CHAPTER

Sensors, actuators and oscilloscope diagnostics

4.1 Introduction

The issues and diagnostic techniques used for sensors and actuators are common to many systems. For example, the testing procedure for an inductive engine speed sensor on a fuel injection system is the same as for an inductive speed sensor on an antilock brake system (ABS) system. Testing sensors to diagnose faults is usually a matter of measuring their output signal. In some cases, the sensor will produce this on its own (e.g. an inductive sensor). In other cases, it will be necessary to supply the correct voltage to the device to make it work (e.g. Hall sensor). It is normal to check that the vehicle circuit is supplying the voltage before proceeding to test the sensor.

Key fact

Testing sensors to diagnose faults is usually a matter of measuring their output signal.

At the beginning of the sections on sensors and actuators, a table is included listing the device, equipment, test method(s), results of the tests and, in most cases, a reference to a scope waveform. A waveform is often the recommended method of testing. The waveform shown will either be the output of a sensor or the signal supplied to an actuator.

Note: Any figures given are average or typical values. Refer to a good reference source such as a workshop manual or ‘data book’ for specific values.

Author’s Note: The waveforms in this chapter were captured using the PicoScope® automotive oscilloscope. I am most grateful to the PicoTech team for supplying information and equipment to assist in the production of this chapter (http://www.picoscope.com).

4.2 Sensors

4.2.1 Introduction and sensor diagnostics

A sensor is a device that measures a physical quantity and converts it into a signal which can be read by an electronic control unit (ECU), an observer or an instrument. For accuracy, most sensors are calibrated against known standards. Most vehicle sensors produce an electrical signal, so checking their output on an oscilloscope is often the recommended method. However, many can also be checked using a multimeter (Table 4.1).

Key fact

Most vehicle sensors produce an electrical signal.

Advanced Automotive Fault Diagnosis.

Table 4.1 Sensor diagnostic methods

Sensor	Equipment	Method(s)	Results	Scope waveform
Thermistor Coolant sensor Air intake temperature sensor Ambient temperature sensor Etc.	Ohmmeter	Connect across the two terminals, or if only one, from this to earth	Most thermistors have a negative temperature coefficient (NTC). This means the resistance falls as temperature rises. A resistance check should give readings broadly as follows: 0 °C = 4500 Ω 20 °C = 1200 Ω 100 °C = 200 Ω	Figure 4.14
Inductive Crankshaft speed and position ABS wheel speed Camshaft position	Ohmmeter AC voltmeter	A resistance test with the sensor disconnected AC voltage output with the engine cranking	Values vary from approx. 200–400 Ω on some vehicles to 800–1200 Ω on others. The 'sine wave' output should be approx. 5V (less depending on engine speed)	Figure 4.2 Figure 4.4 Figure 4.7
Hall effect Ignition distributor Engine speed Transmission speed Wheel speed Current flow in a wire (ammeter amp clamp)	DC voltmeter Logic probe Do NOT use an ohmmeter as this will damage the Hall chip	The voltage output measured as the engine or component is rotated slowly. The sensor is normally supplied with a 5V or a 10–12V	This distributor switches between 0 and approx. 8V as the Hall chip is magnetised or not. Others switch between 0 and approx. 4V A logic probe will read high and low as the sensor output switches	Figure 4.3 Figure 4.17 Figure 4.19
Optical Ignition distributor Rotational speed	DC voltmeter	The device will normally be supplied with a stabilised voltage. Check the output wire signal as the device is rotated slowly	Clear switching between low and high voltage	N/A
Variable resistance Throttle potentiometer Flap-type airflow sensor Position sensor	DC voltmeter Ohmmeter	This sensor is a variable resistor. If the supply is left connected then check the output on a DC voltmeter With the supply disconnected, check the resistance	The voltage should change *smoothly* from approx. 0V to the supply voltage (often 5V) Resistance should change smoothly	Figure 4.8 Figure 4.10
Strain gauges MAP sensor Torque stress	DC voltmeter	The normal supply to an externally mounted manifold absolute pressure (MAP) sensor is 5V. Check the output as manifold pressure changes either by snapping the throttle open, road testing or by using a vacuum pump on the sensor pipe	The output should change between approx. 0 and 5V as the manifold pressure changes. As a general guide 2.5V at idle speed	N/A

(Continued)

Table 4.1 (Continued)

Sensor	Equipment	Method(s)	Results	Scope waveform
Variable capacitance	DC voltmeter	Measure the voltage at the sensor	Small changes as the input to the sensor is varied – this is no difficult to assess because of very low capacitance values	N/A
Accelerometer Knock sensors	Scope	Tap the engine block lightly (13 mm spanner) near the sensor	Oscillating output that drops back to zero If the *whole* system is operating, the engine will slow down if at idle speed	Figure 4.21
Hot wire Airflow	DC voltmeter or duty cycle meter	This sensor includes electronic circuits to condition the signal from the hot wire. The normal supply is either 5 or 12 V. Measure the output voltage as engine speed/load is varied	The output should change between approx. 0 and 5 V as the airflow changes 0.4–1 V at idle is typical. Or depending on the system in use the output may be digital	Figure 4.12
Oxygen Lambda sensor EGO sensor HEGO sensor	DC voltmeter	The lambda sensor produces its own voltage a bit like a battery. This can be measured with the sensor connected to the system	A voltage of approx. 450 mV (0.45 V) is the normal figure produced at lambda value of one The voltage output, however, should vary smoothly between 0.2 and 0.8 V as the mixture is controlled by the ECU	Figure 4.24 Figure 4.26 Figure 4.27
Acceleration switch Dynamic position	DC voltmeter	Measure the supply and output as the sensor is subjected to the required acceleration	A clear switching between say 0 and 12 V	N/A
Rain and other unknown types	DC voltmeter	Locate output wire – by trial and error if necessary and measure dry/wet output (splash water on the screen with the sensor correctly fitted in position)	A clear switching between distinct voltage levels	N/A

4.2.2 Inductive sensors

Inductive-type sensors are used mostly for measuring the speed and position of a rotating component. They work on the very basic principle of electrical induction (a changing magnetic flux will induce an electromotive force in a winding). The output voltage of most inductive-type sensors approximates to a sine wave. The amplitude of this signal depends on the rate of change of flux. This is determined mostly by the original design as in the number of turns,

Key fact

The amplitude of an inductive sensor signal depends on the rate of change of flux.

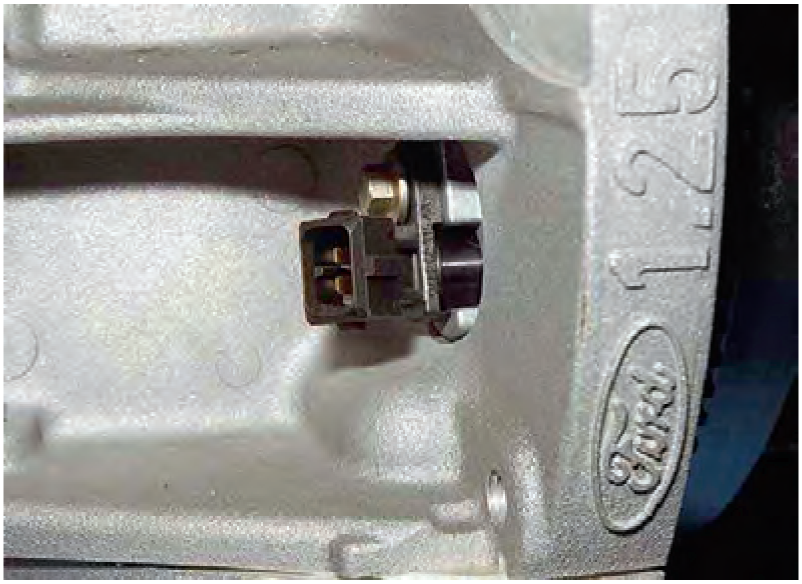


Figure 4.1 Crank sensor in position near the engine flywheel

magnet strength and gap between the sensor and the rotating component. Once in use, though, the output voltage increases with the speed of rotation. In the majority of applications, it is the frequency of the signal that is used.

4.2.2.1 Crankshaft and camshaft sensors

Inductive-type crank and cam sensors work in the same way. A single tooth, or toothed wheel, induces a voltage into a winding in the sensor. The cam sensor provides engine position information as well as which cylinder is on which stroke. The crank sensor provides engine speed. It also provides engine position in many cases by use of a 'missing' tooth (Figure 4.1).

In this particular waveform, we can evaluate the output voltage from the crank sensor. The voltage will differ between manufacturers, and it also increases with engine speed. The waveform will be an alternating voltage signal.

If there is a gap in the trace, it is due to a 'missing tooth' on the flywheel or reluctor and is used as a reference for the ECU to determine the engine's position. Some systems use two reference points per revolution (Figure 4.2).

Definition

CID: Cylinder identification.

The camshaft sensor is sometimes referred to as the cylinder identification (CID) sensor or a 'phase' sensor and is used as a reference to time sequential fuel injection.

The voltage produced by the camshaft sensor will be determined by several factors, these being the engine's speed, the proximity of the metal rotor to the pick-up and the strength of the magnetic field offered by the sensor. The ECU needs to see the signal when the engine is started for its reference; if absent, it can alter the point at which the fuel is injected. The driver of the vehicle may not be aware that the vehicle has a problem if the CID sensor fails, as the drivability may not be affected. However, the MIL should illuminate.

The characteristics of a good inductive camshaft sensor waveform is a sine wave that increases in magnitude as the engine speed is increased, and usually provides one signal per 720° of crankshaft rotation (360° of camshaft rotation).

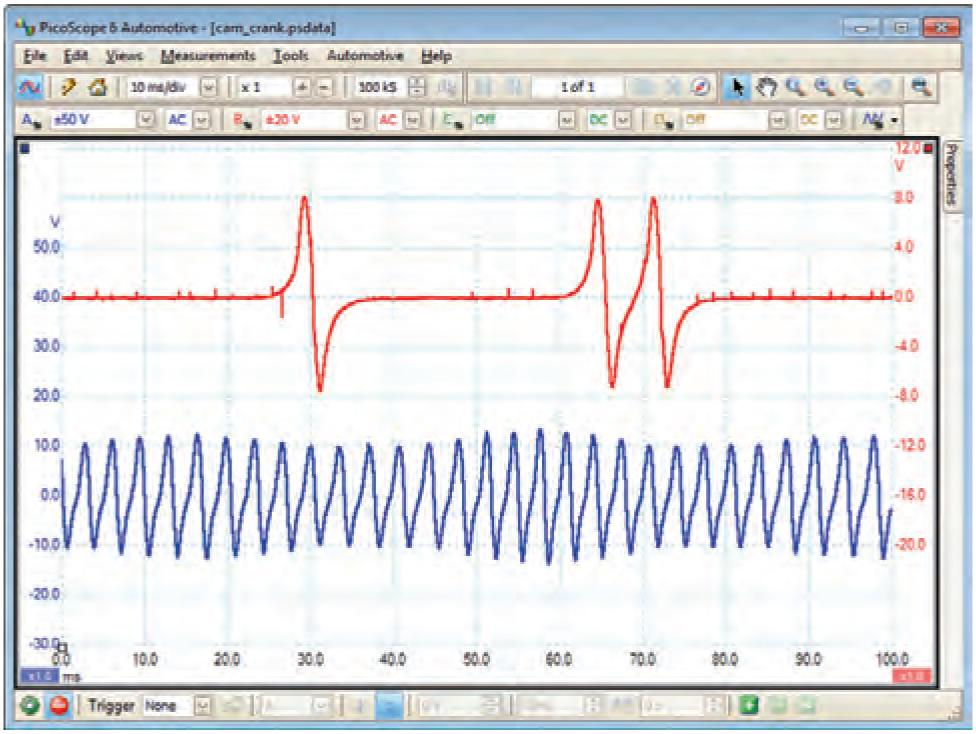


Figure 4.2 Crank and cam sensor output signals

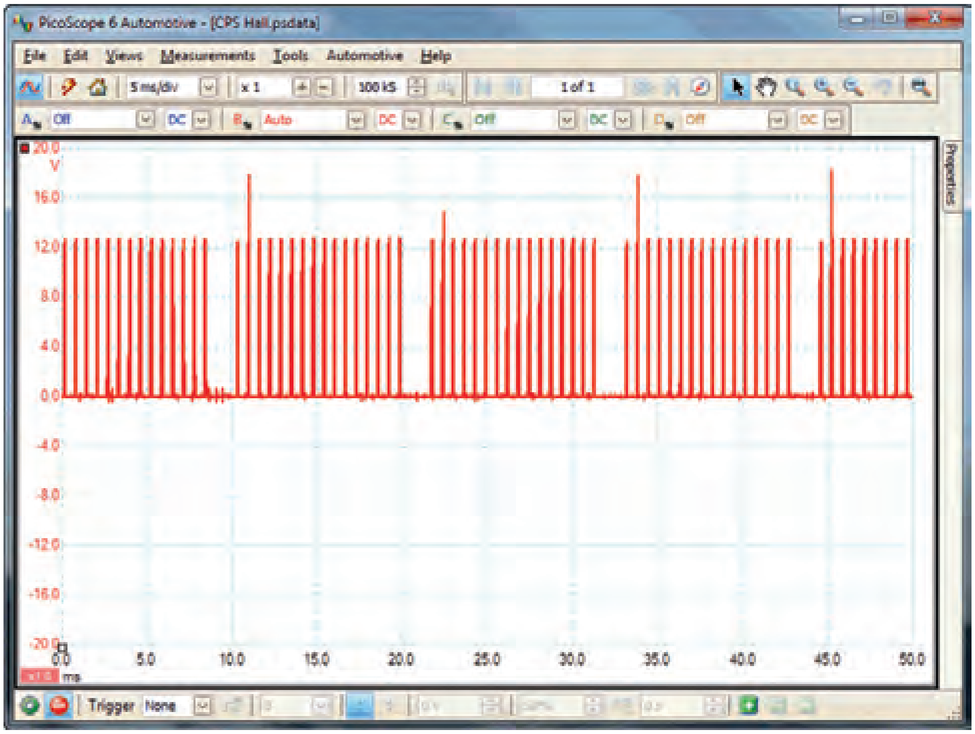


Figure 4.3 Hall effect crank sensor

The voltage will be approximately 0.5 V peak to peak while the engine is cranking, rising to around 2.5 V peak to peak at idle.

Some crankshaft sensors (CAS) are now Hall effect types and will therefore show a broadly square wave signal (Figure 4.3).

Figure 4.4 ABS wheel speed sensors (Source: Bosch Press)

4.2.2.2 ABS speed sensor

The ABS wheel speed sensors have become smaller and more efficient in the course of time. Recent models not only measure the speed and direction of wheel rotation but can be integrated into the wheel bearing as well (Figure 4.4).

ABS relies upon information coming in from the sensors to determine what action should be taken. If, under heavy braking, the ABS ECU loses a signal from one of the road wheels, it assumes that the wheel has locked and releases that brake momentarily until it sees the signal return. It is therefore imperative that the sensors are capable of providing a signal to the ABS ECU. If the signal produced from one wheel sensor is at a lower frequency than the others, the ECU may also react (Figures 4.5 and 4.6).

The operation of most ABS sensors is similar to that of a crank angle sensor (CAS). A small inductive pick-up is affected by the movement of a toothed wheel, which moves in close proximity. The movement of the wheel next to the sensor results in a 'sine wave'. The sensor, recognisable by its two electrical connections (some may have a coaxial braided outer shield), will produce an output that can be monitored and measured on the oscilloscope. Some are now Hall effect types so expect to see a square wave output.

Definition

CAS: Crank angle sensor

4.2.2.3 Inductive distributor pick-up

Not used on modern cars, but there are still plenty out there! The pick-up is used as a signal to trigger the ignition amplifier or an ECU. The sensor normally has two connections. If a third connection is used, it is normally a screen to reduce interference.

As a metal rotor spins, a magnetic field is altered, which induces an AC voltage from the pick-up. This type of pick-up could be described as a small alternator because the output voltage rises as the metal rotor approaches the winding,

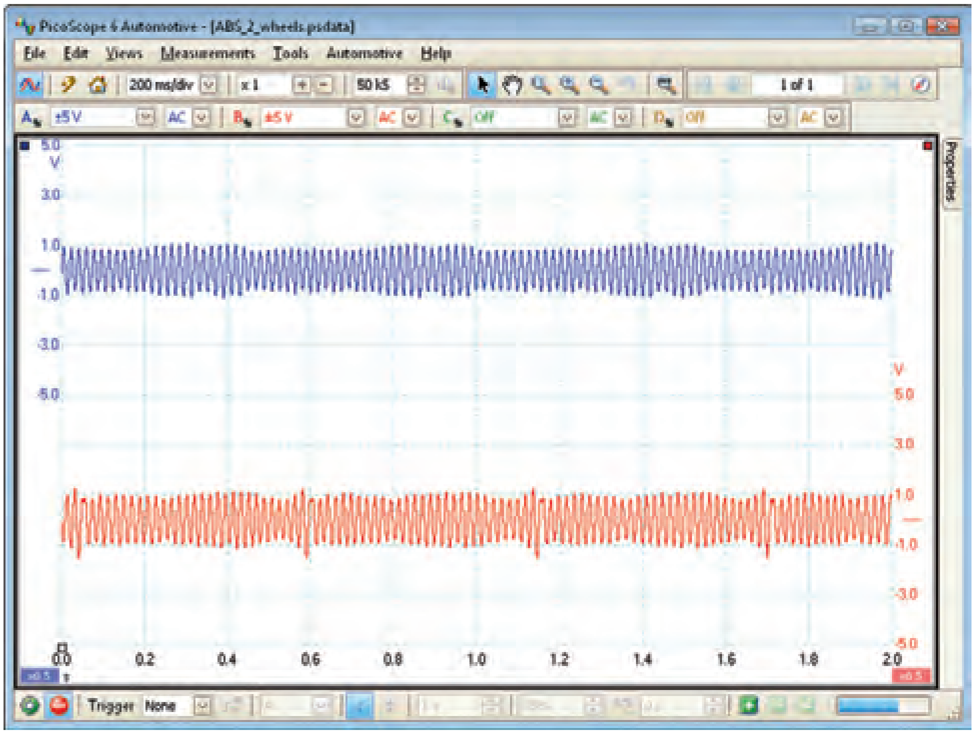


Figure 4.5 ABS speed sensor waveform

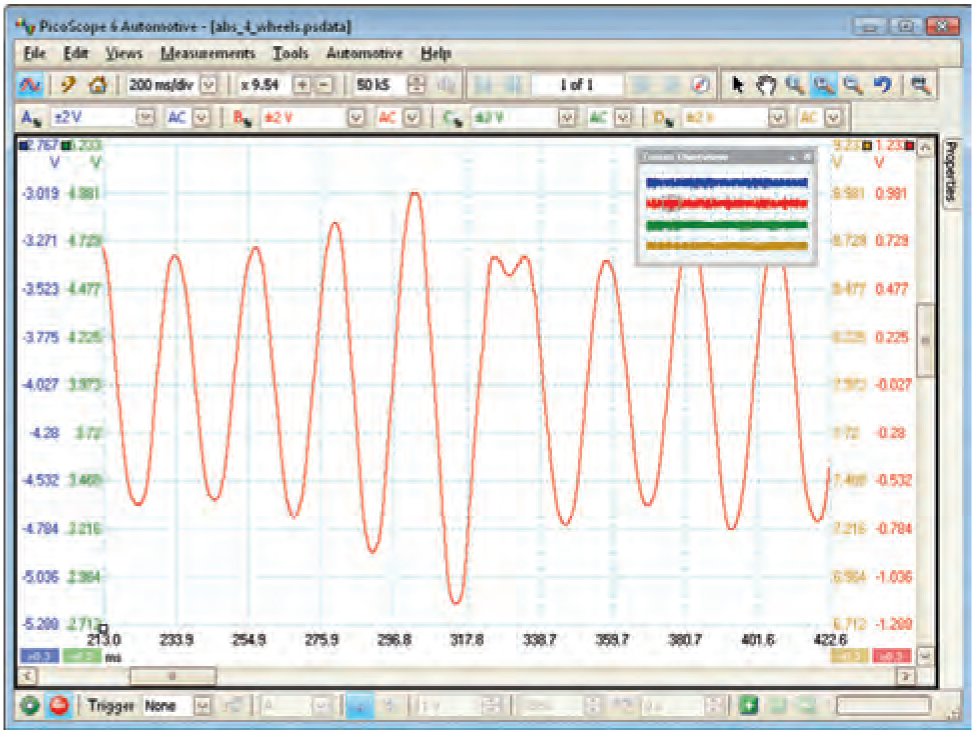


Figure 4.6 ABS speed sensor waveform zoomed in to show the effect of a broken tooth

sharply dropping through zero volts as the two components are aligned and producing a voltage in the opposite direction as the rotor passes. The waveform is similar to a sine wave; however, the design of the components is such that a more rapid switching is evident (Figure 4.7).

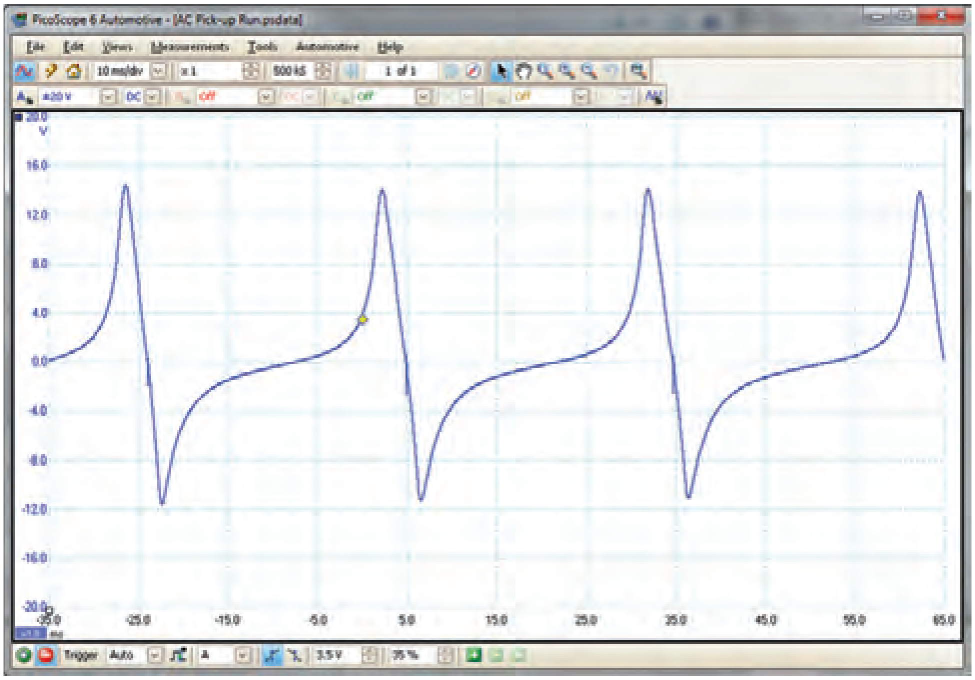


Figure 4.7 Inductive pick-up output signal (engine running)

The voltage produced by the pick-up will be determined by three main factors:

- Engine speed – the voltage produced will rise from as low as 2–3 V when cranking, to over 50 V at higher engine speeds.
- The proximity of the metal rotor to the pick-up winding – an average air gap will be in the order of 0.2–0.6 mm (8–14 thou), a larger air gap will reduce the strength of the magnetic field seen by the winding and the output voltage will be reduced.
- The strength of the magnetic field offered by the magnet – the strength of this magnetic field determines the effect it has as it ‘cuts’ through the windings, and the output voltage will be reduced accordingly.

A difference between the positive and the negative voltages may also be apparent as the negative side of the sine wave is sometimes attenuated (reduced) when connected to the amplifier circuit, but will produce perfect AC when disconnected and tested under cranking conditions.

4.2.3 Variable resistance

The two best examples of vehicle applications for variable resistance sensors are the throttle position sensor and the flap-type airflow sensor. Although variable capacitance sensors are used to measure small changes, variable resistance sensors generally measure larger changes in position. This is due to lack of sensitivity inherent in the construction of the resistive track. The throttle position sensor is a potentiometer in which, when supplied with a stable voltage, often 5 V, the voltage from the wiper contact will be proportional to throttle position. The throttle potentiometer is mostly used to indicate rate of change of throttle position. This information is used when implementing acceleration enrichment or overrun fuel cut-off. The output voltage of a rotary potentiometer is proportional to its position.

Key fact

Variable capacitance sensors are used to measure small changes; variable resistance sensors generally measure larger changes in position.

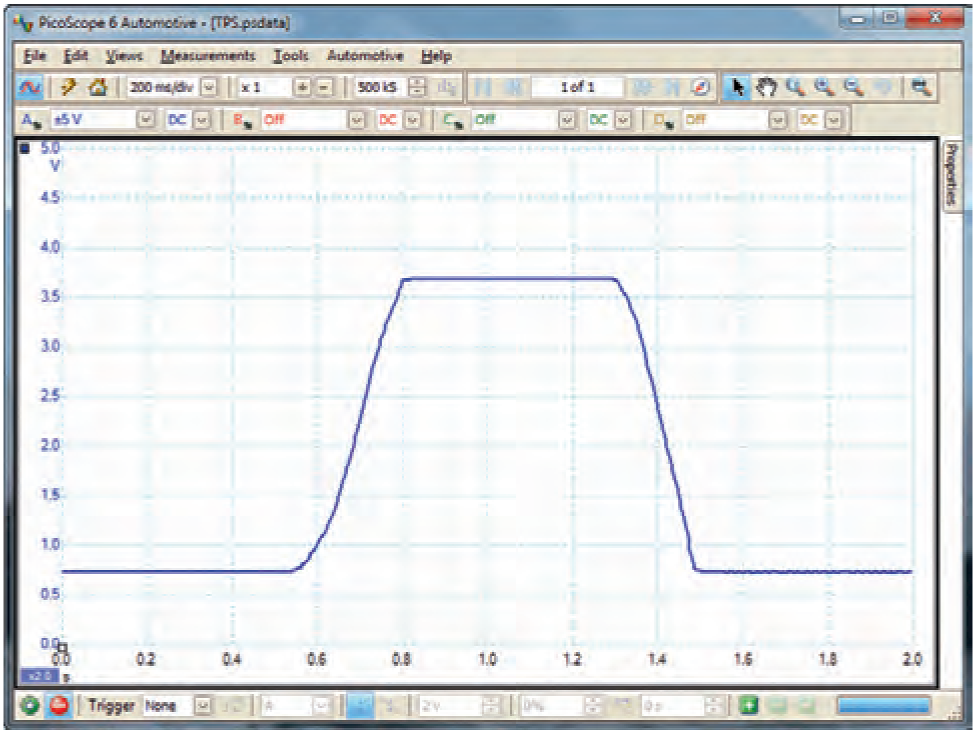


Figure 4.8 Throttle pot output voltage

The airflow sensor shown in Figure 4.9 works on the principle of measuring the force exerted on the flap by the air passing through it. A calibrated coil spring exerts a counter force on the flap such that the movement of the flap is proportional to the volume of air passing through the sensor. To reduce the fluctuations caused by individual induction strokes, a compensation flap is connected to the sensor flap. The fluctuations therefore affect both flaps and are cancelled out. Any damage due to back firing is also minimised due to this design. The resistive material used for the track is a ceramic metal mixture, which is burnt into a ceramic plate at very high temperature. The slider potentiometer is calibrated such that the output voltage is proportional to the quantity of inducted air. This sensor type is not used on modern vehicles as more accurate measurement is possible using other techniques.

4.2.3.1 Throttle position potentiometer

This sensor or potentiometer is able to indicate to the ECU the exact amount of throttle opening due to its linear output.

The majority of modern management systems use this type of sensor. It is located on the throttle butterfly spindle. The 'throttle pot' is a three-wire device having a 5 V supply (usually), an earth connection and a variable output from the centre pin. As the output is critical to the vehicle's performance, any 'blind spots' within the internal carbon track's swept area, will cause 'flat spots' and 'hesitations'. This lack of continuity can be seen on an oscilloscope (Figure 4.8).

A good throttle potentiometer should show a small voltage at the throttle closed position, gradually rising in voltage as the throttle is opened and returning back to its initial voltage as the throttle is closed. Although many throttle position sensor voltages will be manufacturer specific, many are non-adjustable and the voltage will be in the region of 0.5–1.0 V at idle, rising to 4.0 V (or more) with a

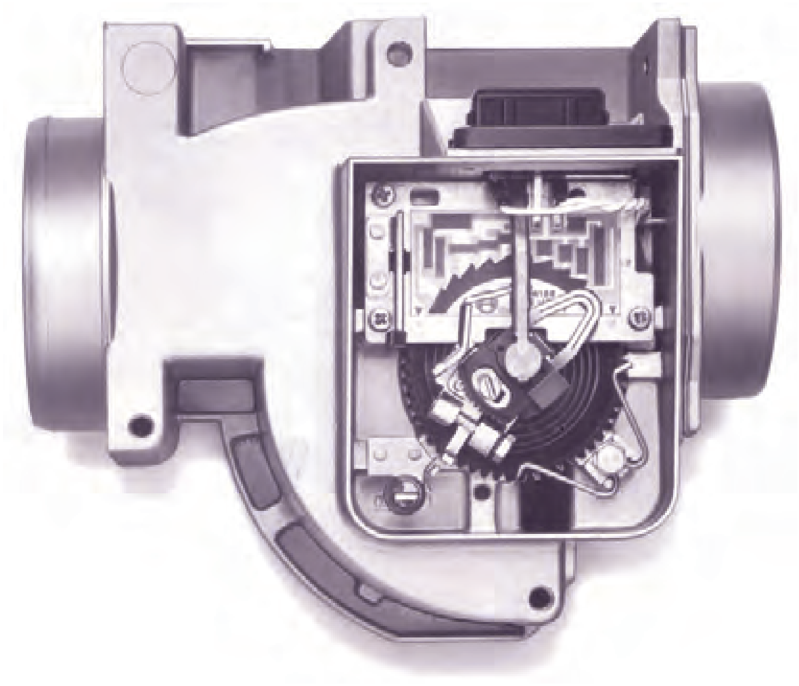


Figure 4.9 Vane- or flap-type airflow sensor(Source: Bosch)

fully opened throttle. For the full operational range, an oscilloscope time scale around two seconds is used.

4.2.3.2 Airflow meter – air vane

The vane-type airflow meter is a simple potentiometer that produces a voltage output that is proportional to the position of a vane. The vane in turn positions itself proportional to the amount of air flowing (Figure 4.9).

The voltage output from the internal track of the airflow meter should be linear to flap movement; this can be measured on an oscilloscope and should look similar to the example shown in Figure 4.10.

The waveform should show approximately 1.0 V when the engine is at idle; this voltage will rise as the engine is accelerated and will produce an initial peak. This peak is due to the natural inertia of the air vane and drops momentarily before the voltage is seen to rise again to a peak of approximately 4.0–4.5 V. This voltage will, however, depend on how hard the engine is accelerated, so a lower voltage is not necessarily a fault within the airflow meter. On deceleration, the voltage will drop sharply as the wiper arm, in contact with the carbon track, returns back to the idle position. This voltage may in some cases 'dip' below the initial voltage before returning to idle voltage. A gradual drop will be seen on an engine fitted with an idle speed control valve (ISCV) as this will slowly return the engine back to base idle as an anti-stall characteristic.

A time base of approximately two seconds plus is used; this enables the movement to be shown on one screen, from idle through acceleration and back to idle again. The waveform should be clean with no 'drop-out' in the voltage, as this indicates a lack of electrical continuity. This is common on an airflow meter with a dirty or faulty carbon track. The problem will appear as a 'flat spot' or hesitation when the vehicle is driven, this is a typical problem on vehicles with high mileage that have spent the majority of their working life with the throttle in one position.

