Stability Analysis of the Freeway Ramp Metering Control Strategy ALINEA

Elias B. Kosmatopoulos & Markos Papageorgiou

Abstract— **Local, traffic-responsive ramp metering is a well-known control measure aiming at ameliorating traffic conditions on interurban or metropolitan freeways and motorways. A particular control strategy (ALINEA) has been successfully implemented in a number of freeways where it was shown to improve dramatically the traffic conditions and to be superior to other competing approaches. Stability analysis based on linearized traffic-flow dynamics and the assumption that the flow and occupancy are related through a static "fundamental diagram" has shown that the actual occupancy converges to the desired one as long as the actual occupancy does not exceed a critical value.**

This paper investigates the global stability properties of the ALINEA ramp-metering control strategy and establishes its convergence properties. More precisely, Lyapunov stability theory is used to show that, under realistic assumptions regarding traffic conditions and for any realistic - but otherwise time-varying or dynamic - flow-occupancy relationship and no matter if the actual occupancy at certain time-instants exceeds the critical occupancy: (i) for all choices of controller parameters the ALINEA ramp metering strategy preserves closed-loop stability and moreover forces the actual occupancy to reach the desired one, (ii) under appropriate choice of controller parameters the ALINEA ramp metering strategy guarantees convergence of the tracking error (defined as the difference between the actual occupancy and the desired one) to small bounded sets.

I. Introduction

Freeways and motroways had been originally conceived so as to provide the possibility of fast travel without delays, both in metropolitan and in interurban areas. However, in the last decades, the number and extent of freeway congestions have been steadily increasing, leading to considerable delays, increasing fuel consumption and environmental pollution, and decreasing road safety. Congestions are caused either by high demand that exceeds the freeway capacity (daily *recurrent* congestions) or by capacity-reducing incidents (*nonrcurrent* congestions). In presence of congestion, the freeway throughput becomes lower than capacity, thus rendering the utilization of the expensive infrastructure non-optimal [1]. One proposed way of ameliorating this situation is ramp metering by use of traffic lights at the freeway on-ramps. This control measure aims at limiting access to the freeway mainstream so as to achieve and maintain capacity flow.

This research was partly supported by the European Commission's IST (Information Society Technologies) Programme under the project RHYTHM (IST-2000-29427). All statements of the paper are under the sole responsibility of the authors and do not necessarily reflect European Commission's policies or views.

The authors are with Dynamic Systems & Simulation Laboratory, Dept. of Production & Management Engineering, Technical University of Crete, Chania, GR 73100.

Ramp metering control strategies have been proposed at several levels of sophistication (open-loop/closed-loop, local/centralized, linear/nonlinear, etc.; the interested reader is referred to [2] for an overview of the proposed approaches) but the high majority of implemented and operating systems are of the local, traffic-responsive type. Hundreds of ramp metering installations, mainly in U.S.A. but increasingly also in Europe and elsewhere, have provided valueable experience on the benefits of this kind of control measure.

From a control engineering point of view, the control laws implemented in the vast majority of existing ramp metering systems may be characterized as naive or heuristic. More precisely, these control laws attempt a sort of feedforward disturbance rejecton that renders them particularly sensitive to a variety of changing conditions. The first local ramp metering control strategy that has been based on straightforward application of classical feedback control theory is ALINEA [3], [4]. ALINEA has been successfully implemented to a number of freeways (motorways) where it was shown to improve dramatically the traffic conditions and to be superior to other competing approaches [5]. Stability analysis based on linearized traffic-flow dynamics and the assumption that the traffic flow and occupancy are related through a static "fundamental diagram" have shown that the actual occupancy converges to the desired one as long as the actucal occupancy does not exceed a "critical" value, beyond which the linearized traffic dynamics become unstable. Surprisingly, simulations using realistic traffic flow models as well as field implementations of ALINEA have shown that the actual occupancy converges to the desired one even when the critical value is exceeded. It is worth noticing that in real-life the flow and occupancy are related through a more complex, dynamic, time-varying relationship rather than the static "fundamental diagram" assumed in the linearization-based analysis of [3], [4].

