Autonomous vehicle scheduling in networks

Jeroen van Riel

August 2025

Contents

From here on, we will denote the state as vectors $x = (x_1, x_2)$, with position x_1 and velocity x_2. To support the characterization of optimal trajectories in the next subsection, we introduce some auxiliary functions to describe parts of a trajectory where $u(t) \in \{-\omega, \omega\}$. Without any additional constraints, the time it takes to fully accelerate from rest to maximum velocity is given by $1/\omega$. Similarly, it takes this same amount of time to fully decelerate from maximum velocity to rest.

Given some initial state $x \in \mathbb{R} \times [0, 1]$ and start time a and end time b such that $a \leq b$ and $x_2 + (b - a)\omega \leq 1$, let the acceleration trajectory $x^+[x, a, b] : [a, b] \to \mathbb{R}^2$ be defined as

$$x^+[x, a, b](\tau) := \begin{pmatrix} x_1 + x_2(\tau - a) + \omega(\tau - a)^2/2 \\ x_2 + \omega(\tau - a) \end{pmatrix}.$$

Similarly, for x, a, b satisfying $a \leq b$, $x_2 - (b - a)\omega \geq 0$, let the deceleration trajectory $x^-[x, a, b] : [a, b] \to \mathbb{R}^2$ be defined as

$$x^-[x, a, b](\tau) := \begin{pmatrix} x_1 + x_2(\tau - a) - \omega(\tau - a)^2/2 \\ x_2 - \omega(\tau - a) \end{pmatrix}.$$

Furthermore, we use the following notation for trajectories with minimum or maximum speed. Let $x^0[p, a, b](\tau) = (p, 0)$ model a vehicle that is standing still and let $x^1[p, a, b](\tau) = (p + \tau - a, 1)$ model a vehicle that drives at full speed, where $\tau \in [a, b]$ in both cases.

We will now examine four cases in which two of these partial trajectories can be joined by a deceleration part. We will later use these results to characterize optimal trajectories for our optimal control problem.

Lemma 1 **($x^1 \to x^0$).** *Let $x^1[p, a, b]$ and $x^0[q, c, d]$ be two trajectories. Considering τ_1 and τ_2 as variables in the equation*

$$x^-[x^1[p, a, b](\tau_1), \tau_1, \tau_2](\tau_2) = x^0[q, c, d](\tau_2),$$

it has solution $\tau_2 = q - p + a + 1/2\omega$ and $\tau_1 = \tau_2 - 1/\omega$, whenever $\tau_1 \in [a, b]$ and $\tau_2 \in [c, d]$.

Proof. The expanded system of state equations is given by

$$\begin{cases} p + \tau_1 - a + (\tau_2 - \tau_1) - \omega(\tau_2 - \tau_1)^2/2 = q, \\ 1 - \omega(\tau_2 - \tau_1) = 0. \end{cases}$$

The second equation yields $\tau_2 - \tau_1 = 1/\omega$, which after substituting back in the first equation yields $p - a + \tau_2 - 1/2\omega - q = 0$, from which the stated solution follows. $\square$

To keep the expressions for the case of joining $x^1 \to x^+$ a little bit simpler, we first consider a full line joining to a acceleration trajectory of full length $1/\omega$.

Lemma 2. *Consider some full acceleration trajectory $x^+[(p, 0), a, a + 1/\omega]$ and the line through $(\lambda, 0)$ with slope 1. Whenever λ, which can be interpreted as a time epoch, satisfies $\lambda \in [a - p - 1/2\omega, a - p + 1/2\omega]$, then the equation*

$$x^+[(p, 0), a, a + 1/\omega](\tau) = x^-[(q, 1), q + \lambda, q + \lambda + 1/\omega](\tau),$$

with τ and q considered as variables, has a unique solution

$$\begin{aligned} \tau &= a + 1/\omega - \sqrt{\frac{a - p + 1/2\omega - \lambda}{\omega}}, \\ q &= 2\tau - a - 1/\omega - \lambda, \end{aligned}$$

so the joining deceleration is given by $x^-[(q, 1), q + \lambda, \tau]$

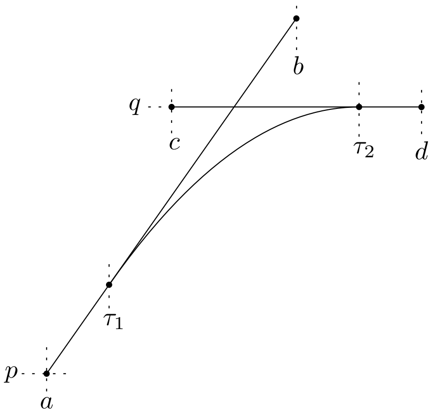

Figure 1: $x^1 \to x^0$

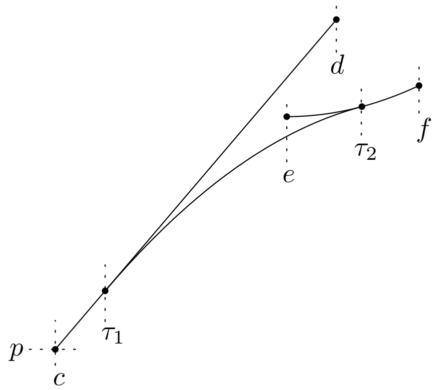

Figure 2: $x^1 \to x^+$

Proof. First of all, the expanded system of state equations is given by

$$\begin{cases} p + \omega(\tau - a)^2/2 = q + (\tau - q - \lambda) - \omega(\tau - q - \lambda)^2/2, \\ \omega(\tau - a) = 1 - \omega(\tau - q - \lambda). \end{cases}$$

We use the second equation to express q in terms of τ, which yields

$$q = 2\tau - 1/\omega - a - \lambda,$$

which we substitute in the first equation to derive the equation

$$\omega\tau^2 - 2(1 + \omega a)\tau + \omega a^2 + a + p + \lambda + 1/2\omega = 0.$$

This is a quadratic equation in τ, with solutions

$$\tau = a + 1/\omega \pm \sqrt{\frac{a - p + 1/2\omega - \lambda}{\omega}},$$

of which only the smallest one is valid, because $\tau \leq a + 1/\omega$. Furthermore, we see that τ is defined as a real number when

$$a - p + 1/2\omega - \lambda \geq 0 \iff \lambda \leq a - p + 1/2\omega.$$

The other requirement is that $\tau \geq a$, which is equivalent to

$$1/\omega \geq \sqrt{\frac{a - p + 1/2\omega - \lambda}{\omega}} \iff \lambda \geq a - p - 1/2\omega.$$

□

Lemma 3 **($x^1 \to x^+$).** *Consider partial trajectories $x^1[p, c, d]$ and $x^+[x, e, f]$.*

Proof. First of all, observe that $x^1[p, c, d]$ lies on the line with slope 1 through $(\lambda, 0) := (c-p, 0)$ and $x^+[x, e, f]$ lies on the full acceleration curve $x^+[(x_1 - x_2^2/(2\omega), 0), e - x_2/\omega, e - x_2/\omega + 1/\omega]$, see Figure 2. Now apply Lemma 2 to $p = x_1 - x_2^2/(2\omega)$, $a = e - x_2/\omega$ and $\lambda = c - p$ yields some solutions τ and q. Let $\tau_2 := \tau$ and let τ_1 denote the time where the line and x^+ join, given by $\tau_1 = \lambda + q$. Now we simply check whether this solution is also feasible for the smaller trajectories. We must have $\tau_1 \in [c, d]$ and $\tau_2 \in [e, f]$. □

Lemma 4. *Consider the acceleration trajectory $x^+[(p, 0), a, b]$ and the horizontal line through $(0, q)$. Let $\tau_1 = a + \sqrt{(q - p)/\omega}$ and $\tau_2 = a + 2\sqrt{(q - p)/\omega}$. If τ_1 satisfies $\tau_1 \in [a, b]$, then both trajectories are joined by deceleration trajectory $x^-[x^+[(p, 0), a, b](\tau_1), \tau_1, \tau_2]$*

Proof. Consider the following equation

$$x^-[x^+[(p,0),a,b](\tau_1),\tau_1,\tau_2](\tau_2) = (q,0).$$

The expanded system of state equations is given by

$$\begin{cases} p + \omega(\tau_1 - a)^2/2 + (\omega(\tau_1 - a))(\tau_2 - \tau_1) - \omega(\tau_2 - \tau_1)^2/2 = q, \\ \omega(\tau_1 - a) - \omega(\tau_2 - \tau_1) = 0. \end{cases}$$

From the second equation, we derive $\tau_1 - a = \tau_2 - \tau_1$. Plugging this back in the first equation yields the quadratic equation $p + \omega(\tau_1 - a)^2 = q$ with solutions $\tau_1 = a \pm \sqrt{(q-p)/\omega}$, of which only the larger one is valid. Finally, the second equation gives $\tau_2 = 2\tau_1 - a$. □

Lemma 5 **($x^+ \to x^0$).** *Consider partial trajectories $x^+[x,c,d]$ and $x^0[q,e,f]$.*

Proof. Observe that $x^+[x,c,d]$ lies on the full acceleration curve $x^+[(x_1 - x_2^2/(2\omega), 0), c - x_2/\omega, c - x_2/\omega + 1/\omega]$. Hence, we can apply Lemma 4 with $p = x_1 - x_2^2/(2\omega)$, $a = c - x_2/\omega$, which yields some solutions τ_1 and τ_2, which are feasible solutions if $\tau_1 \in [c,d]$ and $\tau_2 \in [e,f]$. □

Lemma 6. *Consider full acceleration trajectories $x^+[(p,0),a,b]$ and $x^+[(q,0),c,d]$.*

Proof. Consider the equation

$$x^-[x^+[(p,0),a,b](\tau_1),\tau_1,\tau_2](\tau_2) = x^+[(q,0),c,d](\tau_2),$$

expanded to the system of equations

$$\begin{cases} p + \omega(\tau_1 - a)^2/2 + \omega(\tau_1 - a)(\tau_2 - \tau_1) - \omega(\tau_2 - \tau_1)^2/2 = q + \omega(\tau_2 - c)^2/2, \\ \omega(\tau_1 - a) + \omega(\tau_2 - \tau_1) = \omega(\tau_2 - c). \end{cases}$$

□

Lemma 7 **($x^+ \to x^+$).** *Consider partial trajectories $x^+[x,a,b]$ and $x^+[y,c,d]$.*

Proof. □

Let $D[t_0, t_1]$ denote the set of all state trajectories $x : [t_0, t_1] \to \mathbb{R}^2$ satisfying the double integrator dynamics $\dot{x}_1 = x_2, \dot{x}_2 = u$, the speed constraint $x_2 \in [0, 1]$ and control constraint $u \in [-\omega, \omega]$. Let $\bar{D}[t_0, t_1] \subset D[t_0, t_1]$ denote the set of trajectories x satisfying $x(t_1) = (0, 1)$. Let $\widetilde{D}[t_0, t_1] \subset \bar{D}[t_0, t_1]$ denote those that further satisfy $x(t_0) = (p_0, 1)$. Furthermore, for each of the previous three sets, we use the notation $D_1[t_0, t_1]$ to refer to the set of corresponding trajectories of the position component x_1 only.

