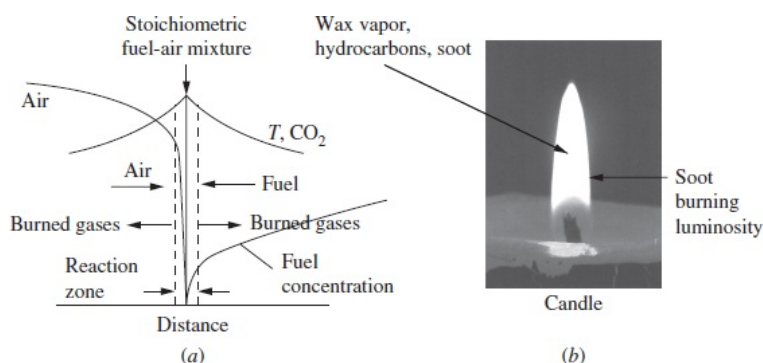


and the pressure rise it produces, would be much too slow. Because engine turbulence scales with piston speed, the combustion rate also scales with piston speed and thus is fast enough throughout the full crankshaft speed range desired for satisfactory engine operation (see [Sec. 8.2.2](#)). However, the turbulent flow primarily affects the gross or larger-scale characteristics of the flame (of order 1 mm) such as its average burning velocity. Locally, the thin reaction sheet flame propagates forward into the adjacent unburned mixture at the laminar flame speed.

Combustion in the compression-ignition or diesel engine occurs through a different sequence of processes. Fuel is injected directly into the engine cylinder toward the end of the compression process just before the desired start of combustion. The injected liquid fuel jets atomize, the fuel droplets move through the high-temperature high-pressure air in the combustion chamber, vaporize, and the fuel vapor mixes with air. The high temperature fuel-vapor air mixture thus produced then ignites spontaneously and burns rapidly. Combustion continues as combustible mixture is prepared by the above mixing process. Thus, for much of the diesel's combustion process, the fuel-burning rate is controlled by the rate at which fuel mixes with air. Such "non-premixed flames" are usually termed *diffusion flames* because the (gaseous) fuel and the air must diffuse into each other and then react in a thin reaction sheet close to where they reach the stoichiometric mixture proportions (see [Fig. 9.3](#)). A candle flame, also illustrated, is a common example of a diffusion flame. Because the in-cylinder engine flow is turbulent, the diesel combustion process occurs primarily as a turbulent (unsteady) diffusion flame. The spontaneous ignition due to the high air temperature in the cylinder when injection commences, and that fuel-air mixing controls the burning rate, are the distinguishing features of the diesel combustion process (the theme of [Chap. 10](#)).

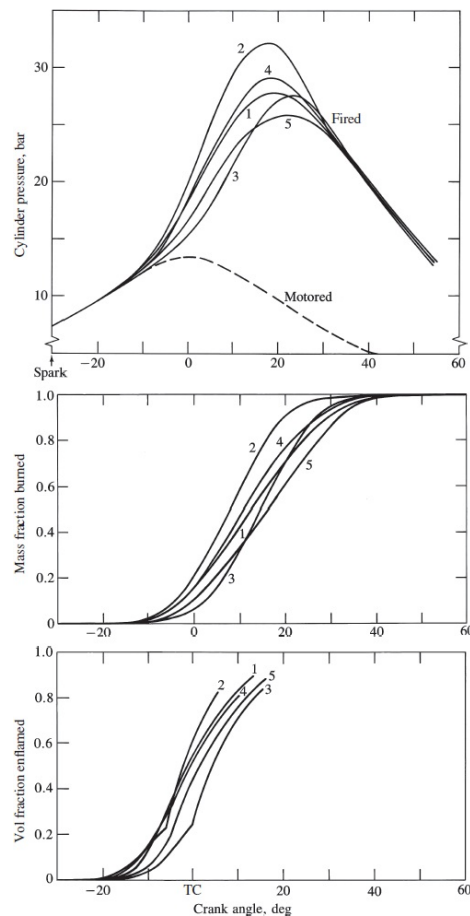
Figure 9.3 Non-premixed or diffusion flame, the basic diesel combustion process. (a) Diffusion flame structure. (b) Example of diffusion flame (candle flame): fuel-rich mixture (wax vapor, partly reacted hydrocarbons, some air, soot particles) inside the luminous flame; flame luminosity results from soot particles burning up in the thin flame reaction zone.



9.1.2. SI Engine Combustion Process

Additional features of the SI engine combustion process are evident from the data in [Fig. 9.4](#), taken from several consecutive cycles of an operating spark-ignition engine. The cylinder pressure, fraction of the charge mass which has burned (determined from the pressure data), and fraction of the cylinder volume enflamed by the front (determined from photographs like [Fig. 9.2](#)) are shown, all as a function of crank angle.⁴ Following spark discharge, there is a period during which the energy release from the developing flame is too small for the pressure rise due to combustion to be discerned. Then, as the flame continues to grow and propagate across the combustion chamber, the pressure steadily rises above the value it would have in the absence of combustion. The pressure reaches a maximum after TC but before the cylinder charge is fully burned, and then decreases as the cylinder volume continues to increase during the remainder of the expansion stroke.

Figure 9.4 Cylinder pressure, mass fraction burned, and volume fraction enflamed for five consecutive cycles in a spark-ignition engine as a function of crank angle. Ignition timing 30° BTC, wide-open throttle, 1044 rev/min, $\phi = 0.98$.⁴



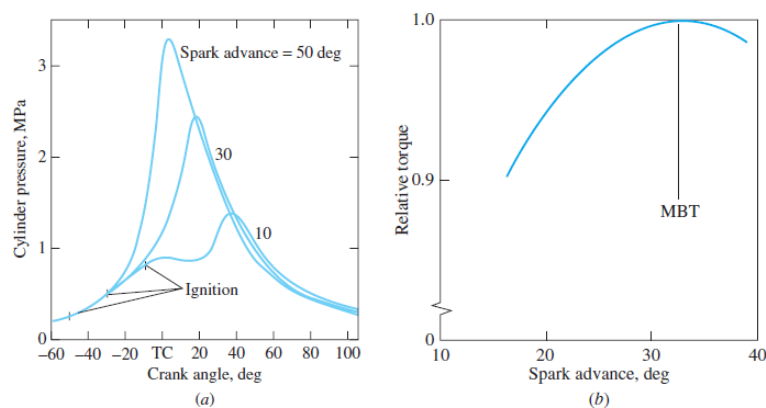
The flame development and subsequent propagation obviously vary, cycle-by-cycle, since the shape of the pressure, volume fraction enflamed, and mass fraction burned curves for each cycle differ significantly. This is because flame growth depends on local mixture motion and composition. These quantities vary in successive cycles in any given cylinder and may vary cylinder-to-cylinder. Especially significant are mixture motion and composition in the vicinity of the spark plug at the time of spark discharge since these govern the early stages of flame development. Cycle-by-cycle and cylinder-to-cylinder variations in this combustion process are important: see [Sec. 9.4.1](#).

Note that the volume fraction enflamed curves rise more steeply than the mass fraction burned curves. In large part, this is because the density of the unburned mixture ahead of the flame is about four times the density of the burned gases behind the flame. Also, there is some unburned mixture behind the visible front to the flame: even when the entire combustion chamber is fully enflamed, some 25% of the mass has still to burn. From this description it is plausible to divide the combustion process into four distinct phases: (1) spark ignition; (2) early flame development; (3) flame propagation; and (4) flame termination. Our understanding of each of these phases will be developed in this chapter.

The combustion event must be properly phased relative to top-center piston position to obtain the maximum power or torque. The combined duration of the flame development and propagation process is typically between 40 and 60 crank-angle degrees. Combustion starts before the end of the compression stroke, continues through the early part of the expansion stroke, and ends after the point in the cycle at which the peak cylinder pressure occurs. The pressure-versus-crank-angle curves shown in [Fig. 9.5a](#) allow us to understand why engine torque (at given engine speed and intake manifold conditions) varies as spark timing is varied relative to piston position, and thus crank angle. If the start of the combustion process is progressively advanced before TC, the compression stroke work transfer (which is *from* the piston to the cylinder gases) increases, as does heat transfer to the combustion chamber walls. If the end of the combustion process is progressively

delayed by retarding the spark timing, the peak cylinder pressure occurs later in the expansion stroke and is reduced in magnitude. These changes reduce the expansion stroke work transfer *from* the in-cylinder gases to the piston. The optimum timing, which gives the maximum brake torque—called *maximum brake torque*, or *MBT*, *timing*—occurs when the magnitudes of these two opposing trends just offset each other. Timing which is advanced or retarded from this optimum gives lower torque. The optimum spark setting will depend on the rate of flame development and propagation, the length of the flame travel path across the combustion chamber, and the details of the flame termination process after it reaches the wall. These depend on engine design and operating conditions, and the properties of the fuel, air, burned gas mixture. Figure 9.5b shows the effect of variations in spark timing on brake torque for a typical naturally-aspirated spark-ignition engine. MBT timing is the appropriate reference point for the actual spark timing, which should be defined relative to MBT timing (e.g., 10 degrees retard from MBT).

Figure 9.5 (a) Cylinder pressure versus crank angle for over-advanced spark timing (50°), MBT timing (30°), and retarded timing (10°). (b) Effect of spark advance (before TC) on brake torque at constant speed and (A/F), at wide-open throttle. MBT is maximum brake torque timing.⁵



Empirical rules for relating the mass burning profile and maximum cylinder pressure to crank angle at MBT timing or 50% mass burned fraction are often used. For example, with optimum spark timing: (1) the maximum pressure occurs at some 13 to 15° after TC; (2) half the charge is burned at about 8° after TC. In practice, the spark is often retarded to give a 1 or 2% reduction in brake torque from the maximum value, to lower peak cylinder pressures and burned gas temperatures. Substantial spark retard is used at times to achieve certain practical benefits (e.g., during cold engine starting to achieve fast catalyst light off, or at high load in turbocharged engines to hold peak pressure to acceptable levels).

So far we have described normal combustion in which the spark-ignited flame moves *steadily* across the combustion chamber until the charge is fully consumed. However, several factors—for example, fuel composition, certain engine design, and operating parameters—may prevent this normal combustion process from occurring. Two types of abnormal combustion have been identified: spontaneous ignition and knock.

Knock is the most important abnormal combustion phenomenon. Its name comes from the noise that results from the spontaneous autoignition of a portion of the fuel, air, residual gas mixture ahead of the advancing flame. As the flame propagates across the combustion chamber, the unburned mixture ahead of the flame—called the *end gas*—is compressed, causing its pressure, temperature, and density to increase. Some of the end-gas fuel air mixture may undergo chemical reactions prior to normal combustion. The products of these reactions may then *autoignite*: that is, spontaneously and rapidly release their chemical energy. When this happens, the end gas burns very rapidly, releasing its energy at a rate 5 to 25 times that characteristic of normal combustion. This causes high-frequency pressure oscillations inside the cylinder that produce the sharp metallic ringing noise called knock.

The presence or absence of knock reflects the outcome of a race between the advancing flame front and the autoignition chemical reactions in the unburned end gas. Knock will not occur if the flame front consumes the end gas before these reactions have time to cause the fuel-air mixture to autoignite. Knock will occur if the precombustion reactions produce autoignition before the flame front arrives. Thus, *spontaneous ignition* is an important abnormal combustion phenomenon.

Another form, *surface ignition*, is ignition of the fuel-air charge by a hot surface such as overheated valves or spark plugs, by glowing combustion-chamber deposits, or by any other hot spot in the engine combustion chamber. This may occur before the spark plug ignites the charge (preignition) or after normal ignition (postignition). Uncontrolled combustion is most evident and its effects are most severe when it results from preignition, since this can lead to heavy knock. However, even when surface ignition occurs after the spark plug fires (postignition), the spark discharge no longer has complete control of the combustion process. Thus, surface ignition may result in knock. Knock which occurs following normal spark ignition is called *spark knock* to distinguish it from knock which has been preceded by surface ignition.

As spark-ignition engine peak in-cylinder pressures have increased, a new form of preignition and knock has been observed: a sporadic preignition process leading to very high peak gas pressures. This "super, or mega, knock" process occurs sporadically (every 30,000 or so cycles), and lasts for one or a few cycles. Initiated by local autoignition, well before the spark discharge occurs, it generates these high pressures because it initiates an overly advanced combustion process leading to very high unburned-mixture pressures and temperatures that cause knock.

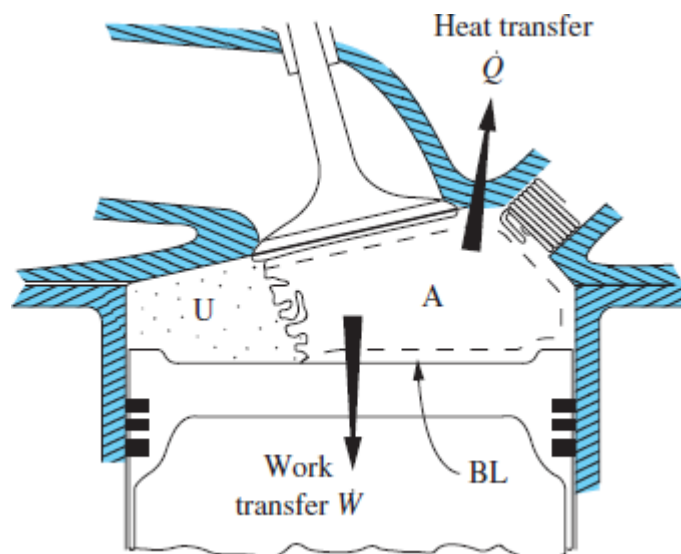
These abnormal combustion phenomena are reviewed in more detail in [Sec. 9.6](#).

9.2. THERMODYNAMICS OF SI ENGINE COMBUSTION

9.2.1. Burned and Unburned Mixture States

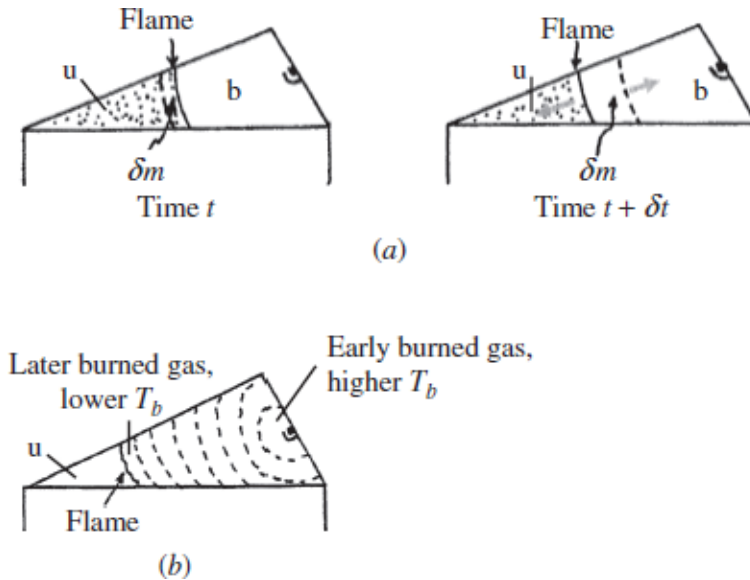
Consider the schematic of the engine cylinder while combustion is in progress shown in Fig. 9.6. Work transfer occurs between the cylinder gases and the piston (to the gas before TC; to the piston after TC). Heat transfer occurs to the chamber walls, primarily from the hot burned gases. At the temperatures and pressures typical of spark-ignition engines, it is a reasonable approximation to assume that the volume of the reaction zone where combustion is actually occurring is a negligible fraction of the chamber volume even though the thickness of the turbulent flame may not be negligible compared with the chamber dimensions (see Sec. 9.3.2). With normal engine operation, at any point in time or crank angle, the pressure throughout the cylinder is essentially uniform.

Figure 9.6 Schematic of turbulent premixed flame in the engine cylinder during combustion: unburned gas (U) to left of flame, burned gas to right. A denotes adiabatic burned-gas core, BL denotes thermal boundary layer in burned gas, $\dot{W}$ is work-transfer rate to piston, $\dot{Q}$ is heat-transfer rate to chamber walls.



Because combustion occurs through a flame propagation process, the changes in state and the motion of the unburned and burned gas are much more complex than the ideal cycle analysis in Chap. 5 suggests. The gas pressure, temperature, and density change as a result of changes in volume due to piston motion. Then, during combustion, the cylinder pressure increases due to the release of the fuel's chemical energy. As shown in Fig. 9.7a, as each element of fuel-air mixture burns and its temperature increases, its density decreases by about a factor of four. This combustion-produced gas expansion compresses the unburned mixture ahead of the flame and displaces it toward the combustion chamber walls. The combustion-produced gas expansion also compresses those parts of the charge which have already burned, and displaces them back toward the spark plug. During the combustion process, the unburned gas elements move away from the spark plug and the burned gas elements move back toward the spark plug. Further, elements of the unburned mixture which burned at different times have different pressures and temperatures just prior to combustion, and therefore end up at different states after combustion. The thermodynamic state and composition of the burned gas is, therefore, non-uniform. A first law analysis of the spark-ignition engine combustion process enables us to quantify these gas states.

Figure 9.7 Schematic of combustion- produced pressure rise and gas motion: (a) burning of mass element δm causes that element to expand, compressing and displacing unburned mixture (u) ahead of the flame and burned gas (b) behind the flame; (b) unmixed model for burned gas temperature distribution resulting from this sequential layer-by-layer burning process.



The conditions in the burned and unburned gas are then determined by conservation of mass:

$$\frac{V}{m} = \int_0^{x_b} v_b dx + \int_{x_b}^1 v_u dx$$

(9.1)

and conservation of energy:

$$\frac{U_0 - W - Q}{m} = \int_0^{x_b} u_b dx + \int_{x_b}^1 u_u dx$$

(9.2)

where V is the cylinder volume, m is the mass of the cylinder contents, v is the specific volume, x_b is the mass fraction burned, U_0 is the internal energy of the cylinder contents at some reference point θ_0 , u is the specific internal energy, W is the work done on the piston, and Q is the heat transfer to the walls. The subscripts u and b denote unburned and burned gas properties, respectively. The work and heat transfers are

$$W = \int_{V_0}^V p dV \quad Q = \int_{\theta_0}^{\theta} \left(\frac{\dot{Q}}{360N} \right) d\theta$$

(9.3)

where

$\dot{Q}$ is the instantaneous heat-transfer rate to the chamber walls.

To proceed further, models for the thermodynamic properties of the burned and unburned gases are required. Several categories of models are described in [Chap. 4](#). Accurate calculations of the state of the cylinder gases require an equilibrium model (or good approximation to it) for the burned gas and an ideal gas mixture model (of frozen composition) for the unburned gas (see [Table 4.2](#)). However, useful illustrative results can be obtained by assuming that the burned and unburned gases are different ideal gases, each with constant specific heats,⁶ that is,

$$pv_b = R_b T_b \quad u_b = c_{v,b} T_b + h_{f,b} \quad (9.4)$$

$$pv_u = R_u T_u \quad u_u = c_{v,u} T_u + h_{f,u} \quad (9.5)$$

Combining Eqs. (9.1) to (9.5) gives

$$\frac{pV}{m} = x_b R_b \bar{T}_b + (1 - x_b) R_u \bar{T}_u \quad (9.6)$$

and

$$\frac{U_0 - W - Q}{m} = x_b (c_{v,b} \bar{T}_b + h_{f,b}) + (1 - x_b) (c_{v,u} \bar{T}_u + h_{f,u}) \quad (9.7)$$

where

$$\bar{T} = \frac{1}{x_b} \int_0^{x_b} T_b dx \quad \text{and} \quad \bar{T}_u = \frac{1}{1 - x_b} \int_{x_b}^1 T_u dx$$

are the mean temperatures of the burned and unburned gases. Equations (9.6) and (9.7) may now be solved to obtain

$$x_b = \frac{pV - p_0 V_0 + (\gamma_b - 1)(W + Q) + (\gamma_b - \gamma_u) m c_{v,u} (\bar{T}_u - T_0)}{m [(\gamma_b - 1)(h_{f,u} - h_{f,b}) + (\gamma_b - \gamma_u) c_{v,u} \bar{T}_u]} \quad (9.8)$$

and

$$\bar{T}_b = \frac{R_u \bar{T}_u}{R_b} + \frac{pV - m R_u \bar{T}_u}{m R_b x_b}$$

(9.9)

If we now assume that the unburned gas is initially uniform and undergoes isentropic compression, then

$$\frac{\bar{T}_u}{T_0} = \left(\frac{p}{p_0} \right)^{(\gamma_u - 1)/\gamma_u}$$

(9.10)

This equation, with Eqs. (9.8) and (9.9), enables determination of both x_b and

$\bar{T}_b$ from the thermodynamic properties of the burned and unburned gases, and known values of p , V , m , and

$\dot{Q}$. Alternatively, if x_b is known then p can be determined. Mass fraction burned and cylinder gas pressure are uniquely related.

While Eq. (9.9) defines a mean burned gas temperature, the burned gas is not uniform. As illustrated in Fig. 9.7, mixture which burns early in the combustion process is further compressed after combustion as the remainder of the charge is burned.

Mixture, which burns late in the combustion process, is compressed prior to combustion and, therefore, ends up at a different final state. A temperature gradient exists across the burned gas with the earlier burning portions at the higher temperature.^{7,8}

Two limiting models bracket what occurs in practice: (1) a *fully mixed* model, where it is assumed that each element of mixture which burns mixes instantaneously with the already burned gases (which therefore have a uniform temperature), and (2) an *unmixed* model, where it is assumed that no mixing occurs between gas elements which burn at different times.

In the fully mixed model, the burned gas is assumed to be uniform,

$T_b = \bar{T}_b$, and the equations given above fully define the state of the cylinder contents. In the unmixed model, shown in Fig. 9.7b, the assumption is made that no mixing occurs between gas elements that burn at different times, and each burned gas element is therefore isentropically compressed (and then expanded) after it combusts.^[1] Thus:

$$\frac{T_b(x'_b, x_b)}{T_b(x'_b)} = \left[\frac{p(x_b)}{p(x'_b)} \right]^{(\gamma_b - 1)/\gamma_b}$$

(9.11)

where

$T_b(x'_b, x_b)$ is the temperature of the element which burned at the pressure

$p(x'_b)$, when the pressure is

$p(x_b)$, and

$$T_b(x'_b) = \frac{h_{f,u} - h_{f,b} + c_{p,u}T_u(x'_b)}{c_{p,b}}$$

(9.12)

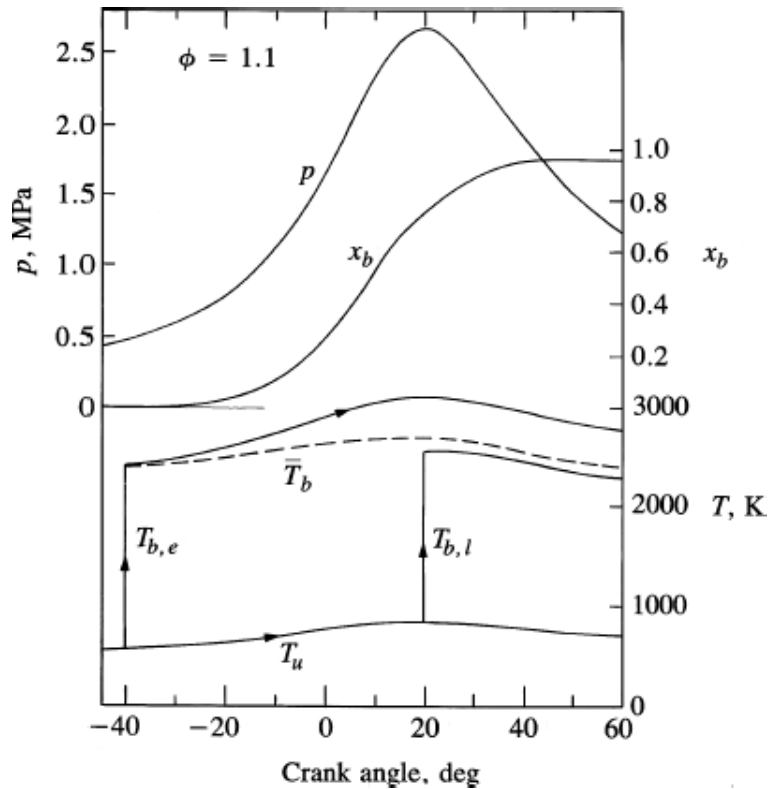
is the temperature resulting from isenthalpic combustion of the unburned gas at

$T_u(x'_b)$,

$p(x'_b)$. An example of the temperature distribution computed with this model is shown in Fig. 9.8. A mixture element that burns right at the start of the combustion process reaches, in the absence of mixing, a peak temperature after combustion about 400 K higher than an element that burns toward the end of the combustion process. The mean burned gas temperature is closer to the lower of these temperatures. These two models approximate, respectively, to situations where the time scale

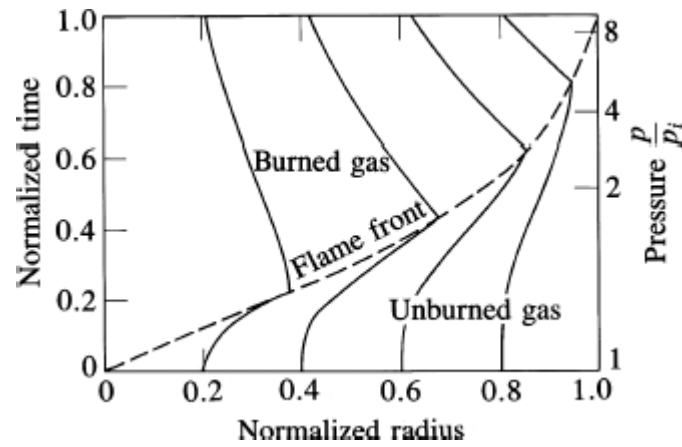
that characterizes the turbulent mixing process in the burned gases is (1) much less than the overall burning time (for the fully mixed model) or (2) much longer than the overall burning time (for the unmixed model). The real situation, based on burned gas temperature measurements,⁸ lies in between.

Figure 9.8 Cylinder pressure, mass fraction burned, and gas temperatures as functions of crank angle during combustion. T_u is unburned gas temperature, T_b is burned gas temperature, the subscripts e and l denote early and late burning gas elements, and $\bar{T}_b$ is the mean burned gas temperature.⁶



As has already been mentioned, the expansion of a gas element which occurs during combustion compresses the gas ahead of the flame and moves it away from the spark plug. At the same time, previously burned gas is compressed and moved back toward the spark plug. Defining this motion in an engine requires sophisticated flow models because the combustion chamber shape is rarely symmetrical and the spark plug is often not centrally located. However, the gas motion in a spherical or cylindrical combustion bomb with central ignition, which can readily be computed, illustrates the features of the combustion-induced motion in an engine. Figure 9.9 shows calculated gas element trajectories for a stoichiometric methane-air mixture, initially at ambient conditions, as a laminar flame with a constant burning velocity propagates outward from the center of a spherical container. Applying this gas motion model to an engine, it can be concluded that an instrument at a given location, initially views earlier burned gas (of higher temperature) and that as more of the charge burns, the instrument will view later burned gas of progressively lower temperature.

Figure 9.9 Particle trajectories in unburned and burned gas as flame propagate outward at constant velocity from the center of a spherical combustion bomb. Stoichiometric methane-air mixture initially at 1 bar and 300 K.



More accurate calculations relating the mass fraction burned, gas pressure, and gas temperature distribution are often required. The accuracy of such calculations depends on the accuracy with which the time-varying heat loss to the chamber walls can be estimated (see [Sec. 12.4.3](#)), also whether flows into and out of crevice regions are significant (see [Sec. 8.6](#)), as well as the accuracy of the models used to describe the thermodynamic properties of the gases. More accurate models for the thermodynamic properties are: an equilibrium model for the burned gas, and specific heat models, which vary with temperature for each of the components of the unburned mixture (see [Sec. 4.1](#) and [4.7](#)). In the absence of significant crevice effects, Eqs. (9.1) and (9.2) can be written as

$$\frac{V}{m} = v_b x_b + v_u (1 - x_b) \quad (9.13)$$

$$\frac{U_0 - W - Q}{m} = u_b x_b + u_u (1 - x_b) \quad (9.14)$$

where

$$v_b = \frac{1}{x_b} \int_0^{x_b} v_b dx \quad \text{and} \quad v_u = \frac{1}{1 - x_b} \int_{x_b}^1 v_u dx$$

and similar definitions hold for

u_b and

u_u . For a given equivalence ratio, fuel and burned gas fraction:

$$h_u = h_u(T_u) \quad h_b = h_b(T_b, p)$$

(9.15a,b)

$$pv_u = \left(\frac{\tilde{R}}{M_u}\right) T_u \quad pv_b = \left(\frac{\tilde{R}}{M_b}\right) T_b$$

(9.16a,b)

and

$$u_u = \bar{h}_u - pv_u \quad u_b = \bar{h}_b - pv_b$$

(9.17a,b)

To simplify the calculations it is convenient to assume that, for the burned gas,

$$u_b = u_b(\bar{T}_b, p) \text{ and}$$

$v_b = v_b(\bar{T}_b, p)$. This corresponds to the fully mixed assumption described above. The effect of neglecting the temperature distribution in the calculation of mass fraction burned is small. In addition, the heat losses from the unburned gas can usually be neglected. The unburned gas is then compressed isentropically.

$\bar{T}_u$ is specified for some initial state of the unburned gas (where $x_b = 0$) by p_0 , V_0 , M_u , and the mass of charge m . Then, since for any isentropic process,

$$\left(\frac{\partial T}{\partial p}\right)_s = \frac{v - (\partial h / \partial p)_T}{(\partial h / \partial T)_p}$$

(9.18)

and

$\bar{T}_u$ can be determined.

Equations (9.13) to (9.18) constitute a set of nine equations for the nine unknowns

v_w

v_b ,

u_w

u_b ,

$\bar{h}_w$

$\bar{h}_b$,

$\bar{T}_w$

$\bar{T}_b$, and x_b or p . One convenient solution method is to eliminate x_b from Eqs. (9.13) and (9.14) to obtain

$$\frac{(V/m) - v_u}{v_b - v_u} - \frac{(U/m) - u_u}{u_b - u_u} = f(\bar{T}_b, \bar{T}_u) = 0$$

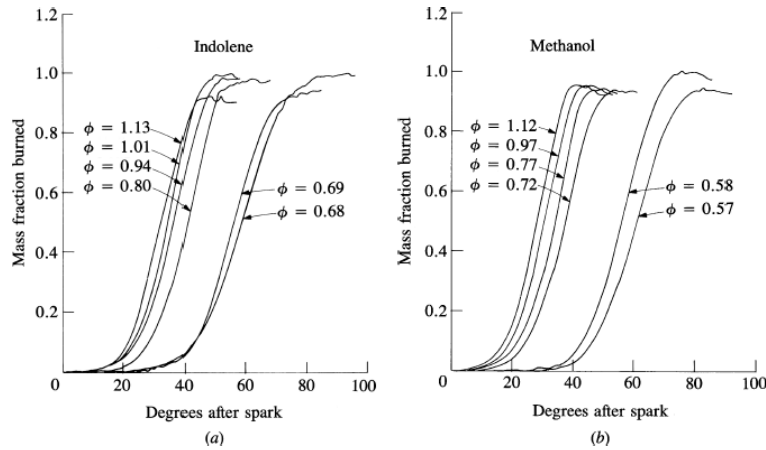
(9.19)

where

$U = U_0 - W - Q$. $\bar{T}_u$ can be determined from Eq. (9.18). Equation (9.19) can then be solved using an appropriate iterative technique for

$\bar{T}_b$, and x_b can be obtained from Eq. (9.13). Some examples of mass fraction burned curves obtained from measured pressure data, with gasoline and methanol fuels, are shown in Fig. 9.10. With accurate pressure versus crank-angle records, values of final mass fraction burned should be close to but lower than unity, usually by about 5% due to incomplete combustion of the fuel within the cylinder: see Fig. 3.9 and Fig. 11.35.

Figure 9.10 Mass fraction burned curves determined from measured cylinder pressure data using two-zone combustion model: (a) gasoline; (b) methanol. ϕ = fuel/air equivalence ratio.⁹



More accurate burned gas temperature calculations need to account for the presence of a thermal boundary layer (of order 1-mm thick) around the combustion chamber walls (see Sec. 12.6.5). The burned gas region in Fig. 9.6 can be divided into an adiabatic core and a boundary layer that grows in thickness with time. In the adiabatic core, in the absence of mixing between gas elements that burn at different times, burned gas is compressed and then expanded isentropically. The burned gas temperature distribution can be calculated as follows. Given the pressure versus crank-angle data, the unburned mixture state can be determined using Eq. (9.18) above. Each small element of unburned mixture burns in an essentially constant-enthalpy constant-pressure process. So the burned state of an element of unburned charge, which burns at $p = p_i$, can be obtained from the relation

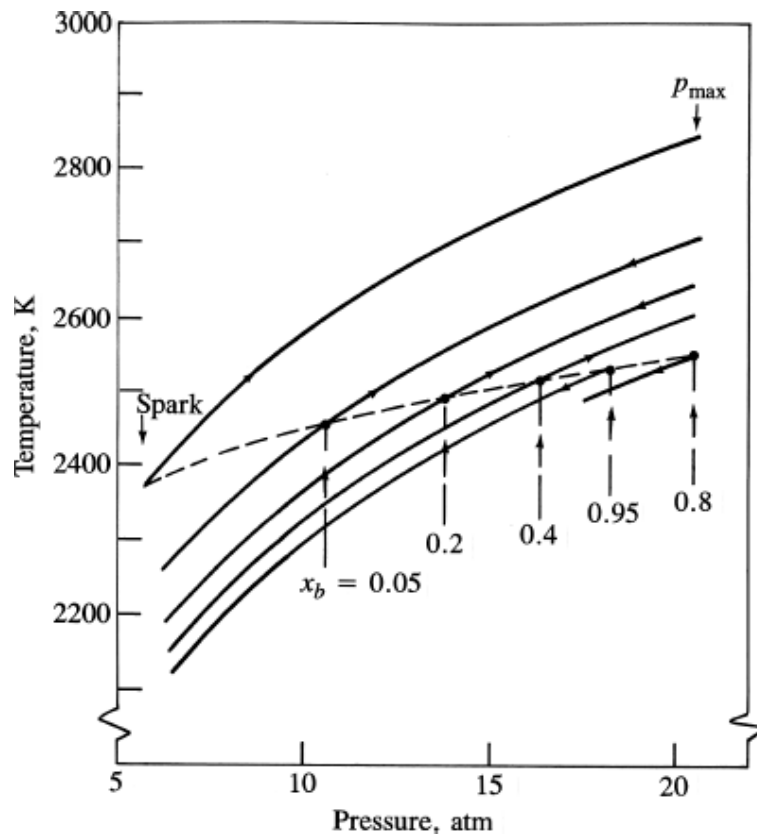
$$h_b(T_{b,i}, p_i) = h_u(T_{u,i}, p_i)$$

After combustion, this element which burned at $p = p_i$ is compressed and expanded along the isentropic:

$$s_b(T_b, p) = s_b(T_{b,i}, p_i)$$

An example of the temperature distribution computed in this manner for this unmixed model in the burned gas adiabatic core is shown in Fig. 9.11. The element ignited by the spark is compressed to the highest peak temperature at $p_{\max}$. The temperature difference across the bulk of the charge ($0.05 < x_b < 0.95$) is about 200 K.

Figure 9.11 Calculated temperature distribution in the adiabatic core of the burned gas zone for the unmixed model assuming thermodynamic equilibrium. $\phi = 1.0$. Dashed line is temperature of each element just after it burns.



9.2.2. Analysis of Cylinder Pressure Data

Cylinder pressure changes with crank angle as a result of cylinder volume change, combustion, heat transfer to the chamber walls, flow into and out of crevice regions, and leakage. The first two of these effects are the largest. The effect of volume change on the pressure can readily be accounted for. Thus, combustion rate information can be obtained from accurate pressure data provided models for the remaining phenomena can be developed at an appropriate level of approximation. The previous section has developed the fundamental basis for such calculations.

Cylinder pressure is usually measured with piezoelectric pressure transducers. This type of transducer contains a quartz crystal. One end of the crystal is exposed through a diaphragm to the cylinder pressure; as the cylinder pressure increases, the crystal is compressed and generates an electric charge which is proportional to the pressure. A charge amplifier is then used to produce an output voltage proportional to this charge. Accurate cylinder pressure versus crank-angle data can be obtained with these systems provided the following steps are carried out: (1) the correct reference pressure is used to convert the measured pressure signals to absolute pressures; (2) the pressure versus crank-angle (or volume) phasing is accurate to within about 0.2° ; (3) the clearance volume is estimated with sufficient accuracy; (4) transducer temperature variations (which can change the transducer calibration factor) due to the variation in wall heat flux during the engine cycle are held to a minimum. Log p versus log V plots can be used to check the quality of cylinder pressure data. The first three of the above requirements can be validated using log p -log V diagrams for a motored engine. If the effects of thermal cycling of the pressure transducer are significant, the expansion stroke on the log p -log V plot for a firing engine shows excessive curvature.^{10, 11}

Figure 9.12 shows pressure-volume data from a firing spark-ignition engine on both a linear p - V and a log p -log V diagram. On the log p -log V diagram the compression process is a straight line of slope about 1.3. The start of combustion can be

identified by the departure of the curve from the straight line. The end of combustion can be located in similar fashion: the expansion stroke following combustion is essentially linear with slope 1.33. Since both the compression of the unburned mixture prior to combustion and the expansion of the burned gases following the end of combustion are close to adiabatic isentropic processes (for which

$$pV^\gamma = \text{constant};$$

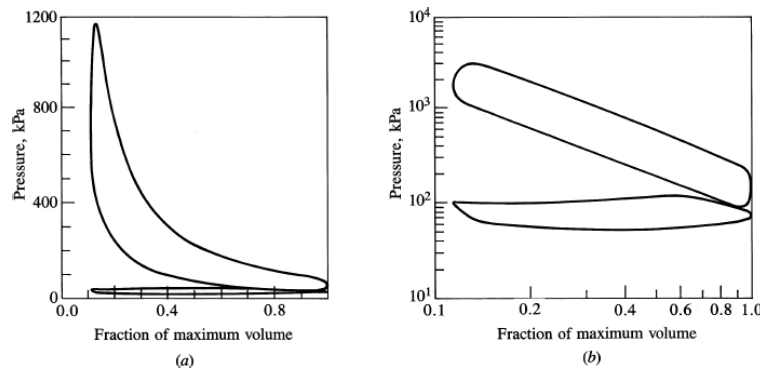
$\gamma = c_p/c_v$), the observed behavior is as expected. Extensive studies show that the compression and expansion processes are well fitted by a polytropic relation:

$$pV^n = \text{constant}$$

(9.20)

The exponent n for the compression and expansion processes is $1.3 (\pm 0.05)$ for most fuels. It is comparable to the average value of γ_u for the unburned mixture over the compression process, but is larger than γ_b for the burned gas mixture during expansion due to heat loss to the combustion chamber walls (see Figs. 4.13 and 4.16).

Figure 9.12 (a) Pressure-volume diagram; (b) log p -log $(V/V_{\max})$ plot. 1500 rev/min, MBT spark timing, imep = 513 kPa, $\phi = 0.8$, $r_c = 8.72$, propane fuel.¹⁰



Log p -log V plots such as Fig. 9.12b approximately define the start and end of combustion, but do not provide a mass fraction burned profile. One well-established technique for estimating the mass fraction burned profile from the pressure and volume data is that developed by Rassweiler and Withrow.² They correlated cylinder pressure data with flame photographs, and showed how Eq. (9.20) could be used to account for the effect of cylinder volume change on the pressure during combustion. Assuming that the unburned gas filling the volume V_u ahead of the flame at any crank angle during combustion has been compressed polytropically by the advancing flame front, then the volume $V_{u,0}$ that mass occupied at time of spark is

$$V_{u,0} = V_u \left(\frac{p}{p_0} \right)^{1/n}$$

(9.21)

Similarly, the burned gas behind the flame filling the volume V_b would, at the end of combustion, fill a volume $V_{b,f}$ given by

$$V_{b,f} = V_b \left(\frac{p}{p_f} \right)^{1/n}$$

(9.22)

The mass fraction burned x_b is equal to

$1 - (V_{u,0}/V_0)$ and to

$V_{b,f}/V_f$, where V_0 and V_f are the total cylinder volume at time of spark (just before combustion) and at the end of combustion, respectively. Since

$V = V_u + V_b$, Eqs. (9.21) and (9.22) then give:

$$x_b = \frac{p^{1/n}V - p_0^{1/n}V_0}{p_f^{1/n}V_f - p_0^{1/n}V_0} \quad (9.23)$$

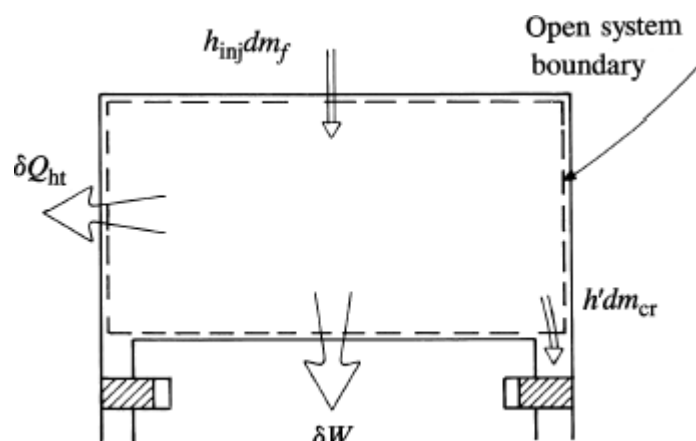
This method is widely used, though it contains several approximations. Heat-transfer effects are included only to the extent that the polytropic exponent n in Eq. (9.22) differs appropriately from γ . The pressure rise due to combustion is proportional to the amount of fuel chemical energy released rather than the mass of mixture burned. Also, the polytropic exponent n is not the same before, and after combustion. Selecting an appropriate value for n (whether n is assumed to be constant or to vary through the combustion process) is one difficulty in applying this pressure data analysis procedure.

The effects of heat transfer, crevices, and leakage can be explicitly incorporated into cylinder pressure data analysis by using a "heat release" approach based on the first law of thermodynamics. A significant advantage of such an approach is that the pressure changes can be related directly to the amount of fuel chemical energy released by combustion, while retaining the simplicity of treating the combustion chamber contents as a single zone. Figure 9.13 shows the appropriate open-system boundary for the combustion chamber.¹² The first law for this open system is

$$\delta Q_{ch} = dU_s + \delta Q_{ht} + \delta W + \sum h_i dm_i \quad (9.24)$$

The change in sensible energy of the charge dU_s is separated from that due to combustion-produced change in composition: the term δQ_{ch} represents the "chemical energy" released by combustion. The work δW is piston work and equal to $p dV$. δQ_{ht} is heat transfer to the chamber walls. The mass flux term represents flow across the system boundary. It includes direct fuel injection and the flow into and out of crevice regions (see Sec. 8.6).

Figure 9.13 Open system boundary for engine combustion chamber for heat-release analysis.



The accuracy with which this energy balance can be made depends on how adequately each term in the above equation can

be quantified. Assuming that U_s is given by $mu(T)$, where T is the mean charge temperature and m is the mass within the system boundary, then

$$dU_s = mc_v(T)dT + u(T)dm$$

Note that this mean temperature determined from the ideal gas law is close to the mass-averaged cylinder temperature during combustion since the molecular weights of the burned and unburned gases are essentially the same. Crevice effects can usually be modeled adequately by flow into and out of a single crevice volume at the cylinder pressure, with the gas in the crevice at a substantially lower temperature. Leakage to the crankcase (blowby) can usually be neglected. Then Eq. (9.24), on substituting for dU_s and $dm_i (= dm_{cr} = -dm)$, becomes

$$\delta Q_{ch} = mc_v dT + (h' - u)dm_{cr} + pdV + \delta Q_{ht}$$

(9.25)

where

$dm_{cr} > 0$ when flow is out of the cylinder into the crevice

$dm_{cr} < 0$ when flow is from the crevice to the cylinder, and

h' is evaluated at cylinder conditions when $dm_{cr} > 0$ and at crevice conditions when $dm_{cr} < 0$.

Use of the ideal gas law (neglecting the change in gas constant R) with Eq. (9.25) then gives

$$\delta Q_{ch} = \left(\frac{c_v}{R}\right) Vdp + \left(\frac{c_v}{R}\right) pdV + (h' - u + c_v T)dm_{cr} + \delta Q_{ht}$$

(9.26)

This equation can be used in several ways. When the heat or energy release term, δQ_{ch} , is combined with the heat-transfer and crevice terms, the combination is termed *net heat release*—the combustion energy release less heat lost to the walls. It is equal to the first two terms on the right-hand side of Eq. (9.26) which, together, represent the sensible energy change and work transfer to the piston. While heat losses during combustion are a small fraction of the fuel energy (10 to 15%), the distributions of heat release and heat transfer with crank angle are different; heat transfer becomes more important as the combustion process ends and average gas temperatures peak. The net heat-release profile obtained from integrating the first two terms on the right-hand side of Eq. (9.26), normalized to give unity at its maximum value, is often interpreted as the burned mass fraction (or, more correctly, the energy-release fraction) versus crank-angle profile.

Use of Eq. (9.26) requires a value for $c_v/R [= 1/(\gamma - 1)]$. The ratio of specific heats γ for both unburned and burned gases decreases with increasing temperature and varies with composition (see Figs. 4.13, 4.16, and 4.18). As the mean charge temperature increases during compression and combustion and then decreases during expansion, γ should vary. An approximate approach, modeling $\gamma(T)$ with a linear function of temperature fitted to the appropriate curves in Figs. 4.13, 4.16, and 4.18 and with γ constant during combustion, has been shown to give adequate results.^{13,14}

The convective heat-transfer rate to the combustion chamber walls can be calculated from the relation

$$\frac{dQ_{ht}}{dt} = Ah_c(T - T_w)$$

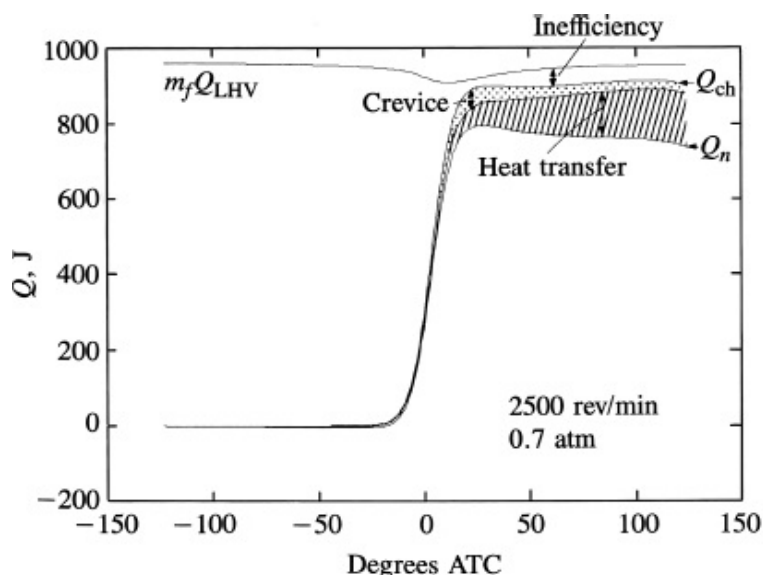
where A is the chamber surface area, T is the mean gas temperature, T_w is the mean wall temperature, and h_c is the heat-transfer coefficient (averaged over the chamber surface area). h_c can be estimated from engine heat-transfer correlations (see Sec. 12.4.3). Since crevice effects are usually small, a sufficiently accurate model for their overall effect is to consider a single aggregate crevice volume where the gas is at the same pressure as the combustion chamber, but at a different temperature, T_w . Since these crevice regions are narrow, an appropriate assumption is that the crevice gas is at the wall temperature. Inserting this crevice model into Eq. (9.26), with $\gamma(T) = a + bT$, gives the *chemical energy- or gross heat-release rate*:

$$\frac{dQ_{ch}}{d\theta} = \frac{\gamma}{\gamma-1} p \frac{dV}{d\theta} + \frac{1}{\gamma-1} V \frac{dp}{d\theta} + V_{cr} \left[\frac{T'}{T_w} + \frac{T}{T_w(\gamma-1)} + \frac{1}{bT_w} \ln \left(\frac{\gamma-1}{\gamma'-1} \right) \right] \frac{dp}{d\theta} + \frac{dQ_{ht}}{d\theta}$$

(9.27)

An example of the use of Eq. (9.27) to analyze an experimental pressure versus crank-angle curve for a spark-ignition engine is shown in Fig. 9.14. The integrated heat release is plotted against crank angle. The lowest curve shown is the net heat release. The addition of heat transfer and crevice models gives the chemical energy release. The curve at the top of the figure is the mass of fuel within the combustion chamber times its lower heating value. It decreases slightly as p_{max} is approached due to flow into the crevices. The difference between the final value of Q_{ch} and $(m_f Q_{LHV})$ should equal the combustion inefficiency (which is a few percent of $m_f Q_{LHV}$). The combustion inefficiency can be determined from the exhaust gas composition (see Sec. 4.9.4). Inaccuracies in the cylinder pressure data and the heat-release calculation will also contribute to this difference. An important advantage of a heat-release analysis that relates the pressure changes to the amount of fuel chemical energy within the cylinder is that this error can be determined. In the example in Fig. 9.14, the measured combustion inefficiency was close to the amount shown.

Figure 9.14 Results of heat-release analysis showing the effects of heat transfer, crevices, and combustion inefficiency.¹²



Two-zone models (one zone representing the unburned mixture ahead of the flame and one the burned mixture behind the flame) are also used to calculate the mass fraction burned profile from measured cylinder pressure data. Figure 9.10 shows results from such an analysis, using the methodology described in Sec. 9.2.1.⁹ The advantage of a two-zone analysis is that the thermodynamic properties of the cylinder contents can be quantified more accurately. The disadvantages are that the unburned and the burned zone heat-transfer areas must both now be estimated, and a model for the composition of the gas

flowing into the crevice region must be developed. Due to this complexity, crevice models are usually omitted despite the fact that their impact can be significant.

9.2.3. Combustion Process Characterization

The mass fraction burned profiles as a function of crank angle in each individual cycle shown in [Fig. 9.4](#) and the chemical energy- or gross heat-release curve in [Fig. 9.14](#) have a characteristic S-shape. The rate at which fuel-air mixture burns increases from a low value immediately following the spark discharge, to a maximum about halfway through the burning process, and then decreases to close to zero as the combustion process ends. It proves convenient to use these mass fraction burned or energy-release fraction curves to characterize different stages of the spark-ignition engine combustion process by their duration in crank-angle degrees, thereby defining the fraction of the engine cycle that they occupy. The flame development process, from the spark discharge which initiates the combustion process to the point where a small but measurable fraction of the charge has burned, is one such stage. It is influenced primarily by the mixture state, composition, and gas motion in the vicinity of the spark plug (see [Sec. 9.3](#)). The stage during which the major portion of the charge burns as the flame propagates to the chamber walls comes next. This stage is obviously influenced by conditions throughout the combustion chamber. The final stage, where the remainder of the charge burns to completion, cannot as easily be quantified because energy-release rates are comparable to other energy-transfer processes that are occurring.

The following definitions are commonly used to characterize these energy-release aspects of combustion:

Flame-development angle $\Delta\theta_d$. The crank-angle interval between the spark discharge and the time when a small but significant fraction of the cylinder mass has burned or fuel chemical energy has been released. Usually this fraction is 10%, though other fractions such as 1% and 5% have been used.^[2]

Rapid-burning angle $\Delta\theta_b$. The crank-angle interval required to burn the bulk of the charge. It is defined as the interval between the end of the flame-development stage (usually mass fraction burned or energy-release fraction of 10%) and the end of the flame-propagation process (usually mass fraction burned or energy-release fraction of 90%).^[3]

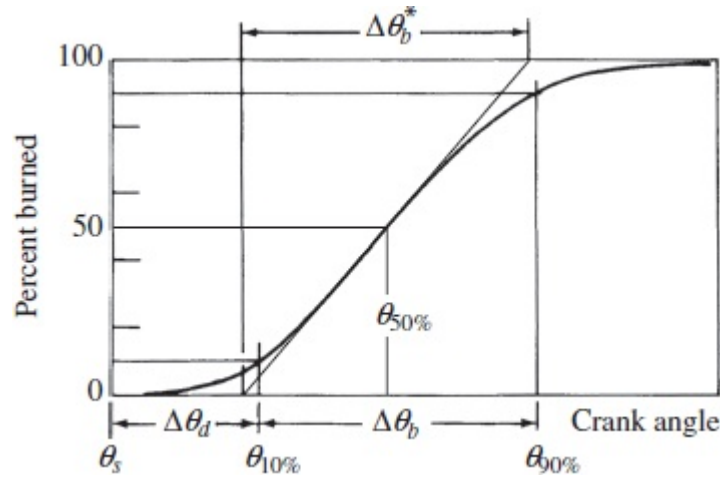
Overall burning angle $\Delta\theta_o$. The duration of the overall burning process. It is the sum of $\Delta\theta_d$ and $\Delta\theta_b$.

[Figure 9.15](#) illustrates these definitions on a mass fraction burned, or fraction of fuel energy released, versus crank-angle plot. While the selection of the 10% and 90% points is arbitrary, such a choice avoids the difficulties involved in determining accurately the shape of the curve at the start and end of combustion. These angles can be converted to times (in seconds) by dividing by $6N$ (with N in revolutions per minute). (Examples of

$\Delta\theta_{0-10\%}$ and

$\Delta\theta_{10-90\%}$ can be found in [Refs. 15 to 18.](#))

Figure 9.15 Definition of flame-development angle, $\Delta\theta_d$ and rapid-burning angle, $\Delta\theta_b$ (or $\Delta\theta_b^*$) on mass fraction burned versus crank-angle curve.



The flame development angle

$\Delta\theta_d$ can be estimated from the following scaling relationship:¹⁶

$$\frac{\Delta\theta_d}{\Delta\theta_{d,\text{ref}}} = \left(\frac{N}{N_{\text{ref}}} \right)^{1/3} \left(\frac{S_{L,\text{ref}}}{S_L} \right)^{2/3}$$

(9.28)

where N is engine speed, and S_L is the laminar flame speed (see [Sec. 9.3.3](#)) and includes the effects of relative air/fuel ratio, burned gas fraction, and mixture temperature and pressure.

When mass fraction burned profiles are available (and as indicated above, they can readily be obtained from cylinder pressure versus crank-angle data), then the crank angle at which 50% of the in-cylinder mass has burned proves to be a useful combustion phasing parameter. [Figure 9.15](#) identifies this 50% burned angle, $\theta_{50\%}$: it is used to characterize the *phasing* of the combustion process and determine a *combustion retard* parameter, $(\theta_{50\%} - \theta_{50\% \text{MBT}})$. It represents the extent (in crank-angle degrees) that the center of the combustion process has been shifted from its location at maximum torque: for a wide range of engines and their operating conditions, $\theta_{50\%}$ at MBT timing occurs at 5 to 7° ATC.

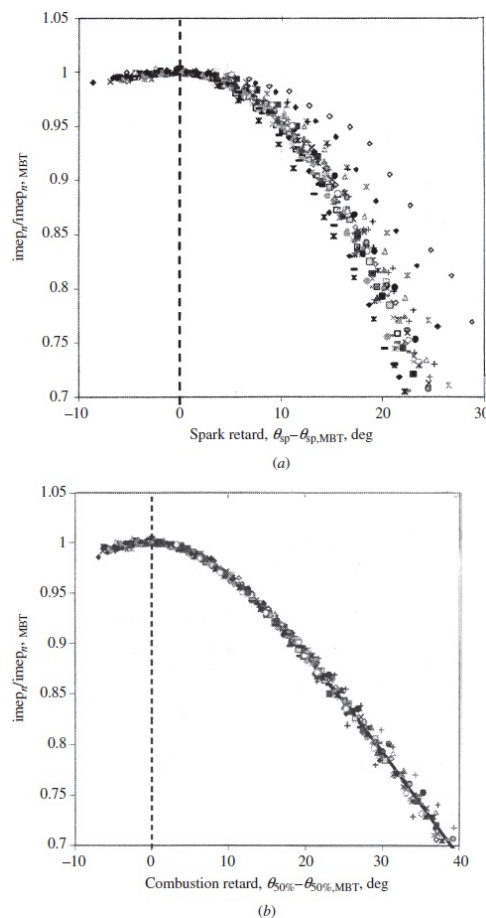
The ability of this combustion retard parameter to correlate the effects of retarding the combustion process (through spark retard from MBT timing) is illustrated in [Fig. 9.16](#). An extensive set of mass fraction burned profiles for many different fuels, relative air/fuel ratios, compression ratios, intake pressures (different engine loads), and spark timings for a naturally-aspirated SI engine are shown.¹⁵ [Figure 9.16a](#) shows net imep (normalized by the net imep at MBT timing) as a function of spark retard from MBT spark: [Figure 9.16b](#) shows the same normalized net imep data as a function of combustion retard, $(\theta_{50\%} - \theta_{50\% \text{MBT}})$. The correlation of net imep with spark retard is good, but the few data sets on the upper right of the band of data, where the relative air/fuel ratio is very lean ($\lambda = 1.6$) and combustion is much slower, do not fit as well. In contrast, the same data with net imep plotted against combustion retard in [Fig. 9.16b](#) show an extremely good correlation. The relationship between relative net imep and the extent to which the center of the combustion process is retarded from its optimum crank-angle location appears to be a universal one for naturally-aspirated SI engines. The correlation is well fitted by the equation

$$\frac{\text{imep}_n}{\text{imep}_{n,\text{MBT}}} = 1 - 0.168 \left\{ \left[1 + 4.44 \times 10^{-3} (\theta_{50\%} - \theta_{50\% \text{MBT}})^2 \right]^{0.5} - 1 \right\}$$

(9.29)

Here, $\theta_{50\%}$ is the location of the 50% mass fraction burned point in degrees ATC and $\theta_{50\%MBT}$ is the same parameter at MBT spark timing. Note that 5 crank-angle (CA) degrees of retard results in a 1% loss of imep_n , and a 10-degrees CA retard results in a 3% imep_n loss. Both of these losses are small. Equivalent spark retards show slightly larger net imep reductions. Net imep is the appropriate "load" parameter to correlate, since combustion phasing affects the pressure levels in the cylinder and, thus, the indicated work.

Figure 9.16 Net imep relative to net imep_{MBT} as a function of combustion retard from the optimum (MBT) phasing: (a) as a function of spark timing relative to MBT spark timing; (b) as a function of $\theta_{50\%}$ relative to $\theta_{50\%MBT}$.¹⁵



In turbocharged SI engines at higher loads where the compressor is supplying a significantly boosted airflow, the situation is more complex. As spark (and combustion) is retarded, the exhaust gas temperature and enthalpy increase. The turbocharger turbine is thus driven more strongly by the exhaust flow, and the turbocharger speed and boost pressure increase. This increases the net imep and torque thus partly offsetting the spark-retard-induced loss of net imep. For a turbocharged SI engine operating with significant boost, the corresponding curves to those of a naturally-aspirated SI engine in Fig. 9.16 are flatter. For example, at 1.4 bar (absolute) boost and 10 degrees spark retard, the drop in net imep is only 1%. At lighter loads, where the turbocharger has little effect on the engine's operation, spark retard effects are similar to naturally-aspirated engine behavior.

A functional form often used to represent the mass fraction burned versus crank-angle curve $x_b(\theta)$ is the Wiebe function:

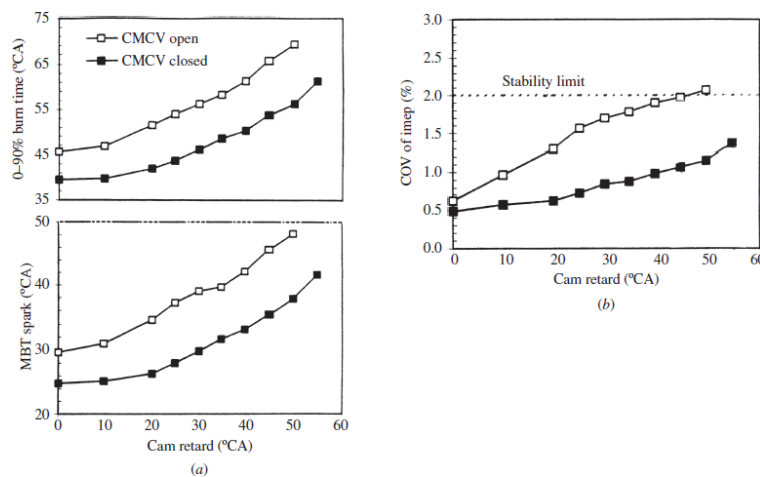
$$x_b = 1 - \exp \left[-a \left(\frac{\theta - \theta_0}{\Delta\theta} \right)^{m+1} \right]$$

(9.30)

where θ is the crank angle, θ_0 is the start of combustion, $\Delta\theta$ is the total combustion duration ($x_b = 0$ to $x_b = 1$), and a and m are adjustable parameters. Varying a and m changes the shape of the curve significantly. Actual mass fraction burned curves have been fitted with $a = 5$ and $m = 2$.

An example of how these burn angles can be used to characterize SI engine combustion is shown in Fig. 9.17. The figure shows the 0 to 90% burn time ($\theta_{0-10\%} + \theta_{10-90\%}$) for a 5.4-liter three-valve per cylinder V-8 SI engine with several features designed to improve the combustion characteristics. These features were: (1) a charge motion control valve (CMCV), a specially shaped "throttle" plate in each intake port which, when used at low engine speeds and light loads, provides a high level of in-cylinder charge motion to enhance the burn rate; (2) variable cam timing (VCT) which retards both the intake and exhaust valve events equally at part load to reduce the pumping losses, increase expansion stroke work, and delay the valve overlap (see Sec. 13.6); and (3) a centrally located spark plug to increase knock resistance so a high compression ratio (9.8) could be used.¹⁸

Figure 9.17 (a) Burn angle, spark to 90% burned, $\Delta\theta_{0-90}$, and MBT spark timing, and (b) coefficient of variation (%) of net imep, as a function of cam retard for a 5.4-liter three-valve per V-8 SI engine with several features designed to improve the combustion characteristics. Engine operating at 1500 rev/min and 2.62 bar bmep with a stoichiometric mixture.¹⁸ Increasing cam retard increases the residual gas fraction in the unburned mixture. CMCV, charge motion control valve.



The data (Fig. 9.17a) show that, as the "dual equal" VCT increasingly retards both valve opening and closing angles, the spark timing must be advanced significantly as the combustion duration ($\Delta\theta_{0-90\%}$) steadily lengthens due to the increasing levels of mixture dilution with residual and exhaust gases, and due to the reduced compression temperature at spark due to late intake valve closing. The CMCV closing increases the burn rate significantly. Typically, $\Delta\theta_{0-10\%}$ and $\Delta\theta_{10-90\%}$ are comparable in magnitude so are some half of the $\Delta\theta_{0-90\%}$. Figure 9.17b shows the impact of cam retard on cycle-by-cycle combustion variability. The coefficient of variation (COV)^[4] of the cycle-by-cycle net imep (determined from analyzing individual cycle cylinder pressure data) needs to be less than about 2% for acceptably smooth engine operation at lighter loads (see Sec. 9.4.1). COV of net imep, combustion duration, and spark advance are related. Note that a spark advance of more than about 40° BTC at lightly loaded engine conditions usually results in substantial combustion variability.¹⁵

-
- [1] This model applies to burned gas regions of the chamber away from the walls. Heat transfer to the walls results in a thermal boundary layer on the walls, which grows with time as shown in [Fig. 9.6](#). The gas in the boundary layer is not isentropically compressed and expanded.
 - [2] This angle is sometimes called the *ignition delay*. Since the flame starts to propagate outward immediately following the spark discharge there is no delay, and the terminology used here is preferred (see [Sec. 9.3](#)).
 - [3] An alternative definition for $\Delta\theta_b$ uses the maximum burning rate to define an angle or time characteristic of the bulk charge burning process⁴ (see [Fig. 9.15](#)). $\Delta\theta_b$ and $\Delta\theta_b^*$ are usually closely comparable.
 - [4] COV of a data set is the standard deviation divided by the mean value.

9.3. FLAME STRUCTURE AND SPEED

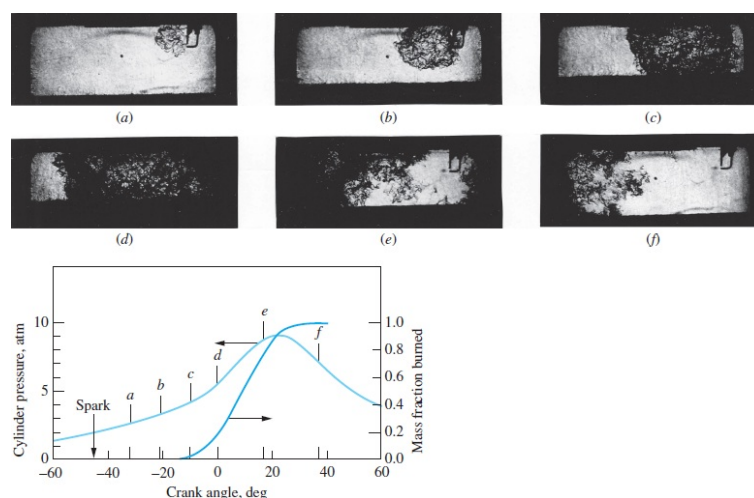
9.3.1. Overall Observations

The combustion process in the spark-ignition engine takes place in a turbulent flow field. This flow field is produced by the high shear flows set up during the intake process and modified during compression, as described in [Chap. 8](#). The importance of the turbulence to the engine combustion process was recognized long ago through experiments where the intake event, and the turbulence it generates, was eliminated by closing the valves and delaying ignition to subsequent cycles: the rate of flame propagation decreased substantially. Understanding the structure of this engine flame as it develops from the spark discharge and the speed at which it propagates across the combustion chamber, and how that structure and speed depend on charge motion, turbulence levels, charge composition, and chamber geometry, are critical to engine optimization. This section reviews experimental evidence that describes the overall features of the flame development and propagation processes.

Direct flame photographs such as those in [Fig. 9.2](#) indicate the location and shape of the actual flame reaction zone which radiates in the blue region of the visible spectrum. An irregular front is apparent. Further insight into the structure of the flame can be obtained from photographs taken with techniques that are sensitive to density changes in the flow field, such as schlieren and shadowgraph. With these techniques, a parallel light beam is passed through the combustion chamber. Portions of the beam, which pass through regions where density gradients normal to the beam exist, are deflected due to the refractive index gradients that result from the density gradients. In the schlieren technique, the beam is focused on a knife edge; the deflected parts of the beam are displaced relative to the knife edge and produce lighter or darker regions when subsequently refocused to produce an image. With the shadowgraph technique, the parallel beam emerging from the combustion chamber is observed directly; deflected parts of the beam produce lighter and darker regions in the image. With these techniques, details of flame structure can be discerned.

[Figure 9.18](#) shows a set of photographs from one engine cycle, from a high-speed schlieren movie taken in a special visualization spark-ignition engine. Also shown are the cylinder pressure versus crank-angle data, and the mass fraction burned profile calculated from the pressure data using the method of Rassweiler and Withrow² (see [Sec. 9.2.2](#)). This engine had a square-cross-section cylinder with two quartz walls to permit easy optical access, but otherwise operated normally.¹⁹ Visualization of the flame is especially important during the early stages of flame development when the pressure rise due to combustion is too small to be detected.

Figure 9.18 Sequence of movie frames from one engine cycle in a square-cross-section single-cylinder engine with two glass walls, and corresponding pressure and mass fraction burned curves. 1400 rev/min, 0.5 atm inlet pressure.



These photographs show how the flame, roughly spherical in shape, grows steadily from the time of spark discharge. The effect of turbulence is already visible in the convoluted flame surface in Fig. 9.18a. The volume enflamed behind the front continues to grow in a roughly spherical manner except where the flame reaches the chamber walls, as seen in Figs. 9.18b and c. The mass fraction burned and the associated pressure rise due to combustion become significant by the time the flame front has traversed two-thirds to three-quarters of the cylinder bore. (Note that the fraction of the cylinder filled with enflamed charge is less than is suggested by the photos because the front of the flame is approximately spherical and the cylinder has a square cross section.) Maximum cylinder pressure occurs close to the time the flame front makes contact with the far wall, as seen in Fig. 9.18e. Finally, the unburned mixture (in the corners of the square cross-section chamber) ahead of, and within the front, burns out and the density gradients associated with the flame reaction zone disappear, clearing the field of view as shown in Figs. 9.18e and f.

A useful relationship between the mass fraction burned ($x_b = m_b/m$), and the volume fraction occupied by the burned gas ($y_b = V_b/V$), can be obtained from the identities

$$m = m_u + m_b \text{ and } V = V_u + V_b$$

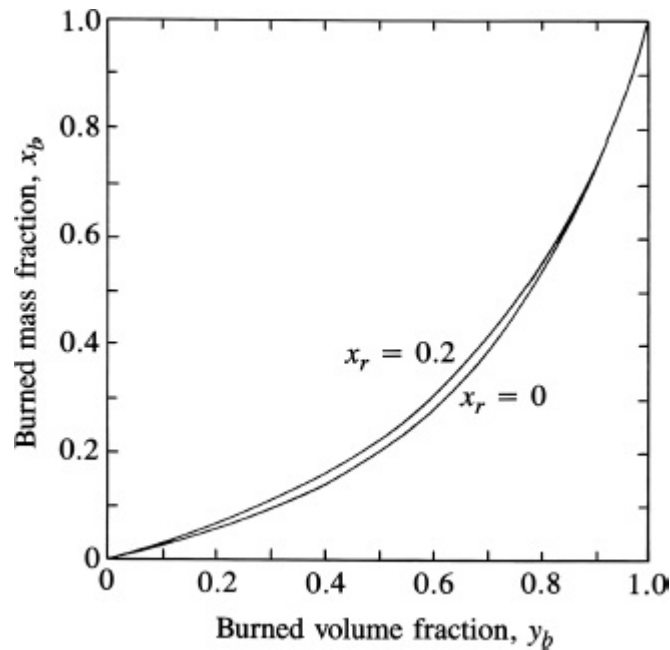
along with the ideal gas law:

$$x_b = \left[1 + \frac{\rho_u}{\rho_b} \left(\frac{1}{y_b} - 1 \right) \right]^{-1}$$

(9.31)

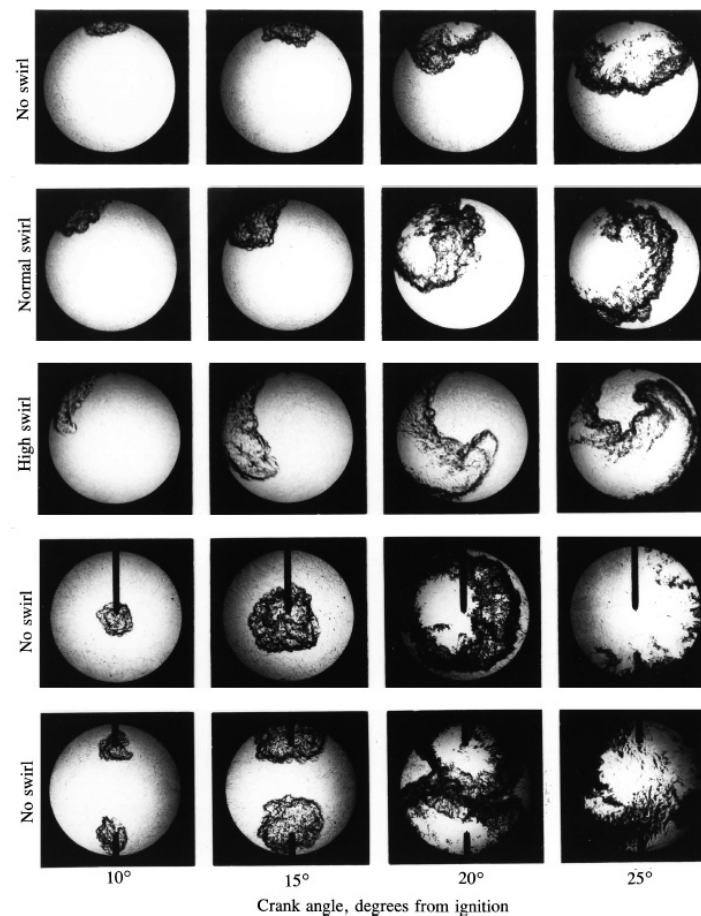
While the density ratio (ρ_u/ρ_b) does depend on the equivalence ratio, burned gas fraction in the unburned mixture, gas temperature, and pressure, its value is close to 4 for most spark-ignition engine operating conditions. Thus, the plot of x_b against y_b has a universal form, as shown in Fig. 9.19. This curve is an important aid in interpreting flame geometry information.

Figure 9.19 Relation between mass fraction burned x_b and volume fraction burned y_b ; x_r is residual mass fraction.⁴



These overall features of the developing and propagating flame are common to almost all engine geometries and operating conditions. Figure 9.20 shows shadowgraph photographs of the flame at fixed crank-angle intervals after ignition, taken through a transparent cylinder head with different geometric and flow configurations.²⁰ The approximately spherical development of the flame from the vicinity of the spark plug, except where it intercepts the chamber walls, is evident for side and center ignition with one plug, and for ignition with two plugs in the absence of any intake generated swirl. With normal levels of swirl, the flame center is convected with the swirling flow, but the flame front as it grows is still approximately spherical in shape. Only with unusually high levels of swirl and aerodynamic stabilization of the flame at the spark plug location does the flame become stretched out and distorted by the flow in a major way.

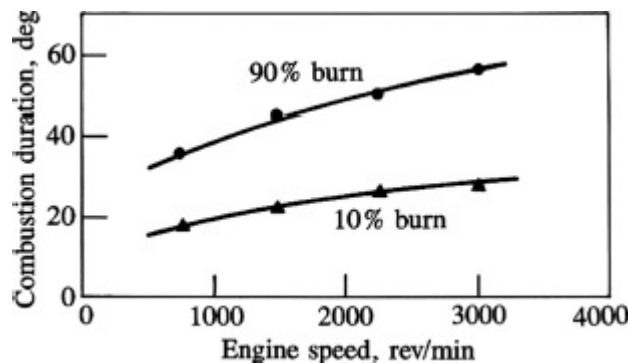
Figure 9.20 Laser shadowgraph photographs of engine combustion process taken in single-cylinder engine with transparent cylinder head. From top to bottom: side plug without swirl; side plug with normal swirl; side plug with high swirl; central plug without swirl; two plugs without swirl.²⁰



At any given flame radius, the geometry of the combustion chamber surfaces and the spark plug location govern the flame front surface area—the area of that portion of the approximately spherical surface corresponding to the leading edge of the flame, contained by the piston, cylinder head, and cylinder wall. The larger this surface area, the greater the mass of fresh charge that can cross this surface and enter the flame zone. The photos in [Fig. 9.20](#) illustrate the importance of flame area. The center plug location gives approximately twice the flame area of the side plug geometry at a given flame radius, and thus burns about twice as fast (the fraction of the cylinder volume enflamed is about twice the size, at a fixed crank-angle interval after spark). The arrangement with two spark plugs on opposite sides of the chamber is not significantly different in enflamed volume and flame front area from the single center plug until the flame fronts are intersected by the cylinder wall.

Mixture burning rate is strongly influenced by engine speed. It is well established that the duration of combustion *in crank-angle degrees* only increases slowly with increasing engine speed. If this were not the case, engines would not be able to operate over the wide range of speeds (600 to 6000 rev/min) that they do. [Figure 9.21](#) shows how the interval between the spark discharge and 10% mass fraction burned, the flame development angle $\Delta\theta_d$, and the interval between the spark and 90% mass fraction burned, the overall burning angle $\Delta\theta_d + \Delta\theta_b$ (see [Sec. 9.2.3](#)), vary with engine speed.¹⁶ The two intervals both increase by a factor of 1.6 for a factor of 4 increase in engine speed; that is, the burning rate throughout the combustion process increases almost, though not quite, as rapidly as engine speed. Additionally, at a given engine speed, increasing in-cylinder gas velocities (e.g., with intake generated swirl or tumble) increases the burning rate: the flame size for the swirling flows in [Fig. 9.20](#) is larger than for the quiescent case with the same plug location at the crank-angle intervals after spark as shown. Increasing engine speed and introducing swirl both increase the levels of turbulence in the engine cylinder at the time of combustion (see [Sec. 8.6](#)). Increased turbulence increases the rate of development and propagation of the turbulent premixed engine flame.

Figure 9.21 Effect of engine speed on flame-development angle (0 to 10% burned) and overall burning angle (0 to 90% burned). $\phi = 1.0$, intake pressure = 0.54 atm, spark = 30° BTC.¹⁶

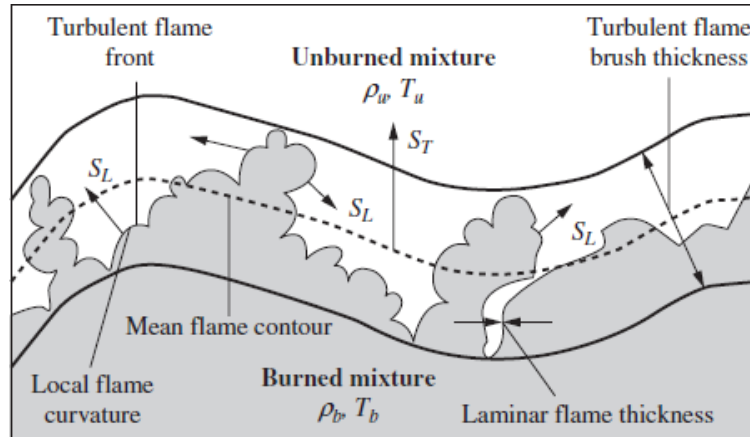


It is also well established that unburned mixture composition and state affect the burning rate. Reducing the inlet pressure (and maintaining the ratio of exhaust to inlet pressure fixed to hold the residual gas fraction constant) increases both the flame development and rapid burning angles. The relative air/fuel ratio or fuel/air equivalence ratio affects the burning rate. Both flame development and burning angles show a minimum for slightly rich mixtures ($\phi \approx 1.1$) and increase significantly as the mixture becomes substantially leaner than stoichiometric.^{16, 17} The burned gas fraction in the unburned mixture, due to the residual gas fraction *and* any recycled exhaust gases, affects the burning rate: increasing the burned gas fraction slows down both flame development and propagation. Fuel composition changes can be significant also. Mixtures of isooctane or conventional gasolines with air and burned gases (at identical conditions) have closely comparable burning rates; propane, methane, methanol, and ethanol mixtures exhibit a modest differences in burning rate; and hydrogen-air mixtures exhibit a substantial difference. The basic combustion chemistry of the fuel, air, burned gas mixture influences the combustion process. However, the relative importance of combustion chemistry effects depends on combustion chamber design and burn rate. Faster burning engines (which have higher turbulence) are less sensitive to changes in mixture composition, pressure, and temperature than are slower burning engines (which have lower turbulence). The effects of chamber geometry, gas motion, and gas composition and state are interrelated.²¹

9.3.2. Flame Structure

Figure 9.22 shows a schematic of the turbulent SI engine flame illustrating its structure and the relevant physical parameters. As indicated, the turbulent flame is a wrinkled thin reaction-sheet "laminar flame." The schematic shows the average leading and trailing boundaries of the flame (solid lines) with the wrinkled flame sheet in between, and the mean flame location (dashed line). The turbulent flame moves forward into the unburned mixture ahead of the flame—largely convected by the "random" turbulent motion (and any mean motion) of the gases within the combustion chamber. Relative to the unburned mixture ahead of the flame, this mean flame contour moves at the turbulent flame speed, S_T . Locally, each portion of this wrinkled reaction-sheet flame, as indicated in the schematic, moves orthogonally into the unburned mixture ahead of it at the local laminar flame speed, S_L . The wrinkled laminar flame sheet is thin: the turbulent flame brush—the average distance separating the leading and trailing locations of the wrinkled laminar flame—is thick.^{22, 23}

Figure 9.22 Schematic of the structure of the turbulent SI engine flame. Note the relevant flame-defining parameters.^{22, 23}



Laminar premixed flames are characterized by their laminar flame speed S_L and a laminar flame thickness δ_L (see Sec. 9.1.1). The laminar flame speed is the velocity at which the flame propagates into quiescent unburned mixture ahead of the flame. There are several ways to define the thickness δ_L of a laminar flame.²⁴ Dimensional arguments give the most commonly used definition. It is based on the thermal diffusivity $\alpha (= k/\rho c_p)$ as indicated in Fig. 9.1, or on molecular diffusivity D_L (see Sec. 4.8) of radical species through the flame:

$$\delta_L \approx \alpha / S_L \approx D_L / S_L$$

since the Lewis number (= Schmidt number/Prandtl number = α/D_L) is close to unity.

Turbulent flames are also characterized by the root mean square value of the velocity fluctuations, the turbulence intensity u' [Eq. (8.3)], and the various length scales of the turbulent unburned-mixture motion ahead of the flame. These various turbulence length scales are the integral length scale l_I [Eq. (8.8)] (a measure of the size of the large energy-containing structures of the flow) and the Kolmogorov scale l_K [Eq. (8.11)] which defines the smallest structures of the flow where kinetic energy is dissipated via molecular viscosity. Several other dimensionless parameters characterize turbulent premixed flames. The dimensionless parameter used to define the turbulence is the turbulent Reynolds number, $Re_T = u' l_I / \nu$. For homogeneous and isotropic turbulence (i.e., no preferred direction), the integral and Kolmogorov scales are related by Eq. (8.14):

$l_K / l_I = Re_T^{-3/4}$. A characteristic turbulent eddy turnover time τ_T can be defined as

$$\tau_T = \frac{l_I}{u'}$$

A characteristic chemical reaction time is the residence time in a laminar flame:

$$\tau_L = \frac{\delta_L}{S_L}$$

The ratio of the characteristic eddy turnover time to the laminar burning time is called the *Damköhler number*.

$$Da = \frac{\tau_T}{\tau_L} = \left(\frac{l_I}{\delta_L} \right) \left(\frac{S_L}{u'} \right)$$

(9.32)

It is an inverse measure of the influence of the turbulent flow on the chemical processes occurring in the flame. Other ratios are of interest. The ratio δ_L/l_K is a measure of the stretch or local distortion to which a laminar flame is subjected by the turbulent flow. Unless $l_K/\delta_L \gg 1$, the concept of a localized flame region has little significance. The ratio u'/S_L is a measure of the relative impact of the turbulence on the flame.

An additional consequence of the turbulence is the local stretching or straining of the reaction sheet that this flame wrinkling produces. Flame staining is quantified by the product of the Lewis number, Le , and the Karlovitz Stretch Factor, K . The Lewis number is the ratio of thermal diffusivity α to molecular diffusivity D : that is, $Le = \alpha/D$. The Karlovitz Stretch Factor is the product of the turbulent strain rate and the transit time τ_L for flow through the laminar flame:

$$K = \left(\frac{1}{A_L} \cdot \frac{dA_L}{dt} \right) \tau_L = \left(\frac{u'}{\ell_T} \right) \left(\frac{\delta_L}{S_L} \right)$$

(9.33)

where A_L is the laminar flame area, u' is the turbulence intensity, ℓ_T is the Taylor microscale of the turbulence [see Eq. (8.12)], δ_L is the laminar flame thickness, and S_L is the laminar flame speed. The higher the value of the product LeK , the more important is the impact of flame stretch in reducing the laminar burning velocity. Under normal engine operating conditions, the Lewis number (for a stoichiometric mixture) is close to 1.0, and K is of order 0.1 so the effect is small. For lean mixtures, Le increases above unity, and higher engine speeds increase the amount of stretch. At the lean operating limit of the engine (see Sec. 9.4.3), these combined effects are the likely cause of misfire—the extinguishing of the flame.

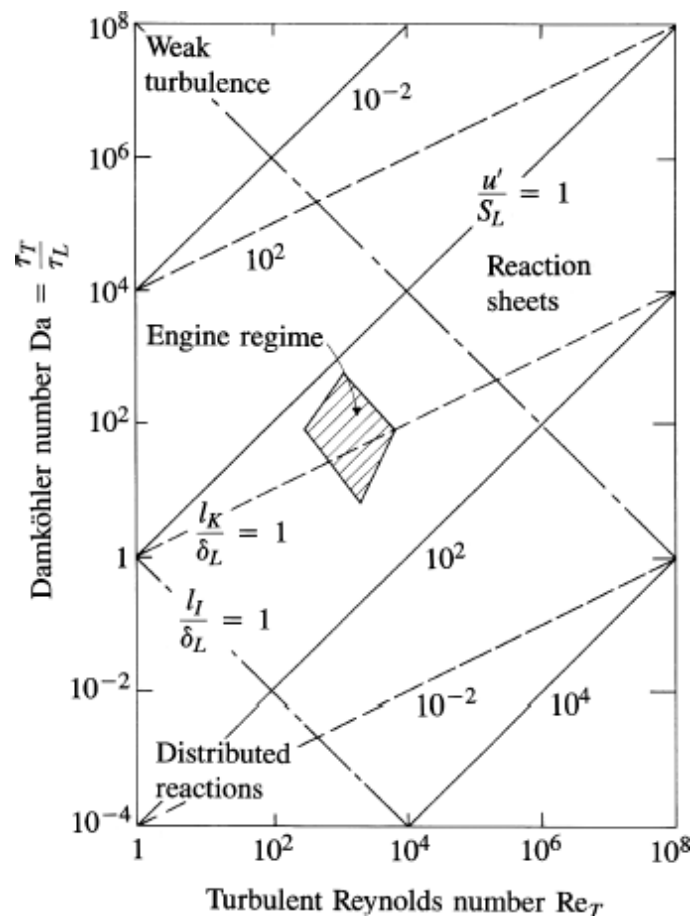
Table 9.1 gives values for the parameters, velocities, and length scales that characterize this engine flame under typical operating conditions. Note that u' and S_T increase with increasing engine speed (which is why fast-enough combustion can be achieved over the full range of engine speeds—600 to 6000 rev/min). Turbulence length scales decrease with increasing speed, as expected from basic turbulence relationships. Adding air to a stoichiometric mixture (which makes it lean) or extra residual and/or EGR (which dilutes the unburned mixture) reduces the laminar flame speed by decreasing the temperature and concentration gradients in the laminar flame: see Fig. 9.1.

Table 9.1 Typical values of quantities which characterize SI engine flames* (Speed = 1500 rev/min; stoichiometric mixture)

Turbulence intensity, u'	2 m/s
Turbulent Reynolds no., Re_T	300
Damköhler no., Da	20
Karlovitz Stretch Factor, K	0.2
Integral scale, l_I	2 mm
Taylor microscale, l_T	0.7 mm
Kolmogorov scale, l_K	0.03 mm
Laminar flame thickness, δ_L	0.02 mm
Laminar flame speed, S_L	0.5 m/s
Ratio u'/S_L	4
Ratio S_T/S_L	4
Mean flame radius of curvature	2 mm
*Values taken from Refs. 22, 24–26.	

Figure 9.23 is a plot of Damköhler number versus turbulent Reynolds number. It shows different regimes of turbulent flames. It has been assumed that $\alpha \approx D_L \approx \nu$ (the kinematic viscosity) and that the relationships for homogeneous isotropic turbulence are valid. Two regimes—distributed reactions and reaction sheets—are normally identified. In the distributed reaction regime, chemical reactions proceed in distributed reaction zones and thin-sheet flames do not occur. A sufficient condition for this regime is $l_I \ll \delta_L$. In the reaction sheet regime, propagating reaction fronts are wrinkled and convoluted by the turbulence. A sufficient condition for the existence of reaction sheets is $l_K \gg \delta_L$. For $Re_T > 1$, there is a region in Fig. 9.23 where $l_I > \delta_L > l_K$: the characteristics of flames in this regime are unclear.

Figure 9.23 Different turbulent flame regimes on a plot of Damköhler number versus turbulent Reynolds number: u' is turbulence intensity; S_L is laminar flame speed; l_I , l_K , and δ_L are the integral scale, Kolmogorov scale, and laminar flame thickness, respectively.²⁷

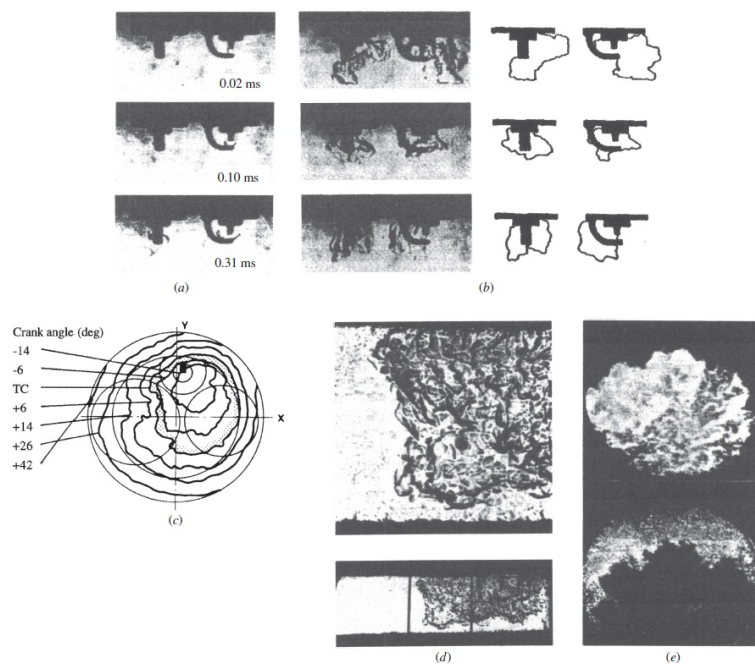


Values of Da and Re_T for a typical spark-ignition engine operation (the cross-hatched region in Fig. 9.23) lie predominantly in the reaction sheet flame regime. Engine operation at high speed (the lower right boundary) and low load (the lower left boundary) gives values of Da and Re_T which fall below the $l_K/\delta_L = 1$ line. This is largely due to the low values of laminar flame speed that result from the high amounts of residual gas and EGR under these conditions (see Sec. 9.3.4). Whether the flame structure under these conditions is significantly different is not known. Observations of engine flames to date, described below, lie above the $l_K/\delta_L = 1$ line, within the reaction sheet regime. One would expect the structure of the flame in a spark-ignition engine, once developed, to be that of a thin reaction sheet wrinkled and convoluted by the turbulent flow.

Detailed observations have been made of flame structure from ignition to the flame extinguishing at the far cylinder wall. A flame develops from the spark discharge, which initiates combustion as follows. In the breakdown phase of ignition (first 100 nsec), a cylindrical discharge between the spark plug electrodes of some 1 mm diameter is established. As electrical energy is fed into the discharge, the plasma reaches temperatures of order 60,000 K, the discharge expands and exothermic chemical reactions rapidly occur, initiating a propagating flame.²³ Figure 9.24a shows how this development of a flame kernel occurs, with three schlieren photographs taken at 20, 100, and 310 μ s after spark discharge onset, of the spark plug electrode gap in a special visualization engine. This sequence, of duration approximately 300 μ s, corresponds to about 2 crank-angle degrees. The outer boundary of this developing flame kernel is roughly spherical, except where impeded by the plug electrodes, and is smooth with modest irregularities corresponding to a thin reaction zone with high-temperature gases inside. The extremely high discharge-generated gas temperatures result in rapid heat conduction to the surrounding gas and the electrodes. When the flame radius is small (~ 1 mm), flame curvature and stretch effects on the growth rate are significant. Then, as this developing sheet-like flame grows, it interacts with the turbulent flow field in the vicinity of the spark plug: the flame's outer surface becomes increasingly convoluted, and the flame center can be convected away from the plug in a direction and with a

velocity that can vary substantially cycle-by-cycle, as seen in Fig. 9.24b.²⁸ Both the bulk flow (through flame kernel convection), and the turbulence intensity (and corresponding turbulence length scales) in the vicinity of the spark discharge affect this stage of the flame development process.

Figure 9.24 Images of SI engine flames using schlieren photography, flame radiation and laser-induced fluorescence techniques. (a) Photos of spark discharge and flame kernel development from a single cycle; stoichiometric operation, 1400 rev/min, mixture pressure and temperature at ignition 2 bar and 625 K. (b) Three consecutive cycles 0.84 ms (7 crank-angle degrees) after spark onset, showing the developing turbulent flame and its outer boundary; engine operating conditions as in (a).²⁸ (c) Flame contours illustrating flame propagation in one cycle: flame outer boundaries obtained from high-speed schlieren cinematography, and with multi-ionization probes; note the approximately spherical outer shape of the flame; 3000 rev/min, bmep = 5 bar, A/F = 13, spark 45° BTC, peak pressure 20 bar.²⁹ (d) Schlieren photo of a fully developed turbulent wrinkled-reaction-sheet SI engine flame: lower photo shows full square-cross-section combustion chamber; upper photo shows center portion enlarged. Operating conditions: stoichiometric, propane fuel, 1400 rev/min, 0.5 bar intake pressure.^{30,31} (e) Simultaneous images of flame radiation (top) and planar laser-induced-fluorescence of unburned hydrocarbon fuel (bottom); image timing is 4° BTC; imaging plane is through the disc-shaped clearance volume.³²



The increasingly wrinkled flame surface moves outward—due to both the flame propagation into the unburned mixture ahead of the flame *and* the expansion of each element of gas as it burns (Sec. 9.2.1)—in an approximately spherical manner from its center, as shown in Fig. 9.24c.²⁹ The flame contours shown define the fraction of the combustion chamber volume enflamed. Note that (see Sec. 9.2.1) with a centrally located spark plug (as in four-valve per cylinder SI engines) in a disc-shaped combustion chamber, only about 10% of the charge has burned when the flame radius is half the cylinder radius.

The wrinkled structure of the flame continues to develop as it propagates across the chamber. Evidence, largely from schlieren photographs and studies of flame structure with laser diagnostics, shows that early in the burning process the flame is a thin, moderately wrinkled but simply connected, reaction sheet between unburned and burned gas. The thickness of this flame front is about 0.1 mm, essentially the thickness of a laminar flame under the prevailing conditions. The scale of the wrinkles is typically 1 or 2 mm at engine speeds of 1000 to 2000 rev/min. As the flame propagates across the chamber, the thickness of the reaction sheet front remains roughly constant, the sheet-like flame becomes more convoluted, and the scale of the wrinkles tends to decrease with time.³⁰

Additional insight into the structure of the developed engine flame can be obtained by examining the flame's leading edge. [Figures 9.24d](#) and [e](#) show results from three different experimental techniques often used to visualize engine flames in their turbulent stage. [Figure 9.24d](#) is a schlieren image of a turbulent engine flame in a square-cross-section visualization engine.³¹ The flame shown was initiated at the spark plug electrode gap in the upper right-hand side of the lower photo and is propagating downward toward the piston crown and to the left toward the far cylinder wall. A portion of this photo is shown enlarged at the top. Schlieren photographs are integrated images along the optical path of the parallel light beam. Given that this flame is propagating outward in a roughly spherical manner, the schlieren image of the flame moving to the left can be interpreted as a tangential view of the spherical turbulent wrinkled reaction-sheet flame, where those sheet regions that have moved ahead of the mean sheet or flame location appear as "hills," stacked beside and behind each other "on the horizon." That portion of the schlieren image that is viewed perpendicular to the plane of the mean flame location appears as an aerial view of "hilly terrain," where the "bottoms of the valleys" appear as lines. A feature of the wrinkled flame sheet model, where the turbulence distorts and convolutes the reaction sheet, is that the local laminar forward-diffusion of the sheet rounds the leading flame sheet regions and forms cusps at the trailing flame sheet regions (see [Fig. 9.22](#)).

[Figure 9.24e](#) shows simultaneous images obtained from flame radiation, and from planar laser-induced fluorescence (LIF) from fuel vapor in the unburned mixture in an optically accessible spark-ignition engine with a disc-shaped combustion chamber.³² The flame radiation technique produces an outer flame image similar to that produced by the schlieren technique. Two-dimensional planar imaging of flames has become a well-developed and valuable technique for examining the details of flame structure. The example shown, which identifies the unburned mixture region in a plane through the clearance height, shows finger-like regions of unburned gas that have penetrated into the turbulent flame, and similar regions of burned mixture that have moved ahead of the mean reaction-sheet location. These two-dimensional planar images usually show a more irregular boundary between the burned and unburned regions than is apparent with the schlieren visualization technique, which integrates density-gradient-produced refractive index changes along the optical path.

As explained earlier in this section, the turbulent spark-ignition flame is in the wrinkled reaction-sheet regime of turbulent combustion. The turbulent Reynolds number is in the range of 100 to 1000. The Damköhler number, which is the ratio of the turbulent eddy turnover time to the residence time in the laminar flame, is of order 100 (the fast chemistry regime), and the smallest turbulent eddy size is larger than the laminar flame thickness so the reaction sheet structure is little affected by the turbulence. The turbulence wrinkles and distorts the flame, increasing the area of thin reaction sheet contained within the turbulent flame zone (often called the flame brush) by a factor that is substantial and that increases with increasing turbulence. Most important, this increasing reaction-sheet area increases the mass-burning rate. An opposing effect that wrinkling also produces is the stretching of the flame sheet, which slows down the molecular diffusive processes within the flame. Whether this flame-stretching effect is significant depends on strain rate and mixture composition.

The turbulence intensity— u' the root mean square value of the fluctuating in-cylinder flow velocity—at time of combustion, has a value about equal to the half the mean piston speed. Hence, the ratio of reaction sheet area (often called the laminar flame area) to the turbulent flame frontal area (for which *flame envelope area* is a useful term) increases with increasing engine speed. It is primarily for this reason that the mass burning rate, which is proportional to the reaction sheet area, increases almost linearly with engine speed, thereby maintaining the crank-angle interval required to burn from 10 to 90% of the in-cylinder charge almost constant over the full speed range of the engine. This ratio of laminar to turbulent flame area typically has a value of about 10 at 2000 rev/min.

The final stage of combustion, when the flame approaches the wall and extinguishes, has been much less extensively examined. There is a substantial literature on laminar flame quenching at a cool wall leading to the formation of quench layers—a thin layer of unburned fuel-air mixture adjacent to the wall ([Sec. 11.4.3](#)). It is often assumed that the local (laminar) regions of the wrinkled turbulent flame that come in contact with the wall quench as just described, while the remainder of the turbulent flame brush burns out as if the wall were not there. This inherently assumes that the unburned gas thermal boundary layer thickness is small relative to the turbulent flame brush thickness. However, as the flame approaches the wall, only successively smaller scales of turbulence are available to affect the flame so the final burn-up process will occur at a slower rate than would occur in the developed turbulent flame.

In terms of the engine *combustion system*—the chamber geometry and plug location, and the geometry of the cylinder head which determines the in-cylinder flow during combustion, the primary quantities of interest are:

- The mass burning rate of the unburned mixture, which through energy conservation, is directly linked to cylinder pressure and hence engine performance;
- The cycle-by-cycle variation in the mass burning rate because this, via torque fluctuations, determines the stability of engine operation;
- The burned gas heat transfer areas, which have a significant, though secondary effect on the energy balance.

In summary, at a more fundamental level, the factors that affect the mass-burning rate are the following:

- The mean flame frontal (or envelope) area contained within the boundaries of the combustion chamber;
- The relevant bulk and turbulence flow parameters during combustion—the local mean flow, the local turbulence intensity, and the flame-wrinkling length scales;
- The local unburned mixture composition and state, which determine the local laminar burning velocity.

9.3.3. Laminar Burning Speeds

An important intrinsic property of a combustible fuel, air, burned gas mixture is its laminar burning velocity. This burning velocity is defined as the velocity, relative to and normal to the flame front, with which unburned gas moves into the front and is transformed to products under laminar flow conditions. Details of the structure of this laminar flame, shown in [Fig. 9.1](#), explain the significance of this quantity. A flame is the result of a self-sustaining chemical reaction occurring within the flame front where unburned mixture has been heated sufficiently to spontaneously combust. The flame consists of two major regions: a preheat and conduction zone, and a reaction zone. In the preheat zone, the temperature of the unburned mixture is raised mainly by heat conduction from the reaction zone: little reaction or energy release occurs and the temperature gradient is concave upward. Upon reaching the mixture's ignition temperature, exothermic chemical reactions begin. The conversion of chemical energy to thermal energy results in a zone where the temperature gradient is concave downward called the reaction zone. The thicknesses of the preheat and reaction zones can be calculated for one-dimensional flames from conservation equations of mass and energy. The thickness of the preheat zone $\delta_{L,ph}$ is

$$\delta_{L,ph} = \frac{4.6\bar{k}}{c_p\rho_u S_L}$$

(9.34)

where
 $\bar{k}$ and

c_p are the mean thermal conductivity and specific heat at constant pressure in the preheat zone, and S_L is the laminar burning velocity. Thus, the factors which govern the laminar burning velocity of a specific unburned mixture—the velocity at which this flame structure propagates relative to the unburned gas ahead of it—are the temperature and species concentration gradients within the flame, and the mixture's transport and thermodynamic properties.

Laminar burning velocities at pressures and temperatures typical of unburned mixture in engines are usually measured in spherical closed vessels by propagating a laminar flame radially outward from the vessel center. The laminar burning velocity is then given by

$$S_L = \frac{dm_b/dt}{A_f \rho_u}$$

(9.35)

where the mass burning rate is determined from the rate of pressure rise in the vessel and A_f is the flame area. Because the laminar flame thickness [e.g., given by Eq. (9.33)] under engine conditions is of order 0.2 mm and is therefore much less than characteristic vessel dimensions, in applying Eq. (9.34) the flame can be treated as negligibly thin. Laminar burning velocities for methane, propane, isooctane, methanol, ethanol, gasoline, and hydrogen—premixed with air—at pressures, temperatures, and equivalence ratios which occur in engines have been measured using this technique.^{33–37} Also, the effect of a burned gas diluent on laminar burning velocity with gasoline-air mixtures has been determined.³⁶ Correlations derived from these data are the most accurate means available for estimating laminar burning velocities for mixtures and conditions relevant to spark-ignition engines.

The effect of the mixture fuel/air equivalence ratio on laminar burning velocity for several hydrocarbon fuels and alcohols is shown in Fig. 9.25. The burning velocity peaks slightly rich of stoichiometric. The values for isooctane and gasoline are closely comparable. Data at higher pressures and temperatures have been fitted to a power law of the form:

$$S_L = S_{L,0} \left(\frac{T_u}{T_0} \right)^\alpha \left(\frac{p}{p_0} \right)^\beta$$

(9.36)

where $T_0 = 298$ K and $p_0 = 1$ atm are the reference temperature and pressure, and $S_{L,0}$, α and β are constants for a given fuel, equivalence ratio, and burned gas diluent fraction. For propane, isooctane, ethanol, and methanol, these constants can be represented by

$$\alpha = 2.18 - 0.8(\phi - 1)$$

(9.37a)

$$\beta = -0.16 + 0.22(\phi - 1)$$

(9.37b)

and

$$S_{L,0} = B_m - B_\phi(\phi - \phi_m)^2$$

(9.38)

where ϕ_m is the equivalence ratio at which $S_{L,0}$ is a maximum with value B_m . Values of ϕ_m , B_m , and B_ϕ are given in Table 9.2.³⁴

Figure 9.25 Laminar burning velocity for several fuels as a function of fuel/air equivalence ratio, at 1 atm and 300 K. Lines are least-squares polynomial fits to data.^{33,34,35}

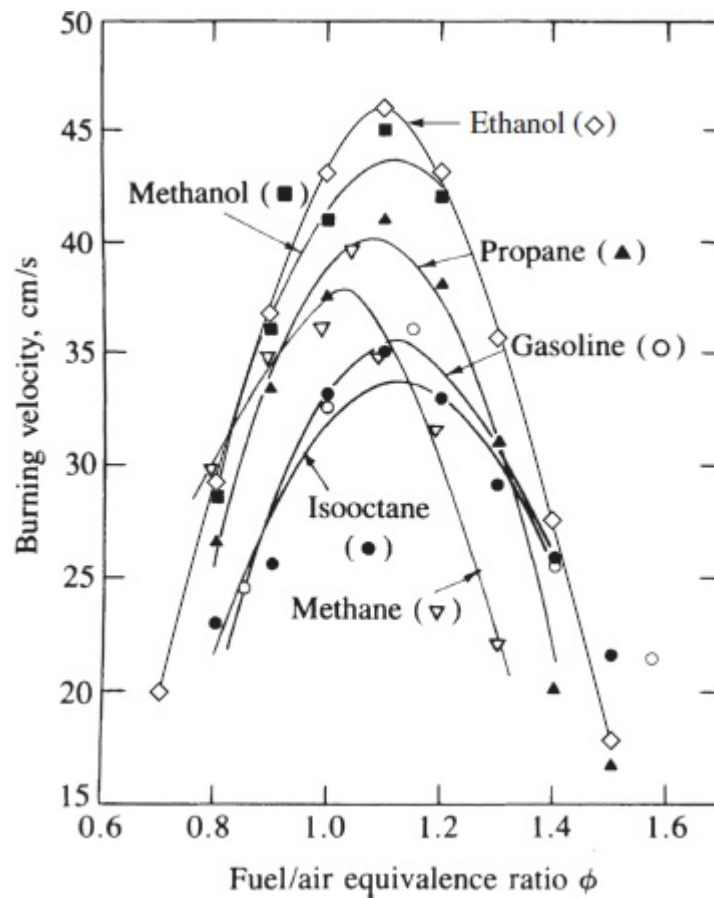


Table 9.2 Parameters ϕ_m , B_m , and B_ϕ for Eq. (9.38)

Fuel	ϕ_m	B_m , cm/s	B_ϕ , cm/s	Ref.
Gasoline	1.21	30.5	-55	36
Isooctane	1.13	26.3	-85	34
Propane	1.08	34.2	-139	34
Ethanol	1.10	40.0	-200	37
Methanol	1.11	36.9	-140	34

Note: Values of $S_{L,0}$ given by Eq. (9.38) are obtained from least-squares fits of Eq. (9.36) to data over the range $p = 1$ to 8 atm, $T_u = 300$ to 700 K. They do not correspond exactly to the laminar flame speed data at 1 atm and 298 K in Fig. 9.25.

For gasoline (a reference gasoline with average molecular weight of 107 and an H/C ratio of 1.69), additional data were available and were correlated by³⁶

$$\alpha_g = 2.4 - 0.271\phi^{3.51}$$

(9.39a)

$$\beta_g = -0.357 + 0.14\phi^{2.27}$$

(9.39b)

For methane, simple equations such as (9.37a, b) do not adequately correlate the data over the range of p and T_u relevant to engines. However, laminar burning velocity data from a spherical constant-volume bomb experiment have been obtained along an unburned gas isentropic path, as the pressure in the bomb rises during combustion. Variation in laminar burning velocity along such unburned gas isentropes does correlate with a power law:

$$S_{L,s} = S_{L,0} \left(\frac{\rho_u}{\rho_{u0}} \right)_s^\varepsilon$$

(9.40)

Values for $S_{L,0}$ and ε from the literature are summarized in Table 9.3.

Table 9.3 Parameters for methane-air laminar burning velocity correlation [Eq. (9.40)]

ϕ	p_i atm	$S_{L,0}^*$ cm/s	ε
1.0	0.5	49	0.51
1.0	1.0	35	0.2
0.8–1.2	1–8	†	0.17–0.19
*At 298 K initial temperature.			
†See Fig. 9.25.			

The presence of burned gas in the unburned cylinder charge due to residual gases and any recycled exhaust gases causes a substantial reduction in the laminar burning velocity. Any burned gas in the unburned mixture reduces the heating value per unit mass of mixture and thus, reduces the adiabatic flame temperature. It acts as a diluent. The effect of increasing burned gas or diluent fraction on laminar flame speed is shown in Fig. 9.26. The diluent used was a mixture of CO_2 and N_2 , chosen to match the heat capacity of actual gasoline-air combustion products.^[1] The proportional reduction in laminar burning velocity is essentially independent of the unburned mixture equivalence ratio, pressure, and temperature over the range of interest in engines. The data in Fig. 9.26 are correlated by the relation:

$$S_L(x_b) = S_L(\tilde{x}_b = 0)(1 - 2.06)\tilde{x}_b^{-0.77}$$

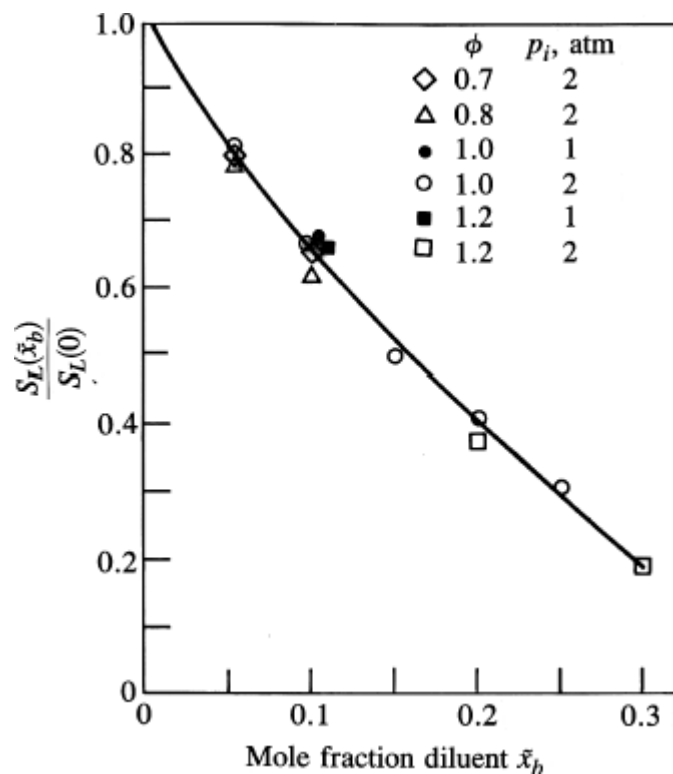
(9.41)

where

$\tilde{x}_b$ is the mole fraction of burned gas diluent. Other studies corroborate the magnitude of this burned gas effect.³⁶

Figure 9.26 Effect of burned gas mole fraction

$\tilde{x}_b$ in unburned mixture on laminar burning velocity. Fuel: gasoline.³⁶



Note that for equal heat capacity added to the unburned mixture, burned gases have a much larger effect on laminar burning velocity than does excess air. For example, the laminar burning velocity of a stoichiometric mixture as it is leaned to $\phi = 0.8$ is reduced by 23%. Adding the same heat capacity of stoichiometric burned gases (which requires a burned gas mole fraction of 0.175) reduces the laminar burning velocity by 55%.³⁶ Allowance for the burned gas fraction in estimating laminar burning velocities for spark-ignition engines is most important.

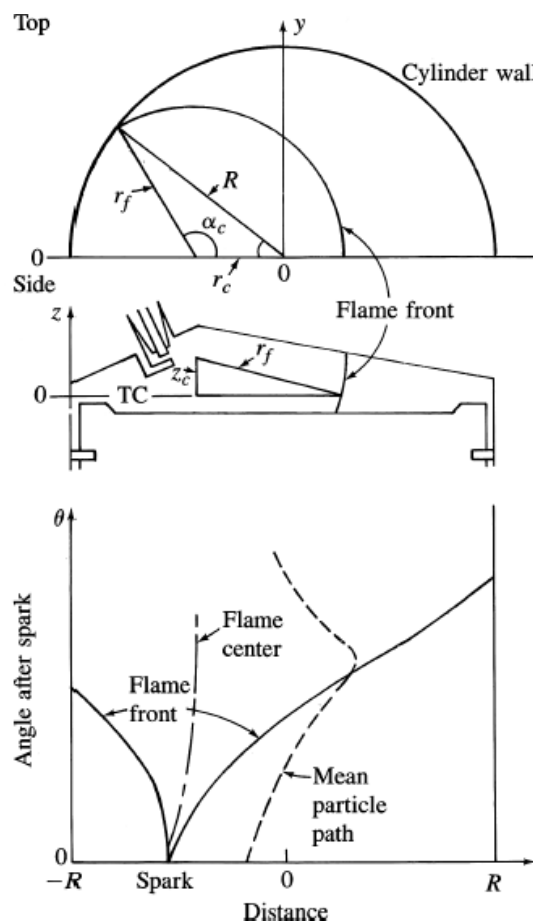
The above correlations define the laminar burning velocity (of a given fuel) as a function of unburned mixture thermodynamic properties and composition only. It has been assumed that flame thickness and curvature effects are negligible. Our interest in laminar burning velocity is twofold: first, it is used to define the characteristic chemical reaction time of the mixture in Eq. (9.32); second, a presumed consequence of the wrinkled thin-reaction-sheet turbulent-flame structure is that, locally, the sheet propagates at the laminar burning velocity. The above correlations adequately characterize a quiescent burning process. However, laminar flame propagation can be influenced by the local flow field in the unburned gas. If the flame thickness is less than the Kolmogorov scale, the primary effect is one of straining, which affects both the flame area (usually referred to as flame stretching for an area increase) and the local (laminar) burning velocity. Local straining can affect the laminar burning velocity and can cause flame extinction. The laminar burning velocity decreases with increasing strain rate, and the Lewis number of the unburned mixture has a significant influence on this rate of decrease. The Lewis number is the ratio of diffusivities of heat and mass. For stoichiometric mixtures it is close to one; it increases above unity as the unburned fuel-air mixture is leaned out. Thus, the local flow field may have a discernable effect on the local burning velocity of the thin laminar-like reaction-sheet flame, especially for lean or dilute mixtures.

9.3.4. Flame Propagation Relations

Combining the heat-release or mass burning rate analysis of Sec. 9.2.2 with an analysis of flame geometry and structure data allows us to quantify the behavior of spark-ignition engine flames. Flame photographs (such as those in Figs. 9.2, 9.18, and 9.20 and 9.24) effectively define the position of the front or leading edge of the turbulent engine flame. The leading "surface"

of the enflamed zone, under normal engine conditions, is close to spherical: only in the presence of high swirl (Fig. 9.20) does substantial distortion of the flame shape occur. The schematic in Fig. 9.22 lays out the two primary approaches to analyzing flame propagation: the wrinkled laminar flame sheet approach, and the turbulent brush model. To a good approximation, the surface that defines the leading edge of the flame (ahead of which only unburned mixture exists) is a portion of the surface of a sphere. Figure 9.27 indicates the geometrical parameters which define this flame surface: r_c , α_c , z_c , the coordinates of the flame center, r_f the radius of the best-fit sphere to the flame front outer boundary, and the geometry of the combustion chamber walls. The flame is initiated at the spark plug; however, the center may move away from the plug during the early stages of its development as shown. Here, we define the *flame front area* A_f as the spherical surface of radius r_f coinciding with the leading edge of the flame contained within the combustion chamber, and the *enflamed volume* V_f as the volume within the chamber behind this flame front.

Figure 9.27 Schematic of spherical flame front in engine combustion chamber identifying parameters which define flame geometry. (From Beretta et al.⁴)



9.3.4.1. Wrinkled Flame Sheet Approach

The thermodynamic analysis of cylinder pressure data allows us to define additional geometrical parameters. The *burned gas radius* r_b is the radius of the spherical surface within the combustion chamber which would contain all the burned gas behind it, that is,

$$V_b(r_b, r_c, \alpha_c, z_c) = V_b(p, \theta)$$

(9.42)

The *spherical burning area* A_b is the area of this spherical surface, that is,

$$A_b = \frac{\partial V_b(r_b, r_c, \alpha_c, z_c)}{\partial r_b}$$

(9.43)

The *laminar burning area* A_L is the surface area the flame would have if it burned locally at the laminar flame speed, that is,

$$A_L = \frac{dm_b/dt}{\rho_u S_L}$$

(9.44)

where S_L is the laminar flame speed in the unburned mixture ahead of the flame (see [Sec. 9.3.3](#)).

Several velocities can be defined. The *mean expansion speed* of the *front* u_f is given by

$$u_f = \frac{dV_f/dt}{A_f}$$

(9.45)

where A_f is the surface area of the "best-fit" sphere through the leading edge of the flame enclosed within the combustion chamber, and V_f is the volume behind that leading flame surface into which unburned mixture is entrained. The *mean expansion speed* of the *burned gas* u_b is

$$u_b = \frac{dV_b/dt}{A_b}$$

(9.46)

The *burning speed* S_b is defined by

$$S_b = \frac{dm_b/dt}{\rho_u A_b}$$

(9.47)

S_b is the (turbulent) flame speed S_T with which the flame frontal or envelope surface (of area A_f) propagates into the unburned gas. The *mean gas speed* just ahead of the flame front u_g is

$$u_g = u_b - S_b$$

(9.48)

Note that combining Eqs. (9.44) and (9.47) gives the relation

$$S_b A_b (= S_T A_b) = S_L A_L$$

Also, it follows from Eqs. (9.31), (9.46), and (9.47) that

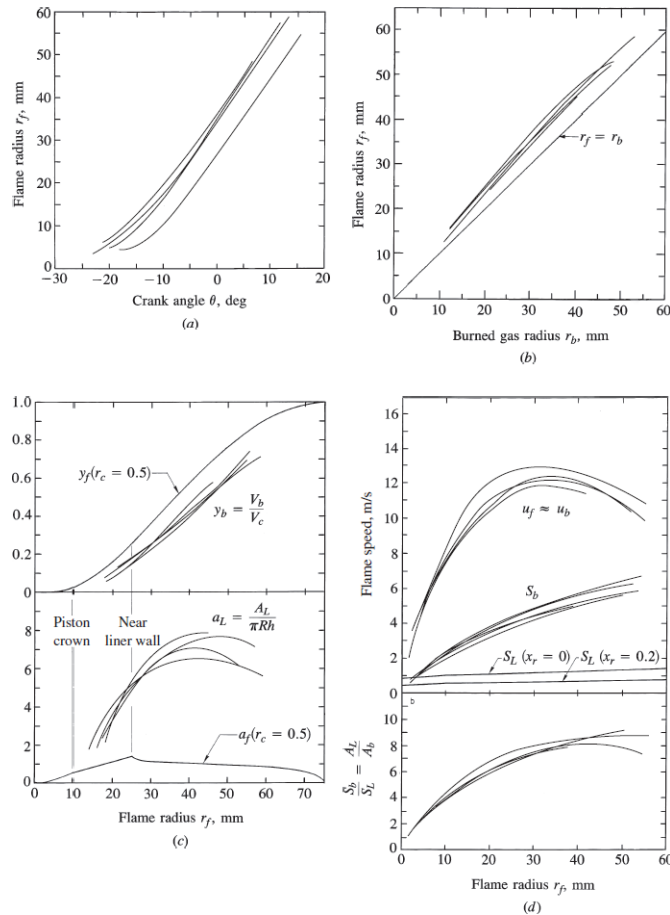
$$\frac{u_b}{S_b} = \frac{\rho_u}{\rho_b} (1 - y_b) + y_b = \frac{\rho_u / \rho_b}{[(\rho_u / \rho_b) - 1] x_b + 1}$$

(9.49)

where S_b and S_T are synonymous. As x_b and $y_b \rightarrow 0$, u_b/S_b approaches the expansion ratio, ρ_u/ρ_b . As x_b and $y_b \rightarrow 1$, u_b/S_b approaches unity.

