Introduction to Traffic Signal Timing

Course No: C02-094
Credit: 2 PDH

Jeffrey W. Buckholz, PhD, P.E., PTOE

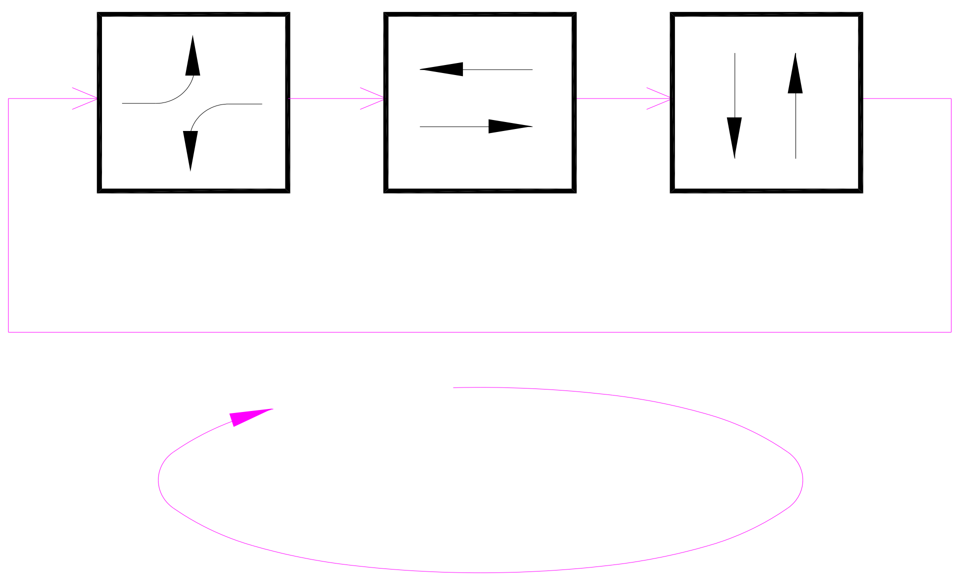
CED engineering.com
PDH FOR THE PROFESSIONAL

Continuing Education and Development, Inc.

P: (877) 322-5800
info@cedengineering.com

This Traffic Signal Operations Course presents information regarding the timing of traffic signals. The material is intended to illustrate the concepts involved in traffic signal timing for modern fully actuated traffic signal controllers. Only basic signal timing concepts at isolated (non-coordinated) intersections are covered; advanced concepts such as coordinated signal timing and traffic signal preemption timing are not addressed. However, this is not an introductory course and knowledge of basic traffic signal operation is assumed.

The cycle of a traffic signal is a complete sequence of signal phases and the time it takes a traffic signal to display this complete sequence is known as the "cycle length".

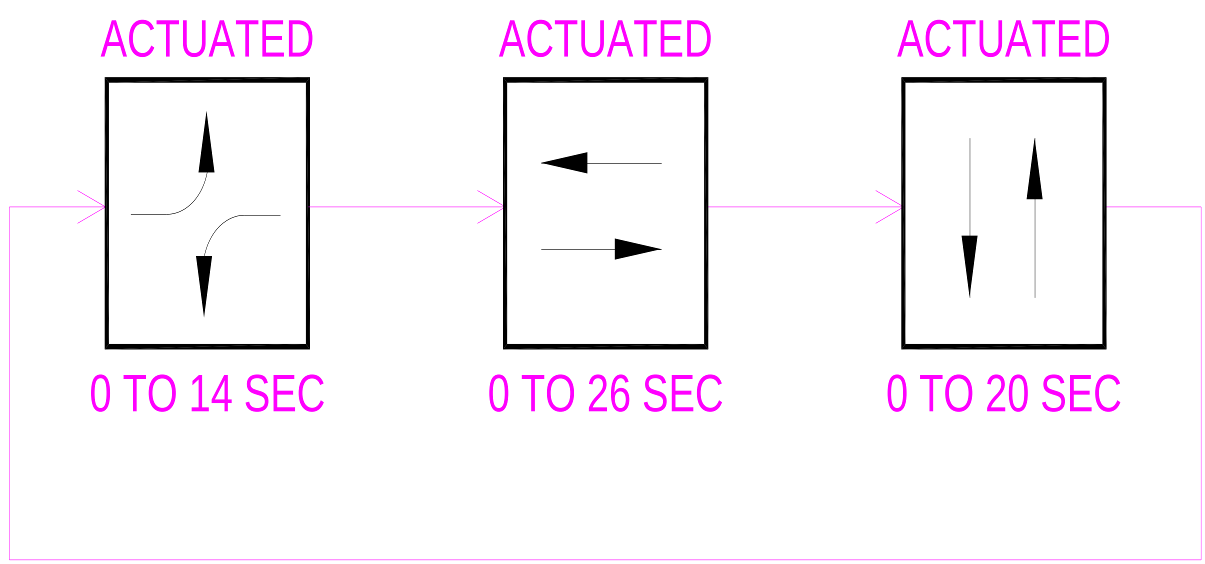


CYCLE = 1 COMPLETE LOOP

If the three phase combinations in the above diagram each last for 30 seconds, then the cycle length for this signal is 90 seconds. A cycle is sometimes referred to as a "dial" since, in old electro-mechanical controllers, a rotating dial was used to control the timing of the various signal displays. The amount of time during which a particular phase has a green indication (either ball or arrow) is known as the "green interval" for that phase

