On the coincidence of system optimum and user equilibrium for a widely used family of cost functions

Paolo Pellizzari and Annamaria Sorato *

Dept. of Applied Mathematics
University "Ca' Foscari" of Venice
Dorsoduro 3825/E, 30123 Venice, Italy.

Abstract. In a simple two-node, one origin-destination network with multiple links, we characterize the coincidence of system optimum, that minimizes the total cost of agents with user equilibrium, that equalizes the cost in each (used) link. If cost functions are, up to a constant, homogeneous of the same degree then the system optimum and the user equilibrium are the same if and only if the freeflows are costant. Some examples show that the hypotheses are not redundant.
Keywords. Traffic flows, system optimum, user equilibrium.

JEL Classification: R41
MSC Classification: 90B20, 60K30

1 Introduction

We consider a model of traffic flow on a network whose roads are used by non-cooperative agents that travel from an origin to a destination, minimizing their own travel time or another suitable cost function. It is well know that the selfish user equilibrium in a traffic network can differ substantially from the globally optimal flow that makes the incurred total (or average) cost minimal. This is due to the different perspective used by the travellers: selfish agents select their route to minimize their own travel time with no regard to the congestion burden they put on others. As this externality is not taken into account, the resulting user equilibrium can be largely different from a system optimum, that is an efficient global allocation minimizing the total cost in the network. This feature is important for the clever design of new or improved networks, in the sense that the behaviour of the users must be taken into consideration, as the Braess paradox strikingly shows, [3] or chapter 6 in [1] for a colloquial presentation whose terminology is repeated in this paper. Along the same research line, other curious and puzzling effects are documented in [4] and [5] where it is shown that

* We thank Marco Li Calzi for drawing our attention on homogeneous cost functions and two anonymous referees for useful remarks. All errors are (sadly) ours:
paolop@unive.it, amsorato@unive.it

polluting emissions can increase even when traffic demand decreases or when a link with zero emissions is added to certain networks.

Recent research work has compared the two equilibria, trying to quantify the extent of the increment of total travelling cost due to the selfish "lack of regulation", as it is dubbed in [8]. Clearly, a system optimum can be obtained enforcing the users to select the appropriate paths, but this has huge coordination costs and would inevitably cause reactions by the agents that are routed along costly (or time consuming) roads. It is in fact very unlikely that an agent tolerates higher costs for the sake of social benefit, when other "equal" users have smaller costs. Enforcement seldom being a useful and realistic policy, efforts have aimed to suggest alternative paths to travellers by radio broadcasts or various signalling devices and adoption of pricing policies.

As theoretical and experimental work, [9], stress the ubiquitousness of sub-optimal selfish equilibria, we are interested in this paper in exploring situations in which the system optimum and the user equilibrium are the same. In detail, we investigate the properties a traffic network should possess to ensure that the two equilibria coincide. This is of obvious interest to a central planner that might modify the network (for example adding or deleting links and adjusting the link costs) in order to exploit the greedy behaviour of the users to achieve maximal efficiency with no visible action (other that proper network construction or modification) or enforcement. To the best of our knowledge, determination of conditions such that system optima are the same than user equilibria has received no attention in the literature.

We address the problem in a simplified setting with only two nodes, one origin-destination pair and multiple paths, that in this case can also be thought as edges or links. The situation is of interest in cases where a heavy traffic demand from the origin to the destination splits in various routes with no common intersection. The simple structure of the network indeed allows Selten and coauthors to develop an experimental study in an identical framework that is studied also in [7].

Our main results shows that from same degree homogeneity of the cost functions up to a constant, it follows that positive system optima are also user equilibria if and only if there are constant freeflows (the cost incurred when the flow on the link is null).

The rest of paper is organized as follows. Section 2 describes the model of a traffic problem on a network in a rather standard way, defines the user equilibrium (UE) and the system optimum (SO) and gives some useful lemmas. In the following Section, we state our main result for a two-nodes network with homogeneous costs up to a constant. Finally, Section 4 is devoted to provide some examples and remarks.