The variation of the above quantities during the engine combustion process, coupled with the photographs and discussion in Sec. 9.3.2, provides substantial insight into the flame development and propagation process. Figure 9.28 shows results from an analysis of cylinder pressure data and the corresponding flame front location information (determined from high-speed movies through a window in the piston) of several individual engine combustion cycles.⁴ The flame radius initially grows at a rate that increases with time and exhibits substantial cycle-by-cycle variation in its early development (Fig. 9.28a). Later ($r_f \gtrsim 30$ mm) the growth rate, which approximates the expansion speed u_b , reaches an essentially constant value. The flame radius r_f is initially equal to the burned gas radius r_b ; it increases above r_b as the flame grows and becomes increasingly distorted by the turbulent flow field (Fig. 9.28b). Eventually $r_f - r_b$ goes to an essentially constant value of about 6 mm for $r_b \gtrsim 30$ mm. This difference, $r_f - r_b$, is approximately half the thickness of the turbulent flame brush: see Fig. 9.22.

Figure 9.28 Variation of flame geometry and velocity parameters during four individual combustion cycles in an SI engine at 1044 rev/min, $\phi = 0.98$, 1 atm inlet pressure: (a) flame radius r_f versus crank angle; (b) flame radius r_f versus burned gas radius r_b ; (c) normalized enflamed volume y_f , burned volume y_b , normalized flame front area a_f , and laminar area a_L versus flame radius; (d) front expansion speed u_b , burning speed S_b , and laminar flame speed S_L versus flame radius. (From Beretta et al.⁴)



Normalized enflamed and burned volumes, and normalized flame front area and wrinkled laminar-flame burning area, are shown in Fig. 9.28c. Volumes are normalized by the clearance volume and areas by πRh , where h is the average clearance height and R the cylinder radius. Discontinuities occur in the flame area a_f at the points where the flame front contacts first the piston face and then the near cylinder wall. The laminar flame area A_L is initially close to the flame area A_f , and then increases rapidly as the flame grows beyond 10 mm in radius (Fig. 9.28d). During the rapid burning combustion phase the value of y_f is significantly greater than y_b . During this phase, the laminar area grows to exceed the flame surface or envelope area by almost an order of magnitude. These observations indicate the existence of substantial pockets of unburned mixture behind the leading edge of the flame.

The ratio of the volume of the unburned mixture within the turbulent flame zone ($V_f - V_b$) to the reaction-sheet area within the flame zone ($A_L - A_f$) defines a characteristic length

$$\ell_T = \frac{V_f - V_b}{A_L - A_f}$$

(9.50)

which can be thought of as the scale of the pockets of unburned mixture within the flame. For the data set of Fig. 9.28, ℓ_T is approximately constant and about 1 mm.³⁰

These flame geometry results would be expected from the previous photographic observations of how the flame grows from a small approximately spherical smooth-surfaced kernel shortly after ignition to a highly wrinkled reaction-sheet turbulent flame of substantial overall thickness. Initially, the amount of unburned gas within the enflamed volume is small. During the rapid burning phase of the combustion process, however, a significant fraction (some 25%) of the gas within the turbulent flame brush is unburned.

The front expansion speed u_f , burning speed S_b , and laminar flame speed S_L are shown in Fig. 9.28d. The expansion speed increases as the flame develops to a maximum value that is several times the mean piston speed of 3 m/s. The burning speed S_b (the turbulent flame speed S_T) increases steadily from a value close to the laminar flame speed at early times to almost an order of magnitude greater than S_L during the rapid burning phase. During this rapid burning phase, since $(r_f - r_b)$ is approximately constant, the flame front expansion speed and the mean burned gas expansion speed are essentially equal. The difference between $u_b \approx u_f$ and S_b is the unburned gas speed u_g just ahead of the flame front. Note that the ratio $u_b/S_b (\approx u_f/S_b)$ decreases monotonically from a value close to the expansion ratio (ρ_u/ρ_b) at spark, to unity as the flame approaches the far wall, as required by Eq. (9.49).

The effect of flame propagation on the flow field in the unburned mixture ahead of the flame is important because it is the turbulence just ahead of the flame that determines the local burning velocity. Note that the mean flow varies cycle-by-cycle, the turbulence is not homogeneous, and the flame motion and shape show significant cyclic variations. The mean velocity normal to the front increases steadily from shortly after ignition, as the combustion-produced gas expansion displaces unburned mixture toward the walls. Whether the turbulence intensity increases due to the compression of the unburned mixture, which occurs during combustion, is unclear. Rapidly imposed distortions of a turbulent flow field produced by combustion-generated gas expansion would lead to an increase in turbulence intensity.

The variation of burning speed with engine speed and thus with the turbulence intensity, has also been carefully examined. A linear correlation between S_b and u' results for the rapid burning combustion phase.³⁸

Once the flame front reaches the far cylinder wall, the front can no longer propagate, however, combustion continues behind the front until all the unburned mixture entrained into the enflamed region is consumed. This final burning or termination phase of the combustion process can be approximated by an exponential decay in the mass-burning rate with a characteristic time constant τ_b (which varies with engine speed, and is of order 1 ms). Since these "peninsulas or islands" of unburned mixture behind the leading edge of the flame have a characteristic scale l_T , based on the laminar flame area [Eqs. (9.44) and (9.50)], it follows that

$$l_T = \tau_b S_L$$

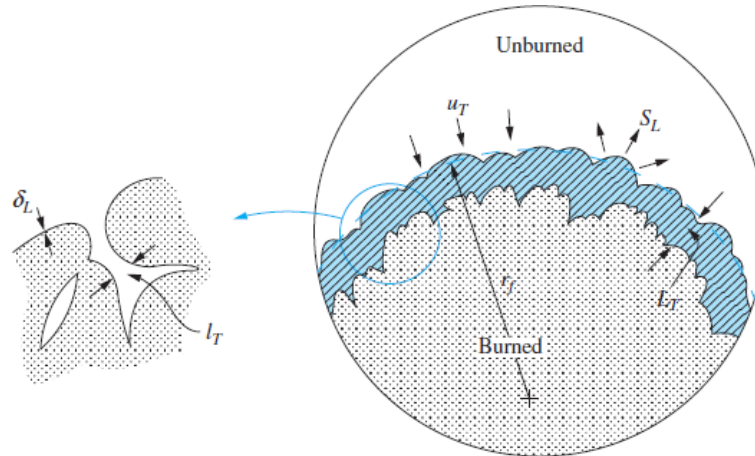
(9.51)

In summary, the above flame data analysis shows that the relationships between r_f and r_b , V_f and V_b , A_f and A_L , u_f , S_b or S_T and S_L are distinctly different in the three phases of combustion: (1) the development phase, where a growing, increasingly-wrinkled, thin reaction-sheet evolves into a turbulent flame from the essentially spherical small flame kernel established by the spark discharge; (2) the rapid-burning phase, where this thick "developed" turbulent flame propagates across the combustion chamber to the far wall during which most of the charge mass is burned; and (3) the termination phase after the flame front has reached the far wall and propagation of the front is no longer possible, when the remaining unburned mixture within the flame burns up. The burning velocity, in the rapid-burning phase of the combustion process, scales with turbulence intensity, which in turn scales with engine speed.

9.3.4.2. Turbulent Flame Brush Approach

The second approach to describing flame propagation is based on *entrainment* of unburned mixture into the "thick" turbulent flame brush, illustrated in Fig. 9.29. As unburned mixture crosses the flame surface a portion of it burns immediately; then a substantial fraction burns *within* the turbulent flame brush. This part of the entrained unburned mixture forms peninsulas, even islands, each surrounded by a thin reaction-sheet flame. These inclusions then burn up at a rate determined by their characteristic size and local laminar burning velocity S_L , as shown. Combustion in the turbulent flame can be thought of as an unburned mixture entrainment process, followed by a local diffusion-controlled burning process.

Figure 9.29 Schematic of turbulent premixed spark-ignition engine flame, illustrating the entrainment-based burning law of Eqs. (9.52) to (9.54). The approximately spherical front of the "thick" flame (dashed line) diffuses outward at the laminar flame speed S_L . Fresh mixture also crosses this front at a characteristic velocity u_T due to turbulent convection. Schematic on left shows flame structure; δ_L is the reaction-sheet thickness, l_T is the characteristic scale of wrinkles in the sheet.



This conceptual model, illustrated in Fig. 9.29, results in two rate equations: one for the mass of unburned-mixture entrained into the flame front (or equivalently, entrained as that front moves into the unburned mixture) and a second for the mass burning rate (Keck et al. ^{4,30,39}). Turbulent convection brings fuel-air mixture into the flame at velocity, u_T . Molecular diffusion of the laminar thin reaction-sheet flame propagates it forward at the laminar flame speed, S_L . The two phenomena effectively add. Thus, the rate of entrainment of unburned mixture into the flame is

$$\frac{dm_e}{dt} = \rho_u A_f S_L + \rho_u A_f u_T (1 - e^{-t/\tau_b})$$

(9.52)

where u_T is the characteristic speed of the turbulent convective transport process. The first term represents molecular (laminar) diffusion through the flame front. The second term represents turbulent convection characterized by u_T . The bracketed term allows for the early flame development stage as the initially laminar flame kernel transitions to a turbulent flame. τ_b is a characteristic burning time for unburned mixture within the turbulent flame brush given by

$$\tau_b = \ell_T / S_L$$

(9.53)

The rate at which mixture burns is then,

$$\frac{dm_b}{dt} = \rho_u A_f S_L + \frac{m_e - m_b}{\tau_b}$$

(9.54)

where $(m_e - m_b)$ represents the mass of unburned mixture already entrained into the thick turbulent flame (the cross-hatched region in Fig. 9.29) that has yet to burn:

$$(m_e - m_b) = \rho_u(V_f - V_b) = \rho_u(A_L - A_f)\ell_T$$

(9.55)

where ℓ_T , V_f , V_b , A_L , and A_f are defined earlier in this section.

To summarize, as illustrated in Fig. 9.29, this approach combines the turbulent convective and laminar (molecular diffusive) propagation of the approximately spherical front of the "thick" turbulent flame into the unburned mixture ahead of the flame, with the local, distributed, laminar-like burn-up of the mixture entrained into the flame. Thus the name: "entrainment" or "eddy-burning" combustion model.

Note that for a quiescent mixture, $u_T \rightarrow 0$ or $\ell_T \rightarrow \infty$, and $S_b \rightarrow S_L$. Initially, at $t \approx 0$, $S_b \approx S_L$. In the developed (quasi-steady-state) turbulent flame combustion phase,

$$S_b \approx u_T + S_L$$

(9.56)

During the final burning stage after the flame front reaches the wall, A_f decreases exponentially to zero as does $(m_e - m_b)$, the remaining unburned mass within the flame.

To apply Eqs. (9.52) and (9.54), the quantities u_T and τ_b or $(\ell_T = \tau_b S_L)$ must be evaluated. Two approaches have been taken: (1) use of empirical correlations for these variables, derived from engine flame data (such as those described in Sec. 9.3.4) and (2) use of more fundamental models to predict these quantities.

Keck derived the following correlations for u_T and ℓ_T based on the application of Eqs. (9.52) and (9.54) to several sets of engine combustion data:

$$u_T = 0.08 u_i \left(\frac{\rho_u}{\rho_i} \right)^{1/2} \quad \text{and} \quad \ell_T = 0.8 L_{iv} \left(\frac{\rho_i}{\rho_u} \right)^{3/4}$$

(9.57)

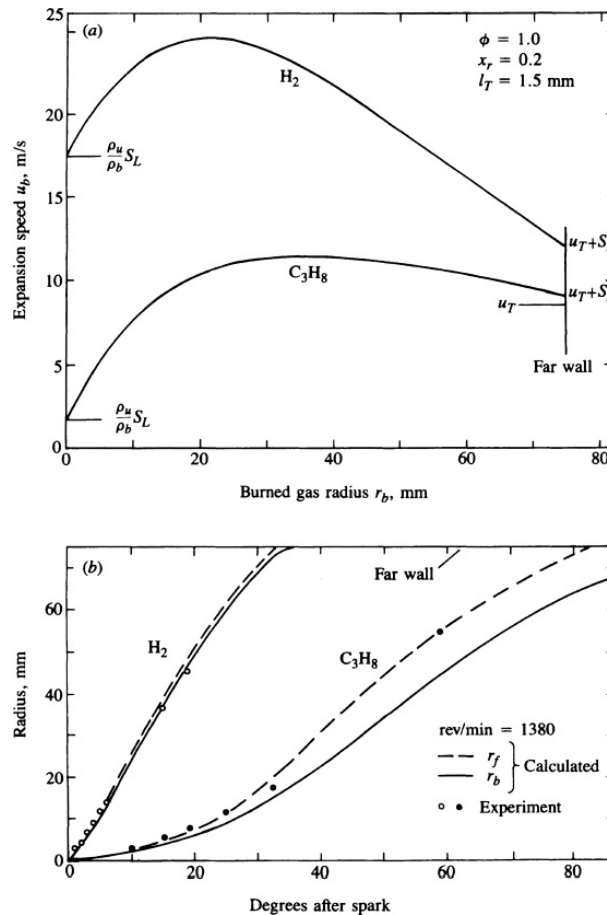
u_T was found to be proportional to

$\sqrt{\rho_u}$ (at time of spark) and to correlate well with mean inlet gas speed

$u_i = \eta_v (A_p / A_{iv}) \bar{S}_p$, where η_v is volumetric efficiency, A_p is piston area, A_{iv} is the maximum open area of the inlet valve, $\bar{S}_p$ is mean piston speed, and ρ_i is intake air density. ℓ_T appears to scale with intake valve lift, L_{iv} ; it decreases with increasing density at a rate proportional to $\rho_u^{-3/4}$. While u_T and ℓ_T are not constant during the combustion process, their variation is modest.³⁰

A quantitative comparison of predicted and measured flame radius as a function of time is shown in Fig. 9.30 for hydrogen and propane fuel-air mixtures which exhibit widely different behavior: the figure indicates both the behavior and validity of this thick flame entrainment model. Predicted burned gas expansion speeds u_b [see Eq. (9.46)] are shown in Fig. 9.30a as a function of burned gas radius; the parameters u_T and ℓ_T were chosen to fit the propane data. Figure 9.30b shows that the measured flame front radii, r_f , are in good agreement with the predicted flame and burned gas radii, r_f and r_b , for these two fuels. The initial expansion speed of hydrogen is about 10 times that of propane. Since $r_f \approx r_b$ for early times and $S_b \approx S_L$, this ratio is expected. As r_b becomes large, $(r_f - r_b) \rightarrow u_T \tau_b$. Note that $\tau_b (= \ell_T / S_L)$ is several times smaller for hydrogen mixtures than for propane mixtures since hydrogen's laminar flame speed is much higher.

Figure 9.30 (a) Calculated burned gas expansion speed u_b during combustion for stoichiometric hydrogen-air and propane-air mixtures as a function of burned gas radius r_b . (b) Comparison of experimentally measured (points) and calculated (dashed curve) flame radii r_f for these mixtures as a function of crank angle. Also shown (solid curve) is the burned gas radius r_b .³⁰



An adaption of this approach developed by Tabaczynski et al.³⁹ used basic turbulent flame propagation theory to argue that ℓ_T in Fig. 9.29 is essentially the Taylor microscale l_M , which is a function of the integral length scale l_I and the turbulent Reynolds number as indicated by Eq. (8.15). From these turbulence assumptions they argued that the flame propagates with a velocity $u' + S_L$, where u' is the local turbulence intensity. Thus, in Eqs. (9.52) and (9.53), u_T and τ_b are given by

$$u_T \approx u' \quad \text{and} \quad \tau_b = \frac{l_T}{S_L} \approx \frac{l_M}{S_L}$$

(9.58)

where l_M , the microscale, is determined from the integral scale and the turbulent Reynolds number via Eq. (8.15) assuming that the turbulence is homogenous and isentropic. The task therefore becomes one of evaluating u' and l_I .

9.3.5. Combustion with Direct Fuel Injection

This chapter has described combustion in spark-ignition engines operating with essentially "premixed" fuel-vapor/air mixtures. The tacit assumption is that the unburned mixture within the cylinder is sufficiently close to homogeneous so that mixture nonuniformities are not important. This is a valid assumption with port fuel injection where the fuel for each cylinder is

injected into the intake port largely onto the intake valve head and port walls in a warmed-up engine at normal operating conditions. At more extreme operating conditions, such as approaching maximum power and when the engine is cold, this close-to-homogeneous mixture assumption is less valid but the engine's combustion process is still only modestly affected. With direct fuel injection into each of the engine's cylinders, the mixture preparation process is significantly different from the port fuel injection process. (See [Chap. 7](#), which discusses SI engine mixture preparation.)

One direct-fuel-injection approach— injection of fuel into each cylinder largely during the intake stroke—does produce a close to homogeneous fuel-vapor/air mixture within the cylinder (as is intended) and combustion is then an essentially premixed turbulent flame propagation process as has been discussed already in this chapter. While this early-direct-injection in-cylinder mixture is "less homogeneous" than mixture resulting from port fuel injection, the flame propagation differences are small. The air pollutant emission differences may be more significant.

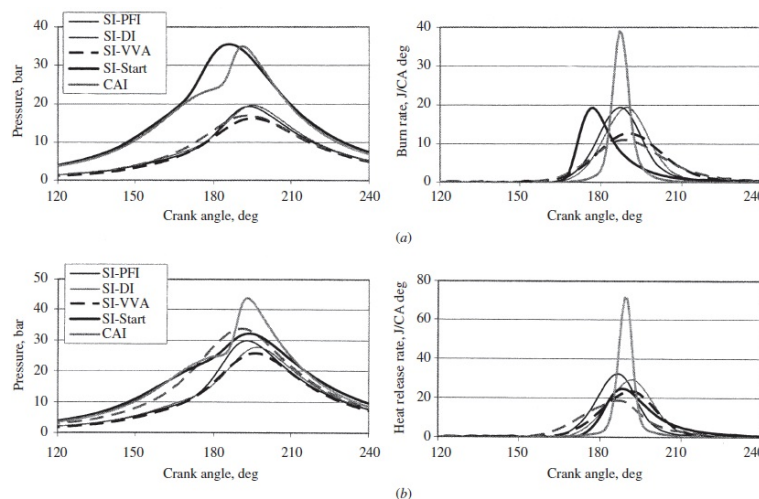
However, direct fuel injection into each engine's cylinder provides the opportunity with *late* fuel injection (during the latter part of compression stroke) to operate the engine *fuel lean* by stratifying the fuel-vapor/air mixture within the combustion chamber so that mixture adjacent to the spark plug is close to stoichiometric (and thus burns rapidly) and the excess air is located around this contained ready-to-burn fuel-vapor/air "cloud". In this lean stratified operating mode, the spark-ignited combustion process is different.

Thus, direct-fuel-injection approaches are categorized as *homogeneous* or *stratified*. In this section, we explore the engine combustion differences under stratified operating conditions. The operating regime where stratified operation would improve efficiency is at engine loads and speeds that are less than about half the maximum values. From mid- to high-loads (and speeds) these stratified DI SI engine concepts transition to homogeneous operation, with the overall relative air/fuel ratio shifting to stoichiometric to realize high specific power and more effective emissions control. The aspects of part-load stratified operation that contribute to higher efficiency are:

- With lean operation overall, compared with stoichiometric, the intake manifold pressures are higher so pumping losses are reduced.
- With excess air, the average burned gas specific heat ratio γ_b is higher (due to lower concentrations of CO_2 and H_2O ; see [Fig. 4.18](#) and [Table 4.7](#)), and average burned-gas temperatures are lower (further increasing γ_b); this increases the expansion stroke work extracted.
- The lower (average) burned gas temperature also results in reduced heat losses to the combustion chamber walls.⁴⁰

[Figure 9.31](#) shows pressure versus crank-angle measurements and corresponding burn rate profiles for several different gasoline engine concepts at light and mid-load. The data come from a single-cylinder engine that could embody the different concepts compared, as well as engine simulation studies.⁴¹ The concepts compared in [Fig. 9.31](#) are: spark-ignited port fuel injection (SI-PFI)—stoichiometric mixture turbulent flame propagation in a port-fuel-injected engine; spark-ignited direct fuel injection (SI-DI)—the same combustion process in a direct-injection engine operated in stoichiometric homogeneous mode; spark-ignited variable valve actuation (SI-VVA)—homogeneous DI operation with late intake valve closing so the engine operates essentially unthrottled at these loads; spark-ignited direct-injection stratified charge (SI-STRAT)—direct fuel injection with a fuel-lean mixture overall, in stratified combustion mode (the primary topic, here); and controlled autoignition (CAI)—a different combustion mode with a lean and dilute largely premixed mixture that ignites spontaneously through control of unburned mixture temperature near TC.^[2] For each of these five engine setups, fuel injection timing and spark timing were adjusted to be close to optimum fuel consumption without compromising combustion stability and the achievement of low engine-out emissions (e.g., through use of substantial EGR).

Figure 9.31 Cylinder pressure and burn rate profiles for several different gasoline engine combustion systems, all tested in the same single-cylinder engine: spark-ignited port fuel injection (SI-PFI), spark-ignited direct fuel injection (SI-DI), spark-ignited variable valve actuation system with high internal EGR (SI-VVA), spark-ignited direct-injection stratified charge (SI-STRAT), and a controlled autoignition concept (CAI) (see Sec. 10.7). For each setup, fuel injection timing and spark timing were adjusted to give close to optimum fuel consumption without compromising combustion stability and emissions. At 2000 rev/min: (a) 3 bar nimep; (b) 5 bar nimep⁴¹. See text.



The SI-PFI and SI-DI, both operating with a homogeneous stoichiometric mixture, show closely comparable pressure and burning rate profiles (at the higher load, Fig. 9.31b, the DI combustion profile is slightly retarded relative to MBT timing and thus has a lower maximum pressure). The SI-VVA has a slower, more spread out, burn rate profile due to lower in-cylinder turbulence levels because it operates unthrottled and at lower peak pressure levels (there is less overall unburned mixture compression).⁴¹

The stratified DISI engine, however, shows a significantly different burning profile that is not symmetric about the maximum burn rate point. The lighter-load condition shows the greater asymmetry. Fuel-vapor/air mixtures burn fastest when close to stoichiometric. Very lean mixtures will not burn at normal engine temperatures and pressures; the laminar flame speeds are much too low (see Fig. 9.25). Thus stratified engine operation requires the creation of a close-enough to stoichiometric fuel-air-residual gas mixture "cloud" in which the spark plug is embedded at the crank angle when the spark discharge initiates combustion. This significantly richer-than-average fuel-air cloud is achieved through use of the appropriate fuel injection technology, injection timing, fuel spray development—atomization, vaporization, and mixing with air, between start of injection and ignition, aided by both intake- and compression-generated in-cylinder air motion and piston-crown and cylinder-head wall shaping (see Sec. 7.7.2). This close-to-stoichiometric "cloud" then burns rapidly as a "premixed" turbulent flame propagates through it. The flame slows down as it approaches the leaner outer regions of the cloud, and then extinguishes. The air (containing some fuel vapor) then mixes with the burned gases produced by combustion of the cloud, completing the burn up of rich burned regions within the cloud and of the fuel that "overmixed" with the excess air during the clouds development, to complete combustion and "homogenize" the burned gases during expansion. At lighter loads, with less fuel relative to the air, this rapid cloud premixed-flame-burning ends sooner, and the burn rate profile is more asymmetric as the figure indicates.

In summary, stratified-charge DI engine combustion occurs in two stages. Following spark ignition, a partially premixed flame propagates through the predominant mixture region where the fuel-air-burned gas (residual plus EGR) distribution has a high laminar burning velocity. This is followed by the still-rapid burnout of locally rich regions behind the flame front as oxygen from leaner regions and some of the surrounding air mix into the hot post-flame gases. Then a slower mixing-controlled combustion occurs of large fuel-rich pockets, such as that resulting from the last injected fuel, which enters the combustion chamber at lower velocity with poorer atomization, and thus slower fuel-air mixing. At lighter loads, the flame may never fill the combustion chamber. As a consequence, at lighter loads, the engine-out hydrocarbon emissions rise significantly.^{40, 42}

The CAI-controlled auto-ignition gasoline engine combustion concept—also shown in [Fig. 9.31](#)—is different from the standard spark-ignited premixed flame propagation process, and the diesel engine's diffusion flame ([Chap. 10](#)). A premixed lean or dilute fuel-vapor/air/residual mixture is created in-cylinder through early fuel injection with enough high-temperature trapped residual for the mixture temperature when combustion commences to be significantly higher than in a standard spark-ignited engine. At the appropriate time in the engine cycle, spontaneous ignition and combustion occur in an essentially volume burning (not flame propagation) manner. The figure shows a rapid symmetric burn rate profile (which limits this CAI combustion process to lighter engine loads). This controlled autoignited combustion mode is discussed in [Sec. 10.7](#).

[1] The water in actual residual and exhaust gas was omitted, and the CO₂ level augmented to compensate. A mixture of 80% N₂ and 20% CO₂, by volume, was used.

[2] Controlled autoignition combustion is discussed in [Sec. 10.7](#).

9.4. CYCLIC VARIATIONS IN COMBUSTION, PARTIAL BURNING, AND MISFIRE

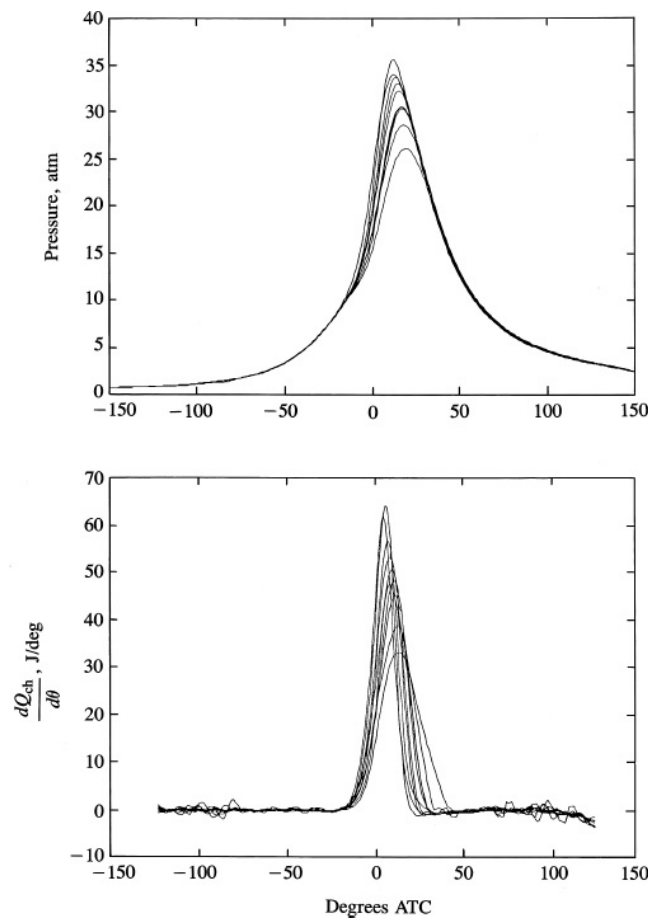
9.4.1. Observations and Definitions

Observation of cylinder pressure versus time measurements in a spark-ignition engine, for successive operating cycles, shows that substantial variations on a cycle-by-cycle basis exist. Since the pressure development is uniquely related to the combustion process, substantial variations in the combustion process on a cycle-by-cycle basis are occurring. In addition to these variations in each individual cylinder, there can be significant differences in the combustion process and pressure development between the cylinders in a multi-cylinder engine. Cyclic variations in the combustion process are caused by variations in mixture motion within the cylinder at the time of spark, cycle-by-cycle, variations in the amounts of air and fuel fed to the cylinder each cycle, and variations in the mixing of fresh mixture and residual gases within the cylinder each cycle which are especially important in the vicinity of the spark plug. Variations between cylinders are caused by differences in these same phenomena, cylinder-to-cylinder.

Cycle-by-cycle variations in the combustion process are important for two reasons. First, since the optimum spark timing is set for the "average" cycle, faster-than-average cycles effectively have over-advanced spark-timing and slower-than-average cycles effectively have retarded timing, so reductions in torque and efficiency result. Second, it is the extremes of the cyclic variations that limit engine operation. At high load, the fastest burning cycles with their over-advanced spark timing are most likely to knock. Thus, the fastest burning cycles determine the engine's fuel octane requirement and limit its compression ratio (see [Sec. 9.6.3](#)). The slowest burning cycles, which are effectively retarded relative to optimum timing, are most likely to burn incompletely. Thus these cycles set the practical lean operating limit of the engine or limit the amount of exhaust gas recycle (used for NO emissions control), which the engine will tolerate. Due to cycle-by-cycle variations, the spark timing and average air/fuel ratio must always be compromises, which are not necessarily the optimum for the average cylinder combustion process. Also, variations in cylinder pressure have been shown to correlate with variations in brake torque, which directly relate to vehicle driveability.

An example of the cycle-by-cycle variations in cylinder pressure, and the variations in mixture burning rate that cause them, are shown in [Fig. 9.32](#). Pressure, and gross heat-release rate calculated from the cylinder pressure using Eq. (9.27), for several successive cycles at a mid-load mid-speed point are shown as a function of crank angle. The maximum heat-release rate and the duration of the heat release or burning process vary by a factor of two from the slowest to the fastest burning cycle shown. The peak cylinder pressure varies accordingly. The faster burning cycles have substantially higher values of maximum pressure than do the slower burning cycles; with the faster burning cycles, peak pressure occurs earlier, closer to TC.

Figure 9.32 Measured cylinder pressure and calculated gross heat-release rate for 10 cycles in single-cylinder spark-ignition engine operating at 1500 rev/min, $\phi = 1.0$, $p_{\text{inlet}} = 0.7$ atm, MBT timing = 25° BTC.¹²



The heat-release rate data in Fig. 9.32 show that there are cycle-by-cycle variations in the early stages of flame development (from zero to a few percent of the total heat release) and in the major portion of the combustion process—the rapid-burning phase—as indicated by the variations in the maximum burning rate.

As the mixture becomes leaner with excess air, or more dilute with a higher burned gas fraction from residual gases or exhaust gas recycle, the magnitude of cycle-by-cycle combustion variations increases. Eventually, some cycles become sufficiently slow burning that combustion is not completed by the time the exhaust opens: a regime where *partial burning* occurs in a fraction of the cycles is encountered. For even leaner or more dilute mixtures, the *misfire* limit is reached. At this point, the mixture in a fraction of the cycles fails to ignite. While spark-ignition engines will continue to operate with a small percentage of the cycles in the partial-burn or misfire regimes, such operation is obviously undesirable from the perspective of efficiency, hydrocarbon emissions, torque variations, and roughness. The partial-burn and misfire regimes are discussed in Sec. 9.4.3.

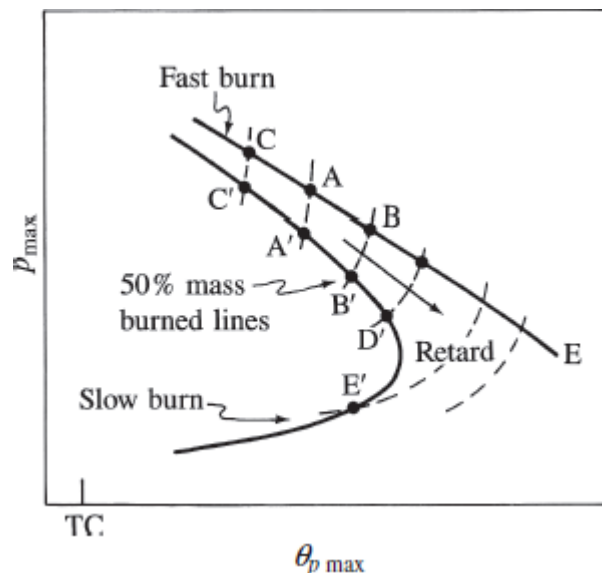
Various measures of cycle-by-cycle combustion variability are used. It can be defined in terms of variations in the cylinder pressure between different cycles, or in terms of variations in the details of the burning process, which cause the differences in pressure. The following quantities have been used:

- *Pressure-related parameters.* The maximum cylinder pressure p_{max} ; the crank angle at which this maximum pressure occurs $\theta_{p \text{ max}}$; the maximum rate of pressure rise $(dp/d\theta)_{\text{max}}$; the crank angle at which $(dp/d\theta)_{\text{max}}$ occurs; the indicated mean effective pressure [which equals $\int p \, dV/V_d$; see Eqs. (2.14), (2.15), and (2.19)].

- **Burn-rate-related parameters.** The maximum heat-release rate (net or gross; see Sec. 9.2.2); the maximum mass burning rate; the flame development angle $\Delta\theta_d$, and the rapid burning angle $\Delta\theta_b$; the crank angle at which 50% of the mixture is burned $\theta_{50\%}$ (see Sec. 9.2.3).
- **Flame front position parameters.** Flame radius, flame front area, enflamed or burned volume, all at given times; flame arrival time at given locations.

Pressure-related quantities are easiest to determine; however, the relation between variations in combustion rate and variations in cylinder pressure is complex.⁴³ Equation (9.26) defines the factors that govern this relationship. Because the rate of change of pressure is substantially affected by the rate of change of cylinder volume as well as rate of burning, changes in the phasing of the combustion process relative to TC (e.g., which result from changes in flame development angle), as well as changes in the shape and magnitude of the heat-release rate profile, affect the pressure. Figure 9.33 illustrates how the magnitude of the maximum cylinder pressure $p_{\max}$, and the crank angle at which it occurs $\theta_{p\max}$ vary as the crank angle at which combustion effectively starts (e.g., θ at which 1% of the cylinder mass has burned) and the burning rate is varied. Curve CABE shows how $p_{\max}$ and $\theta_{p\max}$ vary for a fixed fast-burning heat-release profile (the duration of the heat-release process and its maximum value are held constant), as the phasing of this combustion process relative to TC is varied. A corresponds to MBT timing where the start of combustion is phased to give maximum brake torque, B corresponds to retarded timing, and C to over-advanced timing. C'A'B'D'E' is a similar curve for a slow-burning heat-release profile. A' corresponds to MBT timing, and B' and D' to increasingly retarded timing. Note that with a sufficiently slow-burning heat-release profile—beyond D'— $\theta_{p\max}$ decreases as the burn process is increasingly retarded. This occurs when the rate of increase of pressure due to combustion becomes so low that it is more than offset after TC by the pressure decrease due to volume increase: eventually for extremely slow and late burning, the maximum pressure approaches the motored pressure at TC. The dashed lines show the constant start-of-combustion timing *relative* to MBT for each burn rate curve. Note that $\theta_{p\max}$ for constant relative timing varies little as the heat-release profile or burn rate varies.⁴³

Figure 9.33 Schematic of variation in maximum cylinder pressure and crank angle at which it occurs, in individual cycles. CABE typical of fast heat-release process; C'A'B'D'E' typical of slow heat-release process. (From Matekunas.⁴³)



We can now explain the effects of variations in the heat-release profile (both in the development stage of the burning process which effectively changes the location of the start of combustion, and in the rate of burning throughout the process) on $p_{\max}$ and $\theta_{p\max}$, when the spark timing occurs at a fixed crank angle. For a fixed burning rate profile (duration of burn and maximum burning rate), as the start of combustion is delayed to be closer to TC, $p_{\max}$ decreases and $\theta_{p\max}$ initially increases (A to B or A' to B'). This is the effect of a change in relative timing or phase of the burning process due to a slower initial rate of flame development with fixed spark timing. If, in addition to flame development being slower the heat-release rate throughout

the burning process is lower, then that combustion process is even more retarded from the optimum and $p_{\max}$ decreases further, and $\theta_{p \max}$ increases further, to their values at D'. The effect of a faster initial flame development and faster burning rate, with fixed spark timing, is the opposite. The magnitudes of the changes in $p_{\max}$ and $\theta_{p \max}$ depend, obviously, on the extent of the cyclic variations; they also depend on whether the average burn process is fast or slow. For fast-burning engines, a larger fraction of the heat release occurs near TC when the chamber volume is changing relatively slowly. Thus pressure variations are mainly due to combustion variations. With slow-burning engines, where a significant fraction of the energy release occurs well after TC, the effect of volume change also becomes significant and augments the effect of combustion variations. For large variations and a slow average burning process, $p_{\max}$ can fall below E', and $\theta_{p \max}$ then decreases. A fast-burning combustion process significantly reduces the magnitude of cycle-by-cycle combustion variations and thus their impact on engine performance.^{15, 44, 45}

We can now evaluate the various measures of combustion variability. The maximum pressure variation has been shown to depend on both changes in phasing and burning rate. The magnitude of this variation depends on whether the combustion chamber is faster or slower burning, on average. It also depends on whether the burning process is substantially retarded relative to MBT. It depends, too, on cyclic cylinder fuel and air charging variations. Thus the interpretation of variations in $p_{\max}$ [or in the maximum rate of pressure rise $(dp/d\theta)_{\max}$] in terms of variations in the rate and phasing of the burning process must be done with care. The location of maximum pressure $\theta_{p \max}$ also depends on relative phasing of combustion and on the burn rate profile. In addition, for slow-burning chambers and retarded timing (around E') variations produce little change in $\theta_{p \max}$. However, for fast-burning chambers, with MBT or only slightly retarded timing, the location of peak pressure depends essentially on the phasing of each combustion process relative to its MBT phasing, and is independent of charging variations. For these reasons, $\theta_{p \max}$ is a useful measure of variability in combustion event phasing.²⁵

One important measure of cyclic variability, derived from pressure data, is the *coefficient of variation in indicated mean effective pressure*. It is the standard deviation in imep divided by the mean imep, and is usually expressed in percent:

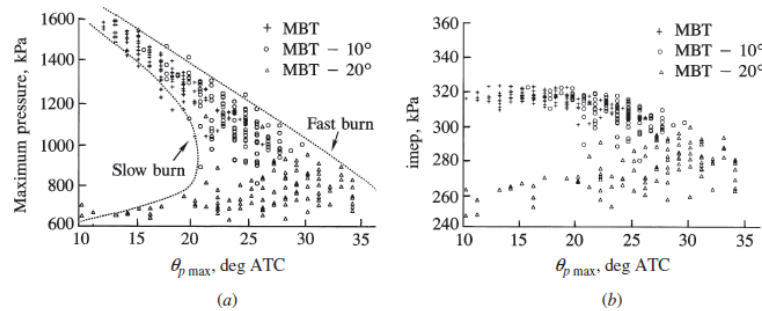
$$\text{COV}_{\text{imep}}(\%) = \frac{\sigma_{\text{imep}}}{\text{imep}} \times 100$$

(9.59)

It defines the cyclic variability in indicated work per cycle (on a gross or net basis), and it has been found that vehicle driveability problems usually result when COV_{imep} exceeds about 2 to 5%, depending on engine operating conditions.

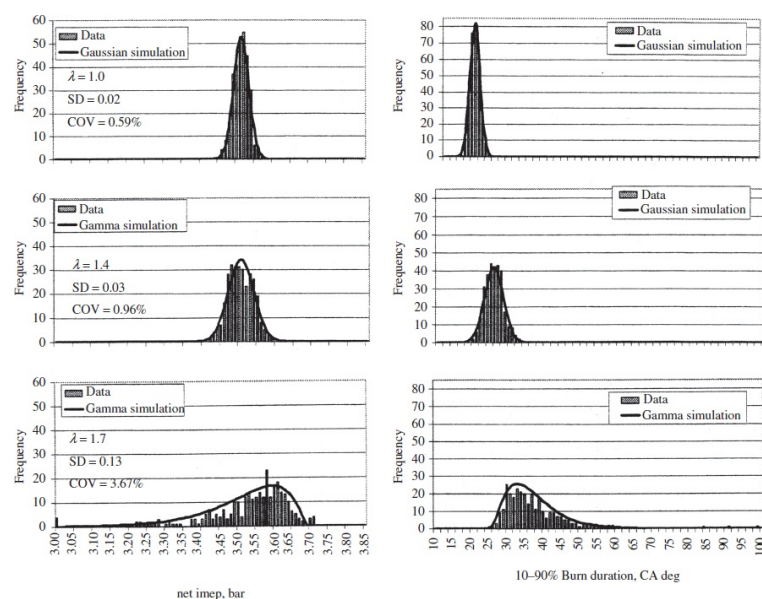
Figure 9.34 illustrates the relationships between $p_{\max}$, $\theta_{p \max}$, and imep for 120 cycles of an engine cylinder at fixed operating conditions and three different spark timings.⁴³ The MBT timing data show a spread in imep at a fixed value of $\theta_{p \max}$. This imep data band is relatively flat and is centered around $\theta_{p \max} \approx 15^\circ$; only at later values of $\theta_{p \max}$ does imep fall off significantly. The vertical spread in imep around $\theta_{p \max} = 15^\circ$ is due to variations in the amount of fuel entering the cylinder each cycle; normal variations in the burn profile under these conditions, which effectively change the phasing of the combustion process, produce only modest reductions in imep. For early $\theta_{p \max}$ (the extreme upper left of Fig. 9.34a), the variations in $p_{\max}$ are also due mainly to these fuel-charging variations, cycle-by-cycle: these are the fastest burning cycles with the most advanced phasing. As $\theta_{p \max}$ increases, the dispersion increases as cyclic variations in phasing and burning rate have increasing impact.

Figure 9.34 (a) Individual-cycle maximum pressure versus crank angle at which $p_{\max}$ occurs. (b) Individual-cycle indicated mean effective pressure versus $\theta_{p_{\max}}$.⁴³



The shapes of the frequency distributions in individual-cycle pressure data [e.g., in $p_{\max}$, $(dQ/d\theta)_{\max}$, $\theta_{p_{\max}}$] and in burn rate data such as $\Delta\theta_d$, $\Delta\theta_b$, $(dQ/d\theta)_{\max}$ depend on whether the combustion process is fast and "robust" (e.g., with close-to-stoichiometric mixtures at higher loads at optimum timing—well away from the lean operating limit of the engine) or slower and less repeatable, closer to the lean or dilute-mixture operating limit. Under robust combustion conditions these distributions are close to normal distributions.¹⁵ When the combustion process is much slower, the cyclic variability becomes large and the distribution becomes skewed toward the slower burning cycles which have low imep (due to the substantial retard of these slower cycles). Examples of the net imep and 10 to 90% burn angles, which illustrate these frequency distribution trends, are shown in Fig. 9.35. A standard gasoline engine was operated at a typical light load, 3.5 bar net imep at 1500 rev/min. With a stoichiometric mixture and fast burn, the COV_{imep} is small (0.6%) and the distributions in $imep_n$ and $\Delta\theta_{10-90}$ are narrow and symmetric. As the relative air/fuel ratio increases from 1.0 through 1.4 to 1.7, the burn rate in these increasingly leaner mixture decreases, and the spread in both $imep_n$ and $\Delta\theta_{10-90}$ becomes larger and, then asymmetric. COV_{imep} increases from its small value at stoichiometric (0.6%), to about 1.0%, to close to 4%. With the leanest mixture shown, some cycles become very slow burning and the net imep decreases significantly as a consequence. Slow-burning cycles are effectively retarded, and much of the mixture burning takes place in an ever expanding combustion chamber volume, which slows the burning process even more since unburned mixture pressures and temperatures are decreasing.^{15, 46}

Figure 9.35 Frequency distributions of net imep and 10 to 90% burn durations for different relative air/fuel ratios ($\lambda = 1.0, 1.4, 1.7$): MBT spark timing, engine operating at 1500 rev/min, net imep = 3.5 bar with indolene, at $r_c = 9.8$. SD, standard deviation; COV, coefficient of variation.¹⁷



Cylinder pressure data are often averaged over many cycles to obtain the mean cylinder pressure at each crank angle. The primary use of this average pressure versus crank-angle data is in calculating the average indicated mean effective pressure (which is a linear function of p). Since combustion parameters are not linearly related to the cylinder pressure [see Eq. (9.27)], analysis of the *average* pressure data will not necessarily yield accurate values of average combustion parameters. The error will be most significant when the combustion variability is largest. It is best to determine mean combustion parameters by averaging their values obtained from a substantial number of individual cycle analysis results. The number of cycles, which must be averaged to obtain the desired accuracy depends on the extent of the combustion variability. For example, while 40 to 100 cycles may define imep to within a few percent when combustion is highly repeatable, several hundred cycles of data may be required when cyclic combustion variations are large.¹⁰

9.4.2. Causes of Cycle-by-Cycle and Cylinder-to-Cylinder Variations

Cycle-by-cycle combustion variations are evident from the beginning of the combustion process. Analysis of flame photographs from many engine cycles taken in special research engines with windows in the combustion chamber has shown that dispersion in the fraction of the combustion chamber volume inflamed is present from the start of combustion (e.g., see Refs. 3, 29, and 37). Dispersion in burning rate is also evident throughout the combustion process (see Figs. 9.4 and 9.32). Three factors have been found to influence this dispersion:

1. The variation in gas motion in the cylinder during combustion, cycle-by-cycle;
2. The variation in the amounts of fuel, air, residual, and recycled exhaust gas supplied to a given cylinder each cycle;
3. Variations in mixture composition within the cylinder each cycle—especially near the spark plug—due to variations in mixing between air, fuel, residual gas and recycled exhaust gas.

The relative importance of these factors is not yet fully defined, and depends on engine design and operating variables. Table 9.4 provides a framework for discussing how the detailed components of these three overall factors lead to variations in the three stages of the SI engine combustion process: early flame development, flame propagation, and the final (late) stage of burning.

Table 9.4 Factors affecting SI engine combustion

Phases of combustion:
1. Early flame development
2. Flame propagation
3. Late stage of burning
Factors affecting SI engine cycle-to-cycle variations:
a. Spark energy deposition into gas (1)
b. Flame kernel motion (1)
c. Heat losses from kernel to spark plug (1)
d. Local turbulence characteristics near plug (1)
e. Local mixture composition near plug (1)
f. Overall charge components – air, fuel, residual, EGR (2, 3)
g. Average turbulence in the combustion chamber (2, 3)
h. Large scale features of the in-cylinder flow (3)
i. Flame geometry interaction with the combustion chamber walls (3)

The variation in the velocity field within the engine cylinder throughout the cycle, and from one cycle to the next, has been reviewed in [Sec. 8.2.2](#). Toward the end of the compression stroke, the ensemble-averaged rms velocity fluctuation is of comparable magnitude to the mean piston speed, and may be larger than the mean flow velocity if there is no strongly directed local mean flow pattern (see [Figs. 8.8](#) and [8.9](#)). This ensemble-averaged velocity fluctuation combines both cycle-by-cycle variation in the mean flow and the turbulent velocity fluctuations. During compression, these two components are of comparable magnitude (see [Fig. 8.9](#)). The available database indicates that substantial variations in the mean flow exist, cycle-by-cycle, both in the vicinity of the spark plug *and* throughout the combustion chamber. Velocity variations contribute to variations in the initial motion of the flame center as it grows from the kernel established by the spark, and in the initial growth rate of the flame: they can also affect the burning rate once the flame has developed to fill a substantial fraction of the combustion chamber. Variations in gas motion near the spark plug convect the flame in its early stages in different directions and at different velocities, cycle-by-cycle. This affects the flame's interaction with the cylinder walls, changing the flame area development with time. Variations in the turbulent velocity fluctuations near the spark plug will result in variations in the rate at which the small initially laminar-like flame kernel develops into a turbulent flame. Variations in the mean flow throughout the chamber will produce differences in flame front shape: also, they may produce differences in turbulence, which affect the propagation velocity of the front.

It is well known that the fuel, air, and recycled exhaust gas flows into each cylinder of a multi-cylinder engine are not identical. These flow rate differences are typically a few percent (see [Secs. 7.5](#) and [7.6](#)). It is also known that the flow patterns within the different cylinders are not necessarily identical due to differences between the individual intake manifold runner and port geometries in many production engines. All these factors contribute to cylinder-to-cylinder variations in the combustion process: there can be significant differences in the mean burn rate parameters as well as in the cyclic variations in these parameters. Also, the limited data available on the variation in mixture composition within each cylinder for each cycle

indicate that cyclic charging variations in individual cylinders are comparable in magnitude to cylinder-to-cylinder differences (i.e., a few percent). Also, the amount of residual gas left in the cylinder may vary somewhat, cycle-by-cycle. At higher loads, where the combustion process is more repeatable (and always completed relatively early in the expansion stroke), and the residual gas fraction is smaller (see [Sec. 6.4](#)), variations in the total amount of residual are not expected to be significant. At light loads (particularly at idle), where combustion variability is much higher, variations in the residual gas mass and its composition may become important.

In addition, the in-cylinder mixing of fuel, air, recycled exhaust, and residual is not complete: nonuniformities in composition exist within the cylinder at the start of combustion. Composition variations, cycle-by-cycle, in the vicinity of the spark plug electrode gap will affect the early stages of flame development, especially as the flame grows through the laminar-like burning phase following the creation of a small flame kernel by the spark discharge (see [Sec. 9.5.1](#)). The standard deviation of the cycle-by-cycle air/fuel ratio fluctuations can be a few percent of the mean (A/F). Significant fluctuations, cycle-by-cycle, in the mixing of fresh mixture with the in-cylinder residual gas have been noted. Nonuniformities in EGR distribution between cylinders and in EGR mixing within the cylinder can also occur and increase the variations in burned gas fraction locally at the spark gap, cycle-by-cycle.⁴⁸

Studies in a multi-cylinder production SI engine, where the fuel/air ratio nonuniformities and the nonuniform mixing of fresh mixture with residual gas were removed in turn as contributors to cycle-by-cycle variations (by comparing premixed propane operation with carbureted operation with gasoline, and by removing residual gas by purging with nonfiring cycles), indicate that the three contributing factors to cyclic combustion variations—velocity variations, fuel/air ratio variations, and residual gas mixing variations—are of comparable importance under light-load operating conditions.

This explanation of the underlying causes for cycle-by-cycle variations is consistent with the description of the turbulent flame propagation process in [Sec. 9.3.4](#). Conditions in the *vicinity* of the spark plug will influence the initial stages of the flame propagation process—establishing a stable kernel and its development into a turbulent flame. During the developed flame propagation phase, the *average* conditions in the bulk gas within the combustion chamber will be the determining factors since the flame front spans the chamber, effectively averaging out local nonuniformities. By conditions are meant the turbulent velocity fluctuations and length scales in the flow, proportions of fuel, air, and burned gas in the mixture, and the mixture state.

The flame development angle $\Delta\theta_d$ (the time to burn a few large eddies and establish a developed turbulent flame) can be expressed as¹⁶

$$\Delta\theta_d = C \left(\frac{l_i}{u'} \right)^{1/3} \left(\frac{l_M}{S_L} \right)^{2/3}$$

(9.60)

The turbulent flow field influences $\Delta\theta_d$ through l_i , u' , and l_M , the integral scale, the turbulence intensity, and the microscale, respectively. The mixture composition influences $\Delta\theta_d$ through the laminar flame speed S_L . There is, therefore, a variability in the flame development period, since all of these quantities can vary in the vicinity of the spark plug on a cycle-by-cycle basis.

The pressure development during the rapid-burning developed turbulent flame propagation phase, when the flame spans the combustion chamber, depends on the average rate of burning in the flame. Thus, variations in the turbulent flow field and mixture composition across the gas entering the flame front are averaged out and are not important. However, variation in chamber-average quantities are significant. The following expression for the maximum burning rate:

$$\left(\frac{dm_b}{dt} \right)_{\max} = \frac{C m_f (h^*/B) (\rho_u^*/\rho_i)^{10/9} [(u'^* S_L^*)/h_i]}{\nu^{*1/3}}$$

(9.61)

Here, m_f is the mass of fuel in the chamber, h is the instantaneous (mean) clearance height, B is the bore, ρ the density, u' the average turbulence intensity across the flame front, and ν the kinematic viscosity: the $*$ denotes the value at the time of the maximum burning rate and the subscript i the value at spark; C is a constant depending on engine design. This equation indicates that cycle-by-cycle variations in the maximum burning rate can result from variations in the overall flow pattern within the combustion chamber (which vary u^*) and from variations in the amount of fuel (m_f) that enters the cylinder each cycle.^[1] Also, it can be seen that variations in the flame development process will result in variations in the maximum burning rate because the crank angle at which the maximum burning rate occurs is shifted, and all the starred parameters in Eq. (9.59) will have different values.

From the discussion of flame development and structure in Sec. 9.3, and Eqs. (9.58) and (9.59) above, we would expect that mixture conditions and motion leading to slower flame development rates (longer flame development angles, $\Delta\theta_d$)—lower turbulence intensities and more dilute mixtures—would also give lower burning rates (longer rapid burning angles, $\Delta\theta_b$). Data from many different engines and a wide range of operating cases show that this is the case. The standard deviation of $\Delta\theta_b$ for a given chamber and operating condition usually increases as the average burning process becomes slower.

One final factor of importance is how variations in flame development and burning rate affect engine torque. With fixed spark timing, such variations in the combustion process cycle-by-cycle result in slower flame developing and burning cycles being retarded and faster developing and burning cycles being over-advanced. The curve of torque versus combustion timing (relative to optimum timing, Fig. 9.16) is almost independent of the burning rate; that is, a given magnitude retard (of say 10° CA) relative to optimum timing gives almost the same reduction in torque for a fast burn as it does for a slow burn. This is because the burning process, for optimum timing, is centered at about 8 to 10° ATC independent of the burn rate, and retard or advance shifts this "center" by equal amounts for all burn rates. One of the major advantages of fast-burning SI engines is now apparent. The magnitude of the variations in the flame development process and subsequent flame propagation rate are decreased as the burning rate is increased. Thus, these smaller combustion variations in fast-burn engines, which correspond to modest retard and over advance in non-average burn rate cycles, have little effect on torque. In contrast, the larger combustion variations of slow-burning engines result in more significant cyclic torque variations.

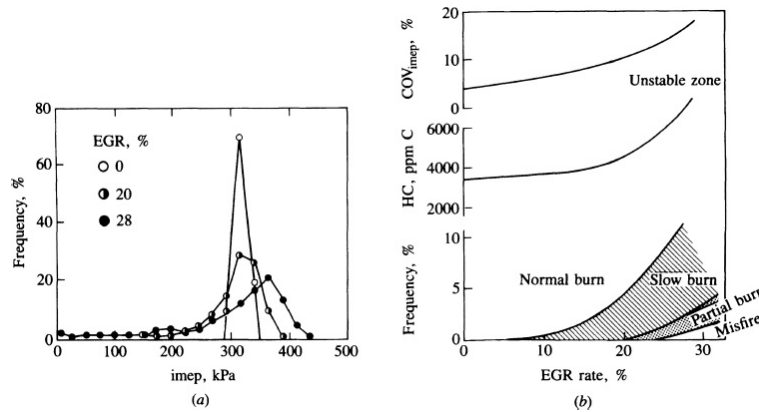
9.4.3. Partial Burning, Misfire, and Engine Stability

As the unburned mixture in a spark-ignition engine is leaned out with excess air or is diluted with increasing amounts of burned residual gas and exhaust gas recycle, the flame development period, the duration of the rapid burning phase, and the cycle-by-cycle fluctuations in the combustion process all increase. Eventually a point is reached where engine operation becomes rough and unstable, and hydrocarbon emissions increase rapidly. The point at which these phenomena occur effectively defines the engine's *stable operating limit*.^[2] These phenomena result from the lengthening of all stages of the combustion process as the unburned mixture is diluted. With increasing dilution, first a fraction of the cycles burns so slowly that combustion is only just completed prior to exhaust valve opening. Then as burning times lengthen further, in some cycles there is insufficient time to complete combustion within the cylinder: also, flame extinguishment before the exhaust valve opens and before the flame has propagated across the chamber may start to occur in some cycles. Finally, misfiring cycles where the mixture never ignites may start to occur. The proportion of partial burning or non-burning cycles increases rapidly if the mixture is made even more lean or dilute: the point is eventually reached where the engine will not run at all.

The impacts on engine stability of increasing combustion variability, due to increasingly lean (excess air) operation and/or higher amounts of exhaust gas recycle, at part-load, are shown in Figs. 9.35 and 9.36. Figure 9.36a shows the distributions of individual-cycle indicated mean effective pressure values for 0, 20, and 28% EGR. Without EGR at these conditions, the spread in imep is narrow. Increasing EGR widens the distribution significantly and cycles with low imep, and eventually zero imep, occur. Figure 9.36b shows how the COV of imep and hydrocarbon emissions increase as EGR is increased. Slow burn, then partial burn, and then misfire cycles occur with increasing frequency. In the slow-burn cycles, combustion is complete but

ends after 80° ATC and the indicated mean effective pressure is low (between 85 and 46% of the normal-burn value). Imep in partial-burn cycles was less than 46% of this value. In misfiring cycles, imep < 0.⁴⁸

Figure 9.36 (a) Frequency distributions in indicated mean effective pressure at different EGR rates for the engine studied; 0% EGR gave good engine stability, 10% acceptable stability, and above 20% poor stability. (b) Coefficient of variation in imep, HC emissions, and percentage of normal, slow, partial-burn, and misfire cycles. Engine conditions: 1400 rev/min, $\phi = 1.0$, MBT timing, imep = 324 kPa.⁴⁸



An explanation of the combustion phenomenon that occurs at an engine's stable operating limit has been developed by Quader.⁵⁰ It involves the following terms:

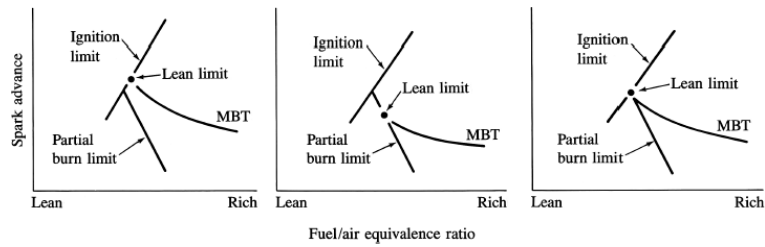
Ignition-limited spark timing or the **ignition limit**. The spark timing (advanced from MBT timing) at which *misfire* (i.e., failure of flame initiation) first occurs at a given mixture composition in a given small but arbitrary fraction of cycles (e.g., 0.5 to 1%).

Partial-burn-limited spark timing or the **partial-burn limit**. The spark timing (retarded from MBT) at which incomplete flame propagation occurs at a given mixture composition in a given small percent of the cycles (again, this frequency is selected arbitrarily for experimental convenience).

Lean or dilute misfire limit at MBT spark The leanest mixture stoichiometry at which the engine could operate stably at MBT spark timing with a misfire frequency below a specified value (again, this frequency, usually a percent or less of the cycles, is selected arbitrarily for convenience). A *dilute misfire limit*, the maximum amount of exhaust gas recycle that can be absorbed at a given stoichiometry for stable engine operation, can be similarly defined.

Engine experiments have defined the locations of the ignition limit line and the partial-burn limit line; they are shown schematically in Fig. 9.37. At a given spark timing, on this spark timing versus fuel/air equivalence ratio plot, progressive leaning of the mixture fed to the engine will lead to the onset of misfire or to the onset of partial burning, depending on the location of the lines and the spark timing selected. The individual figures show the possible interactions of the MBT timing line—leaner mixtures require greater advance—with the ignition limit and partial-burn limit lines. At MBT timing, the partial-burn limit may or may not be reached prior to misfire or the ignition limit. It will depend on the engine and ignition system design and operating conditions. For retarded spark timings relative to MBT, partial burning and not failure of flame initiation is the primary cause of unstable engine operation.

Figure 9.37 Schematics of three possible combinations of ignition limit, partial-burn limit, and MBT timing curves as a function of fuel/air equivalence ratio. (From Quader.^{49, 50})



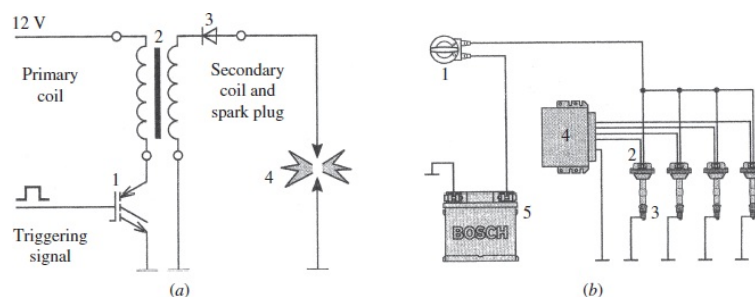
From the above it is clear that flame initiation is a necessary but not sufficient condition for complete combustion. Too slow flame development and propagation following successful ignition is usually the factor which limits engine operation with dilute mixtures. Experiments have shown that a limited crank-angle interval, of some 75 degrees (depending on engine geometry and spark plug location) is available during the engine cycle when conditions are favorable for the complete flame propagation process to be completed. Outside of this crank-angle interval, the mixture pressure and temperature, and the turbulence intensity, are too low to sustain a sufficiently rapid rate of combustion.

- [1] Variations in the total amount of air, recycled exhaust gas, and residual, in the chamber could also, for some operating regimes, be significant.
- [2] This limit has often been called the *lean operating limit*. Since what limits engine operation in practice is excessive torque fluctuations, cycle-by-cycle, and high hydrocarbon emissions, resulting from the use of mixtures made overly dilute with either air or burned gases (or with both), *stable operating limit* is a more appropriate term.

9.5. SPARK IGNITION

In spark-ignition engines, the electrical discharge produced between the spark plug electrodes by the ignition system starts the combustion process close to the end of the compression stroke. The high-temperature plasma kernel created by the spark develops into a self-sustaining and propagating flame front—a thin reaction sheet where the exothermic combustion chemical reactions occur. The function of the ignition system is to initiate this flame propagation process, in a repeatable manner, cycle-by-cycle, over the full load and speed range of the engine at the appropriate point in the engine cycle. Shadowgraph and schlieren photographs of the kernel created by the discharge between the plug electrodes, the growth of that kernel, and its transition to a propagating flame have already been presented in Figs. 9.24a and b, and described in the accompanying text. A spark will arc from one plug electrode to the other when a sufficiently high voltage is applied between them. Ignition systems commonly used to provide this spark are: battery ignition systems where the high voltage is obtained with an ignition coil (coil ignition systems); battery systems where the spark energy is stored in a capacitor and transferred as a high-voltage pulse to the spark plug by means of a special transformer, capacitive-discharge ignition (CDI) systems; and magneto ignition systems where the magneto—a rotating magnet or armature— generates the current used to produce a high-voltage pulse. The dominant system in automotive applications is the Transistorized Coil Ignition (TCI) system. The electrical energy is stored in the inductance of a coil and is released over about 2 ms when the spark discharge is required. Figure 9.38 shows the basic coil ignition system layout for a system with individual coil and plug arrangement. Before ignition is desired, the ignition driver stage switches a current from the electrical system through to the primary winding of the ignition coil. While the primary current circuit is closed (the dwell period) this current builds up a magnetic field in the coil. At the time of ignition, the primary winding current is interrupted producing a high voltage across the coil's secondary winding as the magnetic-field energy is discharged. This voltage (if high enough) causes the gas between the spark plug electrodes to break down and thus produces the spark discharge discharge.⁵¹

Figure 9.38 Example of ignition system with single spark coils. (a) Circuit layout: (1) ignition driver stage, (2) ignition coil, (3) activation arc diode (suppresses the activation spark), (4) spark plug. (b) System layout: (1) ignition lock, (2) ignition coils, (3) spark plugs, (4) control unit, (5) battery. (Courtesy Robert Bosch GmbH and SAE.⁵¹)



The next section reviews our basic understanding of electrical discharges in inflammable gas mixtures relevant to engine ignition (Sec. 9.5.1), the operating characteristics of conventional engine ignition systems (Sec. 9.5.2), and briefly, some alternative approaches to generating a propagating flame (Sec. 9.5.3).

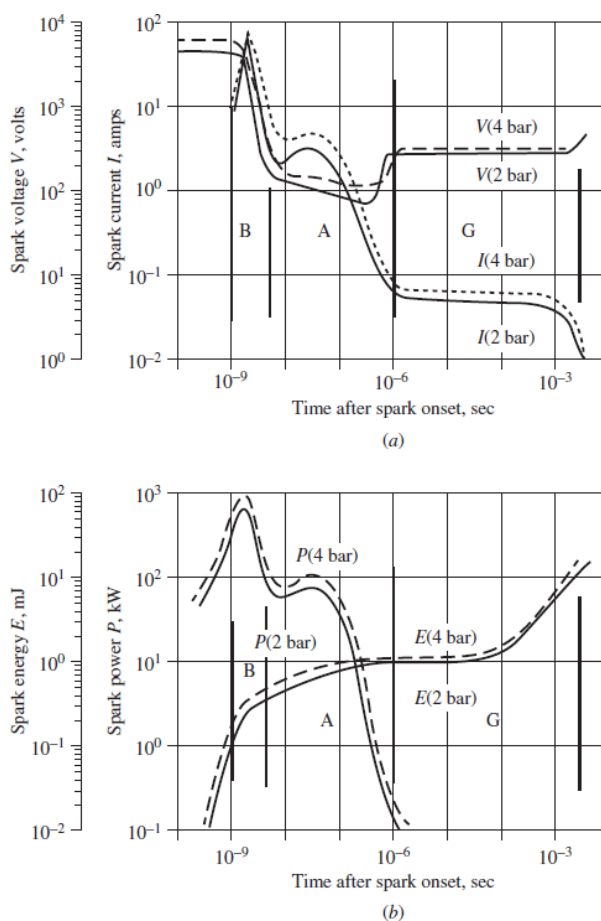
9.5.1. Ignition Fundamentals

A discharge can arc from one spark-plug electrode to the other only if a sufficiently high voltage is applied. In a typical spark discharge, the electrical potential across the electrode gap is increased until breakdown of the intervening mixture occurs. Electrons, accelerated across the gap by the high-applied electric field, rapidly excite and ionize gas molecules. These ionizing "streamers" then propagate from one electrode to the other. The impedance of the gap decreases drastically when a streamer

reaches the opposite electrode and the current through the gap increases rapidly. This stage of the discharge is called the *breakdown phase*. It is followed by the *arc phase*, where the thin cylindrical plasma arc expands largely due to heat conduction and diffusion and, with inflammable mixtures, initiates the exothermic reactions which lead to a propagating flame kernel (Fig. 9.24a). This is usually followed by a *glow discharge phase* where, depending on the details of the ignition system, the energy storage device (e.g., the ignition coil) will dump its energy into the discharge circuit.⁵²

Figure 9.39 shows the behavior of the spark-discharge voltage and current as a function of time for a TCI system. Typical values of discharge voltage V , current I , are shown; actual values depend on the details of the electrical components. The breakdown phase is characterized by a high-voltage (~ 10 kV), high-peak current (~ 200 A), and an extremely short duration (≤ 10 ns). A narrow (~ 40 μ m diameter) cylindrical ionized gas channel is established very early. The electrical energy supplied is transferred almost without loss to this plasma column. The temperature and pressure in the column rise very rapidly to values up to about 60,000 K and a few hundred atmospheres, respectively. A strong shock or blast wave propagates outward, the channel expands and, as a result, the plasma temperature and pressure fall. Some 30% of the plasma energy is carried away by the shock wave; however, most of this is regained since spherical blast waves transfer most of their energy to the gas within a small (~ 2 mm diameter) sphere into which the breakdown plasma soon expands.⁵²

Figure 9.39 (a) Voltage and current variation with time and (b) power and energy variation with time for a commercial transistorized coil spark-ignition system, in air at 300 K supplying 30 mJ to a 1-mm spark plug gap. The pressure of 4 bar roughly corresponds to engine conditions at spark onset.²³ Note the three phases of the discharge—breakdown (B), arc (A), and glow discharge (G).



A breakdown phase always precedes arc and glow discharges: it creates the electrically conductive path between the electrodes. The arc phase voltage is low (< 100 V), though the current can be as high as the external circuit permits. In contrast to the breakdown phase where the gas in the channel is fully dissociated and ionized, in the arc phase the degree of dissociation may still be high at the center of the discharge but the degree of ionization is much lower (about 1%). Voltage

drops at the cathode and anode electrodes are a significant fraction of the arc voltage, and the energy deposited in these electrode sheath regions, which is conducted away by the metal electrodes, is a substantial fraction of the total arc energy (see Table 9.4). The arc requires a hot cathode spot, so heating and evaporation of the cathode material occur. The arc increases in size due primarily to heat conduction and mass diffusion. Due to these energy transfers the gas temperature in the arc is limited to about 5000 K: the temperature and degree of dissociation decrease rapidly with increasing distance from the arc axis. Currents less than 200 mA, a large electrode voltage drop at the cathode (300 to 500 V), a cold cathode, and less than 0.01% ionization are typical for the glow discharge. Energy losses are higher than in the arc phase, and peak equilibrium gas temperatures are about 3000 K.⁵²

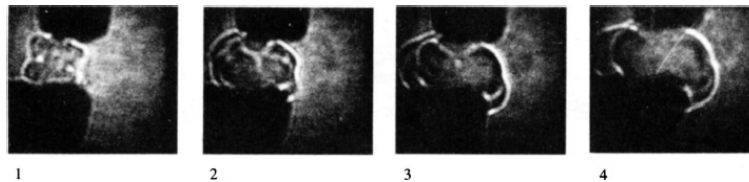
About 0.2 mJ of energy is required to ignite a quiescent stoichiometric fuel-air mixture at normal engine conditions by means of a spark. For substantially leaner and richer mixtures, and where the mixture flows past the electrodes, an order of magnitude greater energy may be required. TCI systems deliver from 30 to 100 mJ of electrical energy to the spark, with most of this in the glow discharge phase. Due to the physical characteristics of the discharge modes discussed above, only a fraction of the energy supplied to the spark gap is transmitted to the gas mixture. The energy balance for the breakdown, arc, and glow phases of the discharge is given in Table 9.5. Radiation losses are small throughout. The end of the breakdown phase occurs as a hot cathode spot develops, turning the discharge into an arc; heat losses to the electrodes then become substantial. The breakdown phase reaches the highest power level (~ 1 MW), but the energy supplied is small (<1 mJ). The glow discharge has the lowest power level (~ 10 W) but the highest energy (30 to 100 mJ), due to its long discharge time. The arc phase lies between (see Fig. 9.39).²³ The different energy transfer capabilities for breakdown, arc, and glow discharges arise primarily from the differences in heat losses to the electrodes, as explained above. These losses increase with increasing supplied energy. In arc and glow discharges, increases in either discharge time or discharge current (or in both) always lead to substantial decreases in energy-transfer efficiency. If the glow discharge current is increased above about 100 mA, the discharge changes to the arc mode and heavy electrode erosion will result. Thus there are practical limits to the arc and glow discharge currents; also, the time available for ignition in the engine limits increases in discharge time.

Table 9.5 Energy distribution for breakdown, arc, and glow discharges[†]

	Breakdown, %	Arc, %	Glow, %
Radiation loss	<1	5	<1
Heat loss to electrodes	5	45	70
Total losses	6	50	70
Plasma energy	94	50	30
[†] Typical values, under idealized conditions with small electrodes.			
Source: Maly and Vogel. ⁵²			

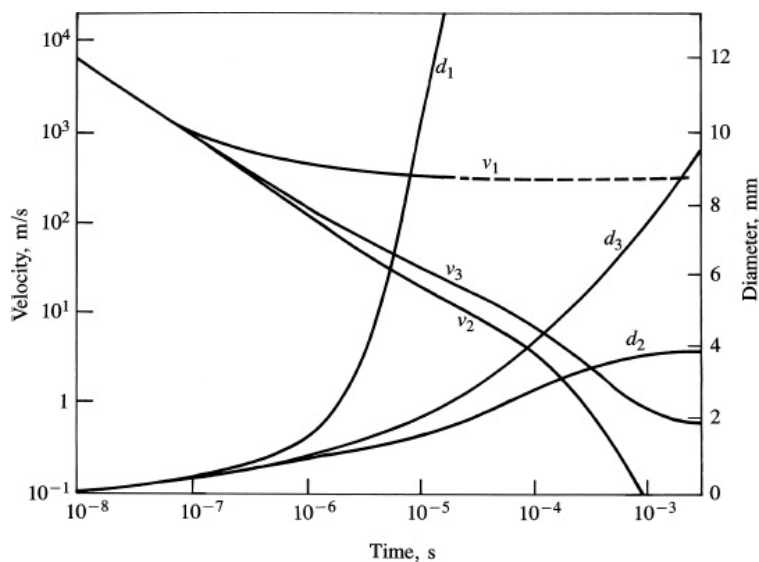
Figure 9.40 shows shadowgraph photos of the spark-generated flame kernel, which illustrate the context for this discussion of the engine spark discharge. The first photo (on the left) shows that by the time the glow discharge commences, the breakdown and arc dischargers have already initiated a 1- to 2-mm diameter laminar-like flame kernel. This then steadily grows in size as the flame propagates outward into the unburned mixture in the vicinity of the electrodes. The photos indicate that heat losses from hot flame kernel to the plug electrodes will be significant.

Figure 9.40 Shadowgraph photographs of spark-generated kernel. First photograph on left shows the kernel between the two spark-plug electrodes (dark in photo), about $35\ \mu\text{s}$ after breakdown; $40\ \mu\text{s}$ between photos. Stoichiometric mixture, 1100 rev/min. (Courtesy A. Douad, Institut Francais du Petrole.)



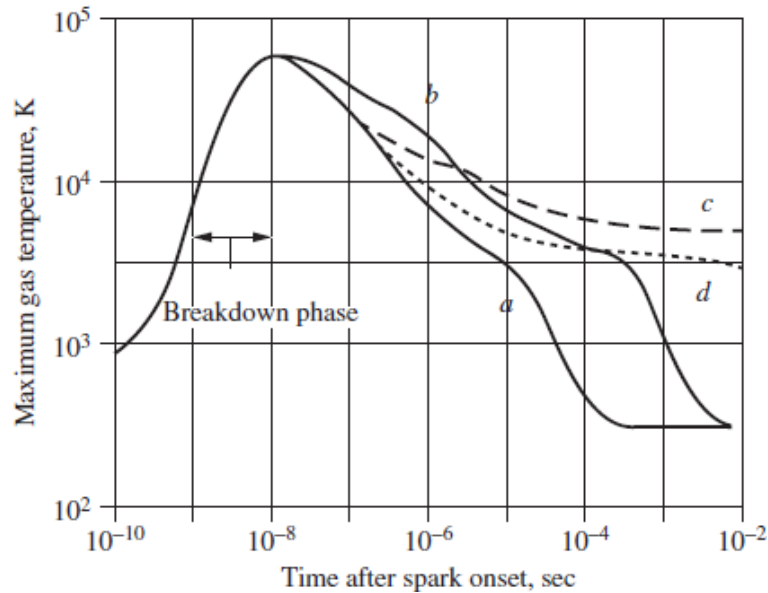
The initial expansion velocities in the discharge are much higher than those in self-propagating flames. **Figure 9.41** shows the expansion velocities and diameters of the volumes activated by a 3-mJ 100- μs discharge from a CDI system as a function of time after spark onset. The curves shown are (1) for the shock wave following a discharge in air at 1 atm; (2) the growth in plasma region in air at 1 atm; (3) the electrical and chemical plasma region following a discharge in a stoichiometric methane-air mixture at 1 atm. The initially strong shock wave attenuates rapidly to the local sound speed. Up to times of order microseconds, the effects of fuel combustion chemistry are small. The change in slope of the velocity curves for the plasma at about $100\ \mu\text{s}$ indicates the transition from an expansion caused by the initial high pressure in the breakdown discharge, to expansion resulting from heat conduction and diffusion.

Figure 9.41 Diameters (d) and expansion velocities (v) of volumes activated by a capacitive-discharge ignition system: 3 mJ electrical energy, 100 μs duration. Subscripts denote: 1, shock wave in air at 1 atm; 2, plasma in air at 1 atm; 3, electrical and chemical plasma in stoichiometric methane-air mixture at 1 atm.⁵²



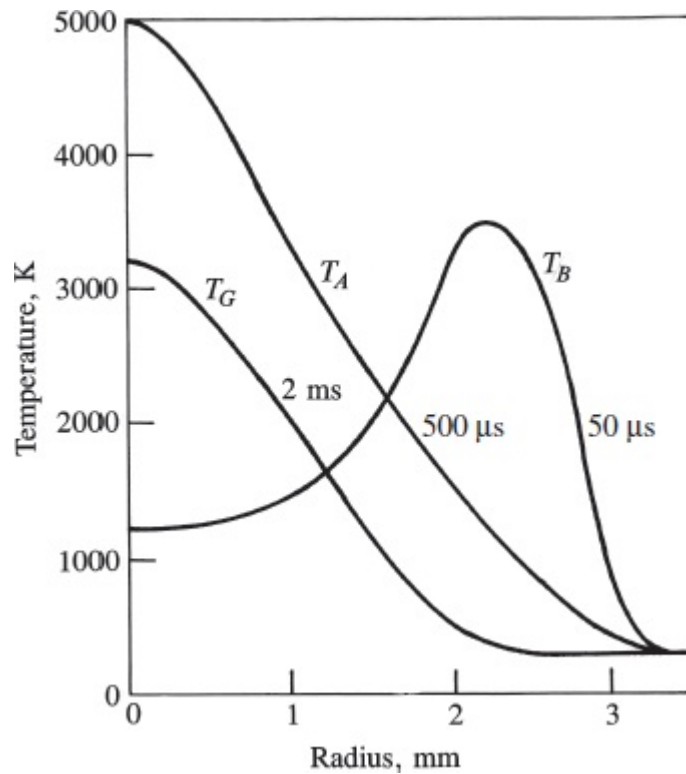
The temperature distributions within the three different types of discharge provide additional insight: see **Fig. 9.42**. During the breakdown discharge, on a time scale of 10 nanoseconds, the temperature rises to 60,000 K. Increasing the breakdown energy does not produce higher kernel temperatures; instead the channel diameter increases producing a larger plasma volume. The kernel temperatures then decrease to the order of 10,000 K on a microsecond time scale as the plasma region expands behind the shock wave. Arc and glow discharges, because their power inputs are much lower, do not increase the kernel temperatures; rather, they extend the cooling period on a microsecond and millisecond time scale, respectively.²³

Figure 9.42 Maximum measured gas temperatures as a function of time in different discharge modes and ignition systems: (a) CDI, 3 mJ, 100 μ s; (b) breakdown, 30 mJ, 60 ns; (c) CDI plus superposed current (2 A), arc, 30 mJ, 230 μ s; (d) CDI plus superposed current (60 mA), glow discharge, 30 mJ, 770 μ s.²³



The characters of the temperature profiles that each of these three types of discharge create in air, with essentially the same total electrical energy input (30 mJ), are different (see Fig. 9.43). The radial profiles in the mid-plane of the discharge are shown. The expansion-wave-induced expansion of the plasma behind the shock wave generated by the breakdown discharge produces a larger plasma, earlier, with a steep temperature front. Thus it creates favorable conditions for transferring heat and radicals to the surrounding unburned mixture. In addition cold gas flows into the central region of the plasma, due to the boundary layers which the rapidly expanding flow sets up on the electrodes, effectively insulating the hottest plasma region from the cold electrode surfaces.⁵² The arc and glow discharges, each preceded by a much lower energy initial breakdown process, show a significantly slower expansion rate and the more gradual temperature profile produced by heat conduction and diffusion.

Figure 9.43 Radial temperature profiles at selected times after spark onset for ignition systems with different electrical energies and discharge times in air at 1 atm, T_B , breakdown discharge, 30 mJ energy, 60 ns duration; T_A , capacitive discharge, 3 mJ energy, 100 μ s duration with superimposed current arc of 2 A, 30 mJ energy, for 230 μ s; T_G , capacitive-discharge system, 3 mJ energy, 100 μ s duration with superimposed constant-current glow discharge of 60 mA, 30 mJ energy, 770 μ s duration.⁵²



Chemical reactions can be observed spectroscopically a few nanoseconds after spark onset. They are initiated by the very high radical density in the breakdown plasma where all the heavy particles are present as highly excited atoms and ions. Since the kernel temperatures are much too high to allow the species present in normal combustion products to exist, combustion reactions take place at the outer plasma surface where the conditions are ideal for rapid chemical activity (temperatures of a few thousand kelvins). The chemical energy released is added to the plasma energy and becomes evident when the plasma velocity falls below about 100 m/s (see Fig. 9.41). At this point, the interior of the plasma still consists of a fully dissociated reacted gas mixture with much of its energy stored in radicals. An indication of the gas composition across the steep temperature profile at the plasma interface can be obtained from Table 9.6, which shows the equilibrium composition of C_8H_{18} -air combustion products over the relevant temperature range. Note, however, that the gas will not be in equilibrium. The different radicals have different diffusivities with the hydrogen atom some five times that of other species. Thus the H radical will diffuse furthest into the as-yet unreacted mixture. On the high-temperature side of the inflammation zone, the large number of radical particles transfers their energy to the mixture molecules within a few collisions. On the low-temperature side of the zone, above-equilibrium concentrations of combustion-initiating radicals (O, H, and OH) build up. In addition, a high heat flux into the region occurs, from the plasma core, by conduction down the steep temperature gradient. As a consequence of these conditions, reactions will occur and energy will be released more rapidly than in a normal flame.⁵²

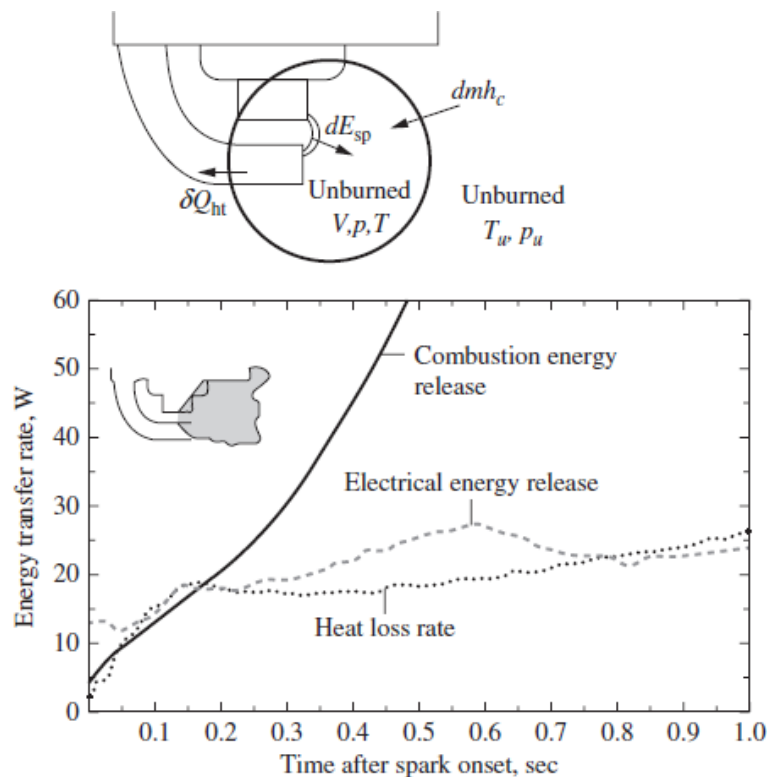
Table 9.6 Equilibrium composition of stoichiometric isooctane-air combustion products

Species	Temperature, K [†]			
	2000	3000	4000	5000
CO	2.4(-3)	6.1(-2)	9.4(-2)	9.0(-2)
CO ₂	1.2(-1)	5.7(-2)	5.0(-3)	3.5(-4)
H	2(-5)	9.7(-3)	1.2(-1)	1.9(-1)
H ₂	6.1(-4)	1.5(-2)	2.4(-2)	3.5(-3)
H ₂ O	1.4(-1)	1.0(-1)	1.1(-2)	1.2(-4)
N	<i>n</i>	1(-5)	6.6(-4)	1.2(-2)
NH	<i>n</i>	<i>n</i>	2(-5)	6(-5)
NO	5.7(-4)	1.5(-2)	2.9(-2)	1.7(-2)
N ₂	7.3(-1)	6.9(-1)	5.7(-1)	5.1(-1)
O	1(-5)	8.6(-3)	9.8(-2)	1.6(-1)
OH	4.5(-4)	2.2(-2)	3.3(-2)	5.6(-3)
O ₂	1.1(-3)	2.3(-2)	1.7(-2)	2.2(-3)
[†] At 4 atm pressure: mole fractions, 9.0(-2) = 9.0×10^{-2} , <i>n</i> = $<5 \times 10^{-6}$.				

The characteristics of the breakdown phase of the discharge have the greatest impact on inflammation. The size of the activated volume at a given time interval after spark initiation, the temperature difference across the kernel outer boundary, and the velocity of the boundary are all substantially increased by increasing the breakdown phase energy. Additional energy input during the arc and glow discharge phases has a more modest effect on these critical kernel properties.

These energy balances for the spark-generated flame kernel growth process, over the 10 μ s to fraction of ms time frame are illustrated in Fig. 9.44. These are experimentally based results, and the various energy fluxes are defined in the flame kernel, spark plug, and gap schematic. The flame outline is a schlieren photo taken some 0.8 ms after spark onset, well into the glow-discharge phase. The electrical energy released (and dissipated) in the flame kernel during this period is about 20 W. The heat loss rate to the plug electrodes, which depends on the portion of the flame surface "wetted" by the plug surfaces (which varies cycle-by-cycle), is of comparable magnitude (and opposite in sign). The combustion energy release rate—as the flame propagates outward—is initially comparable to the electrical and heat loss energy transfer rates. As the flame grows, the rate of combustion energy release within the kernel then outpaces the electrical and heat loss components. We would expect (over the first 0.2 ms) cyclic variations in the early flame kernel growth to be significant.

Figure 9.44 Energy flows (in W) for spark-ignited flame kernel: electrical energy dissipation rate, heat loss rate to spark plug electrodes and insulator, fuel chemical energy release rate, as a function of time after spark onset. Spark timing 65° BTC, relative A/F 1.25.²⁸



Models of the plasma unburned-mixture interface have been developed in attempts to quantify the complex flame initiation and propagation phenomena described above (e.g., Refs. 23, 53, 54, and 55). They are based on the requirement that the energy release due to chemical reaction in the plasma front (and production of radical species) exceed the losses due to conduction and diffusion to the unburned gas ahead of the front. Such models provide a theoretical basis for the well-known fact that the minimum energy required to ignite a premixed fuel-air mixture depends strongly on mixture composition and state. The ignition energy has a minimum for slightly rich-of-stoichiometric mixtures; the minimum energy required for successful ignition increases rapidly as the mixture is leaned out. While the initial plasma kernel growth (up to 10 to 100 μs) is not greatly affected by the mixture equivalence ratio, the inflammation process and the thickness and rate of propagation of the resulting flame are strongly affected. The lean side of the minimum (and the effect of EGR dilution) is of more practical interest. Because the chemical energy density of the mixture and flame temperature decrease as the mixture is leaned out or diluted, the flame speed decreases and the flame becomes thicker. Thus more time is available for heat losses from the inflammation zone, less energy is available to offset these losses, and the rate of energy transfer into the zone decreases. The consequence is that, as the mixture is leaned out, the approximately spherical discharge-created plasma must grow to a larger size before inflammation will occur. Substantially more energy must therefore be supplied to the discharge.⁵²

In engines, the mixture is not quiescent: mean and fluctuating velocities in the range 1 to 10 m/s exist in the clearance volume at TC (see Sec. 8.2.2). On the time scale of the breakdown discharge phase (10 ns), this fluid motion is not important. In the arc and glow discharge phases, however, the arc may be convected by the flow out of the plug electrode gap, and lengthened accordingly. For flow velocities below about 15 m/s, a steady increase in discharge channel length occurs. For higher velocities, an increasing number of re-ignitions occur, so the discharge energy is distributed into many separate channels. As the channel lengthens, the ratio of total discharge voltage to anode plus cathode voltage drop increases substantially, and the relative importance of heat losses to the electrodes decreases. Thus more energy is transferred to the gas—the energy-transfer efficiency increases. However, as the channel is lengthened, the energy transferred is spread over a larger volume. Depending on flow velocity, and mixture and discharge conditions, increasing velocity may increase or decrease the minimum ignition energy or the lean ignition limit for a specific ignition system. Both the mean flow velocity and turbulence levels are

important. With a TCI system, increasing mean flow velocity up to the point where re-ignitions start to occur extends the lean limit. With breakdown discharge systems, the lean ignition limit decreases as flow velocity increases. With capacitive-discharge systems at low flow velocities, the flow has little impact: at high flow velocities the minimum ignition energy increases.

Maly and Vogel⁵² have summarized the fundamental aspects of spark-discharge ignited flames as follows:

- Of the total electrical energy supplied to the spark, only that fraction contained within the outer surface layer of the plasma (of thickness of the order of the inflammation zone) is available for initiating the flame propagation process. The energy density and the temperature gradient in this layer depend on the discharge mode. Highest energy densities and temperature gradients are achieved if the ignition energy is supplied in the shortest time interval.
- A minimum radius of the spark-generated plasma is required for inflammation of the fuel-air mixture to occur. This radius increases rapidly as the mixture is leaned out (or diluted); it decreases with increasing pressure and increasing plasma expansion velocity.
- After inflammation, burning rates are proportional to flame surface area. Thus discharges and plasma geometries that produce the largest inflammation zone surface area, most rapidly, are advantageous.
- The time over which the ignition energy can be used effectively for inflammation decreases as the initial flame velocity increases. Ignition energy supplied after inflammation has occurred will have only a modest impact on flame propagation.

9.5.2. Standard Ignition Systems

The ignition system must provide sufficient voltage across the spark plug electrodes to initiate the discharge and supply enough energy to the discharge to ignite the combustible mixture adjacent to the plug electrodes under all operating conditions. It must create this spark at the appropriate time in each cylinder's operating cycle (normally toward the end of the compression stroke). Usually spark timing is set to give maximum brake torque for that specific operating condition, though the largest torque available may be constrained by maximum pressure limits, or knock control requirements. For a given engine design, this optimum spark timing varies as engine speed, inlet manifold pressure, and mixture composition vary. Thus, in most applications, and especially automotive applications, the system must have means for automatically changing the spark timing as engine speed and load vary.

With an equivalence ratio best suited for ignition (close-to-stoichiometric) and with homogeneous mixture distribution, spark energies of order 1 mJ and durations of a few microseconds would suffice to initiate the combustion process. In practice, circumstances are less ideal. The air, fuel, and recycled exhaust are not uniformly distributed between cylinders; the mixture of air, fuel, recycled exhaust gas, and residual gas within each cylinder is not homogeneous. Also, the pressure, temperature, and density of the mixture between the spark plug electrodes at the time the spark is needed affect the voltage required to produce a spark. These vary significantly over the load and speed range of an engine. The spark energy and duration, therefore, have to be sufficient to initiate combustion under the most unfavorable conditions expected in the vicinity of the spark plug over the complete engine operating range. Usually, if the spark energy exceeds about 50 mJ and the duration is longer than 0.5 ms, reliable ignition is obtained.

In addition to the spark requirements determined by mixture quality, pressure, temperature, and density, there are others determined by the state of the spark plugs. The erosion of the plug electrodes over extended mileage increases the gap and requires a higher breakdown voltage. Also, spark plug fouling due to deposit buildup on the spark plug insulator can result in side-tracking of the spark. When compounds formed by the burning of fuel, lubricating oil, and their additives are deposited on the spark plug insulator, these deposits provide an alternative path for the spark current. If the resistance of the spark plug deposits is sufficiently low, the loss of electrical energy through the deposits may prevent the voltage from rising to that required to break down the gas. The influence of side-tracking on spark generation decreases with lower source impedance of the high-voltage supply, and therefore with a higher available energy.

The fundamental requirements of the high-voltage ignition source can be summarized as: (1) a high ignition voltage to break down the gap between the plug electrodes; (2) a low source impedance to achieve a rapid voltage rise; (3) a high-energy storage capacity to create a spark kernel of sufficient size; and (4) sufficient duration of the discharge to ensure ignition. There are several commonly used concepts that partly or fully satisfy these requirements.⁵¹

9.5.2.1. Coil Ignition Systems

Inductive coil ignition systems have been used in spark-ignition engines for many years. TCI systems with their advantageous performance characteristics and low cost dominate. Figure 9.38 shows the general circuit arrangement of a coil ignition system. The system includes an ignition driver stage, an ignition coil with a primary and secondary winding, a spark plug connected to the high-voltage connection point of the wiring. The circuit functions as follows. Before the spark is needed, the ignition driver stage switches a current from the electrical system through the primary winding of the coil: this builds up a magnetic field in the ignition coil's iron core. At time of ignition, the current through the primary winding is interrupted. The decay of magnetic flux in the coil induces a voltage in both the primary and secondary windings. The voltage induced in the secondary winding is routed to the spark plug to produce the ignition discharge.

Typical current and voltage waveforms are shown in Fig. 9.39. The primary current for any given dwell time t is given by

$$I_p = \frac{V_0}{R}(1 - e^{-Rt/L_p})$$

(9.62)

where I_p is the primary current, V_0 is the supply voltage, R is the total primary circuit resistance, and L_p is the primary circuit inductance. The primary current requires time to build up. At low speeds the time of contact closure is sufficient for the primary current to reach the maximum permitted by the circuit resistance; at high speeds the primary current may not reach its maximum. Thus, only at higher engine speeds does the term e^{-Rt/L_p} becomes significant. When the primary current falls to zero, a voltage of order 15 kV is induced in the secondary winding. The peak value of this voltage is the maximum voltage that can be produced by the system and is called the *available voltage* V_a . The maximum energy transferred to the secondary system is given by

$$E_{s,\max} = \frac{1}{2}C_s V_a^2$$

where C_s is the total capacitance of the secondary circuit. Hence, the available voltage of the system is given by

$$V_a = \left(\frac{2E_{s,\max}}{C_s} \right)^{1/2}$$

(9.63)

If all the energy stored in the primary circuit of the coil, $\frac{1}{2}L_p I_p^2$, is transferred to the secondary,

$$V_a = I_p \left(\frac{L_p}{C_s} \right)^{1/2}$$

(9.64)

When the coil is connected to a spark plug, the secondary voltage will rise to the breakdown potential of the spark plug, and a discharge between the plug electrodes occurs. During this first breakdown phase, as explained earlier, on the 10 nanosecond timescale a current of some 10 A flows as the voltage reduces to a lower value (some 100 V) over about 1 μ s as the energy is dissipated in the arc phase. The value of the voltage which caused breakdown to occur is called the *required voltage* of the spark plug. The interval during which the discharge occurs is called the *spark duration*. The available voltage of the ignition system must always exceed the required voltage of the spark plug to ensure breakdown. The spark must then possess sufficient energy and duration to initiate combustion under all conditions of operation.

In automotive applications, the need for reduced ignition system maintenance, extended spark plug life, improved ignition of lean and dilute mixtures, and increased reliability has led to the widespread use of transistorized coil ignition systems which provide a higher output voltage and which use electronic triggering to realize the required timing without wear or adjustment. The higher output voltage is required because spark plugs are now set to wider gaps (about 1 mm) to extend the ability to ignite the fuel/air burned gas mixture over a wider range of engine operation, and because during the extended mileage between spark plug replacement, electrode erosion further increases the gap. In automotive applications an available ignition voltage of some 30 kV is now usually provided. In addition to higher voltage, longer spark duration (about 2 ms) has been found to extend the engine operating conditions over which satisfactory ignition is achieved.

9.5.2.2. Capacitive-Discharge Ignition (CDI) Systems

With this type of system, a capacitor rather than an induction coil is used to store the ignition energy. The capacitance and charging voltage of the capacitor determine the amount of stored energy. The ignition transformer steps up the primary voltage, generated at the time of spark by the discharge of the capacitor through a thyristor, to the high voltage required at the spark plug. The CDI trigger box contains the capacitor, thyristor power switch, charging device (to convert battery voltage to the charging voltage of 300 to 500 V by means of pulses via a voltage transformer), pulse shaping unit, and control unit. The principal advantage of CDI is its insensitivity to electrical shunts in the high-voltage ignition circuit that result from spark plug fouling. Because of the fast capacitive discharge, the spark is strong but short (0.1 to 0.3 ms). This can lead to ignition failure at operating conditions where the mixture is very lean or dilute.

9.5.2.3. Magneto Ignition System

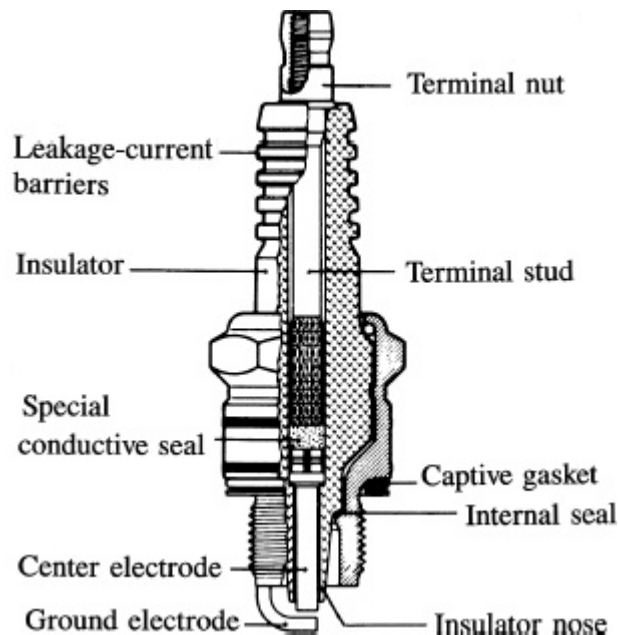
With this type of ignition system, a magneto supplies the ignition voltage for the spark discharge independent of a battery or generator. Magneto ignition is used in small-displacement four-stroke and two-stroke engines. The system operates as follows. A time-varying magnetic flux is set up in the ignition armature as the rotating permanent magnets on the pole wheel generate a current in the closed primary winding. To generate the ignition voltage, the primary current flow is interrupted and the flux collapses rapidly producing a high-voltage pulse in the winding which is connected to the spark plug electrode. The current can be interrupted with contact breakers or with a transistor. Since the flux generated by the rotating pole wheel depends on engine speed, the magnitude of the ignition voltage varies with speed.

9.5.2.4. Spark Plug Design

The function of the spark plug is to provide an electrode gap across which the high-voltage discharge occurs that ignites the compressed mixture of fuel vapor, air, and burned (residual) gas in the combustion chamber. In addition, it must provide a gas-tight conducting path from the high-voltage source to the electrode gap. [Figure 9.45](#) shows a typical spark plug design. There are three principal components: electrodes, an insulator, and a shell. The insulating material must have adequate thermal shock resistance, tensile and compressive strength, and impact strength. It must also have low porosity to limit absorption of combustion gases and high resistivity to prevent leakage of high-voltage charge at both ambient and normal

operating temperatures. Alumina is usually used as the insulator material. The electrodes are normally made of high-nickel alloys to withstand the high ignition voltage, high temperatures, and corrosive gases with minimum erosion. In the standard J-plug shown, the center-electrode surface temperatures can reach about 800°C under normal operating conditions. The ground electrode temperature can reach about 1000°C.⁵⁷

Figure 9.45 Cutaway drawing of standard J-type spark plug. (Courtesy Robert Bosch GmbH and SAE.⁵⁶)



Several aspects of the J-plug electrodes impact its performance in an operating engine. Performance is often characterized as the lean (λ) or dilute (EGR percentage) limit at which the standard deviation (σ) or coefficient of variance (COV) in imep reaches a significant level—e.g., a COV_{imep} of 5% or more, a higher value than is tolerable in a production engine. The orientation of the ground electrode (in relation to the mean flow direction at the plug at time of ignition) affects this limit. Convecting the developing flame kernel from the electrode gap across the plug diameter toward the ground electrode increases the combustion variability: the heat losses from the flame kernel to the ground electrode are highest with this orientation. Also, the cross-sectional shape and size of the ground electrode and center electrode affect the lean limit and combustion variability through the variability in kernel heat loss to the electrodes and the efficiency of the transfer of electrical energy to the spark-generated flame.⁵⁸

The dilution of fuel-air mixture with air (leaner operation), or with exhaust gas recycle, have analogous but not identical effects on ignition. Dilution with EGR degrades ignition system performance (as measured by variability in imep, as a result of combustion variability) more than equivalent dilution with excess air. The dilute operating limits at lighter engine loads with air are almost twice as far beyond stoichiometric in terms of percent dilution, than the extension with exhaust gas recycle beyond zero EGR, due to the greater impact (per unit mass of fuel) of EGR on the mixture burning rate. With modest dilutions ($\leq 10\%$ EGR) there is little difference in the ignition system behavior for EGR and air dilution.⁵⁹

The location of the special conductive seal within the plug insulator shell (see Fig. 9.45) affects the *heat rating* of the spark plug. For a "hot" plug, an insulator with a long conical nose is used; for a "cold" plug a short-nosed insulator is used. The length of the heat conduction path from the insulator nose to the shell is changed in this way to vary and control the temperature of the exposed part of the insulator. The spark plug insulator tip temperature increases with increasing engine speed. It is desirable to have the tip hotter than about 500°C to prevent insulator fouling at low speed and load. High-speed high-load tip temperatures must be kept below about 1100°C to prevent surface ignition. Normally, the gap between the center and ground electrodes is 0.8 to 1.2 mm. For extremely dilute mixtures this is usually increased. Magneto ignition systems use smaller gaps (~ 0.5 mm). High-compression-ratio racing engines use still smaller gaps (0.3 to 0.4 mm).

9.5.3. Alternative Ignition Approaches

A large number of methods for initiating combustion in spark-ignition engines with electrical discharges, in addition to those described in the previous section, have been proposed and examined. These include different designs of spark plug, use of more than one plug, use of higher power, higher energy, or longer-duration discharges, multi-spark systems, and ignition systems that initiate the main combustion process with a high-temperature reacting jet—plasma-jet and flame-jet ignition systems. The standard ignition systems normally ignite the unburned fuel, air, burned gas mixture within the cylinder and perform satisfactorily under conditions away from the lean or dilute engine stable operating limit. Thus, these alternative ignition approaches have the goal of extending the engine's stability limit (and/or of reducing the cyclic combustion variability near the stability limit), usually by achieving a faster initial burning rate than can be obtained with the more conventional systems. This section describes the more interesting of these alternative ignition approaches.

9.5.3.1. Alternative Spark-Discharge Approaches

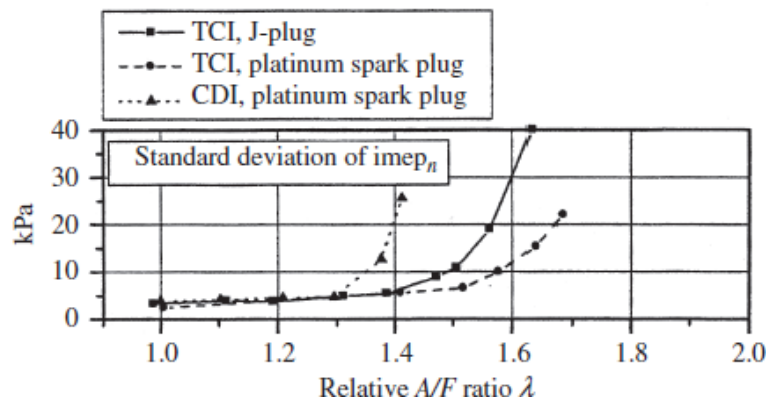
There are many different designs of spark plugs. These use different geometry electrodes, gap widths, and gap arrangements. Electrode arrangements are often categorized as air-gap (e.g., the standard J-type plug design), surface-gap (where the discharge occurs on the insulator surface), and surface-air-gap (a hybrid of these two).

The effects of several plug electrode design features on the engine's stable lean operating limit are illustrated in Fig. 9.46. Ignition system effects are important when misfire due to the quenching effect of the spark plug electrodes determines the stable operating limit. Thus, smaller spark plug center-electrode diameters, larger electrode gap widths, and use of a small diameter platinum center electrode, all extend the lean stability limit to leaner (or more dilute) mixtures for normal spark timings. Figure 9.47 shows how some of the plug designs in Fig. 9.46 affect the standard deviation (and thus COV) of imep as the engine is leaned out. For stoichiometric and up to about 20% lean operation, at this load and speed, the plug and ignition system changes do not affect the combustion variability. The level of dilution at which the imep standard deviation starts to increase rapidly, however, is significantly affected. Note that for the leaner ($\lambda > 1.3$) mixtures, the CDI system with comparable stored energy (140 mJ) to the TCI system (150 mJ), but a shorter discharge time (0.25 ms instead of 1.7 ms) is not as effective. The plug with thinner electrodes and platinum center electrode has the leanest stable operating limit due to its higher energy-into-the-plasma transfer efficiency, and its lower heat losses to the electrodes due to their smaller surface area in contact with the flame kernel.⁵⁹ Multi-gap plugs, designed to produce several discharges which together form a longer arc, have also been used to generate a larger initial flame kernel and thereby extend the lean limit.

Figure 9.46 Several different spark plug designs (with electrode dimensions) and their lean operating limit (λ_{max}) at 2000 rev/min, 280 kPa net imep, moderate tumble, $COV_{imep} = 7\%$.⁵⁹

Name	Production J-Plug	J-Plug with Deeper Spark Location	Spark Plug with Circular Electrodes	Surface-Gap Spark Plug	Platinum Plug with Thin Electrodes
Electrode Gap [mm]	0.8	0.8	1	1.55	1.3
Spark Location x [mm]	3.5	9	5 - 2 - 0	4	3
Drawing					
Lean Limit (λ_{max}) at $\sigma_{imep} = 20$ kPa ($COV_{imep} = 7\%$) 2000 rev/min 280 kPa net imep	1.57	1.50	1.47	1.51	1.67

Figure 9.47 Standard deviation of net imep as a function of relative A/F, λ , at 2000 rev/min, 280 kPa net imep, and moderate tumble in-cylinder flow: TCI J-type plug, TCI thin platinum-center-electrode plug, and CDI platinum electrode plug.⁵⁹



Use of more than one plug, at separate locations in the combustion chamber and fired simultaneously, is also an option. The advantages are twofold. First, the effective flame area in the early stages of flame development is increased substantially: by about a factor of two for two widely spaced plugs. Second, the variations in flow velocity and mixture composition in the vicinity of the (multiple) plugs produce less combined variability in the initial mixture-burning rate than occurs with a single plug. Studies of heat-release rates, flame development angles, and rapid-burning angles, and imep and torque fluctuations, have defined the effects of both increasing the number of ignition sites from 1 to 12 and of changing their geometric location. These results confirm that increasing the number of simultaneously developing flame kernels increases the *initial* mixture-burning rate, as anticipated. It also extends the lean stable operating limit and reduces cyclic combustion variability under conditions where slow and occasional partial-burning cycles would occur with fewer spark plug gaps.⁶⁰

Surface gap plugs (or semi-surface gap plugs) are available, and are considered an option, especially for direct-injection spark-ignition engines. An example of such a spark plug is shown in Fig. 9.46. Two (or more) ground electrodes are used, partly overlapping the insulator tip so the discharge occurs along the insulator surface. The surface discharge provides self-cleaning of the insulator tip which in direct-injection engines, especially when these are operated in stratified mode (see Sec. 9.3.5) where carbon fouling of the insulator is more likely to occur. A potential problem with surface plugs is damage to the insulator tip due to the spark current channeling across the surface. With semi-surface multiple-electrode plug designs (where the spark takes both a gas-spark and surface-spark path) channeling problems can be overcome. The self-cleaning and dispersed sparking features of these plugs stabilize stratified-charge combustion. This type of plug is claimed to be advantageous for the more demanding ignition requirements of direct-injection engines, which operate at higher pressures.⁶¹

Plug and ignition systems that provide multiple spark discharges in sequence are another option. It is thought that each "restrike" of the discharge results in a fresh arc, thus increasing the efficiency with which electrical energy from the ignition coil is transferred to the spark-generated kernel.

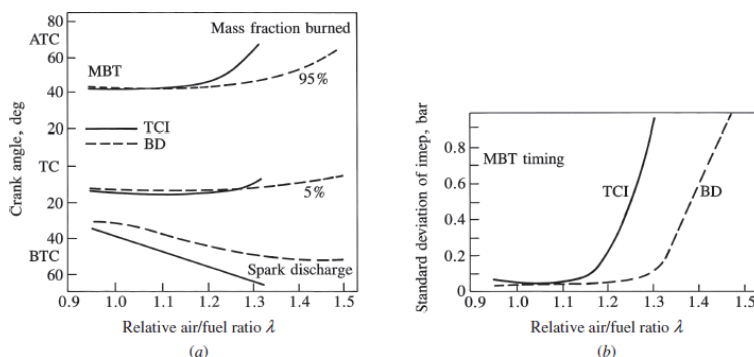
In a broader context, the spark-plug electrodes, with modifications to the ignition system electrical circuit, can be used as an ion current sensor. Once the spark-ignition function has ended, a voltage is applied between the plug electrodes and, by measuring the current (carried by ions) in the hot combustion gases in the electrode gap, the electrical conductivity of that gas can be determined. Thus misfiring engine cycles can be detected since then the gas conductivity is very low due to the low temperature of the unburned mixtures. Also, knock can be detected. The oscillations in peak in-cylinder gas pressure cause oscillations in the temperature of the burned gas that surrounds the spark plug, which cause oscillations in the gas's electrical conductivity, and therefore in the ion current. Thus, there is potential for the ignition system and its spark plug to not only ignite the mixture in the cylinder reliably, but also act as a misfire and knock detector.

Many studies have examined the effects of higher-energy discharges on engine operation near the lean operating limit. It is useful to differentiate between higher current discharges and longer duration discharges: most high-energy (conventional-type) ignition systems have both these features. The results of these studies show that away from the lean or dilute stable

operating limit, increasing the discharge current or duration has no significant effect on engine operating characteristics. The higher current does, as would be expected, result in a larger flame kernel during the initial inflammation process and thereby modestly reduces the spark advance required for MBT with a given combustion chamber and set of operating conditions. However, both higher currents and longer duration discharges do extend the lean engine stability limit (and also the dilute, with EGR, stability limit). Hydrocarbon emissions data indicate that longer discharges have a greater impact than higher currents on extending the misfire limit. The observed increase in fuel consumption as the lean operating limit is approached, results from the rapidly increasing cycle-by-cycle combustion variability, which reduces the average imep and torque. Higher breakdown and arc discharge energies increase electrode erosion.

The discussion of discharge fundamentals in [Sec. 9.5.1](#) showed that depositing energy into the discharge during the initial short breakdown phase resulted in faster flame kernel growth than did depositing the same energy over longer times. In such *breakdown* ignition systems, a capacitor is connected in parallel with the spark plug electrodes, and a low-impedance discharge path allows the energy stored in the capacitor to be discharged into the gap very rapidly. The anticipated effects on the engine's combustion process are observed. Away from the lean engine stability limit, the primary impact is a reduction in the flame development period due to the more rapid initial flame kernel growth. Thus MBT timing is less advanced with these breakdown systems than with conventional systems, as shown in [Fig. 9.48](#). The lean limit can also be extended, and acceptable engine stability obtained (i.e., tolerable cycle-by-cycle combustion variations) for leaner or more dilute engine operation, as shown in [Fig. 9.48b](#).^{62, 63}

Figure 9.48 (a) MBT spark timing and location of 5 and 95% mass fraction burned points for conventional transistorized coil ignition (TCI) system (43 mJ energy, 2 ms duration) and breakdown system (BD) (43 mJ energy, ~10 ns duration) as function of relative air/fuel ratio λ ($=1/\phi$). (b) Standard deviation in indicated mean effective pressure as a function of relative air/fuel ratio λ for TCI and BD systems.⁶²

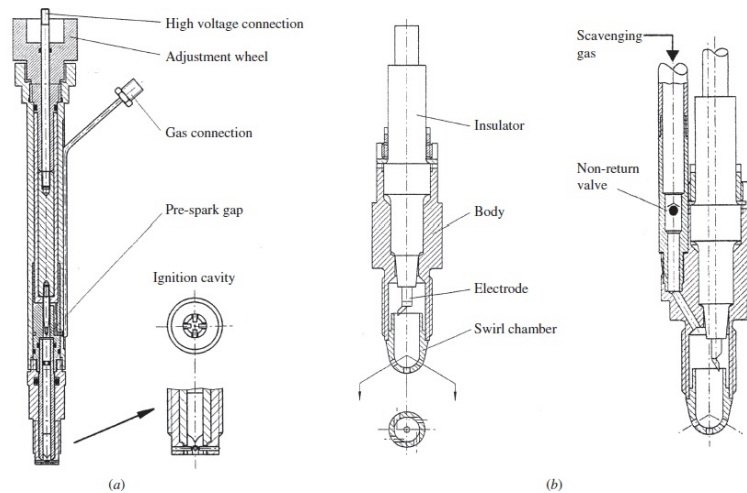


9.5.3.2. Plasma-Jet Ignition

In a plasma-jet ignitor, the spark discharge occurs within a cavity that surrounds the plug electrodes, which connects with the combustion chamber via an orifice (see [Fig. 9.49a](#)). The electrical energy supplied to the plug electrodes is substantially increased above values used in normal ignition systems by allowing a capacitor to discharge at a relatively low voltage and high current, through the spark generated in a conventional manner with a high-voltage low-current ignition system. Stored energies of several hundred mJ are used, and this energy is discharged in some 20 μ s. This high-power discharge creates a high-temperature plasma so rapidly that the pressure in the cavity increases substantially, causing a supersonic jet of plasma to flow from the cavity into the main combustion chamber. The plasma enters the combustion chamber as a turbulent jet, preceded by a hemispherical blast wave. The gas dynamic effects of the blast wave are dissipated by the time combustion starts. The initial combustion rate is typically an order of magnitude faster than with the standard TCI system. Ignition in the main combustion chamber takes place in the turbulent plasma jet; the flame starts out as a turbulent flame in contrast to the flame with conventional ignition systems which is initially laminar-like. The penetration of the jet depends on its initial momentum; it thus depends on the amount of energy deposited, cavity size, and orifice area. If the cavity is filled prior to

ignition with a hydrocarbon (or a mixture of hydrocarbons), the ignition capabilities are enhanced due to the large increase in hydrogen atoms created in the plasma.^{59, 64}

Figure 9.49 (a) Plasma-jet ignition system and (b) swirl chamber spark plug (unscavenged swirl chamber on left, scavenged swirl chamber on right).⁵⁹



The effects of plasma-jet ignitors on engine combustion are analogous to those of breakdown ignition systems. The flame development period is significantly shortened: thus ignition at MBT timing occurs closer to TC. The lean/dilute mixture levels where cycle-by-cycle combustion variability becomes large are significantly extended, as are the engine's lean or dilute stable operating limit.⁵⁹

9.5.3.3. Flame-Jet Ignition

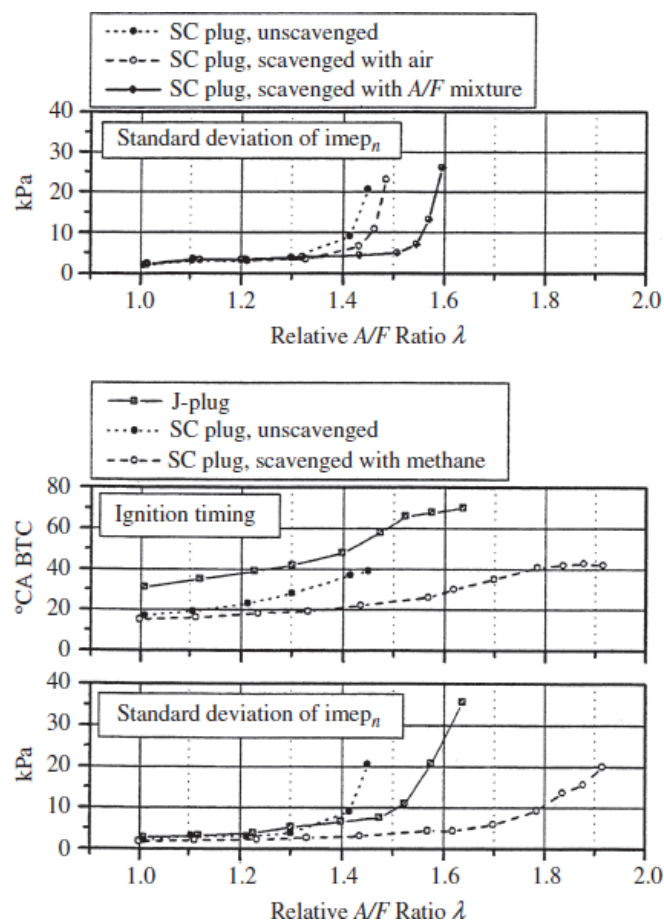
With this type of system, ignition occurs in a prechamber cavity, which is physically separated from the main chamber above the piston and is connected to it via one or more orifices or nozzles. See the swirl chamber and scavenged chamber plug designs in Fig. 9.49b. Locating the spark plug gap within a prechamber results in increased velocities and gas temperatures (through heat transfer from the prechamber walls) in the mixture surrounding the plug electrodes. As the flame develops in this cavity, the pressure of the gases in the prechamber rises, forcing gas out into the main chamber through the orifice (or orifices) as one or more burning turbulent jets. The jet, or with a larger orifice the "cloud," penetrates into the main chamber igniting the unburned mixture in the main chamber, thereby initiating the primary combustion process. Ignition within the cavity is usually achieved with a standard spark discharge. The function of the prechamber or cavity is to transform the initial flame around the spark plug electrodes into one or more pressure-driven flame jets (or cloud) into the main chamber, which have a substantial surface area that can ignite extremely lean or dilute mixtures in a repeatable manner. Many different systems for achieving this goal have been developed; some of these have been used in production spark-ignition engines.

In the simpler types of flame-jet ignition concept, the cavity is unscavenged and there is no prechamber fuel metering system. The function of the prechamber cavity is to increase the initial growth rate of the flame immediately following spark discharge. Another approach is to incorporate a (small) cavity into the spark plug. Systems with prechamber volumes varying from 20% of the clearance volume to about 1% have been developed. The flow pattern produced within the prechamber by the flow into the cavity during compression, the location of the spark plug electrodes within the cavity, and of the nozzle or orifice, are critical design issues. A problem with these systems is that the prechamber is never completely scavenged by fresh mixture between cycles, so the burned gas fraction in the unburned mixture within the prechamber is always higher than the burned gas fraction in the unburned main-chamber mixture.

