Mechanical Feature

Understanding Exhaust Gas Recirculation Systems

Posted 11/5/1997

By Henry Guzman

Exhaust gas recirculation (EGR) systems were introduced in the early '70s to reduce an exhaust emission that was not being cleaned by the other smog controls. Oxides of nitrogen (NOx) are formed when temperatures in the combustion chamber get too hot. At 2500 degrees Fahrenheit or hotter, the nitrogen and oxygen in the combustion chamber can chemically combine to form nitrous oxides, which, when combined with hydrocarbons (HCs) and the presence of sunlight, produces an ugly haze in our skies known commonly as smog.

How to reduce NOx NOx formation can be reduced by:

- Enriching the air fuel (A/F) mixture to reduce combustion temperatures. However, this increases HC and carbon monoxide (CO) emissions.
- Lowering the compression ratio and retarding ignition timing; but this leads to reduced performance and fuel economy.
- Recirculating some exhaust gases.

How EGR systems work The EGR valve recirculates exhaust into the intake stream. Exhaust gases have already combusted, so they do not burn again when they are recirculated. These gases displace some of the normal intake charge. This chemically slows and cools the combustion process by several hundred degrees, thus reducing NOx formation.

The design challenge The EGR system of today must precisely control the flow of recirculated exhaust. Too much flow will retard engine performance and cause a hesitation on acceleration. Too little flow will increase NOx and cause engine ping. A well-designed system will actually increase engine performance and economy. Why? As the combustion chamber temperature is reduced, engine detonation potential is also reduced. This factor enabled the software engineers to write a more aggressive timing advance curve into the spark timing program. If the EGR valve is not flowing, onboard diagnostics (OBD) systems will set a code and the power control module (PCM) will use a backup timing curve that has less advance to prevent engine ping. Less timing advance means less performance and economy. Do your customer a favor and fix those EGR codes that you may have previously deemed as unimportant.

Evolution of the EGR systems The first EGR valves appeared in 1973 on GM cars. Bolted to the intake manifold next to the carburetor, it has ports to the intake and exhaust manifolds. It has a diaphragm that pulls open a valve stem, which allows exhaust to enter the intake manifold when ported vacuum is applied to it. Ported vacuum increases with throttle opening. A thermal vacuum switch prevents vacuum from reaching the EGR during cold engine starts. This system had many problems. It would often open too soon or too much, which caused a hesitation on acceleration as massive amounts of recirculated exhaust hit the combustion chamber. Many people simply disconnected it when it began to cause problems because they did not understand its importance or design. By 1975, if you unplugged an EGR valve, you'd have a driveability complaint of engine ping. Manufacturers and technicians of that era experimented with vacuum orifice restrictors and vacuum delay valves to try to find a happy medium between clean air and performance.

Closed loop systems By 1981, closed loop computer controls were in place. EGR flow was now more carefully controlled with dual diaphragm and back-pressure EGR valves. Modulating the vacuum to the EGR valve's pull, open diaphragm controlled the flow of recirculated ex- haust. Called by various names such as amplifiers, transducers and modulators, both remote and integral vacuum modulated devices were used. The flow of vacuum was further controlled by solenoids that blocked the vacuum ports until certain criteria were met such as engine temperature, rpm and manifold absolute pressure (MAP).

As the manufacturers began to use these complex schemes with vacuum amplifiers, delay valves and solenoids, they added a lot of "spaghetti" to the engine compartment. Plastic vacuum connections would break and rot with age and were not very reliable. Vacuum diagrams were invented and became essential to the smog driveability technicians of the day. As these systems evolved, they had fewer parts and less vacuum tubing. This was achieved by the use of pulse width modulated EGR solenoids. The PCM controlled EGR flow through the use of these solenoids to modulate vacuum to the EGR valve instead of just turning it on or off periodically.

What is pulse width modulation? Let's take a moment to discuss how computers think so we can better understand this common form of PCM control. Computers are binary. The machine language they operate in consists of only two variables: on or off, true or false, high or low. That's the only way a PCM can think. As a result, computer controlled outputs are always on or off, high (system voltage) or low (ground). Therefore, a computer output is always a square wave, or an on-off step when viewed on a lab scope. The high portion of the waveform will usually be battery voltage or PCM voltage of approximately 5 volts, with a few exceptions where the PCM operates at a different voltage.