The purpose of this paper is to investigate further the stability properties of the ALINEA ramp-metering control strategy and to establish its global convergence properties. More precisely, in this paper we take advantage of Lyapunov stability theory to show that, under realistic assumptions regarding traffic conditions and for any realistic - but otherwise time-varying or dynamic - flow-occupancy relationship and no matter if the actual occupancy at certain time-instants exceeds the critical value: (i) for all choices of controller parameters the ALINEA ramp metering strategy preserves closed-loop stability and moreover forces the actual occupancy to reach the desired one, (ii) under appropriate choice of controller parameters the ALINEA ramp metering strategy guarantees convergence of the tracking error (defined as the difference between the actual occupancy and the desired one) to small bounded sets.

II. Ramp Metering and the ALINEA Control Strategy

A. Traffic Flow Dynamics

Let

- q_{out} and q_{in} are the mainstream traffic volumes or flows (veh/h) downstream and upstream of the ramp, respectively.
- o denotes the mainstream occupancy rate downstream of the ramp; an occupancy rate measures the time occupancy (in %) of a detector placed below the highway pavement.
- r is the on-ramp traffic volume or flow that may be controlled using ordinary traffic lights; to this end one may employ an one-car-per-green policy, or on a n-cars-per-green policy (with $n = 2, 3$ or more), or a fixed traffic cycle subdivided into a green and a red phase of controllable duration.

The conservation of vehicles gives

$$\dot{\rho} = \frac{1}{\delta}\left(q_{in} + r - q_{out}\right) \tag{2.1}$$

where ρ is the traffic density (veh/km) defined as the number of vehicles included in the stretch, divided by the lenght δ of the stretch. Because traffic density is not readily measurable (it requires video detectors), it is convenient to replace ρ in (2.1) by o, using the approximate relationship $\rho = \alpha o$ where $\alpha = \mu/(100\lambda)$, μ being the number of lanes of the mainstream and λ being the mean effective vehicle length (as "seen" by the underground electromagnetic loop detector; see [6] for more details); this leads to the following model

$$\dot{o} = \frac{1}{\alpha\delta}\left(q_{in} + r - q_{out}\right) \tag{2.2}$$

Obviously, q_{out} depends on the occupancy o. The most popular model in the traffic dynamics literature to describe the relationship between q_{out} and o (or ρ) is a static function $q_{out} = Q(o)$. The function $Q(\cdot)$ (which is known as the "fundamental diagram" of traffic dynamics) obtains its maximum flow (capacity) q_{cap} at o_{cr}, the critical occupancy. Assuming a static flow-occupancy relationship, it is easy to check that the linearized dynamics of (2.2) are stable for $o < o_{cr}$ while they are unstable for $o > o_{cr}$, which practically means that once the actual occupancy becomes larger than the critical one, the traffic flow dynamics tend to increase the occupancy leading to congested traffic conditions.

B. Ramp Metering

Ramp metering refers to techniques for calculating suitable ramp volumes r so as to keep the downstream traffic conditions at - or close to - a desired level. In the case of traffic cycle realization, r is converted to a green phase duration g by use of

$$g = \frac{r}{r_{sat}} C$$

where C is the fixed traffic cycle duration and r_{sat} is the ramp capacity flow (or saturation flow) that may be fixed or estimated in real time, based on ramp flow measurements filtered over some past cycles. g is usually constrained by $g \in [g_{min}, g_{max}]$, with $g_{max} \le C$ and $g_{min} \ge 0$ denoting the maximum and minimum, respectively, allowable green time. Typically $g_{min} > 0$ to avoid ramp closure.

In the case of n-cars-per-green realization, one typically has a constant-duration green time that permits exactly n-vehicles to pass. The ramp volume r is controlled in this case by varying the red phase duration between a minimum and a maximum value.

C. ALINEA

Assuming that the control input r is updated every Δt time-units (typically $\Delta t = C = 20 \ldots 60$ seconds), we obtain, after some algebraic manipulations, the linearized (around a nominal point) and discretized (with sampling interval equal to Δt) version of (2.2) (see [3])

$$\Delta o(k+1) = \beta \Delta o(k) + \frac{1-\beta}{\nabla \hat{Q}}\left[\Delta q_{in}(k) + \Delta r(k)\right] \tag{2.3}$$

Δ is used to denote the difference between the actual value and the nominal one (e.g., $\Delta r = r - r_{nom}$, where r_{nom} denotes the nominal value of r); $\nabla \hat{Q} = \frac{\partial Q}{\partial o}(\hat{o})$ whereby the nominal occupancy o_{nom} is taken equal to the desired occupancy $\hat{o}$; finally $\beta = \exp(-\frac{\nabla \hat{Q}}{\alpha\delta} T)$ results from the discretization. Notice that β may be neglected in (2.3), if the ratio $\delta/\Delta t$ is sufficiently small.