The ignition and combustion characteristics of the swirl-chamber (SC) plugs, such as those in Fig. 9.49b, are shown in Fig. 9.50 where they are compared to the standard J-plug system. The size of prechamber volume is 2% of the cylinder clearance

volume. The SC plug center electrode has a platinum pin welded to it. The ignition spark arcs from the pin to the prechamber wall. The pressure differences at this part-load low-speed condition between pre- and main-chamber, as combustion transitions from the first of these regions to the latter, are 100 to 200 kPa. Comparisons of combustion phasing with the standard ignition system, as the engine is leaned out, show that the optimum (MBT) spark timing for the flame-jet swirl chamber plug is significantly less advanced than MBT spark for a standard J-plug system due to a faster initial burning rate as the flame develops. However, the crank angles of the middle and end of combustion are closely comparable.

Figure 9.50 Combustion characteristics (standard deviation of net imep and ignition timing) as a function of relative air/fuel ratio of swirl chamber spark plug: unscavenged, and scavenged with air, air-fuel mixture, and methane. J-plug characteristics shown for comparison. 2000 rev/min, 280 kPa net imep, moderate tumble.⁵⁹



The lean operating limits of the different SC plugs depend on whether these plugs are unscavenged, or scavenged with various scavenging gases. The stable lean operating limit is defined by the rapid rise in standard deviation of imep caused by the rapidly worsening cycle-by-cycle combustion variability as the mixture is increasingly leaned out. In unscavenged SC plug operation, this limit is less lean than with the standard J-plug TCI system. Scavenging improves the lean limit: modestly with air scavenging, more with air-fuel mixture, and most with methane. Scavenging removes the high concentration of residual burned gas left in the swirl chamber after combustion, and the different scavenging gases, air, air-fuel mixture, and fuel, in this sequence produce mixture that is successively easier to ignite and burn in the prechamber.

With close to stoichiometric operation, combustion variability impacts are small. With significantly diluted mixtures, the more rapid and more repeatable combustion initiation process in the pre- and main-chamber does provide benefits. There is a history of prechamber stratified engines where, in addition to the main fuel-injection process which occurs into the fresh charge in the intake system, at the time of spark a readily-ignitable rich-of-stoichiometric mixture was set up in the prechamber. With this approach, the prechamber volume was usually 20 to 25% of the clearance volume. A separate small

intake valve fed very rich mixture into the prechamber during the intake process, while the main fuel metering system fed lean mixture to the main intake valve. During intake, the prechamber is completely scavenged by the rich intake stream. During compression, the lean mixture flowing from the main chamber to the prechamber brings the prechamber mixture equivalence ratio to slightly rich-of-stoichiometric at the time of spark discharge.

9.5.3.4. Summary

Overall, several of these alternative ignition concepts extend the engine's lean stable operating limit relative to conventional engines with standard ignition systems by several air/fuel ratios. However, with stoichiometric operation (even with 10 or so percent EGR) "more powerful" ignition systems have little impact on the stability of the engine's combustion process. The primary reason is that under these conditions, combustion variability, cycle-by-cycle, is modest and "improved ignition" is not needed. That is, if ignition is already satisfactory a better ignition system is of limited value.

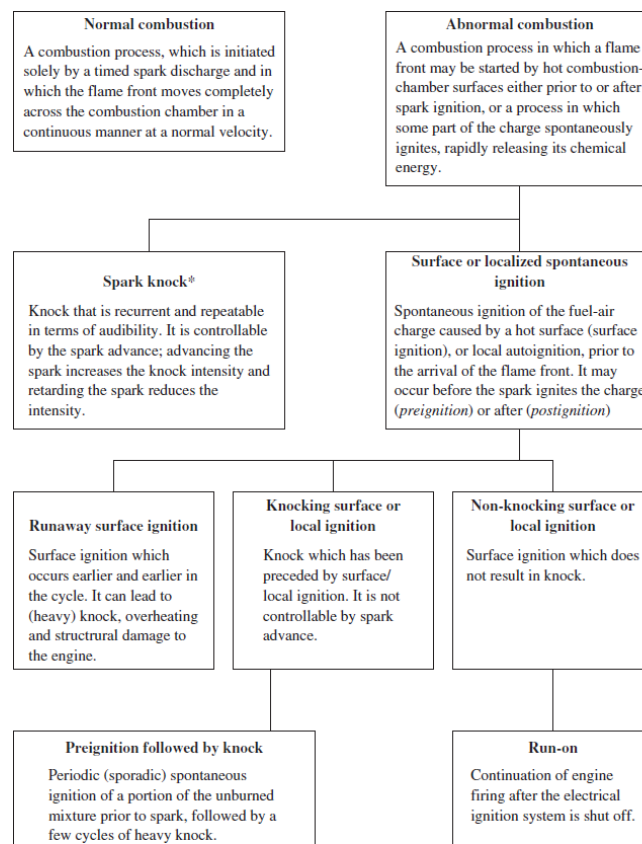
9.6. ABNORMAL COMBUSTION: SPONTANEOUS IGNITION AND KNOCK

9.6.1. Description of Phenomena

Abnormal combustion in SI engines reveals itself in many ways. Of the various abnormal combustion processes which are important in practice, two major phenomena are knock and surface ignition. These abnormal combustion phenomena are of concern because: (1) when severe, they can cause major engine damage; and (2), even if not severe, they are regarded as objectionable by the engine or vehicle operator. *Knock* is the name given to the noise which is transmitted through the engine structure when spontaneous ignition of a significant portion of the end-gas—the fuel, air, residual gas, mixture ahead of the propagating flame—occurs. When this abnormal combustion process takes place, there is usually an extremely rapid release of chemical energy in the end-gas, causing very high local pressures and the propagation of pressure waves of substantial amplitude across the combustion chamber. *Surface ignition* is ignition of the fuel-air mixture by a hot spot on the combustion chamber walls such as an overheated valve or spark plug, or glowing combustion chamber deposit, that is by a means other than the normal spark discharge. It can occur before the occurrence of the spark (*preignition*) or after (*postignition*). Following surface ignition, a flame develops at each surface-ignition location and starts to propagate across the chamber in an analogous manner to what occurs with normal spark ignition.

Because the spontaneous ignition phenomenon that causes knock is governed by the temperature and pressure history of the end gas, and therefore by the phasing and rate of development of the combustion process, various combinations of these two phenomena—localized spontaneous ignition and knock—can occur. These are categorized in [Fig. 9.51](#). When autoignition occurs repeatedly during otherwise normal combustion events, the phenomenon is called *spark-knock*. Repeatedly here means occurring more than occasionally: the knock phenomenon varies substantially cycle-by-cycle, and between the cylinders of a multi-cylinder engine, and does not necessarily occur every cycle (see below). Spark-knock is controllable by the spark advance: advancing the spark increases the knock severity or intensity and retarding the spark decreases knock. Since surface ignition usually causes a more rapid rise in end-gas pressure and temperature than occurs with normal spark ignition (because the flame either starts propagating sooner, or propagates from more than one source), knock is a likely outcome at higher engine loads following the occurrence of surface or local ignition. To identify whether or not surface ignition causes knock, the terms knocking surface ignition and non-knocking surface ignition are used. Knocking surface ignition usually originates from preignition caused by glowing combustion chamber deposits: the severity of knock generally increases the earlier that preignition occurs. Knocking surface ignition cannot normally be controlled by retarding the spark timing since the spark-ignited flame is not the cause of knock. Non-knocking surface ignition is usually associated with surface ignition that occurs late in the operating cycle.

Figure 9.51 Definition of the various normal and abnormal spark-ignition engine combustion phenomena: spark knock, spontaneous surface and local ignition, sporadic preignition, and their relationships.



*Knock: The noise associated with the spontaneous ignition of the fuel-air mixture ahead of the advancing flame front. This spontaneous ignition and the resulting very rapid energy release from the fuel-air mixture this end-gas ahead of the flame generates strong pressure waves.

As in-cylinder pressures have risen due to increasing engine compression ratios and engine turbocharging, a new knock phenomenon has been observed: *sporadic preignition*. Intermittently, a local region of the in-cylinder unburned mixture autoignites toward the end of the compression stroke before the spark discharge occurs. The subsequent chemical energy release can result in abnormally high pressures and unburned mixture temperatures which can lead to heavy engine knock. Effectively, in a few cycles (once every 10^4 or so cycles), this preignition-initiated combustion is very over advanced, and extremely high cylinder pressures result. This stochastic preignition lasts for one, or a few cycles, and then stops. The probable cause of this phenomenon is that oil from the region between the cylinder liner and the piston top land is periodically transported into the cylinder forming a local region of autoignitable hydrocarbon-air mixture. This localized spontaneous ignition of mixture in the cylinder is a sporadic, seemingly stochastic, preignition process that initiates an over-advanced burning process in the rest of the combustion chamber, which leads to heavy knock.

Runaway surface ignition is surface ignition that occurs earlier and earlier in the cycle. It is usually caused by overheated spark plugs or valves or other combustion chamber surfaces. It is the most destructive type of surface ignition since it can lead to serious overheating and structural damage to the engine.

After some additional description of surface-ignition phenomena, the remainder of [Sec. 9.6](#) will focus on knock. This is because surface ignition is a problem that can be solved by appropriate attention to engine design, and fuel and lubricant quality. In contrast, spark-knock is an inherent constraint on engine performance and efficiency since it limits the maximum compression ratio (and boost levels in turbocharged engines) that can be used with any given fuel.

Of all the engine surface-ignition phenomena identified in [Fig. 9.51](#), preignition is potentially the most damaging. Any process that advances the start of combustion ahead of the timing that gives maximum torque will result in higher unburned mixture

temperatures and pressures, and thus heavier knock. It also results in higher heat rejection because of the increasing burned gas pressures and temperatures that result. Higher heat rejection causes higher component temperatures, which in turn, can advance the preignition point (with surface preignition) even further until critical components can fail. The surfaces which can cause preignition are those least well cooled and where deposits build up and provide additional thermal insulation: primary examples are spark plugs, exhaust valves, and deposit build-up on the piston crown and cylinder head. Under normal conditions, using suitable heat-range spark plugs, preignition is usually initiated by an exhaust valve covered with deposits formed from the fuel, and the lubricant which penetrates into the combustion chamber. Cooler exhaust valves and reduced oil consumption usually alleviate this problem. Engine design features that minimize the likelihood of preignition are: appropriate heat-range spark plugs, removal of sharp asperities, rounded metal edges, and well-cooled exhaust valves with sodium-cooled valves as an extreme option.

Knock primarily occurs at loads approaching wide-open-throttle operating conditions. It is thus a direct constraint on engine performance. It also constrains engine efficiency, since by effectively limiting the temperature and pressure of the end-gas at higher loads, it limits the engine compression ratio. The occurrence and severity of knock depend on the knock resistance of the fuel and on the knock suppressing features of the engine. The ability of a fuel to resist knock is measured by its octane number (ON): higher ONs indicate greater resistance to knock (see [Sec. 9.6.3](#)). Gasoline octane ratings can be improved by refining processes such as catalytic cracking and reforming, which convert low-octane hydrocarbons to high-octane hydrocarbons. The octane-number requirement of an engine depends on how its design and the conditions under which it is operating, affect the temperature and pressure of the end-gas ahead of the flame, and on the time required to burn the cylinder charge. An engine's tendency to knock, as defined by its *octane requirement*—the octane rating of the fuel required to avoid knock—is increased by factors that produce higher temperatures and pressures or lengthen the burning time. Thus knock is a constraint that depends on both the quality of available fuels and on the ability of the engine designer to achieve the desired normal combustion behavior while maximizing the engine's performance.

The pressure variation in the cylinder during knocking combustion indicates in more detail what actually occurs. [Figure 9.52](#) shows the cylinder pressure variation in three individual engine cycles, with normal combustion, light knock, and heavy knock, respectively.⁶⁵ When knock occurs, high-frequency pressure fluctuations are observed whose amplitude decays with time. [Figures 9.52a](#) and [b](#) have the same operating conditions and spark advance. About one-third of the cycles in this engine at these conditions had normal, smoothly varying, cylinder pressure records as in [Fig. 9.52a](#). Knock of varying severity occurred in the remaining cycles. With light knock, knock occurs late in the burning process, at close to the peak pressure, and the amplitude of the pressure fluctuations is small ([Fig. 9.52b](#)). Note that the peak pressure in the center figure is higher and earlier than that in the normal cycle; it is a faster-burning cycle. With heavy knock, illustrated here with more advanced spark timing and by selecting a high intensity knocking cycle, knock occurs closer to top-center earlier in the combustion process and the initial amplitude of the pressure fluctuation is much larger. These pressure fluctuations produce the sharp metallic noise called "knock." They are the result of the essentially spontaneous release of much of the end-gas fuel's chemical energy. This produces a substantial *local* increase in gas pressure and temperature, thereby causing a strong pressure wave to propagate away from the end-gas region across the combustion chamber. This strong pressure wave, the expansion wave that accompanies it, and the reflection of these waves by the chamber walls create the oscillatory pressure versus time records shown in [Figs. 9.52b](#) and [c](#). Note that once knock occurs, the pressure distribution across the combustion chamber is no longer uniform: transducers located at different points in the chamber will record different pressure levels at a given time until the wave propagation phenomena described above have been damped out.

Figure 9.52 Cylinder pressure versus crank-angle traces of individual cycles with (a) normal combustion, (b) light knock, and (c) heavy knock. 4000 rev/min, wide-open-throttle, 381-cm³ displacement single-cylinder engine.⁶⁵

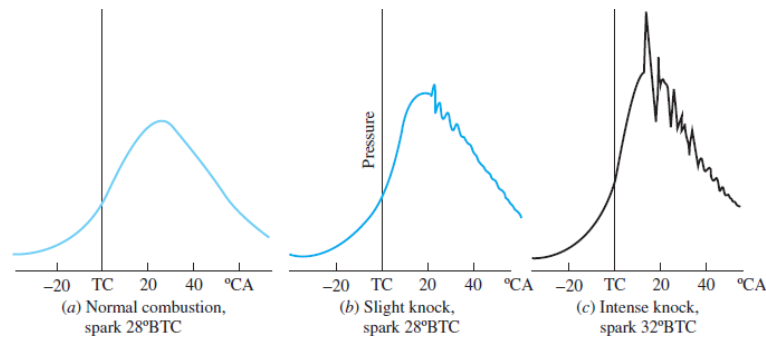
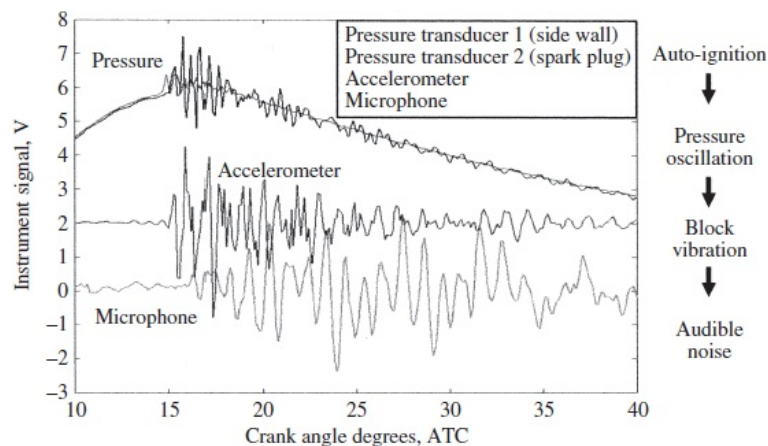


Figure 9.53 illustrates this sequential process—end-gas autoignition leading to audible noise. If the chemical energy release during autoignition of the end gas is large and rapid enough, then pressure oscillations of a few or more bar in amplitude are set up inside the cylinder. As explained above, these pressure waves originate from the high local pressure generated in the end-gas by autoignition.^[1] Pressure measurements at various locations in the combustion chamber vary in amplitude and phasing, and have several different specific frequencies—see below. These pressure waves cause the block (and cylinder head) to vibrate as indicated in **Fig. 9.53** by the output from an accelerometer on the block outer wall. The vibrating block (and head) radiates noise which is picked up by a microphone about 10 cm from the cylinder head outer wall (this air gap explains the delay in the knock sound wave arriving at the microphone). The in-cylinder pressure waves take one or a few oscillations to establish themselves: they then decay since the driving process—autoignition of the end gas—is over.

Figure 9.53 How autoignition-produced pressure oscillations within the engine cylinder produce block wall vibration (measured by an accelerometer) which produces noise (microphone, some 10 cm away from block wall).⁶⁶



The interpretation of the in-cylinder pressure signal depends on the pressure transducer location and the acoustic modes excited by the autoigniting end-gas region. Draper⁶⁷ analyzed the resonant acoustic modes of a disk-shaped combustion chamber. More recent analysis (e.g., [Ref. 68](#)) adds to his findings. The lowest frequency mode is the first circumferential mode at about 6 kHz for normal automotive size engine cylinders ($B \approx 85$ mm): this is the basic knock frequency that we hear. Next in frequency is the second circumferential mode at about 10 kHz, a frequency that is barely audible to the human ear. Several modes have frequencies in the 18 to 20 kHz range.⁶⁸

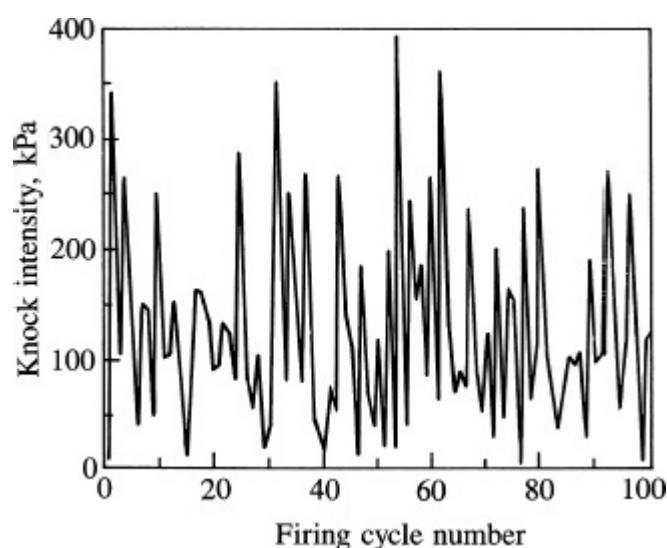
Obvious questions at this point are how should we define knock onset, and once knock occurs what is its severity or intensity? The combined physical and chemical process is complex and many engine and fuel factors are involved, as also are ambient conditions. The transition from normal non-knocking combustion to knocking combustion is just that—a transition. As spark

timing is advanced, the pressure levels reached during combustion increase. As a consequence, the temperatures attained in the end gas increase as does their maximum value (at peak pressure, some three quarters of the way through the combustion process). At a certain end-gas temperature and pressure, autoignition starts to occur releasing a small fraction of the chemical energy in the fuel molecules ahead of the flame as it propagates across the combustion chamber. Prior to knock onset, the rate of energy release is slow. As the spark timing continues to be advanced, the autoignition energy release rate and amount increase until it is rapid enough to generate high local pressures that set up strong pressure waves. As the amplitude of these pressure waves increase, block and cylinder head vibrations generate sound waves that become audible, and then louder as the autoignition process speeds up and spreads, further increasing the pressure levels generated in the end-gas region. Thus several measures—amplitude of in-cylinder pressure oscillations, extent of block and head vibration, noise level generated—can be used to define knock onset (often called borderline or audible knock), as described below.

Several methods of knock detection and characterization have been used. The human ear is a surprisingly sensitive knock detector and is routinely used in determining the *octane requirement* of an engine—the required fuel quality the engine must have to avoid knock. Knock detectors used for knock control systems normally respond to the vibration-driven acceleration of parts of the engine block caused by knocking combustion pressure waves. A high-intensity flash is observed when knock occurs; this is accompanied by a sharp increase in ionization. Optical probes and ionization detectors have therefore been used. The spark plug can serve as an ionization detector. For more detailed studies of knock in engines, the piezoelectric pressure transducer is the most useful monitoring device. Often the transducer signal is filtered so that the pressure fluctuations caused by knock are isolated.

The amplitude of the pressure fluctuation is a useful measure of the intensity of knock because it depends on the amount of end-gas which ignites spontaneously and rapidly, and because engine damage due to knock results from the high gas pressures (and temperatures) in the end-gas region. Use of this measure of knock severity or intensity shows there is substantial variation in the extent of knock, cycle-by-cycle. Figure 9.54 shows this pressure-fluctuation intensity in one hundred consecutive cycles in a given cylinder of a multi-cylinder engine operating at fixed conditions for knocking operation. The intensity varies randomly from essentially no knock to heavy knock.⁶⁹ Cylinder-to-cylinder variations are also substantial due to variations in compression ratio, mixture composition and conditions, burn rate, and combustion chamber cooling. One or more cylinders may not knock at all while others may be knocking heavily.

Figure 9.54 Knock intensity (maximum amplitude of bandpass-filtered pressure signal) in 100 individual consecutive cycles. One cylinder of V-8 engine, 2400 rev/min, wide-open throttle.⁶⁹



Since the knock phenomenon produces a nonuniform state in the cylinder, and since the details of the knock process in each cycle and in each cylinder are different, a fundamental definition of knock intensity or severity is difficult. The ASTM-CFR method for rating fuel octane quality (see Sec. 9.6.3) by the severity of knocking combustion uses the time derivative of

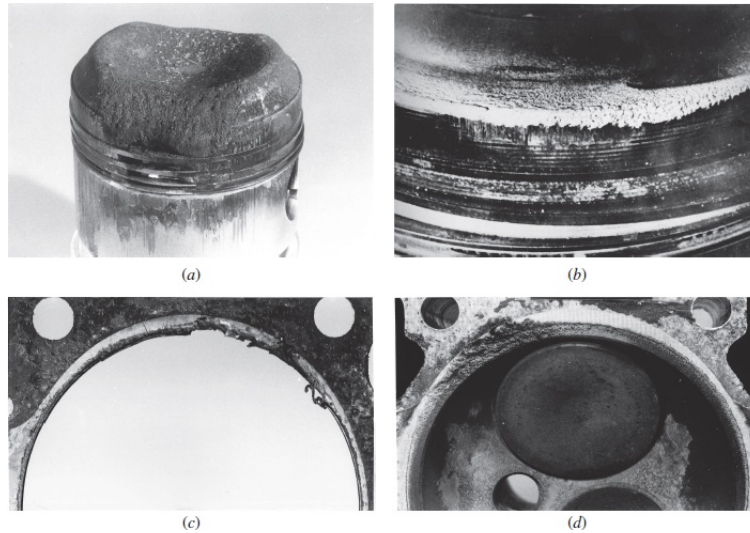
pressure during the cycle. Cylinder pressure is determined with a pressure transducer. The low-frequency component of pressure change due to normal combustion is filtered out and the rate of pressure rise is averaged over many cycles during the pressure fluctuations following knock. This approach obviously provides only an average relative measure of knock intensity. The maximum rate of pressure rise has been used to quantify knock severity. An accelerometer mounted on the engine can give indications of relative knock severity provided that it is mounted in the same location for all tests. The most precise measure of knock severity is the maximum amplitude of the pressure oscillations that occur with knocking combustion. The cylinder pressure signal (from a high-frequency-response pressure transducer) is filtered with a band-pass filter so that only the component of the pressure signal that corresponds to the fluctuations occurring after knock remains. The filter is set for the first circumferential mode of gas vibration in the cylinder (in the 3 to 10 kHz range, depending on bore and chamber geometry). The maximum amplitude of the pressure oscillations gives a good indication of the severity of knock. The knock intensities in the individual cycles shown in [Fig. 9.54](#) were determined in this manner. Note that because the pressure fluctuations are the consequence of a wave propagation phenomenon, the location of the pressure transducer in relation to the location of the knocking end-gas, and the shape of the combustion chamber, will affect the magnitude of the maximum recorded pressure-fluctuation amplitude.

The impact of knock depends on its intensity and duration. Trace knock has no significant effect on engine performance or durability. Heavy knock can lead to extensive engine damage. In automobile applications, a distinction is usually made between "acceleration knock" and "constant-speed knock." Acceleration knock is primarily an annoyance, and due to its short duration is unlikely to cause damage. Constant-speed knock, however, can lead to two types of engine damage. It is especially a problem at high engine speeds where knock is masked by other engine noises and is not easily detected. Heavy knock at constant speed can lead to:

- Preignition, if significant deposits are present on critical combustion chamber components. This could lead to runaway preignition.
- Runaway knock—spark-knock occurring earlier and earlier, and therefore more and more intensely. This soon leads to severe engine damage.
- Gradual erosion of regions of the combustion chamber walls, even if runaway knock does not occur.

The engine can be damaged by knock in different ways: piston ring sticking; breakage of the piston rings and lands; failure of the cylinder head gasket; cylinder head erosion; piston crown and top land erosion; piston melting and holing. Examples of component damage due to preignition and knock are shown in [Fig. 9.55](#).^{70–72}

Figure 9.55 Examples of component damage from abnormal engine combustion: (a) piston holing by preignition;⁷¹ (b) piston crown erosion after 10 hours of high-speed knocking;⁷⁰ (c) cylinder head gasket splitting failure due to heavy knock;⁷¹ (d) erosion of aluminum cylinder head along the top of the cylinder liner due to heavy knock.⁷¹



The mechanisms that cause this damage are thought to be the following. Preignition damage is largely thermal as evidenced by fusion of spark plugs or pistons. When knock is very heavy, substantial additional heat is transferred to the combustion chamber walls and rapid overheating of the cylinder head and piston results. Under these conditions, knock is not stable: the overheating increases the engine's octane requirement, which in turn increases the intensity of knock. It becomes heavier and heavier, and the uncontrolled running away of this phenomena can lead to engine failure in minutes. This damage is due to overheating of the engine: the piston and rings seize in the bore. The damage due to heavy knock over extended periods—erosion of piston crowns and (aluminum) cylinder heads in the end-gas region—is due primarily to the oscillating high gas pressures in this region. Extremely high-pressure pulses of up to 180 atm due to heavy knock can occur locally in the end-gas region, at around 6 kHz and in the 15 to 20 kHz frequency range. These high local fluctuating pressures, combined with the higher-than-normal local surface temperatures which occur with the higher knocking heat fluxes, weaken the material. Pitting and erosion due to fatigue with these excessive mechanical stresses, and breakage of rings and lands, can then occur.^{70–73}

In practice, while knock onset continues to constrain spark-ignition engine design, and thus engine performance and efficiency (with a given fuel), actual knock damage to engines is increasingly rare. Several factors contribute to this trend; for example, use of knock sensors, more sophisticated control of engine operation, and tighter tolerances on fuel antiknock characteristics. However, knock continues to be a fundamental limitation since knock is viewed by vehicle users as objectionable, and the transition from light knock to heavy knock (which could cause engine damage) can occur with modest changes in engine operating condition (e.g., a few degrees change in spark timing).

9.6.2. Knock Fundamentals

Knock originates through autoignition of portions of the unburned mixture ahead of the flame when one or more localized regions in the end gas are compressed to sufficiently high pressures and temperatures to spontaneously react. Initially, the rate of chemical energy release as the fuel molecules start to oxidize is relatively slow, but that energy release rate then accelerates causing a rapid rise in local pressure and temperature setting up strong pressure waves which impact the surrounding mixture and flow. Thus "autoignition" does not occur uniformly throughout the end-gas, it originates in local regions where the composition (concentrations of fuel vapor, air, burned residual gases) and temperature are such that the induction times (time for the oxidation chemistry to get started) are shorter due to these local inhomogeneities and variations in mixture thermal state. These local "easier to ignite" regions are termed exothermic centers. The transition of these

exothermic centers into knock depends on their size, the thermal and concentration gradients at their boundaries, and their location in the end gas.²³

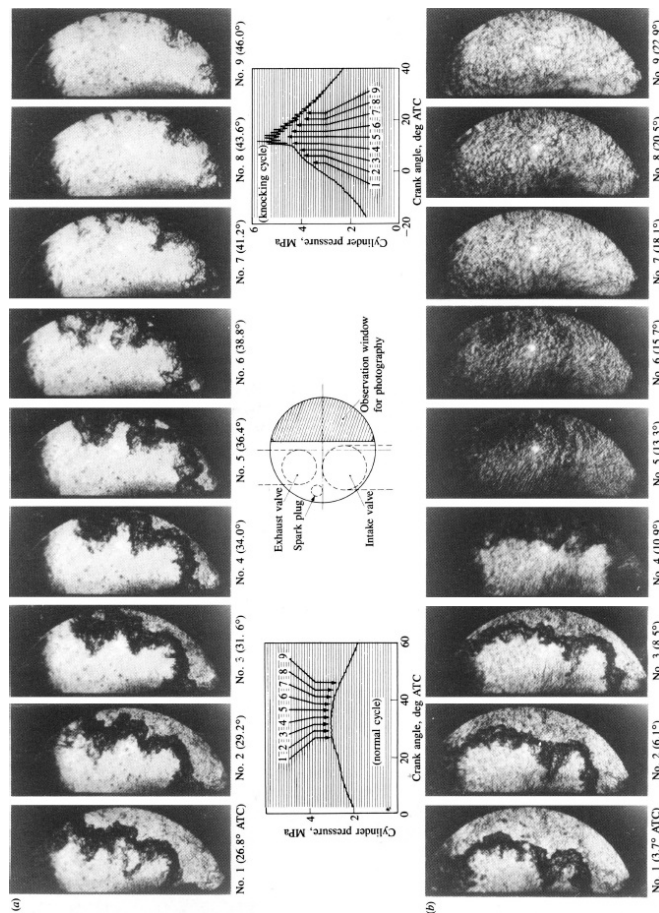
Three modes of knocking combustion are observed, following autoignition of these exothermic centers:

1. *Deflagration*, the common mode, generally resulting from small exothermic centers with steep gradients, creating weak pressure oscillations: after ignition, flame propagation from these centers is similar to the normal SI engine flame propagation process.
2. *Thermal explosion*, characterized by larger centers with flatter gradients: pressure oscillations are moderate, and knock damage, if any occurs, is light.
3. *Developing detonation*, the most violent knock mode: occurring rarely, but generating large pressure oscillations with potential for significant surface damage, characterized by medium-sized exothermic centers with intermediate gradients which lead to strong pressure waves driven by the chemical energy release that these waves induce.

This engine phenomenon "knock" includes also the propagation of strong pressure waves across the chamber, chamber resonance, and transmission of sound through the engine structure. That detonation waves may be part of the intense knock phenomenon, has led many to call knock "detonation." However, the more general term "knock" is preferred, since the phenomenon includes more than autoignition-initiated rapid end-gas energy release.

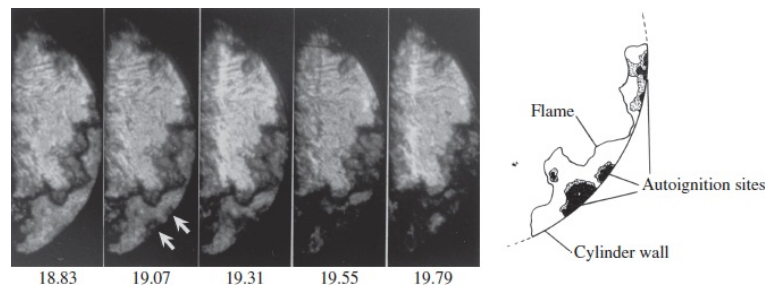
Photographic studies of knocking combustion have been an important source of insight into the fundamentals of the phenomenon over the past 50 years. Figure 9.56 shows two sets of schlieren photographs, one from a cycle with normal combustion and the other from a cycle with knock.⁷⁴ Set (a) were taken in an overhead valve engine with a disc-shaped combustion chamber, with a window which permits observation of the chamber opposite to the spark plug as shown. A reflecting mirror on the piston crown permits use of the schlieren technique, which identifies regions where changes in gas density exist. Operating conditions, except for spark advance, were the same for both cycles. In the normal combustion sequence, the turbulent flame front moves steadily through the end-gas as combustion goes to completion. The cylinder pressure varies smoothly throughout this process. When the spark is advanced by 15°, the end-gas temperature and pressure are increased significantly and knock occurs. In the second sequence of photographs, (b), the initial flame propagation process (photographs 1 to 3) is like that of the normal combustion sequence: then almost the entire region ahead of the flame appears dark (photograph 4). Between photographs 3 and 4 substantial changes in the density and temperature throughout most of the end-gas region have occurred. Examination of the cylinder pressure trace shows that this corresponds closely to the time when the pressure recorded by the pressure transducer rises rapidly. Immediately after this, pressure oscillations (at 6 to 8 kHz) are detected. In photograph 5 the flame is no longer visible: rapid end-gas expansion has disrupted the flame region and pushed it back out of the field of view. Subsequent photographs are alternatively lighter and darker, indicative of changing local density fields as pressure waves propagate back and forth across the chamber.

Figure 9.56 Schlieren photographs from high-speed movies of (a) normal flame propagation through the end gas and (b) knocking combustion (autoignition occurs in photograph 4), with corresponding cylinder pressure versus crank-angle traces. Disc-shaped combustion chamber with details of window shown in insert, 1200 rev/min, 80% volumetric efficiency, $(A/F) = 12.5$. Spark timing: (a) 10° BTC, (b) 25° BTC.⁷⁴



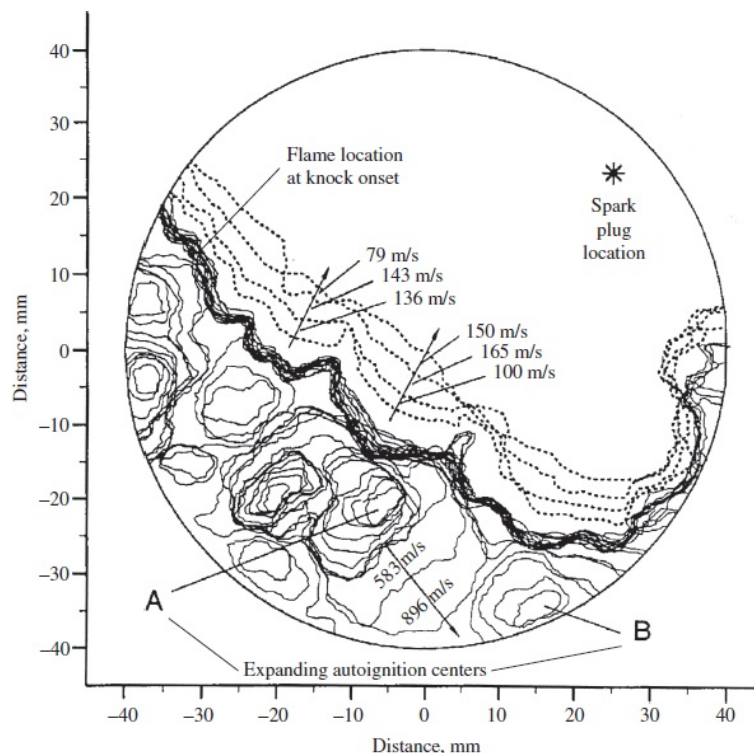
More extensive studies of this type, which relate photographs from high-speed movies or videos of the combustion process to the cylinder pressure development, indicate that locations where ignition of one or more portions of the end-gas first occur, and the subsequent rate with which the ignition process develops throughout the rest of the end-gas, vary substantially cycle-by-cycle and with the intensity of the knocking process. **Figure 9.57** shows five shadowgraph photographs from a knocking engine cycle in a research engine similar to that shown in **Fig. 9.56**. The photographs are $33 \mu\text{s}$ apart; the total sequence shown lasts 1° . The first image shows the flame front prior to onset of knock. The second image shows the onset of autoignition with the appearance of dark regions (developing exothermic centers) near the wall (two identified by arrows) where substantial density gradients resulting from local energy release exist. The third, fourth, and fifth photographs show the spread of these ignited regions with time through the remaining end-gas. The exact location where autoignition occurred was identified with a photo digitizing system, which ranked regions of the photograph by their brightness or darkness. The digitized version of the second photograph is also shown in **Fig. 9.57**. Additional smaller regions of autoignition are evident near the wall and in the vicinity of the flame front. The location of autoignition sites varied with engine operating conditions. The rate of spread of the autoignited end-gas region also varied significantly. Under heavy knocking conditions, the entire end-gas region became ignited very rapidly and high-amplitude pressure oscillations occurred. Under trace knock conditions, autoignition could occur; yet the spread of the autoignited region be sufficiently slow so that only modest pressure oscillations occur.

Figure 9.57 Five shadowgraph photographs of a knocking combustion cycle identifying location of autoignition sites (arrows). Crank angle of each photo indicated: 33 μ s between frames. Operating conditions and engine details as in Fig. 9.56. Photo digitized picture of second photograph (on right) shows additional details of the autoignition sites in the end-gas region.⁷⁵



The end-gas autoignition in these exothermic centers and the resulting spread of the burning process are illustrated by the data in Fig. 9.58. Data from simultaneous LIF images of formaldehyde, formed in the end-gas during the early stages of autoignition which burns up as exothermic centers (ETCs) develop and release their chemical energy, and images from ultra-high-speed schlieren movies which track the motion of the exothermic center boundaries have been combined. Of the several ETCs shown, most are like B where the center develops into a deflagration propagating outward at subsonic speed. However, center A transitions into a developing detonation propagating outward on its lower right-hand side at supersonic speeds, driven by the shock-wave-induced rapid unburned end-gas mixture chemical energy release.

Figure 9.58 Tracings from an ultra-high-speed schlieren movie (750,000 frames/sec) showing formation and outward propagation rates of exothermic autoignition enters in the end-gas. Most centers (e.g., B) develop into deflagrations whereas the lower right-hand side of A transitions into a developing detonation wave. Note the irregular distribution of ECTs, and the pushing back of the regular flame front by the fast end-gas pressure rise.²³

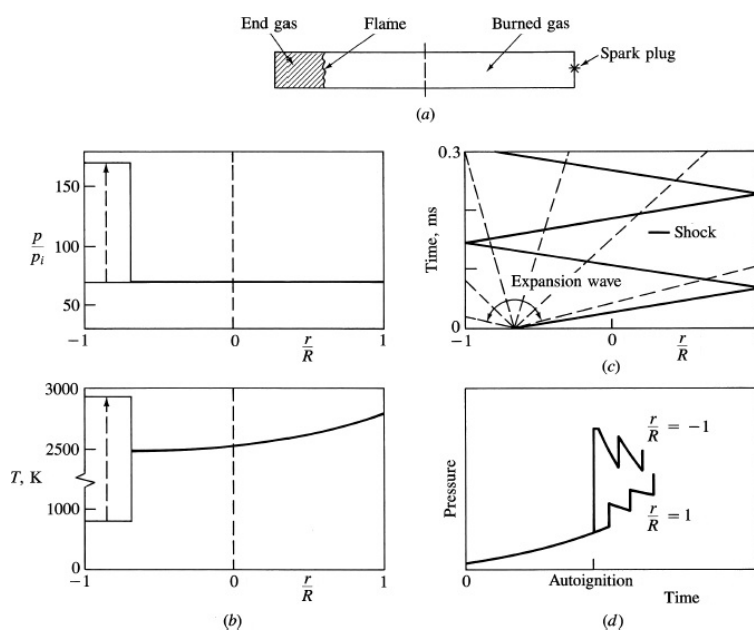


When the above-described end-gas ignition process occurs rapidly, the gas pressure in the end-gas region rises substantially due to the rapid release of the end-gas fuel's chemical energy. A strong pressure or shock wave propagates from the outer

edge of this high-pressure end-gas region across the chamber at supersonic velocity, and an expansion wave propagates into the high-pressure region toward the near wall. The shock wave and expansion wave reflect off the walls of the chamber, producing standing waves of substantial amplitude. The amplitude of the pressure oscillations builds up as the standing waves are established, and then decays as the wave-induced gas motion is damped out. The frequency of the pressure oscillations decreases with time as the initially supersonic pressure waves decay to sound waves. The pressure signal detected with a transducer during a knocking combustion cycle will depend on the details of the end-gas ignition process, the combustion chamber geometry, and the location of the transducer in relation to the end-gas region.

The pressure variation across the cylinder bore, due to knock, is illustrated in Fig. 9.59 for a the disc-shaped combustion chamber with the spark plug located in the cylinder wall as shown in Fig. 9.59a. Figure 9.59b shows the pressure and temperature distribution across the combustion chamber due to the normal flame propagation process, at the time rapid release of the end-gas chemical energy suddenly occurs raising its pressure and temperature as shown. A strong pressure wave will now propagate to the right and an expansion wave to the left, as shown in the distance-time diagram in Fig. 9.59c. These waves reflect off the walls and interact. The pressures at the cylinder wall, in the end-gas region and on the opposite side of the chamber at the spark plug, develop as shown in Fig. 9.59d. In the end-gas region the pressure rises extremely rapidly when knock occurs, to a value considerably higher than that recorded on the opposite side of the chamber where the pressure rises more gradually. Standing waves are then set up and the amplitudes of the oscillations decay as the waves are damped out.

Figure 9.59 Illustration of how cylinder pressure distribution develops following knock onset: (a) schematic of disc-shaped combustion chamber at time of knock; (b) pressure and temperature across the diameter of a disc-shaped engine combustion chamber, before and after end-gas autoignition (assumed to occur very rapidly, at constant end-gas density); (c) schematic of shock and expansion wave pattern following end-gas autoignition on distance-time plot; (d) pressure variation with time at cylinder wall in the end-gas region, and at the opposite side of the cylinder.



The resonant acoustic modes of the combustion chamber define the frequencies of these pressure oscillations. Figure 9.60 shows some of these acoustic modes for a simple circular-cylindrical cavity (which approximates an engine combustion chamber) for which analytical solutions are known.⁶⁷ The first and second circumferential modes and the first radial and first axial modes of such a cylindrical cavity are depicted in Fig. 9.60a, b, c, and d, respectively. The dashed lines indicate nodes and the plus and minus signs identify the relative phases of the anti-nodes. The frequencies of the axial modes depend on crank angle, but the circumferential and radial modes do not. The lowest frequency (for automobile engines, about 6 kHz) is the first circumferential mode. The circumferential modes are always paired, and the paired nodal lines are orthogonal and for

a right circular cylinder have the same frequency. With actual engine cylinder-head geometries, there is an asymmetry and the two frequencies can differ by up to 10%. Linear combinations of the modes in a pair can be used to create a mode with the same shape and frequency as either, but with the node lines rotated. The majority of the pressure wave energy is in the lowest mode.

Figure 9.60 Schematic diagram of some of the acoustic modes of a right cylindrical cavity. Typical frequencies for a standard gasoline SI engine are shown.⁶⁸

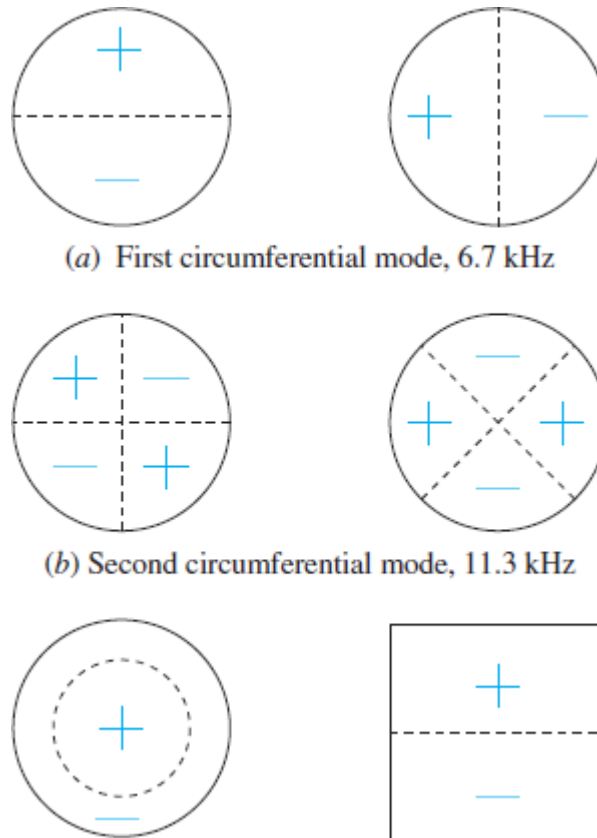
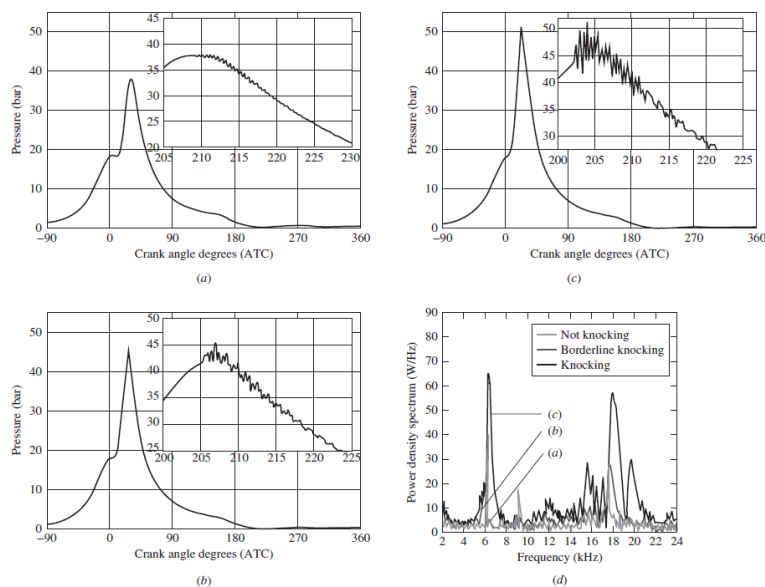


Figure 9.61 shows examples of cylinder pressure traces for individual cycles in a gasoline SI engine at fixed operating conditions, (1500 rev/min, high-load) at different spark timings. The engine had a pent-roof, four-valve, central spark-plug location cylinder head. The piston was flat topped. The bore was 83 mm, the compression ratio 9.8. Figure 9.61b shows a cycle at a spark timing corresponding to borderline knock. Figures 9.61a and c show cycles with spark timings retarded and advanced by 2° CA, respectively. Figure 9.61d shows the fast Fourier transform (FFT) power density spectrum for these three cycles. As the engine transitions through borderline knock, two ranges of frequencies are excited. The lower frequency is around 6 kHz and corresponds to the frequency of the first circumferential mode of the chamber ($f \approx a / (2B)$ where a is the sound speed in the burned gases, and B is the effective cylinder diameter). The single peak at this frequency is well defined, and the peak power density increases as knock becomes more severe. A second range of frequencies is observed between 15 kHz and 22 kHz, and there is no single well-defined peak. This could be the first axial mode whose frequency varies with piston position, and thus crank angle. It could be higher order circumferential modes.⁶⁶ The higher frequency modes do not contribute to the audible noise knock generates because these frequencies are above the range of human hearing.

Figure 9.61 Cylinder pressure traces for typical cycles of a gasoline [toluene reference fuel (95 RON)], SI engine operating at 1500 rev/min and high load, with a stoichiometric mixture, with a spark timing corresponding to borderline knock (b); also with spark timing retarded by 2° such that the engine is not knocking (a); and with spark timing advanced by 2° so it knocks more heavily (c). The region around peak pressure for each cycle is enlarged to show the extent of the pressure oscillations. (d) The fast Fourier transform of the three cycles shown.⁶⁶



In summary, the fundamental explanation of knock described above is based on the concept of local autoignition of the fuel-air mixture in the end-gas. *Autoignition* is the term used for a rapid combustion reaction which is not initiated by any external ignition source. In the basic combustion literature this phenomenon is sometimes called an "explosion." Autoignition of a gaseous fuel-air mixture occurs when the chemical energy released by reactions as thermal energy is larger than the thermal energy lost to the surroundings by heat transfer: as a result, the temperature of the mixture increases rapidly due to the exponential temperature dependence of the rates of the chemical reactions involved. The state at which such spontaneous ignition occurs is often characterized by the *self-ignition temperature*, and the resulting self-accelerating event where the gas pressure and temperature increase rapidly is termed a *thermal explosion*.

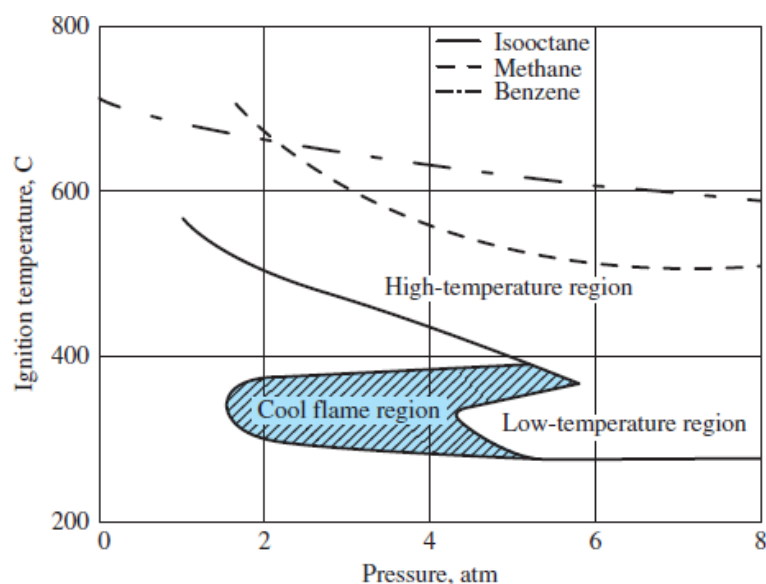
In complex reacting systems such as exist in combustion, the "reaction" is not a single- or even a few-step process; the actual chemical mechanism consists of a large number of simultaneous, interdependent reactions called *chain reactions*. In such chains there is an *initiating* reaction where highly reactive intermediate species or *radicals* are produced from stable molecules (fuel and oxygen). This step is followed by *propagation* reactions where radicals react with the reactant molecules to form products and other radicals to continue the chain. The process ends with *termination* reactions where the chain propagating radicals are removed. Some propagating reactions produce two reactive radical molecules for each radical consumed. These are called *chain-branching* reactions. When, due to chain-branching, the number of radicals increases sufficiently rapidly, the reaction rate becomes extremely fast and a *chain-branching explosion* occurs. While the terms thermal and chain-branching explosions have been introduced separately, in many situations the self-accelerations in temperature and radicals occur simultaneously and the two phenomena are combined.⁷⁶

Let us now turn to the autoignition chemistry of hydrocarbon-air mixtures. The process by which a hydrocarbon is oxidized can exhibit four different types of behavior, or a sequential combination of them, depending on the pressure and temperature of the mixture: slow reactions; single or multiple cool flames (slightly exothermic reactions); two-stage ignition (cool flame followed by a hot flame); single-stage ignition (hot flame). Slow reactions are a low-pressure, low-temperature (<200°C) phenomenon not normally occurring in engines. At 300 to 400°C one or more combustion waves can form, accompanied by faint blue light emission; the reactions are quenched, however, when only a small fraction of the reactants have reacted and the temperature rise is only tens of degrees. These are called *cool flames*. Depending on conditions and the fuel, a cool flame

may be followed by a "hot flame" or high-temperature explosion where the reaction accelerates rapidly after ignition. This is termed *two-stage ignition*. As the temperature of the mixture increases, a transition from two-stage to single-stage ignition occurs. While all hydrocarbons exhibit induction intervals which are followed by a very rapid reaction rate, some hydrocarbon compounds do not exhibit the cool flame or two-stage ignition behavior.

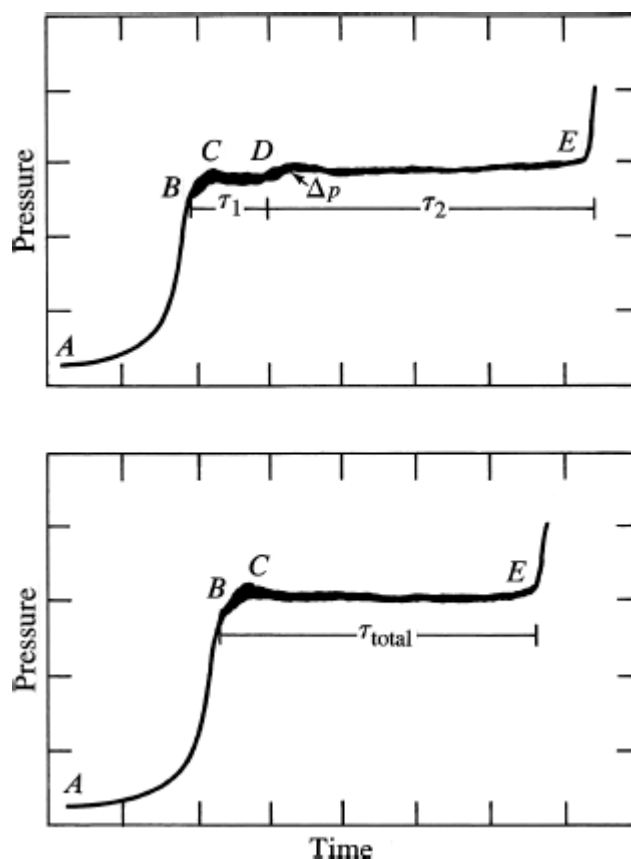
Figure 9.62 shows these ignition limits for isooctane, methane, and benzene. For isooctane, ignition in the low-temperature region is by a two-stage process: there is a first time interval before the cool flame appears and then a second time interval from the appearance of the cool flame to the hot-flame combustion process. Ignition in the high-temperature region is by a continuous one-stage process. Cool flame phenomena vary significantly with hydrocarbon structure. Normal paraffins give strong cool flames, whereas branched-chain paraffins are more resistant. Olefins give even lower luminosity cool flames with longer induction periods. Methane shows only the high-temperature ignition limit, as indicated by Fig. 9.62. Benzene, also, does not exhibit the cool flame phenomenon, and other aromatics give hardly detectable luminosity. It is thought that some compounds knock by a low-temperature two-stage ignition mechanism, some via a high-temperature single-stage ignition mechanism, and for some fuels both mechanisms may play a role.

Figure 9.62 Ignition diagrams for isooctane, methane, and benzene. Two-stage ignition occurs in the low-temperature region; the first stage may be a cool flame. Single-stage ignition occurs in the high-temperature region.

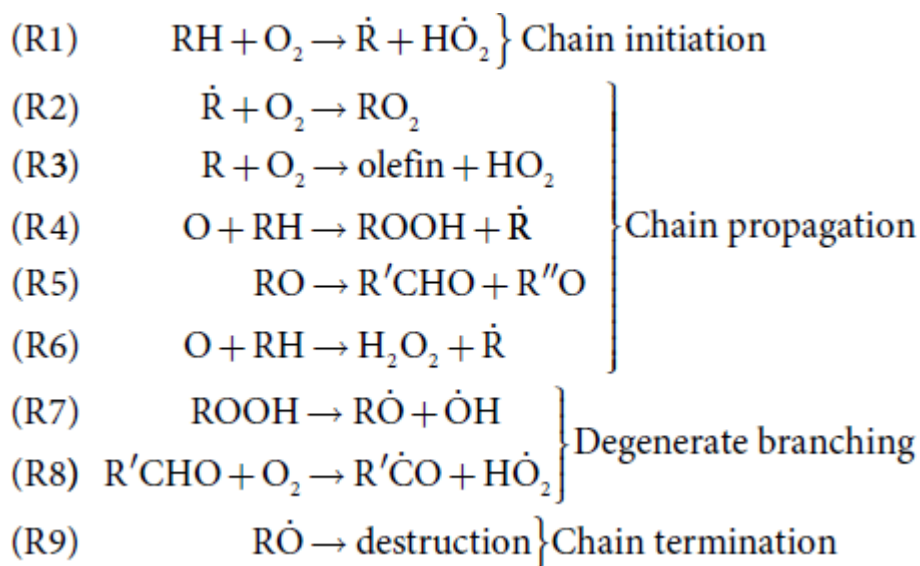


Examples of these two mechanisms in rapid-compression machine experiments, where a homogeneous isooctane-air mixture was compressed to different final conditions in a piston-cylinder apparatus and allowed to autoignite, are shown in Fig. 9.63. $A \rightarrow B \rightarrow C$ is the piston-motion-produced compression. The top trace shows a well-defined cool flame at D , preceding hot ignition at E . The lower trace, at a higher temperature, shows a single-stage ignition process.

Figure 9.63 Pressure records of the autoignition of isooctane-air mixture ($\phi = 0.9$) in a rapid compression machine. A $\rightarrow$ B $\rightarrow$ C is the compression process. Top: two-stage ignition (at D, then at E) at post-compression pressure of 1.86 MPa and temperature of 686 K. Bottom: single-stage ignition at E at post-compression conditions of 2.12 MPa and 787 K. Vertical scale: 690 kPa/division. Horizontal scale: 1 ms/division.⁷⁷



An outline of the basic hydrocarbon (RH) oxidation process due to Semenov is as follows:⁷⁶ it serves as a framework for discussing autoignition chemistry.



(9.65)

The dot denotes a radical; each prime denotes a free bond on the organic radical R.

Reaction R1 is slow and explains the induction period in hydrocarbon combustion. R2 is fast and of near-zero activation energy. R3 leads to olefins known to occur in the oxidation of saturated hydrocarbons. R4 and R5 yield the main intermediates. The degenerate branching comes about from the delay in decomposition of the reactive species in R7 and R8. As one radical is used up to form the reactants in R7 and R8, the multiple radicals do not appear until these reactants decompose.^{76,78}

The following indicates the relevance of the above mechanism to knock in engines. End-gas sampling studies have identified products of slow combustion reactions of isooctane; these principally include olefins, cyclic ethers, aldehydes ($R'CHO$), and ketones ($R''CO$).⁷⁹ Such studies have shown increasing concentrations of peroxides (predominantly H_2O_2 with traces of organic peroxides) with isoparaaffinic fuels which show two-stage ignition behavior. Higher temperature, single-stage ignition fuels such as benzene and toluene gave no detectable peroxide. Aldehydes and ketones have been measured in significant and increasing concentrations in motored engines where the peak cycle temperature was steadily increased. In motored engines, the occurrence of cool flames, the two-stage nature of the autoignition ignition process at intermediate compression temperatures, and the transition to a single-stage ignition process at very high compression ratios (i.e., motored-engine gas temperatures with a peak much higher than normal to simulate end-gas conditions in a firing engine) have also been demonstrated.⁸⁰

Two types of models of this autoignition process have been developed and used: (1) empirical induction-time correlations and (2) chemical mechanisms which embody many or all of the features of the "full" hydrocarbon oxidation process outlined in Eq. (9.65).

Induction-time correlations are derived by matching an Arrhenius function to measured data on induction or autoignition times, for given fuel-air mixtures, over the relevant mixture pressure and temperature ranges. It is then assumed that knock-initiating autoignition occurs when

$$\int_{t=0}^{t_i} \frac{dt}{\tau} = 1$$

(9.66)

where τ is the induction time at the instantaneous temperature and pressure for the mixture, t is the elapsed time from the start of the end-gas compression process ($t = 0$), and t_i is the time at autoignition. This equation^[2] can be derived by assuming that the overall rate of production of the critical species in the induction period chemistry, for a given mixture, depends only on the gas state and that the concentration of the critical species required to initiate autoignition is fixed (i.e., independent of the gas state).⁸¹

A number of empirical relations for induction time for individual hydrocarbons and blended fuels have been developed from fundamental or engine studies of autoignition. These relations have the form

$$\tau = Ap^{-n} \exp\left(\frac{B}{T}\right)$$

(9.67)

where A , n , and B are fitted parameters that depend on the fuel. The most extensively tested correlation is that proposed by Douaud and Eyzat.⁸²

$$\tau = 17.68 \left(\frac{ON}{100}\right)^{3.402} p^{-17} \exp\left(\frac{3800}{T}\right)$$

(9.68)

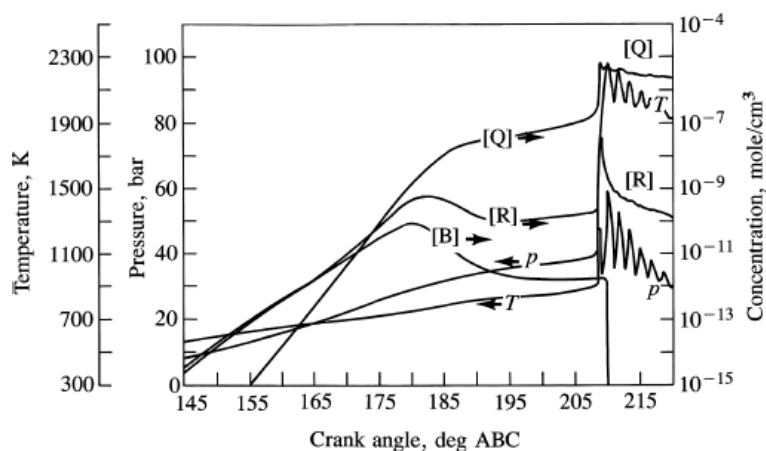
where τ is in milliseconds, p is absolute pressure in bar, and T is in kelvins. ON is the research octane number of the fuel (see [Sec. 9.6.3](#)). If the temperature and pressure time history of the end-gas during an individual cycle are known, Eqs. (9.64) and (9.66) together can be used to determine whether autoignition occurs before the normally propagating flame consumes the end-gas. Use of this approach for predicting knock onset can be surprisingly accurate.⁸³

An important requirement with any model of the end-gas autoignition process is characterizing the end-gas temperature. During intake, the combustion chamber walls are hotter than the entering gases: thus, heat is transferred from the walls to the in-cylinder mixture. Midway during the compression stroke, the mixture temperature rises to levels above the wall temperature and a thermal boundary layer will build up adjacent to the wall, as heat is now transferred to the wall. Unburned mixture away from the wall, however, will be compressed essentially adiabatically (see [Sec. 12.6.5](#)). In addition, any unburned mixture which for some portion of time during intake and compression has been in close proximity to the exhaust valve and piston (which run at higher temperatures than the water-cooled wall regions) could be at higher temperatures than adiabatically compressed mixture due to substantial heat transfer to the unburned mixture. Thus, the end-gas temperature is not uniform and the distribution of temperature is complex. Often, the *mean* unburned mixture temperature is used to characterize its state. Alternatively, the *core* temperature corresponding to adiabatic compression of mixture from conditions at the start of compression is used. In the absence of substantial heating by the exhaust valve and piston, the core temperature is a better representation of the maximum unburned mixture at any point in the cycle.

More complex and complete chemical models of the autoignition process have been developed for specific paraffinic hydrocarbon and alcohols, which are important components in real fuels. However, detailed models are not yet available for fuels with multiple components such as gasolines. A generalized kinetic model for hydrocarbon oxidation based on a degenerate branched-chain mechanism, known as the Shell model, has been developed and tested with some success. The model uses generic chemical entities representative of a variety of individual species which undergo a set of generalized reactions. This is justified by the broadly similar (though complex) ignition behavior of a variety of different fuel molecules and the similar kinetics exhibited by the organic radicals of the same type in the hydrocarbon oxidation process [Reaction set (9.65)].^{77, 84}

The Shell model is based on a generic eight-step degenerate chain-branching reaction scheme. The scheme involves the fuel (RH), oxygen, radicals formed from the fuel (R), products (P), intermediate product (Q), and degenerate-branching agent (B). The rate constants are either fixed at values consistent with the literature or fitted so that measured induction times (such as those illustrated in [Fig. 9.63](#)) are adequately predicted. An example of results obtained with this scheme is shown in [Fig. 9.64](#). It shows the calculated pressure, temperature, and species concentrations of R, Q, and B in the end-gas region of an operating spark-ignition engine, leading up to knock at 209° ABC. In this example, the autoignition model has been incorporated in a multidimensional model of the flow and flame propagation processes within the combustion chamber (see [Sec. 15.5](#)).⁸⁵

Figure 9.64 Pressure, temperature, and composition in the end-gas, before, during, and after autoignition, predicted by the Shell model in a spark-ignition engine combustion process. B, Q, R defined in text.⁸⁵



9.6.3. Fuel Factors

The tendency to knock depends on engine design and operating variables which influence end-gas temperature, pressure, and the time spent at high values of these two properties before flame arrival. Thus, for example, the tendency to knock is decreased through reductions in the end-gas temperature that follow from decreasing the inlet air temperature, or retarding the spark from MBT timing. However, knock is a phenomenon that is governed by both engine and fuel factors: its presence or absence in an engine depends significantly on the "antiknock" quality of the fuel.

Individual hydrocarbon compounds vary substantially in their ability to resist knock, depending on their molecular size and structure. Their tendency to knock has been characterized by the *critical compression ratio* of an engine: i.e., the compression ratio at which, under specified operating conditions, the specific fuel compound will exhibit incipient knock. Knocking tendency is related to molecular structure^[3] as follows:^{86,87}

Paraffins

- Increasing the length of the carbon chain increases the knocking tendency.
- Compacting the hydrocarbon molecule by incorporating side chains (thereby shortening the length of the basic chain) decreases the tendency to knock.
- This adds methyl groups (CH_3) to the side of the basic carbon chain, in the second from the end to the center position, reducing the number of weaker C-H bonds: see below.

Olefins

- The introduction of one double bond has little antiknock effect; two or three double bonds generally result in appreciably less knocking tendency.
- Exceptions to this rule are acetylene (C_2H_2), ethylene (C_2H_4), and propylene (C_3H_6), which knock much more readily than the corresponding saturated hydrocarbons.

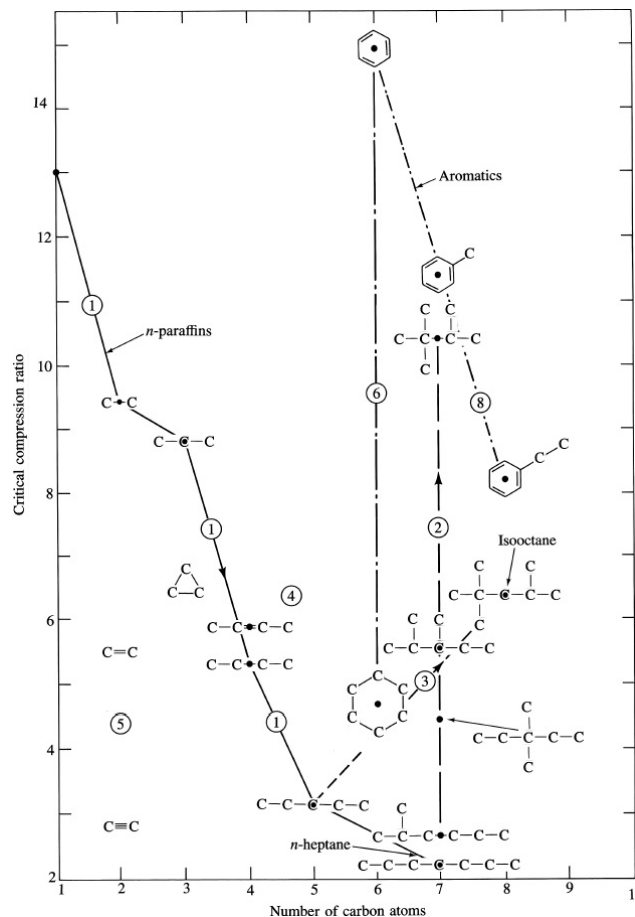
Napthenes and aromatics

- Napthenes have significantly greater knocking tendency than have the corresponding size aromatics.

- Introducing one double bond has little antiknock effect; two and three double bonds generally reduce knocking tendency appreciably.
- Lengthening the side chain attached to the basic ring structure increases the knocking tendency in both groups of fuels, whereas branching of the side chain decreases the knocking tendency.

Figure 9.65 identifies the magnitude of these trends on a plot of the critical compression ratio against the number of carbon atoms in the molecule. The strong dependence of knocking tendency on fuel molecular size and structure is apparent. The location of the various hydrocarbon compounds on Fig. 9.65 can be explained in part through the hydrogen to carbon atom bond arrangements for each molecule. The weakest hydrogen bonds (and thus the most easily broken as the hydrocarbon molecule starts to react by losing hydrogen atoms) are those between a hydrogen atom and a carbon atom with two C-to-C bonds. The *n*-heptane molecule (which has a critical compression ratio of 2.2) has ten of these weaker bonds. Isooctane (2,2,4-trimethylpentane), with only two of these bonds, has a critical compression ratio of 6.4, and resists autoignition well.

Figure 9.65 Critical compression ratio (for incipient knock at 600 rev/min and 450 K coolant temperature) as a function of number of carbon atoms in hydrocarbon molecule illustrating the effects of changes in molecular structure. (Developed from Lovell.⁸⁶)



Practical fuels are blends of a large number of individual hydrocarbon compounds from all the classes of hydrocarbons: alkanes (paraffins), cyclanes (naphthenes), alkenes (olefins), and aromatics (see Sec. 3.3, and alcohols). A more practical measure of a fuel's resistance to knock is obviously required: such a property is the fuel's *octane number*. It determines whether or not a fuel will knock in a given engine under given operating conditions: the higher the ON, the higher the resistance to knock. ON is not a single-valued quantity since it depends on engine design and operating conditions during testing. The standard ON scale is based on two hydrocarbons which define the ends of the scale. By definition, normal heptane ($n\text{-C}_7\text{H}_{16}$) has a value of zero and isooctane (C_8H_{18} , 2,2,4-trimethylpentane) has an ON of 100. These hydrocarbons were chosen because of the great difference in their ability to resist knock and the fact that isooctane had a higher resistance

to knock than the gasolines available at the time the scale was established. Blends of these two hydrocarbons to form *Primary Reference Fuels* define the knock resistance of intermediate ONs: for example, a blend of 10% *n*-heptane and 90% isooctane has an ON of 90. A fuel's ON is measured by determining the PRF, the blend of these two hydrocarbons that matches that fuel's knock resistance in a specific spark-ignition engine under the same operating conditions.

Several standard octane rating methods for fuels have been developed. Two of these—the research method (ASTM D-2699)^[4] and the motor method (ASTM D-2700)—are carried out in a standardized single-cylinder engine. In the motor method, the engine operating conditions are more severe; that is, the conditions are more likely to produce knock. In addition, road octane rating methods have been developed to define the antiknock quality of fuels in cars operated on the road or on chassis dynamometers. The engine used in the ASTM research and motor methods is the single-cylinder engine developed under the auspices of the Cooperative Fuel Research Committee in 1931—the CFR engine.^[5] This test engine is a robust four-stroke overhead-valve engine with an 82.6-mm (3.25-in.) bore and 114.3-mm (5-in.) stroke. The compression ratio can be varied from 3 to 30 while the engine is operating, with a mechanism which raises or lowers the cylinder and cylinder head assembly relative to the crankcase. The engine is equipped with several fuel reservoirs, with a selector valve, so the engine can be operated on any of the three fuels: two reference fuels (usually blends of *n*-heptane and isooctane) and the fuel being rated. The engine operating conditions of the research and motor methods are summarized in [Table 9.7](#). The test conditions were chosen to represent the engine operating regime where knock is most severe.

Table 9.7 Operating conditions for research and motor methods

	Research method	Motor method
Inlet temperature	52°C (125°F)	149°C (300°F)
Inlet pressure	Atmospheric	
Humidity	3.6–7.2 g H ₂ O/kg dry air	
Coolant temperature	100°C (212°F)	
Engine speed	600 rev/min	900 rev/min
Spark advance	13°BTC (constant)	19–26° BTC
		(varies with compression ratio)
Air/fuel ratio	Adjusted for maximum knock	

With the fuel under test, the fuel/air ratio is adjusted for maximum knock. The compression ratio is then adjusted to produce knock of a standard intensity. The level of knock obtained with the test fuel is bracketed by two blends of the reference fuels not more than two octane numbers apart (with one knocking more and one less than the test fuel). The octane number of the gasoline is then obtained by interpolation between the knock-meter scale readings for the two reference fuels and their octane numbers. For fuels below 100 ON, the *primary reference fuels* are blends of isooctane and *n*-heptane; the percent by volume of isooctane in the blend is the octane number. For fuels above 100 ON, the antiknock quality of the fuel is determined in terms of isooctane plus milliliters of the antiknock additive, tetraethyl lead, per U.S. gallon.^[6]

The octane ratings of several individual hydrocarbon compounds and common blended fuels are summarized in [App. D, Table D.4](#). Practically all fuels exhibit a difference between their research and motor octane numbers. The motor method of determining ON uses more severe operating conditions than the research method (higher inlet mixture temperature, higher speed, more advanced spark timing). Thus, the motor octane number (MON) is lower than the research octane number (RON). The numerical difference between these two octane numbers is called the *fuel's sensitivity*.

$$\text{Fuel sensitivity } S = \text{RON} - \text{MON}$$

(9.69)

Fuel sensitivity varies with the composition of the fuel, and thus the source of crude petroleum and the refining process details. The primary reference fuels themselves (mixtures of isooctane and *n*-heptane), by definition have the same octane numbers by both the research and motor methods. Since the primary reference fuels are paraffins, we would expect other paraffins to have little or no sensitivity. In contrast, olefins and aromatics have high sensitivity, as do alcohols. Generally, therefore, straight-run gasolines containing high percentages of saturated hydrocarbons have low sensitivity, while cracked or reformed gasolines containing larger percentages of unsaturated hydrocarbons have high sensitivity. Regular (U.S.) grade unleaded gasoline typically has a RON of at least 91 and a MON of about 83, giving a sensitivity of 8 to 10 octane numbers. Standard grade European gasolines have higher RON and MON values, 95 and 85, respectively.

Research and motor octane number fuel ratings are made in a single-cylinder engine run at constant speed, wide-open throttle, and fixed spark timing. These methods do not always predict how a fuel will behave in an automobile engine operated under a variety of speed, load, and weather conditions. Several methods of rating a gasoline in actual vehicles, either on the road or on chassis dynamometers which duplicate outdoor road conditions have, therefore, been developed. These methods determine the fuel's *road octane number*. More broadly, a measurable engine parameter related to autoignition^[7] can be correlated at fixed engine operating conditions with a fuel's knock resistance through⁸⁸

$$z = a(\text{RON}) + b(\text{MON}) + c$$

Note that *a*, *b*, and *c* are experimentally derived constants that depend on the engine used and its operating conditions (e.g., speed and load). More generally, the propensity of a fuel to autoignite and knock under real world conditions is quantified by its *octane index*, OI. The OI of a fuel is obtained in an engine operating under specified conditions, by determining the primary reference fuel composition that matches the knock onset behavior of the actual fuel. This OI can be related to the fuel's RON and MON using a weighting factor *K*:^{88, 89}

$$\text{OI} = (1 - K)\text{RON} + K \text{ MON} = \text{RON} - KS$$

(9.70)

K is assumed to be independent of the fuel, and depends only on the engine operating conditions. Note that if *K* = 0 the OI corresponds to the fuel's RON: *K* = 1 corresponds to OI = MON. In the United States, the OI posted on retail gasoline pumps is (RON + MON)/2 termed the *antiknock index*. Thus *K* = 0.5 is used to characterize a fuel's "on-the-road" ON.

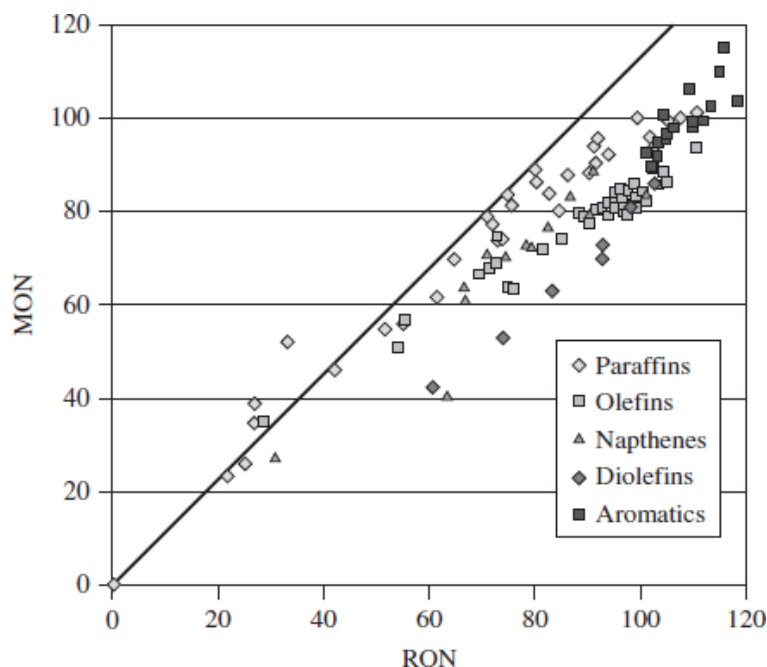
Engines, vehicles, and fuels have all changed significantly since the mid-1900s. Recent analyses have evaluated this octane weighting factor *K* over the past 60 years using data from vehicle-fleet studies where, usually, the engine-in-vehicle fuel octane requirement to avoid knock was determined from maximum vehicle acceleration tests. The basic RON test was adopted in 1930. It was found that it did not adequately predict the antiknock performance of fuels in actual use. The MON test, with its thermally more demanding conditions, was then developed to better reflect the actual antiknock behavior of fuels. Thus, around 1930, *K* was close to unity. A historical review⁹⁰ indicates that *K*, on average, has steadily decreased over time to about 0.5 in 1950, 0.25 in 1970, and is now close to zero (with slightly negative values for some engines and operating conditions). Thus RON is now a better predictor of actual fuel antiknock behavior than combinations of RON and MON, and is used in Europe for this purpose.

It has also become clear that gasolines with higher sensitivity *S* [see Eq. (9.69)] with the same RON value usually have better knock resistance (a higher OI) than fuels of lower sensitivity. The key to understanding this is the unburned mixture temperature behavior, often characterized by the engine's intake air temperature at knock onset. For the RON and MON tests

this air temperature is 52 and 149°C, respectively: for modern naturally-aspirated engines it is about 30°C, and for turbocharged engines about 40°C, significantly lower than the standard test values.

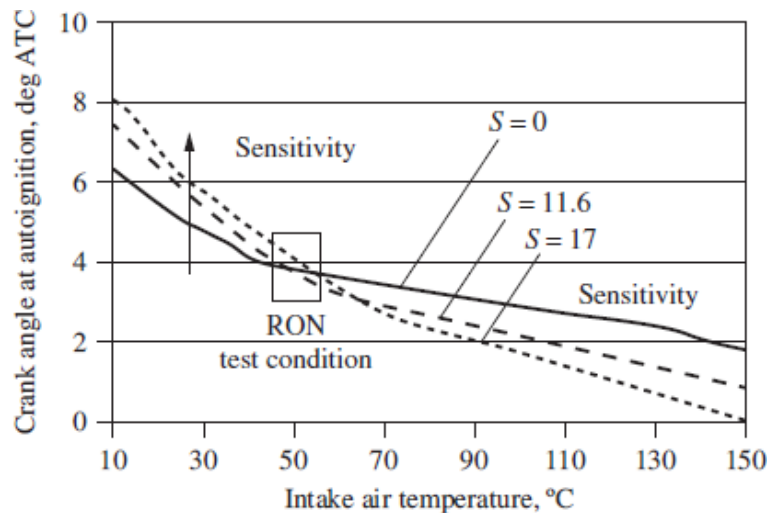
Most fuel components with high sensitivities have more complex molecular structures compared to paraffins. However, there is no simple relationship between sensitivity and the fuel's molecular structure. [Figure 9.66](#) plots MON values against RON values for several different classes of individual hydrocarbons. Non-sensitive fuels lie on the 45° line $\text{MON} \approx \text{RON}$. These tend to be paraffins. Sensitive fuels lie below this line—MON is lower than RON. The higher sensitivity fuels are olefins, aromatics, diolefins, and napthenes. Alcohols (not included in the plot) have high sensitivities.^[8] Note the large variations in RON and MON for each class of fuels.⁹¹

Figure 9.66 Motor octane number (MON) plotted versus research octane number (RON) for a variety of hydrocarbons including paraffins, olefins, napthenes, diolefins, and aromatics. Non-sensitive fuels are located on the diagonal black line and sensitive fuels are located below that line. Note there are large variations in RON and MON for each class of hydrocarbons.⁹¹



Studies of the effects of mixture temperature on the autoignition fuel chemistry explain why increasing fuel sensitivity raises the fuel's knock resistance in gasoline engines. Models with detailed chemical kinetic mechanisms for isooctane, n-heptane, toluene, and ethanol oxidation (with large numbers of species and reactions; see [Sec. 14.5.6](#)) were coupled with engine operating cycle simulation-generated end-gas temperature and pressure profiles to explore the autoignition behavior of several fuels of increasing sensitivity blended from the above compounds. The crank angle at which autoignition occurred was computed as a function of intake air temperature. Results for three 96 RON fuels with different sensitivities are shown in [Fig. 9.67](#). The air temperature in the RON test is 52°C and in the MON test is 149°C. Modern engines operate with air temperatures in the 30 to 40°C range. With air temperature below the RON test value (where the behavior of the different sensitivity fuels is closely comparable), the highest sensitivity fuel ($S = 17$) shows greater resistance to knock: the zero sensitivity fuel (PRF 96) has the lowest knock resistance. At air temperatures above 50°C or so, the sensitivity trend reverses—the least sensitive fuel shows the greatest knock resistance.⁹¹

Figure 9.67 The computed crank angle of autoignition (in degrees ATC) as a function of intake air temperature for three fuels of different sensitivity, S : a primary reference fuel (RON 96), a toluene reference fuel (78% toluene, 22% n -heptane mixture with RON = 96 and $S = 11.6$) and a 63.5% ethanol, 36.5 n -heptane blend with RON = 96 and $S = 17$. The RON octane rating test condition is indicated by the box.⁹¹



For modern gasoline engines, the maximum end-gas temperature under knock-limited conditions is in the 800 to 950 K range. At the low end of this range, in the low-temperature combustion region where the autoignition chemistry is slower, higher sensitivity fuels perform better. At the higher end of this range ($T_{u,max} \geq 900$ K), the high-temperature combustion region where chemistry is significantly faster, less sensitive fuels perform slightly better at resisting knock. In the center of the transition regime, in between (the critical knock onset region with today's gasolines) which corresponds to the "negative" temperature gradient ignition regime (see Fig. 9.62), autoignition induction times of fuels with different sensitivities but the same RON are similar. At end-gas temperatures below this mid-point (≈ 825 K) the steeper induction time dependence on temperature of the more sensitive gasolines, as the autoignition chemistry shifts from the negative-temperature region to the low temperature region, results in slower chemistry and thus greater knock resistance for the more sensitive fuels.⁹¹

Refiners and automobile manufacturers are interested in the octane requirement of engines in vehicles on the road. The *octane requirement* (usually abbreviated to OR) of an engine or vehicle-engine combination is defined as the minimum fuel octane number that will resist knock, at MBT spark timing, over a given engine speed and load range.

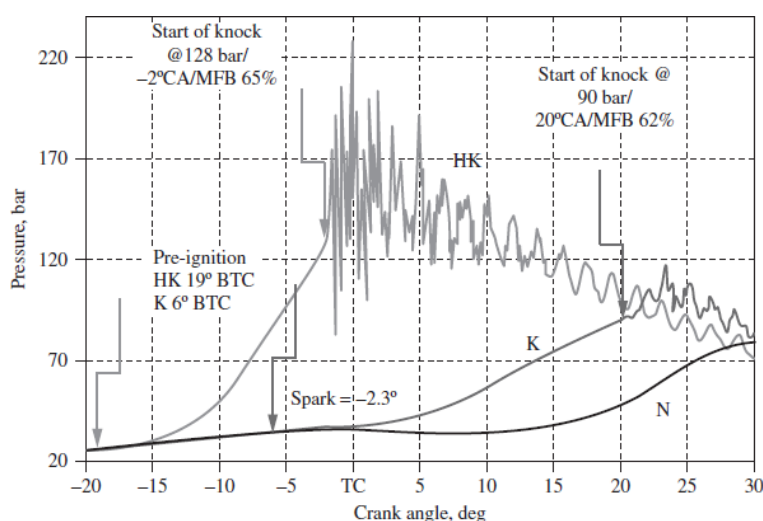
9.6.4. Sporadic Preignition and Knock

As in-cylinder pressures have risen due to increasing engine compression ratios and turbocharging, an additional knock phenomenon has been observed, initiated by a seemingly random preignition process. Sporadically, a local region of the in-cylinder unburned mixture spontaneously ignites toward the end of the compression stroke well before the spark discharge occurs. This preignition initiates combustion well before the intended start of the flame propagation process at time of spark discharge. As a result, in-cylinder gas pressures (and temperatures) rise sooner, attaining much higher values close to TC piston position than with a normal combustion event. The "end gas" under these conditions can autoignite rapidly leading to extremely heavy knock.^[9] Due to this preignition, the combustion process in that engine cycle is very "over-advanced." This phenomenon has been broadly observed as in-cylinder pressures have risen, and its avoidance necessitates constraining the pressure levels at which gasoline spark-ignition engines (especially turbocharged engines) can operate.

The phenomenon is sporadic, occurring once every 20,000 to 40,000 cycles, and lasts for only one or a few sequential cycles. The knock intensity produced (peak to peak amplitude of the knock pressure oscillations) can be up to one- to two-hundred bar, much higher than the knock intensity resulting from spark knock.

Figure 9.68 shows an example of this preignition leading to heavy knock phenomenon.⁹² In-cylinder pressure records from a turbocharged gasoline engine, operating at high-load, 2000 rev/min and 2 bar intake pressure, with significantly retarded spark timing (2.3° CA before TC) are shown. A normal non-knocking cycle (N) without preignition has a low peak (and late) pressure (80 bar) as shown that is close to knock onset conditions for this engine. The next cycle shown (K) is one with fairly heavy knock (KI ~ 24 bar) with preignition occurring at 6° CA BTC. The third cycle (HK) shows preignition occurring much earlier at 19° CA BTC leading to "super knock" with a knock intensity (peak to peak) of about 150 bar.

Figure 9.68 Three individual in-cylinder pressure cycles in a turbocharged gasoline engine operating at high load (2 bar manifold air pressure) and 2000 rev/min with retarded spark timing (2.3° CA before TC). A normal combustion cycle (N) with trace knock has a low (~ 80 bar) and late (30° ATC) peak pressure. A mid-intensity knocking cycle (K), where preignition occurs at 6° BTC, has a pressure at knock onset of 90 bar, and a heavy knock (with early preignition at 19° BTC) cycle, HK, reaches a pressure of 128 bar at knock onset. MFB is mass fraction burned.⁹²



The pressures and unburned mixture (end-gas) temperatures at the autoignition point that produces knock for these three cycles are shown in Table 9.7. The difference in knock intensities between these three cycles cannot be explained by the different unburned mixture mass fractions available for autoignition. Rather they depend on the process and thus speed by which the autoignition pressure waves propagate outward, causing other unburned gas regions to autoignite and establish the strong pressure waves of these intense knock phenomena. The normal cycle (where knock just occurs) is a relative slow deflagration pressure wave process, the normal knock initiation process. The mid-knock intensity cycle is a developing-detonation pressure wave process where the chemical energy release the wave initiates is driving and intensifying the pressure wave. The very high knock intensity cycle is the result of a strong developing detonation wave where the high pressure wave initiates an increasingly large chemical energy release which drives the wave faster and increases its magnitude.

A plausible explanation for the preignition process that initiates this super-knock phenomenon is that it results from vaporized oil in the in-cylinder fuel-air mixture. Oil vapor when mixed with gasoline and air has a lower self-ignition temperature than a gasoline/air mixture and thus will autoignite at lower unburned mixture temperatures. Oil is present as a thin film on the cylinder liner, and between the liner and piston. Gasoline, impinging on the liner absorbs oil from the film, and when the piston travels upward during the compression stroke, fuel-air mixture can collect in the crevice volume above and below the top ring in the piston lands. Thus oil-fuel mixture in vapor or liquid form can accumulate in this region: eventually enough mixture may accumulate and exit the crevice, and be convected into the unburned in-cylinder gases away from the combustion chamber wall. That mixture in that specific cycle (or few cycles, until this accumulated oil plus fuel mixture is gone), then spontaneously ignites. The origin of this early preignition phenomenon is the formation of a local region of oil vapor, gasoline vapor, air, and burned gas (likely during compression) that autoignites due to its low octane rating. This early initiation of combustion results in much higher-than-normal unburned mixture pressures and temperatures that then autoignite the end-gas when only half to

two-thirds of the in-cylinder gasoline and air has burned, resulting in very heavy knock as seen in [Fig. 9.68](#). Experiments designed to test this explanation (such as reducing crevice volumes to increase the frequency of this preignition initiated super-knock phenomenon, use of less volatile oils) provide confirmation.⁹³

9.6.5. Knock Suppression

Knock suppression would allow SI engine compression ratios, and boost levels in turbocharged engines, to be increased thereby improving engine performance and efficiency. Strategies for suppressing or delaying knock onset include use of fuel additives, high-octane fuel blending compounds, and engine modifications, as summarized below. Note that fuel additives are widely used to counteract fuel-related deterioration in engine performance. Detergent additives are used to ensure that the intake system (fuel injectors, intake valves) and the combustion system (spark plugs, combustion chamber walls) remain free of deposits and function as intended. Corrosion inhibitors are included in gasolines to form a thin protective film on metal surfaces to prevent water/moisture induced corrosion. Antioxidants are added to fuels to improve their storage capability.⁵¹

9.6.5.1. Knock-Suppressing Fuel Additives and Blending Compounds

Modern gasolines contain a number of chemical additives designed to improve fuel quality. Gasoline base stocks are also blended with oxygen containing organic compounds to raise their octane rating. The most effective antiknock agents are lead alkyls. Tetraethyl lead (TEL), $(\text{C}_2\text{H}_5)_4\text{Pb}$, was first introduced in 1923. Tetramethyl lead (TML), $(\text{CH}_3)_4\text{Pb}$, was introduced in 1960. Since TML boils in the mid-range of a gasoline (110°C) (whereas TEL boils at the high end, 200°C), the introduction of TML permitted better distribution of octane amongst the cylinders of an engine. In 1959, a manganese antiknock compound (methylcyclopentadienyl manganese tricarbonyl), known as MMT, was introduced as a supplementary antiknock agent for TEL. It is also an antiknock agent in its own right. In the 1970s, unleaded gasolines were introduced and now dominate. Two factors influenced the elimination of lead alkyls: concern about the toxicological aspects of lead in the urban environment, and the use of catalytic devices for emission control that are poisoned by lead. MMT is no longer used as an antiknock additive in unleaded gasoline due to its role as a deposit in plugging exhaust catalytic converters.

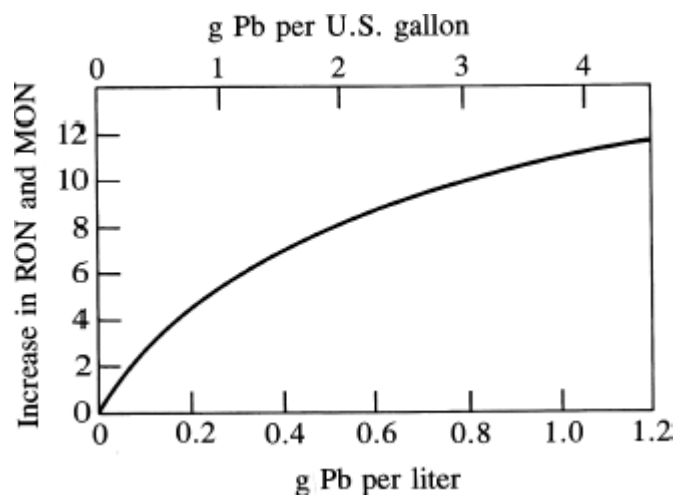
Two (sometimes three) grades of (unleaded) gasoline with different octane ratings are usually available at refueling stations. In the United States the lower and higher rated fuels are labeled "regular" and "premium" with antiknock indexes of 87 and 93. In Europe average gasoline octane numbers are two to three ONs higher.^[10] Gasolines with higher octane ratings than these could be produced at the refinery, but historically the cost of the additional refining required to do this, relative to the in-vehicle benefit has been judged marginal.

This refining limitation on petroleum-derived (unleaded) gasolines has increased interest in other methods of boosting the octane rating of gasolines. The use of oxygenates—alcohols and ethers—as gasoline extenders, which due to their excellent antiknock quality increase the fuel's octane rating, is common. A brief review of the mechanism by which these additives and blending compounds improve the knock resistance of gasoline follows.

Lead is the most effective antiknock element known. In the form of lead alkyls, it is stable and fuel-soluble. It is generally agreed that the alkyls decompose before they exert their antiknock action. The decomposed material—lead oxide, PbO —either as a vapor or as a fog (a dispersion of fine particles), inhibits the preflame chain-branching reactions which lead to autoignition of the fuel-air charge, thus slowing the overall reaction rate. However, lead has little effect on two-stage ignition until after the cool flame. The individual alkyls vary in antiknock behavior as a function of fuel composition and combustion conditions. The average effect of various amounts of TEL in a large number of regular gasolines is summarized in [Fig. 9.69](#): the effectiveness of each successive increment of TEL, added steadily, decreases. The addition of about 0.8 g lead per liter (3 g Pb per U.S. gallon), provides an average gain of about 10 octane numbers, though lead's effectiveness varies with the chemical composition of the base fuel. When a gasoline containing a lead alkyl is burned in a spark-ignition engine, it produces nonvolatile combustion products. These deposit on the walls of the combustion chamber and on the spark plug, causing lead-fouling of the spark plug electrodes and tracking across the plug insulator, and corrosion of the hot exhaust

valve. Commercial lead-containing antiknock fluids, therefore, contain scavenging agents—combinations of ethylene dibromide and ethylene dichloride—which transfer the lead oxides that would otherwise deposit into volatile lead-bromine compounds which are largely exhausted with the combustion gases.

Figure 9.69 Gasoline octane number increase resulting from use of antiknock additive tetraethyl lead. Increase varies with fuel composition: average values shown.



The use of oxygenates as extenders or substitutes for gasoline is increasing. In some cases, this is because the oxygenate can be produced from nonpetroleum sources (e.g., ethanol from biomass) and thus may offer strategic or economic benefits. In other cases, the good antiknock blending characteristics of oxygenates can aid in meeting octane quality requirements where regulations prevent or limit lead alkyl use. Several oxygenates have been used as automotive fuels: those of greatest interest are ethanol (C_2H_5OH), methanol (CH_3OH), tertiary butyl alcohol (TBA, C_4H_9OH), and methyl tertiary butyl ether (MTBE).

Ethanol is being made from biomass, especially sugar cane and corn grain. [Table 9.8](#) lists the knock-suppressing characteristics of these compounds and their physical and chemical characteristics relative to gasoline. The *blending octane value* of these oxygenated compounds is not necessarily the same as the compound's research octane number and depends on the composition (and thus the octane rating) of the gasoline with which the compound is blended. MTBE-gasoline blends have good stability in the presence of water, and MTBE has little effect on fuel vapor pressure, and materials compatibility. TBA is susceptible to water extraction and loss. Ethanol is technically attractive as a high-octane supplement or substitute for gasoline and is used in 5 or 10% blends in gasolines, as well as a stand-alone fuel as E85 (nominally 85% ethanol, 15% gasoline).

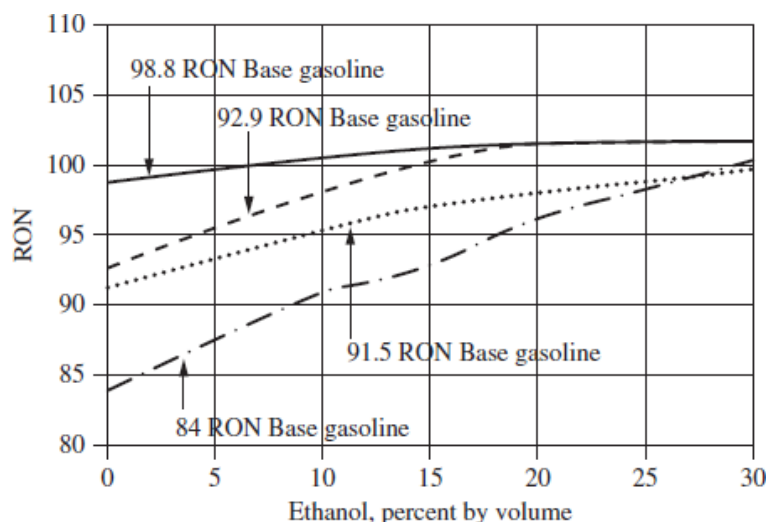
Table 9.8 End-gas conditions at knock onset (see [Fig. 9.68](#))

Cycle	End-gas autoignition			Mass fraction burned
	CA	p_{a_i} bar	T_{a_i} K	
Knock onset N	25° ATC	70	800	Unknown
Knock onset K	20° ATC	90	850	0.62
Early preignition/heavy knock HK	2° BTC	128	926	0.65

Fuel blending of these oxygenated hydrocarbon compounded with a base gasoline can be illustrated as follows with ethanol. Because ethanol raises the octane and volatility of blends, it can be blended with a lower octane, reduced volatility base

gasoline to ensure that the final gasoline-ethanol blend meets the fuel's specifications and regulations. Figure 9.70 shows how the blending of ethanol with various base gasolines (referred to as blendstock for oxygenate blending, BOB) increases the research octane number of the final fuel. Note that as the RON of the blendstock increases, the slope of each increasing RON line decreases. The blending RON of ethanol derived from the linear portion of the curves in Fig. 9.70 (up to about 15% ethanol) decreases from about 145 to 125 as the base gasoline octane increases from 84 to 98 RON. Note also that a saturation effect occurs with ethanol percentages above about 30%.

Figure 9.70 Research octane number as a function of volumetric content of ethanol for four different base gasolines with RON values of 84, 91.5, 92.9, and 98.8.⁸³



Methanol, because it can be made from natural gas, coal, or biomass materials, has potential as an additive and alternative fuel. Its high blending-octane quality (about 130 RON, 95 MON), when used in low-concentration (~5%) methanol-cosolvent-gasoline blends, means it could be used to upgrade low octane gasoline. Problems with methanol include poor solubility in gasoline in the presence of water, high toxicity, an energy content per unit volume about half that of gasoline; high latent heat of vaporization (which can contribute to poor drivability), incompatibility with many commonly used metals and elastomers, blending effects on gasoline volatility which may force the displacement of significant volumes of butane. Some of these problems can be partially reduced by using cosolvents such as TBA or isobutanol. Use of ethanol and methanol as stand-alone fuels in specially designed engines permits advantage to be taken of their high octane rating via higher compression ratios and boost levels. As stand-alone fuels, however, their significantly lower energy content than gasoline is a disadvantage.

The octane requirement of an engine-vehicle combination usually increases somewhat during use, due primarily to the buildup of deposits on the cylinder head and piston crown. While these deposits increase the engine's compression ratio modestly, their largest effect is to increase the temperature of the surface of the combustion chamber in contact with the inducted mixture. This *increases* heat transfer to the fresh mixture during induction and *decreases* heat transfer from the unburned charge during compression. End-gas temperatures are therefore higher, thus increasing the likelihood of knock. As the combustion chamber deposits stabilize (over some 15,000 km of driving), the engine's octane requirement typically increases by up to about 5 octane numbers before it stabilizes. The advent of cleaner gasolines and improved deposit controlling additives is reducing this deposit related octane requirement increase.

9.6.5.2. Engine Knock Control Strategies

Autoignition in the end-gas ahead of the flame, the initiating process for engine knock, depends on the maximum unburned mixture temperature and pressure, and on the time available before flame arrival, for a given fuel octane rating. Here we focus on those spark-ignition engine variables that affect the unburned mixture's temperature and pressure, and time available for autoignition to develop, that can be changed and thus affect knock onset. As is well known, knock constrains engine

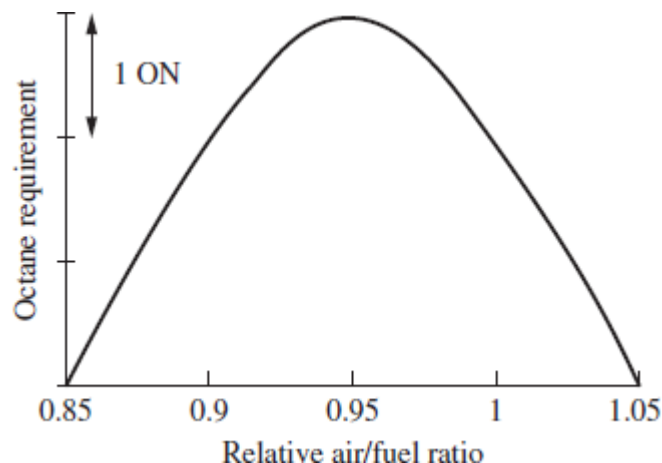
compression ratio in all spark-ignition engines, and also the boost level in turbocharged SI engines, and thus limits the maximum engine mean effective pressure and fuel conversion efficiency.

The engine variables that determine knock onset, and that can be used to control knock are: spark timing (relative to MBT timing), relative air/fuel ratio, intake air temperature, coolant temperature, dilution of the mixture with cooled exhaust gas recirculation, and cooling of the in-cylinder charge through evaporation of the liquid fuel. A review of the literature supplemented by studies of these effects in more modern engines⁹⁵ provides a guide to the effects of several important parameters on the octane requirement (the minimum octane number fuel needed to avoid knock) of gasoline-fueled spark-ignition engines. The results are given in [Table 9.10](#).

Spark retard is the most commonly used control approach to reduce the engine's octane requirement. Retarding the phasing of the combustion process by retarding spark from its MBT timing, decreases the peak cylinder pressure (which occurs later ATC), and as a consequence, decreases the maximum unburned mixture temperature. Many studies have indicated that each crank-angle degree of spark retard decreases the fuel RON required to avoid knock by 1 octane number. Since the loss in torque, bmep, and efficiency is modest for up to 5 degrees of spark retard from MBT timing (see [Fig. 9.16](#)), and only becomes significant as retard increases beyond about 10 degrees, spark retard is widely used. Also its effect on knock avoidance is immediate.

Spark-ignition engines are often operated about 20% fuel rich of stoichiometric at wide-open throttle to increase torque, lower the thermal loading on the exhaust catalyst system (and the turbine in turbocharged engines), and reduce the engine's octane requirement. [Figure 9.71](#) shows the effect of relative air/fuel ratio changes on an engine's octane requirement, based on a review of the relevant literature.⁹⁵ The octane requirement is highest at around 0.95 relative air/fuel ratio, about 5% fuel-rich of stoichiometric. The octane requirement decreases on either side of this maximum—both for richer mixtures, and for less-rich mixtures through stoichiometric to lean mixtures. The sensitivity of these decreases is about 3 ON per 0.1 relative air/fuel ratio, or 2 ON per unit air/fuel ratio. Note that while WOT operation with rich mixtures has historically been used in gasoline engines for many years, it does result in high air pollutant emissions. For this reason, many modern engines operate stoichiometric at WOT so the three-way catalyst continues to be effective. However, this results in higher exhaust gas and catalyst temperatures (due to HC, CO, and H₂ oxidation within the catalyst), and higher turbine inlet temperatures.

Figure 9.71 Effect of relative air/fuel ratio (λ) on engine octane requirement. $\lambda < 1$, rich mixture.⁹⁵



Unburned in-cylinder mixture temperature is a critical variable in the autoignition process leading to knock. It is affected by the intake air temperature, the amount and temperature of the in-cylinder residual gas, and coolant temperature (which affects cylinder head, piston and liner temperatures, and thus heat transfer to the unburned mixture—which is into the mixture through the first half of the compression process). [Table 9.9](#) gives the octane sensitivity of these temperature-related variables: the most important is the intake air temperature.

Table 9.9 Oxygenate properties

	Methanol	Ethanol	TBA	MTBE	Gasoline
Research octane no. (RON)	106	109	~105	~114	91–98
Blending octane no. (RON)	130	129	~110	~125	NA
Weight percent oxygen	50	35	22	18	NA
Stoichiometric (A/F)	6.5	9.0	11.2	11.7	14.5
Specific gravity	0.796	0.794	0.791	0.746	0.74
Lower heating value, MJ/kg	20.0	26.9	32.5	35.2	42
Latent heat of vaporization, MJ/k [†]	1.16	0.84	0.57	0.34	0.35
Boiling temperature, °C	65	78	83	55	27–227
TBA, tertiary butyl alcohol; MTBE, methyl tertiary butyl ether.					
Note that blending octane number values depend on the octane of the base blendstock gasoline and amount of oxygenate. Source: Table D.4 ; Motor Gasolines Technical Review; Chevron Products Company, <i>Motor Gasolines Technical Review</i> , Chevron Corporation, 1996 and 2009. ⁹⁴					
[†] Depends on temperature.					

Direct fuel injection into the cylinder helps suppress knock. The OR decreases by about 4 relative to the equivalent PFI value. As the fuel spray evaporates with DI, its heat of vaporization is drawn from the in-cylinder charge, thus cooling that charge. With gasoline, the maximum in-cylinder charge cooling effect is just under 20 K. Studies that compared the charge cooling effect of gasoline direct in-cylinder fuel injection with intake-port fuel injection indicate that for a gasoline (without any ethanol) the difference in charge cooling is 13 to 14 K which is about 70% of the maximum gasoline evaporative cooling effect. This suggests that some in-cylinder fuel-evaporative charge cooling occurs with port-fuel-injection systems—about a quarter of the maximum cooling effect. Direct injection of gasoline would be expected to realize close to the maximum cooling effect, since close to full evaporative of the fuel occurs prior to combustion.

Intake manifold pressure is also an important variable. Increased in-cylinder pressure accelerates the autoignition chemical reactions. Data indicate that the octane requirement sensitivity to manifold pressure is higher at WOT for naturally-aspirated engines (some 3 or 4 ON per 10 kPa change in MAP) than for turbocharged engines with MAP values approaching 2 bar absolute (about 2 ON per 10 kPa change in MAP). (In part this may be due to maximum pressure constraints in TC engines.)

Note that the octane requirement sensitivities in [Table 9.10](#) will depend somewhat on engine design and operating details and also on whether the engine is turbocharged. In the engine development and design process the many variables that affect the knock requirement of the engine are traded-off against each other so as to optimize part-load engine brake fuel-conversion efficiency, and wide-open-throttle torque (or bmep), for the intended fuel octane rating.

Table 9.10 Effect of engine operating variables on engine's octane requirement

Variable	Octane number requirement change per unit change in variable
Compression ratio	+ 5 ONR/unit r_c increase
Spark retard	– 1 ONR/1° retard
Air/fuel ratio*	– 2 ONR/unit A/F
Intake air pressure†	+ 2–4 ONR/10 kPa increase
Exhaust back pressure	+ 1 ONR/30 kPa increase
Intake air temperature	+ 1 ONR/7 deg C increase
Coolant temperature	+ 1 ONR/10 deg C increase
Dilution cooled EGR	– 3–4 ONR/10% mass dilution
Type of fuel injection	– 4 ONR/DI relative to PFI
Altitude	– 1.4 ONR/300 m rise (0–1800 m elevation)
	– 2.5 ONR/300 m rise (1800–3600 m elevation)
*Octane requirement peaks at 5% (relative air/fuel ratio) rich of stoichiometric and decreases for richer (and leaner) mixtures.	
† Higher value for naturally-aspirated gasoline engines; lower value for turbocharged engines.	

One strategy used to suppress knock onset is the use of *cooled exhaust gas recycle* (EGR). This is mostly used in turbocharged gasoline engines where knock significantly constrains engine compression ratio and boost level. The use of EGR at high load dilutes the engine's in-cylinder fuel/air/burned gas mixture, which normally reduces the torque a given displacement engine produces. However, this knock suppression effect allows a higher compression ratio and boost level to be used which more than offsets the charge dilution effect.

Dilution of the unburned mixture with EGR has the following effects on engine operation. The inert component of the fuel/air/burned gas mixture—nitrogen plus CO_2 plus H_2O , with a stoichiometric mixture—is increased thus slowing down the flame development and rapid burning angles (see Fig. 9.15) due to the lower laminar flame speed (see Sec. 9.3.3). The slower burn requires earlier spark timing to obtain maximum brake torque, but with MBT spark timing, the impact on knock onset is modest.

The explanation of the knock suppressing effect of cooled EGR is based on the fundamental physical variables—unburned mixture pressure and temperature—in the autoignition induction-time correlation, Eq. (9.68). The highest values of T_u and p depend on both the piston-motion-produced compression process coupled with the combustion-produced pressure rise which additionally compresses the unburned mixture.⁹⁶ To a good approximation the unburned mixture is compressed adiabatically by the prevailing pressure, thus

$$\frac{\dot{T}_u}{T_u} = \frac{(\gamma_u - 1)}{\gamma_u} \frac{\dot{p}}{p}$$

(9.71)

The pressure rise rate may be obtained from the "heat" or chemical-energy release First Law analysis of the engine combustion process where the engine cylinder is treated as a uniform one-zone system [Eq. (9.27) with $V_{cr} \rightarrow 0$]:

$$\frac{\dot{p}V + p\dot{V}}{(\gamma - 1)} = \dot{Q} - p\dot{V}$$

which, when combined with Eq. (9.71) gives

$$\frac{\dot{T}_u}{T_u} = \frac{(\gamma_u - 1)}{\gamma_u} \left[(\gamma - 1) \frac{\dot{Q}}{pV} - \gamma \frac{\dot{V}}{V} \right]$$

(9.72)

Note that γ is the average ratio of specific heats for the assumed uniform gases in this one-zone energy release approach:

$\dot{Q}$ is the net heat release,

$$\dot{Q}_{ch} - \dot{Q}_{ht}$$

The effect of increasing EGR on γ_u , the ratio of specific heats of the unburned mixture, is shown in Fig. 9.72: when EGR is increased from 0 to 36%, the $(\gamma_u - 1)/\gamma_u$ factor in Eqs. (9.71) and (9.72) increases by 2.5%. However, with a stoichiometric mixture and constant fuel mass in the cylinder, as the amount of cooled exhaust gas recirculated increases, the in-cylinder pressure increases since, with increasing EGR for fixed fuel mass, the intake manifold pressure must increase. Cycle simulation results show that *prior* to spark, the higher γ_u increases the unburned mixture temperature; but that after combustion starts, the higher pressures decrease the first term in the square brackets in Eq. (9.72)^[11] and more than offset the effect of increasing in $(\gamma_u - 1)/\gamma_u$. So peak unburned mixture temperatures are reduced by cooled EGR. Thus the autoignition induction time is longer and knock onset is suppressed. Figure 9.73 shows the change in peak unburned mixture temperature as a function of EGR for two different loads—one relevant to naturally-aspirated, the other to turbocharged engines. The decrease in peak end-gas temperature is 5 K at 10% EGR and 12 K at 20% EGR.⁹⁶

Figure 9.72 Ratio of unburned mixture specific heats (c_p/c_v) as a function of dilution and gas temperature.⁹⁷

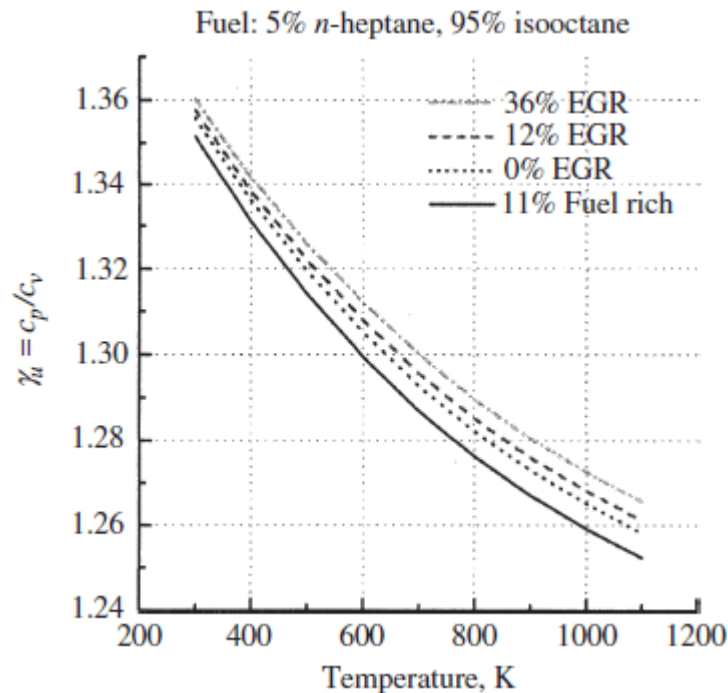
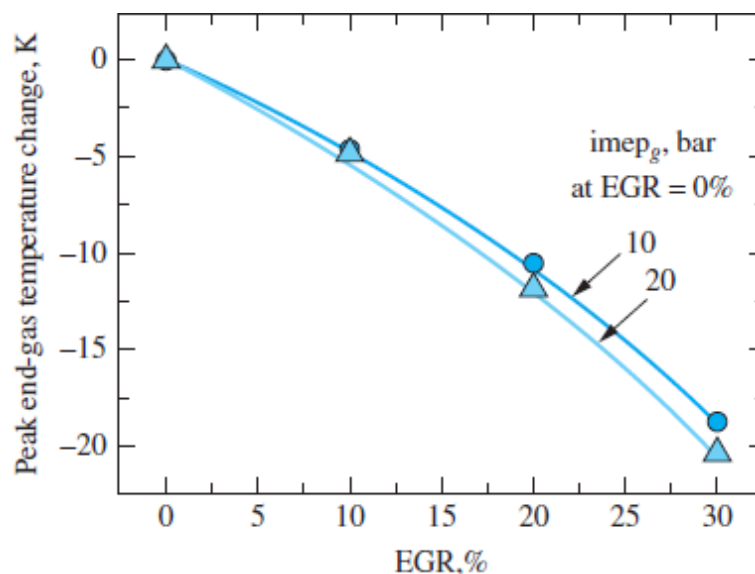


Figure 9.73 Change in maximum end-gas (unburned mixture) temperature as a function of amount of cooled EGR: at constant fuel mass flow rate. Two curves are for $imep_g = 10$ and 20 bar at EGR = 0%.⁹⁶



Use of EGR for knock suppression in turbocharged gasoline engines requires that the recycled gases are cooled in a heat exchanger and fed into the intake via a control valve upstream of the compressor. Since the exhaust gases contain unburned hydrocarbons and fine particulates with some 12% water, as well as trace amounts of lubricating oil, deposit build-up and corrosion in these EGR system components are issues.

[1] Note that during flame propagation, ρ_u/ρ_b is about 4: thus the volume of the end-gas region corresponding to say 20% of the mixture unburned is usually crescent shaped, and its location is defined by the flame propagation process.

-
- [2] Often called the Livengood and Wu autoignition integral after its originators.⁸¹
- [3] See [Sec. 3.3](#) for a review of hydrocarbon structure and its nomenclature.
- [4] ASTM denotes American Society for Testing and Materials; the letter and number define the specific testing code.
- [5] The Cooperative Fuel Research Committee is now the Coordinating Research Council, Inc.
- [6] The octane number of the fuel is calculated from $ON = 100 + 28.28T/[1.0 + 0.736T + (1.0 + 1.472T - 0.0352I_6T^2)^{1/2}]$, where T is milliliters of tetraethyl lead per U.S. gallon. Tetraethyl lead, $(C_2H_5)_4Pb$, contains 64 weight percent lead; 1 ml of TEL contains 1.06 grams of lead.
- [7] For example, z can be the knock limited spark advance or knock limited torque.
- [8] Ethanol, RON = 109, MON = 90, $S = 19$; methanol, RON = 109, MON = 89, $S = 20$.
- [9] This phenomenon is sometimes called "super, or mega knock."
- [10] Research octane numbers are a more accurate predictor of a fuel's resistance to knock. In the United States, regular and premium gasolines have RON values of close to 92 and 98, respectively. In Europe (where RON rather than AKI is used to label gasoline or petrol), the standard and premium fuels have RON values of 95 and 98.
- [11] Q remains essentially unchanged.

9.7. PROBLEMS

9.1. The table gives relevant properties of three different spark-ignition engine fuels. The design and operating parameters of a four-cylinder 1.6-liter displaced volume engine are to be optimized for each fuel over the engine's operating load and speed range. You may assume that for each engine-fuel combination, the gross indicated fuel conversion efficiency and imep at any operating condition are given by 0.85 times the fuel-air cycle efficiency at those conditions. Also, assume that for every five research octane number increase above 95 (the gasoline value) in fuel antiknock rating, the compression ratio can be increased by one unit. For gasoline, the engine compression ratio is 9, so if the octane number of the fuel is 100, a compression ratio of 10 can be used.

- At part-throttle operation—at an intake pressure of 0.5 bar and a speed of 2500 rev/min—estimate the gross indicated fuel conversion efficiency and specific fuel consumption for each engine-fuel combination. Each engine-fuel combination operates at the lean limit given.
- At the appropriate equivalence ratio for maximum power, with 1 bar inlet pressure at the same speed (2500 rev/min), the volumetric efficiencies are as shown. Explain these volumetric efficiency values. Each fuel is fully vaporized in the inlet manifold and the inlet mixture temperature is held constant for all fuels. Manifold and valve design remains the same.
- Estimate the ratios of the maximum indicated mean effective pressure, at the same conditions as in (b), of the natural gas and methanol fueled engines to the maximum imep of the gasoline engine.
- If the methanol engine burns 33% faster than the gasoline engine (i.e., its combustion event takes three-quarters of the time), sketch a *carefully drawn qualitative* cylinder pressure versus crank angle curve for the two engines, from halfway through the compression stroke to halfway through the expansion stroke. Put both curves on the same graph. Conditions are as in (b). Show the motored *and* firing pressure curves for each engine. The spark timing should be set for maximum brake torque for each engine. Show the location of spark timing, the location of maximum cylinder pressure, and the approximate location of the end of combustion.

Fuel properties

Fuel	Formula	Research octane number	λ at lean operating limit	Stoichiometric air/fuel ratio	Heating value, MJ/kg	η_v
Natural gas	CH ₄	120	1.4	17.2	50	0.78
Methanol	CH ₃ OH	105	1.25	6.4	20	0.75
Gasoline	(CH ₂) _n	95	1.1	14.9	44	0.85

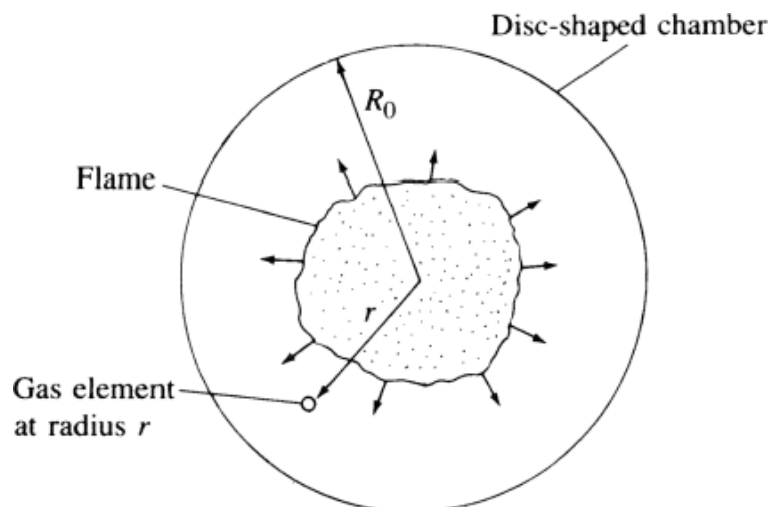
λ = relative air/fuel ratio.

