chapter 5

Sensors

5.1 Linear Optical Sensors

The encoder functioning as a feedback device is one of the basic components of motion control systems. There are three different sensor technologies used in linear servo applications, i.e., inductive, magnetic and optical. The earliest linear encoders utilized in high precision machines, e.g., in metal-cutting industry, were optical. Although other techniques are now also available, optical encoders are still predominant in industrial applications. In linear motor drives, where precision actuation and measurement is involved, most designers employ an incremental optical encoder as a well-accepted part of the electromechanical drive system. The typical optical encoder makes use of a graduated *scale* which is scanned by a movable optical *readhead* . The most important advantage of optical encoders is their easily achievable non-contact operation which eliminates friction and wear and permits reliable high speed performance in workshop environments. Linear optical encoders are capable of achieving very *high resolution* , in some cases comparable to the *laser interferometer technology* . Their accuracy is a few orders of magnitude higher than that of similar magnetic or inductive linear encoders. This is possibly due to the superior precision of *interpolation* performed on much smaller scale grating periods. The interpolation is a self-subdivision process of the signal representing the scale period.

5.1.1 Incremental Encoders

There are two basic methods of generating optical encoder signals. In the first method the *transmitted light* is processed (Fig. 5.1a), while the second method employs the *reflected light* (Fig. 5.1b).

The simplest configuration of an optical encoder is described below. The light emitted by an LED either travels through, or reflects off the scale. It is then directed through an identical index grating and onto photo detectors, which generate electrical currents. Twin signals s_1 and

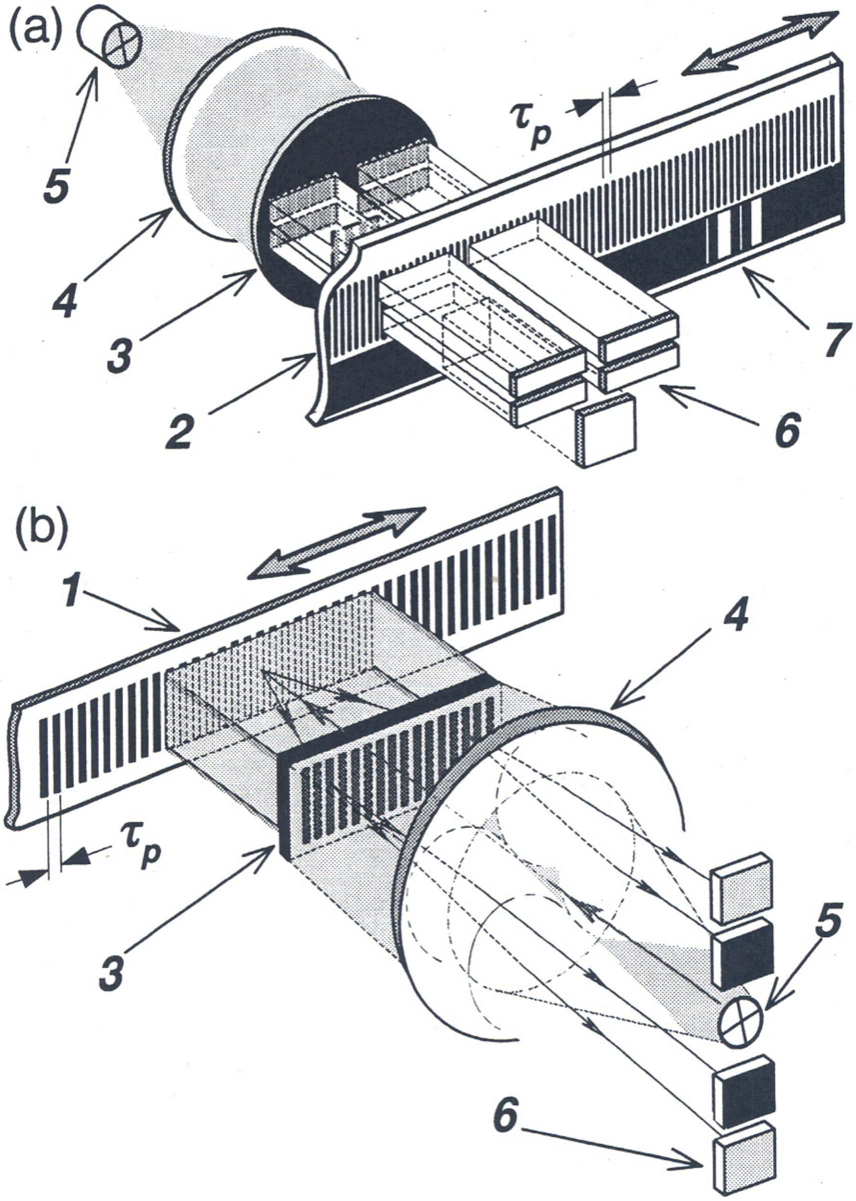


Figure 5.1 Typical scanning methods: (a) transmitted light method, (b) reflected light method. 1 -scale reflective tape, 2 - scale transparent glass,3 - scanning reticle, 4 - condenser lens, 5 - light source (LED), 6 -photodetectors,7 - reference mark.

s_2, 90° out of phase representing sine and cosine waveforms are generated. These operation principles of the optical incremental encoder are illustrated in Fig. 5.2. The photo detector reads the maximum luminous intensity when transparent slots of the scale fully align with transparent slots of the scanning reticle. The light source and the photo detectors move along the glass scale grating. In consequence, the transparent slots of the scanning reticle change periodically their position relative to the stationary slots of the scale. Therefore, the light intensity detected by the photo sensors (photo elements) changes its value from maximum to zero according to a sinusoidal function (Fig. 5.2). Because the photo elements are displaced at the distance of one quarter of the grating period τ_p of the scale, when one photo element detects maximum, the other one reads only half of that. This displacement effectively shifts the two signals s_1 and s_2 by 90° apart in the time domain. The 90° electrical separation or one quarter of the period between the two signals is referred to as the *quadrature* . Signals in the quadrature permit determination of the motion direction and speed at the same time allowing additional resolution through the *edge counting* .