In its simplest form, the ALINEA ramp metering strategy uses a simple I-type regulator [3], [4] given as follows

$$r(k) = r(k-1) + K\left[\hat{o} - o(k)\right] \tag{2.4}$$

where K denotes the positive regulator gain. Applying (2.4) to (2.3) and assuming that β is negligible, we obtain the closed-loop transfer function

$$H(z) = \frac{o(z)}{\hat{o}} = \frac{\frac{K}{\nabla \hat{Q}}}{z - 1 + \frac{K}{\nabla \hat{Q}}}$$

Note that the above transfer function is stable for all positive K, provided that $\nabla \hat{Q}$ is positive; moreover, if this is the case, an optimal dead-beat regulator is obtained by setting $K = \nabla \hat{Q}$. $\nabla \hat{Q}$ is positive only in the LHS of the fundamental diagram (that is, for $\hat{o} < o_{cr}$). It is easy to check that the linearization on the RHS of the fundamental diagram results in an unstable closed-loop transfer function. In other words, the stability analysis based on a linearized version of the traffic flow dynamics results in a stable closed-loop transfer function as long as the occupancy o and its desired value $\hat{o}$ are less than the critical occupancy, while, if this is not the case, the closed-loop transfer function becomes unstable. Similar results hold in the case where β is not negligible and/or a PI-type regulator is used instead of the I-type regulator (2.4) (see [3], [4] for more details).

It is well-established in the control engineering literature that stability analysis based on linearized versions of the plant dynamics can only deliver local stability results, i.e., the results obtained are valid only as long as the plant states remain very close to their nominal values. Moreover, the stability analysis of [3], [4] that was briefly presented above, is based on the assumption of a static flow-occupancy relationship. However, actual field data indicate that flow and occupancy are actually related through a more complex, time-varying, dynamic relationship. Figure 4 shows a speed-flow plot attempting to reflect and classify real freeway measurements.

In our analysis, we will make no assumption about the specific relationship between flow and occupancy (other than that it is a bounded function that is zero at zero). The key idea behind our analysis is to consider $Q(\cdot)$ as a bounded disturbance. In order to illustrate the approach, assume a continuous-time version of ALINEA that uses only proportional feedback, that is

$$r = K(\hat{o} - o) \tag{2.5}$$

Substituting the above equation into (2.2) we obtain

$$\frac{d}{dt}(\hat{o} - o) = -cK(\hat{o} - o) - c(q_{in} - Q) \tag{2.6}$$

where $c = \frac{1}{\alpha\delta} > 0$. It is not difficult to see that if q_{in}, Q are bounded (but otherwise time-varying and dynamic), then the solutions $\hat{o} - o(t)$ of the above differential equation converge to a bounded subset around zero whose radius is proportional to $1/K$. In other words, the control law (2.5) forces the solutions of (2.2) to converge to values arbitrarily close to the desired occupancy. Based on the above simple idea, we present in the next section the stability analysis for the ALINEA ramp metering control algorithm, by incorporating the control constraints imposed on the control input r as well as the hybrid (continuous-time plant dynamics, discrete-time controller dynamics) nature of the closed-loop system.

III. Lyapunov Stability Analysis

Our first step is to introduce a more realistic model for the traffic flow dynamics than the model used in (2.2). Since in real life the occupancy cannot become negative and, moreover, cannot exceed a maximum level, we use the following traffic flow model instead of (2.2)

$$\dot{o} = \mathcal{P}^{o}_{[0,o_{max}]}\left\{\frac{1}{\alpha\delta}(q_{in} + r - Q)\right\} \tag{3.1}$$

The above equation for the traffic flow dynamics is the same as the model in (2.2) derived from the flow-conservation law, with the difference that in (3.1) the projection operator $\mathcal{P}_{[0,o_{max}]}$ is used to guarantee that the solutions $o(t)$ of (3.1) remain always in the set $[0, o_{max}]$, where o_{max} is the maximum occupancy (100%). The projection operator is defined as follows

$$\mathcal{P}^{x}_{[x_{min},x_{max}]}(y) \begin{cases} y & \text{if } x \in (x_{min}, x_{max}) \\ y & \text{if } x = x_{min} \text{ and } y > 0 \\ y & \text{if } x = x_{max} \text{ and } y < 0 \\ 0 & \text{if } x = x_{min} \text{ and } y \leq 0 \\ 0 & \text{if } x = x_{max} \text{ and } y \geq 0 \end{cases} \tag{3.2}$$

Obviously the projection operator in (3.1) guarantees that $o(t) \in [0, o_{max}]$ for all t. For the real system this could be interpreted as follows: (i) if $o = o_{max}$ then q_{in} is sufficiently reduced if necessary; (ii) if $o = 0$ then Q is sufficiently reduced if necessary; both for obvious physical reasons.

