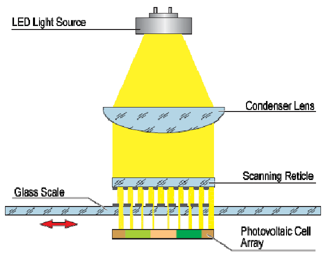
13745 Stockton Ave
Chino, CA 91710
(909) 614-4522

Four Steps to Understanding Scales

Weekly we receive calls and emails from customers asking, "What's the difference between linear scales?" The answer to this question is critical to your expected manufacturing quality outcome in addition to the overall cost of ownership. A Tech Authority is positioned to offer you the best scale solution to match your machining process, budget, and tolerances. Here are four easy steps to understand and help you make an educated decision.

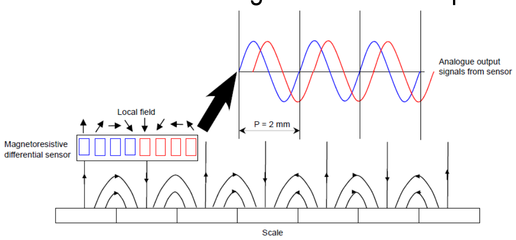
Step 1 - Understand Signal Generation

The basic principle behind the function of all incremental scales is the same; producing counts by detecting transitions. Glass scales detect light and dark transitions, magnetic scales detect North and South polar field transitions, and some use a principle called electromagnetic induction, which measures the change in density of steel balls inside a tube. No matter the method used, the signal generated is always a sine wave, which is divided into smaller signals, producing a fine measuring step required for precision measurement. Dividing the sine wave into smaller steps however, produces a hidden error called "short wave error," which can potentially be greater than the overall error of the entire scale. Knowing this, we can understand why the resolution of a scale doesn't mean nearly as much as the accuracy at which the signals are produced. Let's take a look at some of the most common signal generation technologies.



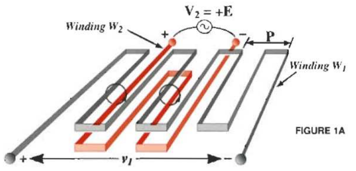
Optical scales are the most common type of scale technology used today. Optical scales are generally glass and are cemented to a support, usually an aluminum extrusion. The scanning unit contains a light source, photocells, and a second graduated piece of glass called the scanning reticle. This scanning reticle sits a short distance from the scale glass. A parallel beam of light produced by the light source and lens, passes through windows on the scanning reticle, through the glass scale, and onto a set of photosensors. When the scanning unit moves, the scale modulates the light beam, creating sinusoidal outputs from the photosensor. An additional window on the scanning reticle has a random graduation that, when aligned with an identical pattern on the scale, creates a reference signal. Encoders described above are limited to measuring lengths of about 3 m or so. Encoders that measure longer distances contain scales made from steel tape. A typical tape is about 0.3 mm thick and contains a grating of highly reflective gold lines with a pitch of about 40µm.

Magnetic scales frequently use a scale consisting of a thin ferromagnetic bar or rod of magnetic material fastened at both ends. The bar is imprinted with magnetic domains. A read head senses magnetic fields as the bar moves to measure position. Ferromagnetic bars are typically magnetized with a period of 0.2 mm (0.008 in.). Sensing electronics basically consist of two magnetic heads spaced so that signals are 90°apart. Each head contains a low-reluctance magnetic yoke wound with primary and secondary coils. Primary coils are powered with alternating current and as the magnetic bar moves past the sensing unit, it changes the magnetic flux through each yoke. This modulates the electrical coupling between the coils because of the nonlinear magnetic qualities of the yoke material. Thus, the amplitude of the output signals change as the scale bar moves resulting in amplitude modulated sine waves. The resulting signal is derived from the two output signals as the phase changes with the relative movement of the bar and reading head. Interpolation circuits can detect this change to gauge position. However, these types of systems have seen only limited use because of the necessary high interpolation factor.

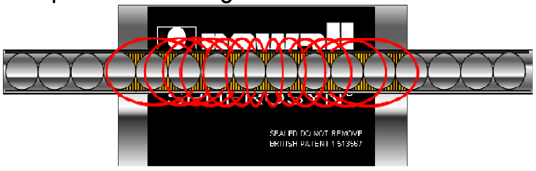


13745 Stockton Ave
Chino, CA 91710
(909) 614-4522

Inductosyn scales are similar in operation to the magnetic systems. These devices contain a scanning plate (slider) and a scale. The scale is a meander pattern that is supplied with a carrier frequency. The slider contains two coils spaced 90° apart. The coils inductively pick up the signal and produce an alternating voltage amplitude that is proportional to the sine and cosine of the slider position within a scale pitch. Tracking electronics count signals received from the slider coils to gauge motion. These systems also incorporate interpolation so that a scale pattern with a 2mm pitch produces a 1µm linear resolution. However, inductosyns are not as widely used as other systems. The mounting and adjusting of the scales can be tedious; and the system must be protected from contaminants.