Linear motors employed in the x-y positioning stages and used in harsh factory or workshop environments require precise resolution with high reliability. The combination of a reflective, flexible scale tape placed along the track, and the read head moving over the tape offers many unique features.

The scale is made out of the steel ribbon 5 to 10 mm wide and 0.2 mm thick, which has relatively low stiffness. Other materials, such as glass, mylar, or non-ferrous metal tapes can also be used. The scale is grated with alternating reflective strips (often made out of gold) and light absorbing spaces. The grating period ranges from 100 to less than 20 μ_m, and after interpolation, resolution up to 0.1 μ_m is possible. The scale can be secured to the most commonly used materials (metals, composites and ceramics) by means of a double-sided, elastic adhesive tape to accommodate the thermal expansion of the base. However, the mounting surface should be relatively smooth, clean and parallel to the axis of motion with the scale ends rigidly fixed to the axis of substrate. The location of stationary scale and moving read head mounted on a positioning stage is shown in Fig. 5.3. The differential movement between the scale and the substrate should be close to zero, even in the presence of large temperature gradients. Usually, the scale tape is protected by varnish coating to facilitate easy cleaning. Scale tapes are generally supplied on a reel for 'cut-to-suit' convenience. For comparison, glass scales reach maximum length of 3 m in a single piece. Usually, the incremental tapes are installed together with the reference marks. Limit switches are separately installed next to the scale itself. These home position and/or zero point indicators for the end of travel are also sensed by the

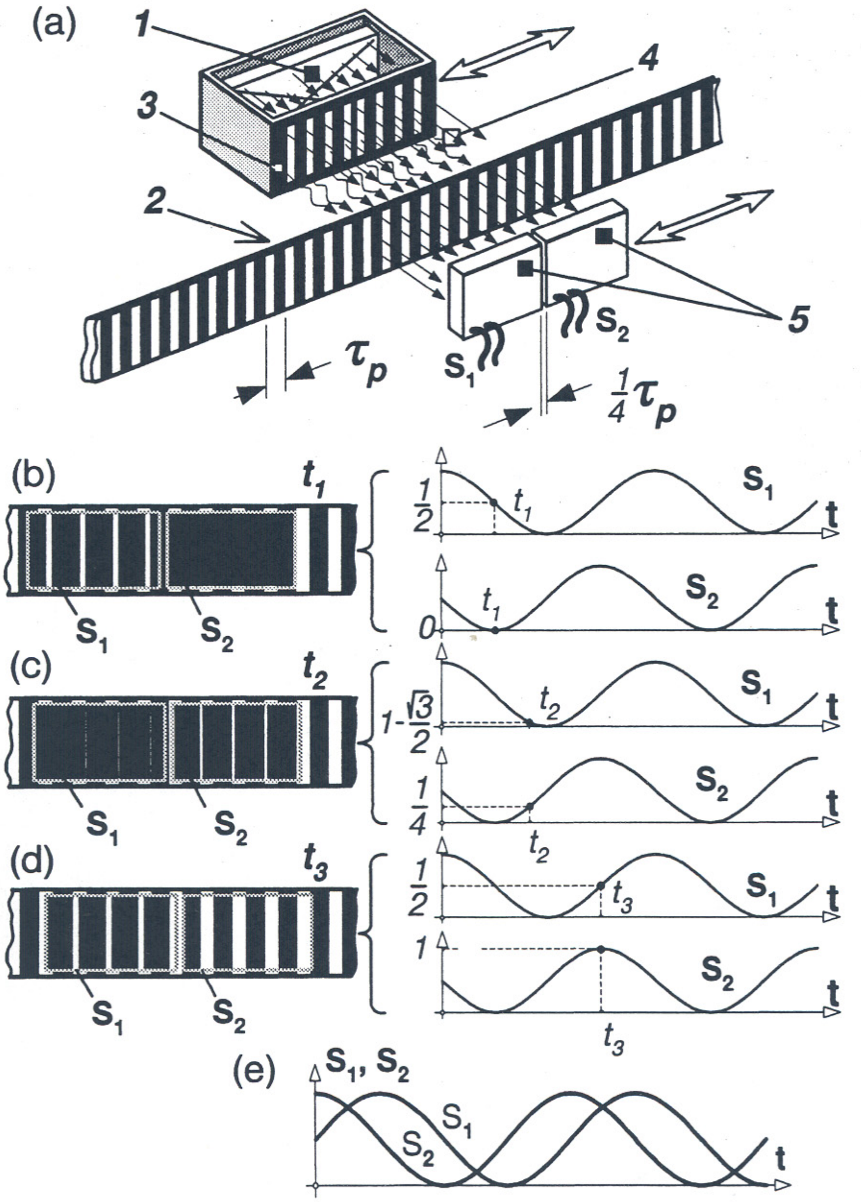


Figure 5.2 Generation of photo detector signals: (a) photoelectric scanning principles; (b),(c), (d) instants of relative position of scale gratings and scanning reticle; (e) variation of signals. 1 - light source (LED), 2 - glass scale grating, 3 - scanning reticle, 4 - light rays, 5 – photo detectors.

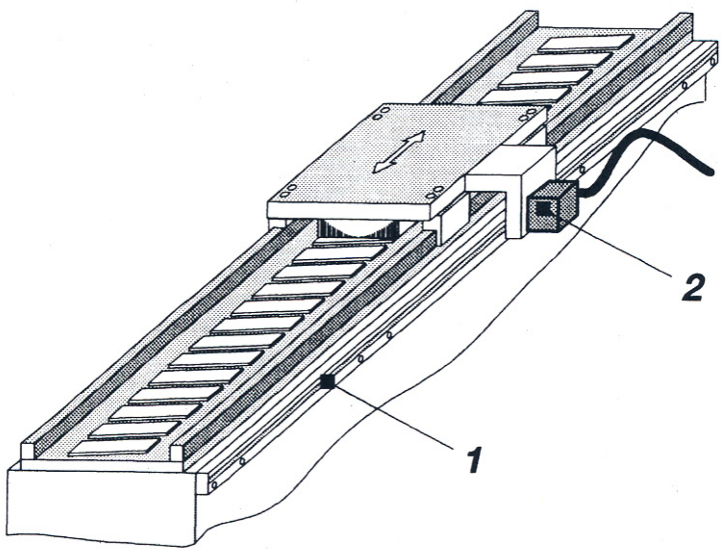


Figure 5.3 Typical location of the linear encoder installed in the positioning stage driven by a linear motor. 1 - scale, 2 – read head.

scanning read head. The signals are synchronized with the incremental channels to guarantee repeatability.