Lemma 8. *Let $g, h \in D_1[a, b]$ such that $h''(t) = -\omega$ for all t. If $g(a) \geq h(a)$ and $g'(a) > h'(a)$, then $g(t) > h(t)$ for $t \in (a, b]$.*

Proof. For any $t \in [a, b]$, we have $g'(t) \geq g'(a) - \omega(t - a) > h'(a) - \omega(t - a) = h'(t)$. $\square$

Lemma 9. *Let $g, h \in D_1[a, b]$ such that $h''(t) = \omega$ for all t. If $g(a) \leq h(a)$ and $g'(a) < h'(a)$, then $g(t) < h(t)$ for $t \in (a, b]$.*

Proof. For any $t \in [a, b]$, we have $g'(t) \leq g'(a) + \omega(t - a) < h'(a) + \omega(t - a) = h'(t)$. $\square$

Lemma 10. *Consider $g, h \in D_1[a, b]$ such that $h''(t) = -\omega$ for all t. Suppose there is some touching time $c \in (a, b]$ such that $g(c) = h(c)$, then we have*

$$g(a) < h(a) \implies g'(c) > h'(c).$$

Proof. Let $g(a) < h(a)$ and consider the smallest time $c \in (a, b]$ such that $g(c) = h(c)$. Suppose that $g'(c) = h'(c)$. Let $f := h - g$, so we have $f(t) > 0$ for $t \in [a, c)$. Using the Taylor expansion of f around c with $\tau > 0$, we derive

$$\begin{aligned}
0 < f(c - \tau) &= f(c) - f'(c)\tau + f''(c)\tau^2/2 + o(\tau^2) \\
&= (h''(c) - g''(c))\tau^2/2 + o(\tau^2) \\
&= (-\omega - g''(c))\tau^2/2 + o(\tau^2).
\end{aligned}$$

Now dividing by τ^2 and taking the limit $\tau \downarrow 0$ gives $g''(c) < -\omega$, which contradicts $g \in D_1[a, b]$. Hence, we have $g'(c) \neq h'(c)$. Observe that $g'(c) < h'(c)$ would mean that $g(c - \epsilon) \geq h(c - \epsilon)$ for some $\epsilon > 0$, contradicting our assumption that c was smallest. Therefore, we have shown that $g'(c) > h'(c)$. Hence, it follows from Lemma 8 that there can only be a single touching time c, which completes the proof. $\square$

Lemma 11. *Consider $g, h \in D_1[a, b]$ such that $h''(t) = \omega$ for all t. Suppose there is some touching time $c \in (a, b]$ such that $g(c) = h(c)$, then we have*

$$g(a) > h(a) \implies g'(c) < h'(c).$$

Proof. Let $g(a) > h(a)$ and consider the smallest time $c \in (a, b]$ such that $g(c) = h(c)$. Suppose that $g'(c) = h'(c)$. Let $f := h - g$, so we have $f(t) < 0$ for $t \in [a, c)$. Using the Taylor expansion of f around c with $\tau > 0$, we derive

$$\begin{aligned}
0 > f(c - \tau) &= f(c) - f'(c)\tau + f''(c)\tau^2/2 + o(\tau^2) \\
&= (h''(c) - g''(c))\tau^2/2 + o(\tau^2) \\
&= (\omega - g''(c))\tau^2/2 + o(\tau^2).
\end{aligned}$$

Now dividing by τ^2 and taking the limit $\tau \downarrow 0$ gives $g''(c) > \omega$, which contradicts $g \in D_1[a, b]$. Hence, we have $g'(c) \neq h'(c)$. Observe that $g'(c) > h'(c)$ would mean that $g(c - \epsilon) \leq h(c - \epsilon)$ for some $\epsilon > 0$, contradicting our assumption that c was smallest. Therefore, we have shown that $g'(c) < h'(c)$. Hence, it follows from Lemma 9 that there can only be a single touching time c, which completes the proof. $\square$

We first analyze the special case when the crossing time of the lead boundary coincides with the scheduled crossing time of the current vehicle.

Lemma 12 (Entry). *Let $\bar{x} \in \bar{D}[t_0, t_1]$ be some lead boundary with control function $\bar{u}$ with alternating intervals $\bar{F}_i, \bar{D}_i, \bar{S}_i, \bar{A}_i$. Assume that problem*

$$
\begin{aligned}
x^* := \arg\max_x \int_{t_0}^{t_1} & x_1(t)dt \\
& x \in \widetilde{D}[t_0, t_1], \\
& x_1(t) \leq \bar{x}_1(t) \text{ for all } t \in [t_0, t_1],
\end{aligned}
$$

is feasible, then an optimal trajectory x^* is determined by the control (t_{d1} and τ_1 in proof)

$$
u(t) = \begin{cases} 0 & \text{for } t \in (0, t_{d1}), \\ -\omega & \text{for } t \in (t_{d1}, \tau_1), \\ \bar{u}(t) & \text{for } t \in (\tau_1, t_1). \end{cases}
$$

Proof.

- Assume feasibility (which means that $\underline{x}$ exists, analyze the exact conditions somewhere separately). Hence $t_0 \leq t_1 + p_0$ (otherwise $\bar{x}(t_1)$ would not be reachable). The case $t_0 = t_1 + p_0$ is trivial, because the optimal trajectory is to simply continue at full speed. When $t_0 < t_1 + p_0$, we need at least one interval of deceleration.

- Let τ_1 denote the earliest possible entry time, which is the smallest time t such that $x(t) = \bar{x}(t)$. This time is obviously well-defined, because $x(t_1) = \bar{x}(t_1)$ by assumption. Next, we do a case distinction on where τ_1 is on the lead vehicle boundary in terms of the type of alternating interval. For each case, we construct a feasible trajectory x^* with one deceleration of length d until τ_1.

- Suppose $\tau_1 \in \bar{S}_i$ for some i, then $x_2^*(\tau_1) = \bar{x}_2(\tau_1) = 0$. This implies that $u^*(t) = -\omega$ for $t \in (\tau_1 - 1/\omega, \tau_1)$ and $u^*(t) = 0$ for $t \in (t_0, \tau_1 - 1/\omega)$. Hence, the position at the entry time needs to satisfy

  $$
  \begin{aligned}
  \bar{x}_1(\bar{t}_{si}) = x_1^*(\tau_1) &= p_0 + \tau_1 - 1/\omega - t_0 + p^-(1/\omega) \\
  &= p_0 + \tau_1 - t_0 - 1/2\omega,
  \end{aligned}
  $$

  from which we derive that

  $$
  \tau_1 = t_0 - p_0 + 1/2\omega + \bar{x}_1(\bar{t}_{si}).
  $$

  Furthermore, we have $t_{d1} = \tau_1 - 1/\omega$.

- Suppose $\tau_1 \in \bar{A}_i$ for some i, then $x_1^*(\tau_1) = \bar{x}_1(\tau_1) = \bar{x}_1(\bar{t}_{ai}) + p^+(\tau_1 - \bar{t}_{ai}; \bar{x}_2(\bar{t}_{ai}))$ and $x_2^*(\tau_1) = \bar{x}_2(\tau_1) = \bar{x}_2(\bar{t}_{ai}) + (\tau_1 - \bar{t}_{ai})\omega$. This means that the duration of the initial deceleration d must be given by

  $$
  d = \frac{1 - x_2^*(\tau_1)}{\omega} = \frac{1 - \bar{x}_2(\bar{t}_{ai})}{\omega} - \tau_1 + \bar{t}_{ai}.
  $$

  Observe that, starting from the initial position, the current vehicle drives at full speed for a duration of $\tau_1 - d - t_0$ before decelerating, so the position of the entry point is given by

  $$
  \begin{aligned}
  x_1^*(\tau_1) &= p_0 + \tau_1 - d - t_0 + p^-(d), \\
  &= p_0 + \tau_1 - t_0 - \omega d^2/2.
  \end{aligned}
  $$

We can find τ_1 by equating $x_1^*(\tau_1) = \bar{x}_1(\tau_1)$, which yields the quadratic equation

$$
\begin{aligned}
&p_0 + \tau_1 - t_0 - \omega d^2/2 = \bar{x}_1(\bar{t}_{ai}) + \bar{x}_2(\bar{t}_{ai})(\tau_1 - \bar{t}_{ai}) + \omega(\tau_1 - \bar{t}_{ai})^2/2, \\
&\implies \tau_1 = \bar{t}_{ai} - \frac{\bar{x}_2(\bar{t}_{ai})}{\omega} + \frac{1}{\omega} \\
&\qquad - \frac{1}{2}\sqrt{\frac{2}{\omega^2} + \frac{2(\bar{x}_2(\bar{t}_{ai}))^2}{\omega^2} + 4\left(\frac{p_0 - t_0 - \bar{x}_1(\bar{t}_{ai}) + \bar{t}_{ai}(1 + \bar{x}_2(\bar{t}_{ai}))}{\omega}\right)} \\
&t_{d1} = \tau_1 - d = 2\tau_1 - \bar{t}_{ai} + \frac{\bar{x}_2(\bar{t}_{ai}) - 1}{\omega}.
\end{aligned}
$$

- If $\tau_1 \in \bar{F}_i \cup \bar{D}_i$, then we show that $\tau_1 = t_{fi}$, so this case reduces to $\tau_1 \in \bar{A}_i$. Suppose there exists a trajectory $x \in \widetilde{D}[t_0, t_1]$ and let τ_1 denote its earliest entry time. First of all, if $\tau_1 \in \bar{F}_i$ for some i, then it is obvious that $\tau_1 = \bar{t}_{fi}$. Next, suppose $\tau_1 \in (\bar{t}_{di}, \bar{t}_{si}]$, so $x_1(\bar{t}_{di}) < \bar{x}_1(\bar{t}_{di})$, then applying Lemma 10 yields $x_2(\tau_1) > \bar{x}_2(\tau_1)$, contradicting the fact that τ_1 was an entry time.

- Show that x_1^* is an upperbound for any $x_1 \in \widetilde{D}_1[t_0, t_1]$. Let τ_0^* denote the start of the first deceleration of x^*. Let x be some feasible solution, then $x_1(t) \leq \bar{x}_1(t) = x_1^*(t)$ for $t \in [\tau_1, t_1]$. We show that $x_1(t) \leq x_1^*(t)$ on $t \in [t_0, \tau_1]$ as well. Let τ_0 denote the start of the first deceleration of x. Suppose $\tau_0 > \tau_0^*$, then we have $x_1(\tau_0) > x_1^*(\tau_0)$ and $x_2(\tau_0) = 1 > x_2^*(\tau_0) = 1 - \omega(\tau_0 - \tau_0^*)$, so by applying Lemma 8 for x^* and x on $[\tau_0, \tau_1]$, we have $x_1(\tau_1) > x_1^*(\tau_1) = \bar{x}_1(\tau_1)$, which violates the feasibility of x. Suppose $\tau_0 < \tau_0^*$, then we obviously must have $x_1(\tau_0^*) < x_1^*(\tau_0^*)$. Suppose that $c = \inf\{t \in (\tau_0^*, \tau_1) : x_1(t) > x_1^*(t)\}$ exists, i.e., x crosses x^* at c for the first time. Application of Lemma 10 gives $x_2(c) > x_2^*(c)$. But then it follows from Lemma 8 that $x_1(\tau_1) > x_1^*(\tau_1) = \bar{x}_1(\tau_1)$, which violates the feasibility of x.

- Now x^* is optimal, because $\int_{t_0}^{t_1} x_1^*(t)dt \geq \int_{t_0}^{t_1} x_1(t)dt$ for any $x \in \widetilde{D}_1[t_0, t_1]$. □

We now consider the general case in which the scheduled crossing time does not need to coincide with the end of the lead vehicle boundary. Invoke Lemma 12 to obtain $\bar{x}'$ and argue that it is a valid upperbound for any feasible trajectory, so we can use it instead of $\bar{x}$ without changing the problem.

Lemma 13 (Exit). *Let $\bar{x}' \in \widetilde{D}[t_0, t_1]$ be some lead boundary. Let $t_2 > t_1$, and assume that problem*

$$
\begin{aligned}
x^* = \arg\max_x \int_{t_0}^{t_1} & x_1(t)dt \\
& x \in \widetilde{D}[t_0, t_2], \\
& x_1(t) \leq \bar{x}_1'(t) \text{ for all } t \in [t_0, t_1],
\end{aligned}
$$

is feasible, then an optimal trajectory x^* is given by ...

Proof.

- We define the boundary $\bar{x}''$ by $\bar{x}''(t) = 1/2\omega$ for $t \leq t_1 - 1/\omega$ and $\bar{x}''(t) = x^+[(-1/2\omega, 0), t_1 - 1/\omega, t_1](t)$ for $t \in [t_1 - 1/\omega, t_1]$. We argue that $\bar{x}''$ is an upperbound for all feasible trajectories. Let $x \in \widetilde{D}[t_0, t_2]$ be any trajectory that has $x_1(t) > -1/2\omega$ for some $t \leq t_0 - 1/\omega$, so that $x_1(t_0 - 1/\omega) > -1/2\omega = \bar{x}_1''(t_0 - 1/\omega)$. Now because $x \in \widetilde{D}[t_0, t_1]$, we have $x(t_2) = \bar{x}''(t_2) = (0, 1)$, but Lemma 11 yields $x_2(t_2) < \bar{x}''(t_2) = 1$, a contradiction.