Once the PCM receives its inputs, such as rpm, throttle angle, coolant temperature and MAP, it then calculates a response based on the software program that is embedded into it. Next, it makes its decision and sends a command in the form of a pulse width modulated signal to turn the EGR solenoid on and off rapidly. The EGR solenoid has two vacuum nipples. One side gets either manifold or ported engine vacuum. The other nipple goes to the EGR valve. Its default position is to block vacuum to the EGR valve. A vent is incorporated to bleed off vacuum when the solenoid is being pulsed. Vacuum flows to the EGR in rapid on-off pulses as the solenoid is commanded by the PCM.

OBD I systems With each succeeding year, the EGR designs became more refined. The California Air Resources Board (CARB) liked GM and Chrysler's onboard diagnostic systems. In 1988, CARB required that all cars sold in California be equipped with an onboard diagnostic system and a "check engine" light to notify the driver of emission system failure. By this time, all manufacturers had to have an EGR system that was capable of alerting the driver if it was not working. OBD I diagnostics and trouble codes were added in to flag opens, shorts and sticking solenoids.

OBD II EGR systems OBD II requires that the EGR system be monitored for abnormally low or high flow rate malfunctions. The EGR is considered malfunctioning when an EGR component fails or a fault in the flow rate results in the vehicle exceeding the Federal Test Procedure (FTP) by 1.5 times. FTP is the government-mandated drive cycle smog test that all new cars must pass and adhere to.

The diagnostic executive, also called the diagnostic task manager by Chrysler, controls the EGR monitor. The executive is an OBD II software agent given the task of managing all the onboard monitors and the scan tool interface. The executive coordinates the sequencing and actuation of all the monitor's test routines. There are eight main monitors whose sole function is to directly monitor and test the components assigned to them to ensure they meet FTP standards for life. These monitors are:

- Catalyst monitor
- EGR monitor
- EVAP monitor
- Fuel system monitor
- Misfire monitor
- Oxygen monitor
- Oxygen heater monitor
- Secondary air injection monitor

A closer look at the EGR monitor Monitor tests are both intrusive and non-intrusive. An example of an intrusive test is when the EGR monitor cycles the EGR valve during a condition when it normally would be closed. In some cases, the customer may feel an intrusive test as a slight miss.

The method of testing used by the EGR monitor varies according to the manufacturer, but there are three main types.

One method includes looking for a change in manifold pressure as the EGR valve is actuated on and off.

A second method involves cycling the EGR valve and looking for a change in short-term fuel trim. When the EGR valve is opened, it displaces some of the air fuel mixture. When the EGR valve is closed, more oxygen enters the combustion chamber, which then leans the mixture somewhat. The O2 sensor will respond with a lean signal to the PCM, which in turn increases pulse width. This is called short-term fuel trim compensation. The EGR monitor looks to see that all these things are occurring as they should. It repeats the tests and averages the results. Before the EGR monitor can begin its testing, it must first receive clearance from the diagnostic executive. The executive ensures that there are no conflicting conditions that would invalidate the EGR monitor's tests. For example, if the car had a lazy O2 sensor, fuel trim compensation to the EGR opening and closing would be inaccurate. Therefore, there are many safeguards built into OBD II to prevent this type of occurrence from happening. OBD II also has rationality checks. In other words, it uses deductive logic and constantly compares its inputs against each other to make sure all are in sync with one another. After the EGR monitor gets the OK to run its tests, it uses strict enabling criteria to ensure accurate testing such as:

- Engine temperature more than 170 F.
- Ambient air temperature more than 20 F.
- Engine run time more than three minutes since 170 F.
- Engine speed 2248-2688 (auto. trans.), 1952-2400 (manual trans.).
- Manifold absolute pressure from 5-20 hg.
- Short Term Adaptive Fuel Trim is adjusting pulse width by less than +7 percent and more than -8 percent.
- TP sensor from 0.6 to 1.8 volts.
- Vehicle speed sensor more than 40 mph.

The above is used for illustrative purposes only. Refer to your manual or CD-ROM information system for specifics to the car you are working on.