9.2. The figure shows the flame propagating radially outward from the center of a disc-shaped combustion chamber. Combustion in such a device has many features in common with spark-ignition engine combustion. The chamber diameter is 10 cm and the height is 1.5 cm. For this problem, the flame can be thought of as a thin cylindrical sheet. Its radius increases approximately linearly with time. The volume of the chamber is constant. The fuel is a typical hydrocarbon fuel; the mixture is stoichiometric; the initial temperature is room temperature.

- Sketch qualitative but *carefully proportioned* graphs of the following quantities versus time from the start of combustion to the end of combustion:

1. The ratio of actual pressure in the chamber to the initial pressure
 2. The ratio of average density of the gas *ahead* of the flame to the initial density
 3. The ratio of the average density of the gas *behind* the flame to the initial density
- b. On a qualitative but *carefully proportioned* graph of r/R_0 versus time show how the radial positions of gas elements, initially at r/R_0 : 0, 0.5, and 1.0 before combustion, change during the combustion process as the flame propagates radially outward from the center ($r = 0$) to the outer wall ($r = R_0$).

Note: Accurate numerical calculations are not required to answer this question. You will be graded on the amount of physical insight your diagrams and the supporting brief explanations of your logic communicate. You should write down any equations or approximate numerical values for relevant quantities that help explain your reasoning.



9.3. In a spark-ignition engine, a turbulent flame propagates through the uniform premixed fuel-air mixture within the cylinder and extinguishes at the combustion chamber walls.

- a. Draw a carefully proportioned qualitative graph of cylinder pressure p and mass fraction burned x_b as a function of crank angle θ for $-90^\circ < \theta < 90^\circ$ for a typical SI engine at wide-open throttle with the spark timing adjusted for maximum brake torque. Mark in the crank angles of spark discharge, and of the flame development period (0 to 10% burned) and end of combustion, on both p and x_b versus θ curves relative to the top-center crank position.
- b. Estimate approximately the fraction of the cylinder volume occupied by burned gases when the mass fraction burned is 0.5 (i.e., halfway through the burning process).
- c. A simple model for this turbulent flame is shown on the left in the figure below. The rate of burning of the charge dm_b/dt is given by

$$\frac{dm_b}{dt} = A_f \rho_u S_T$$

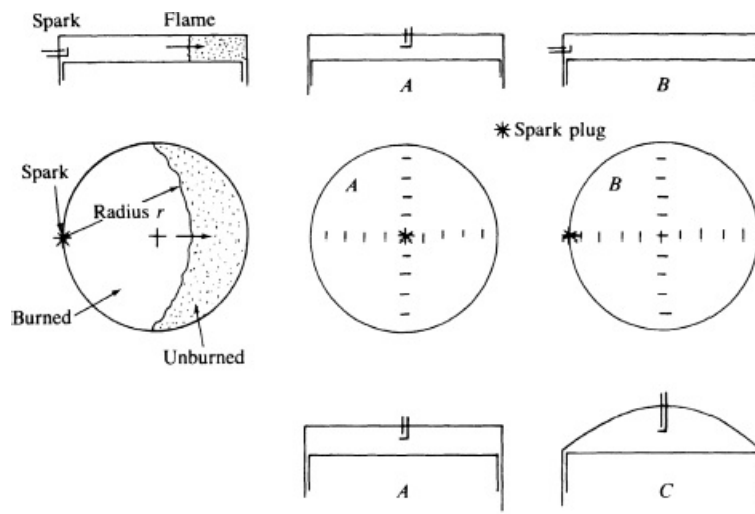
where A_f is the area of the flame front (which can be approximated by a portion of a cylinder whose axis is at the spark plug position), ρ_u is the unburned mixture density, and S_T is the turbulent flame speed (the speed at which the front moves relative to the unburned mixture ahead of it). The rate of mass burning is influenced therefore by combustion chamber geometry (through A_f) as well as those factors that influence S_T (turbulent intensity, fuel/air ratio, residual gas fraction, and EGR). Compare combustion chambers A and B shown in the figure. Sketch the *approximate* location of the flame front when 50% of the mass has been burned. (A careful qualitative sketch is sufficient; however, provide a quantitative justification for your sketch.) Sketch the mass fraction burned versus crank angle curves for these two combustion chambers on the same graph, each with its spark timing set for maximum brake torque. You may assume

the value of S_T is the same for A and B.

d. Compare combustion chambers A and C in the figure which have the same flame travel distance but have different chamber shapes. Which chamber has

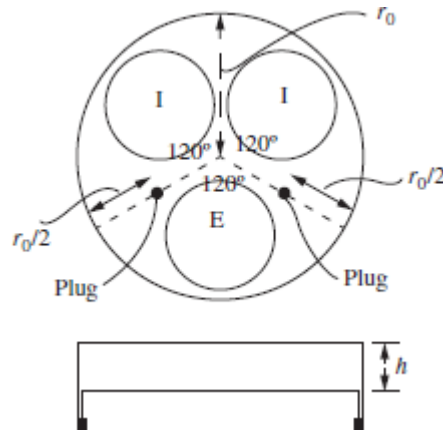
1. The faster rate of mass burning during the first half of the combustion process?
2. The faster rate of mass burning during the second half of the combustion process?
3. The more advanced timing for maximum brake torque?

Explain your answers.



9.4. Compare the three-valve two-plug combustion chamber shown, with a four-valve chamber with one spark plug on the cylinder axis. Make appropriate simple assumptions about the flame shape, as it grows outward from the spark(s).

- a. For the four-valve center plug, plot flame envelope area (flame outer boundary) $A_f/2\pi r_0$ against r_f/r_0 through the combustion process. On the same graph, sketch carefully the equivalent curve for the three-valve two-plug design assuming both plugs fire at the same time: simple geometric calculations are required to short out the physics.
- b. Plot the mass fraction burned curves versus CAD for these two combustion chambers on a qualitative but carefully proportioned graph from 45 CAD BTDC to 45 CAD ATDC. Show on the graph the spark discharge location, the TDC, the location of maximum in-cylinder pressure and the approximate end of combustion, with the spark timing for both chambers set at MBT timing.
- c. With the two-plug chamber, the burn rate can be slowed down by firing one plug later than the other. Estimate approximately the mass fraction burned after which further delaying the second plug firing will have no impact on the burn rate.



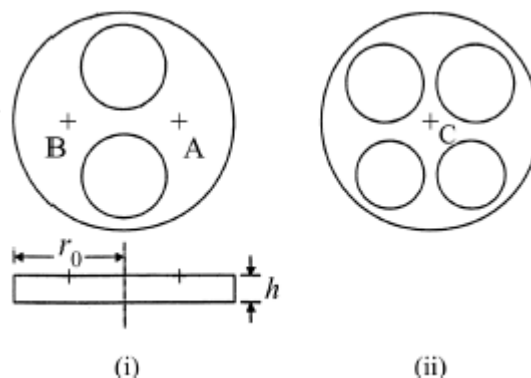
9.5. In the disk-shaped combustion chamber of a spark-ignition engine, the flame propagates radially outward from the spark plug at the center (on the cylinder axis). The flame can be modeled as a thin smooth sheet at the location within the wrinkled turbulent flame, which on average separates the unburned and burned gases.

- Derive an equation for the mass fraction burned $x_b = m_b/m$ as a fraction of r_f/r_0 (r_f = flame radius and r_0 = cylinder radius) and the density ratio $r_d = \rho_u/\rho_b$. Plot this curve for an appropriate value of r_d .
- Draw a carefully proportioned graph of x_b versus CAD for the engine operating at part-load with MBT spark timing (30 CAD BTC), indicating the approximate mass fraction burned at TC and 10 CAD ATC. Note that x_b at $\theta_{p \max}$ is about 0.75.

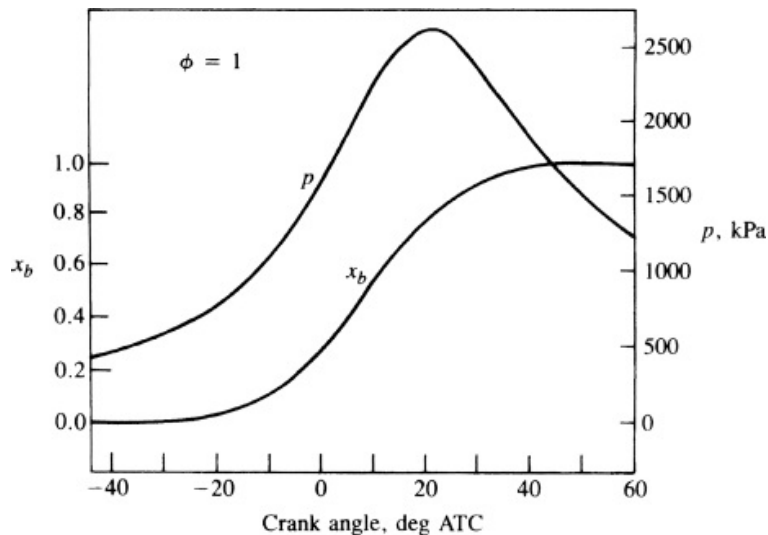
Make an appropriate estimate of $p_{\max}$. The pressure in the cylinder at IVC (50 CAD ABC) is the inlet pressure, i.e., 0.5 bar. The cylinder volume ratio, volume at IVC to volume at $p_{\max}$ is 7.3. The mixture is stoichiometric and the compression ratio is 10. Make reasonable assumptions of any other quantities you may need.

9.6. The cylinder head of a two-valve-per-cylinder spark-ignition engine with a single spark plug [at A, $r_0/2$ from the cylinder axis, in Fig. (i)] is to be modified to increase its burn rate. The options are to add a second spark plug at location B, or use a 4-valve head with a central spark plug [Fig. (ii)]. Assume that the clearance height h is constant, and is the same for both chambers (they are disk-shaped chambers).

- Assuming that each flame propagates outward in all directions at constant velocity from each spark plug, plot a carefully proportioned graph of flame outer (spherical) envelope area A_f divided by $(2\pi r_0 h)$ ($2\pi r_0 h$), versus flame radius r_f/r_0 for the original one plug, the two plug, and the center plug options. R_0 is the cylinder radius. Identify 1.0 on each axis. Sketches and simple geometric calculations are helpful; complex calculations are not required.
- Based on these graphs, explain how the optimum spark timing will differ for these three chambers. For the original two-valve chamber (spark plug at A), the optimum spark timing is 30° before TC and the 50% burned point is 10° after TC. Will the two-plug chamber (A and B), and center plug chamber C, have more advanced or less advanced timing? Explain your logic. Assume the local flame velocity for each chamber is constant and the same for all chambers.



9.7. The graph gives the pressure-crank angle curve for a spark-ignition engine running at $\phi = 1.0$. The mass fraction burned x_b is also shown. Estimate the temperature of the reactants at a number of crank angles and plot a graph of T_u versus θ . Assume the reactants in the cylinder are at 333 K and 1 bar pressure at the start of compression. It is necessary to make simplifications in order to do this and you should explain clearly what other assumptions you make; they should be compatible with [Prob. 9.8](#) below.



9.8. If the combustion takes place progressively through the burning of a large number of very thin zones of gas (see [Fig. 9.7](#)) and there is no mixing between the zones, determine:

- the temperature at -30° , just after combustion, of the zone which burns at -30° ;
- the temperature at 0° , just after combustion, of the zone which burns at 0° ;
- the temperature at $+30^\circ$, just after combustion of the zone which burns at $+30^\circ$;
- the temperature of the products in these three zones at $+30^\circ$.

Plot your results versus crank angle to show whether there is a spatial distribution of temperature in the cylinder after combustion.

Note: Each thin unburned gas element burns at essentially constant pressure and is subsequently compressed and/or expanded. Use charts ([Chap. 4](#)) or an equilibrium computer code. The unburned gas state is given by [Prob. 9.7](#).

9.9. Explain why SI engine torque varies, at fixed speed and inlet mixture conditions, as the spark timing is varied from very advanced (e.g., 60° before TC) to close to TC. What is the "best" spark timing? Explain how it varies with engine speed and load.

9.10.

- Explain the causes of the observed variation in cylinder pressure versus crank angle and imep in spark-ignition engines, cycle-by-cycle.
- What impacts do these cylinder pressure variations have on engine operation?

9.11.

- Describe briefly what occurs when a spark-ignition engine "knocks."
- SI engine knock is primarily a problem at wide-open throttle and lower engine speeds. Explain why this is the case.

- c. With a knock sensor, the normal knock control strategy is to retard the spark timing when knock is detected, until knock no longer occurs. Explain why this strategy is effective and is preferred over other possible approaches (e.g., throttling the inlet, adding EGR).
- d. In a knocking engine, the crank angle at which autoignition occurs and the magnitude of the pressure oscillations which result vary substantially, cycle-by-cycle. Suggest reasons why this happens.

9.12. Knock in spark-ignition engines is an abnormal combustion condition. Almost everyone who rides a motorcycle or drives a car experiences this phenomenon at some time, and usually changing into a lower gear will take the engine away from this condition. Use your experience of what changes in other variables do, and consult this text and other sources to complete a table with the dependent variables shown at the top of the columns.

The independent variables are: speed, compression ratio, chamber surface/volume ratio, spark plug distance from cylinder axis, percent EGR, inlet mixture temperature, inlet mixture pressure, (F/A) , wall temperature, air swirl, squish motion, fuel octane number. In these columns, show the corresponding influence on the dependent variables by a "+" for an increase and a "-" for a decrease. Show the effect of increase in engine system independent variables on: cylinder pressure and temperature, flame speed, total burn time, autoignition induction period, tendency to knock. Provide in the extreme right-hand column brief comments to explain your answers.

Effect of increase in engine independent variable	Dependent variables						
	Cylinder		Flame speed	Total burn time	Induction period	Tendency to knock	Explanation
	Pressure	Temperature					
Speed, rev/min Etc.							

9.13.

- a. The electrical energy stored in a typical ignition system coil is 50 mJ. Almost all this energy is transferred from the coil during the glow discharge phase. If the glow discharge lasts for 2 ms, use the data in Fig. 9.39 to estimate the glow discharge voltage and current.
- b. Only a fraction of this energy is transferred to the fuel-air mixture between the spark plug electrodes. Estimate the energy transferred during the breakdown and glow phases of the discharge, using the data in Fig. 9.39 and Table 9.5.
- c. Overall about one-tenth of the coil energy is transferred to the fuel-air mixture. What fraction of the cylinder contents' chemical energy ($m_f Q_{LHV}$) does this correspond to at a typical part-load condition ($p_i = 0.5$ atm, $\phi = 1.0$)? Assume 500 cm³ per cylinder displaced volume. If the average burned gas temperature within the flame kernel just after spark is 3500 K and the cylinder pressure is 6 atm, what radius of kernel has fuel chemical energy equal to the electrical energy transferred to the kernel?

9.14. An approximate way to calculate the pressure in the end-gas just after knock occurs is to assume that all the end-gas (the unburned gas ahead of the flame) burns instantaneously at constant volume. We assume that the inertia of the burned gases prevents significant gas motion while the end-gas autoignites.

For the pressure data in figure in Problem 9.7, assume autoignition occurs at 10 crank angle degrees after the top-center position. Determine the maximum pressure reached in the end-gas after knock occurs. From the mass fraction burned and approximate average burned gas conditions at this time, estimate the volume occupied by the end-gas as a fraction of the cylinder volume just before autoignition occurs. Use the charts (Chap. 4) or an equilibrium gas composition computer code.

9.15. At spark timing (30° BTC) in an automobile spark-ignition engine (with bore = stroke = 85 mm and $r_c = 9$) at 2000 rev/min, operating on gasoline, the cylinder pressure is 7.5 atm and the mixture temperature is 650 K. The fuel-air mixture

is stoichiometric with a residual gas fraction of 8%. The rapid burning angle $\Delta\theta_b$ (10 to 90% mass burned) is 35° . Estimate (a) the mean piston speed; (b) the average flame travel speed based on $\Delta\theta_b$ (the spark plug is located 15 mm from the cylinder axis); (c) the turbulence intensity at TC [see Eq. (8.25)]; (d) the laminar flame speed at spark; (e) the turbulent burning speed S_b at TC using Eqs. (9.56) and (9.58) (appropriate firing and motored pressures are given in Fig. 9.4a); (f) the mean expansion speed of the burned gases u_b at TC. Discuss briefly the relative magnitudes of these velocities.

9.16. The following combustion chamber design changes increase the mass burning rate in a spark-ignition engine at fixed compression ratio, bore, speed, and inlet mixture conditions. Explain how each change affects the burning rate.

- Reducing the amount of EGR.
- Using two spark plugs per cylinder instead of one.
- Generating swirl within the cylinder using a mask on the cylinder head (see Fig. 8.4).
- Using a combustion chamber with higher clearance height near the spark plug and a more central plug location.

9.17. The following factors contribute to cycle-by-cycle variations in the combustion process of an internal combustion engine, and hence in each cycle's cylinder pressure profile and torque pulse. Provide a *brief* description of why variations in each factor result in variations in the combustion process and how that causes variations in cylinder pressure cycle-by-cycle.

Causes of cycle-by-cycle combustion variations:

- Flame kernel motion
- Electrical energy deposited in spark-generated kernel
- Heat losses from kernel to spark plug
- Local turbulence characteristics near plug
- Amount of fuel/air/EGR fed to cylinder
- Local mixture composition near plug
- Large scale features of the in-cylinder flow
- Average turbulence in the cylinder

9.8. REFERENCES

1. Nakamura, H., Ohinouye, T., Hori, K., et al.: "Development of a New Combustion System (MCA-JET) in Gasoline Engine," SAE paper 780007, *SAE Trans.*, vol. 87, 1978.
2. Rassweiler, G. M., and Withrow, L.: "Motion Pictures of Engine Flames Correlated with Pressure Cards," *SAE Trans.*, vol. 83, pp. 185–204, 1938. Reissued as SAE paper 800131, 1980.
3. Nakanishi, K., Hirano, T., Inoue, T., and Ohigashi, S.: "The Effects of Charge Dilution on Combustion and Its Improvement—Flame Photograph Study," SAE paper 750054, *SAE Trans.*, vol. 84, 1975.
4. Beretta, G. P., Rashidi, M., and Keck, J. C.: "Turbulent Flame Propagation and Combustion in Spark Ignition Engines," *Combust. Flame*, vol. 52, pp. 217–245, 1983.
5. Amann, C. A.: "Cylinder-Pressure Measurement and Its Use in Engine Research," SAE paper 852067, 1985.
6. Lavoie, G. A., Heywood, J. B., and Keck, J. C.: "Experimental and Theoretical Investigation of Nitric Oxide Formation in Internal Combustion Engines," *Combust. Sci. Technol.*, vol. 1, pp. 313–326, 1970.
7. Rassweiler, G. M., and Withrow, L.: "Flame Temperatures Vary with Knock and Combustion-Chamber Position," *SAE Trans.*, vol. 36, pp. 125–133, 1935.
8. Lavoie, G. A.: "Spectroscopic Measurement of Nitric Oxide in Spark-Ignition Engines," *Combust. Flame*, vol. 15, pp. 97–108, 1970.
9. LoRusso, J. A., and Tabaczynski, R. J.: "Combustion and Emissions Characteristics of Methanol, Methanol-Water, and Gasoline-Methanol Blends in a Spark-Ignition Engine," SAE paper 769019, *Proc. Eleventh Intersociety Energy Conversion Engineering Conference*, Lake Tahoe, NV, pp. 122–132, Sept. 12–17, 1976.
10. Lancaster, D. R., Kreiger, R. B., and Lienesch, J. H.: "Measurement and Analysis of Engine Pressure Data," SAE paper 750026, *SAE Trans.*, vol. 84, 1975.
11. Davis, R. S., and Patterson, G. J., "Cylinder Pressure Data Quality Checks and Procedures to Maximize Data Accuracy," SAE paper 2006-01-1346, 2006.
12. Gatowski, J. A., Balles, E. N., Chun, K. M., Nelson, F. E., Ekchian, J. A., and Heywood, J. B.: "Heat Release Analysis of Engine Pressure Data," SAE paper 841359, *SAE Trans.*, vol. 93, 1984.
13. Chun, K. M., and Heywood, J. B.: "Estimating Heat-Release and Mass-of-Mixture Burned from Spark-Ignition Engine Pressure Data," *Combust. Sci. Technol.*, vol. 54, pp. 133–144, 1987.
14. Cheung, H. M., and Heywood, J. B.: "Evaluation of a One-Zone Burn-Rate Analysis Procedure Using Production SI Engine Pressure Data," SAE paper 902749, *SAE Trans.*, vol. 102, 1993.
15. Ayala, F. A., Gerty, M. D., and Heywood, J. B.: "Effects of Combustion Phasing, Relative Air-Fuel Ratio, Compression Ratio, and Load on SI Engine Efficiency," SAE paper 2006-01-0229, 2006.
16. Hires, S. D., Tabaczynski, R. J., and Novak, J. M.: "The Prediction of Ignition Delay and Combustion Intervals for a Homogeneous Charge, Spark Ignition Engine," SAE paper 780232, *SAE Trans.*, vol. 87, 1978.
17. Ayala, F. A., and Heywood, J. B.: "Lean SI Engines: The Role of Combustion Variability in Defining Lean Limits," SAE paper 2007-24-0030, 2007.
18. Stein, R. A., Chou, T., and Lyjak, J. C., "The Combustion System of the Ford 5.4L 3-Valve Engine," *Proc.* Sept. 22–25, 2003, Ann Arbor, MI, 2003.

19. Gatowski, J. A., Heywood, J. B., and Deleplace, C.: "Flame Photographs in a Spark-Ignition Engine," *Combust. Flame*, vol. 56, pp. 71–81, 1984.
20. Witze, P. O.: "The Effect of Spark Location on Combustion in a Variable-Swirl Engine," SAE paper 820044, *SAE Trans.*, vol. 91, 1982.
21. Heywood, J. B.: "Combustion Chamber Design for Optimum Spark-Ignition Engine Performance," *Int. J. Vehicle Design*, vol. 5, no. 3, pp. 336–357, 1984.
22. Wirth, M., Keller, P., and Peters, N.: "A Flamelet Model for Premixed Turbulent Combustion in SI Engines," SAE paper 932646, 1993.
23. Maly, R. R., and Herweg, R.: "Spark Ignition and Combustion in Four-Stroke Gasoline Engines," in *Flow and Combustion in Reciprocating Engines*, C. Arcoumanis and T. Kamimoto (eds), Ch. 1, pp. 1–66, Springer-Verlag, Berlin Heidelberg, 2009,
24. Gillespie, L.: "Aspects of Laminar and Turbulent Burning Velocity Relevant to SI Engines," SAE paper 2000-01-0192, 2000.
25. Merdjani, S., and Sheppard, C. G. W.: "Gasoline Engine Cycle Simulation Using the Leeds Turbulent Burning Velocity Correlations," SAE paper 932640, 1993.
26. Ziegler, G. F. W., Zettlitz, A., Meinhardt, P., Herweg, R., Maly, R., and Pfister, W.: "Cycle-Resolved Two-Dimensional Flame Visualization in a Spark-Ignition Engine," SAE paper 881634, *SAE Trans.*, vol. 97, 1988.
27. Abraham, J., Williams, F. A., and Bracco, F. V.: "A Discussion of Turbulent Flame Structure in Premixed Charges," SAE paper 850345, in *Engine Combustion Analysis: New Approaches* P-156 1985.
28. Pischinger, S., and Heywood, J. B.: "A Study of Flame Development and Engine Performance with Breakdown Ignition Systems in a Visualization Engine," SAE paper 880518, *SAE Trans.*, vol. 97, 1988.
29. Spicher, U., and Backer, H.: "Correlation of Flame Propagation and In-Cylinder Pressure in a Spark Ignited Engine," SAE paper 902126, 1990.
30. Keck, J. C.: "Turbulent Flame Structure and Speed in Spark-Ignition Engines," *Proceedings of the Nineteenth International Symposium on Combustion*, The Combustion Institute, Pittsburgh, pp. 1451–1466, 1982.
31. Namazian, M., Hansen, S. P., Lyford-Pike, E. J., Sanchez-Barsse, J., Heywood, J. B., and Rife, J.: "Schlieren Visualization of the Flow and Density Fields in the Cylinder of a Spark-Ignition Engine," SAE paper 800044, *SAE Trans.*, vol. 89, 1980.
32. Winklhofer, E., Philipp, H., Fraidl, G., and Fuchs, H., "Fuel and Flame Imaging in SI Engines," SAE paper 930871, 1993.
33. Metghalchi, M., and Keck, J. C.: "Laminar Burning Velocity of Propane-Air Mixtures at High Temperature and Pressure," *Combust. Flame*, vol. 38, pp. 143–154, 1980.
34. Metghalchi, M., and Keck, J. C.: "Burning Velocities of Mixtures of Air with Methanol, Iso-octane, and Indolene at High Pressure and Temperature," *Combust. Flame*, vol. 48, pp. 191–210, 1982.
35. Milton, B. E., and Keck, J. C.: "Laminar Burning Velocities in Stoichiometric Hydrogen and Hydrogen-Hydrocarbon Gas Mixtures," *Combust. Flame*, vol. 58, pp. 13–22, 1984.
36. Rhodes, D. B., and Keck, J. C.: "Laminar Burning Speed Measurements of Indolene-Air-Diluent Mixtures at High Pressures and Temperature," SAE paper 850047, 1985.
37. Egolfopolous, N. N., Du, D. X., Law, C. K.: "A Study on Ethanol Oxidation Kinetics in Laminar Premixed Flames, Flow Reactors and Shock Tubes," 24th Symposium (International) on Combustion, the Combustion Institute, Pittsburgh, p. 833, 1992.
38. Groff, E. G., and Matekunas, F. A.: "The Nature of Turbulent Flame Propagation in a Homogeneous Spark-Ignited Engine," SAE paper 800133, *SAE Trans.*, vol. 89, 1980.

39. Tabaczynski, R. J., Trinker, F. H., and Shannon, B. A. S.: "Further Refinement and Validation of a Turbulent Flame Propagation Model for Spark-Ignition Engines," *Combust. Flame*, vol. 39, pp. 111–121, 1980.
40. Zhou, F., Harrington, D. L., and Lai, M.-C.: *Automotive Gasoline Direct-Injection Engines*, SAE Inc., Warrendale, PA, 2002.
41. Kulzer, A., Christ, A., Rauscher, M., Sauer, C., Würfel, G., and Blank, T.: "Thermodynamic Analysis and Benchmark of Various Gasoline Combustion Concepts," SAE paper 2006-01-0231, 2006.
42. Fansler, T. D., and Drake, M. C.: "Flow, Mixture Preparation and Combustion in Direct-Injection Two-Stroke Gasoline Engines," in *Flow and Combustion in Reciprocating Engines*, C. Arcoumanis and T. Kamimoto (eds), Ch. 2, pp. 67–136, Springer-Verlag, Berlin Heidelberg, 2009.
43. Matekunas, F. A.: "Modes and Measures of Cyclic Combustion Variability," SAE paper 830337, *SAE Trans.*, vol. 92, 1983.
44. Young, M. B.: "Cyclic Dispersion in the Homogeneous-Charge Spark-Ignition Engine—A Literature Survey," SAE paper 810020, *SAE Trans.*, vol. 90, 1981.
45. Ozdor, N., Dulger, M., and Sher, E.: "Cyclic Variability in Spark Ignition Engines, A Literature Survey," SAE paper 940987, 1994.
46. Kuroda, H., Nakajima, Y., Sugihara, K., Takagi, Y., and Muranaka, S.: "The Fast Burn with Heavy EGR, New Approach for Low NO_x and Improved Fuel Economy," SAE paper 780006, *SAE Trans.*, vol. 87, 1978.
47. Heywood, J. B.: "Combustion and its Modeling in Spark-Ignition Engines," The Third International Symposium on Diagnostics and Modeling of Combustion in Internal Combustion Engines (COMODIA 94), Yokohama, Japan, July 1994.
48. Matsui, K., Tanaka, T., and Ohigashi, S.: "Measurement of Local Mixture Strength at Spark Gap of S.I. Engines," SAE paper 790483, *SAE Trans.*, vol. 88, 1979.
49. Quader, A. A.: "Lean Combustion and the Misfire Limit in Spark Ignition Engines," SAE paper 741055, *SAE Trans.*, vol. 83, 1974.
50. Quader, A. A.: "What Limits Lean Operation in Spark Ignition Engines—Flame Initiation or Propagation?" SAE paper 760760, *SAE Trans.*, vol. 85, 1976.
51. Bosch, R.: *Automotive Handbook*, 8th edition, Robert Bosch GmbH, Stuttgart, Germany, SAE, 2011.
52. Maly, R., and Vogel, M.: "Ignition and Propagation of Flame Fronts in Lean CH₄-Air Mixtures by the Three Modes of the Ignition Spark," *Proceedings of Seventeenth International Symposium on Combustion*, pp. 821–831, The Combustion Institute, Pittsburgh, 1976.
53. Maly, R.: "Ignition Model for Spark Discharges and the Early Phase of Flame Front Growth," *Proceedings of Eighteenth International Symposium on Combustion*, pp. 1747–1754, The Combustion Institute, Pittsburgh, 1981.
54. Heywood, J. B., and Sher, E.: *The Two-Stroke Cycle Engine: Its Development, Operation, and Design*, Taylor & Francis, Philadelphia, SAE, 1999.
55. Pischinger, S., and Heywood, J. B.: "A Model for Flame Kernel Development in a Spark-Ignition Engine," 23rd International Symposium on Combustion, pp. 1033–1040, The Combustion Institute, Pittsburgh, 1990.
56. Bosch, R.: *Automotive Handbook*, 4th edition, Robert Bosch GmbH, Stuttgart, Germany, SAE, 1976.
57. Yoshida, K., Nakayama, K., and Suzuki, T., "Development of Next-Generation High Ignitability Spark Plug," SAE paper 2008-01-0467, 2008.
58. Lee, Y. G., Grimes, D. A., Boehier, J. T., Sparrow, J., and Flavin, C.: "A Study of the Effects of Spark Plug Electrode Design on 4-Cycle Spark-Ignition Engine Performance," SAE paper 2000-01-1210, 2000.

59. Geiger, J., Pischinger, S., Böwing, R., Koss, H.-J., and Thiemann, J., "Ignition Systems for Highly Diluted Mixtures in SI-Engines," SAE paper 1999-01-0799, 1999.
60. Nakamura, N., Baika, T., and Shibata, Y.: "Multipoint Spark Ignition for Lean Combustion," SAE paper 852092, 1985.
61. Miyashita, N., Matsubara, Y., Iwaa, K., and Ishikawa, M.: "Spark Plugs for Gasoline Direct Injection Engines," SAE paper 2001-01-1200, 2001.
62. Ziegler, G. F. W., Wagner, E. P., Saggau, B., and Maly, R. R.: "Influence of a Breakdown Ignition System on Performance and Emission Characteristics," SAE paper 840992, *SAE Trans.*, vol. 93, 1984.
63. Anderson, R. W., and Asik, J. R.: "Lean Air-Fuel Ignition System Comparison in a Fast-Burn Engine," SAE paper 850076, 1985.
64. Dale, J. D., and Oppenheim, A. K.: "Enhanced Ignition for I.C. Engines with Premixed Gases," SAE paper 810146, *SAE Trans.*, vol. 90, 1981.
65. Douaud, A., and Eyzat, P.: "DIGITAP-An On-Line Autoignition and Processing System for Instantaneous Engine Data-Applications," SAE paper 770218, 1977.
66. Mittal, V., Revier, B. M., and Heywood, J. B.: "Phenomena that Determine Knock Onset in Spark-Ignition Engines," SAE paper 2007-01-0007, 2007.
67. Draper, C. S.: "Pressure Waves Accompanying Detonation in the Internal Combustion Engine," *Jnl. Of the Aeronautical Sciences*, vol. 5, p. 219 et seq, 1938.
68. Scholl, D., Davis, C., Russ, S., and Berash, T.: "The Volume Acoustic Modes of Spark-Ignited Internal Combustion Engines," SAE paper 980893, 1998.
69. Leppard, W. R.: "Individual-Cylinder Knock Occurrence and Intensity in Multicylinder Engines," SAE paper 820074, 1982.
70. Cornetti, G. M., DeCristofaro, F., and Gozzelino, R.: "Engine Failure and High Speed Knock," SAE paper 770147, 1977.
71. Renault, F.: "A New Technique to Detect and Control Knock Damage," SAE paper 820073, 1982.
72. *The Relationship between Knock and Engine Damage*, Tentative Code of Practice, CEC Report M-07-T-83, Co-ordinating European Council (CEC), London, England, 1984.
73. Eng, J. A.: "Characterization of Pressure Waves in HCCI Combustion," SAE paper 2002-01-2859, 2002.
74. Nakajima, Y., Nagai, T., Iijima, T., Yokoyama, J., and Nakamura, K.: "Analysis of Combustion Patterns Effective in Improving Anti-Knock Performance of a Spark-Ignition Engine," *JSAE Rev.*, vol. 13, pp. 9–17, 1984.
75. Nakagawa, Y., Takagi, Y., Itoh, T., and Iijima, T.: "Laser Shadowgraphic Analysis of Knocking in S.I. Engine," SAE paper 845001; also in *XX FISITA Congress Proceedings*, vol. P–143, 1984.
76. Glassman, I.: *Combustion*, 2nd edition, Academic Press, Orlando, FL, 1987..
77. Halstead, M. P., Kirsch, L. J., Prothero, A., and Quinn, C. P.: "A Mathematical Model for Hydrocarbon Autoignition at High Pressures," *Proc. R. Soc., ser. A*, vol. 346, pp. 515–538, 1975.
78. Benson, S. W.: "The Kinetics and Thermochemistry of Chemical Oxidation with Application to Combustion and Flames," *Prog. Energy Combust. Sci.*, vol. 7, pp. 125–134, 1981.
79. Alperstein, M., and Bradow, R. L.: "Investigations into the Composition of End Gases from Otto Cycle Engines," SAE paper 660410, *SAE Trans.*, vol. 75, 1966.
80. Wheeler, R. W.: "Abnormal Combustion Effects on Economy," in *Fuel Economy in Road Vehicles Powered by Spark-Ignition Engines*, J. C. Hilliard and G. S. Springer (eds), Ch. 6, pp. 225–276, Plenum Press, New York, 1984.

81. Livengood, J. C., and Wu, P. C.: "Correlation of Autoignition Phenomenon in Internal Combustion Engines and Rapid Compression Machines," *Proceedings of Fifth International Symposium on Combustion*, p. 347, Reinhold, Pittsburgh, 1955.
82. Douaud, A. M., and Eyzat, P.: "Four-Octane-Number Method for Predicting the Anti-Knock Behavior of Fuels and Engines," SAE paper 780080, *SAE Trans.*, vol. 87, 1978.
83. Kasseris, E., and Heywood, J. B., "Charge Cooling Effects on Knock Limits in SI DI Engines Using Gasoline/Ethanol Blends: Part 2-Effective Octane Numbers," SAE paper 2012-01-1284, 2012.
84. Halstead, M. P., Kirsch, L. J., and Quinn, C. P.: "Autoignition of Hydrocarbon Fuels at High Temperatures and Pressures—Fitting of a Mathematical Model," *Combust. Flame*, vol. 30, pp. 45–60, 1977.
85. Schäpertöns, H., and Lee, W.: "Multidimensional Modelling of Knocking Combustion in SI Engines," SAE paper 850502, 1985.
86. Lovell, W. G.: "Knocking Characteristics of Hydrocarbons," *Ind. Engng. Chem.*, vol. 40, pp. 2388–2438, 1948.
87. Lichty, L. C.: *Combustion Engine Processes*, McGraw-Hill, 1967.
88. Kalghatgi, G. T.: "Auto-Ignition Quality of Practical Fuels and Implications for Fuel Requirements of Future SI and HCCI Engines," SAE paper 2005-01-0239, 2005.
89. Mittal, A., and Heywood, J. B.: "The Relevance of Fuel RON and MON to Knock Onset in Modern SI Engine," SAE paper 2008-01-2414, 2008.
90. Mittal, V., and Heywood, J. B.: "The Shift in Relevance of Fuel RON and MON to Knock Onset in Modern SI Engines Over the Last 70 Years," SAE paper 2009-01-2622, 2009.
91. Mittal, V., and Heywood, J. B.: "The Underlying Physics and Chemistry behind Fuel Sensitivity," SAE paper 2010-01-0617, 2010.
92. Kalghatgi, G. T., Bradley, D., Andrae, J., and Harrison, A. J.: "The Nature of "Super-knock" and its Origins in SI Engines," paper C684028, *Conference on Internal Combustion Engines: Performance, Fuel Economy and Emissions*, ed. Institution of Mechanical Engineers, Chandos, London, 2009.
93. Zahdeh, A., Rothenberger, P., and Schäfer, J.: "Diagnosing Engine Combustion Using High Speed Photography in Conjunction with CFD," AVL Optical Diagnostics Conference, Baden-Baden, 2008.
94. Chevron Products Company, *Motor Gasolines Technical Review*, Chevron Corporation, 1996, and 2009.
95. Russ, S.: "A Review of the Effect of Engine Operating Conditions on Borderline Knock," SAE paper 960497, 1996.
96. Hoepke, B., Jannsen, S., Kasseris, E., and Cheng, W. K.: "EGR Effects on Boosted SI Engine Operation and Knock Integral Correlation," SAE paper 2012-01-0707, 2012.
97. Grandin, B., Angström, H.-K., Stålhammar, P. and Olofsson, E.: "Knock Suppression in a Turbocharged SI Engine by Using Cooled EGR," SAE paper 982476, 1998.

10.0. CHAPTER PRELIMINARIES

ESSENTIAL FEATURES OF PROCESS

TYPES OF DIESEL COMBUSTION SYSTEMS

DIESEL ENGINE COMBUSTION

FUEL SPRAY BEHAVIOR

IGNITION DELAY

MIXING-CONTROLLED COMBUSTION

ALTERNATIVE COMPRESSION-IGNITION COMBUSTION APPROACHES

PROBLEMS

REFERENCES

10.1. ESSENTIAL FEATURES OF PROCESS

Combustion in compression-ignition or diesel engines occurs through a different sequence of processes than does combustion in a spark-ignition engine. The essential features of the diesel combustion process can be summarized as follows. Fuel is injected by the fuel-injection system into the engine cylinder toward the end of the compression stroke, just before the desired start of combustion (SOC). [Figures 1.26 to 1.30](#) illustrate the major components of diesel fuel-injection systems. The liquid fuel, injected at high velocity, usually as several jets through small orifices or nozzles in the injector tip, atomizes into small drops and penetrates into the high-temperature high-pressure air in the combustion chamber. The fuel vaporizes and the fuel vapor mixes with air. The high-temperature fuel-vapor air mixture thus produced ignites spontaneously, after a delay period of a few crank-angle degrees, and starts burning. The cylinder pressure increases as combustion of the evolving fuel-vapor air mixture continues. The consequent compression of the unburned portion of the charge shortens the delay before ignition for the fuel and air which has mixed to within combustible limits, which then burns rapidly. It also reduces the evaporation time of the remaining liquid fuel. Injection continues until the desired amount of fuel has entered the cylinder. Atomization, vaporization, fuel-air mixing, and combustion continue until essentially all the fuel has passed through each process. In addition, mixing of the air remaining in the cylinder with burning and already burned gases continues throughout the combustion and expansion processes. Thus, for much of the diesel's combustion process, the fuel-burning rate is controlled by the rate at which fuel mixes with air.

Such "non-premixed flames" are usually termed *diffusion flames* because the (gaseous) fuel and the air must diffuse into each other and then react. Most of the fuel chemical energy is released in a thin reaction sheet close to where the fuel vapor and air reach the stoichiometric mixture proportions (as shown in [Fig. 9.2](#)). Because the in-cylinder flow is turbulent, the diesel combustion process occurs primarily as an unsteady diffusion flame. Spontaneous ignition, resulting from the high air temperature and pressure in the cylinder when injection commences, and the fact that fuel-air mixing controls the burning rate, are the distinguishing features of the diesel combustion process.

It will be clear from this summary that the compression-ignition combustion process is extremely complex.^[1] The details of the process depend on the characteristics of the fuel, on the design of the engine's combustion chamber and fuel-injection system, and on the engine's operating conditions. It is an unsteady, heterogeneous, three-dimensional combustion process, with complex chemistry. While a well-grounded conceptual understanding of diesel engine combustion has been developed, to date an ability to describe many of the critical individual processes in a quantitative manner is incomplete.

Some important consequences of this combustion process on engine operation are the following:

- Since injection commences just before combustion starts, there is no knock limit as in the spark-ignition engine resulting from spontaneous ignition of the premixed fuel and air in the end-gas. Hence a higher engine compression ratio can be used in the compression-ignition engine, improving its fuel conversion efficiency relative to the SI engine.
- Since injection timing is used to control combustion timing, the delay period between the start of injection (SOI) and start of combustion (SOC) must be kept short (and reproducible). A short delay is also needed to hold the maximum cylinder gas pressure below the maximum the engine can tolerate. Thus, the spontaneous ignition characteristics of the fuel-air mixture must be held within a specified range. This is done by requiring that diesel fuel have a cetane number (a measure of the ease of ignition of that fuel in a typical diesel environment; see [Sec. 10.6.2](#)) above a certain value.
- Since the combustion process occurs within and around the fuel sprays, engine torque can be controlled by varying the amount of fuel injected per cycle, with the engine's airflow essentially unchanged. Thus, the engine can be operated unthrottled, so pumping work requirements are low: this improves part-load mechanical efficiency relative to the SI engine.
- As the amount of fuel injected per cycle is increased, problems with air utilization during combustion eventually lead to the formation of excessive amounts of soot which cannot be burned up prior to exhaust. This excessive soot or black smoke in the exhaust constrains the fuel/air ratio at maximum engine power to values some 30% lean of stoichiometric. Hence, the

maximum indicated mean effective pressure (in a naturally-aspirated engine, or at a given boost level in a turbocharged engine) is lower than values for an equivalent spark-ignition engine operating at stoichiometric.

- Because the diesel always operates with lean fuel/air ratios (and at part load with very lean fuel/air ratios), the effective value of γ ($= c_p/c_v$) during the expansion process is higher than in a spark-ignition engine. This gives a higher fuel conversion efficiency than the spark-ignition engine, for a given expansion ratio (see [Sec. 5.5.3](#)).

A major challenge in diesel combustion chamber design is achieving sufficiently rapid mixing between the injected fuel and the air in the cylinder to complete combustion in the appropriate crank-angle interval relatively close to top-center. The foregoing discussion indicates (and more detailed analysis will confirm) that mixing rates control the fuel-burning rate. Commercial diesel engines are made in a range of cylinder sizes, with cylinder bores varying from about 70 to 900 mm. The mean piston speed at maximum rated power is approximately constant over this size range (see [Sec. 2.2](#)), so the maximum rated *engine* speed will be inversely proportional to the stroke [see Eq. (2.9)]. For a fixed crank-angle interval for combustion (of order 40 to 50 degrees CA to maintain high fuel conversion efficiency), the time available for combustion will, therefore, scale with the stroke. Thus, at the small end of the diesel engine size range, the mixing between the injected fuel and the air must take place on a time scale some 10 times shorter than in engines at the large end of this range. It would be expected, therefore, that the design of the engine combustion chamber, the inlet port and valve, and the fuel-injection system would have to change substantially over this size range to provide the fuel and air motions inside the cylinder required to achieve the desired fuel-air mixing rate. As engine size decreases, more vigorous air motion is required while less fuel jet penetration is necessary. It is this logic, primarily, that leads to the different diesel combustion chamber designs and fuel injection systems found in practice over the large size range of production diesel engines.

[1] Note that the term *compression-ignition*, widely used to characterize the ignition of fuel-air mixture in diesel engines, is really a much abbreviated characterization of the actual process that occurs. That process is the spontaneous ignition of fuel-vapor, air, mixture within the combustion chamber close to top center (TC), due to its high temperature and pressure (which are a consequence of the compression process) shortly after fuel injection commences).

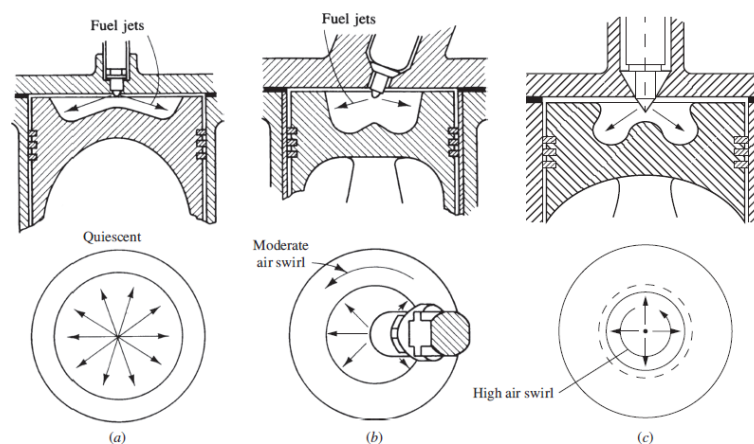
10.2. TYPES OF DIESEL COMBUSTION SYSTEMS

Traditionally, diesel engines have been divided into two basic categories according to their combustion chamber design: (1) *direct-injection (DI) engines*, which have a single open combustion chamber into which fuel is injected directly and (2) *indirect-injection (IDI) engines*, where the chamber is divided into two regions and the fuel is injected into the "prechamber" which is connected to the main chamber above the piston crown via a nozzle, or one or more orifices. IDI engine designs were used extensively in smaller engine sizes, but have largely been replaced by high-speed DI engine designs. Within each category there are several different chamber geometry, air-flow, and fuel-injection arrangements.

10.2.1. Direct-Injection Systems

Figure 10.1 shows three basic DI combustion systems, categorized by their chamber geometries, air motions, and fuel spray arrangement. In the largest-size engines, where mixing rate requirements are least stringent, quiescent DI systems of the type shown in Fig. 10.1a are used. The momentum and energy of the injected fuel jets are sufficient to achieve adequate fuel distribution and rates of mixing with the air. Additional organized air motion is not required. The combustion chamber shape is usually a shallow bowl in the crown of the piston, and a central multi-hole injector is used.

Figure 10.1 Common types of direct-injection diesel engine combustion systems: (a) quiescent chamber with multi-hole nozzle typical of larger engines; (b) bowl-in-piston chamber with moderate swirl and multi-hole nozzle; (c) re-entrant bowl-in-piston chamber with high swirl and multi-hole nozzle: (b) and (c) are used in medium and smaller DI engine size ranges, respectively.

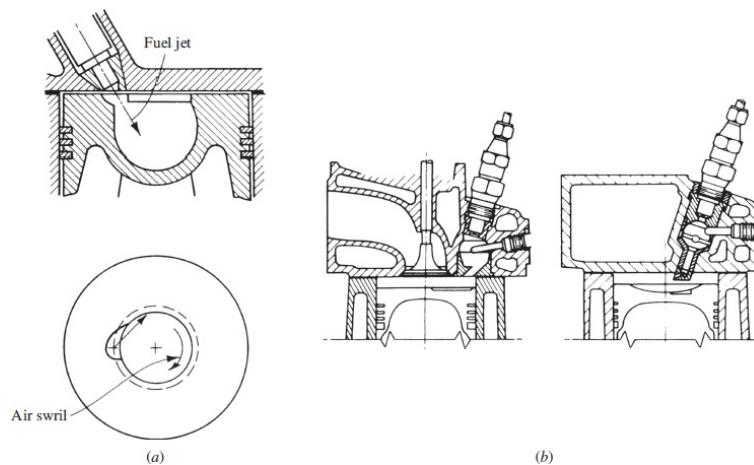


As engine size decreases, increasing amounts of air swirl are used to achieve faster fuel-air mixing rates. Air swirl is generated by suitable design of the inlet port (see Sec. 8.3); the swirl rate can then be increased as the piston approaches TC by forcing the air toward the cylinder axis into a bowl-in-piston type of combustion chamber. Figures 10.1b and c show the two types of DI engine with swirl, in common use. Figure 10.1b shows a medium-sized DI engine with swirl, with a centrally located multi-hole injector nozzle. Here, the design goal is to hold the amount of liquid fuel which impinges on the piston bowl walls to a minimum. Figure 10.1c shows a small high-speed DI engine with a re-entrant bowl-in-piston combustion system, which uses high swirl and a multi-hole injector nozzle. These two types of designs are used in medium-size (10- to 15-cm bore) diesels and, with increased swirl, in small-size (7- to 9-cm bore) diesels.

10.2.2. Other Diesel Combustion Systems

Figure 10.2a shows the M.A.N. wall-distribution combustion system that was developed for commercial vehicle engines to control the harsh combustion noise produced by earlier direct-injection diesels. With this Meurer or M-system, the single-hole fuel-injection nozzle sprays the fuel against the wall of the deep "hemispherical" bowl-in-piston combustion chamber, tangentially with the highly swirling in-bowl air motion. This results in a significantly slower fuel evaporation rate than that occurs in comparable multi-spray air distribution DI approach. With the now common use of turbocharging, and the advances in DI fuel-injection technology that have taken place, any benefits of this M system have largely disappeared.

Figure 10.2 (a) Direct-injection: MAN, Meurer (or M) combustion system with deep bowl in piston, air swirl, and single-hole nozzle. (b) Two common types of small indirect-injection diesel engine combustion systems: left, swirl prechamber; right, turbulent prechamber.



Historically, as diesel combustion systems developed, use of air swirl to enhance fuel-air mixing rates, despite amplification as the air is forced into the piston bowl during compression, did not provide sufficiently high fuel-air mixing rates for small high-speed diesels such as those used in automobiles. Indirect-injection or divided-chamber combustion systems were used instead, where the vigorous charge motion required during fuel injection is generated in a prechamber, during the compression stroke. Two broad classes of IDI systems evolved: (1) swirl chamber systems and (2) prechamber systems, as illustrated in Fig. 10.2b. During compression, air is forced from the main chamber above the piston into the auxiliary chamber, through the nozzle or orifice (or set of orifices). Thus, toward the end of compression, a vigorous flow in the auxiliary chamber is set up; in swirl-chamber systems, the connecting passage and chamber are shaped so that this flow within the auxiliary chamber rotates rapidly.

Fuel was injected into the auxiliary chamber at lower injection-system pressure than was then typical of DI systems, through a pintle nozzle as a single spray: see Fig. 1.18. Combustion starts in the auxiliary chamber; the pressure rise associated with combustion forces fluid back into the main chamber where the jet issuing from the nozzle entrains and mixes with the main chamber air. The glow plug shown on the right of the prechambers in Fig. 10.2b is a cold-starting aid. The plug is heated prior to and during the engine's starting process to ensure ignition of fuel early in the engine cranking process. Such indirect-injection, swirl or pre-chamber, engines are no longer mainstream diesel engine technologies.

10.2.3. Comparison of Different Combustion Systems

The number of different types of combustion chambers proposed and tried since the beginnings of diesel engine development is substantial. Over the years, however, through increased understanding of the physical and chemical processes involved and competition in the marketplace, only a limited number of different designs have survived. The important characteristics of those types of diesel engines now most commonly used are summarized in Table 10.1. The numbers for dimensions and operating characteristics are typical values for each different type of diesel combustion system.

Table 10.1 Characteristics of diesel combustion systems

System	Direct injection systems				Other combustion chambers	
	Quiescent multi-spray	Medium swirl multi-spray	High swirl multi-spray	High swirl "M"	Swirl chamber	Prechamber
Size	Largest	Medium	Smallest	Medium	Smallest	Smallest
Cycle	2-/4-stroke	4-stroke	4-stroke	4-stroke	4-stroke	4-stroke
TC/S/NA*	TC/S	TC/NA	NA/TC	TC/NA	NA/TC	NA/TC
Maximum speed, rev/min	120–2100	1800–3500	2500	4000–4500	3600–5000	4500
Bore, mm	900–150	150–100	100–75	130–100	95–70	95–70
Stroke/bore	3.5–1.2	1.3–1.0	1.1–0.9	1.2–0.9	1.1–0.9	1.1–0.9
Compression ratio	12–15	15–17	18–22	16–18	20–24	22–24
Chamber	Open or shallow dish	Bowl-in-piston	Re-entrant deep bowl-in-piston	Deep bowl-in-piston	Swirl pre-chamber	Single/multi-orifice prechamber
Air-flow pattern	Quiescent	Medium swirl	Highest swirl	High swirl	Very high swirl in prechamber	Very highly turbulent in prechamber
Number of nozzle holes	Multi 6–10	Multi 4–8	Multi 4	Single	Single	Single
Injection pressure	Very high	High	High	Medium	Lowest	Lowest
*Turbocharged/supercharged/naturally-aspirated.						

The larger slowest speed engines used for power generation or marine applications utilize open quiescent chambers in the piston crown, which are close to disc shaped: the motion of the fuel jets is responsible for distributing and mixing the fuel with air. The largest of these are often two-stroke cycle engines. In the next size range, in large truck and locomotive engines, a quiescent chamber consisting of a shallow dish or bowl in the piston crown is used. The air utilization in these engines is low, but they are invariably supercharged and/or turbocharged to obtain high power density.

In the DI category, as engine size decreases and maximum speed rises, swirl is increasingly used to obtain high-enough fuel-air mixing rates. The swirl is generated by suitably shaped inlet ports, and is amplified during compression by forcing most of the air toward the cylinder axis into the deep bowl-in-piston combustion chamber. High-speed smaller versions of the DI diesel have been successfully developed over the past 35 or 50 years for the automobile market. These use high intake-generated swirl, augmented by the radial inward displacement of the air, and the vigorous squish flow the deep re-entrant bowl-in-piston combustion chamber produces, to achieve fast enough fuel-air mixing and combustion rates to operate up to engine speeds of 4500 to 5000 rev/min (lower than maximum speeds for gasoline SI engines, but producing acceptable specific maximum engine power by turbocharging to high boost levels).

Table 10.1 lists the characteristics of three other diesel combustion systems for historical reasons: they are no longer widely used. An alternative to the multi-hole-nozzle swirl system used in medium and heavy-duty truck engines is the M system (or wall-wetting system), where most of the fuel from the single-hole pintle nozzle is placed on the wall of the spherical bowl in

the piston crown. Historically, in the smallest engine sizes, IDI systems were used to obtain the vigorous auxiliary-chamber air motion required for high fuel-air mixing rates. Two different types of auxiliary-chamber geometries were used: a swirl chamber and a turbulence-generating prechamber as shown in [Fig. 10.2b](#).

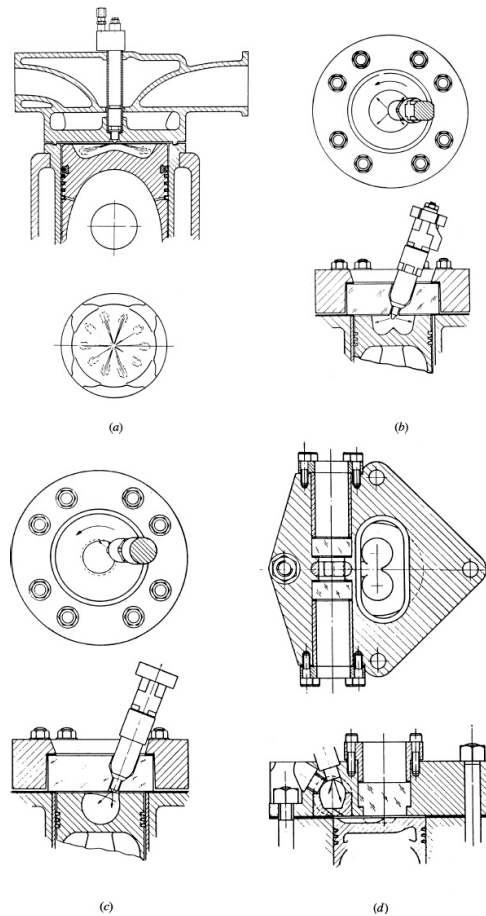
10.3. DIESEL ENGINE COMBUSTION

Studies of diesel engine combustion using optical diagnostics to visualize fuel spray and flame behavior, combined with analyses of engine cylinder pressure data, have led to a widely accepted descriptive model of the compression-ignition engine combustion process. The concept of *heat-release rate* is important to understanding diesel combustion. Its basic definition is the rate at which the chemical energy of the fuel is released by the combustion process. It is usually calculated from cylinder pressure versus crank-angle data, as the energy release rate required to create the measured pressure, using the techniques described in [Sec. 10.4](#).^[1] The combustion model defines several distinct phases of diesel combustion, each controlled by different physical and chemical processes. Although the relative importance of each phase does depend on the combustion system used, and on engine operating conditions, four phases are common to all diesel engines as discussed below.

10.3.1. Optical Studies of Diesel Combustion

Various optical techniques have been used to visualize the diesel combustion process. Many of these studies have been carried out in combustion chambers close to those used in practice, under normal engine operating conditions. Sequences of images over time provide valuable information on the evolution of the combustion process in the different types of engines. [Figure 10.3](#) shows four combustion chamber systems that were studied photographically several decades ago which helped develop earlier combustion models. These are: (a) a quiescent chamber typical of the large diesel engines used for industrial, marine, and rail traction applications (only the burning of a single fuel spray of the multi-spray combustion system could be studied); (b) a medium-speed DI engine with swirl and four fuel jets centrally injected; (c) an M.A.N. "M" DI system; and (d) a Ricardo Comet V swirl-chamber IDI system.¹

Figure 10.3 Four diesel combustion chambers used to obtain the photographs of the compression-ignition combustion processes shown in Fig. 10.4: (a) quiescent DI chamber; (b) multi-hole nozzle DI chamber with swirl; (c) M.A.N. "M" DI chamber; (d) Ricardo Comet IDI swirl chamber.^{2, 3}



The combustion sequences in Fig. 10.4, originally recorded on color film, show the following features:

Fuel sprays. The fuel droplets reflect light from spot lamps and define the extent of the liquid fuel spray prior to complete vaporization.

Premixed burning. These regions are of too low a luminosity to be recorded with the exposure level used. The addition of a copper additive to the fuel gives these normally blue flames a green color bright enough to render them visible.

Diffusion flame. The burning high-temperature carbon particles in this flame (which separates fuel on the inside of the flame from the air outside) provide more than adequate luminosity and appear as yellow-white. As the flame cools, the radiation from the oxidizing particles changes color through orange to red.

Over-rich mixture chemistry. The appearance of brown regions, usually surrounded by a white diffusion flame, indicates excessively rich mixture regions where substantial soot particle production is occurring. Where this fuel-rich soot-laden cloud contacts air, there is a hot (white) diffusion flame.

Figure 10.4 Sequence of photographs (originally in color) from high-speed movies taken in the special visualization diesel engines shown in Fig. 10.3: (a) combustion of single spray burning under large DI engine conditions; (b) combustion of four sprays in DI engine with counterclockwise swirl; (c) combustion of single spray in M.A.N. "M" DI diesel; (d) combustion in prechamber (on left) and main chamber (on right) in Ricardo Comet IDI swirl chamber diesel. 1250 rev/min, imep = 827 kPa (120 lb/in.²)^{1,2} (Courtesy: Ricardo Consulting Engineers.)

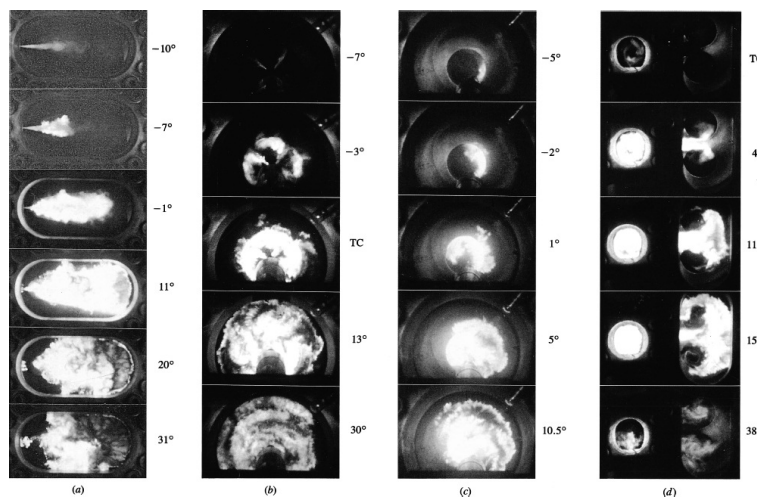


Table 10.2 summarizes the characteristics of these different regions, discernable in the photographs shown in the Fig. 10.4.

Table 10.2 Interpretation of diesel engine combustion photographs^{1,2}

Shade/color*	Interpretation
Dark grey	Background; the gas (air in early stages, combustion products later) is transparent and not glowing
Medium grey (green)	Early in combustion process; low luminosity "premixed" rendered visible by copper added to fuel. Later; burned gas above about 1800°C
Bright white (yellow-tinged)	Carbon particle burnup in diffusion flame, 2000–2500°C
Less-bright white (yellow, orange-red)	Carbon burnup in diffusion flame at lower temperatures; last visible in film at 1000°C
Darker grey regions within white regions (brown)	Soot clouds forming in very fuel-rich mixture regions. Where these meet air (grey) there is always a white fringe of hot flame
*Original photos are in color. ¹	

Figure 10.4a shows a sequence of photographs from one combustion event of the single spray, burning under conditions typical of a large quiescent DI engine. The fuel spray is shown penetrating into the chamber. Ignition occurs at -8° in the fuel-air mixture left behind on the edge of the spray not far from the injector. A flame then spreads rapidly along the outside of the spray to the spray tip where some of the fuel, which has had a long residence time in the chamber, burns with a low-luminosity flame (colored green by the copper fuel additive in the original photos). The diffusion flame engulfing the remainder of the spray is brilliant white with a yellow tinge from the burning of the soot particles which have been formed in the fuel-rich spray core. At this stage (-1°), about 60% of the fuel has been injected. The remainder is injected into this enflamed region, producing a very fuel-rich zone apparent as the brown cloud (11°). This soot cloud moves to the outer region of the chamber (11 to 20°): white-yellow flame activity continues near the injector, probably due to combustion of ligaments of fuel, which issued from the injector nozzle as the injector needle was seating. Combustion continues well into the expansion stroke (31°).

This sequence shows that fuel distribution is highly nonuniform during the combustion process in this type of DI engine. Also, the air between the individual fuel sprays of the quiescent open-chamber diesel mixes with each burning spray relatively slowly, contributing to poor air utilization with this type of combustion chamber.

Figure 10.4b shows a combustion sequence from the DI engine with swirl (the chamber shown in **Fig. 10.3b**). The inner circle corresponds to the deep bowl in the piston crown, the outer circle to the cylinder liner. The fuel sprays (of which two are visible without obstruction from the injector) first appear at -13° . At -7° they have reached the wall of the bowl; the tips of the sprays have been deflected slightly by the anticlockwise swirl. The frame at -3° shows the first ignition. Bright luminous flame zones are visible, one on each spray. Out by the bowl walls, where fuel vapor has been blown around by the swirl, larger burning regions (tinged green in the original photos) indicating the presence of premixed combustion can be seen. The fuel downstream of each spray is next to ignite, burning yellow-white due to the soot formed in the richer mixture. Flame propagation back to the injector follows extremely rapidly and at TC the bowl is filled with flame. At 5° ATC the flame spreads out over the piston crown toward the cylinder wall due to combustion-produced gas expansion and the reverse squish flow (see **Sec. 8.5**). The darker grey (brown) regions (13°) are soot-laden fuel-rich mixture originating from the fuel which impinges on the wall. The last frame (30° ATC) shows the gradual diminution of the soot-particle-laden regions as they mix with the excess air and burn up. The last dull-red burning visible on the film is at about 75° ATC, well into the expansion stroke.

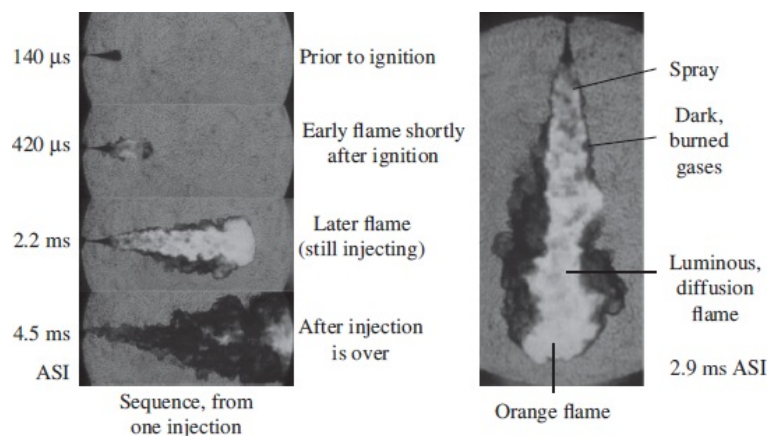
Figure 10.4c shows the combustion sequence for the M-type DI engine.⁴ In the version of the system used for these experiments, the fuel was injected through a two-hole nozzle which produces a main jet directed tangentially onto the walls of the spherical cup in the piston crown, and an auxiliary spray which mixes a small fraction of the fuel directly with the swirling airflow. At -5° the fuel spray is about halfway round the bowl. Ignition has just occurred of fuel adjacent to the wall, which has mixed sufficiently with air to burn. The flame spreads rapidly (-2° , 1°) to envelop the fuel spray and is convected round the cup by the high swirl air flow. By shortly after TC the flame has filled the bowl and is spreading out over the piston crown. A soot cloud is seen near the top right of the picture at 5° ATC which spreads out around the circumference of the enflamed region. There is always a rim of flame between the soot cloud and the cylinder liner as excess air is mixed into the flame zone (10.5°). The flame is of the carbon-burning diffusion-flame type throughout; little premixed combustion (green in the original photos) is seen, even at the beginning of the burning process.

Figure 10.4d shows the combustion sequence in a swirl chamber IDI engine of the Ricardo Comet V design. The swirl chamber (on the left) is seen in the view of the lower drawing of **Fig. 10.3d** (with the connecting passageway entering the swirl chamber tangentially at the bottom left to produce clockwise swirl). The main chamber is seen in the plan view of the upper drawing of **Fig. 10.3d**. Two sprays emerge from the injector nozzle after the SOI at -11° . The smaller, radial auxiliary spray is sharply deflected by the high air swirl. Frame 1 shows how the main spray follows the contour of the chamber; the auxiliary spray has evaporated and can no longer be seen. Green-colored emission is first apparent at -1° from combustion of the vaporized fuel from the auxiliary spray and indicates premixed burning. Combustion then spreads to the main spray (TC), becoming a yellow-white carbon-particle-burning flame with a green fringe. At 4° ATC the swirl chamber appears full of carbon-burning flame, which is being blown through the throat and into the recesses in the piston crown by the combustion-generated pressure rise in the prechamber. The flame jet impinges on the piston recesses entraining the air in the main chamber, leaving grey patches (green in the original) where all the carbon has burned out (4° , 11° , 15° ATC). A darker grey (brown) soot cloud is emerging from the throat. By 15° ATC this soot cloud has spread around the cylinder, with a bright yellow-white flame at its periphery as this soot finds excess air and burns up. The yellow-white flame becomes yellow and then orange-red as the soot particles and burned gases cool on expansion. By 38° ATC most of the flame is burnt out.

Figure 10.5 shows additional details of the diesel spray development and combustion process. A high-pressure diesel fuel-injection system was used to inject a single fuel spray (similar to one spray exiting a multi-hole nozzle) into a constant-volume combustion vessel containing a quiescent gas (with 21% oxygen) at a temperature and pressure corresponding to diesel engine end-of-compression conditions.^[2] Visible light emission from shadowgraph optical images of the jet was recorded with appropriate optical equipment. **Figure 10.5** shows a sequence of four spray images from one experiment, as the injected fuel atomizes, penetrates, entrains air, vaporizes, ignites, and burns. The figure caption and labels describe the individual spray and flame images. An enlargement of one image toward the end of injection identifies more clearly the liquid fuel containing (early) region of the spray, the lift-off of the diffusion flame from the nozzle hole exit, the (orange/yellow) turbulent

unsteady diffusion flame with fuel-rich partly reacted fuel combustion products on the inside of the flame which is made luminous by the burn up of soot particles within the flame, and air mixed with the hot burned gases diffusing outward from the diffusion flame seen as the turbulent irregularly bounded darker region identified by the shadowgraph optical technique. Overall, the injected liquid's momentum propels the fuel droplets into the combustion chamber like a "shotgun blast," entraining high-temperature air into the developing spray, causing the fuel drops to evaporate (the liquid spray core only penetrates partway). The initial momentum continues to move the now largely gaseous spray across the combustion chamber, entraining more hot air as it progresses. A flame is first visible fairly close to the injector nozzle and rapidly spreads around the spray. Its color (white-yellow-orange) indicates that the luminosity is emitted from small soot particles formed in the (fuel-rich) partly reacted mixture inside the spray as these particles diffuse into the flame and burn up. The irregularities in the diffusion flame surface (and in the boundary of the hot burned gases diffusing outward into the air surrounding the diffusion flame) clearly indicate that all these processes are turbulent.

Figure 10.5 Shadowgraph and flame radiation emission images of combustion of a single diesel spray in a constant volume combustion facility at typical diesel engine charge conditions at TC. Sequence of four images: spray prior to ignition, early spray and diffusion flame, later spray and flame, spray and flame after injection has ended. Separate image on right shows hot burned gas (dark) diffusing outward from the diffusion flame. ASI, time after start of fuel injection. (Courtesy: Dr. Dennis Siebers, Sandia National Laboratories.)

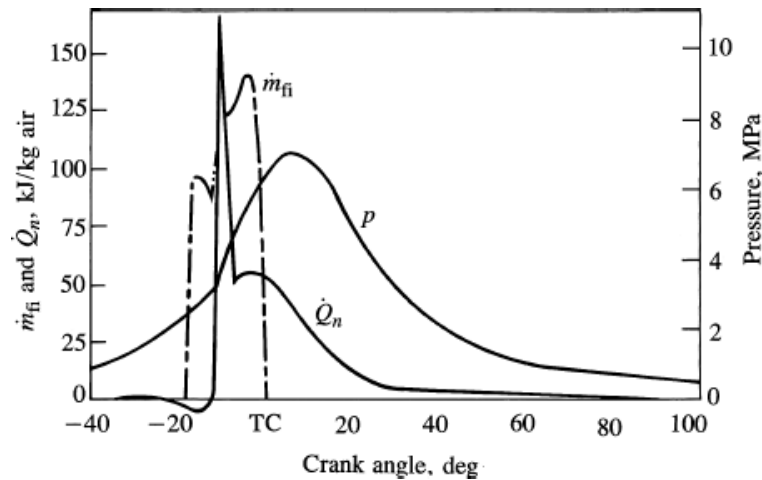


10.3.2. Combustion in Direct-Injection Multi-Spray Systems

Figure 10.6 shows typical data for cylinder pressure (p), fuel injection rate, and net heat release rate through the compression and expansion strokes of a DI diesel. The engine had central fuel injection through a multi-hole nozzle into a bowl-in-piston combustion chamber. The rate of fuel injection is obtained from the fuel-line pressure, cylinder pressure, nozzle geometry, and needle-lift profiles by considering the injector as one or more flow restrictions.⁵ There is a delay of several degrees between the SOI and SOC [identified by the change in slope of the $p(\theta)$ curve]. The pressure rises rapidly for a few crank-angle degrees, then more slowly to a peak value some 5° after start-of-injection. Injection usually continues well into the combustion process. A rate-of-heat-release diagram,^[3] corresponding to this rate of fuel injection and the cylinder pressure data, is also shown in Fig. 10.6. The general shape of the rate-of-heat-release curve is typical of this type of multi-spray DI engine. There is a short *ignition delay* between the SOI at -20° and the SOC at about -15° . The slight loss of "heat" during the delay period is due to heat transfer into the liquid fuel, vaporizing and heating it, and heat transfer to the walls. During the combustion process the burning proceeds in three distinguishable stages. In the first stage, the rate of burning is rapid, lasts for only a few crank-angle degrees, and produces an increasing and then decreasing "spike" in the heat-release rate. It corresponds to the period of rapid cylinder pressure rise. The second stage produces an increasing and then decreasing heat-release rate creating a second more rounded profile, as shown in Fig. 10.6. This is the main heat-release period and usually lasts about 30 to 40° . Up to about 90% of the total fuel energy is released in these first two periods. The "tail" of the heat-release rate diagram during which the

remainder of the fuel's chemical energy is released results from the mixing of the excess air that was not involved in the main combustion process with the burned gases, and equilibrating.

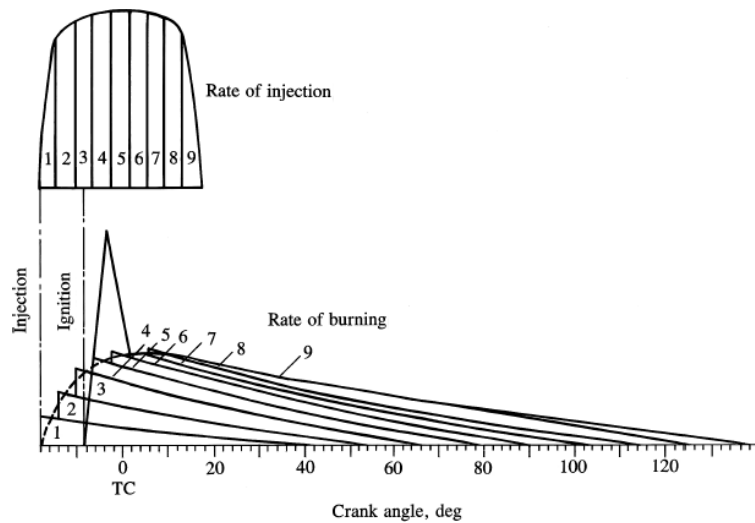
Figure 10.6 Cylinder pressure p , rate of fuel injection and net heat-release rate $\dot{Q}_n$ calculated from p , for a small DI diesel engine, as functions of crank angle: 1000 rev/min, normal injection timing, $b_{mep} = 620 \text{ kPa}$.⁵



From studies of rate-of-injection and heat-release diagrams such as those in Fig. 10.6, over a range of engine loads, speeds, and injection timings, Lyn and Austen^{5,6} developed the following summary. First, the total burning period is significantly longer than the fuel injection period. Second, the absolute burning rate increases proportionally with increasing engine speed; thus on a crank-angle basis, the burning interval remains (for a given amount of fuel injected) essentially constant. Third, the magnitude of the initial peak of the burning-rate diagram depends strongly on the duration of the ignition delay period, being higher for longer delays. These considerations, coupled with engine combustion photographic studies, lead to the following overall framework for diesel combustion.

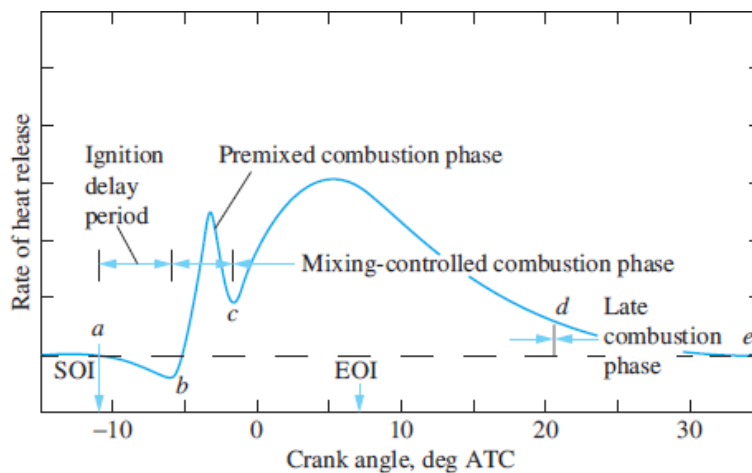
Figure 10.7 shows schematically the rate-of-injection and rate-of-burning diagrams, where the injected fuel as it enters the combustion chamber has been divided into a number of elements. The first element which enters mixes with air and becomes "ready for burning" (i.e., mixes to within combustible limits), as shown conceptually by the lowest triangle along the abscissa in the rate-of-burning figure. While some of this fuel element mixes rapidly with air, the rest will mix more slowly. The second and subsequent elements will mix with air in a similar manner, and the total "ready-for-burning" diagram, enclosed by the dashed line, is obtained. The total area of this diagram is equal to that of the rate-of-injection diagram. Ignition does not occur until after the delay period is over, however. At the ignition point, some of the fuel already injected has mixed with enough air to be within the combustible limits. That "premixed" fuel-air mixture (the shaded region in Fig. 10.7) is then added to the mixture which becomes ready for burning after the end of the delay period, producing the high initial rate of burning as shown. Such a heat-release profile is generally observed with this type of naturally-aspirated DI diesel engine. When the ignition delay is long, the amount of premixed "ready-to-burn" mixture is larger, leading to a high peak rate of heat release. When the delay is short (e.g., in a warmed-up turbocharged DI diesel) the amount of premixed mixture is, relatively, much less and this premixed rate of heat release spike is much lower.^[4]

Figure 10.7 Schematic of relationship between rate of fuel injection and rate of fuel burning or chemical energy release.⁶



We therefore identify the following stages of the overall compression-ignition diesel combustion process on a heat-release-rate diagram for a DI engine (Fig. 10.8).

Figure 10.8 Typical DI engine heat-release-rate diagram identifying the different phases of the diesel combustion process.



An *ignition delay* (ab). The period between the start of fuel injection into the combustion chamber and the SOC determined from the change in slope on the $p-\theta$ diagram, or from a heat-release analysis of the $p(\theta)$ data, or from a luminosity detector.

A *premixed or rapid combustion phase* (bc). Here, chemical energy release from the fuel which has mixed with air to within combustible limits during the ignition delay period occurs rapidly in a few crank-angle degrees. When this burning mixture is added to the fuel which becomes ready for burning and burns during this phase, the high heat-release rates characteristics of this phase result. This rapid energy release comes from combustion of the fuel-rich mixture within each spray.

A *mixing-controlled combustion phase* (cd). Once the fuel and air which premixed during the ignition delay has released some of its chemical energy, the burning rate (or heat-release rate) is then controlled by the rate at which mixture becomes available for burning. While several processes are involved—liquid fuel atomization, vaporization, mixing of fuel vapor with air, rich-mixture chemical reactions within the spray—the rate of burning in this diffusion-flame phase of diesel combustion is controlled primarily by the fuel vapor-air mixing process.

A *late combustion phase (de)*. Heat release continues at a lower and decaying rate into the expansion stroke. There are several causes. A small fraction of the fuel may not yet have burned. A fraction of the fuel energy is present in soot and fuel-rich combustion products and can still be released. The cylinder charge is nonuniform and mixing of burned gases with air during this period promotes more complete combustion and less-dissociated product gases. The kinetics of the final burnout processes become slower as the temperature of the cylinder gases fall during expansion.

10.3.3. Heat-Release-Rate Analysis

Cylinder pressure versus crank-angle data over the compression and expansion strokes of the engine operating cycle can be used to obtain quantitative information on the progress of combustion. Suitable methods of analysis which yield the rate of release of the fuel's chemical energy, usually called the *heat release rate* or rate of fuel burning, through the diesel engine combustion process will now be described. The methods of analysis are similar to those described in [Sec. 9.2.2](#) for spark-ignition engines and start with the first law of thermodynamics for an open system which is quasi static (i.e., uniform in pressure and temperature). The first law for such a system (see [Fig. 9.11](#)) is

$$\frac{dQ}{dt} - p \frac{dV}{dt} + \sum_i \dot{m}_i h_i = \frac{dU}{dt}$$

(10.1)

where dQ/dt is the heat-transfer rate across the system boundary into the system, $p(dV/dt)$ is the rate of work transfer done by the system due to system boundary displacement, $\dot{O}_i$ is the mass flow rate into the system across the system boundary at location i (flow out of the system would be negative), h_i is the enthalpy of flux i entering or leaving the system, and U is the energy of the material contained inside the system boundary.

The following factors make the application of this equation to diesel combustion difficult:

- Fuel is injected into the cylinder. Liquid fuel is added to the cylinder, which vaporizes and mixes with air to produce a fuel/air ratio distribution which is nonuniform and varies with time.
- The composition of the burned gases (also nonuniform) is not known.
- The accuracy of available correlations for predicting heat transfer in diesel engines is not well defined (see [Chap. 12](#)).
- Crevice regions (such as the volumes between the piston, rings, and cylinder wall) constitute a few percent of the clearance volume. The gas in these regions, however, is cooled to close to the wall temperature, increasing its density and, therefore, the relative importance of these crevices. Thus, crevices increase heat transfer and contain a non-negligible fraction of the in-cylinder air during combustion at a density significantly higher than in the rest of the combustion chamber.

Due to difficulties in dealing with these problems, methods of heat-release or burn-rate analysis give only approximate answers.

10.3.3.1. Combustion Efficiency

In both heat-release and fuel mass burned estimations, an important factor is the completeness of combustion. Air utilization in diesels is limited by the onset of black smoke in the exhaust. The smoke is soot particles, which are mainly carbon. Soot and other incomplete combustion products such as unburned hydrocarbons and carbon monoxide represent a combustion inefficiency. Normally, combustion efficiency is calculated from the combustibles in the exhaust. However, torque loss is more

appropriately estimated from the inefficiency resulting from soot, CO and HC present in the cylinder at about the end of combustion; only modest additional torque is generated beyond the first one-third of the expansion stroke. Based on exhaust concentrations, the magnitude of that inefficiency is small. At full load conditions, if only 0.5% of the fuel supplied was present in the exhaust as black smoke, the result would be unacceptable. Soot emissions from the cylinder contain less than 0.1% of the fuel's chemical energy. Typically, hydrocarbon emissions are the order of 0.2% of the fuel. The energy corresponding to the exhausted carbon monoxide is about 0.5%. Thus, the combustion inefficiency [Eq. (4.67)] is usually less than 1%: the combustion efficiency is usually greater than about 99% (see Fig. 3.9). While these emissions are important in terms of their air-pollution impact (see Chap. 11), from the point of view of energy conversion it is a good approximation to regard combustion and heat release in diesels as essentially complete.

10.3.3.2. Heat-Release Analysis: Direct-Injection Engines

For this type of engine, the cylinder contents are a single open system. The only mass flows across the system boundary (while the intake and exhaust valves are closed) are the fuel injection and the crevice flows. An approach, which incorporates the crevice flow, has been described in Sec. 9.2.2; crevice flow effects will be omitted here. Equation (10.1) therefore becomes

$$\frac{dQ}{dt} - p \frac{dV}{dt} + m_f h_f = \frac{dU}{dt}$$

(10.2)

Two common methods are used to obtain combustion information from pressure data via Eq. (10.2). In both approaches, the cylinder contents are assumed to be at a uniform temperature at each instant in time during the combustion process. This commonly used method yields the net fuel energy- or heat-release rate; the other method yields a fuel mass burning rate. The term *apparent* is often used to describe these quantities since both are approximations to the real quantities, which cannot be determined exactly.

If U and h_f in Eq. (10.2) are taken to be the sensible internal energy of the cylinder contents and the sensible enthalpy of the injected fuel, respectively,^[5] then dQ/dt becomes the difference between the chemical energy or "heat" released by combustion of the fuel (a positive quantity) and the heat transfer from the system (in engines, the heat transfer is from the system, and by thermodynamic convention is a negative quantity). Since $h_{s,f} \approx 0$, Eq. (10.2) becomes

$$\frac{dQ_n}{dt} = \frac{dQ_{ch}}{dt} - \frac{dQ_{ht}}{dt} = p \frac{dV}{dt} + \frac{dU_s}{dt}$$

(10.3)

The *net heat-release rate*, dQ_n/dt , which is the difference between the *gross heat-release rate* dQ_{ch}/dt and the heat-transfer rate to the walls dQ_{ht}/dt , equals the rate at which work is done on the piston plus the rate of change of sensible internal energy of the cylinder contents.

If we further assume that the contents of the cylinder can be modeled as an ideal gas, then Eq. (10.3) becomes

$$\frac{dQ_n}{dt} = p \frac{dV}{dt} + m c_v \frac{dT}{dt}$$

(10.4)

From the ideal gas law, $pV = mRT$, with R assumed constant, it follows that

$$\frac{dp}{p} + \frac{dV}{V} = \frac{dT}{T}$$

(10.5)

Equation (10.5) can be used to eliminate T from Eq. (10.4) to give

$$\frac{dQ_n}{dt} = \left(1 + \frac{c_v}{R}\right) p \frac{dV}{dt} + \frac{c_v}{R} V \frac{dp}{dt}$$

or

$$\frac{dQ_n}{dt} = \frac{\gamma}{\gamma - 1} p \frac{dV}{dt} + \frac{1}{\gamma - 1} V \frac{dp}{dt}$$

(10.6)

Here γ is the ratio of specific heats, c_p/c_v . An appropriate range for γ for diesel heat-release analysis is 1.3 to 1.35; Eq. (10.6) is often used with a constant value of γ within this range. More specifically, we would expect γ for diesel engine heat-release analysis to have values appropriate to air temperatures at end-of-compression-stroke prior to combustion (≈ 1.35) and to burned gases at the overall equivalence ratio following combustion (≈ 1.26 – 1.3). The appropriate values for γ during combustion, which would give most accurate heat-release information, are not well defined.^{7,8}

More complete methods of heat-release analysis based on Eq. (10.2) have been proposed and used. These include more sophisticated models for the gas properties before, during, and after combustion, and for heat transfer and crevice effects.⁸ However, it is also necessary to deal with the additional issues of: (1) mixture nonuniformity (fuel/air ratio nonuniformity and burned and unburned gas nonuniformities); (2) accuracy of any heat-transfer model used (see Chap. 12); and (3) the effects of the crevice regions. These additional phenomena must be dealt with at an equivalent level of accuracy for more complex heat-release models to be worthwhile. For many engineering applications, Eq. (10.6) is adequate for diesel engine combustion analysis.

Additional insight can be obtained by incorporating a model for the largest of the effects omitted from Eq. (10.6), the heat transfer rate dQ_{ht}/dt (see Chap. 12): we thereby obtain a close approximation to the gross heat-release rate. The integral of the gross heat-release rate over the complete combustion process should then equal (to within a few percent since the analysis is not exact) the mass of fuel injected m_f times the fuel's lower heating value, Q_{LHV} : that is,

$$Q_{ch} = \int_{t_{start}}^{t_{end}} \frac{dQ_{ch}}{dt} dt = m_f Q_{LHV}$$

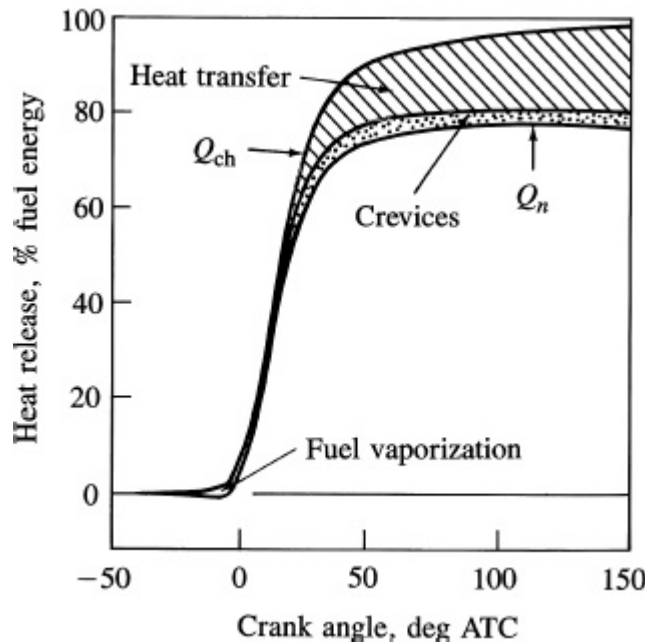
(10.7)

Of course, Eqs. (10.1) to (10.4), (10.6), and (10.7) also hold with crank-angle θ as the independent variable instead of time t .

Figure 10.9 illustrates the relative magnitude of gross and net heat release, heat transfer, crevice effects, and heat of vaporization and heating up of the fuel for a turbocharged DI diesel engine operating in the mid-load, mid-speed range. The net heat release is the gross heat release due to combustion, less the heat transfer to the walls, crevice effects, and the effect of fuel vaporization and heatup. (It was omitted above by neglecting the mass addition term in dU/dt .) This last term is

sufficiently small to be neglected. The enthalpy of vaporization of diesel fuel is less than 1% of its heating value: the energy change associated with heating up and vaporizing the fuel from injection temperature to typical end-of-compression temperatures is about 3% of the fuel heating value. The heat transfer integrated over the duration of the combustion period is 10 to 25% of the total heat released.

Figure 10.9 Gross and net heat-release profiles during combustion, for a turbocharged DI diesel engine at mid-load, mid-speed, showing relative magnitude of heat transfer, crevice, and fuel vaporization and heat-up effects.



10.3.3.3. Fuel Mass Burning Rate Analysis

If the internal energies of the fuel, air, and burned gases in Eq. (10.1) are evaluated relative to a consistent datum (such as that described in Sec. 4.5.2), then this equation can be used to obtain an apparent *fuel mass burning rate* from cylinder pressure versus crank-angle data. (With such a species energy datum the "heat release" is accounted for in the internal energy and enthalpy terms.) Following Krieger and Borman,⁹ Eq. (10.2) can be written as

$$\frac{d}{dt}(mu) = p \frac{dV}{dt} + \frac{dQ}{dt} + h_f \frac{dm}{dt}$$

(10.8)

Here Q is the heat transfer to the gas within the combustion chamber (i.e., $Q = -Q_{ht}$), m is the mass within the combustion chamber, and dm/dt has been substituted for $\dot{m}_f$.

Since the properties of the gas mixture in the cylinder during combustion (assumed to be uniform and in chemical equilibrium at the pressure p and average temperature T) are in general a function of p , T , and the equivalence ratio ϕ :

$$u = u(T, p, \phi) \text{ and } R = R(T, p, \phi)$$

Therefore,

$$\frac{du}{dt} = \frac{\partial u}{\partial T} \frac{dT}{dt} + \frac{\partial u}{\partial p} \frac{dp}{dt} + \frac{\partial u}{\partial \phi} \frac{d\phi}{dt}$$

(10.9a)

$$\frac{dR}{dt} = \frac{\partial u}{\partial T} \frac{dT}{dt} + \frac{\partial R}{\partial p} \frac{dp}{dt} + \frac{\partial R}{\partial \phi} \frac{d\phi}{dt}$$

(10.9b)

Also,

$$\phi = \phi_0 + \left(\frac{m}{m_0} - 1 \right) \frac{1 + (F/A)_0}{(F/A)_s}$$

(10.10)

and,

$$\frac{d\phi}{dt} = \frac{1 + (F/A)_0}{(F/A)_s m_0} \frac{dm}{dt}$$

(10.11)

where (F/A) is the fuel/air ratio; the subscript 0 denotes the initial value prior to fuel injection and the subscript s denotes the stoichiometric value. It then follows that

$$\frac{1}{m} \frac{dm}{dt} = \frac{-(RT/V)(dV/dt) - (\partial u / \partial p)(dp/dt) + (1/m)(dQ/dt) - CB}{u - h_f + D(\partial u / \partial \phi) - C[1 + (D/R)(\partial R / \partial \phi)]}$$

(10.12)

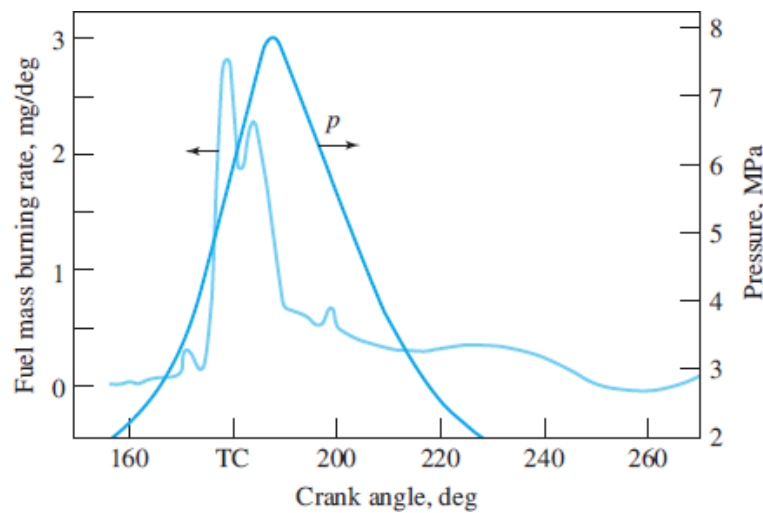
where

$$\begin{aligned} B &= \frac{1}{p} \frac{dp}{dt} - \frac{1}{R} \frac{\partial R}{\partial p} \frac{dp}{dt} + \frac{1}{V} \frac{dV}{dt} \\ C &= \frac{T(\partial u / \partial T)}{1 + (T/R)(\partial R / \partial T)} \\ D &= \frac{[1 + (F/A)_0]m}{(F/A)_s m_0} \end{aligned}$$

Equation (10.12) can be solved numerically for $m(t)$ given m_0 , ϕ_0 , $p(t)$, and appropriate models for the working fluid properties (see Sec. 4.7) and for the heat-transfer term dQ/dt (see Chap. 12).

Figure 10.10 shows cylinder pressure data for an open chamber DI diesel and fuel mass burning rate $dm/d\theta$ calculated from that data using the above method. The heat-transfer model of Annand was used (see Sec. 12.4.3). The result obtained is an *apparent* fuel mass burning rate. It is best interpreted, after multiplying by the heating value of the fuel, as the fuel chemical-energy or heat-release rate. The *actual* fuel-burning rate is unknown because not all the fuel "burns" with sufficient air available locally to produce products of *complete* combustion. About 60% of the fuel has burned in the first one-third of the total combustion period. The integral of the fuel mass burning rate over the combustion process should equal the total fuel mass burned; in this case it is 3% less than the total fuel mass injected. Note that chemical energy continues to be released well into the expansion process. The accuracy of this type of calculation then decreases, however, since errors in estimating heat transfer significantly affect the apparent fuel-burning rate.

Figure 10.10 Cylinder pressure p and fuel mass burning rate calculated from p , as a function of crank angle, using the Krieger and Borman method⁹ for a DI diesel engine at 3200 rev/min and full load.



Krieger and Borman also carried out sensitivity analyses for the critical assumptions and variables. They found that the effect of dissociation of the product gases was negligible. This permits a substantial simplification of Eq. (10.12). With no dissociation, $u = u(T, \phi)$, and

$R = \tilde{R}/M$ can be taken as constant, since the molecular weight M changes little. Then,

$$\frac{dm}{dt} = \frac{[1 + (c_v/R)]p(dV/dt) + (c_v/R)V(dp/dt) - (dQ/dt)}{h_f + (c_v/R)(pV/m) - u - D(\partial u/\partial \phi)}$$

(10.13)

where D , as before, is $[1 + (F/A)_0]m/[(F/A)_s m_0]$. Given the uncertainties inherent in the heat-transfer model and the neglect of nonuniformities and crevices, Eq. (10.13) represents an adequate level of sophistication.

The other sensitivity variations studied by Krieger and Borman were: shifting of the phasing of the pressure data 2° forward and 2° backward; translating the pressure data ± 34 kPa; changing the heat transfer $\pm 50\%$; and increasing the initial mass 5%. The initial mass change had a negligible effect on the fuel burning rate calculations. The heat-transfer changes of $\pm 50\%$ changed the mass of fuel burned by about $\pm 5\%$. The change in phasing of the pressure data was more significant. It should be stressed that *accurate* (in magnitude and phasing) pressure data are an important requirement for useful heat-release or fuel-mass-burning rate analysis.

10.3.3.4. Heat Release Rate: Indirect-Injection Engines

In IDI diesel engines, the pressures in the two chambers, main and auxiliary, are not the same during the combustion process. Since combustion starts in the auxiliary or prechamber, the fuel energy release in the prechamber causes the pressure there to rise above the main chamber pressure. Depending on combustion chamber design and operating conditions, the prechamber pressure rises to be 0.5 to 5 bar above that in the main chamber. This pressure difference causes a flow of fuel, air, and burning and burned gases into the main chamber, where additional energy release now occurs. The analysis of the DI diesel was based on uniform pressure throughout the combustion chamber. For IDI engines the effect of the pressure difference between the chambers must be included. The methodology for carrying out a heat release rate analysis for IDI diesel engines can be found in the first edition of this text,¹ and the references quoted. It will not be discussed here since IDI diesels are no longer a widely used technology.

10.3.4. Conceptual Model of DI Diesel Combustion

Over the past 20 or so years, much information has been gained through use of optical laser diagnostics to measure the distributions of various fuel and chemical species that play key roles in the diesel combustion and pollutant formation processes, and of gas temperatures. When coupled with detailed flow and combustion chemistry modeling, this information has enabled a more complete and integrated description of the DI diesel combustion process to be developed. Based on summaries by Dec,¹⁰ and Flynn et al. at Cummins Inc. and U.S. National Laboratories,¹¹ we will now review that more detailed conceptual model.

Figure 10.11 shows a sequence of illustrative drawings of a developing fuel spray and flame in a multi-spray DI diesel engine without swirl. The crank angles noted on each schematic are after the start of injection (ASI). The needle lift profile for the injector which defines the start and end of fuel injection, and apparent heat-release-rate diagram for these drawings, are plotted in **Fig. 10.12** so they can be keyed to the fuel injection process and fuel chemical energy release. **Figure 10.13** shows more details of the spray and flame structure in the "quasi-steady" mixing-controlled phase of combustion. The different shading intensities indicate the regions of the fuel spray where the several different processes involved in this type of diesel combustion occur.

Figure 10.11 Schematic temporal sequence showing how DI diesel combustion evolves from start of liquid fuel injection through the quasi-steady mixing-controlled burning stage. The location in the spray and surrounding flame of the different processes is indicated. The timing (crank-angle degrees after start of injection, ASI, $1^\circ = 139 \mu\text{s}$) and spray penetration (mm) correspond to combustion in a medium-sized diesel engine, at a typical operating condition. The injection duration extends beyond the 10° sequence shown.¹⁰

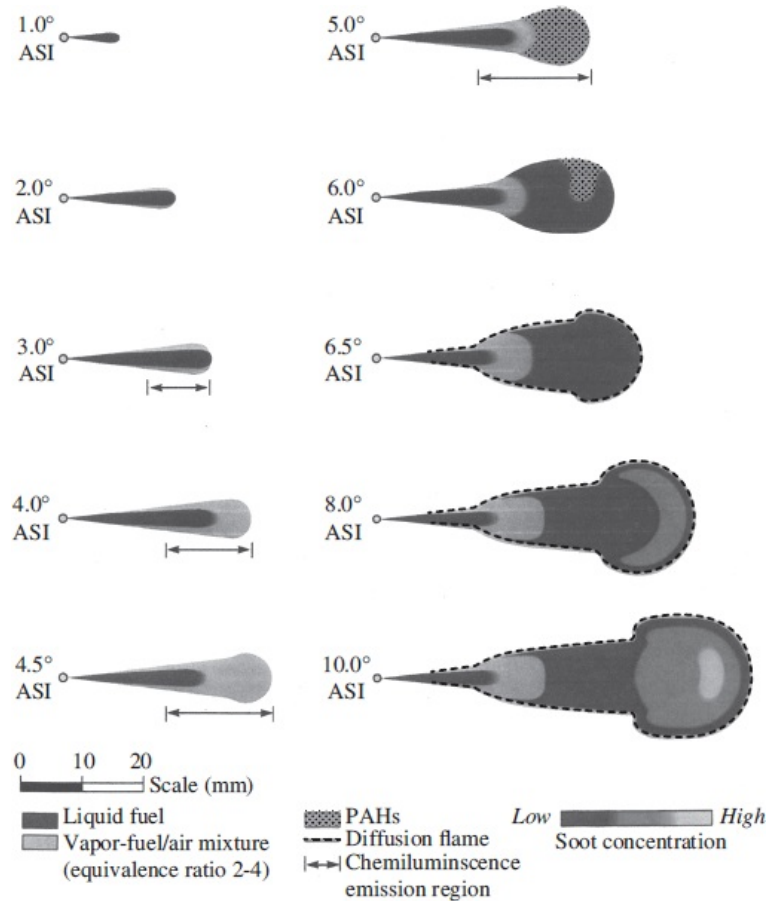


Figure 10.12 Fuel injector needle lift profile which defines the beginning and end of injection, and the net heat-release rate, versus crank angle for the engine conditions of the schematics shown in Fig. 10.11.¹¹

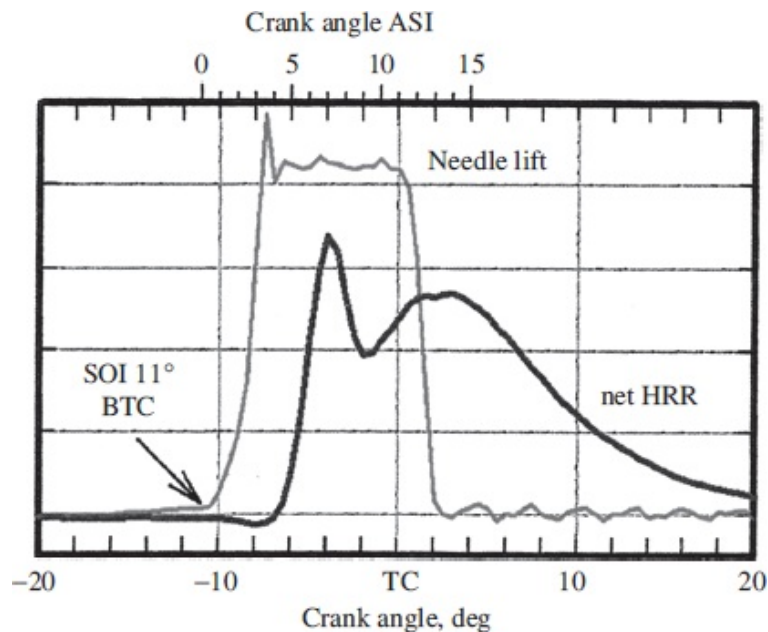
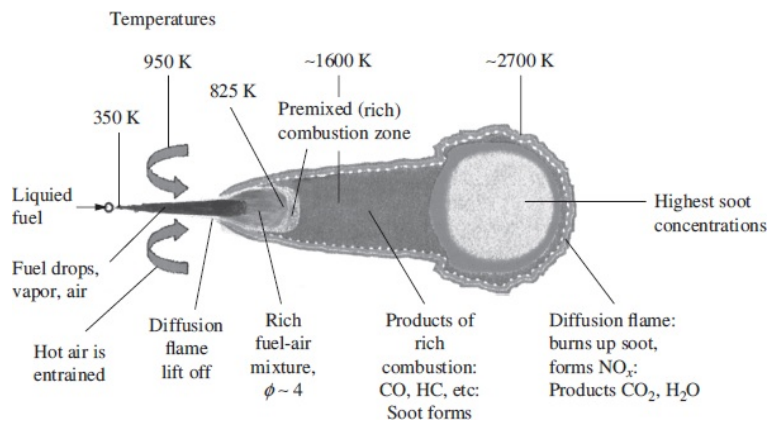


Figure 10.13 Key features, and gas temperature distribution, in a DI diesel fuel spray: injection, atomization, air entrainment, premixed-combustion process, soot formation region, diffusion-flame location, and lift-off distance from nozzle orifice. (Developed from Ref. 11.)



As the liquid fuel exits the nozzle, liquid ligaments and larger drops rapidly atomize to fine drops ($\sim 10 \mu$ diameter) whose momentum carries them into the combustion chamber as a "jet." This early spray region contains the fuel as drops which heat up and vaporize as their motion entrains hot air into the spray. A fuel-vapor air mixture forms within this droplet-containing region and further expands the spray as the spray continues to penetrate (and slow down), up to 4.5° after start of injection (ASI). The ignition chemistry then spontaneously commences in the leading region of the spray (where the fuel has been within the combustion chamber longest) where mixing of air entrained into the spray with evaporated fuel has brought the mixture fuel-air equivalence ratio from very rich values to about 2 to 4 (still well-rich of stoichiometric). The diesel fuel molecules then start to break into fragments, which combine to form polycyclic aromatic hydrocarbons (PAHs), which in turn grow to form increasing amounts of soot particles of increasing size.

At about the crank angle of the peak heat-release-rate of the premixed combustion phase (some 6 to 6.5° ASI in this example), the mixing-controlled diffusion flame forms and rapidly surrounds the jet. The "fuel" for this diffusion flame is the fuel-rich mixture of partly reacted diesel-fuel vapor, entrained air, fuel-fragments, PAHs and soot, all inside the diffusion flame.

The oxygen for this diffusion flame to complete combustion comes from the air that diffuses into the flame from the outside. Thus, [Fig. 10.13](#) summarizes this description of the diesel combustion process during its quasi-steady phase while fuel injection continues and all these processes are occurring in sequence as indicated (after about 8 to 10° ASI in [Fig. 10.11](#)).

[Figure 10.13](#) also shows the temperature distribution within the spray in this quasi-steady stage. The liquid fuel jet entering the chamber atomizes, entrains hot air (plus a modest amount of residual gas) at around 950 K, and vaporizes, producing cooler rich fuel-air mixture ($\phi \sim 4$) at about 650 K. This mixture enters the spray region within the diffusion flame sheath, entraining recirculated partially oxidized rich combustion products, and raising its temperature to about 825 K. Then, low-temperature oxidation reactions occur releasing enough chemical energy to raise the temperature to about 1150 K. At which point oxidation becomes more rapid, this rich mixture passes through a fuel-rich premixed flame consuming the locally available oxygen, releasing 10 to 15% of the fuel's chemical energy, and attaining temperatures of order 1600 K. As the fuel passes through this rich premixed reaction zone or flame, hydrocarbon fragments form which become the building blocks for soot particles within the evolving rich fuel spray. These build up into particle nuclei which become heated by radiation from the flame, grow into soot particles and agglomerate. The highest levels of soot and largest particles are thus found in the leading region of the spray. These soot-particle-containing gases move into the diffusion flame reaction region around the spray where any remaining fuel vapor, fuel fragments, PAHs, and soot particles become oxidized to CO_2 and H_2O using oxygen that diffuses into this flame sheath from the air outside. The highest temperatures are reached in this diffusion flame where close-to-stoichiometric combustion occurs. It is here, within this high-temperature diffusion flame sheath (where there is a slight excess of oxygen) that the air pollutant nitric oxide (NO) largely forms: see [Sec. 11.2.4](#). These products of complete combustion then (in a net sense) diffuse outward from the thin diffusion flame reaction zone mixing with excess air, become leaner, with their temperature and degree of dissociation of CO_2 and H_2O decreasing and releasing chemical energy in the process.¹¹

Key features of this more complete phenomenological diesel spray combustion model (in a warmed-up engine) are the limited penetration of fuel droplets into the combustion chamber, the two stage nature of the combustion process—a fuel-rich premixed flame within the spray followed by a diffusion flame surrounding the spray where the rich premixed combustion products release most of their chemical energy as they fully oxidize. These flame regions are *lifted off* (separated from) the injector nozzle hole allowing the spray to entrain hot air into the spray core to create the rich mixture within the spray. The substantial amount of additional oxygen required for complete combustion diffuses into the spray-enveloping diffusion flame sheath from the outside.

Two important practical areas to be incorporated into the above conceptual model of diesel spray combustion are air swirl, and multiple fuel injections. Increasing levels of air swirl are used to speed up the fuel-air mixing process as engine size—stroke and bore—are reduced and maximum engine speed rises. This swirling air motion, orthogonal to each initial fuel spray axis, forces additional air into the spray; this decreases the spray's fuel/air equivalence ratio more quickly, and thus achieves a less-rich mixture within the spray for the first "premixed" stage of the autoignition process to occur sooner. The higher the swirl level ([Sec. 8.3](#)), the more these effects are augmented. Swirl impacts on diesel combustion are explored more fully in [Sec. 10.6](#).

Fuel-injection technology has progressed to the point where multiple injection pulses are feasible. Pilot injection— injection of a modest fraction of the total amount, injected just prior to the main injection—has long been used to reduce combustion noise and NO_x emissions. Pilot injection shortens the ignition delay by creating a hotter local environment into which the main fuel spray is injected. Post-main-injection pulses are now used to reduce particulate emissions, raise the exhaust gas temperature to initiate the burn-up of particulate matter collected in the exhaust system particulate trap, or provide hydrocarbons for the diesel oxidation catalyst to combust to increase the exhaust gas temperature, or provide reducing species in the exhaust. Obviously, each of the several injections per cycle occurs in the environment created by the previous injections. The topic of multiple-injection effects on diesel combustion is explored more fully in [Sec. 10.7.1](#).

[1] Note that this latter definition (usually called *net* heat release) combines the combustion energy release with heat losses to the cylinder walls.

-
- [2] These gas conditions in the combustion vessel were realized by burning an appropriate mixture of nitrogen, oxygen, acetylene, and hydrogen in the vessel before diesel fuel injection occurred, allowing time for equilibrium and, if necessary, cooling, to achieve the desired gas temperature, pressure, density, and oxygen concentration.
- [3] The heat-release rate plotted here is the net heat-release rate (see [Sec. 10.4](#)). It is the sum of the change of sensible internal energy of the cylinder gases and the work done on the piston. It differs from the rate of fuel energy released by combustion by the heat transferred to the combustion chamber walls, and the thermal energy absorbed by fuel vaporization. The heat loss to the walls is 10 to 25% of the fuel heating value in smaller engines; it is less in larger engine sizes. This net heat release can be used as an indicator of the fuel's chemical energy release when the heat loss is small.
- [4] This two-step model illustrates why the initial premixed rapid-burning part of the heat release rate profile occurs. However, its second step—becomes ready for burning—involves several stages as discussed more fully in [Sec. 10.3.4](#).
- [5] That is, $U = U_s = U(T) - U(298 \text{ K})$ and $h_f = h_{s,f} = h_f(T) - h_f(298 \text{ K})$; see [Sec. 5.5](#) for definitions.

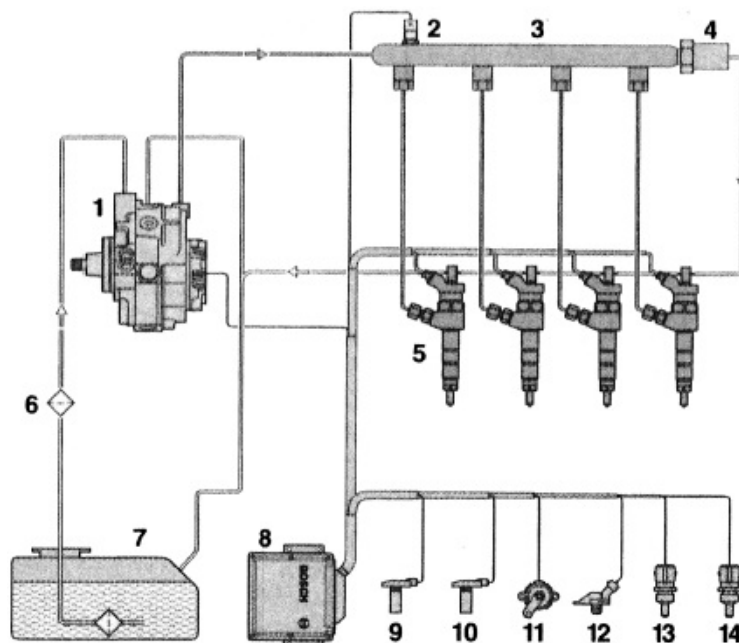
10.4. FUEL SPRAY BEHAVIOR

10.4.1. Fuel Injection

The fuel is injected into the cylinder of a diesel engine through a nozzle with a large pressure difference across the nozzle orifice. The cylinder pressure at injection is typically in the range of 50 to 100 bar (atm). Fuel injection pressures in the range 1000 to over 2000 bar (atm) are used depending on the engine size and type of combustion system employed. These large pressure differences across the injector nozzle are required so that the injected liquid fuel jet will enter the chamber at sufficiently high velocity to (1) atomize into small-sized droplets to enable rapid evaporation, and (2) traverse the combustion chamber in the time available so as to fully utilize the air charge.

Examples of common diesel fuel-injection systems were described briefly in [Sec. 1.7](#) and illustrated in [Figs. 1.26](#) and [1.27](#). The task of the fuel-injection system is to meter the appropriate quantity of fuel for the given engine speed and load to each cylinder, each cycle, and inject that fuel at the appropriate time in the cycle at the desired rate with the spray configuration required for the particular combustion chamber employed. It is important that injections begin and end cleanly, and avoid any secondary injections. To accomplish this task, fuel is drawn from the fuel tank by a supply pump, and forced through a filter to the injection pump. The injection pump sends fuel under pressure to the nozzle pipes which carry fuel to the injector nozzles located in each cylinder head. Excess fuel goes back to the fuel tank. Individual components can be combined to form different diesel fuel injection systems. In some of these (with axial-and radial-piston distributor pumps, and at times in the common-rail system), the fuel supply pump function is integrated into the high-pressure pump. [Figure 10.14](#) shows the layout of a common-rail fuel injection for a four-cylinder passenger car diesel engine with the set of engine sensors it requires.¹²

Figure 10.14 Layout of a common-rail fuel-injection system for a passenger car diesel engine with delivery-controlled high-pressure pump: injection system components and sensors for electronic control unit (ECU).¹² (1) high-pressure pump; (2) rail-pressure sensor; (3) "common rail" fuel rail; (4) pressure limiter; (5) injectors; (6) fuel filter; (7) fuel tank with first stage filter; (8) ECU; (9) engine-speed sensor; (10) phase sensor; (11) pedal-travel sensor; (12) boost-pressure sensor; (13) air-temperature sensor; (14) engine-temperature sensor. (Courtesy Robert Bosch GmbH and SAE.)



In-line injection pumps are used in engines in the 40- to 100-kW per cylinder maximum power range. They contain a plunger and barrel assembly for each engine cylinder. The fuel metering and injection process is illustrated in Fig. 10.15. An engine-driven cam moves the plunger in the delivery direction, and a spring forces it back to its initial position. The plunger stroke is fixed. The plunger fits sufficiently accurately within the barrel to seal without additional sealing elements, even at high pressures and low speeds. The amount of fuel delivered is altered by varying the *effective* plunger stroke. This is achieved by means of a control rod or rack (see Fig. 10.16) which moves in the pump housing and rotates the plunger with its helical groove via a ring gear or linkage lever on the control sleeve. The high-pressure chamber above the plunger is connected with the chamber below the plunger by a vertical groove in the plunger. Delivery ceases when the plunger helix exposes the intake port (port opening), thus connecting the plunger chamber with the suction gallery. When this takes place depends on the rotational position of the plunger. In the case of a lower helix, delivery always starts (port closing) at the same time, but ends sooner or later depending on the rotational position of the plunger. With a plunger with an upper helix, port closing (start of delivery) not port opening is controlled by the helix and is varied by rotating the plunger. Figure 10.16 illustrates how the plunger helix controls fuel delivery.¹²

Figure 10.15 Barrel and plunger with helical groove used to control fuel delivery in the in-line diesel fuel injection pump.¹² (1) from fuel gallery; (2) to nozzle; (3) barrel; (4) plunger; (5) lower helix; (6) vertical (stop) groove. (Courtesy Robert Bosch GmbH and SAE.)

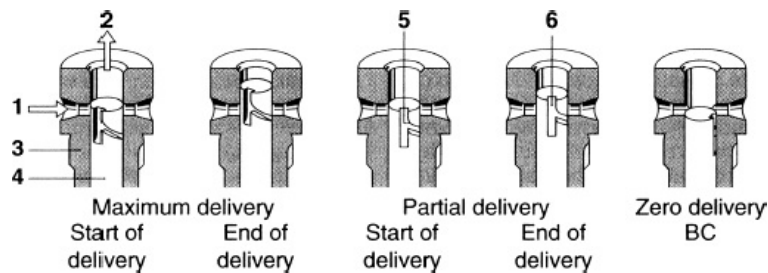
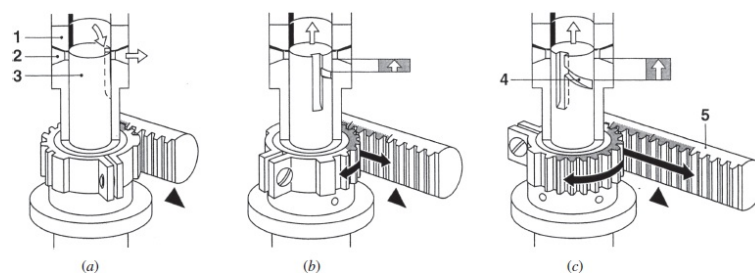


Figure 10.16 Toothed control rack used to rotate plunger (with its helical groove; see Fig. 10.15) to control fuel delivery as a function of engine load:¹² (a) zero delivery; (b) partial delivery; (c) maximum delivery. (1) Pump barrel; (2) inlet port; (3) pump plunger; (4) helix; (5) control rack. (Courtesy Robert Bosch GmbH and SAE.)