- Next, we construct a trajectory x^* by joining $\bar{x}'$ and $\bar{x}''$ using a deceleration part x^-. Argue that such x^* must always exist. Argue that x^* is unique.

- We argue that x^* is an upperbound for any feasible trajectory. First of all, $\bar{x}'$ is an upperbound, so $x_1(\tau_1) \leq \bar{x}'_1(\tau_1)$. Suppose there is some $c \in (\tau_1, \tau_2)$ such that x crosses x^-, so $x_1(c) = x_1^-(c)$ and $x_2(c) > x_2^-(c)$, then by Lemma 8, we have $x_1(\tau_2) > x_1^-(\tau_2) = \bar{x}''_1(\tau_2)$, which contradicts the fact that $\bar{x}''$ is an upperbound.

- Now x^* is optimal, because $\int_{t_0}^{t_2} x_1^*(t)dt \geq \int_{t_0}^{t_2} x_1(t)dt$ for any $x \in \widetilde{D}_1[t_0, t_2]$.

$\square$

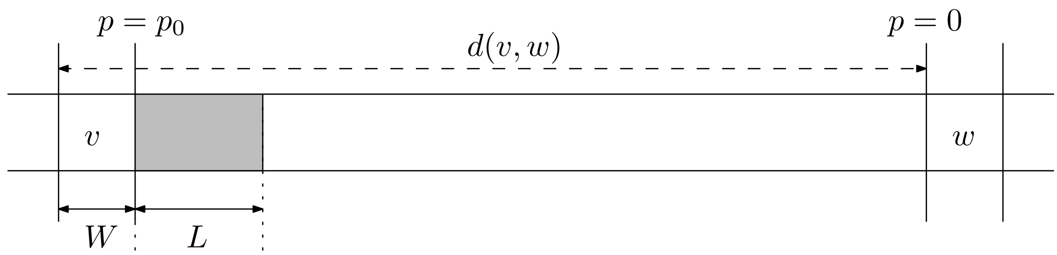

Figure 3: Tandem of two intersections v and w with lane of length $d(v, w)$. The grey rectangle represents some vehicle that just left intersection v. We assume that vehicles must drive at maximum speed as long as they occupy any intersection, so deceleration is only allowed from the shown position onwards.

1 Platoons in tandem of two intersections

When considering multiple vehicles driving between intersection, we must take into account the capacity of the lane segment between them, because the fact that only a limited number of vehicles can drive or wait at the same time on a lane between intersections may cause network-wide effects. The capacity of lanes between intersections is intimately related to the trajectories of vehicles, which we first want to understand better. We will be using an optimal control formulation with the objective that keeps the vehicles as close as possible to the next intersection at all times, which is similar to the MotionSynthesize problem considered in [1]. This problem can be solved using direct transcription, which works well enough if we just want to simulate the behavior of the system. However, we show that it is possible to explicitly formulate the optimal controller and explain how to compute trajectories without using time discretization.

1.1 Problem formulation

Before we turn to the general case of networks of intersection, we will first investigate the trajectories of vehicles in a tandem of two intersections as depicted in Figure 3. Let v denote the left intersection and w the right intersection and assume that vehicles drive from left to right. Furthermore, we will call the road segment strictly between both intersection areas the *lane*. In the following discussion, let $p(t)$ and $v(t)$ denote the position and velocity of the vehicle, respectively, and we call $x(t) = (p(t), v(t))$ its *state*. Let the length and width of a vehicle i be denoted by L_i and W_i, respectively. We measure the position of a vehicle at the front bumper. We fix position $p = 0$ at the stop line of intersection w. We make the following additional assumptions about our lane model.

Assumption 1. *Vehicles are represented as rectangles, all having the same length $L_i = L$ and width $W_i = W$. Lanes are axis-aligned and have width W, such that when lanes intersect, the intersection area is a square. Vehicles are not able to overtake other vehicle.*

Assumption 2. *All vehicles satisfy the same double integrator dynamics $\dot{p} = v$, $\ddot{p} = u$ with velocity bounds $0 \leq v \leq v_{\max}$ and symmetric control bounds $-\omega \leq u \leq \omega$. We assume that the maximum speed is $v_{\max} = 1$, which is without loss of generality, because we can always achieve this by appropriate scaling of positions and the acceleration bound ω.*

Assumption 3. *Vehicles drive at full speed when entering an intersection and keep driving at full speed as long as they occupy an intersection.*

Now assume that some vehicle is scheduled to start entering the lane at time $t_0 < 0$ and has to start entering w at time $t = 0$. Let $p_0 = -d(v, w) + W$ denote the position from where the vehicle starts to enter the lane. We will refer to the vehicle driving in front of the current vehicle as the *lead vehicle*. Let $\bar{p}(t)$ denote the position of rear bumper of the lead vehicle, assuming there is one. To avoid collision with this vehicle, the current vehicle must satisfy the *lead* constraint $p(t) \leq \bar{p}(t)$ at all times. We try to keep the vehicle as close to w as possible at all times. This yields the optimal control problem

$$\begin{aligned} \max_u \quad & \int_{t_0}^{0} p(t)dt \\ \text{s.t.} \quad & \dot{p} = v, \ \ddot{p} = u, \\ & p(t_0) = p_0, \quad p(0) = 0, \\ & v(t_0) = 1, \quad\ \ v(0) = 1, \\ & -\omega \leq u(t) \leq \omega, \\ & \ \ 0 \leq v(t) \leq 1, \\ & \ \ p(t) \leq \ \bar{p}(t). \end{aligned} \tag{1}$$

It is straightforward to solve problem (1) by using direct transcription. Maybe show some example solutions at this point to further motivate the definitions in the next paragraph?

To support the upcoming discussion, we define the following auxiliary functions to describe parts of optimal trajectories where $u(t) \in \{-\omega, \omega\}$. Without any additional constraints, the vehicle dynamics imply that the time it takes to fully accelerate from rest to maximum velocity is given by $d_f := 1/\omega$. Similarly, it also takes d_f time to fully decelerate from maximum velocity to rest. The corresponding full acceleration and deceleration trajectories are given, respectively, by

$$\begin{aligned} p^+(t) &:= \omega t^2/2 && \text{for } 0 \leq t \leq d_f, \\ p^-(t) &:= t - \omega t^2/2 && \text{for } 0 \leq t \leq d_f. \end{aligned}$$

Due to the symmetry of the control bounds, p^- can also be expressed as

$$p^-(t) = p^+(d_f) - p^+(d_f - t) \qquad \text{for } 0 \leq t \leq d_f.$$

Now suppose that we start accelerating with initial velocity $0 \leq v_0 \leq 1$ and accelerate for τ time, which must satisfy $\tau \leq (1 - v_0)/\omega$, then the distance traveled is given by

$$\begin{aligned} p^+(\tau; v_0) &:= p^+(v_0/\omega + \tau) - p^+(v_0/\omega) \\ &= v_0\tau + \omega\tau^2/2. \end{aligned}$$

Similarly, for deceleration with initial velocity, we have again from symmetry

$$\begin{aligned} p^-(\tau; v_0) &:= p^+(\tau; v_0 - \omega\tau) \\ &= v_0\tau - \omega\tau^2/2. \end{aligned}$$

1.2 Optimal control

The aim of this section is to provide an explicit parameterization of optimal solutions of problem (1). Before we proceed, we first rewrite it to the following standard form of optimal control problems

$$
\begin{aligned}
\max \quad & \int_{t=t_0}^{t_f} F(x(t), u(t), t)dt \\
\text{s.t.} \quad & \dot{x}(t) = f(x(t), u(t), t), \quad x(t_0) = x_0, \\
& a(x(t_f), t_f) \geq 0, \\
& b(x(t_f), t_f) = 0, \\
& g(x(t), u(t), t) \geq 0, \\
& h(x(t), t) \geq 0,
\end{aligned} \tag{2}
$$

where x is the vector of state variables and u is the control input. Here, the constraints g are called mixed state constraints, because they involve both state and control variables, while constraints h are called pure state constraints. In this subsection, we will write the state as $x = (x_1, x_2)$ with position x_1 and velocity x_2. The control function u again corresponds to acceleration. The initial state is (p_0, v_0) and the target state is $(0, 1)$. We write v_0, because we will be considering optimal trajectories with $v_0 < 1$ as an intermediate step of the analysis. Let $\bar{x}$ denote the state trajectory of the rear bumper of the lead vehicle. With this new notation, optimal control problem (1) is equivalent to (2) for $v_0 = 1$ and setting[1]

$$
\text{(2a)} \left\{ \begin{aligned}
t_f &= 0, \\
F(x, u, t) &= x_1, \\
f(x, u, t) &= (x_2, u), \\
x_0 &= (p_0, v_0), \\
b(x, t) &= (x_1, x_2 - 1), \\
g(x, u, t) &= \omega^2 - u^2, \\
h_1(x, t) &= x_2 - x_2^2, \\
h_2(x, t) &= \bar{x}_1(t) - x_1.
\end{aligned} \right\} \text{(2b)}
$$

We refer to problem (2a) as the *single vehicle variant*, because ignoring the lead constraint is equivalent to considering a single vehicle in the system. For both problem variants, we will denote an instance using the tuple $z = (t_0, p_0, v_0)$.

In the upcoming analysis, we will encounter control functions that switch between no acceleration, full deceleration and full acceleration, to which we might refer as *bang-off-bang* control. Furthermore, we will see that optimal solutions have a particularly simple structure, illustrated in Figure 4, to which we refer as *alternating control*. The following definition makes this notion precise and introduces the notation that we use to completely characterize optimal trajectories.

Definition 1. *Let $\{x(\cdot), u(\cdot)\}$ be a feasible solution pair for problem (2). Suppose there exists a partition of the control time interval $[t_0, t_f]$, denoted by*

$$t_0 = t_{f1} \leq t_{d1} \leq t_{s1} \leq t_{a1} \leq t_{f2} \leq t_{d2} \leq t_{s2} \leq t_{a2} \leq \cdots \leq t_{f,n+1} = t_f,$$

such that we have the following consecutive intervals

$$
\begin{aligned}
F_i &:= [t_{f,i}, t_{d,i}] \quad \textit{(full speed)}, & S_i &:= [t_{s,i}, t_{a,i}] \quad \textit{(stopped)}, \\
D_i &:= [t_{d,i}, t_{s,i}] \quad \textit{(deceleration)}, & A_i &:= [t_{a,i}, t_{f,i+1}] \quad \textit{(acceleration)},
\end{aligned}
$$

[1] Instead of using quadratic expressions, constraints g and h_1 could have each been written as two linear constraints, e.g., we could have chosen $g(x, u, t) = (\omega - u, \omega + u)$. However, this would violate the constraint qualification conditions of the theorem that we use in Section 1.2.1.

