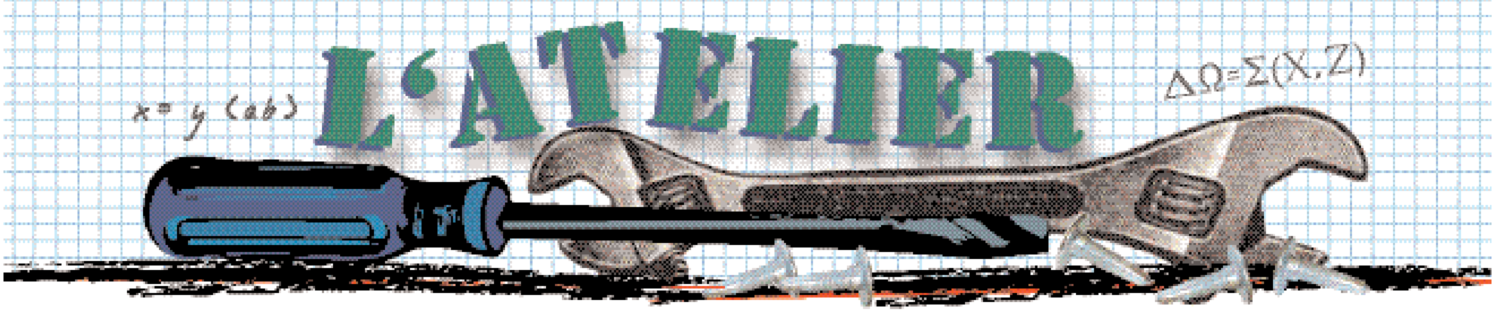


IGNITION SYSTEMS

by Kevin Cameron

The electronics revolution has changed our lives and one of the greatest examples of this reality lay in the dramatic change we have witnessed in engine performance, cleanliness and economy in the past 30 years. The rudimentary breaker points that used to adorn our engines in the early 70's are but a distant memory and in their place we have the latest in acronym-laden electronic brains to baby sit every operating parameter of our vehicle powerplants. This month, Kevin Cameron walks us through the evolution of engine ignition systems and explains how development in this key area has transformed the modern engine.

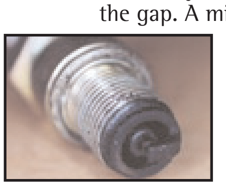
Early Days

From a beginning in the 1860s internal combustion engine ignitions evolved rapidly in the direction of higher speed, with ignition systems passing from flame port to hot-tube (a kind of glow-plug, kept hot by an external burner), and then to electric spark ignition, which had no "speed limit".

High voltage is required to jump a spark plug gap – more than is available from a standard vehicle 12V battery system. In all ignition systems some kind of pulse transformer coil is necessary to step up this or some other voltage to the necessary 15,000 – 30,000V. Such pulse transformers consist of two wire coils wound on an iron core: a primary with only a few turns of heavy wire, and a secondary with hundreds or thousands of turns of fine wire. As 12V current from a battery, or higher voltage from some other source is sent through the primary coil, a magnetic field is generated, intensified by the presence of the iron core. When primary current is suddenly cut off by some device (mechanical contact-breaker, transistor switch, &tc.) this magnetic field collapses. As it does so, its energy induces a voltage in each turn of the secondary coil (which is connected to the spark plug). This voltage, summed over the secondary's many turns, is very high – roughly equal to the primary voltage multiplied times the ratio of primary to secondary wire turns. This voltage is high enough to ionize the plug gap, filling it with ions that cause it to conduct. The rush of electrons across the gap is a spark.

Spark plug fouling and the electronic remedy

With so many wire turns in ignition coil secondaries (a typical turns ratio might be 200, that is to say it has 200 times more secondary windings than primary), a problem arose in certain applications – especially for two-stroke engines. As the magnetic field collapses, the high inductance of the coil's many turns opposes the rise of voltage, slowing the rise of secondary voltage. If the spark plug's insulator has a layer of carbon on it – from, for example, two hours of part-throttle trolling in an outboard motor – the rising spark voltage may leak away across the carbon layer so fast that it cannot rise high enough to jump the gap. A misfire is the result.