A traffic signal under **pre-timed control** has a set sequence of non-actuated phases with fixed green intervals that, for a given cycle length, do not vary. Different cycle lengths may be used during different times of the day but the green interval for each phase combination will not change for a given cycle length. A traffic signal under **semi-actuated control** has a mixture of actuated phases, the green time for which varies with the level of traffic demand, and non-actuated phases. The major street thru phases are typically the non-actuated phases and these phases receive a fixed amount of green time. Any of the actuated phases can be skipped during a given cycle if no vehicle demand is present, but the non-actuated phases are always served.

A traffic signal under **fully-actuated control** has no fixed cycle length since the green intervals for all of the various phases can vary with the level of traffic demand and since any phase can be skipped if no demand is present for that phase.



The cycle length will vary from cycle to cycle with a maximum cycle length for this example of 60 seconds. Fully-actuated controllers offer the most sophisticated timing features and capabilities so the remaining discussion centers on these types of controllers.

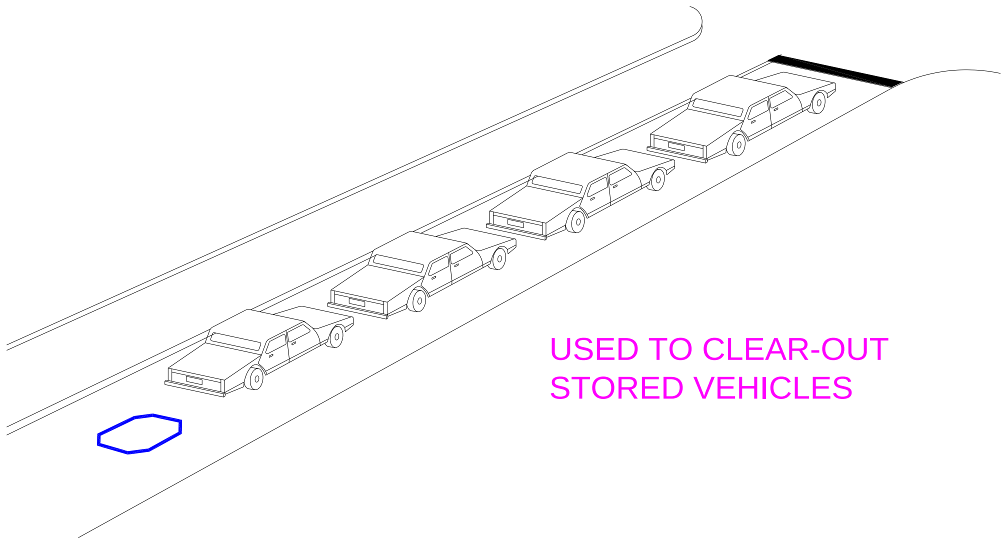
BASIC TIMING INTERVALS

When a signal phase receives the green, the **initial interval** for the phase begins timing. A phase will always receive at least the amount of time programmed into the initial interval. The clock within the controller that keeps track of the initial interval is known as the initial interval timer.

	PHASE							
	1	2	3	4	5	6	7	8
Approach	NB	SB	-	WB	SB	NB	-	EB
Movement	LT	THRU	-	ALL	LT	THRU	-	ALL
Initial	5	20	0	10	5	20	0	10
Passage	2	3	0	3	2	3	0	3
Yellow	3	4.5	0	4.5	3	4.5	0	4.5
All-Red	1	2	0	1	1	2	0	1
MAX I	15	40	0	25	25	40	0	25
MAX II	25	35	0	35	15	35	0	35
Walk	0	7	0	0	0	7	0	0
Ped. Clear	0	15	0	0	0	17	0	0

This timing chart is for an intersection where 6 NEMA phases are in use (1, 2, 4, 5, 6 and 8). The initial interval for phase 4 (the westbound movement) is 10 seconds. The clock within the controller that keeps track of the initial interval is known as the initial interval timer. In this example, the initial interval timer counts down from 10 to 0 every time phase 4 is displayed. The green indication for a phase cannot be terminated until the initial interval counts down to zero.

The purpose of the initial interval depends on whether stop line or upstream detection is being used. When the detection area (of the loop, video camera, infrared unit, etc.) is located upstream from the stop line, the purpose of the initial interval is to provide enough green time to service all vehicles that have queued between the stop line and the detection area.



