CHAPTER 26

Signal Coordination for Arterials and Networks: Undersaturated Conditions

26.1 Basic Principles of Signal Coordination

In situations where signals are close enough together so that vehicles arrive at the downstream intersection in platoons, it is necessary to coordinate their green times so that vehicles may move efficiently through the *set* of signals. It serves no purpose to have drivers held at one signal watching wasted green at a downstream signal, only to arrive there just as the signal turns red.

In some cases, two signals are so closely spaced that they should be considered one signal. In other cases, the signals are so far apart that they may be considered isolated intersections. However, vehicles released from a signal often maintain their grouping for well over 1,000 feet. Common practice is to coordinate signals less than a mile apart on major streets and highways.

26.1.1 A Key Requirement: Common Cycle Length

In coordinated systems, all signals must have the same cycle length. This is necessary to ensure that the beginning of green occurs at the same time relative to the green at the upstream and downstream intersections. There are some exceptions, where a critical intersection has such a high volume that it may require a double cycle length, but this is done rarely and only when no other solution is feasible.

26.1.2 The Time-Space Diagram and Ideal Offsets

The time-space diagram is a plot of signal indications as a function of time for two or more signals. The diagram is scaled with respect to distance, so that one may easily plot vehicle positions as a function of time. Figure 26.1 is a time-space diagram for two intersections. Standard conventions are generally used in such figures: A green signal indication is shown by a blank or simple line (——), yellow by a shaded line (////////), and red by a solid line (▬▬). In many cases, such diagrams show only effective green and effective red, as shown in Figure 26.1. This figure illustrates the path (trajectory) that a vehicle takes as time passes. At $t = t_1$, the first signal turns green. After some lag, the vehicle starts and moves down the street. It reaches the second intersection at some time $t = t_2$. Depending on the indication of that signal, it either continues or stops.

The difference between the two green initiation times (i.e., the difference between the time when the upstream intersection turns green and the downstream intersection turns green) is referred to as the *signal offset* or simply the *offset*. In Figure 26.1, the offset is defined as t_2 minus t_1. Offset is usually expressed as a positive number between zero and the cycle length. This definition is used throughout this and other chapters in this text.

Other definitions of offset are used in practice. For instance, offset is sometimes defined relative to one reference upstream signal, and sometimes it is defined relative to a standard zero. Some signal hardware uses "offset" defined in terms of red initiation, rather than green; other hardware uses the end of green as the reference point. Some hardware uses offset in seconds; other hardware uses offset as a percentage of the cycle length.

The "ideal offset" is defined as exactly the offset such that, as the first vehicle of a platoon just arrives at the downstream signal, the downstream signal turns green. It is usually assumed that the platoon was moving as it went through the upstream intersection. If so, the ideal offset is given by:

$$t_{Ideal} = {}^{L}/_{S} \qquad (26\text{-}1)$$

where: t_{Ideal} = ideal offset, s
L = distance between signalized intersections, ft
S = vehicle speed, ft/s

If the vehicle were stopped, and had to accelerate after some initial startup delay, the ideal offset could be represented by Equation 26-1 plus the startup time at the first intersection (which would usually add 2 to 4 seconds). In general, the startup time would only be included at the *first* of a series of

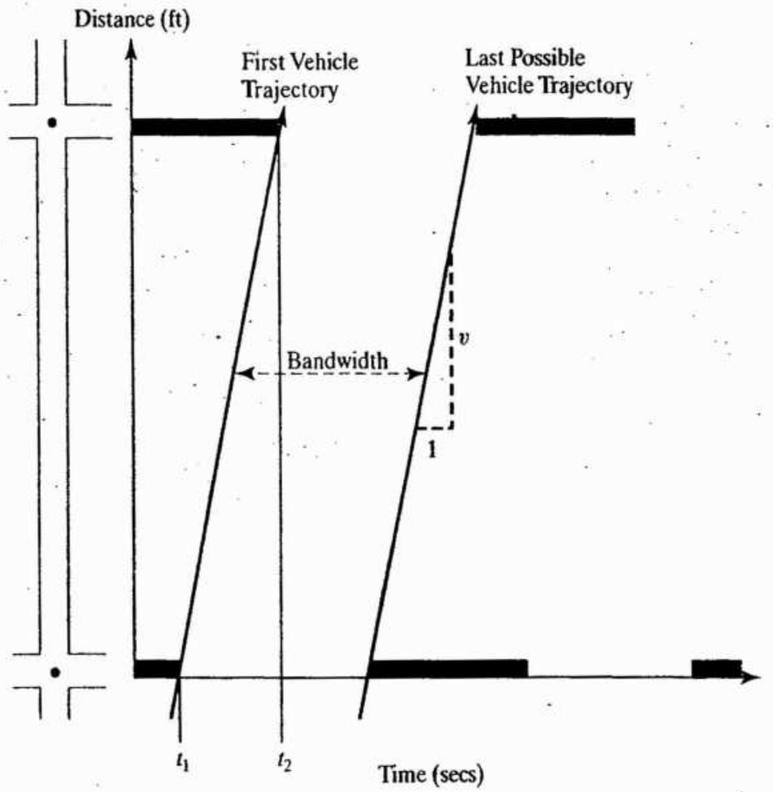


Figure 26.1: Illustrative Vehicle Trajectory on a Time-Space Diagram

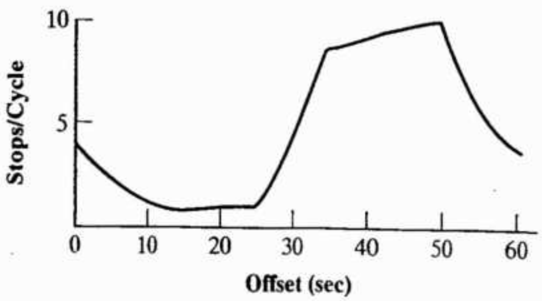
signals to be coordinated, and often not at all. Usually, this will reflect the ideal offset desired for maximum bandwidth, minimum delay, and minimum stops. Even if the vehicle is stopped at the first intersection, it will be moving through most of the system.

Figure 26.1 also illustrates the concept of *bandwidth,* the amount of green time that can be used by a continuously moving platoon of vehicles through a series of intersections. In Figure 26-1, the bandwidth is the entire green time at both intersections because several key conditions exist:

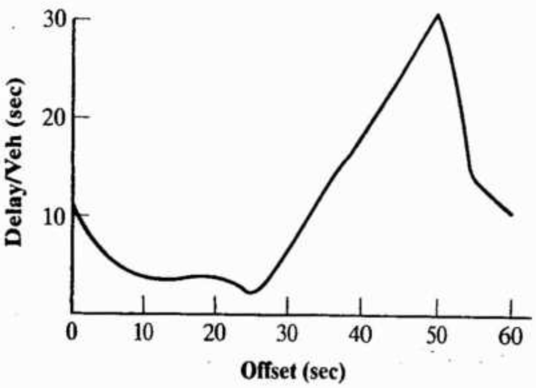
- The green time at both intersections are the same.
- The ideal offset is illustrated.
- There are only two intersections.

In most cases, the bandwidth will be less, perhaps significantly so, than the full green time.

Figure 26.2 illustrates the effect of offset on stops and delay for a platoon of vehicles leaving one intersection and passing through another. In this example, a 25-second offset is ideal because it produces the minimum delay and the minimum number of stops. The effect of allowing a poor offset to exist is clearly indicated: Delay can climb to 30 seconds per vehicle, and the stops to 10 per cycle. Note that the penalty for deviating from the ideal offset is usually not equal in positive and negative deviations. An offset of (25 + 10) = 35 seconds causes much more harm than an offset of (25 − 10) = 15 seconds, although both are 10 seconds from the ideal offset. Figure 26.2 is illustrative because each situation would have similar but different characteristics.



(a) Stops



(b) Delay

600 ft block
600 vph in two lanes
all through traffic
free speed 24 mph
50-50 split
60 sec. cycle length

Figure 26.2: Illustration of the Effects of Offset on Stops and Delay

26.2 Signal Progression on One-Way Streets

Signal progression on a one-way street is relatively simple. For the purpose of this section, it will be assumed that a cycle length has been chosen and that the green allocation at each signal has been previously determined.

26.2.1 Determining Ideal Offsets

Consider the one-way arterial shown in Figure 26.3, with the link lengths indicated. Assuming no vehicles are queued at the signals, the ideal offsets can be determined if the platoon speed is known. For the purpose of illustration, a desired platoon speed of 60 ft/s is used. The cycle length is 60 seconds, and the effective green time at each intersection is 50% of the

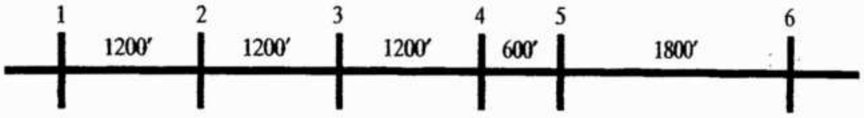


Figure 26.3: Case Study in Progression on a One-Way Street

Table 26.1: Ideal Offsets for Case Study

Signal	Relative to Signal	Ideal Offset
6	5	1,800/60 = 30 s
5	4	600/60 = 10 s
4	3	1,200/60 = 20 s
3	2	1,200/60 = 20 s
2	1	1,200/60 = 20 s

cycle length, or 30 seconds. Ideal offsets are computed using Equation 26-1 and are illustrated in Table 26.1.

Note that neither the cycle length nor the splits enter into the computation of ideal offsets. To see the pattern that results, the time-space diagram should be constructed according to the following rules:

1. The vertical should be scaled so as to accommodate the dimensions of the arterial, and the horizontal so as to accommodate at least three to four cycle lengths.
2. The beginning intersection (Number 1, in this case) should be scaled first, usually with main street green (MSG) initiation at $t = 0$, followed by periods of green and red (yellow may be shown for precision). See Point 1 in Figure 26.4.
3. The main street green (or other offset position, if MSG is not used) of the next downstream signal should be located next, relative to $t = 0$ and at the proper distance from the first intersection. With this point located (Point 2 in Figure 26-4), fill in the periods of effective green and red for this signal.
4. Repeat the procedure for all other intersections, working one at a time. Thus, for Signal 3, the offset is located at point 3, 20 seconds later than Point 2, and so on.

Figure 26.4 has some interesting features that can be explored with the aid of Figure 26.5.

First, if a vehicle (or platoon) were to travel at 60 fps, it would arrive at each of the signals just as they turn green; this is indicated by the solid trajectory lines in Figure 26.5. The solid trajectory line also represents the speed of the "green wave" visible to a stationary observer at Signal 1, looking downstream. The signals turn green in order, corresponding to the planned speed of the platoon, and give the visual effect of a wave of green opening before the driver. Third, note that there is a "window" of green in Figure 26.5, with its end indicated by the dotted trajectory line, which is also the trajectory of the *last* vehicle that could travel through the progression without

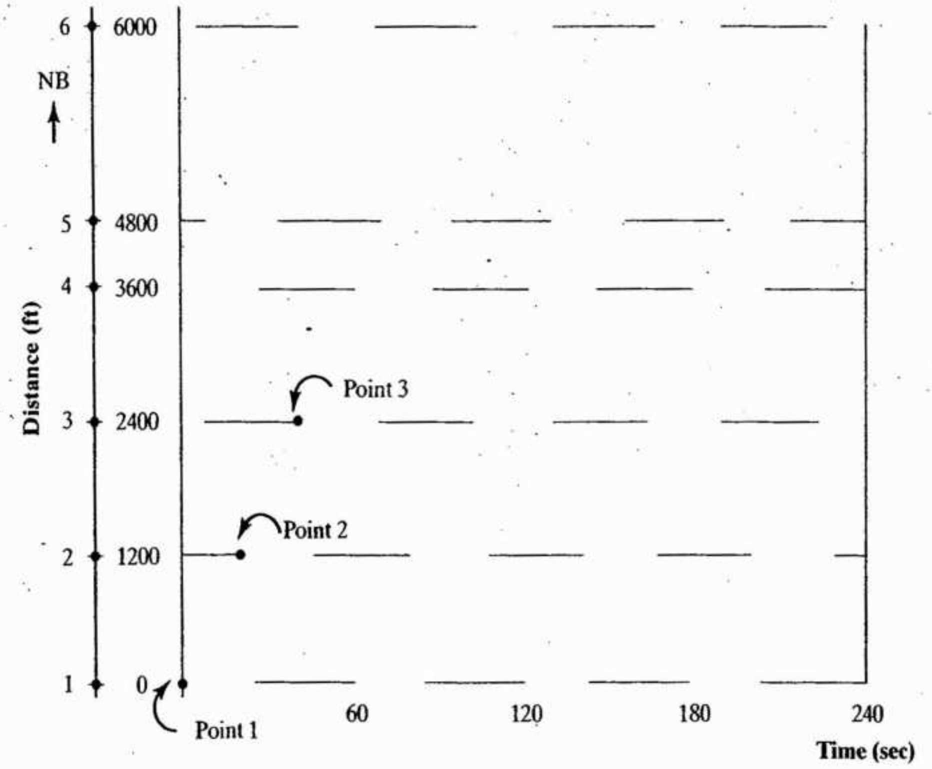


Figure 26.4: Time-Space Diagram for the Case Study

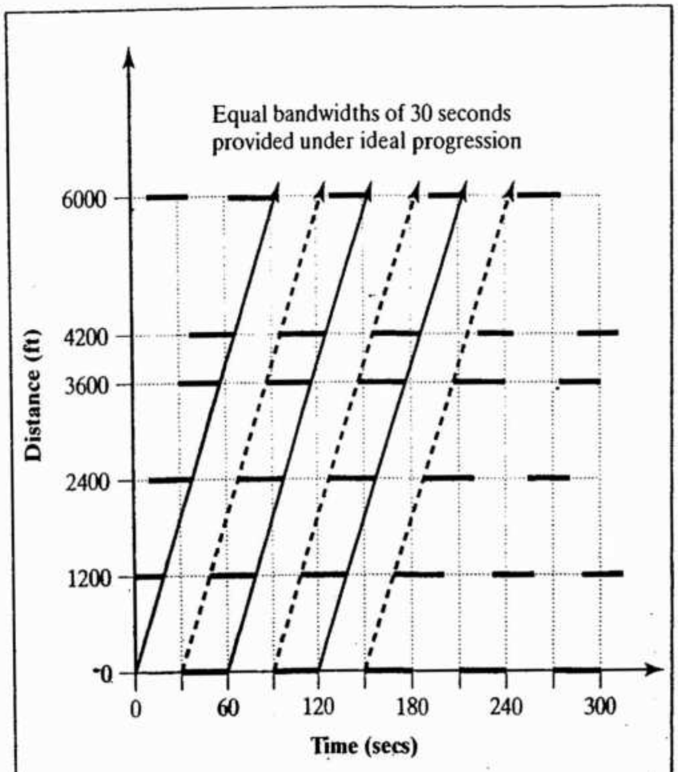


Figure 26.5: Vehicle Trajectory and "Green Wave" in a Progressed Movement

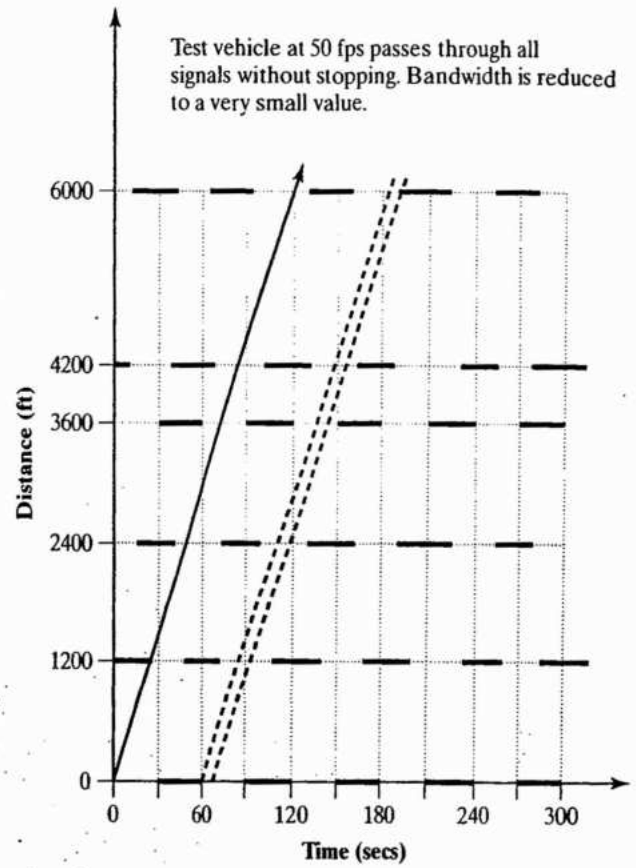


Figure 26.6: The Effect of a 50 ft/s Platoon Speed on Progression

stopping at 60 ft/s. This "window" is the bandwidth, as defined earlier. Again, in this case it equals the green time because all signals have the same green time and have ideal offsets.

26.2.2 Potential Problems

Consider what would happen if the actual speed of vehicle platoons in the case study was 50 ft/s, instead of the 60 ft/s anticipated. The green wave would still progress at 60 ft/s, but the platoon arrivals would lag behind it. The effect of this on bandwidth is enormous, as shown in Figure 26.6. Only a small window now exists for a platoon of vehicles to continuously flow through all six signals without stopping.

Figure 26.7 shows the effect of the vehicle traveling faster than anticipated (70 ft/s in this illustration). In this case, the vehicles arrive a little too early and are delayed; some stops will have to be made to allow the "green wave" to catch up to the platoon.

In this case, the effect on bandwidth is not as severe as in Figure 26.6. In this case, the bandwidth impact of *underestimating* the platoon speed is (60 ft/s instead of 70 ft/s) is not as severe as the consequences of *overestimating* the platoon speed (60 ft/s instead of 50 ft/s).

26.3 Bandwidth Concepts

Bandwidth is defined as the time difference between the first vehicle that can pass through the entire system without stopping and the last vehicle that can pass through without stopping, measured in seconds.

The bandwidth concept is very popular in traffic engineering practice because the windows of green are easy visual images for both working professionals and public presentations. The most significant shortcoming of designing offset plans to maximize bandwidths is that internal queues are often overlooked in the bandwidth approach. There are computer-based maximum bandwidth solutions that go beyond the historical formulations, such as PASSER [*1*] and Tru-Traffic TS/PP [*2*]. We have used the latter in professional practice and show an example in Section 26.7.

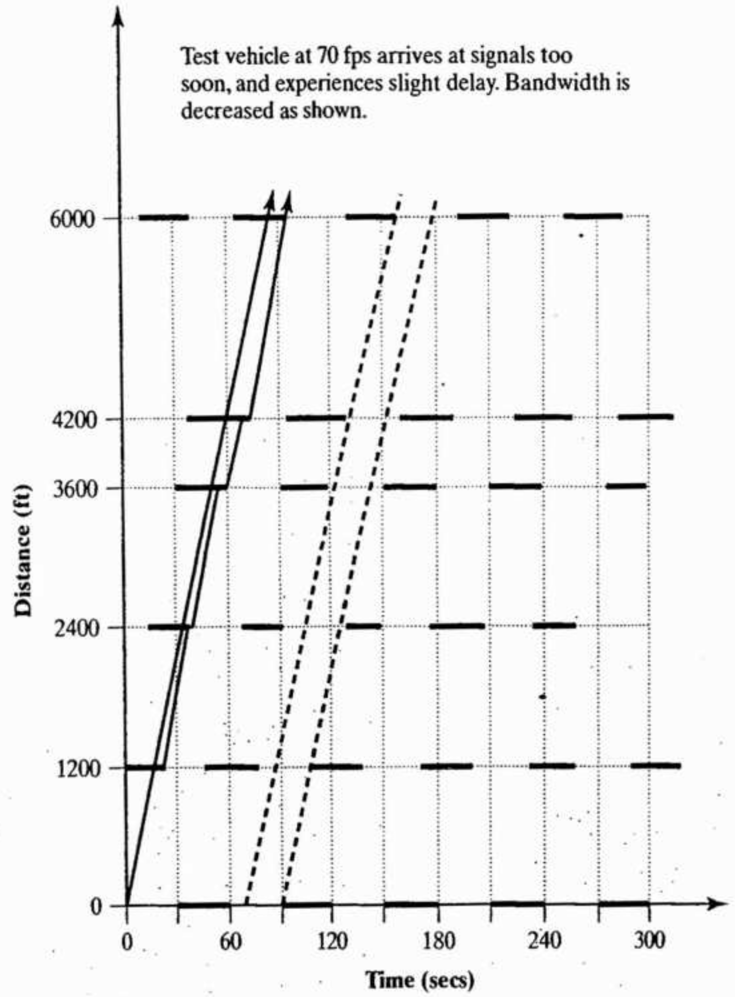


Figure 26.7: The Effect of a 70 ft/s Platoon Speed on Progression

26.3.1 Bandwidth Efficiency

The efficiency of a bandwidth is defined as the ratio of the bandwidth to the cycle length, expressed as a percentage:

$$EFF_{BW} = \left(\frac{BW}{C}\right) * 100\% \qquad (26\text{-}2)$$

where: EFF_{BW} = bandwidth efficiency (%)
BW = bandwidth (s)
C = cycle length (s)

A bandwidth efficiency of 40% to 55% is considered good. The bandwidth is limited by the minimum green in the direction of interest.

Figure 26.8 illustrates the bandwidths for one signal-timing plan. The northbound efficiency can be estimated as (17/60) * 100% = 28.4%. The southbound bandwidth is obviously terrible; there is no bandwidth through the defined system. The northbound efficiency is only 28.4%. This system is badly in need of retiming, at least on the basis of the bandwidth objective. Just looking at the time-space diagram, one might imagine sliding the pattern at Signal 4 to the right and the pattern at Signal 1 to the left, allowing some coordination for the southbound vehicles.