The principle of scanning process can be exemplified by the LIDA linear encoder which was introduced to the market in 1977 by *Heidenhain GmbH* , Traunreut, Germany [49]. The operation of LIDA encoder is explained in Fig. 5.1b. The redhead travels along the scale, while a light beam emitted by an LED source is directed onto the incremental scale grating through a condenser lens and four scanning reticles. Then, it is reflected, and after passing back through the reticles, it is focused onto the photoelectric cells. Photo sensors detect changes in the light intensity caused by the interaction between the scale and reticle gratings. The four sinusoidal scanning signals, corresponding to the changes in the light intensity, are produced by the sensors. These waveforms with 90° phase shift enable formulation of symmetrical encoder zero output signals.

To improve the accuracy, especially in the high precision positioning applications within microelectronics industry, encoders frequently use the interferential scanning principles. The diffracted (interferential) light method is required in the optical encoders with grating periods $\tau_p < 8\,\mu$m. These devices employ a reflection-type diffraction grating fixed the carrier (Fig. 5.4). An infrared LED emits light onto angular scale facets where it is scattered back into the read head through the transparent grating. The periodic pattern on the scale and the periodic indexing of the grating produce sinusoidal interference fringes at the photo detector plane. The fringes move across the detector plane as read head moves along the scale. The arrangement of interlaced groups of photo detectors positioned in repeating patterns generates electric signals related to the fringe movement. The read head electronics processes these signals and generates two sinusoidal waveforms of equal amplitude with the phase shift of 90°. In RGH encoders manufactured by *Renishaw plc* the signal is averaged from over 80 facets in the detector plane. Therefore, the loss of a number of scale facets has only a marginal effect on the signal's amplitude and does not affect the counting process. Furthermore, because the signal is often subjected to disturbing effects (contamination or minor damage to the scale), the filtering and averaging process ensures its stability. In essence, the electronics within the read head eliminates signals which do not match the scale period of 20 μm.

The electronics embedded in the read head converts scanned incremental signals into analog or digital sinusoidal waveforms in the quadrature (Fig. 5.5a). The signal period is equal to the scale pitch. The wave formats follow industry standard outputs: micro current (in μ A) or voltage (1 V peak-to-peak). The read head generates incremental square pulse trains in the quadrature which conform to the standard EIA/RS422 differential line drive output (Fig. 5.5b). These fine resolution digital waveforms are obtained by the subdivision of the analog signal passed from the read head optics. In this context, the resolution is defined as the distance between consecutive edges of the digitized pulse trains. Commercially available read heads typically achieve the resolution of 5.0, 1.0, 0.5 or 0.1 μ m. The read head interpolation is ratio metric, i.e., it is independent of the signal amplitude.

For digital output read heads, the recommended counter clock frequency for a given traversing speed is:

$$f = \frac{v_{tr}}{\tau_p} k_{sf} \qquad (5.1)$$

where v_{tr} is the traversing speed, τ_p is the read head resolution, f is the counter clock frequency of interpolation electronics, and k_{sf} , is the safety factor, typically k_{sf} = 4. If v_{tr} is in m/s and τ_p is in μ m, the counter clock frequency f is in MHz.

Specification of incremental self-adhesive scale RGS-S produced by *Renishaw plc* , Gloucestershire, U.K., is shown in Table 5.1 and Fig. 5.6. Specifications of analog and digital read heads RGH series used in conjunction with RGS-S tape are presented in Table 5.2 and Table 5.3. The edge separation characteristics typical for digital read heads are shown in Fig. 5.7.