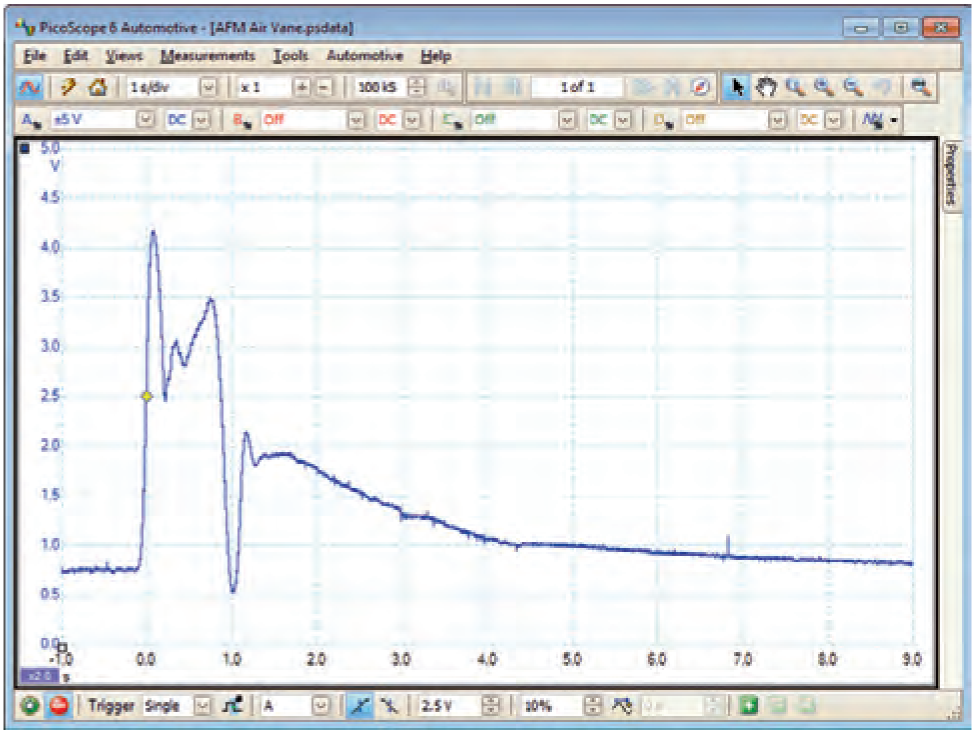


Figure 4.10 Air vane output voltage

4.2.4 Hot wire airflow sensor

The advantage of this sensor is that it measures air mass flow. The basic principle is that as air passes over a hot wire, it tries to cool the wire down. If a circuit is created such as to increase the current through the wire, then this current will be proportional to the airflow. A resistor is also incorporated to compensate for temperature variations. The 'hot wire' is made of platinum and is only a few millimetres long and approximately 70 μm thick. Because of its small size, the time constant of the sensor is very short, in fact in the order of a few milliseconds. This is a great advantage as any pulsations of the airflow will be detected and reacted to in a control unit accordingly.

The output of the circuit involved with the hot wire sensor is a voltage across a precision resistor. The resistance of the hot wire and the precision resistor are such that the current to heat the wire varies between 0.5 and 1.2 A with different air mass flow rates. High-resistance resistors are used in the other arm of the bridge and so current flow is very small. The temperature-compensating resistor has a resistance of approximately 500 Ω, which must remain constant other than by way of temperature change. A platinum film resistor is used for these reasons. The compensation resistor can cause the system to react to temperature changes within about three seconds.

The output of this device can change if the hot wire becomes dirty. Heating the wire to a very high temperature for one second every time the engine is switched off prevents this, by burning off any contamination. In some air mass sensors, a variable resistor is provided to set idle mixture. The nickel film airflow sensor is similar to the hot wire system. Instead of a hot platinum wire, a thin film of nickel is used. The response time of this system is even shorter than the hot wire. The advantage which makes a nickel thick-film thermistor ideal for inlet air

Key fact

A nickel thick-film thermistor is ideal for inlet air temperature sensing because of its very short time constant.

Figure 4.11 Hot wire air mass meter (Source: Bosch Press)

temperature sensing is its very short time constant. In other words, its resistance varies very quickly with a change in air temperature.

4.2.4.1 Airflow meter – hot wire

Figure 4.11 shows a mass airflow sensor from Bosch. This type has been in use since 1996. As air flows over the hot wire, it cools it down, and this produces the output signal. The sensor measures air mass because the air temperature is taken into account due to its cooling effect on the wire.

The voltage output should be linear to airflow. This can be measured on an oscilloscope and should look similar to the example shown in Figure 4.12. The waveform should show approximately 1.0 V when the engine is at idle. This voltage will rise as the engine is accelerated and air volume is increased producing an initial peak. This peak is due to the initial influx of air and drops momentarily before the voltage is seen to rise again to another peak of approximately 4.0–4.5 V. This voltage will, however, depend on how hard the engine is accelerated; a lower voltage is not necessarily a fault within the meter.

On deceleration, the voltage will drop sharply as the throttle butterfly closes, reducing the airflow, and the engine returns back to idle. The final voltage will drop gradually on an engine fitted with ISCV as this will slowly return the engine back to base idle as an anti-stall characteristic. This function normally only effects the engine speed from around 1200 rpm back to the idle setting.

A time base of approximately two seconds plus is used because this allows the output voltage on one screen, from idle through acceleration and back to idle again. The ‘hash’ on the waveform is due to airflow changes caused by the induction pulses as the engine is running.

4.2.5 Thermistors

Thermistors are the most common device used for temperature measurement on the motor vehicle. The principle of measurement is that a change in temperature

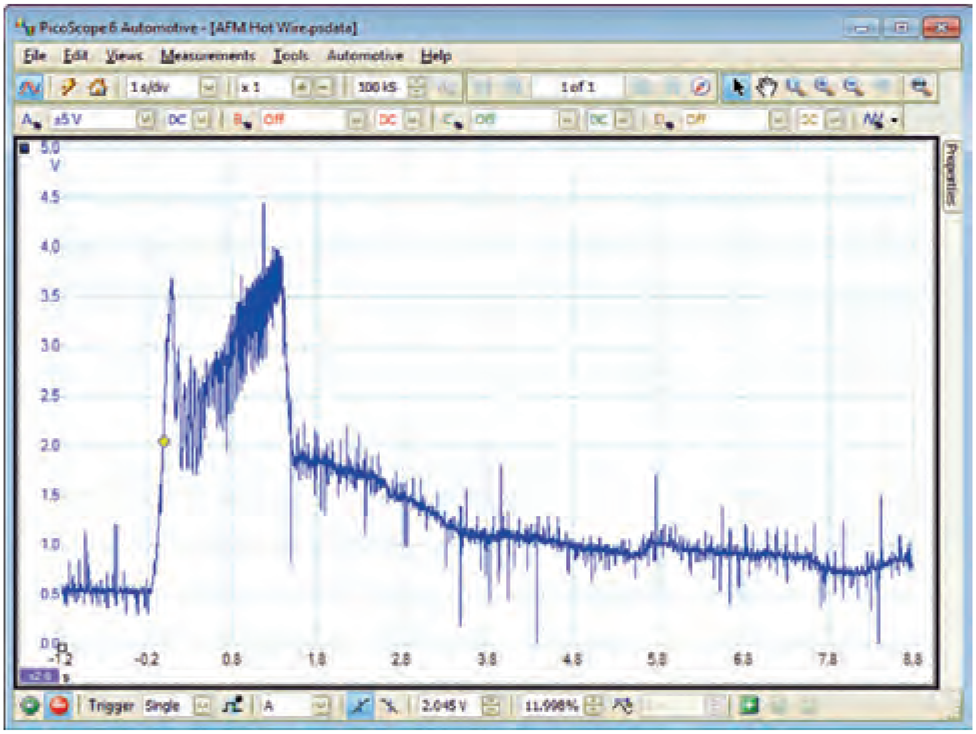


Figure 4.12 Air mass hot wire waveform

Figure 4.13 Temperature sensor

will cause a change in resistance of the thermistor and hence an electrical signal proportional to the temperature being measured. Most thermistors in common use are of the negative temperature coefficient (NTC) type. The actual response of the thermistors can vary but typical values for those used on motor vehicles will vary from several kilo-ohms at 0 °C to few hundred ohms at 100 °C. The large change in resistance for a small change in temperature makes the thermistor ideal for most vehicles uses. It can also be easily tested with simple equipment. Thermistors are constructed of semiconductor materials. The change in resistance with a change in temperature is due to the electrons being able to break free more easily at higher temperatures.

4.2.5.1 Coolant temperature sensor

Most coolant temperature sensors (CTS) are NTC thermistors; their resistance decreases as temperature increases. This can be measured on most systems as a reducing voltage signal.

The CTS is usually a two-wire device with a voltage supply of approximately 5 V (Figure 4.13).

The resistance change will therefore alter the voltage seen at the sensor and can be monitored for any discrepancies across its operational range. By selecting a time scale of 500 seconds and connecting the oscilloscope to the sensor, the output voltage can be monitored. Start the engine and in the majority of cases the voltage will start in the region of 3–4 V and fall gradually. The voltage will depend on the temperature of the engine (Figure 4.14).

The rate of voltage change is usually linear with no sudden changes to the voltage, if the sensor displays a fault at a certain temperature, it will show up in this test.

Definition

NTC: Negative temperature coefficient.

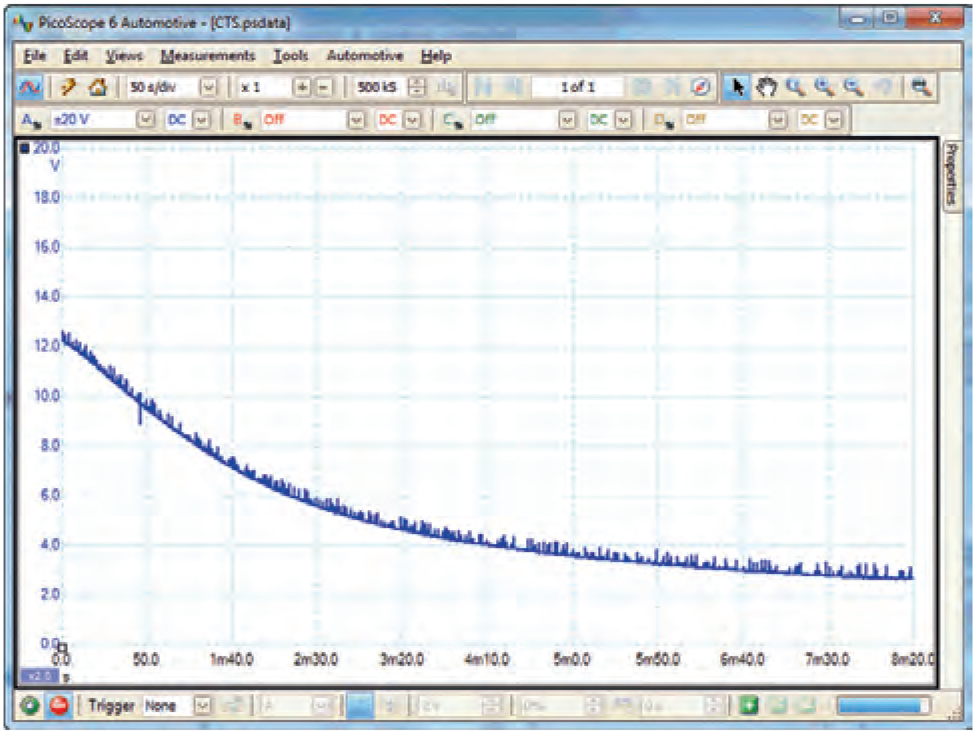


Figure 4.14 Decreasing voltage from the temperature sensor

4.2.6 Hall effect sensors

If a conductor is carrying a current in a transverse magnetic field, then a voltage will be produced at right angles to the supply current. This voltage is proportional to the supply current and to the magnetic field strength (Figure 4.15).

Many distributors employ Hall effect sensors, but they are now also used as rotational sensors for the crank and ABS, for example. The output of this sensor is almost a square wave with constant amplitude. The Hall effect can also be used to detect current flowing in a cable, the magnetic field produced round the cable being proportional to the current flowing. The Hall effect sensors are becoming increasingly popular because of their reliability and also because they produce a constant amplitude square wave in speed measurement applications and a varying DC voltage for either position sensing or current sensing.

The two main advantages are that measurement of lower (or even zero) speed is possible and that the voltage output of the sensors is independent of speed.

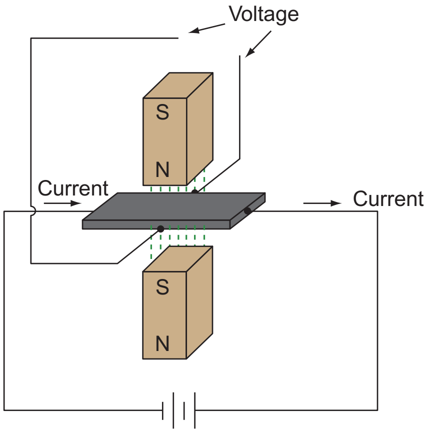


Figure 4.15 Hall effect

4.2.6.1 Hall effect distributor pick-up

Hall sensors are used in a number of ways. The ignition distributor was very common but they are not used now (Figure 4.16).

This form of trigger device is a simple digital ‘on/off switch’ which produces a square wave output that is recognised and processed by the ignition control module or engine management ECU (Figure 4.17).

The trigger has a rotating metal disc with openings that pass between an electromagnet and the semiconductor (Hall chip). This action produces a square wave that is used by the ECU or amplifier.

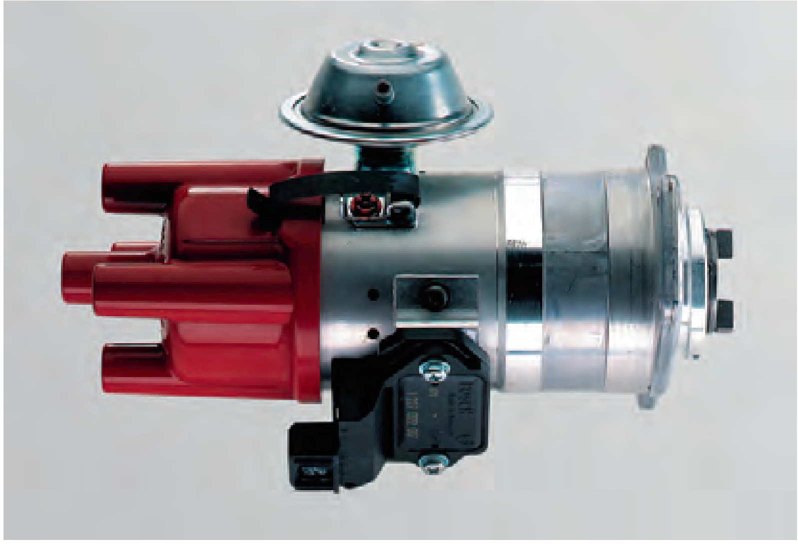


Figure 4.16 Distributors usually contain a Hall effect or inductive pulse generator (Source: Bosch Press)

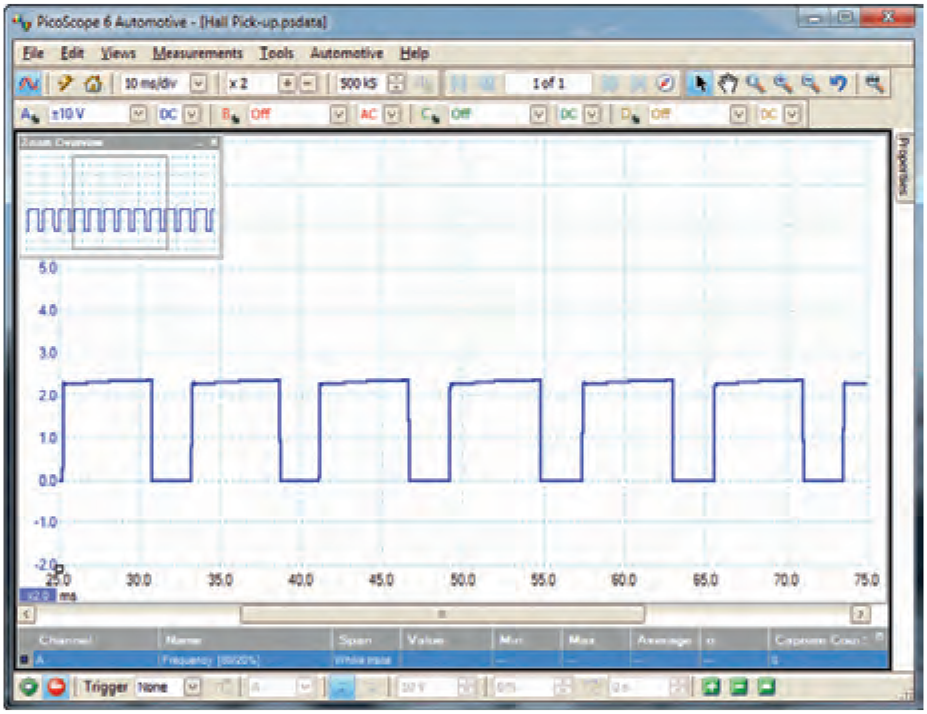


Figure 4.17 Hall output waveform

4.2.6.2 ABS Hall sensor

The sensor when used by ABS for monitoring wheel speed and as transmission speed sensors works using the same effect.

The sensor will usually have three connections: a stabilised supply voltage (often 4 or 5 V), an earth and the output signal. The square wave when monitored on an oscilloscope may vary a little in amplitude; this is not usually a problem as it is the frequency that is important. However, in most cases, the amplitude/voltage will remain constant. Table 4.2 provides a list of technical data for the sensor shown in Figure 4.18.

Table 4.2 Hall sensor data

Supply voltage	4.5 V_{DC}
Nominal sensing distance	1.5 mm
Current (typical)	10 mA
Current (max)	20 mA
Weight	30 g
Temperature range	−30 to +130 °C
Tightening torque	6 Nm
Output	PNP
Output sink voltage	0.4 V_{max}
Trigger type	Ferrous

Figure 4.18 Hall effect ABS or CAS sensor

4.2.6.3 Road speed sensor (Hall effect)

To measure the output of this sensor, jack up the driven wheels of the vehicle and place on axle stands on firm level ground. Run the engine in gear and then probe each of the three connections (+, − and signal) (Figure 4.19).

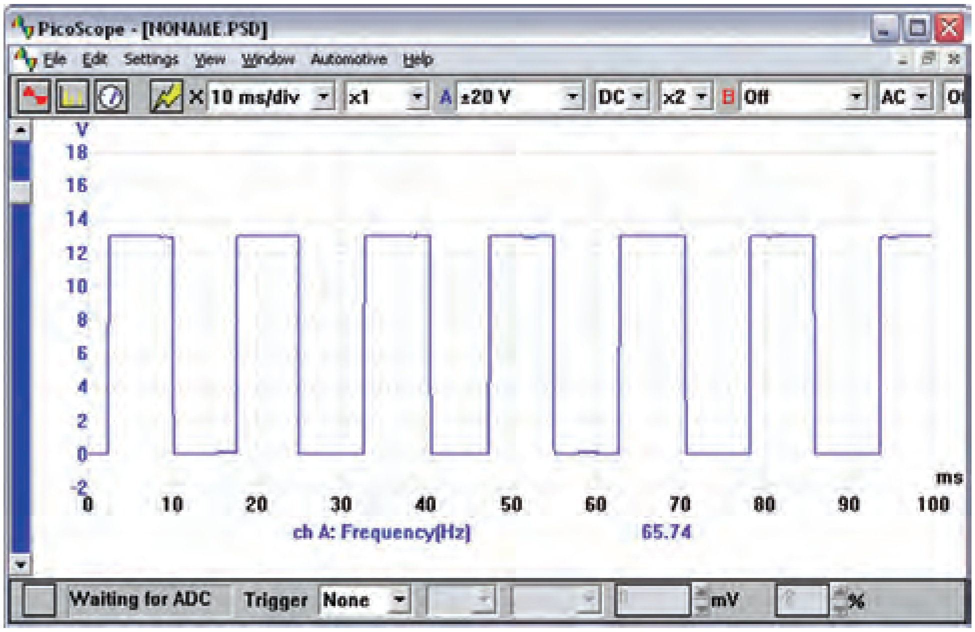
As the road speed is increased, the frequency of the switching should be seen to increase. This change can also be measured on a multimeter with frequency capabilities. The sensor will be located on either the speedometer drive output from the gearbox or to the rear of the speedometer head if a speedo cable is used. The signal is used by the engine ECU and, if appropriate, the transmission ECU. The actual voltage will vary with sensor design.

Definition

Piezoelectric effect: The production of electrical potential in a substance as the pressure on it changes.

4.2.7 Piezo accelerometer

A piezoelectric accelerometer is a seismic mass accelerometer using a piezoelectric crystal to convert the force on the mass due to acceleration into



X 10 ms/div x1 A ±20 V DC x2 B Off AC

ch A: Frequency(Hz) 65.74

Waiting for ADC Trigger None mV %

Figure 4.19 Hall effect road speed sensor waveform

Figure 4.20 Knock sensor

an electrical output signal. The crystal not only acts as the transducer but as the suspension spring for the mass. The crystal is sandwiched between the body of the sensor and the seismic mass and is kept under compression. Acceleration forces acting on the seismic mass cause variations in the amount of crystal compression and hence generate the piezoelectric voltage.

4.2.7.1 Knock sensor

The sensor, when used as an engine knock sensor, will also detect other engine vibrations. These are kept to a minimum by only looking for ‘knock’, a few degrees before and after top dead centre (TDC). Unwanted signals are also filtered out electrically.

The optimal point at which the spark ignites the air/fuel mixture is just before knocking occurs. However, if the timing is set to this value, under certain conditions knock (detonation) will occur. This can cause serious engine damage as well as increase emissions and reduce efficiency.

A knock sensor is used by some engine management systems (Figure 4.20). When coupled with the ECU, it can identify when knock occurs and retard the ignition timing accordingly.

The frequency of knocking is approximately 15 kHz. As the response of the sensor is very fast, an appropriate time scale must be set, in the case of the example waveform a 0 to 500 ms timebase and a −5 to +5 V voltage scale. The best way to test a knock sensor is to remove the knock sensor from the engine and to tap it with a small spanner – the resultant waveform should be similar to the example shown in Figure 4.21.

Note: When refitting the sensor, tighten to the correct torque setting as overtightening can damage the sensor and/or cause it to produce incorrect signals.

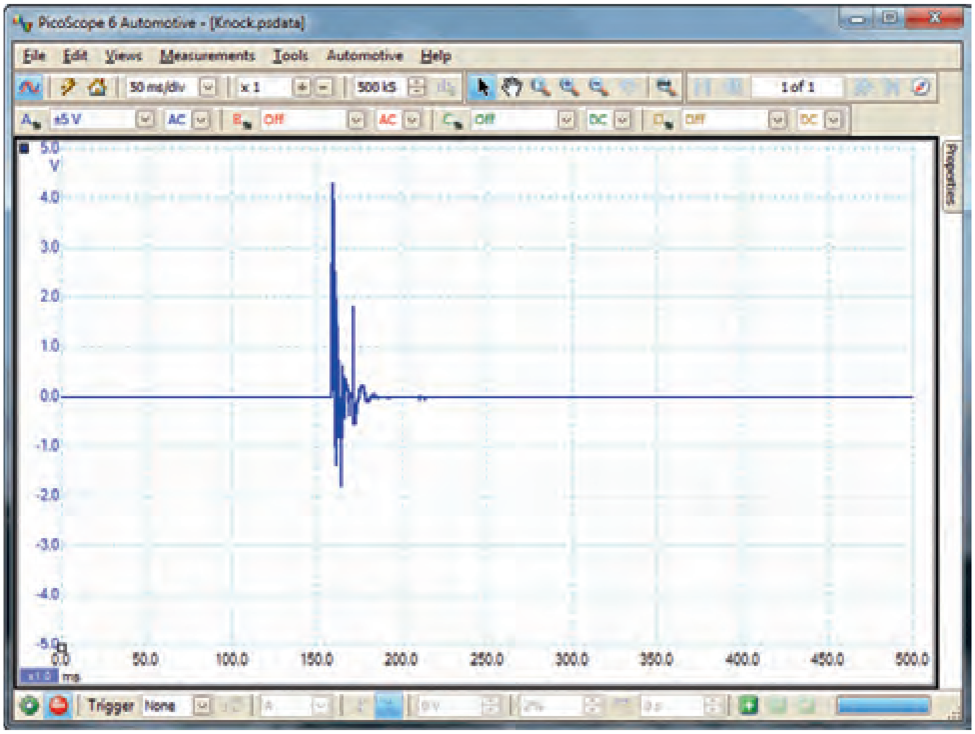


Figure 4.21 Knock sensor output signal

4.2.8 Oxygen sensors

The vehicle application for an oxygen sensor is to provide a closed loop feedback system for engine management control of the air/fuel ratio. The amount of oxygen sensed in the exhaust is directly related to the mixture strength or air/fuel ratio. The ideal air/fuel ratio of 14.7:1 by mass is known as a lambda (λ) value of 1 (Figure 4.22).

Exhaust gas oxygen (EGO) sensors are placed in the exhaust pipe near to the manifold to ensure adequate heating. The sensors operate best at temperatures over 300 °C. In some cases, a heating element is incorporated to ensure that this temperature is reached quickly. This type of sensor is known as a heated exhaust gas oxygen sensor (HEGO). The heating element (which consumes approximately 10 W) does not operate all the time to ensure that the sensor does not exceed 850 °C, at which temperature damage may occur to the sensor. This is why the sensors are not fitted directly in the exhaust manifold. The main active component of most types of oxygen sensors is zirconium dioxide (ZrO_2). This ceramic is housed in gas permeable electrodes of platinum. A further ceramic coating is applied to the side of the sensor exposed to the exhaust gas as a protection against residue from the combustion process. The principle of operation is that at temperatures in excess of 300 °C, the ZrO_2 will conduct the negative oxygen ions. The sensor is designed to be responsive very close to a lambda value of one. As one electrode of the sensor is open to a reference value of atmospheric air, a greater quantity of oxygen ions will be present on this side. Because of electrolytic action, these ions permeate the electrode and migrate through the electrolyte (ZrO_2). This builds up a charge rather like a battery.

The size of the charge is dependent on the oxygen percentage in the exhaust. The closely monitored closed loop feedback of a system using lambda sensing

Key fact

Most lambda sensors operate best at temperatures over 300°C.

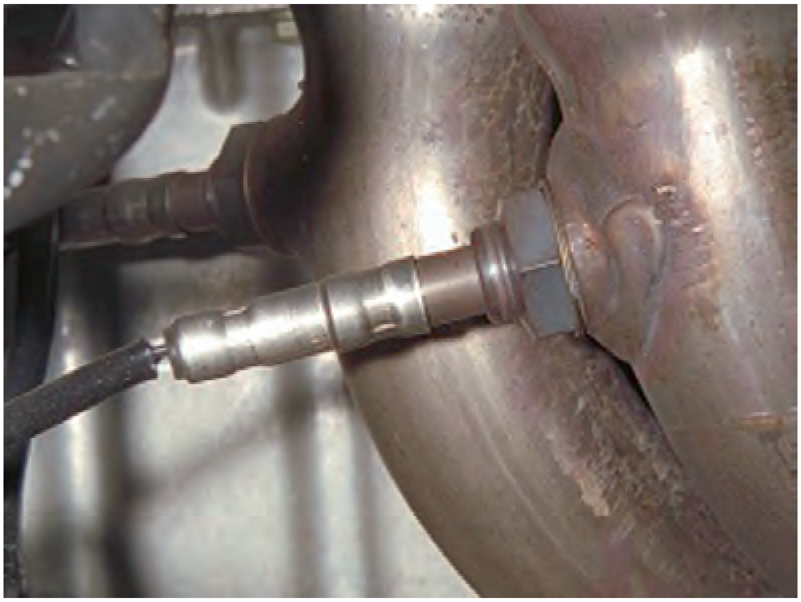


Figure 4.22 Lambda sensor in the exhaust downpipe

Figure 4.23 Titania knock sensor in position

allows very accurate control of engine fuelling. Close control of emissions is therefore possible.

4.2.8.1 Oxygen sensor (Titania)

The lambda sensor, also referred to as the oxygen sensor, plays a very important role in the control of exhaust emissions on a catalyst equipped vehicle (Figure 4.23).

The main lambda sensor is fitted into the exhaust pipe before the catalytic converter. The sensor will have four electrical connections. It reacts to the oxygen content in the exhaust system and will produce an oscillating voltage between 0.5 (lean) and 4.0 V, or above (rich) when running correctly. A second sensor to monitor the catalyst performance may be fitted downstream of the converter.

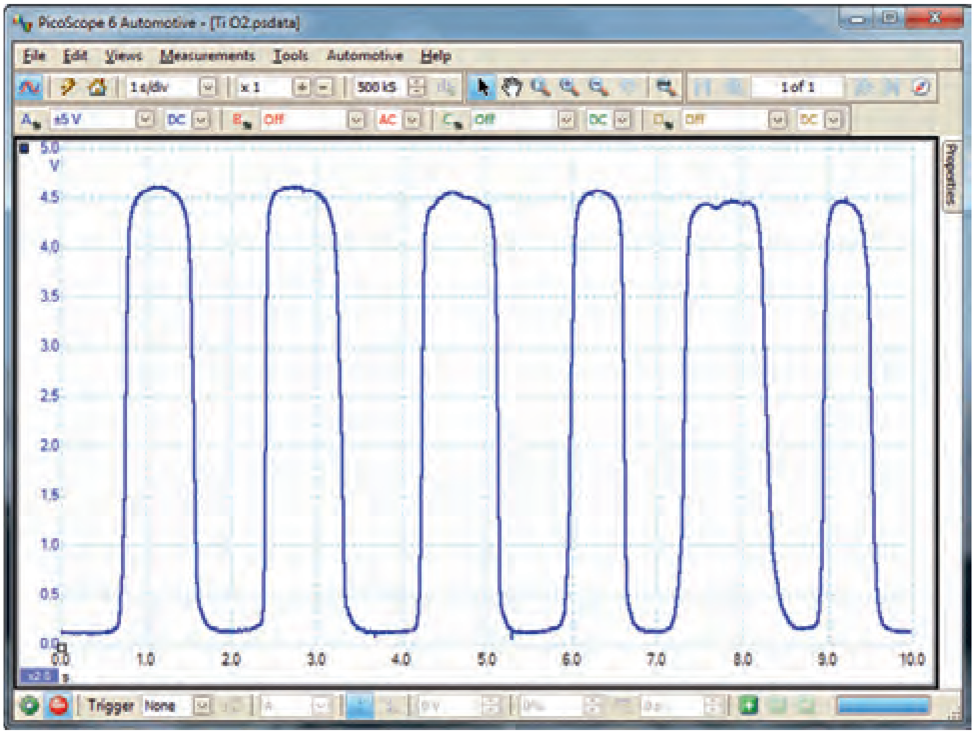


Figure 4.24 Titania lambda sensor output

Titania sensors, unlike Zirconia sensors, require a voltage supply as they do not generate their own voltage. A vehicle equipped with a lambda sensor is said to have ‘closed loop’, this means that after the fuel has been burnt up during the combustion process, the sensor will analyse the emissions and adjust the engine’s fuelling accordingly.

Titania sensors have a heating element to assist the sensor reaching its optimum operating temperature. The sensor when working correctly will switch approximately once per second (1 Hz) but will only start to switch when at normal operating temperature. This switching can be seen on the oscilloscope, and the waveform should look similar to the one in the example (Figure 4.24).

4.2.8.2 Oxygen sensor (Zirconia)

The sensor will have varying electrical connections and may have up to four wires. It reacts to the oxygen content in the exhaust system and will produce a small voltage depending on the air/fuel mixture seen at the time. The voltage range seen will, in most cases, vary between 0.2 and 0.8 V. The 0.2 V indicates a lean mixture and a voltage of 0.8 V shows a richer mixture (Figure 4.25).

Lambda sensors can have a heating element to assist the sensor reaching its optimum operating temperature. Zirconia sensors when working correctly will switch approximately once per second (1 Hz) and will only start to switch when at normal operating temperature. This switching can be seen on the oscilloscope, and the waveform should look similar to the one in the example waveform (Figure 4.26).

Many vehicles now have a pre- and post-cat lambda sensor. Comparing the outputs of these two sensors is a good indicator of catalyst operation and condition (Figure 4.27).