26.3.2 Bandwidth Capacity

In terms of vehicles that can be put through the system of Figure 26.7 without stopping, the northbound bandwidth can carry 17/2.0 = 8.5 vehicles per lane per cycle in a nonstop path through the defined system, assuming that the saturation headway is 2.0 s/veh. Thus the northbound direction can handle 8.5 veh/cycle * 1 cycle/60 sec * 3,600 sec/hr = 510 veh/h/ln very efficiently if they are

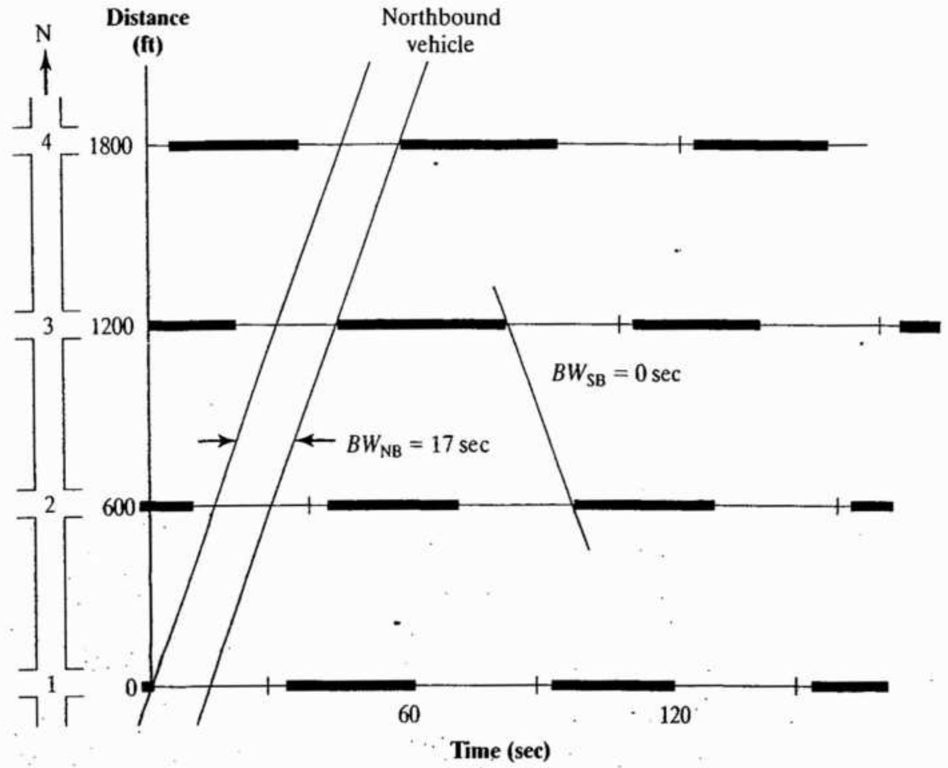


Figure 26.8: Bandwidths on a Time-Space Diagram

organized into eight-vehicle platoons when they travel through this system.

If the per lane demand volume is less than 510 vphpl and if the flows are well organized (and if there is no internal queue development), the system will operate well in the northbound direction, even though better timing plans might be obtained.

In general terms, the number of vehicles that can pass through a defined series of signals without stopping is called the *bandwidth capacity*. The illustrated computation can be described by the following equation:

$$c_{BW} = \frac{3600 * BW * NL}{C * h} \tag{26-3}$$

where: c_{BW} = bandwidth capacity, veh/h
BW = bandwidth, s
NL = number of through lanes in the indicated direction
C = cycle length, s
h = saturation headway, s

Equation 26-3 does not contain any factors to account for nonuniform lane utilization and is intended only to indicate some limit beyond which the offset plan will degrade, certainly resulting in stopping and internal queuing. It should also be noted that bandwidth capacity is *not* the same as lane group capacity. Where the bandwidth is less than the full green time, there is additional lane group capacity outside of the bandwidth.

26.4 The Effect of Queued Vehicles at Signals

To this point, it has been assumed there is no queue standing at the downstream intersection when the platoon (from the upstream signal arrives). This is generally not a reasonable assumption. Vehicles that enter the traffic stream between platoons will progress to the downstream signal, which will often be "red." They form a queue that partially blocks the progress of the arriving platoon. These vehicles may include stragglers from the last platoon, vehicles that turned into the block from unsignalized intersections or driveways, or vehicles that came out of parking lots or parking spots. The ideal offset must be adjusted to allow for these vehicles, so as to avoid unnec-essary stops. The situation without such an adjustment is depicted in Figure 26.9, where it can be seen

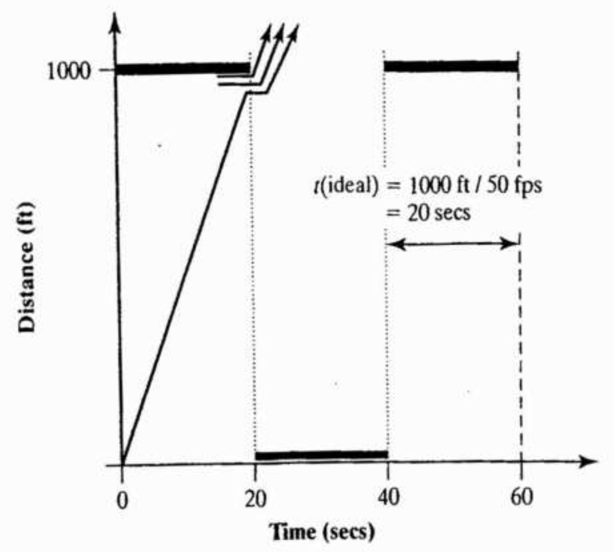


Figure 26.9: The Effect of Queued Vehicles at a Signal

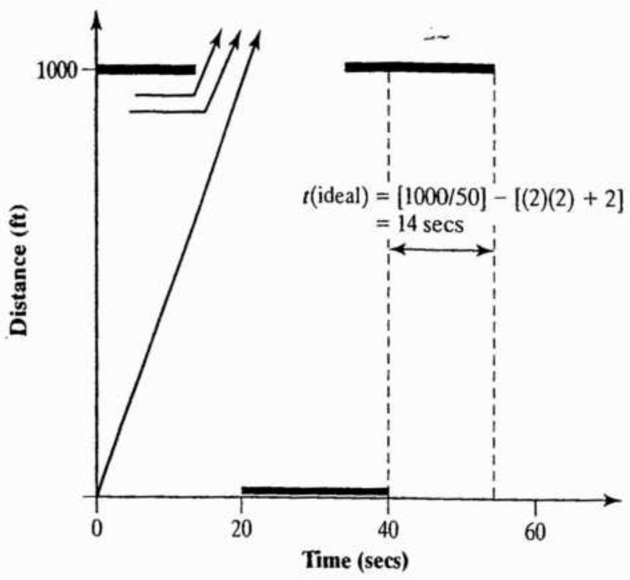


Figure 26.10: Adjustment in Offset to Accommodate Queued Vehicles

that the arriving platoon is delayed behind the queued vehicles as the queued vehicles begin to accelerate through the intersection.

To adjust for the queued vehicles, the ideal offset is adjusted as follows:

$$t_{adj} = \frac{L}{S} - (Qh + \ell_1) \qquad (26\text{-}4)$$

where: t_{adj} = adjusted ideal offset, s
L = distance between signals, ft
S = speed, ft/s
Q = number of vehicles queued per lane, veh
h = discharge headway of queued vehicles, s/veh
ℓ_1 = start-up lost time, s

The lost time is counted only at the first downstream intersection, at most: If the vehicle(s) from the preceding intersection were themselves stationary, their startup causes a shift that automatically takes care of the startup at subsequent intersections.

Offsets can be adjusted to allow for queue clearance before the arrival of a platoon from the upstream intersection. Figure 26.10 shows the situation for use of the modified ideal offset equation.

Figure 26.11 shows the time-space diagram for the case study of Figure 26.4, given queues of two vehicles per lane in all links. Note that the arriving vehicle platoon has smooth flow, and the lead vehicle has 60 ft/s travel speed. The visual image of the "green wave," however, is much faster, due to the need to clear the queues in advance of the arriving platoon.

The "green wave" or the progression speed, as it is more properly called, is traveling at varying speeds as it moves down the arterial. The "green wave" will appear to move ahead of the platoon, clearing queued vehicles in advance of it. The progression speed can be computed for each link as shown in Table 26.2.

Note, however, that the bandwidth, and therefore the bandwidth capacity, is now much smaller. Thus, by clearing out the queue in advance of the platoon, more of the green time is used by queued vehicles, and less is available to the moving platoon.

The preceding discussion assumes the queue is known at each signal. In fact, this is not an easy number to know. However, if we know there is a queue and know its approximate size, the link offset can be set better than by pretending that no queue exists.

Consider the sources of the queued vehicles:

- Vehicles turning in from upstream side streets during their green (which is main street red)
- Vehicles leaving parking garages or spaces
- Stragglers from previous platoons.

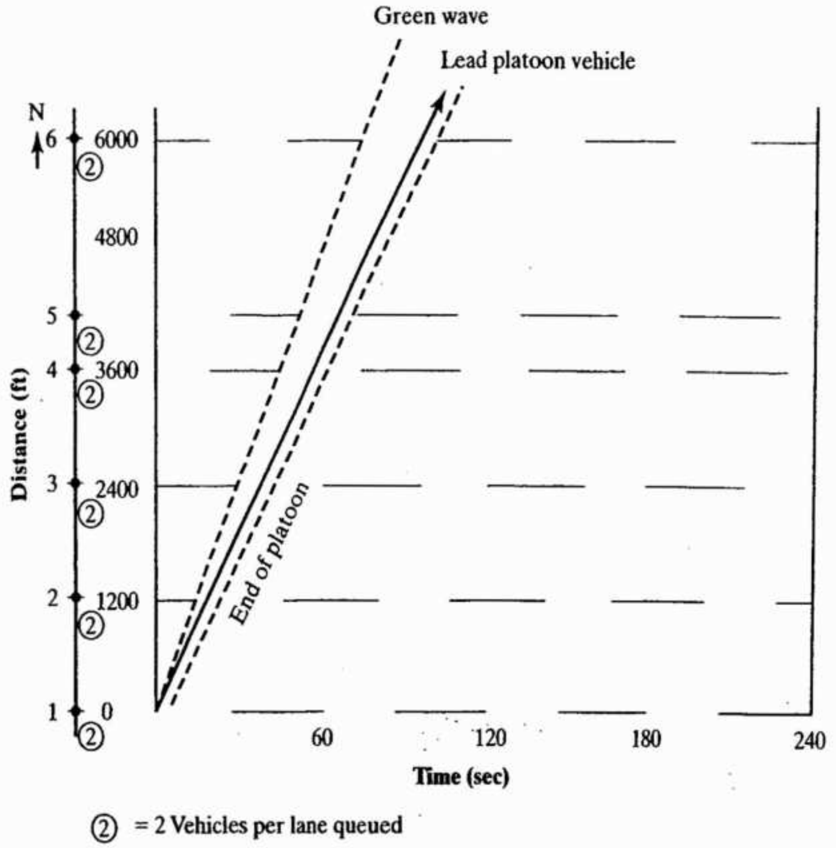


Figure 26.11: Effect of Queue Clearance on Progression Speed

There can be great cycle-to-cycle variation in the actual queue size, although the average queue size may be estimated. Even at that, queue estimation is a difficult and expensive task. Even the act of adjusting the offsets can influence the queue size. For instance, the arrival pattern of the vehicles from the side streets may be altered. Queue estimation is therefore a significant task in practical terms.

26.5 Signal Progression for Two-Way Streets and Networks

The task of progressing traffic on a one-way street has been relatively straightforward. To highlight the essence of the problem on a two-way street, assume the arterial shown in Figure 26.5 is a two-way street rather than a one-way street. Figure 26.12

Table 26.2: Progression Speeds in Figure 26-11

Link	Link Offset (s)	Speed of Progression (ft/s)
Signal 1 → 2	(1,200/60) − (4 + 2) = 14	1,200/14 = 85.7
Signal 2 → 3	(1,200/60) − (4) = 16	1,200/16 = 75
Signal 3 → 4	(1,200/60) − (4) = 16	1,200/16 = 75
Signal 4 → 5	(600/60) − (4) = 6	600/6 = 100
Signal 5 → 6	(1,800/60) − (4) = 26	1,800/26 = 69.2

Total Offset = 78 sec

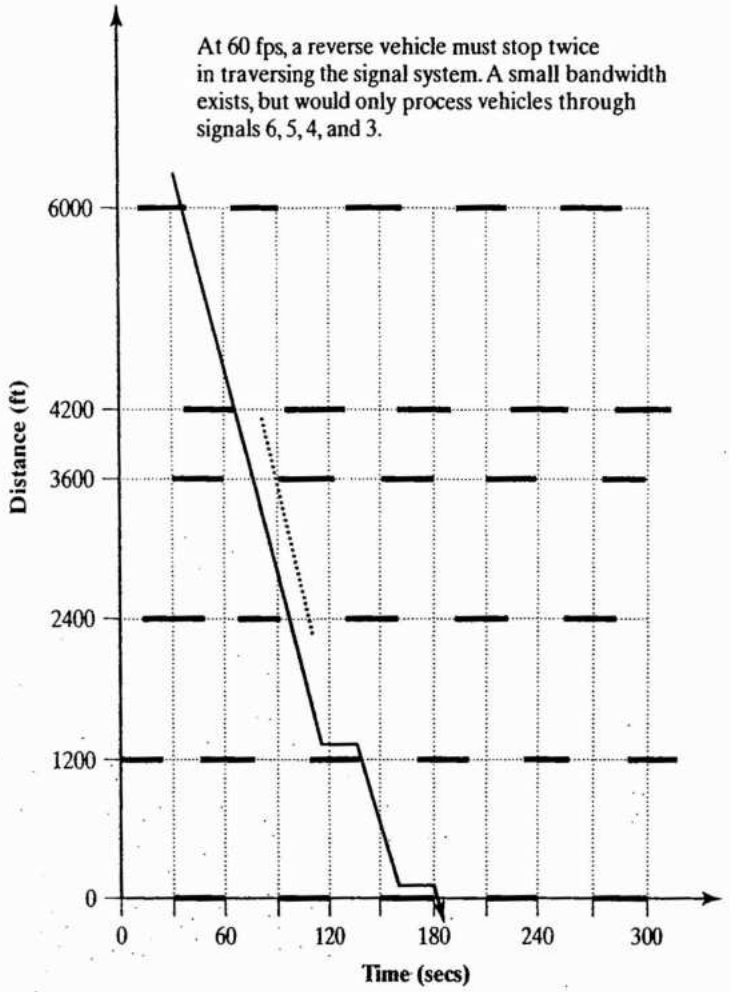

At 60 fps, a reverse vehicle must stop twice in traversing the signal system. A small bandwidth exists, but would only process vehicles through signals 6, 5, 4, and 3.

Figure 26.12: Case Study: The Southbound Result of a Northbound Progression

ows the trajectory of a *southbound* vehicle on this arterial. he vehicle is just fortunate enough not to be stopped until ignal 2, but is then stopped again for Signal 1, for a total of stops and 40 seconds of delay. There is no bandwidth, mean- g it is not possible to have a vehicle platoon pass along the terial nonstop.

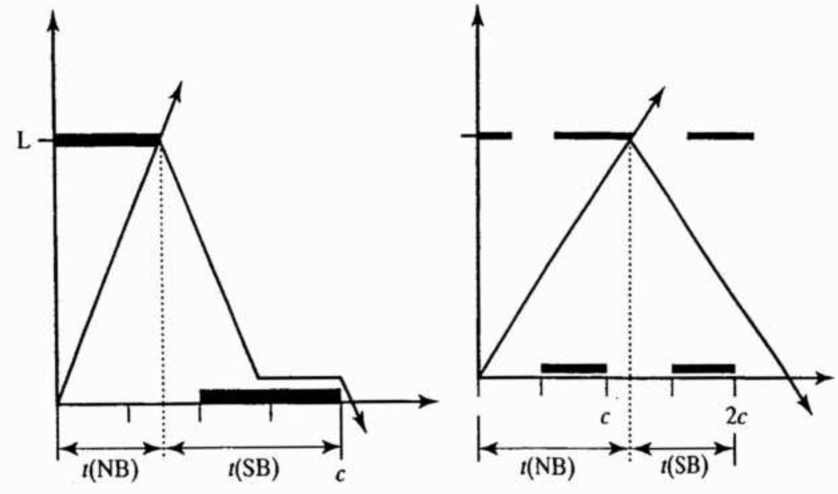
Of course, if the offsets or the travel times had been dif- rent, it might have been possible to have a southbound ndwidth through all six signals.

6.5.1 Offsets on a Two-Way Street

ote that if any offset were changed in Figure 26.12 to commodate the southbound vehicles, then the north- und bandwidth would suffer. For instance, if the offset at gnal 2 were decreased by 20 seconds, then the pattern at that signal would shift to the left by 20 seconds, resulting in a "window" of green of only 10 seconds on the northbound, rather than the 30 seconds in the original display (Figure 26.5).

The fact that the offsets on a two-way street are interrelated presents one of the most fundamental problems of signal optimization. Note that inspection of a typical time-space diagram yields the obvious conclusion that the offsets in two directions add to one cycle length, shown in Figure 26.13 (a). However, for longer blocks, the offsets might add to two (or more) cycle lengths, shown in Figure 26.13 (b).

Figure 26.13 illustrates both actual offsets and travel times, which are not necessarily the same. Although the engineer might desire the ideal offset to be the same as the travel times, this is not always the case. Once the offset is specified in one direction, it is automatically set in the other.



(a) Offsets Add to One Cycle Length (b) Offsets Add to Two Cycle Lengths

Figure 26.13: Offsets on a Two-Way Street Are Not Independent

The general expression for the two offsets in a link on a two-way street can be written as:

$$t_{1i} + t_{2i} = nC \tag{26.5}$$

where: t_{1i} = offset in direction 1 (link i), s
t_{2i} = offset in direction 2 (link i), s
n = integer value
C = cycle length, s

To have $n = 1$ (Figure 26.13a), $t_{1i} \leq C$; to have $n = 2$ (Figure 26-13b), $C < t_{1i} \leq 2C$.

Any actual offset can be expressed as the desired "ideal" offset, plus an "error" or "discrepancy" term:

$$t_{actual(i,j)} = t_{ideal(i,j)} + e_{ij} \tag{26.6}$$

where j represents the direction and i represents the link. In a number of signal optimization programs that are used for two-way arterials, the objective is to minimize some function of the discrepancies between the actual and ideal offsets.

26.5.2 Network Closure

The relative difficulty of finding progressions on a two-way street, compared to on a one-way street, might lead one to conclude that the best approach is to establish a system of one-way streets, to avoid the problem. A one-way street system has a number of advantages, not the least of which is elimination of left turns against opposing traffic. One-way streets simplify network signalization, but they do not eliminate closure problems, and they carry other practical disadvantages. See Chapter 29 for additional discussion of one-way streets.

Figure 26.14 illustrates network closure requirements. In any set of four signals, offsets may be set on three legs in one direction. Setting three offsets, however, fixes the timing of all four signals. Thus setting three offsets fixes the fourth.

Figure 26.15 extends this to a grid of one-way streets, in which all of the north-south streets are independently specified. The specification of one east-west street then "locks in" all other east-west offsets. Note that the key feature is that an open

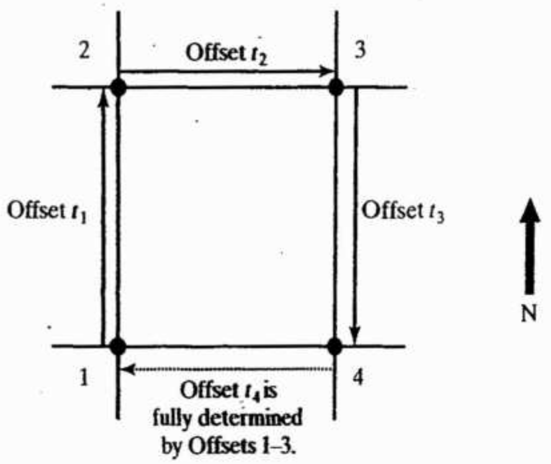


Figure 26.14: Network Closure Illustrated

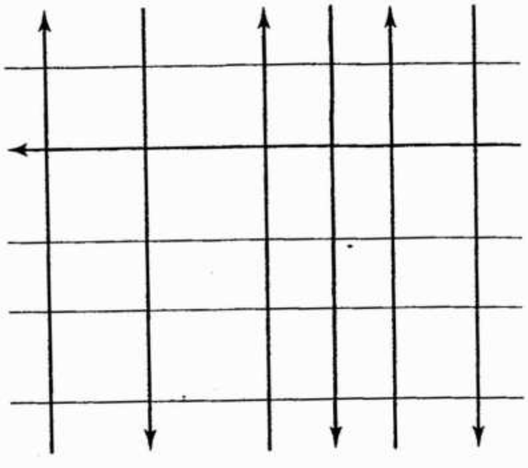


Figure 26.15: Impact of Closure on a Rectangular Street Grid

tree of one-way links can be completely independently set, and that it is the closing or "closure" of the open tree that presents constraints on some of the links.

To develop the constraint equation, refer to Figure 26.14 and walk through the following steps, keying to the green in all steps:

Step 1: Begin at Intersection 1 and consider the green initiation to be time $t = 0$.

Step 2: Move to Intersection 2, noting that the link offset t_1 specifies the time of green initiation at this intersection relative to its upstream neighbor. Thus green starts at Intersection 2 facing northbound at $t = 0 + t_1$.

Step 3: Recognizing that the westbound vehicles get released after the NS green is finished, green begins at Intersection 2 facing west at:

$$t = 0 + t_1 + g_{NS,2}$$

Step 4: Moving to intersection 3, the link offset in Link B specifies the time of green initiation at Intersection 3 relative to Intersection 2. Thus the green begins at Intersection 3, facing west at

$$t = 0 + t_1 + g_{NS,2} + t_2$$

Step 5: Similar to Step 3, the green begins at Intersection 3, but facing south, after the EW green is finished at time

$$t = 0 + t_1 + g_{NS,2} + t_2 + g_{EW,3}$$

Step 6: Moving to Intersection 4, the green begins in the southbound direction after the offset t_3 is added:

$$t = 0 + t_1 + g_{NS,2} + t_2 + g_{EW,3} + t_3$$

Step 7: Turning at Intersection 4, it is the NS green that is added to be at the start of green facing east.

$$t = 0 + t_1 + g_{NS,2} + t_2 + g_{EW,3} + t_3 + g_{NS,4}$$

Step 8: Moving to Intersection 1, it is t_4 that is relevant to be at the start of green facing east:

$$t = 0 + t_1 + g_{NS,2} + t_2 + g_{EW,3} + t_3 + g_{NS,4} + t_4$$

Step 9: Turning at Intersection 1, green will begin in the north direction after the EW green finishes:

$$t = 0 + t_1 + g_{NS,2} + t_2 + g_{EW,3} + t_3 + g_{NS,4} + t_4 + g_{EW,1}$$

This will bring us back to where we started. Thus this is either $t = 0$ or a multiple of the cycle length.

The following relationship results:

$$nC = 0 + t_A + g_{NS,2} + t_B + g_{EW,3} + t_C + g_{NS,4} + t_D + g_{EW,1} \tag{26-7}$$

where the only caution is that the g values should really include the change and clearance intervals.