2 The model

We consider a directed network $G = (V, E)$ with vertex set V and edge set E. We restrict our attention to the simple case of unique source-target pair,

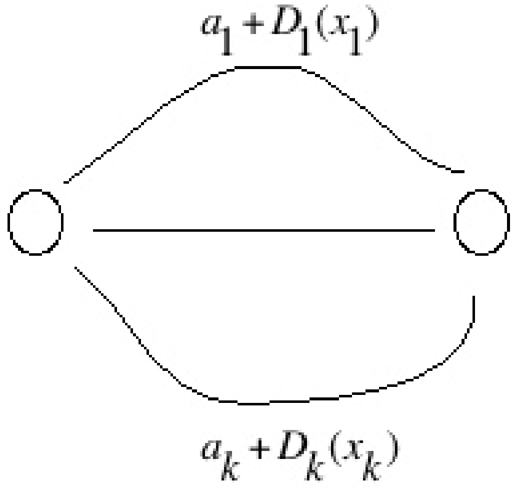

Fig. 1. The network G and some paths linking the unique origin-destination pair. Agents travel from left to right, incurring some congestion-dependent cost.

setting $V = \{O, T\}$ and assume there are k edges (roads) linking O to T. The contribution of each single agent travelling on G is negligible and we assume that unitary traffic is flowing on the network. Define a flow $\mathbf{x}$ as a $k-$tuple of nonnegative real values $\mathbf{x} = (x_1, x_2, \ldots, x_k)$ such that $\sum_{j=1}^{k} x_j = 1$, where each x_j is the fraction of (total) load travelling along the $j-$th edge. This description is, up to a normalization, equivalent to the widely used approach that prescribes some total travel demand N in such a way that q_j travellers are on $j-$th path and $\sum_j q_j = N$. Each edge $e \in E = \{e_1, e_2, \ldots, e_k\}$ is equipped with a differentiable and nondecreasing cost function $C_j(x)$. We assume that the form of the cost function on one edge takes the form $C_j(x) = a_j + D_j(x)$, with $D_j(0) = 0$ and $a_j \geq 0$, in such a way that the cost can be split in two parts: the constant a_j, often named *freeflow*, is the cost incurred when null flow is travelling on the arc, and the term $D_j(x)$ penalizing heavy load. Trivially, the functions D_j inherit the smoothness properties of the costs C_j and are differentiable and nondecreasing. Figure 1 shows a stylized graph of the network.

There are two interesting situations in such a network: a selfish user equilibrium, where agents equally spread congestion on all used links and an efficient system optimum that minimizes the sum (or, equivalently, the mean) of the incurred costs.

Definition 1 (User Equilibrium). *A nonnegative flow* $\overline{\mathbf{x}} = (\overline{x}_1, \ldots, \overline{x}_k)$ *is a User Equilibrium (UE) if for all* $i, j \in \{1, \ldots, k\}$ *such that* $\overline{x}_i > 0$ *and* $\delta \in [0, \overline{x}_i]$, *we have* $\sum_h \overline{x}_h = 1$ *and*

$$C_i(\overline{x}_i) \leq C_j(\overline{x}_j + \delta).$$

Loosely speaking, for each pair of edges, one's cost increases moving from one edge to another. In particular, if $\overline{\mathbf{x}}$ is a UE then all used edges have the same cost, [6].

Lemma 1. *Let $\overline{\mathbf{x}}$ be an UE. Then*

$$C_j(\overline{x}_j) = \lambda,$$

for all j such that $\overline{x}_j > 0$.

Definition 2 (System Optimum). *A flow $\mathbf{x}^* = (x_1^*, \ldots, x_k^*)$ is a System Optimum (SO) if it solves the optmization problem*

$$\min_{\substack{x_1,\ldots,x_k \\ \sum_j x_j = 1 \\ x_1,\ldots,x_k \geq 0}} \quad \mathbf{C}(\mathbf{x}) = \sum_{j=1}^{k} x_j C_j(x_j). \tag{1}$$

The existence of a UE was established in [2], assuming that the cost function of each network link is continuous and nondecreasing in the flow. Moreover, boundedness of the feasible region and continuity of cost functions ensure, by Weierstrass theorem, that a SO exists.