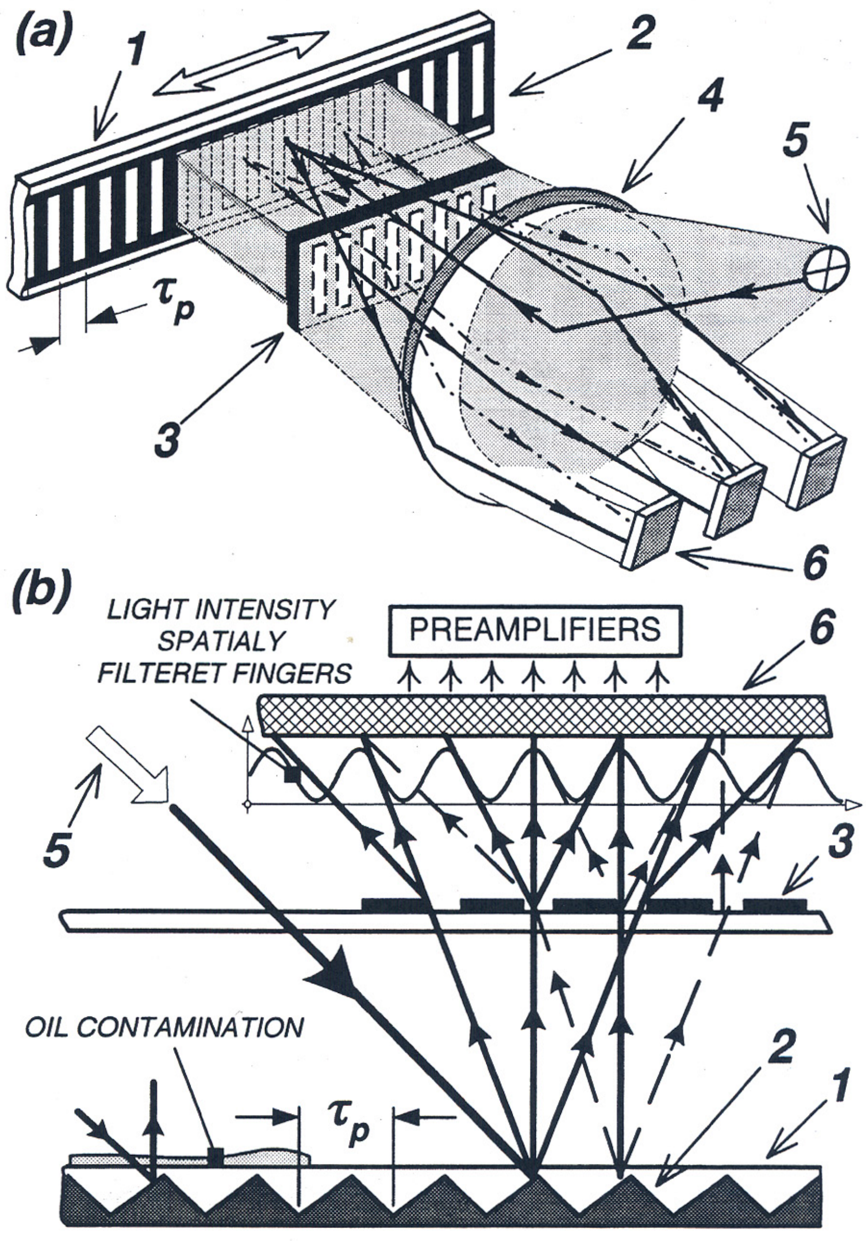
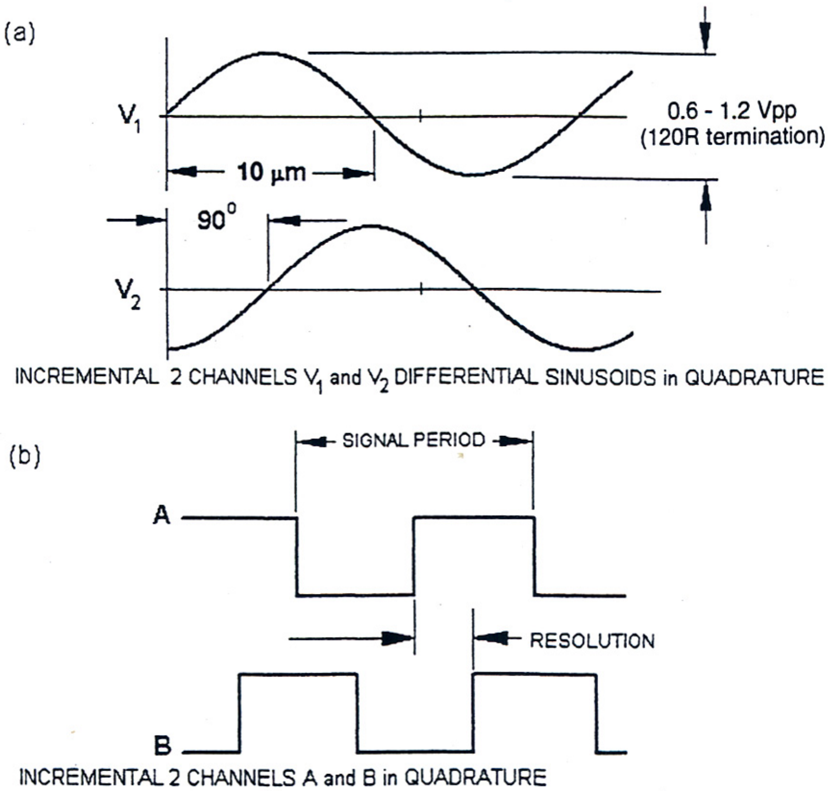


Figure 5.4 Interferential measuring: (a) photo electronic scanning, (b) optical filtering principle. 1 - scale, 2 - scale facets, 3 - phase gratings (read head window), 4 - condenser lens, 5 - oblique illumination from LED, 6 – photo detectors.



READHEAD MODEL	PERIOD (µm)	RESOLUTION (µm)
RGH22 D	20	5
RGH22 X	4	1
RGH22 Z	2	0.5
RGH22 Y	0.4	0.1

Figure 5.5 Incremental signals readheads manufactured by Renishaw: (a) analog readhead, (b) digital readhead. Courtesy of *Renishaw plc*, Gloucestershire, U.K.

Table 5.1 Self-adhesive scale RGS-S for RG2 encoder system manufactured by *Renishaw plc*, Gloucestershire, U.K.

Parameter	Specification
Scale type	Reflective gold plated steel tape with lacquer coating and self-adhesive backing
Scale pitch	20 μ m
Available lengths	Continuous length up to 50 m Longer than 50 m by special order
Measuring lengths	User selectable 'cut-to-requirements' at the place of installation
Accuracy	Typical 15 μ m/m without compensation
Linearity	±3 μ m/m, ±1 μ m/60mm
Substrate materials	Metals, ceramics and composites with expansion coefficient less than 22 μ m/m/°C
Reference mark	Magnetic actuator RGM22S epoxy mounted. One or more at user selected locations. Repeatability of position within: • temperature range ±10°C from installation • speed < 250 mm/s Screw mounted option RGM22SB available

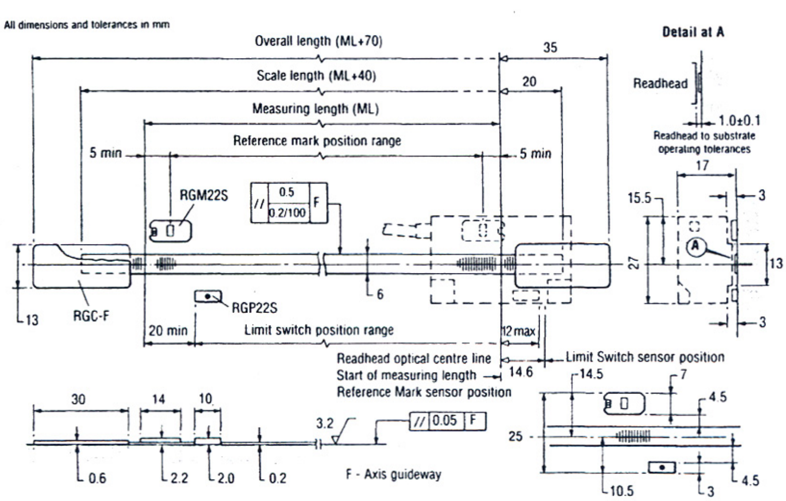


Figure 5.6 Dimensions of RGS-S scale and associated components. Courtesy of *Renishaw plc*, Gloucestershire, U.K.