Note that Equation 26-7 is a more general form of Equation 26-5, for the two-way arterial is a special case of a network. The interrelationships stated in Equation 26-7 are constraints on freely setting all offsets. In these equations one can trade off between green allocations and offsets. To get a better offset in Link 4, one can adjust the splits as well as the other offsets.

Although it is sometimes necessary to consider networks in their entirety, it is common traffic engineering practice to decompose networks into noninterlocking arterials whenever possible. Figure 26.16 illustrates this process.

Decomposition works well where a clear center of activity can be identified, and where few vehicles are expected to pass through the center without stopping (or starting) at or near the center. As the discontinuity in all progressions lies in and directly around the identified center, large volumes passing through can create significant problems in such a scheme.

In summary, if offsets are set in one direction on a two-way street, then the reverse direction is fixed. In a network, you can set any "open tree" of links, but links that close the tree already have their offsets specified.

You are advised to check the literature for the optimization programs in current use. At the present time, the dominant program in the United States seems to be Synchro [3]; several states specify its use in signal optimization and in traffic impact assessments and/or accept Synchro output for levels of

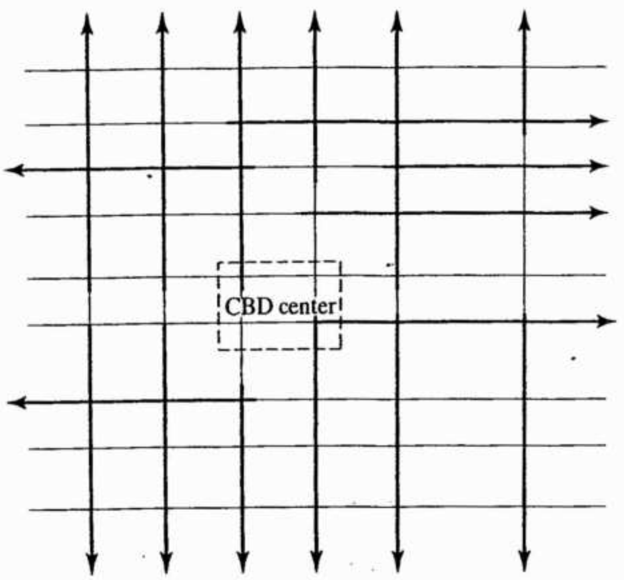


Figure 26.16: Decomposing a Network Into Noninterlocking Arterial Segments

service, equivalent to the Highway Capacity Manual (*HCM*) [*4*] for their purposes.[1] Section 26.7 contains an illustration of Synchro and its use.

TRANSYT [*5*] has been used over the years for arterials and networks and PASSER II [2] for certain arterials. The current version of the HCS+ software [*6*] for intersections is linked to TRANSYT-7F.

The programs just cited are for signal optimization, including determination of offsets. Another group of models exist that simulate and display the results, generally with both 2D and 3D visualizations. These models include VISSIM [*7*], AIMSUN [*8*], SimTraffic[2] [*3*], CORSIM [*9*], and PARAMICS [*10*].

With all of these models, it is important that the user make sure the default parameters reflect the reality of the jurisdiction or area in which they are being used. For instance, the discharge headway may be different than the 1.9 s/veh used in the *HCM*, work zone capacities may be handled differently than the *HCM* or local practice, and so forth.

26.5.3 Finding Compromise Solutions

The engineer usually wishes to design for maximum bandwidth in one direction, subject to some relation between bandwidths in the two directions. Sometimes, one direction is completely ignored. Much more commonly, the bandwidths in the two directions are designed to be in the same ratio as the flows in the two directions.

There are computer programs that do the computations for maximum bandwidth that are commonly used by traffic engineers, as mentioned earlier. Thus it is not worthwhile to present an elaborate manual technique here. However, to get a feel for the basic technique and trade-offs, a small "by hand" example is shown.

Refer to Figure 26.17, which shows four signals and decent progression in both directions. For purposes of illustration, assume it is given that a signal with 50-50 split must be located midway between intersections 2 and 3. Figure 26.18 shows the possible effect of inserting the new signal into the system. It would appear there is no way to include this signal without destroying one or the other bandwidth, or cutting both in half.

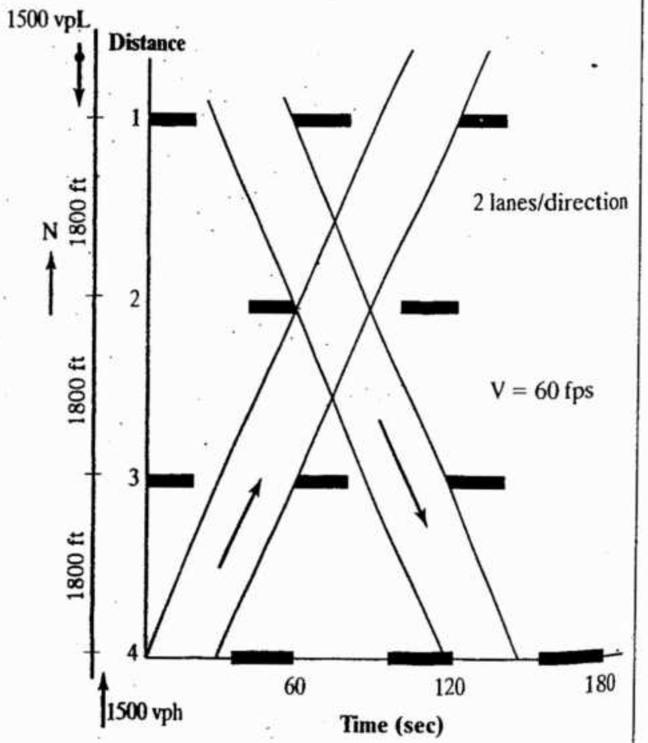


Figure 26.17: Case Study: Four Intersections with Good Two-Way Progression

---

[1] This is not to say that Synchro produces the identical answers as the *HCM* in all cases. We are simply reporting the state of the practice, namely that a number of states treat the Synchro outputs as having the same weight as if they came from the *HCM*.

[2] SimTraffic is available bundled with Synchro, and Synchro can directly feed SimTraffic. However, they are fundamentally different modeling approaches and can produce different estimates of level of service.

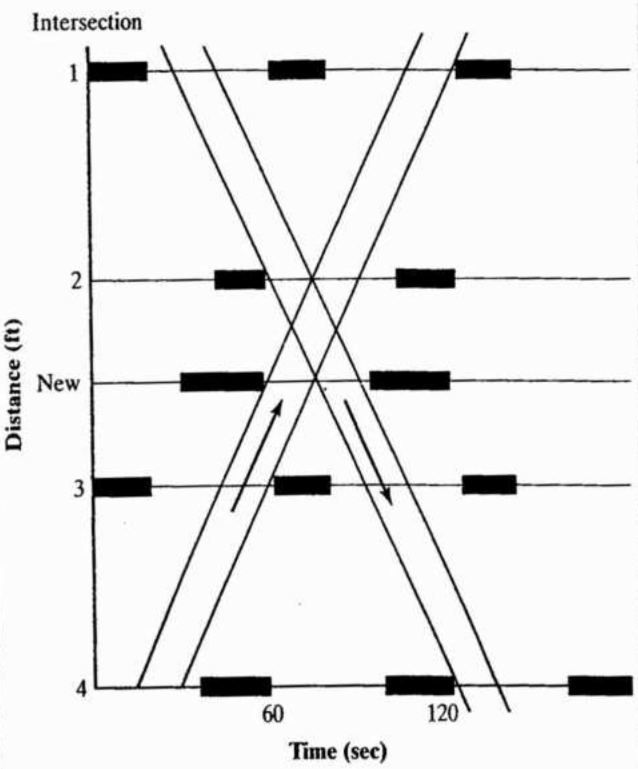


Figure 26.18: Case Study: The Effect of Inserting a New Signal into the System

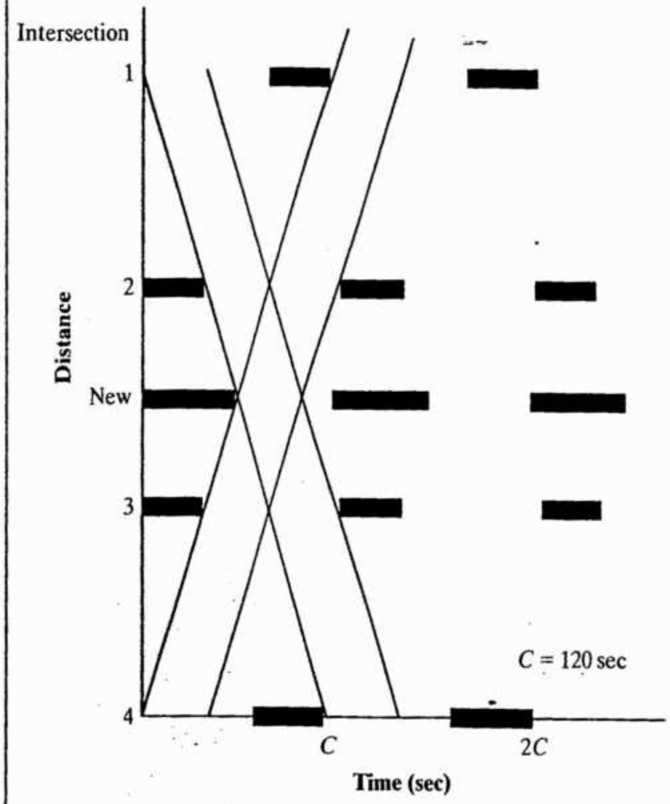


Figure 26.19: Case Study: A Solution with $C = 120$ Seconds

To solve this problem, the engineer must move the offsets around until a more satisfactory timing plan develops. A change in cycle length may even be required.

Note that the northbound vehicle takes $3600/60 = 60$ s to travel from Intersection 4 to Intersection 2, or—given $C = 60$ seconds—one cycle length. If the cycle length had been $C = 120$ seconds, the vehicle would have arrived at Intersection 2 at $C/2$, or half the cycle length. If we try the 120-second cycle length, then a solution presents itself.

Figure 26.19 shows one solution to the problem, for $C = 120$ seconds, which has a 40-second bandwidth in both directions for an efficiency of 33%. The 40-second bandwidth can handle $(40/2.0) = 20$ vehicles per lane per cycle. Thus if the demand volume is greater than $3{,}600(40)(2)(2.0)(120) = 1{,}200$ veh/h, then it will not be possible to process the vehicles nonstop through the system.

As indicated in the original information (see Figure 26.17), the northbound demand is 1500 veh/h. Thus there will be some difficulty in the form of excess vehicles in the platoon. They can enter the system but cannot pass Signal 2 nonstop. They will be "chopped off" the end of the platoon and be queued vehicles in the next cycle. They will be released in the early part of the cycle and arrive at Signal 1 at the beginning of red. Figure 26.20 illustrates this, showing that these vehicles then disturb the next northbound through platoon.

Note the Figure 26.20 illustrates the limitation of the bandwidth approach when internal queuing arises, disrupting the bandwidth. The figure also shows the southbound platoon pattern, suggesting that the demand of exactly 1,200 veh/h might give rise to minor problems of the same sort at Signals 3 and 4.

If one were to continue a trial-and-error attempt at a good solution, it should be noted that:

- If the green initiation at Intersection 1 comes earlier in order to help the main northbound platoon avoid the queued vehicles, the southbound platoon is released sooner and gets stopped or disrupted at Intersection 2.
- Likewise, shifting the green at Intersection 2 cannot help the northbound progression without harming the southbound progression.
- Nor can shifting the green at Intersection 3 help the southbound progression without harming the northbound progression.

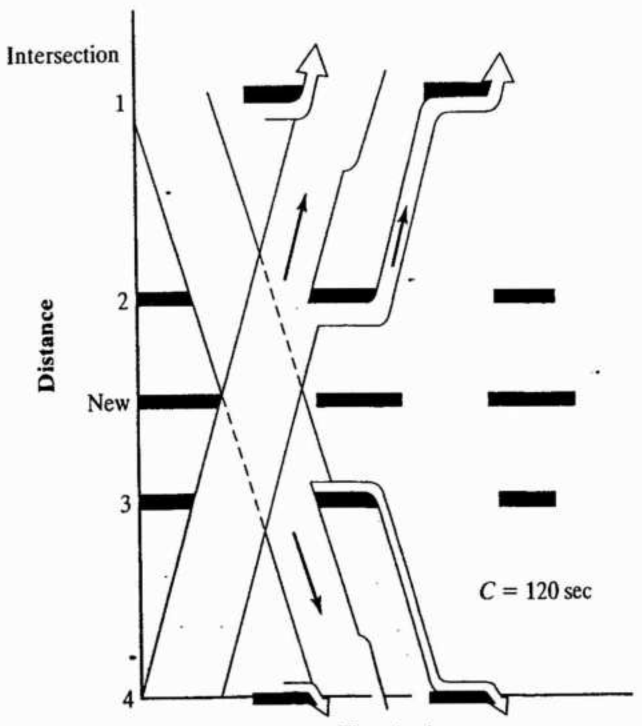


Figure 26.20: Case Study: Effect on Platoons with Demand Volume 1,500 Veh/h

- Some green can be taken from the side street and given to the main street.
- It is also possible that the engineer may decide to give the northbound platoon a more favorable bandwidth because of its larger demand volume.

This illustration showed insights that can be gained by simple inspection of a time-space diagram, using the concepts of bandwidth, efficiency, and an upper bound on demand volume that can be handled nonstop.

26.6 Common Types of Progression

26.6.1 Progression Terminology

The sole purpose of this section is to introduce some common terminology:

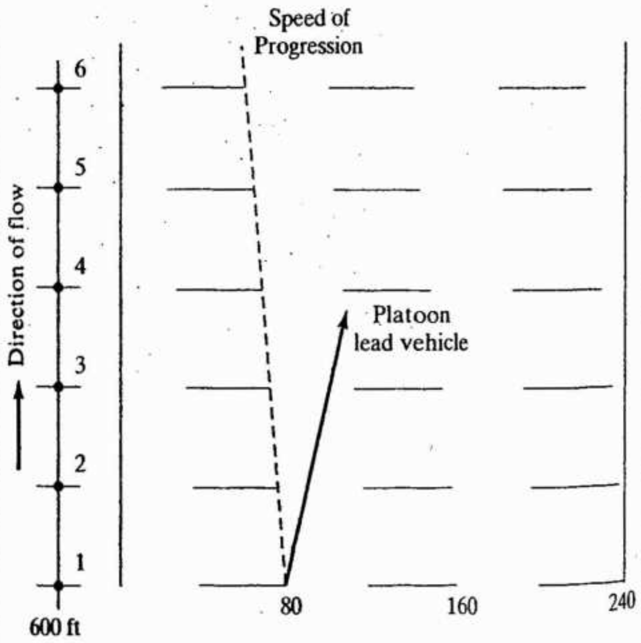
- Simple progression
- Forward progression
- Flexible progression
- Reverse progression

Simple progression is the name given to the progression in which all signals are set so that a vehicle released from the first intersection will arrive at all downstream intersections just as the signals at those intersections initiate green. That is, each offset is the ideal offset, set by Equation 26-4 with zero queue. Of necessity, simple progressions are effective only on one-way streets or on two-way streets on which the reverse flow is small or neglected.

Because the simple progression results in a green wave that advances with the vehicles, it is often called a *forward progression,* taking its name from the visual image of the advance of the green down the street.

It may happen that the simple progression is revised two or more times in a day, so as to conform to the direction of the major flow, or to the flow level (because the desired platoon speed can vary with traffic demand). In this case, the scheme may be referred to as a *flexible progression.*

Under certain circumstances, the internal queues are sufficiently large that the ideal offset is negative; that is, the downstream signal must turn green before the upstream signal, to allow sufficient time for the queue to start moving before the arrival of the platoon. Figure 26.21 has link lengths of 600 feet, platoon speeds of 60 ft/s, and internal queues averaging 7 vehicles per lane at each intersection. The visual image of such a pattern is of the green marching upstream, toward the drivers in the platoon. Thus it is referred to as a *reverse progression.*



$$t_{\text{ideal}} = \frac{600}{60} - (7)(2) = -4 \text{ sec}$$

Figure 26.21: Illustration of a Reverse Progression

Figure 26.21 also illustrates one of the unfortunate realities of so many internal queued vehicles: The platoon's lead vehicle only gets to Signal 4 before encountering a red indication. As he platoon passes Signal 3, there are only 12 seconds of green o accommodate it, resulting in all vehicles beyond the sixth i.e., 12/2 = 6) being cut off at Signal 3.

In the next several sections, common progression systems that can work extremely effectively on two-way arterials and streets are presented. As you will see, these systems rely on having uniform block lengths and an appropriate relationship among block length, progression speed, and cycle length. Because achieving one of these progressions has major benefits, the traffic engineer may wish to set he system cycle length based on progression requirements, ntroducing design improvements at intersections where the system cycle length would not provide sufficient capacity. Rather than increase the system cycle length to accommodate he needs of a single intersection, redesign of the intersection should be attempted to provide additional capacity at the desired system cycle length.

26.6.2 The Alternate Progression

For certain uniform block lengths, and all intersections with a 50-50 split of effective green time, it is possible to select a feasible cycle length such that:

$$\frac{C}{2} = \frac{L}{S} \qquad (26\text{-}8)$$

where: C = cycle length, s
L = block length, ft
S = platoon speed, ft/s

n this situation, the progression of Figure 26.22 can be obtained. There is no limit to the number of signals that may be included in the progression.

The name for this pattern is derived from the "alternate" appearance of the signal displays: As the observer at Signal 1 ooks downstream, the signals alternate—red, green, red, green, and so forth.

The key to Equation 26-8 is that the ideal offset in either direction (with zero internal queues) is L/S. That is, the travel ime to each platoon is exactly half the cycle length, so that he two travel times add up to the cycle length.

The efficiency of an alternate system is 50% in each direction because all of the green is used in each direction. The bandwidth capacity for an alternate progression is found using Equation 26-3, and noting that the bandwidth, BW, is equal to half the cycle length, C. If a saturation headway of 2.0 s/veh is assumed, then:

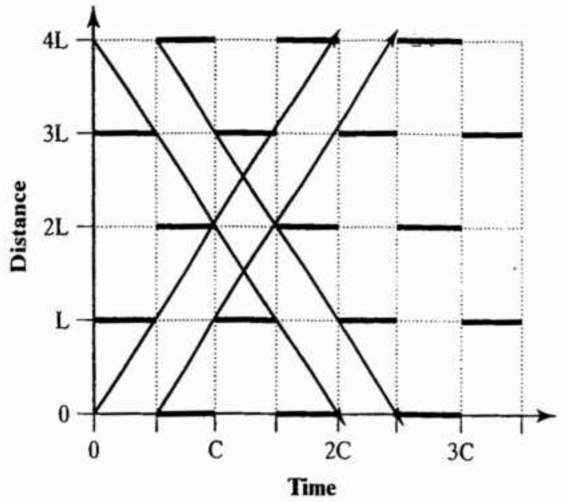


Figure 26.22: Alternate Progression Illustrated

$$c_{BW} = \frac{3600 * BW * NL}{h * C} = \frac{3600 * 0.5C * NL}{2.0 * C} = 900NL$$

where all terms are as previously defined. This is an approximation based on the assumed saturation headway of 2.0 s/veh. The actual saturation headway may be determined more accurately using the *Highway Capacity Manual* procedure for intersection analysis.

Note that if the splits are not 50-50 at some signals, then (1) if they favor the main street, they simply represent excess green, suited for accommodating miscellaneous vehicles, and (2) if they favor the side street, they reduce the bandwidths.

As a practical matter, note the range of the block lengths for which alternate patterns might occur. Using Equation 26-8, appropriate block lengths are computed for platoon speeds of 30 and 50 mi/h (that is, 45 and 75 ft/s), and cycle lengths of 60 and 90 seconds. The results are shown in Table 26.3. These

Table 26.3: Come Illustrative Combinations for Alternate Progression

Cycle Length (s)	Platoon Speed (fps)	Matching Block Length (ft)
60	45	1,350
60	75	2,250
90	45	2,025
90	75	3,375

results are illustrative; other combinations are clearly possible as well. All of these signal spacings imply a high-type arterial, often in a suburban setting.

26.6.3 The Double-Alternating Progression

For certain uniform block lengths with 50-50 splits, it is not possible to satisfy Equation 26-8, but it is possible to select a feasible cycle length such that:

$$\frac{C}{4} = \frac{L}{S} \tag{26-9}$$

In this situation, the progression illustrated in Figure 26.23 can be obtained.

The key is that the ideal offset in either direction (with zero internal queues) over *two* blocks is one half of a cycle length, so that two such travel times (one in each direction) add up to a cycle length. There is no limit to the number of signals that can be involved in this system, just as there was no limit with the alternate system.

The name of the pattern is derived from the "double alternate" appearance of the signal displays; that is, as the observer at Signal 1 looks downstream, the signals alternate in pairs—green, green, red, red, green, green, red, red, and so forth.

The efficiency of the double alternate signal system is 25% in each direction because only half of the green is used in each direction. The upper limit on the bandwidth capacity may be approximated by assuming a 2.0 s/veh saturation headway and noting that the BW is a quarter of C.

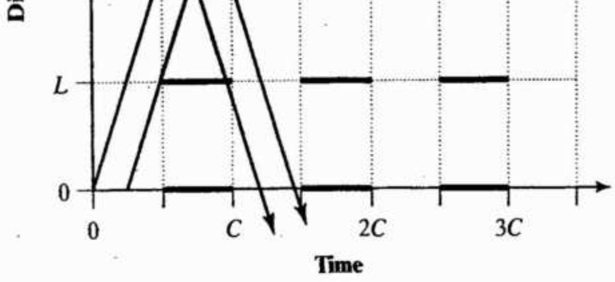


Figure 26.23: Double Alternate Progression Illustrated

Table 26.4: Illustrative Combinations for Double-Alternate Progression

Cycle Length (s)	Platoon Speed (fps)	Matching Block Length (ft)
60	45	675
60	75	1,125
90	45	1,012
90	75	1,688

As with the alternate system, if the splits are not 50-50 at some signals, then (1) if they favor the main street, they simply represent excess green, suited for accommodating miscellaneous vehicles, and (2) if they favor the side street, they reduce the bandwidths.

$$c_{BW} = \frac{3600 * BW * NL}{h * C} = \frac{3600 * 0.25C * NL}{2.0 * C} = 450NL$$

Table 26.4 shows some illustrative combinations of cycle length, platoon speed, and block lengths for which a double alternating progression would be appropriate. Other combinations, of course, are possible as well. Some of these signal spacings represent a high-type arterial. With the shorter cycle lengths, however, some urban facilities could also have the necessary block lengths.