In the next section we provide sufficient and necessary conditions for an UE to be also a SO and viceversa. The following lemma will be useful in the sequel.

Lemma 2. *Let the cost functions be such that, for every $j = 1, \ldots, k$*

$C_j(x) > a_j$, for all $x > 0$;
$C_j(0) = a, \forall j$.

Then

A1. The UE $\overline{\mathbf{x}}$ is such that $\overline{x}_j > 0, j = 1, \ldots, k$, i.e. all links are used by the agents. In brief, $\overline{\mathbf{x}} > 0$.
A2. The SO $\mathbf{x}^*$ is such that $x_j^* > 0, j = 1, \ldots, k$, i.e. nonnegative constraints are not binding. In brief, $\mathbf{x}^* > 0$.

Proof. Let $\overline{\mathbf{x}} = (\overline{x}_1, \ldots, \overline{x}_k)$ be an UE. By contradiction, we assume without loss of generality that $\overline{x}_1 = 0$ and $\overline{x}_2 > 0$. Noting that $D_2(\overline{x}_2) > 0$, by continuity of D_1, there exists $\overline{\delta}$ such that

$$\forall x \in]0, \overline{\delta}[, \quad 0 < D_1(x) < D_2(\overline{x}_2).$$

Hence, there exists a positive $\delta < \overline{x}_2$ such that

$$C_2(\overline{x}_2) > C_1(\overline{x}_1 + \delta) = C_1(\delta),$$

contradicting the assumption that $\overline{\mathbf{x}}$ is UE.

Let us now assume that $\mathbf{x}^* = (x_1^*, \ldots, x_k^*)$ be a SO. By contradiction, we assume without loss of generality that $x_1^* = 0$ and $x_2^* > 0$. We will show that $\mathbf{x}^*$ cannot be optimal showing that moving some amount ϵ of flow from link 2 to 1 reduces the objective function. In other words, we claim that

$$\mathbf{C}(\epsilon, x_2^* - \epsilon, x_3^*, \ldots, x_k^*) < \mathbf{C}(\mathbf{x}^*) \text{ for some } \epsilon > 0.$$

Consider the sum of the first two terms in the summation (1)

$$x_1^* C_1(x_1^*) + x_2^* C_2(x_2^*) = x_2^*(a + D_2(x_2^*))$$

to be compared to

$$\epsilon C_1(\epsilon) + (x_2^* - \epsilon) C_2(x_2^* - \epsilon) = \epsilon(a + D_1(\epsilon)) + (x_2^* - \epsilon)(a + D_2(x_2^* - \epsilon)).$$

Using the definition of the C_j's and a Taylor expansion, the right hand side yields

$$\begin{aligned} &\epsilon a + \epsilon \left(D_1(0) + \epsilon D_1'(0) + o(\epsilon)\right) + \\ &\quad + x_2^* a + x_2^* \left(D_2(x_2^*) - \epsilon D_2'(x_2^*) + o(\epsilon)\right) \\ &\quad - \epsilon a - \epsilon \left(D_2(x_2^*) - \epsilon D_2'(x_2^*) + o(\epsilon)\right). \end{aligned}$$

Simplifying and omitting higher order terms we obtain

$$x_2^*(a + D_2(x_2^*)) - \epsilon \Big(\underbrace{x_2^* D_2'(x_2^*)}_{\geq 0} + \underbrace{D_2(x_2^*)}_{>0} \Big) < x_2^*(a + D_2(x_2^*)),$$

contradicting the optimality of $\mathbf{x}^*$. □

Some remarks are in order. A1 rules out the existence of links that are not used by selfish agents. This does not mean that the links with null flow in an UE are not used in a SO. Indeed, it is well known that there are situations where A2 holds even though A1 does not, as some social benefit can often be obtained rerouting a portion of the traffic on links that would not have been used on a selfish basis.