The third type of EGR monitoring design includes monitoring an EGR position sensor and a back-pressure sensor. Some Fords use a differential pressure feedback sensor that reads exhaust back-pressure upstream and downstream of the EGR valve to determine its flow rate and operation.

While OBD I systems would usually flag an inoperative EGR system, OBD II systems are given the task of determining the correct amount of EGR flow to keep the car running clean.

Next month, we will get into diagnosis, testing and repair techniques for all the different types of EGR systems. I will also cover pattern failures of all types, including mechanical problems such as plugged EGR passages that can cause rpm specific misfire concerns.

Henry Guzman is an ASE master tech with L1 certification. He has 20 years of experience working as a technician on foreign and domestic cars.

Mechanical Feature

EGR Systems: Operation and Diagnosis

Posted 12/18/1997

By Henry Guzman

Last month, we covered the reasons why exhaust gas recirculation (EGR) systems were created, which is primarily to control and reduce nitrogen oxide (NOx) formation, and secondarily to improve engine performance with ping control. We covered the evolution of EGR systems and ended with a look at how onboard diagnostic (OBD-II) EGR systems operate.

This month, we will detail the operation and testing of EGR systems. It's helpful to remember that if a hard trouble code has been set, the power control module (PCM) has determined that a gross emission failure has occurred. Gross failure is defined as excessive exhaust emissions more than 1 1/2 times what the car was designed to produce, according to the Federal Test Procedure (FTP). This can happen if too little or too much EGR is flowing. In the case of an EGR code, that would be 1 1/2 times the NOx that is allowable. Even if the car appears to run well, excess smog is being produced. An EGR failure will not only cause excess NOx, back up strategies will often entail enriching the mixture or retarding the timing. This in turn increases HC and CO emissions. Even if you don't see it on a 4-gas exhaust analysis test, the emissions are being created and the catalytic converter will not always cover them up.

EGR valves and components

To diagnose EGR systems properly, it's important to understand how they work and what kind of communication they have with PCMs. Knowing this will help you understand a flow chart and wiring diagram, and to come up with a test strategy that you are able to complete with your available tools and equipment.

Let's start with that golden oldie, the single diaphragm EGR valve. It consists of a spring-loaded diaphragm that is connected to a pintle and seat by a slender steel shaft. Normally closed by spring tension, as it receives ported vacuum, the diaphragm rises, which pulls the pintle off its seat and enables exhaust to flow into the valve's chamber and then on to the intake manifold. To test this component, use a hand-held vacuum pump connected to the vacuum nipple to raise and hold the diaphragm. About eight inches of vacuum should do the trick. The valve should hold vacuum and raise the pintle in a linear fashion. When the engine is idling, pumping it up should stall the engine. This type of valve may or may not have vacuum modulation. Remember, vacuum modulation to the EGR is a vital ingredient of good driveability and precise NOx control. This type of EGR valve is used with a thermal vacuum switch and maybe an inline vacuum delay valve.

The positive back pressure EGR valve can be identified by the letter "P" stamped next to the part number and date code. A back pressure valve is easy to spot because its pintle shaft is much thicker than the single diaphragm type. This is so because the shaft is hollow. The hollow design allows exhaust gases to flow into the shaft and push up on it. When positive back pressure in the exhaust system is sufficient, the shaft raises up and seals the built-in control valve. Once the control valve is closed, it allows applied vacuum to pull up on the diaphragm. Without back pressure to lift the hollow shaft and close the control valve opening, the EGR valve will not hold vacuum. It is bled off to the atmosphere. This design thus modulates EGR flow by modulating the applied vacuum. As engine load increases, so does engine back pressure, which causes the control valve inside the EGR to trap vacuum and open up. To test this valve, bring the engine up to 2,000 rpms to create back pressure, then apply vacuum. EGR should open and cause a 100 rpm drop or more. Exhaust leaks or a modified exhaust system can create havoc here. Adding dual exhaust or headers on a car designed for a single exhaust will reduce back pressure and set a Code 32 on GM cars. Positive back pressure EGR valves are used in simple vacuum controlled systems, as well as more complex pulse width modulated applications.