Table 5.2 Analog read heads manufactured by *Renishaw plc*, Gloucestershire, U.K.

Parameter	Specifications	
Type	RGH22C 12 μ A differential.	RGH22B 1 V (peak-to-peak) differential
Signals	Incremental 2 channels I_1 and I_2 differential sinusoids in quadrature (90° phase shift). Signal period 20 μ m.	Incremental 2 channels V_1 and V_2 differential sinusoids in quadrature (90° phase shift). Signal period 20 μ m.
Output	7 to 16 μ A	0.6 to 1.2 V (peak-to-peak)
Reference	Differential pulse 10 μ A duration 126°	Differential pulse Duration 126°
Power supply	5v ±5%, 120 mA (typical)	
Speed	1 m/s at 50 kHz maximum	5 m/s at 250 kHz maximum
Temperature	-20 to +70°c storage 0 to +55°c operating	
Humidity	10 to 90% RH non-condensing	
Sealing	IP54	
Operating acceleration	30g	
Shock acceleration	100g (11 ms , one half of sinusoid)	
Vibration Under operation	10g at 55 to 2000Hz	
Mass	Read head: 45g, cable: 32g/m	
Cable	Available lengths 0.5,1.0,1.5,3.0 and 5.0 m Flexible life> 10^7 cycles at 50 mm bend radius for integral cable and> 10^6cycles at 75 mm bend radius for extension cable14 core, double shield, outer diameter 7.2 mm	

Table 5.3 Specifications of digital readhead manufactured by *Renishaw plc*, Gloucestershire, U.K.

Parameter	Specifications
Output signal	Square differential line driver to EIA RS422 Incremental channels A and B in quadrature (90° phase shift)
Signal period	20 μ m for D type 4 μ m for X type 2 μ m for Z type 0.4 μ m for Y type Resolution for all models 0.25xperiod
Alarm period	Separate alarm channel or three state alarm Incremental channels force an open circuit for reliable operation when signal is too low
Power supply	5v ±5%, 120 mA (typical) 150 ma for Y type only
Operating acceleration	30g
Shock acceleration	100g (11 ms , 1/2 sine)
Vibration Under operation	10g at 55 to 2000Hz(ICE 68-2-6)
Temperature	-20 to +70°c storage 0 to +55°c operating
Humidity	10 to 90% RH non-condensing
Mass	Read head: 45g, cable: 32g/m
Signal terminations	Standard RS422A line receiver circuitry RC filter is recommended Resistance 120 Ω Capacitor 4.7 nF for cable length < 25 m and 10 nF for cable length> 25 m
Cable	Available lengths 0.5,1.0,1.5,3.0 and 5.0 m Flexible life> 10^7 cycles at 50 mm bend radius for integral cable and> 10^6cycles at 75 mm bend radius for extension cable 14 core, double shield, outer diameter 7.2 mm

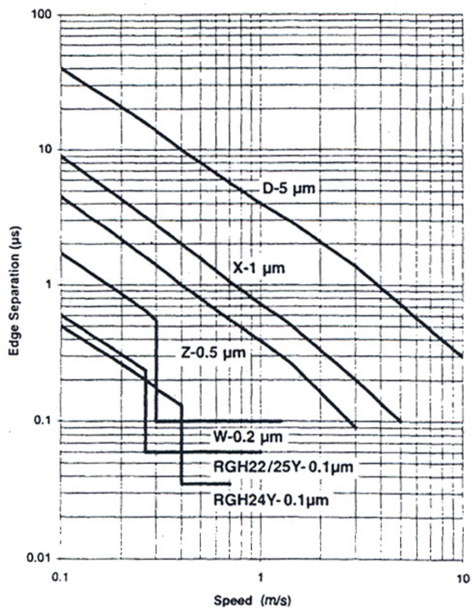


Figure 5.7 Edge separation for RGH readheads. Courtesy of *Renishaw plc*, Gloucestershire, U.K.

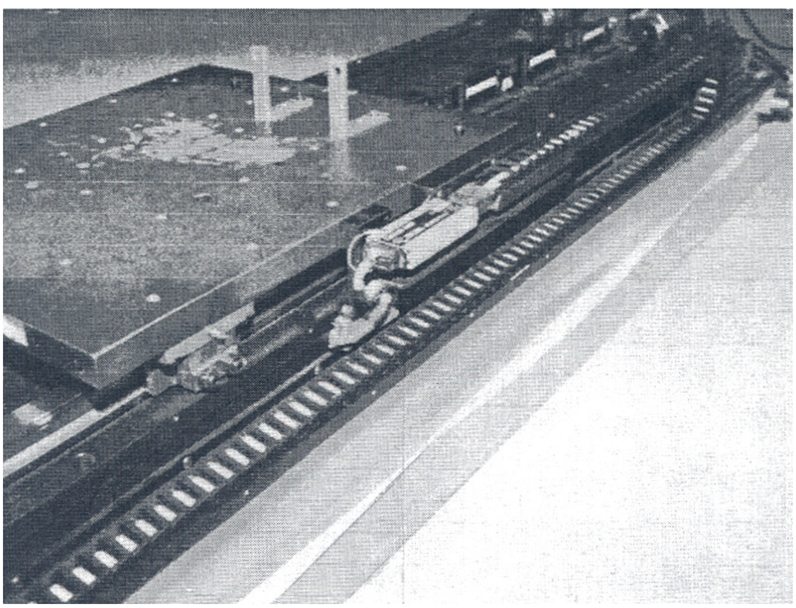


Figure 5.8 Readhead and reflective scale arrangement on the positioning stage with a linear motor. Photo courtesy of *United Technologies Research Center*, East Hartford, CT, U.S.A.