The ALINEA ramp metering algorithm is a discrete-time feedback control method, that updates the control signal r every Δt time-units, where Δt denotes the controller sampling time (i.e., $r(t)$ remains constant in the intervals $[(i-1)\Delta t, i\Delta t)$ for all positive integers i). Before we present the ALINEA algorithm we define for each time-instant t, the integer $\kappa(t)$ satisfying

$$t = \kappa(t)\Delta t + \mathrm{mod}(t, \Delta t)$$

where $\mathrm{mod}(\cdot)$ is the modulus function. Moreover, we define

$$[t] \stackrel{\triangle}{=} \kappa(t)\Delta t$$

i.e., for each t, $[t]$ denotes the most recent time-instant at which $r(t)$ has been updated.

Using the above definitions, the ALINEA ramp metering strategy becomes

$$r(t) = \begin{cases} r_{min} \text{if } r([t-\Delta t]) + K(\hat{o} - o([t])) < r_{min} \\ \quad \text{or } o([t]) > \bar{o}_{max} \\ r([t-\Delta t]) + K(\hat{o} - o([t])) \text{if } r([t-\Delta t]) \\ \quad +K(\hat{o} - o([t])) \in [r_{min}, r_{max}] \text{ and } o([t]) \leq \bar{o}_{max} \\ r_{max} \text{if } r([t-\Delta t]) + K(\hat{o} - o([t])) > r_{max} \\ \quad \text{and } o([t]) \leq \bar{o}_{max} \end{cases} \tag{3.3}$$

where, $\hat{o}$ denotes the desired occupancy, $K > 0$ denotes the controller gain, r_{min}, r_{max} denote the minimum and maximum allowable on-ramp flow, respectively, and $\bar{o}_{max} < o_{max}$ is a positive design constant that is set close to maximum occupancy o_{max}.

Remark 1: In the stability analysis that follows, we make no assumption that K is a constant. All the results of these paper are applicable to either constant or time-varying controller gains K. ⋄

Remark 2: The form of ALINEA as originally proposed in [3], [4] does not include the condition where $r(t) = r_{min}$ if $o([t]) > \bar{o}_{max}$. The reason why this extra condition is included in (3.3) is to avoid situations where the motorway is saturated at its maximum (i.e., $o(t) > \bar{o}_{max}$) and ALINEA produces large $r(t)$. This may happen in cases where the controller gain K is small, in which case the reaction of ALINEA is slow (notice that the larger the controller gain K, the more aggressive ALINEA's control behavior). The extra condition $r(t) = r_{min}$ if $o([t]) > \bar{o}_{max}$ simplifies the stability analysis significantly, as it permits to omit the analysis of cases where $o(t) \approx o_{max}$ and $r(t) > r_{min}$ in (3.1), that is, cases where, although the traffic is saturated at its maximum, ALINEA allows significant amounts of vehicles to enter through the ramp. $\diamond$

We will make the following assumptions:

(A1) q_{in}, Q are bounded, satisfying $Q \leq Q_{max}, q_{in} \leq q_{max}$, where Q_{max}, q_{max} are positive reals.

(A2) $Q \geq 0$ with $o = 0 \Rightarrow Q = 0$.

(A3) $q_{in} + r_{min} \geq 0$.

(A4) $r_{max} - Q(t) + q_{in}(t) \geq c_1 > 0$, $r_{min} - Q(t) + q_{in}(t) \leq -c_2 < 0$ for all t, where c_1, c_2 are positive constants.

Note that we make no assumption regarding the specific form of Q. In our analysis, Q can be any function of o and t as long as it is bounded and it is zero when the occupancy is zero. Almost all the flow-occupancy models (fundamental diagrams) proposed in the literature (either through mathematical analysis or experimentation with real data) satisfy assumptions (A1), (A2) and thus are covered by our analysis.

