising solution: adjusting the fundamental diagram so that speeds drop below free-flow even at subcritical densities. This technique is discussed in detail in the next section, where we show that equilibria can be made unique (except for an arbitrarily small interval of low demand) with an arbitrarily small perturbation of travel times. While the Highway Capacity Manual (HCM) (Transportation Research Board, 2010, Exhibit 11-6) suggests travel times begin to decrease when density exceeds 1000-2000 passenger cars per lane, we actually propose introducing a perturbation at a much smaller density value. We believe this distinction is acceptable for the following reason. The HCM and DTA serve different purposes — if average speed is nearly constant with density over a particular range, for the purposes of operational analysis there can be considerable advantage in approximating the speed as constant, simplifying the model and avoiding the impression of false precision. However, as we have shown in this paper, in DTA this assumption can produce an infinite number of equilibria and an unbounded price of anarchy. Perturbing the fundamental diagram can address this issue while introducing a change in travel times which can be made as small as desired, as we show in the next section. We feel the benefits of reducing the size of the equilibrium set outweigh the minor loss in accuracy in travel time calculation.

4 Piecewise-linear fundamental diagrams

In light of the computational advantages piecewise linear fundamental diagrams hold, we propose introducing a second piece to the uncongested portion of the fundamental diagram, as in Figure 4. Adding an additional piece to the fundamental diagram poses no major difficulties to the primary solution methods for the LWR equations, merely adding one additional term to the minimization in the sending flow equation in the cell-transmission or link-transmission models, or one additional wave speed to check in Newell's method or Daganzo's variational method.

We demonstrate this by modifying the previous example in this way, beginning with an illustrative example before addressing more fully the question of how the additional piece should be added. Figure 4 shows a piecewise linear fundamental diagram for links 2 and 3, with an added line segment in the uncongested region of slope 2, with the original triangular fundamental diagram superimposed. Note that the critical density and backward wave speed have changed in order to keep the capacity and jam density unchanged; specifically, the critical density increased from $W/6$ to $W/5$. For this diagram, link speeds are decreasing in k whenever $k > W/10$, with a unique speed corresponding to each such density value. When the steady inflow rate to link i is r_i with this fundamental diagram, a routine application of Newell's method yields the relation

$$N_i^{\downarrow}(t) = \min\{r_i t - r_i/2 + W/20, r_i t - r_i/3\} \tag{19}$$

where $i \in \{2, 3\}$ or

$$N_i^{\downarrow}(t) = \begin{cases} r_i t - r_i/2 + W/20 & \text{if } r_i \geq 3W/10 \\ r_i t - r_i/3 & \text{if } r_i < 3W/10 \end{cases} \tag{20}$$

while $N_i^{\uparrow}(t) = r_i t$; thus if $r_i \geq 3W/10$ we calculate $t_i^{\uparrow}(n) = n/r_i$, $t_i^{\downarrow}(n) = n/r_i + 1/2 - W/20r_i$ and