26.6.4 The Simultaneous Progression

For very closely spaced signals, or for rather high vehicle speeds, it may be best to have all the signals turn green at the same time. This is called a simultaneous system because all the signals turn green simultaneously. Figure 26.24 illustrates a simultaneous progression.

The efficiency of a simultaneous system depends on the number of signals involved. For N signals:

$$EFF(\%) = \left[\frac{1}{2} - \frac{(N-1)*L}{S*C}\right]*100\% \tag{26-10}$$

For four signals with $L = 400$ feet, $C = 80$ seconds, and $S = 45$ ft/s, the efficiency is 16.7%. For the same number of signals with $L = 200$ feet, it is 33.3%.

Simultaneous systems are advantageous only under a limited number of special circumstances. The foremost of these special circumstances is very short block lengths. The simultaneous system has an additional advantage, however,

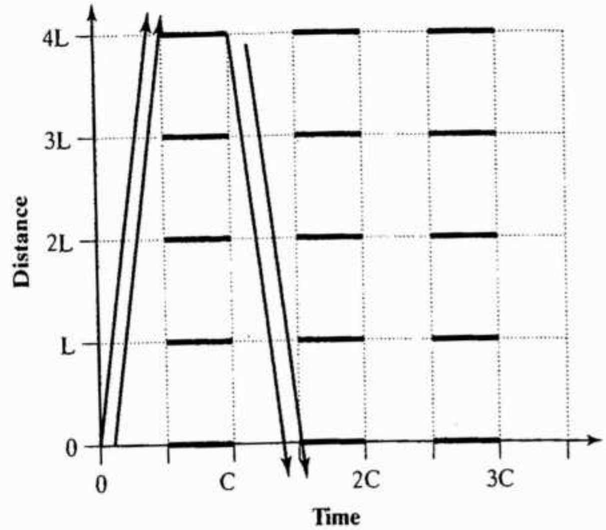


Figure 26.24: Simultaneous Progression Illustrated

that is not at all clear from a bandwidth analysis: Under very heavy flow conditions, it forestalls breakdown and spillback. This is so because (1) it allows for vehicle clearance time at the downstream intersection where queues inevitably exist during heavy flow, and (2) it cuts platoons off in a way that generally prevents blockage of intersections. This works to the advantage of cross traffic. Specific plans for controlling spillback under heavy traffic conditions are discussed in Chapter 27.

26.6.5 Insights from the Importance of Signal Spacing and Cycle Length

It is now clear that:

- All progressions have their roots in the desire for ideal offsets.
- For certain combinations of cycle length, block length, and platoon speed, some very satisfactory two-way progressions can be implemented.
- Other progressions can be designed to suit individual cases, using the concept of ideal offset and queue clearance, trial-and-error bandwidth based approaches, or computer-based algorithms.

A logical first step in approaching a system is simply to ride the system and inspect it. As you sit at one signal, do you see the downstream signal green, but with no vehicles being processed? Do you arrive at signals that have standing queues but were not timed to get them moving before your platoon arrived? Do you arrive on the red at some signals? Is the flow in the other direction significant, or is the traffic really a one-way pattern, even if the streets are two way?

It is very useful to sketch out how much of the system can be thought of as an "open tree" of one-way links. This can be done with a local map and an appreciation of the traffic flow patterns. A distinction should be made among:

- Streets that are one way.
- Streets that can be treated as one way, due to the actual or desired flow patterns.
- Streets that must be treated as two way.
- Larger grids in which streets (one way and two way) interact because they form unavoidable "closed trees" and are each important in that they cannot be ignored for the sake of establishing a "master grid" that is an open tree.
- Smaller grids in which issue is not coordination but rather local land access and circulation, so that they can be treated differently. Downtown grids may well fall into the latter category, at least in some cases.

The next most important issue is the cycle length dictated by the signal spacing and platoon speed. Attention must focus on the combination of cycle length, block length, and platoon speed, as shown earlier in this chapter.

Figure 26.25 shows the three progressions of the preceding sections—alternate, double alternate, and simultaneous—on the same scale. The basic "message" is that as the average signal spacing decreases, the type of progression best suited to the task changes.

Figure 26.26 illustrates a hypothetical arterial that comes from a low-density suburban environment with a larger signal spacing, into the outlying area of a city, and finally passes through one of the city's CBDs. As the arterial changes, the progression used may also be changed, to suit the dimensions.[3]

Note that the basic lesson here is that a system can sometimes be best handled by breaking it up into several smaller systems. This can be done with good effect on even smaller systems, such as 10 consecutive signals, of which a

---

[3] Of course, if the flow is highly directional—as may well be from the suburbs in the morning—then these suggestions are superseded by the simple expedient of treating the streets as one-way streets and imposing a simple forward progression, with queue clearance if needed.

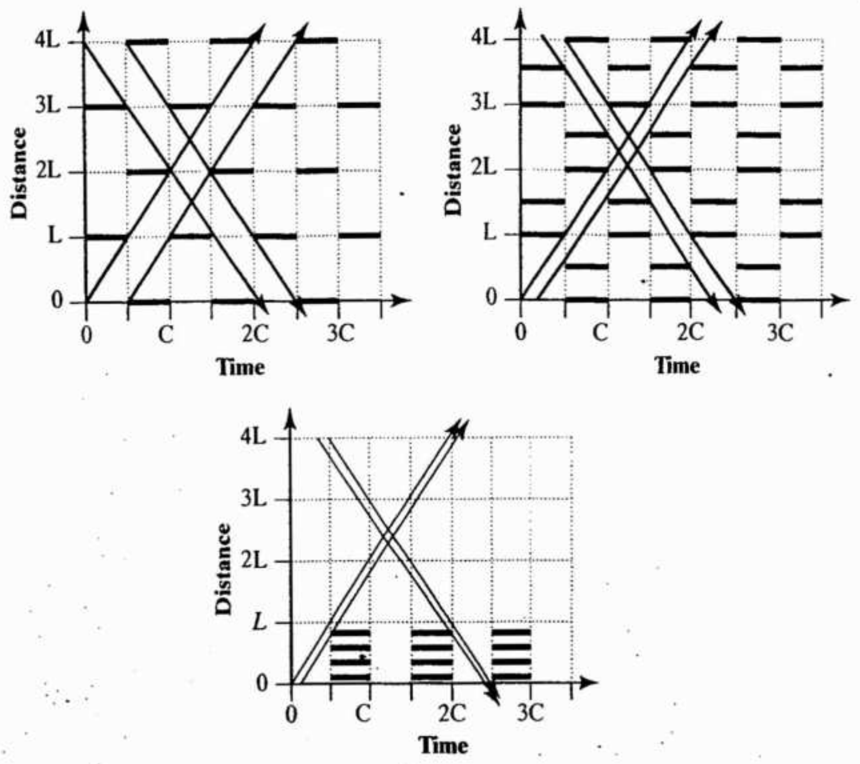


Figure 26.25: Comparison of Scales on Which Standard Patterns Are Used

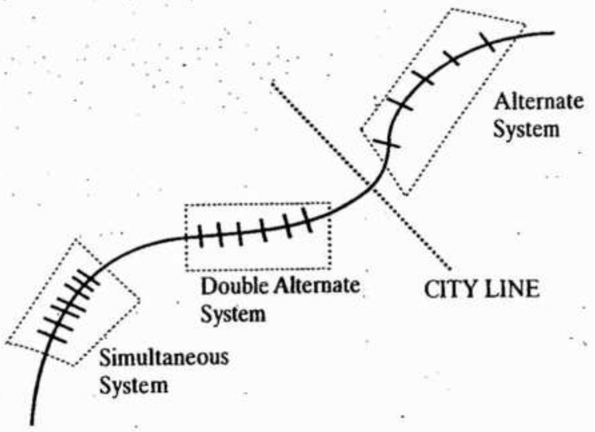


Figure 26.26: Hypothetical Use of Several Patterns Along the Same Arterial

contiguous 6 are spaced uniformly and the other 4 also uniformly, but at different block lengths. Note that to the extent that block lengths do no exist perfectly uniformly, these plans can serve as a basis from which adaptations can be made. Note also that the suitability of the cycle length has been significant. It is often amazing how often the cycle length is poorly set for system purposes.

26.7 Software for Doing Signal Progression

The purpose of this section is to *illustrate* the use of software that is commercially available to provide ease of computation in determining progressions. You cannot expect to become an effective user of either program from these brief illustrations. But it is common (and advisable) that such computer programs be used in at least the second course in a traffic engineering sequence. The user manuals and lecturer presentation, focused on case study or project applications, can provide the needed background. (In our own courses, we also depend on the students climbing the learning curve quickly as part of a project-based course.)

The first program that is illustrated, TruTraffic TS/PP, is especially interesting because it is based on the bandwidth approach that is considered "dated" by those who have used the programs that focus on delay-optimized signal settings that take into account internal queueing. But the simple fact is that bandwidth-based solutions are less data intensive, very suitable in many applications, and easy to manage.

The second program that is illustrated (Synchro) is the most commonly used delay- or stops-based optimization program in the United States. As already noted, a number of states accept its output as equivalent to *HCM* results, at least for sets of signals on an arterial or network. Synchro has an imbedded macro model of flow profiles that it uses to estimate delays and stops as it iteratively finds a solution that minimizes an overall "objective function" that is expressed in terms of stops and delays.

26.7.1 Bandwidth-Based Solutions

Bandwidth-based solutions remain an important tool for traffic engineers, particularly for off-peak periods when demand is relatively small and/or when flow is highly directional. Although prior examples in this chapter emphasized achieving windows of green along the entire arterial, bandwidth solutions also can be used to move platoons along relatively long arterials in a set of bandwidths, with the breaks between bandwidths occurring where it is logical or suitable to stop and re-form platoons—for instance, just upstream of a set of closely spaced signals, so that the platoons that might overflow the short block spacings are not stopped in that section of the arterial. Bandwidth solutions are also used effectively to discourage speeding, encourage adherence to the speed limit, and identify green that can be allocated to increased pedestrian walking times [8].

Figure 26.27 shows the output of TruTraffic TS/PP for the arterial addressed in Figure 26.4, considering it as a two-way arterial on which we wish to achieve equal bandwidths (16 seconds) in the two directions if at all possible.

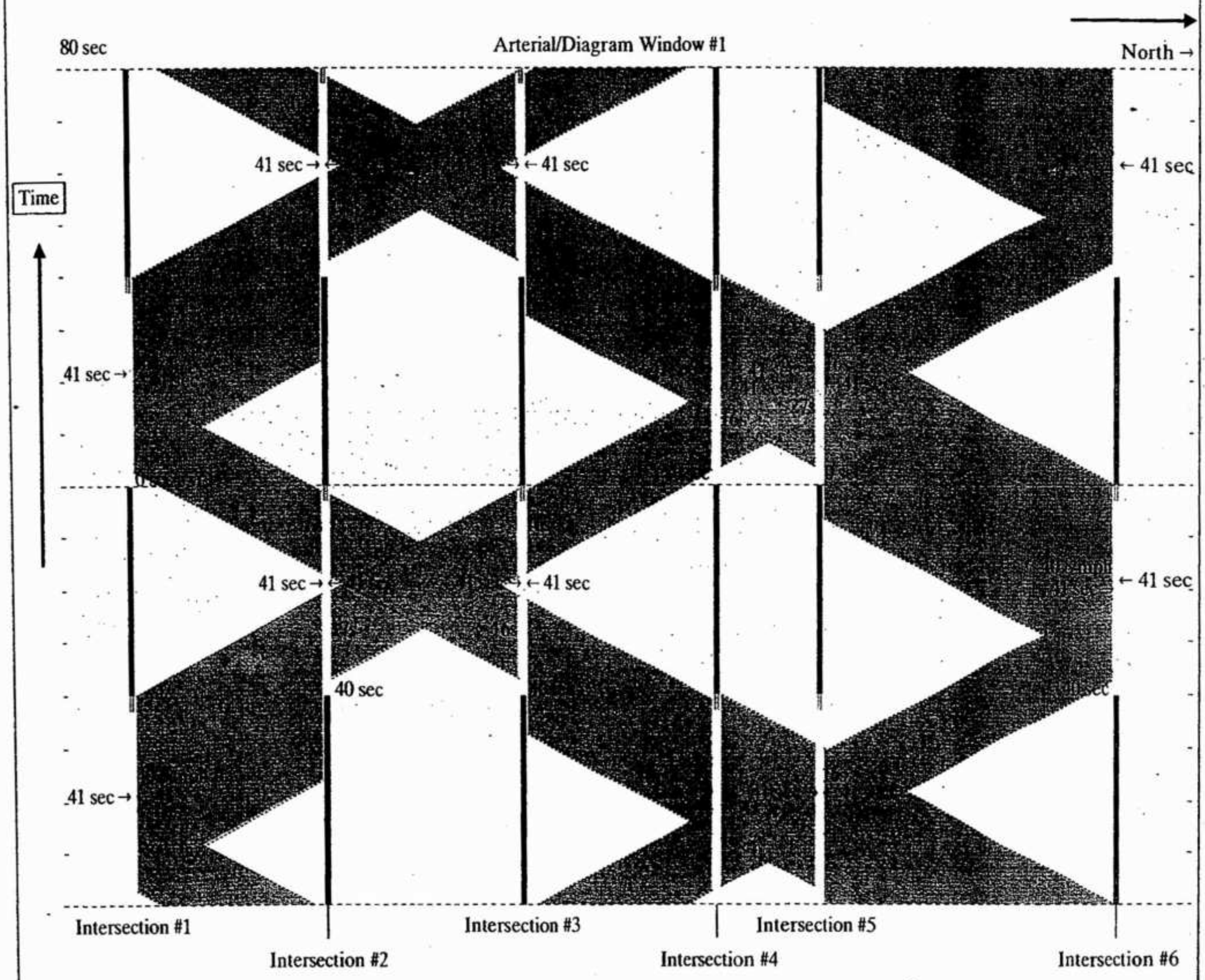


Figure 26.27: A Solution for Equal Bandwidths in Both Directions, Using TruTraffic TS/PP (Figure 26.4 Arterial)

Because of the way in which the output is printed (generally, on longer paper), the directions are somewhat reversed from the prior time-space diagrams. In this chapter arrows have been added, to emphasize the directions in which time and distance both increase. From this figure, note that:

- The bandwidths are shown in seconds, and partial bandwidths are shown when breaks are required. The "message" in such cases is that the platooned vehicles can travel only so many blocks without stopping;
- To visualize a vehicle going faster than the design speed of the bandwidth, it may be easiest for the reader to turn the display so that time is on the bottom, given the reversal in the figure.

Figure 26.28 shows the same illustration, except the target is to get a NB bandwidth of at least 20 seconds, with "as good as possible" in the SB.

The bandwidths shown in Figure 26.28 vary as opportunity presents itself, indicating that the band can be wider in some segments along the arterial than in others; this is shown in both

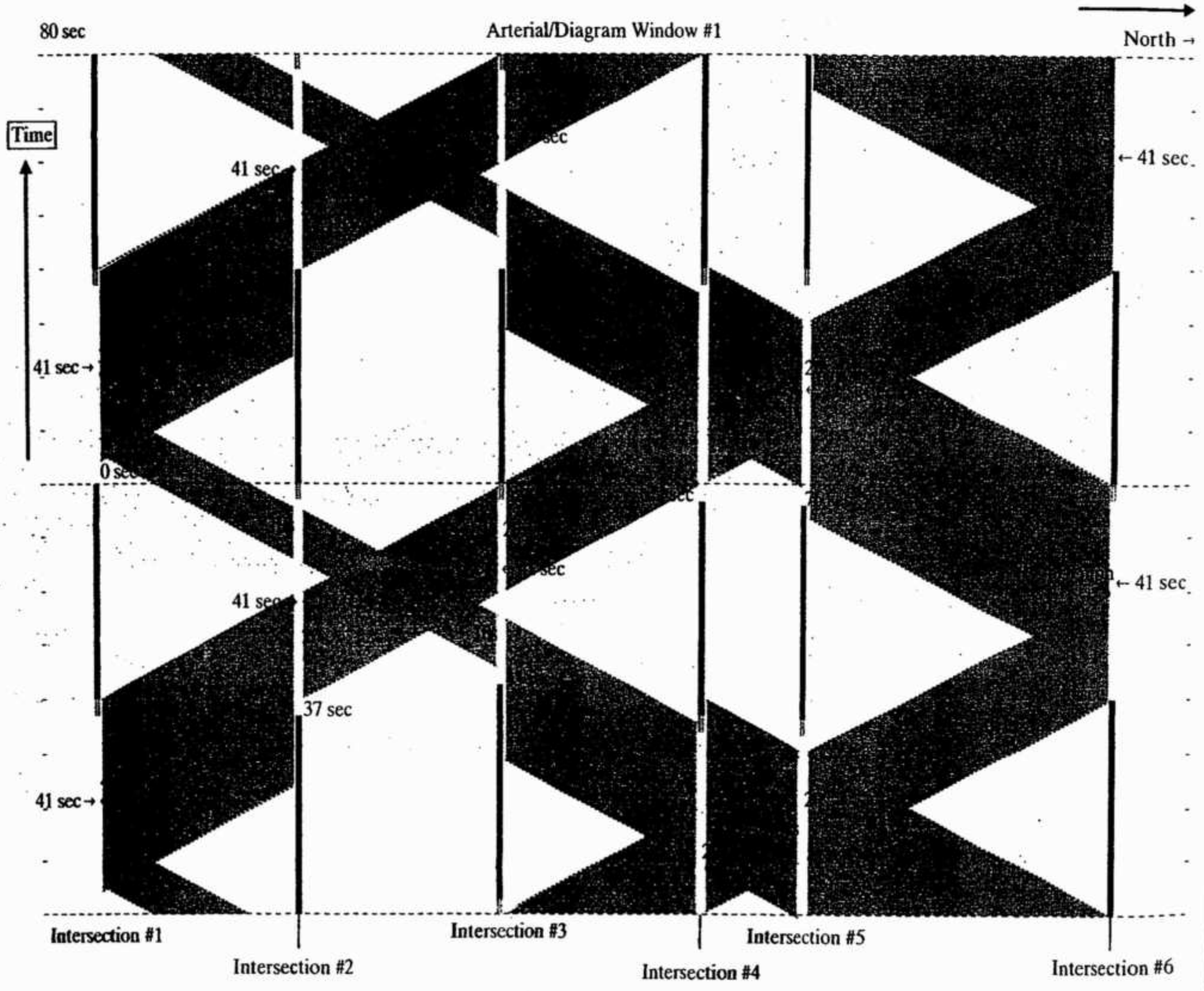


Figure 26.28: A Solution for Preferred NB Bandwidth, Two-Way Street, Using TruTraffic TS/PP (Figure 26-4 Arterial)

Note: The associated text describes the indicated items.

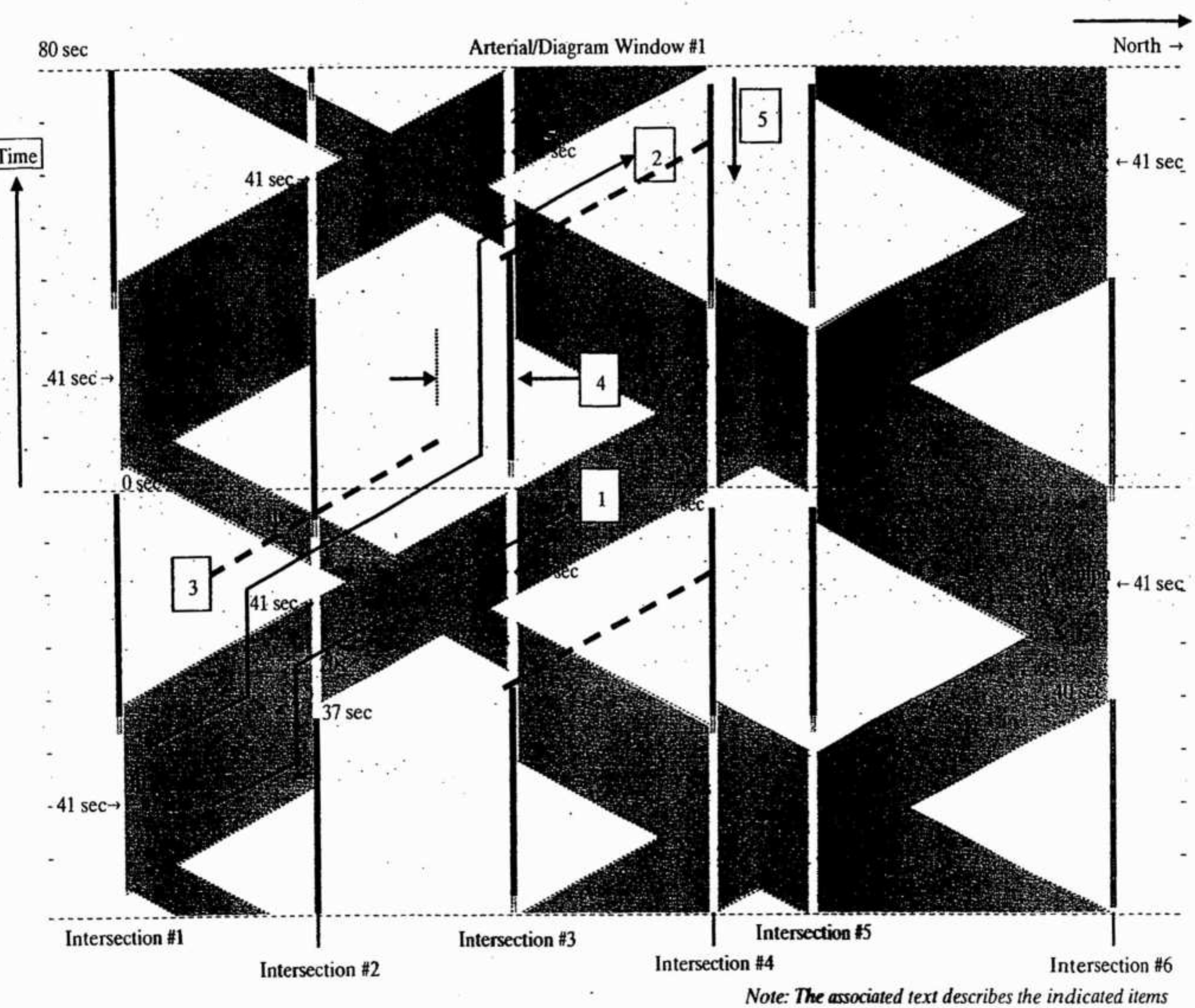
directions. TruTraffic also allows the user to turn this and other features "on" or "off," so that

a. One can emphasize a fixed bandwidth that can move the entire distance (or some large subsegment), rather than that shown in Figure 26.28; for instance, in the northbound direction, the band can be reduced to 20 seconds in Figure 26.28, emphasizing the best band that can get through with no one stopping.
b. One can also restart the bands at any intersection, designating a new start point.