Assumptions (A3), (A4) are "controllability" assumptions. If these assumptions do not hold, then no ramp metering policy can stabilize the system around the desired occupany $\hat{o}$.

The stability analysis of ALINEA will be made possible through the Lyapunov function V, defined as follows

$$V = \frac{1}{2}(\hat{o} - o)^2 = \frac{1}{2}\tilde{o}^2$$

where $\tilde{o} \stackrel{\triangle}{=} \hat{o} - o$ denotes the tracking error. Using (3.1) we obtain

$$\begin{aligned} \dot{V} &= -\dot{o}\tilde{o} \\ &= -\mathcal{P}^{o}_{[0,o_{max}]}\left\{\frac{1}{\alpha\delta}\left(q_{in} + r - Q\right)\right\}\tilde{o} \end{aligned}$$

Let t_s denote a time-instant at which $\tilde{o}(t)$ changes sign, that is

$$\mathrm{sgn}(\tilde{o}(t_s^-)) = -\mathrm{sgn}(\tilde{o}(t_s^+))$$

and consider a time-interval (t_s, T) at which the sign of $\tilde{o}$ remains constant, i.e.,

$$\mathrm{sgn}(\tilde{o}(t)) = \mathrm{sgn}(\tilde{o}(t_s^+)),\ \forall t \in (t_s, T)$$

We have for all $t \in [t_s, T)$

$$\dot{V} = -\mathrm{sgn}(\tilde{o}(t))\mathcal{P}^{o}_{[0,o_{max}]}\left\{\frac{1}{\alpha\delta}\left(q_{in} + r(t) - Q\right)\right\}|\tilde{o}(t)|$$

Consider any $t \in (t_s, T)$. We have two cases:

Case 1:

$$\mathrm{sgn}(\tilde{o}(t)) = \mathrm{sgn}(\tilde{o}(t_s^+)) = +1$$

i.e., the occupancy $o(t)$ is smaller than the desired occupancy $\hat{o}$. Since $K(\hat{o}(t) - o(t)) = K\tilde{o}(t) > 0$, we have from (3.3)

$$r(t) = \min\left\{r_{max}, r([t_s]) + K\sum_{i=\kappa(t_s)}^{\kappa(t)}|\tilde{o}([t_s] + (i - \kappa(t_s))\Delta t)|\right\} \tag{3.4}$$

Using the above equation, the fact that $o(t) < \hat{o} < o_{max}$, assumptions (A2), (A3) and the definition of the projector operator (3.2) it can be seen that $\dot{V}$ is negative definite if

$$\min\left\{r_{max}, r([t_s]) + K\sum_{i=\kappa(t_s)}^{\kappa(t)}|\tilde{o}([t_s] + (i - \kappa(t_s))\Delta t)|\right\} > Q(t) - q_{in}(t)$$

But, since $\tilde{o}(t) > 0$, $\forall t \in (t_s, T)$, we have that $K\sum_{i=\kappa(t_s)}^{\kappa(t)}|\tilde{o}([t_s] + (i - \kappa(t_s))\Delta t)|$ is an increasing function[1] on (t_s, T), and therefore (by taking into account the fact that q_{in}, Q are bounded), there exists a $t_p \in (t_s, T)$ such that, $\forall t \in [t_p, T)$

$$K\sum_{i=\kappa(t_s)}^{\kappa(t)}|\tilde{o}([t_s] + (i - \kappa(t_s))\Delta t)| > Q(t) - q_{in}(t) - r([t_s])$$

Combining the above inequality with (A1), (A3), (A4) and (3.4), we obtain $\dot{V}(t) < 0$ for all $t \geq t_p, t < T$ and therefore, using standard Lyapunov stability arguments, $\tilde{o}(t)$ is strictly decreasing on (t_p, T), which implies that $\tilde{o}(t)$ will eventually cross zero.

(C1) In the case $o(t) < \hat{o}$, the control law (3.3) forces $o(t)$ to reach $\hat{o}$. In other words, assuming that $\tilde{o}$ at t_s crosses zero and becomes positive after t_s, then $\tilde{o}(t)$ may stay positive for some period of time, but will eventually start decreasing until it reaches 0 again.