The previous result basically shows that if the a_j's are constant then both $\overline{\mathbf{x}}$ and $\mathbf{x}^*$ have strictly positive components (i.e. A1 and A2 always hold). In the following section we give our main result, stating that in the presence of homogeneous D_j's, positive SO and UE coincide if and only if the freeflows are constant.

3 Identity of SO and UE

This section is devoted to characterize the networks that admit coincident SO and UE's.

Proposition 1. *Assume the functions D_j be homogeneous of degree $p > 0$ on $\mathbf{R}^+$, nondecreasing and not identically null. Then $\mathbf{x} > 0$ is a SO and a UE if and only if $a_j \equiv a, j = 1, \ldots, k$ i.e. the a_j's are constant.*

Proof. The assumptions of homogeneous, nondecreasing and not identically null D_j's imply that (see [10])

$$D_j(x) = \beta_j x^p, \beta_j > 0, j = 1, \ldots, k.$$

In particular, as each term $x_j C_j$ is a convex function, it follows that the objective function $\mathbf{C}(\mathbf{x})$ in (1) is convex. Hence, being convex the optimization problem for the SO, the Kuhn-Tucker (KT) first order conditions are necessary and sufficient for a solution, see [11].

Assume $\mathbf{x} > 0$ is both a SO and an UE. Then by Lemma 1 we have

$$a_j + D_j(x_j) = \lambda, j = 1, \ldots, k \quad \sum_{j=1}^{k} x_j = 1. \tag{2}$$

As $\mathbf{x}$ is SO also KT conditions must hold, hence

$$\begin{aligned} & a_j + D_j(x_j) + x_j D_j'(x_j) - \xi - \mu_j = 0, \\ & \sum_{j=1}^{k} x_j = 1, \\ & x_j \geq 0, \mu_j \geq 0, \\ & x_j \mu_j = 0, \\ & \forall j = 1, \ldots, k, \end{aligned} \tag{3}$$

where $\xi, \mu_1, \ldots, \mu_k$ are suitable multipliers. Using Euler's theorem for homogeneous functions and observing that $\mathbf{x} > 0$ implies $\mu_j = 0, j = 1, \ldots, k$, equation (3) can be rewritten as

$$a_j + (1 + p) D_j(x_j) - \xi = 0.$$

Recalling from (2) that $D_j(x_j) = \lambda - a_j$ and substituting in the previous equation gives

$$a_j + (1 + p)(\lambda - a_j) - \xi = 0, \quad j = 1, \ldots, k.$$

Hence

$$a_j = \frac{\lambda(1 + p) - \xi}{p},$$

which is a constant independent of j.

Assume now that $a_j \equiv a$ and let $\mathbf{x}^*$ be a SO: we want to show that it is also a UE and $\mathbf{x}^* > 0$. As before, the KT conditions for the problem (1) hold at $\mathbf{x}^*$:

$$a + (1 + p) D_j(x_j^*) - \xi - \mu_j = 0, \tag{4}$$

$$\sum_{j=1}^{k} x_j^* = 1,$$
$$x_j^* \geq 0, \mu_j \geq 0,$$
$$x_j^* \mu_j = 0,$$
$$\forall j = 1, \dots, k,$$

By Lemma 2, $\mathbf{x}^* > 0$ and this in turn gives $\mu_j = 0, j = 1, \dots, k$. Then $D_j(x_j^*) = (\xi - a)/(1+p)$ for all j and $\mathbf{x}^*$ is also an UE by Lemma 1.