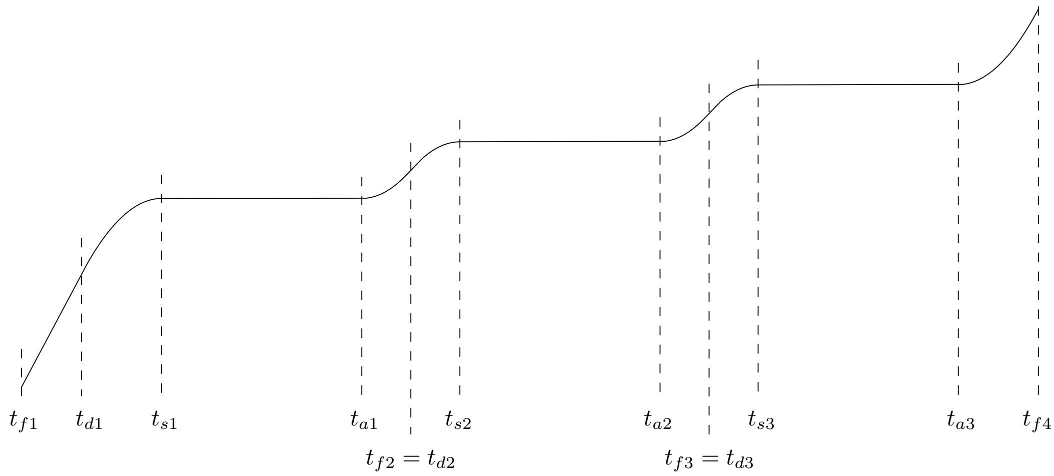

Figure 4: Some example of an alternating vehicle trajectory x. The particular shape of this trajectory is due to two preceding vehicles, which causes the two "bumps" at the times where these vehicles exit the lane. The trajectory of the current lead vehicle and previous lead vehicles that gave rise to this particular shape are not shown for clarity.

then $\{x(\cdot), u(\cdot)\}$ *is called* alternating on $[t_0, t_f]$ *if the control function* u *satisfies*

$$u(t) = \begin{cases} -\omega & \textit{when } t \in \cup_i D_i, \\ \omega & \textit{when } t \in \cup_i A_i, \\ 0 & \textit{otherwise,} \end{cases}$$

such that the state trajectory x *satisfies* $0 \leq x_2 \leq 1$ *and*

$$\begin{aligned} x_2(t) &= 1, \quad \textit{for all} \;\; t \in \cup_i F_i, \\ x_2(t) &= 0, \quad \textit{for all} \;\; t \in \cup_i S_i. \end{aligned}$$

Note that alternating trajectory is a stronger notion than bang-off-bang control, as we explicitly require two consecutive bangs to have opposite signs.

Theorem 1. *Consider optimal control problem (2a) with* $v_0 = 1$ *and lead vehicle trajectory* $\bar{x}(\cdot)$ *such that* $\{\bar{x}(\cdot), \bar{u}(\cdot)\}$ *is alternating on* $[\bar{t}_0, \bar{t}_f]$*, for some control function* $\bar{u}$ *and times* $\bar{t}_0 \leq t_0$ *and* $\bar{t}_f \leq t_f$*, then any optimal pair* $\{x^*(\cdot), u^*(\cdot)\}$ *must be alternating.*

The proof of this result will also show how optimal solutions can be computed. We will first consider situations in which the lead constraint h_2 is inactive, such that it can essentially be ignored. Next, we generalize this resulting lemma slightly to also allow non-maximal initial velocities $v_0 < 1$. This lemma is then used characterize the optimal trajectories of the original problem with constraint h_2.

Since the system dynamics are time-homogeneous, i.e., they do not depend on time, we can apply Bellman's optimality principle to show optimality of subproblems as stated in the following result.

Lemma 14. *Let* $x^* : [t_0, 0] \to (-\infty, 0]$ *be an optimal trajectory for problem (1). For any time* $\tau \in [t_0, 0]$*, consider the subproblem with* $t_0' = t_0 + \tau$ *and* $p_0' = p_0 + x^*(\tau)$*, then Bellman's principle of optimality implies that the optimal trajectory is now simply given by the truncation of* x^* *to* $[\tau, 0]$*, which we will denote by* $x^*|_{[\tau,0]}$*.*

1.2.1 Inactive lead constraint

We now consider the case when the lead constraint h_2 can be ignored, which might occur when the preceding vehicle is sufficiently far away, or the current vehicle is the only vehicle in the system. In this case, we consider the problem (2a). We prove Theorem 1 in this special case as Lemma 15 below. When dealing with optimal control problems, Pontryagin's Maximum Principle (PMP) provides a family of necessary conditions for optimality. We are going to apply such a PMP-style necessary condition that deals with mixed and pure state constraints to characterize the optimal control. A common approach to deal with pure state inequality constraints h is to introduce a multiplier η and append ηh to the Hamiltonian, which is known as the direct adjoining approach. Instead, we use Theorem 5.1 from [2] (see also equations (4.29) in [3]), which is a so-called indirect adjoining approach, because ηh^1 is appended to the Hamiltonian instead, where h^1 is defined as

$$h^1 = \frac{dh}{dt} = \frac{\partial h}{\partial x} f + \frac{\partial h}{\partial t}.$$

For our current problem, the first derivative of pure state constraint h_1 with respect to time t is given by

$$h_1^1(x, u, t) = u - 2x_2 u,$$

which shows that h_1 is first-order, because the control u appears after differentiating once.

Lemma 15. *Consider problem (2a) with $v_0 = 1$. Let the acceleration/deceleration duration and the start of the deceleration period, respectively, be given by*

$$\begin{aligned} d &= \min\left\{1/\omega, \sqrt{(p_0 - t_0)/\omega}\right\}, \\ t_d &= \omega d^2 - 2d + t_0 - p_0. \end{aligned}$$

The problem is feasible if and only if $t_0 \leq p_0$ and $t_d \geq 0$ and in that case the optimal control function is given by

$$u(t) = \begin{cases} -\omega & \text{for } t_d < t < t_d + d, \\ \omega & \text{for } -d < t < 0, \\ 0 & \text{otherwise.} \end{cases}$$

Proof. The initial velocity is $x_2(t_0) = 1$, so without any further deceleration, the earliest possible time of arrival is $t_0 - p_0$, so there is no feasible control whenever $t_0 > p_0$, showing the necessity of the first feasibility condition. Whenever we have $t_0 = p_0$, the optimal controller is simply $u(t) = 0$. In the rest of the proof, we assume that $t_0 < p_0$. In that case, there needs to be at least one period of deceleration and thus also at least one period of acceleration to satisfy the boundary condition $x_2(t_0) = x_2(0) = 1$.

By adjoining $\mu g + \eta h_1^1$ to the Hamiltonian $H(x, u, \lambda, t) = F(x, u, t) + \lambda f(x, u, t)$ we obtain the Lagrangian in so-called Pontryagin form

$$\begin{aligned} L(x, u, \lambda, \mu, \eta, t) &= H(x, u, \lambda, t) + \mu g(x, u, t) + \eta h_1^1(x, u, t) \\ &= x_1 + \lambda_1 x_2 + \lambda_2 u + \mu(\omega^2 - u^2) + \eta(u - 2x_2 u). \end{aligned}$$

Let $\{x^*(t), u^*(t)\}$ denote an optimal pair for problem (2a), which must satisfy the necessary conditions in Theorem 5.1 from [2]. The Hamiltonian maximizing condition requires

$$\lambda_2 u^*(t) \geq \lambda_2 u \quad \text{at each } t \in [t_0, 0],$$

for all $-\omega \leq u \leq \omega$ such that $u - 2x_2^* u \geq 0$ whenever $x_2^* = 0$ or $x_2^* = 1$. Hence, the optimal controller must satisfy

$$u^*(t) = \begin{cases} -\omega & \text{when } \lambda_2(t) < 0, \; x_2^*(t) > 0, \\ \omega & \text{when } \lambda_2(t) > 0, \; x_2^*(t) < 1, \\ 0 & \text{otherwise.} \end{cases} \tag{3}$$

Observe that the Lagrange multipliers $\mu(t)$ and $\eta(t)$ must satisfy the complementary slackness conditions

$$\begin{aligned} \mu(\omega^2 - (u^*)^2) &= 0, \quad \mu \geq 0, \\ \eta(x_2 - x_2^2) &= 0, \quad \eta \geq 0 \text{ and } \dot{\eta} \leq 0. \end{aligned}$$

Whenever $u^* \neq 0$ on any open interval, then it is clear that we must have $0 < x_2 < 1$ on that interval, so the second complementary slackness condition requires that we have $\eta = 0$. Hence, we have $\eta u^* = 0$ almost everywhere. Next, we investigate the costate trajectory λ, which must satisfy the adjoint equations

$$\dot{\lambda} = -L_x(x^*, u^*, \lambda, \mu, \eta, t) \iff \begin{cases} \dot{\lambda}_1 = -1, \\ \dot{\lambda}_2 = -\lambda_1 + 2\eta u^* = -\lambda_1, \end{cases}$$

so we have $\dot{\lambda}_2(t) = t + c_1$ and $\lambda_2(t) = \frac{1}{2}t^2 + c_1 t + c_2$ for some constants c_1, c_2. In view of (3) and the initial condition $x_2(t_0) = 1$, this shows that $u^*(t)$ has at most one period of deceleration and at most one period of acceleration. As we argued above, we need at least one period of deceleration and acceleration, so $\lambda_2(t)$ must have exactly two zeros. Let t_d and t_a denote the start of the deceleration and acceleration periods, respectively. We show that these are uniquely determined by the necessary optimality conditions when the problem is feasible, fixing a unique optimal control function $u^*(t)$. Let $d \geq 0$ denote the duration of the deceleration, which is also the duration of the acceleration, because of the boundary conditions $x_2(t_0) = x_2(0) = 1$. Observe that $d \leq 1/\omega$ to satisfy velocity constraint h_1.

First, we show that t_a is uniquely determined in terms of d, using the necessary condition

$$\frac{\partial L}{\partial u}\Big|_{u=u^*(t)} = 0 \iff \lambda_2 - 2\mu u^* - 2\eta x_2^* + \eta = 0. \tag{4}$$

Suppose $t_a + d < 0$, which means $d = 1/\omega$, then for $t \in (t_a + 1/\omega, 0)$, we have $x_2^*(t) = 1$ and $u^*(t) = 0$, so the complementary slackness conditions require $\mu(t) = 0$, such that the necessary condition gives $\lambda_2(t) = \eta(t)$. But this would mean that we have $\dot{\eta}(t) = \dot{\lambda}_2(t) > 0$, which contradicts the necessary condition $\dot{\eta}(t) \leq 0$, hence $t_a = -d$.

Next, we show that t_d is uniquely defined in terms of d. Because the vehicle drives at full velocity initially, we have $x_1^*(t_d) = p_0 + t_d$. During deceleration, the vehicle moves exactly $p^-(d; 1) = p^-(d)$, so we have

$$\begin{aligned} x_1^*(t_d + d) &= x_1^*(t_d) + p^-(d) \\ &= p_0 + t_d - t_0 + d - \omega d^2/2. \end{aligned}$$

The vehicle is stationary in between $t_d + d$ and t_a, so we have $x^*(t_a) = x^*(t_d + d)$. Observe that the vehicle moves the same distance $p^+(d; 1 - \omega d) = p^-(d)$ during acceleration due to the symmetric control bounds, which yields

$$\begin{aligned} 0 = x_1^*(0) &= x_1^*(t_a) + p^-(d) \\ &= p_0 + t_d - t_0 + 2p^-(d) \\ &= p_0 + t_d - t_0 + 2d - \omega d^2. \end{aligned}$$

Hence, we have $t_d = \omega d^2 - 2d + t_0 - p_0$.