The question in hand is, while $\tilde{o}(t)$ stays positive, what is the largest value it can take. To answer this question, consider again that $t \in (t_s, T)$. In this case one of the following two conditions is satisfied:

(a)[$r(t) < r_{max}$]: In this case, we have

$$\tilde{o}(t) = \tilde{o}([t])\left(1 - \frac{K}{\alpha\delta}(t - [t])\right) + \bar{q}(t, [t])$$

where

$$\bar{q}(t,[t]) = -\frac{1}{\alpha\delta}\left(\int_{[t]}^{t} q_{in}(\tau)d\tau + r([t])(t - [t]) - \int_{[t]}^{t} Q(\tau)d\tau\right)$$

Note that, since q_{in}, r, Q are nonnegative and bounded quantities, there exists a positive constant c_+, independent of K, t, t_s, such that

$$\bar{q}(t,[t]) \leq c_+(t - [t]) \tag{3.5}$$

Therefore

$$|\tilde{o}(t)| \leq |\tilde{o}([t])|\left|1 - \frac{K}{\alpha\delta}(t - [t])\right| + c_+(t - [t]) \tag{3.6}$$

[1] Note that, although the term $K\sum_{i=\kappa(t_s)}^{\kappa(t)}|\tilde{o}([t_s] + (i - \kappa(t_s))\Delta t)|$ remains constant at the time-intervals $([t_s] + \kappa(t)\Delta t, (\kappa(t)+1)\Delta t)$, it strictly increases at the time-instants $[t_s] + \kappa(t)\Delta t$, $\forall t \in (t_s, T)$. In what follows, we will say that a function of time is *strictly decreasing (increasing) wrt* Δt *on the time-interval* (t_1, t_2) if it is non-increasing (resp. non-decreasing) for $t \in ((i-1)\Delta t, i\Delta t)$ but it is stirctly decreasing (resp. increasing) at $t = i\Delta t$ for all i that are positive integers and satisfy $((i-1)\Delta t, i\Delta t) \subset (t_1, t_2)$.

Assume that K satisfies the following inequality

$$\frac{(1-\varepsilon)\alpha\delta}{\Delta t} \leq K \leq \frac{(1+\varepsilon)\alpha\delta}{\Delta t} \Longrightarrow \left|1-\frac{K}{\alpha\delta}\Delta t\right| \leq \varepsilon \quad (3.7)$$

where $\varepsilon \in [0,1)$ is a nonnegative constant. Then, it can be seen that we obtain

$$|\tilde{o}([t]+\Delta t)| \leq |\tilde{o}([t])|\varepsilon + c_+\Delta t \quad (3.8)$$

Using the fact that $\varepsilon \in (0,1)$ we obtain

$$|\tilde{o}([t])| < \frac{1}{1-\varepsilon}c_+\Delta t \quad (3.9)$$

Moreover, using the above inequality, (3.7) and (3.6) we obtain

$$|\tilde{o}(t)| < \left(1+\frac{1}{1-\varepsilon}\right)c_+\Delta t \quad (3.10)$$

Finally, from the definition of V and (3.8) we have

$$\begin{aligned} V([t+\Delta t]) &= \frac{1}{2}\left(|\tilde{o}([t]+\Delta t)|\right)^2 \\ &\leq \frac{1}{2}\left(|\tilde{o}([t])|\varepsilon + c_+\Delta t\right)^2 \\ &\leq \frac{1}{2}\left(|\tilde{o}([t])|^2\varepsilon^2 c_+^2\Delta t^2\right) \\ &\leq \varepsilon^2 V([t]) + \frac{1}{2}c_+^2\Delta t^2 \quad (3.11) \end{aligned}$$

(b) [$r(t) = r_{max}$]: From (A4) we have that $\dot{V}$ is negative definite, that is

$$\begin{aligned} \dot{V} &= -\frac{1}{\alpha\delta}\left(q_{in}(t) + r_{max} - Q(t)\right)|\tilde{o}(t)| \\ &\leq -\frac{c_1}{\alpha\delta}|\tilde{o}(t)| \leq -\frac{c_1}{\alpha\delta o_{max}}|\tilde{o}(t)|^2 \\ &= -\frac{2c_1}{\alpha\delta o_{max}}V(t) \quad (3.12) \end{aligned}$$

and thus $|\tilde{o}(t)|$ is strictly decreasing after $r(t)$ becomes equal to r_{max} (in the last inequality of (3.12) we made use of $|\tilde{o}(t)| \leq o_{max}$, $\forall t$ which results from (3.1)).
Combining the analysis of (a) and (b), we conclude:
(C2) Suppose that at $t = t_s$, $\tilde{o}(t)$ crosses zero and becomes positive for $t \in (t_s, T)$. Then, if K satisfies (3.7), $|\tilde{o}(t)|$ satisfies

$$|\tilde{o}(t)| < \left(1+\frac{1}{1-\varepsilon}\right)c_+\Delta t, \forall t \in (t_s, T)$$