Finally, assume that $\overline{\mathbf{x}}$ is an UE (and still $a_j \equiv a$). Strict positiveness of $\overline{\mathbf{x}}$ is an immediate consequence of Lemma 2. Then $D_j(\overline{x}_j) = \overline{\lambda}$ for all j. We want to show that,

$$(\overline{\mathbf{x}}, \xi, \mu_1, \dots, \mu_k) = (\overline{x}_1, \dots, \overline{x}_k, \xi, 0, \dots, 0), \quad \text{for some } \xi, \tag{5}$$

is a full solution of KT conditions for a SO. Plugging (5) into the KT conditions gives

$$a + (1+p)D_j(\overline{x}_j) - \xi = 0,$$
$$\sum_{j=1}^{k} \overline{x}_j = 1,$$
$$\overline{x}_j \geq 0, \mu_j \geq 0,$$
$$\overline{x}_j \mu_j = 0,$$
$$\forall j = 1, \dots, k,$$

that is trivially satisfied if we set $\xi = (1+p)\overline{\lambda} + a$. By sufficiency of KT conditions, we can conclude that $\overline{\mathbf{x}}$ is a SO. □

4 Examples and discussion

This section provides some simple (counter)examples to demonstrate that all the assumptions are required and gives some conclusive remarks.

Example 1. Let $C_1(x) = 10 + 10x, C_2(x) = 30 + 10x$. Then it is immediate to show that the UE and SO are identical, $\overline{\mathbf{x}} = (1, 0) = \mathbf{x}^*$. Though UE coincides with SO, $\mathbf{x}^*$ is not positive and Proposition 1 does not hold being $a_1 = 10 \neq 30 = a_2$.

It is very easy to find networks with non constant freeflows that have different SO and UE, but even in the (rare) cases when $\mathbf{x}^* = \overline{\mathbf{x}}$, Example 1 shows that $\mathbf{x}^*$ is not necessarily positive.

The following example stresses that the assumption of homogeneity of *same* degree p is essential.

Example 2. Let $C_1(x) = 10+10x, C_2(x) = 25/2+10x^2$. Some computations show that

$$\overline{\mathbf{x}} = \left(\frac{1}{2}, \frac{1}{2}\right) = \mathbf{x}^* > 0,$$

but nevertheless the a_j's are different.

We conclude the paper with a final consideration. Homogeneity of the D_j's is a rather acceptable and widely used assumption on the cost functions. With some additional technical hypotheses, we characterize the identity of SO and UE completely in terms of constant freeflows. Most of the traffic networks have however widely different freeflows and this is in total agreement with the well known fact that coincidence of UE and SO is a very rare event. Proposition 1 could hopefully suggest proper behaviour when links of an existing network are modified or added by a central planner. If efficiency is desidered, then an effort should be done in order to achieve costant freeflows as this situation would 'force' unaware selfish agents to adopt socially optimal flows without any explicit form of imposition.

References

1. Batten, D. (2000). *Discovering artificial economics*. Westview Press.
2. Beckmann, C., McGuire, C., and Winsten, C. (1956). *Studies in the economics of transportation*. Yale University Press.
3. Braess, D. (1956). über ein paradoxon der verkehrsplanung. *Unternehmensforschung*, 12:258–268.
4. Nagurney, A. (2000). Congested urban transportation networks and emission paradoxes. *Transportation Research Part D: Transport and Environment*, 5(2):145–151.
5. Nagurney, A. (2001). Paradoxes in networks with zero emission links: implications for telecommunications versus transportation. *Transportation Research Part D: Transport and Environment*, 6(4):283–296.
6. Nagurney, A. and Dong, J. (2002). *Supernetworks: decision-making for the information age*. Edwar Elgar Publishing.
7. Roughgarden, T. (2001). Designing networks for selfish users is hard. Technical report, "Department of Computer Science, Cornell University".
8. Roughgarden, T. and Tardos, E. (2002). How bad is selfish routing? *J. ACM*, 49(2):236–259.
9. Selten, R., Schreckenberg, M., Pitz, T., Chmura, T., and Kube, S. (2003). Experiments and simulations on day-to-day route choice-behaviour. Technical report, CESifo.
10. Simon, C. and Blume, L. (1994). *Mathematics for economists*. Norton.
11. Sundaram, R. K. (1996). *A first course in optimization theory*. Cambridge University Press, Cambridge.