4.2. The Queue Accumulation Polygon

We will now introduce a concept that is fundamental to the analysis of traffic signal operations, and one which will be invoked repeatedly throughout many of the sessions that follow. It is very important that you understand the modeling of queue accumulation and discharge, and the queue accumulation polygon (QAP) is the basis of just about all of the analytical models described in the literature. Fortunately, it’s a fairly simple concept.

Let’s begin by representing one cycle of operation graphically as shown below:

The red and green portions are shown with the symbols g and r representing the effective green and red time, respectively. Remember from Session 2 that the effective green time is the total phase time (green + yellow + all-red) minus the lost time, L. The rest of the time in the cycle is effective red time. This definition simplifies the figure above by eliminating the need to show the yellow, all-red and lost time intervals.

Now, let’s make this the horizontal axis of a graph that plots the number of vehicles in the queue at all points in the cycle as shown below:

Assumptions:

- No initial queue at the beginning of the red
- Constant arrival rate throughout the cycle.
We must assume that the cycle starts with no queue, and that the arrival rate of vehicles on this approach is uniform throughout the cycle. We will deal later with the complications that arise when these assumptions are violated.

It is clear that the number of vehicles in the queue will increase at a constant rate, as determined by the traffic volume on the approach, and will reach its maximum value at the end of the effective red. Let’s go back to our simple example (again, the northbound approach) and compute the maximum number of vehicles in the queue. What we already know at this point is that the traffic volume is 500 vph, the cycle length is 90 sec and the effective green time is 33.3 sec. As shown in the box at the right, we can compute the arrival rate in vehicles per second by dividing the hourly volume (500) by the number of seconds in an hour (3600). We can compute the effective red time by subtracting the effective green time (33.3) from the cycle length (90). We can then compute the number of vehicles in the queue at the end of the red by multiplying the arrival rate (0.139) by the length of the effective red (56.7). The queue will start to decrease as vehicles depart on the green, so the number in the queue at the end of the red will be the maximum number for the cycle.

Now, let’s examine the departure of vehicles from the queue, as shown below. During the effective green period, the number of vehicles in the queue will decrease at a rate determined by the difference between the departure rate (i.e., saturation flow rate) and the arrival rate. Remember, vehicles continue to arrive on the green at the same rate that they did on the red.

Uniform delay is given by the area contained within the queue accumulation polygon.

---

**Max Queue Size (northbound)**

\[
q = \frac{500}{3600} = 0.139 \text{ veh per sec} \\
g = 33.3 \text{ sec} \\
r = 90 - 33.3 = 56.7 \text{ sec} \\
N_{\text{max}} = 0.139 \times 56.7 = 7.88 \text{ veh}
\]
This completes the QAP which, in this case is simply a triangle, shaded as shown above. You might argue that calling it a polygon might be glorifying it somewhat, but remember, this is a very basic example. When we introduce the complications of multi-phase operation and permitted left turns later, you will see some very complex shapes. Just be thankful that, for the moment, it’s only a triangle.

Now, it is a principle of operations research that delay in a system between any two points in time is the integral of the storage within the system expressed as a function of time. If you insist on the full treatment:

\[ d = \int_{t_1}^{t_2} s(t)dt \]

Where \( d \) = the delay in the system between the two points in time and \( s(t) \) is the storage within the system as a function of time.

If you prefer the simple explanation, it means that the total delay to vehicles on each cycle will be given by the area contained within the QAP. That’s good news because we all have known for many years how to compute the area of a triangle as half the base times the altitude, and we have already figured out the altitude.

The computation of the base of the QAP is illustrated in the box at the right. Notice from the QAP figure that part of the effective green time is included in the base and part is not. The part that is included is defined as the *saturated green*, \( g_s \). The portion left over after the queue has been serviced is defined as the *unsaturated green time*, \( g_u \). These two components together make up the whole effective green time.

It is clear from the figure above that the base of the QAP is simply the sum of the red time, \( r \), and the saturated green time, \( g_s \). We already know \( r \), and we have enough information to compute \( g_s \) by dividing the queue size at the beginning of the effective green by the rate at which vehicles depart from the queue.

We must be careful to use the net departure rate here, because vehicles continue to arrive during the green period. We must first express the saturation flow rate in vehicles per second instead of vehicles per hour to keep the units consistent. We do this by dividing the hourly saturation flow rate (1800) by the number of seconds in an hour (3600). The net departure rate is simply the saturation flow rate per second (0.5) minus the arrival rate per second (0.139). Knowing this, we can easily determine the saturated green time by dividing the initial queue at the beginning of the effective green (7.88) by the net departure rate (0.361). Now we can determine the base of the triangle by adding the saturated green time (21.8) to the effective red time (56.7).
At this point, the area contained within the QAP is within easy reach. We simply take half the base (78.5) and multiply by the altitude (7.88). The resulting number (309.4) sounds like an awful lot of delay, so we need to examine what this number really means. The base of the QAP triangle is given in seconds, and the altitude is given in vehicles. So the area will be given in vehicle-seconds. The vehicle-second is the basic delay unit and is defined as the result of one vehicle waiting for one second. For example, 60 vehicle-seconds could be the result of 60 vehicles each waiting for one second. It could also be the result of one vehicle waiting for one minute, two vehicles waiting for 30 seconds each, etc.

Since the QAP represents one cycle of operation, its area actually represents the number of vehicle-seconds of delay per cycle. The problem with this representation is we don’t really know whether the computed number represents a lot of delay or not. We would do better to express the result in terms of the average number of seconds of delay experienced by each vehicle. The box on the right illustrates these computations.

We must first determine the number of vehicles processed on each cycle by multiplying the arrival rate per second (0.139) by the cycle length (90). The result (12.5) looks very familiar. You will remember that we arrived at the same number by a different method in the last section when we were trying to compute the probability of failure.

The delay per vehicle is now a trivial computation. We just divide the total delay per cycle (309.4) by the number of vehicles per cycle (12.5) to arrive at a unit delay of 24.8 seconds per vehicle.

Well, all of this was fun, but we would probably rather do it next time using a formula. Fortunately, there is such a formula, generally known as Weber’s uniform delay equation. The details are shown in the box at the right, along with an example calculation for the northbound movement. Note that the answer given by the formula agrees with our step-by-step solution. The 0.1 second numerical difference between the two results is due to accumulated round off.

Notice that the term uniform delay has been introduced here. This is because of the assumption of a uniform arrival rate. We will examine the effect of random arrivals next.

This exercise has illustrated the principle of the queue accumulation polygon and its application to one approach of the sample intersection. What about the other approaches? You guessed it! See the assignment.