Finally, we show that d is uniquely determined whenever the problem is feasible. Observe that d determines a feasible control if and only if $0 \leq t_d$ and $t_d + d \leq t_a$, where the last inequality must be equality whenever $d < 1/\omega$ in order to satisfy (3). Using the definitions of t_d and t_a, inequality $t_d + d \leq t_a$ can be rewritten to

$$d^2 \leq (p_0 - t_0)/\omega. \tag{5}$$

Suppose the problem has $1/\omega^2 \leq (p_0 - t_0)/\omega$, and suppose $d < 1/\omega$, then $d^2 < 1/\omega^2 \leq (p_0 - t_0)/\omega$, which is not allowed by (3), so $d = 1/\omega$. Otherwise, when $1/\omega^2 > (p_0 - t_0)/\omega$, then we must have $d < 1/\omega$ and equality in (5), which yields $d = \sqrt{(p_0 - t_0)/\omega}$. Observe that $1/\omega^2 \leq (p_0 - t_0)/\omega$ is equivalent to $1/\omega \leq \sqrt{(p_0 - t_0)/\omega}$, which allows us to conveniently write these two solutions as a single expression $d = \min\{1/\omega, \sqrt{(p_0 - t_0)/\omega}\}$. □

Remark 1. The feasibility constraint $t_0 \leq p_0$ gives a lower bound on the amount of time in which the vehicle can travel across the lane to the next intersection. Therefore, we will refer to it as the *travel constraint.*

Lemma 16. *Consider problem (2a) with $v_0 = 0$. Define the candidate durations*

$$\begin{aligned}
a^{(1)} &= \sqrt{-p_0/\omega - 1/2\omega^2}, \\
a_1^{(2)} &= \frac{p_0 + 1/2\omega}{1 + t_0\omega} - \frac{1 + t_0\omega}{4\omega}, \\
d^{(2)} &= -t_0/2 - 1/2\omega, \\
a_2^{(2)} &= -a_1^{(2)} + 1/2\omega - t_0/2.
\end{aligned}$$

Furthermore, define the candidate subproblem $z_1 := (t_1, p_1, 1) := (t_0 + 1/\omega, p_0 + 1/2\omega, 1)$ and let $d(z_1)$ and $t_d(z_1)$ denote the duration and start, respectively, of the deceleration period as given by Lemma 15 for subproblem z_1. Problem (2a) is feasible if and only if $t_0 \leq -1/\omega$ and $p_0 \leq -1/2\omega$ and one of the cases holds in the definition

$$(a_1, t_d, d, a_2) := \begin{cases} (1/\omega, t_1 + t_d(z_1), d(z_1), d(z_1)) & \text{if } t_1 \leq p_1 \text{ and } t_d(z_1) \geq 0, \\ (a^{(1)}, t_0 + a^{(1)}, a^{(1)}, 1/\omega) & \text{if } a^{(1)} \in [0, 1/\omega] \text{ and } a^{(1)} < d^{(2)}, \\ (a_1^{(2)}, t_0 + a_1^{(2)}, d^{(2)}, a_2^{(2)}) & \text{if } a_1^{(2)}, d^{(2)}, a_2^{(2)} \in [0, 1/\omega], \end{cases}$$

and in that case, the optimal control function is given by

$$u(t) = \begin{cases} \omega & \text{for } t_0 < t < t_0 + a_1, \\ -\omega & \text{for } t_d < t < t_d + d, \\ \omega & \text{for } -a_2 < t < 0, \\ 0 & \text{otherwise.} \end{cases}$$

Proof. The necessary optimality conditions of the current problem lead to the same Hamiltonian maximization condition and adjoint equations as in Lemma 15. Since $v_0 = 0$, there must be at least one acceleration period, followed by an optional deceleration and an optional acceleration. Let a_1, d, a_2 denote the durations of these three periods, respectively. From (3) follows that the first acceleration happens on $(t_0, t_0 + a_1)$. From (4) follows that the second acceleration must happen on $(-a_2, 0)$, by the same argument as in Lemma 15. Let t_d denote the start of the deceleration period. Furthermore, let $s_1 := t_d - t_0 - a_1$ and $s_2 := -a_2 - t_d - d$ denote the times between the two consecutive pairs of period, such that $t_0 + a_1 + s_1 + d + s_2 + a_2 = 0$. Suppose that $t_0 > -1/\omega$, then any trajectory satisfies $x_2(0) < 1$, which is infeasible, so the first feasibility condition is necessary. Observe that any trajectory must accelerate for a duration of at least $1/\omega$, which causes a displacement of at least $p^+(1/\omega) = 1/2\omega$. Hence, a feasible trajectory only exists if $p_0 \leq -1/2\omega$.

Full first acceleration. Suppose there exists a trajectory x with $a_1 = 1/\omega$. In this case, we can leverage Lemma 15 as follows. Observe that $x_2(t_1) = 1$, so due to Lemma 14, the truncated trajectory $x|_{[t_1,0]}$ is a solution to subproblem $z_1 = (t_1, x_1(t_1), 1)$. Therefore, z_1 must satisfy the two feasibility conditions of Lemma 15. Furthermore, we must have $a_2 = d$.

Next, we show that when a trajectory x exists with $a_1 = 1/\omega$, then there cannot exist another feasible trajectory with $a_1 < 1/\omega$ that satisfies the necessary optimality conditions, so x must be optimal. Assume there exists an optimal trajectory x' with $a_1' < 1/\omega$, then

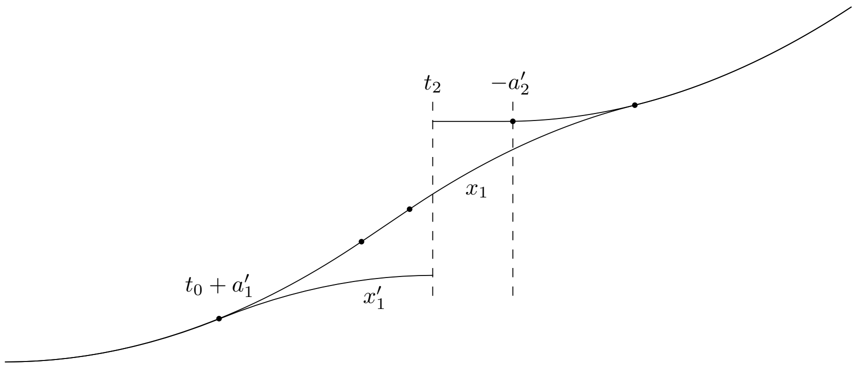

Figure 5: Illustration of contradiction (7) for proving that a trajectory x' with $a_1' < 1/\omega$ does not exist when a trajectory x exists with full first acceleration $a_1 = 1/\omega$. The dots mark the control switching times for both trajectories. Hence, the left-most and right-most dot correspond to an inflection point of x' and x, respectively.

$s_1' = 0$ and $d' \leq a_2'$. First, suppose x is such that $d = 1/\omega$, then the total travel distance satisfies

$$\begin{aligned} x_1'(0) &= p_0 + p^+(a_1') + p^-(d') + p^+(a_2') \\ &< p_0 + p^+(1/\omega) + s_1 + p^-(1/\omega) + p^+(1/\omega) = x_1(0) = 0, \end{aligned}$$

which shows that x' is infeasible. Otherwise, x is such that $d < 1/\omega$, so $s_2 = 0$. First, we show that $a_2' > d$, which holds trivially when $s_2' > 0$, because then we have $a_2' = 1/\omega$. When $s_2' = 0$, we derive it using

$$\begin{aligned} t_0 + a_1' + d' + a_2' &= t_0 + a_1 + s_1 + 2d = 0, \\ \implies 2a_2' &\geq d' + a_2' > s_1 + 2d \geq 2d. \end{aligned} \tag{6}$$

It also follows from (6) that $t_2 := t_0 + a_1' + d' < t_0 + a_1 + d$. Because $a_1' < a_1$, we have $x_1'(t_2) < x_1(t_2)$. Furthermore, we have $x_1'(-a_2') > x_1(-a_2')$. But this leads to the contradiction

$$x_1'(t_2) < x_1(t_2) \leq x_1(-a_2') < x_1'(-a_2') = x_1'(t_2). \tag{7}$$

Partial first acceleration. Suppose there exists a trajectory x with $a_1 < 1/\omega$, then $s_1 = 0$ and $t_d = t_0 + a_1$. First, suppose $s_2 > 0$, then $d = a_1 \in [0, 1/\omega]$ and $a_2 = 1/\omega$, so this happens if only if $t_0 + 2a_1 + 1/\omega < 0$, which is equivalent to $a_1 < d^{(2)}$. Due to the symmetric control bounds, we have $p^+(a_1) = p^-(a_1; \omega a_1)$ and we can find a_1 using the total travel distance equation

$$\begin{aligned} 0 = x_1(0) &= p_0 + 2p^+(a_1) + p^+(1/\omega) \\ &\implies a_1 = \sqrt{-p_0/\omega - 1/2\omega^2}. \end{aligned}$$

Next, suppose $s_2 = 0$, then $a_1, d, a_2 \in [0, 1/\omega]$ can be determined by solving a system of three equations. We have the total time equation $t_0 + a_1 + d + a_2 = 0$ and the total acceleration equation $a_1 - d + a_2 = 1/\omega$, from which we can derive

$$\begin{aligned} a_1 + a_2 &= d + 1/\omega = -t_0 - d \\ &\implies d = -1/2\omega - t_0/2 \\ &\implies a_2 = -a_1 + 1/2\omega - t_0/2. \end{aligned}$$

We can find a_1 from the total travel distance equation

$$\begin{aligned} 0 = x_1(0) &= p_0 + p^+(a_1) + p^-(d; \omega a_1) + p^+(a_2; \omega(a_1 - d)) \\ &\implies a_1 = \frac{p_0 + 1/2\omega}{1 + t_0\omega} - \frac{1 + t_0\omega}{4\omega}. \end{aligned}$$

Finally, we show that when a trajectory x exists such that $a_1 < 1/\omega$ and $s_2 > 0$, we show that there cannot exist another optimal trajectory x' with $s_2' = 0$. Suppose such x' exists, then it satisfies $a_1' < 1/\omega$, as shown in the previous section of the proof. Therefore, we have $(a_1', d', a_2') = (a_1^{(2)}, d^{(2)}, a_2^{(2)})$, so feasibility of x' implies $a_2^{(2)} = a_1' \leq 1/\omega$. It is straightforward to verify that this inequality can be rewritten to

$$a_2^{(2)} = d^{(2)}/2 - (a^{(1)})^2/2d^{(2)} + 1/\omega \leq 1/\omega,$$

from which we derive that $a^{(1)} \geq d^{(2)}$. However, this contradicts $a^{(1)} = a_1 < d^{(2)}$, which must hold because we assumed that x was such that $s_2 > 0$. $\square$

1.2.2 Active lead constraint

We now consider problem (2b) when lead constraint h_2 becomes active somewhere along the optimal trajectory. Throughout the following discussion, let $\bar{x}$ denote the trajectory of the lead vehicle's rear bumper and assume that there is some control function $\bar{u}$ such that $\{\bar{x}(\cdot), \bar{u}(\cdot)\}$ is alternating. Let the corresponding switching times be denoted as $\bar{t}_{\cdot i}$. The lead constraint is of second order, because the control function u only appears after differentiating twice, as seen in

$$\begin{aligned} h_2(x, t) &= \bar{x}_1 - x_1, \\ h_2^1(x, t) &= \bar{x}_2 - x_2, \\ h_2^2(x, u, t) &= \bar{u} - u. \end{aligned}$$

Proof sketch.

- The costate trajectory component λ_2 is exactly like in Lemma 15 when we assume there are no discontinuities, so this would violate h_2, so we must have at least one discontinuity, which can only happen at h_2 entry time.
- This discontinuity must take λ_2 from negative to positive. This can be used to show that there is only one possible location for the discontinuity to happen.
- First deceleration must start such that h_2 is hit. We derive this *entry time* τ_1, which is later used to formulate our so-called *buffer constraints* on the crossing times.
- Now given the "shape" of the trajectory, we can formulate a system of equations to determine the lengths of acceleration/deceleration at the end of the trajectory, after the *exit time*.

Proof of Theorem 1. The Hamiltonian is still the same as in Lemma 15, but we need to add a new term to the Lagrangrian, which is now

$$\begin{aligned} L(x, u, \lambda, \mu, \eta_1, \eta_2, t) &= H(x, u, \lambda, t) + \mu g(x, u, t) + \eta_1 h_1^1(x, u, t) + \eta_2 h_2^2(x, u, t) \\ &= x_1 + \lambda_1 x_2 + \lambda_2 u + \mu(\omega^2 - u^2) + \eta_1(u - 2x_2 u) + \eta_2(\bar{u} - u). \end{aligned}$$

Due to the Hamiltonian maximization principle, the optimal control function u^* must satisfy the following form

$$u^*(t) = \begin{cases} 0 & \text{when } \lambda_2(t) > 0,\ x_2^*(t) = 1, \\ 0 & \text{when } \lambda_2(t) < 0,\ x_2^*(t) = 0, \\ -\omega & \text{when } \lambda_2(t) < 0,\ x_2^*(t) > 0, \\ \omega & \text{when } \lambda_2(t) > 0,\ x_2^*(t) < 1,\ x_1 < \bar{x}_1, \\ \bar{u}(t) & \text{when } \lambda_2(t) > 0,\ x_2^*(t) < 1,\ x_1 = \bar{x}_1. \end{cases}$$

The adjoint equations are still the same as in the proof of Lemma 15. Hence, whenever there exists an optimal controller such that h_2 does never become active, it is equal to the controller from Lemma 15. Therefore, from here on we assume that constraint h_2 becomes active, so let the corresponding first *entry time* be denoted by τ_1. We now argue that λ_2 must have a discontinuity at τ_1.