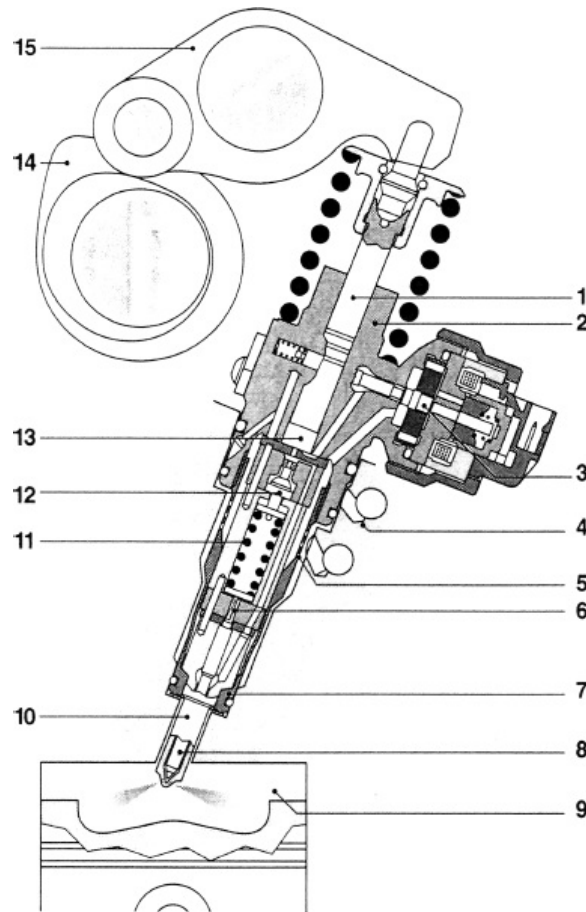


Distributor-type fuel-injection pumps are normally used in multi-cylinder engines up to about 50 kW per cylinder maximum power with injection pressures up to 2000 bar (atm), at speeds to about 2500 rev/min. These pumps have only one plunger and barrel. The pump plunger is made to describe a combined rotary and stroke movement by the rotating cam plate. The fuel is accurately metered to each injection nozzle in turn by this plunger which simultaneously acts as the distributor. Such units are more compact and cheaper than in-line pumps but cannot achieve such high injection pressures. The distributor-type fuel-injection pump is combined with the automatic timing device, governor, and supply pump to form a single unit.

Also used extensively on both smaller and larger engines are unit injectors where the pump and injector nozzle are combined into a single unit. An example of a unit injector and its driving mechanism is shown in Fig. 10.17. Fuel, supplied to the injector through a fuel-distributing manifold, enters the cavity or plunger chamber ahead of the plunger through a metering orifice.

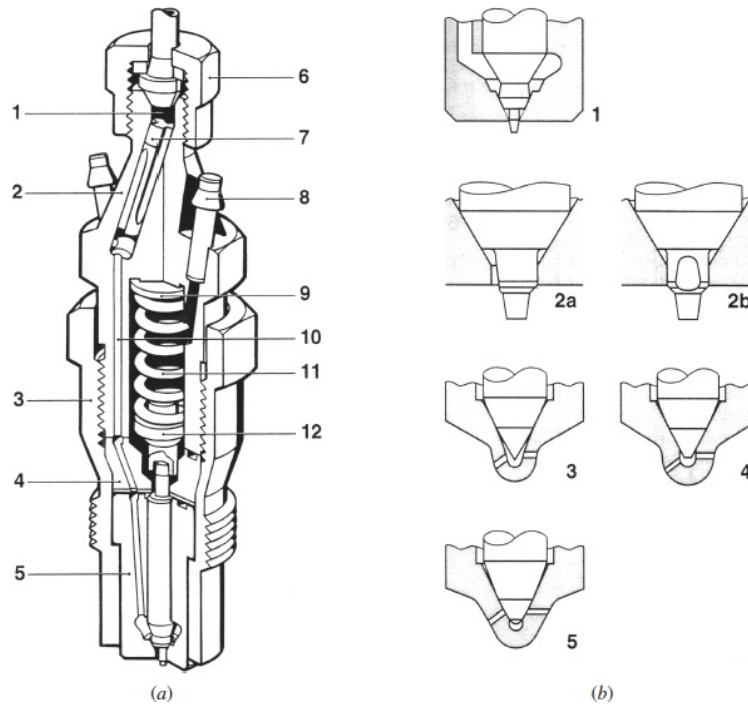
When fuel is to be injected, the cam via the rocker arm pushes down the plunger, closing the metering orifice and compressing the fuel, causing it to flow through check valves and discharge into the cylinder through the injector nozzle holes.

Figure 10.17 Example of a unit injector and its driving mechanism used in passenger car diesel engines.¹² (1) Pump plunger; (2) pump body; (3) solenoid-valve needle; (4) fuel return (low-pressure stage); (5) inlet passage; (6) hydraulic stop (damping unit); (7) retaining nut; (8) nozzle needle; (9) engine combustion chamber; (10) integrated nozzle; (11) nozzle compression spring; (12) accumulator bypass plunger; (13) plunger chamber; (14) camshaft; (15) roller rocker arm. (Courtesy Robert Bosch GmbH and SAE.)



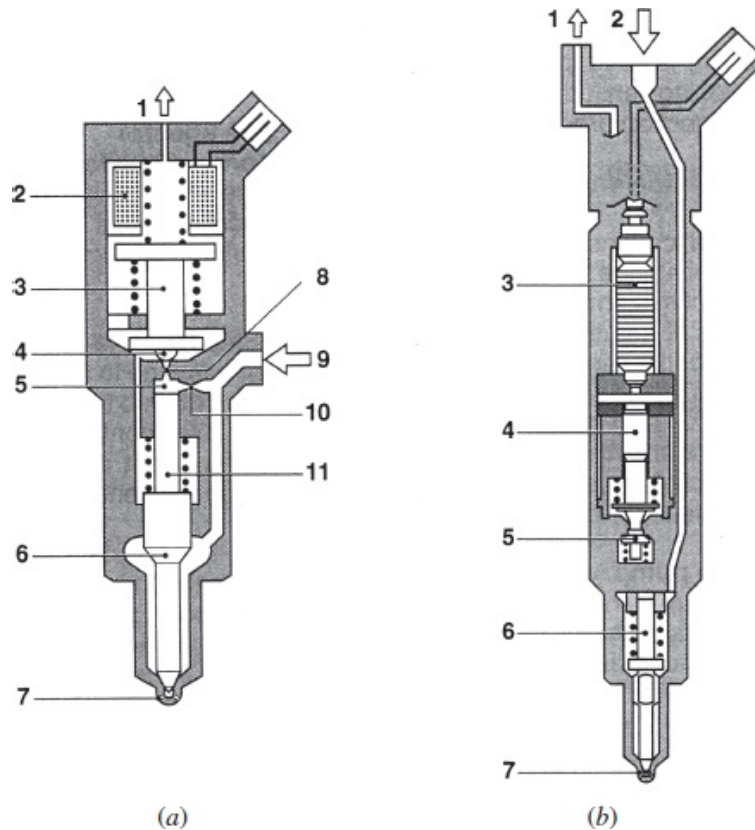
The most important part of the injection system is the nozzle. Examples of different nozzle types and a nozzle holder assembly are shown in Fig. 10.18. The nozzles shown are fluid-controlled needle valves where the needle is forced against the valve seat by a spring. The pressure of the fuel in the pressure chamber above the nozzle aperture opens the nozzle by the axial force it exerts on the conical surface of the nozzle needle. Needle valves are used to prevent dribble from the nozzles when injection is not occurring. It is important to keep the volume of fuel left between the needle and nozzle orifices (the sac volume) as small as possible to prevent any fuel flowing into the cylinder after injection is over, to control hydrocarbon emissions (see Sec. 11.4.4). Multi-hole nozzles are used with most DI diesel engine systems. Pintle nozzles, where the needle projects into and through the nozzle hole, are used in IDI engine systems. The shape of the pin on the end of the nozzle needle controls the spray pattern and fuel-delivery characteristics. Auxiliary nozzle holes are sometimes used to produce an auxiliary smaller spray to aid ignition and starting.

Figure 10.18 (a) Layout of injector nozzle and nozzle holder assembly with throttling pintle: (1) inlet; (2) nozzle-holder body; (3) nozzle-retaining nut; (4) shim; (5) nozzle; (6) union nut with high-pressure line; (7) edge filter; (8) leak-fuel port; (9) pressure-adjusting shims; (10) pressure passage; (11) compression spring; (12) pressure pin. (b) Examples of different nozzle shapes and nozzle hole arrangements: (1) throttling-pintle nozzle; (2) throttling-pintle nozzle with flat-cut pintle (2a side view, 2b front view); (3) hole-type nozzle with conical blind hole; (4) hole-type nozzle with cylindrical blind hole; (5) sac-less nozzle.¹² (Courtesy Robert Bosch GmbH and SAE.)



Electronic control of injection is now widely used. In an electronic injector, such as that shown in [Fig. 10.19a](#), a solenoid-operated control valve performs the injection timing and metering functions in a fashion analogous to the ports and helices of the mechanical injector. Solenoid valve closure initiates pressurization and injection, and opening causes injection pressure decay and end of injection. Duration of valve closure determines the quantity of fuel injected. The unit shown uses camshaft/rocker arm driven plungers to generate the injection pressure, and employs needle-valve nozzles of conventional design. Increased flexibility in fuel metering and timing and simpler injector mechanical design are important advantages.¹

Figure 10.19 Examples of electronically controlled injectors. (a) Solenoid-valve injector: (1) fuel return; (2) solenoid coil; (3) solenoid armature; (4) valve ball; (5) valve control chamber; (6) nozzle-needle pressure shoulder; (7) injection orifice; (8) outlet restrictor; (9) high-pressure port; (10) inlet restrictor; (11) valve plunger. (b) Piezo injector: (1) fuel return; (2) high-pressure port; (3) piezo-actuator module; (4) hydraulic booster; (5) valve; (6) nozzle needle; (7) injection orifice.¹² (Courtesy Robert Bosch GmbH and SAE.)



A newer option is piezoelectric injectors (Fig. 10.19b), where a piezo-actuator drives the nozzle needle rather than solenoid-valve controlled fuel pressure. Here, an electrical pulse through multiple thin quartz crystals creates the force (pressure), which through a hydraulic booster moves the needle. This reduces needle size and weight, and friction within the injector, providing these advantages: more compact and lighter injectors; multiple injections; more precise control of each injection with shorter time intervals between injections. Injectors with pressures of close to 200 MPa produce better fuel atomization, and can provide up to five injections per cylinder per cycle. Use of piezo-injectors in diesel engines reduces engine noise and air-pollutant emissions significantly, and increases power output and efficiency.

Accurate predictions of fuel behavior within the injection system require sophisticated hydraulic models. Approximate estimates of the injection rate through injector nozzles can be made as follows. If the pressure upstream of the injector nozzle can be estimated or measured, and assuming the flow through each nozzle is quasi steady, incompressible, and one dimensional, the mass flow rate of fuel injected through the nozzle $\dot{m}_f$ is given by

$$\dot{m}_f = C_D A_n \sqrt{2\rho_f \Delta p}$$

(10.14)

where A_n is the nozzle minimum area, C_D the discharge coefficient, ρ_f the fuel density, and Δp the pressure drop across the nozzle. If the pressure drop across the nozzle and the nozzle open area are essentially constant during the injection period, the mass of fuel injected is then

$$m_f = C_D A_n \sqrt{2\rho_f \Delta p} \frac{\Delta\theta}{360N}$$

(10.15)

where $\Delta\theta$ is the nozzle open period in crank-angle degrees and N is engine speed. Equation (10.15) illustrates the dependence of injected amounts of fuel on key injection system and engine parameters.

10.4.2. Overall Spray Structure

The fuel is introduced into the combustion chamber of a DI diesel engine through one or usually several nozzles or orifices with a large pressure differential between the fuel supply line and the cylinder. Different designs of nozzle are used (e.g., single-orifice, multi-orifice, throttle, or pintle; see Fig. 1.27), depending on the needs and size of the combustion system employed. Diesel injectors usually operate with fuel-injection pressures between 1000 atm and 2500 atm (bar).^[1] At time of injection, the air in the cylinder has a pressure of 50 to 100 atm (bar), temperature about 1000 K, and density between kg/m^3 15 and 35 kg/m^3 . Nozzle hole diameters span the range of 0.1 to 0.3 mm diameter (automotive applications), with length/diameter ratios from 2 to 8. Typical distillate diesel fuel properties are: relative specific gravity of 0.8, viscosity between $3 \text{ kg/m} \cdot \text{s}$ and $10 \text{ kg/m} \cdot \text{s}$ and surface tension about $3 \times 10^{-2} \text{ N/m}$ (at 300 K). Figure 10.20 illustrates the basic features of a typical DI engine fuel spray. As the liquid jet leaves the nozzle it atomizes into fine drops whose momentum creates a spray, and spreads out as the spray entrains the surrounding air. The initial jet velocity is greater than 100 m/s. The outer surface of the liquid jet breaks up into drops of order 10- μm diameter or less, close to the nozzle exit as the liquid column leaving the nozzle rapidly disintegrates. As injected fuel moves away from the nozzle, the mass of air entrained into the spray increases, the spray diverges, its width increases, and its velocity decreases. The fuel drops evaporate as this hot-air entrainment process proceeds. The tip of the spray penetrates further into the combustion chamber as injection continues, but at a decreasing rate. Figure 10.21 shows photographs of a diesel fuel spray injected into quiescent air in a rapid-compression machine which simulates diesel conditions.¹³ Two different photographic techniques, back lighting and shadowgraph,^[2] have been used to distinguish the liquid-containing core of the jet and the extent of the fuel vapor region of the spray which surrounds the liquid core. The region of the jet closest to the nozzle (until injection ceases at 3.3 ms) contains liquid drops and ligaments; the major region of the spray consists of a substantial fuel vapor plus air cloud around this narrow core which contains the remaining liquid fuel.

Figure 10.20 Schematic of diesel fuel spray, defining (a) its major features and (b) its key parameters: spray axis, CL; penetration distance S ; full core angle θ ; projected spray area (grey) A_s .

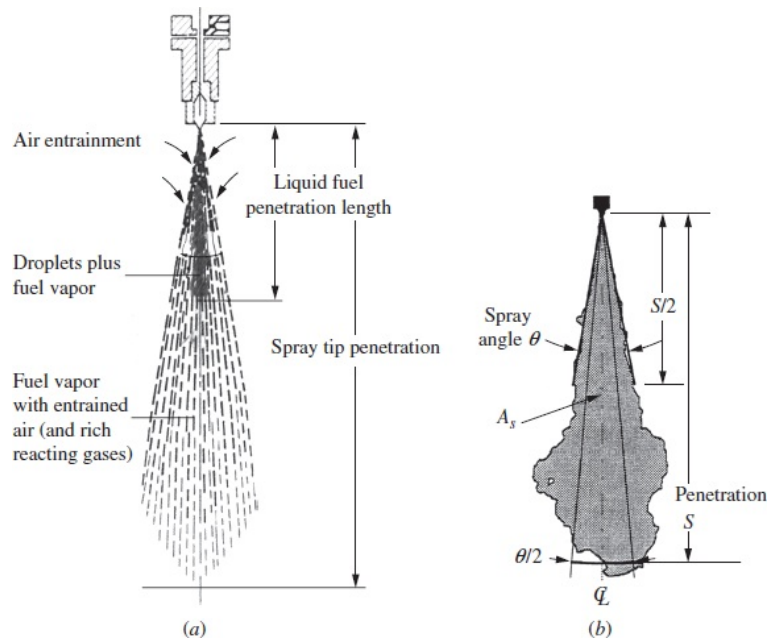
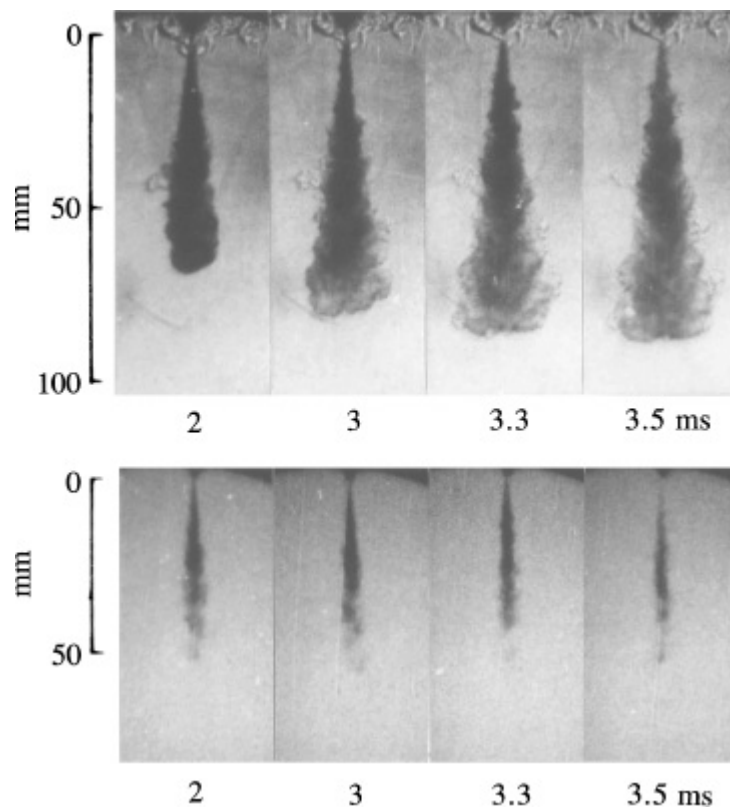


Figure 10.21 Shadowgraph and back-illuminated photographs of evaporating spray injected into nitrogen at 3.4 MPa and 670 K in rapid-compression machine. Times in milliseconds are after start of injection: injection duration 3.3 ms. Top (shadowgraph) photographs show full fuel vapor and liquid region. Bottom (back-illuminated) photographs show only the liquid-containing core.¹³

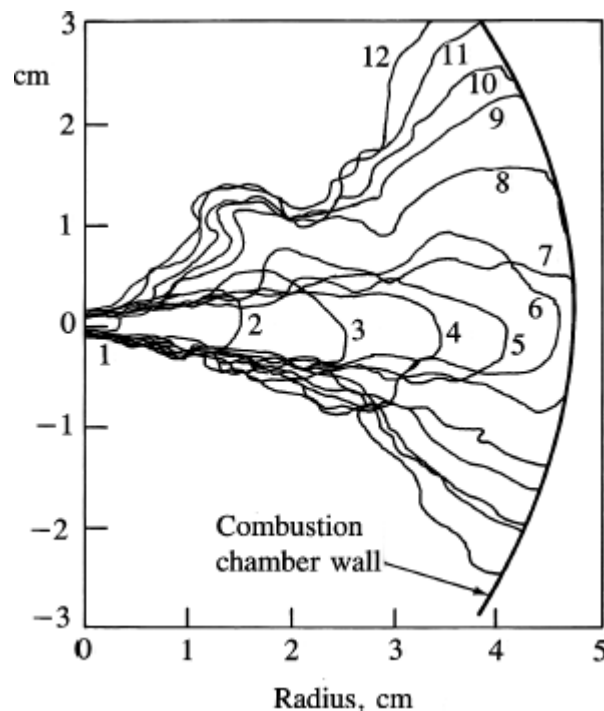


So far we have discussed a vaporizing fuel spray prior to the onset of combustion. Within this fuel-rich, fuel vapor plus entrained hot air spray, the autoignition chemistry occurs which raises the mixture temperature within the spray in the

"premixed rapid-burning" first phase of the combustion process. This then sets up the "mixing-controlled" second phase of the diesel burning process where a diffusion flame (lifted off the injector nozzle exit) surrounds the spray. These spray combustion phases are discussed in more detail in [Secs. 10.5](#) and [10.6](#).

Different spray configurations are used in the different diesel combustion systems described earlier in this chapter. The simplest configuration involves multiple sprays injected into quiescent air in the largest-size diesels ([Fig. 10.1a](#)). [Figures 10.20](#) and [10.21](#) illustrate the essential features of each spray under these circumstances until interactions with the wall occur. Each liquid fuel jet atomizes into drops and ligaments as fuel exits the nozzle orifice (or shortly thereafter). The spray immediately starts entraining air, spreads out and slows down as the mass flow in the spray increases. The droplets on the edges of the spray evaporate first, creating a fuel vapor-air mixture sheath around the liquid-containing core. The highest velocities are on the jet axis. The equivalence ratio is highest on the centerline and fuel-rich throughout, and drops to zero (unmixed air) at the spray boundary. Once the sprays have penetrated to the outer regions of the combustion chamber, they interact with the chamber walls. The spray is then forced to split and flow tangentially along the wall. Eventually the sprays from multi-hole nozzles interact with one another. [Figure 10.22](#) shows a diesel fuel spray interacting with the cylindrical outer wall of a disc-shaped combustion chamber in a rapid-compression machine under typical diesel-injection conditions. The cylinder wall causes the spray to split with about half flowing circumferentially in either direction. Adjacent sprays then interact forcing the flow radially inward toward the chamber axis.¹⁴

Figure 10.22 Outlines of the outer vapor boundary of a diesel fuel spray from 12 successive frames of a rapid-compression-machine high-speed shadowgraph movie showing interaction of vaporizing spray with the cylindrical wall of the combustion chamber. Injection pressure 60 MPa. Time between frames 0.14 ms.¹⁴



Most of the other combustion systems in [Figs. 10.1](#) and [10.2](#) use air swirl to increase fuel-air mixing rates. A schematic of the spray pattern which results when a fuel jet is injected radially outward into a swirling flow is shown in [Fig. 10.23](#). Because there is now relative motion in both radial and tangential directions between the initial jet and the air, the structure of the jet is more complex. As the spray entrains air and slows down it becomes increasingly bent toward the swirl direction; for the same injection conditions it will penetrate less with swirl than without swirl. An important feature of the spray is the large vapor-containing region downstream of the liquid-containing core. [Figure 10.24](#) shows schlieren photographs of four fuel jets injected on the axis of an IDI diesel engine prechamber with high clockwise swirl. The photograph on the left, with high sensitivity, shows the outer boundary of the fuel vapor region of the spray; the low-sensitivity photograph on the right locates

the liquid phase regions of the spray.¹⁵ The interaction between the swirl and both liquid and vapor spray regions is evident, and changes how the spray interacts with the chamber wall.

Figure 10.23 Schematic of fuel spray injected radially outward from the chamber axis into a swirling airflow. Distribution of fuel-air equivalence ratio (ϕ) within the spray is indicated.

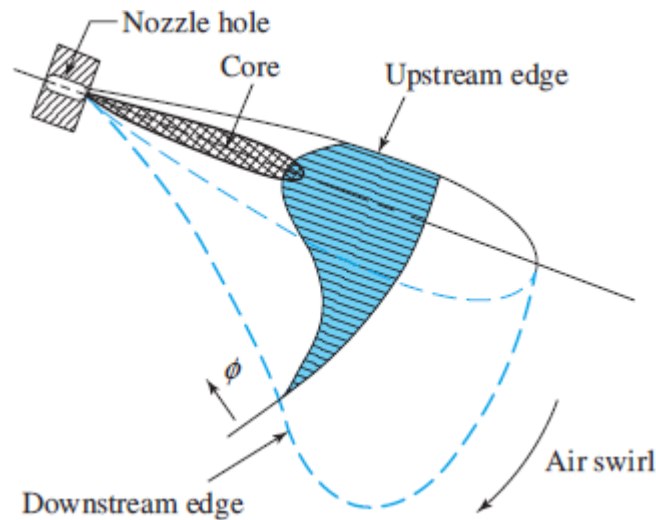
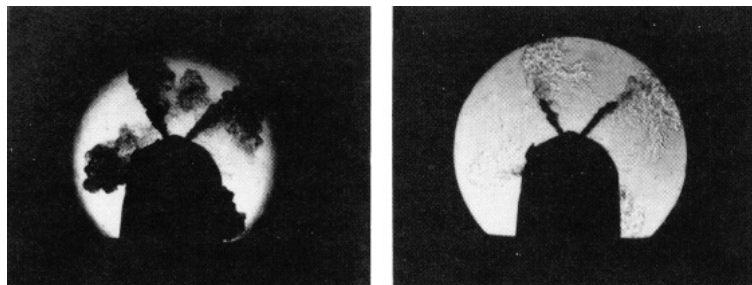


Figure 10.24 Schlieren photographs of vaporizing sprays injected into the swirling airflow in a transparent prechamber of special IDI diesel. Left: high-sensitivity image showing boundaries of the vapor regions of spray. Right: low-sensitivity image showing liquid-containing core (dark) in relation to vapor regions (mottled).¹⁵



To couple the spray-development process with the ignition phase of the combustion, it is important to know which regions of the spray contain the fuel injected early in the injection process since it is these regions that are likely to autoignite first. Each spray develops as follows. At the start of injection liquid fuel enters the quiescent air charge, almost immediately atomizes into small drops that move outward from the nozzle, and slow down rapidly as air is entrained into the spray and accelerated. This start-up process forms a vortex or "puff" at the head of the spray. The injected fuel which follows encounters less resistance; thus drops from that fuel overtake the drops from first-injected fuel forcing them outward toward the periphery of the spray. At the tip of this developing spray, the drops with their surrounding fuel vapor and entrained air meet high aerodynamic resistance and slow down, but the spray continues to penetrate the air charge because this retarded mixture at the tip of the spray is continually replaced by new higher-momentum later-injected drops.¹⁶ As Figs. 10.21 and 10.24 indicate, these drops evaporate quickly.

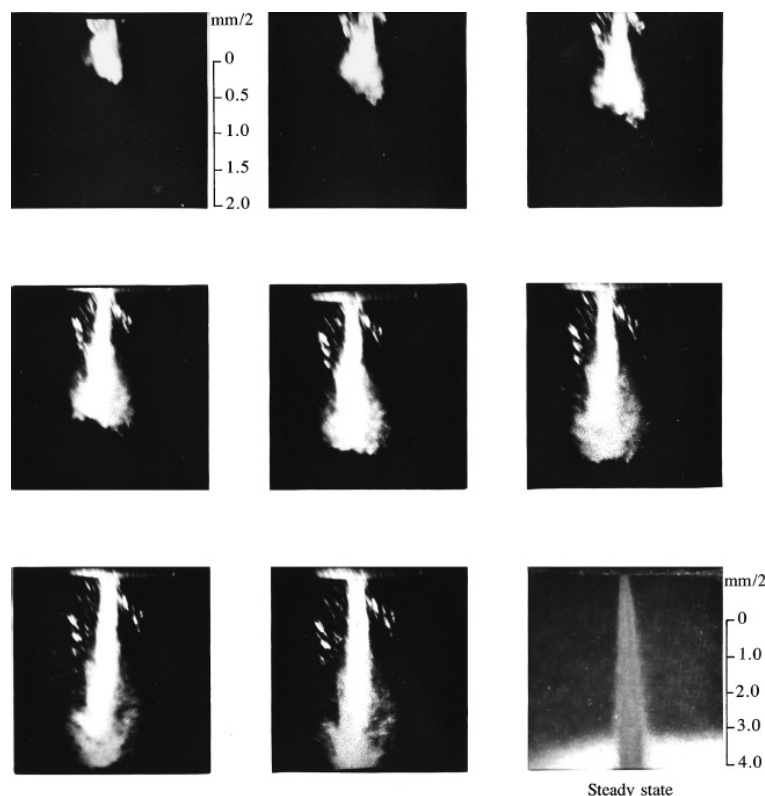
10.4.3. Atomization and Spray Development

Under diesel engine injection conditions, the liquid fuel jet exiting the nozzle breaks up rapidly and forms a cone-shaped spray. This type of behavior falls into the atomization breakup regime, and produces droplets with sizes much less than the

nozzle exit diameter. This behavior is different from other modes of liquid jet breakup. At low jet velocity, in the Rayleigh regime, the liquid jet breakup is due to the unstable growth of surface waves caused by surface tension and results in drops larger than the jet diameter. As jet velocity is increased, forces due to the relative motion of the jet and the surrounding air augment the surface tension force, and lead to drop sizes of the order of the jet diameter. This is called the first wind-induced breakup regime. A further increase in jet velocity results in breakup characterized by divergence of the jet spray after an intact or undisturbed length downstream of the nozzle. In this second wind-induced breakup regime, the unstable growth of short-wavelength waves induced by the relative motion between the liquid and surrounding air produces droplets whose average size is less than the jet diameter. Further increases in jet velocity lead to the atomization regime, where the breakup of the outer surface of the jet occurs at (or before) the nozzle exit plane and results in droplets whose average diameter is much smaller than the nozzle diameter. Aerodynamic interactions at the liquid/gas interface appear to be one major component of the atomization mechanism in this regime.^{16–18}

A sequence of very short time exposure photographs of the emergence of a liquid jet from a nozzle of 0.34 mm diameter and $L_n/d_n = 4$, into high-pressure nitrogen at ambient temperature, is shown in Fig. 10.25. The figure shows how the spray tip penetrates and the spray spreads during the early part of its travel.¹⁸ Data such as these have been used to examine the dependence of the spray development on gas and liquid density, liquid viscosity, and nozzle geometry.^{18–22} The effects of the most significant variables, gas/liquid density ratio, and nozzle geometry on initial jet spreading angle are shown in Fig. 10.26. For a given geometry (cylindrical hole with sharp corners), the initial jet spreading or spray angle increases with increasing gas/liquid density ratio as shown in Fig. 10.26a for non-vaporizing sprays, typical density ratios for diesel injection conditions are between 0.015 and 0.05. Of several different nozzle geometry parameters examined, the length/diameter ratio proved to be the most significant (see Fig. 10.26b).

Figure 10.25 Photographs showing the emergence (and steady state image, bottom right) of a high-pressure liquid spray from a nozzle orifice. Time between frames 2.1 μ s. Liquid: water; gas: nitrogen at 1380 kPa, Δp across nozzle = 11 MPa, nozzle diameter = 0.34 mm.¹⁹



The roles of various parameters on the spreading angle of a diesel spray through their effect on the fuel atomization and turbulent air-entrainment processes are not yet well understood. Figure 10.26 compares non-vaporizing spray development

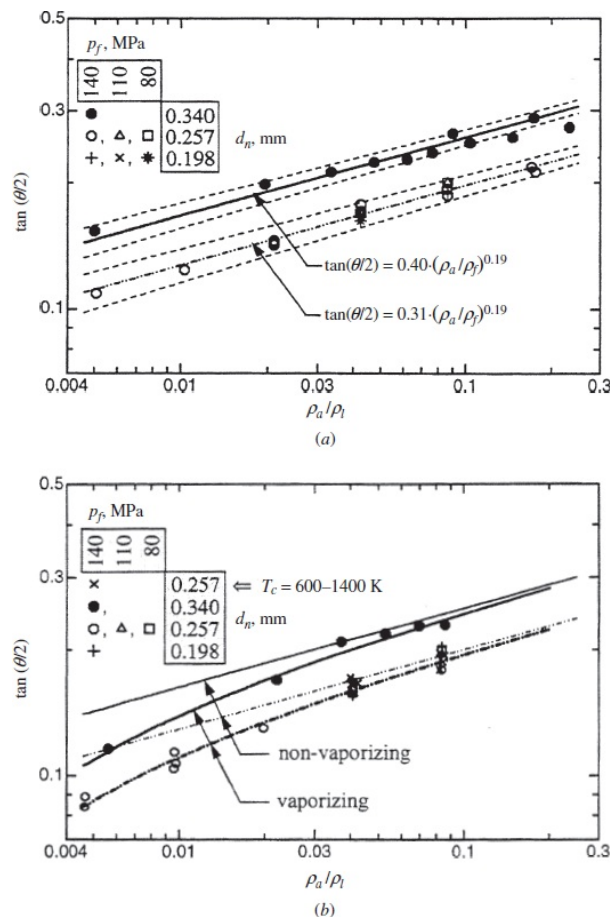
(fuel injected into a cold ambient gas) with vaporizing (hot ambient gas—but without combustion), as a function of the gas to liquid density ratio. The spreading angles shown are the spreading angles of the quasi-steady portion of the fuel jet; i.e., its upstream portion. The jet spreading angle (defined by the tangent of $\theta/2$; see Fig. 10.20b) has been correlated empirically by:²²

$$\tan(\theta/2) = c_1 \left[\left(\frac{\rho_a}{\rho_f} \right)^{0.19} - c_2 \left(\frac{\rho_f}{\rho_a} \right)^{0.5} \right]$$

(10.16)

where ρ_a and ρ_f are ambient gas and liquid fuel densities, c_1 is a constant depending on the nozzle geometry (between 0.26 and 0.4), and c_2 is zero for non-vaporizing sprays and 0.0043 for vaporizing sprays.^{21,23} Equation (10.16) is the basis for the correlating lines in Fig. 10.26: Note that the behavior of vaporizing and non-vaporizing sprays is different (even at the same ρ_a/ρ_f ratio). Vaporization of the fuel droplets within the spray cools the entrained air through the fuel's heat of vaporization requirement. This cooling results in a higher density mixture in the jet relative to the non-vaporizing situation which, based on jet momentum conservation, results in slower penetration. A competing effect is that air entrainment is reduced by vaporization resulting in slower jet deceleration and reduced spreading angles.²²

Figure 10.26 (a) Measured spray angles of non-vaporizing sprays versus density ratio ρ_a/ρ_f = gas density/liquid density for several nozzle geometries and liquid injection pressures of 80 to 140 MPa. Nozzle diameters $d_n = 0.2$ to 0.34 mm. (b) Measured spray angles versus density ratio for a range of nozzle geometries. $d_n = 0.2$ to 0.34 mm diameter, same range of injection pressures. Correlating Eq. (10.16) also shown.^{21,23}



Jet breakup trends can be summarized as follows. The initial jet divergence angles increase with increasing gas density. Divergence begins progressively closer to the nozzle as gas density increases. Jet divergence angles increase with decreasing fuel viscosity; divergence begins at the nozzle exit once the liquid viscosity is below a certain level. Nozzle design affects the onset of the jet atomization regime. Jet divergence angles decrease with increasing nozzle length. For the same length, rounded inlet nozzles produce less divergent jets than sharp-edged inlet nozzles. The initial jet divergence angle and intact spray length are quasi steady with respect to changes in operating conditions which occur on time scales longer than about 20 μs .¹⁹ Note that while all these results were obtained under conditions where evaporation was not occurring, the atomization and initial spray-development processes are not that significantly affected by evaporation (see [Sec. 10.4.6](#)).

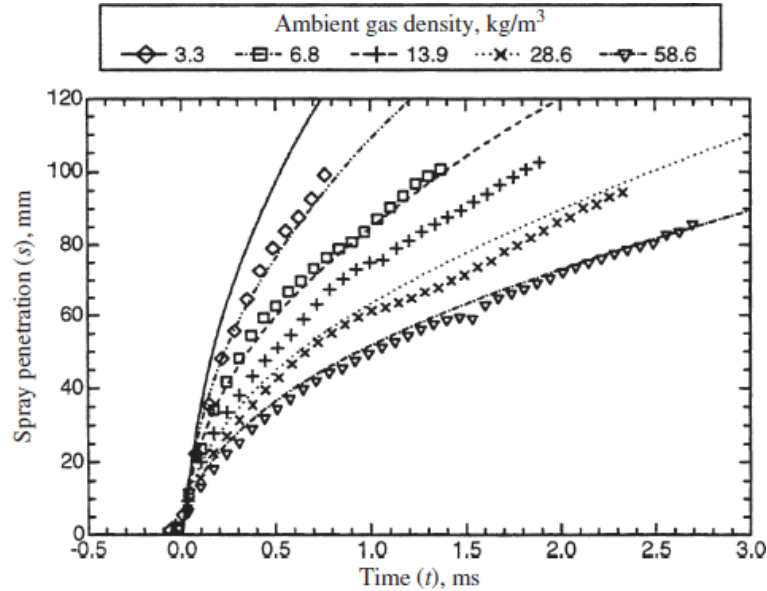
10.4.4. Spray Penetration

The speed and extent to which the fuel spray penetrates across the combustion chamber has an important influence on air utilization and fuel-air mixing rates. In some engine designs, where the walls are hot and high air swirl is present, fuel impingement on the walls is desired. However, in multi-spray DI diesel combustion systems, over-penetration results in impingement of liquid fuel on cool surfaces which, especially with little or no air swirl, lowers mixing rates and increases emissions of unburned and partially burned species. Yet under-penetration results in poor air utilization since the air on the periphery of the chamber does not then contact the fuel. Thus, the penetration of liquid fuel sprays under conditions typical of those found in diesel engines has been extensively studied.

Many correlations based on experimental data and turbulent gas jet theory has been proposed for fuel spray penetration. These predict the penetration S of the fuel spray tip across the combustion chamber for injection into quiescent air as occurs in larger DI engines, as a function of time. More recent work^{21,22} has compared vaporizing sprays with non-vaporizing sprays. For short times and penetrations, and long times and penetrations, the differences in jet penetration are small. In between, the differences are more significant.

Studies have examined spray tip penetration as a function of time, following the start of a diesel injection process in high-pressure constant-volume vessels.²² The data shown in [Fig. 10.27](#) illustrate the sensitivity of the spray tip position as a function of time to the ambient gas state for fuel jets injected into quiescent gas at a temperature of 1000 K. These data show the impact of changes in ambient gas density (ρ_a), the parameter that defines the ambient resistance to liquid penetration. Injection pressure, especially at short time scales, is also significant: higher injection pressures result in faster penetration due to the higher initial liquid- fuel momentum. See Eq. ([10.14](#)).

Figure 10.27 Measured penetration data as a function of time for vaporizing fuel sprays (symbols). Lines show predicted penetrations from Eq. (10.17) for non-vaporizing sprays. The nozzle orifice diameter, pressure drop across the orifice, and ambient gas temperature were 241 μm , 137 MPa, and 1000 K.²¹



A scaling law for non-vaporizing (isothermal) spray penetration developed by Naber and Siebers,²¹ expressed in dimensionless form is:

$$\tilde{t} = \frac{\tilde{S}}{2} + \frac{\tilde{S}}{4} \sqrt{1 + 16\tilde{S}^2} + \frac{1}{16} \ln \left(4\tilde{S} + \sqrt{1 + 16\tilde{S}^2} \right)$$

(10.17)

where the dimensionless time and penetration length are given by

$$\tilde{t} = t/t^+ \quad \text{and} \quad \tilde{S} = S/x^+$$

(10.18a,b)

t^+ and x^+ are the characteristics time and length scales of the fuel spray defined by:

$$t^+ = \frac{d_n \sqrt{\tilde{\rho}}}{\tan(\alpha/2) U_f} \quad \text{and} \quad x^+ = \frac{d_n \sqrt{\tilde{\rho}}}{\tan(\alpha/2)}$$

(10.19a,b)

The effective orifice diameter d_n , density ratio $\tilde{\rho}$, and liquid fuel velocity at nozzle exit U_f , are given by:

$$d_n = \sqrt{C_a d}, \quad \tilde{\rho} = \rho_f / \rho_a,$$

(10.20a,b)

and

$$U_f = C_v \sqrt{2(p_f - p_a) / \rho_f}$$

(10.20c)

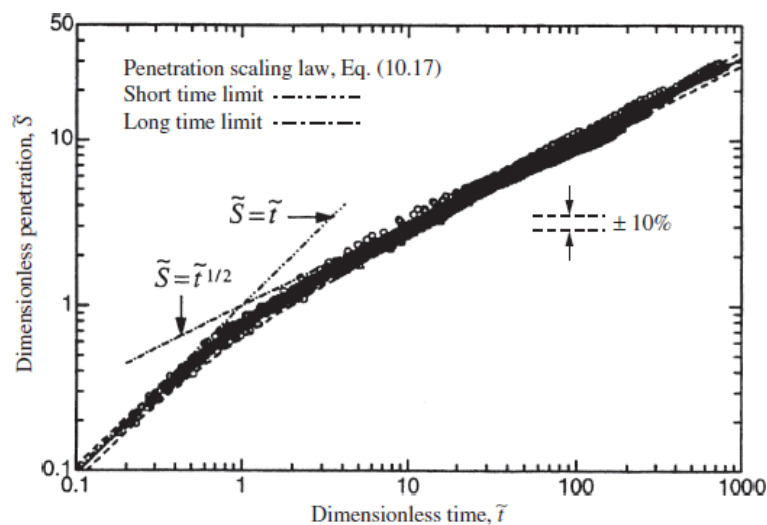
d is the fuel nozzle orifice diameter, C_a and C_v are the orifice area contraction coefficient and velocity coefficient, respectively, and p_f and p_a are the fuel injector pressure and ambient gas pressure.

Figure 10.28 compares non-vaporizing fuel spray penetration data with the dimensionless correlation [Eq. (10.17)]. Note the short-time and long-time penetration limits have a square-root-of-time, and a linear-time dependence, respectively, as shown. The transition occurs at

$$\tilde{t} = 1 \text{ and}$$

$\tilde{S} = 0.7$ where the spray behavior transitions from being dominated by the injected fuel to being dominated by entrained gas.

Figure 10.28 Comparison of measured non-vaporizing fuel spray penetration data with the penetration scaling law [Eq. (10.17)]. The data cover a wide range of conditions: the correlation predicts the data well.^{21,22}



An inverse version of Eq. (10.17), shown below as Eq. (10.21), gives the penetration distance $\tilde{S}$ as a function of time directly:

$$\tilde{S} = \tilde{t} / (1 + \tilde{t}^{n/2})^{1/n}$$

(10.21)

where $n = 2.2$. This agrees closely with Eq. (10.17) in the near-field and far-field limits, and is within 5% in the transition region.²²

Figure 10.29 shows the effect of combustion on spray development. The data were taken in the same constant-volume vessel with a single injected fuel spray. The ambient gas here contains oxygen to enable combustion. The lines are fits to the vaporizing spray data, where the ambient gas was not (1000 K), as with the combusting spray data but contained no oxygen.

The combusting sprays (especially at the higher ambient gas densities) spread more quickly due to the chemical energy release within the spray which decreases the spray density (by raising the mixture temperature), thus causing the mixture within the spray to expand and move forward faster.

Figure 10.29 Penetration data for combustion fuel sprays as a function of time for a range of ambient gas densities. The orifice diameter, pressure drop across the injector orifice, and ambient gas temperature are the same as in Fig. 10.27. Curves are fits to vaporizing fuel spray penetration data in Fig. 10.27.^{21,22}

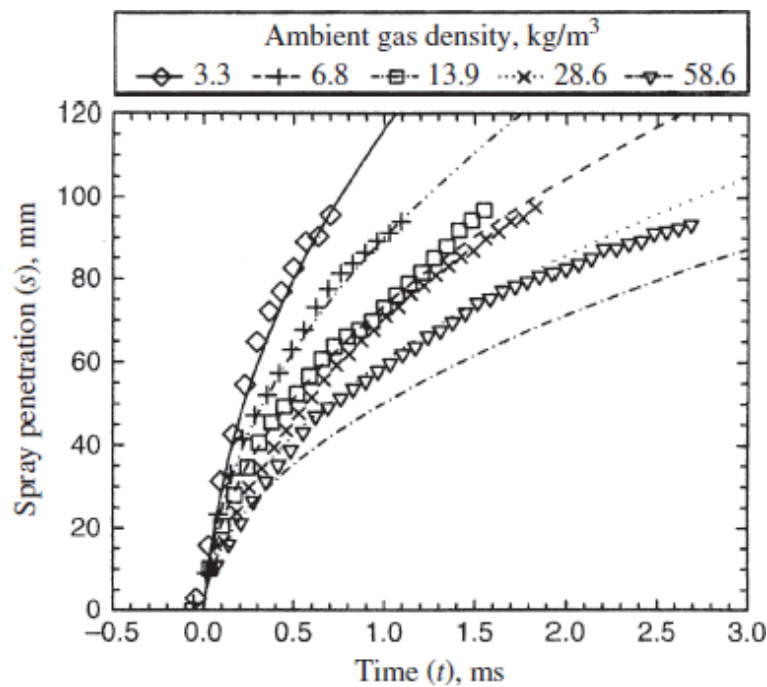


Figure 10.30 compares penetration data from vaporizing fuel sprays and combusting (and vaporizing) sprays on dimensionless

$\tilde{S}$ versus

$\tilde{t}$ plots, with the correlation for non-vaporizing sprays, Eq. (10.17). Under vaporizing but non-combusting conditions, Fig. 10.30a, the dimensionless spray penetration is slower by as much as 20% under the lower ambient density conditions (see Fig. 10.27). Fuel droplet evaporation cools the mixture within the spray thereby increasing its density and, through conservation of momentum, slowing it down. This is partly offset by reduced entrainment into the spray (as evidenced by the decrease in jet spreading angle; see Fig. 10.26b), which would speed the spray up. The vaporizing and combusting spray behavior, Fig. 10.30b, (which of course is the situation in an actual engine) is similar to that of the vaporizing but non-combusting spray up to dimensionless time

$\tilde{t}$ up to about 10 (

$\tilde{S}$ about 2). Combustion has no impact until spontaneous ignition occurs and chemical energy release is initiated. The times of ignition for the higher three ambient gas density conditions in Fig. 10.29 are 0.6, 0.35, and 0.25 ms for the mid to highest density range, respectively. For times beyond these ignition times, combustion has a growing impact. The spray then decelerates increasingly more slowly as evidenced by the data beyond

$\tilde{t}$ of about 10 to 20 (when ignition typically occurs). Ultimately, as

$\tilde{t}$ increases above about 30, the combusting jet penetrates about 20% further than would an equivalent non-vaporizing jet.²²

This is likely due to the decreasing density of the within-the-spray mixture as hot combustion products produced within the diffusion flame around the spray diffuse into (and outward) from the spray boundary (see photos, Fig. 10.5). Reducing the spray density, since spray momentum is conserved, increases the spray mixture velocity. Flame height measurements of burning fuel sprays indicate that combustion is typically complete by

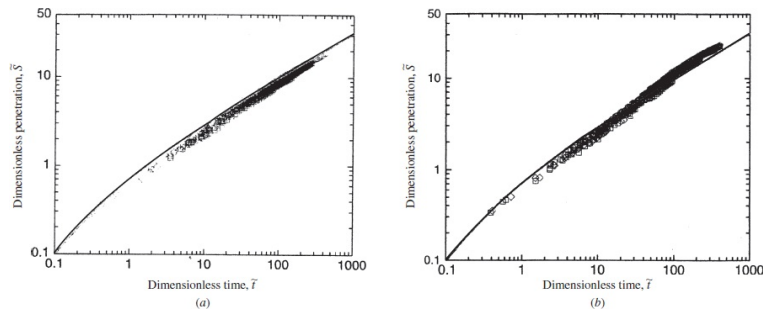
$\tilde{t}$ of about 270 and

$\tilde{S}$ of 16. For

$\tilde{t}$ greater than this end-of-diffusion-flame time, the combusting spray and the vaporizing spray trends should converge.

Figures 10.30a and b indicate that is occurring.

Figure 10.30 (a) Comparison between spray penetration measurements for vaporizing sprays under typical diesel conditions and the penetration scaling law for non-vaporizing sprays, Eq. (10.17), in dimensionless penetration versus dimensionless time coordinates. (b) Similar comparison between combusting spray data, and non-vaporizing spray correlation. Conditions for two sets of data are the same except that ambient gas in the combusting experiments had 21% oxygen.²²



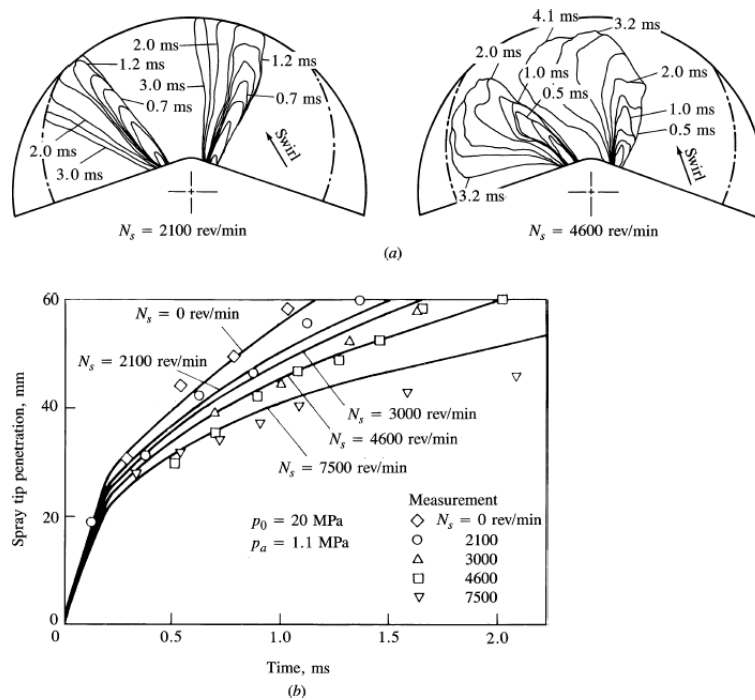
In medium duty, and especially in high-speed light-duty DI diesel engines, air swirl is used to increase the fuel-air mixing rates. Increasingly higher swirl (rotation rate about the cylinder axis), created during intake and enhanced as the in-cylinder air is forced radially inward toward the cylinder axis during compression with reduced bowl diameter/bore ratios and deeper bowls (see Fig. 10.1), is used as cylinder size decreases and maximum engine speed increases. This enhanced air swirl affects the behavior of the four to six fuel sprays injected outward into the piston bowl from the cylinder axis.

The effect of air swirl on spray penetration is important. Figure 10.31 shows how the spray shape and location change as swirl is increased: it also shows how spray tip penetration varies with time and swirl rate.²³ These authors related their data on spray tip penetration with swirl, S_s , through a correlation factor to the equivalent penetration S , without swirl [given by an earlier equivalent to Eq. (10.17)]:

$$\frac{S_s}{S} = \left(1 + \frac{\pi R_s N S}{30 v_j} \right) \quad (10.22)$$

Here R_s is the swirl ratio which equals the swirl rate in revolutions per minute N_s divided by the engine speed N (rev/min), and v_j is the initial fuel jet velocity (meters per second). The curves in Fig. 10.31b correspond to Eq. (10.22). Swirl both reduces the radial penetration of the spray and spreads out the spray in the swirl direction, more rapidly.

Figure 10.31 (a) Measured outer boundary of sprays injected into swirling airflow and (b) spray tip penetration as a function of time for different swirl rates. Solid lines show Eq. (10.22).²³



10.4.5. Droplet Size Distribution

Previously in [Sec. 10.4](#), we have discussed the overall characteristics of the diesel fuel spray—its spreading rate and penetration into the combustion chamber. While the distribution of fuel via the spray trajectory throughout the combustion chamber is important, atomization of the liquid fuel into many small rapidly moving drops is also necessary to create a large surface area across which liquid fuel can evaporate. Here we review how the drop sizes in the spray depend on injection parameters, and the air and fuel properties. Since the measurement of droplet characteristics in an operating diesel engine is extremely difficult, most results have come from studies of fuel injection into constant-volume chambers filled with high-pressure quiescent air, often at room temperature.

During the injection period, the injection conditions such as injection pressure, nozzle orifice area, and injection rate may vary. Consequently, the droplet size distribution at a given location in the spray may also change with time during the injection period. In addition, since the details of the atomization process are different in the spray core and at the spray edge, and the trajectories of individual drops depend on their size, initial velocity, and location within the spray, the drop size distribution is likely to vary with position within the spray. None of these variations has yet been adequately quantified.

The aerodynamic theory of jet breakup in the atomization regime summarized in [Sec. 10.4.3](#) (based on work by G. I. Taylor) leads to the prediction that the initial average drop diameter D_d is proportional to the length of the most unstable surface waves:¹⁶

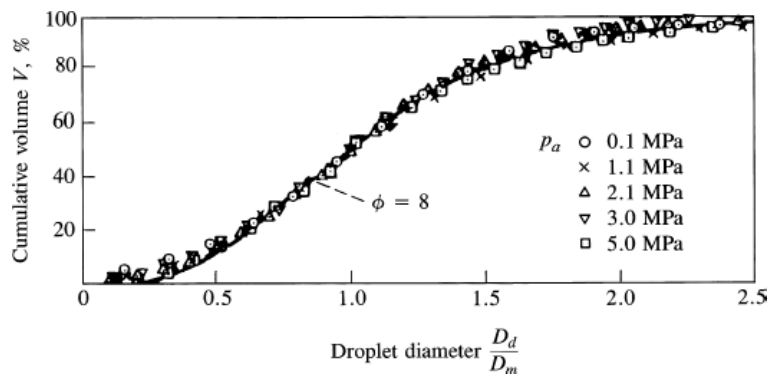
$$D_d = C \frac{2\pi\sigma}{\rho_g v_r^2} \lambda^* \quad (10.23)$$

Here σ is the liquid-fuel surface tension, ρ_g is the gas density, v_r is the relative velocity between the liquid and gas (taken as the

mean injection velocity v_j), C is a constant of order unity, and λ^* is the dimensionless wavelength of the fastest growing wave. λ^* is a function of the dimensionless number $(\rho_l/\rho_g)(Re_j/We_j)^2$ where the jet Reynolds and Weber numbers are given by $Re_j = \rho_l v_j d_n / \mu_l$ and $We_j = \rho_l v_j^2 d_n / \sigma$ and d_n is the nozzle orifice diameter. λ^* goes to $3/2$ as this number increases above unity. Near the edge of the spray close to the nozzle, this equation predicts observed drop size trends with respect to injection velocity, fuel properties, nozzle L/d , and nozzle diameter, though measured mean drop sizes are larger by factors of 2 to 3.²⁵ However, within the dense early region of the spray, secondary atomization phenomena—coalescence and breakup—occur which will change the droplet size distribution and mean diameter. The downstream drop size in the conical sprays used in diesel-injection systems is markedly influenced by both drop coalescence and breakup. Eventually a balance is reached as coalescence decreases (due to the expansion of the spray) and breakup ceases (due to the reduced relative velocity between the drops and the entrained gas).²⁶

Measurements of droplet size distributions within simulated diesel sprays indicate how droplet size varies with location. Figure 10.32 shows several normalized cumulative drop size distributions superposed, which define a close-to-universal curve.

Figure 10.32 Normalized drop-size cumulative frequency distribution in spray injected into ambient-temperature air for air pressures from 0.1 to 5 MPa. Throttling pintle nozzle with nozzle opening pressure of 9.9 MPa. D_d is initial average drop diameter [see Eq. (10.23)]; median drop diameter $D_m = 1.22 D_{SM}$.²⁴



Drop size distributions are usually expressed as frequency distributions of drop volume. If $f(D_d)dD_d$ is the probability of finding a drop with diameter between D_d and $D_d + dD_d$, then

$$\int_0^{\infty} f(D_d) dD_d = 1 \quad (10.24)$$

and the volume distribution is

$$\frac{1}{V} \frac{dV}{dD_d} = \frac{f(D_d) D_d^3}{\int_0^{\infty} f(D_d) D_d^3 dD_d} \quad (10.25)$$

With high-pressure DI diesel fuel injection systems [e.g., 1400 bar injection pressure, 200 mm³ (160 mg) fuel delivered per cycle, 8 nozzle holes per cylinder with 0.2 mm orifice diameter], for a 12.9-liter medium-duty diesel engine rated at 290 to 300

kW at 2000 rev/min, and with fuel exiting the nozzle at a velocity of about 250 m/s, the average initial drop sizes is about 10 μm diameter. As the spray moves away from the injector orifices, the average drop size increases presumably due to some droplet coalescence and vaporization of the smaller drops in the distribution. The average drop size decreases as we move radially outward from the spray axis, presumably due to higher local concentrations of (hotter) entrained air, and higher droplet evaporation rates.

To characterize the spray, expressions for drop size distribution and mean diameter are needed. An appropriate and commonly used mean diameter is the *Sauter mean diameter*:

$$D_{SM} = \left(\int D_d^3 dn \right) / \left(\int D_d^2 dn \right)$$

(10.26)

where dn is the number of drops with diameter D_d between $D_d - dD_d/2$ and $D_d + dD_d/2$. The integration is usually carried out by summing over an appropriate number of drop size groups. The Sauter mean diameter is the diameter of the drop that has the same surface/volume ratio as that of the total distribution.

Various expressions for the distribution of drop sizes in liquid sprays have been proposed. One, by Hiroyasu and Kadota,²⁴ is based on the chi-square statistical distribution fit to the available experimental data. Figure 10.32 shows that the chi-square distribution with a degree of freedom equal to 8 fits well to the measurements. Here, D_m is the median drop diameter which for this chi-square curve is $1.22D_{SM}$. The non-dimensional expression for drop size distribution in sprays injected through hole nozzles, pintle nozzles, and throttling pintle nozzles given by the chi-square distribution is

$$\frac{dV}{V} = 13.5 \left(\frac{D_d}{D_{SM}} \right)^3 \exp \left[-3 \left(\frac{D_d}{D_{SM}} \right) \right] d \left(\frac{D_d}{D_{SM}} \right)$$

(10.27)

An empirical expression for the Sauter mean diameter D_{SM} (in micrometers) for typical diesel fuel properties given by Hiroyasu and Kadota²⁴ is

$$D_{SM} = A(\Delta p)^{-0.135} \rho_a^{0.121} V_f^{0.131}$$

(10.28)

where Δp is the mean pressure drop across the nozzle in megapascals, ρ_a is the air density in kilograms per cubic meter, and V_f is the amount of fuel delivered per cycle per cylinder in cubic millimeters per stroke. A is a constant which equals 25.1 for pintle nozzles, 23.9 for hole nozzles, and 22.4 for throttling pintle nozzles. Other expressions for predicting D_{SM} can be found in Ref. 27.

The effects of injection pressure, nozzle geometry and size, air conditions, and fuel properties on Sauter mean diameter in sprays obtained with diesel fuel-injection nozzles have been extensively studied. Various immersion, photographic, and optical techniques for making such measurements have been used. Some of the major effects are illustrated in Figs. 10.33 and 10.34, which show average Sauter mean diameters determined optically from studies of steady diesel fuel sprays in a pressurized vessel. Figure 10.33 shows that nozzle size effects the mean drop size in the expected direction. Nozzle length/diameter ratio is also shown to be important: an $L_n/d_n = 4$ gives the minimum mean drop size at low and intermediate injection pressures. This L_n/d_n also corresponds to the minimum value of spray breakup length and to the maximum spray

cone angle. Fuel viscosity and surface tension also affect mean drop size as shown in Fig. 10.34, with the effects being most significant at lower injection pressures.

Figure 10.33 Effect of fuel-injection pressure and nozzle geometry and size on Sauter mean drop diameter: (a) effect of nozzle length/diameter ratio L_n/d_n and injection pressure and (b) effect of nozzle diameter d_n and injection pressure.²⁸

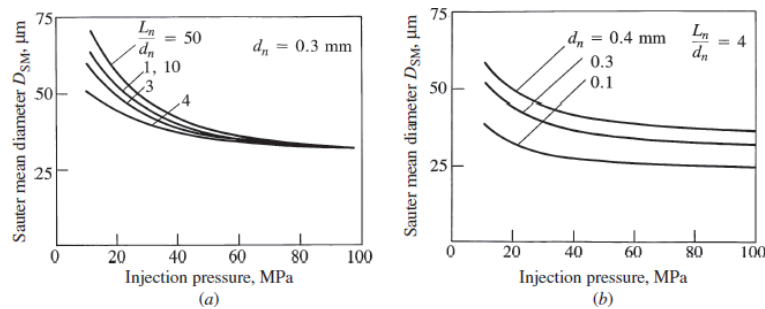
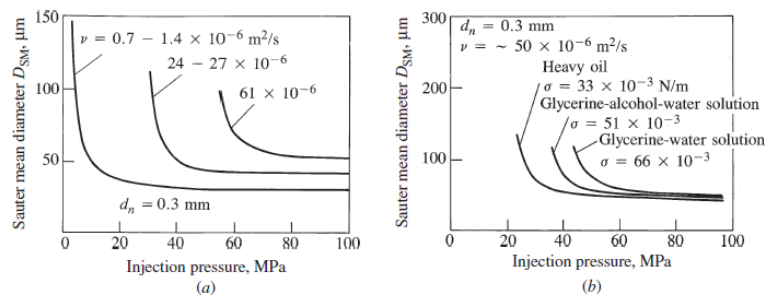


Figure 10.34 Effect of (a) liquid viscosity ν and (b) liquid surface tension σ on Sauter mean drop diameter as a function of injection pressure. Air conditions: 3 MPa and ambient temperature.²⁷



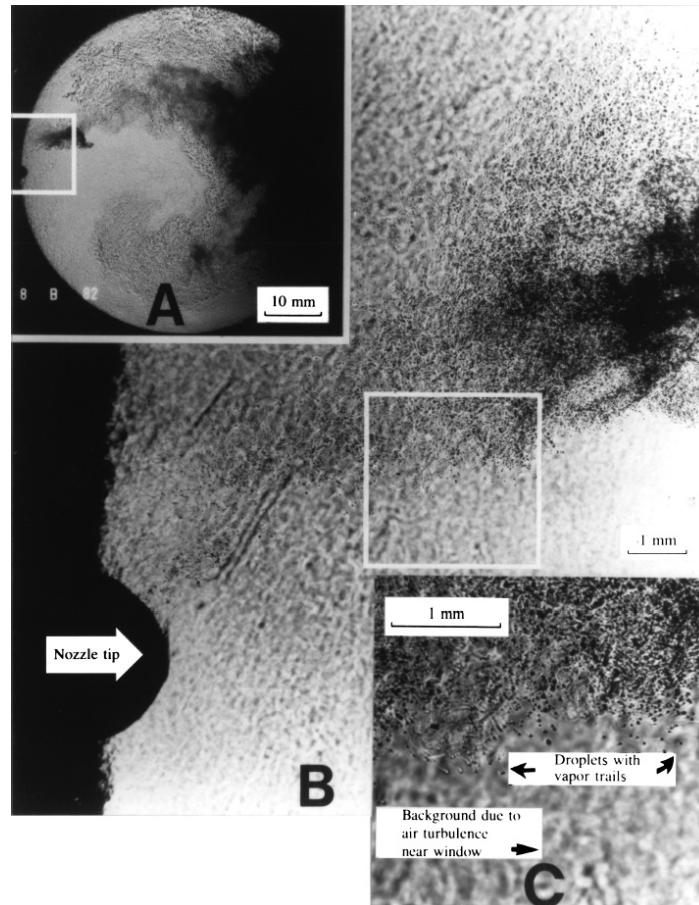
10.4.6. Spray Evaporation

The injected liquid fuel, atomized into small drops near the nozzle exit to form a spray, must evaporate before it can mix with air and burn. Figure 10.21 showed the basic structure of an evaporating fuel spray under conditions typical of a large DI diesel engine. Back illumination showed that a core exists along the axis of the spray where the liquid fuel ligaments or drops are sufficiently dense to attenuate the light beam. Once the start-up phase of the injection process is over, the length of this liquid-containing core remains essentially constant until injection ends. This core is surrounded by a much larger vapor-containing spray region which continues to penetrate deeper into the combustion chamber, the core extends only partway to the spray tip.

Additional insight into the physical structure of evaporating sprays can be obtained from the schlieren photographs taken just after the end of injection in a prechamber engine with air swirl, shown in Fig. 10.35. The lowest magnification picture (A) shows the overall structure of the spray. The only liquid-containing region evident is that part of the core nearest the nozzle which shows dark on the left of the photograph. The spreading vapor region of the spray, carried around the chamber by the swirling airflow, appears mottled due to local turbulent vapor concentration and temperature fluctuations. The dark region within the spray vapor region is due to soot formed where the fuel vapor concentration is sufficiently high. It is probable that, after the breakup length, the dense liquid core of the spray is primarily composed of droplets but the concentration is so high along the optical path that the light beam is fully extinguished. However, the last part of the core close to the nozzle tip (B) disperses sufficiently for individual features to be resolved. The small black dots are liquid fuel drops in the size range of 20 to 100 μm . Fuel drop vapor trails can be observed in the highest magnification photo (C) corresponding to various stages of evaporation. These range from drops showing little surrounding vapor to vapor trails with little liquid remaining at the head.

The vapor trails show random orientations relative to the spray axis, presumably due to local air turbulence. The process of droplet evaporation under normal engine operating conditions appears to be rapid relative to the total combustion period.¹⁵

Figure 10.35 Shadowgraph photographs at three magnifications taken just after the end of injection of a diesel-fuel spray into swirling airflow in the prechamber of a diesel engine with optical access. Nozzle hole diameter = 0.25 mm.¹⁵

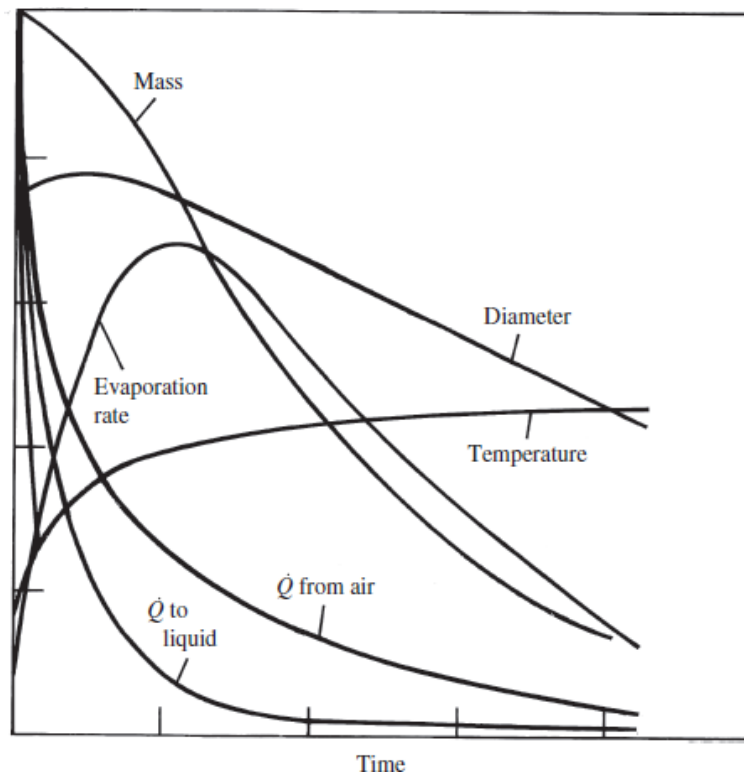


Let us now examine the drop evaporation process in more detail. Consider a liquid drop at close to ambient temperature injected into air at typical end-of-compression engine conditions. Three phenomena will determine the history of the drop under these conditions:

1. Deceleration of the drop due to aerodynamic drag
2. Heat transfer to the drop from the air
3. Mass transfer of vaporized fuel away from the drop

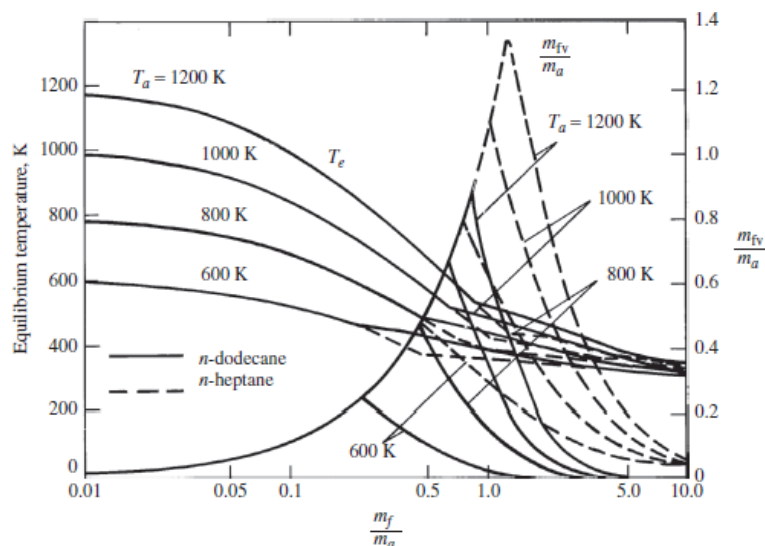
As the droplet temperature increases due to heat transfer from the hot surrounding air, the fuel vapor pressure increases and the evaporation rate increases. As the mass transfer rate of vapor away from the drop increases, so the fraction of the heat transferred to the drop surface which is available to increase further the drop temperature decreases. As the drop velocity decreases, the convective heat-transfer coefficient between the air and the drop decreases. The combination of these factors gives the behavior shown in Fig. 10.36 where drop mass, temperature, velocity, vaporization rate, and heat-transfer rate from the air are shown schematically as a function of time following injection. Analysis of individual fuel drops 25 μm in diameter, injected into air at typical diesel conditions, indicates that evaporation times are usually less than 1 ms.^{29,30}

Figure 10.36 Schematic of variation of mass, diameter, temperature, evaporation rate, heat-transfer rate from air, and heat-transfer rate to liquid drop core as function of time during evaporation process of individual drop in diesel environment at the time of injection.²⁹



Such an analysis is relevant to drops that are widely separated (e.g., at the edge of the spray). In the spray core, where drop number densities are high, the evaporation process has a significant effect on the temperature and fuel-vapor concentration in the air within the spray. As fuel vaporizes, the local air temperature will decrease and the local fuel vapor pressure will increase. Eventually thermodynamic equilibrium would pertain: this process is usually called adiabatic saturation.²⁹ Calculated thermodynamic equilibrium temperatures for diesel spray conditions are plotted in Fig. 10.37 as a function of the fuel/air mass ratio for *n*-dodecane and *n*-heptane. The initial liquid fuel temperature was 300 K. The ratio of fuel vapor to air mass at these equilibrium conditions is also shown. To the left of the peaks in the m_{fv}/m_a curves, only fuel vapor is present. To the right of these peaks, liquid fuel is also present because the vapor phase is saturated.³⁰ Liquid fuel vaporization causes substantial reductions in gas temperature. While this equilibrium situation may not be reached within the spray, these results are useful for understanding the temperature distribution within an evaporating spray.

Figure 10.37 Adiabatic-saturation conditions for equilibrium mixtures formed by injecting *n*-dodecane and *n*-heptane, initially liquid at 300 K, into air at initial temperature T_a between 600 K and 1200 K and initial density 6.5 kg/m^3 . Equilibrium mixture temperature (T_e) and ratio of fuel vapor mass (m_{fv}) to air mass (m_a) shown as function of ratio of total fuel mass m_f to m_a . Fuel vapor only present to left of peaks in m_{fv}/m_a curves: liquid fuel also present to right of peaks.³¹



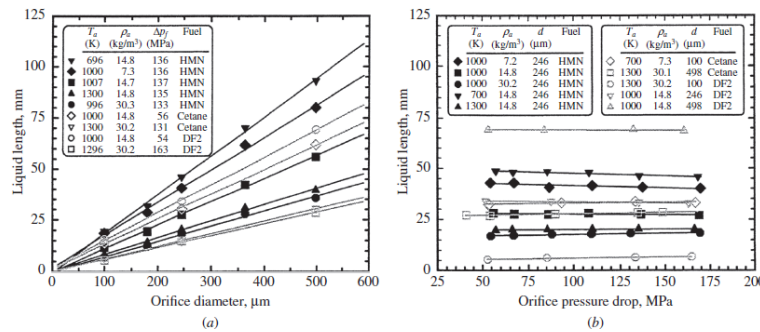
To quantify accurately the fuel vaporization rate within a diesel fuel spray requires the solution of the coupled conservation equations for the liquid droplets and the air within the combustion chamber. Various phenomenological models and computational fluid dynamic models have been developed for this purpose (see Secs. 14.4.3 and 14.5.5). In the more sophisticated of these, the spray is assumed to be composed of discrete computational particles each of which represents a group of droplets of similar size, temperature, etc. The distribution functions in droplet size, velocity, temperature, etc. produced by the fuel injector are statistically sampled and the resulting discrete particles are followed along lagrangian trajectories as they interact and exchange mass, momentum, and energy with the surrounding gas. Drops interact directly with each other via collisions and indirectly via evaporation by modifying the ambient vapor concentration and gas temperature. Studies with such models indicate that, under normal diesel engine conditions, 70 to 95% of the injected fuel is in the vapor phase at the start of combustion. Evaporation is more than 90% complete after about 1 ms. However, only 10 to 35% of the vaporized fuel has mixed to within flammability limits in a typical medium-speed DI diesel engine. Thus combustion is largely mixing-limited, rather than evaporation-limited.³² Of course, under cold-starting conditions, evaporation becomes a major constraint.

The extent of the liquid-containing core of the spray is an important factor. Initially, the fuel drops move rapidly into the engine's combustion chamber thus promoting entrainment of air into the spray, and fuel-air mixing. Figure 10.21 shows that this liquid core initially penetrates increasingly into the combustion chamber within the overall spray. It then ceases to grow in length, while the spray continues to penetrate due to its momentum, because the liquid fuel drops that constitute this liquid core have now evaporated. This maximum core penetration distance, relative to the distance along the spray axis from nozzle orifice to combustion chamber wall, is an important fuel injection system parameter. Overpenetration of liquid fuel results in wall wetting leading to slower fuel vaporization and to higher hydrocarbon and particulate emissions. The maximum liquid fuel penetration distance—the so-called liquid length within the fuel spray—thus needs to be characterized. It depends on the basic spray-initiating parameters (drop size and velocity, and ambient gas density) as well as the fuel's vaporization parameters (volatility characteristics, fuel and ambient gas temperatures). Optical techniques are usually used to determine liquid (core) lengths for sprays injected into closed vessels under diesel-like conditions.

Figure 10.38 shows the effects of orifice diameter and orifice pressure drop for heptamethylnonane (HMN), cetane, and No. 2 diesel fuel, for a range of conditions relevant to diesel operation. The liquid length rises linearly with orifice size, but is

essentially independent of orifice pressure drop. Both ambient air density which impacts the deceleration of the droplets and the spray, and ambient air temperature which impacts the droplet evaporation process, have significant effects.

Figure 10.38 Measured liquid lengths—the maximum liquid fuel penetration distances—as a function of (a) nozzle orifice diameter and (b) orifice pressure drop, for different ambient gas temperatures and densities. Fuels are HMN, cetane, and No. 2 diesel fuel. Lines are least square fits to the data.³²



A mixing-limited fuel vaporization process has been shown to explain these results.²² The fuel injection rate is proportional to the injection velocity, which varies with the square root of the orifice pressure and the orifice diameter squared (through Eq. [10.15]). The total air entrained into the spray up to any axial location is proportional to that axial distance, the orifice diameter, and the injection velocity.²³ These dependences indicate that to entrain enough hot air to fully vaporize the fuel, the axial extent of the liquid core will increase in proportion to the increase in orifice diameter, as the data indicates. With respect to the fuel injection velocity, these dependencies mean that the amount of air entrained up to any given axial liquid-core tip location, and the fuel injection rate, scale with injection velocity. Thus the *liquid length* is independent of the orifice pressure drop, as Fig. 10.38b indicates.²²

A scaling law for liquid-phase fuel penetration can be derived from an energy balance for a one-dimensional model of an evaporating spray, in addition to spray mass and momentum balances.³² At the maximum liquid core penetration (the liquid length), just enough energy has been transferred from the hot entrained air to the liquid drops to fully vaporize the fuel. It is also assumed that at the liquid length, saturated fuel vapor conditions pertain, and that air entrainment rates for the vaporizing spray differ little from those of a non-vaporizing spray. That is, the spray penetration length scaling law, Eq. (10.17) and its inverse Eq. (10.21) defines the overall spray behavior.³² That liquid-length scaling law, in dimensionless form is,

$$\tilde{L} = 0.47 \sqrt{\left(\frac{2}{B(T_a, p_a, T_f)} + 1 \right)^2 - 1}$$

(10.29)

where

$\tilde{L} = L_t/x^+$ with x^+ defined by Eq. (10.19b), and

$$B(T_a, p_a, T_f) = \frac{Z_a(T_s, p_a - p_s)p_s M_f}{Z_f(T_s, p_s)(p_a - p_s)M_a} = \frac{h_a(T_a, p_a) - h_a(T_s, p_a - p_s)}{h_f(T_s) - h_f(T_f, p_a)}$$

(10.30)

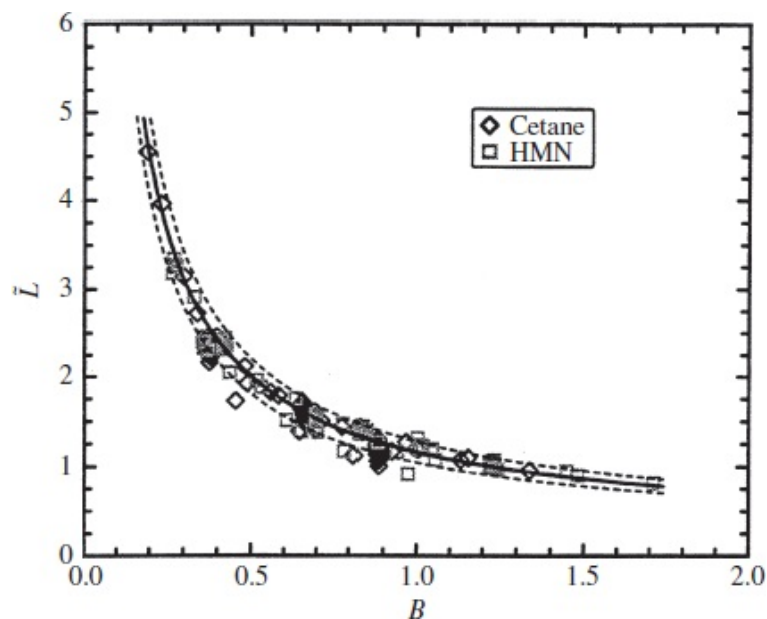
The term B results from the energy balance on the fuel spray, and an equation of state. The expressions in Eq. (10.30) are for single component fuels. Equation (10.30) is an implicit equation for the saturated fuel/ambient-gas mixture temperature T_s , at the liquid length, in terms of the fuel and ambient gas properties, and the initial fuel and ambient gas conditions. The

temperature T_a is the ambient gas temperature. The pressures p_a and p_s are the ambient gas pressure and the saturation (i.e., partial) pressure of the fuel at the liquid length. Thus the partial pressure of the entrained ambient gas in the fuel jet is $p_a - p_s$. The (saturated) fuel vapor pressure p_s is related to the saturation temperature T_s through the saturation pressure-temperature relationship of the fuel. The variables h_f and h_a , and M_f and M_a , are the specific enthalpies and molecular weights of the fuel and ambient gas, while Z_f and Z_a are the compressibilities of fuel vapor and the ambient gas at saturated conditions. (Note that at such conditions, gases deviate from the ideal gas behavior—see App. B—and the ideal gas law is then written as $p = \rho \tilde{Z} RT / M$.) The fuel at the nozzle orifice exit is assumed to be at the injector tip temperature T_f and ambient gas pressure p_a . The enthalpy different in the numerator of the right-hand side of Eq. (10.30) is the specific enthalpy transferred from the entrained gas to the fuel: the enthalpy difference in the denominator is the specific enthalpy required to heat up and vaporize the liquid fuel.

Figure 10.39 shows a comparison of the liquid length scaling correlation, Eq. (10.30), with data, as a function of B . The conditions are representative of diesel in-cylinder ambient conditions at TC, given in Ref. 21, using HMN and cetane fuel. B varies from about 0.2 to 1.7 as the dimensionless liquid length (L_l/x^+) decreases from about 5 to 1. The correlation fits the data well.

Figure 10.39 Dimensionless liquid length [

$\tilde{L} = L_l/x^+$ see Eq. (10.29)] as a function of B [see Eq. (10.30)], an expression that results from an energy balance of the fuel spray and its entrained air, and equations of state for fuel vapor and air.³²

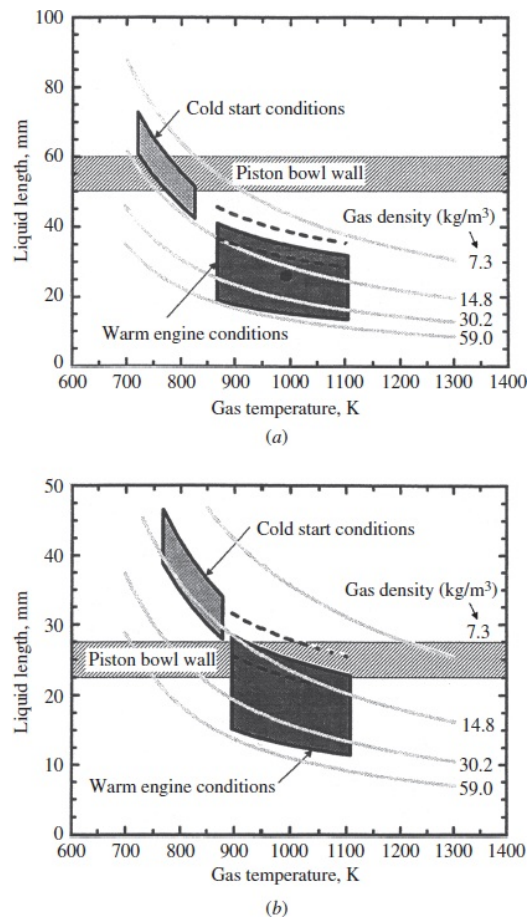


The relevance of this liquid fuel core penetration length for (larger) heavy-duty DI diesels and (smaller) high-speed light-duty DI diesels has been examined and provides useful insights into the behavior of evaporating sprays under both warmed-up and cold-start engine conditions. Table 10.3 gives the assumed geometry and operating details of the two diesel engines considered: a 12-liter six-cylinder large truck engine and a 2-liter four-cylinder high-speed passenger car engine. Their bowl-in-piston combustion chambers correspond to those shown in Figs. 10.1a and c. Figure 10.40 shows liquid lengths for these two engines, predicted with the correlation Eq. (10.29) as a function of TC in-cylinder ambient gas temperature (typically 1000 K). The upper and lower edges of the darker shaded regions correspond to boost pressures of 0.7 and 3 bar, respectively. The lighter shaded regions correspond to cold start conditions with ambient gas temperatures at TC of about 800 K. With the smaller bore engine, liquid impingement in the bowl wall is more likely to occur. Note, however, that the smaller high-speed engine will have high swirl, which would reduce the actual liquid fuel penetration somewhat below that given by the no-swirl correlation.

Table 10.3 Parameters used in liquid-length scaling law predictions for heavy- and light-duty DI diesel engines³²

	Heavy-duty	Light-duty	Units
Bore	135–140	80–85	mm
Bowl diameter/bore	0.75	0.5	—
Compression ratio	15	19	—
Orifice diameter	194	160	μm
Orifice C_a	0.84	0.84	—
Warm operation:			
Intake temperature	60–160	45–120	°C
Intake pressure	0.7–3.0	0.7–2.3	bar
Fuel temperature	100	100	°C
Cold start:			
Intake temperature	0–40	0–40	°C
Intake pressure	0.7–1	0.7–1	bar
Fuel temperature	0–40	0–40	°C

Figure 10.40 Liquid lengths predicted for (a) heavy-duty DI diesel and (b) (smaller) high-speed DI diesel engines defined in Table 10.3. The dark gray regions are for warmed-up engine operating conditions; the light gray regions denote cold start conditions. Typical piston bowl wall distances from the fuel injector nozzle orifice are indicated. The light gray lines are for constant in-cylinder gas densities, at TC of compression.³²



[1] Typical numbers: pump and distribution system, up to 1500 bar; common rail system, up to 1700 bar; unit injectors, up to 2500 bar; piezoelectric injectors, up to 1800 bar.

[2] The back lighting identifies regions where sufficient liquid fuel (as ligaments or drops) is present to attenuate the light. The shadowgraph technique responds to density gradients in the test section, so it identifies regions where fuel vapor exists.

10.5. IGNITION DELAY

10.5.1. Definition and Discussion

The ignition delay in a diesel engine is defined as the time (or crank angle) interval between the start of injection and the start of combustion. The SOI is usually taken as the time when the injector needle lifts off its seat. The SOC is more difficult to determine precisely. It is often identified from the change in slope of the heat-release rate (from negative to positive) calculated from the cylinder pressure data using the techniques described in [Sec. 10.3.3](#). This occurs shortly after detectable heat (chemical energy) release begins, which offsets the slight decline in cylinder pressure due to fuel evaporation within the developing spray. However, hydrocarbon reaction chemistry is apparent within the spray before this point, evidenced by the onset of luminosity (chemiluminescence). This luminosity effectively defines the onset of the breakup of the fuel molecules and the initiation of partial fuel oxidation, and precedes the point of pressure rise by some 100 μs , about 1 crank-angle degree.

Both physical and chemical processes must take place before a noticeable fraction of the chemical energy of the injected liquid fuel can be released. The physical processes are: the atomization of the liquid fuel into droplets close to the injection nozzle exit; the vaporization of some of these fuel droplets; the mixing of fuel vapor with air that has been entrained into the spray. The chemical processes are the pre-major-heat-releasing combustion reactions of the fuel, air, residual gas mixture, that lead to autoignition. These processes are affected by engine design and operating variables, and fuel characteristics, as this section ([10.5](#)) examines.

Good atomization requires high fuel-injection pressure, small injector hole diameters, appropriate fuel viscosity, and high cylinder air pressure at the time of injection (see [Sec. 10.4.3](#)). The rate of vaporization of the fuel droplets depends on the size of the droplets, their distribution, and their velocity, the pressure and temperature inside the combustion chamber, and the volatility of the fuel. The rate of fuel-air mixing is controlled largely by injector and combustion chamber design. Some cylinder-head and piston-crown shapes are designed to amplify swirl and create turbulence in the air charge during compression. IDI engine designs use a prechamber or swirl chamber to create the vigorous air motion necessary for rapid fuel-air mixing. In DI engines, injector design features such as the number and spatial arrangement of the injector holes determine the injected fuel distribution. The details of each nozzle hole affect the spray cone angle. The penetration of the spray depends on the size of the fuel droplets, the injection pressure, the air density, and the air-flow characteristics. The arrangement of the sprays, the spray cone angle, the extent of spray penetration, and the airflow all affect the rate of air entrainment into the spray. These physical aspects of fuel-injection and fuel-spray behavior are reviewed in [Sec. 10.4](#).

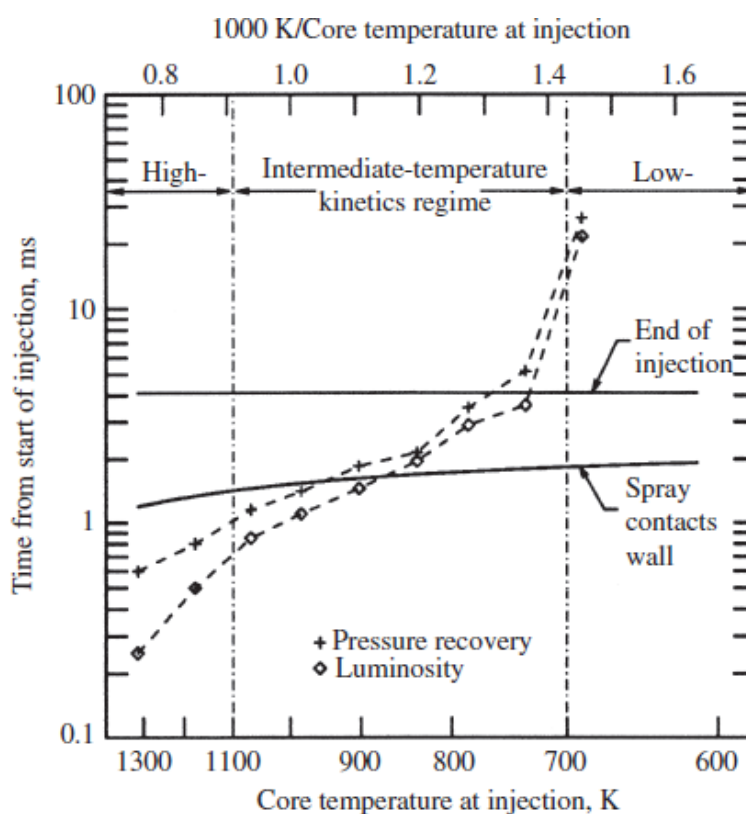
The chemical component of the ignition delay is controlled by the precombustion reactions of the fuel. A fundamental discussion of autoignition or spontaneous hydrocarbon oxidation in premixed fuel-air mixtures is given in [Sec. 9.6.2](#). Since the diesel engine combustion process is heterogeneous, its spontaneous ignition process is especially complex. Ignition occurs in regions where fuel vapor has mixed with some air, through the cracking of large hydrocarbon molecules to smaller molecules (partial oxidation reactions). These chemical processes depend on the composition of the fuel and the cylinder charge temperature and pressure, as well as the physical processes described above which govern the distribution of fuel within each spray.

Since the spontaneous ignition characteristics of the fuel affect the ignition delay, this property of a fuel is very important in determining diesel engine operating characteristics such as fuel conversion efficiency, smoothness of operation, misfire, smoke emissions, noise, and ease of starting. The ignition quality of a fuel is defined by its *cetane number*. Cetane number is determined by comparing the ignition delay of the fuel with that of primary reference fuel mixtures in a standardized engine test (see [Sec. 10.5.2](#)). For low cetane fuels with too long an ignition delay, most of the fuel will be injected before ignition occurs, which results in very rapid burning rates once combustion starts with high rates of pressure rise and high peak pressures. Under extreme conditions, when the autoignition of much of the injected fuel occurs after a long delay (e.g., under

cold-starting conditions), this produces an audible "knocking sound," often referred to as diesel knock. For higher cetane number fuels, with shorter ignition delays, ignition occurs before most of the fuel has been injected. The rates of heat release and pressure rise are then controlled primarily by the rate of injection and fuel-air mixing, and smoother engine operation results.

The timing relationships between the various measures of autoignition onset and the fuel injection process, as a function of the in-cylinder air temperature at SOI (close to TC), are shown in Fig. 10.41. (Note that 1 ms at 1800 rev/min corresponds to about 10 crank-angle degrees.) Typical end-of-compression temperatures in a warmed-up engine are close to 1000 K. The luminosity onset precedes the pressure recovery point at 1000 K by about 0.2 ms (about 2 degrees CA). Under cold-start conditions (750 to 800 K air temperature), the pressure rise lag is significantly longer. The mixture temperature values where the autoignition chemistry transitions from low- to intermediate- to high-temperature reaction schemes are indicated.³³ Section 10.5.3 will further discuss these details.

Figure 10.41 Timing of several fundamental features of the diesel spray autoignition process as a function of in-cylinder air temperature at time of injection. Symbols show measured times (in a constant volume combustion vessel) of luminosity onset and pressure recovery (rise) occurrence. Solid lines show approximate times for the spray to impinge on the chamber walls, and of end of injection. Ignition did not occur for gas temperatures below about 750 K.³³



Pilot injection (the injection of 5 to 10% of the fuel some 10 to 20 degrees before the start of the main injection) is a strategy enabled by the development of multi-pulse fuel injection technology (see Sec. 10.7.1) that is used to shorten the ignition delay. The spontaneous ignition and combustion of the pilot fuel creates a hotter local gas environment around the fuel injector into which the main injection sprays penetrate. The higher temperature of the ambient gas (mainly air), which is then entrained into the main-injection sprays, raises the within-the-spray mixture temperature thus speeding up the premixed-mixture ignition chemistry.

10.5.2. Fuel Ignition Quality

The ignition quality of a diesel fuel is defined by its *cetane number*. The method used to determine the ignition quality in terms of cetane number is analogous to that used for determining the antiknock quality of gasoline in terms of octane number. The cetane number scale is defined by blends of two pure hydrocarbon reference fuels. Cetane (*n*-hexadecane, $C_{16}H_{34}$), a hydrocarbon with high ignition quality, represents the top of the scale with a cetane number of 100. An iso-cetane, 2-2-4-4-6-8-8 heptamethyl-nonane (HMN), also $C_{16}H_{34}$, which has a very low ignition quality, represents the bottom of the scale with a cetane number of 15.^[1] Thus, cetane number (CN) is given by

$$CN = \text{percent } n\text{-cetane} + 0.15 \times \text{percent HMN}$$

(10.31)

The engine used in cetane number determination is a standardized single-cylinder, variable compression ratio, engine with special loading and accessory equipment and instrumentation. The engine, the operating conditions, and the test procedure are specified by ASTM Method D 613.³⁴ The operating requirements include: engine speed, 900 rev/min; coolant temperature, 100°C; intake air temperature, 65.6°C (150°F); injection timing, 13° BTC; injection pressure, 10.3 MPa (1500 lb/in.²). With the engine operating under these conditions, on the fuel whose cetane number is to be determined, the compression ratio is varied until combustion starts at TC: that is, an ignition delay period of 13° (2.4 ms at 900 rev/min) is produced. The above procedure is then repeated using reference fuel blends. Each time a reference fuel is tried, the compression ratio is adjusted to give the same 13° ignition delay. When the compression ratio required by the actual fuel is bracketed by the values required by two reference fuel blends differing by less than five cetane numbers, the cetane number of the fuel is determined by interpolation between the compression ratios required by the two reference blends.

Because of the expense of the cetane number test, many correlations which predict ignition quality based on the physical properties of diesel fuels have been developed. A calculated *cetane index* (CCI) is often used to estimate ignition quality of diesel fuels (ASTM D976³⁵). It is based on API gravity and the mid-boiling point (temperature 50% evaporated). It is applicable to straight-run fuels, catalytically cracked stocks, and blends of the two. Its use is suitable for most diesel fuels and gives numbers that correspond quite closely to cetane number. A *diesel index* is also used. It is based on the fact that ignition quality is linked to hydrocarbon composition: *n*-paraffins have high ignition quality, and aromatic and napthenic compounds have low ignition quality. The aniline point (ASTM D611³⁶), the lowest temperature at which equal volumes of the fuel and aniline become just miscible, is used together with the API gravity to give the diesel index:^[2]

$$\text{Diesel index} = \text{aniline point } (^{\circ}\text{F}) \times \text{API gravity}^j / 100$$

(10.32)

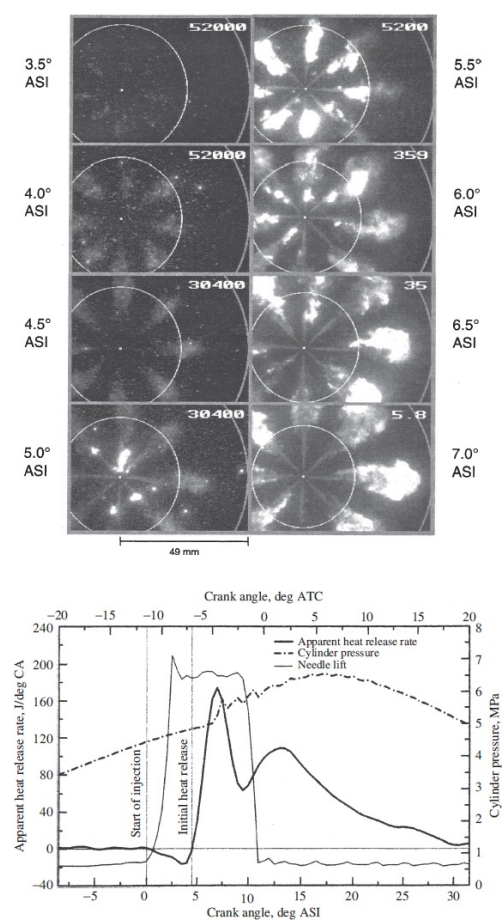
The diesel index depends on the fact that aromatic hydrocarbons mix completely with aniline at comparatively low temperatures, whereas paraffins require considerably higher temperatures before they are completely miscible. Similarly, a high API gravity denotes low specific gravity and high paraffinicity and, again, good ignition quality. The diesel index usually gives values slightly above the cetane number. It provides a reasonable indication of ignition quality in many (but not all) cases.

10.5.3. Autoignition and Premixed Burn

Basic studies in constant-volume combustion vessels, in steady-flow reactors, and in rapid-compression machines have been used to examine the various stages by which autoignition of fuel-air mixtures under conditions relevant to diesel engines occurs. In some of these studies the fuel and air were premixed; in many, fuel injection was used. Studies with fuel injection into constant-temperature and pressure environments, similar to conditions in a diesel engine at TC of compression, have identified the key steps in the process. The start of the diesel combustion process—spontaneous autoignition, followed by a

short and rapid "premixed-burning" process—is illustrated by the natural flame emissions images in Fig. 10.42 with the diagram showing the fuel injection profile (injector needle lift), apparent heat release rate and the cylinder pressure versus crank angle.³⁷ Faint (low intensity) chemiluminescence emission is apparent at 3.5° after start of injection (ASI), from the early injected fuel vapor/entrained air mixture in the leading region of these eight fuel sprays. These regions are beyond the "liquid length," the maximum penetration of the liquid fuel containing spray core (see Sec. 10.4.6). This chemiluminescence (emitted because fuel molecule breakdown and partial oxidation is occurring) becomes much stronger (4.0° ASI and 4.5° ASI) and clearly visible throughout the downstream region of all the sprays. Note that the apparent heat release rate is now rising sharply: the premixed burn is getting underway.

Figure 10.42 Sequence of natural flame emissions images taken through a window in the piston crown of an optical-access DI diesel engine with eight fuel sprays and 140-mm bore, operating at naturally-aspirated engine conditions and 1200 rev/min. Timing of each image, after start of injection (ASI), and relative light intensity gain (upper right of each image) are indicated. Bottom of figure: apparent heat-release rate, cylinder pressure, injector needle lift profiles versus crank angle.³⁷



The early stages of the (weaker) chemiluminescence are primarily the result of contributions from formaldehyde (CH_2O) and CH emissions [see reactions in Eq. (9.63), an outline of the hydrocarbon autoignition chemistry], corresponding to the "cool-flame" and "preignition glow" phenomena in rich mixtures. And the leading regions of these diesel sprays contain relatively well-mixed rich fuel-vapor/entrained air mixture in the fuel-air equivalence ratio range of 2 to 4.^{37, 38}

The hydrocarbons autoignition initiating mechanism is similar to the chemistry that initiates spark-ignition engine knock (see Sec. 9.6.2). At sufficiently high temperatures, a hydrocarbon air mixture undergoes a sequence or chain of reactions that produces an ever increasing pool of radicals (highly reactive species) through radical chain propagating and branching reactions. As the bonds in the basic hydrocarbon molecule are successively broken, chemical energy is released at a first

slowly then rapidly increasing rate. Radical removing (chain-terminating) reactions also occur, offsetting the radical build-up to varying extent as chemistry and energy release proceed. See the reaction set in Eq. (9.63) which further delineates this process.

The much brighter emission images in Fig. 10.42 (from about 5° to 6° ASI) result from soot particle radiation. Fuel molecule breakdown in the rich fuel-vapor/air mixture within the spray, in the early chemiluminescence phase, leads to the formation of PAHs which are precursors in the soot formation process (see Secs. 11.5.3 and 11.5.4 on diesel particulate formation). These PAHs form rapidly, grow into nuclei, and grow and agglomerate into soot (primarily carbon) particles. Significant chemical energy (heat) release is now occurring (6° ASI, Fig. 10.42).

The final images, 6.5° and 7° ASI, in Fig. 10.42 show the development of diffusion flames around the individual sprays. These diffusion flames, where oxidation of the fuel is completed, are located at the spray boundaries with "fuel" (very rich mixture consisting of fuel vapor, partially reacted fuel molecules, PAHs, soot particles, CO, and H₂) on the inside of the flame and air on the outside, constitute the second mixing-controlled phase of diesel combustion where the majority of the heat release occurs. By extrapolating the rounded portion of the heat release rate curve (which is this second mixing-controlled phase) at its left-hand end down to the horizontal axis, one sees that these diffusion flames should become apparent at 6° to 6.5° ASI. Note that these developing burning-spray images correspond to the illustrative schematic already shown in Fig. 10.11. Autoignition usually commences a few crank-angle degrees after start of fuel injection, close to TC, in the rich fuel-vapor/entrained-air mixture in the leading region of each developing spray, begins with the breakdown of fuel molecules through high-enough velocity molecular collisions to initiate this bond-breaking process. It then continues and transitions into the (rapid) premixed burning phase as the fuel hydrocarbon fragments build-up into PAHs, which continue to grow in molecular size into soot particles releasing significant chemical energy in the process and increasing the temperature within each spray (see Fig. 10.13).³⁷

A final point to note relates to the magnitude of the peak premixed-burn heat-release rate and its relative timing (see Fig. 10.42 for an illustrative example). As in-cylinder gas conditions such as ρ_g and T_g change, and as a consequence the ignition delay shortens or lengthens, so the premixed-burn peak heat-release rate decreases or increases. With shorter delays, the amount of rich fuel-vapor/entrained-air mixture within the sprays ready for spontaneous autoignition to occur decreases so its chemical energy available for rapid release decreases. The mixing-controlled burning heat-release-rate profile then absorbs the additional fuel to be burned in this second mode without major change. However, as the ignition delay lengthens significantly (and the premixed-burn fuel amount increases substantially) then the mixing-controlled burning phase is more noticeably affected.³⁷

Ignition delay data from actual and simulated engine experiments have usually been correlated by equations of the form:

$$\tau_{id} = Ap^{-n} \exp\left(\frac{E_A}{\bar{R}T}\right)$$

(10.33)

In general, τ_{id} is a function of mixture temperature, pressure, equivalence ratio, and fuel properties (though no accepted form for the variation with equivalence ratio is yet established). In the studies summarized and referenced in the first edition of this text,¹ the fuel was injected into a uniform air environment where the pressure and temperature only changed due to the cooling effect of the fuel-vaporization and fuel-heating processes. In an engine, pressure and temperature change during the delay period due to the compression resulting from piston motion. To account for the effect of changing conditions on the delay the following empirical integral relation has been used:

$$\int_{t_{si}}^{t_{si} + \tau_{id}} \left(\frac{1}{\tau}\right) dt = 1$$

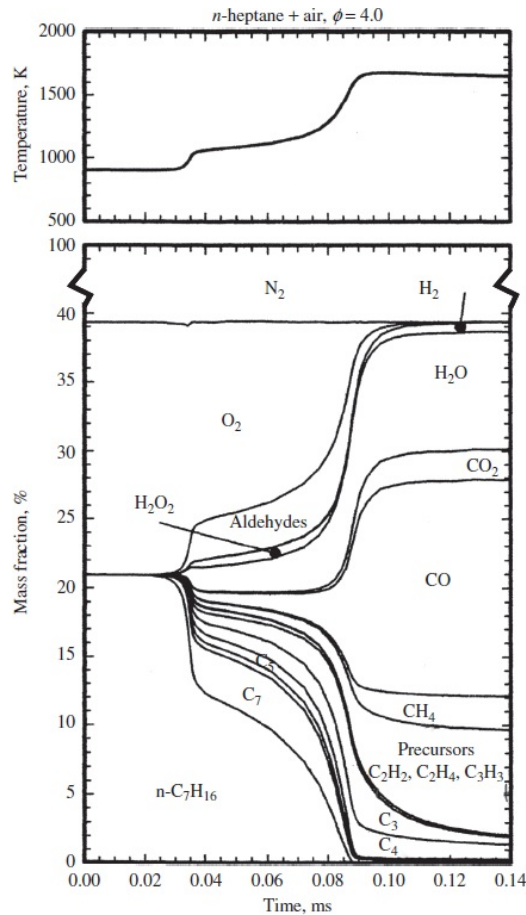
(10.34)

where t_{si} is the time of start of injection, τ_{id} is the ignition delay period, and τ is the ignition delay at the conditions pertaining at time t . Whether the variation in conditions is significant depends on the amount of injection advance before TC that is used and the length of the delay.

The first edition of this text¹ compared several published Arrhenius equations for ignition delay of the form of Eq. (10.33). See also Ref. 39. These earlier correlations no longer seem useful for several reasons. In-cylinder pressures and temperatures at TC continue to change as turbocharging dominates and boost levels increase. Such conditions are difficult to realize in out-of-engine experiments. Increasing injection pressures and high swirl in smaller high-speed DI engines are reducing the contribution to the delay that these enabling physical processes contribute. Also, using a single Arrhenius expression to represent the overall chemistry of the autoignition process is too simplistic given the complex chemistry actually involved.

An illustration of that chemistry is shown in Fig. 10.43 where a simpler hydrocarbon molecule—heptane (C_7H_{16}), which has a cetane number of 60 (close to that of diesel fuel which is about 50), was used diesel fuel surrogate. The detailed chemical kinetic model (using 550 species and 2454 reactions) is briefly described in the reference.¹¹ The chemistry calculation is at a starting temperature of 900 K, pressure of 83 bar, and fuel-air equivalence ratio of 4, appropriate conditions for an autoignition, and premixed-burn simulation in a fuel spray in a turbocharged diesel engine. Autoignition commences at about 0.03 msec (about 3° ASI) and the low-temperature first stage (and modest) energy release occurs (temperature rise of some 150 K). Just after 0.07 msec (7° ASI), the second stage of the ignition process (high-temperature) chemistry begins which rapidly raises the temperature in the leading half of the spray plume to rise to 1650 K (see also Fig. 10.14), through a larger energy release (raising the temperature by some 400 K or more). After 0.1 to 0.15 msec (1 to 1.5° ASI), the rich mixture within the spray contains H_2O , CO_2 , CO , CH_4 , soot precursors (acetylene and polyacetylenes), and lower carbon atom hydrocarbons (C_3 and C_4) along with some 60% nitrogen. By energy fraction (mass fraction $\times$ heating value), the combustibles H_2 , CO , CH_4 , soot precursors, and lower carbon number hydrocarbons are 11, 21, 16, 40, and 12%, respectively, of the *unreleased* rich-mixture chemical energy. Some 25% of the fuel's full heating value has been released at this point. This is the "fuel" on the fuel side of the diffusion flame, during the mixing-controlled combustion phase.¹¹

Figure 10.43 Results of chemical kinetic analysis of *n*-heptane vapor reaction with air, at $\phi = 4$ and initial mixture temperature of 900 K and pressure 83 bar (typical TC diesel engine conditions). The evolving species concentrations and gas temperature in this rich mixture within each diesel spray are shown over time (0 to 0.14 ms).¹¹



Such chemical-kinetic autoignition modeling studies indicate a strong dependence of reaction time scales on ambient temperatures. For example, an increase from 750 K to 1000 K reduces reaction times from about 1 to 0.1 ms. Ignition delays of some few degrees crank angle, 0.2 to 0.3 ms, are needed for normal diesel engine operation.¹¹

10.5.4. Physical Factors Affecting Ignition Delay

The physical variables that control the development of the fuel spray, and the air-charge state (its pressure, temperature, and velocity), will influence the ignition delay. These quantities depend on the design of the fuel-injection system and combustion chamber, and the engine operating conditions. The injection system variables affecting the fuel-spray development are injection timing, quantity, velocity, rate, drop size, and spray form or type. The relevant charge conditions depend on the combustion system employed, the details of the combustion chamber design, inlet air pressure and temperature, compression ratio, injection timing, residual gas amount and state, coolant and oil temperature, and engine speed. Data on these interactions are available for different types of diesel engines. The trends observed with the different diesel combustion systems are generally similar, though details may differ. In this section the ignition delay trends during fully warmed-up engine operation are discussed. In modern diesel engines, ignition delays are short: a few crank-angle degrees. Thus, changes in delay in warmed-up engines are now less significant. However, the dependence of the ignition delay on engine design and operating variables during engine starting and warm-up is important, and may be different from fully warmed-up behavior due to lower air temperatures and pressures.

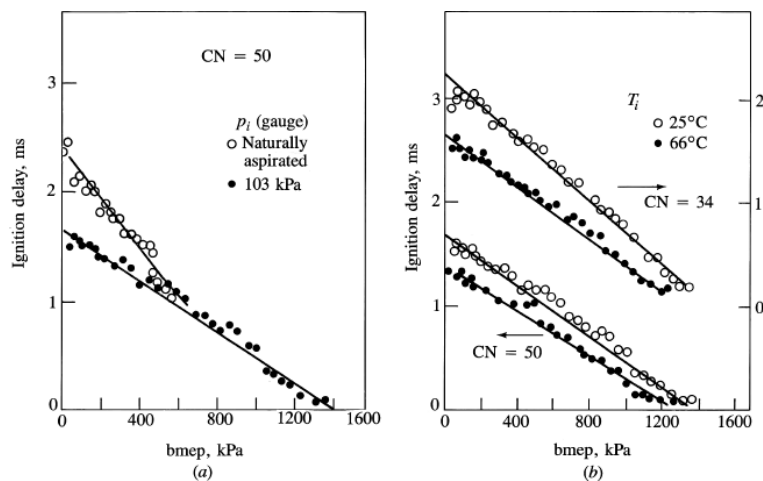
10.5.4.1. Injection Timing

At normal engine conditions (low to medium speed, fully warmed engine), the minimum delay occurs with the start of injection at about 10° BTC. The increase in the delay with earlier or later injection timing occurs because the air temperature and pressure change. If injection starts earlier, the initial air temperature and pressure are lower so the delay will increase. If injection starts later (closer to TC) the temperature and pressure are initially slightly higher but then decrease as the delay proceeds. The most favorable conditions for ignition lie in between.

10.5.4.2. Injection Quantity or Load

Figure 10.44 shows the effect of increasing injection quantity or engine load on ignition delay. The delay decreases approximately linearly with increasing load for this DI engine. As load is increased, the residual gas temperature and the wall temperature increase, as does heat transfer to the air charge during the first half of compression. This results in higher charge temperature (and, to a lesser extent, charge pressure) at injection, thus shortening the ignition delay. When adjustment is made for this increasing temperature, it is found that increasing the quantity of fuel injected has little significant effect on the delay period under normal operating conditions. Under engine starting conditions, however, the delay increases significantly due to the larger drop in mixture temperature: engine components are cold, there is no in-cylinder air heating, and significantly more air cooling during compression. This latter result should be expected since evaporation of the injected fuel further reduces mixture temperature.

Figure 10.44 Effect of inlet air pressure and temperature on ignition delay versus load in a small DI diesel at 1980 rev/min. (a) Engine naturally-aspirated, and turbocharged with 1 atm boost (2 bar absolute); inlet air temperature $T_i = 25^\circ\text{C}$; 50 cetane number fuel. (b) Engine naturally-aspirated; $T_i = 25$ and 66°C ; 34 and 50 cetane number fuel.⁴⁰



10.5.4.3. Drop Size, Injection Velocity, and Rate

These quantities are determined by injection pressure, injector nozzle hole size, nozzle type, and geometry, but they have little impact on the ignition delay. At normal operating conditions, increasing injection pressure produces only modest decreases in the delay. Doubling the nozzle hole diameter at constant injection pressure to increase the fuel flow rate (by a factor of about 4) and increase the drop size (by about 30%) had no significant effect on the delay. Studies of different nozzle hole geometries showed that the length/diameter ratio of the nozzle was not significant; nor did changes in nozzle type (multi-hole, pintle, pintaux) cause any substantial variation in delay at normal engine conditions.³⁸

10.5.4.4. Intake Air Temperature and Pressure

[Figure 10.44](#) also shows values of ignition delay for diesel fuels for different charge temperatures and pressures at the time of injection. The intake air temperature and pressure will affect the delay via their effect on charge conditions during the delay period. The figure shows the effects of inlet air pressure and temperature as a function of engine load. The fundamental ignition data available show a strong dependence of ignition delay on charge temperatures below about 1000 K at the time of injection. Above about 1000 K, the data suggest that the charge temperature is no longer as significant. Through this temperature range there is an effect of pressure at the time of injection on ignition delay: the higher the pressure the shorter the delay, with the effect decreasing as charge temperatures increase and delay decreases. Since air temperature and pressure during the delay period are such important variables, other engine variables that affect the relation between the intake air state and the charge state at the time of injection will influence the delay. Thus, an increase in the compression ratio will decrease the ignition delay, and injection timing will affect the delay (as discussed above), largely due to the changes in charge temperature and pressure at the time of injection.

10.5.4.5. Engine Speed

Increases in engine speed at constant load result in a slight decrease in ignition delay when measured in milliseconds; in terms of crank-angle degrees, the delay increases almost linearly. A change in engine speed changes the temperature/time and pressure/time relationships. Also, as speed increases, injection pressure increases. The peak compression air temperature increases with increasing speed due to smaller heat loss during the compression stroke.⁴⁰

10.5.4.6. Combustion Chamber Wall Effects

The impingement of the spray on the combustion chamber wall obviously affects the fuel evaporation and mixing processes. Impingement of the fuel jet on the wall occurs, to some extent, in almost all of the smaller, higher speed engines (see [Fig. 10.40](#)). Engine and combustion bomb experiments have been carried out to explore the effect of wall impingement on the ignition delay. There does not seem to be a consistent trend. When the evolving fuel spray impinges on the combustion bowl wall, there will be some cooling of the spray tip. However, with short delays (in warmed-up engines) autoignition usually occurs before any significant spray-wall interaction.

10.5.4.7. Swirl Rate

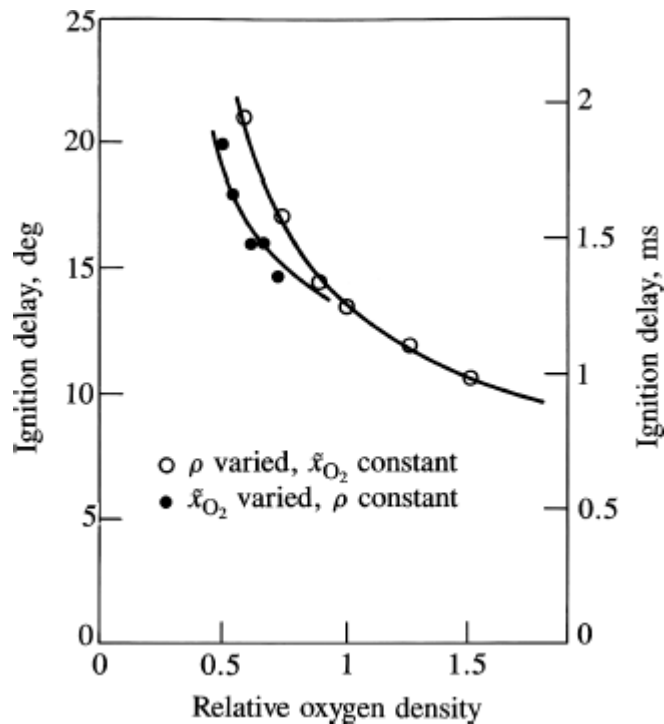
Changes in swirl rate change the fuel evaporation and fuel-air mixing processes (see [Figs. 10.23](#) and [10.24](#)). They also increase wall heat transfer during compression and, hence, reduce the charge temperature at injection. While engine studies of the effect of swirl rate on ignition delay have been made, only limited data are available. At normal operating engine conditions, the effects of swirl rate changes on the delay are small. Under engine starting conditions (low engine speeds and compression temperatures), the effect is more important, presumably due to the higher rates of evaporation and mixing obtained with swirl.

10.5.4.8. Oxygen Concentration

The oxygen concentration in the charge into which the fuel is injected would be expected to influence the delay. The oxygen concentration is changed, for example, when exhaust gas is recycled to the intake for the control of oxides of nitrogen emissions (see [Sec. 11.2.4](#)). Results of a study carried out in a single-cylinder DI engine operated at a constant air/fuel ratio (30:1), manifold temperature, injection timing, and speed (1800 rev/min), where the oxygen concentration was varied by

recirculating known amounts of cooled exhaust, are shown in Fig. 10.45.⁴¹ Oxygen density is normalized by the naturally-aspirated no-recirculation test value. As oxygen concentration is decreased, ignition delay increases.

Figure 10.45 Effect of oxygen density in gas on ignition delay in single-cylinder DI engine of 1.3-dm³ displacement with $r_c = 15$ at 1800 rev/min. Oxygen density changed by recycling exhaust gas at constant inlet density and by varying inlet pressure from 0.5 to 3 atm with air.⁴¹



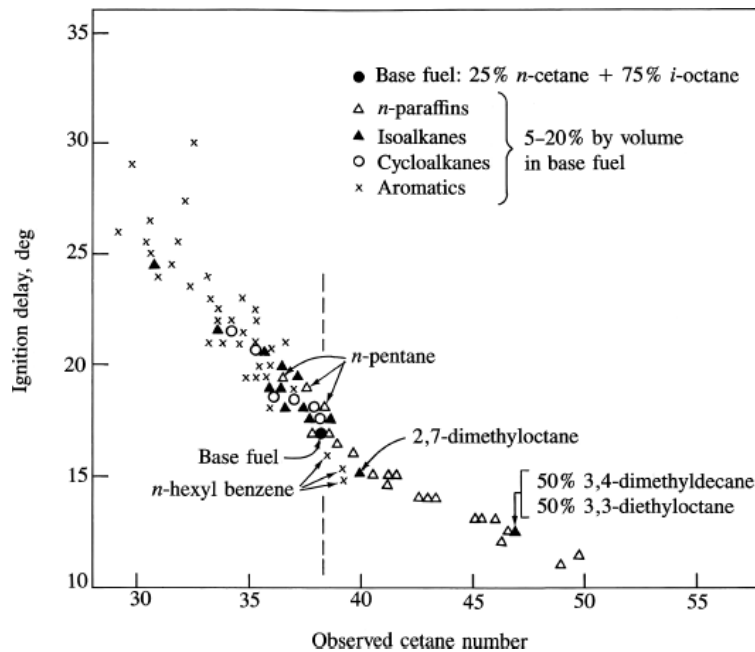
10.5.5. Effect of Fuel Properties

10.5.5.1. Diesel Fuel

Since both physical and chemical processes take place during the ignition delay, the effects of changes in the physical and chemical properties of fuels on the delay period have been studied. The chemical characteristics of the fuel are much the more important. The ignition quality of the fuel, defined by its cetane number, will obviously affect the delay. The dependence of cetane number on fuel molecular structure is as follows. Straight-chain paraffinic compounds (normal alkanes) have the highest ignition quality (autoignite more quickly), a trend which improves as the chain length increases. Aromatic compounds have poor ignition quality as do the alcohols (hence, the difficulties associated with using methanol and ethanol, possible alternative fuels, in compression-ignition engines). Figure 10.46 illustrates these effects. A base fuel was blended with pure paraffinic (normal, iso-, and cycloalkanes), aromatic, and olefinic hydrocarbons of various carbon numbers, by up to 20% by volume. The base fuel, a blend of 25% *n*-hexadecane and 75% isooctane, had a cetane number of 38.3. The figure shows that the resulting ignition delays correlate well as a function of cetane number at constant compression ratio and engine operating conditions. Addition of normal alkanes (excluding *n*-pentane and lower carbon number alkanes) improves the ignition quality. As the chain length of the added paraffin gets longer (higher carbon number) the cetane number improvement increases. Isoalkanes, depending on the degree of branching, degrade ignition quality (unless the branching is concentrated at one end of the molecule, when these types of isoalkanes improve ignition quality). Cycloalkanes and aromatics generally reduce the cetane number, unless they have a long *n*-alkane chain attached to the ring. The cetane number of a fuel (a measure of its ability to autoignite) generally varies inversely with its octane number (a measure of its ability to resist autoignition; see Fig.

9.65 for the effect of hydrocarbon structure on knock). The cetane number of commercial diesel fuel is normally in the range of 40 to 55.