Case 2:

$$\mathrm{sgn}(\tilde{o}(t)) = \mathrm{sgn}(\tilde{o}(t_s^+)) = -1$$

i.e., the occupancy $o(t)$ is larger than the desired occupancy $\hat{o}$. Since $K(\hat{o}(t) - o(t)) = K\tilde{o}(t) < 0$, we have from (3.3)

$$\begin{aligned} r(t) &= \mathcal{I}_{sat}([t])r_{min} + (1-\mathcal{I}_{sat}([t]))\max\{r_{min}, r([t_s]) \\ &\quad -K\sum_{i=\kappa(t_s)}^{\kappa(t)}|\tilde{o}([t_s]+(i-\kappa(t_s))\Delta t)|\} \quad (3.13) \end{aligned}$$

where $\mathcal{I}_{sat}(t)$ is an indicator function satisfying $\mathcal{I}_{sat}(t) = 1$ if $o(t) > \bar{o}_{max}$ and $\mathcal{I}_{sat}(t) = 0$, otherwise. At this point, we will add an extra assumption.
(A5) $o_{max} - \bar{o}_{max} > c_+\Delta t - \hat{o}$.
This assumption will be utilized in our analysis as follows: first, notice that $\tilde{o}(t)$ satisfies (3.5). Using (3.5), (3.5) and (A5) it can be seen that $o(t) < o_{max}$ for all $t \in (t_s, T)$ for all positive K. Using (3.13), the facts that $o(t) > \hat{o}$ and $r([t]) = r_{min}$ if $o([t]) > \bar{o}_{max}$, the definition of the projector operator (3.2) and $o(t) < o_{max}$ resulting from assumption (A5), it can be seen that $\dot{V}$ is negative definite if either $r([t]) = r_{min}$ or

$$\max\left\{r_{min}, r([t_s]) - K\sum_{i=\kappa(t_s)}^{\kappa(t)}|\tilde{o}([t_s]+(i-\kappa(t_s))\Delta t)|\right\} > Q(t) - q_{in}(t)$$

But, since $|\tilde{o}(t)| > 0$, $\forall t \in (t_s, T)$, we have that $K\sum_{i=\kappa(t_s)}^{\kappa(t)}|\tilde{o}([t_s]+(i-\kappa(t_s))\Delta t)|$ is a strictly increasing function wrt Δt on (t_s, T), and therefore (taking into account the boundedness of q_{in}, Q), there exists a $t_n \in (t_s, T)$ such that, $\forall t \in [t_n, T)$

$$K\sum_{i=\kappa(t_s)}^{\kappa(t)}|\tilde{o}([t_s]+(i-\kappa(t_s))\Delta t)| > Q(t) - q_{in}(t) - r([t_s])$$

Combining the above inequality with (A1), (A3), (A4) and (3.13), we obtain $\dot{V}(t) < 0$ for all $t \geq t_n, t < T$ and therefore, using standard Lyapunov stability arguments, $\tilde{o}(t)$ is strictly decreasing on (t_n, T), which implies that $\tilde{o}(t)$ will eventually cross zero.
(C3) In the case $o(t) > \hat{o}$, the control law (3.3) forces $o(t)$ to reach $\hat{o}$. In other words, assuming that $\tilde{o}$ crosses zero at t_s and becomes negative after t_s, then $\tilde{o}(t)$ may stay negative for some period of time, but will eventually start increasing until it reaches 0 again.
Similarly to the analysis performed in Case 1, we can see that,

- In the case $r(t) > r_{min}$, $\tilde{o}(t)$ satisfies (3.5), and, therefore, using the same analysis as in Case 1, if K satisfies (3.7) then $|\tilde{o}|$ and V satisfy (3.9), (3.10) and (3.11).
- In the case $r(t) = r_{min}$, it follows form (**??**) that $\dot{V}(t)$ is negative definite and therefore $|\tilde{o}(t)|$ is a strictly decreasing function of time.