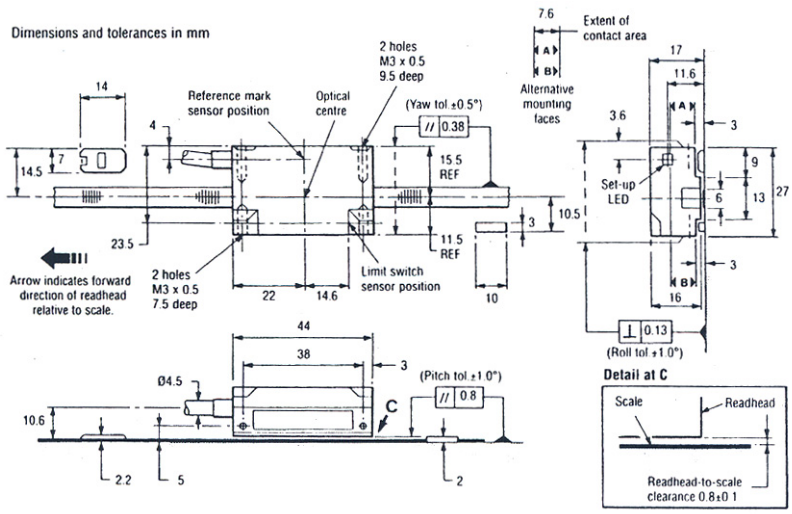


Figure 5.9 Dimensions of the analog readhead RGH. Courtesy of *Renishaw plc*, Gloucestershire, U.K.

The optical readhead working with reflective scale is shown in Fig. 5.8.

The commonly used scale RGS-S and one of the RGH22 series read-heads comprise the RG2 system, i.e., non-contact, optical encoder designed for position feedback solutions (*Renishaw*). The readhead can be chosen with either sinusoidal or square wave output. The selected type depends on the application, electrical interfacing, and required resolution. Typically, the RG2 encoders are employed in linear motor-driven machines such as tool presetters, measuring and layout equipment and other high speed systems in which interpolation is provided by subsequent electronics. In the environments subjected to severe radio frequency interference (RFI), the RGH22B readhead with analog differential output voltage is preferred to the RGH22C model having the output current signal.

Specifications of the two-coordinate PP 281 R encoder manufactured by *Heidenhain, GmbH* , Traunreut, Germany, are listed in Table 5.4.

The *x*-*y* motion stages applied in clean-room environments, e.g., semiconductor industry or ultra-precision machine tools, such as grinders for ferrite components and diamond lathes for optics are equipped with integrated two-coordinate encoders. The incremental *x*-*y* encoder contains a two-dimensional phase grating structure on a glass substrate (Fig. 5.10).

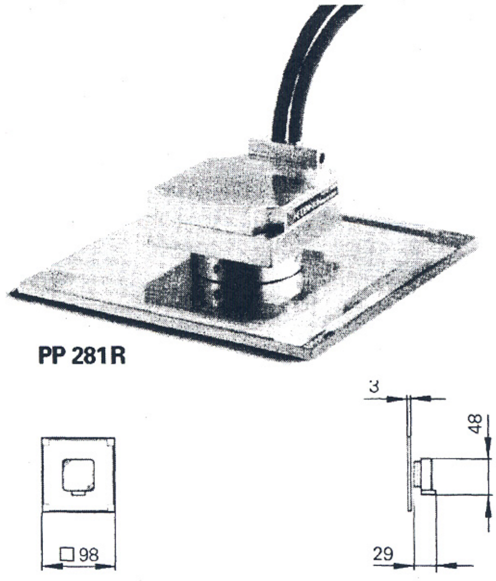


Figure 5.10 Two-coordinate incremental encoder PP281R. Photo courtesy of *Heidenhain, GmbH*, Traunreut, Germany.

Table 5.4 Specifications of PP 281 R two-coordinate incremental encoder manufactured by *Heidenhain, GmbH*, Traunreut, Germany.

Parameter	Specification
Grating period	$\tau_p = 8\ \mu$m
Coefficient of thermal expansion	8 pikomillimeter /K
Accuracy	$\pm 1\ \mu$m
Measuring range	68 mm X 68 mm other ranges also available
Vibration	less than 80 m/s^2 at 55 to 2000 Hz
Shock vibration	less than 100 m/s^2(11 mS)
Operating temperature	0 to +50°C
Mass of grid plate	75g
Mass of APE and cable	120g
Mass of scanning head	170g
Power supply	5 V ±10%, 100 mA (without load)
Output signal	1 v (peak-to-peak)
Signal period	4 μm

The measurement in a plane is possible through an interferential scanning method. Two reference marks, one in each measurement direction, serve to define accurately zero positions. The 8 μm grating period with fine interpolation and high uniformity of scanning is capable of 10 nm resolution.

5.1.2 Absolute Encoders.

Typically, *absolute encoders* are utilized in the devices inactive for long periods of time or moving at low speeds. They are also applied to the systems where linear position must be maintained regardless of power interruptions, or where safe and failure-free operation is required. Primarily, machine tools and robotics applications make use of the absolute position encoders. These devices supply a whole output word with unique binary code pattern representing each position. This code is derived from independent tracks on the linear scale detected by individual photo detectors. The output from these detectors would then be *high* or *low* depending on the code pattern read off the linear scale for the particular position. Absolute encoders are similar to incremental devices; however, they contain more sensors. The overall complexity depends on the generated size of the word. The longer the *logic* word, the more complex and expensive the system. For each *bit* in the output signal the encoder uses one track of the code scale. Therefore, a 10 bit encoder has 10 tracks to detect the light passing through them. For higher number of tracks it may be necessary to use multiple sources of light to assure an adequate illumination. The principle of operation of the linear absolute encoder is illustrated in Fig. 5.11. Although the information read from data tracks can be converted into position signals using many different codes, natural binary code (NBC), gray, gray excess, and binary coded decimal (BCD) codes are most common.

The NBC derives the numerical value from exponents with base 2. For example, the number 179 is expressed as $1 \times 2^7 + 0 \times 2^6 + 1 \times 2^5 + 1 \times 2^4 + 0 \times 2^3 + 0 \times 2^2 + 1 \times 2^1 + 1 \times 2^0$. In other words, the NBC value for 179 is 10110011.

