ENCODERS AND THEIR TYPES

An **encoder** is a sensor of mechanical motion that generates digital signals in response to motion. As an electro-mechanical device, an encoder is able to provide motion control system users with information concerning position, velocity and direction. There are two different types of encoders: linear and **rotary**. A linear encoder responds to motion along a path, while a rotary encoder responds to rotational motion. An encoder is generally categorized by the means of its output. An incremental encoder generates a train of pulses which can be used to determine position and speed. An absolute encoder generates unique bit configurations to track positions directly.

Basic Types of Encoders

Linear and rotary encoders are broken down into two main types: the absolute encoder and the incremental encoder. The construction of these two types of encoders is quite similar; however they differ in physical properties and the interpretation of movement.

Incremental Encoder

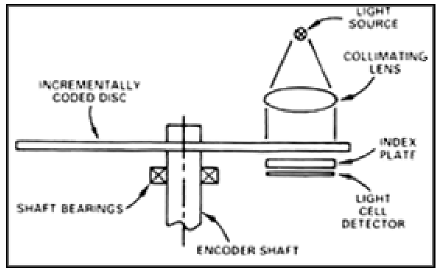
An Incremental rotary encoder is also referred to as a quadrature encoder. This type of encoder utilizes sensors that use **optical**, mechanical or **magnetic** index counting for angular measurement.

How do Incremental Encoders Work?

Incremental rotary encoders utilize a transparent disk which contains opaque sections that are equally spaced to determine movement. A light emitting diode is used to pass through the glass disk and is detected by a photo detector. This causes the encoder to generate a train of equally spaced pulses as it rotates. The output of incremental rotary encoders is measured in pulses per revolution which is used to keep track of position or determine speed.

A single-channel output is commonly implemented in applications in which direction of movement is not significant. Instances in which direction sensing is important, a 2-channel, quadrature, output is used. The two channels, A and B, are commonly 90 electrical degrees out of phase and the electronic components determine the direction based off the phase relationship between the two channels. The position of an incremental encoder is done by adding up all the pulses by a counter.

A setback of the incremental encoder is count loss which occurs during power loss. When restarting, the equipment must be referenced to a home position to reinitialize the counter. However, there are some incremental encoders, like those sold at Anaheim Automation, which come equipped with a third channel called the index channel. The index channel produces a single signal pulse per revolution of the encoder shaft and is often used as a reference marker. The reference marker is then denoted as a starting position which can resume counting or position tracking.



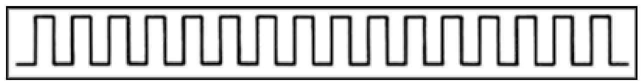


Figure 5: Pulse Train Produced from Incremental Encoder

NOTE: Incremental rotary encoders are not as accurate as absolute rotary encoders due to the possibility of interference or a misread.

Absolute Encoder

An absolute encoder contains components also found in incremental encoders. They implement a photodetector and LED light source but instead of a disk with evenly spaced lines on a disc, an absolute encoder uses a disk with concentric circle patterns.

How do Absolute Encoders Work?

Absolute encoders utilize stationary mask in between the photodetector and the encoder disk as shown below. The output signal generated from an absolute encoder is in digital bits which correspond to a unique position. The bit configuration is produced by the light which is received by the photodetector when the disk rotates. The light configuration received is translated into gray code. As a result, each position has its own unique bit configuration.

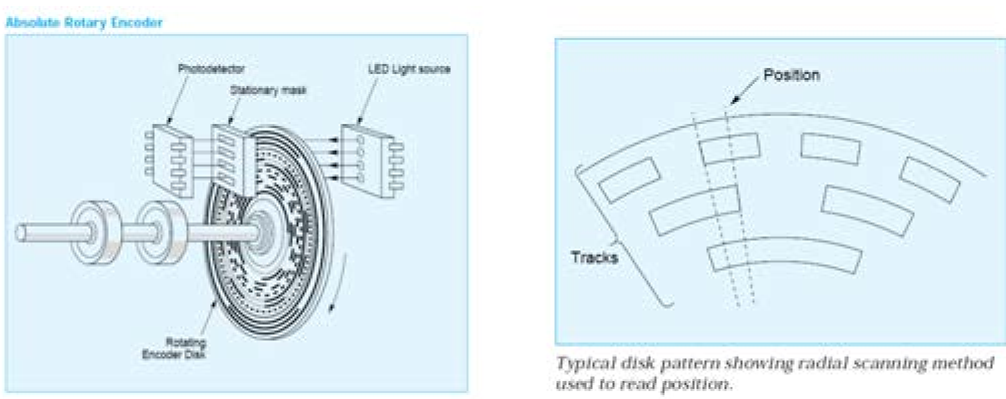


Figure 6: Components of Absolute Encoder

Linear Encoder

A linear encoder is a sensor, transducer or reading-head linked to a scale that encodes position. The sensor reads the scale and converts position into an analog or digital signal that is transformed into a digital readout. Movement is determined from changes in position with time. Both optical and magnetic linear encoder types function using this type of method. However, it is their physical properties which make them different.

How do Optical Linear Encoders Work?