Figure 10.46 Effect of type of hydrocarbon structure on ignition quality of fuels in DI diesel combustion process at constant compression ratio and engine operating conditions.⁴²



Cetane number is controlled by the source of crude oil, by the refining process, and by additives or ignition accelerators. Just as it is possible to reduce the tendency to knock or autoignite in spark-ignition engine fuels by adding antiknock agents, so there are additives that improve the ignition quality of compression-ignition engine fuels. Generally, substances that increase the tendency to knock enhance ignition, and vice versa. Ignition-accelerating additives include organic peroxides, nitrates, and nitrites. The most important of these commercially are the alkyl nitrates (isopropyl nitrate, primary amyl nitrates, primary hexyl nitrates, octyl nitrate). Typically, about 0.5% of these additives by volume in a distillate fuel gives about a 10 cetane number increase in a fuel's ignition quality, though their effectiveness may depend on the composition of the base fuel. The incremental effect of increasing amounts of ignition-accelerating additives on cetane number decreases.⁴⁰ Usually, the ignition delay obtained with cetane improved blends are found to be equivalent to those obtained with natural diesel fuels of the same cetane number. Two potential practical uses for ignition accelerators are in upgrading the ignition characteristics of poorer quality diesel fuel and (in much larger amounts) making possible the use of alcohols in compression-ignition engines.

The physical characteristics of diesel fuel do not significantly affect the ignition delay in fully or partially warmed-up engines. Tests with fuels of different front-end volatility (over the cetane number range 38 to 53), and with substantially different front-end ignition quality for the same average cetane number, showed no discernible differences. Fuel viscosity variations over a factor of 2.5 also showed no significant effect.⁴⁰ Thus, in a warmed-up engine, variations in fuel atomization, spray penetration, and vaporization rate over reasonable ranges do not appear to influence the duration of the delay period significantly (see also Sec. 10.4.6 on fuel spray evaporation).

10.5.5.2. Biodiesel

An alternative to petroleum-derived diesel fuel is *biodiesel*. The term refers to a vegetable oil or animal-fat based fuel consisting of long-chain alkyl (methyl or ethyl) esters.^[3] These are made by chemically reacting lipids (e.g., vegetable oils, animal fats) with methanol or ethanol. While similar in behavior to diesel fuel, there are differences. The vegetable oils used in

biodiesel production differ in the composition of their fatty acids, and differences in the pretreatment of these oils and the biodiesel production process also cause variations in properties.

Biomass sources used to produce vegetable oils for biodiesel production include rapeseed, soybeans, sunflower seed, palm and coconut oils. Highest yields come from palm oil (4500–5000 liters/hectare) with coconut about half that, and the others listed above in the 500 to 1000 l/ha range. Present expectations are that an upper bound on biodiesel productions is about 10% of petroleum diesel.

Biodiesel is currently mostly blended with conventional diesel to minimize the impact of its somewhat different characteristics. Five to ten percent blends (B5 and B10) are commonly used. B100 can be used in diesels but engine modifications would be required.

The characteristics of biodiesel that could affect combustion in diesel engines are its physical and chemical properties. Its molecular size (C_{16} – C_{18}) is in the diesel fuel range (C_{10} – C_{20}), but its boiling point is higher. Its viscosity is higher than that of diesel fuel. It has about 10% lower energy content per unit mass or volume (lower heating value varies with composition but is about 38 MJ/kg). Its cetane number (46 to 60) is higher than that of standard diesel fuel (40 to 55). It has better lubricating properties than diesel due to its higher viscosity. It is, however, a more aggressive solvent, and is less resistant to oxidation so its "shelf life" is significantly more limited. Its use blended in modest amounts with petroleum-based diesel fuel is thus a pragmatic approach.^{43,44}

10.5.6. Correlations for Ignition Delay in Engines

Many correlations have been proposed for predicting ignition delay as a function of engine and air charge variables. These usually have the form of Eq. (10.33) and have been based on data from more fundamental experiments in combustion vessels and flow reactors. An important factor in assessing the appropriateness of any correlation is how it is to be used to predict the magnitude of the delay. If an equation for predicting the complete delay process (including all the physical and chemical processes from injection to combustion) is required, then the data show that such a simple form for the equation is unlikely to be adequate for the full range of engine condition. Much more detailed chemical-kinetic models for the autoignition process of a premixed fuel-air mixture during the delay period are being used in conjunction with models for the physical processes of fuel evaporation and fuel-air mixing to predict the evolving spray processes.

An empirical formula, developed by Hardenberg and Hase⁴⁵ for predicting the duration of the ignition delay period in DI engines, has been shown to give reasonable agreement with experimental data over a wide range of engine conditions (see Fig. 10.47).⁴⁶ This formula gives the ignition delay (in crank-angle degrees) in terms of charge temperature T (kelvins) and pressure p (bar) during the delay (taken as TC conditions) as

$$\tau_{id}(CA) = (0.36 + 0.22\bar{S}_p) \exp \left[E_A \left(\frac{1}{\tilde{R}T} - \frac{1}{17,190} \right) + \left(\frac{21.2}{p - 12.4} \right)^{0.63} \right]$$

(10.35)

where

$\bar{S}_p$ is the mean piston speed (meters per second) and

$\tilde{R}$ is the universal gas constant (8.3143 J/mol · K). E_A (joules per mole) is the apparent activation energy, and is given by

$$E_A = \frac{618,840}{CN + 25}$$

(10.36)

where CN is the fuel cetane number. The apparent activation energy decreases with increasing cetane number. The delay in milliseconds is given by

$$\tau_{\text{d}}(\text{ms}) = \frac{\tau_{\text{d}}(\text{CA})}{0.006N}$$

where N , engine speed, is in revolutions per minute. Values for T and p can be estimated using a polytropic model for the compression process:

$$T_{\text{TC}} = T_i r_c^{n-1}, p_{\text{TC}} = p_i r_c^n$$

(10.37a,b)

where n is the polytropic exponent, r_c is the compression ratio, and the subscript i denotes intake manifold conditions. Values of the polytropic exponent are given in Fig. 10.48 for a direct-injection diesel under warm and cold engine operating conditions.^{41,46}

Figure 10.47 Comparison of engine ignition delays predicted with Eq. (10.35) with corresponding measured values.⁴⁶

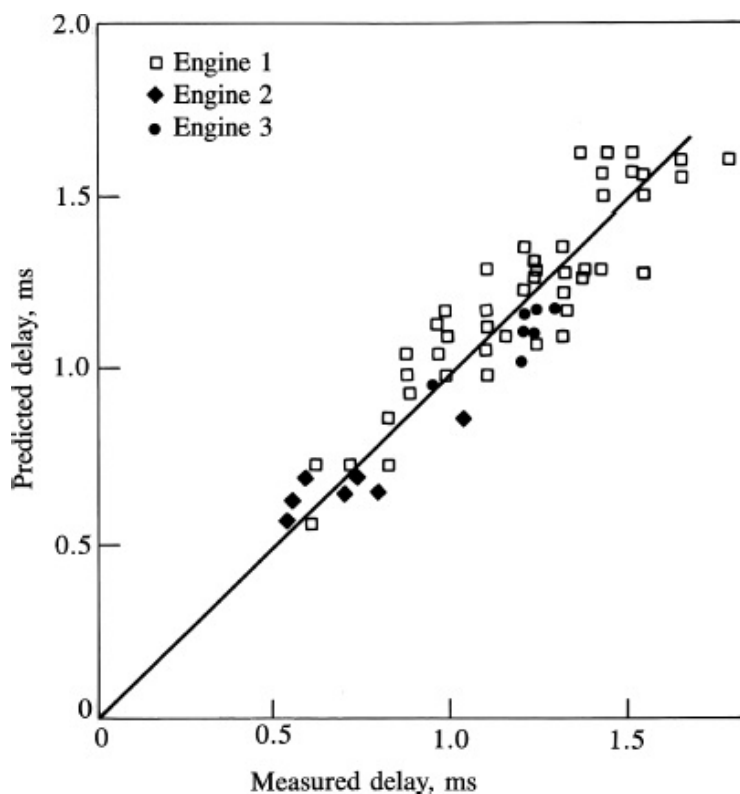
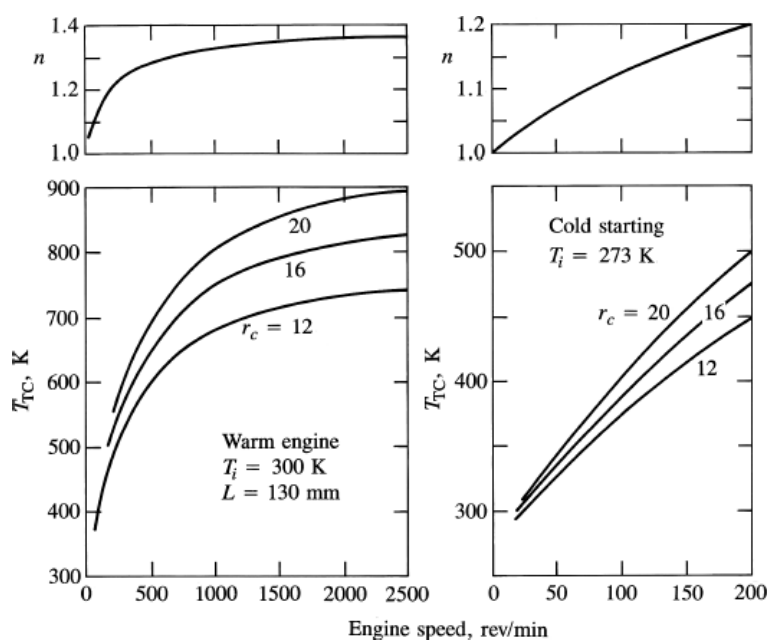


Figure 10.48 Exponent n for polytropic model of compression process in Eq. (10.37) and corresponding end-of-compression air temperature at TC. Warm and cold DI diesel engine with 130-mm stroke.⁴⁵



- [1] In the original procedure α -methyl-naphthalene ($C_{11}H_{10}$) with a cetane number of zero represented the bottom of the scale. Heptamethyl-nonane, a more stable compound, has replaced it.
- [2] API (American Petroleum Institute) gravity is based on specific gravity and is calculated from: API gravity (degrees) = $141.5 / (\text{specific gravity at } 60^\circ\text{F}) - 131.5$.
- [3] These are usually fatty acid methyl ester (FAME) and fatty acid ethyl ester (FAEE).

10.6. MIXING-CONTROLLED COMBUSTION

10.6.1. Background

Earlier sections of this chapter have described our current understanding of the individual processes, which together make up the total fuel injection-mixing-burning sequence—atomization, vaporization, spray development, air entrainment, ignition, and combustion. The photos in [Fig. 10.5](#) illustrate details of the diesel combustion process, and the schematics in [Fig. 10.11](#) identify the several steps involved. The interaction mixing-controlled phase of combustion is the area of focus of this section. The photographs of the compression-ignition combustion process in different types of diesel engines in [Figs. 10.4](#) and [10.5](#) show how the diesel diffusion flame forms, following ignition and start of the premixed-burned phase, spreads rapidly, and envelops the spray. Depending on the spray configuration, the visible flame may then fill almost the entire combustion chamber. The diffusion flame and spray geometries are closely related. Mixing processes are also critical during the ignition delay period: while the duration of the delay period is not influenced in a major way by the rates of the spray processes which together control "mixing," the amount of fuel mixed with air to within combustible limits during the delay (which affects the rate of pressure rise once ignition has occurred) is obviously directly related to mixing rates.

However, while it is well accepted that diesel combustion is normally controlled by the fuel-air mixing rate, the mechanisms by which the diffusion flame forms, spreads rapidly, and then stabilizes are not well understood. The difficulties are twofold. First, the spray geometry in real diesel combustion systems is extremely complex. Second, the phenomena, which must be described (especially the development of the unsteady turbulent diffusion flame) are inadequately understood.

Current capabilities for modeling these phenomena are reviewed in [Chap. 14](#). Thermodynamic-based models of the diesel combustion process with atomization, vaporization, and spray development described by empirical or phenomenological turbulent-jet-based submodels have been developed and used to predict burning rates. These are described in [Secs. 14.4.3](#) and [14.4.4](#) and show reasonable but not precise agreement with data. Fluid-mechanic-based models of airflow, fuel spray behavior, and combustion have been and are still being developed (see [Sec. 14.5](#)). While realistic air-flow predictions are now feasible, predictions of spray behavior are less well developed and combustion-rate predictions are still exploratory.

In the following sections, the evidence linking spray characteristics to flame structure and fuel burning rates is summarized.

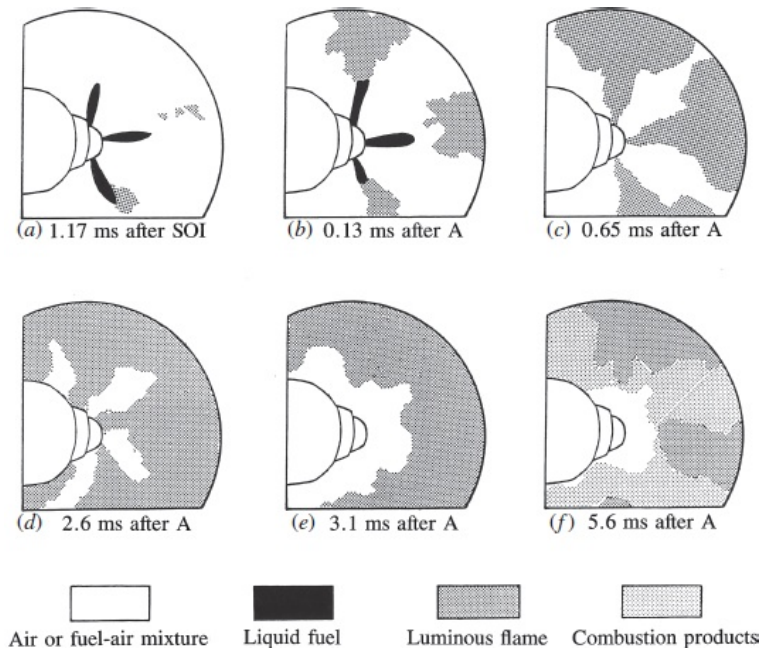
10.6.2. Spray and Flame Structure

The structure of each fuel spray is that of a narrow liquid-containing core (densely filled with drops of order $20\text{ }\mu\text{m}$ in diameter) surrounded by a much larger gaseous-jet region containing fuel vapor (see [Fig. 10.21](#)). The fuel concentration in the core is extremely high: local fuel/air equivalence ratios near the nozzle of order 10 have been measured during the injection period. Fuel concentrations within the spray decrease with increasing radial and axial position at any given time, and with time at a fixed location once injection has ended. The fuel distribution within the spray is controlled largely by turbulent-jet mixing processes. Fuel vapor concentration contours determined from interferometric studies of unsteady vaporizing diesel-like sprays, confirm this gaseous turbulent-jet-like structure of the spray, with its central liquid-containing core which evaporates relatively quickly once fuel injection ends.⁴⁷

[Figure 10.42](#) shows how the autoigniting of each spray and the subsequent development of the diffusion flame around the edge of the penetrating spray occur. [Figure 10.49](#) shows how this process continues under conditions typical of DI quiescent-chamber diesel engines. It shows tracings of the liquid spray core and the initiation and spread of the diffusion flame boundaries, taken from high-speed movies of the injection and combustion processes with central injection of five fuel jets into a disc-shaped chamber in a rapid-compression machine.⁴⁸ Such studies show that diffusion flame first forms around the edge of the spray (which may by then have interacted with the combustion chamber walls). This diffusion flame then moves

rapidly back along the spray boundary toward the injector nozzle. As time progresses, combustion products diffuse outward (and inward), from the diffusion flame enveloping the spray, as indicated in Fig. 10.49. This outward diffusion of (hot) burned gas is apparent in the combined natural light emission and shadowgraph image of a single diesel fuel spray in Fig. 10.5.

Figure 10.49 Tracings of the boundary of liquid fuel core in the spray and flame, from high-speed movies of diesel combustion taken in a rapid-compression machine, looking down on piston through transparent head. First occurrence of luminous flame at A, 1.17 ms after start of injection. End of injection at D.⁴⁸

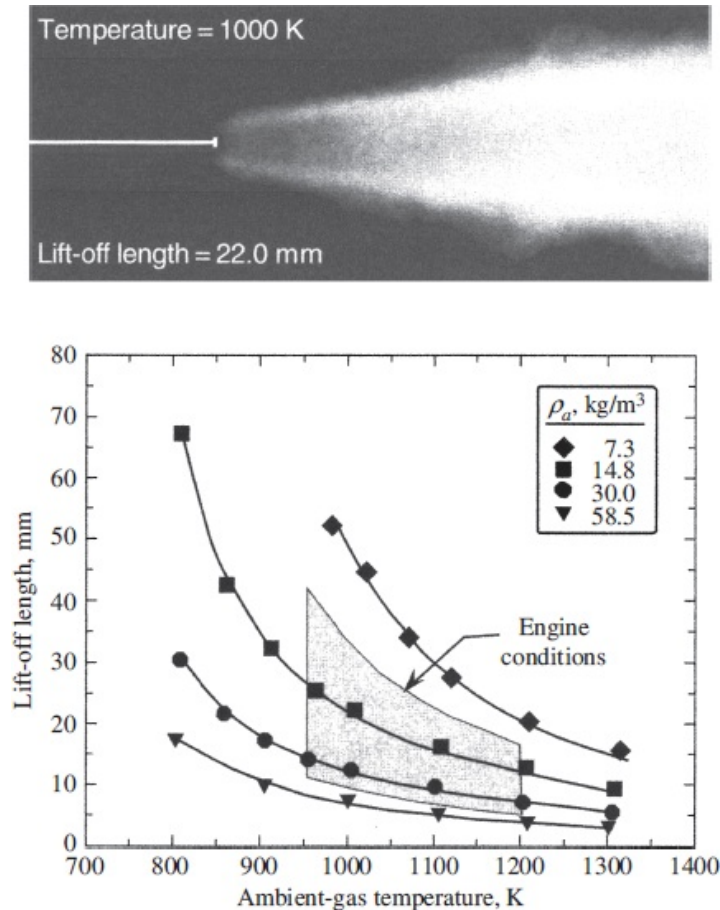


With modern high-injection-pressure (1000 bar plus) fuel systems, which produce high-velocity (of order 500 m/s) fuel drops, the flame does not move all the way back to the injector. It stabilizes during the quasi-steady phase of the spray's development some distance away from the nozzle exit. This phenomenon is called *flame lift-off*; it is illustrated in Fig. 10.13. As shown, its importance is that substantial entrainment of hot air into the spray occurs along this lift-off length.^[1] This significantly affects the fuel-air equivalence ratio within the spray, which impacts the mixture temperature and its autoignition chemistry. For moderate loads, in a quiescent DI diesel engine, some 20% of the air needed for complete combustion of the fuel is entrained into the spray over this lift-off distance.

Thus, the initiation of the diffusion flame and its spread around the developing spray, the essential component of the mixing-controlled combustion phase in diesel engines, occurs as follows. As the premixed-burn combustion phase progresses, the fuel-rich mixture within the leading portion of the spray becomes hotter through its chemical energy release, and due to the rising pressure within the cylinder (which also increases the air temperature). Within one or a few crank-angle degrees, one or more (localized) diffusion flames form with their energy release further increasing gas temperatures. These initial diffusion flames spread rapidly around the outer boundaries of the turbulent sprays where the fuel-air mixture is slightly rich of stoichiometric where reaction rates are fastest. The flame moves back toward the injector nozzle orifice to the location where the local jet velocity is high enough that its convective resistance to flame spreading stabilizes the diffusion flame, a short distance (the flame lift-off distance, typically some 20 mm) from the jet or spray's origin.

Figure 10.50 shows a time-averaged image of a burning diesel fuel spray in an ambient environment at 1000 K with gas density of 14.8 kg/m^3 , injection pressure of 138 MPa, and orifice diameter of $180 \text{ }\mu\text{m}$. The flame lift-off length, indicated by the white bar, was 22 mm. The experiments were done in a closed vessel.²² The lower part of the figure shows how this lift-off length varies with ambient-gas temperature and density at the conditions listed above (with 21% oxygen in the ambient gas).

Figure 10.50 Example of time-averaged image of a burning diesel fuel spray under diesel TC-of-compression conditions showing diffusion flame lift-off-distance (22 mm). See text for details of the conditions, in quiescent combustion vessel (upper figure). Lower figure, lift-off lengths as a function of ambient gas conditions (temperature and density). Oxygen concentration in ambient gas 21%, pressure drop across orifice 138 MPa, orifice diameter 180 μm . Grey region indicates typical conditions for a quiescent DI diesel.^{22,49}



The changes in lift-off length with respect to these and other parameters cause changes in the air entrained into each spray. The quantity of entrained air prior to lift-off can be estimated. This amount can be expressed as a percentage of the total air required to fully burn the fuel being injected: the stoichiometric air requirement. This amount defines how rich the mixture is that undergoes the autoignition and premixed-burn first phase of diesel combustion. That region is shown in the annotated schematic of an individual spray in Fig. 10.13: it lies downstream of the rich unburned-mixture region, which is just downstream of flame lift-off, and ahead of the products of rich combustion region, which is further downstream. This percentage, ξ_{st} is given by:

$$\xi_{st}(\%) = \frac{100}{\phi} = 100 \left[\frac{(1 + 16\tilde{L}_{LO}^2)^{1/2} - 1}{2(F/A)_s} \right]$$

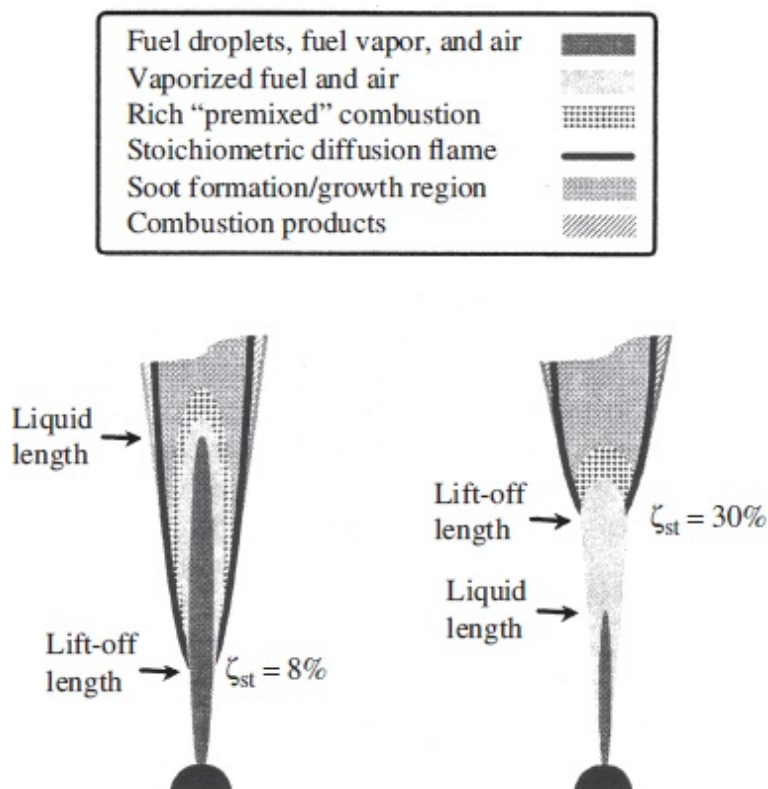
(10.38)

where $\tilde{L}_{LO}$ is the lift-off length/ x^+ , and x^+ is the characteristic length scale of the fuel spray defined by Eq. (10.19b).²²

The variables that affect the diffusion flame lift-off length are: ambient gas temperature and density, oxygen concentration in the ambient gas (air or air plus EGR), orifice diameter, orifice pressure drop, local fuel-air equivalence ratio. The stable lift-off location downstream of the nozzle orifice is a balance between the downstream (turbulent) flow in the outer regions of the spray, and the upstream diffusion-flame area expansion speed. This balances the downstream spray momentum convecting the flame away from the nozzle against the upstream flame propagation, which depends on the local rate of chemical energy release. Effectively, the spreading quenches at the lift-off distance. With an easier-to-quench mixture state, flame lift-off is farther from the injector nozzle: with a harder-to-quench mixture state, the flame stabilizes closer to the nozzle.^{22, 49}

This flame lift-off length discussion when connected with the earlier discussion of the liquid containing core of the spray (quantified by its liquid length, Sec. 10.4.6), and the premixed- and diffusion-phases of combustion, suggests that changes in the relative spatial relationship between air entrainment, fuel vaporization, and combustion zones can occur and substantially change conditions within the spray. Figure 10.51 shows two extremes: on the left the liquid length is substantially longer than the (short) flame lift-off length; and on the right the reverse occurs with a short liquid length and long lift-off distance. On the right, the amount of (hot) air entrained is almost one-third of the stoichiometric air requirement, the premixed rapid combustion process will be less rich (a ϕ of order 3), and the amount of soot formed subsequently will be less (see Fig. 10.56). On the left, the entrained air is less than 10%: ϕ at the lift-off point will be very rich ($\phi \sim 10$), and would be expected to become less-rich moving downstream through the elongated premixed-combustion region as additional air is entrained. Here the increased richness would likely increase the soot yield significantly.

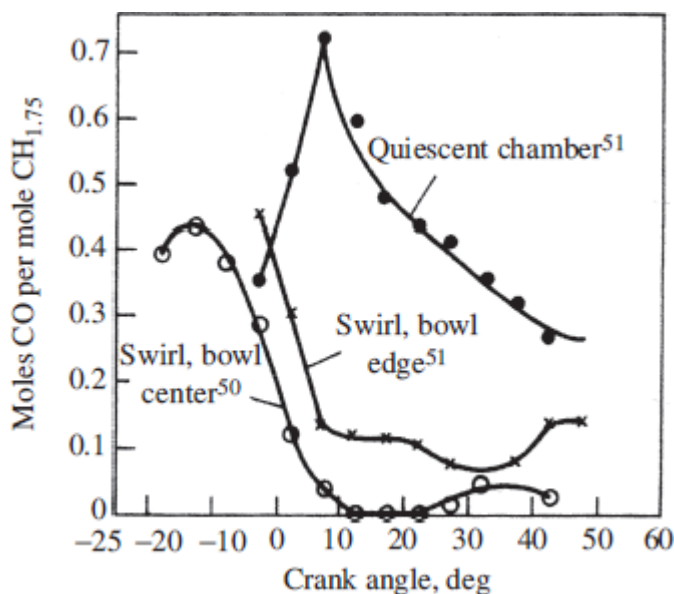
Figure 10.51 Schematics showing how the relative spatial relationship between fuel vaporization and the location of the DI diesel combustion zones may vary. See discussion in text. On the left, ambient gas conditions are 1100 K and 23 kg/m³, the orifice pressure drop is 40 MPa, and orifice diameter is 250 μ m. On the right, ambient gas conditions are 1000 K and 20 kg/m³, orifice pressure drop is 200 MPa, and orifice diameter is 100 μ m.⁴⁹



The role of air swirl in promoting more rapid fuel-air mixing in medium-size and smaller diesel engines is evident from gas-sampling studies in engines with these different combustion systems. The variation of gas species and unburned hydrocarbon concentrations within critical regions of the combustion chamber have been mapped out, and these data show that during the

earlier stages of injection and combustion, the boundaries of the individual sprays can be identified as they are convected around the combustion chamber bowl by the swirl. The fuel distribution within the combustion chamber is highly nonuniform. However, additional air has mixed into the spray due to swirl to bring the fuel/air equivalence ratios within the spray, in the outer regions of the chamber, closer to stoichiometric values.⁵⁰ This substantially different character of the spray with swirl is clear from the data in Fig. 10.52. While the local ϕ values obviously depend on the relation of the sample valve location to spray position at any given crank angle, the much lower values of moles CO per mole $\text{CH}_{1.75}$ (and thus of equivalence ratio) with swirl relative to values in a quiescent chambers, during injection and the early stages of combustion, clearly indicate that swirl enhances fuel-air mixing rates substantially. As combustion ends, the fuel distribution within the combustion chamber, at least on a gross geometric scale, becomes more uniform. However, early in the combustion process, the high CO levels found in these different diesel combustion systems, as shown in Fig. 10.52, indicate that the burned gases are only partially reacted. With quiescent chambers this is largely due to lack of oxygen. With swirl, however, significant oxygen is present; whether the high CO with swirl is due to kinetic limitations or to smaller-scale mixture non-homogeneities is unclear.

Figure 10.52 Time-resolved gas-composition data obtained from rapid-acting sampling valves from within the combustion chambers of quiescent and high-swirl bowl-in-piston DI diesel engines. CO concentrations (in moles CO per mole fuel, $\text{CH}_{1.75}$) are shown: (1) on the spray centerline with larger quiescent chamber diesel engine; (2) at edge of deep bowl, and within shallow bowl, in engines with swirl.^{50, 51}



10.6.3. Fuel-Air Mixing and Burning Rates

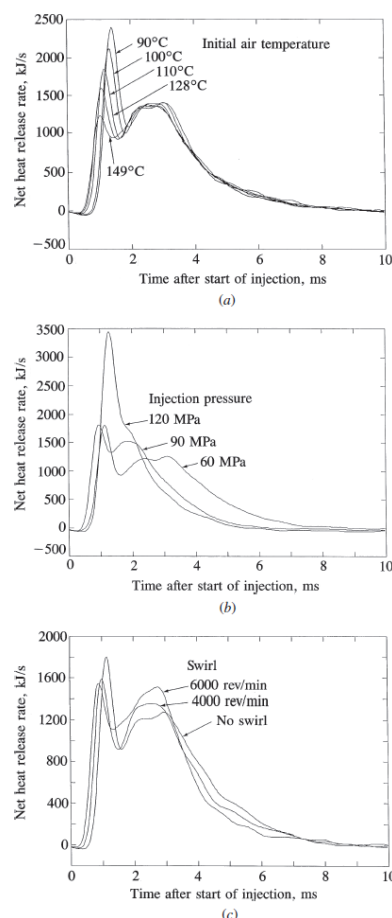
The energy-release profile during diesel combustion obtained from heat-release analyses of cylinder pressure data identifies two main stages of combustion (see Fig. 10.8). The first is the premixed-combustion phase, when the fuel which has mixed sufficiently with air within the spray to form an ignitable fuel-rich mixture during the delay period, releases part of its chemical energy. The second is the mixing-controlled combustion phase, where a diffusion flame surrounds the spray and completes the fuel oxidation process. Experimental evidence from heat-release analysis indicates that the majority of the fuel (usually more than 80%) burns during this second mixing-controlled phase.

That the fuel-burning or heat-release rate during the main diffusion flame combustion process is predominantly mixing controlled is supported by many types of evidence. Estimates of the rate at which fuel-air mixture with composition within the combustible limits is produced in diesel sprays under typical engine conditions, based on a variety of turbulent-jet models of the spray, show that mixing rates and burning rates are comparable in magnitude. Estimates of characteristic times for the

turbulent-jet mixing processes in diesel combustion chambers show these to be comparable to the duration of the heat-release process, and much longer than characteristic times for fuel evaporation and the thin diffusion flame chemical kinetics.

Then, measured diesel heat-release profiles during this main part of combustion show trends with engine design and operating parameter changes that correspond to fuel-air mixing being the primary controlling factor. Examples of heat-release profiles measured in rapid-compression-machine studies of diesel combustion, shown in Fig. 10.52, illustrate this clearly. The rapid-compression machine had a disc-shaped chamber of 10 cm diameter with a 3.1-cm clearance height at the end of a compression process through a volume ratio of 15.4; a five-hole centrally located fuel-injector nozzle was used. Figure 10.53a shows the heat-release profiles for different initial (prior to compression) air temperatures, which produce different ignition delays. Longer delays allow more fuel to mix to within combustible limits during the delay, so the peak premixed heat-release rate increases. However, the mixing-controlled-phase heat-release-rate magnitudes are essentially the same because the spray-mixing processes are little affected by these changes in air temperature. Figures 10.53b and c show that heat-release rates throughout the combustion process are increased by increased fuel-injection rate (achieved by increasing the fuel-injection pressure) and by swirl. Both these changes increase the fuel-air mixing rates within the fuel spray and therefore increase the heat-release rate during the mixing-controlled combustion phase.

Figure 10.53 Net heat-release rates, as a function of time after start of injection, calculated from cylinder pressure data from rapid-compression-machine studies of DI diesel combustion. (a) Effect of varying initial air temperature: 4000 rev/min swirl, injection pressure 60 MPa, 100°C initial temperature corresponds, via Eq. (10.37a), to about 900 K. (b) Effect of varying injection pressure: no swirl. (c) Effect of varying swirl: injection pressure 60 MPa.¹⁴



Diesel engine heat-release-rate trends, as design and operating variables are changed, can be related to mixing rates in analogous fashion. Table 10.4 summarizes the trends that have been investigated. The directional effects of changes in engine parameters on the ignition delay period and the fuel-air mixing rate are all consistent with the measured changes in

premixed and mixing-controlled heat-release rates. The controlling role of fuel-air mixing in the diesel engine fuel spray on combustion is clear.

Table 10.4 Effects of engine design and operating variables on heat-release rates¹

Parameter varied	Effect on			
	τ_{id}	$\dot{m}_m$	$\dot{Q}_p$	$\dot{Q}_m$
Injection rate ↑	0	↑	↑	↑
Turbocharger boost ↑	↓	0	↓	0
Compression ratio ↓	↑	0	↑	0
Number of injector holes ↑	0	↑	↑	↑
Injection advance ↑	↑	0	↑	0
Swirl ↑	0	↑	↑	↑
Intake-air temperature ↓	↑	0	↑	0
Injection pressure ↑	0	↑	↑	↑
Speed ↑	0	↑	↑	↑
τ_{id} , ignition delay; $\dot{m}_m = (dm/dt)_m$, fuel-air mixing rate; $\dot{Q}_p = (dQ/dt)_p$, heat-release rate during premixed-combustion phase; $\dot{Q}_m = (dQ/dt)_m$, heat-release rate during mixing-controlled combustion phase; ↑ increase; ↓ decrease; 0 minor effect.				
Source: From Plee and Ahmad. ³⁹				

[1] The spray and flame tracings shown in Fig. 10.49 do not show this flame lift-off. This is likely due to the fuel injection details in this older study resulted in short lift-off distances.

10.7. ALTERNATIVE COMPRESSION-IGNITION COMBUSTION APPROACHES

Modern diesel engine injection equipment permits multiple fuel injection pulses and control of the fuel injection rate during each pulse. So the basic diesel combustion process of a fuel-rich "premixed flame" within each individual fuel spray, followed by the burning to completion of this rich mixture in a diffusion flame around each spray, can be modified.

The various approaches of interest can be categorized into two broad groups: multiple fuel injections intended to improve the basic diesel combustion process, and alternative combustion concepts which are significantly different from the standard mode. [Table 10.5](#) provides a framework for categorizing these alternative compression-ignition combustion strategies, and a summary of their key features and operating characteristics. Additional discussion of diesel combustion with multiple fuel injections, and of what are often called *advanced combustion concepts* where the fuel-air mixing and burning processes differ significantly from the normal diesel engine processes, follows in [Secs. 10.7.1](#) and [10.7.2](#), respectively.

Table 10.5 Summary of alternative compression-ignition combustion and injection strategies

Feature	Diesel, with multiple injections	Advanced combustion	
		Premixed	Late
Primary objective:	Reduce the noise, emissions of standard diesel combustion	Achieve low burned gas temperatures (with attendant emissions reduction benefits)	
Approach:	Pilot injection to improve main diesel combustion process. Additional injections to control emissions (via exhaust treatment devices)	Early fuel injection, lean or dilute mixture, long induction period, premixed fuel-air mixture	Late fuel injection, high EGR levels and mixing rates to achieve faster, after TC, combustion process
Key features:	Common rail high-pressure injection system with piezoelectric injector	Early injection (intake or early compression), well- or pre-mixed fuel-air mixture, spontaneous ignition after long induction period	Close to TC injection, high levels of cooled EGR, high swirl, fast fuel-air mixing, moderate ignition delay
Combustion characteristics:	More flexible and controlled diesel combustion (premixed burning and diffusion flame)	Lean and/or dilute premixed autoignition process with low burned-gas temperatures	Dilute, fast-burning; spontaneously ignited, premixed and mixing-controlled burning processes, early in the expansion stroke
Efficiency improvement potential:	Moderate	High	Moderate
Emissions levels:	Improved PM and NO _x compared to standard diesel	Lowest PM and NO _x emissions	Lower NO _x , PM improved, compared to standard diesel
Fuels:	Standard diesel fuel	Needs low cetane number fuel	Standard diesel fuel

As [Table 10.5](#) indicates, use of multiple injections has several objectives. The different injection pulses can modify the in-cylinder gas state to reduce the ignition delay of the primary main combustion process (through use of pilot injection), enhance the burning rate of the latter portion of the main injected fuel pulse (split injection), or reduce the in-cylinder particulate loading (post injection). As use of diesel exhaust after treatment devices to oxidize hydrocarbons, reduce nitrogen oxides, and filter and burn up particulates expands, late fuel injection pulses are being used to release additional fuel chemical

energy at appropriate times in the engine cycle to raise gas temperatures locally, in the bulk charge, or in the exhaust.⁵² Figure 10.54 illustrates these injection strategies: it also indicates the relative magnitude (proportion) of these several fuel injection pulses.

Figure 10.54 Schematic of multiple fuel injection strategy options identifying the objectives of the different injection pulses. DOC, diesel oxidation catalyst.⁵²

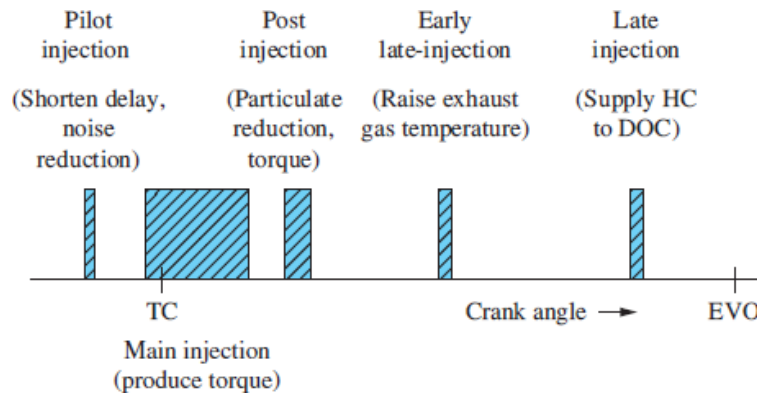


Table 10.5 divides the advanced combustion strategies into two classes with substantive differences in basic approach: (1) early injection of (some) fuel so the fuel-air mixture burned is closer to being *premixed* and (2) where injection is retarded to achieve *late* (and usually fast) combustion, largely to reduce in-cylinder pollutant formation.

10.7.1. Multiple-Injection Diesel Combustion

Figure 10.54, as briefly discussed in the preceding section, illustrates several different multiple fuel-injection strategies in relation to the main fuel injection pulse, which normally occurs around TC. The primary role of each injection event is indicated as is their relative injection duration.

10.7.1.1. Pilot Injection

As mentioned, pilot injection is increasingly used to shorten the ignition delay between start of injection, and start of combustion of the main injection of the standard diesel injection strategy. Its objective is primarily to reduce engine noise (see Sec. 15.7.1). The combustion of the pilot injection fuel with its entrained air raises the temperature of the mixture around the injector, thus reducing the ignition delay of the main-injection fuel. The shorter the delay, the less rapid the initial burning rate of the main fuel pulse. Because there is less time to "prepare" the early injected portion of the main injection for rapid combustion, the initial combustion-generated rate of pressure rise is reduced resulting in lower pressure-rise-rate generated noise levels. This pilot-injection effect is most advantageous at the lighter engine loads and prior to the engine fully warming up. At higher engine loads and speeds, the in-cylinder gases at time of injection are hotter so the ignition delays are inherently shorter so the impact of pilot injection is significantly reduced.

A second effect of pilot injection is that combustion of the pilot fuel adds a diluent—burned gases—to the "air" surrounding the injector. Thus, the flame temperatures in the diffusion flame surrounding each main-injection fuel spray are reduced, because the oxygen combusted in these diffusion flames has this additional dilution beyond the nitrogen in the air and the burned gas residual from the previous cycle. This reduces the NO_x formation rate in the diffusion flame (see Sec. 11.2.4): it may also reduce the soot oxidation rate in the diffusion flame and adversely affect the particulate emissions (Sec. 11.5.5).

Typically, pilot fuel injection amounts are a few percent (5 to 15%) of the total fuel injected, depending on load, and its injection is started some 15 to 35 degrees CA before TC. Pilot fuel injection results in significant reductions in the ignition delay—by a

factor of about two.

10.7.1.2. Post Injection

Post or after injection primarily impacts particulate emissions. Post injection with a relatively short delay (some 10 to 15° CA) between end of main injection and start of post injection pulses usefully reduces soot emissions (up to 40%). Particulate emissions are usually increased with post injections closer to the end of main injection, and with significantly later post injections. Use of post injections is primarily a strategy for reducing particulate emissions at lighter loads and lower engine speeds.

The late injection pulses illustrated in [Fig. 10.54](#) are used to enable the performance of exhaust system aftertreatment devices. Fuel injection (and its combustion) part-way through the expansion stroke raises the exhaust gas temperature and is used to initiate the particulate burn-up or regeneration process in particulate traps/filters. Injection late in the expansion stroke is used when a reducing agent (hydrocarbons) is needed to reduce NO_x in a lean exhaust gas in a *selective catalytic reduction* (SCR) system, which is used for NO_x removal (see [Sec. 11.6.3](#)).

10.7.2. Advanced Compression-Ignition Combustion Concepts

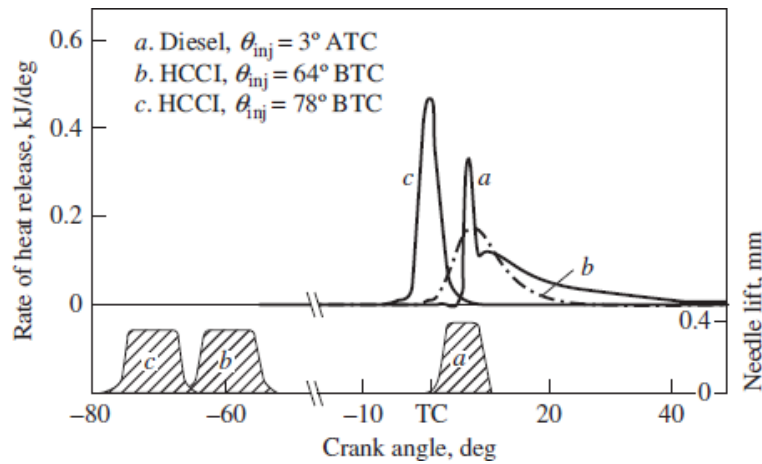
Spontaneous fuel-air mixture ignition (compression ignition) can be used to initiate engine combustion using a variety of fuels, and fuel injection and mixing strategies. These alternative approaches are being explored for their high efficiency, and low particulate and oxides of nitrogen emissions potential. In the late 1970s, homogeneous charge two-stroke cycle engines were successfully operated with compression ignition.⁵³ Najt and Foster⁵⁴ analyzed homogeneous-charge compression ignition (HCCI) fundamentals, concluding that the ignition process is controlled by the low-temperature hydrocarbon oxidation kinetics while the main energy releasing process is controlled by the high-temperature hydrocarbon oxidation kinetics. See [Sec. 9.6.2](#) for a more detailed discussion of the autoignition chemistry of hydrocarbon-air mixtures.

The low temperature oxidation reactions in engines start to occur at about 800 K. Intermediate species such as aldehydes are produced through abstraction of H-atoms from the fuel molecules. These peroxides react with OH to produce a growing number of active radicals, converting chemical energy into thermal energy at an accelerating rate. The more direct oxidation path then commences at temperatures between 950 and 1000 K, and intermediates such as olefins and alkyl radicals are increasingly rapidly consumed. Depending on the relative rates of the low- and high-temperature paths, which are governed by the fuel's composition and the molecular structure of the individual hydrocarbons involved and mixture temperature, a two-stage ignition process may occur (when the low-temperature oxidation path is slower). Or, the two processes may effectively merge into one continuous process.⁵⁵

In 1984, Oppenheim⁵⁷ (see also [Sec. 10.5.3](#)) proposed the concept of HCCI where the unburned mixture conditions necessary for compression ignition in engines were achieved by controlling temperature (and dilution) by varying the amounts of residual gases and EGR. It also became clear that the fuel's autoignition characteristics were a critical factor in this controlled autoignition combustion process.

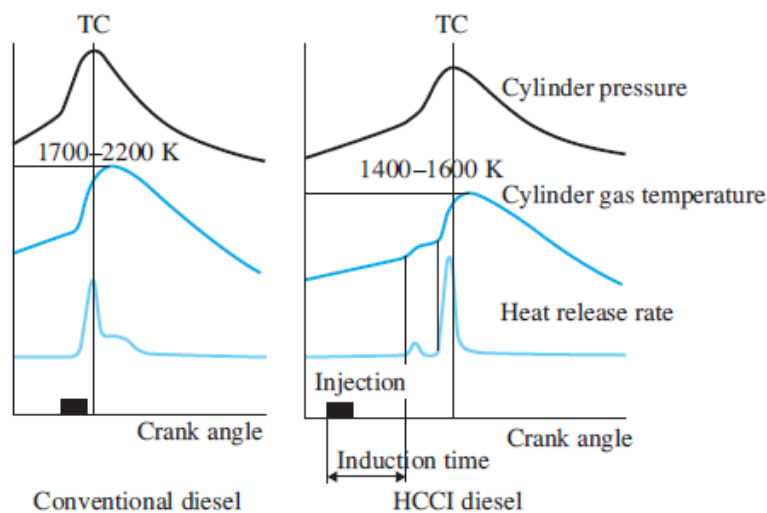
[Figure 10.55](#) illustrates the basic features of this HCCI combustion process, by comparing a conventional diesel combustion process with the HCCI process. The HCCI combustion process occurs in a (fairly) well-mixed lean charge achieved by early injection, in this case of a fuel with a long chemical induction time (ignition delay). The smooth two-stage heat-release-rate profile of HCCI is clear, and the HCCI profiles contrast with the conventional diesel heat-release rate profile of a rapid/premixed initial burning phase followed by the mixing-controlled diffusion-flame phase.

Figure 10.55 Comparison of (a) the heat-release rate profile of a standard diesel combustion process with start of injection (SOI) 3° ATC and start of combustion (SOC) a few degrees later, with two HCCI engine cases: (b) HCCI heat-release rate profile with SOI at 64° BTC and SOC shortly after TC; (c) with SOI at 78° BTC and SOC 4° BTC.⁵⁶



The advantages of spontaneously igniting a homogeneous lean and/or dilute mixture are the low burned gas temperatures that result which avoid the fuel-air equivalence ratio (ϕ) and temperature (T) regimes where soot formation and where NO formation occur. As shown in Fig. 10.56, the cylinder gas temperatures (average in-cylinder values) for the HCCI combustion are below about 1600 K, and are relatively uniform. With conventional diesel combustion, in-cylinder gas temperatures are significantly higher (and within the diffusion flame that surrounds each fuel spray, see Fig. 10.13, reach about 2700 K), thus forming significant soot within the rich reacting spray as well as NO in the close-to-stoichiometric diffusion flame. The NO_x emissions reduction potential of the HCCI type of combustion process is illustrated by the relative engine-out NO_x emission levels of the different combustion process examples shown in Fig. 10.55. The base diesel case (a) had NO_x levels of 400 ppm: the HCCI cases (b) and (c) had levels of 250 ppm, and 20 ppm, respectively.⁵⁶

Figure 10.56 Profiles of cylinder pressure, average gas temperature, and heat-release rate versus crank angle for a standard diesel combustion process, and an HCCI combustion process, are shown. The much longer ignition delay with HCCI combustion enables more complete mixing between fuel vapor and air, and thus eliminates the mixing-controlled phase of diesel combustion.⁵⁵



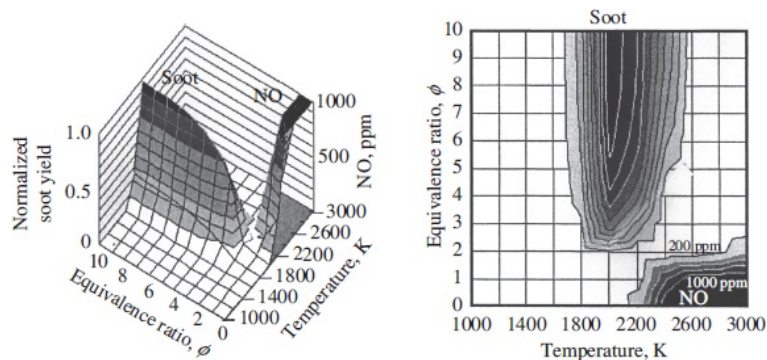
Many different approaches to achieving these objectives through spontaneous ignition of well-mixed lean (and dilute) mixtures have been proposed and examined. These are predominantly *low temperature combustion* (LTC) approaches for the reasons that Fig. 10.56 illustrates. They can be grouped in two categories as "premixed" or "late" combustion modes (see

Table 10.5). These correspond to *early* fuel injection approaches which allow much more time for premixing the fuel and air, and *late* injection which, due to the lengthening of the ignition delay, results in better fuel-air mixing than with the standard diesel injection and combustion processes. However, with late injection, dilution with EGR is usually needed to further reduce the peak burned-gas temperatures, and both of these changes require a speeding up of the combustion process to maintain high engine efficiency.

Many different labels have been proposed for the various approaches: for example, premixed charge compression ignition (PCCI), reaction controlled compression ignition (RCCI), modulated kinetics (MK) (a late combustion concept). Each of these titles abbreviates the essential feature of these advanced combustion concepts—more complete mixing (pre-mixing) of the charge, control of combustion initiation and progression through managing the kinetics of the combustion chemistry, and reducing the burned gas (flame) temperatures through dilution with air, residual gas, or EGR. See Ref. 55 for an overview.

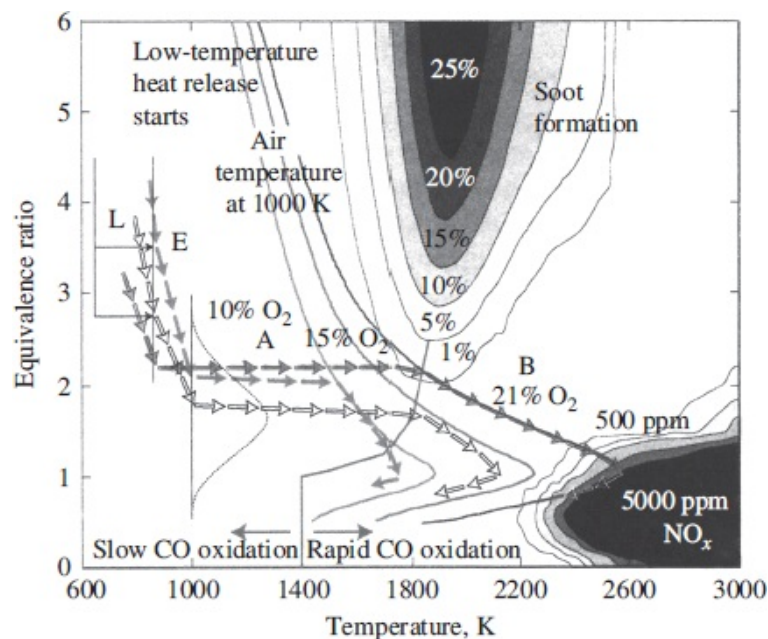
The fuel-air equivalence ratio temperature paths of an early injection and late injection low-temperature combustion concept identify both the opportunities and problems of these advanced combustion approaches. Figure 10.57 shows the gas temperatures that result from burning a homogeneous mixture of diesel-like fuel and air at the specified fuel-air equivalence ratio and temperature of 1000 K. For very rich mixtures ($\phi \geq 2$) this temperature can exceed the equilibrium burned gas temperature by several hundred degrees K before endothermic reactions lower the temperature to its equilibrium value. The soot formation region (richer than $\phi \approx 2$ with T between 1600 K and 2200 K) with percentages of fuel carbon converted to soot, and the NO formation region, are shown in Fig. 10.58. Also shown are lines of 21% (air), 15%, and 10% oxygen (corresponding to increasing levels of EGR). A typical path for a fuel parcel in a conventional diesel combustion process (with air) ignites at a ϕ in the 2 to 4 range, and enters the soot formation region as rich reacting fuel-air mixture; then in the diffusion flame ($\phi \approx 1$, $T = 2700$ K) the burned gases are in the NO formation region. A and B indicate the premixed (rapid) burning phase and mixing-controlled (diffusion) burning phase of the diesel combustion process.

Figure 10.57 Fuel-air equivalence ratio versus burned gas temperature plot relevant to diesel combustion analysis showing soot formation (rich mixtures) and NO formation (high temperature, near stoichiometric) regions.⁵⁸



A path characteristic of an early-injection high-dilution combustion concept is shown in Fig. 10.58 by the filled-in arrows (E). Typical operating conditions for this path are 10% oxygen concentration (65% EGR) and SOI at 20° crank-angle BTC. For the late-combustion example (L, open arrows), the SOI was close to TC and the O₂ concentration was about 15%. The longer delay (relative to conventional diesel combustion, A, solid heavier arrows) gives a "less rich" mixture composition at SOC, but temperatures are lower during the bulk of the combustion process due to the EGR dilution. There is still a diffusion-controlled burning process with the late-injection case but its NO formation impact is reduced due to the increased dilution and its later occurrence, and thus lower peak flame temperature. The temperature boundary between more rapid and less rapid CO oxidation regimes is indicated. This is important, because for rich mixtures ($\phi > 1$) the products of premixed combustion (CO, H₂, and other organic compounds) must be fully oxidized to utilize the fuel's chemical energy completely.

Figure 10.58 Paths for two low-temperature combustion modes in the fuel-air equivalence ratio (ϕ) and burned gas temperature (T) plane: early injection (E), late injection (L), and standard diesel combustion (A). High EGR contours (15, 10% O_2) are also shown.⁵⁹



Successful implementation of these various HCCI and low-temperature combustion concepts does lead to higher engine efficiencies and lower particulate and NO emissions in those regions of the engine-operating map where these advanced combustion concepts can be used. However, these regions are generally limited to lighter loads (mep values well below the maximum) by the very rapid HCCI combustion rate, which follows ignition, which leads to knock-like pressure oscillations. With extremely lean operation (and at high dilution rates with EGR) misfire becomes a problem resulting in high levels of engine-out hydrocarbon and carbon monoxide emissions. Also, engine control depends on managing the charge temperature and composition in each cycle appropriately, and is made difficult by the strong dependence of spontaneous ignition of fuel-air mixtures to these variables, especially temperature. Use of "partially premixed" fuel-air-residual gas mixtures with temperature and composition stratification alleviates such challenges to a degree, but these limits to a broad operating range still remain.

A basic issue is that engines usually need to operate over a wide speed range (e.g., SI engines 600 to 6000 rev/min). Diesels and spark-ignition engines employ a combustion process whose rate scales fundamentally with the in-cylinder fluid turbulence, which scales with engine speed. Kinetically-controlled engine combustion concepts do not have that inherent combustion-rate scaling with engine speed.

10.8. PROBLEMS

- 10.1.** Describe the sequence of processes which must occur before the liquid fuel in the injection system in a direct-injection compression-ignition engine is fully burned.
- 10.2.** Small high-swirl direct-injection CI engines have fuel conversion efficiencies which are about 10% higher than values typical of equivalent indirect-injection engines. (IDI engines were used previously.) What combustion-system-related differences between DI and IDI diesels contribute to this higher efficiency?
- 10.3.** In a diesel engine, because the fuel distribution is nonuniform the burned gas temperature is nonuniform. Consider small fuel-air mixture elements initially at 1000 K and 6.5 MPa at top-center with a range of equivalence ratios. Each element burns at essentially constant pressure. Calculate (using the charts in [Chap. 4](#), or an appropriate chemical equilibrium thermodynamic computer code) the burned gas temperature for mixture equivalence ratios of 0.4, 0.6, 0.8, 1.0, 1.2. (Assume the fuel is isooctane.) Note: the reaction zone in the main diesel combustion phase (a diffusion flame) is stoichiometric.
- 10.4.** The levels of combustible species in the exhaust of a direct-injection diesel engine are: HC, 0.8 g/kW·h; CO, 3 g/kW·h; hydrogen, negligible; particulates, 0.7 g/kW·h. If the specific fuel consumption is 210 g/kW·h, calculate the combustion efficiency.
- 10.5.** Consider a turbocharged 8-liter six-cylinder direct-injection diesel engine (see [Fig. 15.34](#)) operating at maximum power (210 kW), 2500 rev/min and an equivalence ratio of 0.6. Bore = 114 mm, stroke = 130 mm, $r_c = 16.5$. Estimate the following:
- Mass of air in each cylinder per cycle, mass and volume (as liquid) of diesel fuel injected per cylinder per cycle.
 - Estimate the spray penetration as a function of time, the liquid core length (of an evaporating spray), the lift-off length, and the average drop size. Use equations from the relevant sections of [Chap. 10](#). You will need to make several assumptions about parameters that define the engines operating conditions and fuel spray details. List and briefly explain your assumptions.
 - Assuming all fuel droplets are the same size as the average drop, how many drops are produced per injection? If these drops are uniformly distributed throughout the air charge at TC, what is the approximate distance between drops? (Neither of these assumptions is correct; however, the calculations illustrate the challenges of the fuel-air mixing process.)
- 10.6.** Estimate the following quantities for a typical direct-injection diesel fuel spray. The injection pressure is 500 bar; the cylinder pressure during injection is 50 bar.
- Assuming that the flow through the nozzle orifice is incompressible and quasi-steady, estimate the liquid fuel velocity at the orifice exit. At this velocity, how long would the fuel take to reach the cylinder wall? The bore is 125 mm.
 - Each nozzle orifice diameter d_n is 0.34 mm and $L_n/d_n = 4$. Determine the spray spreading angle and plot spray tip penetration versus time.
 - Use Eq. (10.28) to estimate the initial average drop size assuming that the injection process in (a) above continues for 1 millisecond and the injector nozzle has four orifices.
- 10.7.** Diesel fuel is injected as a liquid at room temperature into air at 50 bar and 800 K, close to TC at the end of compression. If the overall equivalence ratio is 0.7, estimate the reduction in average air temperature which would occur when the fuel is fully vaporized and uniformly mixed. Assume such mixing takes place at constant volume prior to any combustion.

10.8. Using Eq. (10.35) estimate the ignition delay in milliseconds and crank-angle degrees for these operating conditions: low swirl DI diesel at 600 rev/min; high swirl DI diesel 1800 rev/min, with lower and higher compression ratio, respectively: other details as in Table 10.3. The fuel cetane number is 45; stroke = 135 and 83 mm, respectively. Discuss differences in the predicted ignition-delay values.

10.9. The compression ratio of truck diesel engines must be set at about 16 so that the engine will start when cold. Using Eqs. (10.35) to (10.37) develop a graph of τ_{id} (in degrees) as a function of compression ratio for $r_c = 12$ to 20. Assume $p_i = 1$ bar, $T_i = 255$ K, $n = 1.13$, speed = 100 rev/min, bore = stroke = 120 mm, fuel cetane number = 45. If the ignition delay must be less than 20° CA for satisfactory starting, what compression ratio is required?

10.10. Data and equations in this chapter can be used to estimate the fraction of the fuel injected into a direct-injection diesel engine which burns in the premixed phase. Use this information to explain why, for turbocharged DI diesel engines the ignition delay and relative amount of fuel burned during the premixed combustion phase is less important than it is for naturally-aspirated engines.

10.9. REFERENCES

1. Heywood, J. B.: *Internal Combustion Engine Fundamentals*, 1st edition, McGraw-Hill, New York, 1988.
2. Alcock, J. F., and Scott, W. M.: "Some More Light on Diesel Combustion," *Proc. Auto. Div., Instn Mech. Engrs*, No. 5, pp. 179–191, 1962–1963.
3. Scott, W. M.: "Understanding Diesel Combustion through the Use of High Speed Moving Pictures in Color," SAE paper 690002, *SAE Trans.*, vol. 78, 1969.
4. Neitz, A., and D'Alfonso, N.: "The M.A.N. Combustion System with Controlled Direct Injection for Passenger Car Diesel Engines," SAE paper 810479, 1981.
5. Austen, A. E. W., and Lyn, W.-T.: "Relation between Fuel Injection and Heat Release in a Direct-Injection Engine and the Nature of the Combustion Processes," *Proc. Instn Mech. Engrs*, No. 1, pp. 47–62, 1960–1961.
6. Lyn, W.-T.: "Study of Burning Rate and Nature of Combustion in Diesel Engines," in *Proc. Ninth International Symposium on Combustion*, pp. 1069–1082, The Combustion Institute, Pittsburg, PA, 1962.
7. Cheng, W., and Gentry, R.: "Effects on Charge Non-Uniformity on Diesel Heat Release Analysis," SAE paper 861568, 1986.
8. Gatowski, J. A., Balles, E. N., Chun, K. M., Nelson, F. E., Ekchian, J. A., and Heywood, J. B.: "Heat Release Analysis of Engine Pressure Data," SAE paper 841359, *SAE Trans.*, vol. 93, 1984.
9. Krieger, R. B., and Borman, G. L.: "The Computation of Apparent Heat Release for Internal Combustion Engines," ASME paper 66-WA/DGP-4, in *Proc. Diesel Gas Power*, ASME, 1966.
10. Dec, J. E.: "A Conceptual Model of DI Diesel Combustion Based on Laser-Sheet Imaging," SAE paper 970873, *SAE Trans.*, vol. 106, pp. 1319–1348, 1997.
11. Flynn, P. F., Durrett, R. P., Hunter, G. L., zur Loye, A.O., Akinyemi, O.C., Dec, J.E., and Westbrook, C.K.: "Diesel Combustion: An Integrated View Combining Laser Diagnostics, Chemical Kinetics, and Empirical Validation," SAE paper 1999-01-0509, *SAE Trans.*, vol. 108, pp. 587–600, 1999.
12. Bosch, R.: *Automotive Handbook*, 6th edition, Robert Bosch GmbH, SAE, 2004.
13. Kamimoto, T., Kobayashi, H., and Matsuoka, S.: "A Big Size Rapid Compression Machine for Fundamental Studies of Diesel Combustion," SAE paper 811004, *SAE Trans.*, vol. 90, 1981.
14. Balles, E.: "Fuel-Air Mixing and Diesel Combustion in a Rapid Compression Machine," Ph.D. Thesis, Department of Mechanical Engineering, MIT, Cambridge, MA, June 1987.
15. Browne, K. R., Partridge, I. M., and Greeves, G.: "Fuel Property Effects on Fuel/Air Mixing in an Experimental Diesel Engine," SAE paper 860223, 1986.
16. Bracco, F. V.: "Modeling of Engine Sprays," SAE paper 850394, 1985.
17. Kuo, T., and Bracco, F. V.: "Computations of Drop Sizes in Pulsating Sprays and of Liquid-Core Length in Vaporizing Sprays," SAE paper 820133, *SAE Trans.*, vol. 91, 1982.
18. Reitz, R. D., and Bracco, F. V.: "Mechanism of Atomization of a Liquid Jet," *Phys. Fluid*, vol. 25, no. 10, pp. 1730–1742, 1982.
19. Reitz, R. D., and Bracco, F. V.: "On the Dependence of Spray Angle and Other Spray Parameters on Nozzle Design and Operating Conditions," SAE paper 790494, 1979.
20. Wu, K.-J., Su, C.-C., Steinberger, R. L., Santavicca, D. A., and Bracco, F. V.: "Measurements of the Spray Angle of Atomizing Jets," *J. Fluids Engng*, vol. 105, pp. 406–413, 1983.

21. Naber, J. D., and Siebers, D. L.: "Effects of a Gas Density and Vaporization on Penetration and Dispersion of Diesel Sprays," SAE paper 960034, *SAE Trans.*, vol. 105, pp. 82–111, 1996.
22. Siebers, D. L.: "Recent Developments on Diesel Fuel Jets Under Quiescent Conditions," in *Flow and Combustion in Reciprocating Engines*, C. Arcoumanis and T. Kamimoto (eds), Ch. 2, Springer-Verlag, Berlin, 2009.
23. Hiroyasu, H., Kadota, T., and Arai, M.: "Supplementary Comments: Fuel Spray Characterization in Diesel Engines," in *Combustion Modeling in Reciprocating Engines*, J. N. Mattavi and C. A. Amann (eds), pp. 369–408, Plenum Press, New York, 1980.
24. Hiroyasu, H., and Kadota, T.: "Fuel Droplet Size Distribution in Diesel Combustion Chamber," SAE paper 740715, *SAE Trans.*, vol. 83, 1974.
25. Wu, K.-J., Reitz, R. D., and Bracco, F. V.: "Measurements of Drop Size at the Spray Edge near the Nozzle in Atomizing Liquid Jets," *Phys. Fluids*, vol. 29, no. 4, pp. 941–951, 1986.
26. Reitz, R. D., and Diwakar, R.: "Effect of Drop Breakup on Fuel Sprays," SAE paper 860469, 1986.
27. Hiroyasu, H.: "Diesel Engine Combustion and Its Modeling," in *Diagnostics and Modeling of Combustion in Reciprocating Engines*, pp. 53–75, COMODIA 85, Proceedings of Symposium, Tokyo, Sept. 4–6, 1985.
28. Arai, M., Tabata, M., and Hiroyasu, H.: "Disintegrating Process and Spray Characterization of Fuel Jet Injected by a Diesel Nozzle," SAE paper 840275, *SAE Trans.*, vol. 93, 1984.
29. El Wakil, M. M., Myers, P. S., and Uyehara, O. A.: "Fuel Vaporization and Ignition Lag in Diesel Combustion," in *Burning a Wide Range of Fuels in Diesel Engines*, *SAE Progress in Technol.*, vol. 11, pp. 30–44, SAE, 1967.
30. Borman, G. L., and Johnson, J. H.: "Unsteady Vaporization Histories and Trajectories of Fuel Drops Injected into Swirling Air," in *Burning a Wide Range of Fuels in Diesel Engines*, *SAE Progress in Technology*, vol. 11, pp. 13–29, SAE, 1967; also SAE paper 598C, 1962.
31. Kamimoto, T., and Matsuoka, S.: "Prediction of Spray Evaporation in Reciprocating Engines," SAE paper 770413, *SAE Trans.*, vol. 86, 1977.
32. Siebers, D. L.: "Scaling Liquid-Phase Fuel Penetration in Diesel Sprays Based on Mixing-Limited Vaporization," SAE paper 1999-01-0528, *SAE Trans.*, vol. 108, no. 3, pp. 703–708, 1999.
33. Edwards, C. F., Siebers, D. L., and Hoskin, D. H.: "A Study of the Autoignition Process of a Diesel Spray via High Speed Visualization," SAE paper 920108, *SAE Trans.*, vol. 101, sec. 3, pp. 187–204, 1992.
34. ASTM Method D 613, *Cetane Number of Diesel Fuels*, American Society of Testing and Materials.
35. ASTM D 976, *Standard Test Method for Calculated Cetane Index of Distillate Fuels*, American Society of Testing and Materials.
36. ASTM D 611, *Standard Test Methods for Aniline Point and Mixed Aniline Point of Petroleum Products and Hydrocarbon Solvents*, American Society of Testing and Materials.
37. Dec, J. E., and Espey, C.: "Chemiluminescence Imaging of Autoignition in a DI Diesel Engine," SAE paper 982685, *SAE Trans.*, vol. 107, sec. 3, pp. 2230–2254, 1998.
38. Higgins, B., Siebers, D., and Aradi, A.: "Diesel-Spray Ignition and Premixed-Burn Behavior," SAE paper 2000-01-0940, *SAE Trans.*, 2000.
39. Plee, S. L., and Ahmad, T.: "Relative Roles of Premixed and Diffusion Burning in Diesel Combustion," SAE paper 831733, *SAE Trans.*, vol. 92, 1983.
40. Wong, C. L., and Steere, D. E.: "The Effects of Diesel Fuel Properties and Engine Operating Conditions on Ignition Delay," SAE paper 821231, *SAE Trans.*, vol. 91, 1982.

41. Andree, A., and Pachernegg, S. J.: "Ignition Conditions in Diesel Engines," SAE paper 690253, *SAE Trans.*, vol. 78, 1969.
42. Glavincevski, B., Gulder, O. L., and Gardner, L.: "Cetane Number Estimation of Diesel Fuels from Carbon Type Structural Composition," SAE paper 841341, 1984.
43. Bosch, R.: *Automotive Handbook*, 8th edition, Robert Bosch GmbH, SAE, 2011.
44. Ryan, T., and Maly, R. R.: "Fuel Effects in Engine Combustion and Emissions," in *Flow and Combustion in Reciprocating Engines*, C. Arcoumanis and T. Yamamoto, T(eds), Ch. 8, Springer-Verlag, Berlin, Heidelberg, 2009.
45. Hardenberg, H. O., and Hase, F. W.: "An Empirical Formula for Computing the Pressure Rise Delay of a Fuel from Its Cetane Number and from the Relevant Parameters of Direct-Injection Diesel Engines," SAE paper 790493, *SAE Trans.*, vol. 88, 1979.
46. Dent, J. C, and Mehta, P. S.: "Phenomenological Combustion Model for a Quiescent Chamber Diesel Engine," SAE paper 811235, *SAE Trans.*, vol. 90, 1981.
47. Lakshminarayan, P. A., and Dent, J. C.: "Interferometric Studies of Vaporising and Combustion Sprays," SAE paper 830244, *SAE Trans.*, vol. 92, 1983.
48. Colella, K. J., Balles, E. N., Ekchian, J. A, Cheng, W. K., and Heywood, J. B.: "A Rapid Compression Machine Study of the Influence of Charge Temperature on Diesel Combustion," SAE paper 870587, 1987.
49. Siebers, D., and Higgins, B.: "Flame Lift-Off on Direct-Injection Diesel Sprays Under Quiescent Conditions," SAE paper 2001-01-0530, 2001.
50. Rhee, K. T., Myers, P. S., and Uyehara, O. A.: "Time- and Space-Resolved Species Determination in Diesel Combustion Using Continuous Flow Gas Sampling," SAE paper 780226, *SAE Trans.*, vol. 87, 1978.
51. Nightingale, D. R.: "A Fundamental Investigation into the Problem of NO Formation in Diesel Engines," SAE paper 750848, *SAE Trans.*, vol. 84, 1975.
52. Ikagemi, M., and Kamimoto, T.: "Conventional Diesel Combustion," in *Flow and Combustion in Reciprocating Engines*, C. Arcoumanis and T. Kamimoto (eds), Ch. 6, Springer-Verlag, Berlin, 2009.
53. Heywood, J. B., and Sher, E.: *The Two-Stroke Cycle Engine: Its Development Operatoin, and Design*, Taylor and Francis, Philadelphia, 1999.
54. Najt, P. M., and Foster, D.E.: "Compression-Ignited Homogeneous Charge Combustion," SAE paper 830264, *SAE Trans.*, 1983.
55. Ohsawa, K., and Kamimoto, T.: "Advanced Diesel Combustion," in *Flow and Combustion in Reciprocating Engines*, C. Arcoumanis and T. Kamimoto (eds), Ch. 7, Springer-Verlag, Berlin, 2009.
56. Hikosaka, N.: "A View of the Future of Automoyive Diesel Engines," SAE paper 972682, 1997: also in *Diesel Emissions and their Control*, W. A. Majewski and M. K. Khair (eds), SAE International, Warrendale, PA, 2006.
57. Oppenheim, A. K.: "The Knock Syndrome—Its Cures and its Victims," SAE paper 841339, 1984.
58. Akihama, K., Takatori, Y., Inagaki, K., Sasaki, S., and Dean, A.M.: "Mechanism of the Smokeless Rich Diesel Combustion by Reducing Temperature," SAE paper 2001-01-0655, 2001.
59. Kook, S., Bae, C., Miles, P.C., Choi, D., and Pickett, L.M.: "The Influence of Charge Dilution and Injection Timing on Low-Temperature Diesel Combustion and Emissions," SAE paper 2005-01-3837, 2005.

11.0. CHAPTER PRELIMINARIES

NATURE AND EXTENT OF PROBLEM

NITROGEN OXIDES

CARBON MONOXIDE

HYDROCARBON EMISSIONS

PARTICULATE EMISSIONS

EXHAUST GAS TREATMENT

PROBLEMS

REFERENCES

11.1. NATURE AND EXTENT OF PROBLEM

Spark-ignition (SI) and diesel engines are major sources of air pollutants. The spark-ignition engine exhaust gases contain oxides of nitrogen (nitric oxide, NO, and small amounts of nitrogen dioxide, NO₂—collectively known as NO_x), carbon monoxide (CO), and organic compounds, which are unreacted or partially reacted fuel hydrocarbons (HC). The relative amounts of these *engine-out emissions* that enter exhaust depend on engine design and operating conditions but are of order: NO_x, 500 to 1000 ppm or 20 g/kg fuel; CO, about 1% or 150 g/kg fuel; and HC, 2000 ppm (as C₁) or 15 g/kg fuel.^[1] Catalytic converters in the exhaust system now reduce these engine-out emissions by over 90%. Piston blowby gases, and fuel evaporation and release to the atmosphere through vents in the fuel system, especially after engine shut-down, were historically additional sources of unburned hydrocarbons. However, in modern engines these nonexhaust sources are controlled by returning the blowby gases from the crank-case to the engine intake system and by venting the fuel tank through a vapor-absorbing carbon cannister which is purged with some of the engine intake air during normal engine operation.

In diesel engine exhaust, concentrations of NO_x are comparable to those from SI engines. Diesel hydrocarbon emissions are significant, though exhaust concentrations are much lower than typical SI engine levels. The hydrocarbons in the exhaust may also condense to form white smoke during engine starting and warm-up. Specific hydrocarbon compounds in the exhaust gases are the source of diesel odor. Diesel engines are an important source of particulate emissions; between about 0.2 and 0.5% of the fuel mass is emitted as small (10 to 500 nm diameter) particles, which consist primarily of soot with additional hydrocarbon material. Diesel engines are not a significant source of carbon monoxide.

Use of alcohol fuels in either of these engines substantially increases aldehyde emissions. Aldehydes become a significant contributor to photochemical smog when these oxygenates are blended with gasoline (e.g., 10% ethanol with gasoline, E10). The dominant fuels, gasoline and diesel, contain sulfur, the amount varying depending on world region. The sulfur is oxidized (i.e., burned) to produce sulfur dioxide, SO₂, of which a fraction can be oxidized to sulfur trioxide, SO₃. The latter combines with water to form a sulfuric acid aerosol. Sulfur in these fuels, over time, degrades the effectiveness of the exhaust catalysts now widely used to substantially reduce engine-out emissions. In many parts of the world, the sulfur content of transportation fuels is tightly regulated, and "sulfur-free fuels" with <10 ppm (by weight) sulfur are required in the major developed-world regions. Fuel quality varies widely in other parts of the world and, historically, levels have been much higher (of order 600 ppm in gasoline, up to 0.5% in diesel).

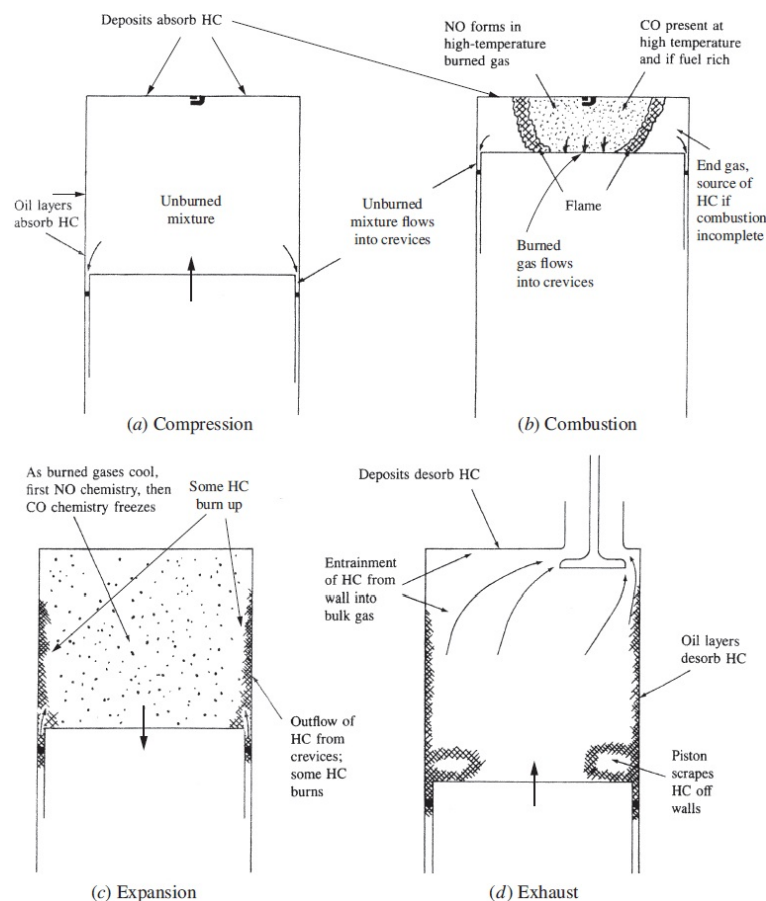
In general, the concentrations of these pollutants in internal combustion engine exhaust gases differ from values calculated assuming chemical equilibrium. Thus the detailed chemical mechanisms by which these pollutants form and the kinetics of these processes are important in determining emission levels. For some pollutant species, for example, carbon monoxide, organic compounds, and particulates, the formation and destruction reactions are intimately coupled with the primary fuel combustion process. Thus an understanding of the formation of these species requires knowledge of the combustion chemistry. For nitrogen oxides and sulfur oxides, the formation and destruction processes are not part of the fuel combustion process. However, the reactions, which produce these species, take place in an environment created by the combustion reactions, so the two processes are still linked. A summary of the mechanisms by which these pollutants form in internal combustion engines provides an introduction to this chapter. In subsequent sections, the details of the basic formation mechanisms of each pollutant and the application of these mechanisms to the combustion process in both spark-ignition and compression-ignition engines will be developed.

As noted above, catalytic converters and particulate traps are now built into engine exhaust systems to *clean up* the exhaust gases before they leave the tailpipe and enter the atmosphere. These exhaust treatment devices will be described in [Sec. 11.6](#). The efficiency of these exhaust treatment devices is very high: the percent of the engine-out emissions removed is usually over 90%. However, to be effective, the temperature of the catalysts used for HC, CO, and NO_x removal must be at about 250°C for the catalyst to be effective. Thus engine-out emissions levels during this catalyst warm-up period are especially important, as is achieving the catalyst *light-off* state rapidly. Once the catalyst or particulate trap has fully warmed-up, the engine's task is to feed exhaust gas of the appropriate composition (e.g., relative air/fuel or fuel/air ratio) and

temperature to realize very high efficiency and low vehicle emissions levels. This is the context for our review of the mechanisms within the engine by which these pollutants form.

The processes by which pollutants form within the cylinder of a conventional spark-ignition engine are illustrated in Fig. 11.1. The schematic shows the combustion chamber during four different phases of the engine operating cycle: compression, combustion, expansion, and exhaust. Nitric oxide (NO) forms in the high-temperature burned gases behind the flame through chemical reactions involving nitrogen and oxygen atoms and molecules: this nitrogen/oxygen system does not attain chemical equilibrium. The higher the burned gas temperature, the higher is the rate of formation of NO. As the burned gases cool during the expansion stroke the reactions involving NO freeze, and leave NO concentrations well in excess of levels corresponding to equilibrium at exhaust conditions. Carbon monoxide also forms during the combustion process. With rich fuel-air mixtures, there is insufficient oxygen to burn fully all the carbon in the fuel to CO_2 ; also in the high-temperature products, even with lean mixtures, dissociation produces significant CO levels. Later in the expansion stroke, the CO oxidation process also freezes as the burned gas temperature falls.

Figure 11.1 Summary of HC, CO, and NO pollutant formation mechanisms in a spark-ignition engine.



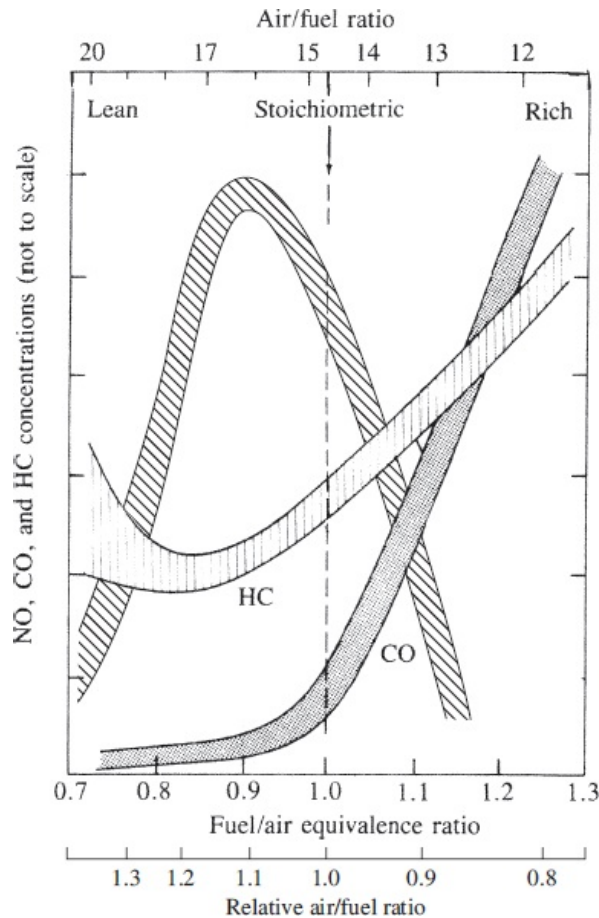
The hydrocarbon emissions have several different sources. During compression and combustion, the increasing cylinder pressure forces some of the gas in the cylinder into crevices—narrow volumes connected to the combustion chamber: the volumes between the piston, rings, and cylinder wall are the largest of these. Most of this gas in the crevices is unburned fuel-air mixture, which escapes the primary combustion process because the entrance to these crevices is too narrow for the flame to enter. This gas, which leaves these crevices during the expansion and exhaust processes, is the major source of hydrocarbon emissions. Another source is flame quenching on the combustion chamber walls. A quench layer containing unburned and partially burned fuel-air mixture is left at the wall when the flame, as it approaches the wall, extinguishes. However, the HC in these thin (≤ 0.1 mm) layers burn up rapidly after flame quenching so this is not a large source. It has been shown that the porous deposits on the walls of engines in actual operation do increase engine HC emissions, due to a flame quenching process. An additional source of hydrocarbons is believed to be the engine oil left in a thin film on the cylinder liner.

This oil layer absorbs and desorbs fuel hydrocarbon components, before and after combustion, respectively, thus permitting a fraction of the fuel to escape the primary combustion process unburned. A final source of HC in engines is incomplete combustion due to bulk quenching of the flame in that fraction of engine cycles where combustion is especially slow (see [Sec. 9.4.3](#)). Such conditions are most likely to occur during transient engine operation when the air/fuel ratio, spark timing, and the fraction of the exhaust recycled for emission control are not properly matched.

These as yet unburned hydrocarbons exit the cylinder by being entrained into the bulk-gas flow during blowdown, and at the end of the exhaust stroke as the piston pushes gas scraped off the wall out of the exhaust valve. Oxidation of the hydrocarbons which escape the primary combustion process by any of the above processes can occur during expansion and exhaust. The amount of oxidation depends on the temperature and oxygen-concentration time histories of these HC as they mix with the bulk gases.

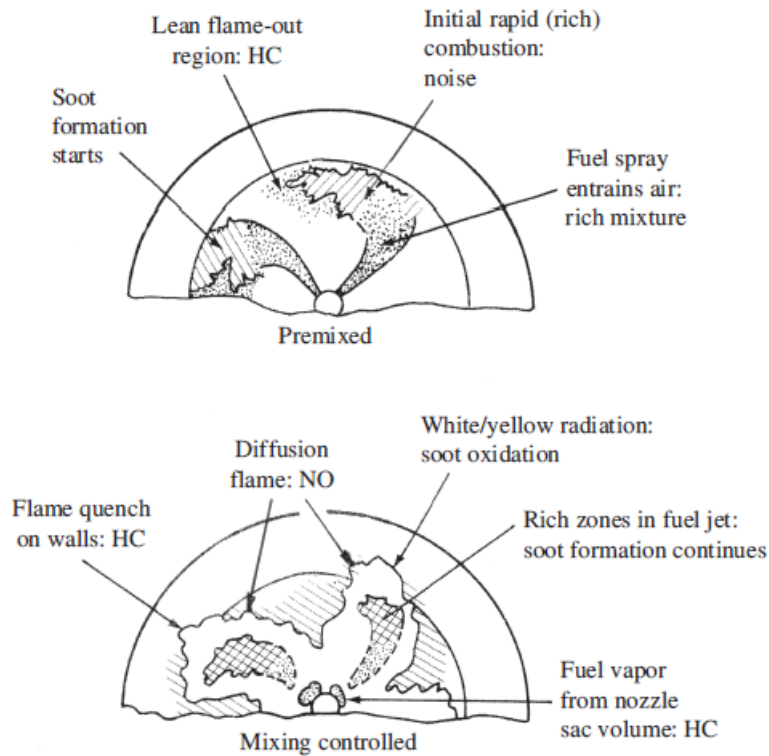
One of the most important variables in determining spark-ignition engine emissions is the relative air/fuel ratio, λ . [Figure 11.2](#) shows how NO, CO, and HC exhaust emissions vary with this parameter. The spark-ignition engine has historically been operated close to stoichiometric, or slightly fuel-rich, to ensure smooth and reliable operation. [Figure 11.2](#) shows that leaner mixtures give lower CO and HC emissions until the combustion quality becomes poor (and eventually misfire occurs), when HC emissions rise sharply and engine operation becomes erratic. However, NO emissions peak about 10% lean of stoichiometric. The shapes of these curves indicate the complexities of emission control. In a cold engine, when fuel vaporization is slow, the fuel flow may be increased to provide an easily combustible slightly-rich mixture in the cylinder. But then CO and HC emissions are higher. At part-load conditions, lean mixtures could be used which would produce lower HC and CO emissions (at least until the combustion quality deteriorates), but NO emissions would be high. Use of recycled exhaust to dilute the engine intake mixture lowers the NO levels, but also deteriorates combustion quality. Exhaust gas recirculation (EGR) is used with stoichiometric mixtures in many engine control systems. Note that the highest power levels are obtained from the engine with slightly rich-of-stoichiometric mixtures and no recycled exhaust to dilute the incoming charge. As we will see, several emission control techniques are required to reduce engine-out emissions of all three pollutants, over all engine operating modes.

Figure 11.2 Variation of HC, CO, and NO concentration in the exhaust of a conventional spark-ignition engine with relative air-fuel ratio and fuel/air equivalence ratio.



In the diesel engine, the fuel is injected into the cylinder just before combustion starts, so throughout most of the critical parts of the cycle the fuel distribution is nonuniform. The pollutant formation processes are strongly dependent on the fuel distribution and how that distribution changes with time due to mixing with hot air. [Figure 11.3](#) illustrates how various parts of the fuel spray and diffusion flame affect the formation of NO, unburned HC, and soot (or particulates) during the "premixed" and "mixing-controlled" phases of diesel combustion in a direct-injection (DI) engine with swirl. Nitric oxide forms in the high-temperature burned gas regions as before, but temperature and fuel/air ratio distributions within the burned gases are now nonuniform and formation rates are highest in the close-to-stoichiometric diffusion flame reaction zone. Soot forms in the rich unburned-fuel-containing core of fuel spray, after the fuel has vaporized through mixing with hot entrained air (and later with burned gas). Soot oxidizes in the diffusion region flame when it contacts oxygen, giving rise to the yellow luminous character of the flame. Hydrocarbons and aldehydes originate in regions where the flame quenches both on the walls and where excessive dilution with air prevents the combustion process from either starting or going to completion. Fuel that vaporizes from the nozzle sac volume during the later stages of combustion is also a source of HC. Combustion-generated noise is controlled by the early part of the premixed (rich) combustion process, the initial rapid heat release immediately following the ignition delay.

Figure 11.3 Summary of pollutant formation mechanisms in a direct-injection diesel engine during "premixed" and "mixing-controlled" combustion phases.



[1] Percentages or parts per million (ppm) are molar fractions times 100 or 10^6 , respectively. Emission Indexes are g pollutant per kg fuel.

11.2. NITROGEN OXIDES

11.2.1. Kinetics of NO Formation

While nitric oxide (NO) and nitrogen dioxide (NO₂) are usually grouped together as NO_x emissions, nitric oxide is the predominant oxide of nitrogen produced inside the engine cylinder. The principal source of NO is the oxidation of atmospheric (molecular) nitrogen. The mechanism of NO formation from atmospheric nitrogen has been studied extensively.¹ It is generally accepted that in combustion of near-stoichiometric fuel-air mixtures, the principal reactions governing the formation of NO from nitrogen (and its destruction) are^[1]



The forward and reverse rate constants (

k_i^+ and k_i^- , respectively) for these reactions have been measured in numerous experimental studies. Recommended values for these rate constants taken from a critical review of this published data are given in [Table 11.1](#). Note that the equilibrium constant for each reaction, $K_{c,i}$ (see [Sec. 3.7.2](#)), is related to the forward and reverse rate constants by $K_{c,i} = k_i^+ / k_i^-$. The rate of formation of NO via reactions (11.1) to (11.3) is given by [see Eqs. (3.48) and (3.51)]

$$\begin{aligned} \frac{d[\text{NO}]}{dt} = & k_1^+ [\text{N}_2] [\text{O}] + k_2^+ [\text{N}] [\text{O}_2] + k_3^+ [\text{N}] [\text{OH}] \\ & - k_1^- [\text{NO}] [\text{N}] - k_2^- [\text{NO}] [\text{O}] - k_3^- [\text{NO}] [\text{H}] \end{aligned} \quad (11.4)$$

Table 11.1 Rate constants for NO formation mechanism¹

Reaction	Rate constant, cm ³ /mol · s	Temperature range, K	Uncertainty, factor of or %
(1) N ₂ + O → NO + N	$7.6 \times 10^{13} \exp [-38,000/T]$	2000–5000	2
(-1) N + NO → N ₂ + O	1.6×10^{13}	300–5000	±20% at 300 K 2 at 2000–5000 K
(2) N + O ₂ → NO + O	$6.4 \times 10^9 T \exp [-3150/T]$	300–3000	±30% at 300–1500 K 2 at 3000 K
(-2) O + NO → O ₂ + N	$1.5 \times 10^9 T \exp [-19,500/T]$	1000–3000	±30% at 1000 K 2 at 3000 K
(3) H + OH → NO + H	4.1×10^{13}	300–2500	±80%
(-3) H + NO → OH + N	$2.0 \times 10^{14} \exp [-23,650/T]$	2200–4500	2

where [] denotes species concentrations in moles per cubic centimeter when k_i have the values given in Table 11.1. The forward rate constant for reaction (11.1) and the reverse rate constants for reactions (11.2) and (11.3) have large activation energies, which result in a strong temperature dependence of NO formation rates.

A similar relation to (11.4) can be written for $d[N]/dt$:

$$\begin{aligned} \frac{d[N]}{dt} = & k_1^+ [N_2] [O] - k_2^+ [N] [O_2] - k_3^+ [N] [OH] \\ & - k_1^- [NO] [N] - k_2^- [NO] [O] - k_3^- [NO] [H] \end{aligned}$$

(11.5)

Since [N] is much less than the concentrations of other species of interest ($\sim 10^{-8}$ mole fraction), the steady-state approximation^[2] is appropriate: $d[N]/dt$ is set equal to zero and Eq. (11.5) used to eliminate [N]. The NO formation rate then becomes

$$\frac{d[NO]}{dt} = 2k_1^+ [N_2] [O] \frac{1 - [NO]^2 / (K [O_2] [N_2])}{1 + k_1^- [NO] / (k_2^+ [O_2] + k_3^+ [OH])}$$

(11.6)

where

$$K = (k_1^+ / k_1^-) (k_2^+ / k_2^-).$$

NO forms in both the flame front and the post-flame gases. In engines, however, combustion occurs at high pressure so the flame reaction zone is extremely thin (~ 0.1 mm) and residence time within this zone is short. Also, the cylinder pressure rises during most of the combustion process, so burned gases produced early in the combustion process are compressed to a higher temperature than they reached immediately after combustion. Thus, NO formation in the post-flame gases almost always dominates any flame-front-produced NO. It is appropriate, therefore to assume that the combustion chemistry and NO

formation processes are decoupled and to approximate the concentrations of O, O₂, OH, H, and N₂ by their equilibrium values at the local pressure and equilibrium temperature.

To introduce this equilibrium assumption it is convenient to use the notation

$R_1 = k_1[N_2]_e[O]_e = k_1^-[NO]_e[N]_e$, where $[]_e$ denotes equilibrium concentration, for the one-way equilibrium rate for reaction (11.1), with similar definitions for

$R_2 = k_2^+[N]_e[O_2]_e = k_2^-[NO]_e[O]_e$ and

$R_3 = k_3^+[N]_e[OH]_e = k_3^-[NO]_e[H]_e$. Substituting $[O]_e$, $[O_2]_e$, $[OH]_e$, $[H]_e$, and $[N_2]_e$ for $[O]$, $[O_2]$, $[OH]$, $[H]$, and $[N_2]$ in Eq. (11.6) yields

$$\frac{d[NO]}{dt} = \frac{2R_1 \{1 - ([NO]/[NO]_e)^2\}}{1 + ([NO]/[NO]_e) R_1 / (R_2 + R_3)}$$

(11.7)

Typical values of R_1 , R_1/R_2 , and $R_1/(R_2 + R_3)$ are given in Table 11.2. The difference between R_1/R_2 and $R_1/(R_2 + R_3)$ indicates the relative importance of adding reaction (11.3) to the mechanism.

Table 11.2 Typical values of R_1 , R_1/R_2 , and $R_1/(R_2 + R_3)$ *

Equivalence ratio	$R_1^†$	R_1/R_2	$R_1/(R_2 + R_3)$
0.8	5.8×10^{-5}	1.2	0.33
1.0	2.8×10^{-5}	2.5	0.26
1.2	7.6×10^{-6}	9.1	0.14
*At 10 atm pressure and 2600 K.			
†Units gmol/cm ³ · s.			

The strong temperature dependence of the NO formation rate can be demonstrated by considering the initial value of $d[NO]/dt$ when $[NO]/[NO]_e \ll 1$. Then, from Eq. (11.7),

$$\frac{d[NO]}{dt} = 2R_1 = 2k_1^+[N_2]_e[O]_e$$

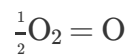
(11.8)

The equilibrium oxygen atom concentration is given by

$$[O]_e = \frac{K_{p(O)}[O_2]_e^{1/2}}{(\tilde{R}T)^{1/2}}$$

(11.9)

where $K_{p(O)}$ is the equilibrium constant for the reaction



and is given by

$$K_{p(\text{O})} = 3.6 \times 10^3 \exp\left(\frac{-31,090}{T}\right) \text{ bar}^{1/2}$$

(11.10)

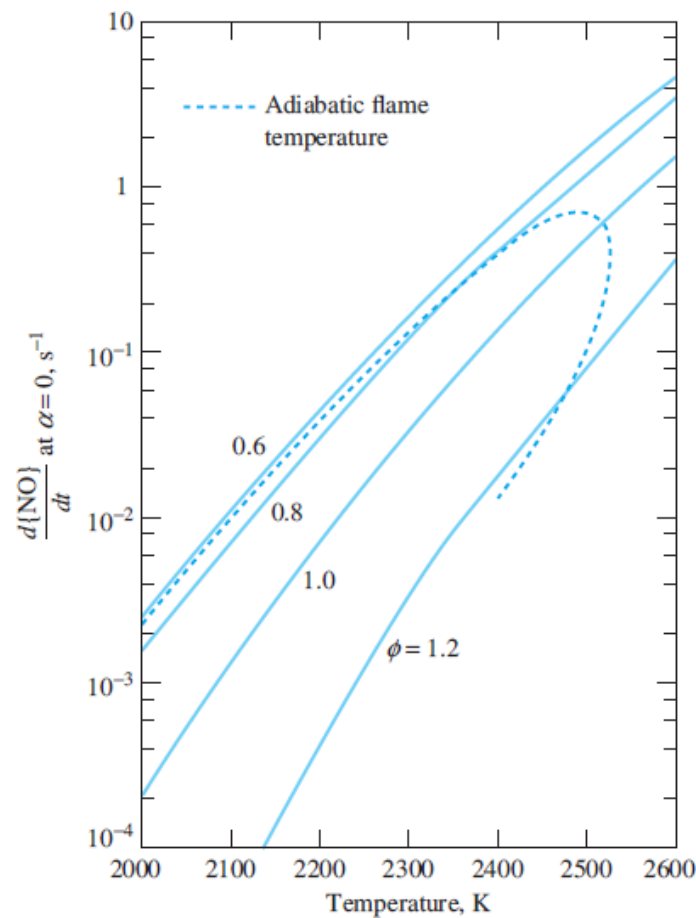
The initial NO formation rate may then be written [combining Eqs. (11.8), (11.9), and (11.10) with k_1^+ from Table 11.1] as

$$\frac{d[\text{NO}]}{dt} = \frac{6 \times 10^6}{T^{1/2}} \exp\left(\frac{-69,090}{T}\right) [\text{O}_2]_e^{1/2} [\text{N}_2]_e \text{ mol/cm}^3 \cdot \text{s}$$

(11.11)

The strong dependence of $d[\text{NO}]/dt$ on temperature in the exponential term is evident. High temperatures and high oxygen concentrations result in high NO formation rates. Figure 11.4 shows the NO formation rate as a function of gas temperature and fuel/air equivalence ratio in post-flame gases. Also shown is the adiabatic flame temperature attained by a fuel-air mixture initially at 700 K at a constant pressure of 15 bar. For adiabatic constant-pressure combustion (an appropriate model for each element of fuel that burns in an engine), this initial NO formation rate peaks about 10% lean of the stoichiometric composition, and decreases rapidly as the mixture then becomes leaner or richer.

Figure 11.4 Initial NO formation rate, mass fraction per second (for $[NO]/[NO]_e \ll 1$) as a function of temperature for different equivalence ratios (ϕ) and 15 bar pressure. Dashed line shows adiabatic flame temperature for kerosene combustion with 700 K, 15 bar air.³



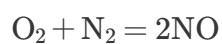
A characteristic time for the NO formation process, τ_{NO} , can be defined by

$$\tau_{NO}^{-1} = \frac{1}{[NO]_e} \frac{d[NO]}{dt} \quad (11.12)$$

$[NO]_e$ can be obtained from the equilibrium constant

$$K_{NO} = 20.3 \times \exp(-21,650/T)$$

for the reaction



as $[NO]_e = (K_{NO}[O_2]_e[N_2]_e)^{1/2}$. Equations (11.11) and (11.12) can be combined to give

$$\tau_{\text{NO}} = \frac{8 \times 10^{-16} T \exp(58,000/T)}{p^{1/2}}$$

(11.13)

where

τ_{NO} is in seconds, T in kelvins, and p in bar. Use has been made of the fact that

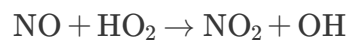
$\tilde{x}_{\text{N}_2} \approx 0.71$. For engine combustion conditions,

τ_{NO} is usually comparable to or longer than the times characteristic of changes in engine conditions so the formation process is kinetically controlled. However, for close-to-stoichiometric conditions at the maximum pressures and burned gas temperatures,

τ_{NO} is of the same order as typical combustion times (1 ms) and equilibrium NO concentrations may be attained. There is good evidence that this NO formation model is valid under conditions typical of those found in engines.⁴

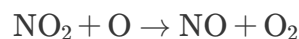
11.2.2. Formation of NO₂

Chemical equilibrium considerations indicate that for burned gases at typical flame temperatures, NO₂/NO ratios should be negligibly small. While experimental data show this is true for spark-ignition engines, in diesels NO₂ can be 10 to 30% of the total oxides of nitrogen, engine-out, emissions.⁵ A plausible mechanism for the persistence of NO₂ is the following.⁶ NO formed in the flame zone can be rapidly converted to NO₂ via reactions such as



(11.14)

Subsequently, conversion of this NO₂ to NO occurs via

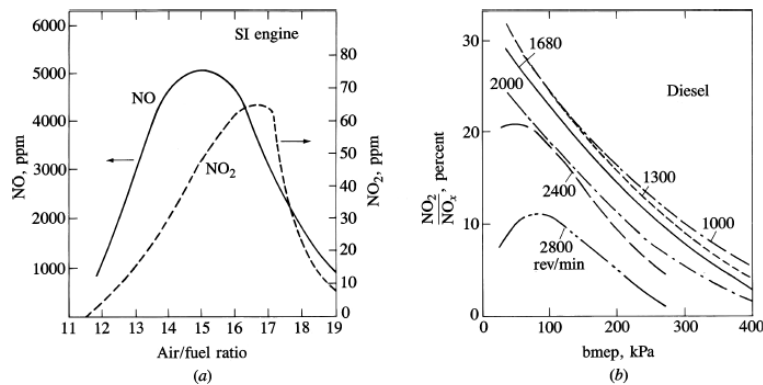


(11.15)

unless the NO₂ formed in the flame is quenched by mixing with cooler fluid. This explanation is consistent with the highest NO₂/NO ratio occurring at light load in diesels, when cooler regions which could quench the conversion back to NO are widespread.⁵

Figure 11.5 shows examples of NO and NO₂ emissions data from a spark-ignition and a diesel engine. The maximum value for the ratio (NO₂/NO) for the SI engine is 2 percent, at a relative air-fuel ratio of about 1.2. For the diesel this ratio is higher, and is highest at light load, when the engine is operating very lean; it also depends on engine speed.

Figure 11.5 (a) NO and NO₂ concentrations in SI engine-out exhaust as function of air/fuel ratio, 1500 rev/min, wide-open throttle; (b) NO₂ as percent of total NO_x in diesel exhaust as function of load and speed.⁵

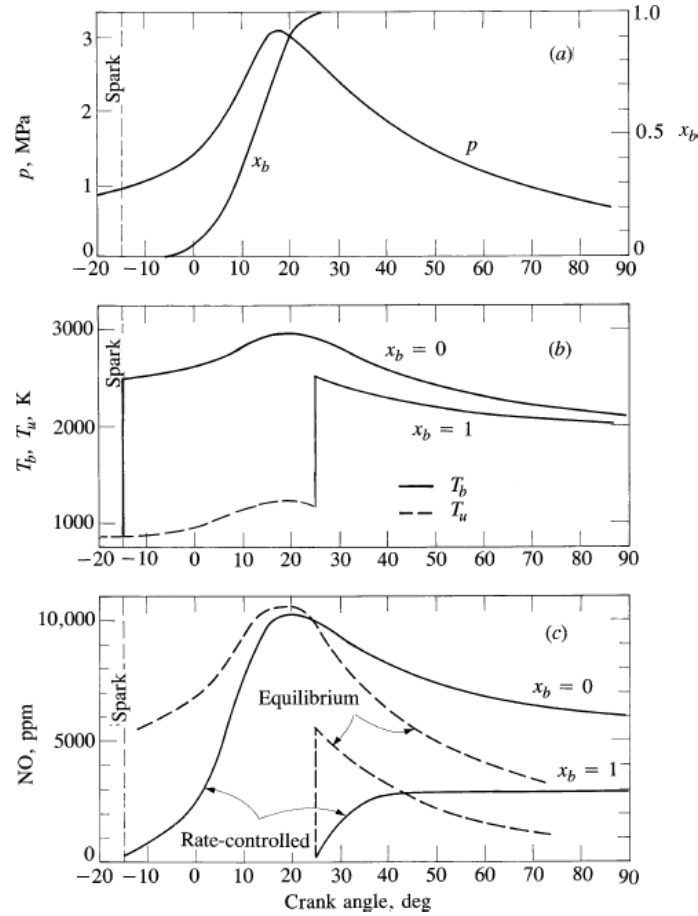


It is customary to measure total oxides of nitrogen emissions, NO plus NO₂, with a chemiluminescence analyzer and call the combination NO_x. It is always important to check carefully whether specific emissions data for NO_x are given in terms of mass of NO or mass of NO₂, which have molecular weights of 30 and 46, respectively.

11.2.3. NO Formation in Spark-Ignition Engines

In conventional spark-ignition engines the fuel and air (and any recycled exhaust) are mixed together during the engine intake process. Vigorous mixing with the residual gas within the cylinder also occurs during the intake. Thus the fuel/air ratio and the amount of diluent (residual gas plus any recycled exhaust) are close to uniform throughout the charge within the cylinder during combustion.^[3] Since the composition is essentially uniform, the nature of the NO formation process within the cylinder can be understood by coupling the kinetic mechanism developed in [Sec. 11.2.1](#) with the burned gas temperature distribution and pressure in the cylinder during the combustion and expansion processes. The temperature distribution which develops in the burned gases due to the passage of the flame across the combustion chamber has been discussed in [Sec. 9.2.1](#). Mixture which burns early is compressed to higher temperatures after combustion, as the cylinder pressure continues to rise; mixture which burns later is compressed primarily as unburned mixture and ends up after combustion at a lower burned gas temperature. [Figures 11.6a and b](#) show measured cylinder pressure data from an operating engine, with estimates of the mass fraction burned (x_b) and the temperatures of a gas element which burned just after spark discharge and a gas element which burned at the end of the burning process. The model used to estimate these temperatures assumed no mixing between mixture elements which burn at different times. This assumption is more realistic than the alternative idealization that the burned gases mix rapidly and are thus uniform (see [Sec. 9.2.1](#)). Using the NO formation kinetic model [Eq. (11.7)] to calculate NO concentrations in these burned gas elements, with the equilibrium concentrations of the species O, O₂, N₂, OH, and H corresponding to the average fuel/air equivalence ratio and burned gas fraction of the mixture, and these pressure and temperature profiles, the rate-limited concentration profiles in [Fig. 11.6c](#) are obtained. Also shown are the NO concentrations that would correspond to chemical equilibrium at these conditions. The rate-controlled concentrations rise from the residual gas NO concentration, lagging the equilibrium levels, then cross the equilibrium levels, and "freeze" well above the equilibrium values corresponding to exhaust conditions. Depending on details of engine operating conditions, the rate-limited concentrations may or may not come close to equilibrium levels at peak cylinder pressure and gas temperature. Also, the amount of decomposition from peak NO levels which occurs during expansion depends on engine conditions as well as whether the mixture element burned early or late.⁷

Figure 11.6 Illustration of SI engine NO formation process: (a) measured cylinder pressure p and calculated mass fraction burned x_b ; (b) calculated temperature of unburned gas T_u and burned gas T_b in early- and late-burning elements; (c) calculated NO concentrations in early- and late-burning elements for rate-controlled formation model and at equilibrium.⁷



Once the NO chemistry has frozen during the early part of the expansion stroke, integration over all elements will give the final average NO concentration in the cylinder which equals the exhaust concentration. Thus, if $\{NO\}$ is the local mass fraction of NO, then the average exhaust NO mass fraction is given by

$$\{\overline{NO}\} = \int_0^1 \{NO\}_f dx_b \quad (11.16)$$

where $\{NO\}_f$ is the final frozen NO mass fraction in the element of charge which burned when the mass fraction burned was x_b . Note that $\{NO\} = [NO]M_{NO}/p$, where $M_{NO} = 30$, the molecular weight of NO. The average exhaust concentration of NO as a mole fraction is given by

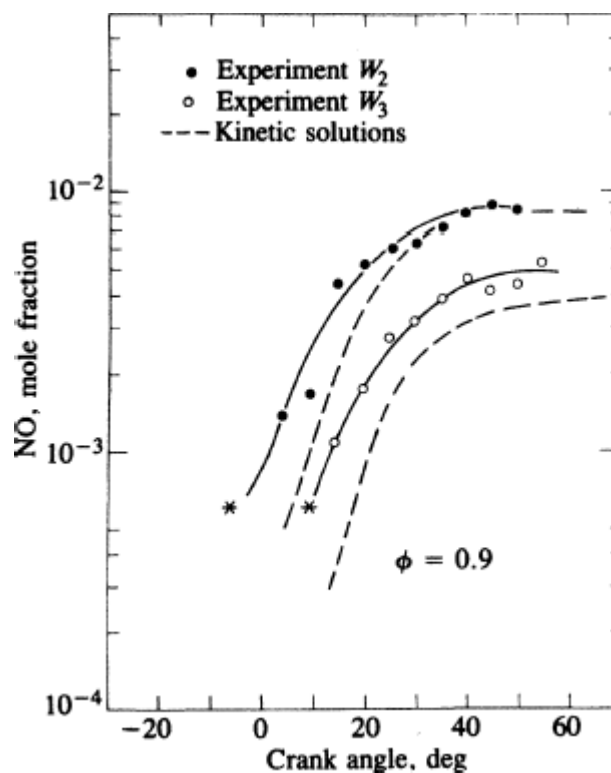
$$\tilde{x}_{NO_{av}} = \{\overline{NO}\} \frac{M_{exh}}{M_{NO}} \quad (11.17)$$

and the exhaust concentration in ppm is

$\tilde{x}_{\text{NO}_{\text{av}}} \times 10^6$. The earlier burning fractions of the charge contribute much more to the exhausted NO than do later burning fractions of the charge: frozen NO concentrations in these early-burning elements can be an order of magnitude higher than concentrations in late-burning elements. In the absence of vigorous bulk gas motion, the highest NO concentrations occur nearest the spark plug.

Substantial experimental evidence supports this description of NO formation in spark-ignition engines. The NO concentration gradient across the burned gas in the engine cylinder, due to the temperature gradient, has been demonstrated using gas sampling techniques^{8,9} and using measurements of the chemiluminescent radiation from the reaction $\text{NO} + \text{O} \rightarrow \text{NO}_2 + h\nu$ to determine the local NO concentration. Figure 11.7 shows NO concentration data as a function of crank angle, taken by Lavoie¹⁰ through two different windows in the cylinder head of a specially constructed L-head engine where each window was a different distance from the spark plug. The stars indicate the estimated initial NO concentration that results from mixing of the residual gas with the fresh charge, at the time of arrival of the flame at each window. The observed NO mole fractions rise smoothly from these initial values and then freeze about one-third of the way through the expansion process. NO levels observed at window W_2 , closest to the spark plug, were substantially higher than those observed at window W_3 . The dashed lines show calculated NO concentrations obtained using the NO formation kinetic model with an "unmixed" thermodynamic analysis for elements that burned at the time of flame arrival at each window. Since the calculated values started from zero NO concentration at the flame front (and not the diluted residual gas NO level indicated by the star), the calculations initially fall below the data. However, the difference between the two measurement locations and the frozen levels are predicted with reasonable accuracy. Thus, the rate-limited formation process, freezing of NO chemistry during expansion, and the existence of NO concentration gradients across the combustion chamber have all been observed.

Figure 11.7 Spectroscopically measured NO concentrations through two windows W_3 and W_2 in special L-head SI engine (W_2 is closer to spark than W_3). The asterisks mark estimated initial conditions and flame arrival times. The dashed lines are calculated rate-limited concentrations for portions of the charge burning at these flame arrival times, with zero initial NO concentration.¹⁰



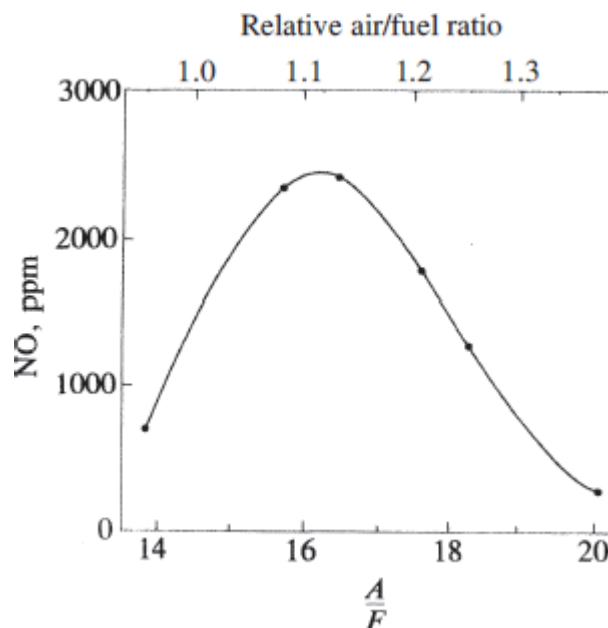
The most important engine variables that affect NO emissions are the relative air/fuel ratio (or fuel/air equivalence ratio), the burned gas fraction of the *in-cylinder* unburned mixture, and spark timing. The burned gas fraction depends on the amount of diluent such as recycled exhaust gas (EGR) used for NO_x emissions control, as well as the residual gas fraction. Fuel

properties will also affect burned gas conditions; the effect of normal variations in gasoline properties is modest, however. The effect of variations in these parameters can be explained with the NO formation mechanism described above: changes in the time history of temperature and oxygen concentration in the burned gases during the combustion process and early part of the expansion stroke are the important factors.¹¹

11.2.3.1. Equivalence Ratio

Figure 11.8 shows the effect of variations in the fuel/air ratio on NO emissions. Maximum burned gas temperatures occur at $\lambda \approx 0.9$; however, at this relative air/fuel ratio oxygen concentrations are low. As the mixture is further enriched, burned gas temperatures fall. As the mixture is leaned out, increasing oxygen concentration initially offsets the falling gas temperatures and NO emissions peak at $\lambda \approx 1.1$. Detailed predictions of NO concentrations in the burned gases suggest that the concentration versus time histories under fuel-lean conditions are different in character from those for fuel-rich conditions. In lean mixtures, NO concentrations freeze early in the expansion process and little NO decomposition occurs. In rich mixtures, substantial NO decomposition occurs from the peak concentrations present when the cylinder pressure is a maximum. Thus in lean mixtures, gas conditions at the time of peak pressure are especially significant.⁷

Figure 11.8 Variation of exhaust (engine-out) NO concentration with A/F and relative air/fuel ratio. Spark-ignition engine, 1600 rev/min, $\eta_v = 50\%$, MBT timing.¹²



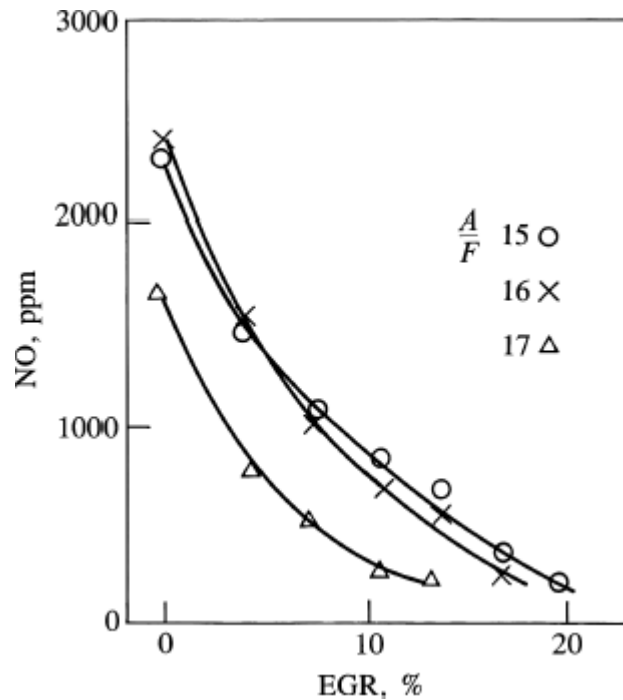
11.2.3.2. Burned Gas Fraction

The unburned mixture in the cylinder contains fuel vapor, air, and burned gases. The burned gases are residual gas from the previous cycle and any exhaust gas recycled to the intake for NO_x emissions control. The residual gas fraction is influenced by load, valve timing (especially the extent of valve overlap), and, to a lesser degree, by speed, air/fuel ratio, and compression ratio as described in Sec. 6.4. The burned gases act as a diluent in the unburned mixture; the absolute temperature reached after combustion varies inversely with the burned gas mass fraction. Hence, increasing the burned gas fraction reduces NO emissions levels. However, it also reduces the combustion rate and, therefore, makes stable combustion more difficult to achieve (see Sec. 9.4).

Figure 11.9 shows the effect of increasing the burned gas fraction by recycling exhaust gases to the intake system just below the throttle plate. Substantial reductions in NO concentrations are achieved with 10 to 15% EGR, which is about the maximum

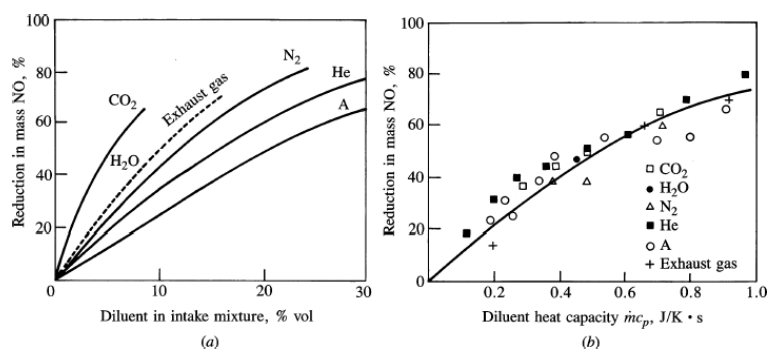
amount of EGR the engine will tolerate under normal part-throttle conditions. Of course, increasing the EGR at fixed engine load and speed increases the inlet manifold pressure, while fuel flow and airflow remain approximately constant.

Figure 11.9 Variation of exhaust (engine out) NO concentration with percent recycled exhaust gas (EGR). Spark-ignition engine, 1600 rev/min, $\eta_v = 50\%$, MBT timing.¹²



The primary effect of the burned gas diluent in the unburned mixture on the NO formation process is that it reduces flame temperatures by increasing the heat capacity of the cylinder charge, per unit mass of fuel. **Figure 11.10** shows the effect of different diluent gases added to the engine intake flow, in a single-cylinder engine operated at constant speed, fuel flow, and air flow.¹³ The data in **Fig. 11.10a** show that equal volume percentages of the different diluents produce different reductions in NO emissions. The same data when plotted against diluent heat capacity (diluent mass flow rate $\times$ specific heat, c_p) collapse to close to a single curve.