Fouled spark plug (left) versus normal or optimal (right)

The answer was to speed up the coil's rate of voltage rise, but doing so in a typical contact-breaker system was difficult. The obvious solution was to increase primary voltage so that coil turns ratio could be reduced – but mechanical switches handle high voltage poorly. Engineers remedied this at first by using the breaker points to control a transistor switch, and later substituted a non-contact trigger to operate the transistor. By sending 400 volts instead of 12 into the coil primary, its turns ratio could be reduced by the ratio of 12 to 400 (or a factor of 33) – enough to cut coil inductance and speed up rise time. If this power was supplied to the coil primary from a capacitor, such systems were called capacitor-discharge ignitions, or CDIs. High voltage for capacitor charging could be generated in either of two ways:

(1) In a magneto-type CDI, high voltage pulses were generated as needed by a rotating magnet generator,

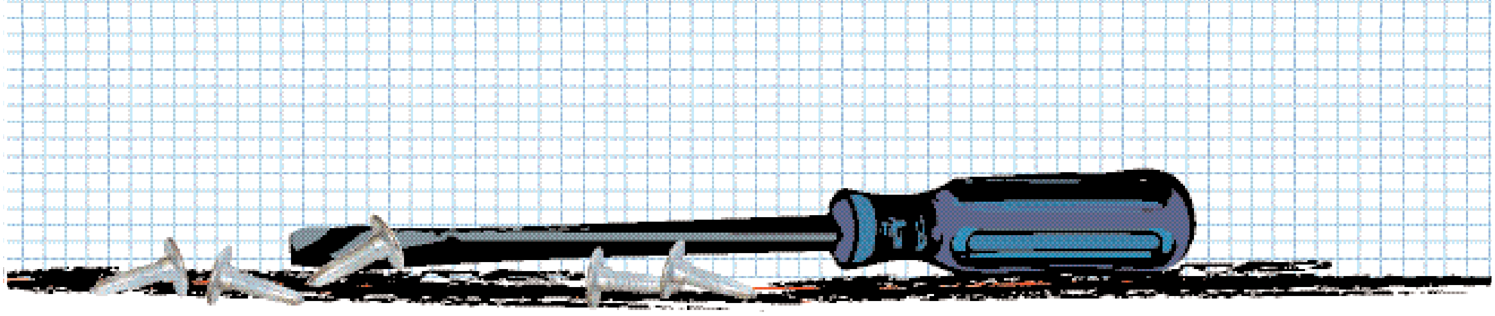
(2) In a battery CDI, battery voltage was chopped into pulsed AC, then stepped up to high voltage through a transformer, then rectified back to DC and charged into the capacitor(s). On some systems, the hum of this inverter could be heard as the ignition was switched on.

Spark Timing

Soon it was discovered that an engine could not start and run efficiently on the same ignition timing. Retarding the spark to TDC prevented kick-back (that is, the engine attempting to spin backwards) during starting, but operation at normal engine speed required an earlier timing - to give peak combustion pressure at the correct piston position – about 11-degrees ATDC. For this reason, early autos and motorcycles had manual spark timing levers – retard for starting, advance for running. When electric starting arrived, this starting retard was automated by flyweights and springs that advanced the breaker cam once the engine started.

Ignition Timing – Two-Stroke vs. Four-Stroke

A four-stroke engine runs well on essentially fixed spark timing across a wide rpm range. This is so because (1) as an engine speeds up, its combustion flame speed increases



in proportion, driven by increasing charge turbulence, and (2) a four-stroke engine, not being dependent upon exhaust-pipe pumping, has near-constant cylinder filling, so there is little flame speed variation from this cause. For these reasons, a four-stroke ignition curve may give 0 degrees at starting, 15 or so degrees at idle, then jumps up to somewhere between 30 and 40 degrees in the operating range. Honda research, performed in 1964, established that this near-constant ignition timing requirement remains true up to at least 27,000-rpm.

Contrast the two-stroke engine, whose breathing varies with the strong pumping action of pressure waves in its exhaust system. Flame speed varies with cylinder filling, just as the speed of a forest fire increases as trees are more closely-spaced. Flame speed, therefore, increases as exhaust action stuffs more charge into the cylinder. Ideally, a two-stroke should have more ignition advance at low revs, when it is not “on the pipe”, and progressively less advance as the pipe stuffs more and more charge into the cylinder.