The light source and lens produce a parallel beam of light which pass through four windows of the scanning reticle. The four scanning windows are shifted 90 degrees apart. The light then passes through the glass scale and is detected by photosensors. The scale then transforms the detected light beam when the scanning unit moves. The detection of the light by the photosensor produces sinusoidal wave outputs. The linear encoder system then combines the shifted signals to create two sinusoidal outputs which are symmetrical but 90 degrees out of phase from each other. A reference signal is created when a fifth pattern on the scanning reticle becomes aligned with an identical pattern on the scale.

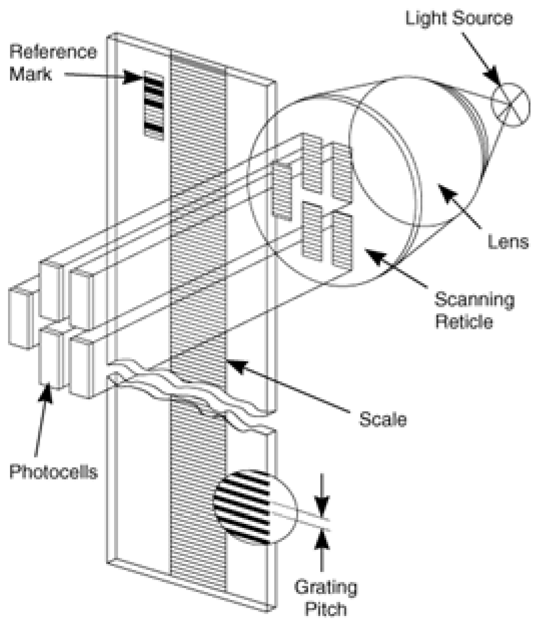
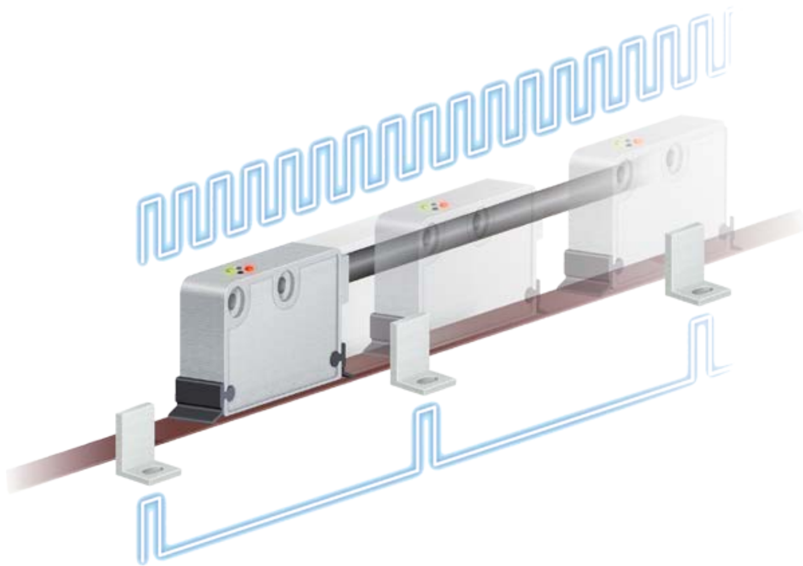


Figure 4: Linear Encoder Components

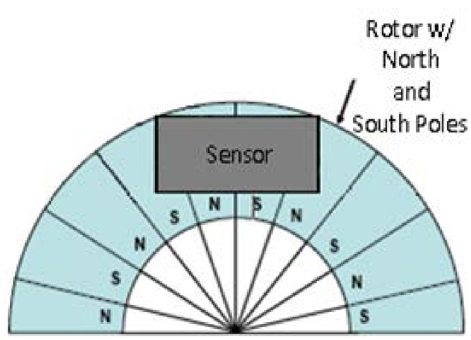
How does a Linear Encoder Work?

A Linear Encoder system uses a magnetic sensor readhead and a magnetic scale to produce TTL or analog output for Channel A and B. As the magnetic sensor passes along the magnetic scale, the sensor detects the change in magnetic field and outputs a signal. This output signal frequency is proportional to the measuring speed and the displacement of the sensor. Since a linear encoder detects change in the magnetic field, the interference of light, oil, dust, and debris have no effect on this type of system; therefore they offer high reliability in harsh environments.



Magnetic Rotary Encoder

A magnetic encoder consists of two parts: a rotor and a sensor. The rotor turns with the shaft and contains alternating evenly spaced north and south poles around its circumference. The sensor detects these small shifts in the position N>>S and S>>N. There many methods of detecting magnetic field changes, but the two primary types used in encoders are: Hall Effect and Magneto resistive. Hall Effect sensors work by detecting a change in voltage by magnetic deflection of electrons. Magneto resistive sensors detect a change in resistance caused by a magnetic field.