The binary code is a *polystrophic code* characterized by multiple bit changes [121]. It requires many bit transitions simultaneously, e.g., counting from 127 to 128 in NBC requires simultaneous transition of 8 bits from 01111111 binary to 10000000 binary. In a practical electronic circuitry, all of these bits cannot be changed at precisely the same time. There is some delay within individual bit transitions. Ambiguity in the simultaneous bit changes, imperfection in the readhead mechanical installation, hysteresis and noise comprise only a few factors which affect the accuracy of the position detection. The potential error in the reading of the most significant bit can result in 180° feedback signal error.

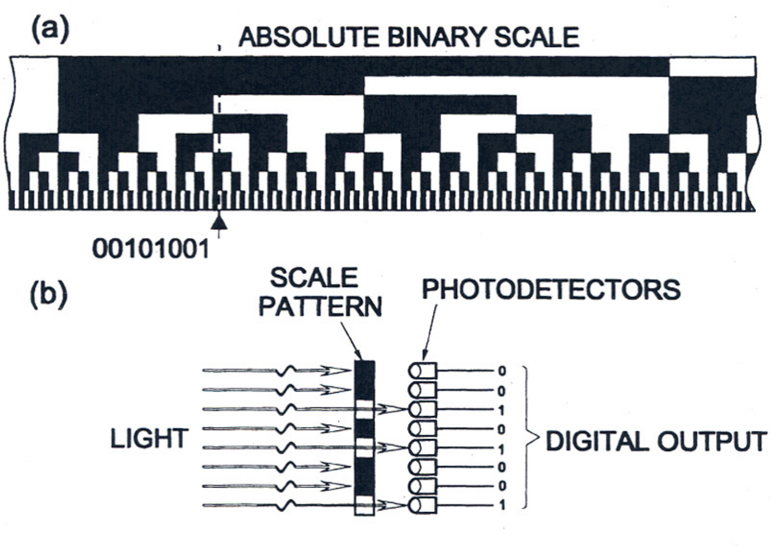


Figure 5.11 Principle of operation of an absolute encoder: (a) absolute linear scale, (b) detection of bits.

More sophisticated scanning methods are used in modern absolute position encoders. Two of them, the V-scan and the U-scan allow for reliable simultaneous bit transitions. In the V-scan method, the sensors are positioned in the V-shape arrangement in two sensor banks (Fig. 5.12). Such a distribution makes room for error tolerances in the encoder system. The less significant bit is used to define in which direction the scale is moving, i.e., what kind of transition is performed (high-low or low-high).

Another non-ambiguous method is the gray code, particularly well suited to optical encoders. In this *monostropic code* , only two neighboring position values differ in exactly one binary digit, i.e. only one track changes at a time. This limits any decision during edge transition to plus or minus one count. Therefore, the maximum error when moving from one position to the next is 1/4 of the grating period of the finest track.

The gray excess code consists of a section from the middle of the gray code pattern. This permits a position value other than 2 and yet remains a unit-distance code (monostropic). An example of the gray excess code is: 4-bits of gray code provide 16 absolute position values, and to solve 10 positions, the first and last 3 values are omitted from the graduation pattern to produce the *10-excess-3 gray code*. In the end, these codes (gray code, gray excess code, or any other appropriate code) are converted by the subsequent electronics (microprocessor) into

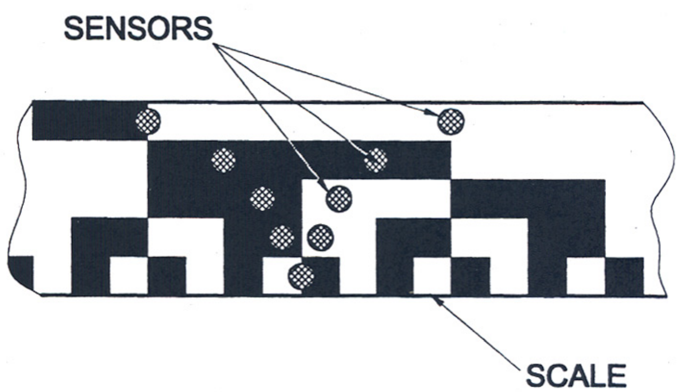


Figure 5.12 Arrangement of sensors in V-scan method.

Table 5.5 Decimal, binary, gray and gray excess codes

Decimal code	0	1	2	3	4	5	6	7
Binary code	0000	0001	0010	0011	0100	0101	0110	0111
Gray code	0000	0001	0011	0010	0110	0111	0101	0100
Gray excess	0010	0110	0111	0101	0100	1100	1101	1111

the NBC. Differences between the binary and gray codes are shown in Table 5.5.

In linear motion systems with single axes and travel distances over several meters, the application of absolute encoder scales has serious limitations in insufficient step resolution. For example, the scale with 12 tracks can generate 12-bit position information. This translates to 4096 unique encodings per scale length, providing approximately 250 μm resolution for 1 m of travel distance. In some cases it may be insufficient. An increase in the number of tracks can overcome this problem, but it results in higher complexity and costs of the encoder system. In practice, the total travel distance is subdivided into sections instead. Each section has the same absolute linear scale. To detect which section is actually scanned, the encoders use *distance-code* reference marks. The span between every second reference mark is constant, but it varies by a resolution step between two consecutive marks. The information as to which section is sampled comes from these reference mark signals. This method can reduce the search interval to 100 mm or less, instead of the entire scale length. In motion systems where the safety and failure free operation is not a priority, incremental instead of the absolute optical scales can be employed for sections working with distance-code reference marks. Absolute position is determined by counting the number of steps from the reference marks; however, the information can be lost in the case of power interruption.

Linear encoders used in the metal cutting industry, i.e. LBM drives for machine tool tables, must meet the following requirements [49]:

- high counting accuracy at high speeds, e.g. from 0.1 to 0.25 m/s in milling of gray-cast iron and aluminum (this translates into wide frequency range of position loop-control and therefore, fast feed- forward control);
- high acceleration capability, typically from 10 to 40 m/s^2 and even higher;
- high maximum rapid-travel speeds, typically from 1 to 1.5 m/s, sometimes even 2 m/s.

