The incremental encoder – operation principals & fundamental signal evaluation possibilities (part 1)

White Paper

By Dipl.-Ing Kamil Pogorzelski and Dr.-Ing Franz Hillenbrand

The use of incremental encoders has become indispensable in virtually all areas of industry. Its uses can be found for making simple displacement measurements on assembly lines, tool and die factories, robots, etc. Other areas include RPM measurements of rotating machinery, location and position measurements, and velocity. Found both in the sensors (incremental), as well as in the measurement modules (encoder interfaces), the signals of the incremental encoder are evaluated for substantial differences. These differences are often deciding factors on the suitability of a measurement system for the desired application. Therefore, the structure and operation of these sensors have become indispensable and the manner in which the signals in the encoder interfaces are processed and evaluated by the company, imc Meßsysteme GmbH, are discussed in this article.

The following principle that is described can be applied to both rotary and translational movements. In the case of rotational motion, an evenly divided rotating disc will display an evenly divided linear scale. In addition, other elements, such as gears, often divide the circumference equally, thus, also work as an encoder. One can denote, that if tracks are divided into a certain grid pattern, then one can obtain, by applying additional tracks with different pitch information, the absolute angle or travel position. The following discussion is restricted to the consideration of single-track encoder, which is determined by summing the individual increments the position. Since the principle for rotary and translational encoders is equal, the following descriptions will only be a discussion on the measurement of angles.

Incremental encoder:

The segments principally serve to generate a signal from the progressive motion of the angular increments. These pulses are counted by the encoder interface and from this, determine the total angle traversed. The angle determination is relative, since the incremental encoder does not tell you exactly where it is located at the beginning of the measurement. The implementation of this principle always comprises a sensor, which converts the passing segments into a voltage signal. This voltage signal can be generated in analog or digital form. As a frequently used method in the automotive industry, this type of incremental encoding works with the induction of gear teeth as the encoder element. By counting the number of teeth on a gear, the changing inductive resistance of a signal is generated by the corresponding number of segments that pass by the sensor. With the appropriate shape of the teeth on the gear, a sin-like voltage waveform can be generated. The big advantage of this type of incremental encoder is that it works without moving contacts and is therefore used in places with strong vibrations and harsher environmental conditions, such as in vehicles.

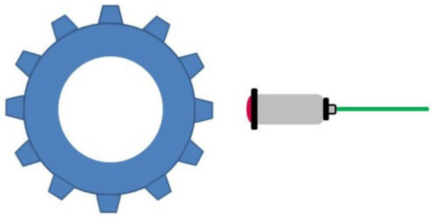


Figure 1: *Incremental encoder according to the inductive principle.*

Another principle is the opto-electronic principle. In the opto-electronic principle, a voltage signal is created either by shining light from an LED through slits on the surface of a disc (transmitted light principle) that are then received by a photocell, or reflecting this LED light on a mirror surface (incident light principle). In this case, the slots or the mirrored surface form the segments. Because of a light scattering effect between the LED and photocell, a sinusoidal voltage waveform arises instead of a triangular one. Together, with the inductive principle, they are the most commonly used incremental encoders. Additional incremental encoders work according to other physical principles.

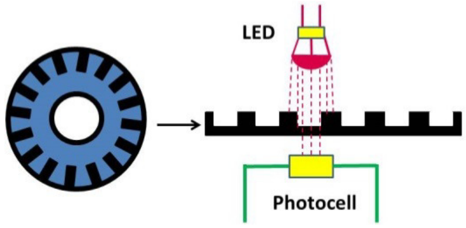


Figure 2: *Incremental encoder according to the optical-electronic principle.*

The counter in an encoder interface reacts to the change of edge of a signal from a previously defined low level (logic 0) to a defined high level (logic 1). With a digitalized signal, this edge change is easier to recognize since there are no undefined intermediate states in the transition from low to high. The digitization of a signal is performed using a comparator circuit.

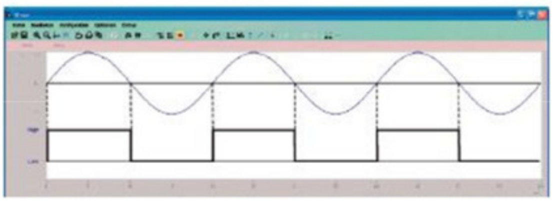


Figure 3: *Digitization of a sinusoidal incremental encoder signal.*

Usually, the analog signals are digitized within the sensor for easier processing. In cases where this does not happen, the difference between the various encoder interfaces becomes clear. To evaluate an analog signal properly, it requires a high-quality amplifier. The following section describes why certain properties of an amplifier are specifically distinguishable in this area.

Amplifier of an encoder interface:

Figure 4: imc CRONOS*flex* HRENC-4 module.