Figure 11.10 (a) Percentage reduction in mass NO emissions with various diluents. (b) Correlation of NO reduction with diluent heat capacity. Spark-ignition engine operated at 1600 rev/min, constant brake load (intake pressure ~ 0.5 atm), with MBT spark timing.¹³



A similar study where the burned gas fraction in the unburned charge was varied by changing the valve overlap, compression ratio, and EGR, separately, showed that, under more realistic engine operating conditions, it is the heat capacity of the total diluent mass in the in-cylinder mixture that is important. Whether the diluent mass is changed by varying the valve overlap,

EGR, or even the compression ratio, was not important.¹⁴ Application of the ideal gas law to the in-cylinder burned gases, first without EGR and then with EGR, at the same crank angle, yields a pair of equations

$$pV = M_{\text{noEGR}}RT_{\text{noEGR}}; \quad pV = M_{\text{EGR}}RT_{\text{EGR}}$$

(11.18)

Together, these illustrate the impact of adding EGR at constant engine torque. At the same crank angle (say close to the end of combustion), the volumes will be the same and the pressures will be closely comparable for the same torque. The gas constant will be little affected by the EGR since the burned gas molecular weight changes are small. With 10% EGR, the in-cylinder mass will be some 10% higher since, at constant torque, the fuel plus air mass will be little different. Thus, temperatures will be about 10% lower which is a significant reduction (see Fig. 11.4). Note that with exhaust gas as the diluent, the impact on NO is through the temperature reduction resulting from the increased in-cylinder mass with EGR, for essentially the same fuel mass. The change in specific heat per unit mass is modest.

11.2.3.3. Excess Air and EGR

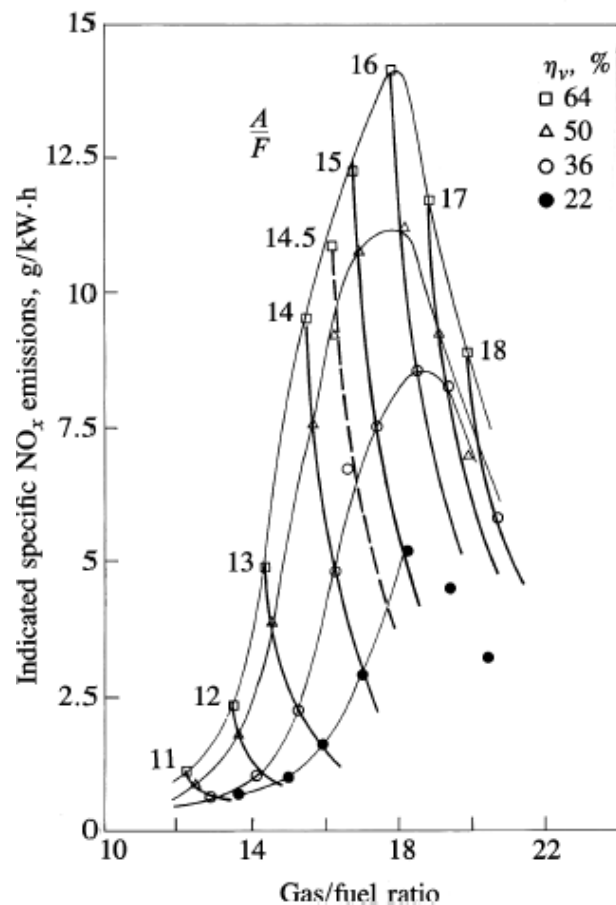
Because of the above, it is possible to correlate the influence of engine operating variables (such as air/fuel ratio, EGR, and engine speed and load) on NO emissions with two parameters which define the in-cylinder mixture composition: the relative air/fuel ratio (or the fuel/air equivalence ratio) and the gas/fuel ratio. The gas/fuel ratio (G/F) is calculated from the air plus residual plus EGR mass by

$$\frac{G}{F} = \frac{\text{total mass in cylinder}}{\text{fuel mass in cylinder}} = \frac{A}{F} \left(1 + \frac{x_b}{1 - x_b} \right)$$

(11.19)

where x_b is the burned gas fraction [Eq. (4.3)]. These together define the relative proportions of fuel, air, and burned gases in the in-cylinder mixture, and hence will correlate NO emissions.^[4] Figure 11.11 shows the correlation of specific NO emissions, from a four-cylinder SI engine, over a range of operating conditions as a function of air/fuel ratio and gas/fuel ratio. Lines of constant air/fuel ratio and volumetric efficiency are shown; the direction of increasing dilution with residual gas and EGR at constant air/fuel ratio is downward and to the right. Excessive dilution results in poor combustion quality, partial burning, and, eventually, misfire (see Sec. 9.4.3). Lowest NO emissions consistent with good fuel consumption (avoiding the use of rich mixtures) are obtained with a stoichiometric mixture, with as much dilution as the engine will tolerate without excessive deterioration in combustion quality.¹⁵ Note that increasing mixture dilution (through higher residual or EGR) at the stoichiometric A/F decreases NO emissions faster than does leaner operation (increasing A/F above the stoichiometric value of about 14.6) at constant residual plus EGR. This is largely due to the fact that when lean, a portion of the residual and EGR can be thought of as "air" since the burned gases contain oxygen.

Figure 11.11 Correlation between gas/fuel ratio (G/F) and indicated specific NO_x emissions at various air/fuel ratios (A/F) and volumetric efficiencies (η_v). Spark-ignition engine operated at 1400 rev/min with spark timing retarded to give 0.95 of maximum brake torque.¹⁵



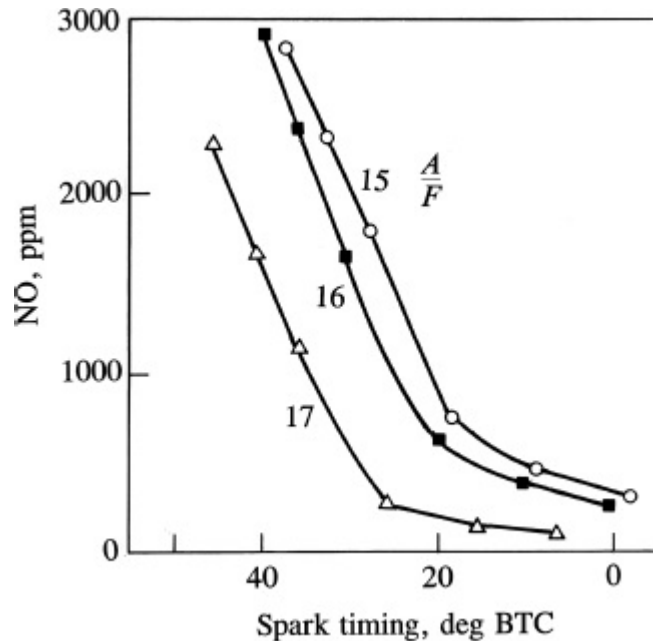
Comparisons between predictions made with the NO formation model (described at the beginning of this section) and experimental data show good agreement with normal amounts of dilution.¹⁶ With extreme dilution, at engine-out NO levels of 100 ppm or less, the NO formed within the flame reaction zone cannot always be neglected. Within the flame, the concentrations of radicals such as O, OH, and H can be substantially in excess of equilibrium levels, resulting in much higher formation rates within the flame than in the post-flame gases. It is believed that the mechanism [reactions (11.1) to (11.3)] and the formation rate equation (11.6) are still valid. However, neglecting the faster flame-front-formation rate of NO due to high concentrations of O, OH, and H may no longer be an appropriate assumption.¹⁷

11.2.3.4. Spark Timing

Spark timing significantly affects NO emission levels. Advancing the timing so that combustion occurs earlier in the cycle increases the peak cylinder pressure (because more fuel is burned before TC and the peak pressure moves closer to TC where the cylinder volume is smaller); retarding the timing decreases the peak cylinder pressure (because more of the fuel burns after TC). Higher peak cylinder pressures result in higher peak burned gas temperatures, and hence higher NO formation rates. For lower peak cylinder pressures, lower NO formation rates result. Figure 11.12 shows typical NO engine-out emission data for a spark-ignition engine as a function of spark timing. NO emission levels steadily decrease as spark timing is retarded from maximum brake torque (MBT) timing and moved closer to TC. Since exact determination of MBT timing is difficult (and not critical for fuel consumption and torque where the variation with timing around MBT is modest), there is always some uncertainty in NO emissions at MBT timing. Often, therefore, an alternative reference timing is used, where spark is retarded from MBT timing to the point where torque is decreased by 1 or 2% from the maximum value (usually about 5 degrees retard).

Great care with spark timing is necessary to obtain accurate NO emissions measurements under MBT-timing operating conditions.

Figure 11.12 Variation of exhaust (engine-out) NO concentration with spark retard. 1600 rev/min, $\eta_v = 50\%$; left-hand end of curve corresponds to MBT timing for each A/F.¹²



11.2.4. NO_x Formation in Compression-Ignition Engines

The kinetic mechanisms for NO and NO₂ formation described in Secs. 11.2.1 and 11.2.2 and the assumptions made regarding equilibration of species in the C-O-H system apply to diesels as well as to spark-ignition engines. The critical difference, of course, is that injection of fuel into the cylinder occurs just before combustion starts, and that nonuniform burned gas temperature and composition result from this nonuniform fuel distribution during the diesel combustion process. The fuel-air mixing and combustion processes are extremely complex. During the "premixed" combustion phase immediately following the ignition delay, rich fuel-air mixture with a fuel-air equivalence ratio between 2 and 4 reacts due to spontaneous ignition within each diesel fuel spray. During the mixing-controlled combustion phase, the burning mixture is close to stoichiometric (the flame structure is that of a turbulent diffusion flame). However, throughout the combustion process mixing between fuel vapor, air, and lean, stoichiometric, and rich burned mixture is occurring, thus changing the composition of gas elements in the cylinder. In addition to these composition (and hence temperature) changes due to mixing, temperature changes due to compression and expansion occur as the cylinder pressure rises and falls.

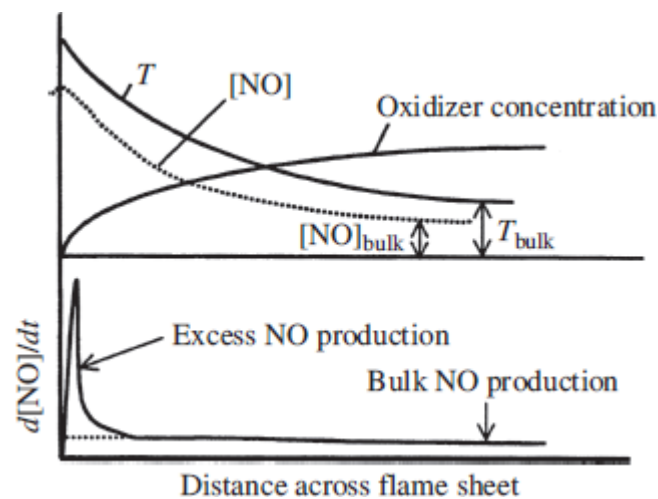
The discussion in Sec. 11.2.1 showed that the critical equivalence ratio for NO formation in high-temperature high-pressure burned gases is close to stoichiometric. Figure 11.4 is relevant here: it shows the initial NO formation rate in combustion products formed by burning a mixture of a typical hydrocarbon fuel with air (initially at 700 K, at a constant pressure of 15 bar). NO formation rates are high, and within a factor of 2 of the maximum value for $0.85 \leq \phi \leq 1.1$. Thus little NO will form during the fuel-rich premixed burning phase; almost all the NO will form in the mixing-controlled diffusion-flame burning process. In this diffusion flame, which surrounds each fuel spray, the fuel-rich partially reacted mixture diffuses into the flame from inside the spray and air diffuses into the flame from the outside. Chemical reaction is centered where the mixture composition is close to stoichiometric, producing the highest NO formation rates. These burned gases then diffuse away from the flame front, both inward into the fuel-rich spray, and outward into the surrounding air (see Fig. 10.5).

The critical time period is when burned gas temperatures are at a maximum. This will be close to the time of peak cylinder

pressure. As combustion proceeds and cylinder pressure increases, unburned mixture temperatures will rise. After the time of peak pressure, burned gas temperatures will decrease as the cylinder gases expand. The decreasing temperature due to expansion *and* due to mixing of the high-temperature gases with air slows down and then freezes the NO chemistry. This second effect (which occurs only in the diesel) means that freezing occurs more rapidly than in the spark-ignition engine, and less decomposition of the NO occurs.

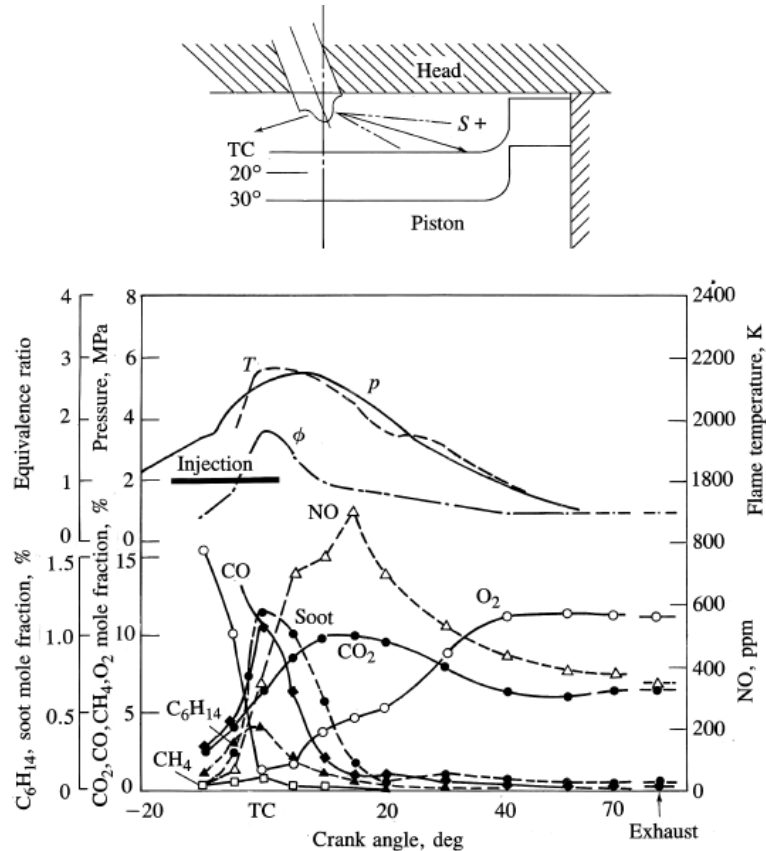
Figure 11.13 illustrates the above discussion of the diesel's NO formation process. It shows profiles of NO concentration, gas temperature, oxidizer concentration, and NO formation rate, with distance, moving outward from the center of the diffusion flame surrounding each diesel fuel spray into the outside air. The peak NO formation rate occurs slightly lean of stoichiometric, (see Fig. 11.4), and dominates the production process.¹⁸

Figure 11.13 Schematic of NO formation in the diffusion flame-sheet surrounding each diesel fuel spray, with temperature, oxygen, and NO concentration profiles. Rich side of flame sheet on left; bulk air in cylinder on right.¹⁸



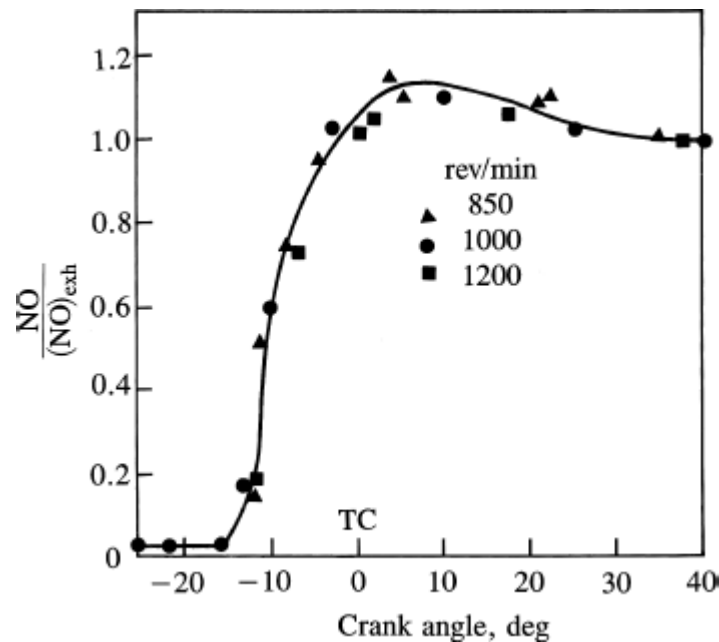
The above description is supported by species concentration data obtained from experiments where gas was sampled from within the cylinder of a normally operating diesel engine and analyzed. Figure 11.14 shows time histories of the major species concentrations, through the combustion process, determined with a rapid-acting sampling valve (1 ms open time) in a quiescent *direct-injection diesel* engine. Concentrations at different positions in the combustion chamber were obtained; the sample valve location for the Fig. 11.14 data is shown. Local NO concentrations rise from the residual gas value following start of combustion, to a peak at the point where the local burned gas equivalence ratio transitions from rich to lean (where the CO₂ concentration has its maximum value). As the local burned gas equivalence ratio becomes leaner due to mixing with excess air, NO concentrations decrease primarily due to dilution with air since formation becomes much slower as dilution occurs. At the time of peak NO concentrations within the bowl (15° ATC), most of the bowl region temperatures fall.

Figure 11.14 Concentrations of soot, NO, and other combustion product species measured at outer edge of the bowl-in-piston combustion chamber (location S) of a quiescent DI diesel with a rapid sampling valve. Cylinder gas pressure p , mean temperature T , and local equivalence ratio ϕ are also shown. Bore = 95 mm, stroke = 110 mm, $r_c = 14.6$. Four-hole nozzle with hole diameter = 0.2 mm.²⁰



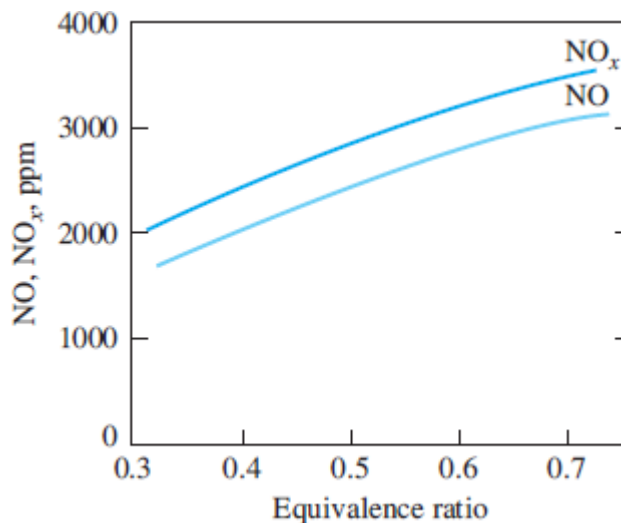
The total amount of NO within the cylinder of this type of direct-injection diesel during the NO formation process has also been measured.¹⁹ At a predetermined time in one cycle, once steady-state warmed-up engine operation had been achieved, the contents of the cylinder were dumped into an evacuated tank by rapidly cutting open a diaphragm which had previously sealed off the tank system. Figure 11.15 shows how the ratio of the average cylinder NO concentration divided by the exhaust concentration varies during the combustion process. NO concentrations reach a maximum shortly after time of peak pressure. There is a modest amount of NO decomposition. Variations in engine speed have little effect on the shape of this curve. The 20 crank-angle degrees after the start of combustion is the critical time period.

Figure 11.15 Ratio of cylinder-average NO concentration at given crank angle (determined from cylinder-dumping experiments) to exhaust NO concentration. DI diesel, equivalence ratio = 0.6, injection timing at 27° BTC.¹⁹



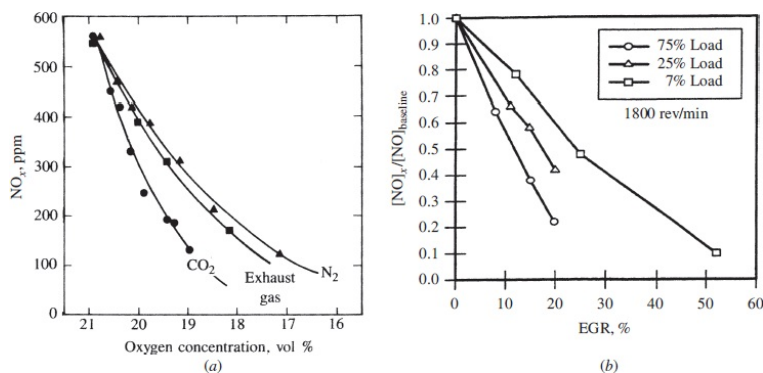
Results from similar cylinder-dumping experiments where injection timing and load (defined by the overall equivalence ratio) were varied also showed that almost all of the NO forms within 20° following the start of combustion. As injection timing is retarded, so the combustion process is retarded: NO formation occurs later, and concentrations are lower since peak temperatures are lower. The effect of the overall equivalence ratio, or load, on NO_x concentrations is shown in Fig. 11.16. At high load, with higher peak pressures, and hence temperatures, and larger area and longer lasting regions of close-to-stoichiometric diffusion-flame-produced burned gas, NO levels increase. Both NO and NO_x concentrations were measured; NO₂ is 10 to 20% of total NO_x (see also Fig. 11.5b). Though NO levels decrease with a decreasing overall equivalence ratio, they do so much less rapidly than do spark-ignition engine NO emissions (see Fig. 11.8) since, in the diesel, the zones where NO formation rates are highest are stoichiometric. Though the amount of fuel injected decreases proportionally as the overall equivalence ratio is decreased, almost all the fuel still burns close to stoichiometric. Thus NO emissions should be roughly proportional to the mass of fuel injected, provided burned gas pressures and temperatures in the diffusion flame do not change greatly.

Figure 11.16 Exhaust NO_x and NO concentrations as a function of overall equivalence ratio or engine load. DI diesel, 1000 rev/min, injection timing at 27° BTC.¹⁹



Diluent added to the intake air (such as recycled exhaust) are effective at reducing the NO formation rate and NO_x emissions. As with spark-ignition engines, the effect is primarily one of reducing the burned gas temperature for a given mass of fuel and oxygen burned. Figure 11.17a shows the effect of dilution of the intake air with N_2 , CO_2 , and exhaust gas on NO_x exhaust levels.²¹ The heat capacity of CO_2 (per mole) at the temperatures relevant to diesel combustion is about twice that of N_2 . That of exhaust gas is slightly higher than that of N_2 . Thus these data show that the effect is primarily one of reduced burned gas temperatures. Note that the composition of the exhaust gas of a diesel varies with load. At idle, there is little CO_2 and H_2O , and the composition does not differ much from that of air. At high load the diluent capacity (mass times specific heat of inert gas) increases since the concentrations of CO_2 and H_2O in the recycled exhaust are substantially higher. Figure 11.17b shows the effect of increasing EGR at higher, mid-, and lower-loads.²² The engine can tolerate high EGR levels (50%) at low load since at very lean operating condition ($\phi \sim 0.2$), a major fraction of this EGR is effectively "air." As load and overall equivalence ratio increases, the mixture becomes less lean, and a larger fraction of the EGR is effectively "stoichiometric burned gas" and less is "air." Thus EGR behavior in the diesel combustion process is significantly different than that in a spark-ignition engine. Only a portion of the EGR acts as an inert diluent in the close-to-stoichiometric region of the diffusion flame where almost all of the NO is formed.

Figure 11.17 (a) Effect of reduction in oxygen concentration by different diluents (exhaust gas, CO_2 , N_2) on NO_x emissions in DI diesel. Bore = 140 mm, stroke = 152 mm, $r_c = 14.3$, speed = 1300 rev/min, fuel rate = $142 \text{ mm}^3/\text{stroke}$, injection timing at 4° BTC.²¹ (b) Effect of recycled exhaust gas (EGR) at light, mid, and high loads: Heavy-duty diesel engine at 1800 rev/min.²²



Confirmation that NO forms in the close-to-stoichiometric burned gas regions, and that the magnitude of the stoichiometric burned gas temperature controls NO_x emissions, is given by the following. Plee et al.^{23,24} have shown that the effects of

changes in intake gas composition (with EGR, nitrogen, argon, and oxygen addition) and temperature on NO_x emissions can be correlated by

$$\text{EI}_{\text{NO}_x} = \text{constant} \times \exp\left(\frac{E}{\tilde{R}T_f}\right)$$

(11.20)

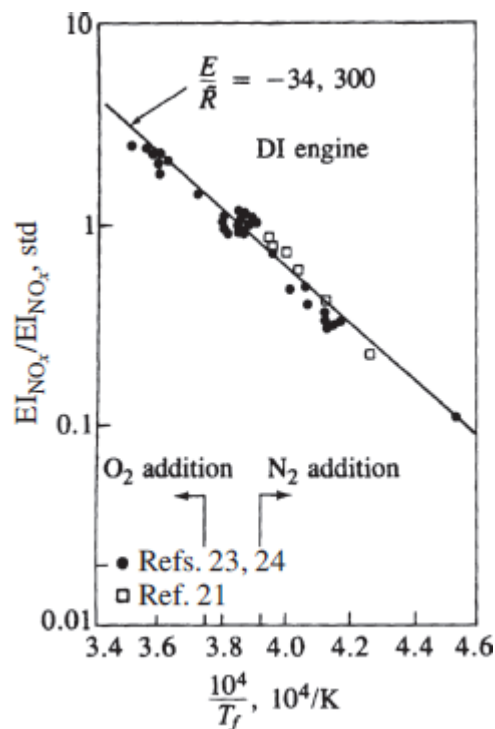
T_f (kelvins) is the stoichiometric adiabatic flame temperature (evaluated at a suitable reference point: fuel-air mixture at top-center pressure and air temperature) and E is an overall activation energy. Figure 11.18 shows

EI_{NO_x} for a range of intake air compositions and temperatures, from two DI engines for several loads and speeds, normalized by the engine NO_x level obtained for standard air, plotted on a log scale against the reciprocal of the stoichiometric adiabatic flame at TC conditions. A single value of $E/\tilde{R}$ correlates

EI_{NO_x} over more than an order of magnitude.²³ While there is some scatter since the model used is overly simple, and load, speed, and other engine design and operating parameters affect the process, the overriding importance of the close-to-stoichiometric diffusion-flame burned gas temperature is clear.

Figure 11.18 Correlation of NO_x emissions index

EI_{NO_x} for a wide range of operating conditions with the reciprocal of the stoichiometric mixture flame temperature for DI diesel engines: Flame temperatures varied by addition of different diluents and oxygen.^{21,23,24} Values of EI_{NO_x} normalized with value at standard conditions.



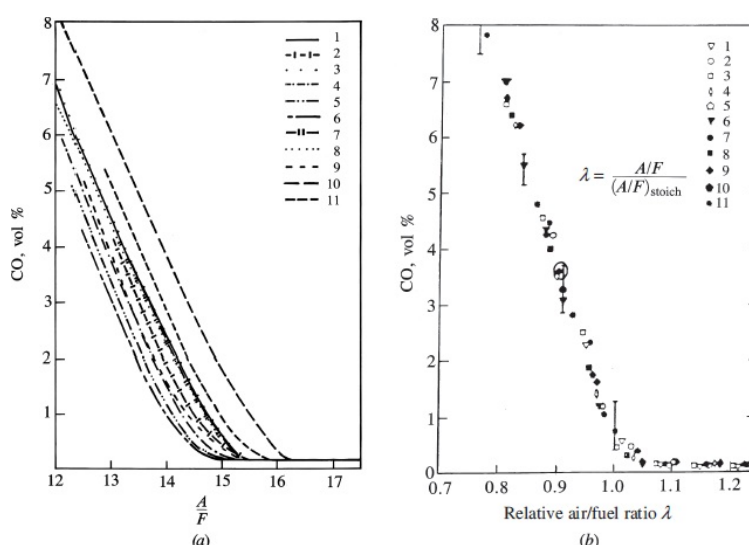
[1] This is often called the extended Zeldovich mechanism¹. Zeldovich was the first to suggest the importance of reactions (11.1) and (11.2). Lavoie et al.² added reaction (11.3) to the mechanism; it contributes significantly.

-
- [2] The steady-state approximation is used when the rate of change in concentration of a chemical species is sufficiently fast that the reactant and product species can maintain a quasi-equilibrium.
- [3] It is well known that the mixture composition within the cylinder is not completely uniform and varies modestly from one cycle to the next. Both these factors contribute to cycle-by-cycle combustion variations. For the present discussion, the assumption of mixture uniformity is adequate.
- [4] Spark timing also affects NO emissions, as discussed next. The above discussion relates to engines run with spark timing at MBT or with torque at a fixed percentage below (and close to) the maximum.

11.3. CARBON MONOXIDE

Carbon monoxide (CO) emissions from internal combustion engines are controlled primarily by the relative air/fuel ratio or fuel/air equivalence ratio. Figure 11.19 shows CO levels in the exhaust of a spark-ignition engine for several different fuel compositions.²⁵ When the data are plotted against the relative air/fuel ratio, they are correlated by a single curve. For fuel-rich mixtures CO concentrations in the exhaust increase steadily with decreasing relative A/F ratio, as the amount of excess fuel increases. For fuel-lean mixtures, CO concentrations in the exhaust vary little with relative A/F and are low, of order 10^{-3} mole fraction.

Figure 11.19 Variation of SI engine CO emissions with 11 fuels of different H/C ratio: (a) with air/fuel ratio and (b) with relative air/fuel ratio λ .²⁵



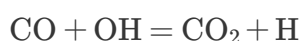
Since spark-ignition engines often operate close to stoichiometric at part load and, depending on calibration, may be fuel-rich at high load (see Sec. 7.1), CO emissions are significant and need to be controlled. Diesels, however, always operate well on the lean side of stoichiometric; CO emissions from diesels are low enough to be unimportant, and will not be discussed further.

The levels of CO observed in spark-ignition engine exhaust gases are lower than the maximum values measured within the combustion chamber, but are significantly higher than equilibrium values for the exhaust conditions. Thus the processes which govern CO exhaust levels are kinetically controlled. In a premixed hydrocarbon-air flame, the CO concentration increases rapidly in the flame zone to a maximum value, which is larger than the equilibrium value for adiabatic combustion of the fuel-air mixture. CO formation is one of the principal reaction steps in the hydrocarbon combustion mechanism, which may be summarized by¹



(11.21)

where R stands for the hydrocarbon radical. The CO formed in the combustion process via this path is then oxidized to CO_2 at a slower rate. The principal CO oxidation reaction in hydrocarbon-air flames is



(11.22)

The rate constant for this reaction is¹

$$k_{\text{CO}}^+ = 6.76 \times 10^{10} \exp\left(\frac{T}{1102}\right) \text{ cm}^3/\text{gmol}$$

(11.23)

It is generally assumed that in the post-flame combustion products in a spark-ignition engine, at conditions close to peak cycle temperatures (2500 to 2800 K) and pressures (15 to 40 atm), the carbon-oxygen-hydrogen system is equilibrated. Thus CO concentrations in the immediate post-flame burned gases are close to equilibrium. However, as the burned gases cool during the expansion and exhaust processes, depending on the temperature and cooling rate, the CO oxidation process [reaction (11.22)] may not remain locally equilibrated.

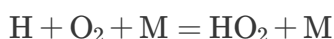
Newhall carried out a series of kinetic calculations for an engine expansion stroke assuming the burned gas at the time of peak cylinder pressure was uniform and in equilibrium.²⁶ Of the reactions important to CO chemistry, only three-body radical-recombination reactions such as



(11.24)



(11.25)



(11.26)

were found to be rate controlling. The bimolecular exchange reactions and the CO oxidation reaction (11.22) were sufficiently fast to be continuously equilibrated. Only during the later stages of the expansion stroke was the CO concentration predicted to depart from equilibrium, as shown in Fig. 11.20. Using this technique to predict average CO levels at the end of expansion, over a range of equivalence ratios (rich to lean), Newhall²⁶ obtained a good match to experimental data (see Fig. 11.21). The kinetically controlled aspects of the CO emissions mechanism have thus been demonstrated.

Figure 11.20 Results of kinetic calculations of CO concentrations, and equilibrium concentrations, during expansion stroke following TC combustion in SI engine; stoichiometric mixture.²⁶

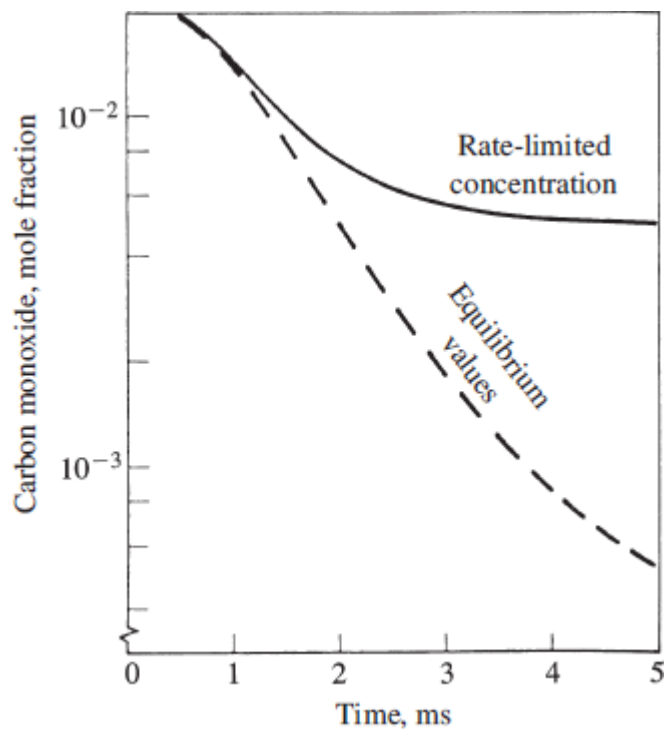
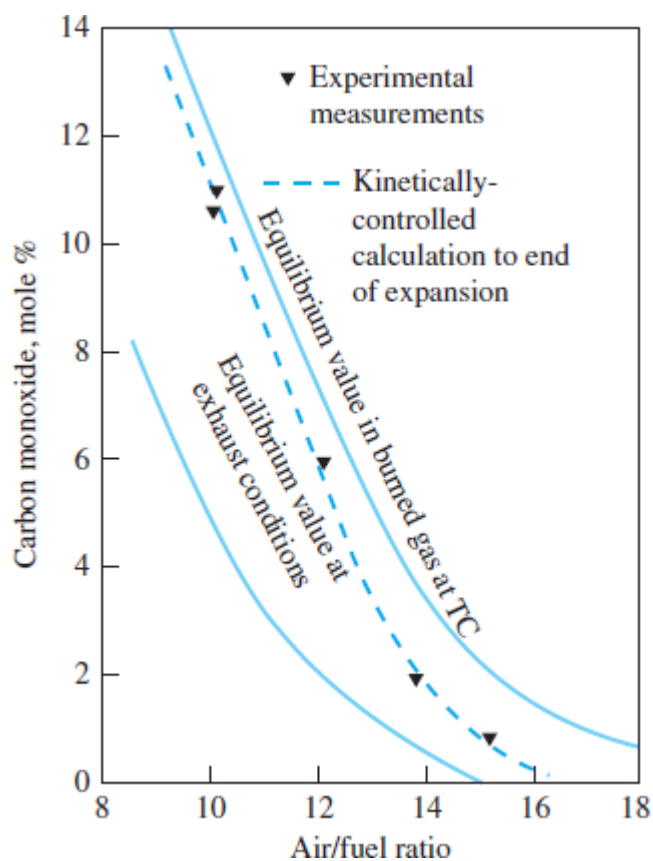


Figure 11.21 Predicted CO concentration at the end of expansion stroke, compared with measured exhaust concentrations, as function of air/fuel ratio. Equilibrium levels at TC combustion and exhaust conditions also shown.²⁶



Conclusions from these detailed studies show that exhaust CO concentrations for fuel-rich mixtures are close to equilibrium concentrations in the burned gases during the expansion process. For close-to-stoichiometric mixtures, the CO engine-out emissions are much higher than CO equilibrium concentrations corresponding to exhaust conditions. For fuel-lean mixtures, measured CO emissions are substantially higher than predictions from kinetically-controlled analysis of the chemistry bulk gas chemistry. One possible explanation of this lean-mixture discrepancy is that only partial oxidation to CO may occur as some of the hydrocarbons emerging during expansion and exhaust from crevices in the combustion chamber, and from oil layers or deposits on the chamber walls, burn up.

As a practical matter, exhaust CO emissions are determined by the relative air/fuel ratio. The degree of control achieved within the engine to date has come from improving mixture uniformity and leaning-out the intake mixture. In multi-cylinder engines, because CO increases rapidly as the inlet mixture becomes richer than stoichiometric, cylinder-to-cylinder variations in equivalence ratio about the mean value are important: nonuniform distribution of fuel and air about stoichiometric increases average emissions. Thus improved cylinder-to-cylinder fuel/air ratio distribution has become essential. Also, because it is necessary to enrich the intake mixture when the engine is cold, CO emissions during engine warm-up may be higher than emissions in the fully warmed-up state. Further, in transient engine operation during acceleration and deceleration, control of fuel metering has had to be improved. Major reductions in CO beyond what can be achieved in the engine are possible with exhaust catalysts, which are reviewed in [Sec. 11.6](#). Oxidation of CO in the exhaust system without use of special exhaust treatment devices does not normally occur to any significant extent because the exhaust gas temperature is too low: CO oxidation reactions effectively freeze as the burned gases pass into the exhaust system.²⁷

11.4. HYDROCARBON EMISSIONS

11.4.1. Background

Hydrocarbons, or more appropriately organic emissions, are the consequence of incomplete combustion of the hydrocarbon fuel. The level of unburned hydrocarbons (HC) in the exhaust gases of gasoline fueled spark-ignition engines is typically 1 to 2% of the fuel: in diesels it is much less. The levels are generally specified in terms of the total hydrocarbon concentration expressed in parts per million carbon atoms (C₁).^[1] While total hydrocarbon emission is a useful measure of combustion inefficiency, it is not necessarily an appropriate index of pollutant emissions impacts. Engine exhaust gases contain a wide variety of hydrocarbon compounds. Table 11.3 shows the average breakdown by class of the hydrocarbons in spark-ignition engine exhaust gases, both with and without a catalytic converter, with gasoline fuel. Some of these hydrocarbons are essentially inert physiologically and some are virtually unreactive from the standpoint of photochemical smog. Others are highly reactive in the smog-producing chemistry. (Also, some hydrocarbons are known carcinogens; see Sec. 11.5.2). Based on their potential for oxidant formation in the photochemical smog chemistry, hydrocarbon compounds are characterized by their reactivity—their ozone forming potential in the atmosphere. Table 11.4 shows one reactivity scale which is commonly used to estimate the reactivity of exhaust gas hydrocarbon mixtures: the *Incremental Reactivity* scale. All hydrocarbons except methane react, given enough time. More detailed breakdowns of the composition of spark-ignition and diesel engine exhaust HC are available in the literature. The values in Table 11.4 for specific organic compounds are in gO₃ (the dominant oxidizing species present in photochemical smog) per gram of NMOG (non-methane organic gases) in the exhaust: the Maximum Incremental Reactive value. Ranges of values are given for the important classes of organic species. Lower molecular weight paraffinic compounds have the lowest values, less than or about 1. Olefins are highest, values in the 4 to 10 range. The aromatics that are in gasoline are also high, 2.7 to about 10. These are called *incremental reactivities* since they are weighted by the concentrations of the main individual hydrocarbon species in the hydrocarbon mix in engine or vehicle exhaust, and then summed to determine the overall atmospheric reactivity of the exhaust gas stream.

Table 11.3 Hydrocarbon composition of spark-ignition engine exhaust (by class)²⁹

	Carbon, percent of total HC			
	Paraffins	Olefins	Acetylene	Aromatics
Engine-out emissions	33	27	8	32
Catalyst-out emissions	57	15	2	26

Table 11.4 Reactivity of different classes of hydrocarbons

Hydrocarbons	MIR value (gO ₃ /g NMOG)*
C ₁ –C ₄ paraffins	0.25–1
C ₅ and higher molecular weight paraffins	1–2
Cyclic paraffins	1–2.5
Ethylene	7.3
Olefins	4–10
Diolefins	10
Benzene	0.4
Toluene	2.7
Higher aromatics	6–10
Methanol	0.6
Ethanol	1.3
Formaldehyde	7.2
Acetaldehyde	5.5
*MIR: Maximum Incremental Reactivity values determined from standardized smog chamber testing. See Carter for more details. ³⁰	

Fuel composition can significantly influence the composition and magnitude of the organic emissions. Fuels containing high proportions of aromatics and olefins produce relatively higher concentrations of reactive hydrocarbons. However, many of the organic compounds found in the exhaust are not present in the fuel, indicating that significant pyrolysis and synthesis occur during parts of the combustion process.

Oxygenates are present in engine exhaust, and are known to participate in the formation of photochemical smog. Some oxygenates are also irritants and odorants. Oxygenates are generally categorized as carbonyls, phenols, and other noncarbonyls. The carbonyls of interest are low-molecular-weight aldehydes and aliphatic ketones. The volatile aldehydes are eye and respiratory tract irritants. Formaldehyde is a major component ($\leq 20\%$ of total carbonyls). Carbonyls account for about 10% of the HC emissions from diesel passenger car engines, but only a few percent of spark-ignition engine HC emissions. Phenols are odorants and irritants: levels are much lower than aldehyde levels. Other noncarbonyls include methanol, ethanol, nitromethane, methyl formate. Whether these are significant with hydrocarbon fuels is unclear. Use of alcohol fuels, as standalone fuels or blended with gasoline, increases oxygenate emissions. Both methanol and aldehyde emissions increase substantially above gasoline-fueled levels with methanol-fueled spark-ignition engines.

11.4.2. Flame Quenching and Oxidation Fundamentals

Flame quenching or extinction occurs at the walls of engine combustion chambers. The cool walls of the chamber act as a sink for the thermal energy and the active radical species generated within the flame. Quenching of the flame occurs under

several different geometrical configurations: the flame may be propagating normal to, parallel to, or at an angle to the wall; the flame may quench at the entrance to a crevice—a thin volume with a narrow entrance to the combustion chamber, such as the region between the piston crown and the cylinder wall. When the flame quenches, it leaves a thin layer or volume of unburned mixture ahead of the flame. Whether this results in hydrocarbon emissions depends on the extent to which these quench region hydrocarbons can subsequently be oxidized.

Flame-quenching processes have been analyzed by relating the heat release within the flame to the heat loss to the walls under conditions where quenching just occurs. This ratio, called a Peclet number (Pe), is approximately constant for any given geometrical configuration. The simplest configuration for study is the two-plate quench process, where the minimum spacing between two parallel plates through which a flame will propagate is determined. The Peclet number for this two-plate configuration is given by:

$$\text{Pe}_2 = \frac{\rho_u S_{Lc_{p,f}} (T_f - T_u)}{k_f (T_f - T_u) / d_{q2}} = \frac{\rho_u S_{Lc_{p,f}} d_{q2}}{k_f}$$

(11.27)

which is approximately constant over a wide range of conditions. ρ , S_L , c_p , T , and k are the density, laminar flame speed, specific heat at constant pressure, gas temperature, and thermal conductivity, respectively, with the subscripts u and f referring to unburned and flame conditions. d_{q2} is the two-plate quench distance. The wall temperature and unburned gas temperature are assumed to be equal. This assumption is also appropriate in the engine context since there is ample time during the compression stroke for a thermal boundary layer to build up on the combustion-chamber walls to a thickness of at least the quench distance.

Lavoie³¹ has developed empirical correlations for two-plate quench distance data for propane-air mixtures: only limited data for liquid hydrocarbon fuels such as isooctane are available. The data in the pressure range of 3 to 40 atm are well fitted by

$$\text{Pe}_2 = \frac{9.5}{\phi} \left(\frac{p}{3} \right)^{0.26 \times \min(1, 1/\phi^2)}$$

(11.28)

where p is the pressure in atmospheres and ϕ is the fuel/air equivalence ratio. The two-plate quench distance d_{q2} is then obtained from Eq. (11.27) and Prandtl number and viscosity relations for the flame conditions (see Sec. 4.8 or Ref. 31). Thus the minimum size crevice region into which a flame will propagate can be determined.

For the process of a flame front quenching on a single wall, there are many possible geometries. The simplest is where the flame front is parallel to the wall and approaches it head on. This one-wall quench distance d_{q1} , defined as the position of closest approach of the reaction zone to the wall, scales with flame properties in a similar way to the two-plate quench distance. Thus, a one-wall Peclet number relation can be formed:

$$\text{Pe}_1 = \frac{\rho_u S_{Lc_{p,u}} d_{q1}}{k_u} \approx 8$$

(11.29)

where the subscript u denotes properties evaluated at unburned gas conditions.

Using the wall temperature as representative of the local unburned gas temperature (because the thermal boundary-layer thickness is greater than typical quench distances), Lavoie showed that

$$\frac{d_{q1}}{d_{q2}} = \frac{Pe_1}{Pe_2} = 0.2$$

(11.30)

is a reasonable fit to the single-wall quench data. Single-wall quench distances are in the range of 0.04 to 0.15 mm. Typical two-wall quench distances for spark-ignition engine conditions are 0.2 to 0.6 mm; these distances indicate the minimum crevice size the flame will enter. However, the largest engine crevice, the gap between the piston crown (topland) and the cylinder liner, has a different geometry from either the two-wall or single-wall quench situation. The piston crown is effectively a flat plate, and the liner is perpendicular to it. The flame is thus propagating into a "corner" and the liner would be expected to provide additionally local flame cooling. This would likely increase the quenching distance for the actual geometry above the two-wall quench distance value. More recent experimental data on crevice volume impacts on efficiency and HC emissions supports this discussion: see Fig. 11.27.³²

While a fraction of the fuel hydrocarbons can escape the primary combustion process unburned, or only partially reacted, oxidation of some of these hydrocarbons can occur during the expansion and exhaust processes. Hydrocarbon oxidation rates have been determined in a number of different studies and several different empirical correlations of the data in the form of overall reaction rate equations have been proposed. A reasonable fit to the experimental data on unburned HC burnup is the rate expression:³¹

$$\frac{d[\text{HC}]}{dt} = -6.7 \times 10^{15} \exp\left(\frac{-18,735}{T}\right) \tilde{x}_{\text{HC}} \tilde{x}_{\text{O}_2} \left(\frac{p}{RT}\right)$$

(11.31)

where [] denotes concentration in moles per cubic centimeter,

$\tilde{x}_{\text{HC}}$ and

$\tilde{x}_{\text{O}_2}$ are the mole fractions of HC and O₂, respectively, t is in seconds, T is in kelvins, and the density term (p/RT) has units of moles per cubic centimeter. The spread in the data about this equation is substantial, however.

Studies of combustion of premixed fuel-air mixtures at high pressure in closed vessels have been useful in identifying the mechanisms by which hydrocarbons escape complete combustion. The residual hydrocarbons left unburned in the vessel following a combustion experiment have been shown to come primarily from crevices in the vessel walls. Thus wall quench hydrocarbons apparently diffuse into the burned gases and oxidize following the quenching event. Analytical studies of the flame quenching process, and postquench diffusion and oxidation with kinetic models of the hydrocarbon oxidation process, are in agreement with this assessment. Flame quenching can be thought of as a two-stage process. The first step is the extinction of the flame a short distance from the cold wall, determined by a balance between conduction of thermal energy from the hot reaction zone to the wall and the thermal energy released in the reaction zone by the flame reactions. The second step is the post-quench HC diffusion and oxidation occurring on a time scale of one or a few milliseconds after quenching. The diffusion and oxidation process ultimately reduces the mass of wall quench hydrocarbons to levels well below its value at the time of quenching.

Closed-vessel combustion experiments have also been used to show that oil layers on the walls of the vessel cause an increase in residual HC levels after combustion is complete. The additional HC that result in these experiments with oil films are primarily fuel molecules, and are proportional to the amount of oil placed on the walls of the reactor and the solubility of the specific fuel in that oil. Absorption of fuel in the oil occurs prior to ignition and the dissolved fuel is then desorbed into the

burned gases after combustion is complete. Thus fuel absorption into and desorption from oil layers on the cylinder liner is a potentially important engine HC mechanism.³³

11.4.3. HC Emissions from Spark-Ignition Engines

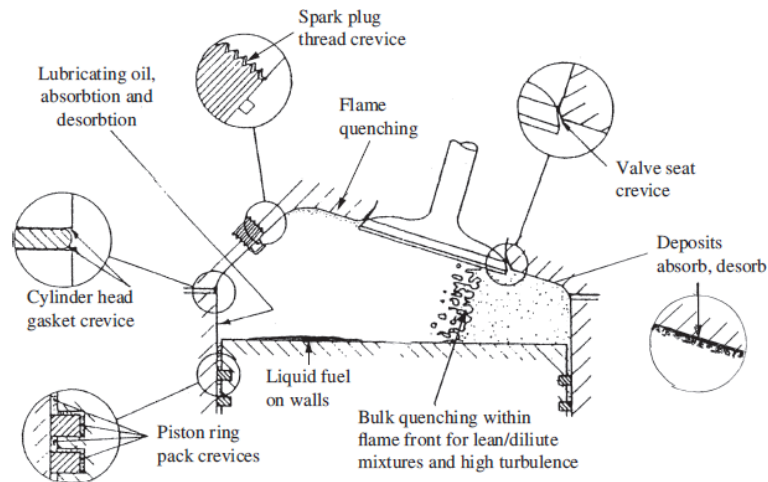
Engine-out hydrocarbon levels in the exhaust of a spark-ignition engine under normal operating conditions are typically in the range 1000 to 3000 ppm C₁. Normally, these HC emissions are between about 1 and 2% of the fuel flow into the engine. As indicated in [Fig. 11.2](#), HC emissions rise steadily as the mixture becomes richer than stoichiometric, and level off at a relative (lean) A/F close to 1.15. With very lean mixtures, as combustion quality deteriorates, HC emissions rise rapidly due to incomplete combustion or misfire in a fraction of the engine's operating cycles. As outlined in [Sec. 11.1](#), there are several sources that contribute to engine HC emissions. By sources, we mean any HC escaping the primary combustion process. Some of these HC may then fully or partly oxidize in the expansion and exhaust processes. While the complete HC emissions process is complex, there is sufficient information available to explain the overall process, and thus how major engine variables influence engine-out HC emission levels.

Several HC emissions formation mechanisms for spark-ignition engines (where the fuel-air mixture is essentially premixed) have been proposed. They are listed in [Table 11.5](#), and illustrated in [Fig. 11.22](#). The largest of these is the filling of crevice volumes with largely unburned mixture which, since the flame quenches near the crevice entrance, escapes the primary combustion process. All these processes together result in some 5 to 10% of the fuel not burning during the normal premixed-flame combustion process. Thus, in addition to being the source of the engine's HC emissions, they constitute a significant loss of torque and efficiency.^{32,34}

Table 11.5 Processes by which gasoline compounds escape burning during normal combustion

- Gasoline vapor-air mixture compressed into the combustion chamber crevice volumes
- Gasoline compounds absorbed in oil layers on the cylinder liner
- Gasoline absorbed by and/or contained within deposits on the cylinder head and piston crown
- Quench layers on the combustion chamber wall left as the flame extinguishes close to the wall
- Gasoline vapor-air mixture left unburned when the flame extinguishes prior to reaching the walls
- Liquid gasoline within the cylinder that does not evaporate and mix with sufficient air to burn prior to the end of combustion
- Leakage of unburned mixture through the (nominally) closed exhaust valve

Figure 11.22 Schematic of SI engine cylinder head and piston during combustion identifying the sources of hydrocarbon emissions.

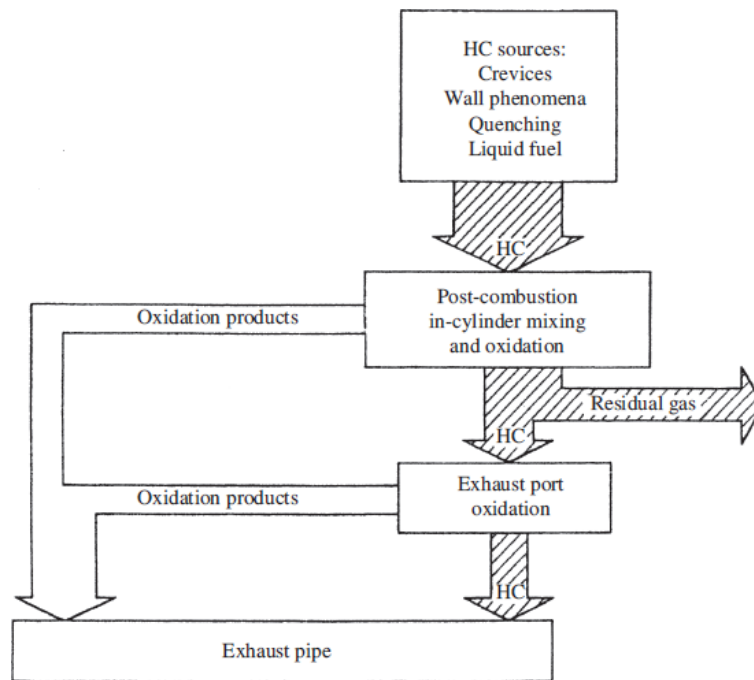


All these processes (except misfire) result in unburned hydrocarbons close to the combustion chamber walls rather than in the bulk of the in-cylinder gases. Thus, the distribution of these HC within the cylinder will not be uniform. For most of these sources, as combustion ends after peak pressure, these HC will flow out of the region where they were "stored", will diffuse into and mix with the hot burned gases in the bulk of the cylinder and, if their temperature becomes sufficiently high, will start to oxidize. [Table 11.6](#) summarizes, for each of these sources the processes by which, the HC that escaped burning during the normal combustion process leave the crevices, oil layers, etc., mix with the burned gases, then oxidize (partially or fully) or exit the cylinder in the exhausting gases. They then may oxidize in the exhaust system, or become HC emissions. [Figure 11.23](#) illustrates the total process leading to unburned HC emissions in the engine exhaust, from the various sources—mechanisms by which fuel escapes burning during the normal combustion process—through whether or not these HC oxidize in the cylinder or exhaust port, and if not, becoming engine-out HC emissions. The figure approximately quantifies the consequences of this many step process through the engine's (four-stroke) operating cycle by the width of each shaded bar. Engine-out HC are about one-third of the HC that escaped burning during combustion. Since engine-out HC^[2] correspond to 1 to 2% of the fuel entering each cylinder each cycle, the loss of torque and efficiency is significant.

Table 11.6 Processes by which these unburned hydrocarbons can oxidize or exit the cylinder

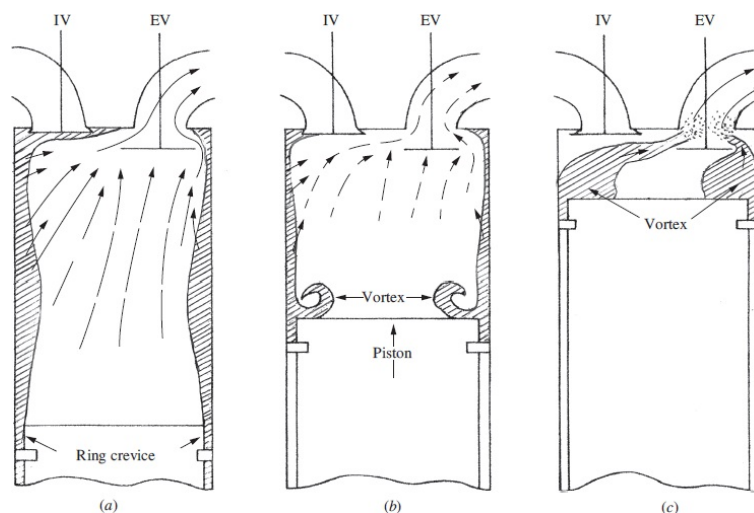
- Outflow of unburned fuel-air mixture from crevice volumes; mixing with burned gases; some of the crevice HC will oxidize.
- Diffusion of hydrocarbon vapor out of oil layers and deposits into the burned gases; some of these HC will oxidize.
- Mixing of wall and bulk quench gases with burned gases; some will oxidize.
- Exhaust blowdown process will carry a fraction of the in-cylinder unburned HC into the exhaust.
- The displacement of gases by the piston during the exhaust stroke will transport an additional fraction of the in-cylinder unburned HC into the exhaust.
- The unburned HC, which exit the cylinder, will mix with hot exhaust gases; a fraction of these HC will oxidize in the exhaust port and manifold.

Figure 11.23 The various steps in the overall SI engine hydrocarbon mechanism: HC sources, post combustion mixing and oxidation, retention, and exhaust. ³⁵ Width of shaded bars indicate HC concentration.



The HC concentration in the exhaust gases varies significantly during the exhaust process. [Figure 11.24](#) illustrates our current understanding. The sources of HC are initially around the outside of the combustion chamber; these hydrocarbons then diffuse from the walls into the bulk gas. Mixing of these HC with the bulk cylinder gases (and some oxidation) occurs during expansion and the exhaust blowdown process. The remaining unburned HC are exhausted in two peaks of approximately equal mass: the first of these coincides with the exhaust blowdown flow (which removes the majority of the mass from the cylinder). Then, during the final stages of the exhaust stroke, piston motion pushes most of the remaining fraction of the cylinder mass (with its high HC concentration) into the exhaust. This leaves a high concentration of HC in the residual gas in the cylinder, and also in the exhaust port and manifold. Approximately one-third of the hydrocarbons left unburned in an engine combustion event are retained in the cylinder and recycled, under typical driving conditions (light load). ^{36,37}

Figure 11.24 Schematic of in-cylinder flow processes by which hydrocarbons, post combustion, diffuse from the various sources into the bulk burned gases and are entrained into the blowdown exhaust flow (a), and then into the piston-driven outflow, (b) and (c).



These several schematics of the total spark-ignition engine HC emission process (Figs. 11.22, 11.23, and 11.24) explain its several sequential steps and parallel paths. We will return to the overall process after discussing the more important contributing mechanisms.

11.4.3.1. Crevice HC Mechanism

The crevices in the combustion chamber walls—small volumes with narrow entrances—into which the flame is unable to penetrate have been shown to be the largest source of unburned HC. The largest of these crevice regions is the volume between the piston, piston rings, and cylinder wall. Other crevice volumes in production engines are the threads around the spark plug, the space around the plug center electrode, crevices around the intake and exhaust valve heads, and the head gasket crevice. Table 8.1 shows the size and relative importance of these crevice regions in one cylinder of a typical production gasoline engine determined from measurements of cold-engine components. Total crevice volume is about 1.5 to 2.5% of the clearance volume and the piston and ring pack crevices are the dominant contributor.³²

The sequence of crevice processes occurring during the engine cycle are the following. As the cylinder pressure rises during compression, unburned mixture is forced into the crevice regions. Since these volumes are thin, they have a large surface/volume ratio so the gas flowing into each crevice cools by heat transfer to close to the wall temperature. During combustion, while the pressure continues to rise, unburned mixture continues to flow into the crevice volumes. When the flame arrives at each crevice, it can either propagate into these crevices and partially burn the fuel and air within the crevice, or it can quench at the crevice entrance. Whether the flame quenches depends on crevice entrance geometry, the composition of the unburned mixture, and its thermodynamic state as described in Sec. 11.4.2. After flame arrival at the crevice entrance, burned gases will flow into each crevice until the cylinder pressure starts to decrease. Once the cylinder pressure reaches its peak and starts to fall, gas flows back from each crevice into the cylinder.

The most important of these crevices, the volume between the piston, piston rings, and cylinder wall, is shown schematically in Fig. 8.32. The overall piston crevice region consists of a series of volumes, connected by flow restrictions such as the ring side clearance and ring gap, whose geometry changes as the ring moves up and down in the ring groove sealing either the top or (mostly) bottom ring surface. The gas flow, pressure distribution, and ring motion are therefore coupled, and their behavior during the compression and expansion strokes has already been discussed in Sec. 8.6. During compression and most of combustion, mass flows *into* the volumes in this total crevice region. Once the cylinder pressure starts to decrease (at MBT spark timing, about 15° ATC), gas flows out of the top of these crevice regions in Fig. 8.27 into the cylinder at low velocity adjacent to the cylinder wall. The important result is that a fraction of the total fuel mass in the cylinder (5 to 10%; see Example 11.1 below) is trapped in these regions at the time of peak cylinder pressure and escapes the primary combustion process. Most of this trapped gas flows back into the cylinder during the expansion process. Depending on spark plug location in relation to the position of the top ring gap, more than half of this gas is likely to be unburned fuel-air mixture. Its potential contribution to unburned HC emissions is obvious.

Example

EXAMPLE 11.1

At the peak pressure point in the cycle, the mass of gas in the crevices is at its maximum. If the volume of the crevices (V_{crev}) is 2% of the cylinder clearance volume (V_c), the crevice gases are at the average wall temperature (e.g., liner 100°C, piston crown 200°C), the in-cylinder gas is 75% burned (burned gas temperature 2700 K, unburned gas at 800 K), and the cylinder volume at 15° ATC is 1.16 times the clearance volume, calculate the mass of gas in the crevices relative to total mass in the cylinder. If the crevice gas is two-thirds unburned mixture, what fraction of the fuel is in the crevices at this point in the cycle?

The appropriate scaling relationship is:

$$\frac{m_{\text{crev}}}{m_{\text{cyl}}} = \frac{V_{\text{crev}}\rho_{\text{crev}}}{V_{\text{cyl}}\rho_{\text{cyl}}} = \frac{V_{\text{crev}}}{V_{\text{cyl}}} \left(\frac{p}{RT_{\text{crev}}} \right) \left(\frac{RT_{\text{cyl}}}{p} \right)$$

where the ideal gas law, $p = \rho RT$, has been used. The average crevice gas temperature is approximately 150°C ($=423\text{ K}$); the average cylinder gas temperature is $(0.75 \times 2700 + 0.25 \times 800)$ which is 2225 K . Thus

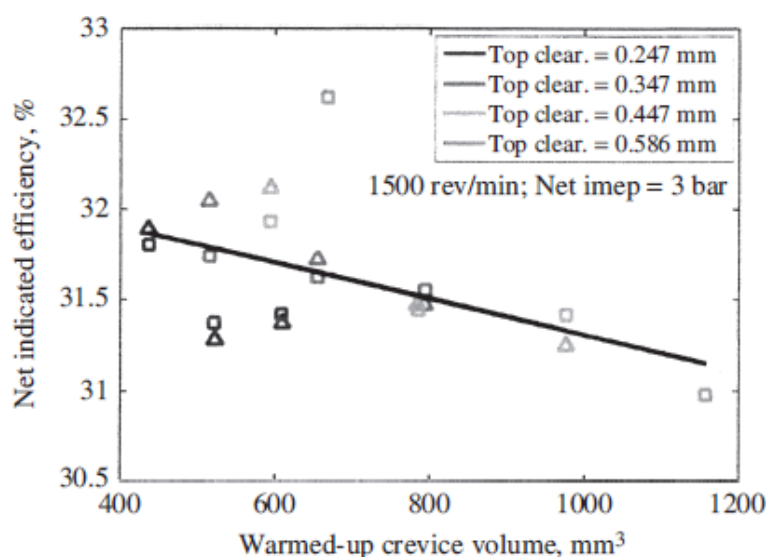
$$\frac{m_{\text{crev}}}{m_{\text{cyl}}} = \frac{2}{100} \times \frac{2225}{423} = 0.105 \text{ or about } 10\%.$$

If the unburned gas fraction is two-thirds, then the fraction of fuel which escapes burning during combustion is 7%.

There is substantial evidence to support the above description of crevice HC phenomena and the piston/ring crevice region in particular. Visualization studies in a special engine have identified the spark plug crevice outflow, low-velocity gas expansion out of the volume above the first ring after the time of peak pressure, and the jet-like flow through the top ring gap later in the expansion process when the pressure difference across the top ring changes sign.³⁸ Gas sampling from the volume above the top ring has shown that the gas composition in this region corresponds to unburned fuel-air mixture until flame arrival at the crevice entrance closest to the sampling valve location. Next, burned gases enter the crevice as the cylinder pressure continues to rise. Then, during expansion as gas flows out of this region, the composition of the gas sampled reverts back toward that of the unburned mixture which entered the crevice region earlier.³⁹ Flame geometry calculations suggest that about two-thirds of the crevice gases are unburned fuel, air, residual gas mixture, though the percentage depends on the spark plug location and the symmetry of flame front propagation.³⁴

Direct evidence that the piston and ring crevice regions are a major contributor to exhaust HC emissions comes from experiments where the volume of this crevice region has been substantially changed. HC emissions increase as the top-land crevice volume increases. The extent of the topland height has been decreased over time (from some 10 to 5 mm), to decrease this crevice volume. So has the radial clearance. HC emissions do scale with piston crevice size and reductions in other crevice volumes, but the relationship is more complex than a linear one due to partial oxidation of the crevice HC and there being other HC sources. The data in Fig. 11.25 confirm that the fuel stored in the piston crevices is significant through its impact on the engine's fuel conversion efficiency.³² The least-squares linear fit to the data indicate that a 1000 mm^3 (1 cm^3) increase in piston top-land crevice results in a 1 to 1.3 percentage points decrease in net indicated fuel-conversion efficiency. This change in efficiency (due to a 1000-mm^3 change in piston top-land crevice volume) corresponds to 3% of the fuel entering the cylinder not burning early enough to produce torque. Allowing for in-cylinder HC oxidation (about two-thirds), HC retention in the residual gas (one-third), and exhaust port oxidation (also about one-third) this would result in 0.4 to 0.5% of the fuel as HC. Thus the piston top-land crevice, some three-quarters of the total cylinder crevice volume, is the origin for about 40% of the engine-out HC. This percentage is in agreement with other estimates.³⁴

Figure 11.25 Effect of crevice volume on net indicated fuel conversion efficiency. Data from single-cylinder SI engine with several different piston ring-pack dimensions, each giving different piston crevice volumes: 1500 rev/min and net imep = 3.6 bar.³²



Just how the flame quenches at crevice entrances is only partly understood. Figure 11.26 compares the warmed-up top-land clearances with the two-plate quench distance [Eqs. (11.27 and 11.28)] for a range of cold piston-radial-clearances, with the predicted two-plate quench distance. In this study,³² the warmed-up radial clearance was 0.08 to 0.09 mm less than the cold clearance, and the comparison suggests that at the higher loads some flame penetration into the crevice may occur. An earlier study directly measured flame penetration: with the specific piston geometries and operating conditions used, flame penetration varied from essentially zero to a few mm (average just over 1 mm) so was not significant, most of the time.⁴⁰ Also, the geometry of the two-plate flame quenching situation is different from that of the piston crevice, as illustrated in Fig. 11.27.³² The vertical cylinder-liner wall in the right-hand sketch (which is cooler than the piston crown) would be expected to cool "half" of the flame close to the walls quicker, so quenching would occur at wider radial piston clearances than the basic two-plate quench studies would indicate. Thus we anticipate that under current engine circumstances, flame quenching close to the piston crevice entrances does occur.

Figure 11.26 Comparison of warmed-up piston-to-liner top-land clearance for the four different pistons tested in the study shown in Fig. 11.25, with the two-plate quench distance, as a function of engine load. (Speed change effects are small.)³²

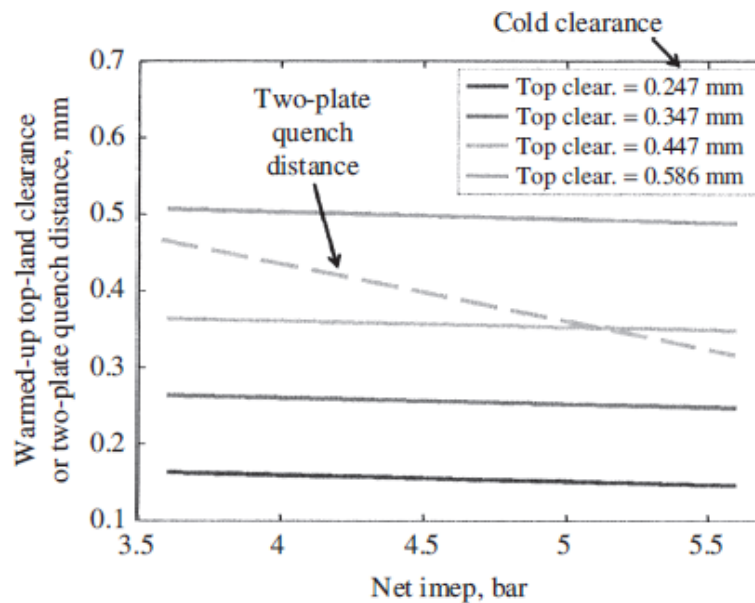
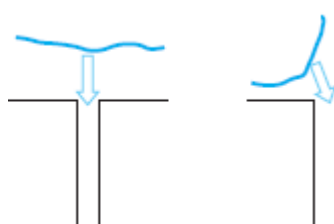


Figure 11.27 The geometries for the two-plate flame quenching process (left), and the piston-liner crevice flame quenching process (right).



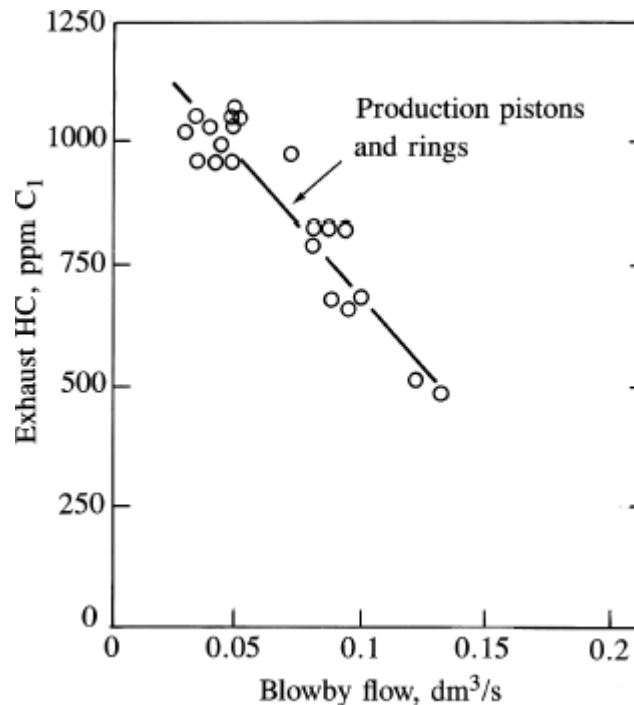
The impact of the various crevice volumes identified in Fig. 11.22 can be inferred from their relative volumes. Table 11.7 compares the actual size of the different crevice volumes, and their percentage of the total, in a modern automobile spark-ignition engine. Through improvements in design, over time, these volumes have been decreased. The top-land height is constrained by piston rim strength considerations. The piston top-land dominates (~ 75%) though the spark plug crevices are about one-fourth of that. (Whether the piston second land volume, between the top and second rings, is a significant contributor is unclear. The gas pressure in this region is only a few atmospheres.) Earlier studies which converted the cylinder head, and the cylinder with piston close to TC, with steel plates to constant volume combustion vessels, then sealing and filling the various crevices in turn, produced HC concentrations after combusting the fuel air mixture inside each vessel which then allowed estimates of the relative contributor from each crevice.⁴¹ These percentages were comparable in ranking and in magnitude to the volume percentages in Table 11.7.

Table 11.7 Volumes of the various crevices (mm^3 and percent of total) in single cylinder of modern spark-ignition engine³²

	Volume (mm^3)	Percent
Spark plug crevice	187	18.0
Valve seat crevice	32	3.1
Head gasket crevice	37	3.6
Piston crevice	782	75.3
Total crevice volume	1038	100.0
Clearance volume (mm^3)	45,400	
Vol. piston crevice/Vol. clearance (%)		1.72
Vol. total crevice/Vol. clearance (%)		2.28

Blowby is the gas that flows from the combustion chamber, past the piston and rings, into the crankcase. It is forced through any leakage paths afforded by the piston-liner-ring assembly in response to high in-cylinder gas pressures during compression, combustion, and expansion. Blowby of gases from the cylinder to the crankcase removes gas from this crevice region and thereby prevents some of the crevice gases from returning to the cylinder. Crankcase blowby gases were once vented directly to the atmosphere and constituted a significant source of HC emissions. The engine's crankcase is now vented to the intake system, the blowby gases are recycled, and this source of HC emissions is now fully controlled. Blowby at a given speed and load is controlled primarily by the greatest flow resistance in the flow path between the cylinder and the crankcase. This is the compression ring gap area. [Figure 8.35](#) shows how blowby increases linearly with this smallest gap area. [Figure 11.28](#) shows how exhaust HC levels decrease as blowby increases and more crevice HC flows to the crankcase. While blowby is usually, measured as a volume flow rate (at standard atmospheric pressure and temperature) out of the engine's crankcase, it is more meaningful to express it as a percentage (by mass) of the total flow rate into the cylinder. That flow in is air plus fuel plus any EGR. For a typical cylinder volume (500 cm^3) at mid-load and low-speed, 0.05 liter/s blowby corresponds to about 1% of this total flow. See light-load example below.

Figure 11.28 Effect of increasing crankcase blowby on exhaust hydrocarbon emissions. Production pistons and rings. SI engine at 1200 rev/min, intake manifold pressure 0.6 atm, A/F = 14.2. Blowby is per cylinder: for a 500-cm³ displaced-volume cylinder, 0.5 liters/s correspond to about 1.2% of the air, fuel, and EGR flow at this mid-load condition.⁴²



Example

EXAMPLE 11.2

A 500-cm³ displaced-volume cylinder operating at a manifold air pressure of 0.5 bar and 1500 rev/min, a typical light-load operating condition, has a blowby of 0.05 dm³ (liters)/s. Estimate this blowby as a percentage of the total mass flow rate into the cylinder.

The air flow into the engine is given by

$$\dot{m}_a = \eta_v \rho_{a,i} V_d N / 2$$

where $\rho_{a,i}$ is the air density at intake air conditions: with this definition, $\eta_v \approx 0.8$. The density of air is obtained from the ideal gas law:

$$\rho_{a,i} = \frac{p}{RT} = \frac{0.5 \times 100}{287 \times 310} = 0.56 \text{ kg/m}^3$$

so

$$\dot{m}_a = 0.8 \times 0.56 \times 500 \times 10^{-6} \times 1500 / 60 \div 2 = 2.8 \times 10^{-3} \text{ kg/s}$$

The blowby flow of 0.05 liter/s, assuming that the blowby gas (a mixture of fuel and air, and burned gases) has a density close to that of air (1.18 kg/m³) at standard pressure and temperature, converts to a mass flow:

$$\dot{m}_{by} = 0.05 \times 10^{-3} \times 1.18 \times 10^3 = 0.059 \text{ g/s}$$

The cylinder intake flow of air, fuel (assumed stoichiometric), and EGR (assumed to be 10 percent of the exhaust flow) is given by

$$\dot{m} = \dot{m}_a + \dot{m}_f + \dot{m}_{EGR} = \dot{m}_a \left(1 + \frac{\dot{m}_f}{\dot{m}_a} \right) / (1 - x_{EGR})$$

Thus:

$$\dot{m}_i = 2.8 \text{ (g/s)} (1 + 0.067) / (1 - 0.1) = 3.3 \text{ g/s}$$

and blowby flow as a fraction of total intake flow is

$$x_{by} = 0.059 / 3.3 = 0.018 (= 1.8\%)$$

Blowby is typically 1 to 2% of the in-cylinder mass, per cycle. This loss of gas from the cylinder has a modest effect on torque (the cylinder pressure is slightly lower due to blowby) but since the blowby is recycled, the impact on efficiency is negligible.

Crankcase blowby gases represent a performance loss. They are a smaller efficiency loss because crankcase gases are now recycled to the engine intake system.

The fate of the crevice HC when they flow back into the cylinder during expansion and exhaust is not well understood. Both low-velocity creeping flows (from the piston top-land crevice) and jet-like flows (from the top ring gap) have been observed (see Fig. 8.34). While the latter could mix rapidly with the high-temperature bulk burned gases, the latter will enter the cooler gases in the cylinder wall boundary layer and (probably) mix and burn much more slowly. Hydrocarbon transport and oxidation processes are discussed more fully below.

11.4.3.2. HC Absorption and Desorption in Engine Oil

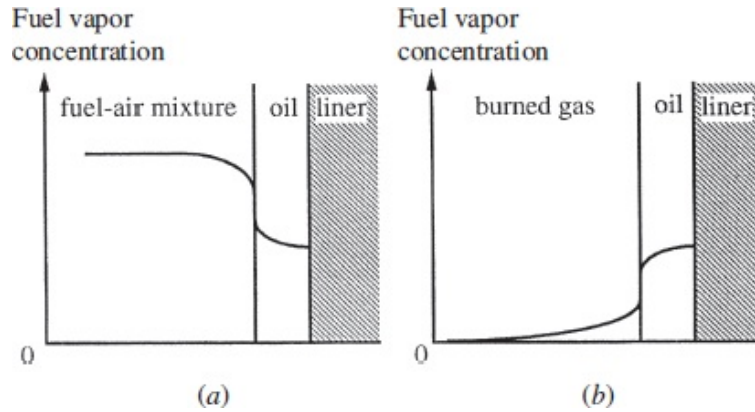
Lubricating oil absorbs gasoline components from the in-cylinder unburned mixture. Absorption and desorption occur within the cylinder each engine cycle thus providing another mechanism for some of the fuel to escape burning during the primary combustion process.

This absorption/desorption mechanism is thought to work as follows. An oil layer a few microns thick on the cylinder liner is uncovered by the descending piston during the intake stroke, and remains exposed to the in-cylinder gases until covered by the piston during the compression stroke. During this time the oil layer is exposed to unburned fuel-vapor, air, residual gas mixture (and possibly, during the intake process, to liquid fuel). Figure 11.29a shows the expected concentration profiles of gasoline vapor components that are soluble in the oil during intake and compression. As a consequence, such vapor components are transferred from the bulk gas into the oil layer. It is usually assumed that Henry's Law, which relates the partial pressure of each vapor component p_f to the dissolved mole fraction of that component in the oil $\tilde{x}_f$, is valid:

$$\tilde{x}_f = \frac{p_f}{H}$$

(11.32)

Figure 11.29 Schematic of HC absorption into, and desorption out of, oil layers on the liner. Gasoline component vapor concentration profile in the in-cylinder gases and in the lubricating oil on the cylinder liner: (a) during intake and compression and (b) post combustion and during exhaust.



where

$\tilde{x}_f$ is the mole fraction of fuel vapor in the oil, p_f is the partial pressure of fuel vapor in the gas, and H is Henry's constant. Henry's constant, which relates these two properties, varies strongly with temperature so the amount of vapor absorbed into the oil is significantly higher at lower oil temperatures and (via each component's partial pressure) at higher gas pressures.

During the expansion and exhaust strokes, the oil layer on the liner, uncovered by the downward moving piston, has higher concentrations of gasoline compounds in the oil than correspond to equilibrium with the close-to-zero levels of gasoline in the burned gases which fill the cylinder: see Fig. 11.29b. Thus hydrocarbons are desorbed out of the oil film and diffuse into the cylinder gases. If they do not then oxidize, they can exit the cylinder and contribute to the engine-out unburned hydrocarbon emissions. If the oil layer is sufficiently thin, and thus diffusion is rapid, Eq. (11.32) can be used to estimate the mole fraction of the fuel dissolved in the oil. Since

$p_f = n_{f,c} \tilde{R}T/V$ (where $n_{f,c}$ is the number of moles of fuel in the cylinder, T is the temperature, and V the cylinder volume) and $x_f = n_f / (n_{f,o} + n_o) = n_{f,o} / n_o$ for $n_o \gg n_{f,o}$ (where $n_{f,o}$ is the number of moles of fuel dissolved in the oil and n_o is the number of moles of oil),⁴³ then

$$\frac{n_{f,o}}{n_{f,c}} = \frac{n_o \tilde{R}T}{HV}$$

(11.33)

Diffusion is sufficiently rapid for Eq. (11.33) to be valid if the diffusion time constant τ_d is much less than characteristic engine times: that is,

$$\tau_d \approx \frac{\delta^2}{D} \ll N^{-1}$$

where δ is the oil layer thickness, D is the diffusion coefficient for fuel vapor in the oil, and N is engine speed. D for a hydrocarbon diffusing through a motor oil is of order 10^{-6} cm²/s at 300 K and of order 10^{-5} cm²/s at 400 K. Oil film thicknesses on the cylinder wall vary during the operating cycle between about 1 μ m and 10 μ m.^{44,45} Thus diffusion times for engine conditions are 10^{-1} to 10^{-3} s; thus for the thinnest oil layers, equilibration would essentially be achieved.

Evidence for oil absorption and desorption at the high end of this range come from experiments where a specially designed spark-ignition engine was operated normally with an oil film on the liner to lubricate the piston rings and skirt, and then oil free.⁴⁶ For stoichiometric operation in the oil-free mode, at light load and low speed, engine out emissions were about 25% less than with normal lubrication. Other studies^{47,48} indicate that in a warmed-up engine the contribution may be less than about 10%. In a cold engine, when the solubility of key fuel components in the oil is significantly higher, the contribution is likely to be higher. Under colder engine conditions, gasoline compounds are present in the oil in the sump (1 to 2%), and at higher levels than this in the oil residing in the ring pack region.⁴⁹ Under these conditions, it is likely that liquid fuel impinges on the liner oil layer, and flows into the ring pack. This additional gasoline in the oil is likely to desorb HC into the burned exhausting gases as the engine warms up and be a source of additional HC emissions.

The different organic compounds in gasoline vary significantly in their solubility in engine oils, which makes accurate predictions of the amount of HC this mechanism produces difficult. However, there is now substantial experimental and theoretical evidence that oil layers within the cylinder, primarily on the cylinder liner, are a source of some 10 to 25 percent of the unburned HC emissions.³⁴

11.4.3.3. Impact of Deposits

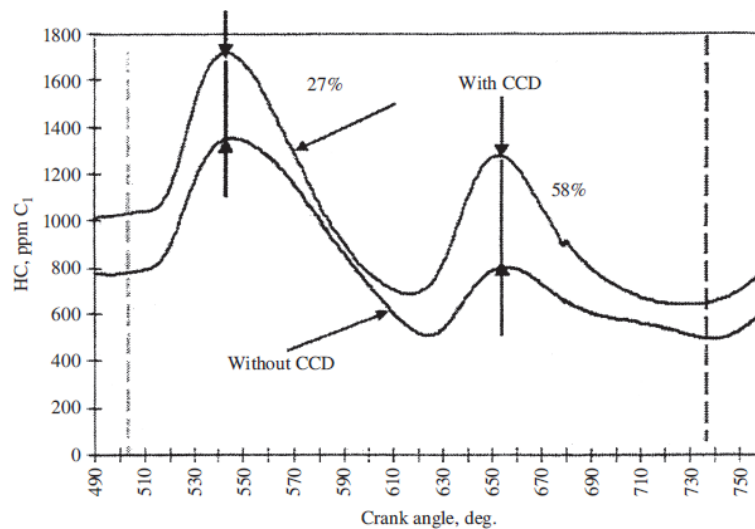
Over extended mileage, deposits build up on the intake valves, and on the combustion chamber walls of spark-ignition engines fueled with standard gasolines. Both these types of deposits affect hydrocarbon emissions. The deposits on the outside of the intake valve, which are porous, absorb some of the liquid fuel injected onto the back (intake side) of the valve with port fuel injection systems, thereby delaying its entry into the cylinder during engine transients. The in-cylinder mixture is therefore leaner than intended during accelerations (throttle openings) and richer than intended during decelerations (throttle closings). Changes in mixture composition affect engine-out HC emissions. Intake valve deposits primarily affect HC emissions during changes in engine load and, during starting and warm-up. They do not significantly affect steady-state HC emissions.

Combustion chamber deposits are known to increase steady-state as well as transient HC emissions. While deposit build-up on the cylinder and piston crown walls can take some 10,000 to 15,000 km to stabilize, their impact on HC emissions occurs much more rapidly. The engine-out HC rise most rapidly from clean-engine values over about 10 hours of operation, then more slowly over the next 20 hours, and stabilize over the next 10 to 20 hours even though deposit thicknesses continue to grow. However, the extent to which combustion chamber deposits in actual use build-up is not well quantified, and may well vary significantly with use patterns as well as the fuel used, oil consumption and oil type (the lubricant is the key component in forming deposits). SI engine deposits are porous in structure. The larger pores apparently act as "distributed crevices," and store HC during the combustion process. The smaller pores with their large surface to volume ratio may absorb significant HC into any engine oil absorbed onto the deposits, which reacts with gasoline combustion-generated radicals to form these deposits. Substantial levels of deposits in modern spark-ignition engines (e.g., levels which cause an increase in octane requirement of 5 to 10 octane numbers) can cause steady-state engine-out hydrocarbon emissions to increase by significant amounts.⁵¹

The above discussion is consistent with data on in-cylinder deposit build-up effects on HC emissions, which show that HC emissions can rise significantly over several thousand miles depending on how mileage was accumulated. [Figure 11.30](#) shows the time-resolved HC concentrations measured with a fast-response FID HC analyzer through the exhaust process (130° ATC EVO to 15° ATC EVC: there are sampling time delays) and compares a clean engine (without combustion chamber deposits) with the same engine with stabilized deposit levels (with CCDs). The trace shows that total HCs are high during the first blowdown high mass flow phase, and later during the latter half of the exhaust stroke. This pattern of the blowdown flow entraining HC originating close to the cylinder head and resulting in significant HC exiting the cylinder in the first third of the exhaust process, and of the piston motion during the latter part displacing high HC gases from the bottom and middle of the cylinder during the displacement exhaust phase and out of the exhaust valve, has been known for many years.³⁶ With CCD, in [Fig. 11.30](#), additional HCs exit early (likely from the fuel absorbed into the cylinder-head deposits) and exit later (likely from piston crown deposits). The time-averaged HC concentration difference due to deposits in this example of an engine with significant deposits is an increase of 20 to 25%. What is not yet adequately documented is the effect on engine-out HC

emissions of combustion chamber deposit levels typically found in modern spark-ignition engines, with today's gasolines and lubricants.

Figure 11.30 Time-resolved HC concentrations through the exhaust process of a warmed-up spark-ignition engine, with and without combustion chamber deposits (CCDs). Percentages are the HC increase at the two peaks above the clean engine (without CCDs) values. 2-liter four-cylinder naturally-aspirated gasoline-fueled engine operated at wide-open throttle, 2500 rev/min, and $\lambda = 0.9$. EVO and EVC at 503° and 730° ATC intake stroke, respectively.⁵⁰



11.4.3.4. Flame Quenching

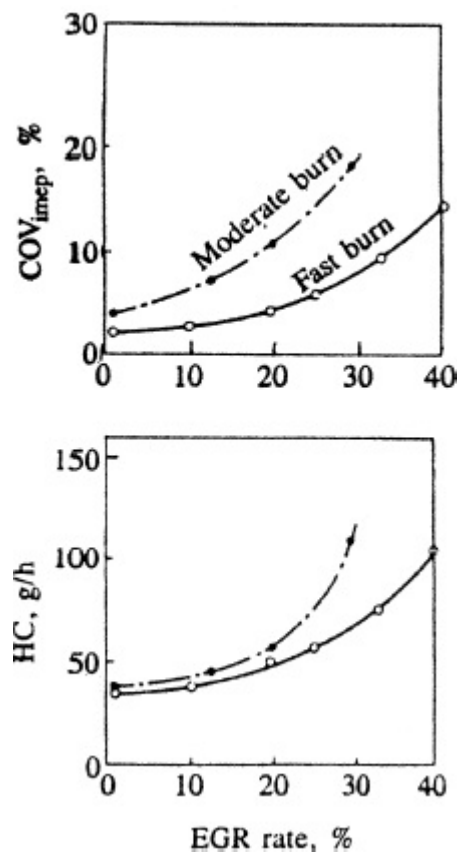
It has long been known that the spark-ignition engine flame extinguishes a short distance from the cooled combustion chamber walls leaving a thin layer of unburned (or partially reacted) fuel-air mixture adjacent to the wall. This is called a quench layer. While originally thought to be a major source of unburned HC emissions, it is now known that the mixture in these quench layers diffuses into the hot combustion products outside the layer and burns up during expansion, under most engine operating conditions.³⁷ It is not yet clear whether under extremely lean or high-EGR engine operation the thicker quench layers that then occur do fully burn up. However, even when the mixture in the quench layers does fully burn up before the engine cycle is completed, flame quenching at the combustion chamber walls still prevents a small fraction of the fuel (about 0.5%) from burning during the normal engine combustion process. Quench layers thus constitute a small loss in engine performance and efficiency.

Flame extinguishment prematurely, before the above-described flame quenching at the chamber walls occurs, is potentially a more significant contributor to engine-out HC emissions. Such extinguishment can occur when an engine is operated close to its dilute operating limit (e.g., with high EGR), or its lean operating limit. Then, in slower burning engine cycles, the flame may approach the chamber walls sufficiently slowly, and late in the expansion stroke, that the pressure and temperature in the unburned mixture ahead of the flame is decreasing with time rather than increasing. The flame can extinguish under these circumstances resulting in a partial burning cycle, which could cause high HC emissions *for that cycle*. The effect on average HC emissions would depend on the relative frequency of occurrence of such partial burning cycles. It is also possible that misfiring cycles occur occasionally close to the EGR or lean limit when the flame initiated by the spark discharge extinguishes early in its development process. Then, close to a full cylinder charge of fuel and air will exit the cylinder unburned.

With a well-designed fast-burn combustion system, under normal steady-state engine operating conditions, flame-quenching phenomena are not expected to contribute significantly to HC emissions. However, during transients, with high levels of EGR, and with retarded timing, partial burning and even quench layers may contribute enough to total HC emissions to warrant attention. Poor combustion evident as partial burning and even misfiring in a small fraction of the engine's operating cycles

are the explanations often given for part of the increase in engine-out HC emissions that occurs as EGR is increased at a fixed load and speed, as shown in Fig. 11.31.³⁴ The top figure shows how combustion slows down and deteriorates as EGR increases from moderate to high levels (COV_{imep} is the standard deviation, cycle-to-cycle, in imep divided by the mean imep. See Sec. 9.4.3. A COV_{imep} below 2 to 5%, depending on operating conditions, is regarded as satisfactory for "stable" combustion. Part of the difference between the slower and faster burning combustion systems shown, at higher EGR levels, is due to the deteriorating combustion process. However, in a normally operating engine, bulk flame quench should not be occurring.

Figure 11.31 Effect of increasing percent exhaust gas recycle (EGR) on the coefficient of variation in indicated mean effective pressure and engine-out HC emissions at 1400 rev/min and 324 kPa imep, stoichiometric operation, MBT spark timing.³⁴



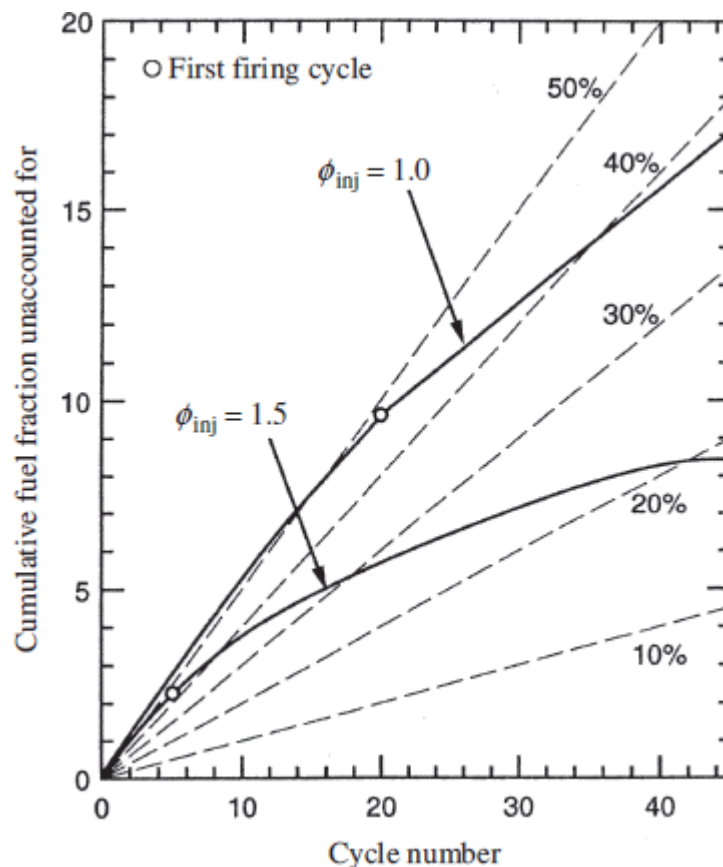
11.4.3.5. Liquid Fuel on the Cylinder Walls

Both port-fuel-injected and direct-injected gasoline engines are in common use. In port-injected engines, the injector usually directs the liquid gasoline toward the back of the intake valve and surrounding intake port surface area. Most of the injected gasoline impinges on these surfaces, though a portion of the fuel will evaporate from the fuel droplets formed in the atomization process. The behavior of the liquid fuel after injection is strongly influenced by the timing of the injection pulse relative to the intake valve opening and closing times. With injection onto a closed valve, the rapid backflow of burned gases from the cylinder immediately following intake valve opening assists the fuel vaporization process off the valve and port surfaces substantially;⁵² under fully warmed-up part-load engine operating conditions, almost all the gasoline fuel enters the cylinder as vapor. With injection toward an open valve, however, more of the fuel will enter the cylinder as liquid, as droplets or ligaments entrained in the intake airflow. A major part of this liquid evaporates and mixes with the in-cylinder air, fuel vapor, and residual gas, to form the normal combustible mixture during the intake and compression process. The remaining liquid, especially the less volatile fuel components may form a liquid film, or may be stored in the deposits, oil layers or crevices in the combustion chamber, and vaporize into the bulk gases during the expansion and exhaust processes. The incomplete

oxidation of this fuel would contribute to the exhaust hydrocarbon emissions. The HC emissions index, Eq. (2.44), for liquid fuel entering the cylinder with port-injection is some five times that for fuel that enters as vapor, though this varies substantially with the valve seat location where the liquid flows in.⁵³

The amount of not-yet evaporated fuel in the cylinder is especially significant during the engine starting process. Because the inlet valve and port, and the intake air, are cold, only a fraction of the injected fuel will evaporate: to achieve rapid and smooth starting, a much larger amount of fuel than would be required with complete vaporization must be injected, and a substantial amount of that fuel does not appear in the combustible charge mixture. This phenomenon is illustrated in Fig. 11.32. The data for this figure was obtained from single-cylinder engine running at a fixed speed (900 rev/min) and intake manifold pressure (0.5 bar). The in-cylinder hydrocarbon vapor concentrations before combustion were measured by a fast-response flame-ionization HC detector. From these HC concentrations values, the fuel mass, which had evaporated to form the combustible mixture, was estimated for each cycle. The remaining fuel mass was labeled as "unaccounted for." The cumulative amount of "unaccounted for" fuel is plotted as a function of the number of cycles from first injection in Fig. 11.32. The first firing cycles are also marked in the figure. Initially, only about 50% of the fuel enters the cylinder and vaporizes. The cumulative mass of unaccounted-for fuel for enrichments typical of engine startup could be some 10 or so injections. Part of the fuel fraction that enters the cylinder without vaporizing and mixing with the combustible charge mixture may find its way to the oil sump. The remaining part, however, may mix with the charge later in the cycle and, if it is not oxidized, would contribute to the engine-out hydrocarbon emissions.

Figure 11.32 "Unaccounted for" fuel accumulation (normalized by the amount of one injection) as a function of cycle number during engine starting and warm-up: 900 rev/min, intake manifold pressure 0.5 bar, spark timing 20° BTC. The circle on each trace marks the first cycle with significant positive imep. Also shown are lines of constant fuel-air equivalence ratio, based on injected fuel, of 1.0 and 1.5.^{34,54}



Comparisons have been made between the HC emissions levels with prevaporized gasoline fueling and with port injection.

For port-fuel-injected gasoline (indolene), the steady-state HC emission levels are only about 10% higher than the prevaporized fuel case. During the warm-up process, however, the HC emissions are substantially higher (by about a factor of 2) than those of the prevaporized fuel case. The time scale to steady state is of order 200 seconds, which is much slower than that for the prevaporized fuel. The higher HC emissions (compared to the prevaporized fuel case) during warm-up are attributed to the presence of the liquid fuel in the crevice/oil layer/deposit regions in the combustion chamber, and are presumably due to two effects: (1) The presence of liquid fuel would mean that a much larger amount of HC is stored in these regions and (2) the transport processes of the liquid fuel into and out-of these "storage" regions are different from those of the fuel vapor. Both effects would contribute to the higher level of HC emissions. That the steady state HC emissions are only about 10% higher than in the prevaporized case is probably due to the fact that the intake manifold and back of the inlet valve have reached high enough temperatures to achieve nearly complete vaporization: thus little gasoline enters the cylinder as liquid.

11.4.3.6. Gasoline Direct-Injection Engines

Direct-injection (DI) fuel systems are increasingly replacing (or supplementing) port-fuel-injection systems. With DI, fuel is injected directly into the combustion chamber through an injector located in each cylinder head. The behavior of the liquid fuel injected into the cylinder has several HC emissions impacts. First, the liquid fuel films formed on the intake port and valve walls with port injection are eliminated. However, with DI there will be much more liquid fuel within the cylinder, especially with a cold engine and during warm up, which radically changes the fuel metering calibration during starting and warm up, and transients. The fuel entering each cylinder in a DI engine is exactly what is injected. In port-injected systems, due to the buildup of liquid fuel films within the intake port, there are fuel transport delays and, during engine start up, significantly more fuel must be injected than enters the cylinder during the first several cycles. Thus DI fuel injection has the potential for reducing the cold-engine fuel enrichment required, and the more complex fuel metering calibration necessitated by port fuel injection. Thus, with DI, engine-out hydrocarbon emissions during these phases of engine operation can be reduced.

For homogeneous-charge DI spark-ignition engines under quasi-steady operating conditions with gasoline, the HC emissions mechanisms are essentially the same as with port-fuel-injection engines. The behavior of the DI fuel spray does need to be carefully managed to avoid any substantial liquid wetting of the cylinder liner and piston crown surfaces. Fuel wetting the walls and forming liquid films does increase the HC emissions from an SI engine. The reason, as explained in the previous subsection, is that vaporizing that liquid fuel on the walls (largely due to heat transfer from the hot burned gases after the flame passes over the film), and then mixing that fuel vapor with sufficient oxygen in those burned gases to burn up that fuel, takes time and thus occurs well into the expansion stroke.⁵⁵

DI gasoline engines that are operated stratified at light- and mid-load generally have higher HC emissions than homogeneous stoichiometric-operating DI engines. Flame quenching near the outer boundary of the close-to-stoichiometric fuel-containing cloud (see [Sec. 9.3.5](#)) occurs due to the very lean mixtures (due to overmixing of fuel with air) in these outer cloud regions. Then, since the temperatures of lean-mixture burned gases are lower than of stoichiometric burned gases, the potential for post-flame oxidation of these HC is significantly reduced. Also, with lean engine operation overall in this stratified DI mode, exhaust gas temperatures are lower (as a consequence of the lower expansion-stroke gas temperatures of lean gases, and of higher engine efficiency) so afterburning of HC in the exhaust port is reduced as also is the oxidation catalyst efficiency (see [Sec. 11.6.3](#)) due to lower catalyst temperatures. Higher HC emissions levels from lean stratified DI gasoline spark-ignition engines is one of their major development challenges, as is the parallel need for catalyst systems for NO_x removal from lean exhaust gases (such as the lean-NO_x trap, [Sec. 11.6.3](#)).⁵⁶

11.4.3.7. Exhaust Valve Leakage

Experiments with fast-response flame-ionization HC detectors in the exhaust port of engines have shown that, periodically, a rise in HC concentrations can occur in the port while the valve is closed. It is believed that exhaust valve leakage can occur intermittently, even with normally functioning exhaust valve systems. A plausible cause is that combustion chamber deposits

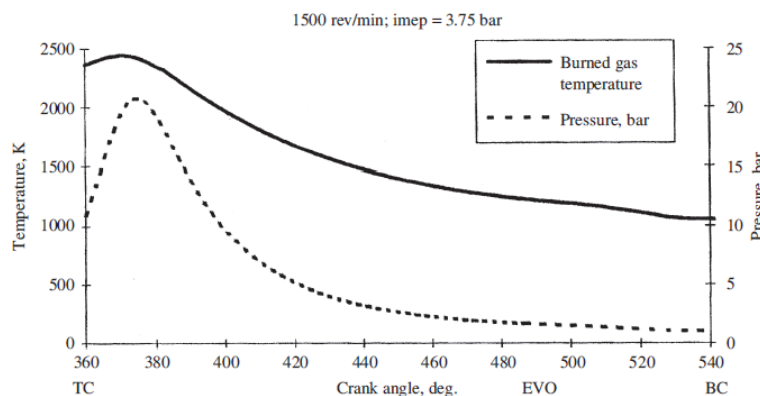
may break-up and the larger particles that result become embedded for short periods in the exhaust valve seat. If leakage occurs, then prior to combustion, unburned mixture can flow into the exhaust. Under most circumstances this source of HC emissions is expected to be small though not necessarily negligible. In poorly maintained older engines, it may be a more significant source: up to 7% of total engine-out HC emissions.

11.4.3.8. Oxidation and Retention in the Cylinder

As summarized in Table 11.6, after the end of the normal combustion process, the HC which through the above-described mechanisms did not burn, can oxidize within the cylinder, or may be retained within the cylinder with the residual burned gas and therefore be recycled, or may be oxidized within the exhaust port and manifold. Each of these three processes significantly reduces the amount of unburned HC emitted by the engine. This second phase of the total HC emissions mechanism, despite its importance, has received much less attention than the "sources" part of the mechanism summarized in Table 11.5 and described above.^[3]

Figure 11.33 shows the variation in the average in-cylinder gas temperature for a typical part-load operating condition during the expansion and exhaust processes. If the unburned HC from the sources described above mix early enough with the burned gases, and provided enough oxygen is present, they will burn. Note that there are two different types of sources. Crevices and flame quenching produce unburned *fuel-air mixture*; here the oxygen required to burn the fuel is already mixed with the fuel vapor. In contrast, oil layers, deposits, and liquid fuel mechanisms produce unburned hydrocarbon vapor, *separated* from the oxygen it will need to burn.

Figure 11.33 Cylinder pressure and average in-cylinder (burned) gas temperature during the expansion and exhaust strokes of the SI engine cycle, 1500 rev/min, 3.8 bar imep, stoichiometric operation. (Note that at peak pressure about 80% of the mixture has burned.)³⁴



The studies of in-cylinder HC oxidation suggest that, locally, the HC in the quench layers on largely deposit-free combustion-chamber walls diffuse into the burned gases during the expansion stroke and oxidize almost completely. Engine cycle simulations suggest that the in-cylinder gas expansion and associated temperature decrease that occurs during the blowdown process that immediately follows exhaust valve opening slows down the in-cylinder oxidation process sufficiently to prevent much additional burn up occurring. At burned gas temperatures below about 1000 K, significant oxidation of HC in the time available in the cylinder is not likely to occur. An assessment of total in-cylinder HC oxidation by Daniel,⁵⁷ who fitted a physical mixing model and an overall chemical oxidation model to a matrix of engine-out HC emissions data at different operating conditions, estimated the amount of this post-flame oxidation. This model fitting process indicated that about one-third of the HC which escape the combustion process do fully oxidize within the cylinder. It is likely that the amount oxidized for each of the different HC sources depends on the composition of the source "efflux," the time of occurrence of that efflux during expansion and exhaust, and the location of the source within the cylinder. Various studies of these questions suggest that the HC from sources that produce fuel-air mixture after normal combustion (crevices, flame quenching) will oxidize significantly more than the HC from sources that produce fuel vapor (oil layers, deposits, liquid fuel films.)

The retention of the still unburned HC within the cylinder is also poorly understood. While the residual gas fraction can be estimated accurately, the HC are not uniformly distributed within the cylinder: since the sources are on the walls of the combustion chamber, the fraction of the in-cylinder HC retained is likely to be larger than the residual fraction. Previous analysis of the shape of the HC concentration and mass flow profiles through the exhaust process^{36, 58} indicates that about half the exhausted HC exit the cylinder during the blowdown process and half toward the end of the exhaust-stroke displacement process. The blowdown presumably entrains much of the unburned HC adjacent to the wall at the top of the cylinder. The exhaust stroke is known to roll up the gas layer adjacent to the cylinder liner (which presumably contains much of the remaining unburned HC from the piston crevices and oil layers) into a vortex (see [Sec. 8.9](#)). [Figure 11.24](#) illustrates our understanding of how the high HC-concentration gases from the various HC sources located around the outside of the combustion chamber move into the bulk gases, and then exit the cylinder during the blowdown and piston displacement phases of the exhaust process.

11.4.3.9. Oxidation in the Exhaust

With a warmed-up engine, the gas temperature in the exhaust port, and even (under some conditions) in the exhaust manifold is hot enough for gas-phase HC oxidation to occur. Prior analyses and measurements confirm that for stoichiometric operation at part-load, with a warmed-up engine, about one-third of the engine-out HC will be burned up, as shown in [Fig. 11.23](#). Under fuel-rich operating conditions, oxidation can be enhanced by air addition into the port region to ensure that adequate oxygen for burnup is available.

To oxidize the hydrocarbons in the gas phase, a residence time of order 50 ms at temperatures in excess of 600°C is required. To oxidize carbon monoxide, temperatures in excess of 700°C are required. Average exhaust gas temperatures at the cylinder exit (at the exhaust valve seat) are about 800°C; average gas temperatures at the exhaust port exit are about 600°C.^[4] [Figure 6.21](#) shows an example of measured cylinder pressure and the gas temperature at the exhaust port exit, and the estimated mass flow rate into the port, during the exhaust process at a part-throttle operating condition. Exhaust port residence time and gas temperatures vary significantly through the process. Precise values of these variables obviously depend on engine operating conditions. Normally, it is only in the exhaust port (and upstream end of the manifold) that any significant gas-phase HC oxidation can occur. Note that spark retard from MBT timing is an especially important variable. For cold-start engine-out HC emissions control, retarding the spark timing to after TC for 10 or more seconds is widely used to create a hotter exhaust flow which increases HC, CO, and H₂ burn up, and more quickly heats up and lights off the catalyst.

11.4.3.10. Secondary Air

Injection of secondary air into the exhaust ports of automobile spark-ignition engines to reduce the hydrocarbon emissions (and also CO emissions) has a long history. It is an effective strategy, but adds an extra system (an air pump, drive, and controls), and thus cost. It does provide significant reductions in cold start HC emissions entering the catalyst, and also speeds up the catalyst light-off process. Both of these benefits help achieve extremely low vehicle-out HC emissions levels. Also, secondary air can be used to compensate for engine operation with a richer-than-stoichiometric mixture during starting and warm-up, which achieves smoother and more stable combustion but significantly increases engine-out HC emissions. An important part of this cold-start strategy is to retard the combustion phasing relative to piston motion, by sparking the fuel-air mixture in the cylinder (operating in the stoichiometric to some 20% rich range) after top center. Such retarded combustion increase the exhaust gas temperature significantly thus enabling much greater HC and CO burn up, with the two effects combining to light off the catalyst much more rapidly.

A study of the effects of secondary air injection on exhaust gas emissions, exhaust system heat transfer, and catalyst light off during cold start, using a variety of techniques to quantify emissions levels and temperatures at several locations from cylinder exit to the catalytic converter outlet, has identified the fundamentals of what occurs. The engine was operated at fast idle (i.e., cold engine) conditions, at fixed λ , with retarded spark timing (4 to 10° ATC), at fixed speed (1450 rev/min) and airflow (11 g/s). In-cylinder lambda was varied from 0.8 to 1.0. Secondary airflow rate ranged from 0 to 100% of the intake

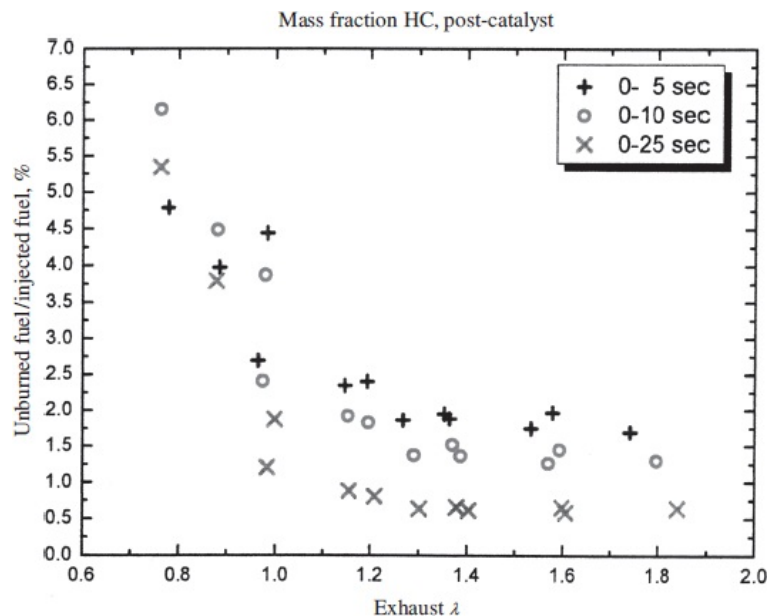
flow in the four-cylinder 2.2 liter naturally-aspirated gasoline engine. Operating 20% rich with a secondary air flow of 100% of the intake flow yielded the highest catalyst feed-gas temperature prior to catalyst light off (at 4.2 seconds). Substantial spark retard is required to initiate thermal oxidation of HC and CO in the exhaust gases. Holding this retarded spark for the duration of the catalyst light off period (~10 sec for an ambient-temperature cold start) is needed to keep the engine-out gases hot and enhance thermal oxidation in the exhaust port and manifold. To maximize this thermal oxidation, the secondary air should be injected as close to the exhaust valve as possible. The chemical energy released in the exhaust port outweighs the cooling effect of the secondary air. With $\lambda = 0.8$, secondary air could release up to 90% of the cylinder-out exhaust gas chemical energy. Overall, the secondary air injection system reduced HC emissions by some 50 to 90% and CO emissions by 35 to 90% over the first 25 sec of operation at the conditions studied.⁵⁹

An overall finding of this study was that post-catalyst HC emissions levels decreased steadily up to an exhaust relative air/fuel ratio of 1.3, with increasing secondary air, as shown in Fig. 11.34. The exhaust λ is defined by

$$\lambda_{\text{exh}} = \left(\frac{\dot{m}_{a,i} + \dot{m}_{a,s}}{\dot{m}_f} \right) / (A/F)_s$$

(11.34)

Figure 11.34 Cumulative mass HC emissions normalized by the amount of injected fuel at post-catalyst location as a function of exhaust lambda (relative air/fuel ratio, including secondary air), Eq. (11.34), for time periods 0 to 5 sec, 0 to 10 sec, and 0 to 25 sec.⁵⁹



where subscripts i and s denote intake and secondary air flows. Beyond this secondary air level there was no additional benefit from adding more air: the post-catalyst HC become essentially independent of the engine's relative air-fuel ratio.

11.4.3.11. Summary of the Total Process

It will be apparent from the above that the HC emissions formation process in spark-ignition engines is complex, and that there are several paths by which small but important amounts of the fuel escape combustion. It is appropriate here to summarize the overall structure of the spark-ignition engine hydrocarbon emissions problem and identify the key factors and engine variables that influence the different parts of that problem. Table 11.8 provides such a summary. The total process is

divided into four sequential steps: (1) the creation of unburned hydrocarbon sources; (2) the oxidation of a fraction of these HCs within the cylinder, following mixing with the hot bulk gases; (3) the flow of a fraction of the unoxidized HCs from the cylinder into the exhaust; and (4) the oxidation in the exhaust system of a fraction of the HCs that exit the cylinder. The detailed processes and the design and operating variables that influence each of these four steps in a significant way are listed.

Table 11.8 Critical steps and important engine variables in HC emissions mechanism

<ol style="list-style-type: none"> 1. Sources of unburned HC <ol style="list-style-type: none"> a. Crevices <ol style="list-style-type: none"> 1. Crevice volume 2. Crevice location (relative to spark plug) 3. Load 4. Crevice wall temperature 5. Mixture composition* b. Oil layers <ol style="list-style-type: none"> 1. Oil consumption 2. Wall temperature 3. Speed c. Incomplete combustion <ol style="list-style-type: none"> 1. Burn rate and variability 2. Mixture composition* 3. Load 4. Spark timing[†] d. Combustion chamber walls <ol style="list-style-type: none"> 1. Deposits 2. Wall roughness 	<ol style="list-style-type: none"> 2. In-cylinder mixing and oxidation <ol style="list-style-type: none"> a. Mixing rate with bulk gas <ol style="list-style-type: none"> 1. Speed 2. Swirl ratio 3. Combustion chamber shape b. Bulk gas temperature during expansion and exhaust <ol style="list-style-type: none"> 1. Speed 2. Spark timing[†] 3. Mixture composition* 4. Compression ratio 5. Heat losses to walls c. Bulk gas oxygen concentration <ol style="list-style-type: none"> 1. Relative (A/F) ratio d. Wall temperature <ol style="list-style-type: none"> 1. Important if HC source near wall 2. For crevices: importance depends on geometry
--	---

<p>3. Fraction HC flowing out of cylinder</p> <p>a. In-cylinder flow during exhaust stroke</p> <ol style="list-style-type: none"> 1. Valve overlap 2. Exhaust valve size and location 3. Combustion chamber shape 4. Compression ratio 5. Speed <p>b. Residual fraction</p> <ol style="list-style-type: none"> 1. Load 2. Exhaust pressure 3. Valve overlap 4. Compression ratio 5. Speed 	<p>4. Oxidation in exhaust system</p> <p>a. Exhaust gas temperature</p> <ol style="list-style-type: none"> 1. Speed 2. Spark timing[†] 3. Mixture composition[*] 4. Compression ratio 5. Secondary air flow 6. Heat losses in cylinder and exhaust <p>b. Oxygen concentration</p> <ol style="list-style-type: none"> 1. Relative (A/F) ratio 2. Secondary air flow and addition point <p>c. Residence time</p> <ol style="list-style-type: none"> 1. Speed 2. Load 3. Exhaust component volume <p>d. Exhaust reactor[‡]</p> <ol style="list-style-type: none"> 1. Oxidation catalyst 2. Three-way catalyst
<p>* Relative air/fuel ratio and burned gas fraction (residual plus recycled exhaust gas).</p>	
<p>† Relative to MBT spark timing.</p>	
<p>‡ See Sec. 11.6.</p>	

The specific processes are listed in [Table 11.8](#) in the most likely order of importance. Through each of these source mechanisms, fuel or fuel-air mixture escapes the primary combustion process. Those hydrocarbons must then survive the expansion and exhaust processes and pass through the exhaust system without oxidation if they are to end up in the atmosphere as HC emissions. The rate of mixing of these HCs with the hot bulk cylinder gases, the temperature and composition of the gases with which these HCs mix, and the subsequent temperature-time and composition-time histories of the mixture, will govern the amount of in-cylinder oxidation that occurs. The distribution of these HCs around the combustion chamber is nonuniform (and changes with time); they are concentrated close to the walls of the chamber. The fraction of these HCs that will exit the chamber during the exhaust process will depend on the details of the in-cylinder flow patterns that take them through the exhaust valve. Overall, the magnitude of the residual fraction will be one major factor, but the residual gas is known to be much richer in HCs than the average exhaust gas. In particular, the flow patterns in the cylinder toward the end of the exhaust stroke, as gas is scraped off the cylinder wall as the piston moves toward the exhaust valve, will be important. Finally, a fraction of the unburned HCs which leave the cylinder through the exhaust valve will burn up within the exhaust system. Gas-phase oxidation in the exhaust ports (and possibly in the hotter parts of the exhaust manifold) is significant. The amount depends on the gas temperature, composition, and residence time. Then, the catalysts in the exhaust system provide a very substantial additional reduction in HC emission levels. Exhaust catalysts and their operating characteristics are described in [Sec. 11.6.3](#).

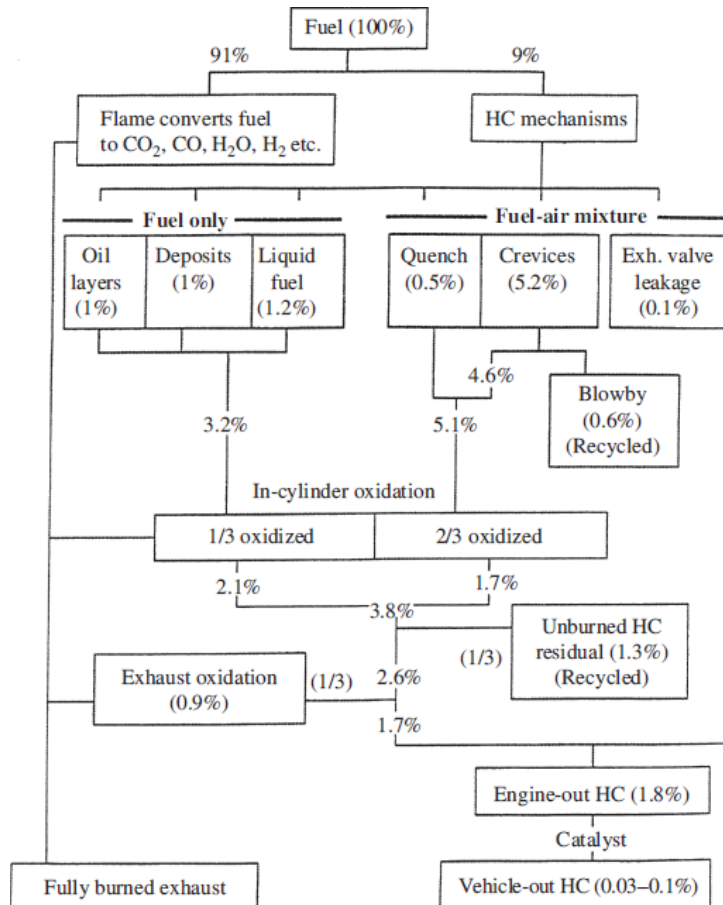
We now quantify the fuel fractions involved in the full HC emissions mechanism, based on our earlier discussion of each step in the process. The overall task is to quantify what fraction of the fuel remains unburned after each stage in the schematic shown in [Fig. 11.23](#). This has been done for a typical part-load operating condition defined by the details in [Table 11.9](#).³⁴

Table 11.9 Part-load operating condition for fuel and HC flow chart

Fuel = gasoline
 Stoichiometric A/F
 Speed = 1500 rev/min
 Inlet pressure = 0.46 atm
 Imep = 3.8 bar, bmep = 2.7 bar
 Spark timing = MBT (20° BTC)
 Warmed-up coolant and oil

The complete flow chart for the gasoline which enters each cylinder each cycle is shown in [Fig. 11.35](#). The fuel within each cylinder prior to any chemical reaction (100%) is at the top: through normal flame-propagation combustion some 90% of it burns producing hot combustion products CO_2 , CO , H_2O , H_2 , N_2 , ... (the left-hand side of the diagram) which encompasses almost all the exhaust gases. The hydrocarbon emissions component of the exhaust prior to catalyst entry (lower right) is about 1.8% of the mass of the fuel, a well-defined number with a spread of about ± 0.5 for typical engine operation. Note that normally about half these HC are compounds that are in the fuel, and half are partial reaction products of fuel molecules. The individual HC sources, and the amounts shown are *emissions indexes*, Eq. (2.44), expressed as percent of the initial fuel in the cylinder. They are categorized as *fuel only* sources and *fuel-air mixture* sources (e.g., oil layers and crevices, respectively). This was done because evidence suggests that when the HCs from an individual source are fuel mixed with air, it is easier to oxidize them in the cylinder (post normal combustion) because the fuel vapor already has the oxygen needed to burn it up mixed with it. If the source outflow is HC alone (e.g., gasoline compounds desorbing from oil layers on the liner) then the HCs involved must mix with oxygen *and* with hot burned gases before they can oxidize. The in-cylinder oxidation ratios (one-third oxidized, or two-thirds oxidized for fuel only and fuel-air mixture, respectively) reflect that distinction.