Key fact

Many vehicles now have a pre- and post-cat lambda sensor

Figure 4.25 Zirconia-type oxygen sensor

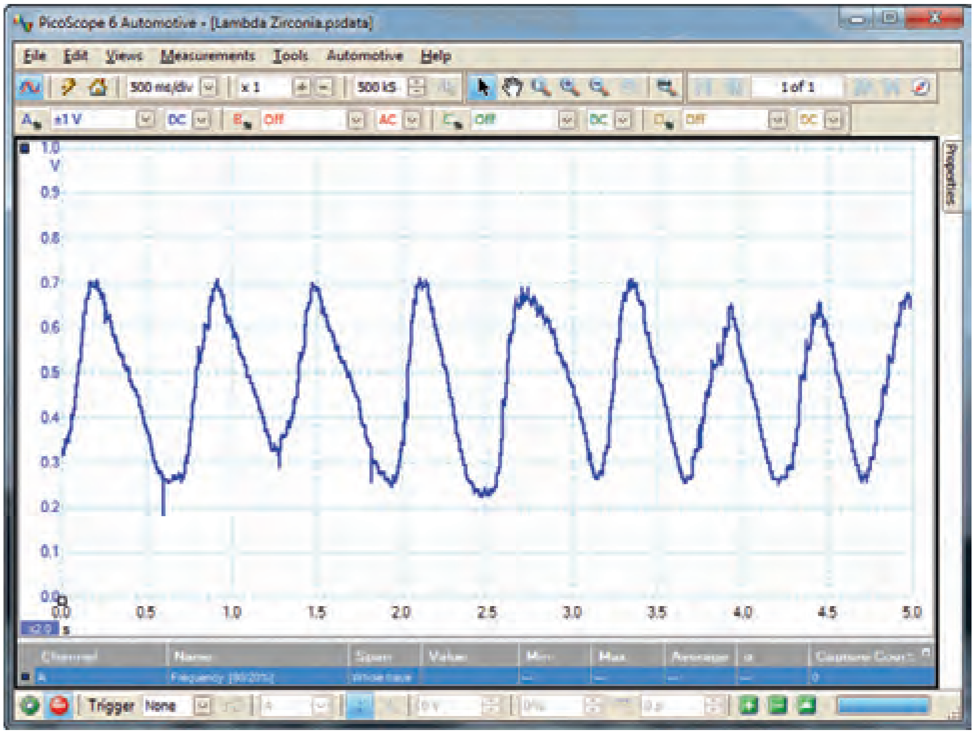


Figure 4.26 Zirconia oxygen sensor output

4.2.9 Pressure sensors

4.2.9.1 Strain gauges

When a strain gauge is stretched its resistance will increase, and when it is compressed its resistance decreases. Most strain gauges consist of a thin layer of film that is fixed to a flexible backing sheet. This in turn is bonded to the part where strain is to be measured. Most resistance strain gauges have a resistance of approximately 100 Ω.

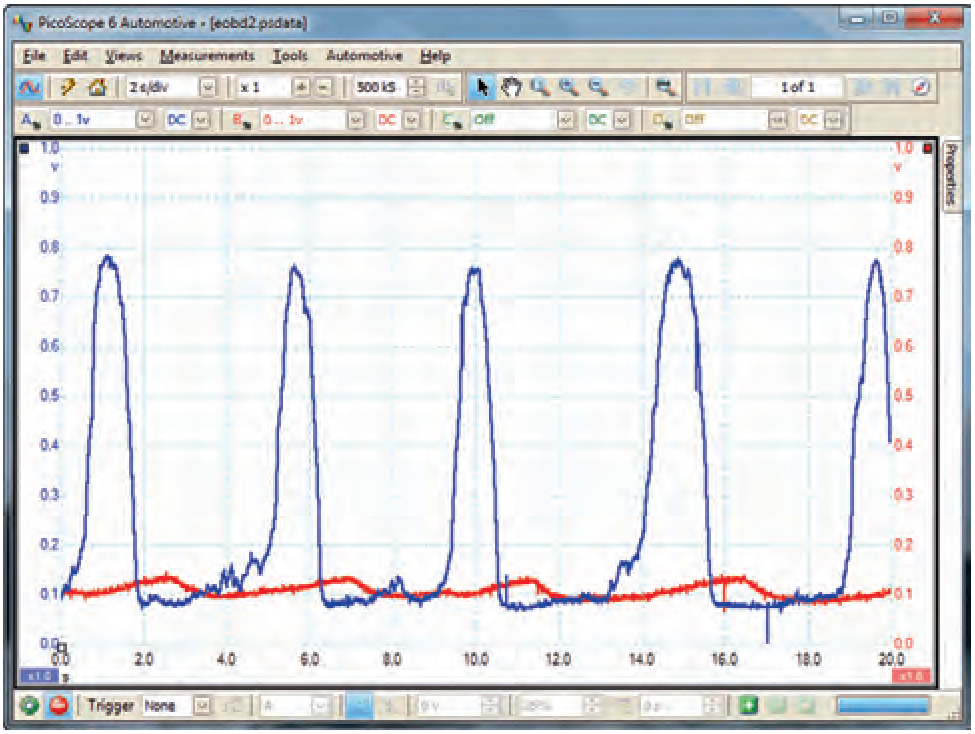


Figure 4.27 Pre-cat signal shown in blue and post-cat in red

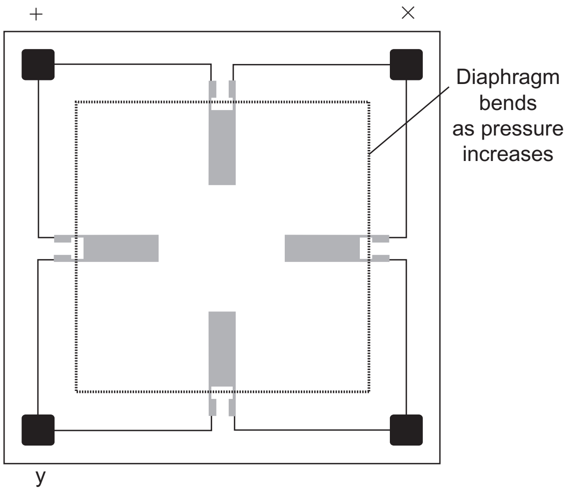
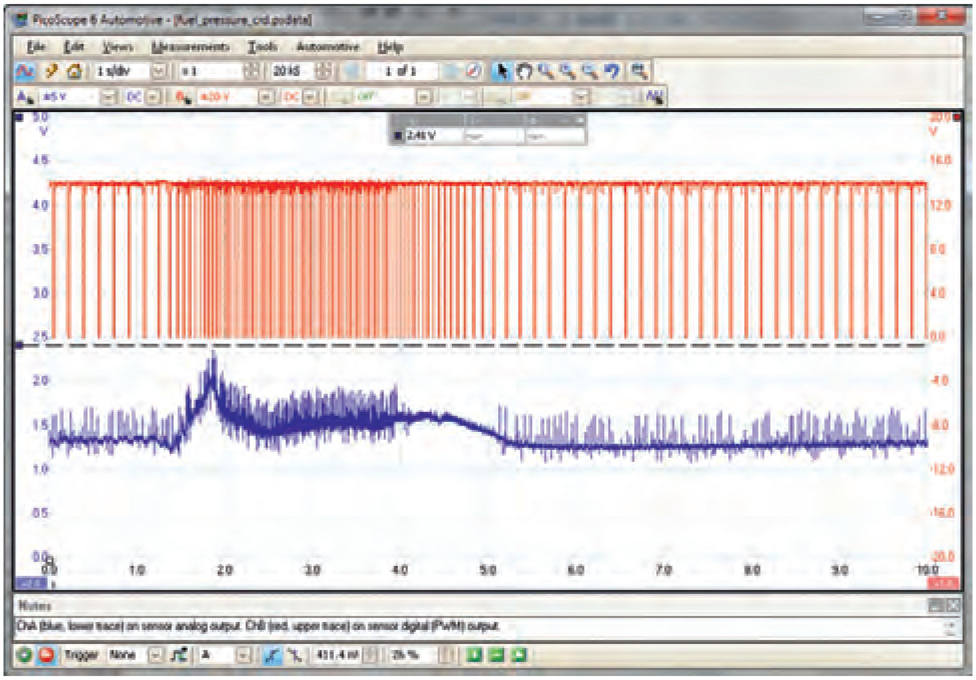


Figure 4.28 Strain gauge pressure sensor

Strain gauges are often used indirectly to measure engine manifold pressure. Figure 4.28 shows an arrangement of four strain gauges on a diaphragm forming part of an aneroid chamber used to measure pressure. When changes in manifold pressure act on the diaphragm, the gauges detect the strain. The output of the circuit is via a differential amplifier, which must have a very high input resistance so as not to affect the bridge balance. The actual size of this sensor may be only a few millimetres in diameter. Changes in temperature are compensated for by using four gauges which when affected in a similar way cancel out any changes.



ChA (blue, lower trace) on sensor analog output. ChB (red, upper trace) on sensor digital (PWM) output.

Figure 4.29 Common rail fuel pressure, the blue trace is from the sensor analogue output the red trace is from sensor digital (PWM) output

Definition

PWM: Pulse width modulation is an adjustment of the duty cycle of a signal or power source, to either convey information over a communications channel or control the amount of power sent to a load (e.g. an actuator).

4.2.9.2 Fuel pressure

Common rail diesel system pressure signals can be tested (remember these systems operate at very high pressure). The pulse width modulation (PWM) signal should be at the same amplitude but the on/off ratio can vary (Figure 4.29).

4.2.9.3 Manifold absolute pressure

Manifold absolute pressure (MAP) is a signal used to determine engine load. There are two main types: analogue and digital. The analogue signal voltage output varies with pressure whereas the digital signal varies with frequency. These sensors use a piezo crystal, strain gauges or variable capacitance sensors or similar. The signals are also processed internally (Figures 4.30 and 4.31).

4.2.10 Variable capacitance

The value of a capacitor is determined by

- surface area of its plates;
- distance between the plates;
- the dielectric (insulation between the plates).

Sensors can be constructed to take advantage of these properties. Three sensors, each using the variable capacitance technique, are shown in Figure 4.32. These are (a) liquid level sensor in which the change in liquid level changes the dielectric value; (b) pressure sensor similar to the strain gauge pressure sensor in which the distance between capacitor plates changes; and (c) position sensor which detects changes in the area of the plates.

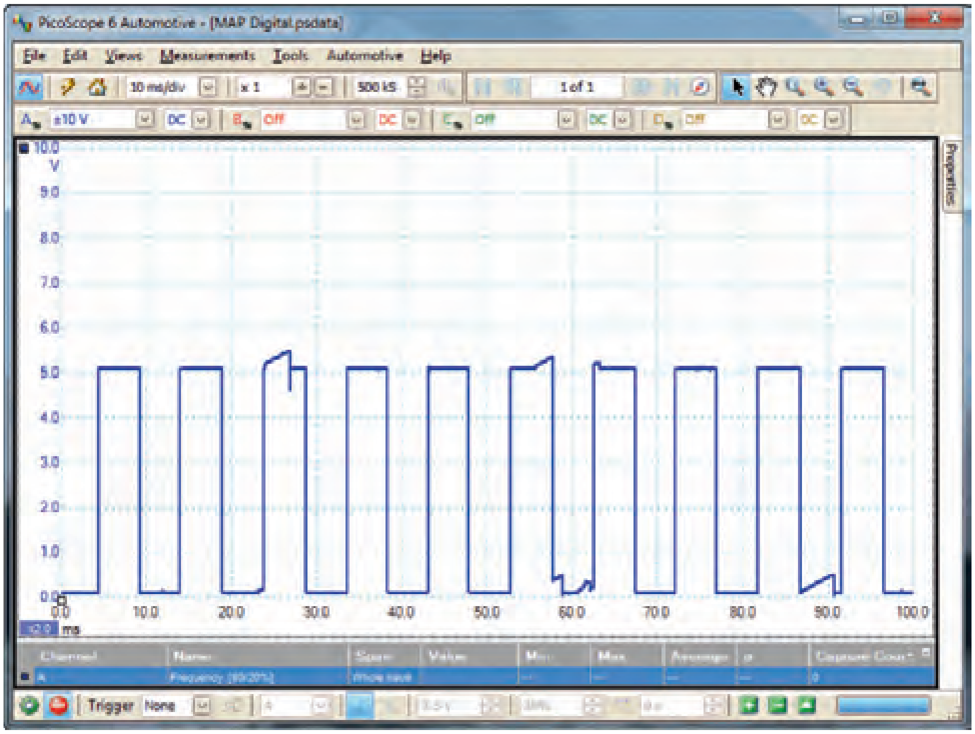


Figure 4.30 Digital MAP sensor

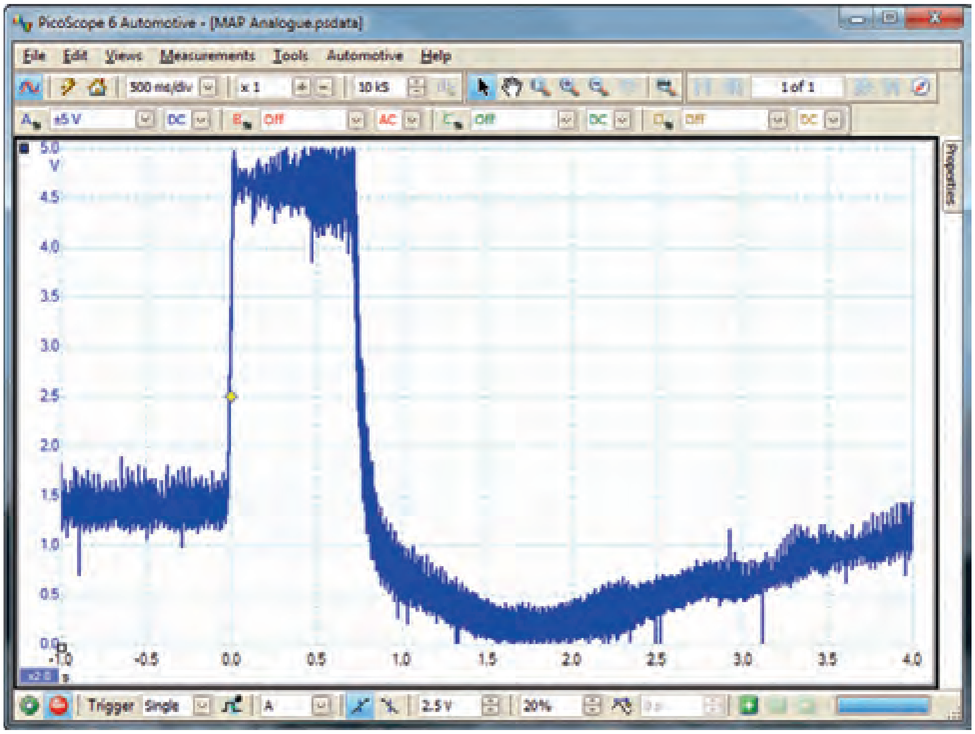


Figure 4.31 Analogue MAP sensor

4.2.10.1 Oil quality sensor

An interesting sensor used to monitor oil quality is now available, which works by monitoring changes in the dielectric constant of the oil. This value increases as antioxidant additives in the oil deplete. The value rapidly increases if coolant

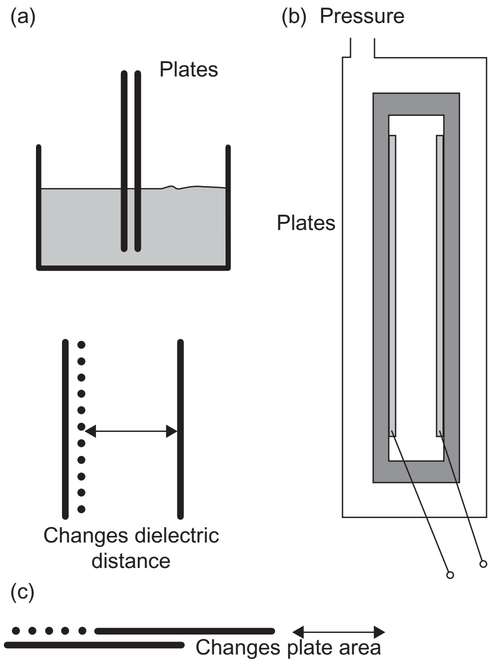


Figure 4.32 Variable capacitance sensors: (a) liquid level, (b) pressure and (c) position

Figure 4.33 Oil quality sensor (Source: Bosch Media)

contaminates the oil. The sensor output increases as the dielectric constant increases (Figure 4.33).

4.2.11 Optical sensors

An optical sensor for rotational position is a relatively simple device. The optical rotation sensor and circuit shown in Figure 4.34 consists of a phototransistor as a detector and a light emitting diode (LED) light source. If the light is focused to a very narrow beam then the output of the circuit shown will be a square wave with frequency proportional to speed.

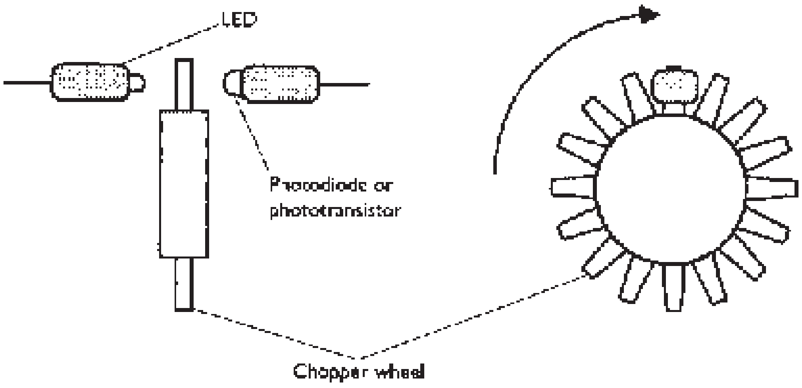


Figure 4.34 Optical sensor

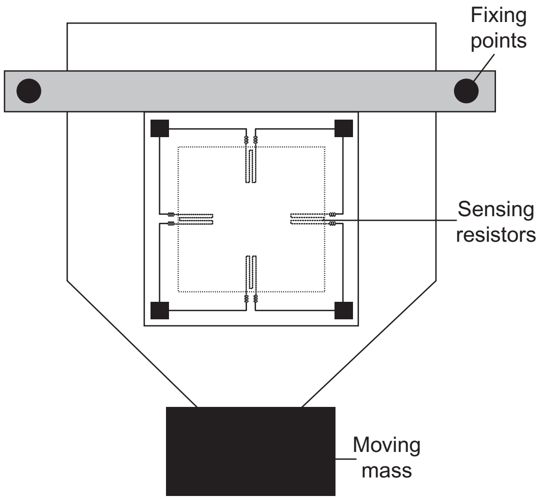


Figure 4.35 Strain gauge accelerometer

4.2.12 Dynamic position sensors

A dynamic position or movement of crash sensor can take a number of forms; these can be described as mechanical or electronic. The mechanical system works by a spring holding a roller in a set position until an impact (acceleration/deceleration) above a predetermined limit provides enough force to overcome the spring and the roller moves, triggering a micro switch. The switch is normally open with a resistor in parallel to allow the system to be monitored. Two switches similar to this may be used to ensure that an airbag is deployed only in the case of sufficient frontal impact.

Figure 4.35 is a further type of dynamic position sensor. Described as an accelerometer, it is based on strain gauges. There are two types of piezoelectric crystal accelerometer, one much like an engine knock sensor and the other using spring elements. A severe change in speed of the vehicle will cause an output from these sensors as the seismic mass moves or the springs bend. This sensor has been used by supplementary restraint systems (SRS). Warning: For safety reasons, it is not recommended to test a sensor associated with an airbag circuit without specialist knowledge and equipment.

Many sensors are now integrated into ECUs (Figure 4.36). This means that they are almost impossible to test – but fortunately have become very reliable!

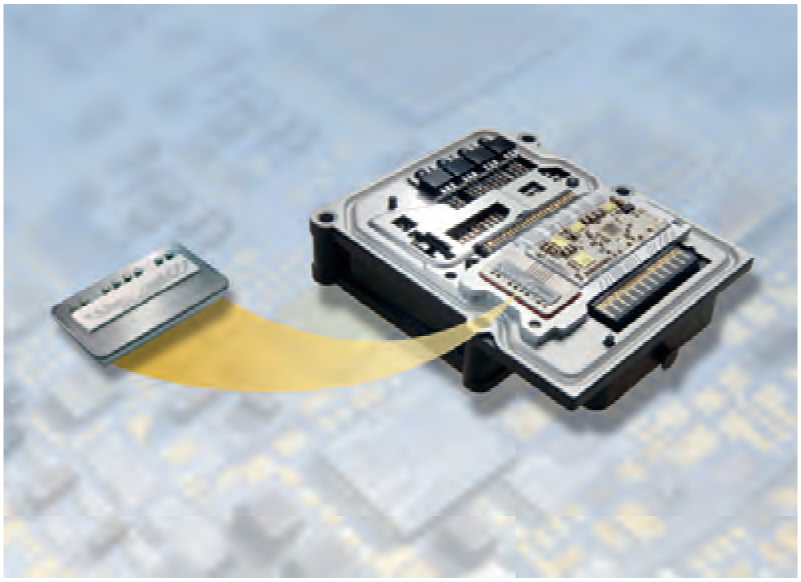


Figure 4.36 Yaw rate and acceleration sensor integrated into an electric stability program (ESP) control unit(Source: Bosch Media)

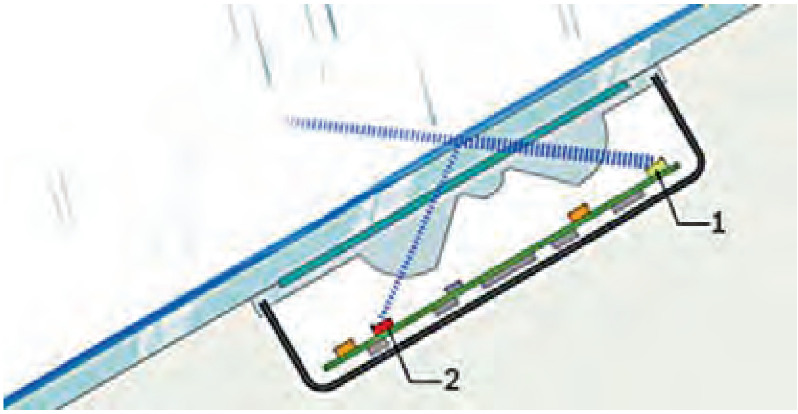


Figure 4.37 Rain sensor principle: 1 – LED; 2 – photo diode

4.2.13 Rain sensor

Rain sensors are used to switch on wipers automatically. Most work on the principle of reflected light. The device is fitted inside the windscreen and light from an LED is reflected back from the outer surface of the glass. The amount of light reflected changes if the screen is wet, even with a few drops of rain (Figures 4.37 and 4.38).

4.3 Actuators

4.3.1 Introduction

There are many ways of providing control over variables in and around the vehicle. ‘Actuators’ is a general term used here to describe a control mechanism. When controlled electrically, they will work either by a thermal or by a magnetic

Figure 4.38 Rain sensor package (Source: Bosch Media)

effect. In this section, the term actuator will generally be used to mean a device which converts electrical signals into mechanical movement (Table 4.3).

4.3.2 Testing actuators

Safety first

When powering a device, use a fused supply (jumper) wire.

Testing actuators can be simple as many are operated by windings. The resistance can be measured with an ohmmeter. A good tip is that where an actuator has more than one winding (e.g. a stepper motor), the resistance of each should be about the same. Even if the expected value is not known, it is likely that if all the windings read the same then the device is in working order.

With some actuators, it is possible to power them up from the vehicle battery. A fuel injector should click, for example, and a rotary air bypass device should rotate about half a turn. Be careful with this method as some actuators could be damaged. At the very least, use a fused supply (jumper) wire.

4.3.3 Motorised and solenoid actuators

4.3.3.1 Motors

Permanent magnet electric motors are used in many applications and are very versatile. The output of a motor is of course rotation, and this can be used in many ways. If the motor drives a rotating 'nut' through which a plunger is fitted on which there is a screw thread, the rotary action can easily be converted to linear movement. In most vehicle applications, the output of the motor has to be geared down, this is to reduce speed and increase torque. Permanent magnet motors are almost universally used now in place of older and less practical motors with field windings. Some typical examples of the use of these motors are listed as follows:

- windscreen wipers;
- windscreen washers;
- headlight lift;
- electric windows;
- electric sunroof;
- electric aerial operation;
- seat adjustment;

Table 4.3 Actuator diagnostic methods

Actuator	Equipment	Method(s)	Results	Scope waveform
Solenoid Fuel injector Lock actuator	Ohmmeter	Disconnect the component and measure its resistance	The resistance of many injectors is approx. 16 Ω (but check data) Lock and other actuators may have two windings (e.g. lock and unlock). The resistance values are very likely to be the same	Figures 4.47 and 4.48
Motor See previous list	Battery supply (fused) Ammeter	Most 'motor' type actuators can be run from a battery supply after they are disconnected from the circuit. If necessary the current draw can be measured	Normal operation with current draw appropriate to the 'work' done by the device. For example, a fuel pump motor may draw up to 10 A, but an idle actuator will only draw 1 or 2 A	N/A
Solenoid actuator (idle speed control)	Duty cycle meter	Most types are supplied with a variable ratio square wave	The duty cycle will vary as a change is required	Figure 4.41
Stepper motor Idle speed air bypass Carburettor choke control Speedometer drivers	Ohmmeter	Test the resistance of each winding with the motor disconnected from the circuit	Winding resistances should be the same. Values in the region of 10–20 Ω are typical	Figure 4.44
Thermal Auxiliary air device Instrument display	Ohmmeter Fused battery supply	Check the winding for continuity; if OK, power up the device and note its operation (for instruments, power these but use a resistor in place of the sender unit)	Continuity and slow movement (several seconds to a few minutes) to close the valve or move as required	N/A
EGR valve	Ohmmeter Fused battery supply	Check the winding(s) for continuity; if OK, power up the device and note its operation	Continuity and rapid movement to close the valve	Figure 4.56

- mirror adjustment;
- headlight washers;
- headlight wipers;
- fuel pumps;
- ventilation fans.

One disadvantage of simple motor actuators is that no direct feedback of position is possible. This is not required in many applications; however, in cases such as seat adjustment when a 'memory' of the position may be needed, a variable resistor–type sensor can be fitted to provide feedback. Three typical motor actuators are shown in Figure 4.39. The two motors on the right are used for window lift. Some of these use Hall effect sensors or an extra brush as a feedback device.

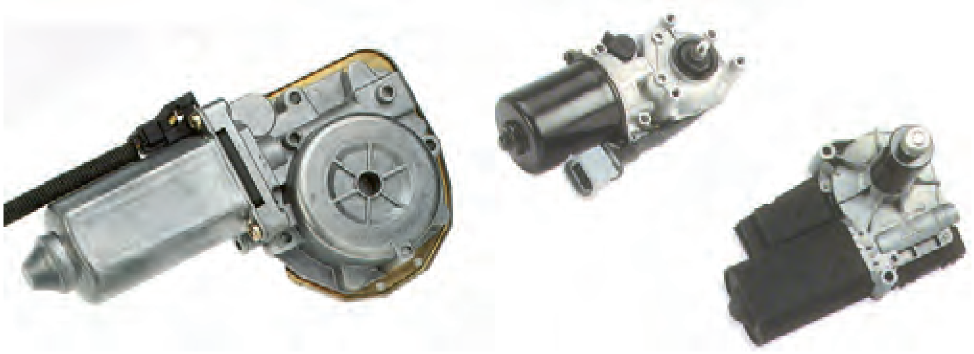


Figure 4.39 Window lift and wiper motors

Figure 4.40 Rotary idle control valve

4.3.3.2 Rotary idle speed control valve

The rotary ISCV will have two or three electrical connections, with a voltage supply at battery voltage and either a single- or a double-switched earth path. The device is like a motor but only rotates about half a turn in each direction.

This device is used to control idle speed by controlling air bypass. There are two basic types in common use. These are single-winding types, which have two terminals, and double-winding types, which have three terminals. Under ECU, the motor is caused to open and close a shutter, controlling air bypass. These actuators only rotate approximately 90° to open and close the valve. As these are permanent magnet motors, the 'single or double windings' refer to the armature.

The single-winding type is fed with a square wave signal causing it to open against a spring and then close again, under spring tension. The on/off ratio or duty cycle of the square wave will determine the average valve open time and hence idle speed. With the double-winding type, the same square wave signal is sent to one winding but the inverse signal is sent to the other. As the windings are wound in opposition to each other, if the duty cycle is 50% then no movement will take place. Altering the ratio will now cause the shutter to move in one direction or the other (Figure 4.40).

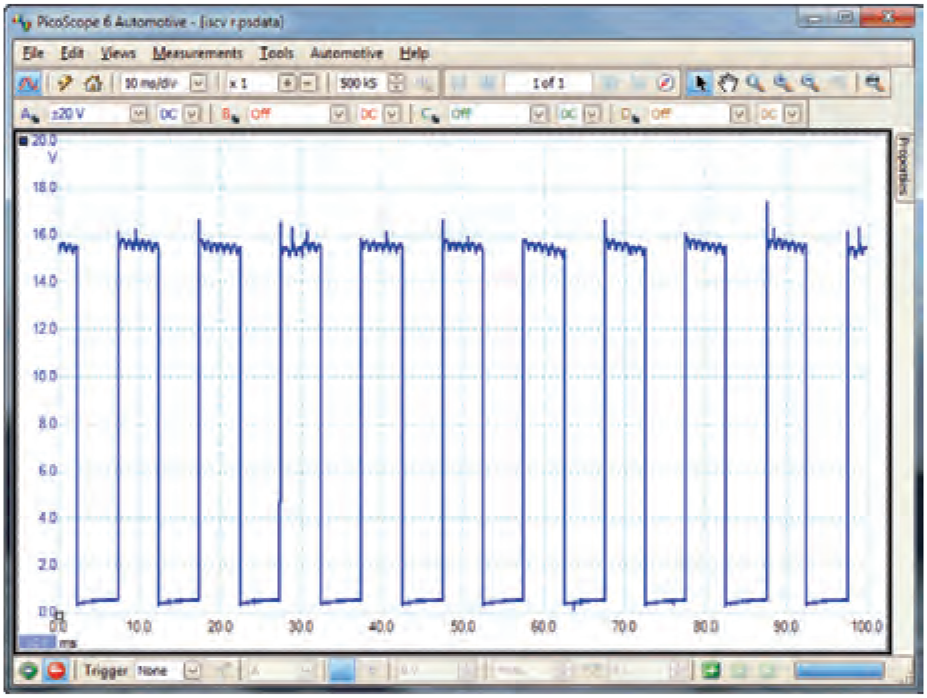


Figure 4.41 Signal supplied to a rotary idle control valve

The rate at which the earth path is switched is determined by the ECU to maintain a prerequisite idle speed according to its programming.

The valve will form an air bypass past the throttle butterfly to form a controlled air bleed within the induction tract. The rotary valve will have the choice of either single or twin earth paths, the single being pulled one way electrically and returned to its closed position via a spring; the double-switched earth system will switch the valve in both directions. This can be monitored on a dual trace oscilloscope. As the example waveform shows, the earth path is switched and the resultant picture is produced. The idle control device takes up a position determined by the on/off ratio (duty cycle) of the supplied signal.

Probing onto the supply side will produce a straight line at system voltage, and when the earth circuit is monitored, a square wave will be seen (Figure 4.41). The frequency can also be measured as can the on/off ratio.

4.3.3.3 Stepper motors

Stepper motors are becoming increasingly popular as actuators in the motor vehicle. This is mainly because of the ease with which they can be controlled by electronic systems. Stepper motors fall into the following three distinct groups, the basic principles of which are shown in Figure 4.42:

- variable reluctance motors;
- permanent magnet (PM) motors;
- hybrid motors.

The underlying principle is the same for each type. All of them have been and are being used in various vehicle applications. The basic design for a permanent magnet stepper motor comprises two double stators. The rotor is often made of barium-ferrite in the form of a sintered annular magnet. As the windings are energised in one direction then the other, the motor will rotate in 90° steps. Half

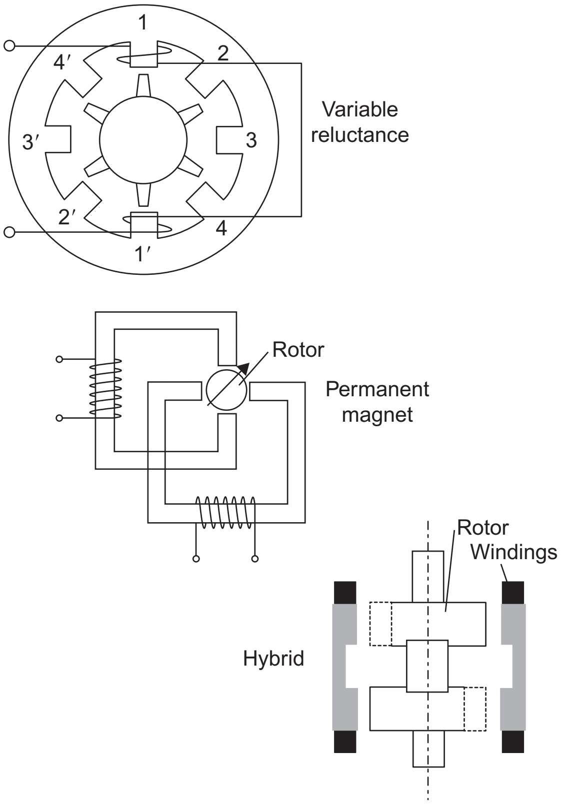


Figure 4.42 Stepper motor principle

step can be achieved by switching on two windings. This will cause the rotor to line up with the two stator poles and implement a half step of 45°. The direction of rotation is determined by the order in which the windings are switched on or off or reversed. The main advantages of a stepper motor are that feedback of position is not required. This is because the motor can be indexed to a known starting point and then a calculated number of steps will move the motor to any suitable position.

