TRANSPORTATION SCIENCE
Vol. 39, No. 4, November 2005, pp. 446–450
ISSN 0041-1655 | EISSN 1526-5447 | 05 | 3904 | 0446

DOI 10.1287/trsc.1050.0127

On a Paradox of Traffic Planning

Dietrich Braess
Faculty of Mathematics, Ruhr-University Bochum, 44780 Bochum, Germany, dietrich.braess@rub.de

Anna Nagurney, Tina Wakolbinger
Department of Finance and Operations Management, Isenberg School of Management, University of Massachusetts, Amherst, Massachusetts 01003 {nagurney@gbfin.umass.edu, wakolbinger@som.umass.edu}

For each point of a road network, let there be given the number of cars starting from it, and the destination of the cars. Under these conditions one wishes to estimate the distribution of traffic flow. Whether one street is preferable to another depends not only on the quality of the road, but also on the density of the flow. If every driver takes the path that looks most favorable to him, the resultant running times need not be minimal. Furthermore, it is indicated by an example that an extension of the road network may cause a redistribution of the traffic that results in longer individual running times.

Key words: traffic network planning; paradox; equilibrium; critical flows; optimal flows; existence theorem
History: Received: April 2005; revision received: June 2005; accepted: July 2005.

Translated from the original German: Braess, Dietrich. 1968. Über ein Paradoxon aus der Verkehrsplanung. *Unternehmensforschung* **12** 258–268.

1. Introduction

The distribution of traffic flow on the roads of a traffic network is of interest to traffic planners and traffic controllers. We assume that the number of vehicles per unit time is known for all origin-destination pairs. The expected distribution of vehicles is based on the assumption that the most favorable routes are chosen among all possible ones. How favorable a route is depends on its travel cost. The basis for the evaluation of cost is travel time.

The road network is modeled by a directed graph for the mathematical treatment. A (travel) time is associated with each link. The computation of the most favorable distribution can be considered solved if the travel time for each link is constant, i.e., if the time is independent of the number of vehicles on the link. In this case, it is equivalent to computing the shortest distance between two points of a graph and determining the corresponding critical (here meaning shortest) path. See Bellman (1958), von Falkenhausen (1963), and Pollack and Wiebenson (1960).

In more realistic models, however, one has to take into account that the travel time on the links will strongly depend on the traffic flow. Our investigations will show that we will encounter new effects compared to the model with flow-independent costs. Specifically, a more precise formulation of the problem will be required. We have to distinguish between flow that will be optimal for all vehicles and flow that is achieved if each user attempts to optimize his own route.

Referring to a simple model network with only four nodes, we will discuss typical features that contradict facts that seem to be plausible. Central control of traffic can be advantageous even for those drivers who think that they will discover more profitable routes for themselves. Moreover, there exists the possibility of the paradox that an extension of the road network by an additional road can cause a redistribution of the flow in such a way that increased travel time is the result.

2. Graph and Road Network

Directed graphs are used for modeling road maps, and the links, the connections between the nodes, have an orientation (Berge 1958, von Falkenhausen 1966). Two links that differ only by their direction are depicted in the figures by one line without an arrowhead.

In general, the nodes are associated with street intersections. Whenever a more detailed description is necessary, an intersection may be divided into (four) nodes with each one corresponding to an adjacent road; see Figure 2 (Pollack and Wiebenson 1960).

We will use the following notation for the nodes, links, and flows. The indices belong to finite sets. Because we use each index only in connection with one variable, we do not write the range of the indices.

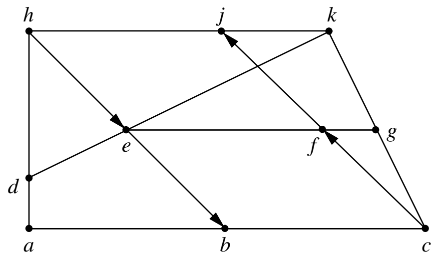

Figure 1

$\{a^i\}$ nodes of the graph.
$\{u_\alpha\}$ oriented links of the graph.
φ_α flow on u_α (vehicles/time).