Before entry time. First of all, suppose τ_1 satisfies $\bar{t}_{fi} \leq \tau_1 \leq \bar{t}_{si}$ for some i, then we show that $\tau_1 = \bar{t}_{fi}$. Use Taylor approximation around τ_1 to show that $\bar{t}_{fi} < \tau_1 \leq \bar{t}_{si}$ would imply that $x_2 < -\omega$ somewhere, which is not possible.

Otherwise, suppose τ_1 satisfies $\bar{t}_{si} \leq \tau_1 \leq \bar{t}_{ai}$ for some i, then $x_2(\tau_1) = \bar{x}_2(\tau_1) = 0$. This implies that $u(t) = -\omega$ for $t \in (\tau_1 - 1/\omega, \tau_1)$ and $u(t) = 0$ for $t \in (t_0, \tau_1 - 1/\omega)$. Hence, the position of the entry point needs to satisfy

$$\begin{aligned}\bar{x}_1(\bar{t}_{si}) = x_1(\tau_1) &= p_0 + \tau_1 - 1/\omega - t_0 + p^-(1/\omega) \\ &= p_0 + \tau_1 - t_0 - 1/2\omega,\end{aligned}$$

from which we derive that

$$\tau_1 = t_0 - p_0 + 1/2\omega + \bar{x}_1(\bar{t}_{si}).$$

Furthermore, we have $t_{d1} = \tau_1 - 1/\omega$ in this case.

Otherwise, suppose τ_1 satisfies $\bar{t}_{ai} \leq \tau_1 \leq \bar{t}_{f,i+1}$ for some i, then $x_1(\tau_1) = \bar{x}_1(\tau_1) = \bar{x}_1(\bar{t}_{ai}) + p^+(\tau_1 - \bar{t}_{ai}; \bar{x}_2(\bar{t}_{ai}))$ and $x_2(\tau_1) = \bar{x}_2(\tau_1) = \bar{x}_2(\bar{t}_{ai}) + (\tau_1 - \bar{t}_{ai})\omega$. This means that the duration of the initial deceleration d must be given by

$$d = \frac{1 - x_2(\tau_1)}{\omega} = \frac{1 - \bar{x}_2(\bar{t}_{ai})}{\omega} - \tau_1 + \bar{t}_{ai}.$$

Observe that, starting from the initial position, the current vehicle drives at full speed for a duration of $\tau_1 - d - t_0$ before decelerating, so the position of the entry point is given by

$$\begin{aligned}x_1(\tau_1) &= p_0 + \tau_1 - d - t_0 + p^-(d), \\ &= p_0 + \tau_1 - t_0 - \omega d^2/2.\end{aligned}$$

We can find τ_1 by equating $x_1(\tau_1) = \bar{x}_1(\tau_1)$, which yields the quadratic equation

$$\begin{aligned}p_0 + \tau_1 - t_0 - \omega d^2/2 &= \bar{x}_1(\bar{t}_{ai}) + \bar{x}_2(\bar{t}_{ai})(\tau_1 - \bar{t}_{ai}) + \omega(\tau_1 - \bar{t}_{ai})^2/2, \\ \implies \tau_1 &= \bar{t}_{ai} - \frac{\bar{x}_2(\bar{t}_{ai})}{\omega} + \frac{1}{\omega} \\ &\quad - \frac{1}{2}\sqrt{\frac{2}{\omega^2} + \frac{2(\bar{x}_2(\bar{t}_{ai}))^2}{\omega^2} + 4\left(\frac{p_0 - t_0 - \bar{x}_1(\bar{t}_{ai}) + \bar{t}_{ai}(1 + \bar{x}_2(\bar{t}_{ai}))}{\omega}\right)} \\ t_{d1} &= \tau_1 - d = 2\tau_1 - \bar{t}_{ai} + \frac{\bar{x}_2(\bar{t}_{ai}) - 1}{\omega}.\end{aligned}$$

After exit time. Now consider the part of the trajectory after leaving the boundary. We leverage the result from Lemma 16. □

We now show how the result of Theorem 1 can be used to compute the trajectories of a sequence of vehicles scheduled to drive through a lane. We keep $p_0 = -d(v, w) + W$ fixed. Let the entry and exit times be denoted, respectively, by $y(i, v)$ and $y(i, w)$, for some vehicle indices $i = 1, \dots, n$. Assume these *crossing times* y enable a feasible sequence of trajectories $x^{(i)}$. We use $x[t_0, p_0, \bar{x}_1](t)$ to denote the optimal trajectory given by Theorem 1, evaluated at time t. Similarly, we write $x[t_0, p_0, \varnothing]$ to denote the solution given by Lemma 15, so when the lead constraint can be ignored. For the very first vehicle that arrives, we compute

$$x^{(1)}(t) = x[y(1, v) - y(1, w),\ p_0,\ \varnothing](t - y(1, w)).$$

For every next vehicle $i \geq 2$, we recursively compute

$$x^{(i)}(t) = x[y(i, v) - y(i, w),\ p_0,\ x_1^{(i-1)}(t - y(j, w)) - L](t - y(j, w)).$$

Include illustration to support this notation and derivation. Discuss where these trajectories are actually defined (domain).

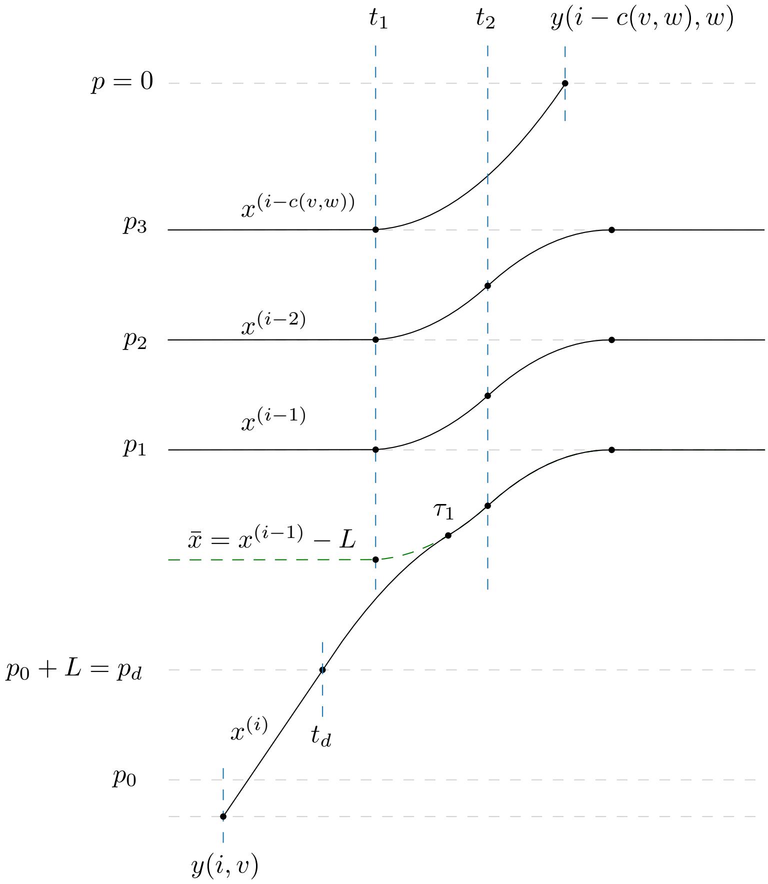

Figure 6: Example with stationary capacity $c(v, w) = 3$.

1.3 Feasibility of crossing times

Theorem 2. *Given crossing times $y(i, v)$ and $y(i, w)$, there exists a feasible trajectory x if and only if the crossing times satisfy the follow constraints $y(i, v) \geq y(i-1, v) + \rho$ and $y(i, w) \geq y(i-1, w) + \rho$ and buffer constraint $y(i, v) \geq \dots$ and travel constraint $y(i, v) + p_0 \leq y(i, w)$.*

Proof. Argue travel constraint. Argue follow constraint. Rest is for buffer constraint.

Necessary. Consider crossing times $y(i, v)$ and define $t_a := y(i - c(v, w), w) - 1/\omega$ and $t_d := t_a + \min\{1/\omega, \sqrt{L/\omega}\}$.

Sufficient. □

We now investigate the limit on the number of vehicles that can occupy the lane when waiting. Imagine that vehicles enter the lane until it is full and then only after all vehicles have come to a full stop, they start to leave the lane by crossing w. We refer to the maximum number of vehicles that can wait in the lane as the *stationary lane capacity.* Suppose that we want to design the tandem network that has a stationary lane capacity of at least $c(v, w)$. Vehicles are required to drive at full speed as long as they occupy any intersection. Therefore, a vehicle crossing v can only start decelerating after $p(t) \geq p_0 + L$, so the earliest position where a vehicle can come to a stop is $p_0 + L + p^+(d_f)$. Because vehicles need to gain maximum speed before reaching w, the position closest to w where a vehicle can wait is $-p^+(d_f)$. Hence,

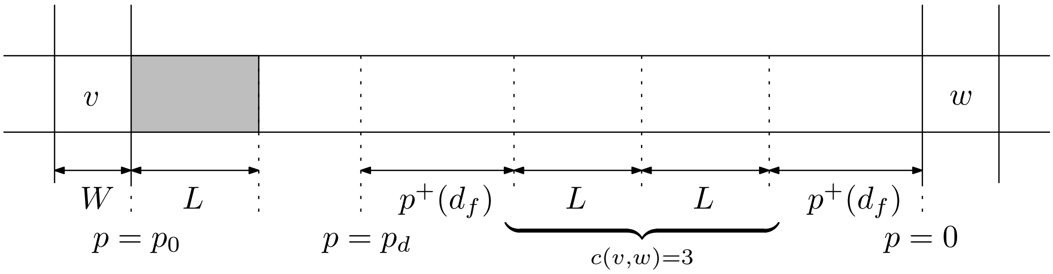

Figure 7: Tandem of intersections with indicated distances used in the derivation of the stationary lane capacity, which is the maximum number of vehicles that can stop and wait in the lane, before they leave.

in order to accomodate for $c(v, w)$ waiting vehicles, the length of the lane must satisfy

$$d(v, w) \geq W + L + 2p^+(d_f) + (c(v, w) - 1)L,$$

as illustrated in Figure 7. Conversely, given the lane length $d(v, w)$, the corresponding stationary lane capacity is given by[2]

$$c(v, w) = \texttt{floor}\left(\frac{d(v, w) - W - 2p^+(d_f)}{L}\right),$$

where $\texttt{floor}(x)$ denotes the largest integer smaller than or equal to x.

It turns out that the fixed locations where vehicles wait in the above scenario are helpful in describing the optimal trajectories, even when vehicles never fully stop. We will denote these fixed *waiting positions* as

$$p_k = -p^*(d_f) - (c(v, w) - k)L,$$

for $k = 1, \ldots, c(v, w)$. Furthermore, let $p_d = p_1 - p^*(d_f)$ denote the position from which vehicles must decelerate in order to stop at the first waiting position p_1. Now consider a vehicle that moves from p_k to the next waiting position p_{k+1}, so it moves exactly distance L. We consider such a start-stop movement, without considering any safe following constraints. By symmetry of the control constraints, the vehicle moves the same distance during both acceleration and deceleration. Furthermore, the vehicle needs to be at rest at the start and end of such trajectory. Hence, it is clear that it takes the same amount of time d_s to accelerate and decelerate. We assume that $d_s < d_f$, which ensures that maximum velocity is never reached during the start-stop movement, which is illustrated in Figure **??**. In this case, it is clear that we must have $L = 2p^*(d_s)$, from which we derive that $d_s = \sqrt{L/a_{\max}}$.

[2]Without Assumption 1, we cannot derive such a simple formula, because it would depend on the specific lengths of those vehicles currently in the system.