If the bandwidth in "a" is capable of handling the existing traffic, this approach has meaning. Referring to Equation 26-3 and using $BW = 20$ seconds and $C = 80$ seconds, one can compute $c_{BW} = 3{,}600(20)(1)/(80)(2.1) = 429$ vph on a *per lane* basis.

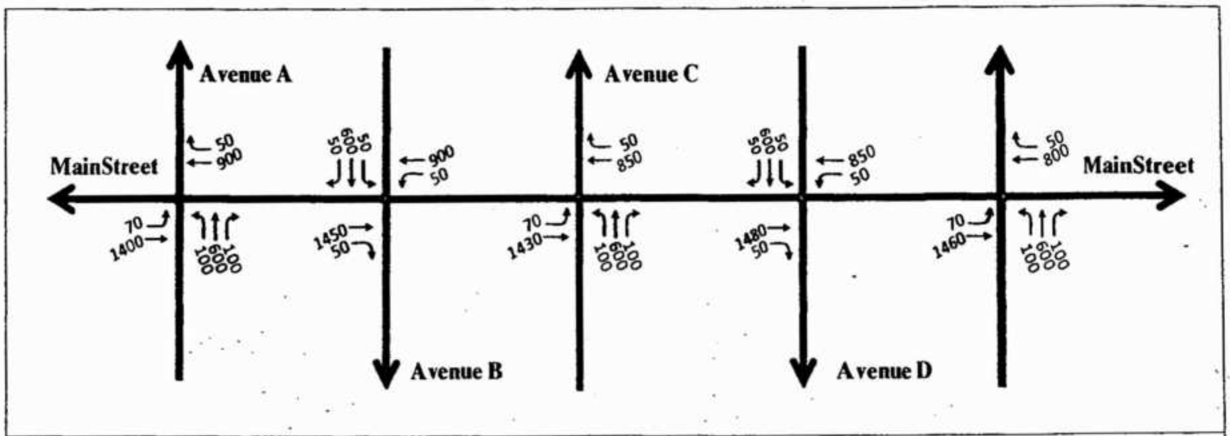
The actual trajectories along the arterial depend on where the platoon started, turns in and out, and some dispersion, even when the demand is less than c_{BW}. In practice, one has to take into account that when northbound "early greens" are given to allow southbound bandwidth, northbound speeding may be allowed because drivers may perceive short-term gain by moving faster.

When the demand exceeds c_{BW} and TruTraffic or such tool is still used, more has to be taken into account. For instance, refer to Figure 26.29: When the northbound platoon fills most of the bandwidth from Intersection 1 to Intersection 2,



Note: The associated text describes the indicated items

Figure 26.29: Some of the Considerations in Using TruTraffic to Design Bandwidths



- All the intersections are spaced at 1500 feet
- 2 lanes per direction on Main St
- left turn bays on Main St, 250 ft, both directions
- RTOR prohibited everywhere
- 60 fps free flow speed, Main Street
- All avenues have 2 moving lanes, but are one-way streets
- PHF = 0.91
- Minimum ped crossing times = 17 sec across side streets, 30 sec Main St

Figure 26.30: Inputs for Illustrative Synchro Case

- The lead part of the platoon must stop at Intersection 2 and then continue as shown by ___.
- The second part of the platoon, designated by ___, is queued temporarily at Intersection 2 and then "clipped" and stopped at Intersection 3. It can extend to the dashed line shown by ___, depending on the demand. *Further, it then becomes the lead part of the next platoon, as shown by the dashed line between Intersections 3 and 4.*
- Notice that because this pattern is repetitive, the same reality happens in each and every cycle, so that the same thing happened *before* path __ as shown by the earlier (i.e., lower) dashed line. Hence, the vehicles in the path ___ were not actually the lead vehicles *and they are temporarily queued* at Intersection 4. This concern or ripple effect exists throughout the system, in both directions;
- At every intersection, but most notably ones close together, one must assure that the queues shown by ___ can fit into the block. In some cases, it is clear that they do, by inspection. In other cases, the width of the stopped section of the band has to be used to estimate the queue size, convert it to an expected length, and see if it approaches or exceeds the block length.[4]
- Given the knowledge of the presence of vehicles as shown by ___, one may wish to consider moving the green initiation at ___ as shown, although this has an adverse effect on the southbound traffic.

The TruTraffic tool has the capability to show the added widths described here, and it is useful in such cases as just described. It is also quite possible to use such a tool (really, the underlying bandwidth principle) effectively even when *apparent* bandwidth solution leads to such displacements.[5] At some point however, delay/stopped-based optimization tools may be easier to use and more appropriate, even with rough estimates of demand (i.e., without all the detail shown in the next section).

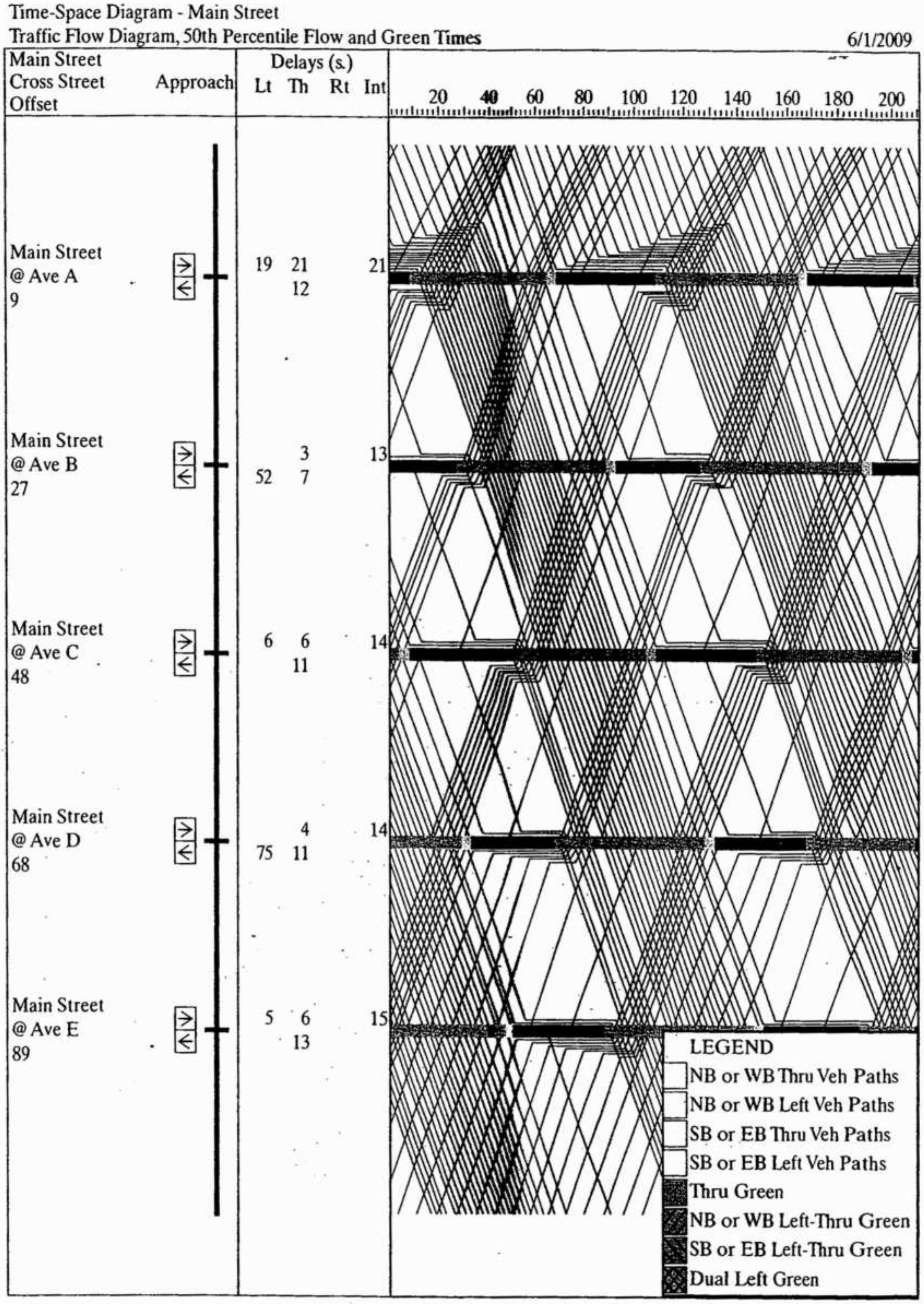
26.7.2 Synchro

Consider the inputs as shown in Figure 26.30. Figure 26.31 shows an illustrative Synchro solution for a cycle length of 100 seconds; the user is provided with choices over a range of

---

[4] The practical result is that very closely spaced intersections operate simultaneously, so that the platoon moves without stopping in that short distance.

[5] Notice that the demand does not have to exceed *or even approach* the capacity of the *signals*. The issue at hand is the demand that can be accommodated by c_{BW}. Hence solutions that involve narrow bands over considerable lengths might look good initially but can induce problems when demand exceeds



Time-Space Diagram - Main Street
Traffic Flow Diagram, 50th Percentile Flow and Green Times
6/1/2009

Main Street Cross Street Offset	Approach	Delays (s.) Lt	Th	Rt	Int
Main Street @ Ave A 9	→	19	21		21
	←		12		
Main Street @ Ave B 27	→		3		13
	←	52	7		
Main Street @ Ave C 48	→	6	6		14
	←		11		
Main Street @ Ave D 68	→		4		14
	←	75	11		
Main Street @ Ave E 89	→	5	6		15
	←		13		

Figure 26.31: Sample Synchro Time-Space Diagram for Illustrative Case

user-specified cycle lengths and selects from them. Table 26.5 shows sections of the standard Synchro output tables.

Although this illustration will not be discussed in detail at this point, it will be used again in Chapter 30.

In practice, engineers often fine-tune the optimized signal offsets during the field installation, to adapt to field conditions (that is, how drivers actually arrive, etc.).

26.8 Closing Comments

This chapter has introduced the basic considerations and concepts of signal coordination for undersaturated flows on one-way and two-way arterials, and in networks.

Although the commercially available computer programs are extremely efficient and cost-effective tools for developing signal timing plans (including offsets, for coordination), you will find that thoughtful reflection on the basic principles will help not only at arriving at good solutions, but also in checking the outputs of the programs. For instance, if one is fortunate enough to have block lengths (L) and desired speeds (S) that happen to approximate a half cycle length ($L/S \sim C/2$), one can expect the solution to look like a standard alternative progression. If it does not, then a significant number of turn-ins or internally generated vehicles (e.g., from a midblock parking lot) may be skewing the offsets, so that queues are cleared and overall delay/stops minimized. But if this potential explanation is absent, another explanation might be that the user made an error with the inputs.

There are now excellent Web sites that can serve as the basis for keeping up to date and aware of the trends related to signal timing and coordination. For instance, the TRB Traffic Signal Systems Committee has its Web site at http://www.signalsystems.org.vt.edu/. FHWA's Office of Operations Web site on publications at http://www.ops.fhwa.dot.gov/publications/publications.htm is an excellent source of key material, much of which can be downloaded. The *Traffic Signal*

Table 26.5: Segments of Synchro Output, Related to Illustrative Case

Detailed Measures of Effectiveness

3. Main Street & Ave A

Volume (vph)	1470	950	800	3220
Control Delay / Veh (s/v)	21	12	31	21
Queue Delay / Veh (s/v)	0	0	0	0
Total Delay / Veh (s/v)	21	12	31	21
Total Delay (hr)	8	3	7	18
Stops / Veh	0.74	0.58	0.83	0.72
Stops (#)	1085	555	665	2305
Average Speed (mph)	19	28	8	19
Total Travel Time (hr)	16	10	9	34
Distance Traveled (mi)	306	270	69	644
Fuel Consumed (gal)	28	18	14	60
Fuel Economy (mpg)	10.9	15.4	4.8	10.7
CO Emissions (kg)	1.97	1.23	1.01	4.20
NOx Emissions (kg)	0.38	0.24	0.20	0.82
VOC Emissions (kg)	0.46	0.28	0.23	0.97
Unserved Vehicles (#)	0	0	0	0
Vehicles in dilemma zone (#)	70	64	40	174

Arterial Level of Service: EB Main Street

							Arterial	Arterial LOS
Ave A	II	41	23.9	20.6	44.5	0.21	16.8	E
Ave B	II	41	29.0	3.3	32.3	0.28	31.7	B
Ave C	II	41	29.0	6.2	35.2	0.28	29.1	B
Ave D	II	41	29.0	3.6	32.6	0.28	31.4	B
Ave E	II	41	29.0	5.7	34.7	0.28	29.5	B
Total	II		139.9	39.4	179.3	1.34	27.0	C

Table 26.5: Segments of Synchro Output, Related to Illustrative Case (*Continued*)

HCM Signalized Intersection Capacity Analysis
3. Main Street & Ave A 6/1/2009

	↗	→	↘	↙	←	↖	↖	↑	↗	↘	↓	↙
Movement	EBL	EBT	EBR	WBL	WBT	WBR	NBL	NBT	NBR	SBL	SBT	SBR
Lane Configurations	↰	↑↑			↑↱			↰↱				
Ideal Flow (vphpl)	1900	1900	1900	1900	1900	1900	1900	1900	1900	1900	1900	1900
Total Lost time (s)	4.0	4.0			4.0			4.0				
Lane Util. Factor	1.00	0.95			0.95			0.95				
Frt	1.00	1.00			0.99			0.98				
Flt Protected	0.95	1.00			1.00			0.99				
Satd. Flow (prot)	1770	3539			3511			3451				
Flt Permitted	0.20	1.00			1.00			0.99				
Satd. Flow (perm)	371	3539			3511			3451				
Volume (vph)	70	1400	0	0	900	50	100	600	100	0	0	0
Peak-hour factor, PHF	0.91	0.91	0.91	0.91	0.91	0.91	0.91	0.91	0.91	0.91	0.91	0.91
Adj. Flow (vph)	77	1538	0	0	989	55	110	659	110	0	0	0
RTOR Reduction (vph)	0	0	0	0	0	0	0	0	0	0	0	0
Lane Group Flow (vph)	77	1538	0	0	1044	0	0	879	0	0	0	0
Turn Type	Perm						Perm					
Protected Phases		4			8			2				
Permitted Phases	4						2					
Actuated Green, G (s)	56.0	56.0			56.0			36.0				
Effective Green, g (s)	56.0	56.0			56.0			36.0				
Actuated g/C Ratio	0.56	0.56			0.56			0.36				
Clearance Time (s)	4.0	4.0			4.0			4.0				
Lane Grp Cap (vph)	208	1982			1966			1242				
v/s Ratio Prot		c0.43			0.30							
v/s Ratio Perm	0.21							0.25				
v/s Ratio	0.37	0.78			0.53			0.71				
Uniform Delay, d1	12.2	17.1			13.8			27.5				
Progression Factor	1.00	1.00			0.76			1.00				
Incremental Delay, d2	5.0	3.1			0.9			3.4				
Delay (s)	17.2	20.2			11.4			30.9				
Level of Service	B	C			B			C				
Approach Delay (s)		20.0			11.4			30.9			0.0	
Approach LOS		C			B			C			A	

Intersection Summary			
HCM Average Control Delay	20.2	HCM Level of Service	C
HCM Volume to Capacity ratio	0.75		
Actuated Cycle Length (s)	100.0	Sum of lost time (s)	8.0
Intersection Capacity Utilization	68.0%	ICU Level of Service	C
Analysis Period (min)	15		

c Critical Lane Group

Timing Manual is available at http://www.ops.fhwa.dot.gov/publications/fhwahop08024/index.htm.

References

1. *PASSER II-90 Microcomputer User's Guide,* distributed by the McTrans Center, Gainsville, FL, 1991.
2. Bullock, G., "Tru-Traffic TS/PP Version 7.0 Time-Space/Platoon-Progression Diagram Generator: Users Manual," April 2007, www.tsppd.com.
3. SYNCHRO Studio 7, Synchro plus SimTraffic and 3D Viewer User Guide, Trafficware Ltd., June 2006, www.trafficware.com.
4. *Highway Capacity Manual,* Transportation Research Board, Washington, DC, 2000.
5. TRANSYT-7F Release 11, Users Guide, distributed by the McTrans Center, University of Florida, 2008, http://mctrans.ce.ufl.edu/featured/TRANSYT-7F/.
6. HCS+ Release 5: Users Manual, distributed by the McTrans Center, University of Florida, 2008, http://mctrans.ce.ufl.edu/hcs/.

7. VISSIM 5.0 User's Guide, PTV, 2007, http://www.ptvamerica.com/vissim.html
8. AIMSUN Users Manual v6, TSS-Transport Simulation Systems, October 2008.
9. TSIS User's Guide, Version 6.0, CORSIM User Guide, distributed by the McTrans Center, University of Florida, 2008.
10. S-Paramics microscopic simulation software, distributed by SIAS Limited (http://www.sias.com/ng/sparamicshome/) and Quadstone Paramics distributed by Quadstone (http://www.paramics-online.com/).
11. Discussion with John Tipaldo, PhD PE, NYCDOT, May 2009, regarding off-peak use of bandwidth solutions.

Problems

26-1. Two signals are spaced at 1,000 feet on an urban arterial. It is desired to establish the offset between these two signals, considering only the primary flow in one direction. The desired progression speed is 40 mph. The cycle length is 60 seconds. Saturation headway may be taken as 2.0 sec/veh, and the startup lost time as 2.0 seconds.

(a) What is the ideal offset between the two intersections assuming that vehicles arriving at the upstream intersection are already in a progression (i.e., a moving platoon) at the initiation of the green?

(b) What is the ideal offset between the two intersections assuming that the upstream signal is the first in the progression (i.e., vehicles are starting from a standing queue)?

(c) What is the ideal offset assuming that an average queue of three vehicles per lane is expected at the downstream intersection at the initiation of the green? Assume the base conditions of part a.

(d) Consider the offset of part a. What is the resulting offset in the opposite (off-peak) direction? What impact will this have on traffic traveling in the opposite direction?

(e) Consider the offset of part a. If the progression speed was improperly estimated, and the actual desired speed of drivers was 45 mi/h, what impact would this have on the primary direction progression?

26-2. Consider the time-space diagram for this problem. For the signals shown:

(a) What is the NB progression speed?

(b) What is the NB bandwidth and bandwidth capacity? Assume a saturation headway of 2.0 s/veh.

(c) What is the bandwidth in the SB direction for the same desired speed as the NB progression speed? What is the SB bandwidth capacity for this situation?

(d) A new development introduces a major driveway that must be signalized between intersections 2 and 3. It requires 15 seconds of green out of the 60-second system cycle length. Assuming that you had complete flexibility as to the exact location of the new driveway, where would you place it? Why?

26-3. A downtown grid has equal block lengths of 750 feet along its primary arterial. It is desired to provide for a progression speed of 30 mi/h, providing equal service to traffic in both directions along the arterial.

(a) Would you suggest a alternate or a double-alternate progression scheme? Why?

(b) Assuming your answer to part a, what cycle length would you suggest? Why?

26-4. Refer to the figure below. Trace the lead NB vehicle through the system. Do the same for the lead SB vehicle. Use a platoon speed of 50 ft/s. Estimate the number of stops and the seconds of delay for each of these vehicles.

26-5. Refer to the figure below. Find the NB and the SB bandwidths (in seconds). Determine the efficiency of the system in each direction and the bandwidth capacity. There are three lanes in each direction. The progression speed is 50 ft/s.

26-6. (a) If vehicles are traveling at 60 fps on a suburban road, and the signals are 2,400 feet apart, what cycle length would you recommend? What offset would you recommend?

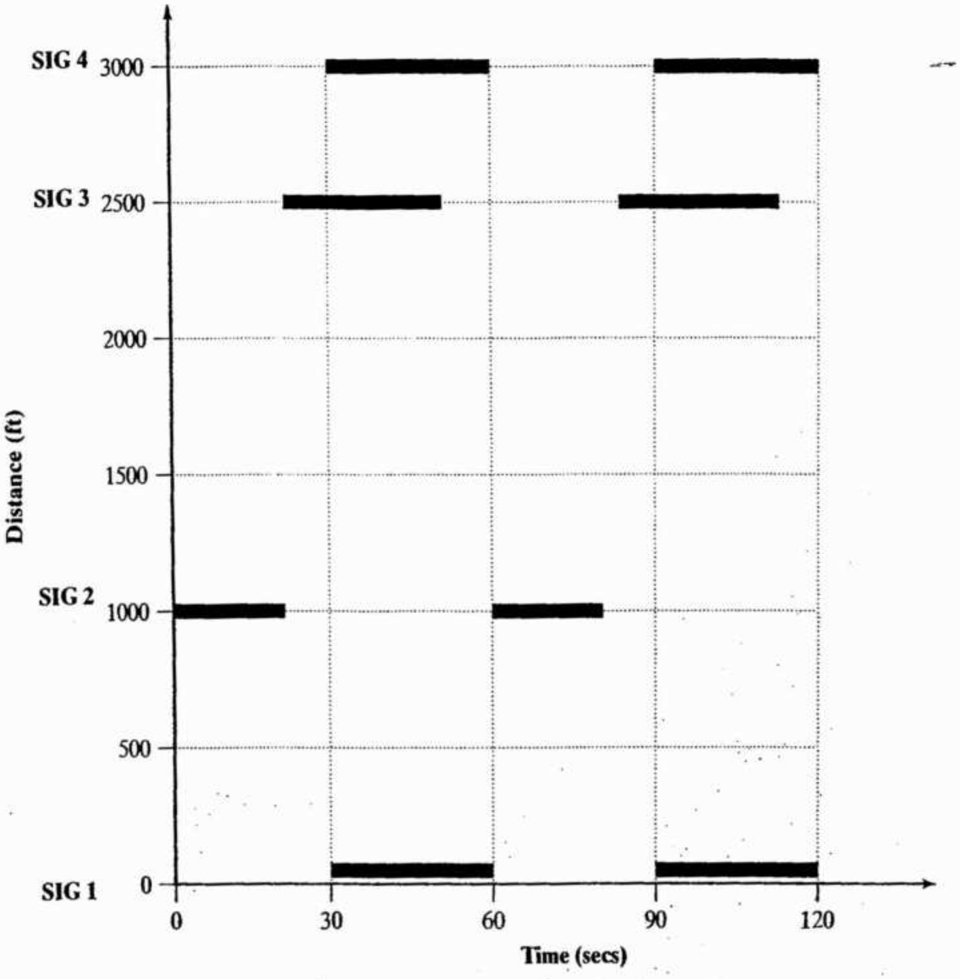
(b) If an unsignalized intersection is to be inserted at 600 feet from one of the signalized intersections, what would you recommend?