For measurement of a sensor that has longer lines and the effective suppression of common mode noise and ground loops, initially a high-impedance differential amplifier is needed. The noise that usually adheres to the analog signals, also requires different filters whose cutoff frequency must be set differently, depending on the maximum signal frequency expected for the incremental encoder. Since this edge change must be counted, within the encoder interface there must be threshold conditions for the conversion of analog signals into countable digital signals. The crucial factor is a freely definable hysteresis of the threshold. This prevents the superimposed interferences that oscillate about the thresholds of the signal are not counted more than once including the multiple counting of spurious peaks or dips, e.g., when performing RPM measurements. In order to ensure the optimal use of different incremental encoders on an encoder interface, the various filters and the hysteresis channels should be set individually. Also, especially with incremental encoders, in which the amplitude of the signal being produced is not constant and the decreasing amplitude of the signal to noise ratio is getting worse, one must ensure an adjustable hysteresis for correct operation of the counter. For example, increasing rotary encoders that use permanent magnets (segments) induce a voltage in a coil (sensor) with the speed of their amplitude by the proportionality between velocity and the induced voltage.

Zero pulse:

Most incremental encoders produce a so-called zero pulse with an excellent increment at the periphery. A measurement system which evaluates these knows after the first zero pulse, the position in which the encoder is located, and also avoids errors in the measurement from phantom pulses.

Most encoder interfaces are able to process this information.

Resolution:

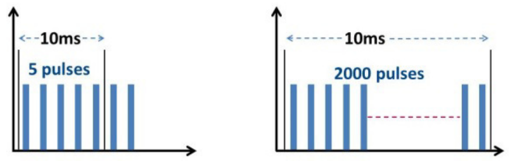
The accuracy with which an incremental encoder measurement can be carried out, doesn't depend solely on the number of segments. In general, the fineness of the angular increments, which is determined by the number of segments, is referred to as the resolution. An aspect that is often not taken into account with incremental encoders is the influence that the number of segments, thus, the number of pulses per revolution, of the measurement's cutoff frequency.

For instance, take the following practical example:

You are measuring the RPM of an internal-combustion engine flywheel. However, what you really want to see is the changes in RPM. An incremental encoder, with only one pulse per revolution, offers at 3000 RPM, a measurement frequency of 50Hz. According to the sampling theorem, only signals with a frequency ≤ 25Hz can be measured. In a four-cylinder engine alone, they receive 4 excitation pulses during one revolution. As a result, the measured RPM change already receives a fundamental frequency of 200Hz. Suppose then, that its harmonics are also measured. An incremental encoder with a resolution in the angle range of tenths or hundredths of a degree is required. Another aspect that determines the resolution of a measurement lies in the way the incoming signals are evaluated in an encoder interface. This applies to measurement values derived from an additional time measurements, such as velocity, frequency or RPM measurements. In terms of the resolution, three possibilities of velocity measurements are compared here in order to illustrate the difference. For each of the three methods of measuring velocity given here, the case of slow and high signal speeds are considered.

Measuring velocity – method 1:

Here, the velocity is easily calculated from the number of pulses received during a sampling interval. The sampling interval is 10ms long. In case 1, an incremental encoder is used in which one pulse per revolution is produced and, in case 2, an incremental encoder in which 400 pulses per revolution is generated.



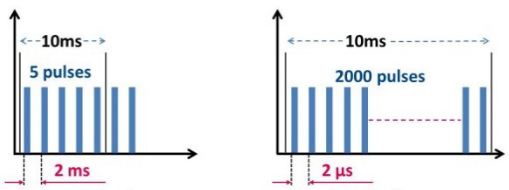
Case 1: Pulse frequency = 500Hz

Case 1: Pulse frequency = 200kHz

By this measurement mode, the resolution is determined by the number of pulses within the sampling interval. The resolution in Case 1 is therefore 1: 5, which corresponds to a resolution of less than 3 bits. However, in Case 2 the resolution increases to 1: 2000, which already corresponds to a resolution of almost 11 bits. In this method, the resolution increases with the number of pulses (segments) within a sampling interval. This results in an extremely limited range of use.

Measuring velocity – method 2:

The most commonly used method to measure velocity is to measure the time between two successive pulses and from this, determine the velocity. Here, pulse counting is omitted. This method is in contrast to the first one and is fully independent on the length of the sampling interval. It depends only on the resolution of the time measurement. If the time is measured by means of a counter, which is started with two successive pulses and stopped, then the counter frequency determines the time resolution. In this example, the counter frequency is 32MHz.



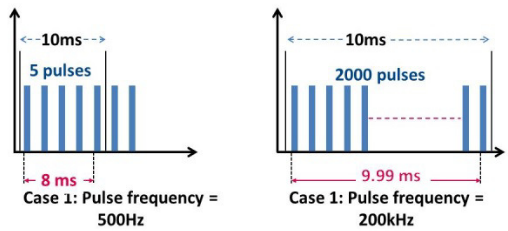
Case 1: Pulse frequency = 500Hz

Case 1: Pulse frequency = 200kHz

The resolution in this method is derived from the pulse frequency and the counter frequency. In Case 1, the resolution is 1: 64,000, which corresponds to a resolution of just under 16 bit. In Case 2, a resolution of 1: 160 is measured, which corresponds to a decrease to less than 8 bits. Using this method for measuring the velocity would clearly bring better results. The weakness of method of measuring velocity becomes visible only at high pulse frequencies.