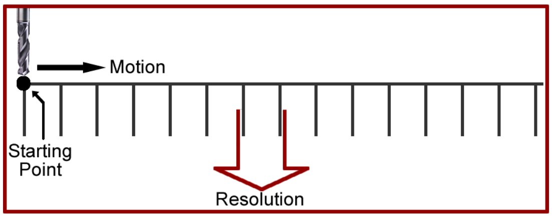


Electromagnetic Induction scales are similar in operation to both the magnetic and inductive technologies seen above. Electromechanical Induction scales consist of high-precision nickel-chrome ball bearings inserted into a stainless steel or carbon fiber tube. The ball bearings are then pre-loaded against a traceable standard to compensate for long wave inaccuracies. Inducing a 10kHz sinusoidal current through a single drive coil within the reader head generates an electromagnetic field. This field interacts with the nickel chrome elements contained in the scale. A set of four pickup coils detect variations in the induced field that are then combined and processed by the electronic circuitry to generate a signal that varies as the head moves along the scale. Depending on the position of the reader head as it passes over each element, the phase shift of this pickup signal relative to the drive signal will vary between 0 and 360 degrees.



Step 2 - Understand Resolution

Resolution is the smallest interval recognized between two distinguishable points. Or think of it this way; resolution is the smallest degree of movement that a scale can detect. Also known as step size, resolution is determined by the feedback device and capabilities of the motion system and theoretical resolution may exceed practical resolution. For example, in a ball-screw-based positioning system, a theoretical resolution of 1µm can be obtained by combining a 10mm/rev screw, a 1000-line encoder, and a x10 multiplier. The actual motion system will never be able to make a single 1µm step due to friction, windup, and mechanical compliance. Therefore, the practical resolution is actually less.



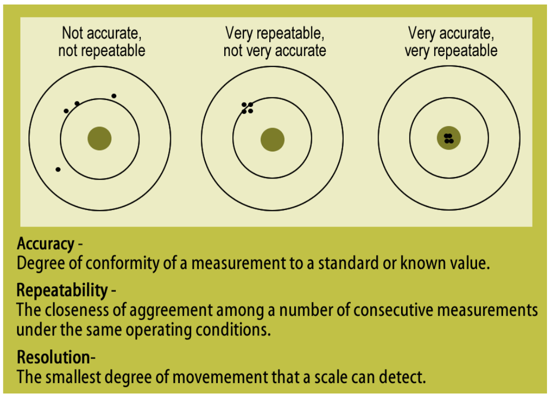
It is important to realize that resolution is only meaningful when measured in terms of the output motion of the component. For example, a translation stage may have a digital readout that displays the position to 6 digits (i.e. 0.000001) which is its *display resolution*, while its smallest incremental movement capability is 0.0001. In this case, the position resolution is 10^{-4} while the display resolution is 10^{-6}.

Step 3 - Understand Accuracy

Accuracy is the degree of conformity to a standard or known value while repeatability is the degree of reproducibility. The analogy used here to explain the difference between accuracy and repeatability is the target comparison. In this analogy, repeated measurements are compared to arrows that are shot at a target. Accuracy describes the closeness of arrows to the bullseye at the target center. Arrows that strike closer to the bullseye are considered more accurate. The closer a system's measurements to the accepted value, the more accurate the system is considered to be.

13745 Stockton Ave
Chino, CA 91710
(909) 614-4522

A Tech Authority
When Precision Counts



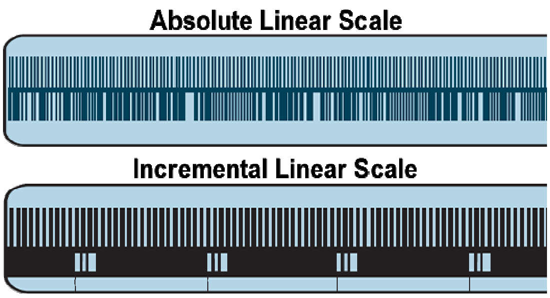
To continue the analogy, if a large number of arrows are shot, repeatability would be the size of the arrow cluster. When all arrows are grouped tightly together, the cluster is considered repeatable since they all struck close to the same spot, even if not necessarily near the bullseye. The measurements are repeatable, though not necessarily accurate.

However, it is not possible to reliably achieve accuracy in individual measurements without repeatability. The average position might be an accurate estimation of the bullseye, but the individual arrows are inaccurate.

Ideally a measurement device is both accurate and repeatable, with measurements all close to and tightly clustered around the known value. The accuracy and repeatability of a measurement process is usually established by repeatedly measuring some traceable reference standard. Such standards are defined in the International System of Units and are maintained by national standards organizations such as the National Institute of Standards and Technology.

Step 4 - Understand Incremental & Absolute

With the **incremental measuring method,** the graduation consists of a periodic grating structure. The position information is obtained by counting the individual increments (measuring steps) from some point of origin. Since an absolute reference is required to ascertain positions, these types of scales are provided with an additional track that bears a reference mark. The absolute position on the scale, established by the reference mark, is gated with exactly one signal period. The reference mark must therefore be scanned to establish an absolute reference or to find the last selected datum.



In some cases this may necessitate machine movement over large lengths of the measuring range. To speed and simplify such “homing sequences,” some scales feature distance-coded reference marks—multiple reference marks that are individually spaced according to a mathematical algorithm. The subsequent electronics find the absolute reference after traversing two successive reference marks—only a few millimeters apart. With distance-coded reference marks, the absolute reference is calculated by counting the signal periods between two unique reference marks.

With the **absolute measuring method,** the position value is available from the encoder immediately upon switch-on and can be called at any time by the subsequent electronics. There is no need to move the axes to find the reference position. The absolute position information is read from the scale graduation, which is formed from a serial absolute code structure. A separate incremental track is interpolated for the position value and at the same time is used to generate an optional incremental signal.

Craig D’Ambrosio
General Manager
A Tech Authority, Inc.
www.atechauthority.com
craig@atechauthority.com
(909) 614-4518