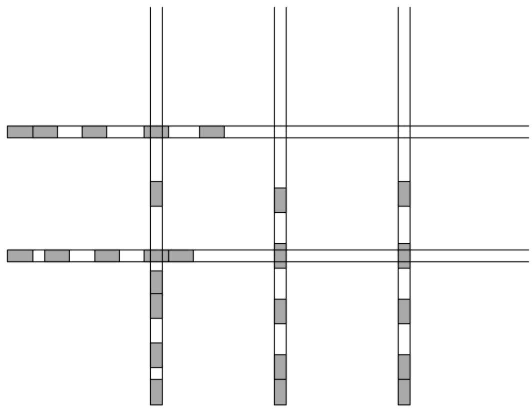

Figure 8: Illustration of some grid-like network of intersections with vehicles drawn as grey rectangles. There are five vehicle routes: two from east to west and three from south to north. Turning at intersections is not allowed.

2 Vehicle scheduling in networks

We now extend the single intersection model to a network of intersections without turning routes, illustrated in Figure 8. We define a directed graph $(\bar{V}, E)$ with nodes $\bar{V}$ and arcs E, representing the possible paths that vehicles can follow. Nodes with only outgoing arcs are *entrypoints* and nodes with only incoming arcs are *exitpoints*. Let V be the set of *intersections*, which are all the nodes with in-degree at least two. Let $d(v, w)$ denote the distance between nodes v and w. For each route index $r \in \mathcal{R}$, we let

$$\bar{V}_r = (v_r(0), v_r(1), \ldots, v_r(m_r), v_r(m_r + 1))$$

be the path that vehicles $i \in \mathcal{N}_r$ follow through the network. We require that the first node $v_r(0)$ is an entrypoint and that the last node $v_r(m_r + 1)$ is an exitpoint and we write

$$V_r = \bar{V}_r \setminus \{v_r(0),\, v_r(m_r + 1)\}$$

to denote the path restricted to intersections. We say that some $(v, w) \in E$ is on path V_r whenever v and w are two consecutive nodes on the path and we write E_r to denote the set of all these edges. We require that routes can only overlap at nodes by making the following assumption.

Assumption 4. *Every arc $(v, w) \in E$ is part of at most one route V_r, such that routes do not share lanes. This ensures that the order of vehicles on each lane is completely determined by the order of vehicles on the corresponding lane.*

We start by considering networks in which all roads are axis-aligned such that intersections always involve perpendicular lanes and where routes are such that no turning is required. For each $v \in V_r$ define the conflict zone $\mathcal{E}_r(v) = (b_r(v), e_r(v))$ and consider the union

$$\mathcal{E}_r = \bigcup_{v \in V_r} \mathcal{E}_r(v)$$

corresponding to the positions of vehicles $i \in \mathcal{N}_r$ for which it occupies an intersection on its path V_r. By reading $\mathcal{E}_i \equiv \mathcal{E}_r$ for $r(i) = r$, the single intersection problem naturally extends to the network case. Like before, the resulting problem can be numerically solved by a direct transcription method.

2.1 General decomposition

The general two-stage decomposition for the single intersection extends rather naturally to the present model. Let for each pair (i, v) of some vehicle $i \in \mathcal{N}$ and an intersection $v \in V_{r(i)}$ along its route, let

$$\inf\{t : x_i(t) \in \mathcal{E}_r(v)\} \quad \text{and} \quad \sup\{t : x_i(t) \in \mathcal{E}_r(v)\}$$

be the crossing time and exit time, which we denote by $y(i, v)$ and $y(i, v)+\sigma(i, v)$, respectively. Instead of a single set of conflicts, we now define for each intersection $v \in V$ in the network the set of conflict pairs

$$\mathcal{D}^v = \{\{i, j\} \subset \mathcal{N} : r(i) \neq r(j), v \in V_{r(i)} \cap V_{r(j)}\}.$$

Now the two-stage approach is to solve

$$\begin{aligned}
\min_{y,\sigma} \quad & \sum_{r \in \mathcal{R}} F(y_r, \sigma_r) \\
\text{s.t.} \quad & y(i, v) + \sigma(i, v) \leq y(j, v) \text{ or} \\
& y(j, v) + \sigma(j, v) \leq y(i, v), && \text{for all } \{i, j\} \in \mathcal{D}^v \text{ and } v \in V, \\
& (y_r, \sigma_r) \in \mathcal{S}_r, && \text{for all } r \in \mathcal{R},
\end{aligned}$$

where $F(y_r, \sigma_r)$ and $\mathcal{S}_r$ are the value function and set of feasible parameters, respectively, of the parametric trajectory optimization problems

$$\begin{aligned}
F(y_r, \sigma_r) = \min_{x_r} \quad & \sum_{i \in \mathcal{N}_r} J(x_i) \\
\text{s.t.} \quad & x_i(t) \in D_i(s_{i,0}), && \text{for } i \in \mathcal{N}_r, \\
& x_i(y(i, v)) = b_r(v), && \text{for } v \in V_r, i \in \mathcal{N}_r, \\
& x_i(y(i, v) + \sigma(i, v)) = e_r(v), && \text{for } v \in V_r, i \in \mathcal{N}_r, \\
& x_i(t) - x_j(t) \geq L, && \text{for } (i, j) \in \mathcal{C} \cap \mathcal{N}_r,
\end{aligned}$$

where we again use subscript r to group variables according to their associated route.

2.2 Decomposition for delay objective

Suppose we use use the crossing at the last intersection as performance measure, by defining the objective function as

$$J(x_i) = \inf\{t : x_i(t) \in \mathcal{E}_r(v_r(m_r))\}.$$

We show how to reduce the resulting problem to a scheduling problem, like we did in the single intersection case. We will again assume Assumption 1 and Assumption 3, so vehicles will always cross intersections at full speed, and all vehicles share the same geometry. Hence, the occupation time $\sigma \equiv \sigma(i, v)$ is the same for all vehicles and intersections. For this reason, we will write the shorthand $y_r \in \mathcal{S}_r$, because σ_r is no longer a free variable.

As a consequence of Assumption 1 and Assumption 3, each lower-level trajectory optimization problem for a given route $r \in \mathcal{R}$ decomposes into a sequence of problems, each corresponding to two consecutive intersection along V_r. This means that $y_r \in \mathcal{S}_r$ is equivalent to $y_{(v,w)} \in \mathcal{S}_{(v,w)}$ for each $(v, w) \in E_r$, where $y_{(v,w)}$ denotes the vector of all variables $y(i, v)$ and $y(i, w)$ for all $i \in \mathcal{N}_r$ and $\mathcal{S}_{(v,w)}$ denotes the set of values of $y_{(v,w)}$ for which a feasible trajectory part can be found. Hence, we will now focus on a tandem of two intersections and investigate the trajectories of vehicles in this with the goal of stating sufficient conditions for $y_{(v,w)} \in \mathcal{S}_{(v,w)}$.

2.3 Crossing time scheduling

Analogously to the single intersection case, we let the earliest crossing time $\beta_t(i, v)$ for vehicle $i = (r, k) \in \mathcal{N}$ at network nodes $v \in \bar{V}_r$ in current disjunctive graph G_t be recursively defined through

$$\beta_t(i, v) = \begin{cases} a(i, v) & \text{if } v \text{ is an entrypoint,} \\ \max_{i \in \mathcal{N}_t^-(j)} \beta_t(i, v) + w(i, j) & \text{otherwise,} \end{cases}$$

where $\mathcal{N}_t^-(j)$ is again the set of in-neighbors of node j in G_t. For empty schedules, it is easily seen that we have $\beta_0(i, v_r(0)) = a(i, v_r(0))$ for entrypoints and we have $\beta_0(i, v_r(l+1)) = \beta_0(i, v_r(l)) + d(v_r(l), v_r(l+1))/v_{\max}$ for $l = 1, \dots, m_r$, so between consecutive intersections on the same route.

There are two natural choices for the objective to optimize in this scheduling setting. We minimize the time each vehicle is in the system, which is equivalent to minimizing the delay at the last intersection of each vehicle's route, written as

$$\text{obj}_1(y) = \sum_{i \in \mathcal{N}} y(i, v_r(m_r)) - \beta_0(i, v_r(m_r)).$$

Alternatively, it also makes sense to minimize the delay at every intersection along each vehicle's route, so we can also minimize

$$\text{obj}_2(y) = \sum_{i \in \mathcal{N}} \sum_{v \in V_{r(i)}} y(i, v) - \beta_0(i, v).$$

3 Learning to schedule

Like in the single intersection case, we try to model optimal solutions using autoregressive models. In the single intersection case, we argued that each schedule is uniquely defined by its route order η. Generalizing this to the case of multiple intersections, we see that a schedule is uniquely defined by the set of route orders η^v at each intersection $v \in V$. Instead of working with such a set of sequences, we will intertwine these sequences to obtain a single *crossing sequence* η, which consists of pairs (r, v) of a route $r \in \mathcal{R}$ at some intersection $v \in V_r$ and we will refer to such a pair as a *crossing*. Of course, this crossing sequence can be constructed in many different ways, because it does not matter in which order the intersections are considered, which is illustrated in Figure 9. Given some problem instance s, we will consider autoregressive models of the form

$$p(\eta|s) = \prod_{t=1}^{N} p(\eta_t|s, \eta_{1:t-1}). \tag{8}$$

These models can also be understood in terms of a step-by-step schedule construction process that transitions from a partial schedule state $s_t = (s, \eta_{1:t})$ to the next state s_{t+1} by selecting some *action* $\eta_{t+1} = (r_{t+1}, v_{t+1})$. We say a crossing is pending when it still has unscheduled vehicles. Similarly, we say an intersection is pending when some of its crossings are still pending. With these definitions, the set of *valid actions* $\mathcal{A}(s_t)$ at some intermediate state s_t is exactly the set of pending crossings. We again emphasize that multiple sequences of actions lead to the same schedule, because the order in which intersections are considered does not matter for the final schedule, which is illustrated in Figure 9. The models that we study can be understood as being parameterized as some function of the disjunctive graph G_t of partial schedule s_t, so an alternative way of writing (8) that emphasizes this is

$$p(\eta = ((r_1, v_1), \ldots, (r_N, v_N)) \mid s) = \prod_{t=1}^{N} p(r_t, v_t|G_{t-1}). \tag{9}$$

Instead of modeling the joint probability distribution $p(r_t, v_t|G_{t-1})$ over crossings, we can also apply the chain rule to factorize it as

$$p(r_t, v_t|G_{t-1}) = p(r_t|v_t, G_{t-1})p(v_t|G_{t-1}). \tag{10}$$

In this case, the model $p(r_t|v_t, G_{t-1})$ can be thought of as predicting a set of actions, one for each intersection.

Intersection visit order. In the general class of autoregressive models for inputs and outputs with sets, of which our model above is a special case (we have a set of sequences as output), it has been noted before that the order in which inputs or outputs are presented to the model during training has a considerable impact on the final model fit [4]. For our model, we also expect to find this effect, so we will investigate the impact of the order in which intersections are visited, determined by $p(v_t|G_{t-1})$. We now propose some heuristic ways to define $p(v_t|G_{t-1})$. Later, we will consider a neural network parameterization. First of all, the simple *random* strategy would be to sample some intersection with pending crossings at each step. In the *boundary* strategy, we keep visiting the same intersection until it is done (when it has no pending crossings anymore), then move to some next intersection. When the network of intersections is free of cycles, we could for example follow some topological order. We use the term “boundary” because this strategy produces trajectories along the boundary of the grid in Figure 9. In the *alternating* strategy, we keep alternating between intersection to keep the number of scheduled vehicles balanced among them. This produces trajectories that can be understood as being close to the “diagonal” of the grid in Figure 9. Again, the order in which we alternate between intersections may again be based on some topological order.