.

When the detection area is located at the stop line, the purpose of the initial interval is to ensure that a reasonable minimum green time is displayed, not a too-short green time that might violate driver expectation or a too-long green time that is wasteful.

Although there is considerable debate amongst traffic professionals as to the appropriate values to use for initial intervals, the following ranges are reasonable:

Major Street Thru Movements:	10 to 20 seconds
Minor Street Thru Movements:	4 to 10 seconds
Turning Movements:	4 to 7 seconds

Because drivers tend to expect more green time for the thru movements on major roads than they do for the thru movements on minor roads, or for turning movements, the main street thru movements usually have the longest initial intervals.

The clock within the controller that keeps track of the **passage interval** is known as the passage interval timer. When the presence of a vehicle is no longer detected, the passage timer begins counting down, beginning at the passage value that has been programmed into the controller for that phase. If another vehicle is detected before the passage timer reaches zero, the passage timer will reset to the passage value. This resetting process is repeated as long as vehicles continue to be detected before the passage interval counts down to zero. The passage interval is reset every time a vehicle is detected. Because continued actuations will reset the passage interval and extend the green interval for a phase, the passage interval is sometimes referred to as the "extension" interval.

	PHASE							
	1	2	3	4	5	6	7	8
Approach	NB	SB	-	WB	SB	NB	-	EB
Movement	LT	THRU	-	ALL	LT	THRU	-	ALL
Initial	5	20	0	10	5	20	0	10
Passage	2	3	0	3	2	3	0	3
Yellow	3	4.5	0	4.5	3	4.5	0	4.5
All-Red	1	2	0	1	1	2	0	1
MAX I	15	40	0	25	25	40	0	25
MAX II	25	35	0	35	15	35	0	35
Walk	0	7	0	0	0	7	0	0
Ped. Clear	0	15	0	0	0	17	0	0

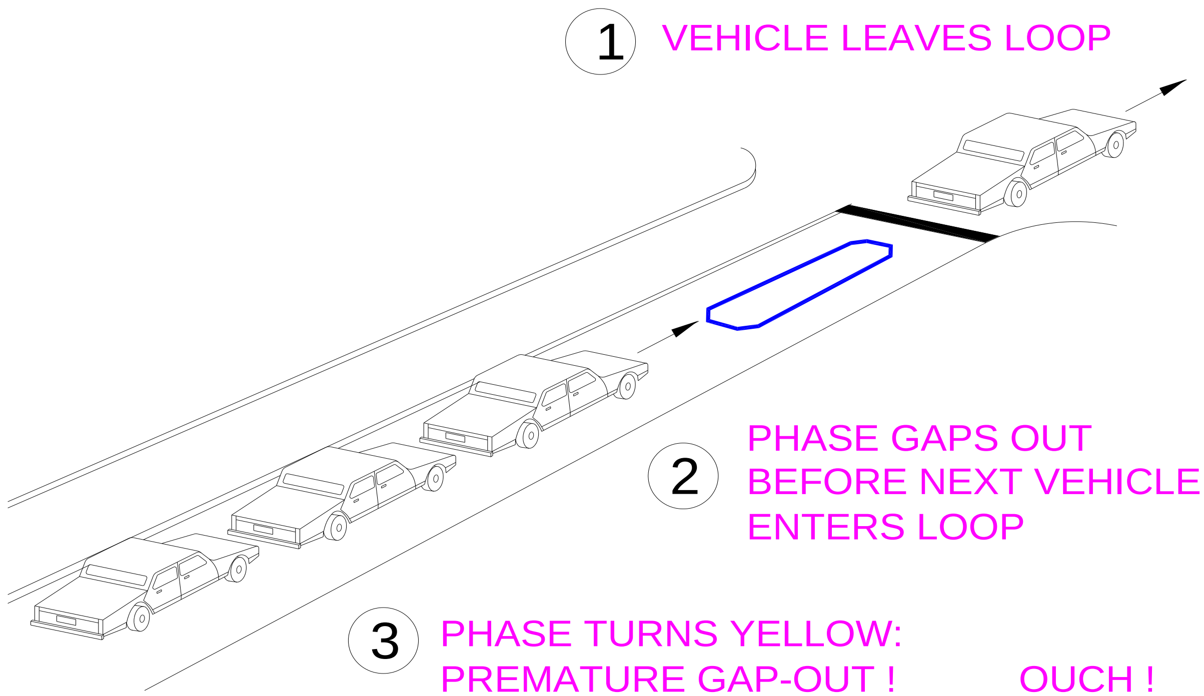
In our example, the passage interval for NEMA phase 5, the southbound left turn, is set at 2 seconds.

If the passage timer reaches zero the phase is said to "gap out" and, unless some other controller function prevents it (such as an initial interval timer that is still counting down), the phase will turn yellow and the right-of-way will be transferred to the next phase where demand is present.