In summary:
(C4) Suppose that at $t = t_s$, $\tilde{o}(t)$ crosses zero and becomes negative for $t \in (t_s, T)$. Then, if K satisfies (3.7), $|\tilde{o}(t)|$ satisfies

$$|\tilde{o}(t)| < \left(1+\frac{1}{1-\varepsilon}\right)c_+\Delta t,\ \forall t \in (t_s, T)$$

The next Theorem summarizes the results of our analysis:

Theorem 1: Consider the traffic flow dynamics (3.1), the control law (3.3) and assume that (A1)-(A5) hold. Then,

(a) For all positive controller gains K, the control law (3.3) guarantees that, if $\tilde{o}(t) \neq 0$ at some time-instant t, $\tilde{o}$ will be forced to cross zero.
(b) If K satisfies (3.7), then, after $\tilde{o}(t)$ crosses zero for the first time, it remains bounded in the subset $\mathcal{C}$ defined as follows

$$\mathcal{C} = \left\{ |\tilde{o}(t)| \in \Re_+ : |\tilde{o}(t)| < \left(1 + \frac{1}{1-\varepsilon}\right) c_+ \Delta t \right\}$$

where $\varepsilon, c_+, \Delta t$ have been defined in the stability analysis.
(c) If K satisfies (3.7) the convergence of $\tilde{o}(t)$ to the subset $\mathcal{C}$ is exponential.

Remark 3: Control law (3.3) corresponds to the I-type regulator version of the ALINEA ramp metering strategy proposed in [3], [4]. The results of this paper can be easily extended to the case where a PI-type regulator is used. $\diamond$

Remark 4: Relation (3.7) defines the "optimal" choices for the controller gain K, in the sense that for K satisfying (3.7) we obtain the smaller worst-case bounds (as defined in the subset $\mathcal{C}$) for the tracking error. Any other choice that does not satisfy (3.7) may lead to tracking errors that are larger than the ones defined in the subset $\mathcal{C}$ as it can be easily seen by using inequality (3.6). Notice also that among all choices for K that satisfy (3.7), the one that corresponds to the smallest worst-case bound for the tracking error is the one for $\varepsilon = 0$.

For a typical freeway example, the length of the stretch $\delta = 0.2$ km, the number of lanes $\mu = 3$ and the sampling interval $\Delta t = 60$ seconds $= 1/60$ hours.Assuming a mean effective vehicle length $\lambda = 6$ meters $= 0.006$ km, we obtain from (3.7) for $\varepsilon = 0$ the regulator gain value $K = 60$, which is very close to the values ($K \approx 70$) chosen in actual field implementations of ALINEA [2]-[5].

$\diamond$

Acknowledgements

The authors express their deep appreciation to Gabriel Gomes for his helful suggestions, corrections and comments on a preliminary version of this paper.

References

[1] C. Chen, Z. Jia, and P. Varaiya, "Causes and cures of highway congestion," *IEEE Control Systems Magazine*, pp. 26-32, Dec. 2001.

[2] M. Papageorgiou and A. Kotsialos, "Freeway ramp metering: An overview," *Proc. of the 2000 IEEE Conference on Intelligent Transportation Systems*, pp. 228-239, Dearborn (MI), USA, Oct. 1-3, 2000.

[3] M. Papageorgiou, H. Hadj-Salem, J.-M, Blosseville, "ALINEA: A local feedback control law for on-ramp metering," *Transportation Research Record*, vol. 1320, pp. 58-64, 1991.

[4] M. Papageorgiou and H. Hadj-Salem, "A low cost tool for freeway ramp metering," *Preprints of the 4th IFAC Symp. Low Cost Automation*, Buenos Aires, Argentina, pp. 54-59, Sept. 13-15, 1995.

[5] M. Papageorgiou, M., H. Haj-Salem, and F. Middleham, "ALINEA local ramp metering: Summary of field results," *Transportation Research Record 1603*, pp. 90-98, TRB, National Research Council, Washington D.C., 1997.

[6] Z. Jia, C. Chen, B. Coifman, and P. Varaiya, " The PeMS algorithms for accurate, real-time estimates of g-factors and speeds from single-loop detectors," *Proc. 4th IEEE Conference on Intelligent Transportation Systems*, pp. 538-543, Oakland, California, USA, 2001.

[7] *Traffic Flow Theory* document of TRB, http://www.tfhrc.gov/its/tft/chap2.pdf.