Many people, raised on four-strokes, are confused when they see a two-stroke timing curve. As an example, Yamaha’s 1980 TZ500 race engine had 36 degrees of ignition timing at 6000-rpm, but only 16-degrees at 12,000. Although this looks “backward” to four-stroke people, it is only what is necessary to keep peak pressure occurring at the same 11-degrees ATDC that peak torque requires.

Effect of Bore Size

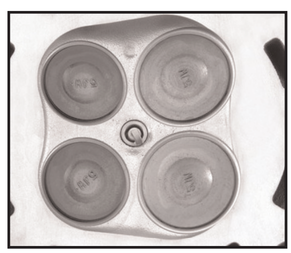
In very large-bore engines such as the latest near-1000-cc snowmobile twins or recent BMW twins, the time required for combustion with a single spark plug may exceed the time detonation takes to develop. To avoid such detonation, such engines may be given one or more extra plugs per cylinder (English 4-stroke designer Al Melling uses 3 plugs per 4V cylinder), fired simultaneously. Much nonsense has been uttered about “colliding shock waves” from multiple ignition. The flame front of normal combustion is no more a shock wave than is the leading edge of a grass fire, and when two such flame fronts come together nothing exciting happens. Shock waves occur when the last remnants of unburned charge, heated by being compressed by expanding combustion gas, ignite spontaneously, after the normal spark has occurred. This abused and chemically-altered end-gas burns at sonic speed, generating a shock wave that can blast loose heat-softened metal from piston or head. This is detonation.

Note also that a hole through the center of a piston indicates pre-ignition. The hole is through the center as this is the part of the piston farthest from the cool cylinder wall, so it heats up fastest, softens, and punches through. Detonation damage occurs on the edges of the piston, typically (in two-strokes) on the hotter exhaust side. In extreme cases, detonation erodes enough metal to expose the piston rings.

Combustion Chamber Differences

Why should two-strokes require less peak-torque ignition timing than four-strokes? A two-stroke has a nearly ideal fast-burn combustion chamber, for it surrounds a central hemispherical pocket chamber with a ring of turbulence-generating squish – the region in which the outer edge of the piston comes very close to the head at TDC. The spark plug is ideally located – in the exact center, minimizing flame travel in all directions. The open central pocket allows charge motion to persist through TDC.

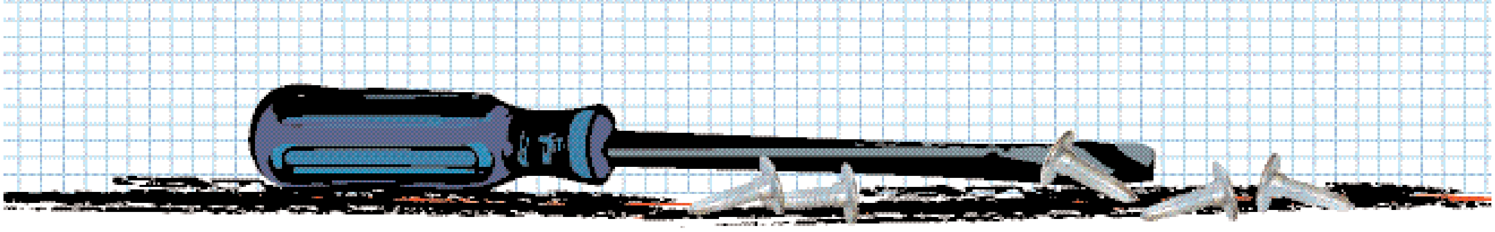
A four-stroke, by comparison, needs most of its cylinder head area for its intake and exhaust valves, leaving very little squish area to generate flame-accelerating charge turbulence. The presence of the valves, with their edges and shape, acts as something of a damper on charge motion, slowing it (and with it, combustion) as the piston approaches TDC. Modern four-valve designs permit a central spark plug location, but older four-strokes with two valves must locate the plug to one side – or use twin plugs.