The length of the passage interval is heavily dependent upon the design of the detection system. Longer passage intervals tend to be associated with short detection areas whereas shorter passage intervals are usually associated with long detection areas. The following range of passage intervals is reasonable:

50-foot, or longer, detection area:	0 to 2 seconds
40-foot detection area:	1 to 2.5 seconds
30-foot detection area:	2 to 3 seconds
20-foot detection area:	3 to 4 seconds
6-foot detection area:	4 to 5 seconds

After the last vehicle leaves the detection area, the passage interval must count down to zero before the controller can transfer the right of way to the next phase. Consequently, shorter passage intervals allow the controller to recognize earlier that demand for the phase has ended, reducing the amount of wasted green time at the end of the phase. However, if the passage interval is set at a value that is too short for the detection area, then premature gap out can occur. Premature gap out occurs when a vehicle leaves the detection area and the passage timer counts down to zero before the next vehicle waiting in line is able to enter the detection area. This causes the signal to gap out even though there may be a substantial number of vehicles still waiting in the queue.



Under certain circumstances, single lane approaches may experience large gaps between successive vehicles - requiring longer than normal passage intervals to avoid premature gap out. Circumstances that might produce larger-than-normal gaps between vehicles include: a significant number of slow trucks, a preponderance of elderly individuals with slow reaction times, adverse weather conditions, or the presence of a steep upgrade.

When long detection areas with short passage intervals are used, the last motorist to leave the loop will see the yellow indication appear almost immediately upon exiting the detection area. Although this constitutes efficient operation, the quick onset of the yellow interval makes some drivers uneasy and may result in periodic complaints, at least until motorists become accustomed to this type of operation.

Before the green can be transferred across the barrier due to gap out, all phases that are currently timing must gap out. For our example, if phases 2 and 6 are timing, both phase 2 and phase 6 must gap out before phase 4 or 8 can be serviced. If the passage timer of one phase reaches zero while the passage timer of the other active phase is still counting down, the right-of-way will not transfer. If the dormant phase subsequently receives another actuation its passage timer will be reset and the phase green will continue to be displayed, even if the previously active phase now counts down to zero. When the controller is programmed such that the passage timer can reset and all active phases must be at zero at the same instance for the right-of-way to transfer, "simultaneous gap out" is said to be in effect. Most modern controller's can be programmed so that once a phase's passage timer has reached zero it cannot be reset. If the controller has been programmed in this way, then simultaneous gap out is said to be inhibited. The end result is that simultaneous gap out provides more green time for the phases that have it active and less green time for other phases at the intersection.

If the passage interval cannot begin counting down until the initial interval has finished timing, "sequential timing" of the initial and passage intervals is said to be in effect. Under sequential timing one passage interval will always be timed, consequently, the minimum green interval for the phase will be equal to the sum of the initial and passage intervals. Using our example, if sequential timing is active then the minimum green time for phase 2 would be 23 seconds. If the passage interval can count down at the same time that the initial interval is counting down, "concurrent timing" of the initial and passage intervals is said to be in effect. Under concurrent timing, the minimum green interval for the phase will be equal to the initial interval (20 seconds for phase 2).

The **maximum interval** keeps the green indication from being extended an unreasonably long period of time if the passage timer fails to gap out. This can happen if there is constant demand for the green on that phase or if a detection failure (such as a stuck detector or a damaged loop) produces a continuous call. The clock within the controller that keeps track of the maximum interval is known as the maximum interval timer. The maximum interval timer begins counting down as soon as vehicle demand is recognized on a conflicting phase. The value at which the timer begins counting is the maximum interval that has been programmed into the controller.

	PHASE							
	1	2	3	4	5	6	7	8
Approach	NB	SB	-	WB	SB	NB	-	EB
Movement	LT	THRU	-	ALL	LT	THRU	-	ALL
Initial	5	20	0	10	5	20	0	10
Passage	2	3	0	3	2	3	0	3
Yellow	3	4.5	0	4.5	3	4.5	0	4.5
All-Red	1	2	0	1	1	2	0	1
MAX I	15	40	0	25	25	40	0	25
MAX II	25	35	0	35	15	35	0	35
Walk	0	7	0	0	0	7	0	0
Ped. Clear	0	15	0	0	0	17	0	0

In our example, 40 is the maximum interval programmed for phase 6.

If the maximum timer reaches zero before the phase has gapped out, the phase will "max out". The signal heading for this phase will provide a yellow indication and the right-of-way will be transferred to the next phase where demand is present.

The maximum interval that each phase should have is dependent upon the traffic volumes handled during that phase. Phases that serve movements with high traffic demand will typically require longer maximum intervals than phases with lesser demand. The end result is that the major street thru phase typically has a higher maximum interval than the cross street thru phase, with left turn phases having the lowest maximum intervals. Since vehicles can turn during the permissive interval at left turn lanes controlled by protected/permissive phasing, protected/permissive phases typically require less of a maximum interval than protected-only left turn phases.