$$\tau_i(n) = 1/2 - W/20r_i \tag{21}$$

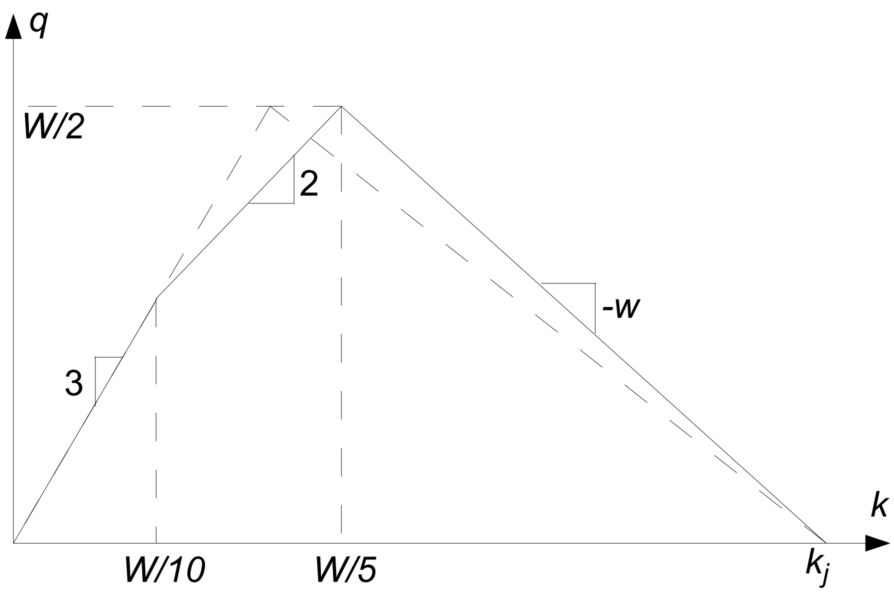

Figure 4: Piecewise linear fundamental diagram for links 2 and 3; capacity and jam density unchanged.

for all vehicles n since there is no bottleneck on the link; and if $r_i < 3W/10$ we simply have $\tau_i(n) = 1/3$.

Assuming p constant, the inflow rates r_2 and r_3 to links 2 and 3 are related to the diverge by $r_2 = \min\{pW, W/2\}$ and $r_3 = \min\{(1-p)W, W/2\}$. Taking $p \in [1/2, 1]$ by symmetry, we have $r_2 = W/2$ and therefore $\tau_2(n) = 2/5$. Substituting the relation for r_3 into (20) we have

$$\tau_3(n) = \begin{cases} \frac{1}{2}\left(1 - \frac{1}{10(1-p)}\right) & \text{if } p \leq 7/10 \\ 1/3 & \text{if } 1/2 \leq p < 7/10 \end{cases} \tag{22}$$

Thus, the only equilibrium with constant splitting proportion occurs when $p = 1/2$ and $\tau_2 = \tau_3 = 2/5$.

While this example used a specific piecewise-linear fundamental diagram for concreteness, any such diagram with more than one increasing piece and the same capacity would also produce a unique equilibrium. The more general case is discussed next.

It is natural to ask how a piecewise-linear fundamental diagram should be chosen; one compelling feature of the triangular diagram is its economy of parameters, being completely specified by free-flow speed, capacity, and jam density. One approach is to calibrate directly to traffic observations, but acquiring enough data to accurately calibrate these diagrams for each link in a large roadway network may be prohibitive. Perhaps a simpler idea is to perturb the triangular diagram slightly, so the piecewise-linear diagram retains its general character, but ensures unique speeds (even if very close to free-flow) over a larger range of densities, as in Figure 5.

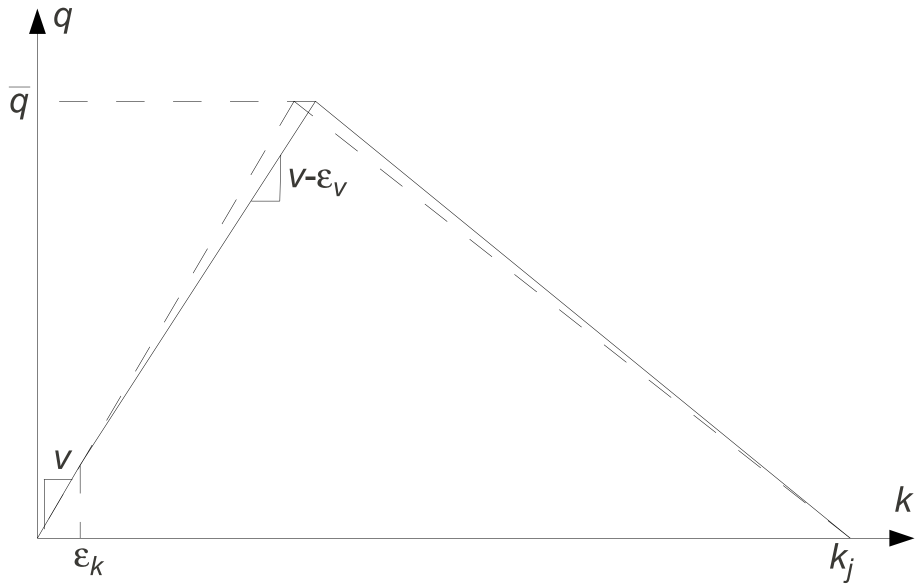

Figure 5: Slightly perturbed fundamental diagram approximating triangular shape.