The stepper motor, when used to control idle speed, is a small electro-mechanical device that allows either an air bypass circuit or a throttle opening to alter in position depending on the amounts that the stepper is indexed (moved in known steps) (Figure 4.43).

Stepper motors are used to control the idle speed when an ISCV is not employed. The stepper may have four or five connections back to the ECU. These enable the control unit to move the motor in a series of 'steps' as the circuits are earthed to ground. These devices may also be used to control the position of control flaps, for example, as part of a heating and ventilation system (Figure 4.44).

The individual earth paths can be checked using the oscilloscope. The waveforms should be similar on each path. Variations to the example shown here may be seen between different systems (Figure 4.45).

4.3.4 Solenoid actuators

The basic operation of solenoid actuators is very simple. The term 'solenoid' actually means 'many coils of wire wound onto a hollow tube'. This is often misused but has become so entrenched that terms like 'starter solenoid', when really it is a starter actuator or relay, are in common use. A good example of a solenoid actuator is a fuel injector.

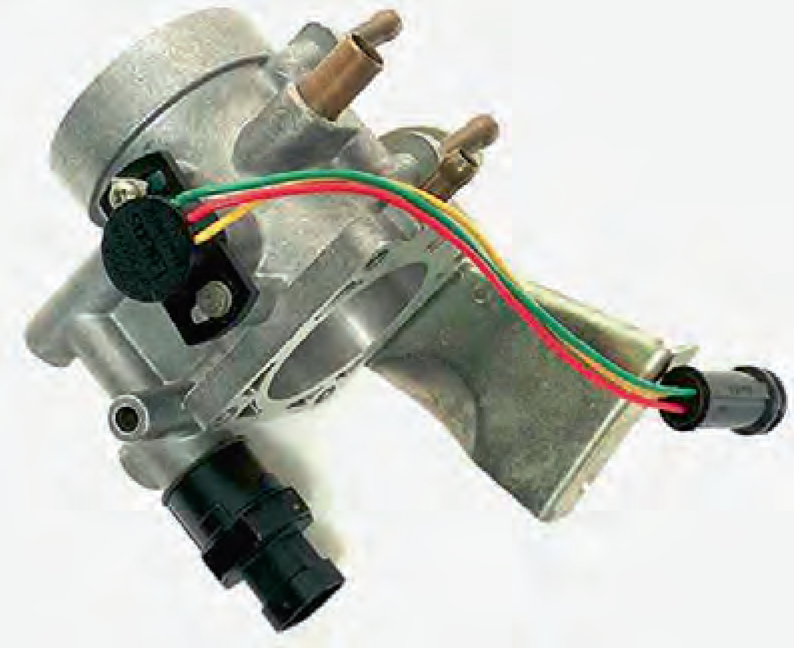
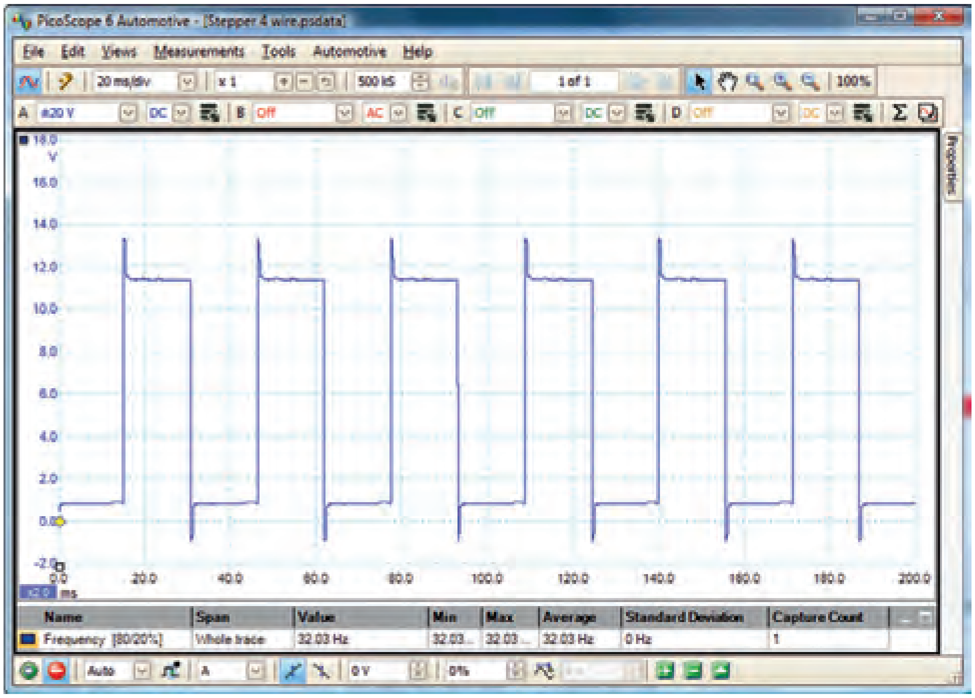


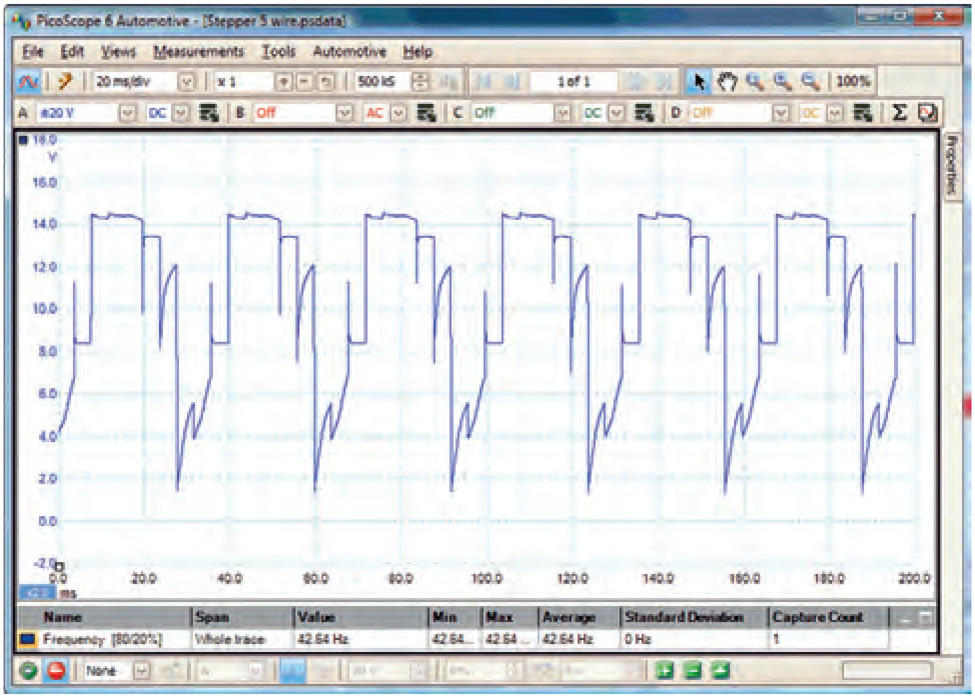
Figure 4.43 Stepper motor and throttle potentiometer on a throttle body



Name	Span	Value	Min	Max	Average	Standard Deviation	Capture Count
Frequency [80/20%]	Whole trace	32.03 Hz	32.03...	32.03...	32.03 Hz	0 Hz	1

Figure 4.44 Stepper motor signals

When the windings are energised, the armature is attracted due to magnetism and compresses the spring. In the case of a fuel injector, the movement is restricted to approximately 0.1 mm. The period that an injector remains open is very small; under various operating conditions, between 1.5 and 10 ms being typical. The time it takes an injector to open and close is also critical for accurate fuel metering. Some systems use ballast resistors in series with the fuel injectors. This allows lower inductance and resistance operating windings to be used, thus speeding up reaction time.



Name	Span	Value	Min	Max	Average	Standard Deviation	Capture Count
Frequency [80/20%]	Whole trace	42.64 Hz	42.64...	42.64 ...	42.64 Hz	0 Hz	1

Figure 4.45 Alternative stepper motor signal

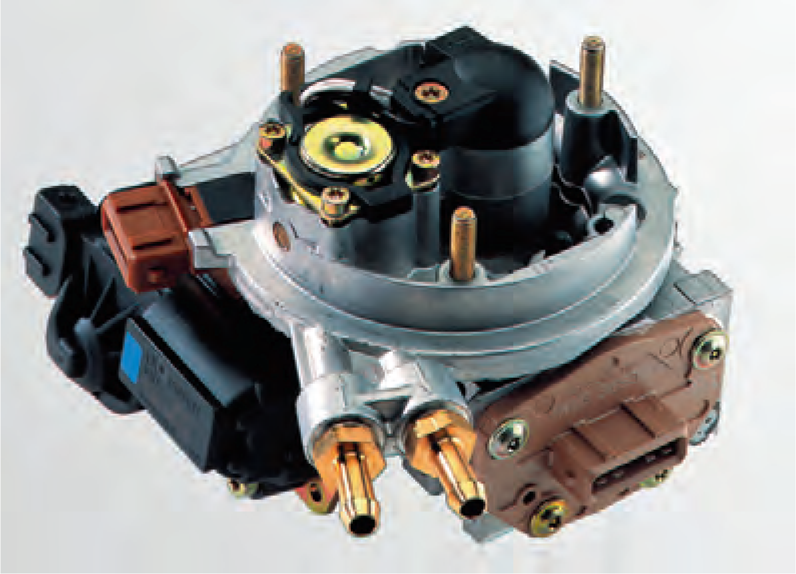


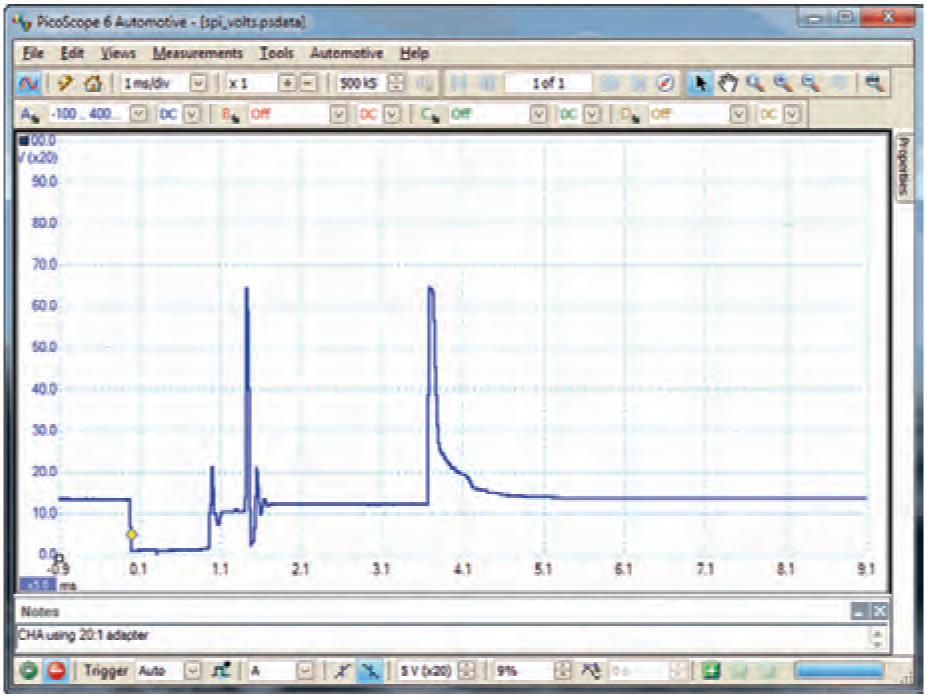
Figure 4.46 Throttle body with a single injector

Other types of solenoid actuators, for example door lock actuators, have less critical reaction times. However, the basic principle remains the same.

4.3.4.1 Single-point injector

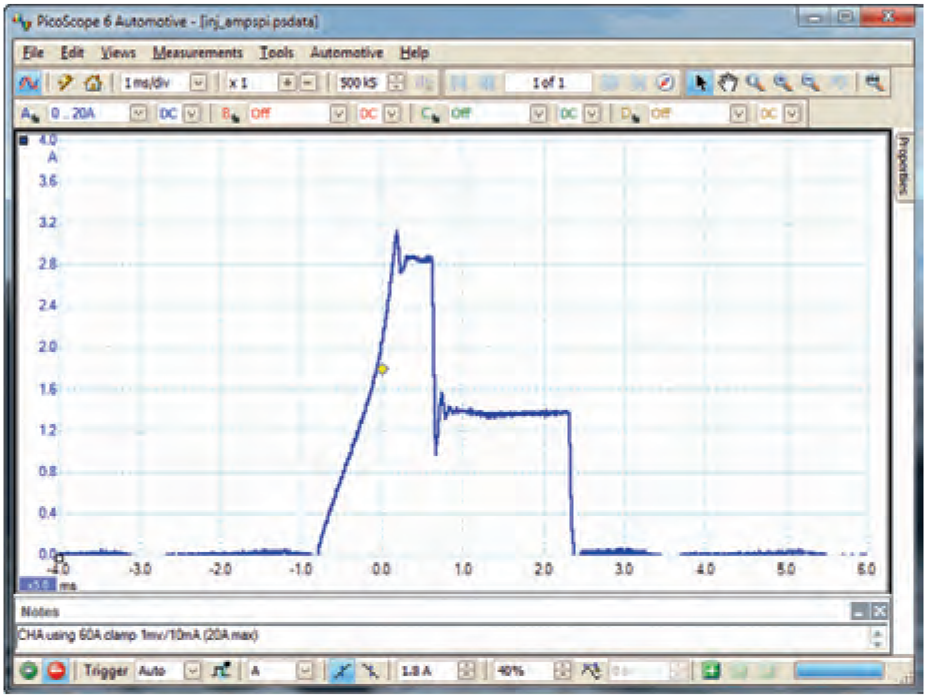
Single-point injection is also sometimes referred to as throttle body injection (Figure 4.46).

A single injector is used (on larger engines two injectors can be used) in what may have the outward appearance to be a carburettor housing.



CHA using 20:1 adapter

Figure 4.47 Single-point injector voltage waveform



CHA using 60A clamp 1mv/10mA (20A max)

Figure 4.48 Single-point injector current waveform

The resultant waveform from the single-point system shows an initial injection period followed by voltage-pulsing of the injector in the remainder of the trace. This ‘current limiting’ section of the waveform is called the supplementary duration and is the part of the injection trace that expands to increase fuel quantity. This shows better in a current rather than voltage waveform (Figures 4.47 and 4.48).

Figure 4.49 Multi-point injectors on the rail. Also shown are the pressure regulator and sensor

4.3.4.2 Multi-point injector

This injector is an electro-mechanical device which is fed by a 12 V supply. The voltage will only be present when the engine is cranking or running because it is controlled by a relay that operates only when a speed signal is available from the engine. Early systems had this feature built into the relay; most modern systems control the relay from the ECU (Figure 4.49).

The length of time the injector is held open will depend on the input signals seen by the ECU from its various engine sensors. The duration of open time or 'injector duration' will vary to compensate for cold engine starting and warm-up periods. The duration time will also expand under acceleration. The injector will have a constant voltage supply while the engine is running and the earth path will be switched via the ECU, the result can be seen in the example waveform (Figure 4.50). When the earth is removed, a voltage is induced into the injector and a spike approaching 60 V is recorded.

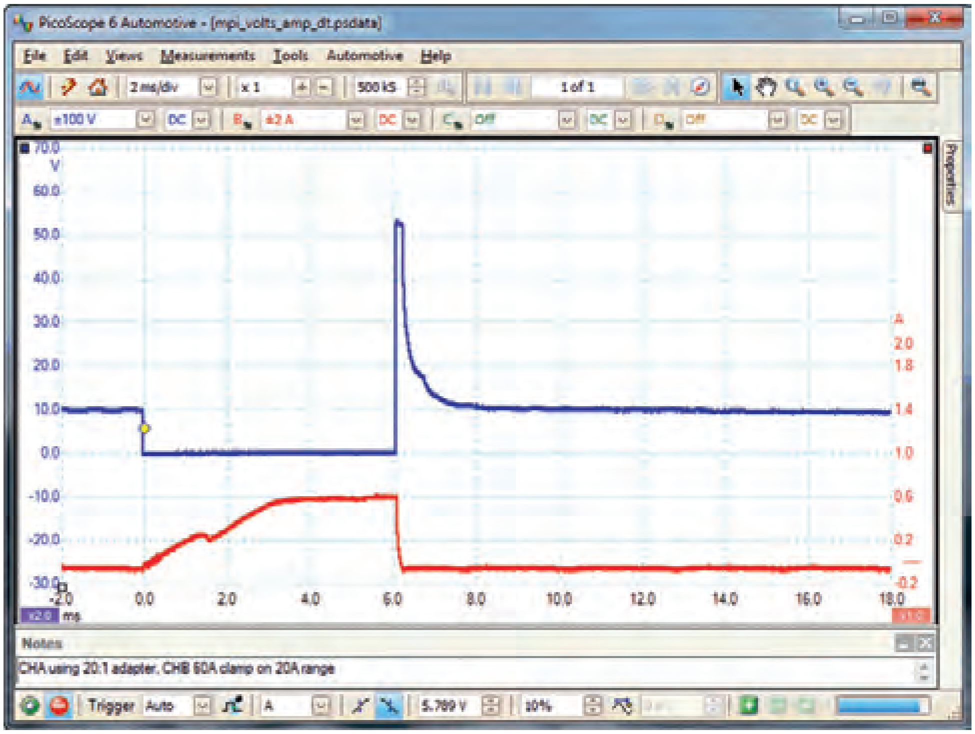
Key fact

The length of time an injector is held open depends on the sensor input signals to the ECU.

The height of the spike will vary from vehicle to vehicle. If the value is approximately 35 V, it is because a zener diode is used in the ECU to clamp the voltage. Make sure the top of the spike is squared off, indicating the zener dumped the remainder of the spike. If it is not squared, this indicates the spike is not strong enough to make the zener fully dump, meaning there is a problem with a weak injector winding. If a zener diode is not used in the computer, the spike from a good injector will be 60 V or more.

Multi-point injection may be either sequential or simultaneous. A simultaneous system will fire all four injectors at the same time with each cylinder receiving two injection pulses per cycle (720° crankshaft rotation). A sequential system will receive just one injection pulse per cycle, which is timed to coincide with the opening of the inlet valve.

Monitoring the injector waveform using both voltage and amperage allows display of the 'correct' time that the injector is physically open. The current waveform (the one starting on the zero line) shows that the waveform is 'split' into two defined areas.



CHA using 20:1 adapter, CHB 60A clamp on 20A range

Figure 4.50 Multi-point injector waveform, red shows the current and blue the voltage signal

The first part of the current waveform is responsible for the electromagnetic force lifting the pintle; in this example, the time taken is approximately 1.5 ms. This is often referred to as the solenoid reaction time. The remaining 2 ms is the actual time the injector is fully open. This, when taken as a comparison against the injector voltage duration, is different to the 3.5 ms shown. The secret is to make sure you compare like with like!

4.3.4.3 Common rail diesel injector

Common rail diesel systems are becoming more erm, *common*, particularly in Europe (Figure 4.51).

It can be clearly seen from the example waveform that there are two distinctive points of injection, the first being the ‘pre-injection’ phase, with the second pulse being the ‘main’ injection phase (Figure 4.52).

As the throttle is opened, and the engine is accelerated, the ‘main’ injection pulse expands in a similar way to a petrol injector. As the throttle is released, the ‘main’ injection pulse disappears until such time as the engine returns to just above idle.

Under certain engine conditions, a third phase may be seen, this is called the ‘post-injection’ phase and is predominantly concerned with controlling the exhaust emissions.

4.3.4.4 Idle speed control valve

This device contains a winding, plunger and spring. When energised, the port opens, and when not, it closes (Figure 4.53).

The electromagnetic ISCV will have two electrical connections: usually a voltage supply at battery voltage and a switched earth.

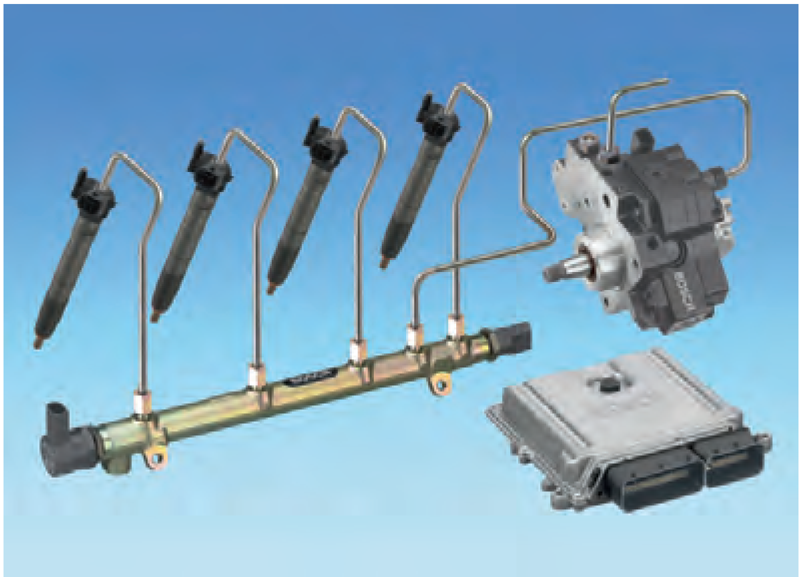
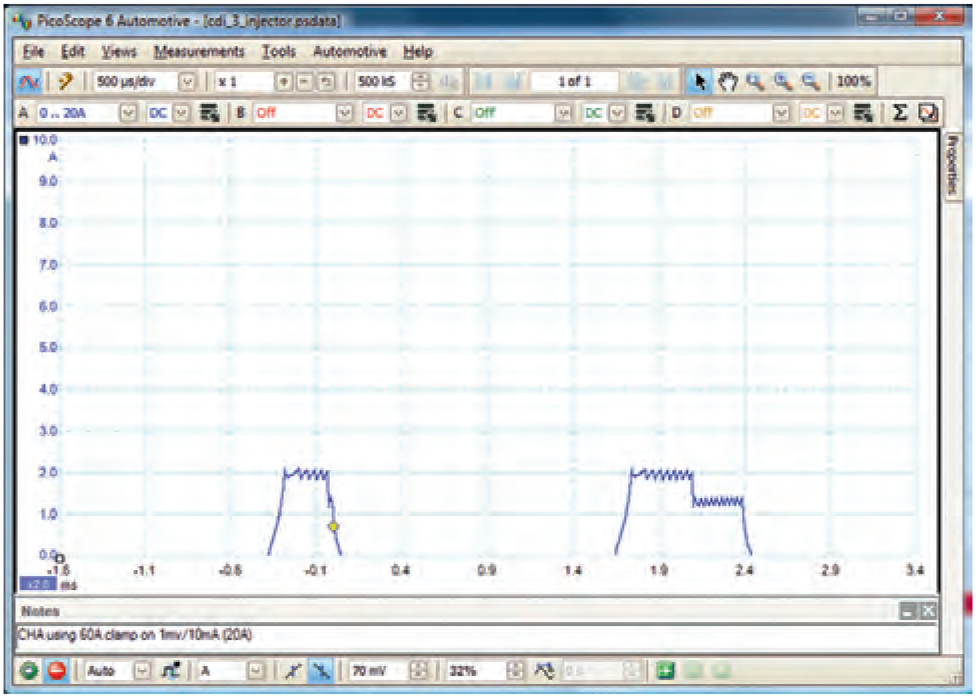


Figure 4.51 Common rail diesel pump, rail, injectors and ECU (Source: Bosch Press)



CHA using 60A clamp on 1mv/10mA (20A)

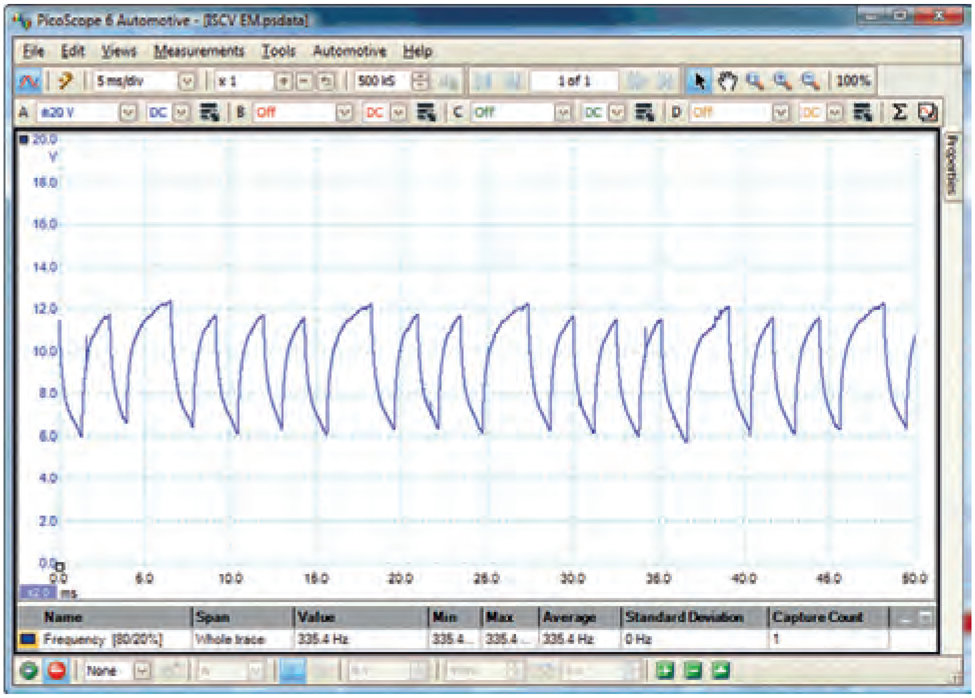
Figure 4.52 CR injector (current) waveform showing pre- and main injection pulses

The rate at which the device is switched is determined by the ECU to maintain a prerequisite speed according to its programming. The valve will form an air bypass around the throttle butterfly. If the engine has an adjustable air bypass and an ISCV, it may require a specific routine to balance the two air paths. The position of the valve tends to take up an average position determined by the supplied signal. Probing onto the supply side will produce a straight line at system voltage (Figure 4.54).

Definition

ISCV: Idle speed control valve.

Figure 4.53 Electromagnetic idle speed control valve



Name	Span	Value	Min	Max	Average	Standard Deviation	Capture Count
Frequency [80/20%]	Whole trace	335.4 Hz	335.4...	335.4...	335.4 Hz	0 Hz	1

Figure 4.54 Signal produced by an electromagnetic idle speed control valve

4.3.4.5 Exhaust gas recirculation valve

Various types of exhaust gas recirculation (EGR) valve are in use based on simple solenoid operation. One development in actuator technology is the rotary electric exhaust gas recirculation (EEGR) valve for use in diesel engine applications. This device is shown in Figure 4.55. It has a self-cleaning action, accurate gas flow control and a fast reaction speed.

Figure 4.55 Rotary EGR valve (Source: Delphi Media)

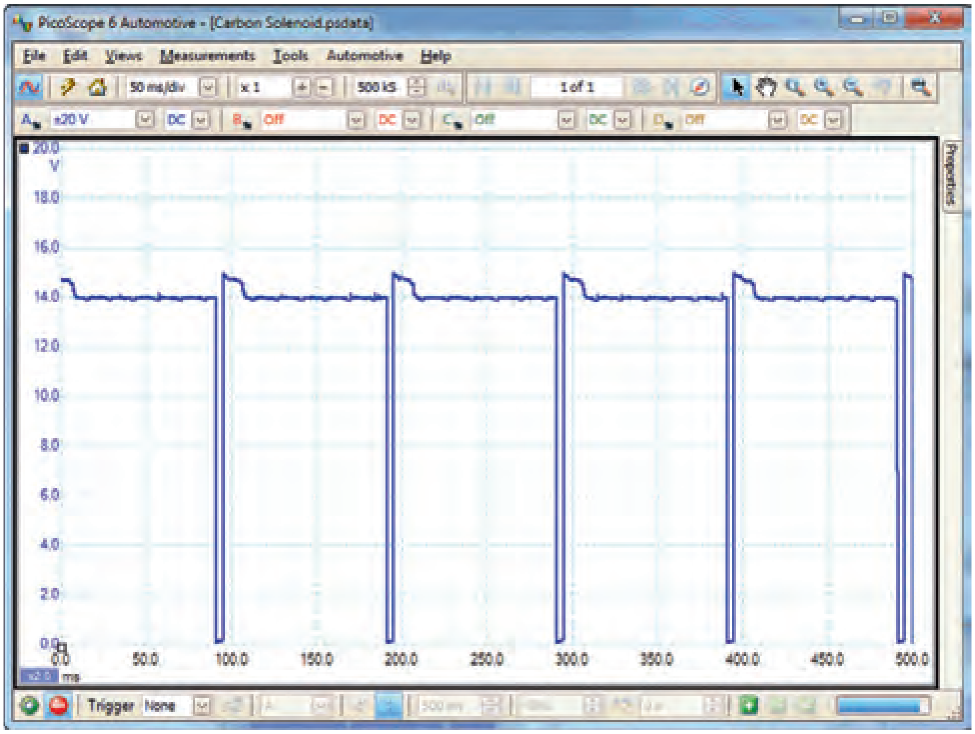


Figure 4.56 Carbon canister control valve signal

4.3.4.6 Carbon canister and other valves

There are a number of valves used that are effectively simple solenoid controlled devices. Measuring on one terminal will usually show battery supply voltage. The other terminal will show battery voltage when switched off and zero (ground or earth) voltage when the valve is switched on (Figure 4.56).

4.3.5 Thermal actuators

An example of a thermal actuator is the movement of a traditional type fuel or temperature gauge needle. A further example is an auxiliary air device used on many earlier fuel injection systems. The principle of the gauge is shown in Figure 4.57.

When current is supplied to the terminals, a heating element operates and causes a bimetallic strip to bend, which moves the pointer. The main advantage of this type of actuator, when used as an auxiliary device, apart from its simplicity, is that if it is placed in a suitable position, its reaction time will vary with the temperature of its surroundings. This is ideal for applications such as fast idle or cold starting control where, once the engine is hot, no action is required from the actuator.

Safety first

Even the earth path of the coil can produce over 350V – take care.

4.4 Engine waveforms

4.4.1 Ignition primary

The ignition primary waveform is a measurement of the voltage on the negative side of the ignition coil. The earth path of the coil can produce over

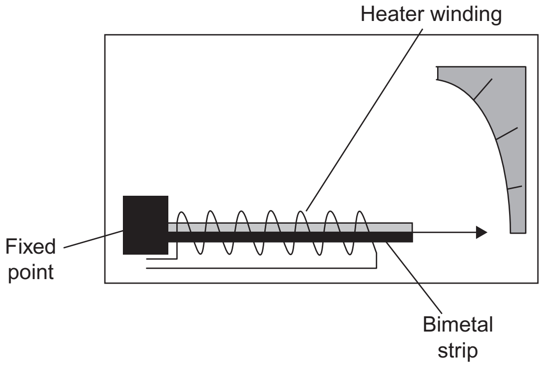


Figure 4.57 Thermal actuator used as a gauge

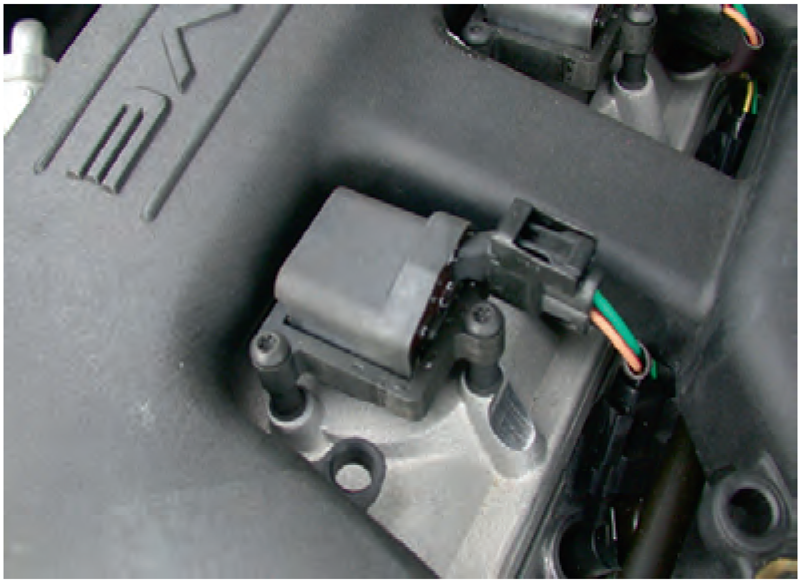


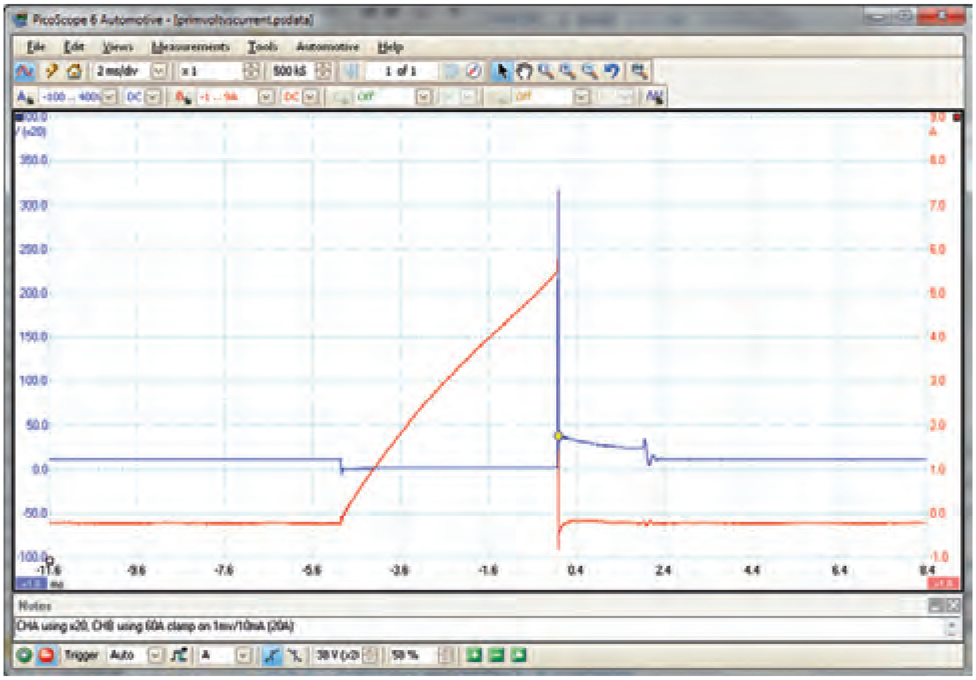
Figure 4.58 Coil on plug (COP) ignition

350 V. Different types of ignition coils produce slightly different traces but the fundamental parts of the trace and principles are the same (Figure 4.58).