Although a wide variation in maximum intervals are used depending on the unique traffic volume characteristics of an intersection, the following range of maximum intervals is fairly typical:

Major Street Thru Phase:	30 to 60 seconds
Minor Street Thru Phase:	25 to 40 seconds
Protected-Only Left Turn Phase:	20 to 30 seconds
Protected/Permissive Left Turn Phase:	15 to 20 seconds

Most modern controllers allow the user to invoke a second set of maximum intervals, referred to as "MAX II" intervals, if so desired. These maximum intervals may be either shorter or longer than the normal maximum (MAX I) intervals depending on their purpose. If MAX II intervals are used during late night operation, then they will be shorter than the MAX I intervals. In our example, phase 5 has a 15 second Max II interval that is considerably shorter than its 25 second Max I interval. However, if MAX II Intervals are used to provide additional time during peak hours, then they will be longer. MAX II intervals are activated by a time clock, either one internal to the controller or one from some external device, such as a time switch.

Some controllers provide more than two maximum intervals (MAX III, MAX IV, etc.) for the ultimate in timing flexibility. Certain controllers also have the ability to sense if a phase is repeatedly maxing-out (which indicates that not all waiting vehicles are making it through on the green) and, if it is, can increase the maximum interval for that phase. The maximum is increased until the phase begins gapping-out instead of maxing-out (which indicates that all waiting vehicles are now being served). When this occurs, the maximum interval will start shrinking towards its original value.

Another progressive feature that some controllers have is the ability to time a different maximum interval when the phase is maxing-out because the detector has failed, rather than in response to actual vehicle demand. When a detector failure is sensed by the controller, the phase is placed on "fail max recall" and a different (typically shorter) maximum interval is invoked. This reduces the considerable wasted green time that typically results when the normal maximum interval is used during detector failure.

PHASE TERMINATION

For signals that are not running under coordination and are not being affected by signal preemption (topics not covered in this course), there are only three ways for a phase to terminate:

1. Gap Out
2. Max Out
3. Enter Program or Conflict Flash

Program flash is said to occur when, at a certain time of the day, the signal purposely goes into flash mode. This typically occurs late at night when traffic volumes drop and the need to assign right-of-way via signalization is no longer needed. Conflict flash occurs automatically when the conflict monitor senses an electrical malfunction that could result in the display of unsafe signal indications, such as conflicting greens. The signal is immediately transferred into flash mode.

At efficiently timed traffic signals during normal operations the signal will tend to max out during peak hours when traffic demands are heavy and gap out during the rest of the day.

Traffic Engineers use the term "cycle failure" (or, more accurately, “phase failure”) to refer to the situation where some of the vehicles waiting in a lane are not able to make it thru the intersection on the first green indication. Repeated phase failures on an approach may be caused by a number of problems, including passage intervals that are set too short for the detection layout that is in place or maximum intervals that are set too short for the level of vehicular demand.

CHANGE AND CLEARANCE INTERVALS

The Manual On Uniform Traffic Control Devices (MUTCD) requires that all yellow change and red clearance intervals be determined using engineering practice.

The purpose of the **yellow change interval** is to alert the driver that the right-of-way is changing due to phase termination and, depending on the vehicle's location relative to the stop line, to permit the driver to stop or to provide enough time for the driver to legally enter the intersection (prior to the display of the red indication).

The minimum length of the yellow change interval is typically calculated using the extended kinematic formula recommended by the Institute of Transportation Engineers (ITE):

Yellow Change Interval = $V_E/[(2 \times a) + (64.4 \times g)] + (V_{85} - V_E)/[a + (32.2 \times g)] + t$

where: V_E = Intersection Entry Speed in feet per second
V_{85} = 85th Percentile Approach Speed in feet per second
a = Deceleration Rate in ft/sec/sec (10 is typically used)
g = Grade of Approach in percent/100 (negative if downhill)
t = Reaction Time in seconds (1.0 is typically used)

The intersection entry speed is measured at the stop line while the 85th percentile approach speed is measured upstream of the intersection. For through movements the 85th percentile approach speed (if unknown) can be estimated by adding 10 fps to the posted speed limit for the approach. For left turn movements, the intersection entry speed (if unknown) can be estimated at 29 fps (20 mph).

Based on this formula, the length of the yellow interval increases with increasing approach speed. The formula also produces longer yellow intervals for approaches located on downgrades and shorter yellow intervals for approaches located on upgrades.

Many modern controllers do not recognize a yellow interval of less than 3 seconds. If the user attempts to program a value of less than 3 seconds in these controllers, the controller will ignore it and use 3 seconds instead.

The MUTCD requires that the duration of a yellow change interval not change on a cycle-by-cycle basis although different yellow change intervals can be used for different signal timing plans.