While several perturbation schemes can be imagined, it seems natural to require the capacity and jam density to remain unchanged while allowing the critical density and backward wave speed to adjust to fit the additional piece, since the former quantities are more often measured in the field or calculated using procedures such as that in the HCM. Such a scheme can be implemented as follows, using several small positive constants (denoted ϵ with appropriate subscripts):

1. For $k \in [0, \epsilon_k]$, $q(k) = vk$, that is, for a small initial portion the fundamental diagram is unchanged.
2. For $k \in [\epsilon_k, k_c + \epsilon_c]$, $q(k) = v\epsilon_k + (v - \epsilon_v)(k - \epsilon_k)$ where ϵ_c is chosen so that $q(k_c + \epsilon_c) = \overline{q}$, that is, on the second piece the flow rises at rate slightly less than free-flow, and continues until capacity is reached.
3. For $k \in [k_c + \epsilon_c, k_j]$, $q(k) = \overline{q}(1 - (k_j - k)/(k_j - k_c - \epsilon_c))$, that is, on the final piece the flow decreases linearly to zero.

Geometric calculations show that the new critical density $k_c + \epsilon_c$ is $(\overline{q} - \epsilon_k \epsilon_v)/(v - \epsilon_v)$, as compared to the triangular critical density $k_c = \overline{q}/v$, and that the new backward wave speed is $\overline{q}(v - \epsilon_v)/(k_j v - \overline{q} + \epsilon_v k_j + \epsilon_k \epsilon_v)$, as compared to the triangular backward wave speed $\overline{q}v/(k_j v - \overline{q})$, and both of these differences shrink to zero as ϵ_v and ϵ_k grow small.

With these fundamental diagrams, the range of demand values producing infinitely many equilibria is greatly decreased. Solving the LWR model to obtain travel times as a function of link inflow rates yields

$$\tau = \begin{cases} 1/v & \text{if} r \leq v\epsilon_k \\ \dfrac{r - \epsilon_k \epsilon_v}{r(v - \epsilon_v)} & \text{otherwise} \end{cases} \tag{23}$$

Assuming $p \in (0, 1)$ (that is, both links 2 and 3 are used), then $\tau_2 = \tau_3$. If $\tau_2 = \tau_3 = 1/v$, then both links are at free flow, and we must have $pD = r_2 \leq v\epsilon_k$ and $(1-p)D = r_3 \leq v\epsilon_k$. Equivalently, we must have $p \leq v\epsilon_k/D$ and $p \geq 1 - v\epsilon_k/D$, and any p satisfying both inequalities is a user equilibrium. It is impossible to satisfy both inequalities unless $v\epsilon_k/D \geq 1/2$, or $D \leq 2v\epsilon_k$, so unless the demand is quite small, this case will not apply.

If $\tau_2 = \tau_3 > 1/v$, then

$$\frac{r_2 - \epsilon_k \epsilon_v}{r_2(v - \epsilon_v)} = \frac{r_3 - \epsilon_k \epsilon_v}{r_3(v - \epsilon_v)} \tag{24}$$

or $r_2 = r_3$ since both the left-hand side and right-hand side are strictly decreasing in their arguments. Therefore the unique equilibrium is $p = 1/2$.

Finally, the cases $p = 0$ and $p = 1$ can be discarded unless $D \leq v\epsilon_k$, in which case any solution is a user equilibrium. Otherwise, $p = 0$ implies $\tau_2 > \tau_3 = 1$ and $p = 1$ implies $1/v = \tau_2 < \tau_3$, neither of which satisfies the Wardrop condition.