We consider traffic networks with stationary flows. It is useful to regard the total flow as the sum of threads with each thread being associated with an origin-destination pair. Each thread corresponds to a path in the graph. We need to consider only paths without cycles.

$\{U_\beta\}$ paths that do not contain links more than once.
Φ_β flow along U_β (vehicles/time).
Φ the vector with the components Φ_β.

The flows on paths and links are related by the arc-path incidence matrix C whose coefficients $c_{\alpha\beta}$ assume only the values 0 or 1 because cycles are excluded:

$$c_{\alpha\beta} = \begin{cases} 1 & \text{if link } u_\alpha \text{ is contained in path } U_\beta \\ 0 & \text{otherwise.} \end{cases} \tag{2.1}$$

Obviously,

$$\varphi_\alpha = \sum_\beta c_{\alpha\beta}\Phi_\beta. \tag{2.2}$$

Of course, all flow variables in a traffic network are nonnegative.

For simplicity, a part of the general considerations will be done only for the special case in which the total flow has a common origin node and a common destination node. Those nodes will be denoted as a^0 and a^∞, respectively. The total flow is given by

$$|\Phi| = \sum_\beta \Phi_\beta. \tag{2.3}$$

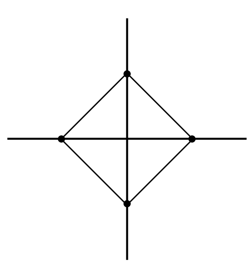

Figure 2

In the general case with more than one origin-destination pair we introduce index sets B_ν such that the groups

$$\{U_\beta;\ \beta \in B_\nu\}$$

contain the paths with the same origin and destination nodes. Here, ν specifies the origin-destination pair. Analogously to (2.3), we have

$$|\Phi_\nu| = \sum_{\beta\in B_\nu} \Phi_\beta. \tag{2.3$'$}$$

Which paths are considered optimal is determined by their respective travel costs. The costs depend on the length of the road, travel time, and other costs (cf. von Falkenhausen 1966, p. 23). The dominating feature is travel time, and for reasons of clarity we identify costs with travel time. In this way, it is also clear that costs depend on the volume of traffic. Moreover, we regard the model as deterministic, and stochastic arguments are deliberately ignored.[1] The following definitions are to be understood in this framework:

$t_\alpha(\varphi)$ travel time required on u_α if u_α carries flow $\varphi = \varphi_\alpha$.
$T_\beta^{ik}(\Phi)$ travel time for getting to a^i from a^k on the path U_β. (If this is impossible, the function value is ∞.) The superscripts i and k will be suppressed if a^i is the destination and a^k is the origin of Φ_β.

Travel time depends on Φ, in particular on the flow on U_β. Because travel time for a path is the sum of the times for its links, we have

$$T_\beta(\Phi) = \sum_\alpha c_{\alpha\beta} t_\alpha(\varphi_\alpha). \tag{2.4}$$

Here, the functional relationship $\varphi_\alpha = \varphi_\alpha(\Phi)$ is given by (2.2). Moreover, we define the most unfavorable time by

$$|T^{ik}(\Phi)| = \max\{T_\beta^{ik}(\Phi);\ \Phi_\beta \neq 0\} \tag{2.5}$$

and

$$|T_\nu(\Phi)| = \max\{T_\beta^{i_\nu k_\nu}(\Phi);\ \beta \in B_\nu,\ \Phi_\beta \neq 0\}, \tag{2.6}$$

where a^{i_ν} and a^{k_ν} are the nodes of the destination of the paths $\{U_\beta,\ \beta \in B_\nu\}$ and of the origins, respectively. The functions t_α are assumed to have the following properties:

I. $t_\alpha \geq 0$.
II. t_α is a nondecreasing function.

[1] For the traffic planner, deviations in traffic densities are not interesting for the design as long as the impact on travel time is small. Therefore, we do not obtain serious errors if individual (stochastic) quantities are replaced by their mean values. It is known that problems arise in models with flow-independent costs because one needs a criterion as to how to assign a portion of flow to nearly optimal (suboptimal) paths (von Falkenhausen 1966).