26-7. You have two intersections 3,000 feet apart and have achieved some success with a 50-50 split, 60-second cycle length, and simultaneous system.

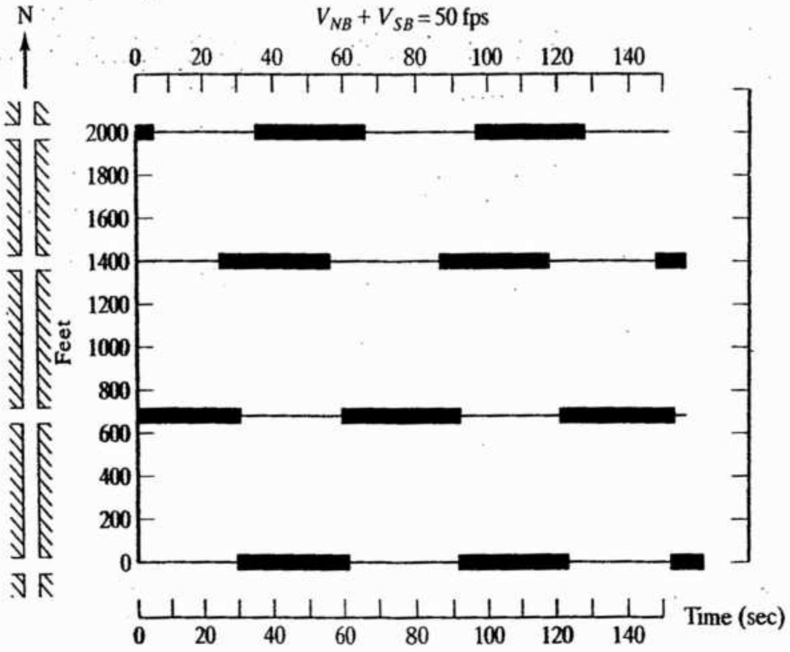
(a) Draw a time-space diagram and analyze the reason for your success.

(b) A developer who owns the property fronting on the first 2000 feet of the subject distance plans a major employment center. She plans a major driveway and asks your advice on its location. What is your recommendation, and why?

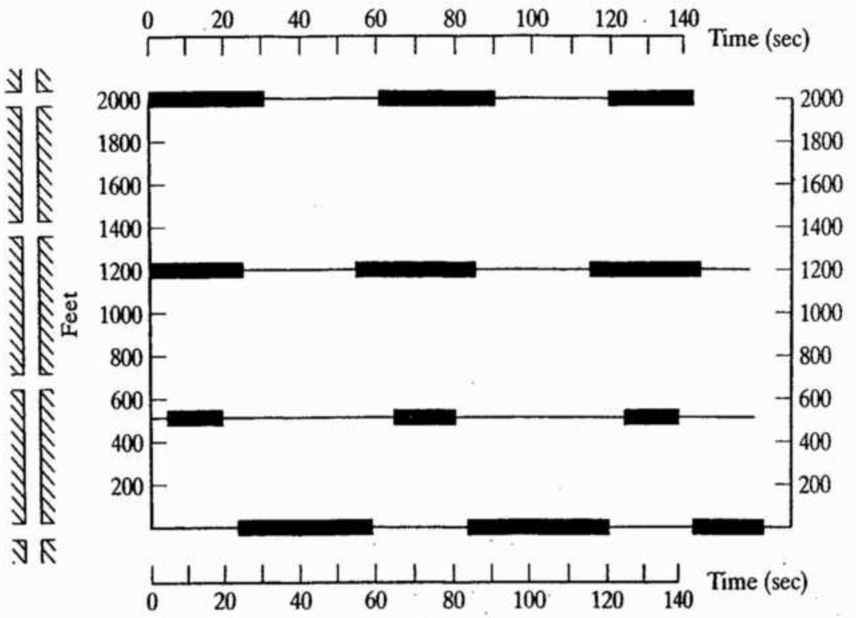
26-8. (a) Consider four intersections, spaced by 500 feet. The platoon speed is 40 ft/s. Recommend a set of offsets for the eastbound direction, considering only the eastbound traffic.



Time-Space Diagram for Problem 26-2



Time-Space Diagram for Problem 26-4



Time-Space Diagram for Problem 26-5

(b) If there are queues of three vehicles at each of the intersections, recommend a different set of offsets (if appropriate).

26-9. (a) Construct a time-space diagram for the following information and estimate the northbound bandwidth and efficiency for platoons going at 50 ft/s:

Signal No.	Offset (sec)	Cycle length	Split (MSG first)
6	16	60	50:50
5	16	60	60:40
4	28	60	60:40
3	28	60	60:40
2	24	60	50:50
1		60	60:40

All of the offsets are relative to the preceding signal. All signals are two phase. There are two lanes in each direction. All block lengths are 1,200 feet.

(b) Estimate the number of platooned vehicles that can be handled nonstop northbound and southbound.

26-10. For the situation in Problem 26-5, design a better timing plan (if possible), under two different assumptions:

(a) Only the northbound flow is important.

(b) The two directions are equally important.

26-11. Find the offset for a link of 1,500 ft, no standing queue at the downstream signal, and a platoon traveling at 40 ft/s. Re-solve if there is a standing queue of eight vehicles per lane.

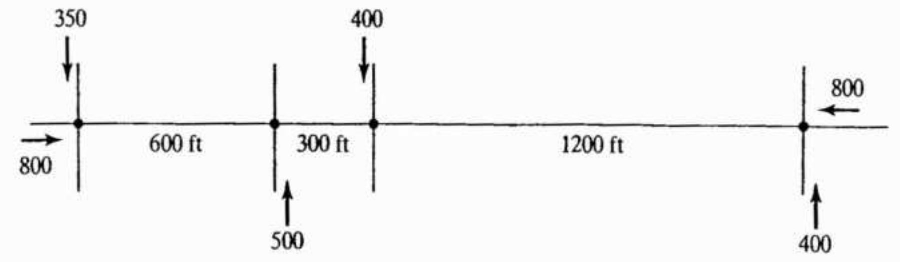
26-12. Develop an arterial progression for the situation shown in the figure below. Use a desired platoon speed of 40 ft/s. For simplicity, the volumes shown are already corrected for turns and PHF.

26-13. Throughout this chapter, the emphasis was on platoons of vehicles moving through the system, with no desire to stop. However, buses travel slower than most passenger cars and must stop. This problem addresses the timing of signals solely for the bus traffic.

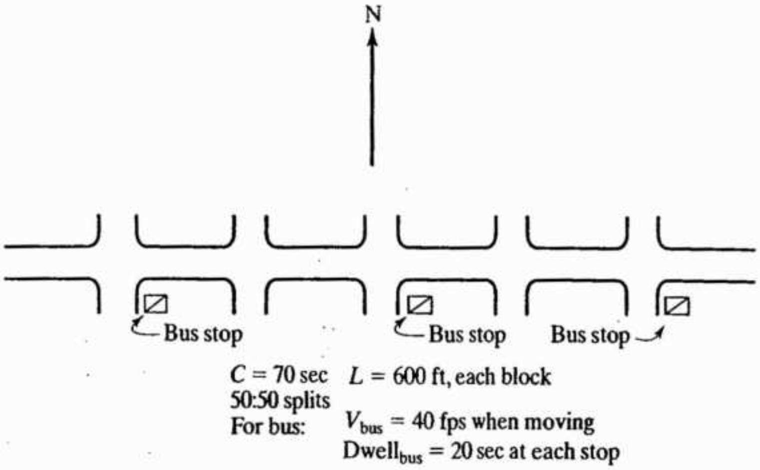
(a) For the situation shown in the figure, time the signals for the eastbound bus. Draw a time-space diagram of the solution.

(b) Now consider the westbound bus. Locate the westbound bus stops approximately every two blocks and adjust the offsets to make the best possible path for the westbound bus, without adversely affecting the eastbound bus. Draw the revised time-space diagram.

(c) Show the trajectories of the eastbound and westbound lead passenger cars going at 60 fps.



Arterial for Problem 26-12

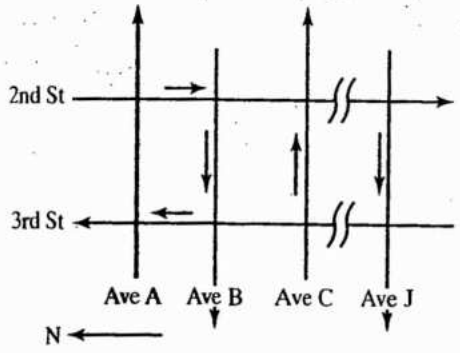


$C = 70$ sec $L = 600$ ft, each block
50:50 splits
For bus: $V_{\text{bus}} = 40$ fps when moving
$\text{Dwell}_{\text{bus}} = 20$ sec at each stop

Arterial for Problem 26-13, with Bus Stops

26-14. Refer to the figure below. Second street is southbound with offsets of +15 seconds between successive signals. Third street is northbound with offsets of +10 seconds between successive signals. Avenue A is eastbound, with a +20-second offset of the signal at Second Street and Avenue A relative to the signal at Third Street and Avenue A. Given this information, find the offsets along Avenues B through J. The directions alternate, and all splits are 60-40, with the 60 on the main streets (2nd and 3rd Streets).

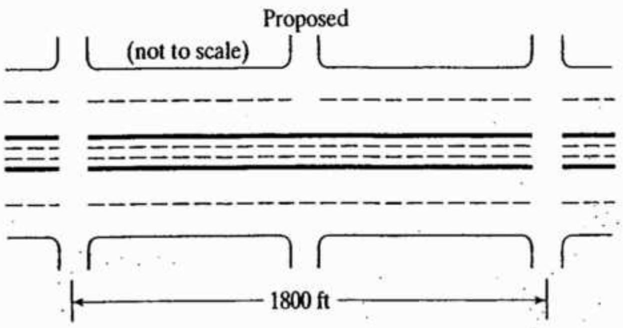
26-15. Given three intersections, spaced 600 feet apart, each with $C = 60$ seconds and 50-50 split, find an offset pattern that equalizes the bandwidth in the two directions. *Hint:* Set the first and the third relative to each other, and then do the best you can with the second intersection. This is a good way to start.



Notes: (1) C-60 sec Splits 60:40
(2) Block lengths 600 ft in all cases
(3) All streets are one-way.

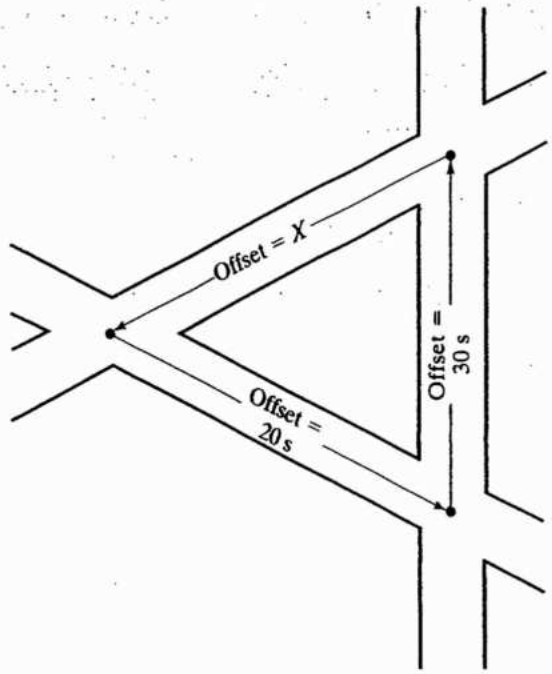
Network for Problem 26-14

26-16. A major development is proposed abutting a suburban arterial as shown in the figure below. The arterial is 60-feet wide, with an additional 5 feet for shoulders on each side, and no parking. There is moderate development along the arterial now. Platoons of vehicles travel at 60 ft/s in each direction. The center lane shown in the figure is for turns only. The proposed development is on the north side, with a major driveway to be added at 900 feet along the arterial, requiring a signal. Evaluate the impact of this development in detail. Be specific, and illustrate your points and recommendations.



Arterial for Problem 26-16

26-17. Refer to the figure. Find the unknown offset X. The cycle length is 80 seconds. The splits are 50-50.



Arterials for Problem 26-17

26-18. Given an arterial with 20 consecutive signals, spaced at 1,500 ft with vehicles moving at 50 ft/s, which coordination scheme is the best: simultaneous, alternate, or double alternate? What cycle length should be used?

26-19. Given the following information for the indicated arterial, with $C = 70$ seconds and 50-50 splits:

(a) Plot the time-space diagram.

(b) Find the two bandwidths. Show them graphically and find the numeric values. If they do not exist, say so.

(c) An intersection is to be placed midway between Intersections 3 and 4, with $C = 70$ seconds and a 50-50 split. Recommend an appropriate offset.

	1		2		3		4
		600 ft		600 ft		600 ft	
offsets:		+20 s		+30 s		+10 s	

CHAPTER 27

Signal Coordination for Arterials and Networks: Oversaturated Conditions

It is well recognized that the oversaturated traffic environment is fundamentally different from the undersaturated environment. In undersaturated networks, capacity is adequate and queue lengths are generally well contained within individual approaches.

The oversaturated environment is characterized by an excess of demand relative to capacity ($v/c > 1.00$) and thus unstable queues that tend to expand over time. This tends to lead to situations in which queues spill back into the upstream intersections and block other flows. This in turn inhibits discharge on one or more approaches, in effect reducing capacity when it is most needed.

Control policies for oversaturated networks thus focus on maintaining and fully exploiting capacity to *maximize productivity* (vehicle throughput) of the system by controlling the inherent instability of queue growth.

27.1 System Objectives for Oversaturated Conditions

For undersaturated flow ($v/c < 1.00$) on arterials and in networks, the primary objective is to attain smooth flow as described in Chapter 26, often expressing this objective in terms of minimizing delay and stops. This primary objective tends to become the "prime directive" and even the obsession.

But when approaches, intersections, or systems become oversaturated, the mission and objective is *fundamentally different.* But many people do not change their mindset, and—even as this is written—most of the computational tools (i.e., signal optimization programs) do not address oversaturation directly.[1]

When networks are congested, the explicit objectives change to:

- *Maximize system throughput.* This is the primary objective. It is achieved by (a) avoiding queue spill back, which blocks intersections and wastes green time; (b) avoiding starvation: the tardy arrival of traffic at the stop bar that wastes green time; and (c) managing queue formation to yield the highest service rate across the stop bar.
- *Fully utilize storage capacity (queue management).* This objective seeks to confine congested conditions to a limited area by managing queue formation in the context of a "feed forward" system.
- *Provide equitable service.* Allocate service to cross street traffic and to left-turners so that all travelers are serviced adequately and the imperative of traffic safety is observed.

Because intersection blockage can so degrade the network, its removal must be the keystone of the prime objective for the traffic engineer.

27.2 Root Causes of Congestion and Oversaturation

Reference [1] addressed the "root causes" of congestion and oversaturation, so that practitioners could put both the apparent problem *and* the physical limits on remedying it into perspective as they try to identify solutions.

Many of the root causes are imbedded in the fabric of our urban areas and the routes connecting them. Consider, for instance, the following list, extended from [1]:

- *Convergence of routes.* It is inevitable that congestion will arise as several routes converge on such natural features as river crossings (bridges, tunnels), paths out of valleys or between surrounding hills, paths to the waterfront, airports, or major attractors. Yet this convergence is imbedded in the history and evolution of our cities, sometimes over centuries or decades;
- *Crossings of major routes.* It tends to be unavoidable that two major arterials cross each other as they carry their respective through movements from the origins to destinations consistent with their design. It is also common that turn movements at these intersections of major routes tend to be significant as people move from one to the other. These intersections become "critical intersections" at which demand exceeds capacity, and there is no true "major" and "minor" street, rendering the adage of giving priority to the major street moot;
- *Natural features, historic sites, and special architecture.* It is common that topography (knolls, hills) have shaped the development of the area, that large parks or commons areas exist, and that monuments in town circles, large libraries or railroad stations, and, more recently, auto-free zones define where the flows are and the paths they must take from one critical intersection to another;
- *Street spacings.* They are also frequently defined by history and uses that predate motor vehicles. In some cases, they are defined by the original surveys (as in the western United States). It is truly serendipitous when street spacings fit the neat "alternate progression" needs implied by the $L/S = C/2$ relation of Chapter 26;
- *Unavailability of alternative routes.* The work in [1] uncovered the apparent paradox that medium-size cities and towns reported more perceived congestion that either their larger or smaller counterparts. A little probing revealed that small areas may have one main street with short peak "hours" and that large areas may have multiple routes between any origin and destination (subject to preceding points, of course) *but* that medium-size areas had significant traffic that still centered around one major arterial or street;
- *Natural variability.* The traditional view of both capacity and demand is that they are both known and invariant, at least for the period of the analysis.[2] But yet both are random around these expected values, leading to situations in which v/c fluctuates well above (and below) its nominal value.

---

[1] Specifically, they do not model or take into account the effects of spillback and intersection blockage. Reference [2] notes that "those methods seemed insensitive to even the existence of growing residual queues, let alone (consideration of them) in the optimization objective. This impression is supported by a cursory review of available tools. . . . No existing popular tools could be identified . . . that optimized specifically to maximize throughput or to manage queues."

[2] It is true that peaking is taken into account by the peak-hour factor (PHF), but this simply increases the value of the constant that is put into the steady-state equations that exist in the literature and key references.

This mixture of history, topography, convergence of routes, relative size, and natural variability creates an environment that virtually defines both the street system and the critical intersections. Added to this is the mixture of modes that are common in most urban areas (and suburban towns)—buses, paratransit, truck deliveries, and more recently package delivery services, bicycles, and a growing emphasis on pedestrian travel as a mode worthy of attention (and design) in its own right.[3]

27.3 Overall Approaches to Address Oversaturation

The *overall approach* stated in [1] as a set of logical steps was as follows:

- Address the root causes of congestion—first, foremost, and continually;
- Update the signalization, for poor signalization is frequently the cause of what looks like an incurable problem;
- If the problem persists, use novel signalization to minimize the impact and spatial extent of the extreme congestion;
- Provide more space, by use of turn bays and parking restrictions;
- Consider both prohibitions and enforcement realistically—is it effort likely to be effective or futile? Will it only transfer the problem to another location? Is it contrary to the fundamental use of the area?
- Take other available steps, such as right-turn-on-red, recognizing that the benefits will generally not be as significant as either signalization or more space;
- Develop site-specific evaluations where there are conflicting goals, such as providing local parking versus moving traffic, when the decision is ambiguous. Explicitly consider the solution in terms of economics.

The last category was intended (for instance) to focus the debate on the use of space, by quantifying the effects and tradeoffs—for example, use for good delivery, bus lane, parking, or through/turning traffic.

The preceding list was constructed in [1] with some allowance for ease of implementation: It is generally easier to change signalization than to remove urban parking, it is generally easier to treat spot locations than entire arterials, and so forth.

Reference [2] expressed the approach in terms of *throughput strategies* (which it considered curative) and *queue management strategies* (which it considered palliative).

For curative *throughput strategies*, the experts consulted in the course of the Reference [2] work identified three categories or themes:

- Make the best use of the physical space available in the intersection.
- Make the best use of the green time in the cycle.
- Minimize the negative effects of other influences.

The specific techniques identified in [2] included:

- Work back from the downstream bottleneck;
- Run nearby intersections on a single controller, so that they stay in lock-step even during transitions;
- Improve the lane utilization;
- Run heavy left turns on a lag phase;
- Find and use the right cycle;
- Service heavy movements more than once in a cycle ("phase reservice");[4]
- Consider the effect of buses;
- Minimize the effect of pedestrian movements;
- Seek all possible available green time, even if exotic signal controller features must be used;
- Consider congested and uncongested movements separately;
- Prevent actuated short greens.

The more detailed advice in [2] on some of these includes (1) in general, it is good to set green times so that one bus per cycle can be serviced on relevant phases, when one considers bus dwell times; (2) pedestrian actuation buttons allow long phase minima to be avoided when the button is not actuated; (3) slow-starting trucks may lead to a short actuated phase, and gap settings have to consider this; (4) lanes get less productive over time, even with a standing queue, so that greens of up to 30 seconds are more productive than longer ones (indeed, it was noted that this is particularly true for through lanes next to left-turn bays, where movements into the turn bay actually leave gaps in the thru flow *if* the green time is longer than the bay length).

---

[3] The 2003 edition of the *MUTCD* required that pedestrian crossing times be computed based on the entire travelled path, rather than to center of the far lane. As of this writing, the 2010 edition of the *MUTCD* is likely to reduce the assumed pedestrian walking time from 4.0 ft/s to 3.5 ft/s, thereby increasing minimum phase durations.

[4] This technique assumes there are some phases for which the v/c is less than 1.00, perhaps due to a combination of demand and minimum pedestrian requirements on phase duration.

For the palliative *queue management strategies,* the experts consulted in [2] identified techniques including:

- Reduce minor splits to encourage diversion (although it was noted that this generally does not work);
- Run nearby intersections on a single controller, so that they stay in lock-step even during transitions;
- Balance the queues for conflicting approaches;
- Prevent queues from spreading congestion;
- Meter traffic into the bottlenecks;
- Prevent downstream queues from backing up into the bottleneck intersections.

The approaches enumerated in [1] and [2] tend to be *operational* in nature, looking for short-term solutions to existing critical problems. There are indeed longer term remedies that can be addressed, mostly falling under the general heading of *transportation demand management.* These include:

- Assuring that alternative *higher occupancy modes* of transportation are available, such as bus routes, light rail, and even heavy rail. But in these alternatives, it is necessary to plan over the long term for *integrated* and *complete* systems, rather than fragmented routes;
- Encouraging *higher occupancy use* of automobiles by providing the support facilities (e.g., park and ride lots), lower tolls, special use lanes, and even incentives (e.g., parking or parking reimbursement);
- Encouraging *temporal shifts in demand* by incentives and disincentives (congestion pricing, for instance, can be viewed both ways), and by strategic planning with large employers.

To a large extent, these are excellent long-term measures that cannot and should not be ignored. The present chapter, however, focuses on shorter term operational issues—the result of inaction or inattention on the broader scale.