Bore/Stroke Ratio Effects

In two-strokes, bore and stroke are nearly equal – a compromise between a stroke short enough for reasonable rpm capability, and a stroke long enough to accommodate ports of adequate area (Do the math – you’ll be surprised. A two-stroke’s port area increases with stroke – the opposite of a four-stroke.). This keeps the bore reasonably small, shortening flame travel distance to a minimum.

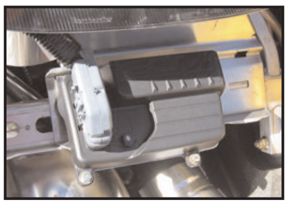
In four-strokes, the need for power-boosting high revs calls for a short stroke, and the need for generous valve area calls for a big bore. Big bore not only increases flame travel distance, but when combined with a desirable high compression ratio it results in a vertically very tight combustion chamber. Because this leaves little room for vigorous charge turbulence, flame speed is further slowed. Production autos have about equal bores and strokes, mainly for emissions reasons. A Formula One engine may have a bore that is 2.25 times its stroke, requiring inefficiently early ignition timings of 60-70 degrees BTDC (and exotic fuels to accelerate combustion). The motorcycle-derived four-strokes now in snowmobiles are a compromise between these



extremes, with bore about 1.5 times stroke. Their ignition timings at peak torque are in the degree range of high 30s to low 40s.

Timing Control

Early ignition systems either provided fixed timing or allowed a mechanical retard for starting only. In their powerbands, engines ran on fixed timing. This was a compromise hardship in two-strokes, which therefore had much too little advance when running “below the pipe”, and too much when “on the pipe”. A clever tuner could obtain an advantage by advancing the timing on an acceleration course, or retarding it on a stop-speed course. In the mid-1970s electronic advance-retard curves began to be used, but as their operation depended on inductance differences in multiple coils, their range of timing change was limited.



MPEM electronic module found on various Ski-Doo sleds

Once the coming of emissions laws required auto makers to more closely control fuel mixture and timing, small digital computers were developed to provide this control. For the first time, this meant that the correct ignition timing for every condition of rpm and load could be provided. As the engine ran, the computer would check the rpm and throttle angle, then look up in a stored data table the correct ignition timing for that condition. On the crank was a trigger wheel which provided a reference timing – let us say at 45 degrees BTDC. The computer would note this position, then count off from that to the desired ignition point, at which it would fire the spark plug. Timing would be re-computed for each firing.

Ignition Maps

The stored timing information took the form of a “map”, for that is just what it looked like if graphed out in three dimensions – a map of hilly terrain. Length and width would be the two variables rpm and throttle angle, while height would be the corresponding timing. These maps were determined by running production engines on the dyno and finding the correct values for scores or hundreds of points. The computer would then interpolate between points.

Other variables affecting timing could be added as well – for example, atmospheric density. Additional functions could be added at will, such as a rev limiter, a schedule of rpm limits rising slowly through break-in, or an automatic spark retard that occurs when detonation is detected by suitable sensors. In racing ignitions, multiple maps could be provided.

Who Controls the Computer?

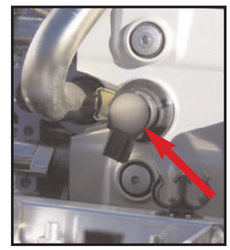
The next step was to make map data addressable by an external computer or programming box. This became necessary for racing use, or for customers who found they could not change timing appropriately when they modified their engines. Over time it became routine to plug in the laptop, pull down the ignition map, and make necessary changes as desired. Subsequently, ignition and fuel systems have been developed which “learn” to provide what a modified engine needs. A fuel system does this by reference to an oxygen sensor in the exhaust, and an ignition may use a detonation sensor in a similar way, “experimenting” through successive approximations to arrive at a correct fuel or ignition map.

Giving Cylinders What They Need

In NASCAR auto racing it was long known that each cylinder receives a slightly different fuel-air mixture, and thus should have individualized ignition curves, or even its own unique set of cam timings. Success with this work has led to engineers beginning to consider all multi-cylinder engines as a cluster of individual single-cylinder engines, each with its own unique requirements. Meeting those requirements has resulted in small but useful gains in power, economy, and reduced emissions.