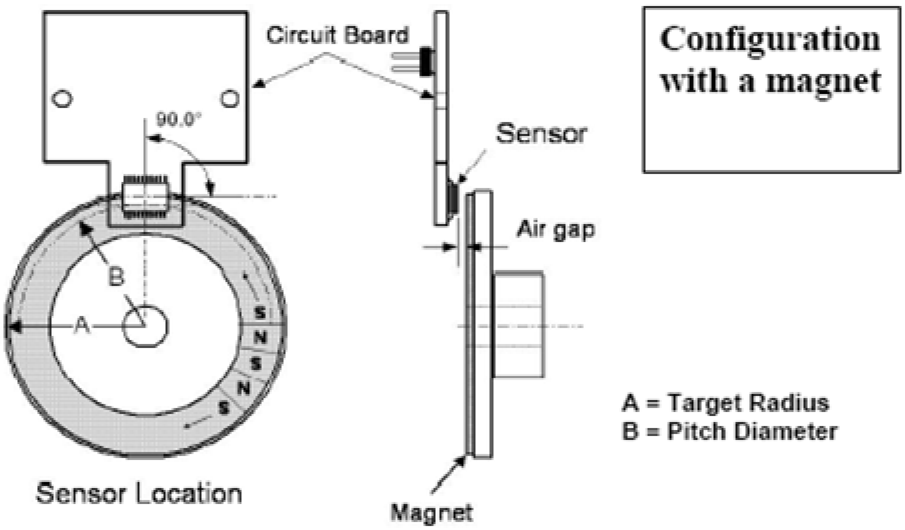
Hall-Effect sensing
The Sensor produces and processes Hall-Effect signals producing a quadrature signal as is common with optical

encoders. The output is generated by measuring magnetic flux distributions across the surface of the chip. The output accuracy is dependent on the radial placement of the IC with respect to the target magnet. The chip face should be parallel to the magnet so the magnet to sensor air gap is consistent across the sensor face.

Magnetic encoders avoid the three vulnerabilities that optical encoders face:

- Seal failures which permit the entry of contaminents
- The optical disk may shatter during vibration or impact
- Bearing failures

Magnetic devices designed effectively eliminate the first two failure modes and offer an opportunity to reduce bearing failures as well. Magnetic encoders do not make errors due to contamination because their sensors detect variations in magnetic fields imbedded in the rotor and oil, dirt and water do not affect these magnetic fields.



Hall-Effect sensors generally have lower cost and are less precise than magnetic resistive sensors. This means that Hall-Effect sensors, when used in an encoder produce more "jitter", or error in the signal caused by sensor variations.

Using Encoders with Arduino

A simple Example:

The following sketch demonstrates how an encoder is read. It simply updates a counter (encoder0Pos) every time the encoder turns by one step, and sends it via serial to the PC.

```
/* Read Quadrature Encoder
   Connect Encoder to Pins encoder0PinA, encoder0PinB, and +5V.

   Sketch by max wolf / www.meso.net
   v. 0.1 - very basic functions - mw 20061220

*/

int val;
int encoder0PinA = 3;
int encoder0PinB = 4;
int encoder0Pos = 0;
int encoder0PinALast = LOW;
int n = LOW;
```

```
void setup() {
  pinMode (encoder0PinA, INPUT);
  pinMode (encoder0PinB, INPUT);
  Serial.begin (9600);
}

void loop() {
  n = digitalRead(encoder0PinA);
  if ((encoder0PinALast == LOW) && (n == HIGH)) {
    if (digitalRead(encoder0PinB) == LOW) {
      encoder0Pos--;
    } else {
      encoder0Pos++;
    }
    Serial.print (encoder0Pos);
    Serial.print ("/");
  }
  encoder0PinALast = n;
}
```

Libraries

AdaEncoder

Utilizes any of the ATmega328P pins via the PinChangeInt library.
The AdaEncoder library was created for working with basic 2-pin quadrature encoders such as the following:
https://www.adafruit.com/products/377
http://www.sparkfun.com/products/9117
From the Introduction:
This library interfaces with 2-pin encoders (2 pins A and B, then a common pin C). It does not indicate every state change, rather, it reports only when the decoder is turned from one detent position to the next. It is interrupt-driven and designed to be fast and easy to use. The interrupt routine is lightweight, and the programmer is then able to read the direction the encoder turned at their leisure (within reason; what's important is that the library is reasonably forgiving). The library is designed to be easy to use (it bears repeating :-)) and it is reasonably immune to switch bounce.

Here's an example with two encoders connected. Encoder a is connected to pins 2 and 3, b is connected to 5 and 6:

```
#include <PinChangeInt.h> // necessary otherwise we get undefined reference errors.
#include <AdaEncoder.h>

#define a_PINA 2
#define a_PINB 3
#define b_PINA 5
#define b_PINB 6
```

```
int8_t clicks = 0;
char id = 0;

void setup()
{
  Serial.begin(115200);
  AdaEncoder::addEncoder('a', a_PINA, a_PINB);
  AdaEncoder::addEncoder('b', b_PINA, b_PINB);
}

void loop()
{
  encoder *thisEncoder;
  thisEncoder = AdaEncoder::genie(&clicks, &id);
  if (thisEncoder != NULL) {
    Serial.print(id); Serial.print(':');
    if (clicks > 0) {
      Serial.println(" CW");
    }
    if (clicks < 0) {
      Serial.println(" CCW");
    }
  }
}
```