In the waveform shown, the horizontal voltage line at the centre of the oscilloscope is at fairly constant voltage of approximately 30–40 V, which then drops sharply to what is referred to as the coil oscillation (Figure 4.59). The length of the horizontal voltage line is the ‘spark duration’ or ‘burn time’, which in this particular case is approximately 1 ms. The coil oscillation period should display a minimum of three to four peaks (both upper and lower). A loss of peaks would indicate a coil problem.

There is no current in the coil’s primary circuit until the dwell period. This starts when the coil is earthed and the voltage drops to zero. The dwell period is controlled by the ignition amplifier or ECU and the length of the dwell is determined by the time it takes to build up to approximately 6 A. When this predetermined current has been reached, the amplifier stops increasing the primary current and it is maintained until the earth is removed from the coil. This is the precise moment of ignition.

The vertical line at the centre of the trace is in excess of 300 V, this is called the ‘induced voltage’. The induced voltage is produced by magnetic inductance.



CHA using x20, CHB using 60A clamp on 1mv/10mA (20A)

Figure 4.59 Primary ignition voltage and current traces

At the point of ignition, the coil's earth circuit is removed and the magnetic flux collapses across the coil's windings. This induces a voltage between 150 and 350 V. The coil's high-tension (HT) output will be proportional to this induced voltage. The height of the induced voltage is sometimes referred to as the primary peak volts.

From the example current waveform, the limiting circuit can be seen in operation. The current switches on as the dwell period starts and rises until the required value is achieved (usually 6–8 A). At this point, the current is maintained until it is released at the point of ignition.

The dwell will expand as the engine revs are increased to maintain a constant coil saturation time. This gives rise to the term 'constant energy'. The coil saturation time can be measured and will remain the same regardless of the engine speed. The example shows a charge time of approximately 3.5 ms.

Safety first

Some coils can produce over 50 000 V – take care.

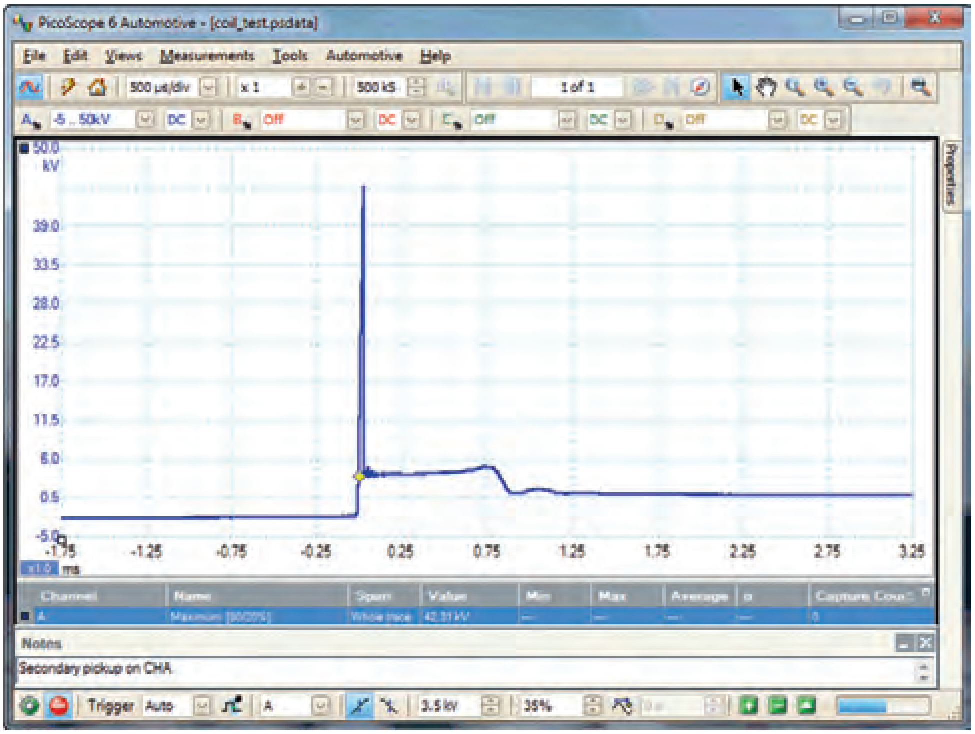
4.4.2 Ignition secondary

The ignition secondary waveform is a measurement of the HT output voltage from the ignition coil. Some coils can produce over 50 000 V. Different types of ignition coils produce slightly different traces but the fundamental parts of the trace and principles are the same (Figure 4.60).

The ignition secondary picture shown in the example waveform is from an engine fitted with electronic ignition. In this case, the waveform has been taken from the main coil lead (king lead). Suitable connection methods mean that similar traces can be seen for other types of ignition system (Figure 4.61).

The secondary waveform shows the length of time that the HT is flowing across the spark plug electrode after its initial voltage, which is required to initially jump the plug gap. This time is referred to as either the 'burn time' or the 'spark duration'. In the trace shown, it can be seen that the horizontal voltage line in the

Figure 4.60 Spark plugs(Source: Bosch Press)



Channel	Name	Span	Value	Min	Max	Average	σ	Capture Count
A	Maximum [80C0%]	Whole trace	42.31 kV	—	—	—	—	0

Notes

Secondary pickup on CHA

Figure 4.61 Ignition secondary trace

centre of the oscilloscope is at fairly constant voltage of approximately 4 or 5 kV, which then drops sharply into the 'coil oscillation' period.

The coil oscillation period should display a minimum of three or four peaks (same as for the primary trace). A loss of peaks indicates that the coil may be faulty. The period between the coil oscillation and the next 'drop down' is when the coil is at rest and there is no voltage in the secondary circuit. The 'drop down' is referred to as the 'polarity peak', and produces a small oscillation in the opposite

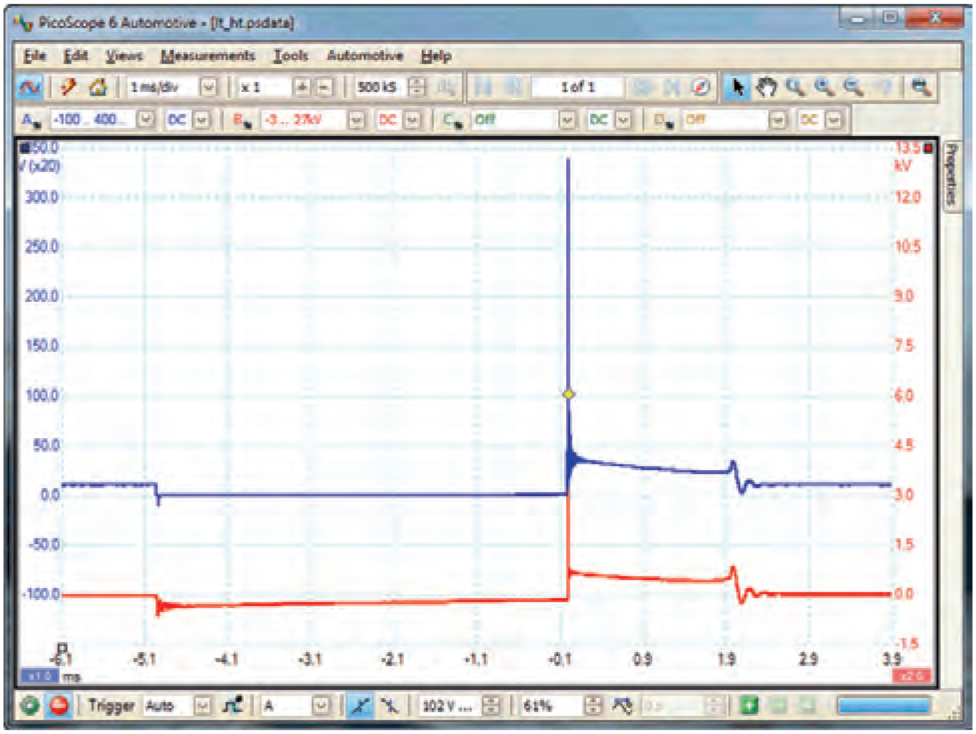


Figure 4.62 Distributorless ignition showing low and high tension (primary and secondary)

direction to the plug firing voltage. This is due to the initial switching on of the coil’s primary current.

The plug firing voltage is the voltage required to jump and bridge the gap at the plug’s electrode, commonly known as the ‘plug kV’. In the example in Figure 4.61, the plug firing voltage is approximately 45 kV.

When the plug kVs are recorded on a distributorless ignition system (DIS) or coil per cylinder ignition system, the voltage seen on the waveform should be in the ‘upright position’. If the trace is inverted, it would suggest that either the wrong polarity has been selected from the menu or in the case of DIS, the inappropriate lead has been chosen. The plug voltage, while the engine is running, is continuously fluctuating and the display will be seen to move up and down. The maximum voltage at the spark plug can be seen as the ‘Ch A: Maximum (kV)’ reading at the bottom of the screen.

It is a useful test to snap the throttle and observe the voltage requirements when the engine is under load. This is the only time that the plugs are placed under any strain and is a fair assessment of how they will perform on the road.

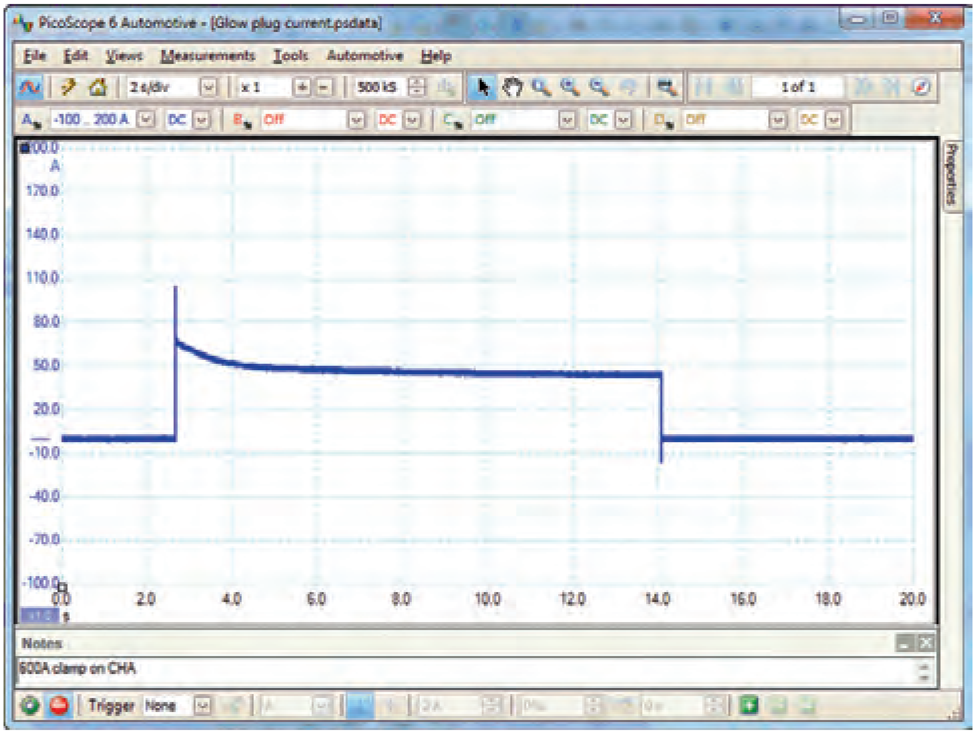
The second part of the waveform after the vertical line is known as the spark line voltage. This second voltage is the voltage required to keep the plug running after its initial spark to jump the gap. This voltage will be proportional to the resistance within the secondary circuit. The length of the line can be seen to run for approximately 2 ms.

Key fact

Fluctuations in voltage on the spark line could indicate poor combustion.

4.4.3 Diesel glow plugs

A diesel glow plug is a simple heater. Measuring its current will indicate correct operation because as temperature increases in a glow plug so does resistance and therefore the current falls after an initial peak (Figure 4.63).



Notes

600A clamp on CHA

Trigger None

Figure 4.63 Glow plug current

Figure 4.64 Alternator (Source: Bosch Media)

4.4.4 Alternator waveform

Checking the ripple voltage produced by an alternator (Figure 4.64) is a very good way of assessing its condition.

The example waveform illustrates the rectified output from the alternator (Figure 4.65). The output shown is correct and there is no fault within the phase windings or the diodes (rectifier pack).

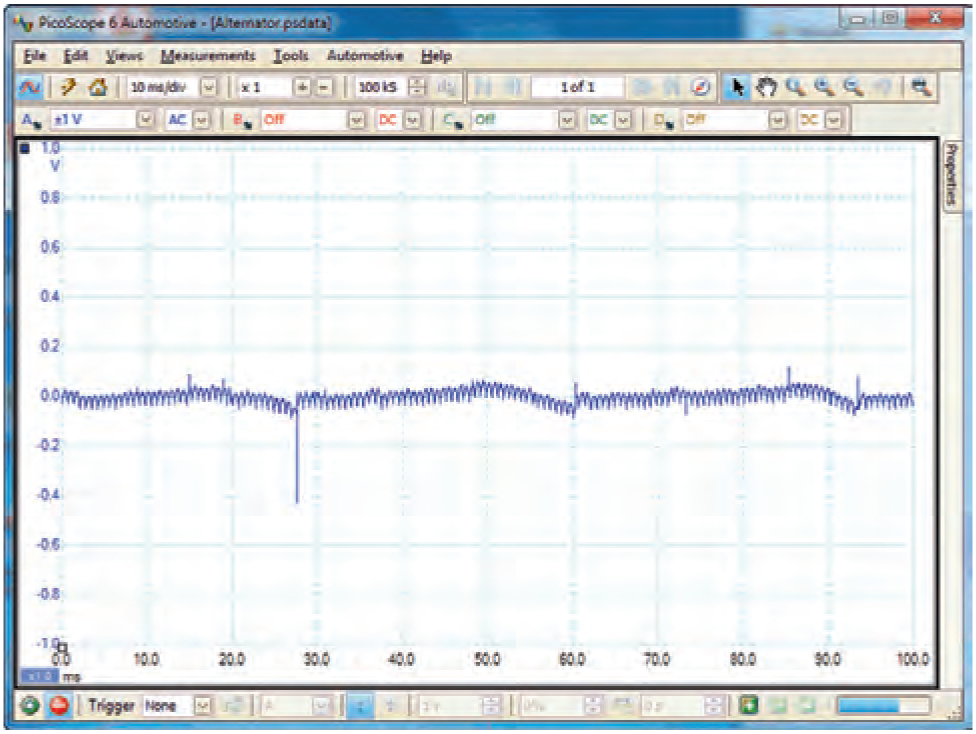


Figure 4.65 Alternator ripple voltage

The three phases from the alternator have been rectified to DC from its original AC and the waveform shows that the three phases are all functioning.

If the alternator is suffering from a diode fault, long downward ‘tails’ appear from the trace at regular intervals and 33% of the total current output will be lost. A fault within one of the three phases will show a similar picture to the one illustrated but is three or four times the height, with the base-to-peak voltage in excess of 1 V.

The voltage scale at the side of the oscilloscope is not representative of the charging voltage, but is used to show the upper and lower limits of the ripple. The ‘amplitude’ (voltage/height) of the waveform will vary under different conditions. A fully charged battery will show a ‘flatter’ picture, while a discharged battery will show an exaggerated amplitude until the battery is charged. Variations in the average voltage of the waveform are due to the action of the voltage regulator.

4.4.5 Relative compression petrol

Measuring the current drawn by the starter motor is useful to determine starter condition but it is also useful as an indicator of engine condition (Figure 4.66).

The purpose of this particular waveform is therefore to measure the current required to crank the engine and to evaluate the relative compressions.

The amperage required to crank the engine depends on many factors, such as the capacity of the engine, number of cylinders, viscosity of the oil, condition of the starter motor, condition of the starter’s wiring circuit and compressions in the cylinders. Therefore, to evaluate the compressions, it is essential that the battery is charged and the starter and associated circuit are in good condition.

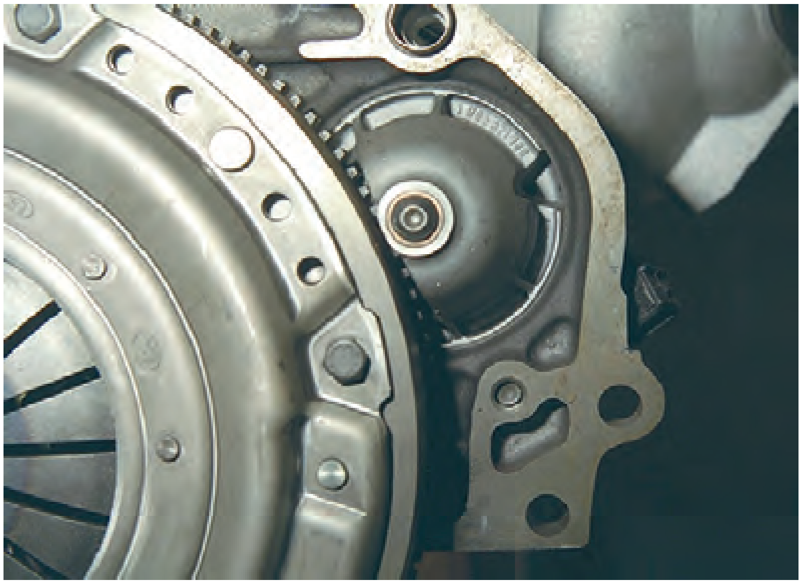


Figure 4.66 Starter and ring gear

The current for a typical four cylinder petrol/gasoline engine is in the region of 100–200 A. In the waveform shown (Figure 4.67), the initial peak of current (approximately 300 A) is the current required to overcome the initial friction and inertia to rotate the engine. Once the engine is rotating, the current will drop. It is also worth mentioning the small step before the initial peak, which is being caused by the switching of the starter solenoid.

The compressions can be compared against each other by monitoring the current required to push each cylinder up on its compression stroke. The better the compression, the higher the current demand and vice versa. It is therefore important that the current draw on each cylinder is equal.

4.5 Communication networks

There are three common multiplexed communication systems in current use. These systems reduce the number of wires needed and also allow information from sensors or different ECUs to be shared across a network. The three main systems are

- CAN;
- LIN;
- FlexRay.

4.5.1 CAN

Controller area network (CAN) is a protocol used to send information around a vehicle on data bus. It is made up of voltage pulses that represent ones and zeros, in other words, binary signals. The data is applied to two wires known as CAN high and CAN low (Figure 4.68).

In this display, it is possible to verify that data is being continuously exchanged along the CAN bus. It is also possible to check that the peak-to-peak voltage

Definition

DLC: Diagnostic/Data link connector

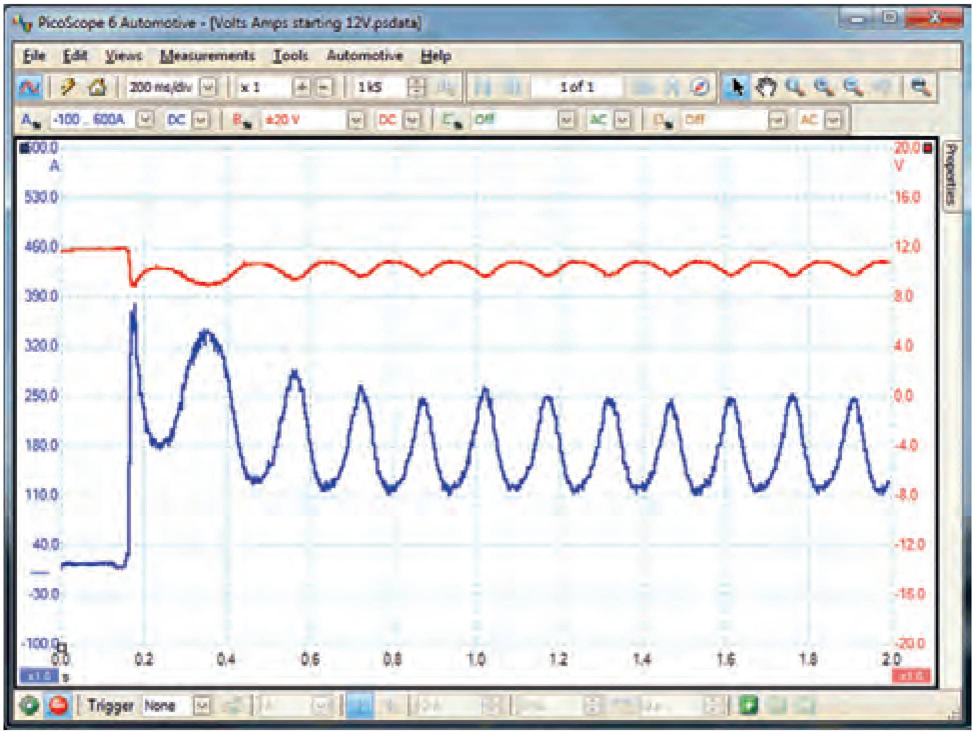
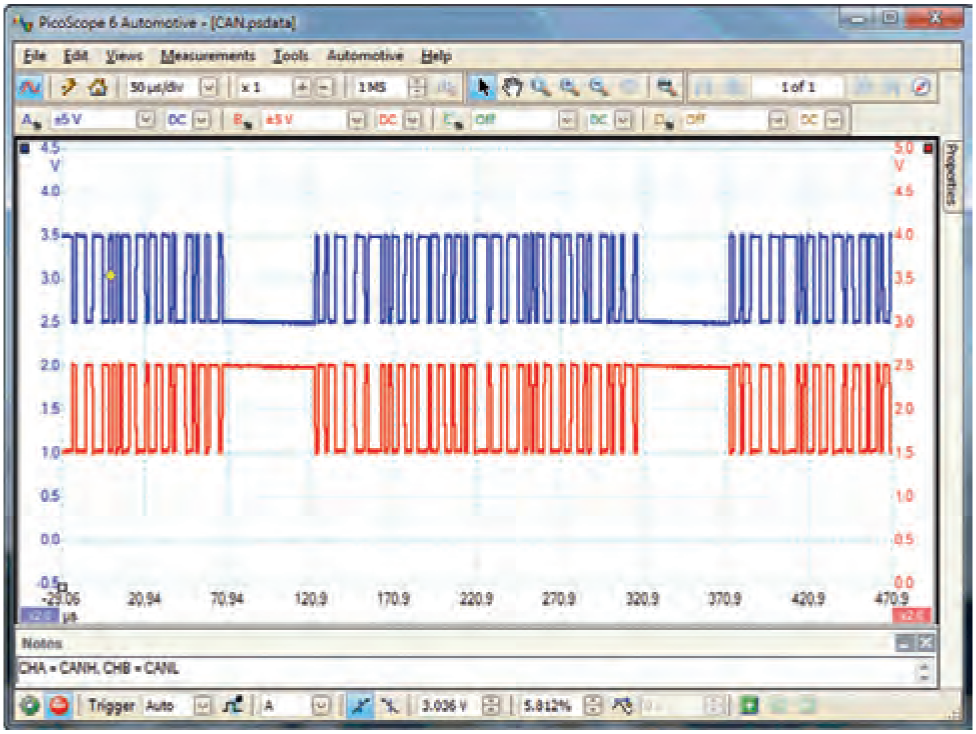


Figure 4.67 Spark ignition engine cranking amps

Figure 4.68 DLC socket – pin 6 is CAN high and pin 14 is CAN low

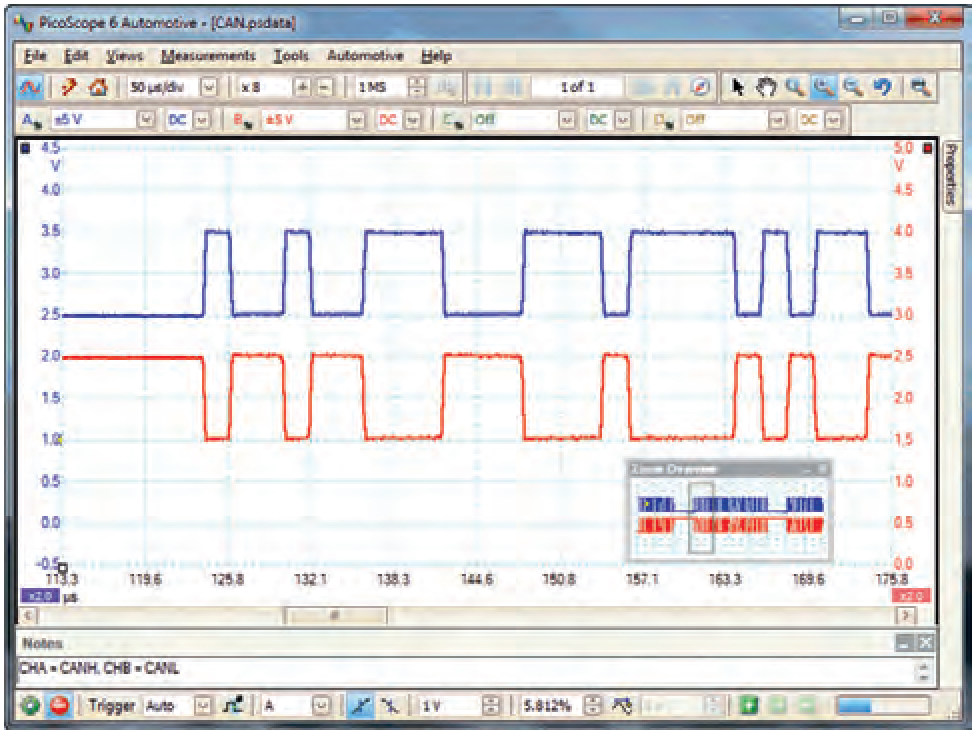
levels are correct and that a signal is present on both CAN lines. CAN uses a differential signal, and the signal on one line should be a coincident mirror image (the signals should line up) of the data on the other line (Figure 4.69).

The usual reason for examining the CAN signals is where a CAN fault has been indicated by on-board diagnostics, or to check the CAN connection to a suspected faulty CAN node. The vehicle manufacturers’ manual should be referred to for precise waveform parameters.



CHA = CANH, CHB = CANL

Figure 4.69 CAN high and low signals on a dual trace scope



CHA = CANH, CHB = CANL

Figure 4.70 CAN signal zoomed in

When the signal is captured on a fast timebase (or zoomed in), it allows the individual state changes to be viewed. This enables the mirror image nature of the signals and the coincidence of the edges to be verified (Figure 4.70).

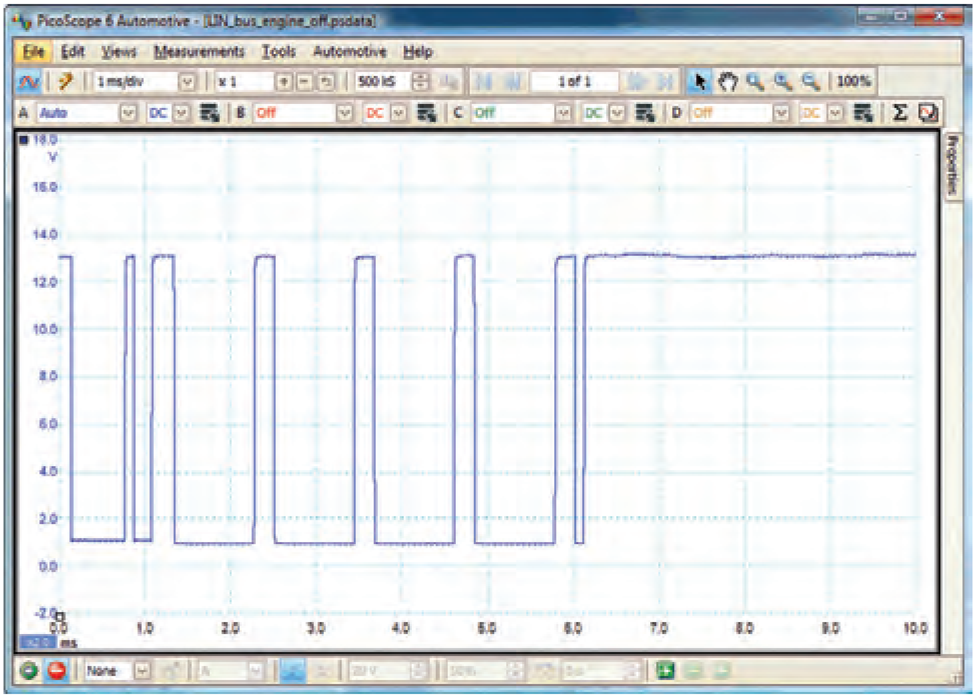


Figure 4.71 LIN waveform

4.5.2 LIN

Local interconnect network (LIN) bus communication is becoming more common on modern CAN bus–equipped vehicles. It is a low-speed, single-wire serial data bus and a sub-bus of the faster, more complex CAN bus. It is used to control low-speed non-safety-critical housekeeping functions on the vehicle, especially windows, mirrors, locks, HVAC units and electric seats.

Definition

LIN: Local interconnect network.

The LIN bus is proving popular because of its low cost and also because it reduces the bus load of the supervising CAN network.

LIN signals can be measured by connecting between earth/ground and the signal wire. It is not possible to decode the signal but a correctly switching square waveform should be shown (Figure 4.71).

4.5.3 FlexRay

FlexRay uses very high speed signals, so it is necessary to use high-speed probes (these are supplied with an advanced diagnostics kit) (Figure 4.72). The FlexRay-high and FlexRay-low pins are usually available at the multi-way connector at each ECU on the network.

It is possible to verify that data is being continuously exchanged on the FlexRay network, that the peak-to-peak voltage levels are correct and that a signal is present on both FlexRay lines. FlexRay uses a differential signal, so the signal on one line should be a mirror image of the data on the other line (Figure 4.73).