Naturally, the more functions are required, the faster the computer must run to perform the necessary computations, so 8-bit machines have given way to 16 and 32-bit processors, in an evolution much like that of desktop machines. Quite ordinary engine controllers are now much faster and have more memory than on-board computers developed for the Space Shuttle.

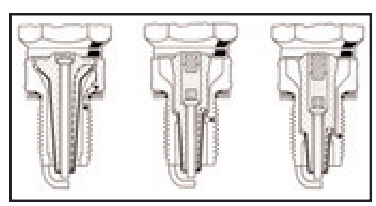
Low Tension Ignition Returns



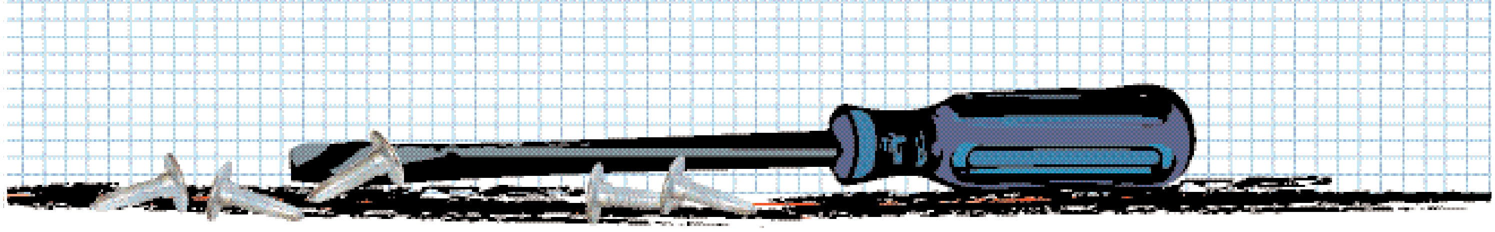
Separate ignition maps for each cylinder required that each cylinder have its own coil. Individual ignition coils are now being integrated into spark plug caps, now called “stick coils”. This just reproduces (50 years later!) the low tension systems used on large aircraft piston engines in the 1950s, in which the coils were mounted very close to the plugs, eliminating the corona and other high voltage problems of the previous era.

Spark Plug Heat Range

Early spark plugs employed insulators made by stacking countless discs of mica onto the central conductor. Such plugs were regarded as desirably easy to “read” in the days when engine tuners determined fuel mixture by plug condition. The onset of WW II required millions of more durable spark plugs, and since that time insulators have been made of alumina porcelain. To prevent fouling, the tip of this insulator must operate between 400 and 750C. Above the higher temperature the plug electrodes may act as a source of ignition before the spark. Because engines differ in their combustion severity and in their duty cycle, spark plugs are manufactured with various degrees of exposure of the central insulator. The longer the exposed insulator, the hotter it will operate under given conditions. Thus, an engine with moderate combustion, or one used at moderate duty cycle, will require a plug with a longer insulator to keep its tip in the above temperature range. If we now soup up that engine (higher compression, porting, improved exhaust pipe, &c) its combustion will become more intense, possibly pushing plug insulator temperature too high. We must now replace the “hot-running” plug with a cooler-running type. So it is that each plug manufacturer produces a booklet, available upon request, listing the correct plug and heat range for each application.



«Hot» plug (left) with substantial exposed insulator. At the extreme right a «cold» plug with much less exposed insulator



Plug Gap Styles

Many plug gap styles are also made. The commonest of these has its insulator end approximately flush with the end of the steel shell, with the ground electrode welded to the shell and bent over to form the plug gap to the center wire projecting from the end of the insulator. When it is desired that the spark be placed deeper into the chamber, a projected-tip gap may be used, in which the insulator projects beyond the end of the steel shell a few mm. Plugs of this type are noted for good throttle response. If there is a clearance problem in the combustion chamber, the antique R-gap, or retracted-gap design may be used. In this type, the end of the center wire is flush with or below the end of the shell, and the side wire is pressed through a hole drilled in the side of the shell. For special applications, a surface-gap plug type has also been made. In these, the shell, insulator, and center wire all form a common surface, and the gap is radial in all directions.