III. t_α is semicontinuous, i.e., $\lim_{\varphi\to\varphi_0;\,\varphi<\varphi_0} t_\alpha(\varphi) = t_\alpha(\varphi_0)$.

The first two assumptions are natural in view of the problem setting. Assumption 3 simplifies the mathematical treatment. In this case, the functions $t_\alpha(\varphi)$ are lower semicontinuous (Natanson 1961); i.e., we have

$$\lim_{\varphi\to\varphi_0} t_\alpha(\varphi) \ge t_\alpha(\varphi_0). \tag{2.7}$$

In Sections 4 and 5 we will assume continuity for even further simplification.

3. Optimality

We will discuss which flow distributions admit travel times for all drivers to be as short as possible when the traffic networks have only one origin-destination pair. The time that is needed to reach the destination in the most unfavorable case measures how well the flows are distributed. This time is given by Equation (2.6). The total flow $|\Phi| = \chi$ is considered fixed, with χ given.

Definition. The flow Φ is *optimal* if the relation

$$|T(\Phi)| \le |T(\Psi)| \tag{3.1}$$

holds for all Ψ with

$$|\Phi| = |\Psi|. \tag{3.2}$$

It is essential in this definition that the value of a flow distribution is guided by the travel time of *all drivers* (in contrast to the setting in the next section). The concept and the results do not differ substantially if the mean value of the travel time

$$\frac{1}{|\Phi|}\sum_\beta \Phi_\beta T_\beta(\Phi) \tag{3.3}$$

and not the maximal time, $|T(\Phi)|$, determines the quality. It cannot be decided by mathematical arguments which specification is more appropriate. This decision must be left to the traffic planners.[2] We will only postulate the following consistency property: It should be impossible to redistribute optimal flows so that each driver achieves a reduction of the costs.

Now we turn to the flows that are optimal according to (3.1).

Theorem. *Assume that $t_\alpha(\varphi)$ is semicontinuous from below whenever $0 \le \varphi \le \chi$. Then, an optimal flow exists with $|\Phi| = \chi$.*

[2] Optimality could be easily defined for traffic with several origins and destinations on the basis of mean values of travel times. On the other hand, can one require that someone be content with a long drive to reduce the mean travel time? Our definition has an advantage that will become clear later. The optimal solution is at least as advantageous for each driver as the equilibrium.

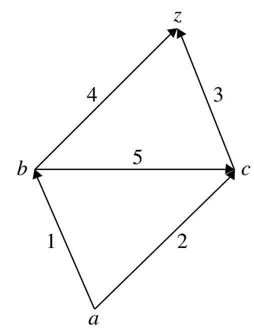

Figure 3

Proof. To prove the theorem, we consider a minimal sequence $\Phi^{(n)}$, i.e., a sequence with $|\Phi^{(n)}| = \chi$ and $\lim_{n\to\infty} |T(\Phi^{(n)})| = \inf\{|T(\Phi)|;\ |\Phi| = \chi\}$. Because the paths do not contain cycles, it follows that $0 \le \Phi_\beta^{(n)} \le \chi$ and $0 \le \varphi_\alpha^{(n)} \le \chi$. Due to the boundedness, a subsequence $\Phi^{(n_\nu)}$ with convergent values for $\Phi_\beta^{(n_\nu)}$ can be selected. Let

$$\lim_{\nu\to\infty} \Phi_\beta^{(n_\nu)} = \Phi_\beta^*.$$

It follows from (2.4), (2.7), and lower semicontinuity that

$$T_\beta(\Phi^*) \le \lim_{\nu\to\infty} T_\beta(\Phi^{(n_\nu)}). \tag{3.4}$$

Let U_β be a path with $\Phi_\beta^* \ne 0$. Then, we have $\Phi_\beta^{(n_\nu)} \ne 0$ for all sufficiently large ν, and consequently for the minimal sequence

$$\begin{aligned}\varliminf_{\nu\to\infty} T_\beta(\Phi^{(n_\nu)}) &\le \lim_{\nu\to\infty} |T(\Phi^{(n_\nu)})| \\ &= \inf\{|T(\Phi)|;\ |\Phi| = \chi\}.\end{aligned} \tag{3.5}$$

The inequalities (3.4) and (3.5) imply that Φ^* is an optimal flow distribution.