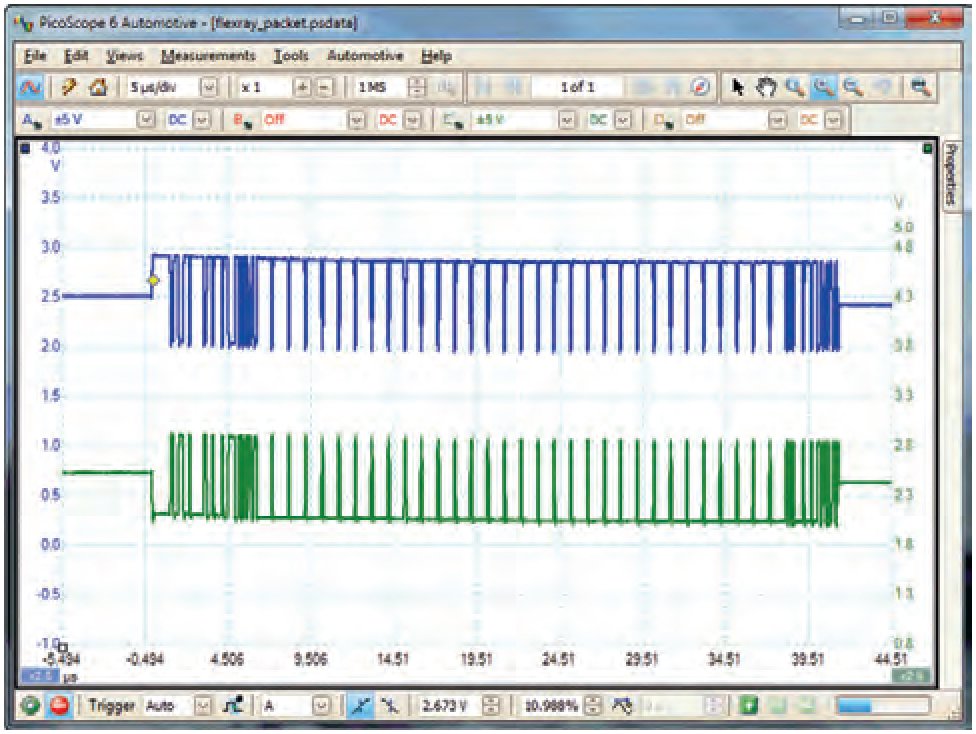


Figure 4.72 FlexRay signal

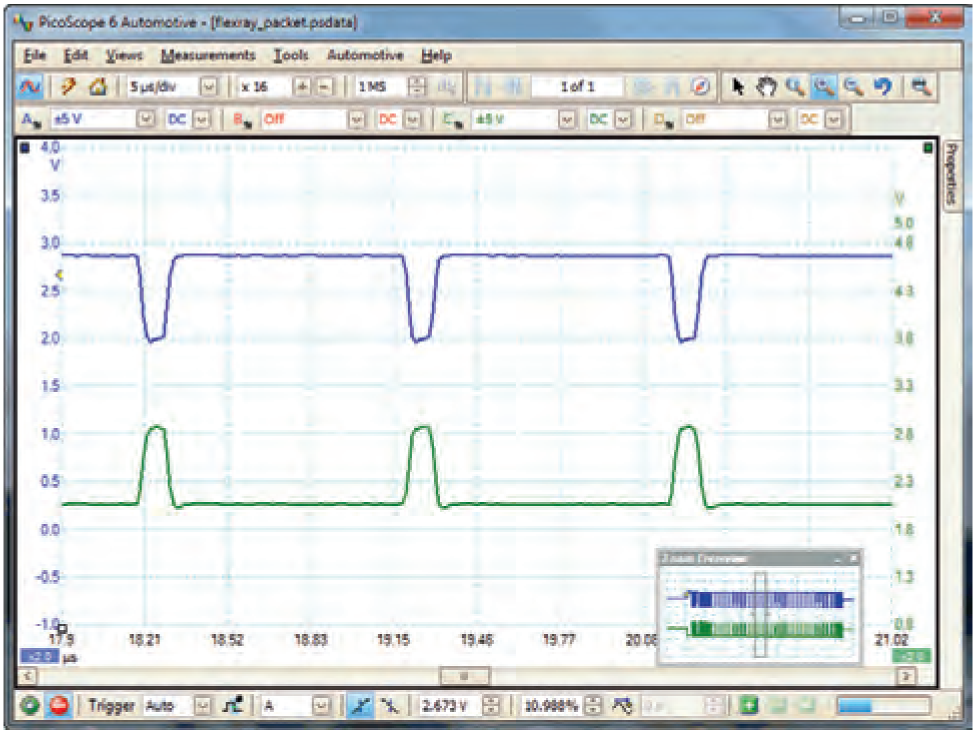


Figure 4.73 A closer view of a FlexRay signal

4.6 Summary

'Scope' diagnostics, particularly for sensors and actuators, is now an essential skill for the technician to develop. As with all diagnostic techniques that use test equipment, it is necessary for the user to know how

1. the vehicle system operates;
2. to connect the equipment;
3. readings should be interpreted.

Remember that an oscilloscope is actually just a voltmeter or ammeter, but it draws a picture of the readings over a set period of time. Learn what good waveforms look like, and then you will be able to make good judgements about what is wrong when they are not so good.

Acknowledgement

I am grateful to PicoTech for permission to use waveforms from their extensive library. Visit http://www.picotech.com for more information.

CHAPTER

On-board diagnostics

5.1 History

5.1.1 Introduction

Originating in the United States, and subsequently followed by Europe, Asia and many others, now governments around the globe have augmented vehicle emissions control legislation. This includes a requirement that all vehicles sold within their territories must support an on-board diagnostic (OBD) system that can be operated to determine the serviceability of the vehicle's emission control systems, sub-systems and components.

Enabled by the increasing advances in electronics and microprocessor software development, this system, now commonly termed as OBD, has been developed over recent years and is now implemented by all major motor vehicle manufacturers. Furthermore, this has been extended to allow diagnosis of non-emission-related vehicle systems.

5.1.2 Vehicle emissions and environmental health

From as early as 1930, the subject of vehicle engine emissions influencing environmental health was very topical in the state of California. Already with a population of 2 million vehicles, scores of people died and thousands became sick due to air pollution–related illnesses (Figure 5.1).

Figure 5.1 Early traffic jam

In 1943, following the outbreak of the Second World War, the population of California had risen to some 7 million people with 2.8 million vehicles travelling over a total of 24 billion miles. Already, smog was apparent and people suffered with stinging eyes, sore throat and breathing difficulties. The local government initiated a study into the cause of the problem. Scientists at CALTEC and The University of California investigated the problem of smog.

In 1945, after the conclusion of the war, Los Angeles began its air pollution control program and established the Bureau of Smoke Control. On 10 June 1947, the then California governor Earl Warren signed the Air Pollution Control Act. By 1950, California's population had reached 11 million people. Total registered vehicles in California exceeded 4.5 million and vehicle miles travelled (VMT) was 44.5 billion. The search for the root cause of smog production went on. Reports of deaths in other countries became apparent, for example, thousands of people died in London of a 'mystery fog' (Figure 5.2).

Definition

VMT: Vehicle miles travelled.

In 1952, Dr Arie Haagen-Smit determined the root cause of smog production. He surmised that engine pollutants, carbon monoxide (CO), hydrocarbons

Advanced Automotive Fault Diagnosis.

Figure 5.2 Smog over Los Angeles

(HC) and various oxides of nitrogen (NO_x) combine to generate the smog, which consists of ozone and carbon dioxide.

Carbon dioxide is a pollutant, which is now said to contribute to global warming and climate change. Ozone, occupying a region of the lower atmosphere, is now known to cause respiratory ill health and lung disease and is also thought to make a much greater contribution to the greenhouse effect than even carbon dioxide.

The state became a centre for environmental activism. Naturally, amidst a public outcry to preserve the local environment, the state began to legislate for controls on motor vehicle emissions. So began an initiative that would span over 50 years, one that would drive change in a world industry and lead the world in the fight for clean air.

5.1.3 History of the emissions control legislation

In 1960, the Motor Vehicle Pollution Control Board was established with a mandate to certify devices proposed to be fitted on cars for sale in California. In addition, the Federal Motor Vehicle Act of 1960 was enacted, requiring federal research to combat motor vehicle engine pollution. Manufacturers made technology improvements, and during this period, California's population reached 16 million. Total registered vehicles approached 8 million and VMT was 71 billion.

In 1961, in an effort to control HC crankcase emissions, the first piece of vehicle emissions control legislation mandating the use of specific hardware was issued. Positive crankcase ventilation (PCV) controls HC crankcase emissions by extracting gases from the crankcase and recirculating them back into the fresh air/fuel charge in the cylinders.

A key turning point in history, in 1966, the California Motor Vehicle Pollution Control Board pioneered the adoption of vehicle tailpipe emissions standards for HC and CO and the California Highway Patrol began random roadside inspections of the smog control devices fitted to vehicles.

The following year, the governor of California, Ronald Reagan, signed the Mulford-Carrell Air Resources Act. This effectively allowed the state of

Table 5.1 California state-wide average emissions per vehicle, 1969

NO_x (g/mile)	HC (g/mile)
5.3	8.6

Table 5.2 California state-wide average emissions per vehicle, 1980

NO_x (g/mile)	HC (g/mile)
4.8	5.5

Definition

CARB: California Air Resources Board.

California to set its own emissions standards. The same year saw the formation of the California Air Resources Board (CARB), which was created from the amalgamation of the Motor Vehicle Pollution Control Board and the Bureau of Air Sanitation.

In 1969, the first California State Ambient Air Quality Standards are extended by California for photochemical oxidants, suspended particulates, sulphur dioxide (SO_2), nitrogen dioxide (NO_2) and CO. California's population reached 20 million people. Total registered vehicles exceeded 12 million and VMT was 110 billion.

Total cumulative California vehicle emissions for HC and NO_x are estimated at 1.6 million tons/year (Table 5.1). In 1970, the US Environmental Protection Agency (EPA) was created. Its primary directive was, and still is, to protect all aspects of the environment. The next seven years witnessed further development of emissions control legislation and increasing employment of vehicle emissions control technology.

In 1971, CARB adopted the first vehicle NO_x standards. The EPA announced National Ambient Air Quality Standards for particulates, HC, CO, NO_2, photochemical oxidants (including ozone) and SO_2.

The first two-way catalytic converters came into use in 1975 as part of CARB's Motor Vehicle Emission Control Program followed by an announcement that CARB will limit lead in gasoline.

In 1977, Volvo introduced a vehicle marketed as 'Smog-Free'. This vehicle supported the first three-way catalytic (TWC) converter to control HC, NO_x and CO emissions.

In 1980, the California population reached 24 million people. Total registered vehicles were in the region of 17 million and VMT was 155 billion (Table 5.2).

Total cumulative California vehicle emissions for NO_x and HCs remain at 1970 levels of 1.6 million tons/year, despite a rise of 45 billion in VMT over those 10 years.

The legislative controls had clearly begun to have a positive effect. Spurred on by this victory, CARB began a program of compliance testing on 'in-use' vehicles in order to determine whether they continue to comply with emission standards as vehicle mileage increases. Vehicle manufacturers commissioned the development of more durable emission control systems.

Introduction of the biennial California Smog Check Program was seen in 1984, the aim of which was to identify vehicles in need of maintenance and to confirm the effectiveness of their emissions control systems (Figure 5.3).

1943: First smog alarm in LA
1950: 4,5 Mio. vehicles in California
1952: Dr Arie Haagen-Smit analyses the reasons for smog development
1960: 8 Mio. vehicles in California
1961: Introduction of crankcase ventilation (PCV)
1966: Federal Clean Air Act
1967: Foundation of CARB, Chairman: Haagen-Smit
1970: Foundation of EPA
1980: 17 Mio. vehicles in California
1988: CARB decides OBD II for 1994 MY
1990: Number of smog days goes down, CARB decides LEV and ZEV - program
1995: 26 Mio. vehicles in California
1996: Ozon pollution 59% below 1965, number of smog days 94% below 1975

Figure 5.3 History of CARB emission legislation activity

The mid-term period of emissions control legislation ended in 1988 with a key announcement, which saw the beginning of on-board diagnostics. The California Clean Air Act was signed and CARB adopted regulations that required that all 1994 and beyond model year cars were fitted with 'on-board diagnostic' systems. The task of these systems was, as it is now:

> *To monitor the vehicle emissions control systems performance and alert owners when there is a malfunction that results in the lack of function of an emissions control system/sub-system or component.*

5.1.4 Introduction of vehicle emissions control strategies

To meet the ever increasing but justifiable and 'wanted' need of vehicle emissions control legislation, vehicle manufacturers were forced to invest heavily in the research and development of Vehicle Emission Control Strategies. Building upon the foundation laid by PCV, the two-way and three-way catalyst, manufacturers further developed emissions control hardware. Such systems included exhaust gas recirculation, secondary air injection, fuel tank canister purge, spark timing adjustment, air/fuel ratio (AFR) control biasing, fuel shut off under negative torque conditions (overrun or cruise down), to name but a few.

This development continued and expanded meaning that these systems demanded an ever-increasing array of sensors and actuators. The resolution of measurement, control of AFR, actuator displacement rates and accuracy of displacement, etc. was way beyond that which could be provided by traditional existing mechanical technologies.

At about this time, an enabler was provided in the form of recent advances in microprocessor technology. The path was clear, the drivers for OBD system monitoring were in force and the enablers were available. On-board diagnostics was born.

5.2 What is on-board diagnostics?

Fundamentally, a contemporary microprocessor-based on-board diagnostics or OBD system is intended to self-diagnose and report when the performance of the vehicle's emissions control systems or components have degraded. This is to the extent that the tailpipe emissions have exceeded legislated levels or are likely to be exceeded in the long term.

When an issue occurs, the OBD system illuminates a warning lamp known as the malfunction indicator lamp (MIL) or malfunction indicator (MI) on the instrument cluster. In the United States, this symbol often appears with the phrase 'Check Engine', 'Check' or 'Service Engine Soon' contained within it. European vehicles tend to use the engine symbol on an orange background (Figure 5.4).

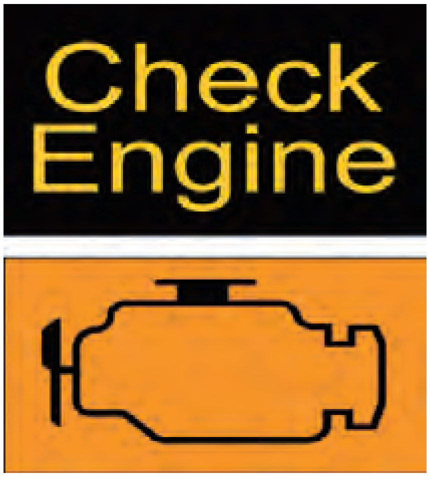


Figure 5.4 Malfunction indicator lamp symbols

When the fault occurs, the system stores a diagnostic trouble code (DTC) that can be used to trace and identify the fault. The system will also store important information that pertains to the operating conditions of the vehicle when the fault was set. A service technician is able to connect a diagnostic scan tool or a code reader that will communicate with the microprocessor and retrieve this information. This allows the technician to diagnose and rectify the fault, make a repair/replacement, reset the OBD system and restore the vehicle emissions control system to a serviceable status.

Key fact

When the fault occurs, the system stores a diagnostic trouble code (DTC).

As vehicles and their systems become more complex, the functionality of OBD is being extended to cover vehicle systems and components that do not have anything to do with vehicle emissions control. Vehicle body, chassis and accessories such as air conditioning or door modules can now also be interrogated to determine their serviceability as an aid to fault diagnosis.

5.2.1 OBD scenario example

While driving, a vehicle owner observes that the vehicle's engine 'lacks power' and 'jumps sometimes'. This is a problem often faced by technicians in that customers often have no engineering or automotive knowledge and use lay terms to describe what is happening with a very complex system. The driver does, however, report that the MIL has been illuminated.

The technician connects a scan tool that can communicate using an industry standard communications protocol. The OBD code memory is checked and data is presented in a way that also conforms to a standard. DTC P1101 with the description 'MAF sensor out of self-test range' is stored in memory, which means that the OBD system component monitor has identified the mass airflow (MAF) sensor circuit voltage as outside an acceptable range (Figure 5.5).

Upon confirming the fault the system was smart, it defaulted to a 'safe' value of MAF, a concept known as failure mode effects management (FMEM), to allow the driver to take the vehicle to a place of repair. While this FMEM value was a good short-term solution, it is not a sufficient substitute for the full functionality of a serviceable MAF sensor.

Definition

FMEM: Failure mode effects management allows the driver to take the vehicle to a place of repair Also described as 'limp home' mode.

Since the MAF sensor determines the MAF going into the engine intake, it will be impossible for the system to run at the optimum AFR for efficient burning of the air/fuel charge within the cylinder. It may be that tailpipe emissions are likely to rise beyond legislated limits.

Also, the MAF sensor is used by other emissions control systems on the vehicle, now that its input is unreliable it follows that those systems are no longer working

Figure 5.5 Hot wire MAF sensor (Source: Bosch)

at their optimum levels and may not work at all. This is the reason for the MIL illumination, which says, in as many words, ‘An emissions control system/sub system or component has become unserviceable!’

Visual inspection of the MAF sensor reveals that it has become damaged beyond repair and needs replacing. This is carried out, the technician clears the DTC from the OBD system memory, resets the system, and takes a short test drive; later the diagnostic scan tool confirms that the DTC is no longer present. The road test also confirms that the previous drive issue is no longer apparent.

5.2.2 Origins of OBD in the United States

The previous example relates to the current situation, but when OBD was first introduced, standards and practices were less well defined. Manufacturers developed and applied their own systems and code descriptions. This state of affairs was obviously undesirable since non-franchised service and repair centres had to understand the various subtleties of each system; this meant having different scan tools, as well as a multitude of leads, manuals and connectors. This made diagnostics unwieldy and expensive. This stage became known as ‘OBD1’, the first stage of OBD introduction.

In the late 1980s, the Society of Automotive Engineers (SAE) defined a list of standard practices and recommended these to the EPA. The EPA acknowledged the benefits of these standards and recommendations, and adopted them. In combination, they changed the shape and application of OBD. The recommendations included having a standard diagnostic connector, a standard scan tool and a communications protocol that the standard scan tool could use to interface with the vehicle of any manufacturer.

Key fact

OBD2 was adopted for implementation by 1 January 1996.

The standard also included mandatory structures and descriptions for certain emission control system/component defects. These were called ‘P0’ codes. Manufacturers were still free to generate their own ‘manufacturer-specific code descriptions’ known as ‘P1’ codes. This phase of implementation became known as OBD2 and was adopted for implementation by 1 January 1996.

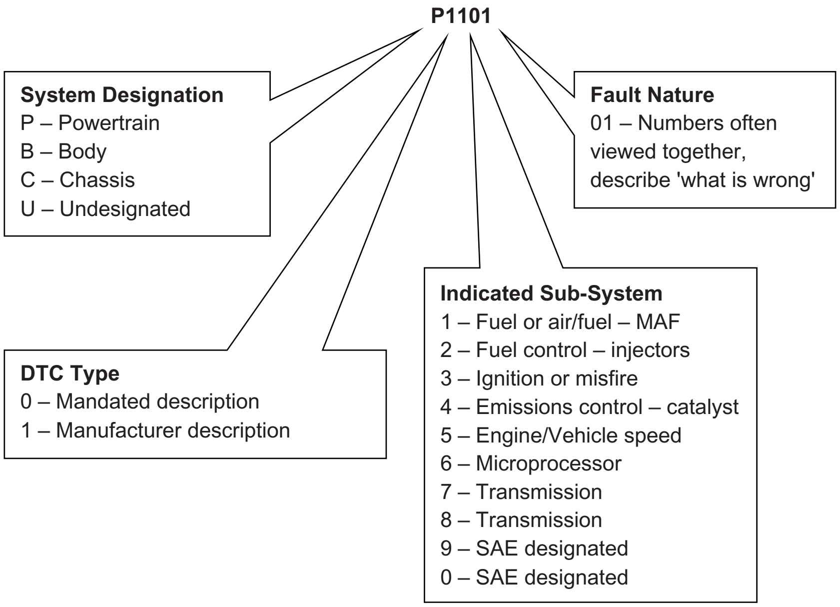


Figure 5.6 P-code composition

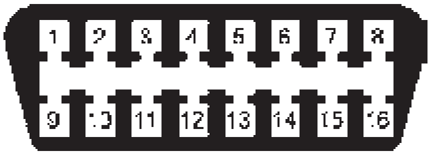


Figure 5.7 Sixteen pin DLC OBD2/ EOBD connector

5.2.3 P-code composition

The DTC is displayed as a five-character alphanumeric code. The first character is a letter that defines which vehicle system set the code, be it powertrain, body or chassis.

- P means powertrain system set the code.
- B means body system set the code.
- C means chassis system set the code.
- U is currently unused but has been ‘stolen’ to represent communication errors.

P codes are requested by the microprocessor controlling the powertrain or transmission and refer to the emissions control systems and their components.

B codes are requested by the microprocessor controlling the body control systems. Collectively these are grouped as lighting, air conditioning, instrumentation or even in-car entertainment or telematics.

C codes are requested by the microprocessor controlling the chassis systems that control vehicle dynamics such as ride height adjustment, traction control, etc.

The four numbers that follow the letter detail information pertaining to what sub-system declared the code. Using the example from before, see Figure 5.6.

An integral feature of the OBD system is its ability to store fault codes relating to problems that occur with the engine electronic control system, particularly faults that could affect the emission control system, and this is one of its primary functions. For the diagnostic technician, this is a powerful feature which can clearly assist in locating and rectifying problems on the vehicle when they occur.

The diagnostic socket used by systems conforming to European OBD (EOBD)/ OBD2 standards should have the following pin configuration (Figure 5.7):

1 ignition positive supply;
2 bus + line, SAE J1850 (PWM);

3. manufacturer’s discretion;
4. chassis ground;
5. signal ground;
6. CAN bus H;
7. K-line;
8. manufacturer’s discretion;
9. manufacturer’s discretion;
10. bus – line (PWM);
11. manufacturer’s discretion;
12. manufacturer’s discretion;
13. manufacturer’s discretion;
14. CAN bus L;
15. L-line or second K-line;
16. vehicle battery positive.

With the introduction of OBD2 and EOBD, this feature was made even more powerful by making it more accessible. Standardisation of the interface connector known as the diagnostic link connector (DLC) and communication protocol allowed the development of generic scan tools, which could be used on any OBD compliant vehicle.

5.2.4 European on-board diagnostics and global adoption

Europe was not immune to the environmental issues associated with smog. A major smog episode occurred in London in December 1952; this lasted for five days and resulted in approximately 4000 deaths. The UK government passed its first Clean Air Act in 1956, which aimed to control domestic sources of smoke pollution.

In 1970, the then European community adopted directive 70/220 EEC – ‘Measures to be taken against Air Pollution by Emissions from Motor Vehicles’. This basically set the foundation for future legislation to curb motor vehicle pollution in Europe. This directive was amended over the next three decades when in October 1998 the amendment 98/69/EC ‘On-Board Diagnostics (OBD) for Motor Vehicles’ was adopted, which added Annex XI to the original 70/220 document. Annex XI details the functional aspects of OBD for motor vehicles in Europe and across the globe. This became known as EOBD.

5.2.5 Summary

A major contributing factor to environmental health issues in the United States was found to be motor vehicle emissions pollution. Scientific studies by government sponsored academic establishments and vehicle manufacturers then took place over several years. Legislative bodies were formed, which later developed and enacted vehicle emissions control legislation that forced vehicle manufacturers to develop control strategies and incorporate them within their production vehicles.

As microprocessor technologies became more advanced and commercially viable, the legislation was augmented to include a self-diagnosing OBD system, which would report when the emissions control system was unserviceable. First attempts by manufacturers to use such a system were applied unilaterally, which resulted in confusion, regenerative work and a poor reception of the OBD (now termed OBD1) concept. A revision of the legislation adopted SAE recommended standards, which resulted in the OBD (now termed OBD2) system becoming largely generic and applicable across the whole range of vehicle manufacturers.

As environmental activism spread across to Europe, vehicle manufacturers realised they had to support a philosophy of sustainable growth. Similar legislation was adopted and EOBD manifested itself in a form very similar to that observed in the United States.

5.3 Petrol/Gasoline on-board diagnostic monitors

5.3.1 Introduction

This section will cover the fundamentals of some of the OBD systems employed on mainstream petrol/gasoline vehicles. The concept of how the OBD system is divided into a series of software-based serviceability indicators, known as 'OBD monitors', is also covered.

5.3.2 Legislative drivers

In Europe, the European Directive 70/220 EEC was supplemented by European Directive 98/69/EC (Year: 1998: OJ Series: L – OJ Number: 350/1). This introduced legislation mandating the use of OBD systems in passenger vehicles manufactured and sold after 1 January 2001.

In the United States, legislation was first introduced in California by the California Air Resources Board (CARB) in 1988 and later federally by the Clean Air Act Amendments of 1990. This meant that the enforcing body, the EPA, requires that states have to develop state implementation plans (SIPs) that explain how each state will implement a plan to clean up pollution from sources including motor vehicles. One aspect of the requirement is the performance of OBD system checks as part of the required periodic inspection.

In order to be compliant with legislation and sell vehicles, manufacturers needed to engineer 'early warning' monitoring sub-systems that would determine when emission control systems had malfunctioned to the extent that tailpipe emissions had (or were likely to in the long term) exceeded a legislated level. OBD 'monitors' were derived for this purpose.

5.3.3 Component monitoring

The emission control systems integral with the vehicle employ many sensors and actuators. A software program housed within a microprocessor defines their actions.

The 'component monitor' is responsible for determining the serviceability of these sensors and actuators. Intelligent component drivers linked to the microprocessor have the ability to enable/disable sensors/actuators and to receive signals. The analogue inputs from the sensors are converted to digital values within the microprocessor.

In combination with these component drivers, the microprocessor possesses the functionality to detect circuit faults on the links between microprocessor and

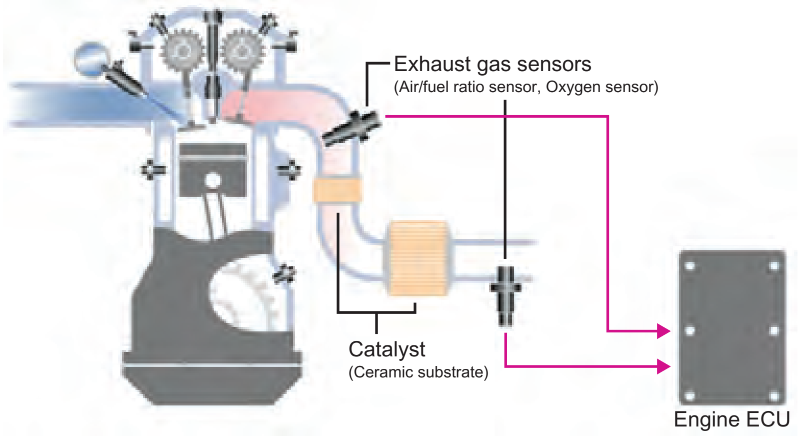


Figure 5.8 Exhaust gas oxygen sensors positioned pre- and post-catalyst (Source: http://www.globaldensoproducts.com)

component. In addition, rationality tests can be performed to determine whether the sensor is operating out of range of its specification.

5.3.4 Rationality testing

Rationality tests can be performed on such sensors as the MAF sensor and throttle body. For example, the MAF is tested by observing its output value in comparison to a 'mapped' value normalised by throttle position and engine speed. The map or table contains expected MAF output values for the engine speed/throttle set point.

Should the MAF output lie outside of an acceptable range (threshold) of values for that engine speed/throttle set point, then a fault is reported.

5.3.5 Circuit testing

The component monitor is capable of monitoring for circuit faults. Open circuits, short circuits to ground or voltage can be detected. Many manufacturers also include logic to detect intermittent errors.

5.3.6 Catalyst monitor

The purpose of the catalyst is to reduce tailpipe/exhaust emissions. The 'catalyst monitor' is responsible for determining the efficiency of the catalyst by inferring its ability to store oxygen. The method favoured most by the majority of manufacturers is to fit an oxygen sensor before and after the catalyst.

As the catalyst's ability to store oxygen (and hence perform three-way catalysis) deteriorates, the oxygen sensor downstream of the sensor will respond to the oxygen in the exhaust gas stream and its signal response will exhibit a characteristic similar to the upstream oxygen sensor (Figures 5.8 and 5.9).

An algorithm within the microprocessor analyses this signal and determines whether the efficiency of the catalyst has degraded beyond the point where the vehicle tailpipe emissions exceed legislated levels. If the microprocessor determines that this has occurred, then a malfunction and a DTC are reported. Repeat detections of a failed catalyst will result in MIL illumination (Figure 5.10).

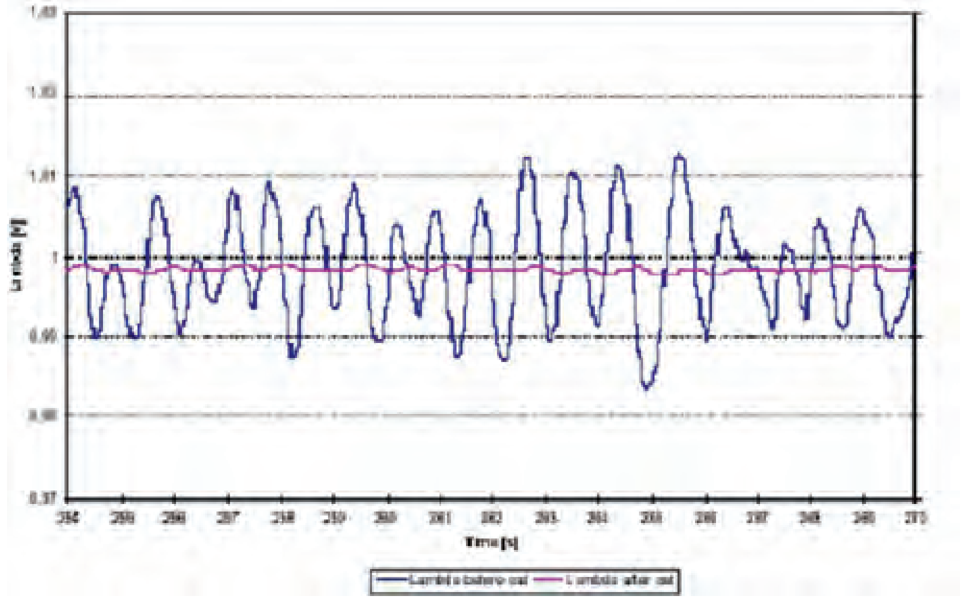


Figure 5.9 Upstream and downstream exhaust gas sensor activity – good catalyst (Source: SAE 2001-01-0933 New Cat Preparation Procedure for OBD2 Monitoring Requirements)

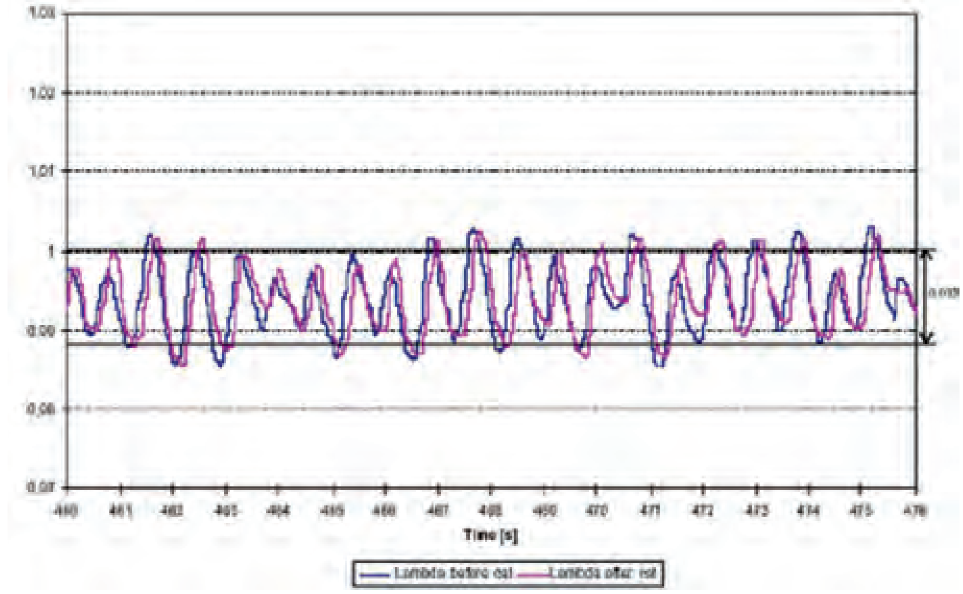


Figure 5.10 Upstream and downstream exhaust gas sensor activity – failed catalyst (Source: SAE 2001-01-0933 New Cat Preparation Procedure for OBD2 Monitoring Requirements)

5.3.7 Evaporative system monitor

The purpose of the evaporative (EVAP) emissions control system is to store and subsequently dispose off unburned HC emissions, thus preventing them from entering the atmosphere. This is achieved by applying a vacuum across the fuel tank. The vacuum then causes fuel vapour to be drawn through a carbon canister in which the HC vapours are collected and stored.

Evaporative emissions control system

1 Line from fuel tank to carbon canister. **2** Carbon canister. **3** Fresh air. **4** Canister-purge valve.
5 Line to intake manifold. **6** Throttle valve.
p_s Intake manifold pressure. p_u Atmospheric pressure. Δp Difference between intake manifold pressure and atmospheric pressure.