EGR solenoids are used with all types of EGR valves, especially back pressure type valves. The EGR solenoid will have two or more vacuum lines and an electrical connector. The solenoid also has an air bleed and sometimes an air filter. Vacuum is bled off through the filter vent. The PCM uses the solenoid to regulate vacuum to the EGR valve. The vacuum can be manifold or ported vacuum. The solenoid is a vacuum switch with inlet and outlet vacuum ports. The PCM calculates intended EGR flow from various other inputs and then sends a pulsed "on/off" signal to the solenoid. No vacuum flows until commanded by the PCM. This signal turns the vacuum on and off in rapid succession. This is called "pulse width modulation." If the filter becomes clogged, the vacuum cannot bleed off and too strong a signal will be sent to the valve. If that happens, the EGR valve will open too much and cause a driveability problem.

Remote vacuum transducers
All manufacturers use them, but they are very popular with Toyotas and other Asian cars. Shaped like a flying saucer with three or more vacuum ports, they modulate vacuum by using manifold and ported vacuum against each other along with an exhaust back pressure input. The result is a carefully controlled vacuum signal to the EGR valve that is mechanically modulated by engine load. Your best bet with these is to study the vacuum diagram on the underhood emissions label. Make sure the vacuum hoses are in good condition and properly routed. Many of these units have air filters also. You can clean them out to prevent too much EGR flow.

The negative back pressure EGR valve is identified by the letter "N" and looks similar to the positive back pressure EGR valve. The valve is opened by a combination of applied engine vacuum to the control valve and negative exhaust system pulses that happen as each exhaust valve closes. As soon as the pintle opens, back pressure is reduced slightly, which opens a control valve vacuum bleed and then the valve quickly closes. In this manner, EGR is modulated by negative exhaust system pulses. To test it, apply vacuum with a hand pump when the engine is off. The valve should open and hold vacuum.

Integrated electronic/mechanical EGR valves
This type of valve has different names with each manufacturer. It is easily identified because it has a single vacuum source inlet and a three-wire electrical connector. Mechanically, it operates like a single diaphragm EGR valve with a twist. It has a pintle position sensor riding atop the EGR diaphragm. This tells the PCM the amount of EGR valve opening as it is actuated. The PCM then commands a pulse width modulated solenoid to apply an appropriate amount of vacuum on-time. GM makes one of these units that has the integral pintle sensor and an integral solenoid with air filter. The only separately serviceable part is the air filter. Ford, Honda and Mazda all use a variation of this design with remotely mounted solenoids. The idea behind the pintle sensor is to give the PCM precise feedback as to exactly where the EGR valve is positioned. The PCM can then modulate the vacuum signal to it accordingly. The pintle position sensor is a potentiometer. Like a throttle position sensor, it is a variable resistor. The wiper arm within the sensor

can wear and develop opens in the sensor return signal. A sweep test with a digital volt/ohm meter (DVOM) or scope can be used to test the sensor. The PCM has an internal "map" of where the pintle sensor should be at any given time. If the sensor's voltage reading is too high or low, a trouble code will be set. With Fords, it is the infamous Code 31. This code could be caused by several different factors.

- If the pintle position sensor (Ford calls it the EVP sensor) is shorted or open, you could have a code set.
- If the EGR valve becomes carboned up and does not seat fully, the EVP sensor gives a high reading and a code is set.
- If the diaphragm of the EGR valve is bad, then it, too, is flagged.

The fix

On Fords and Mazdas, the only sure fix is to replace the sensor and valve as an assembly. Sometimes you can get temporary relief by filing down the pintle sensor stem to lower the sensor return voltage to specs. Or you could add a thicker gasket between the valve and sensor. You can spend a lot of time trying to capture the intermittent failure in the act. This is not recommended. The codes are rarely false. Note that there are two interchangeable sensors; one is gray and the other is black. Key-on-engine-off (KOEO) voltage for the gray sensor is 0.40 volt and for the black sensor is about 0.83 volt. Don't mix them up, or Code 31 won't go away.

On Hondas, the fix is sometimes achieved by cleaning out moisture and crud from the EGR vacuum lines with shop air pressure. Disassembly of the vacuum air box is required for access.