Measuring velocity – method 3:

Another method that the company imc Meßsysteme GmbH introduced into their encoder interfaces (ENC- 4) is to use a combination of timing and counting of the pulses. In this case, it is not the time between two pulses that follow one another that are measured, but instead, the time between the first and the last pulse within a sample interval. Together, with the number of pulses within this sampling interval, the velocity is determined. In this example, the counter frequency is 32MHz.



In Case 1, the resolution is 1: 256,000, thus, just under 18 bits. In case 2, due to the higher pulse frequency, resolution improved. Since the first and last pulses are now significantly closer to the sampling interval threshold, the measured time compared to case 1 increases. This results in a resolution of 1: 319,680, which corresponds to a resolution of 18 bits. The advantage of this method is that it provides high resolution, and above all, stability over a wide range at very low to extremely high pulse frequencies. However, resolution is not the same resolution: To have the knowledge of these different methods, as with the encoder interface and its ability to exploit signals, allows one in the first place to objectively compare and evaluate each other. As has been shown, counter frequency alone is no clear measure of the resolution quality of an encoder interface. Only the matched operation of a high counter frequency, as with the encoder interface used by imc, can a high resolution in every application be guaranteed.

In general, this illustrates that the pure technical data are not necessarily a criterion for the quality of an encoder interface. To provide background knowledge about a product or topic is what the goal of this, and other White Papers is and is what we would like to provide to potential customers. To conclude, a brief description of two-lane incremental encoders is given and be sure to read the second part of this White Paper on further evaluation possibilities of incremental encoder signals.

Quadrature (two-track) encoder:

In most cases, it is not only important to know how fast something is moving, but if it moves forward or backward. For this reason, most incremental encoders feature a second track (Y-track) that generates a staggered pulse one-quarter from the first track (X-track). The direction of rotation can be determined by whichever pulse comes first. Some encoder interfaces, such as the imc CRONOS*flex* HRENC-4, use this original pulse scheme thereby resulting in a pulse multiplication in favor of the resolution. In this case, both tracks generate a pulse from each edge change (High -> Low, low -> High). This results in a pulse multiplication by a factor of 4. Sensors that deliver sinusoidal signals offset by 90° from one another over their two-track analog counterpart are referred to as SinCos encoders. These are a special feature, as they offer the possibility of a huge increase in resolution with its orthogonal system. Exactly how this type of resolution enhancement works and what prerequisites must be met by an encoder interface to realize this can be dealt with in detail in a subsequent White Paper.

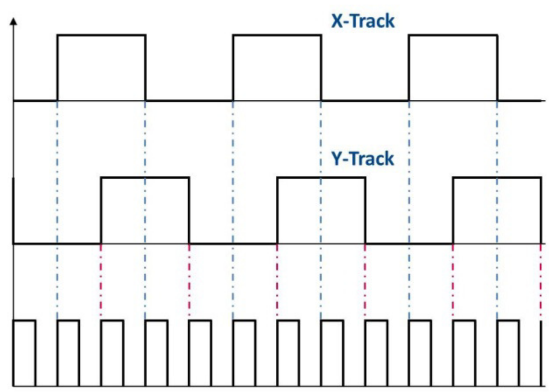


Figure 5: *Pulse multiplication by pulse generation at edge changes of both tracks (X, Y).*

Additional information:

imc Test & Measurement GmbH
Voltastr. 5
13355 Berlin, Germany

Telephone: +49 (0)30-46 7090-0
Fax: +49 (0)30-46 31 576
E-mail: hotline@imc-tm.de
Internet: http://www.imc-tm.com

imc Test & Measurement GmbH is a manufacturer and solution provider of productive test and measurement systems. imc implements metrological solutions for research, development, service and production. imc has particular expertise in the design and production of turnkey electric motor test benches. Precisely outfitted sensor and telemetry systems complement our customer applications.

Our customers from the fields of automotive engineering, mechanical engineering, railway, aerospace and energy use imc measurement devices, software solutions and test stands to validate prototypes, optimize products, monitor processes and gain insights from measurement data. As a solution provider, imc offers their customers an attractive and comprehensive range of services. These include project consulting, contracted measurements, data evaluation, specialist deployment, customer-specific software development and system integration. imc consistently pursues its claim of providing services for “productive testing”.

If you would like to find out more specific information about imc products or services in your particular location, or if you are interested in becoming an imc distributor yourself, please go to our website where you will find both a world-wide distributor list and more details about becoming an imc distributor yourself:

http://www.imc-tm.com/our-partners/