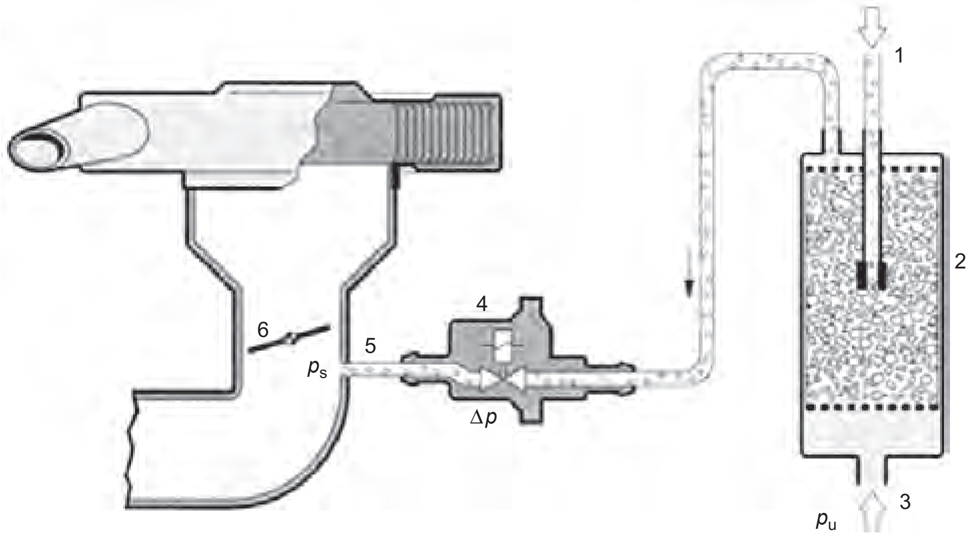


Figure 5.11 Evaporative emissions control system (Source: Bosch)

During certain closed loop fuel control conditions, the microprocessor activates a solenoid-controlled ‘vapour management valve’. This allows the manifold vacuum to draw vapour from the carbon canister along vapour lines, which terminate in the intake manifold. The fuel vapour is then combined and combusted with the standard air/fuel charge; the closed loop fuel control system caters for the additional AFR enrichment to ensure that stoichiometric fuelling continues (Figure 5.11).

The evaporative system monitor is responsible for determining the serviceability of the EVAP system components and detecting leaks in the vapour lines. Most manufacturers check for fuel vapour leaks by employing a diagnostic that utilises a pressure or vacuum test on the fuel system.

European legislation dictates that these checks are not required. However, vehicles manufactured in the United States after 1996 and before 1999 generally employ a system that uses a pressure or vacuum system. This must be able to detect a leak in a hose or filler cap that is equivalent to that generated by a hole, which is 0.040 inch (1 mm) in diameter. Vehicles manufactured after 2000 must support diagnostics that are capable of detecting a 0.020 inch (0.5 mm) hole.

5.3.8 Fuel system monitoring

As vehicles accumulate mileage so also do the components, sensors and actuators of the emissions control systems. MAF sensors become dirty and their response slows with age. Exhaust gas oxygen sensors also respond slower as they are subject to the in-field failure modes such as oil and fuel contamination, thermal stress and general ageing. Fuel pressure regulators perform outside of their optimum capacity; fuel injectors become slower in their response; and partial blockages mean that they deliver less and sometimes more fuel than requested.

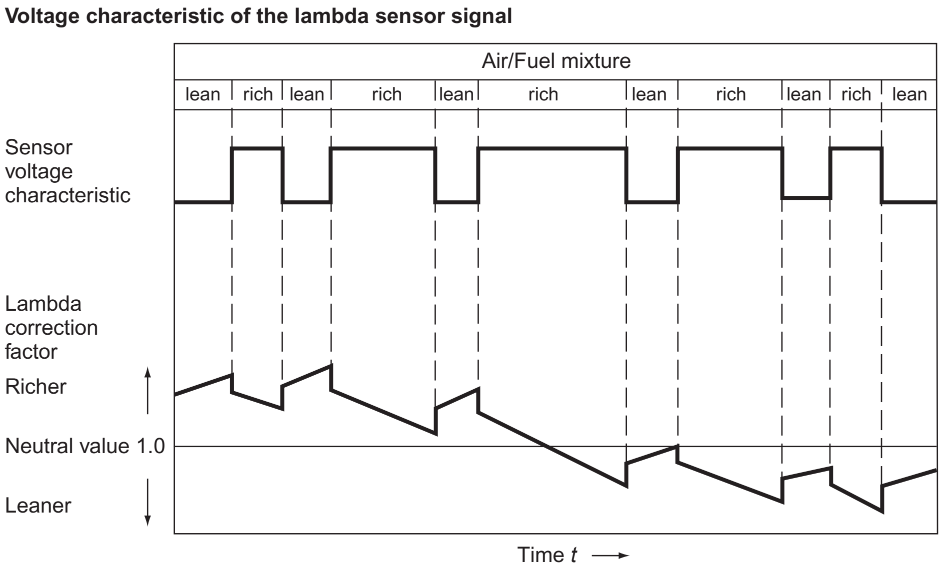


Figure 5.12 Rich AFR lambda sensor signal fuelling error (Source: Bosch)

If this component ageing were not compensated for, it would mean that the fuel system would not be able to maintain normal fuelling around stoichiometric AFR as shown in Figure 5.12. The end result would be the potential to exceed emission limits. A lambda value of one is required in order for the three-way catalyst to work. In addition to this, more severe fuelling errors would cause noticeable effects in the performance of the vehicle leading to customer complaints and potential to damage the manufacturer's brand image.

This compensation strategy is known as adaptive learning. A dedicated piece of software contained within the electronic control unit (ECU) learns these deviations away from stoichiometry, while the fuel system is in closed loop control. They are stored in a memory that is only reset when commanded by a technician and which is also robust to battery changes.

These memory-stored corrections are often termed 'long-term' fuel corrections. They are often stored in memory as a function of air mass, engine speed or engine load.

An exhaust gas oxygen sensor detects the amount of oxygen in the catalyst feed gas and the sensor produces a voltage, which is fed back to the microprocessor. This is then processed to determine the instantaneous or 'short-term' fuel correction to be applied. This is done in order to vary the fuel around stoichiometry and allow three-way catalysis to occur.

The microprocessor then calculates the amount of fuel required using an equation, which is shown here in its most basic form (Figure 5.13).

$$\text{Fuel mass} = \frac{\text{air mass} \times \text{long-term fuel trim}}{\text{short-term fuel trim} \times 14.64}$$

Referring to Figures 5.11 and 5.12, it can be seen that when there is a component malfunction, which causes the AFR in the exhaust stream to be rich, then there is a need to adapt to this to bring the AFR back into the region of stoichiometry. The value of the long-term fuel trim correction must decrease because less fuel is required.

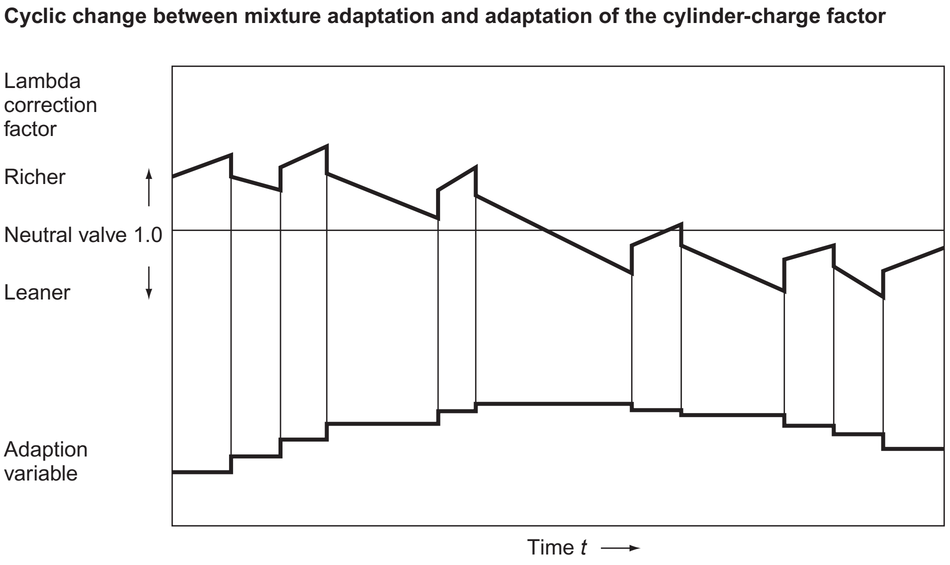


Figure 5.13 Adaptive fuel strategy in operation (Source: Bosch)

Should the situation continue and the problem causing rich AFR become slowly worse, the error adaption will continue with an ever-decreasing value for the long-term fuel trim being applied, learned and stored in memory.

The purpose of the fuel monitor is to determine when the amount of long-term adaptive correction has reached the point where the system can no longer cope. This is also where long-term fuel trim values reach a pre-defined or 'calibrated' limit at which no further adaption to error is allowed. This limit is calibrated to coincide with exhaust tailpipe emissions exceeding legislated levels. At this point and when a short-term fuelling error exceeds another 'calibrated' limit, a DTC is stored, and after consecutive drives, the MI is illuminated.

The opposite occurs, with extra fuel being added, via the long-term fuel trim parameter, should an error occur that causes the AFR at the exhaust gas oxygen sensor to be lean.

5.3.9 Exhaust gas recirculation monitor

As combustion takes place within the engine cylinders, nitrogen and oxygen combine to form various oxides of nitrogen, collectively termed as NO_x. NO_x emissions can be reduced up to a certain point by enriching the AFR, beyond the point at which HC and CO emissions begin to increase. NO_x emissions are generated as a function of combustion temperature, so another way to reduce these is to decrease the compression ratio which leads to other inefficiencies like poor fuel economy.

Most manufacturers employ an emissions control sub-system known as exhaust gas recirculation (EGR). This by definition recirculates some of the exhaust gases back into the normal intake charge. These 'combusted' gases cannot be burnt again so they act to dilute the intake charge. As a result, in-cylinder temperatures are reduced along with NO_x emissions (Figure 5.14).

The EGR system monitor is responsible for determining the serviceability of the sensors, hoses, valves and actuators that belong to the EGR system. Manufacturers employ systems that can verify that the requested amount of exhaust gas is flowing back into the engine intake. Methods can be both

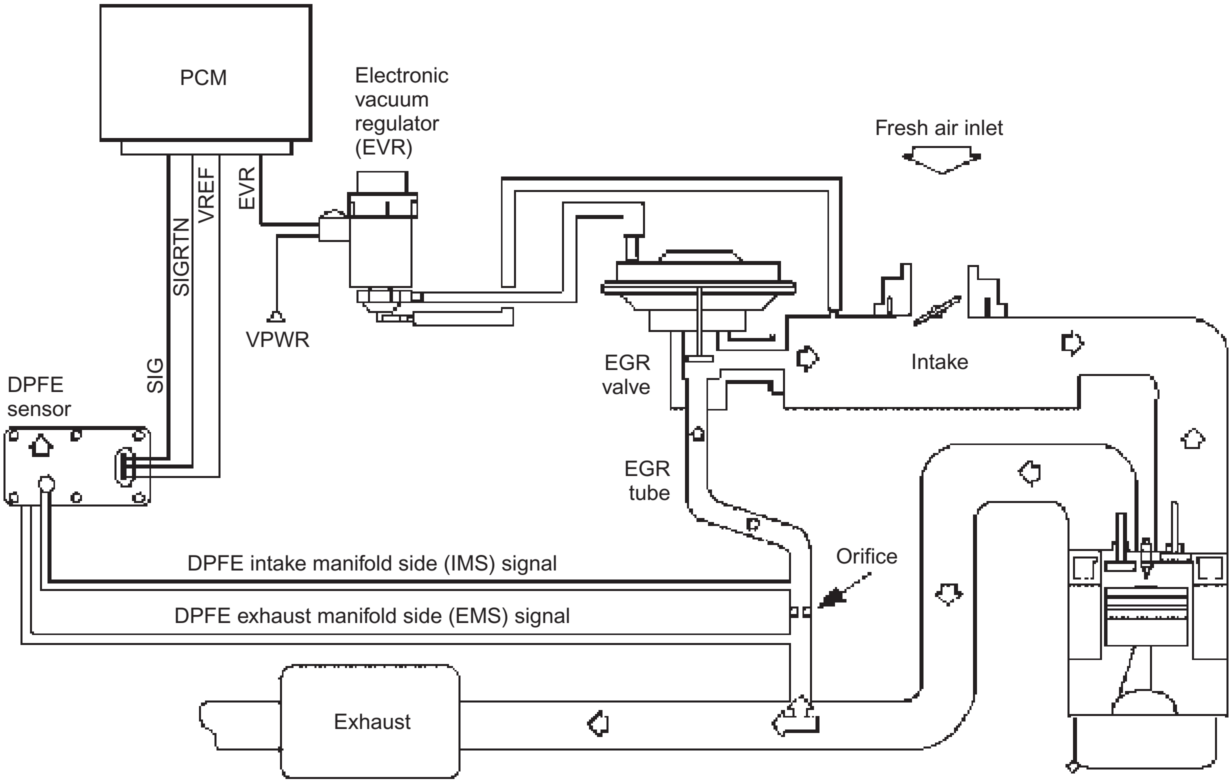


Figure 5.14 EGR system using differential pressure monitoring (Source: Ford Motor Company)

intrusive and non-intrusive, such as a change in manifold pressure as EGR flows and is then shut off.

One method monitors AFR excursions after the EGR valve is opened, and then closed as the AFR becomes lean. Another system employs a differential pressure scheme that determines the pressure both upstream and downstream of the exhaust to determine whether the requested flow rate is in effect. Yet another system employs a temperature sensor, which reports the change in temperature as EGR gases flow past the sensor. The temperature change will be mapped against the amount of EGR flowing, so when an amount of EGR is requested, the flow rate is inferred by measuring the change in temperature.

5.3.10 Secondary air monitor

The exhaust system catalyst is not immediately operative following a start where the engine and exhaust system is cool. Temperature thresholds above which the catalyst is working, and three-way catalysis is occurring, vary as a function of the exhaust gas system package. Typically, this ‘light off’ point occurs at temperatures of approximately 260 °C/500 °F. Some manufacturers employ electrically heated catalysts to reach this temperature rapidly, but these are expensive to manufacture and replace.

Most manufacturers rely on the exhaust gases as a source of heat in order to bring the catalyst up to light off temperature. When the vehicle is started from cold, the AFR is rich; this is required to ensure a stable engine start for cold pull-away. From an emissions perspective, the impact is observed in the production of HC and CO in the exhaust stream because the exhaust system catalyst has not reached light off.

Definition

Catalyst light off temperature is the point at which it starts to operate fully.

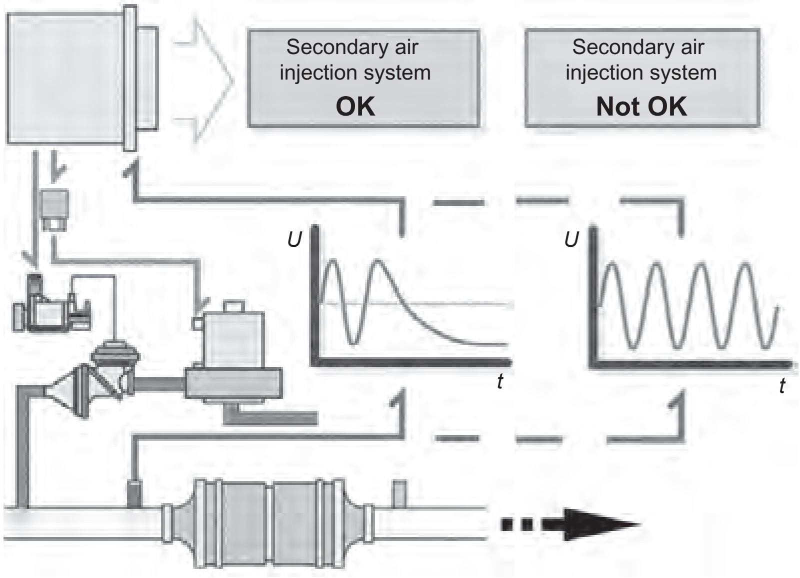


Figure 5.15 Secondary airflow diagnostic monitoring

The secondary air system uses a pump, which adds more air into the exhaust stream at a point before the catalyst follows a cold start. The secondary air combusts the HC in the catalyst, generating heat, which, in turn, promotes light off and further emissions reduction.

Older systems support a belt-driven mechanical pump with a bypass valve when secondary airflow is not required. Modern vehicles employ an electric air pump operated by the engine-management ECU [powertrain control module (PCM)] via relays.

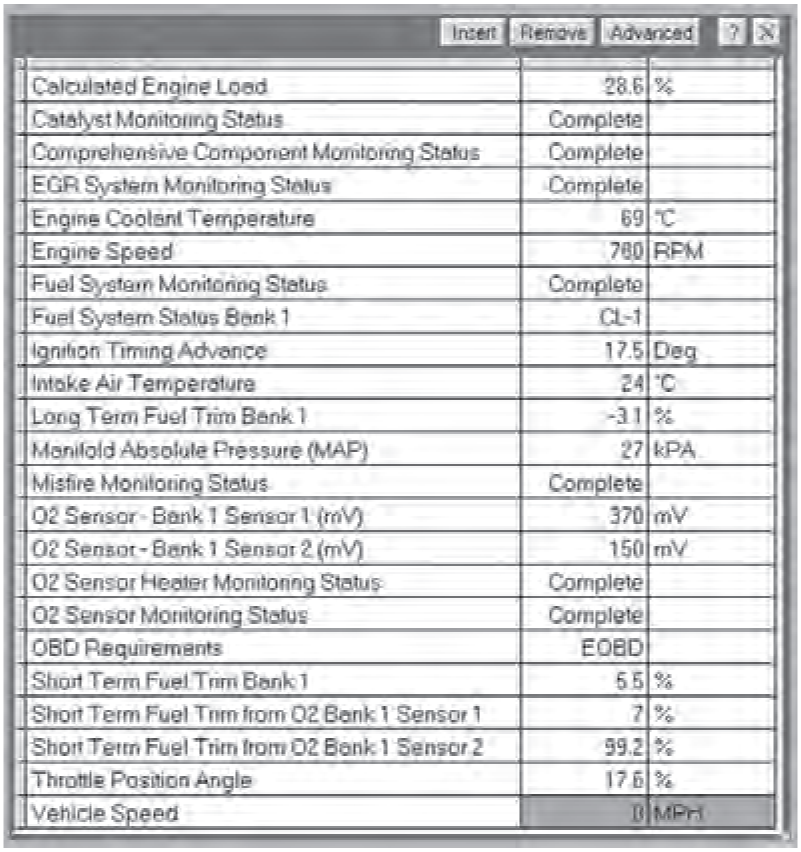
The secondary air monitor is responsible for determining the serviceability of the secondary air system components. Most strategies monitor the electrical components and ensure the system pumps air when requested by the ECU. To check the airflow, the ECU observes the response of the exhaust gas oxygen sensor after it commands the fuel control system to enter open-loop control and force the AFR to become rich. The secondary air pump is then commanded on and the ECU determines the time taken for the exhaust gas oxygen sensor to indicate a lean AFR. If this time exceeds a calibrated threshold, a DTC is stored (Figure 5.15).

5.3.11 Monitors and readiness flags

An important part of any OBD system is the system monitors and associated readiness flags. These readiness flags indicate when a monitor is active. Certain monitors are continuous, for example, misfire and fuel system monitors.

Monitor status (ready/not ready) indicates if a monitor has completed its self-evaluation sequence. System monitors are set to ‘not ready’ if cleared by scan tool and/or the battery is disconnected. Some of the monitors must test their components under specific, appropriate preconditions:

- The evaporative system monitor has temperature and fuel fill level constraints.
- The misfire monitor may ignore input on rough road surfaces to prevent false triggers.
- The oxygen sensor heater must monitor from a cold start.



Calculated Engine Load	28.6	%
Catalyst Monitoring Status	Complete	
Comprehensive Component Monitoring Status	Complete	
EGR System Monitoring Status	Complete	
Engine Coolant Temperature	69	°C
Engine Speed	780	RPM
Fuel System Monitoring Status	Complete	
Fuel System Status Bank 1	CL-1	
Ignition Timing Advance	17.5	Deg
Intake Air Temperature	24	°C
Long Term Fuel Trim Bank 1	-3.1	%
Manifold Absolute Pressure (MAP)	27	kPA
Misfire Monitoring Status	Complete	
O2 Sensor - Bank 1 Sensor 1 (mV)	370	mV
O2 Sensor - Bank 1 Sensor 2 (mV)	150	mV
O2 Sensor Heater Monitoring Status	Complete	
O2 Sensor Monitoring Status	Complete	
OBD Requirements	EOBD	
Short Term Fuel Trim Bank 1	5.5	%
Short Term Fuel Trim from O2 Bank 1 Sensor 1	7	%
Short Term Fuel Trim from O2 Bank 1 Sensor 2	99.2	%
Throttle Position Angle	17.6	%
Vehicle Speed	0	MPH

Figure 5.16 System monitors (marked as 'Complete') and live data shown in scan tool

Most other system monitors are not continuous and are only active under certain conditions. If these conditions are not fulfilled, then the readiness flag for that monitor is set to 'not ready'. Until the readiness flags are set appropriately, it is not possible to perform a test of the OBD system and its associated components (Figure 5.16).

There is no universal drive cycle that is guaranteed to set all the system monitors appropriately for a test of the OBD system. Most manufacturers and even cars have their own specific requirements, and irrespective of this, there are still some specific vehicles that have known issues when trying to set readiness flag status. To allow for this vehicles of model year 1996–2000 are allowed two readiness flags to be 'not ready'. After this, 2001 onwards, one readiness flag is allowed to be 'not ready' prior to a test.

5.4 Misfire detection

5.4.1 Misfire monitor

When an engine endures a period of misfire, at best tailpipe emissions will increase and at worst catalyst damage and even destruction can occur. When misfire occurs, the unburned fuel and air is discharged direct to the exhaust system where it passes directly through the catalyst.

Subsequent normal combustion events can combust this air/fuel charge in something akin to a bellows effect, which causes catalyst temperatures to rise considerably. Catalyst damage failure thresholds are package specific but are in

Key fact

When a misfire occurs, unburned fuel and air pass through the catalyst and can cause damage.

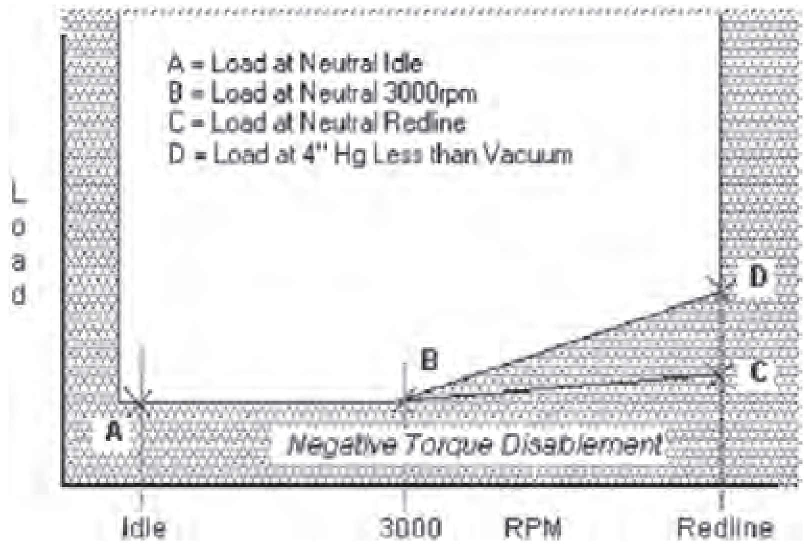


Figure 5.17 Misfire enablement window (Ford Motor Company)

the region of 1000 °C. The catalyst itself is a very expensive service item whether replaced by the customer or the manufacturer under warranty.

The misfire monitor is responsible for determining when misfire has occurred, calculating the rate of engine misfire and then initiating some kind of protective action in order to prevent catalyst damage.

The misfire monitor is in operation continuously within a ‘calibrateable’ engine speed/load window defined by the legislation. The United States requires misfire monitoring throughout the revs range but European legislation requires monitoring only up to 4500 rpm (Figure 5.17).

The crankshaft sensor generates a signal as the wheel rotates and the microprocessor processes this signal to determine the angular acceleration of the crankshaft produced by each engine cylinder when a firing event occurs. When a misfire occurs, the crankshaft decelerates and a cam position sensor identifies the cylinder that misfired.

Processing of the signal from the crank position sensor is not straightforward. A considerable amount of post-processing takes place to filter the signal and disable monitoring in unfavourable conditions. The misfire monitor must learn and cater for the differences in manufacturing tolerances of the crankshaft wheel and so has a specific sub-algorithm for learning these differences and allowing for them when calculating the angular acceleration of the crankshaft (Figure 5.18). These correction factors are calculated during deceleration, with the injectors switched off. They should be re-learned following driveline component changes such as flywheel, torque converter, crankshaft sensor, etc.

The misfire monitor must be able to detect two types of misfire:

- Type A misfire
- Type B misfire.

A type A misfire is defined as that rate of misfire, which causes catalyst damage. When this occurs, the MI will flash at a rate of 1 Hz and is allowed to stop flashing should the misfire disappear. The MI will stay on steady state should the misfire re-occur on a subsequent drive and the engine operating conditions are ‘similar’, that is, engine speed is within 375 rpm, engine load is within 20% and

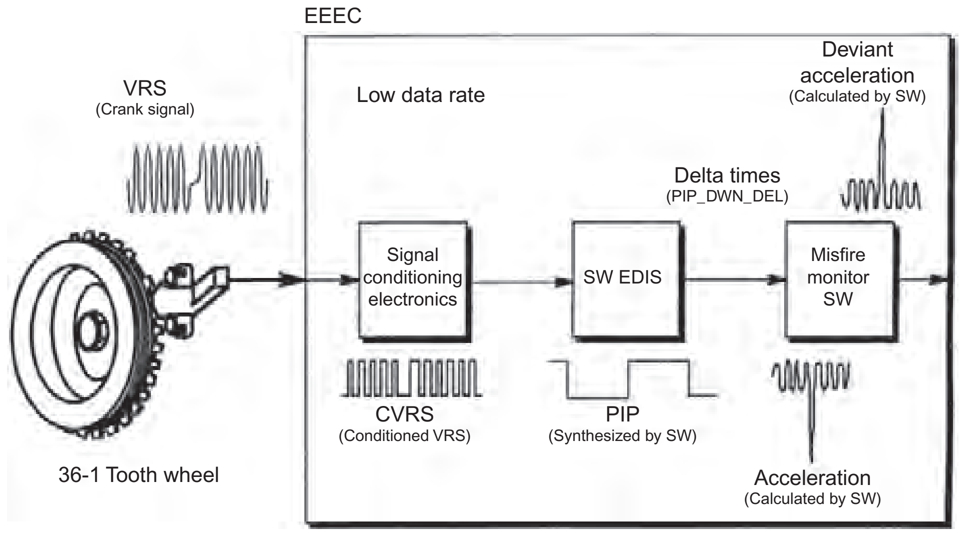


Figure 5.18 Crankshaft mounted wheel and sensor source of angular acceleration (Source: Ford Motor Company)

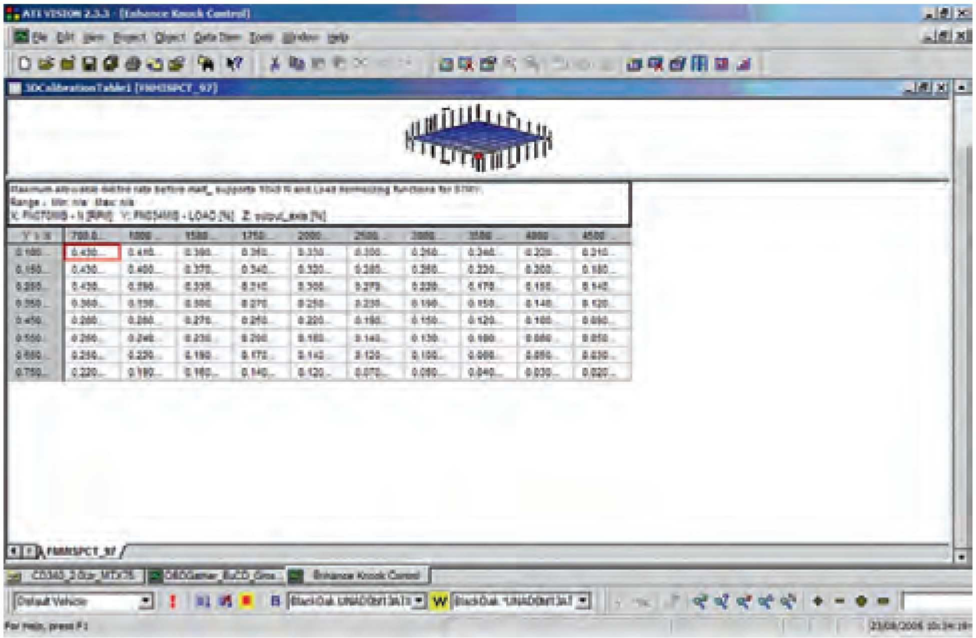


Figure 5.19 System development screen showing type A misfire rates normalised by engine speed and load (Source: Ford Motor Company)

the engine's warm-up status is the same as that under which the malfunction was first detected (and no new malfunctions have been detected).

The rate of misfire that will cause catalyst damage varies as a function of engine speed and load. Misfire rates in the region of 45% are required to damage a catalyst at neutral idle, while at 80% engine load and 4000 rpm, misfire rates in the region of only 5% are needed (Figure 5.19).

A type B misfire is defined as that rate of misfire which will cause the tailpipe emissions to exceed legislated levels. This varies from vehicle to vehicle and is dependent upon catalyst package. MI operation is the same as for standard DTCs.

The above is the most common method but misfires can be detected in a number of different ways as outlined in the following sections.

5.4.2 Crank speed fluctuation

A misfire event in a cylinder results in a lost power stroke. The gap in the torque output of the engine and a consequential momentary deceleration of the crankshaft can be detected using the crankshaft position sensor. By closely monitoring the speed and acceleration of the crankshaft, misfiring cylinders can be detected; this technology is very commonly used in OBD systems to detect non-firing cylinders that can cause harmful emissions and catalyst damage (Figure 5.20).

There are a number of technical challenges that have to be overcome with this technique, the accuracy achieved and reliability of the system is very dependent on the algorithms used for signal processing and analysis. Under certain conditions, misfire detection can be difficult, particularly at light load with high engine speed. Under these conditions, the damping of firing pulses is low due to the light engine load, and this creates high momentary accelerations and decelerations of the crankshaft. This causes speed variation which can be mistakenly taken by the OBD system as a misfire. With this method of misfire detection, careful calibration of the OBD system is necessary to avoid false detection. Another vehicle operation mode which can cause problems is operation of the vehicle on rough or poorly made roads. This also causes rapid crankshaft oscillation that could activate false triggers, and under these conditions the misfire detection must be disabled.

5.4.3 Ionising current monitoring

An ionisation current sensing ignition system consists of one ignition coil per cylinder, normally mounted directly above the spark plug. Eliminating the

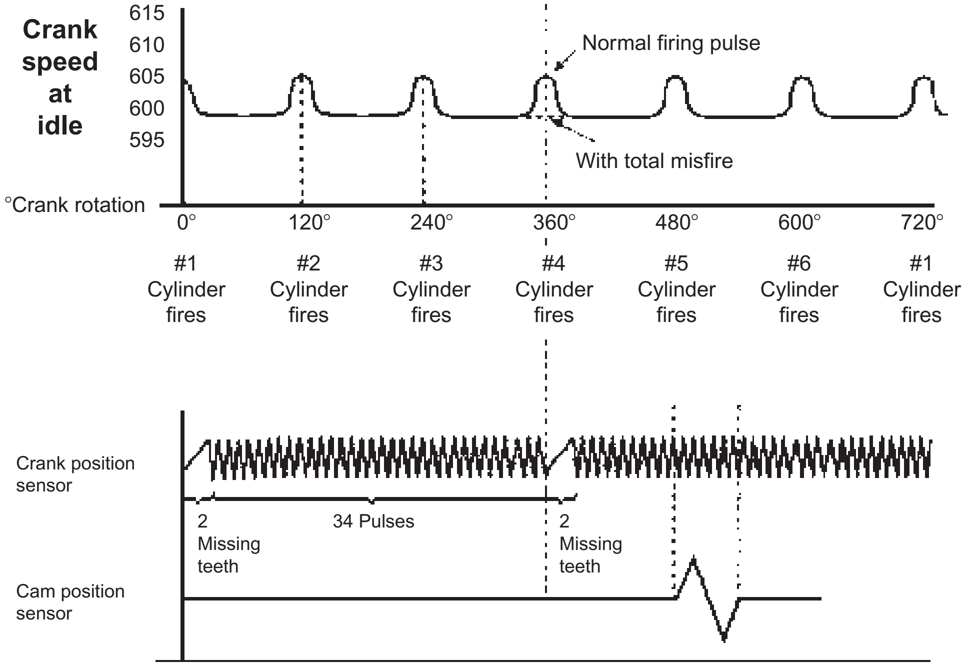


Figure 5.20 Misfire detection via crank sensor

distributor and high-voltage leads helps promote maximum energy transfer to the spark plug to ignite the mixture. In this system, the spark plug is not only used as a device to ignite the air/fuel mixture but is also used as an in-cylinder sensor to monitor the combustion process. The operating principle used in this technology is that an electrical current flow in an ionised gas is proportional to the flame electrical conductivity. By placing a direct current bias across the spark plug electrodes, the conductivity can be measured. The spark current is used to create this bias voltage and this eliminates the requirement for any additional voltage source.