We now turn to a model example with four nodes. For convenience, the link travel times $t_\alpha(\varphi)$ are linear functions. (Moreover, the graph does not contain cycles.)

$$\begin{aligned} t_1(\varphi) = t_3(\varphi) &= 10\varphi, \\ t_2(\varphi) = t_4(\varphi) &= 50 + \varphi, \\ t_5(\varphi) &= 10 + \varphi. \end{aligned} \tag{3.6}$$

(a) If a total flow of $|\Phi| = 2$ is to be guided from a to z, the optimal solution is

$$\Phi_{abcz} = 2, \qquad \Phi_{abz} = \Phi_{acz} = 0, \qquad |T(\Phi)| = 52.$$

(b) If a total flow of $|\Phi| = 6$ is to be guided from a to z, the optimal solution is

$$\Phi_{abcz} = 0, \qquad \Phi_{abz} = \Phi_{acz} = 3, \qquad |T(\Phi)| = 83.$$

(c) If a total flow of $|\Phi| = 20$ is to be guided from a to z, the optimal solution is

$$\Phi_{abcz} = 0, \qquad \Phi_{abz} = \Phi_{acz} = 10, \qquad |T(\Phi)| = 160.$$

Obviously, all the solutions are unique.

4. Critical Flow

Each driver attempts to find for himself the most favorable path. It is assumed that he obtains the information that is necessary for determining the route. Therefore, our approach differs significantly from the approach in a game-theoretic consideration; see also the footnote in Section 2.

We consider once more the model example from the last section. If the volume of traffic is as in cases (a) or (c), then it is most profitable to move in accordance with the optimal flow. This is different in case (b). The optimal flow moves along paths (abz) and (acz). There exists, however, a path for which travel time is lower. Specifically, $T_{abcz} = 70 < 83 = |T(\Phi)|$. Suppose that the vehicles are distributed as the optimal flow. Those drivers to which the link travel times are known would move to path ($abcz$) and destroy the optimality.

If the drivers of two vehicles possess perfect information from experience, they will not choose paths with different travel times. Therefore, we consider the hypothesis as realistic that the traffic flow will be distributed in a manner that will be called critical.

Definition. The flow Φ is a *critical flow*[3] if for all paths U_β

$$\begin{aligned} T_\beta(\Phi) &= |T(\Phi)| \quad \text{if } \Phi_\beta \neq 0,\\ T_\beta(\Phi) &\geq |T(\Phi)| \quad \text{if } \Phi_\beta = 0. \end{aligned} \tag{4.1}$$

Obviously, criticality has the following meaning. The destination will be reached on all paths with nonvanishing flow at the same time.[4] Travel time on paths with no flow is the same or even larger. The analogous property holds for all nodes in between.

Theorem. *Let Φ be a critical flow. Then, a number τ^i exists for each node a^i such that for all paths that pass a^i*

$$\begin{aligned} T_\beta^{(i)}(\Phi) &= \tau^i \quad \textit{if } \Phi_\beta \neq 0,\\ T_\beta^{(i)}(\Phi) &\geq \tau^i \quad \textit{if } \Phi_\beta = 0. \end{aligned}$$

Moreover, we have for all paths U_β that run from a^k to a^i

$$T_\beta^{ik}(\Phi) \geq \tau^i - \tau^k,$$

and, moreover,

$$\tau^\infty = |T(\Phi)|.$$

[3] We restrict ourselves to continuous functions t_α. If only semicontinuity is assumed, relations (4.1) have to be replaced by

$$\begin{aligned} T_\beta^-(\Phi) &\leq |T(\Phi)| \quad \text{if } \Phi_\beta \neq 0,\\ T_\beta^+(\Phi) &\geq |T(\Phi)| \quad \text{if } \Phi_\beta = 0. \end{aligned}$$

Here, $T_\beta^+(\Phi)$ and $T_\beta^-(\Phi)$ denote the upper and lower limits, respectively, where jumps occur. The values coincide with $T_\beta(\Phi)$ at points of continuity. The existence theorem in Section 5 also holds in this more general case.