To summarize: if $W \leq v\epsilon_k$, all solutions are user equilibria; if $v\epsilon_k \leq W \leq 2v\epsilon_k$, any $p \in [1 - v\epsilon_k/W, v\epsilon_k/W]$ is a user equilibrium; and if $W \geq 2v\epsilon_k$, the unique equilibrium is $p = 1/2$. (Figure 6) Furthermore, the maximum possible difference in travel time between the triangular fundamental diagram and perturbed fundamental diagram is

$$\frac{1}{v} - \frac{\bar{q} - \epsilon_k \epsilon_v}{\bar{q}(v - \epsilon_k)} = \frac{1}{v} - \frac{1}{v}\left(1 - \frac{\epsilon_k \epsilon_v}{\bar{q}}\right)\frac{1}{1 - \epsilon_k/v} \tag{25}$$

$$= \frac{1}{v} - \frac{1}{v}\left(1 - \frac{\epsilon_k \epsilon_v}{\bar{q}}\right)\left(1 + \frac{\epsilon_k}{v} + \frac{\epsilon_k^2}{v^2} + \cdots\right) \tag{26}$$

$$= \frac{1}{v} - \frac{1}{v}\left(1 + \frac{\epsilon_k}{v} + \frac{\epsilon_k^2}{v^2} + \cdots\right) + \frac{\epsilon_k \epsilon_v}{\bar{q}v}\left(1 + \frac{\epsilon_k}{v} + \frac{\epsilon_k^2}{v^2} + \cdots\right) \tag{27}$$

$$= \left(\frac{\epsilon_k \epsilon_v}{\bar{q}v} - \frac{\epsilon_k}{v^2}\right)\left(1 + \frac{\epsilon_k}{v} + \frac{\epsilon_k^2}{v^2} + \cdots\right) \tag{28}$$

using the formula for the sum of a geometric series. As $\epsilon_k \to 0$, the terms $\frac{\epsilon_k}{v}, \frac{\epsilon_k^2}{v^2}, \ldots$ shrink to zero, and the change in travel times is $O(\epsilon_k \epsilon_v)$. Thus, the change in travel times and the range of inflow rates W producing multiple equilibria can be made arbitrarily small by choosing small values of ϵ_k and ϵ_v.

5 Conclusion

This paper discussed how diverge and traffic flow models can interact to produce counterintuitive phenomena, as demonstrated by an example in which every feasible assignment is a user equilibrium even though the travel times vary widely, and in which the "price of anarchy" can be made

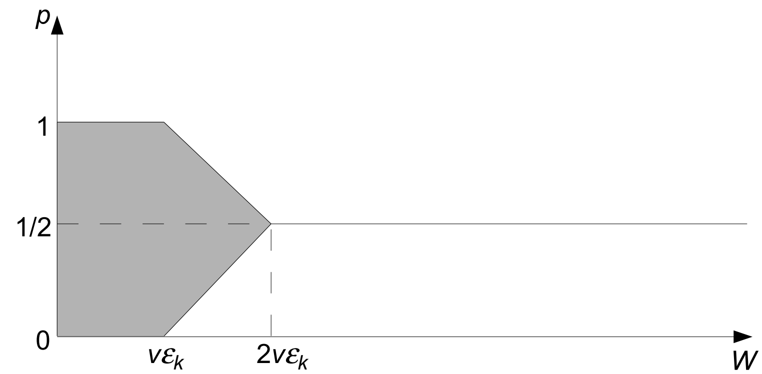

Figure 6: Range of user equilibria as a function of inflow rate W.

arbitrarily high either by increasing the demand or decreasing the link capacities. Our opinion is that this phenomenon is a modeling artifact, rather than something observable in the field. Nevertheless, it poses serious challenges for dynamic traffic assignment algorithms in identifying the "correct" equilibrium solution, and through spillback mechanisms, can propagate unrealistic congestion throughout a network. Simple approaches (such as those based on generating an initial set of paths) may fail to overcome this difficulty, particularly in grid networks.

However, modifying the fundamental diagram by including a second uncongested linear piece can remedy or limit the effects of this phenomenon without compromising the efficiency of LWR solution methods, even when the resulting fundamental diagrams are only slightly perturbed from a triangular one. Further research is needed to quantify the extent to which this phenomenon occurs in larger networks, to investigate other means of resolving or mitigating this modeling issue, and to more generally explore dynamic traffic assignment problems with multiple equilibria.