The purpose of the **red clearance interval** is to provide a buffer of time between the termination of the yellow on one phase and the subsequent display of the green interval on a conflicting phase.

ITE recommends use of the following formula for establishing the red clearance interval:

Red Clearance Interval = $(W+L)/V_E - t_s$

where: V_E = Intersection Entry Speed in feet per second
W = Width of Intersection (distance traversed from stop line to far edge of intersecting street) in feet
L = Length of the Vehicle in feet (20 is typically used)
t_s = Conflicting Vehicle Start-Up Delay in seconds (this is an optional parameter that is often conservatively set to zero)

Although long all-red intervals provide an increased safety buffer between conflicting vehicle movements, there are drawbacks. The use of long all-red intervals results in a considerable amount of wasted time at an intersection. It can also encourage motorists to become more aggressive and "run the red" since they quickly learn that there is a significant dead period before waiting vehicles begin to move.

Some advanced controllers allow the yellow and all-red intervals to be automatically increased by a specified percentage during adverse weather. This provides an additional margin of motorist safety when roads are slick and visibility is reduced.

PEDESTRIAN TIMING

Where signalized pedestrian crossings are used, the value of the WALK interval and flashing DON'T WALK interval must be established. The flashing DON'T WALK interval is also referred to as the "pedestrian change interval”.

The Manual on Uniform Traffic Control Devices (MUTCD) recommends the use of a WALK interval of at least 7 seconds, with 4 seconds being the absolute minimum.

WALK intervals of greater than 7 seconds may be used where the duration of the associated vehicular green interval is long enough to accommodate it. The purpose of the WALK interval is to provide enough time for a group of waiting pedestrians to enter the crosswalk before the flashing DON'T WALK commences.

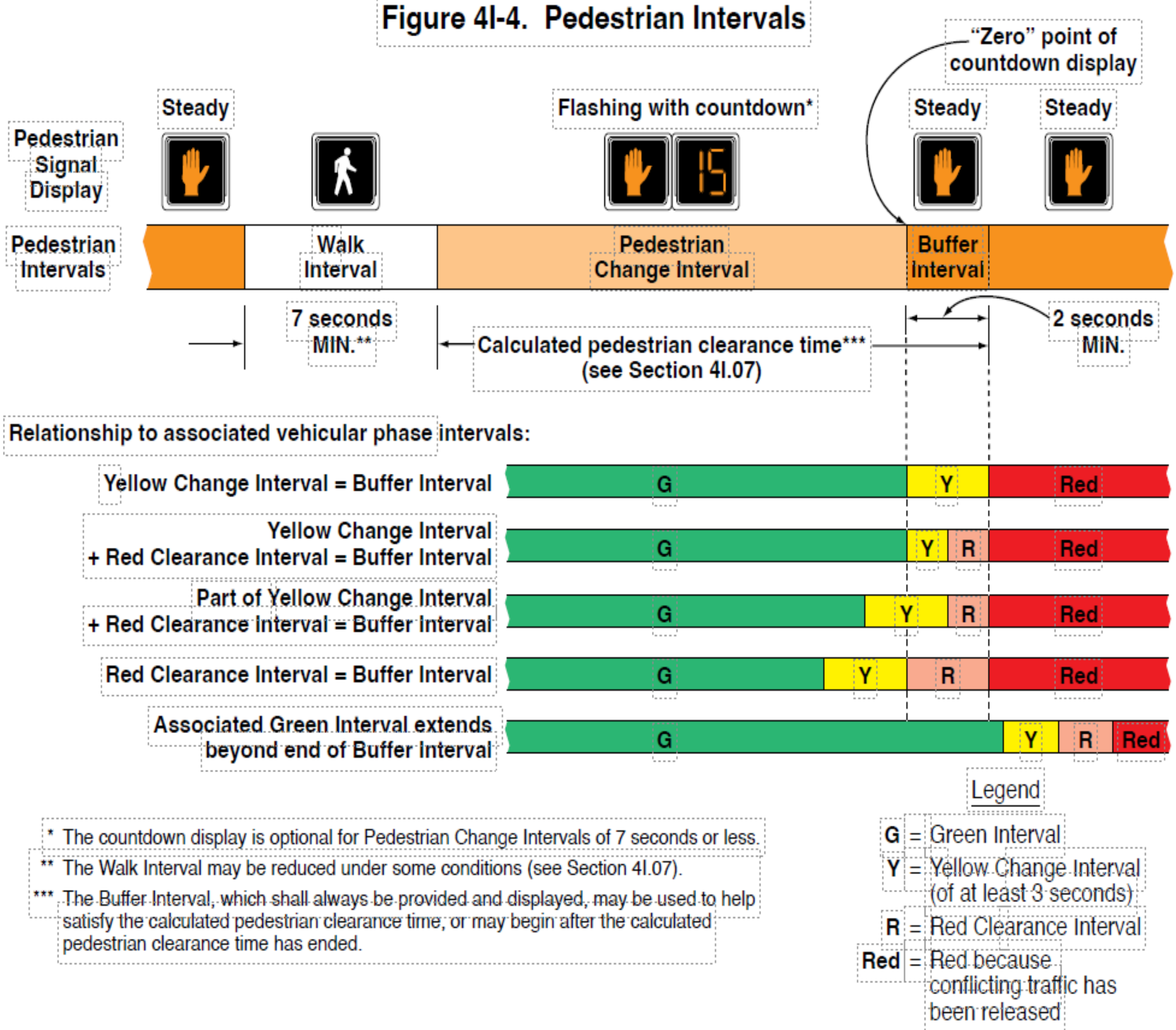
The required “**pedestrian clearance time**” is the time needed by a pedestrian who has just stepped into the crosswalk to reach a safe location at the other end of the crosswalk before the green indication is given to conflicting traffic. According to the MUTCD, the pedestrian clearance time should be long enough to allow a pedestrian walking at **3.5 fps** (feet per second) to travel from one side of the street to the other. For an 84-foot crossing, a pedestrian clearance time of 24 seconds would be required.