Digital EGR valves are unique in several ways. Only GM uses them. They are completely electronic, controlled solely by the PCM. They come in two or three solenoid models, depending on the application. Part of the valve is open to exhaust flow at all times. When the solenoid pulls the pintle open, exhaust leaves the EGR valve chamber and directly enters the intake manifold. This method is different from all other EGR valves. All other EGR valves open to allow exhaust to enter their chamber first, then circulate through the valve on to the intake passage. The benefit of the digital EGR valve is speed and accuracy. It meters EGR flow 10 times faster than a vacuum modulated system. The valve is actuated by an individual quad driver from the PCM for each solenoid the valve has. Battery power is fed through terminal D when the key is turned on through a 15-amp ignition fuse. When the PCM grounds a solenoid, a magnetic field is created that causes the armature to lift open the pintle. The PCM uses this system to actuate each solenoid in increments. The increments are displayed on a scan tool as percentages of total flow. With a bi-directional scan tool, the digital EGR valve can be commanded open in a variety of increments. Don't despair if you do not have a bi-directional scan tool. You can still work with this! It's easy. Simply unplug the four-wire connector. Run a fused 12-volt wire to terminal D and alternately touch each of the other terminals to ground with a test probe. This will cause each solenoid to pull open. You can do this test with the engine idling and check for an rpm drop as you ground each solenoid. If you don't get a good rpm drop on this or any other EGR valve, you may have plugged or restricted EGR passages, which can cause a code to be set.

The linear EGR valve is a high-tech system. It uses a closed loop method for the utmost in EGR control and driveability. All electronic, its built-in pintle-position sensor allows the PCM to continuously monitor "actual pintle position" and adjust it to the "desired pintle position" as a percentage. A generic scan tool will display these parameters just as it does "actual rpm" and "desired rpm." This feature is a boon to troubleshooting. For example, I had a '92 Chevy half-ton drive into our shop running terribly. The Malfunction Indicator Light (MIL) was on steady. My aftermarket scan tool pulled up a Code 32. A quick look at the data stream showed that actual and desired pintle position did not match - at idle, "desired pintle position" was zero; "actual" was about 40 percent. This told me the PCM was trying to close the EGR valve but could not. Twenty minutes later, I had the EGR valve out of the vehicle and found a large chunk of carbon stuck between the pintle and its seat. Spring tension was holding it tightly and, of course, the PCM was squeezing it as it tried to close the valve. I pried the carbon out, reinstalled the valve, and all was well.

Repeat failure
Repeat failure is a common problem. You can recommend a top engine clean to the customer and attempt to clean loose carbon from the upper intake. GM has noted the problem and come up with a software update for the PCM. Essentially, what the update does is periodically command the EGR to 100 percent opening to prevent or flush out carbon chunks. The new prom numbers were in a "special policy" procedure bulletin and not a regular TSB. The special policy number is 96067(A). It refers to '92-'94 S/T, M/L and C trucks with the 4.3 V6 engine and linear EGR valve. The bulletin number is 67-65-38. It refers to '95 C, S/T and M/L trucks with the 4.3 V6 engine and linear EGR valve.

EGR/PCM strategy
GM's Code 32 has been around a long time and can be caused by a variety of reasons. Every three to five years, the PCM strategy on this code changes. Make sure you review the proper flow chart any time you work on one on these. The strategy is slightly different depending on which engine, transmission, body type or year the car is. The most common strategy entails the PCM looking for fuel integrator counts to decrease momentarily when the EGR is commanded open. Why? There is no oxygen in the inert EGR gas, so the integrator subtracts fuel to compensate. For this to happen you must have a good working oxygen (O2) sensor. O2 sensor checks are usually not in the Code 32 flow chart, so be aware. Newer models and other makes look for a change in manifold absolute pressure (MAP) when EGR is flowing, which is a more reliable method.

Mechanical failures
In closing, I'll leave you with a pattern failure seen in '85-'92 Nissans, models 240 SX and Stanza, and some Hondas about the same vintage. They have individual EGR passages running to each cylinder. Eventually, some of these passages (one or two) will plug up with carbon while others stay open. The remaining passages receive 4-runner EGR volume, which is far too much and causes a misfire on the affected cylinders. It happens above idle and under load. Sometimes you can unscrew Allen head access plugs to clean out the passages, other times you must remove the upper intake plenum chamber to do the repair.

Henry Guzman is an ASE master tech with L1 certification. He has 20 years of experience working as a technician on foreign and domestic cars.