One long-term trend deserves special mention in this section. The historic model is of the "hub city" such as New York or Chicago in which the population surrounds the core in rings and travels to it in the journey to work. For several decades, census data have been showing a distinctly different trend, even in these prototypical areas: What would have been called "suburb-to-suburb" work trips are growing, and hub-bound travel is decreasing. The result can be depicted as a set of mini-cities within the same region, depending on true arterials and other facilities to link them. Concurrent with this, the directional patterns become more balanced, so that setting signals for "tidal flow" (80-20 directional splits) in the morning and evening becomes less typical.

27.4 Classification

NCHRP 194 [1] included the words "Three groups of terms must be defined" and then listed (1) congestion-related terms, (2) terms related to the types of oversaturation, and (3) characteristics distinguishing the productivity of an intersection and the perception of congestion. Table 27.1 shows the attempt of that era to define the congestion-related terms.

The question of classification of congestion has been addressed repeatedly in the literature. For instance, Reference [2] notes that [3] and [4] are two of the few works on metering, but that much work was not clearly related to actions that practitioners would take. Reference [2] also recommends a simplified and hierarchical classification, namely:

- Light traffic
- Moderate traffic
- Heavy traffic
- Oversaturated operations

It further notes that "Experts were not generally interested in defining the point of saturation in a precise or scientifically

Table 27.1: Early Definition of Congestion-Related Terms (c. 1978)

	Congested		
	Saturated		
Uncongested	**Stable**	**Unstable**	**Oversaturated**
No queue formation.	Queue formation, but not growing. Delay effects local.	Queue formation and growing. Delay effects still local. A transient state may be only of short duration.	Queue formation and growing to a point where upstream intersection performance is adversely affected.

(*Source:* Reference 1, p. 113.)

Table 27.2: Types of Oversaturation / Congestion

Types of Oversaturation	**Types of Congestion**
NCHRP 194 [1], 1978	**ITE Webinar, October 11, 2007 [5]**
Type I—The critical intersection (CI) has a smaller *g*/*C* ratio than does the upstream intersection. Type II—The CI and the upstream intersection have the same *g*/*C* ratio. However, the capacity of the CI is less than that of the upstream intersection because of factors other than the g/C ratio (such as turning movements and/or physical conditions). Type III—Heavy turn-in movements from the upstream cross street fill up the entire link or a significant part of it during a red phase on the arterial and cause spill back on the arterial. Type IV—This type of oversaturation of a CI results from the signal offset between the CI and its upstream intersection. Type V—This is defined as being a combination of two or more of Types I through IV oversaturation.	Type I—CI has less green time than upstream intersections (typical of CI at junction of crossing arterials). Type II—CI has lower capacity due to additional phases, geometrics, or backups from downstream intersection. Type III—CI has greater demand due to heavy turning at upstream intersection. Type IV—Congestion due to offset relationship with upstream intersection.

elegant way, but were rather interested in conditions that would justify a change in strategy or more detailed adjustment to operation on the ground." We agree with this classification and with the observation.

It was interesting that some of the thoughts from the cited earlier era (c. 1978) are still reflected in current practice and in professional development, including an ITE webinar [5] given in October 2007. Refer to Table 27.2 for an illustration.

These two events—separated by so many years—highlight the reality that engineers are coping with the same fundamental problems, often expressed in the same ways.

The needs of different modes sometimes reinforce each other in selecting a treatment. Whereas tradition often favors longer cycle lengths as demand grows, *shorter* cycle lengths (1) are advocated in [1] and in [5] in order to achieve shorter queues and more phases per hour, but they also (2) allow for lower pedestrian waiting time and better levels of service for pedestrians, an emerging priority in the multimodal environment. The case can also be made that *turn restrictions* aid more than one mode, and that *some bus measures* (e.g., stop location) can aid several modes.

27.5 Metering Plans

One short-term demand management strategy is *metering*. Three forms of metering can be applied within a congested traffic environment, characterized by demand exceeding supply (i.e., *v*/*c* deficiencies): internal, external, and release.

Internal metering refers to the use of control strategies within a congested network so as to influence the distribution of vehicles arriving at or departing from a critical location. The vehicles involved are stored on links defined to be part of the congested system under control, so as to eliminate or significantly limit the occurrence of either upstream or downstream intersection blockage.

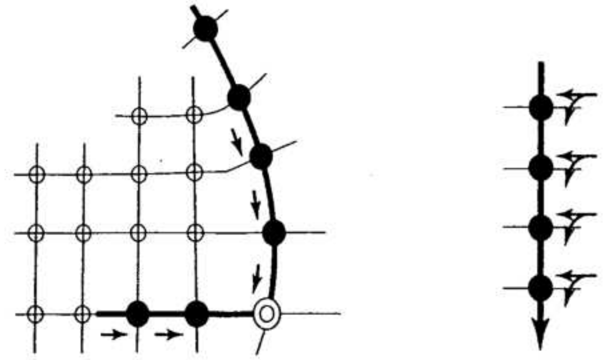
Note that the metering concept does not explicitly minimize delay and stops, but rather it manages the queue formation in a manner that maximizes the productivity of the congested system.

Figures 27.1 and 27.2 show situations in which internal metering might be used: (1) controlling the volume being discharged at intersections upstream of a critical intersection (CI), thus creating a "moving storage" situation on the upstream links, (2) limiting the turn-in flow from cross streets, thus preserving the arterial for its through flow, and (3) metering in the face of a backup from "outside."

External metering refers to the control of the major access points to the defined system, so that inflow rates into the system are limited if the system is already too congested (or in danger of becoming so).

External metering is convenient conceptually because the storage problem belongs to "somebody else," outside the system. However, there may be limits to how much metering can be done without creating major problems in the "other" areas. Figure 27.3 shows a network with metering at the access points.

As a practical matter, there must be a limited number of major access points (such as river crossings, a downtown



(a) g/c reduced in order to lower indicated discharge flows

(b) Cross street (g/C) lowered to preserve actorial for through flow

● Internally metered intersection

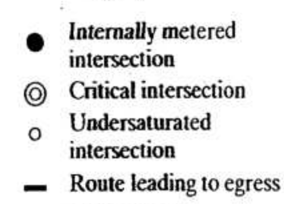
◎ Critical intersection

○ Undersaturated intersection

Figure 27.1: Internal Metering Used to Limit Volume Arriving at Critical Location

surrounded by water on three sides, a system that receives traffic from a limited number of radial arterials, etc.). Without effective control of access, the control points can potentially be bypassed by drivers selecting alternative routes.

Release metering refers to cases where vehicles are stored in such locations as parking garages and lots, from which their release can be controlled (at least in principle). The fact that they are stored "off street" also frees the traffic engineer of the need to worry about their storage and their spill-back potential.



● Internally metered intersection

◎ Critical intersection

○ Undersaturated intersection

▬ Route leading to egress

— Other street

Figure 27.2: Application of Internal Metering in the Face of a Backup from "Outside"

Release metering can be used at shopping centers, megacenters, major construction sites, and other concentrations. Although there are practical problems with public (and property-owner) acceptance, this could even be—and has been—a developer strategy to lower the facility's discharge rates so that adverse impacts are avoided.[5] Such strategies are of particular

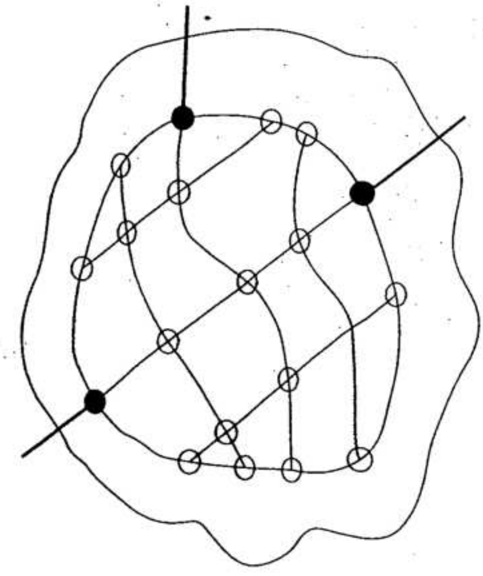


Figure 27.3: Illustration of External Metering

---

[5] Traffic impacts by developers leads to requirements for traffic mitigation, generally at the cost to the developer. The general subject is addressed in Chapter 30.

interest when the associated roadway system is distributing traffic to egress routes or along heavily congested arterials.

27.6 Signal Remedies

It is difficult to overstate how often the basic problem is poor signalization. Once the signalization is improved through reasonably short cycle lengths, proper offsets (including queue clearance), and proper splits, then many problems disappear. Sometimes, of course, there is just too much traffic. At such times *rapid adjustment to splits* to meet short-term relative demand changes (i.e., in competing directions), *equity offsets* to lower the likelihood of spill back and to aid cross flow, a *different concept of splits*, namely to manage the spread of congestion, and *phase reservice* may be appropriate if other options cannot be called on. These options may be used as distinct treatments or as part of a *metering plan* (addressed in the previous section).

27.6.1 Responsive/Adaptive Phase Duration Changes

Locations such as college entrances have short bursts of inflows followed by short bursts of outflows (both in the order of 15 to 25 minutes), directly related to their class schedules. Control in such cases must adapt to the rapidly changing demand, in order to avoid precipitating oversaturation that can promulgate and perpetuate itself.

27.6.2 Shorter Cycle Lengths

An earlier chapter demonstrated that increasing the cycle length does *not* substantially increase the capacity of the intersection (a change of +50% in cycle length could add +5% to 8% in capacity, under favorable conditions). But the favorable conditions include maintaining high discharge rates over long phases, and this does not occur in practice.

Rather, the emphasis should be on the related reality that as the cycle length increases, so do the lengths of stored queue and the length of the discharged platoons, which then arrive at downstream intersections that may have shorter cycle lengths and cannot be stored or processed easily.

Thus the likelihood of intersection blockage increases, with substantial adverse impacts on system capacity. This is particularly acute when short link lengths are involved.

Note that a critical flow of v_i veh/hr/lane nominally discharges v_i *C/3600* vehicles in a cycle. If each vehicle requires D feet of storage space, the length of the downstream link in a congested environment (assuming the downstream signal can process the queue in one cycle but that it will be forced to stop) would have to be:

$$L \geq \left(\frac{v_i C}{3600} \right) \mathrm{D} \qquad (27\text{-}1)$$

where L is the *available* downstream space in feet. This "available" space may be the full link length or by some lower value, perhaps 150 feet less than the true length (to keep the queue away from the discharging intersection or to allow for turn-ins). The engineer will have to decide that, noting that [1] shows results that queues that occupy more than 85% of the link length actually inhibit the discharge rate at the upstream intersection.

Equation 27-1 may be rearranged as:

$$C \leq \left(\frac{L}{D} \right) \left(\frac{3600}{v_i} \right) \qquad (27\text{-}2)$$

Note that v_i in this case is the discharge volume per downstream lane, which may differ from the demand volume, particularly at the fringes of the "system" being considered. Refer to Figure 27.4 for an illustration of this relationship. Note that only rather high flows (≥800 veh/h/ln) and short blocks will create very severe limits on the cycle length. However, these are just the situations of most interest for conditions or extreme congestion. *Note that the discharge volume v_i depends on the upstream demand and (g/C) allocation,* and that this analysis really has to be carried along the arterial.

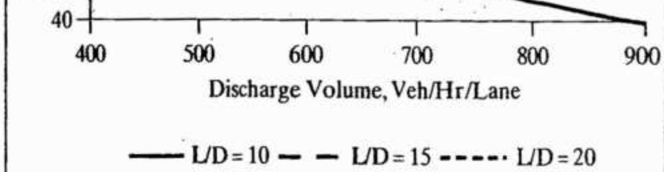


Figure 27.4: Maximum Cycle Length as a Function of Block Length[6]

---

[6] It is recognized that cycle lengths shorter than 60 seconds are rare, but they are shown to emphasize the shape of the curves.

In Figure 27.4, note that D equals approximately 25 feet, so that block lengths with *available* storage of 250 feet to 500 feet are represented (300 to 600 feet, if the 85% rule of thumb cited earlier is used).

The important lesson from Figure 27.4 is that shorter cycle lengths are not only good but *necessary* to manage the size of queues arriving in downstream links.

This analysis assumes that the downstream link can itself discharge the arriving queue in one cycle. To achieve this (for instance, at a critical intersection), it may be necessary to allow the downstream capability determine the upstream discharge, which would have to be achieved by reassigning green time there (to minor movements) or imposing an all-red (i.e., metering). Failing this, the downstream queues will grow and other measures will be needed to avoid spill back.

27.6.3 Equity Offsets

Offsets on an arterial are usually set to move vehicles smoothly along the arterial, which is logical. If no queues exist on the arterial, the ideal offset is L/S, where L is the signal spacing in feet and v is the vehicle speed in feet per second. If a queue of Q vehicles exists, the ideal offset t_{ideal} is:

$$t_{ideal} = \frac{L}{S} - Q * h \qquad (27\text{-}3)$$

where h is the discharge headway of the queue, in seconds. Clearly, as Q increases, t_{ideal} decreases, going from a "forward" progression to a "simultaneous" progression to a "reverse" progression.

Unfortunately, as the queue length approaches the block length, such progressions lose meaning, for it is quite unlikely that both the queue *and* the arriving vehicles will be processed at the downstream intersection. Thus the arrivals will be stopped in any case.

At the same time, the cross-street traffic at the upstream intersection is probably poorly served because of intersection blockages. Figure 27.5 illustrates the time at which t_{ideal} would normally cause the upstream intersection to switch to green (relative to the downstream intersection).

Consider the following case, illustrated in Figure 27.6: Allow the congested arterial to have its green at the upstream intersection until its vehicles just begin to move. Then switch the signal so these vehicles flush out the intersection but no new vehicles continue to enter.

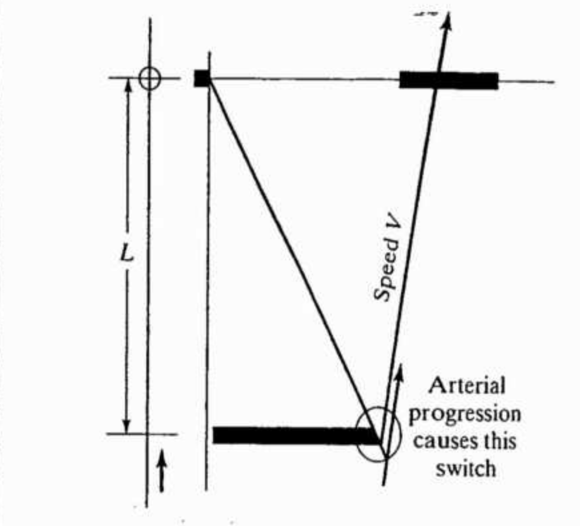


Figure 27.5: Ideal Offset with Queue Filling Block

At the same time, this gives the cross-street traffic an opportunity to pass through a clear intersection. This concept, defined as *equity offset*, can be translated into this equation:

$$t_{equity} = g_i C - \frac{L}{S_{acc}} \qquad (27\text{-}4)$$

where g_i is the *upstream* main street (i.e., the congested intersection) green fraction and S_{acc} is the speed of the "acceleration wave" shown in Figure 27.7.

A typical value is 16 fps. Comparing Figure 27.5 and 27.7, it is clear that equity offset causes the upstream signal to go red just when "normal" offsets would have caused it to switch to green in this particular case. This is not surprising, for the purpose is different: Equity offsets are intended to be fair (i.e., equitable) to cross-street traffic.

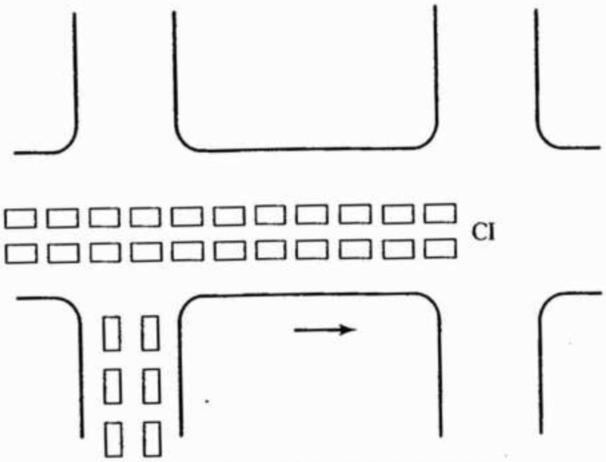
Simulation tests using a microscopic simulation model have shown the value of using equity offsets: Congestion does not spread as fast as otherwise and may not infect the cross streets at all.

Figure 27.8 (a) shows a test network used to test the equity-offset concept. Link 2 is upstream of the critical intersection (CI). For the demands and signal splits shown it is likely to accumulate vehicles, with spill back into its upstream intersection likely. If this occurs, the discharge from Link 1 will be blocked and its queue will grow. In the extreme, congestion will spread.

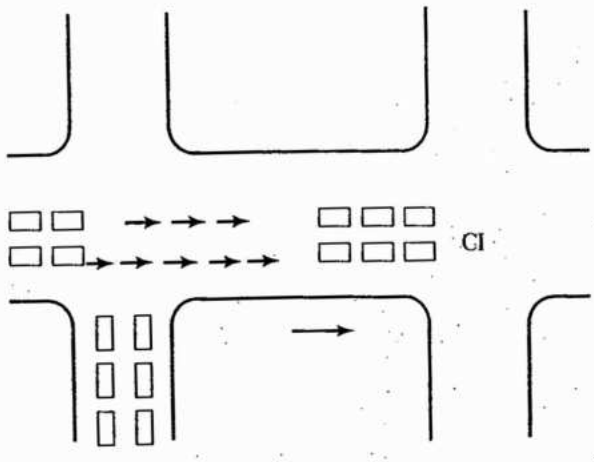
The equity offset is computed as

$$t_{equity} = (0.60)(60) - \frac{600}{16} = -1.5\,s$$

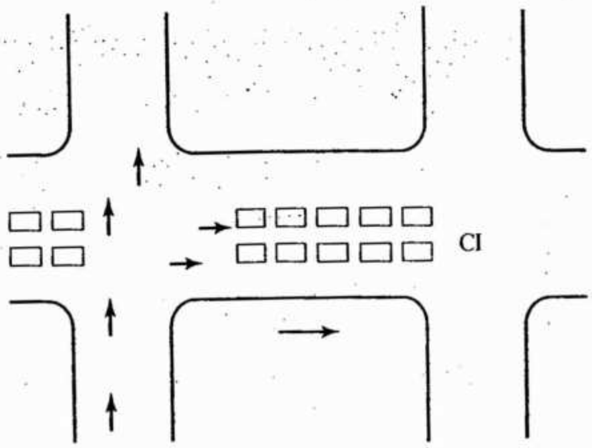
using Equation 27-4. (At 25 feet per vehicle and a platoon speed of 50 ft/s, Equation 27-3 would have yielded



(a) Main-street vehicles block upstream



(b) As main-street vehicles just begin to clear, give green to side street



(c) Cross street are thus allowed to discharge

Figure 27.6: Concept of Equity Offsets to Clear Side Streets

t_{ideal} = (600/50) − (24)(2)= −36 seconds for progressed movement. Of course, progressed movement is a silly objective when 24 vehicles are queued for 30 seconds of green.)

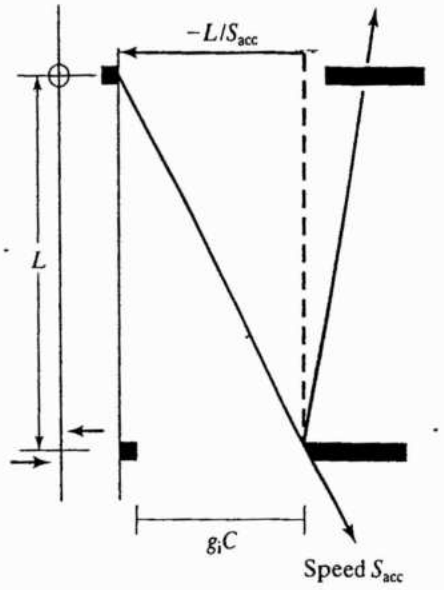


Figure 27.7: Equity Offset to Benefit Cross Street

Figure 27.8 (b) shows the side-street queue (i.e., the Link 1 queue) as a function of the main-street offset. Note that an offset of −36 seconds is the same as an offset of +24 seconds when C = 60 seconds, due to the periodic pattern of the offsets. Figure 27.8 (b) shows the best result for allowing the side street to clear when the equity offset (offset = −1.5 seconds) is in effect, and, in this case, the worst results when the queue-adjusted "ideal offset" (offset = 24 seconds) would have been in effect.

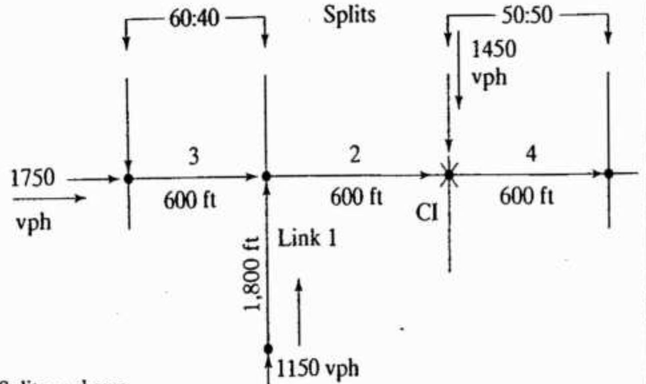
The preceding discussion assumes that the cross-street traffic does not turn into space opened on the congested arterial. If a significant number of cross-street vehicles do turn into the arterial, a modification in the offset is appropriate to assure that the upstream traffic on the congested arterial also has its fair share.

The equity offset concept has been used to keep side-street flows moving when an arterial backs up from a critical intersection (CI). It may also be used to keep an arterial functioning when the cross streets back up across the arterial from their critical intersections.

Figure 27.9 shows another illustration of the concept that offsets can be designed to relieve congestion and the likelihood of upstream intersection blockage.