The MUTCD also recommends that the total of the WALK and pedestrian clearance time be sufficient to allow a person travelling at **3 fps** to cross from a distance 6 feet before the face of the curb (or edge of pavement) to the other side of the street. The total distance for our example would be 84 + 6 = 90 feet, which would require 30 seconds (90 / 3) of total pedestrian crossing time. 30 - 24 = 6, so we would need to program at least 6 seconds of WALK time.

Long pedestrian timings can play havoc within efficient intersection operation. Therefore, to keep the amount of time required for the pedestrian change interval at a reasonable value, many agencies subtract the associated yellow change and red clearance intervals from the required pedestrian clearance time to obtain the required pedestrian change interval. The MUTCD allows this practice, with the last 2 seconds of this change time being referred to as the "buffer interval". Continuing our example, if we have a 4 second yellow interval and a 3 second all-red interval for the associated vehicular phase, the required pedestrian change (flashing DON'T WALK) interval would be 24 - 7 = 17 seconds with the last 2 seconds of the 7 second change interval serving as the minimum buffer interval.

The following figure from the MUTCD summarizes pedestrian interval relationships:

Figure 4I-4. Pedestrian Intervals



* The countdown display is optional for Pedestrian Change Intervals of 7 seconds or less.

** The Walk Interval may be reduced under some conditions (see Section 4I.07).

*** The Buffer Interval, which shall always be provided and displayed, may be used to help satisfy the calculated pedestrian clearance time, or may begin after the calculated pedestrian clearance time has ended.

Countdown pedestrian signals, which are required by the MUTCD when the pedestrian change interval is longer than 7 seconds, display the number of seconds remaining until the termination of the pedestrian change interval.

Countdown pedestrian signals have become very popular with traffic engineers and the public alike due to the valuable information they provide crossing pedestrians.

Pedestrian walking speeds vary substantially depending on the type of pedestrian. The average pedestrian walks at a speed of approximately 4 fps while small school children and the elderly typically have a slower walking speed of about 3.5 fps. The frail elderly can have a very slow walking speed, with 2.5 fps not being an unreasonable value for these individuals. As discussed previously, the MUTCD requires the use of a walking speed no greater than 3.5 fps for calculating the pedestrian clearance time. However, a speed of 4 fps may be used at locations where an "extended pushbutton press function" is provided. This function allows extra crossing time to be requested by slower pedestrians by extended actuation of the pedestrian detector.

Some agencies use long walk and pedestrian clearance intervals at all of their pedestrian crossings, regardless of the type of pedestrians that frequent the crossing. Using long pedestrian times at crossings that do not warrant their use can produce a substantial amount of wasted green time and severely diminish the efficiency of the intersection. The use of unduly long pedestrian times can also have a negative effect on coordinated signal systems by forcing the use of longer-than-desired cycle lengths.

Many modern controllers allow the use of more than one set of basic timing intervals during the day. Initial intervals, passage intervals, change and clearance intervals, and even pedestrian timings can be varied to meet changing conditions at the intersection. Each set of basic timing intervals is referred to as a "service plan".

VOLUME DENSITY OPERATION

Volume density operation is a special type of actuated signal control that, when properly used, often improves the efficiency of the signal. The characteristic of volume density operation that results in this improved efficiency is the gradual reduction of the extension (passage) interval over time; referred to as "gap reduction". As the extension interval shrinks the probability of gap-out increases. This ensures that waiting vehicles will promptly receive the green unless a tightly grouped stream of vehicles is being served. Under volume density control, stragglers will not continue to extend the green.