[4] Each path with nonvanishing flows is therefore a critical value in the spirit of standard optimization problems on graphs (Berge 1958).

An indirect proof is easy. Set

$$\tau^i = \min\{T_\beta^{i0}(\Phi);\ \beta\}.$$

If the inequalities in the theorem would not hold, then there would exist a path from a^0 to a^∞ with $T_\beta < |T(\Phi)|$.

Before we deduce the existence of critical flows and additional properties, we provide a numerical result for the model example above. The unique critical flow in case (b) with total flow $|\Phi| = 6$ is

$$\Phi_{abcz} = \Phi_{abz} = \Phi_{acz} = 2$$

and

$$\tau^a = 0, \qquad \tau^b = 40, \qquad \tau^c = 52, \qquad \tau^z = 92.$$

Hence, we have already obtained the result that the critical flow does not always coincide with the optimal flow. This happens not only for the optimality criterion introduced in the last section, but for each consistent definition because there obviously exist flows such that travel time of *all vehicles* is smaller than 92. Although each driver chooses the most favorable path for himself, none of them achieves the value that each one of them could achieve at optimal flow.

In this framework, we also recognize a paradoxical fact. If the link u_5 is eliminated in the road network, the critical flow coincides with the optimal flow; the distribution of the traffic flow is improved in this case. This means that for real-life traffic practice: *In unfavorable situations an extension of the road network may lead to increased travel times.*

5. An Existence Theorem

The existence of critical flows for a given total flow can be shown for continuous and nondecreasing functions $t_\alpha(\varphi)$. This will be done by the reduction to a convex program. Because $t_\alpha(\varphi)$ is nondecreasing, the associated function

$$f_\alpha(\varphi) = \int_0^\varphi t_\alpha(\psi)\,d\psi \tag{5.1}$$

is convex. We do not restrict our attention to traffic with only one origin and one destination and choose a slightly more general definition.

Definition. The flow Φ is a *critical flow* if for all paths U_β with $\beta \in B_\nu$

$$\begin{aligned} T_\beta(\Phi) &= |T_\nu(\Phi)| \quad \text{if } \Phi_\beta \neq 0,\\ T_\beta(\Phi) &\geq |T_\nu(\Phi)| \quad \text{if } \Phi_\beta = 0. \end{aligned} \tag{5.2}$$

Theorem. *Assume that the functions $t_\alpha(\varphi)$ are continuous and nondecreasing for $0 \leq \varphi \leq \chi = \sum_\nu \chi_\nu$. Then, the* *solutions of the convex program*

$$\begin{aligned}&\sum_\alpha f_\alpha(\varphi_\alpha)=\text{Min!}\\ &\varphi_\alpha=\sum_\beta c_{\alpha\beta}\Phi_\beta,\\ &\sum_{\beta\in B_\nu}\Phi_\beta=\chi_\nu,\\ &\Phi_\beta\ge 0\end{aligned}\tag{5.3}$$

are critical flows.

Proof. The variables φ_α may be temporarily eliminated by substitution. The Kuhn-Tucker conditions (Collatz and Wetterling 1966) (with Lagrange multipliers λ_ν and μ_β for the remaining constraints) are

$$\begin{aligned}&\sum_\alpha c_{\alpha\beta}t_\alpha(\varphi_\alpha)-\lambda_\nu-\mu_\beta=0\quad(\beta\in B_\nu)\\ &\sum_{\beta\in B_\nu}\Phi_\beta=\chi_\nu\\ &\Phi_\beta\cdot\mu_\beta=0\\ &\Phi_\beta\ge 0,\qquad \mu_\beta\ge 0.\end{aligned}\tag{5.4}$$

The equations in the first line imply, due to (2.4),

$$\left.\begin{aligned}T_\beta(\Phi)-\lambda_\nu=0\quad\text{if }\Phi_\beta\ne 0,\\ T_\beta(\Phi)-\lambda_\nu\ge 0\quad\text{if }\Phi_\beta=0.\end{aligned}\right.\quad(\beta\in B_\nu).$$

Hence, $\lambda_\nu=|T_\nu(\Phi)|$, and Φ is a critical flow.