The ion current is monitored, and if no ion-generating flame is produced by the spark, no current flows through the measurement circuit during the working part of the cycle. The ion current versus time trace is very different from that of a cycle when normal combustion occurs, and this information can be used as a differentiator to detect misfire from normal combustion. This method has proven to be very effective at monitoring for misfires under test conditions and also in practice.

The signal the system produces contains misfire information and, in addition, can provide objective knock or detonation information. This can be used for engine control systems where knowledge of the actual combustion process is required (as mentioned above) (Figures 5.21 and 5.22).

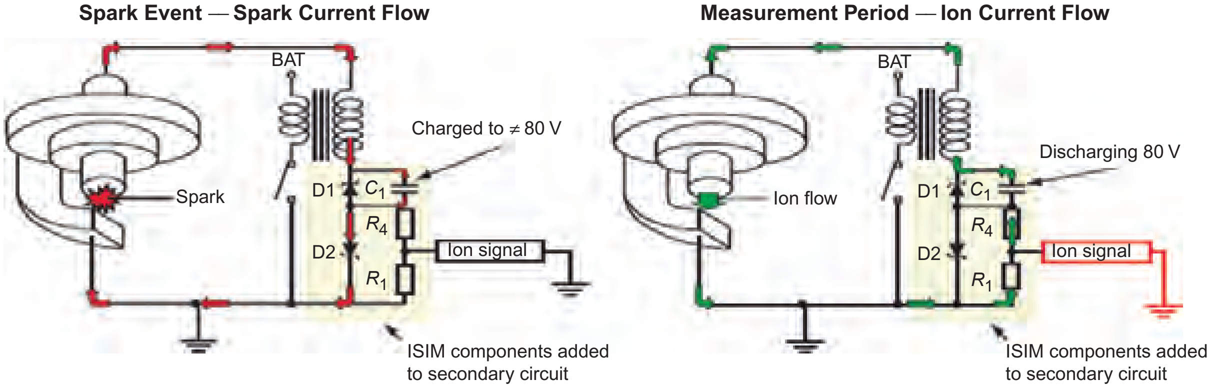


Figure 5.21 Ion-sensing circuit in direct ignition system

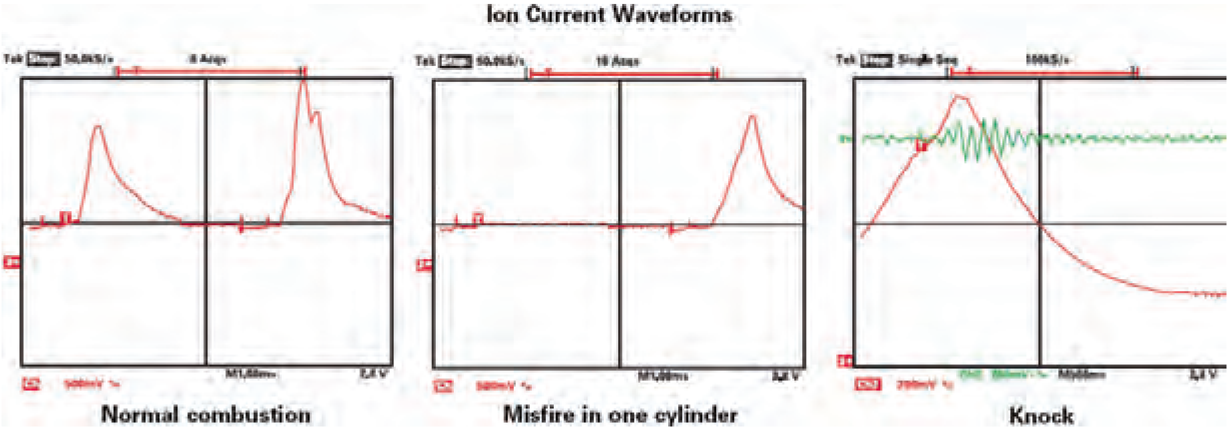


Figure 5.22 Resulting waveforms from the ion-sensing system

Figure 5.23 Cylinder pressure sensor mounted in the engine

5.4.4 Cylinder pressure sensing

This technology has great potential not just for OBD applications but also for additional feedback to the engine-management system about the combustion process due to the direct measurement technique (Figure 5.23). This additional control dimension can be utilised to improve engine performance and reduce emissions further. With respect to misfire detection, this method provides reliable detection of a positive combustion event and can easily detect misfire with utmost reliability.

The major drawback is the availability of suitable sensors that could be installed into the engine at production and would be durable enough to last the life of the engine and provide the required performance expected of sensors in an OBD system. For certain engine applications, sensors are available, and currently combustion sensor technology is under rapid development such that this technical hurdle will soon be overcome.

5.4.5 Exhaust pressure analysis

This solution involves using a pressure sensor in exhaust manifold combined with a Fourier analysis as the first stage of the signal processing. Using a sensor to analyse the gas pulses in the exhaust manifold, it is possible to detect single misfires, and additionally, it is possible to identify which cylinder is misfiring. This method is less intrusive than the above and could potentially be retrofitted at the production stage. A sensor in the exhaust can detect misfiring cylinders but cannot give useful, qualitative information about the combustion process. This technique has been demonstrated as capable of detecting all misfires at engine speeds up to 6000 rpm, for all engine configurations, loads and fuels. Generally, a ceramic capacitive–type sensor has been employed, which has a short response time and good durability.

5.5 OBD summary

OBD monitoring applies to systems which are most likely to cause an increase in harmful exhaust emission, namely

- all main engine sensors;
- fuel system;
- ignition system;
- EGR system.

The system uses information from sensors to judge the performance of the emission controls, but these sensors do not directly measure the vehicle emissions.

An important part of the system, and the main driver information interface, is the ‘check engine’ warning light, also known as the MIL. This is the main source of feedback to the driver to indicate if an engine problem has occurred or is present. When a malfunction or fault occurs, the warning light illuminates to alert the driver. Additionally, the fault is stored in the ECU memory. If normal condition is reinstated, the light extinguishes but the fault remains logged to aid diagnostics. Circuits are monitored for open or short circuits as well as plausibility. When a malfunction is detected, information about the malfunctioning component is stored.

An additional benefit allows the diagnostic technician to be able to access fault information and monitor engine performance via data streamed directly from the ECU while the engine is running (on certain vehicles). This information can be accessed via various scan tools available on the market and is communicated in a standardised format, so one tool (more or less!) works with all vehicles. The data is transmitted in a digital form via this serial interface. Thus, data values are transmitted as data words and the protocol used for this data stream has to be known in order to evaluate the information properly.

The benefits of having an OBD system are that it

- encourages vehicle and engine manufacturers to have a responsible attitude to reducing harmful emissions from their engines via the development of reliable and durable emission control systems;
- aids diagnosis and repair of complex electronic engine and vehicle control systems;
- reduces global emissions by identifying and highlighting immediately to the driver or user emission control systems in need of repair;
- provides ‘whole life’ emission control of the engine.

On-board diagnostics, or OBD, was the name given to the early emission control and engine-management systems introduced in cars. There was no single standard – each manufacturer often uses quite different systems (even between individual car models). OBD systems have been developed and enhanced, in line with United States government requirements, into the current OBD2 standard. The OBD2 requirement applies to all cars sold in the United States from 1996. EOBD is the European equivalent of the American OBD2 standard, which applies to petrol cars sold in Europe from 2001 (and diesel cars three years later).

Key fact

OBD2 (also OBDII) was developed to address the shortcomings of OBD1 and make the system more user friendly for service and repair technicians.

5.5.1 OBD2

Even though new vehicles sold today are cleaner than they have ever been, the millions of cars on the road and the ever-increasing miles they travel each day make them our single greatest source of harmful emissions. While a new vehicle may start out with very low emissions, infrequent maintenance or failure of components can cause the vehicle emission levels to increase at an undesirable rate. OBD2 works to ensure that the vehicles remain as clean as possible over their entire life. The main features of OBD2 are, therefore, as follows:

- malfunction of emission relevant components to be detected when emission threshold values are exceeded;
- storage of failures and boundary conditions in the vehicle’s fault memory;
- diagnostic light (MIL) to be activated in case of failures;
- read out of failures with generic scan tool.

The increased power of micro controllers (CPUs) in ECUs has meant that a number of important developments could be added with the introduction of OBD2. These include catalyst efficiency monitoring, misfire detection, canister purge and EGR flow rate monitoring. An additional benefit was the standardisation of diagnostic equipment interfaces.

For OBD1, each manufacturer applied specific protocols. With the introduction of OBD2, a standardised interface was developed with a standard connector for all vehicles, and a standardised theory for fault codes relating to the engine and powertrain (more about this later). This meant that generic scan tools could

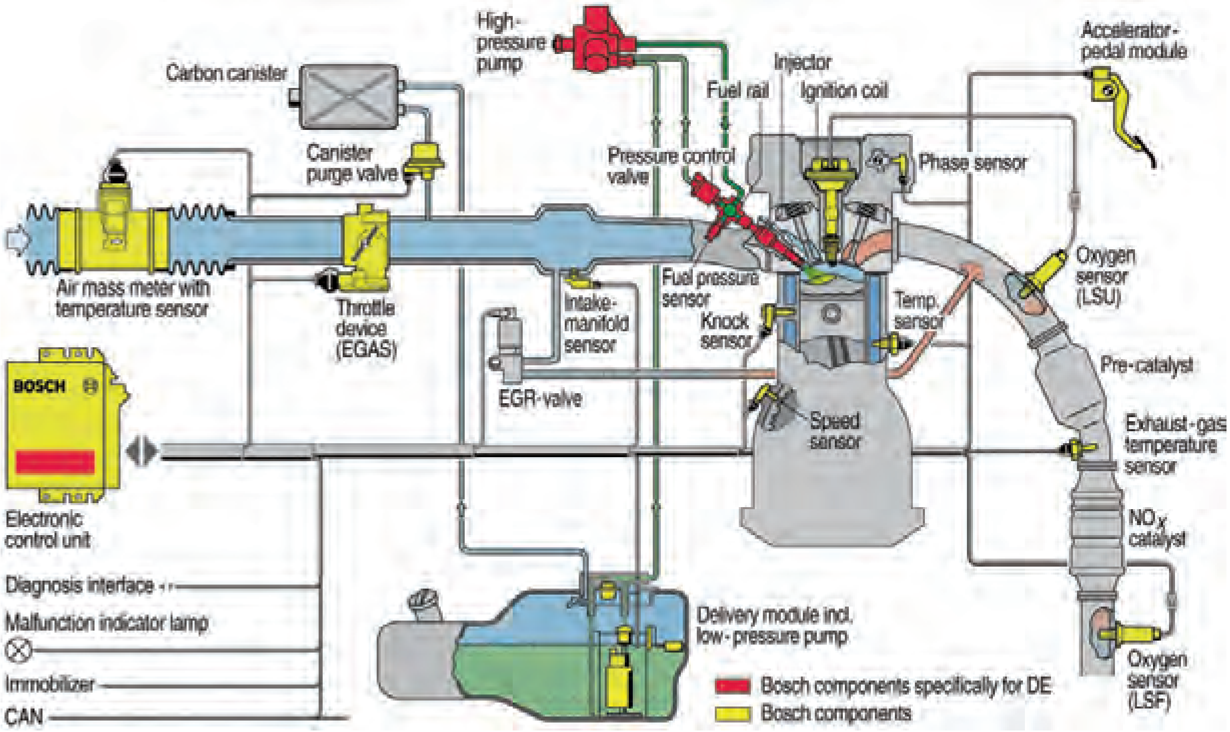


Figure 5.24 OBD2 system showing the main components of a gasoline direct injection system (Source: Bosch Media)

be developed and used in the repair industry by diagnostic technicians to aid troubleshooting of vehicle problems.

Another feature of OBD2 is that the prescribed thresholds at which a fault is deemed to have occurred are in relation to regulated emission limits. The basic monitor function is as follows:

- Monitoring of catalyst efficiency, engine misfire and oxygen sensors function such that crossing a threshold of 1.5 times the emission limit will record a fault.
- Monitoring of the evaporation control system such that a leak greater than the equivalent leak from a 0.04 inch hole will record a fault.

The main features of an OBD2 compliant system (as compared to OBD1) are as follows (Figure 5.24):

- pre- and post-catalyst oxygen sensors to monitor conversion efficiency;
- much more powerful ECU with 32 bit processor;
- ECU map data held on EEPROMS such that they can be accessed and manipulated via an external link; no need to remove ECU from vehicle for software updates or tuning;
- more sophisticated EVAP system, can detect minute losses of fuel vapour;
- EGR systems with feedback of position/flow rate;
- sequential fuel injection with MAP and MAF sensing for engine load.

5.5.2 EOBD

EOBD is an abbreviation of European on-board diagnostics. All petrol/gasoline cars sold in Europe since 1 January 2001, and diesel cars manufactured from 2003, must have OBD systems to monitor engine emissions. These systems were introduced in line with European directives to monitor and reduce emissions from cars. All such cars must also have a standard EOBD diagnostic socket that provides access to this system. The EOBD standard is similar to the US OBD2 standard. In Japan, the JOBD system is used. The implementation plan for EOBD was as follows:

- January 2000 OBD for all new petrol/gasoline vehicle models
- January 2001 OBD for all new petrol/gasoline vehicles
- January 2003 OBD for all new diesel vehicle models PC/LDV
- January 2004 OBD for all new diesel vehicles PC/LDV
- January 2005 OBD for all new diesel vehicles HDV.

The EOBD system is designed, constructed and installed in a vehicle such as to enable it to identify types of deterioration or malfunction over the entire life of the vehicle. The system must be designed, constructed and installed in a vehicle to enable it to comply with the requirements during conditions of normal use.

In addition, EOBD and OBD2 allow access to manufacturer-specific features available on some OBD2/EOBD compliant scan tools. This allows additional parameters or information to be extracted from the vehicle systems. These are in addition to the normal parameters and information available within the EOBD/OBD2 standard. These enhanced functions are highly specific and vary widely between manufacturers.

The monitoring capabilities of the EOBD system are defined for petrol/gasoline (spark ignition) and diesel (compression ignition) engines. The following is an outline:

Spark ignition engines

- Detection of the reduction in the efficiency of the catalytic converter with respect to emissions of HC only.
- The presence of engine misfires in the engine operation region within the following boundary conditions.
- Oxygen sensor deterioration.
- Other emission control system components or systems, or emission-related powertrain components or systems which are connected to a computer, the failure of which may result in tailpipe emission exceeding the specified limits.
- Any other emission-related powertrain component connected to a computer must be monitored for circuit continuity.
- The electronic evaporative emission purge control must, at a minimum, be monitored for circuit continuity.

Compression ignition engines

- Where fitted, reduction in the efficiency of the catalytic converter.
- Where fitted, the functionality and integrity of the particulate trap.
- The fuel injection system electronic fuel quantity and timing actuator(s) is/are monitored for circuit continuity and total function failure.
- Other emission control system components or systems, or emission-related powertrain components or systems which are connected to a computer, the failure of which may result in tailpipe emission exceeding the specified limits given. Examples of such systems or components are those for monitoring and control of air mass flow, air volumetric flow (and temperature), boost pressure and inlet manifold pressure (and relevant sensors to enable these functions to be carried out).
- Any other emission-related powertrain component connected to a computer must be monitored for circuit continuity (Table 5.3).

Definition

EOBD: European on-board diagnostics.

Table 5.3 Emission limits table for comparison

Legislation	**OBD malfunction limit (g/km)**			
	HC	**CO**	**NO_x**	**PM**
EPA	≥1.5 times the applicable federal standard			
EPA – method	Multiplicative relative to limits			
CARB 1 and 2	≥1.5 times the relevant CARB emission limits			
CARB 1 and 2 – method	Multiplicative relative to limits			
EOBD positive ign. 2000	0.40	3.20	0.60	–
EOBD diesel 2003	0.40	3.20	1.20	0.18
EOBD positive ign. 2005	0.20	1.40	0.30	–
EOBD diesel 2008 (for indication only)	0.30	2.40	0.90	0.14
EOBD – method	Absolute limits			

5.5.3 Features and technology of current systems

To avoid false detection, the legislation allows verification and healing strategies. These are outlined as follows:

MIL activation logic for detected malfunctions

To avoid wrong detections, the legislation allows verification of the detected failure. The failure is stored in the fault memory as a pending code immediately after the first recognition but the MIL is not activated. The MIL will be illuminated in the third driving cycle, in which the failure has been detected; the failure is then recognised as a confirmed fault.

MIL healing

The MIL may be deactivated after three subsequent sequential driving cycles during which the monitoring system responsible for activating the MIL ceases to detect the malfunction, and if no other malfunction has been identified that would independently activate the MIL.

Healing of the fault memory

The OBD system may erase a fault code, distance travelled and freeze frame information if the same fault is not re-registered in at least 40 engine warm-up cycles.

Freeze frame

This is a feature that can assist in the diagnosis of intermittent faults. Upon determination of the first malfunction of any component or system, 'freeze frame' engine conditions present at the time must be stored in the computer memory. Stored engine conditions must include, but are not limited to,

- calculated/derived load value;
- engine speed;
- fuel trim values (if available);
- fuel pressure (if available);
- vehicle speed (if available);

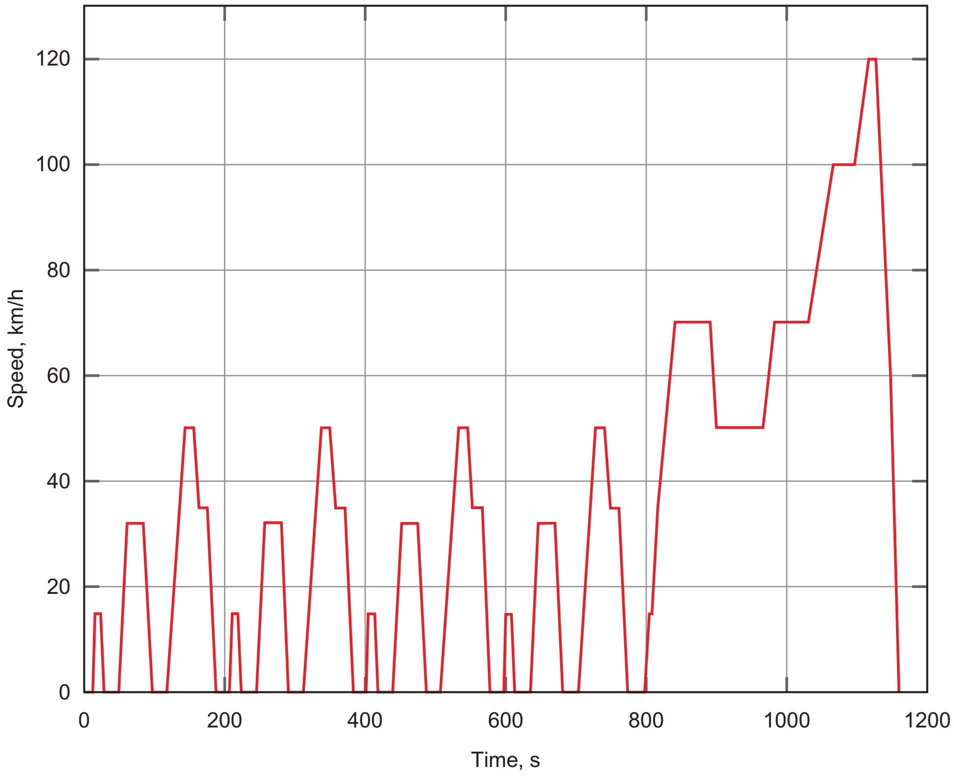


Figure 5.25 New European Driving Cycle (NEDC)

- coolant temperature;
- intake manifold pressure (if available);
- closed or open-loop operation (if available);
- the fault code which caused the data to be stored.

5.6 Driving cycles

5.6.1 Introduction

Even before a vehicle is subjected to OBD systems, it must pass stringent emissions tests. This is done by running the vehicle through test cycles and collecting the exhaust for analysis.

5.6.2 Europe

The New European Driving Cycle (NEDC) is a driving cycle consisting of four repeated ECE-15 driving cycles and an extra-urban driving cycle (EUDC). The NEDC is meant to represent the typical usage of a car in Europe, and is used, among other things, to measure emissions (Figure 5.25). It is sometimes referred to as MVEG (Motor Vehicle Emissions Group) cycle.

The *old* European ECE-15 driving cycle lies between 0 and 800 seconds and represented an urban drive cycle. The section from 800 seconds represents a sub-urban drive cycle, and is now called the New European Driving Cycle.

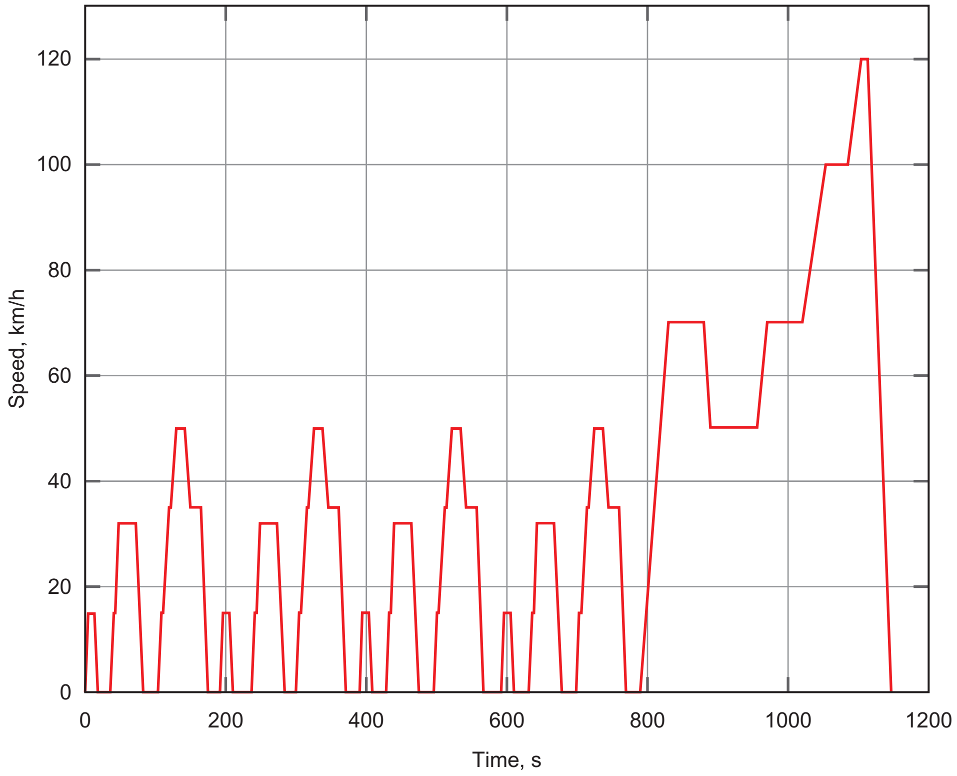


Figure 5.26 Modified New European Driving Cycle (MNEDC)

The cycle must be performed on a cold vehicle at 20 °C (68 °F). The cycles may be performed on a normal flat road, in the absence of wind. However, to improve repeatability, they are generally performed on a rolling road.

Several measurements are usually performed during the cycle. The figures made available to the general public are the following:

- urban fuel economy (first 800 seconds);
- extra-urban fuel economy (800–1200 seconds);
- overall fuel economy (complete cycle);
- CO_2 emission (complete cycle);

The following parameters are also generally measured to validate the compliance to European emission standards:

- carbon monoxide (CO);
- unburnt hydrocarbons (HC);
- nitrogen oxides (NO_x);
- particulate matter (PM).

A further tightening of the driving cycle is the Modified New European Driving Cycle (MNEDC), which is very similar to the NEDC except that there is no warm-up time at the start (Figure 5.26).

5.6.3 United States

In the United States, a cycle known as the Federal Test Procedure FTP-75 is used. This has been added to and became known as the Supplementary Federal Test Procedure (SFTP) (Figure 5.27).

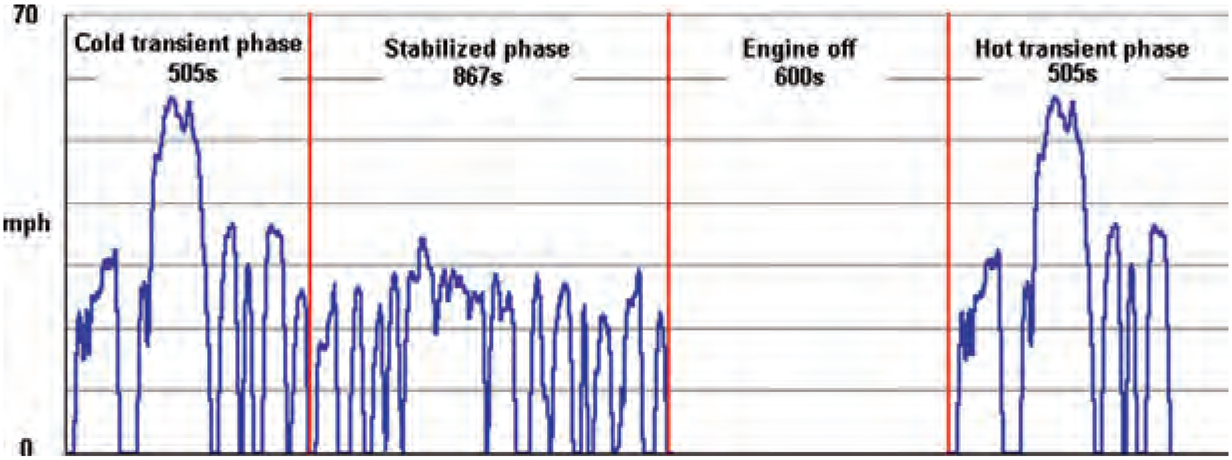


Figure 5.27 US Federal Test Procedure

5.7 Future developments in diagnostic systems

5.7.1 OBD3

Key fact

OBD3 may take OBD2 further by adding remote data transfer

The current generation of OBD is a very sophisticated and capable system for detecting emissions problems. However, it is necessary to get the driver of the vehicle to do something about the problem. With respect to this aspect, OBD2/EOBD is no improvement over OBD1 unless there is some enforcement capability. Plans for OBD3 have been under consideration for some time now. The idea being to take OBD2 a step further by adding remote data transfer.

An OBD3 equipped vehicle would be able to report emissions problems directly back to a regulatory authority. The transmitter, which will be similar to those currently used for automatic toll payments, would communicate the vehicle identification number (VIN) and any diagnostic codes that have been logged. The system could be set up to automatically report an emissions problem the instant the MIL light is on, or alternatively, the system could respond to answer a query about its current emissions performance status. It could also respond via a cellular or satellite link, reporting its position at the same time.

While somewhat ‘Big Brother’, this approach is very efficient. The need for periodic inspections could be eliminated because only those vehicles that reported problems would have to be tested. The regulatory authorities could focus their efforts on vehicles and owners who are actually causing a violation rather than just random testing. It is clear that with a system like this, much more efficient use of available regulatory enforcement resources could be implemented, with a consequential improvement in air quality.

An inevitable change that could come with OBD3 would be even closer scrutiny of vehicle emissions. The misfire detection algorithms currently required by OBD2 only look for misfires during driving conditions that occur during the prescribed driving cycles. It does not monitor misfires during other engine operating modes, like full load. More sophisticated methods of misfire detection (as discussed in Chapter 4) will become common place. These systems can feedback other information to the ECU about the combustion process, for example, the maximum cylinder pressure, detonation events or work done via an indicated mean effective pressure (IMEP) calculation. This adds another dimension to the engine control system allowing greater efficiency and more power from any given engine design by just using more sophisticated ECU control strategy.

Future OBD system will undoubtedly incorporate new developments in sensor technology. Currently, the evaluation is done via sensors monitoring emissions indirectly. Clearly an improvement would be the ability to measure exhaust gas composition directly via on-board measurement (OBM) systems. This is more in keeping with emission regulation philosophy and would overcome the inherent weakness of current OBD systems, that is, they fail to detect a number of minor faults that do not individually activate the MIL or cause excessive emissions but whose combined effect is to cause the production of excess emissions.

The main barrier is the lack of availability of suitably durable and sensitive sensors for CO, NO_x and HC. Some progress has been made with respect to this, and some vehicles are now being fitted with NO_x sensors. Currently, there does appear to be a gap between the laboratory-based sensors used in research and reliable mass produced units that could form the basis of an OBM system. The integration of combustion measurement in production vehicles produces a similar problem.

5.7.2 Diesel engines

Another development for future consideration is the further implementation of OBD for diesel engines. As diesel engine technology becomes more sophisticated, so does the requirement for OBD. In addition, emission legislation is driving more sophisticated requirements for after-treatment of exhaust gas. All of these sub-systems are to be subjected to checking via the OBD system and present their own specific challenges; for example, the monitoring of exhaust after-treatment systems (particulate filters and catalysts) in addition to more complex EGR and air management systems.

5.7.3 Rate-based monitoring

Rate-based monitoring will be more significant for future systems which allow in-use performance ratio information to be logged. It is a standardised method of measuring monitoring frequency and filters out the effect of short trips, infrequent journeys, etc. as factors which could affect the OBD logging and reactions. It is an essential part of the evaluation where driving habits or patterns are not known and it ensures that monitors run efficiently in use and detect faults in a timely and appropriate manner. It is defined as

$$\text{Minimum frequency} = \frac{N}{D}$$

where N = number of times a monitor has run and D = number of times the vehicle has been operated.

5.7.4 Model-based development

A significant factor in the development of future systems will be the implementation of the latest technologies with respect to hardware and software development. Model-based development and calibration of system will dramatically reduce the testing time by reducing the number of test iterations

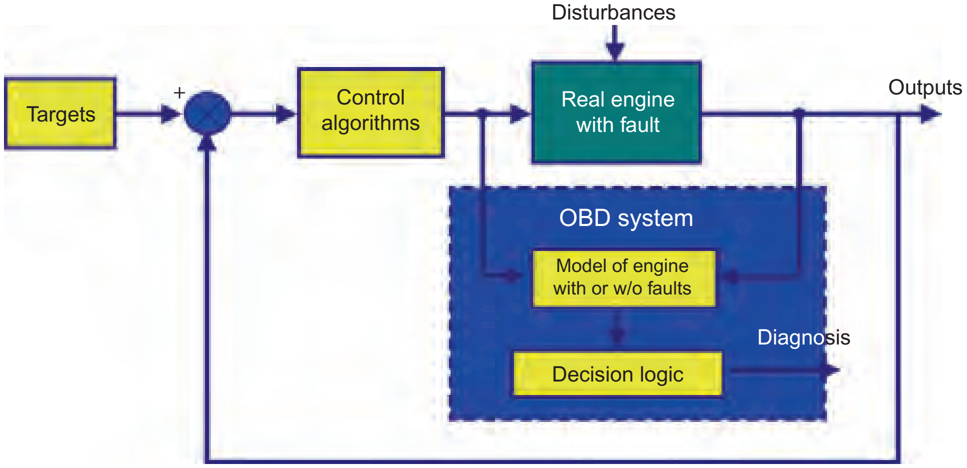


Figure 5.28 Model-based calibration of OBD system

required. This technique is quite common for developing engine-specific calibrations for ECUs during the engine development phase (Figure 5.28).

Hardware-in-loop (HIL) simulation plays a part in rapid development of any hardware. New hardware can be tested and validated under a number of simulated conditions, and its performance verified before it even goes near any prototype vehicle. The following tasks can be performed with this technology:

- full automation of testing for OBD functionality;
- testing of parameter extremes;
- testing of experimental designs;
- regression testing of new designs of software and hardware;
- automatic documentation of results.

5.8 Summary

Clearly, OBD is here to stay and will continue to be developed. It is a useful tool for the technician as well as a key driver towards cleaner vehicles. The creation of generic standards has helped those of us at the 'sharp end' of diagnostics significantly.

OBD has a number of key emission-related systems to 'monitor'. It saves faults in these systems in a standard form so that they can be accessed using a scan tool.

However, with the possibility of OBD3 using the navigation system to report where we are, speed and traffic light cameras everywhere and monitoring systems informing the authorities about the condition of our vehicles, whatever will be next?

Acknowledgement

I am most grateful to Dave Rogers (http://www.autoelex.co.uk) and Alan Malby (Ford Motor Company) for their excellent contributions to this chapter.