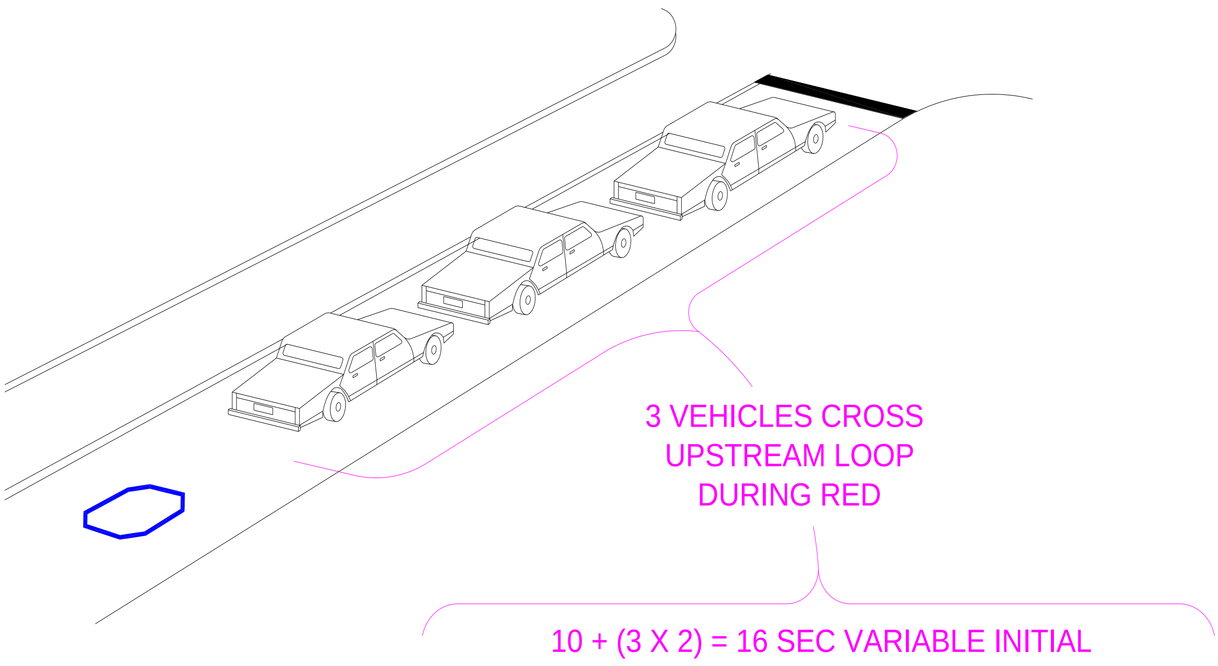
The time from the beginning of the phase's green interval to the point and which its extension begins to shrink is called the "time before reduction". A shorter time before reduction will result in quicker termination of the phase being timed and prompter service to waiting vehicles. The lowest value to which the extension will be allowed to shrink is called the "minimum extension". The extension will shrink, over time, from the value programmed into the passage interval to this value. The time from the point at which the extension begins to shrink to the point at which the extension reaches its minimum value is called the "time to reduce". As with the time before reduction setting, a shorter time to reduce setting will result in quicker termination of the phase being timed and prompter service to waiting vehicles.

Volume density operation is best illustrated with an example. A typical intersection approach using volume density operation might have the following settings:

- Initial Extension 4 seconds
- Minimum Extension 2 seconds
- Time Before Reduction 15 seconds
- Time to Reduce 20 seconds
- Initial Interval 10 seconds
- Seconds per Actuation 2 seconds
- Maximum Variable Initial 18 seconds

With these settings, the extension interval will begin to shrink 15 seconds after the phase turns green. The extension interval will drop a total of 2 seconds (4 seconds initial extension - 2 second minimum extension) over 20 seconds producing a 1 second reduction for every 10 seconds of time elapsed. Consequently, 25 seconds after the phase turns green the extension will have shrunk to 3 seconds and 10 seconds after that (35 seconds after the phase turns green) the extension will have bottomed-out at 2 seconds. This reduction in passage time makes it easier for gap-out to be achieved.

Upstream detection and a variable initial interval are typically used with volume density operation. Every time that a vehicle crosses the upstream detection area during the red or yellow interval of the phase an additional amount of time is added to the initial interval, hence the term "variable initial".



The amount of time, in seconds, that is added to the initial interval each time a vehicle detection occurs is called "seconds per actuation". In this example, 2 seconds is added for each of 3 vehicles arriving to produce a variable initial interval of 16 seconds for this particular cycle. When the approach is serviced, it will be guaranteed at least 16 seconds of green time. The use of a variable initial interval ensures that all vehicles that have accumulated between the upstream loop and the stop line during the phase's red period will be serviced when the phase turns green. The longest variable initial interval that can be obtained is referred to as the "maximum variable initial" (18 seconds in this example). Additional actuations will no longer increase the variable initial interval once this ceiling value has been reached.