Figure 11.35 Flow chart for the gasoline fuel within each cylinder at the end of intake through both the normal combustion process and the total HC emissions mechanism. Sources or processes by which some of the fuel escapes burning during the normal combustion process are shown as boxes at the top of the diagram. In-cylinder oxidation, retention in the cylinder in the residual gas, and oxidation in the exhaust port and manifold, and catalyst are shown in sequence below. Numbers in parenthesis denote the HC emissions index (percent of gasoline in the cylinder at the start of each cycle) after each step in the total process.³⁴



During the exhaust process, of the HCs *not oxidized* within the cylinder (3.8%) one-third is retained within the cylinder (at this part-load condition), and two-thirds are exhausted of which one-third oxidizes within the exhaust ports. Thus from the original 9% of the fuel that escaped burning during normal combustion, some 1.8% remains in the exhaust as engine-out HC emissions.

Note that through blowby to the engine crankcase (0.6%), which is returned to the engine's intake system and the 1.3% remaining in the cylinder with the residual gases, together, 1.9% of this 9% of fuel escaping normal combustion is recycled to the cylinder. [Table 11.10](#) summarizes the percentage of fuel involved in each HC source, and the percent that source contributes to the total engine-out HC emissions. The importance of crevices is clear.

Table 11.10 HC sources: Magnitudes and percent of total engine-out HC emissions

Source	% of fuel escaping normal combustion	% of total HC engine-out emissions
Crevices	5.2	38*
Quench	0.5	5
Oil layers	1.0	16
Deposits	1.0	16
Liquid fuel	1.2	20
Exhaust valve leakage	0.1	5
TOTAL	9.0	100
*Blowby (0.6%) subtracted		

The loss in imep is less than this 9% by the amount that in-cylinder HC oxidation contributes to work transferred to the piston ($\int p dV$) during expansion. This HC--oxidation expansion-stroke work contribution decreases from unity if it occurs during normal combustion at close to TC (which achieves close to the full expansion possible) to close to zero between EVO and BC. Post-normal-combustion in-cylinder oxidation of HCs is likely to contribute, on-average, about one-third of this. Thus, this additional imep contribution would be about $[(3.2 + 5.1) - (2.1 + 1.7)] \div 3 \approx 1.5\%$. So the overall imep loss is $(9.0 - 1.5) \approx 7.5\%$. The fuel conversion efficiency loss benefits from the blowby and residual being "recycled": that is $(0.6 + 1.3) \approx 1.9\%$ for an efficiency loss of $(7.5 - 1.9) \approx 5.6\%$. It is often not realized that the engine-out HC emissions of about 2% actually represent a three to four times larger performance and efficiency loss.

11.4.4. Hydrocarbon Emission Mechanisms in Diesel Engine

11.4.4.1. Background

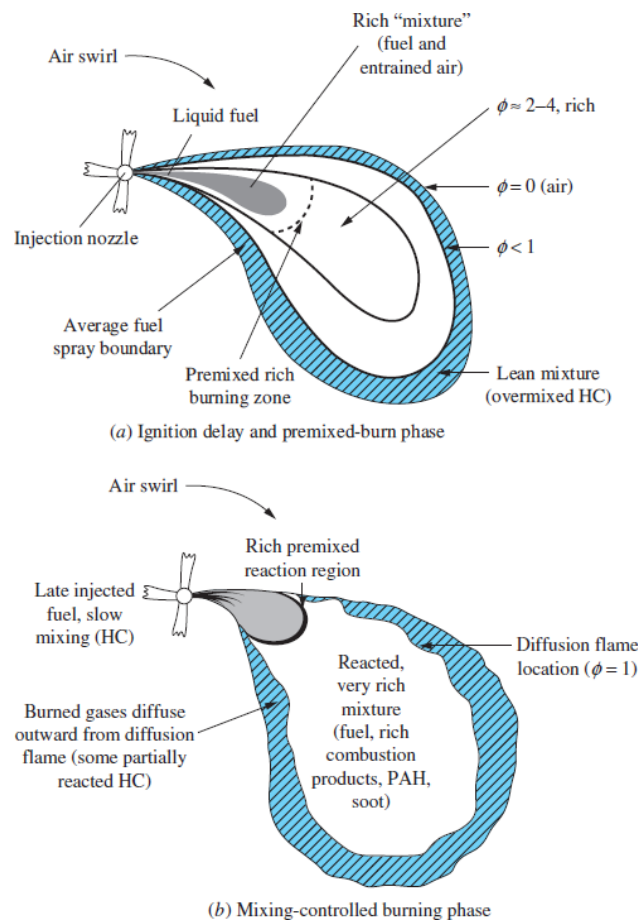
Diesel fuel contains hydrocarbon compounds with higher molecular weights, and hence higher boiling points, than gasoline. Also, substantial pyrolysis of fuel compounds occurs within the fuel sprays during the diesel combustion process. Thus, the composition of the unburned and partially oxidized hydrocarbons in the diesel exhaust is complex, different from that in the spark-ignition engine exhaust, and extends over a larger molecular size range. Gaseous hydrocarbon emissions from diesels are measured using a hot particulate filter (at 190°C) followed by a heated flame-ionization detector. The HC constituents vary from methane to the heaviest hydrocarbons, which remain in the vapor phase in the heated sampling line (which is also maintained at about 190°C). Any hydrocarbons heavier than this are condensed and, with the solid-phase soot, are filtered from the exhaust gas stream upstream of the detector. The particulate emission measurement procedure includes a portion of total engine hydrocarbon emissions. Particulates are collected by filtering from a diluted exhaust gas stream at a temperature of 52°C or less. Those hydrocarbons that condense at or below this temperature are absorbed onto the soot. They are the *extractable fraction* of the particulate: that is, that fraction which can be removed by a powerful solvent, typically between about 15 and 45% of the total particulate mass. This section discusses gaseous hydrocarbon emissions; particulate emissions—soot and extractable material—are discussed in [Sec. 11.5](#).

The complex heterogeneous nature of diesel combustion, where fuel evaporation, fuel vapor, air, and burned gas mixing, and rich premixed combustion followed by diffusion flame burn up, can occur in sequence and simultaneously, has been discussed extensively in [Chap. 10](#). As a result of this complexity, there are several processes that could contribute to diesel

engine hydrocarbon emissions. In [Chap. 10](#) the combustion process was divided into four stages: (1) the ignition delay which is the time between the start of injection and autoignition; (2) the premixed or rapid combustion phase, during which much of the fuel that has mixed in each spray with entrained air during the delay period then partially burns; (3) the mixing-controlled diffusion-flame combustion phase, during which the rate of burning depends on the rate of fuel-air mixing; and (4) the late combustion phase where heat release continues at a low rate governed by the mixing of residual combustibles with excess oxygen and the kinetics of the oxidation process. There are two primary paths by which hydrocarbons can escape this diesel combustion process unburned: the fuel-air mixture can become too lean to autoignite or to support a propagating flame at the conditions prevailing inside the combustion chamber: or, during the primary combustion process, the fuel-air mixture may be too rich to ignite. This fuel can then be consumed only by slower thermal oxidation reactions later in the expansion process which, because the gas temperatures are lower, are significantly slower. Thus, hydrocarbons remain unconsumed due to quenching of the oxidation process.^[5]

[Figure 11.36](#) illustrates schematically how these processes can produce incomplete combustion products. Fuel injected during the ignition delay ([Fig. 11.36a](#)) will vaporize and mix within each fuel spray with air entrained into the spray, producing rich fuel vapor/air mixture ($\phi \approx 2-4$) in the bulk of the spray. Around the edges of these turbulent fuel sprays, more air is available and locally some fuel will "overmix" with air producing lean mixture that will be difficult to autoignite and burn at conditions prevailing inside the combustion chamber (though some of this mixture may burn later if it mixes with high-temperature burned products early enough in the expansion stroke). In the "premixed" burning phase which follows the delay, ignition occurs in the bulk of the spray where the rich mixture is now hot enough to autoignite and release a portion of its chemical energy. Complete combustion of this over-rich mixture depends on further mixing with air in the diffusion flame surrounding each spray, or later still after mixing with already-burned gases before further expansion and cooling occurs. Of all these possible mechanisms, the overlean mixture path is believed to be the most important.²¹

Figure 11.36 Schematic representation of the diesel HC formation mechanism identifying its several components: (a) during the ignition delay and premixed burning phase; (b) during the mixing-controlled diffusion-flame burning phase.



For the fuel injected after the ignition delay period is over (Fig. 11.36b), rapid oxidation of the partly-burned fuel and products of fuel pyrolysis, as these mix with air in the diffusion flame that surrounds each spray, results in complete combustion. However, slow mixing of fuel and pyrolysis products with air, resulting in over-rich mixture or quenching of the combustion reactions, can also result in incomplete combustion products, pyrolysis products, and unburned fuel being present in the exhaust.²¹

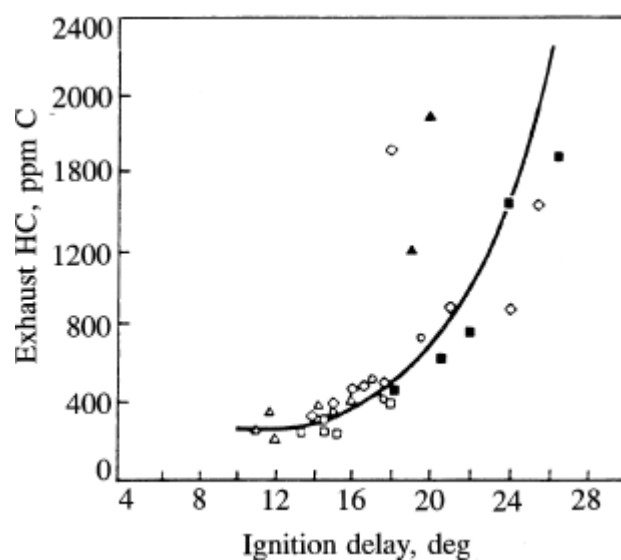
Hydrocarbon emission levels from diesels vary widely with operating conditions, and the different HC formation mechanisms are likely to be more or less important at different operating modes. Engine idling and light-load operation produce significantly higher hydrocarbon emissions than full-load operation. However, when the engine is overfueled, HC emissions increase very substantially. As is explained more fully below, overmixing (overleaning) is an important source of HC, especially under light-load operation. Undermixing, resulting in over-rich mixture during the combustion period, is the mechanism by which some of the fuel remaining in the injector-nozzle sac volume escapes combustion, and is also the cause of the high HC emissions during overfueled. Wall temperatures affect HC emissions suggesting that wall quenching of the combustion chemistry may be important, also.

11.4.4.2. Overleaning

As soon as fuel injection into the cylinder commences, a distribution in fuel/air equivalence ratio across each fuel spray develops. The amount of fuel that is mixed leaner than the lean combustion limit ($\phi \approx 0.3$) increases rapidly with time, as illustrated in Fig. 11.36a. In each fuel spray core, the mixture is significantly rich of stoichiometric. However, fuel close to the

spray boundary has already mixed beyond the lean limit of combustion and will not autoignite. This mixture can only oxidize by relatively slow thermal-oxidation reactions which may not be complete. Within this region, unburned fuel, fuel decomposition products, and partial oxidation products (aldehydes and other oxygenates) will exist; some of these will escape the cylinder without being burned. The magnitude of the unburned HC from these overlean regions will depend on the amount of fuel injected during the ignition delay, the mixing rate with air during this period, and the extent to which prevailing cylinder conditions are conducive to autoignition. A correlation of unburned HC emissions with the length of the ignition delay would be expected. The data in Fig. 11.37 from a direct-injection naturally-aspirated engine show that a good correlation between these variables exists. As the delay period increases beyond its minimum value (due to changes in engine operating conditions) the HC emissions increase at an increasing rate.⁶⁰ Thus, overleaning of fuel injected during the ignition delay period is a significant source of hydrocarbon emissions, especially under conditions where the ignition delay is long.

Figure 11.37 Correlation of exhaust HC concentration with duration of ignition delay for DI diesel engine. Various fuels, engine loads, injection timings, boost pressures, at 2800 rev/min.⁶⁰

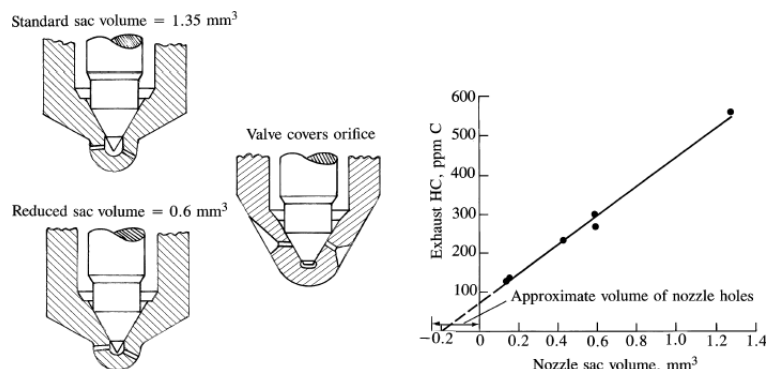


11.4.4.3. Undermixing

Two sources of fuel which enter the cylinder during combustion and which result in HC emissions due to slow or under mixing with air have been identified. One is fuel that leaves the injector nozzle at low velocity, often late in the combustion process. The most important source here is the nozzle sac volume, though secondary injections can increase HC emissions if the problem is severe. The second source is the excess fuel that enters the cylinder under overfueling conditions.

At the end of the fuel-injection process, the injector sac volume (the small volume left in the tip of the injector after the needle seats) is left filled with fuel. As the combustion and expansion processes proceed, this fuel is heated and vaporizes, and enters the cylinder at low velocity through the nozzle holes. This fuel vapor (and perhaps drops of fuel also) will mix relatively slowly with air and may escape the primary combustion process. Figure 11.38 shows HC emissions at the minimum ignition delay for a direct-injection diesel engine as a function of sac volume, along with drawings of some of the injector nozzles used. The correlation between HC emissions (under conditions when the overleaning mechanism is least significant) and sac volume is striking. Not all the fuel in the sac volume is exhausted as unburned hydrocarbons. For example, in Fig. 11.38, a volume of 1 mm³ gives 350 ppm C₁, while 1 mm³ of fuel would give 1700 ppm C₁. The sac volume may not be fully filled with fuel. Also, the higher-boiling-point fractions of the fuel may remain in the nozzle. Significant oxidation may also occur. The sac volume in current production nozzles helps to equalize the fuel pressures immediately upstream of the nozzle orifices. A small sac volume makes this equalization less complete, and exhaust smoke levels deteriorate.

Figure 11.38 Effect of nozzle sac volume on exhaust HC concentration, DI diesel engine, at minimum ignition delay. $V_d = 1$ liter/cylinder, 1700 to 2800 rev/min.⁶⁰



In DI engines, black exhaust smoke limits the full-load equivalence ratio to about 0.7. Under transient conditions as the engine goes through an acceleration process, overfueling can occur. Even though the overall equivalence ratio remains lean, locally over-rich conditions may persist through the expansion stroke and into the exhaust process. HC emissions are unaffected by an increasing equivalence ratio until a critical value of about 0.9 is reached when levels increase dramatically. This mechanism is not significant under normal operating conditions, but can contribute HC emissions under heavy acceleration conditions if overfueling occurs.

11.4.4.4. Quenching and Misfire

Hydrocarbon emissions have been shown to be sensitive to oil and coolant temperature. For example, when these temperatures were increased from 40 to 90°C in a DI diesel, HC emissions decreased by 30%. Since the ignition delay was maintained constant, overmixing phenomena should remain approximately the same. Thus, wall quenching of the flame may also be a significant source of HC, depending on the degree of spray impingement on the combustion chamber walls.

While cycle-by-cycle variation in the combustion process in diesel engines is generally much less than in spark-ignition engines, it can become significant under adverse conditions such as low compression temperatures and pressures, and retarded injection timings. Substantial variations, cycle-by-cycle, in HC emissions are thought to result. In the limit, if misfire (no combustion) occurs in a fraction of the operating cycles, then engine HC emissions rise as the percentage of misfires increases. However, complete misfires in a well-designed and adequately controlled diesel engine are unlikely to occur within its normal operating range.²¹

Engine-out hydrocarbon emissions from diesel engines are low (of order 0.1% of the fuel) over much of the engine's operating range. However, at very light loads (especially engine idle), HC emissions as a fraction of the fuel can be several times higher. The very lean in-cylinder mixture at these loads results in a much larger percentage of the injected fuel overmixing and, as a result, not burning. Table 11.11 provides two examples of HC and particulate emissions (and the volatile organic fraction of that particulate extracted with a solvent—heavier HC that condense in the exhausting and sampling systems) for a heavy-duty 12-liter and 1.9-liter light-duty diesel engine. Both engines were turbocharged and intercooled. The light-duty passenger car engine was fitted with a diesel oxidation catalyst (DOC; see Sec. 11.6.3), which significantly reduces both the vehicle-out HC emissions and the fraction of the particulate that is absorbed volatile organic compounds. The heavy-duty diesel engine, at idle (mode 4) emits about five times the HC as a fraction of the fuel than at low to mid loads. But the fuel injected per cycle is significantly lower (about 2 kg/h for mode 4, compared to 15 to 25 for the other modes). The oxidation catalyst used with the light-duty diesel engine is effective at reducing the HC emissions above about 40% of the maximum torque, but much less effective at light load and idle (modes 4 and 5) when the catalyst temperature is significantly lower. The catalyst also reduces the soluble (volatile) organic fraction of the particulate emissions significantly.⁶¹

Table 11.11 HC, particulate, and soluble organic fraction emissions at different speeds and loads⁶¹

Mode	Speed, rev/min	Load, %	HC, g/kg fuel	TPM, g/kg fuel	SOF, % TPM
Heavy-duty diesel engine*					
1	1700	31	0.94	1.07	32
2	1400	17	1.44	0.87	43
3	1200	36	0.65	0.46	43
4	800	Min.	5.64	1.25	74
Light-duty diesel engine (with oxidation catalyst) [†]					
1	2620	36	0.17	0.50	12
2	2500	76	0.06	0.98	31
3	1900	50	0.07	0.60	6
4	1400	20	1.11	0.38	12
5	1200	Min.	2.44	0.53	22
HC = hydrocarbons; TPM = total particulate mass; SOF = soluble organic fraction of particulate.					
*Six-cylinder 12-liter turbocharged aftercooled DI diesel engine rated at 312 kW at 2100 rev/min, peak torque 2237 Nm at 1200 rev/min.					
[†] Four-cylinder 1.9-liter turbocharged aftercooled DI diesel engine rated at 66 kW. Exhaust includes a diesel oxidation catalyst.					

While basic diesel engine-out HC emissions levels are low, multiple fuel-injection strategies are now being used which, with their several start and end-of-injection processes, will complicate and likely worsen HC emissions (see [Sec. 10.7.1](#)). Also, the advanced diesel combustion modes being explored and developed, where "premixing" of fuel and air well before the normal combustion process occurs, has significantly lower engine-out NO_x emissions but several times higher HC and CO emissions than a conventional diesel (see [Sec. 10.7.2](#)). Also, as exhaust particulate traps, NO_x reduction catalysts, and DOCs are now part of the total diesel engine's emission control system, the hydrocarbons in the exhaust gases play a role in the operation of these exhaust treatment devices. The HC can act as a source of thermal energy as they react in these devices: they also can act as a reducing agent as NO_x is reduced to N₂, or as a source of active species as they oxidize. Depending on the exhaust system devices and regeneration strategies utilized, the HC concentrations may be deliberately increased (e.g., by post-main and/or late fuel injections) to enhance emissions control.

[1] This is because the standard detection instrument, a flame ionization detector (FID), is effectively a carbon atom counter: for example, one propane molecule generates three times the response generated by one methane molecule. Some data in the literature are presented as ppm propane (C₃), or ppm hexane (C₆); to convert to ppm C₁ multiply by 3 or by 6, respectively. Older measurements of hydrocarbon emissions were made with nondispersive infrared (NDIR) detectors which had different sensitivities for the different hydrocarbon compounds. For gasoline-fueled engines, HC emissions determined by FID analyzers are about twice the levels determined by NDIR analyzers,²⁸ though this scaling is not exact.

[2] About half of these HC are fuel compounds; about half are partially reacted and fragmented fuel compounds.

- [3] Note that exhaust valve leakage is a different type of "source" than the others. It allows HC to bypass the in-cylinder oxidation process; this source is normally small.
- [4] Note that there is a significant variation in the temperature of the exhaust gases during the exhaust process. The gas exhausted first is about 100 K hotter than the gas exhausted at the end of the process (see [Sec. 6.5](#)).
- [5] Note that under normal engine operating conditions, the diesel engine combustion inefficiency is small, about 1%; see [Sec. 4.9.4](#) and [Fig. 3.9](#).

11.5. PARTICULATE EMISSIONS

Particulate emissions have long been an issue with diesels. In such engines, as described in [Chap. 10](#), soot is formed within the evolving fuel sprays, and while much of the soot burns up inside the cylinder, a portion of it does not and is exhausted. Then, in the exhaust and during dispersion of the gases exiting the exhaust pipe, the gases cool, nucleate condense, and form small particles. As higher-molecular-weight hydrocarbons continue to condense, they are absorbed into the soot particles. Thus a size distribution of particulates builds up: the larger particles are categorized as *accumulation mode* (particles in the diameter range of 50 to 500 nm); the small particles below about 50 nm in diameter constitute the *nucleation mode*. With port-fuel-injection gasoline engines, the particulate emissions are much lower than in diesels: by a factor of about 70 on a mass basis. However, direct-injection gasoline engine particulate emissions are close to a factor of 10 higher than those from port-fuel-injection equivalent engines, again by mass.

Particulate traps located in the exhaust system of diesel and direct-injection gasoline engines have been developed, and are widely used in diesels. These traps remove almost all the particulates from the exhaust gases using the thin walls of porous ceramic monoliths as a filter. Over time the pressure drop across the particulate trap increases, and the ceramic filter must be regenerated. Usually this requires that its temperature be increased so the particulate layer on the filter burns off. See [Sec. 11.6.4](#) for a more detailed discussion of particulate traps.

11.5.1. Spark-Ignition Engine Particulates

Traditionally, three classes of spark-ignition engine particulate emissions have been defined: inorganic particles (of which lead compounds used to dominate), organic particulates (including soot), and sulfates.

Depending on the sulfur content in the fuel, significant sulfate emissions can occur with oxidation-catalyst equipped engines. Gasolines can contain from 10 up to some 500 ppm by weight sulfur (depending on the extent of sulfur removal in the refinery), which is oxidized within the engine cylinder to sulfur dioxide. This SO_2 can be oxidized by the exhaust catalyst to SO_3 which combines with water at ambient temperatures to form a sulfuric acid aerosol. Levels of sulfate emissions depend on the fuel sulfur content, the operating conditions of the engine, and the details of the catalyst system used. Low sulfur levels in gasolines are important to maintain high exhaust catalyst efficiencies over extended use.

For automobile engines operated with leaded gasolines (which contain about 0.15 g Pb/liter), the particulate emission rates are typically 100 to 150 mg/km. This particulate is dominated by lead compounds: 25 to 60% of the emitted mass is lead.⁶² The particulate emission rates are considerably higher when the engine is cold, following start-up. The exhaust temperature has a significant effect on emission levels. The particle size distribution with leaded fuel is about 80% by mass below 2 μm diameter and about 40% below 0.2 μm diameter. Most of these particles are presumed to form and grow in the exhaust system due to vapor phase condensation enhanced by coagulation. Some of the particles are emitted directly, without settling. Some of the particles either form or are deposited on the exhaust system walls. Many of these are removed when the exhaust flow rate is suddenly increased and, together with rust and scale, account for the increase in mass of particles emitted during accelerations. Only a fraction of the lead consumed in the fuel is exhausted, the remainder being deposited within the engine and exhaust system.

Current port-fuel-injected unleaded-gasoline SI engines have particulate emissions of a few mg/km in automobiles without catalysts. Over the next decade, particulate emissions requirements are likely to be reduced by up to an order of magnitude. The particulate is primarily condensed organic materials, though soot emissions (black smoke) can result from the burn up of overly rich mixtures. Direct-injection gasoline engines have higher particulate emissions than do port-injected engines by up to a factor of 10 on a mass basis, and are similarly higher based on particulate number density. Stratified charge direct-injection engines have much higher particulates (more nucleation particles) than do homogeneous stoichiometric DI engines. The following example illustrates the various ways that are used to characterize a vehicle's (and engine's) particulate

emissions: for example, mass and number of particles emitted per unit distance traveled, particulate number density in the exhaust, particulate emissions index [Eq. (2.44)] g PM/kg fuel.

Example

EXAMPLE 11.3

Particulate emissions are expressed in many ways: mg particulate matter (PM)/km, particle number density PN/cm³, g PM/kg fuel, #PN/km. The 2014–2017 European PN limit for passenger vehicles is 6×10^{11} #/km: convert this particulate emissions level to the other parameters listed above.

Needed inputs are:

- Vehicle fuel economy/consumption 35 miles per gallon gasoline is equivalent to $35 \times 0.42 = 15$ km/liter which corresponds to 6.7 liters/100 km. The fuel mass--consumption rate is 50 g/km or 80 g/mile. The density of gasoline is 0.75 kg/liter = 0.75 g/cm³.
- Stoichiometric gasoline/air mixture: thus the mass of air + fuel (equals the mass of exhaust), relative to mass of fuel, is

$$(m_a + m_f) / m_f = 15.6$$

The density of exhaust gas at standard pressure and temperature ($\approx$ density of air) is 1.2 kg/m³.

- Density of PM (close to the density of soot) is about 2 g/cm³. The average particle size is 80 nm diameter.

Starting with 6×10^{11} #PN/km, find #PN/cm³ of exhaust. The multiplication sequence is:

$$6 \times 10^{11} \left(\frac{\#}{\text{km}} \right) \times \frac{100}{6.7} \left(\frac{\text{km}}{\text{lit. fuel}} \right) \times \frac{1}{10^3} \left(\frac{\text{lit}}{\text{cm}^3} \right) \times \frac{1}{0.75} \left(\frac{\text{cm}^3 \text{fuel}}{\text{g fuel}} \right) \times \frac{1}{15.6} \left(\frac{\text{g fuel}}{\text{g exh}} \right) \times 1.2 \times 10^{-3} \left(\frac{\text{g exh}}{\text{cm}^3 \text{e.}} \right) \\ = \frac{9.2 \times 10^5 \# \text{PN} / \text{cm}^3 \text{exh}}{\quad}$$

To convert 6×10^{11} #PN/km to mg PM/km, the multiplication sequence is:

$$6 \times 10^{11} \left(\frac{\#}{\text{km}} \right) \times \frac{4}{3} \pi \left(\frac{80 \times 10^{-7}}{2} \right)^3 (\text{cm}^3) \times 2 \left(\frac{\text{gPM}}{\text{cm}^3 \text{PM}} \right) \times 10^3 \left(\frac{\text{mg}}{\text{g}} \right) = \frac{0.32 \text{mgPM/km}}{\quad}$$

To convert 0.32 mg PM/km to mg PM/g fuel, the multiplication sequence is:

$$0.32 \left(\frac{\text{mg PM}}{\text{km}} \right) \times \frac{100}{6.7} \left(\frac{\text{km}}{\text{lit fuel}} \right) \times \frac{1}{0.75} \left(\frac{\text{cm}^3 \text{fuel}}{\text{g fuel}} \right) \times 10^3 \left(\frac{\text{lit}}{\text{cm}^3} \right) = \frac{0.0064 \text{mgPM/g fuel}}{\quad}$$

With injection of liquid fuel directly into the cylinder, there are two key questions that impact particulate formation: (1) Is there significant liquid fuel wetting of the piston crown, cylinder head and liner? (2) Is there time for essentially complete evaporation of the fuel? The questions are coupled since slower evaporation results in greater penetration of drops in the fuel spray: with the resulting increase in wall wetting, the evaporation rate and timing become even more critical. The solid component of the particulate is largely soot which forms as gasoline vapor evaporates from the liquid-fuel wall films and

mixes with the hot burned gases (which contain modest amounts of oxygen) thus forming hot enough very fuel-rich regions—the conditions under which soot forms. And, as the engine overall shifts to rich of stoichiometric, accumulation mode PM increases since, with more fuel and less air, these rich hydrocarbon-vapor/burned gas zones are more extensive and richer. As fuel injection timing shifts later from the intake stroke (homogeneous DI) to during the compression stroke (lean stratified DI operation), there is less time for evaporation and mixing, and more wall wetting, so accumulation mode PM increases. As ethanol is added to the gasoline fuel, additional oxygen is available in these rich vaporized-fuel burned-gas regions where soot forms, so its formation rate is reduced. Also, the period during and after engine start up would be expected to be a period when particulate formation rates are higher.

The nucleation mode particulate mostly come from higher molecular weight hydrocarbon components in the gasoline fuel which condense during the exhaust process, and from soot precursor nuclei which are quenched before significant agglomeration into larger particles can occur. (See [Sec. 11.5.4](#) on soot formation mechanisms.)

11.5.2. Characteristics of Diesel Particulates

11.5.2.1. Measurement Techniques

Diesel particulates consist principally of combustion-generated carbonaceous material (soot) on which organic compounds in the exhaust gases have been absorbed. Most particulate material results from incomplete combustion of fuel hydrocarbons; some is contributed by the lubricating oil. The engine-out emission rates are typically 0.03 g/km for light-duty diesels in an automobile. In larger heavy-duty direct-injection engines, particulate emission rates are 0.5 to 1.5 g/brake kW · h. The composition of the particulate material depends on the conditions in the engine exhaust and particulate collection system. At temperatures above 500°C, the individual particles are principally clusters of many small spheres or spherules of carbon (with a small amount of hydrogen) with individual spherule diameters between 15 nm and 30 nm. As temperatures decrease below 500°C, the particles become coated with adsorbed and condensed high molecular-weight organic compounds which include: unburned hydrocarbons, oxygenated hydrocarbons (ketones, esters, ethers, organic acids), and polynuclear aromatic hydrocarbons. The condensed material also includes (depending on the sulfur content in the fuel) species such as sulfur dioxide, and sulfuric acid (or sulfates).

The objective of most particulate measurement techniques is to determine the amount of particulate being emitted to the atmosphere. Techniques for particulate measurement and characterization range from simple smoke meter opacity readings to analyses using exhaust dilution tunnels. Most techniques require lengthy sample-collection periods because the particulate emission rate is usually low. The physical conditions under which particulate measurements are made are critical because the emitted species are unstable and may be altered through loss to surfaces, change in size distribution (through agglomeration), and chemical interactions among other species in the exhaust, at any time during the measurement process (including sampling, storage, or examination). The most basic information is normally obtained on a gravimetric or mass basis: a known fraction of the exhaust flow is sampled from the diluted exhaust; the particulate matter is filtered and weighed (under carefully controlled conditions). Emissions are then expressed in grams per kilometer for a vehicle, grams per kilowatt-hour for an engine, grams per kilogram of fuel or milligrams per cubic meter of exhaust (at standard conditions). Smoke meters measure the relative quantity of light that passes through the exhaust or the relative reflectance of particulate collected on filter paper. They do not measure mass directly. They are used to determine visible smoke emissions and provide an approximate indication of mass emission levels. Visible smoke from heavy-duty diesels at high load is regulated. In the standard mass emission measurement procedure, dilution tunnels are used to simulate the physical and chemical processes the particulate emissions undergo as they enter the atmosphere. In the dilution tunnel, the raw exhaust gases are diluted with ambient air to a temperature of 52°C or less, and a sample stream from the diluted exhaust is filtered to remove the particulate material.

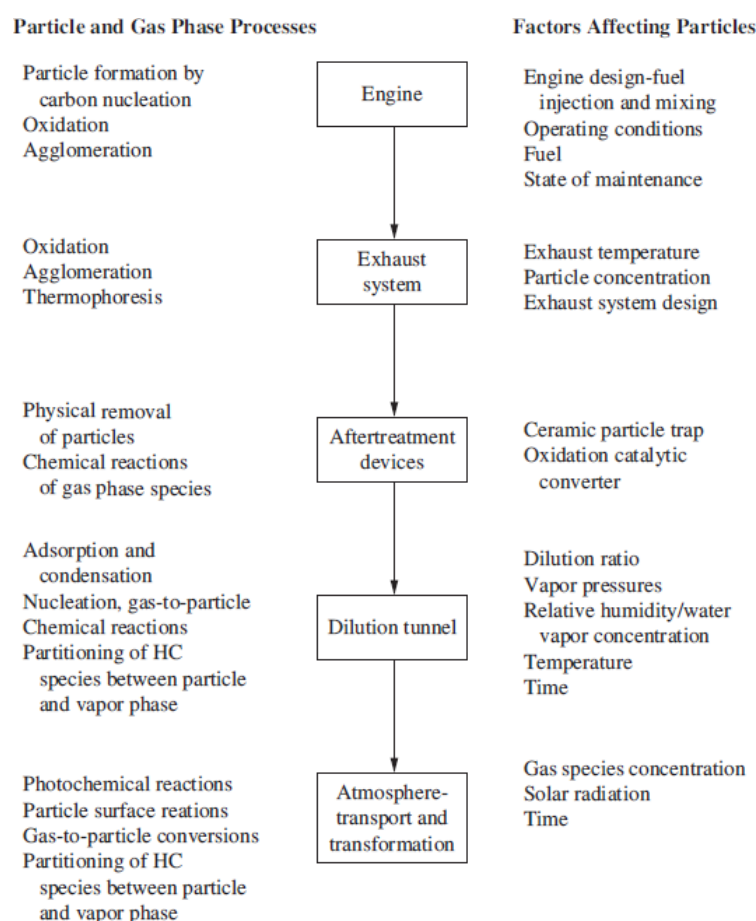
Particulate counting, as part of efforts to reduce the number of fine particles emitted from diesel engines can be done with several instruments. The Condensation Particulate Counter (CPC) has been chosen for determining the number of particles

for compliance with regulation. Particle size and distribution can be obtained with several devices: for example, a Scanning Mobility Particle Sizer (SMPS), an Electrical Low Pressure Impactor (ELPI), and a Differential Mobility Spectrometer (DMS). The *Bosch Automotive Handbook* briefly summarizes these several particulate measurement approaches and devices.⁶³

11.5.2.2. Particulate Composition and Structure

As mentioned above, the particle size distribution in diesel exhaust is bi-modal with a nucleation mode (less than about 50 nm in diameter) and an accumulation mode (from about 50 to 500 nm diameter). [Figure 11.39](#) is a schematic of the particle and gas phase processes, and the factors affecting these processes, from the start of combustion to exhaust discharge into the atmosphere. Dispersion into the atmosphere, and transformation, then continue. Note that in today's engines, this sequence includes the in-cylinder combustion process, what occurs in the exhaust system and the aftertreatment devices (particulate filter and diesel oxidation catalyst) it contains, dilution with air and cooling as the exhaust gases exit the exhaust pipe, and diffusion and continuing dilution in the atmosphere.

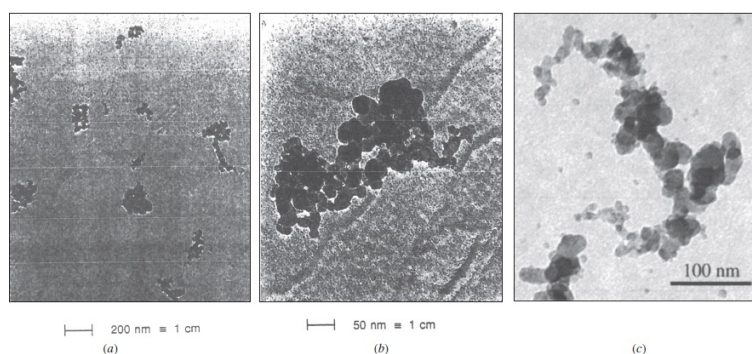
Figure 11.39 Schematic of the particle and gas phase processes and factors effecting the particle size distribution and chemical composition of diesel engine particulate matter emissions into the atmosphere.⁶⁴



The use of a dilution tunnel to characterize the amount and composition of the particulates in the exhaust gas is meant to simulate in a controlled way the actual mixing, cooling and dispersion that occurs in practice as the engine's particulate laden exhaust gases enter the atmosphere. Examples of these particulates are shown in [Fig. 11.40](#). In [Fig. 11.40a](#) and [b](#), the gases were sampled from the combustion chamber at 150° ATC when the exhaust valve is opening. The samples are seen to consist of collections of primary particles (spherules) of 10 to 25 nm diameter, agglomerated into aggregates (hereafter called particles). Individual particles range in size and appearance from clusters to chains of spherules. During combustion, the

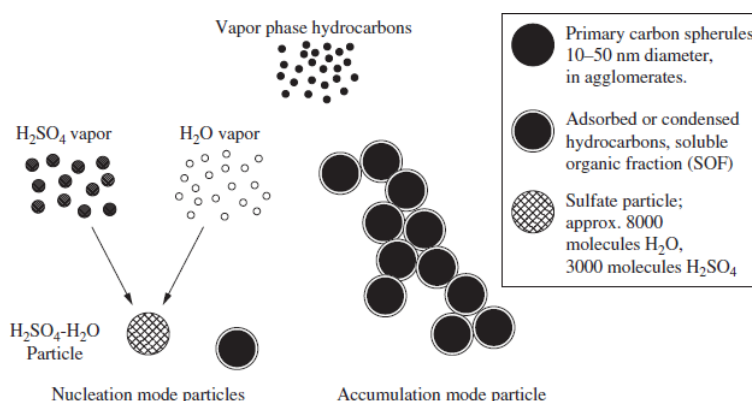
sampled aggregate particles are significantly larger: by EVO their overall size has been reduced yet the agglomerated spherule structure persists. At EVO clusters may contain some 20 to 100 or so spherules. The spherules are combustion-generated soot particles mostly in the 10- to 25-nm size range.⁶⁵ Figure 11.40c shows a typical accumulation-mode particle and several much smaller, spherical, nucleation-mode particles. These were obtained by sampling from a diesel engine exhaust, then diluted, and filtered. An accumulation particle, a chain-like collection of many soot spherules, and several small nucleate mode aerosol particles are illustrated.⁶⁶

Figure 11.40 Photomicrographs of diesel particulates sampled from the in-cylinder gases, or the diluted exhaust. In-cylinder samples extracted at 150°C ATC (near EVO) at 2000 rev/min, 5 bar bmep, 31% EGR. Magnification: (a) 50k and (b) 200k.⁶⁵ (c) Sample from engine exhaust, then diluted and collected on transmission electromicroscopy grid: shows chain-like accumulation-mode particle and several small spherical nucleation-mode particles.⁶⁶



Determination of the particle size distribution from such photomicrographs involves assigning a representative dimension to such complex and irregular aggregates, and introduces uncertainties arising from only having two-dimensional images of particles available. Other approaches based on inertial impactors and electrical aerosol analyzers have been used. The data indicate that the particle size distribution is bimodal. The smaller-size nucleation mode is composed of liquid hydrocarbon drops, individual spherules, and sulfate ($\text{H}_2\text{O}-\text{H}_2\text{SO}_4$) nuclei. The larger-size accumulation mode is of agglomerated spherule particles characterized by number-mean diameters of 100 to 150 nm. Figure 11.41 illustrates this discussion of the different components of diesel exhaust particulates. An additional feature that it adds is that higher molecular weight hydrocarbons are absorbed onto individual spherule and agglomerated particle surfaces. These hydrocarbons range from about C_{20} to C_{30} . Condensed hydrocarbons and sulfates, and individual soot spherules, comprise the small, nucleation mode, particles.⁶⁵

Figure 11.41 Schematic illustration of components of diesel particulate matter: accumulation mode; largely soot spherules, which agglomerate to form larger particles: nucleation mode; condensed hydrocarbon particles, individual soot spherules, and sulfate particles; these compounds are also absorbed onto the agglomerates.⁶⁵



The exhaust particulate is usually partitioned with an extraction solvent into a soluble fraction and a dry-soot fraction. Two commonly used solvents are dichloromethane and a benzene-ethanol mixture. Typically 15 to 30 mass percent is extractable, though the range of observations is larger. Thermogravimetric analysis (weighing the sample as it is heated) produces comparable results. Typical average chemical compositions of the two particulate fractions are given in Table 11.12. Dry soot has a much lower H/C ratio than the extractable material. Although most of the particulate emissions are formed through incomplete combustion of fuel hydrocarbons, engine oil may also contribute significantly. The number-average molecular weight of the extractable material shown in Table 11.12 ranged from about 360 to 400 for a variety of engine conditions. This fell between the average molecular weight of the fuel (199) and that of the lubricating oil (443 when fresh and 489 when aged).⁶⁷ All of the oil contribution appeared in the extractable material. The contributions from the different individual compounds in the fuel have also been studied. All the compounds examined—paraffins, olefins, and aromatics—contributed to the particulate emissions; as a group, aromatics were the greatest contributors.

Table 11.12 Chemical composition of particular matter⁶⁷

	Idle	48 km/h
Extractable composition	$C_{23}H_{29}O_{4.7}N_{0.21}$	$C_{24}H_{30}O_{2.6}N_{0.18}$
H/C	1.26	1.63
Dry-soot composition	$CH_{0.27}O_{0.22}N_{0.01}$	$CH_{0.21}O_{0.15}N_{0.01}$
H/C	0.27	0.21

In addition to the elements listed in Table 11.12, trace amounts of sulfur, zinc, phosphorus, calcium, iron, silicon, and chromium have been found in particulates. Sulfur and traces of calcium, iron, silicon, and chromium are found in diesel fuel: zinc, phosphorus, and calcium compounds are used in lubricating oil additives.⁶⁷ The sulfate (SO_4) associated with diesel particulate matter is collected during dilution tunnel sampling. It originates with the sulfur in the fuel, which is oxidized to SO_2 , then to SO_3 , which combines with H_2O to form H_2SO_4 vapor. Depending on local conditions (temperature and concentration/partial pressure) as the exhaust is diluted and cooled, multiple H_2O and H_2SO_4 vapor molecules collide and form molecular clusters which may then further grow. Alternatively, depending on local gas conditions, the molecular clusters of H_2O and H_2SO_4 may not have sufficient energy to be stable and eventually break up. Whether or not sulfate nucleation mode particles form depends on the sulfur level in the fuel and the thermodynamic state of the exhaust and the kinetics of the system.⁶⁴ In regions of the world where controlling vehicle emissions has made the greatest progress, fuel sulfur levels are low (below 15 ppm S by weight). Historically sulfur levels in distillate diesel fuels have been much higher (up to 5000 ppm and higher), especially in developing world regions. Fuel sulfur levels are, over time, being reduced.^[1]

A lattice image of a portion of a diesel particle is shown in Fig. 11.42; it suggests that a concentric lamellate structure arranged around the center of each spherule. This arrangement of concentric lamellas is similar to the structure of carbon black. This is not surprising; the environment in which diesel soot is produced is similar to that in which oil furnace blacks are made. The carbon atoms are bonded together in hexagonal face-centered arrays in planes, commonly referred to as platelets. As illustrated in Fig. 11.43, the mean layer spacing is 0.36 nm (only slightly larger than graphite). Platelets are arranged in layers to form crystallites. Typically, there are two to five platelets per crystallite, and on the order of 10^3 crystallites per spherule. The crystallites are arranged with their planes more or less parallel to the particle surface. The spherules, 10 to 25 nm in diameter, are fused together to form particles as shown in Fig. 11.41. A single spherule contains 10^5 to 10^6 carbon atoms.^{68,69}

Figure 11.42 Lattice-imaging micrograph of spherules in an accumulation-mode diesel particle.⁶⁸

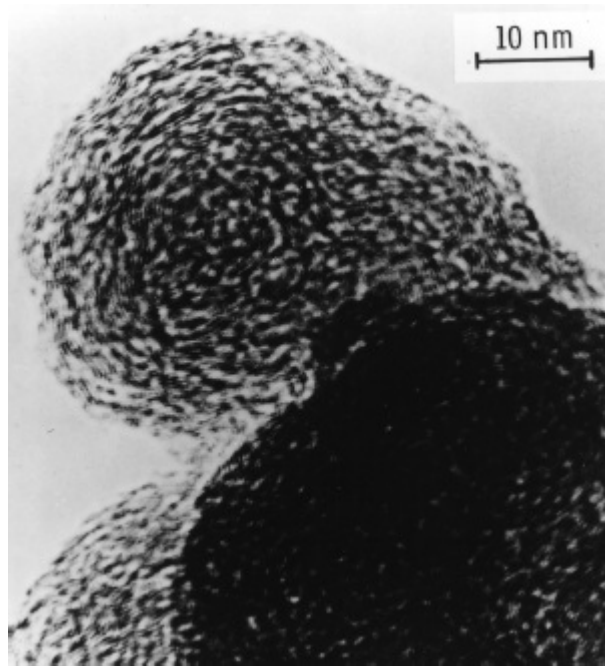
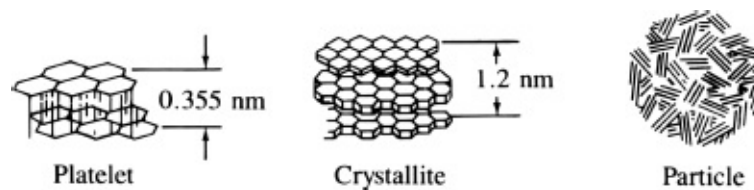


Figure 11.43 Substructure of carbon particles.⁶⁹



A surface area of about $200 \text{ m}^2/\text{g}$ has been measured for diesel soot. A smooth-surfaced 20-nm diameter sphere (a spherule with a density of 2 g/cm^3) would have a surface/mass ratio of $150 \text{ m}^2/\text{g}$, so the measured value is about 1.3 times the superficial area. Approximating a particle of agglomerated spherules by a single sphere of 200 nm diameter gives a surface/mass ratio of $15 \text{ m}^2/\text{g}$.⁶⁹ These data and estimates of superficial area per unit mass indicate that diesel soot has relatively low porosity.

11.5.3. Particulate Distribution within the Cylinder

As the above discussion indicates, the solid component (soot) forms first within the fuel-rich core of the individual fuel sprays, once combustion has been initiated. The soluble organic fraction (SOF) of the particulate (the condensed hydrocarbon component) and nucleation mode particles form later during the exhaust process, and as dilution with air occurs as the exhaust gases leave the tailpipe. Many measurements have been made of the particulate distribution within the combustion chamber of operating diesel engines. The results provide valuable information on the particulate formation and oxidation processes and how these relate to the in-cylinder fuel distribution and heat-release development. Techniques used to obtain particulate concentration data include: use of rapid-acting poppet or needle valves which draw a small gas sample from the combustion chamber at a specific location and time for analysis, optical absorption techniques, and cylinder dumping where the cylinder contents are rapidly emptied into an evacuated tank at a preset time in the cycle. Concentration data taken at specific locations within the combustion chamber during the engine cycle are not necessarily representative of the cylinder contents in general: nor do they represent the time history of a given mass of gas. The fuel distribution, mixing, and heat-release patterns in the cylinder are highly nonuniform during the soot-formation process.

In direct-injection diesel engines, the soot component of the particulate forms within each fuel spray where local average equivalence ratios are very rich (see Secs. 10.3.4 and 10.7.2). Soot concentrations rise rapidly soon after combustion starts. A large fraction of the fuel molecules in each fuel-rich spray core react to form polycyclic aromatics, which grow to form soot particles. The fraction of local fuel carbon found as soot in these very fuel-rich regions is high, some 50%. The sequences of physical and chemical processes that lead to soot formation are illustrated in Figs. 11.44 and 11.45. Figure 11.44 shows how soot develops in each individual fuel spray in a quiescent DI diesel combustion chamber. During the ignition delay period the spray penetrates into the combustion chamber, entraining hot end-of-compression air. The fuel droplets evaporate forming a fuel-rich ($\phi \approx 2-4$) mixture which, after a few crank-angle degrees, starts to react chemically, emit chemiluminescence, and release some 20% of its chemical energy. The autoignition chemistry breaks down the fuel molecules into polyacetylenes, which combine to form polycyclic aromatics. These molecules grow rapidly within the spray to ever larger molecules which form soot particle nuclei which grow into soot spherules and agglomerate into accumulation mode particles, as the mixture within the spray moves away from the nozzle. Thus, as shown, the soot concentration increases along the spray (now surrounded by a diffusion flame) from the rich premixed combustion zone some halfway along the spray toward the head. As fuel-rich, partly reacted, fuel-air mixture diffuses outward into the diffusion flame now surrounding each spray, the soot particles rise in temperature, and with the oxygen available from the air side of the flame, largely burn up.

Figure 11.44 Schematics illustrating the critical soot forming processes, and the regions where PAH and soot form in an individual diesel fuel spray during the ignition delay, premixed rapid burning phase, and mixing-controlled diffusion flame phase of combustion.⁷⁰

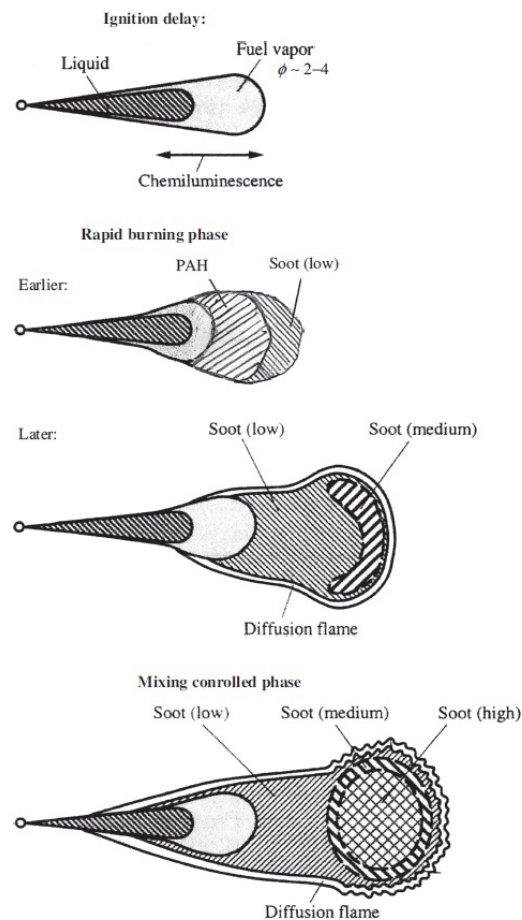


Figure 11.45 The sequence of chemical reactions that occur during fuel spray development, combustion, fuel reaction, and soot formation and oxidation, in a DI diesel engine fuel spray.⁷¹

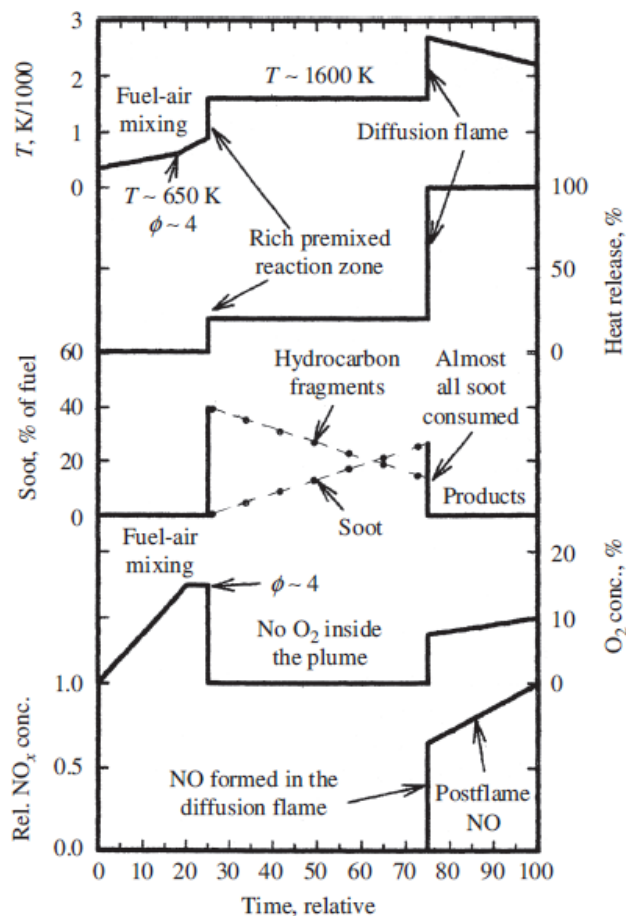


Figure 11.45 illustrates the chemical aspects of this sequence of processes occurring within each fuel spray. The premixed-reaction-zone fuel chemical-energy release (some 20% of the fuel's total energy) raises the gas temperature within the spray to about 1600 K as the fuel molecules fragment and (with many intermediate steps; see Sec. 11.5.4), form PAHs, followed by soot formation, in the fuel-rich within-the-spray mixture. This "rich-mixture" fuel (with its high soot content) then moves into the diffusion flame now surrounding each spray, is heated, mixes with oxygen, and releases almost all its as yet unreleased chemical energy. As this mixing-controlled diffusion-flame combustion process moves to completion, almost all the soot particles are burned up. As noted in Sec. 11.2.4, this is also the primary NO formation region.

11.5.4. Soot Formation Fundamentals

The soot particles, whose characteristics have been described in the above two sections, form primarily from the carbon in the diesel fuel. The formation process starts with a fuel molecule containing 12 to 22 carbon atoms with an H/C ratio of about 2, and ends up with particles typically a hundred nanometers or so in diameter, composed of spherules 10 to 25 nm in diameter each containing some 10^5 carbon atoms, and having an H/C ratio of about 0.1. Much of the information available on the fundamentals of soot formation comes from studies in simple premixed and diffusion flames, stirred reactors, shock tubes, and constant-volume combustion vessels. Also, the production of carbon black requires a high yield of soot from pyrolysis of a hydrocarbon feedstock, and the literature from that field has much to contribute. However, the characteristics of diesel combustion which make it different from these more basic studies—the high gas temperatures and pressures, complex fuel composition, importance of turbulent mixing, the unsteady nature of the process, and the three-dimensional geometry of fuel sprays—make it challenging to interpret fundamental ideas regarding soot formation in the diesel context. There is much about the soot formation process in diesel engines that is incompletely understood.

Soot formation takes place within the combusting fuel spray at temperatures of about 1600 K, pressures of 100 bar, and with much less air locally than is needed to fully burn the fuel. The time available for the formation of soot particles from the fuel hydrocarbons is of the order of milliseconds. The resulting aerosol—dispersed solid-phase particles in a gas—can be characterized by the total amount of condensed phase (often expressed as the soot volume fraction, F_V , the volume of soot/total volume), the number of soot particles per unit volume (N), and the size of the particles (e.g., average diameter d). F_V , N , and d are mutually dependent [e.g., for spherical particles $F_V = (\pi/6)Nd^3$], and any two of these variables characterize the system. It is most convenient to consider N and F_V as the independent variables since they each relate to the "almost-separate" stages of soot particle generation (the source of N) and soot particle growth (the source of F_V).

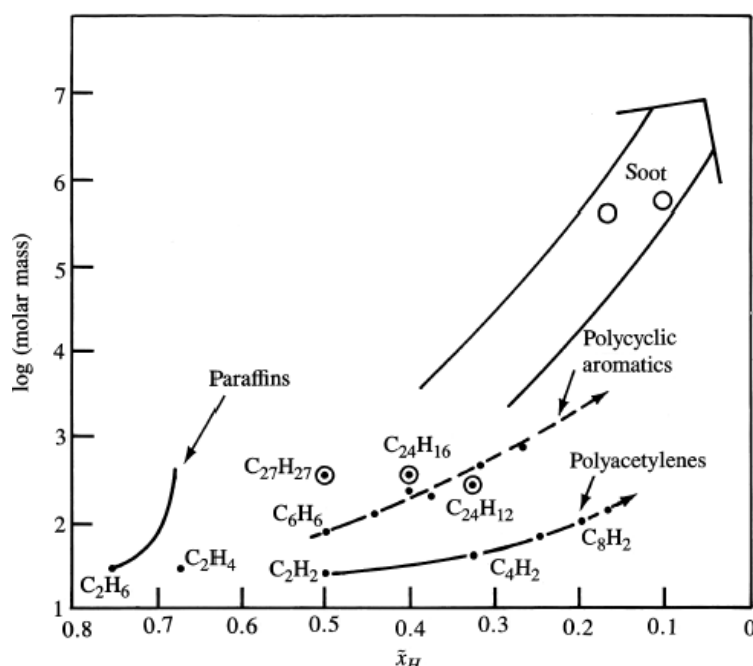
The formation process can be summarized as follows.⁷²

1. Particle formation, where the first condensed-phase material arises from the fuel molecules via their partial oxidation and/or pyrolysis products. These products typically include various unsaturated hydrocarbons, particularly acetylene and its higher analogues ($C_{2n}H_2$), and polycyclic aromatic hydrocarbons (PAH). These two types of molecules are considered the most dominant precursors of soot in flames. The condensation reactions of gas-phase species such as these lead to the appearance of the first recognizable soot particles (often called nuclei). These first particles are very small ($d < 2$ nm) and the formation of large numbers of them involve minimal soot loading in the region of their formation.
2. Particle growth, which includes both surface growth, coagulation, and aggregation. Surface growth, by which the bulk of the solid-phase material is generated, involves the attachment of gas-phase species to the surface of particles and their incorporation into the particulate phase. [Figure 11.46](#), where the log of the molecular weight of a species is plotted against its hydrogen mole fraction

$\tilde{x}_H$, illustrates some important points about this process. Starting with a fuel molecule where

$\tilde{x}_H \geq 0.5$, it is apparent that neither polyacetylene chain growth nor PAH growth would lead to soot particles which have $\tilde{x}_H$ in the range 0.1 to 0.2. What is required is condensation of species with the right hydrogen content, or condensation of species with higher hydrogen content followed by dehydrogenation, or a combination of both these processes. Obviously some polyacetylenes and some PAH can satisfy these requirements, as can saturated platelets (e.g., $C_{27}H_{27}$; see [Sec. 11.5.2](#)). Surface growth reactions lead to an increase in the amount of soot (F_V) but the number of particles (N) remains essentially unchanged. The opposite is true for growth by coagulation, where the particles collide and coalesce, which decreases N with F_V constant. Once surface growth stops, continued aggregation of particles into chains and clusters can occur.

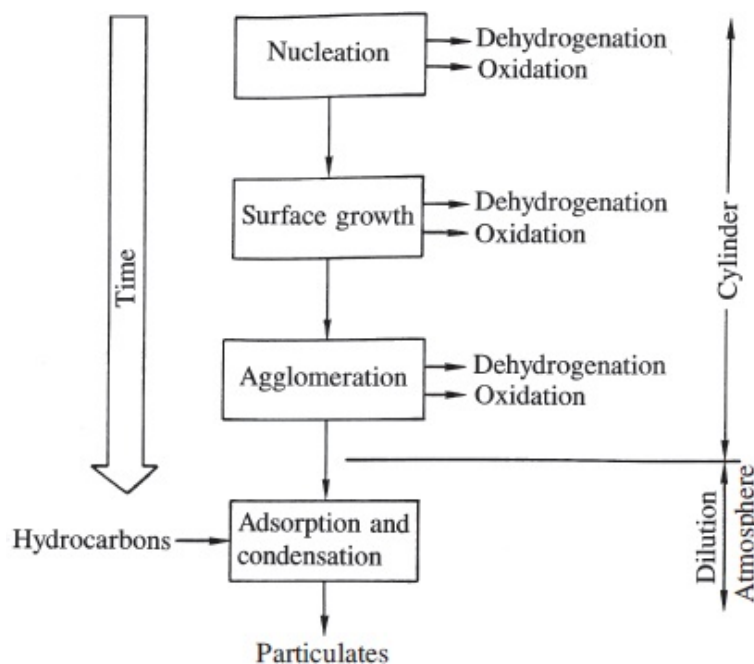
Figure 11.46 Paths to soot formation on plot of species molecular weight M versus hydrogen mole fraction $\tilde{x}_H$.⁷²



These stages of particle generation and growth constitute the soot formation process. At each stage in the process oxidation can occur, where soot or soot precursors are burned in the presence of oxidizing species to form gaseous products such as CO and CO₂. The eventual emission of soot from the engine will depend on the balance between these processes of formation and burnout. The emitted soot is then subject to a further mass addition process as the exhaust gases cool and are diluted with air. Adsorption into the soot particle surface and condensation to form new particles of hydrocarbon species in the exhaust gases occurs in the exhaust system and in the dilution tunnel which simulates what happens in the atmosphere.

Figure 11.47 illustrates the relationship between these processes.⁶⁷ Although they are illustrated as discrete processes, there is overlap, and they may occur concurrently in a given elemental mixture region within the diesel combustion chamber. Of course, also due to the non-homogeneous nature of the mixture and the duration of fuel injection and its overlap with combustion, at any given time different processes are in progress in different regions or packets of fluid. The fundamentals of each of these processes will now be reviewed.

Figure 11.47 Processes leading to net production of diesel particulates.⁶⁷



11.5.4.1. Soot Particle Formation

Empirically, it has been found useful to define the composition of the fuel-oxidizer mixture at the onset of soot formation in flames by the carbon/oxygen ratio. Equilibrium considerations indicate that soot (C_S) should form when, in the overall reaction



(11.35)

m becomes larger than $2y$: i.e., when the C/O ratio exceeds unity. The corresponding fuel/air equivalence ratio is given by

$$\phi = 2 \left(\frac{C}{O} \right) (1 + \delta)$$

(11.36)

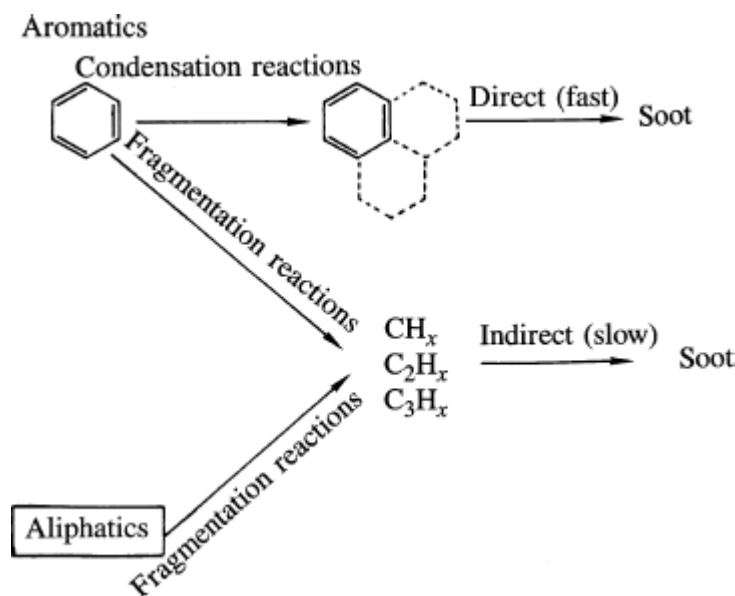
where $\delta = n/(4m)$: ϕ is 3 for $(C/O) = 1$, with $n/m = 2$. The experimentally observed critical C/O ratios are less than unity, however, varying with fuel composition and details of the experimental setup from about 0.5 to 0.8. The critical C/O ratio for soot formation increases with increasing temperature but is only weakly dependent on pressure. Beyond the carbon formation limit, the yield of soot increases rapidly with increasing C/O ratio and is strongly enhanced by increasing pressure.⁷²

It is obvious that soot formation is a nonequilibrium process. Several different theories have been advanced to explain the pyrolysis process—the extensive decomposition and atomic rearrangement of the fuel molecules that culminates in nucleation that occurs in very fuel-rich mixtures. Often-cited mechanisms are thermal cracking that results in fragmentation of fuel molecules into smaller ones, condensation reactions and polymerization that result in larger molecules, and dehydrogenation that lowers the H/C ratio of the hydrocarbons destined to become soot. Three different paths to the production of soot appear to exist, depending on the formation temperature. At the lowest temperatures (≤ 1700 K), only aromatics or highly unsaturated aliphatic compounds of high molecular weight form substantial solid carbon through pyrolysis. At intermediate temperatures

(>1800 K), all normally used hydrocarbon fuels produce soot if burned at sufficiently rich stoichiometry, but appear to do so by following a different path.

A simple mechanistic model for nucleation in the low and intermediate temperature ranges which has considerable experimental support is illustrated in Fig. 11.48.⁶⁷ At low temperatures, an aromatic hydrocarbon can produce soot via a relatively fast direct route that involves condensation of the aromatic rings into a graphite-like structure. Above about 1800 K, however, a slower, less-direct route is favored that entails ring breakup into smaller hydrocarbon fragments. These fragments then polymerize to form larger unsaturated molecules that ultimately produce soot nuclei. Aliphatic molecules can only follow this latter less-direct route. Experimental measurements in flames suggest that polyunsaturated hydrocarbon compounds are involved in nucleation, and acetylenes and polyacetylenes have been detected that decrease in concentration as the mass of carbon formed increases. Such observations fit the indirect path in Fig. 11.48.

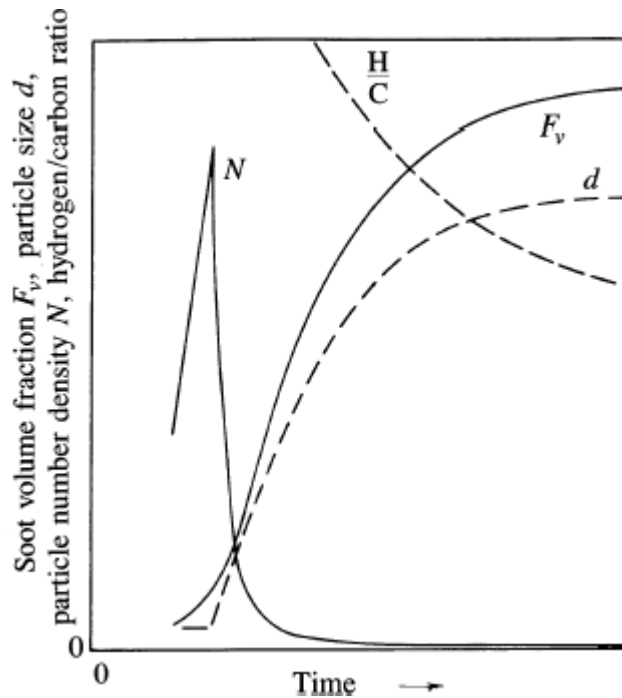
Figure 11.48 Mechanistic model for formation of soot from aromatic and aliphatic hydrocarbon compounds.⁶⁷



11.5.4.2. Soot Particle Growth

Nucleation produces a large number of very small particles with a modest soot loading. The bulk of the solid-phase material is generated by surface growth, which involves the gas-phase deposition of hydrocarbon intermediates on the surfaces of the spherules that develop from the nuclei. A qualitative description of the changes that occur as a function of time in a rich premixed flame during nucleation and surface growth is shown in Fig. 11.49. The soot fraction, F_V , in units of soot volume per unit volume of gas, is related to the number density N and the volume-mean diameter of the soot particles by

Figure 11.49 Variation in soot volume fraction F_v , particle number density N , particle size d , and soot hydrogen/carbon ratio with time in a flame.⁶⁷



$$F_V = \frac{\pi}{6} N d^3$$

(11.37)

d is the actual diameter of the spherules, or the diameter of a sphere of equivalent volume to an agglomerated particle. The rate of change of particle number density with time t can be written

$$\frac{dN}{dt} = \dot{N}_n - \dot{N}_a$$

(11.38)

where

$\dot{N}_n$ is the rate at which fresh nuclei appear and

$\dot{N}_a$ is the rate of agglomeration of spherules or particles that collide and stick. At the peak of the N curve,

$\dot{N}_n = \dot{N}_a$. To the left of the peak,

$\dot{N}_a > \dot{N}_n$ the particle diameter remains essentially constant at the minimum detectable diameter and the (small) rise in soot volume is dominated by nucleation. To the right of the peak in the N curve,

$\dot{N}_a > \dot{N}_n$. The number of agglomerating collisions is high because of the high number density; at the same time nucleation ends because there is enough dispersed surface area for gaseous deposition of hydrocarbon intermediates so the probability of generating new nuclei falls to zero. With nucleation halted slightly to the right of the N -curve peak, all the subsequent increase in soot volume fraction (the majority) stems from surface growth. To the right of the N curve peak, in the case illustrated, the number density falls by three orders of magnitude. This is the result of agglomeration, which is responsible for a portion of the increase in particle diameter. Agglomeration does not contribute to the rise in soot volume fraction, F_v . Surface growth that takes place on nuclei and on spherules is responsible for forming the concentric shells (somewhat distorted and warped) that constitute the outer portions of spherules and which are distinct from the less-organized spherule

center (see Figs. 11.42 and 11.43). Surface growth on agglomerated particles may partly fill in the crevices at the junctures of adjoining spherules to provide the nodular structure evident in Fig. 11.42.⁶⁷

Once particles have formed, inter-particle collisions can lead to agglomeration, thereby decreasing the number of particles and increasing their size. Three types of agglomeration have been identified in soot formation. During the early stages of particle growth, collision of two spherical particles may result in their *coagulation* into a single spheroid. This is easy to visualize in hydrocarbon pyrolysis where the beginnings of a soot particle may have the viscosity of a tarry liquid. Also, when the individual particles are small, rapid surface growth will quickly restore the original spherical shape. This process occurs up to diameters of about 10 nm. On the other hand, if spherules have solidified before collision and surface growth rates have diminished, the resulting particles resemble a cluster in which the original spherules retain much of their individual identity. After surface growth essentially ceases, continued coalescence of the soot particles results in the formation of chainlike structures of discrete spherules. This suggests electrostatic forces are significant. Positive charge measured on these particle chains is claimed to be the cause of their chainlike structure.^{67,69} This latter coalescence once surface growth ceases is termed *aggregation*.

It has been shown experimentally that during coagulation, the rate of decrease of particle number density was proportional to the product of a coagulation coefficient and the square of the number density:

$$-\frac{dN}{dt} = KN^2$$

(11.39)

This is the Smoluchowski equation for coagulation of a liquid colloid. Based on Brownian motion, this equation is applicable when the Knudsen number (ratio of mean free path to particle diameter) exceeds 10. K depends on such factors as particle size and shape, size distribution, and the temperature, pressure, and density of the gas. Equation (11.39) has been used to predict coagulation rates in low-pressure sooting flames.^{69,72} These studies show that under conditions approximating those in engine flames, the fraction of the initial number density N_0 remaining at time t is given approximately by

$$\frac{N}{N_0} \approx (KN_0t)^{-1}$$

(11.40)

Thus as t increases, N/N_0 decreases rapidly. Although these coagulation calculations are simplistic (in that many of the assumptions made are not strictly valid since soot particles are not initially distributed homogeneously within the combusting fuel-rich sprays, they are not monodisperse, and surface growth and oxidation may be taking place during agglomeration), an overall conclusion is that the rate of coagulation of spherules and particles to larger particles is very sensitive to number density. Thus the number of particles decreases rapidly with advancing crank angle in the diesel engine during the early part of the expansion process and agglomeration is essentially complete well before the exhaust valve opens.

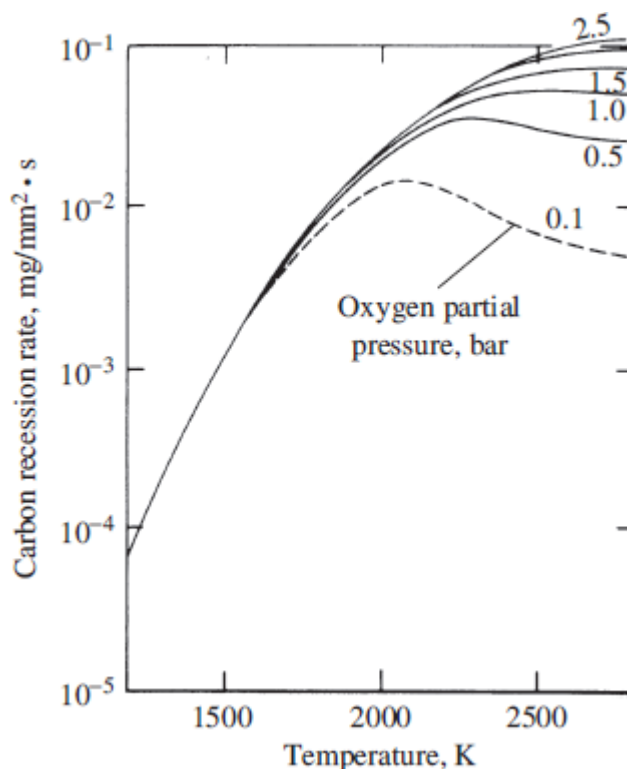
Throughout the soot formation process, the H/C ratio of the hydrocarbons formed in the pyrolysis and nucleation process, and of the soot particles, continually decreases. The H/C ratio decreases from a value of about 2, typical of engine fuels, to of order 1 in the youngest soot particles that can be sampled, and then to 0.2 to 0.3 once surface growth has ceased in the fully agglomerated soot.⁷² The latter stages of this process are indicated in Fig. 11.49. The addition of mass to the soot particles occurs by reaction with gas-phase molecules. The reacting gas-phase hydrocarbons appear to be principally acetylenes, with larger polymers adding faster than the smaller. Small polyacetylenes undergo further polymerization in the gas phase, presumably by the same mechanism leading to nucleation. As a result of preferential addition of the larger polymers, the H/C ratio of the particles decreases toward its steady-state value. Thus most of the polyacetylenes added must be of very high molecular weight, or dehydrogenation must also take place.^{69,72}

11.5.5. Soot Oxidation

In the overall soot formation process, shown schematically in Fig. 11.47, oxidation of soot at the precursor, nuclei, and particle stages can occur. The engine in-cylinder soot-concentration data reviewed in Sec. 11.5.3 indicate that almost all the soot formed within the fuel sprays is oxidized before the exhaust process commences. In the discussion of diesel combustion movies in Sec. 10.3.1, dark (brown) regions were observed in the original color photographs (see Fig. 10.4); these were interpreted as soot particle clouds, and were seen to be surrounded by a diffusion flame which appeared white from the luminosity of the high-temperature soot particles consumed in this flame. As air mixed with this soot-rich region, the diffusion flame eradicated the dark soot clouds as the particles were burned up.

In general, the rate of heterogeneous reactions such as the oxidation of soot depends on the diffusion of reactants to and products from the diffusion flame region as well as the kinetics of the oxidation reactions. For particles less than $1\text{ }\mu\text{m}$ diameter, diffusional resistance is minimal. The soot oxidation process in the diesel cylinder is kinetically controlled, therefore, since particle sizes are smaller than this limit. There are many species in or near the diffusion flame that could oxidize soot: examples are O_2 , O , OH , CO_2 , and H_2O . Earlier reviews of soot oxidation in flames postulated that at high oxygen partial pressures, soot oxidation could be correlated with a semi-empirical formula based on close-to-stoichiometric combustion products; however, oxidation by OH radicals (which are plentiful within the flame) may also be important under high temperature, high pressure, stoichiometric and slightly-lean within-the-flame conditions. Figure 11.50 shows the carbon particle surface recession rate that this semi-empirical formula predicts as a function of temperature and oxygen partial pressure.⁶⁷

Figure 11.50 Specific soot oxidation rate predictions as a function of temperature and oxygen partial pressure, from the Nagle and Strickland-Constable, O_2 based, correlation.⁶⁷



It has been argued on the basis of structural similarities that the rates of oxidation of soot and graphite should be similar. While this is a significant simplification, the semi-empirical formula of Nagle and Strickland-Constable has been shown⁷³ to correlate pyrolytic-graphite oxidation for oxygen partial pressures $p_{\text{O}_2} < 1\text{ atm}$ and temperatures between 1100 K and 2500 K. This formula is based on the concept that there are two types of sites on the carbon surface available for O_2 attack. For the

more reactive type A sites, the oxidation rate is controlled by the fraction of sites not covered by surface oxides (and therefore is of mixed order, between 0 and 1 in

p_{O_2}). Type B sites are less reactive, and react at a rate which is first order in

p_{O_2} . A thermal rearrangement of A sites into B sites is also allowed (with rate constant k_T). A steady-state analysis of this mechanism gives a surface mass oxidation rate w ($\text{gC}/\text{cm}^2 \cdot \text{s}$):

$$\frac{w}{12} = \left(\frac{k_A p_{O_2}}{1 + k_Z p_{O_2}} \right) x + k_B p_{O_2} (1 - x)$$

(11.41)

where x is the fraction of the surface occupied by type A sites and is given by

$$x = \left(1 + \frac{k_T}{p_{O_2}} \right)^{-1}$$

(11.42)

The empirical rate constants determined by Nagle and Strickland-Constable for this model are listed in [Table 11.13](#).

According to this mechanism, the reaction is first order at low oxygen partial pressures, but approaches zero order at higher pressures. At a given oxygen pressure, the rate initially increases exponentially with temperature (equivalent activation energy is k_A/k_Z or 34,100 cal/mol). Beyond a certain temperature the rate decreases as the thermal rearrangement favors formation of less reactive B sites. When, at sufficiently high temperature, the surface is completely covered with B sites, the rate is first order in oxygen partial pressure and increases again with temperature.⁷² Park and Appleton⁷³ have compared this formula with oxidation rate data obtained from pyrolytic graphite samples, carbon black particles, and with the available flame soot oxidation data. The formula correlates the data shown to within a factor of 2. Under diesel engine conditions, the O_2 partial pressure can be high ($\sim$ several atmospheres), as can the temperature of close-to-stoichiometric burned gases (≤ 2800 K). [Figure 11.50](#) shows the carbon particle surface recession rate that this formula predicts as a function of temperature and oxygen partial pressure.⁶⁷

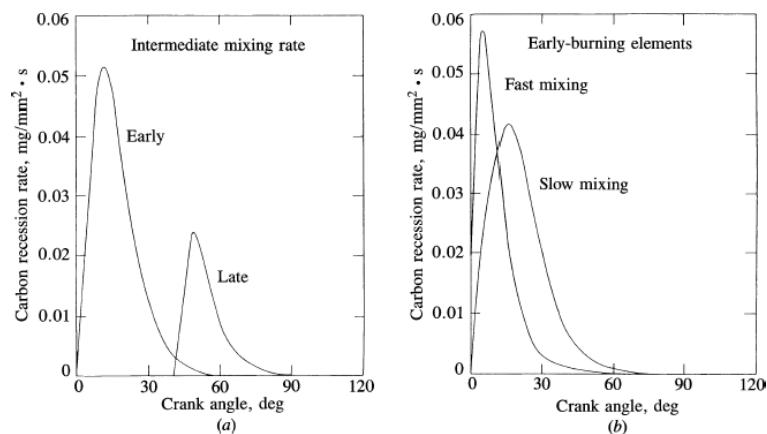
Table 11.13 Rate constants for Nagle and Strickland-Constable soot oxidation mechanism⁷⁴

Rate constant	Units
$k_A = 20 \exp(-15,100/T)$	$\text{g}/\text{cm}^2 \cdot \text{s} \cdot \text{atm}$
$k_B = 4.46 \times 10^{-3} \exp(-7640/T)$	$\text{g}/\text{cm}^2 \cdot \text{s} \cdot \text{atm}$
$k_T = 1.51 \times 10^5 \exp(-48,800/T)$	$\text{g}/\text{cm}^2 \cdot \text{s}$
$k_Z = 21.3 \exp(2060/T)$	atm^{-1}

This Nagle and Strickland-Constable relationship has been used to estimate the amount of soot that can be oxidized in a typical diesel engine. It was assumed that soot was present in stoichiometric combustion products within the diffusion flames, at selected times in the cycle, and that mixing with air leaned out the burned gas mixture at different rates until the overall fuel/air equivalence ratio was reached. The surface recession rate during this process was computed. [Figure 11.51](#) shows sample results at an engine speed of 1600 rev/min and an overall cylinder equivalence ratio of 0.58. Fast, intermediate, and slow mixing occurred in 30, 70, and 140 crank-angle degrees, respectively. The surface recession rate rises to a maximum as

p_{O_2} increases, and then decreases as the falling gas temperature more than offsets the increasing oxygen concentration. While the shape of the recession rate versus time curves depends on the mixing rate, the total amount of carbon burned (the area under each curve in Fig. 11.51(b) is about the same ($0.1 \mu\text{g}/\text{mm}^2$). However, the point in the cycle at which the soot-containing burned gas mixture passes through stoichiometric is much more important. For the late-mixing element shown (mixing lean of stoichiometric at 40° ATC), the total carbon mass oxidized is only 40% of that for the early mixing calculation. This is primarily due to the decreasing gas temperatures as the expansion stroke proceeds, and not the longer time available for burnup.⁷⁴

Figure 11.51 Soot particle burnup rate in diesel combustion environment: (a) in early- and late-burned fuel-air elements with intermediate mixing rate; (b) for fast and slow mixing for early-burning elements.⁶⁷



For a spherical soot particle, the mass burning rate w ($\text{g}/\text{cm}^2 \cdot \text{s}$) can be converted to a surface recession rate using

$$\frac{dr}{dt} = \frac{-w}{\rho}$$

where ρ is the density ($\approx 2 \text{ g}/\text{cm}^3$). The integrated values of $w(t)$ when divided by ρ then give the maximum radius of a soot particle that can be burned up. Integrated values of $0.1 \mu\text{g}/\text{mm}^2$ (estimated for TC start of burn up) correspond to a radius of about 50 nm or diameter of 100 nm. Individual spherule diameters are less than about 30 nm, so soot which mixes with oxygen containing burned gases early in the expansion stroke is likely to be fully burned. Thus the soot present in the exhaust would be expected to come from regions which mix with oxygen too late for the oxidation rate to be sufficient for complete particle burn-up.

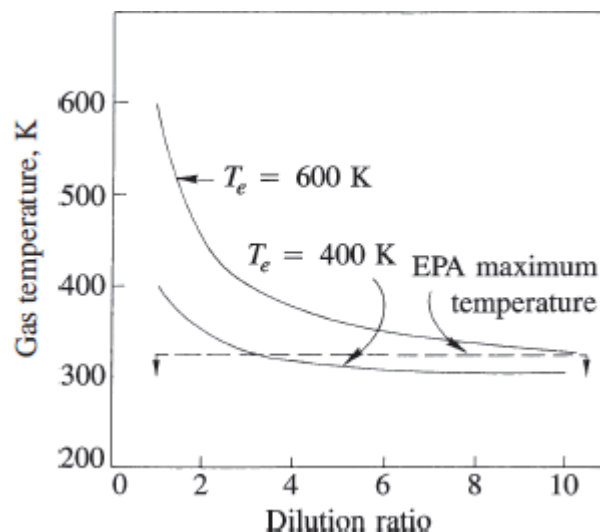
Agglomeration will have an indirect influence on the amount of soot oxidized through its effect on surface area. In the limiting case of a spherical cluster, n monodisperse spherules ($10 \leq n \leq 100$) can be imagined as compacted into a single solid sphere of equal volume. Alternatively, the same n spherules can be imagined compacted into a cylinder of diameter equal to that of the original spherules. Since oxidative attack is essentially an exterior surface phenomenon, the surface/volume ratio is the appropriate measure of the effect of particle shape on soot mass burn-up rate. It can be shown that the surface/volume ratios for the single sphere, single cylinder, and individual spherule, are in the ratio $n^{-1/3}$, $\frac{2}{3}$, and 1, respectively. Thus agglomeration will decrease the relative oxidation rate. In the limit, spherical clusters are less desirable than a chain; the larger the cluster the bigger the relative reduction in surface area. However, the densely packed spherule limit does not appear to be approached in practice. A specific surface area, of about $200 \text{ m}^2/\text{g}$ for diesel soot, has been measured.⁷⁵ A smooth-surfaced 25-nm diameter spherule with a $2\text{-g}/\text{cm}^3$ density has a surface/mass ratio of $150 \text{ m}^2/\text{g}$; the measured value is about one-third higher value, indicating low porosity and an agglomerate structure which is loosely rather than densely packed.⁷⁴

The research literature on soot oxidation (e.g. Refs. 76 and 77) suggest an alternative mechanism: oxidation by OH radical attack. In an atmospheric-pressure flame environment with rich and close-to-stoichiometric combustion products where O_2 mole fractions are low, oxidation by OH radical attack is more significant than oxidation by O or O_2 . Thus, OH radicals could be important in oxidizing soot in the diesel's diffusion flame under close-to-stoichiometric conditions. The maximum soot-particle surface-recession rates in the diesel combustion context occur in the diffusion flame that surrounds each fuel spray. The soot particles formed within each spray (where the local fuel-air equivalence ratio is significantly rich of stoichiometric, $\phi \approx 2-4$) enter the diffusion flame from the "fuel side"; oxygen enters from the air surrounding each spray. The soot particles are heated to close to the stoichiometric flame temperature (≈ 2700 K) and burn. Maximum oxidation rates most likely occur at an equivalence ratio of 0.9 (10% lean), where there is excess oxygen, as the combustion products containing these high-temperature particles diffuse outward from the close-to-stoichiometric diffusion-flame reaction zone and mix with air. As this outward diffusion and mixing with ever more air continues, particles that have not burned up will cool so soot oxidation slows down significantly. These particles become a key component in the engine-out particulate emissions.

11.5.6. Adsorption and Condensation

The final process in the particulate formation sequence illustrated in Fig. 11.47 is adsorption and condensation of hydrocarbons. This occurs after the cylinder gases have been exhausted from the engine, as these exhaust gases are diluted and cooled by mixing with air. In the standard particulate mass emission measurement process this occurs in a dilution tunnel which simulates approximately the actual atmospheric dilution process. A diluted exhaust gas sample is then filtered to remove the particulate. After equilibrating the collection filter at controlled conditions to remove water vapor, the particulate mass is obtained by weighing. In the prescribed EPA procedure, the filter temperature must not exceed 52°C . For a given exhaust gas temperature, the filter (and sample) temperature depends on the dilution ratio, as shown in Fig. 11.52.

Figure 11.52 Effect of exhaust gas dilution ratio on the temperature of the collected particulate sample as a function of engine exhaust temperature T_e .⁶⁷



The effect of the dilution ratio (and the dependent sample temperature) on collected particulate mass, for a standard dilution tunnel where the total sample is partitioned into extractable and nonextractable fractions, is as follows. The nonextractable fraction, the carbonaceous soot generated during combustion, is not affected by the dilution process. With no dilution (a dilution ratio of unity) the difference between the total and nonextractable mass is small; the bulk of the extractable fraction is acquired after the exhaust gas is mixed with dilution air. Extensive studies of this dilution process have shown that both adsorption and condensation occur. Adsorption involves the adherence of molecules of unburned hydrocarbons to the surfaces of the soot particles by chemical or physical (van der Waals) forces. This depends on the fraction of the available particle surface area occupied by hydrocarbons and on the partial pressure of the gaseous hydrocarbons that drives the

adsorption process. As the dilution ratio increases above unity, the effect of decreasing temperature on the number of active absorption sites dominates and the extractable fraction increases. At high dilution ratios, the sample temperature becomes insensitive to the dilution ratio (see Fig. 11.52) but the decreasing hydrocarbon partial pressure causes the extractable mass to fall further. At its maximum, the extractable fraction is about one-third of the total particulate. Condensation will occur whenever the vapor pressure of the gaseous hydrocarbon exceeds its saturated vapor pressure. Increasing dilution decreases hydrocarbon concentrations and hence vapor-pressure. However, the associated reduction in temperature does reduce the saturation pressure. High exhaust concentrations of hydrocarbons are the conditions where condensation is likely to be most significant, and the hydrocarbons most likely to condense are those of low volatility. Sources of low-volatility hydrocarbons are the high-boiling-point end of the fuel, unburned hydrocarbons that have been pyrolyzed but not consumed in the combustion process, and the lubricating oil.⁶⁷

Experiments with a passenger car diesel where the oil was tagged with a radioactive tracer have shown that the oil can contribute between 2 and 25% of the total particulate mass, with the greatest contribution occurring at high speed. About half of the extractable mass was traceable to the oil. All the material traceable to the oil was found in the extractable fraction, indicating that the oil did not participate in the combustion process. However, the oil is not always a significant contributor: in other diesel engines, fuel may be the dominant source of extractable material.⁶⁷

[1] With 3000 ppm S in the diesel fuel, the sulfate contribution to total particulate was about 5%. At 100 ppm S in the fuel, the sulfate contribution to the total particulate was an order of magnitude smaller, 0.4%.⁶⁴

11.6. EXHAUST GAS TREATMENT

11.6.1. Available Options

Our discussion so far has focused on *engine-out* emissions. Substantial further reductions in emissions can be obtained by removing pollutants from the exhaust gases in the engine exhaust system. Effective devices developed to achieve this result include catalytic converters (oxidizing catalysts for HC and CO, reducing catalysts for NO_x, and three-way catalysts for all three pollutants), and traps or filters for particulates. The stringency of current and anticipated vehicle air-pollutant regulations requires major reductions in engine-out emissions, so effective aftertreatment catalyst and filter technologies in the exhaust system are now essential.

The temperature of exhaust gas in a spark-ignition engine can vary from 300 to 400°C during idle to about 900°C at high-power operation. The most common range is 400 to 600°C. Spark-ignition engines operate most of the time at the stoichiometric air/fuel ratio to enable the three-way catalyst technology. Historically, such engines have operated fuel-rich at high torque levels to prevent knock onset and reduce exhaust gas temperatures, but that option is being used much less since it significantly increases HC and CO emissions. Lean operation of SI engines improves the engine's efficiency at part-load, but achieving stable combustion and tight emission's control are challenges. The exhaust gas may therefore contain modest amounts of oxygen (when slightly lean) or more substantial amounts of CO (when rich). In contrast, diesel engines, where load is controlled by the amount of fuel injected, always operate well lean of stoichiometric. The diesel's exhaust gas therefore contains substantial oxygen and is at a lower temperature (200 to 500°C). Removal of gaseous pollutants from the exhaust gases after they leave the engine cylinder can be either thermal or catalytic. In order to oxidize the hydrocarbons in the gas phase without a catalyst, a residence time of order 50 ms and exhaust temperatures in excess of 600°C are required. To oxidize CO, temperatures in excess of 700°C are required. Temperatures high enough for homogeneous thermal oxidation can be obtained by spark retard (with some loss in efficiency) and minimizing heat losses in the exhaust ports and manifold. The residence time can be increased by increasing the exhaust manifold volume to form a *thermal reactor*; however, this approach has not proved successful.

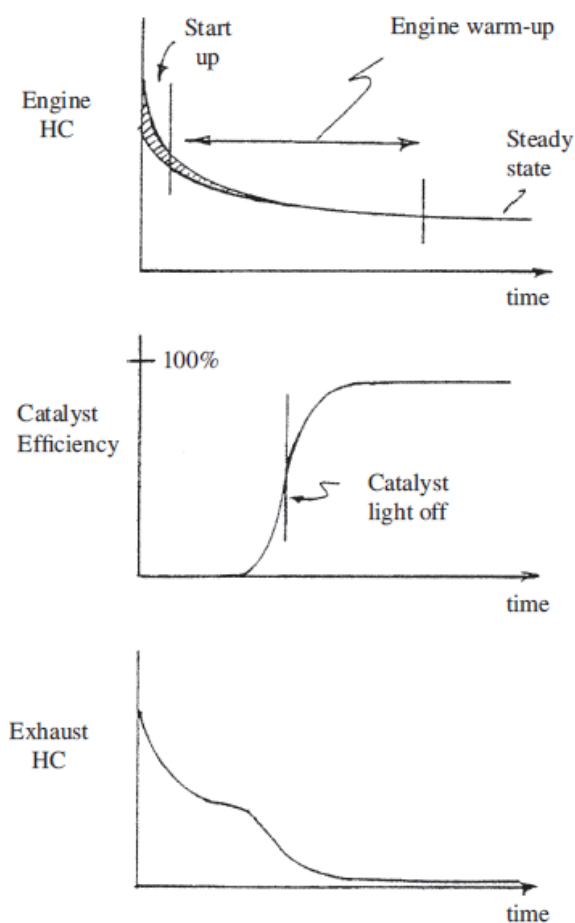
The task of catalytic aftertreatment devices is to oxidize the CO and HC in the exhaust gas to CO₂ and H₂O at temperature levels the exhaust gases can readily provide. With NO_x (predominantly nitric oxide, NO) the task is reduction to N₂, again at or below the available exhaust gas temperature. The temperatures required for gas-phase chemistry to effect these changes, since the *activation energy* (or energy barrier) for breaking these air pollutant molecule bonds is above the thermal energy available at average exhaust gas temperatures. So only modest pollutant reduction in the exhaust system occurs. The catalyst (e.g., a collection of noble metal atoms on a high surface-area substrate) facilitates the desired chemistry by reducing this energy barrier so reactions can occur at significantly lower (and attainable) temperatures. For CO and HC oxidation on the catalyst surface, sufficient oxygen must be available. With a stoichiometric mixture, the air (oxygen) associated with the CO and unburned HC is (on average) available. For NO reduction in spark-ignition engines, reducing agents in the exhaust gas (CO, H₂, and HC) are present. For diesels, with their lean operation, reducing agents such as ammonia NH₃ are added to effect this reduction. Two added features often embodied are that catalysts can be formulated to target and enhance specific (selective) chemical reaction rates. Also, storage capacity for a specific pollutant can be built into the catalyst, with release and removal performed periodically, triggered by changing the exhaust gas composition: for example, in lean NO_x storage-reduction catalysts for SI engines where the engine-out NO stored when the engine operates lean, is released and reduced when the engine is switched for a short period to fuel-rich operation. Catalytic oxidation of CO and hydrocarbons and reduction of NO_x in the exhaust can be achieved at temperatures as low as 250 to 300°C. Thus effective removal of these pollutants occurs over a much wider range of temperatures than can be achieved with gas-phase thermal oxidation of HC and CO.

In spark-ignition engine exhausts, catalysts for CO, HC, and NO removal simultaneously are widely used, and are highly effective. However, these *three-way catalysts* are only effective if the exhaust gas is stoichiometric. Precisely at this condition, there is oxygen available for oxidation and reducing agents (CO, H₂, and HC) are available for NO reduction. In diesels, which

operate lean, NO_x and particulates are the primary pollutants emitted. Diesel oxidation catalysts are frequently used to oxidize hydrocarbons since, if not removed, higher molecular weight HC can be absorbed later, onto soot particles. Selective catalytic reduction (SCR) of NO_x (in an otherwise oxidizing environment) is effected using NH_3 as the primary reducing agent which is added in the form of urea to the exhaust gases. Also, NO_x storage catalysts for lean operating engines are available where NO is converted to NO_2 , and stored as a nitrate. When the storage capacity is full, the engine is operated rich for a short period, the nitrate is decomposed, released as NO_2 , and reduced to N_2 by the HC, CO, and H_2 in the rich exhaust.

For effective operation, these exhaust catalysts must operate at an elevated temperature (a few hundred $^\circ\text{C}$) and with an exhaust gas flow rate (which defines the residence time) and composition matched to the catalyst's size and chemical conversion task. Thus insertion of catalysts into the engine's exhaust requires that the catalyst and engine be designed as two interacting systems (often with sensors and feedback loops). The warm up of the catalyst following a cold engine start is one key area. The catalyst surface must reach a temperature of about 300°C to become effective, and the exhaust gas flow and temperature are the available sources of thermal energy. Figure 11.53 illustrates the warm up of an oxidation catalyst and its evolving impact on engine-out HC emissions. During the starting and warm up of a spark-ignition engine, the engine-out HC emissions are higher than the warmed-up engine HC levels. But the catalyst takes time to warm up. It takes time for the catalyst to *light off*: that is, attain a temperature at which catalytic efficiency is defined as the percentage of a given pollutant removed by the catalyst). Thus the HCs exhausted from the exhaust system are initially the cold engine-out levels: only after catalyst light off do the engine-out levels drop rapidly to a warmed-up engine and catalyst (usually a reduction of over 90%).

Figure 11.53 Illustration of the warm-up behavior of a spark-ignition engine/catalyst system. Catalyst light off takes some 10 to 30s, depending on conditions.



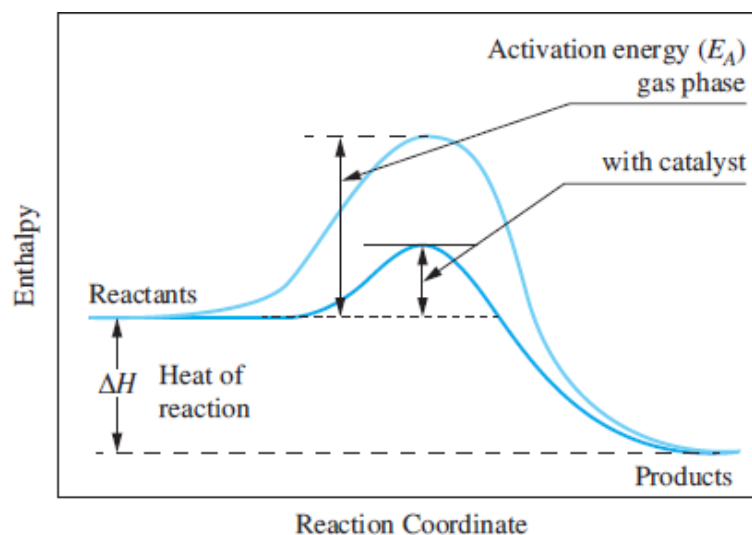
Particulates in the exhaust gas stream can be removed by a trap or filter. Due to the small particle sizes involved, some type of filter is the most effective trapping method. Porous ceramic-monolith catalyst-support technology has proved a good

approach to exhaust gas filtration. Half of the wall-flow monolith channels are plugged at one end, the other half are plugged at the other end. When used as a particulate trap, the porosity of the monolith channel walls has higher and carefully controlled porosity. The channel or cell density (about 200 cells/in², or 30 per cm²) is a factor of 2 to 4 lower than that of catalyst substrates. A filter wall-flow monolith for a diesel automobile would be about 150 mm in diameter, 250 mm in length, and thus have about 5000 flow channels (into only half of which the exhaust gas flows). Such a particulate filter is usually preceded by a diesel oxidation catalyst (DOC), which removes the hydrocarbons in the exhaust gas prior to where they might condense or be absorbed onto the soot particles. Mass accumulates within the trap, and the exhaust manifold pressure thus increases during trap operation. The particulates, once trapped and accumulated, are burned up either by initiating oxidation within the trap by raising the trap temperature by adding to the exhaust's thermal energy by burning extra fuel in the cylinder or in the manifold. Or the trap may contain catalytically active material to initiate burn up of the particulate matter at lower temperature. The operation of particulate traps is reviewed in [Sec. 11.6.4](#).

11.6.2. Catalyst Fundamentals

Following this introduction to automotive catalysts, we now develop several relevant fundamental concepts. [Figure 11.54](#) illustrates the role of the catalyst in enabling reactions in more detail (e.g., between a given hydrocarbon molecule and oxygen in the exhaust gases at the catalyst surface) at temperatures of order 300 to 600°C. Such diagrams indicate the state (enthalpy) of the reactants as they transition to end-products (CO₂ and H₂O in the example of HCs above) as the reaction coordinate, reactants to products, progresses. The heat of reaction (ΔH ; see [Sec. 3.5.1](#) and [Fig. 3.5](#)), the chemical energy released by the reaction, is not changed by the presence of the catalyst. However, the presence of an effective catalyst (e.g., the noble metal platinum) significantly reduces the energy (E_A) required to break the molecular bonds in the reactant molecules. Thus the desired chemistry occurs at the active site on the catalyst surface at a significantly lower temperature than would be needed to obtain sufficiently energetic molecular collisions in the gas phase. The catalyst provides a lower energy barrier (E_A) as reactants are transformed to products, which increases the rates of the reactions involved. The changes in enthalpy ΔH and free energy ΔG , initial to final state, are neither affected by the catalyst nor is the final equilibrium. The rate of reaction varies inversely to the activation energy. The activation energy for the entire reaction E_A , is primarily governed by the slowest of all the reaction steps involved in the overall reactants to products chemical process. This slowest step is the rate-limiting step in the entire process.

Figure 11.54 Reaction energy paths for reactants transitioning to products illustrating the role of the activation energy barrier. For an exothermic reaction (e.g., CO oxidizing to CO₂) where, overall, thermal energy is released (ΔH is positive). The presence of the catalyst significantly reduces the energy barrier to the reaction.



Catalysts are characterized by two basic properties: their *activity* and their *selectivity*. Activity relates to the degree to which the catalyst increases the rate of the relevant chemical reactions. Normally, a higher activity—faster reaction rate—in a given context is more desirable. Reaction rates are proportional to the concentrations of the reacting species, each raised to the power of its stoichiometric coefficient in the relevant chemical reaction (see Sec. 3.7.2). The expressions for and values of the rate constants for the slower reactions are critical. The catalyst enhances the transition to chemical equilibrium: it does not change that equilibrium (see Sec. 3.7.1).

A catalyst can be selective: that is, accelerate certain specific reactions much more than others. A catalyst's selectivity, which is usually designed into a catalyst through the elements incorporated in its formulation, is also usually dependent on its temperature: it enhances specific reactions within a temperature window. Diesel NO_x catalysts are "selective" in this way.

Automotive catalysts are heterogeneous catalysts: that is, the active noble metal catalyst materials are dispersed over a solid surface. Their structure and mode of operation are illustrated in Figs. 11.55 and 11.56. Monoliths (honeycomb like structures) are widely used in automotive applications as supports. As shown in Fig. 11.55, the exhaust is directed through square cross-section channels [e.g., 400 to 600 channels/in², 60 to 90 channels/cm²; some 15 cm (6 in.) long], which expose the exhaust gas to a large surface area for sufficient time. Such a monolith support is extruded. Synthetic cordierite, a combination of magnesia, alumina, and silica, is the material used ($2\text{MgO} \cdot 2\text{Al}_2\text{O}_3 \cdot 5\text{SiO}_2$). High melting point (over 1300°C), low thermal expansion coefficient (hence good resistance to thermal shock), low heat capacity (for rapid warm-up), and adequate strength are its key properties. The monolith is surrounded by a flexible seal and support inside a steel can as shown.

Figure 11.55 Monolith design of catalytic converter for spark-ignition engine emissions control.

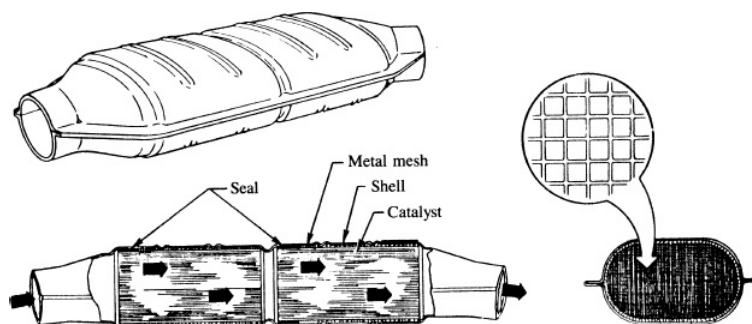
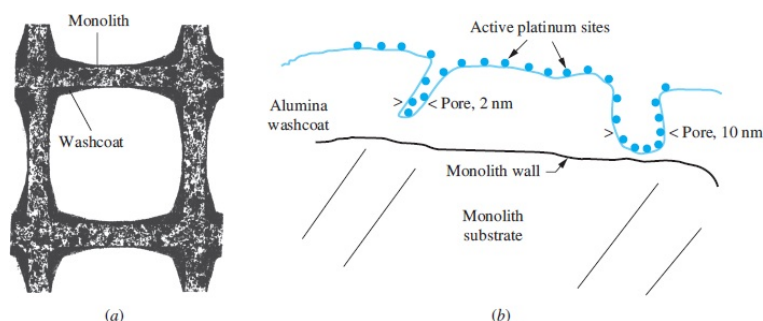


Figure 11.56 (a) Cross-section showing a monolith channel and washcoat (about 1 mm² flow area); (b) schematic of catalytically active (Pt) sites dispersed over the washcoat surface which is deposited on the monolith substrate walls.⁷⁸



The critical catalyst component is the *washcoat*. Commercial ceramic monoliths have large pores and relatively low surface area (e.g., 0.3 m²/g). The washcoat, deposited on the substrate, through its high porosity provides a large surface area (100 m²/g) on which to disperse the noble metal catalyst material. The washcoat material can also enhance the activity of the catalyst through catalyst component/support interactions, and stabilize the precious metals against deactivating mechanisms such as sintering. Figure 11.56(a) illustrates the layout and role of the washcoat. The base material is γ -alumina often with several of these oxides—cerium, lanthanum, iron, nickel, barium, and neodymium—added. The monolith may be dipped in a

washcoat slurry (with the catalytic metals included); or a two-step process may be used where the monolith is first coated with an non-catalytic layer followed by dipping it into a solution containing the catalytic salts which are absorbed onto the washcoat surface, which is then dried and calcined by heating in the air.

The noble metal catalyst atoms are dispersed over this highly porous washcoat surface as illustrated in Fig. 11.56b. The surface of the alumina is rich in surface OH^- groups, which are sites at which the noble metal catalytic material can bond. The smaller the size of the tiny platinum (and palladium, and rhodium) particles (or crystallites), the more dispersed these active sites will be, and the higher the catalytic activity.

The conversion of reactants (e.g., hydrocarbons and oxygen) to products (CO_2 and H_2O) at the catalyst surface involves the following physical and chemical steps. First, bulk diffusion of the reactants within each of the channels toward the walls occurs. Diffusion into the pores of the washcoat toward the internally dispersed catalyst material then follows. Next, the reactant molecules must chemisorb at the surface onto an active (catalytic) site. If the surface temperature is high enough (so the reactant molecules are at or above the peak of the activation energy barrier shown in Fig. 11.54), the reactants undergo the chemical reactions that convert them to products which then desorb from the catalytic sites. Products then diffuse back through the pore network to the outer washcoat surface. Finally, the product species diffuse into the bulk gases flowing through the monolith channels.

These steps—reactant bulk mass transfer, pore diffusion, chemisorption and reaction to products—followed by the reverse sequence for the products, occur at different rates which are dependent on temperature. Chemical reaction rates follow an exponential dependence on temperature, $\exp(-E_A/RT)$, where E_A is the activation energy (see Fig. 11.54). In automotive catalysts, E_A is typically 40 kJ/mol or greater. Pore diffusion is largely dependent on pore size and geometry: the rate at which reactants and products diffuse into the pores varies approximately with the square root of the rate constant for the chemical reaction, thus giving it an effective activation energy of about half that of the chemical reaction, 25 to 40 kJ/mol. The bulk diffusion/mass transfer rates vary approximately as $T^{3/2}$, which corresponds to some 8 to 16 kJ/mol. At low temperatures, the chemical reaction rates are slow relative to diffusion rates. As temperature increases, the steps with the highest activation emerges (and exponential dependence) increase the fastest and pore diffusion becomes limiting. At high temperatures both chemical and pore diffusion rates become sufficiently fast that bulk diffusion with its lessor dependence on temperature becomes rate limiting. Equations that describe the reaction chemistry and species diffusion processes can be used to quantify this sequence of individual processes and thus obtain estimates of species conversion rates in these monolith high-surface-area catalysts.⁷⁹

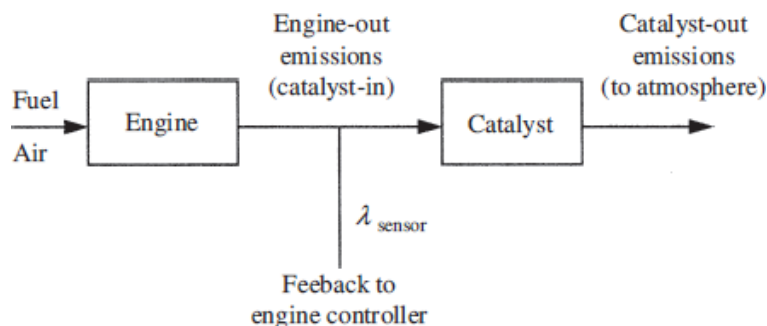
Automotive catalysts have been developed to the point where they are extremely effective at removing air pollutants from engine exhaust gases. Catalyst effectiveness is quantified through its catalyst efficiency which is defined as the ratio of the amount of a particular air pollutant (e.g., HC) removed by the catalyst to the amount of that pollutant entering the catalyst:

$$\eta_{\text{cat,HC}} = \frac{\dot{m}_{\text{HC,in}} - \dot{m}_{\text{HC,out}}}{\dot{m}_{\text{HC,in}}} = 1 - \frac{\dot{m}_{\text{HC,out}}}{\dot{m}_{\text{HC,in}}}$$

(11.43)

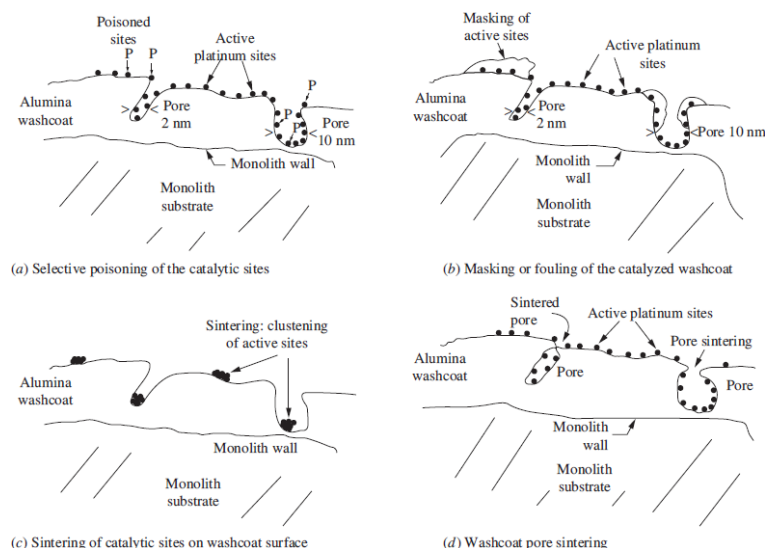
Figure 11.57 illustrates the layout of a gasoline engine plus three-way catalyst system with its relative air/fuel ratio sensor to ensure stoichiometric operation, and provides the context for this catalyst efficiency definition. Catalyst efficiencies are now well-above 90%, and with clean fuels can maintain these high efficiencies over extended mileage. Note that $\dot{m}_{\text{HC,out}}/\dot{m}_{\text{HC,in}}$ is the catalyst *inefficiency*, and is 5% for a 95% catalyst efficiency corresponding to a factor of $100 \div 5 = 20$ reduction in engine-out emissions. Catalyst efficiency increases as temperature increases in an "S-shaped" curve. At the low temperature end of the curve there is little catalyst activity. At some point, the catalyst "lights off" and catalytic chemical activity becomes significant (with HC and CO oxidizing automotive catalysts, significant chemical energy is released within the catalyst which further raises its temperature. A commonly used *light-off temperature* definition is the temperature (T_{50}) at which the catalyst efficiency reaches 50%.

Figure 11.57 Layout of spark-ignition engine plus three-way catalyst system, with its λ -sensor (and feedback to the engine control system), to maintain stoichiometric operation.



While catalysts are not consumed during catalytic reactions, they steadily (though slowly) deactivate with use. The important deactivation modes are poisoning, fouling or masking, and thermal deactivation or sintering. [Figure 11.58](#) illustrates these catalyst deactivation processes. Poisoning is a major cause of deactivation. Contaminants (usually in the fuel or oil consumed) such as lead, phosphorus, sulfur, and other inorganic metals, deactivate the noble metal (e.g., Pt) active sites by forming catalytically inactive alloys, or chemi-absorb onto active sites and block those sites from further catalytic action. The former mode (e.g., with phosphorus) results in permanent deactivation. The latter site blocking (e.g., with sulfur) can be reversed by desorbing the poison from the catalyst surface. [Figure 11.58a](#) illustrates this catalyst poisoning mode: the number of active sites is reduced so activity decreases. Masking or fouling of the catalyst washcoat, illustrated in [Fig. 11.58b](#), covers portions of the washcoat surface by physically depositing high molecular weight HCs from the oil, components of the oil additive package, scale from corrosion, dusts, etc., as a glaze over portions of the washcoat surface. These "deposits" prevent exhaust gas species from reaching active catalytic sites and also narrow and block entrances to the smaller washcoat pores. The third mode is thermal (high temperature) deactivation or sintering, as shown in [Figs. 11.58c](#) and [d](#). Here, under higher than intended surface temperatures, the active sites (fine catalyst crystals) can migrate over the surface, coalesce and form larger crystals at fewer active sites, thus decreasing catalytic activity. At higher temperatures still, the porosity of the alumina-based washcoat may be reduced through "carrier sintering" ([Fig. 11.58d](#)). Note that loss of active sites through selective poisoning and catalyst sintering primarily affects the reaction kinetics: thus higher temperatures are needed to achieve a given activity level but the maximum conversion efficiency may well be unaffected. Catalyst deactivation resulting from increased pore diffusion resistance manifests itself as lower conversion efficiency, and thus changes the efficiency versus temperature curve slope. Masking causes both loss of active (surface) sites and pore blocking. Thus the conversion efficiency curve is shifted toward higher temperatures with a significant reduction in slope. Note that with clean fuels, low oil consumption, and adequate control of catalyst temperatures, the deterioration in catalyst efficiency with extensive use (mileage in a vehicle) can be held within tolerable bounds.^{78,79}

Figure 11.58 Catalyst deactivation processes: (a) selective poisoning (P) of active platinum (Pt) sites; (b) masking or fouling of sections of the washcoat surface; (c) migrating and coalescence of Pt active sites on the washcoat surface due to higher than intended temperatures; (d) washcoat pore sintering due to higher still temperatures, which reduces washcoat porosity.⁷⁸



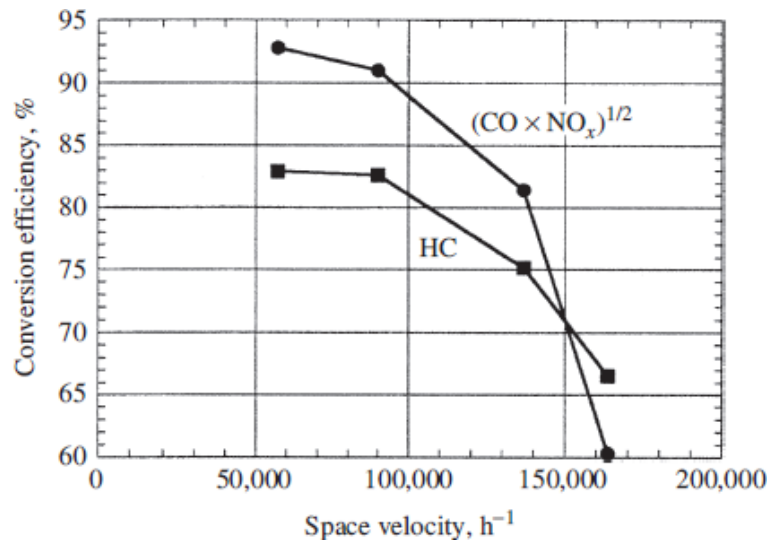
The above discussion implies that the air pollutants to be removed from the exhaust gases must spend sufficient time within the catalyst to complete the sequence of processes that essentially complete removal requires. The catalyst must be sized so that over the operating range of the engine, in a given application, this constraint is satisfied. The residence time within the catalyst is defined in terms of the volumetric flow of exhaust gases and the catalyst volume through the space velocity, SV:

$$SV = \text{volume flow rate of exhaust} / \text{catalyst volume}$$

(11.44)

The volume flow rate is at standard atmospheric pressure and temperature. The catalyst volume is the volume of the honeycomb monolith.^[1] The normal units are reciprocal hours, h^{-1} . In automotive applications, the maximum space velocity is limited to about 50,000 to 100,000 h^{-1} as indicated in Fig. 11.59. The inverse of the space velocity is the exhaust gas residence time within the catalyst. 50,000 h^{-1} corresponds to a 70-ms residence time, which at 2000 rev/min is about two crank revolutions. Too high a flow rate (too short a residence time) results in a reduction in catalyst activity.

Figure 11.59 Effect of space velocity [SV, defined in Eq. (11.44)] on catalyst conversion efficiency. Maximum conversion efficiency achieved at below about 50,000 (h^{-1}).⁸⁰



Note that a spark-ignition engine operating under conditions typical of light-duty vehicle urban driving (about one-quarter the maximum torque, and at low engine speeds), the exhaust gas flow is about one-tenth that at maximum rated-power conditions (WOT and 5000 to 6000 rev/min). Under normal driving conditions, space velocities are much below the full power value. Thus (with appropriate catalyst temperatures) only the front axial portion of the catalyst monolith and washcoat is needed to achieve high catalyst efficiencies. At high loads and speeds, that is at high space velocities, then close to the full catalyst length is needed to remove almost all the engine-out emissions.

The size and geometry of the catalyst substrate also affect the pressure drop across the catalytic converter. The typical monolith substrate converter (see Fig. 11.55) has an entrance section, the monolith substrate section (with some 10,000 channels), and an exit section. The flow losses in the entry and exit header sections depend on exhaust pipe diameter and other geometric details, and are small. The flow blockage the monolith structure causes, creates contraction and expansion losses as the exhaust flows into and out of the monolith. Thus the overall monolith pressure drop, Δp_c , is:

$$\Delta p_c = \Delta p_m + \Delta p_{ch} + \Delta p_{out}$$

(11.45)

The channel pressure drop Δp_{ch} is about 90% of the total: Δp_{in} and Δp_{out} are relatively small.

The flow in the monolith channels is laminar, and the pressure drop is given by the Darcy equation:

$$\Delta p_{ch} = 4f(L/d_{ch})(\rho v^2/2)$$

(11.46)

where f is the friction factor, L and d_{ch} are the length and hydraulic diameter of each channel, ρ is the exhaust gas density, and v the gas velocity in the channels. (The gas velocity is obtained from the engine mass flow rate divided by the gas density and the substrate open area.)

The flow in catalyst monolith channels is always laminar, so the friction factor f is then given by

$$f = K/(\text{Re})$$