Platinum electrode

Iridium electrode

Because an electric field is most intense at sharp edges or small radii, all kinds of patent plug electrodes have been designed to make it easier and surer for the spark to jump the gap. Chief among these is the fine-wire plug. To prevent rapid erosion of the very hot center wire, it is tipped with one of the high-melting-point metals such as platinum or iridium. Such plugs are not cheap but they perform and last well.

In the days of leaded gasoline gap erosion was much faster than at present, when spark plugs easily go 40,000-km or more.

Exotic Plugs

Many claims are made for star-shaped wire, ground electrodes with artistic holes in them, or multiple gaps. Multiple-gap plugs come to us from the aircraft world, where changing plugs on a B-36 bomber meant 336 tedious removals. By spreading gap erosion (aircraft fuel contained 1.6-gm of lead per liter) over multiple gaps, plugs were made to last longer. For the purposes of snowmobile users, the sharp throttle response of wire-gap single-electrode plugs is more attractive than 2500-hour life (that would be three hours on the sled, every day through seven four-month winters. Who even keeps a sled that long?).

Magneto systems fired gaps of .35-mm to .60-mm, and the latter gap was long considered normal for most engines. When emissions laws forced auto makers to run leaner mixtures, larger gaps and higher spark energy were necessary to push sparks through a larger sample of charge, making it more likely that some part of the big spark would pass through an ignitable region of charge. Gaps therefore increased to 1.5-mm.

Although the 14-mm X 1.25mm plug thread remains common, the lack of space in small four-valve combustion chambers has led to production of 12-mm and 10-mm plugs. Honda's RC-166 6-cylinder 250 racer of 1966 had 8-mm plugs! Plug threads are made in various lengths – this is called "reach". The end of the plug shell should come flush with the inside surface of the combustion chamber.

Flamethrowers?

At one time it was thought that future lean-burn engines would require very high-energy ignitions. All sorts of proposals were made for devices to produce fireballs. One such product even made its way to market, only to find that its igniters had a life of only 500 miles. The right way to do this was soon discovered – to stratify the charge and place a conventional spark plug in a zone rich enough to be ignited conventionally.

What about super-duper ignition systems that claim increased power? As Champion Spark Plug engineer the late Bobby Strahlman once said, "If a new ignition system gives you increased power, then your previous system was defective". Adding spark energy beyond what is necessary to produce reliable ignition is a waste of time. The energy of combustion is fixed by chemistry – there is no way to "ignite the charge harder" and get more power than chemistry already put into the fuel-air charge.

Long- vs. Short-Duration Sparks

It has been noticed that magnetos sometime succeed where ultra-modern "flamethrower" CDIs have failed. Why? A magneto spark lasts a long time, so it does a good job of igniting messy mixtures containing fuel droplets and/or rich/lean zones. This is why the first two-stroke CDI systems (circa 1969) worked poorly. A long-duration spark allows a lot of mixture to blow through the plug gap – perhaps bringing a kernel of ideal mixture that will light up. Simple electronic spark systems may produce very short-duration sparks – and if such a spark occurs just as a big blob of fuel or a lean zone is in the gap, you'll have a misfire. Knowing this, designers of electronic systems can extend spark duration to correct this problem.

Multiple Sparks

Another way to accomplish the same thing is to fire each plug many times in a so-called MSD, or Multiple-Spark Discharge system. In these systems the secondary circuit may be designed to oscillate, producing several sparks. These have been popular in taxi engines, which must spent long periods idling, and are run until they are worn out, allowing considerable oil to reach their combustion chambers.

Multiple spark systems have also been useful for the stratified-charge mode in engines such as Evinrude's E-Tec line. In this case the added sparks increase the likelihood that one of them will find a kernel of ideal fuel-air mixture in the confusion of rich and lean zones – and ignite it.

It Never Ends

In future we can expect engine control systems to become self-teaching, able to make and update their own fuel and ignition maps, compensate for normal engine wear, and accommodate themselves to owner modifications. Three years ago I asked a Ducati engineer about such systems. He replied, "Yes, is possible, also now OK for car racing. But this is same size as small luggage, not so good for bike. But smaller one day soon."

In the effort to combine improved performance with conformity with ever-tighter emissions and noise standards, engines are being given a "nervous system" of sensors and actuators, controlled by faster and faster computers. It is impossible to predict how far this process will go, or what it may yet accomplish.