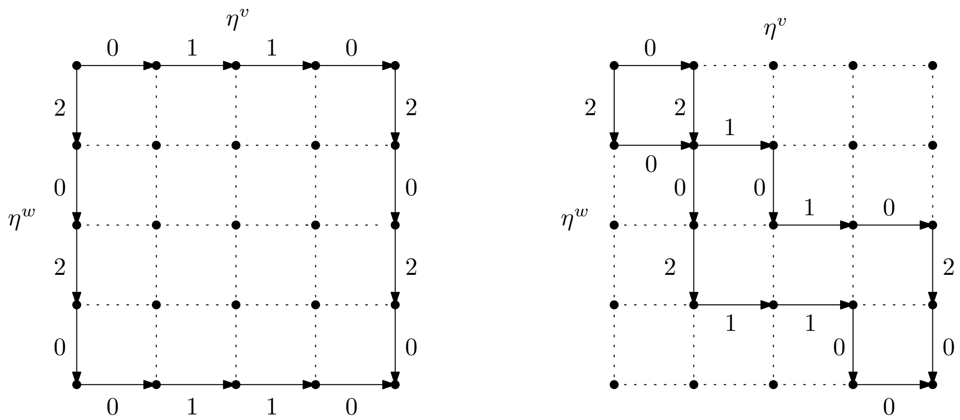

Figure 9: When each route order is locally fixed at every intersection, the global crossing sequence is not uniquely determined, because these local sequences may be merged in any order. Suppose we have a tandem of two intersections and a horizontal arrows correspond to taking the next local action at intersection v, a vertical arrows correspond to taking next local action at intersection w. Each valid crossing sequence corresponds to some path from the top-left to the bottom-right corner. Although any such crossing sequence produces the same schedule, it might be that our autoregressive model fits better on some sequences than others. For example, we might expect that sequences on the boundary of the grid, shown in the left grid, are harder to learn from data than sequences that stay closer to the diagonal, like in the right grid. The intuition is that we need to "look into the future" more to learn the former, while in the latter trajectories, progress in the two intersections is more balanced.

3.1 Model parameterization

3.1.1 Threshold heuristics

It is straightforward to extend the threshold rule to networks of intersections, when assuming a fixed intersection order. Each time some next intersection is visited, we apply the single intersection threshold rule to pick the next route. This is straightforward to do, because we can just consider the disjunctive subgraph induced by the nodes belonging to that intersection to arrive at the single intersection case. Furthermore, the definition of the threshold rule itself does not depend on the network of intersections. This is a desirable property, because it allows us to tune the threshold on small networks and then apply it on larger ones.

3.1.2 Neural constructive heuristic

We will now propose a neural network parameterization of $p(r_t|v_t, G_{t-1})$ and train it based on optimal schedules in a supervised learning setting. The model can be best understood as solving a so-called multi-label classification problem, because it needs to provide a distribution over routes at every intersection. The training data set $\mathcal{X}$ consists of pairs $(G_{t-1}, (r_t, v_t))$, to which we refer to as *state-action* pairs to draw the parallel with the terminology used in the reinforcement learning literature. To obtain these pairs, we sample a collection of problem instances, which are solved to optimality using a MILP solver. For each optimal schedule, we compute the corresponding optimal route order η^v for each intersection. From these, we can construct $\mathcal{X}$ in different ways. For example, for each solved instance, we can randomly select some intersection order μ, which fixes the crossing order η. We can then replay this sequence of actions step-by-step to obtain the corresponding sequence of state-action pairs. The model might become more robust when training on multiple samples of intersections orders per instance. Alternatively, we can consider one of the fixed intersection orders described at the start of this section (random, boundary, alternating). Furthermore, combined with one of the

Table 1: Average scaled optimal objective value computed using MILP and using the threshold heuristic with threshold $\tau = 0$. Each class of problem instances is identified by the number n of vehicle arrivals per route and the grid network size as cols x rows.

n	size	MILP	time	$\tau = 0$ (gap)	random (gap)	boundary (gap)	alternate (gap)
5	2x1	57.27	0.06	65.27 (11.45%)	58.96 (2.95%)	58.72 (2.52%)	58.23 (1.66%)
5	3x1	57.67	0.12	68.34 (15.44%)	59.77 (3.65%)	59.82 (3.74%)	58.72 (1.83%)
5	3x2	57.35	1.38	69.17 (18.32%)	60.88 (6.16%)	60.36 (5.25%)	58.82 (2.56%)

above strategies, we can also employ some kind of fixed lookahead procedure, as illustrated in Figure **??**. At inference time, we use the same intersection order as during training and at each step we greedily select r_t.

We will now describe how the model is parameterized based on recurrent embeddings of the sequences of crossing time lower bounds at each crossing, in a somewhat similarly to how we used the horizons in the single intersection model. Let $k(r, v)$ denote the number of scheduled vehicles at crossing (r, v) and let n_r denote the total number of vehicles on route r. For each crossing (r, v), consider the earliest crossing time of the next unscheduled vehicle, which we denote as

$$T(r, v) = \beta(r, k(r, v) + 1, v).$$

We define the *horizon* of crossing (r, v) to be the sequence of relative lower bounds

$$h(r, v) = (\beta(r, k(r, v) + 2, v) - T, \dots, \beta(r, n_r, v) - T).$$

Each such horizon is embedded using an Elman RNN. To produce a logit for each crossing, these embeddings are fed through a feedforward network, consisting of two hidden layers of size 128 and 64 neurons with relu activation and batch normalization layers in between. See Figure **??** for a schematic overview of the architecture. Next: instead of $h(r, v)$, feed $(T(r, v), h(r, v))$ to the FF.

To train the model, we use the Adam optimizer with learning rate 10^{-4} and a batch size of 40 state-action pairs. We experienced a case of the exploding gradients problem when we did not use the batch normalization layers. To allow a fair comparison of methods accross instances of different sizes, both in terms of the number of vehicles and the size of the network, we adapt both objective as follows. For the first variant, we divide by the total number of vehicles, so we report $\text{obj}_1(y)/|\mathcal{N}|$. Given some problem instance, let N denote the length of a crossing sequence, so it is also the total number of vehicle-intersection pairs (i, v) that need to be scheduled. For the second objective variant, we consider the average delay per vehicle-intersection pair, so we report $\text{obj}_2(y)/N$, which we report in Table 1.

3.2 Reinforcement learning

Instead of using a fixed training set $\mathcal{X}$ of state-action pairs during training, we can fit our model from the perspective of a reinforcement learning problem, which we already alluded to in Section 3.1.2. More precisely, given some scheduling problem instance s, we are dealing with a Deterministic Markov Decision Process (DMDP), where partial disjunctive graphs serve as states and crossings correspond to actions. The potential benefit of using reinforcement learning is that we do not have to fix an intersection order: our hope is that the training procedure will automatically converge to some good intersection order.

The threshold heuristic ($\tau = 0$) can provide a *baseline* for reinforcement learning, reducing the variance of the REINFORCE estimator.

N.B. The step-numbering is different in $G_0, \eta_1, G_1, \eta_2, \dots$ from the common RL notation $S_0, A_0, R_1, S_1, A_1, \dots$, so generally $A_t = \eta_{t+1}$.

References

[1] D. Miculescu and S. Karaman, "Polling-systems-based Autonomous Vehicle Coordination in Traffic Intersections with No Traffic Signals," July 2016.

[2] R. F. Hartl, S. P. Sethi, and R. G. Vickson, "A Survey of the Maximum Principles for Optimal Control Problems with State Constraints," *SIAM Review*, vol. 37, no. 2, pp. 181–218, 1995.

[3] S. P. Sethi, *Optimal Control Theory: Applications to Management Science and Economics*. Cham: Springer International Publishing, 2019.

[4] O. Vinyals, S. Bengio, and M. Kudlur, "Order Matters: Sequence to sequence for sets," Feb. 2016.

A Pontryagin's Maximum Principle

Because the literature does not provide a clear statement of Pontryagin's maximum principle in case of pure state constraints of arbitrary order and our main reference contains a typo, we first state the precise form that we are going to use. We present here the so-called indirect adjoining approach found in [2]. Instead of the general form, we present here a variant specialized to the case when state constraints are of first and second order.

Consider the optimal control problem with state constraints

$$\begin{aligned} \max \quad & \int_{t=t_0}^{t_f} F(x(t), u(t), t) dt \\ \text{s.t.} \quad & \dot{x}(t) = f(x(t), u(t), t), \quad x(t_0) = x_0, \\ & a(x(t_f), t_f) \geq 0, \\ & b(x(t_f), t_f) = 0, \\ & g(x(t), u(t), t) \geq 0, \\ & h(x(t), t) \geq 0, \end{aligned} \tag{11}$$

with objective function $F : \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R} \to \mathbb{R}$, dynamics function $f : \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R} \to \mathbb{R}^n$, mixed constraints $g : \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R} \to \mathbb{R}^s$, pure constraints $h : \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}^q$ and boundary constraint functions a, b mapping $\mathbb{R}^n \times \mathbb{R}$ into $\mathbb{R}^l$ and $\mathbb{R}^{l'}$, respectively. These functions are assumed to be continuously differentiable with respect to all their arguments. Given some solution $\{x^*(t), u^*(t)\}$, we use the notation $F^*[\,t\,] = F(x^*(t), u^*(t), t)$ for evaluation along this trajectory of function F and similarly for other functions.

Let $h_1(x, t)$ and $h_2(x, t)$ be pure state constraints of order one and two, respectively. Let their derivatives be

$$\begin{aligned} h_1^1(x, u, t) &= \frac{dh_1(x(t), t)}{dt}, \\ h_2^1(x, t) &= \frac{dh_2(x(t), t)}{dt}, \quad h_2^2(x, u, t) = \frac{dh_2^1(x(t), t)}{dt}. \end{aligned}$$

The Hamiltonian is given by

$$H(x, u, \lambda, t) = F(x, u, t) + \lambda f(x, u, t)$$

and the Lagrangian is given by

$$L(x, u, \lambda, \mu, \eta_1, \eta_2, t) = H(x, u, \lambda, t) + \mu g(x, u, t) + \eta_1 h_1(x, u, t) + \eta_2 h_2(x, u, t).$$

Let the control region be defined as

$$\begin{aligned}\Omega(x,t) = \{u \in \mathbb{R} \mid\ & g(x,u,t) \geq 0, \\ & h_1^1(x,u,t) \geq 0 \text{ if } h_1(x,t) = 0, \\ & h_2^2(x,u,t) \geq 0 \text{ if } h_2(x,t) = 0\}.\end{aligned}$$

Theorem 3 (Pontryagin's Maximum Principle). *Let $\{x^*(\cdot), u^*(\cdot)\}$ be an optimal pair for the optimal control problem, then there exists a piecewise absolutely continuous costate trajectory $\lambda(\cdot)$, piecewise continuous multiplier functions $\mu(\cdot)$, $\eta_1(\cdot)$ and $\eta_2(\cdot)$, numbers $\zeta^0(\tau_i), \zeta^1(\tau_i), \zeta^2(\tau_i)$ for each point τ_i of discontinuity of $\lambda(\cdot)$, such that the following conditions hold almost everywhere:*

$$\begin{gathered}u^*(t) = \arg\max_{u \in \Omega(x^*(t),t)} H(x^*(t), u, \lambda(t), t), \\ L_u^*[t] = 0, \\ \dot{\lambda}(t) = -L_x^*[t], \\ \mu(t) \geq 0, \mu(t) g^*[t] = 0, \\ \eta_1 \geq 0, \dot{\eta}_1(t) \leq 0, \eta_1 h_1^*[t] = 0, \\ \eta_2 \geq 0, \dot{\eta}_2(t) \leq 0, \ddot{\eta}_2(t) \geq 0, \eta_2 h_2^*[t] = 0.\end{gathered}$$

At each entry time τ, the costate trajectory λ may have a discontinuity of the form

$$\begin{gathered}\lambda(\tau^-) = \lambda(\tau^+) + \zeta^0(\tau)(h_1)_x^*[\tau] + \zeta^1(\tau)(h_2)_x^*[\tau] + \zeta^1(\tau)(h_2^1)_x^*[\tau], \\ H^*[\tau^-] = H^*[\tau^+] - \zeta^0(\tau)(h_1)_t^*[\tau] - \zeta^1(\tau)(h_2)_t^*[\tau] - \zeta^1(\tau)(h_2^1)_t^*[\tau], \\ \zeta^0(\tau) \geq 0, \quad \zeta^0(\tau) h_1^*[\tau] = 0, \\ \zeta^1(\tau) \geq 0, \quad \zeta^1(\tau) h_2^*[\tau] = 0, \\ \zeta^2(\tau) \geq 0, \quad \zeta^2(\tau) h_2^*[\tau] = 0.\end{gathered}$$