References

Anshelevich, E. and S. Ukkusuri (2009). Equilibria in dynamic selfish routing. In *Proceedings of the 2nd International Symposium on Algorithmic Game Theory*, Paphos, Greece, pp. 171–182.

Daganzo, C. F. (1994). The cell transmission model: a dynamic representation of highway traffic consistent with the hydrodynamic theory. *Transportation Research Part B 28*(4), 269–287.

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

Daganzo, C. F. (1998). Queue spillbacks in transportation networks with a route choice. *Transportation Science 32*(1), 3–11.

Daganzo, C. F. (2005a). A variational formulation of kinematic waves: basic theory and complex boundary conditions. *Transportation Research Part B 39*(2), 187–196.

Daganzo, C. F. (2005b). A variational formulation of kinematic waves: solution methods. *Transportation Research Part B 39*(10), 934–950.

Doan, K. and S. Ukkusuri (2012). Price of anarchy in dynamic traffic networks: a computational study. Presented at the Annual Meeting of the Institute for Operations Research and Management Sciences, Phoenix, AZ.

Guo, X., H. Yang, and T. L. Liu (2010). Bounding the inefficiency of logit-based stochastic user equilibrium. *European Journal of Operational Research 201*(2), 463–469.

Han, D. and H. Yang (2008). The multi-class, multi-criterion traffic equilibrium and the efficiency of congestion pricing. *Transportation Research Part E 44*(5), 753–773.

Huang, H. J., T. L. Liu, X. L. Guo, and H. Yang (2011). Inefficiency of logit-based stochastic user equilibrium in a traffic network under ATIS. *Networks and Spatial Economics 11*(2), 255–269.

Jin, W. L. and H. M. Zhang (2003). On the distribution schemes for determining flows through a merge. *Transportation Research Part B 37*(6), 521–540.

Koutsoupias, E. and C. H. Papadimitriou (1999). Worst-case equilibria. In *Proceedings of the 16th Annual Symposium on Theoretical Aspects of Computer Sciences*, pp. 404–413.

Lebacque, J. P. (1996). The Godunov scheme and what it means for first order traffic flow models. In *Proceedings of the 13th International Symposium on Transportation and Traffic Theory*, London, pp. 647–678.

Lighthill, M. J. and G. B. Whitham (1955). On kinematic waves II. A theory of traffic flow on long crowded roads. *Proceedings of the Royal Society A 229*(1178), 314–345.

Newell, G. F. (1993a). A simplified theory of kinematic waves in highway traffic, part I: General theory. *Transportation Research Part B 27*(4), 281–287.

Newell, G. F. (1993b). A simplified theory of kinematic waves in highway traffic, part II: Queueing at freeway bottlenecks. *Transportation Research Part B 27*(4), 289–303.

Newell, G. F. (1993c). A simplified theory of kinematic waves in highway traffic, part III: Multi-destination flows. *Transportation Research Part B 27*(4), 305–313.

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

Nie, Y. M., J. Ma, and H. M. Zhang (2008). A polymorphic dynamic network loading model. *Computer-Aided Civil and Infrastructure Engineering 23*(2), 86–103.

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

Roughgarden, T. A. (2002). *Selfish Routing.* Ph. D. thesis, Cornell University.

Transportation Research Board (2010). *Highway Capacity Manual.* Washington, DC: National Research Council.

Ukkusuri, S. V., L. Han, and K. Doan (2012). Dynamic user equilibrium with a path based cell transmission model for general traffic networks. *Transportation Research Part B 46*(10), 1657–1684.

Wardrop, J. (1952). Some theoretical aspects of road traffic research. *Proceedings of the Institute of Civil Engineers, Part II*, 325–378.

Yperman, I. (2007). *The Link Transmission Model for Dynamic Newtork Loading.* Ph. D. thesis, Katholieke Universiteit Leuven, Belgium.