The existence of a critical flow is immediate. The set of critical flows is convex. A solution is also the unique solution if the function $t_\alpha(\varphi)$ is strictly monotone at $\varphi=\varphi_\alpha$ for at least one link u_α. Therefore, we can guarantee uniqueness of the solution if each path contains at least one link on which $t(\varphi)$ is strictly monotone on the whole domain.

The possibility of characterizing critical flows as the solutions of a minimization problem is connected with a symmetry in the model. Roughly speaking, we can say: Each driver induces the same delay for the other drivers as the other one does for him. This symmetry no longer holds in a more general model.

In particular, the travel time is not equal for all vehicles on a street with several lanes; it depends on the type (class) of the vehicle. The most significant differences are between passenger cars and trucks. This can be partially incorporated into the theory by dividing the flows into groups:

$$\varphi_\alpha=\sum_g\varphi_\alpha^g,\qquad \Phi_\beta=\sum_g\Phi_\beta^g.$$

The travel time of each group also depends on the amount of flow of the other groups on the corresponding link

$$t_\alpha^g(\varphi)=t_\alpha^g(\varphi^1,\varphi^2,\ldots).$$

The arguments that led us to introduce the concept of critical flows can be directly extended. We do not provide the defining relationships because they differ from (5.2) merely by indices referring to groups. Nevertheless, there is an essential difference. If there is more than one group, the equations can no longer be related to a variational principle.[5]

6. Violation of Monotonicity

The critical flows in the model example of Section 3 are optimal in cases (a) and (c). It is obvious from the numerical solutions that we do not have

$$|\Phi|\le|\Psi|\quad\text{implies the relations }\varphi_\alpha\le\psi_\alpha,\ \Phi_\beta\le\Psi_\beta$$

for critical flows (and consequently also not for optimal ones), although one might expect this. There are consequences for the numerical treatment of the problem, in particular, for an approximation procedure that was recently suggested.

Let n be a given natural number. Determine the shortest paths for all origin-destination pairs of interest when there is no traffic on the roads. It is assumed that the nth portion of the flow chooses those routes. Now, the shortest paths for the new situation are evaluated, and another nth portion of the flow is determined. By proceeding in the same way, the total flow will be distributed on the traffic network. The result is considered to be an approximate solution.

If this method is applied to the model example above with $|\Phi|=20$, then path $(abcz)$ is chosen in the first steps of the procedure although this path carries zero flow in the solution. When n is increased, the approximate solutions do not converge to the correct solution. However, one can use the convex program (5.3) for the computation of the critical flows, and here well-known algorithms are available (Collatz and Wetterling 1966). We did not study whether the evaluation of the shortest paths helps to accelerate the codes.

Acknowledgment

The author thanks Prof. Dr. Helmut Werner for the promotion of this research and Prof. Dr. W. Knödel for valuable stimulating remarks.

References

Bellman, R. 1958. On a routing problem. *Quart. Appl. Math.* **16** 87–90.

Berge, C. 1958. *La Theorie des Graphes*. Paris, France.

Collatz, L., W. Wetterling. 1966. *Optimierungsaufgaben*. Springer, Berlin-Heidelberg-New York.

Natanson, I. P. 1961. *Theorie der Funktionen einer reellen Veränderlichen*. Berlin, Germany.

Pollack, M., W. Wiebenson. 1960. Solutions of the shortest route problem—A review. *Oper. Res.* **8** 224–230.

von Falkenhausen, H. 1963. Ein Verfahren zur Prognose der Verkehrsverteilung in einem geplanten Straßennetz. *Unternehmensforschung* **7** 75–88.

von Falkenhausen, H. 1966. Ein stochastisches Modell zur Verkehrsumlegung. Dissertation, Darmstadt, Germany.

[5] We note that there is an analogous situation with the diffusion equations in reactor physics. The equations for models with one group of neutrons are self-adjoint while the multigroup equations are not.