27.6.4 Imbalanced Split

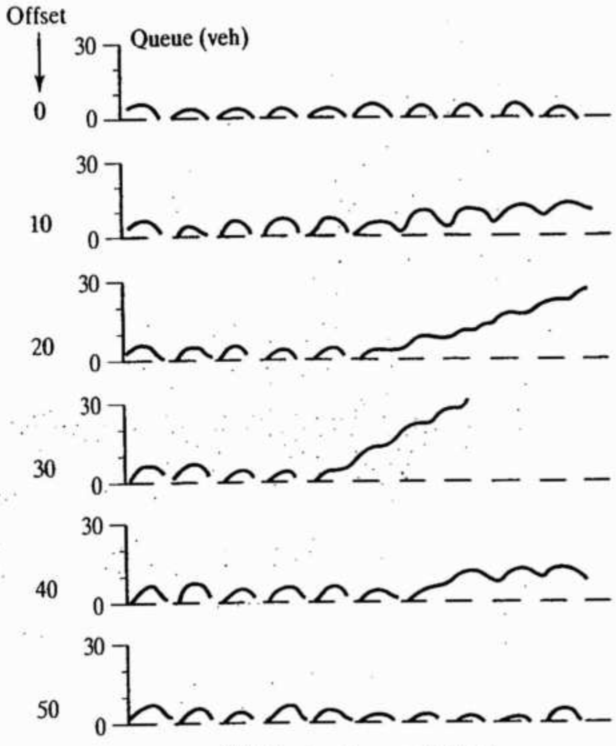
For congested flow, the standard rule of allocating the available green in proportion to the relative demands could be used, but it



Splits as shown
Offset for simple progression:
600/45 fps = 13.3 sec.
Two moving lanes per link
No turns

$C = 60$ sec
Offsets: Simple progression, except for link 2
Link 2: Variable

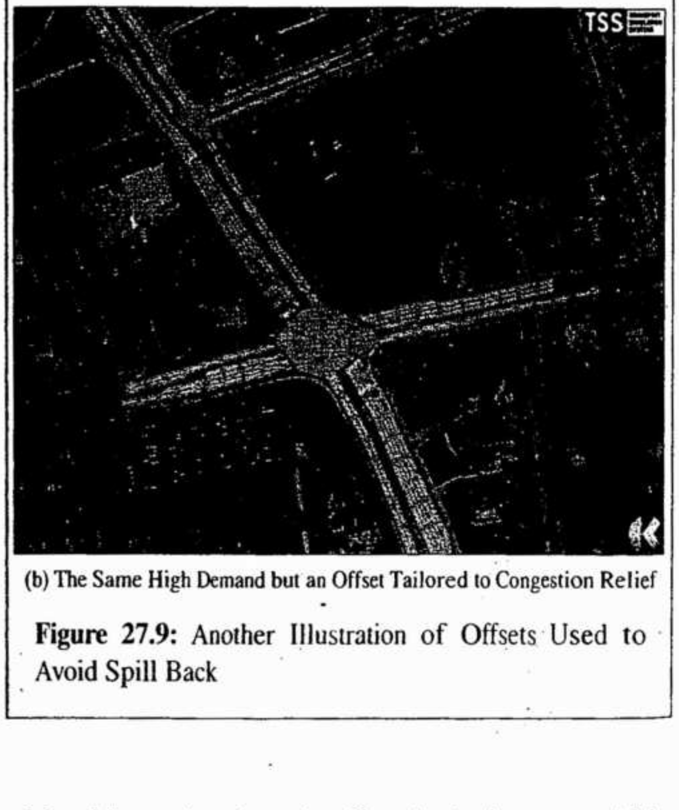
(a) Test network



(b) Side-street queue (Link 1)

Figure 27.8: Equity Offsets Avoid Side-Street Congestion, Despite Spill Backs

(a) High Demand and "Traditional" Offsets, with Consequent Congestion



(b) The Same High Demand but an Offset Tailored to Congestion Relief

Figure 27.9: Another Illustration of Offsets Used to Avoid Spill Back

does not address an important problem. Consider the illustration of Figure 27.10. If the prime concern is to avoid impacting Route 347 and First Ave. (but with little concern for the minor streets in between, if any), it is not reasonable to use a 50:50 split.

Considering that the relative storage available is 750 feet in one direction and 3000 feet in the other, and we wish neither to be adversely affected, the impact could be delayed for the longest time by causing the excess-vehicle queue to grow in proportion to their available storage. The two critical-lane discharge flows f_i would have to be set such that:

$$\frac{d_1 - f_1}{d_2 - f_2} = \frac{L_1}{L_2} \tag{27-5}$$

and:

$$f_1 + f_2 = CAP \tag{27-6}$$

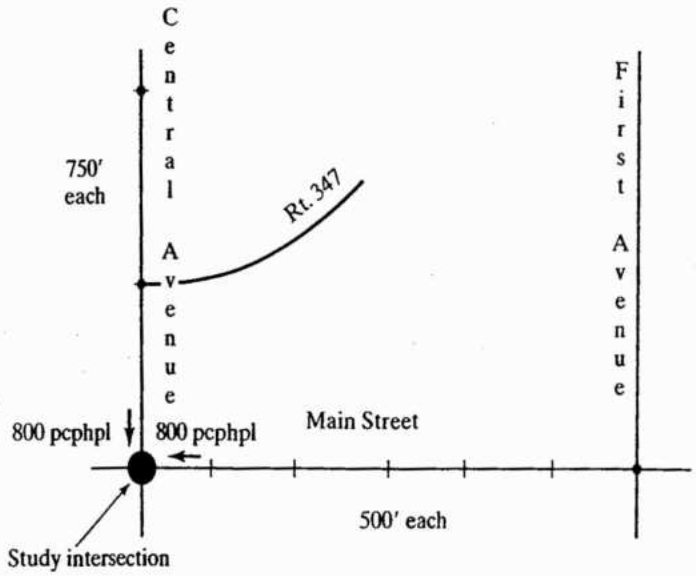


Figure 27.10: An Illustration of Split Determination

(Used with permission of Transportation Research Board, from [1].)

where d_i are the demands (veh/h/ln), L_i is the storage and CAP is the sum of the critical-lane flows (i.e., the capacity figure).

For the illustrative problem, using CAP = 1550 veh/h/ln, the preceding equations result in $f_1 = 954$ veh/h/ln and $f_2 = 759$ veh/h/ln, where direction 1 is the shorter distance. This is a 56:44 split.

Note that in the extreme, if only one direction has a cross route that should not be impacted, much of the green could be given to that direction (other than some minimum for other phases) in order to achieve that end.

27.6.5 Phase Reservice

The term *phase reservice* refers to servicing important phases more than once in a cycle, by going back to them, generally to the disbenefit of side street movements on other phases. The technique is for clearing queues on protected lefts and saturated approaches, but it generally requires that there are undersaturated phases at the intersection, so that one can "catch up" with servicing them on a future cycle if necessary.

Phase reservice can aid the basic objectives of maximizing throughput and queue management. It does require that both drivers and pedestrians become familiar with this sort of operation, so that all concerned are aware that the "normal" sequence of phases cannot be counted on.

27.6.6 Pedestrian Minima Provided Only Upon Request

In areas in which there are relatively little pedestrian traffic, satisfying the pedestrian minimum crossing times on all approaches may lead to phases that waste green time and to longer-than-necessary cycle lengths. In such cases, traffic engineers sometimes do consider invoking pedestrian minima only when there is an actual pedestrian actuation of a pedestrian button.

This of course requires that the intersection have functioning pedestrian buttons in place (as well as pedestrian signals) and that the pedestrians learn that the only way to assure the pedestrian crossing times is to *use* the buttons.[7]

27.7 Variations in Demand and Capacity

In current (and historic) methods, traffic demand and the capacity of the traffic signal are taken as fixed values, rather than intrinsically random.

---

[7] Although there is no factual basis to support the supposition, the authors do speculate that this environment might be an ideal application of the pedestrian "countdown" signals that are coming into use. The pedestrian is rewarded by the action of pushing the button having a definitive and highly visible effect, namely the pedestrian signal changing *and* providing information on time remaining.

For simplicity, consider the equation $D = 1 / (1 - \{v/c\})$, a very oversimplified version of the standard delay equation but sufficient for the illustration. In probability theory, it is clear that

$$E[D] \neq \frac{1}{1 - E[v/c]} \neq \frac{1}{1 - (E[v]/E[c])} \qquad (27\text{-}7)$$

That is, the expected value of a quantity such as delay *cannot* be computed by simply putting the expected values of the input values into the right-hand side of a static equation.

For years, such an approach was *good enough*, considering (1) the amount of data available, (2) the cost of data, and (3) the need for results that were good enough for the applications at hand.

Still, it was always known there were problems—delay data had very high variability, particularly as the v/c ratio increases (plots of delay versus v/c have trends, but also high scatter of the data at high v/c). In addition, it was known that some data (maybe much data) represented the transition between two demand levels, rather than "steady-state" data from different equilibrium conditions. But the need for data—and the cost of it—often competed with such precision. And, besides, the results were *good enough* for the applications at hand.

27.7.1 An Illustration of the Effects of Demand and Capacity Variability on Delay

For present purposes, we consider the case in which the expected value of demand is 1700 veh/h and of capacity is 1900 veh/h. The computations of delay are illustrative, and we do not list values for all of the factors in the *HCM*. The computed v/c ratio is 0.89 in this case, and the computed delay is 51.0 s/veh.

Table 27.3 shows the same computation but with normally distributed values of demand and capacity. A total of 525 samples, or "observations," were generated in a spreadsheet for this illustration.

Figure 27.11 contains histograms of the "actual" observed delays, based on the spreadsheet computations. Note that:

1. There is a significant fluctuation in the delay, even when the standard deviation is as low as 30 vph;
2. If it were possible to use the rule of thumb that 95% of the delay observations fall within 2 standard deviations, the estimated range is 24.4 to 85.2 vph, quite a variation around an observed mean of 54.8 vph;
3. But it is *not* possible to apply this rule of thumb because it is only truly valid for the symmetric normal distribution. As Figure 27.11b shows, when $\sigma = 60$ vph, the delay distribution is *not* symmetric but rather has a long tail—*and* a noticeable set of *apparently* anomalous very high delays.
4. Thus the common rule of thumb would *underestimate* the spread in asymmetric cases.
5. As to the "apparent" anomaly cited in number 3, Figure 27.12 shows the range of v/c ratios that existed in Case 2. Although the average was close to the 0.89 computed from averages, the *range* is from about $v/c = 0.80$ up to over 1.00. These higher values in particular would lead to very long delays, if the intersection approach were long enough to receive and store the vehicles.

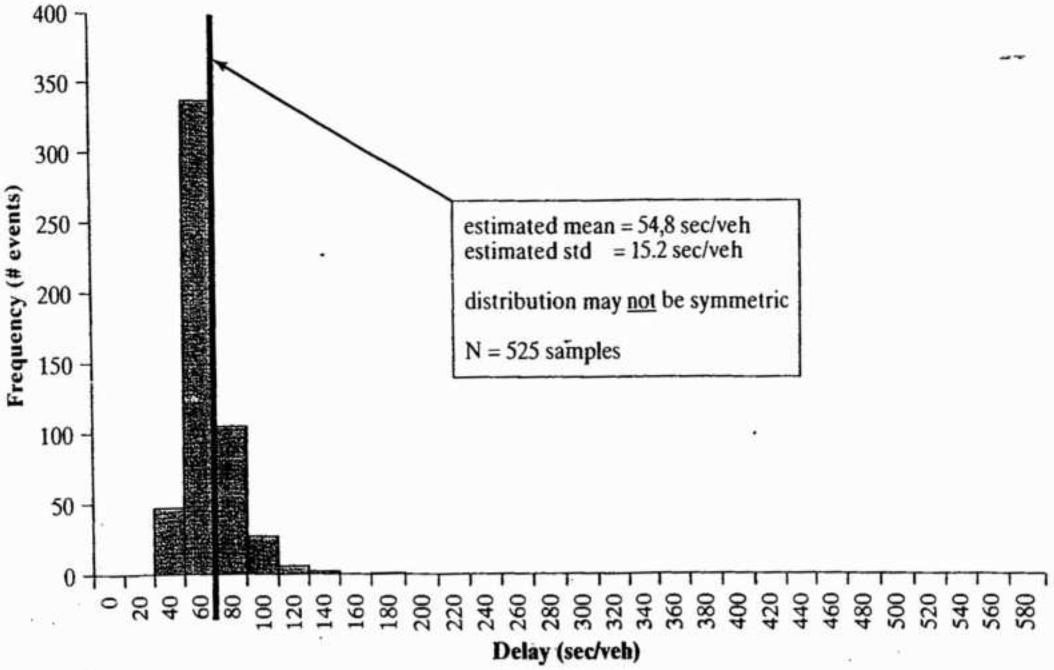
We did additional analyses that showed that even when $\sigma = 30$ s/veh, the asymmetric distribution occurs when $\{v/c\}$ is higher. The long-tailed delay distribution of Figure 27.11.b can be expected routinely and must be taken into account.

27.7.2 Practical Implications

Clearly, high $\{v/c\}$ ratios should be avoided, to avoid the large standard deviations demonstrated in even this simple illustration.

Table 27.3: An Illustrative Study

	demand	capacity	est avg v/c	est std v/c	delay based upon expected values	related std
mean=	1700	1900				
std=						
>Traditional	0	0	0.89	0.00	51.0	
>Case 1	30	30	0.90	0.02	54.8	15.2
>Case 2	60	60	0.90	0.04	70.5	73.8
					sec/veh	sec/veh

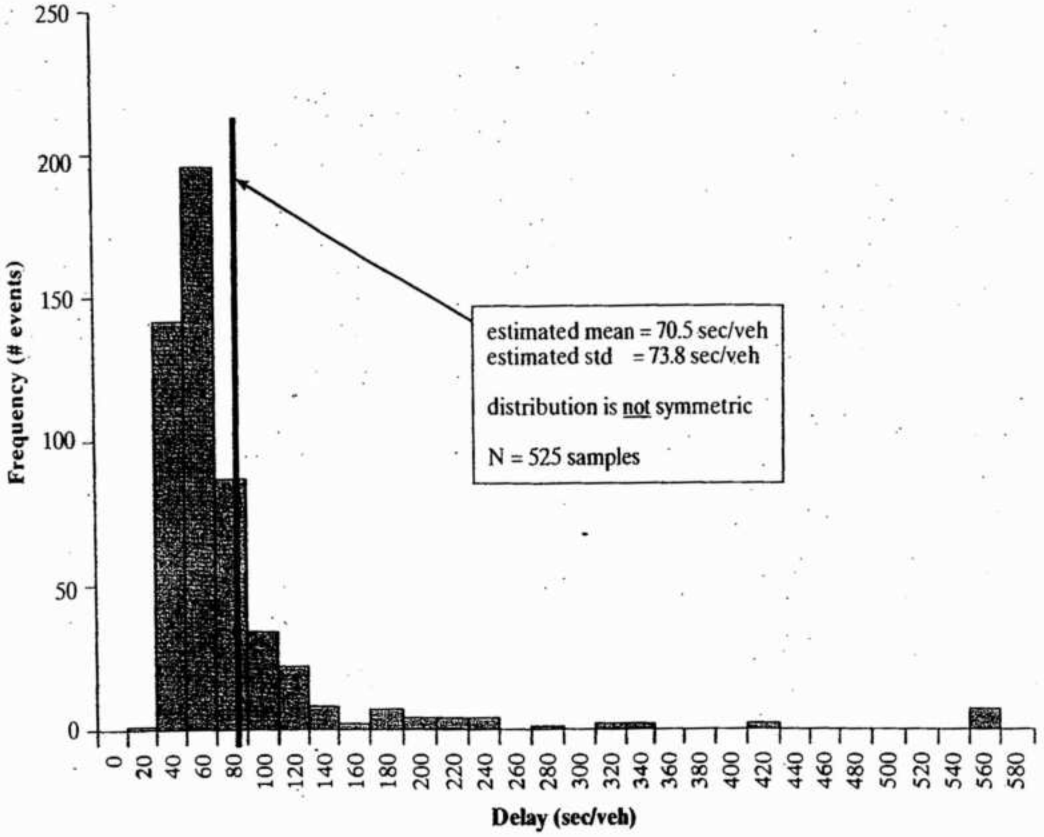


estimated mean = 54,8 sec/veh
estimated std = 15.2 sec/veh

distribution may <u>not</u> be symmetric

N = 525 samples

(a) Case 1: Standard Deviations of 30 vph



estimated mean = 70.5 sec/veh
estimated std = 73.8 sec/veh

distribution is <u>not</u> symmetric

N = 525 samples

(b) Case 2: Standard Deviations of 60 vph

Figure 27.11: Histograms of the "Observed" Delay for Cases 1 and 2

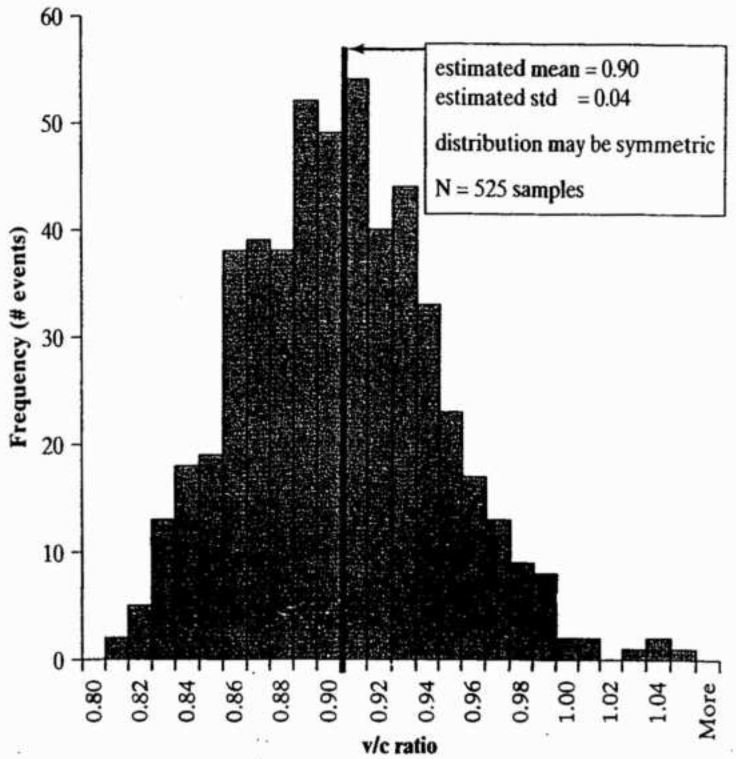


Figure 27.12: The Range of $\{v/c\}$ Values in Illustrative Case 2

But the subject of this chapter is the oversaturated condition, which occurs when v/c exceeds 1.00 systematically—or when randomness pushes it there, and the inability to "catch up" perpetuates it.

Capacity variations can be due to such factors as (a) "natural" variations due to the randomness of discharge headways, loss times, the percentage of different vehicle types in queues, the percentage of turns in a given time period, left-turners at the head of a queue in a shared lane, and even (b) singularities or "outliers" in some of these values due to double parkers, deliveries, buses hanging out of bus stops, and other such factors. Local "demand" variations can be due to true natural variations.

The "bottom line" is that as v/c approaches 1.00, the effects of variation have greater impact: the variability of delay (and thus travel time), the asymmetric tail, and the extreme values are all associated with $\{v/c\}$ being pushed near or above 1.00 in bursts.

27.7.3 A Closing Note on This Topic

Figure 27.11b results in an (estimated) average observed delay of 70.5 s/veh with the 525 observations, which raises two troublesome issues:

- The "observed" average is quite a bit different than the value computed using the averages of the inputs, namely 50.1 s/veh (see Table 27.3, "Traditional"), and the deterministic or "expected value" analysis may have been used past its valid range;
- How much data is going to be desired to estimate the observed average delay with good confidence, versus how much can the budget pay for?

Because the spread in Figure 27.11b is so high, the confidence bound on the estimate of the mean is $\pm\ 1.96\,(73.8)/\sqrt{525}$, or $\pm$ 6.3 s/veh. That is, after $N = 525$ observations, we could only be 95% sure that the true mean was somewhere between 70.5 $\pm$ 6.3 s/veh (i.e., in the range 64.2 to 76.8 s/veh).

27.8 Summary and Further Readings

This chapter addressed oversaturated conditions on surface streets. The problem of congestion and saturation is widespread and is not often approached consistently. Definite measures can be taken, but preventive action addressing the root causes must be given a high priority. Among the possible measures, those relating to signalization generally can have the greatest impact. The nonsignal remedies are in no way to be minimized, particularly those that provide space, whether for direct productivity increases or for removing impediments to the principal flow.

A wealth of information is available at the TRB Traffic Signal Systems Committee Web site [6], including the set of presentations at the Committee's January 2007 Workshop on "Operating Traffic Signal Systems in Oversaturated Conditions" and its July 2006 Mid-Year Meeting that focused on oversaturation. This material was also revisited in the context of this proposal. In the interest of space, the individual materials are not enumerated or commented on at this time. It is refreshing to see the number of case studies of actual applications that appear in this compendium of information.

To assist engineers and planners in applying HCM methodologies, the HCM 2000 includes default values for many of the more difficult-to-obtain input parameters and variables. Field measured default values for use in applications of the HCM were assembled in the research performed in NCHRP Report 599 [7], *Default Values for Highway Capacity and Level of Service Analyses*. Of the 63 default values used in the HCM 2000, that research found 19 values that had a high degree of impact on the service measures. For the HCM 2000 Urban Streets analysis, signal density was found to have a high impact. For signalized intersections, the peak-hour factor, length of analysis period, arrival type, adjusted saturation flow rate, lane width, percent heavy vehicles, and lane utilization were found to have high impact on the estimate of delay. These high-impact variables could also be considered for their impact on the variability in demand and capacity.

References

1. McShane, W.R., et al., "Traffic Control in Oversaturated Street Networks," National Cooperative Highway Research Program, (NCHRP) Report 194, Transportation Research Board, National Research Council, Washington, DC, 1978.
2. Denney, R.W., Head, L., Spencer, K., "Signal Timing Under Saturated Conditions," FHWA Office of Operations, FHWA-HOP-09-008, November 2008.
3. Lieberman, E., and Messer, C., "Internal Metering Policy for Oversaturated Networks," Final Report, NCHRP 3-38(4), 1992.
4. Lieberman, E.B., Chang, J., and Prassas, E.S., "Formulation of Real-Time Control Policy for Oversaturated Arterials," *Transportation Research Record 1727*, TRB, National Research Council, Washington DC, 2000.
5. "Signal Timing for Congested Conditions," ITE Professional Development Program, October 11, 2007 webinar, presented by Woody Hood, Maryland DOT.
6. TRB Traffic Signal Systems Committee (TRB Committee ANB25) Web site, http://www.signalsystems.org.vt.edu/
7. Zegeer, J., Vandehey, M., Blogg, M., Nyguen, K., and Ereti, M., *Default Values for Highway Capacity and Level of Service Analyses*, NCHRP Report 599, Transportation Research Board, National Research Council, Washington, DC, 2008.