Manufacturers have developed two types of constructions which can meet these requirements: (a) exposed and (b) sealed encoders. Exposed encoders are recommended in clean environments without a danger to contaminate the optics. However, in machines either completely encapsulated or using coolant and/or lubricant, sealed encoders are preferred. The advantage of the sealed system lies in the reduction of requirements for finishing the mounting surface. Furthermore, sealed linear encoders are characterized by simple mounting and higher protection rating. On the other hand, the advantages of exposed encoders include higher traversing speed, no friction and better accuracy. Therefore, exposed encoders most often find applications in precision machines, measuring systems, and production equipment for the semiconductor industry. On the other hand, the sealed linear encoders are widely utilized in the metal cutting machines.

The *Heidenhain* LC 181 sealed absolute position encoder data is shown in Table 5.6. The LC 181 absolute encoder generates the absolute position value from seven incremental tracks. The grating periods of the tracks differ in a manner that makes it possible to evaluate the measuring signals of all seven tracks. This allows identification of any location on the scale within the measuring length of 3 m. In addition to the absolute position information, the LC 181 encoder provides sinusoidal incremental signal with its period of 16 μm at 1 V (peak-to-peak).

Table 5.6 Sealed absolute linear encoder LC 181 manufactured by *Heidenhain, GmbH*, Traunreut, Germany.

Parameter	Specifications
Measuring standard	DIADUR glass scale with 7 tracks with different grating periods
Data interface	Synchronous serial (EnDat)
Incremental signal	1 V (peak-to-peak) signal period 16 μm
Accuracy grades	±5 μm, ±3 μm
Measuring steps	1 μm, 0.1 μm
Measuring length	240 to 3040 mm
Length of sealed scale	Measuring length + 119 mm
Width of sealed scale	40 mm
Height of sealed scale	62.5 mm
Height of sealed scale and read head	85 mm

5.2 Linear Magnetic Encoders

5.2.1 Construction

As compared with optical sensors, their magnetic counterparts are characterized by simplicity, reduced sensitivity to contamination, robustness and low cost. Magnetic sensors can work in the presence of heavy liquid and chip build up. Made out of metal, they can withstand more severe vibrations, and are perceived to be more reliable. In addition, these devices have lower power requirements, good performance characteristics and are well suited for large volume manufacturing technology.

Magnetic encoders utilize *magneto-resistive* (MR) sensing elements and *magnetically salient targets* . The magnetically salient target is a long, alternatively magnetized ruler. The MR elements (sensors) change their resistance under the influence of the magnetic flux density and can sense flux densities above 0.005 T [121]. The principle of operation of the magnetic linear encoder with the MR sensors is explained in Fig. 5.13. The MR sensor resistance changes approximately ±1.6% as the magnetic field excited by the passing salient target changes its polarity. Four sensors are electrically connected to a resistive bridge polarized by 5 V d.c. source. The bridge output voltage varies sinusoidally within the amplitude of 0.08 V (peak-to-peak) reflecting changes of sensor resistances. The two magnetic poles affect the sensors in the same way but with opposite polarity. Therefore, when the alternatively magnetized ruler moves one pole pitch τ , the output signal will complete one cycle.

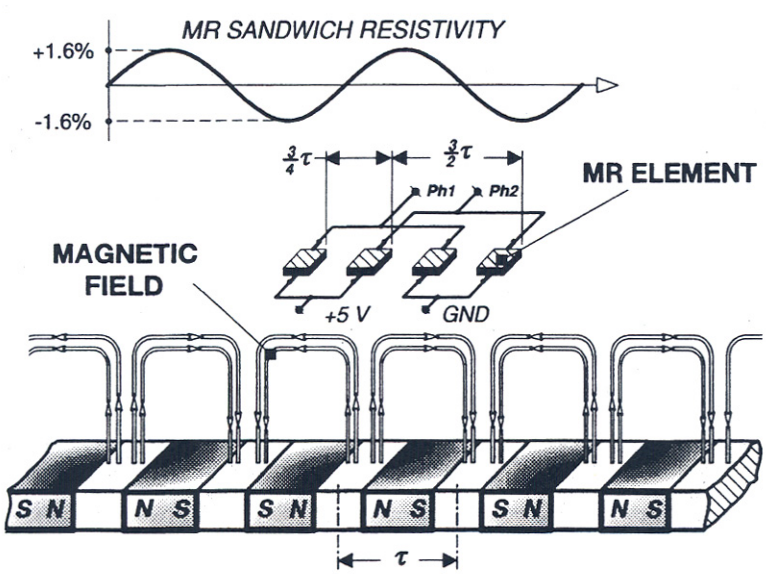


Figure 5.13 Magnetic encoder with magneto-resistive sensors.

Magnetic encoders employing MR sensors are capable of producing output signals with frequencies up to 200 kHz. High frequency response requires very high resolution, which is a function of the air gap size. The smaller the gap, the higher the resolution. This gap size should be approximately 80% of the pole pitch. For example, if a motion system requires the resolution of 0.05 mm (20 kHz frequency with 1 m/s linear speed), the encoder with 4x interpolation should contain the magnetic target with 0.2 mm pole pitch. The air gap in such a system is about 0.15 mm.

Sensing elements and magnetically salient targets can be supplied to the end user as separate components. Sometimes, the adaptation of motion system hardware to serve as a long salient target is possible. In such a system adapted by the user, air gaps between sensors and the target are usually of the order of a few millimeters.

In magnetic encoders with large air gaps the *Hall elements* are more suitable than MR sensors. These are true solid state devices with good operating temperature limits, typically from -40 to 150 °C, long life expectation (20 billion operations), and can work at zero speeds. The linear (analog) Hall element has a wide range of output signal (from 1.5 to 4.5 V) and a reasonable frequency response (100 kHz). Its output voltage is

Table 5.7 Specifications of the IHRM 12P15001 sensor employing magnetically biased Hall element manufactured by *BEl Corporation, Industrial Encoder Division*, Tustin, CA, U.S.A.

Parameter	Specifications
Voltage supply range	8 to 28 V d.c.
Supply current	20 mA
Max. switching current	100 mA
Max. switching frequency	20 kHz
Voltage drop	< 3 V d.c.
Air gap	2.5mm
Temperature range	-40 to 120°C
Temperature coefficient	-3%/K
Short circuit protection	Yes
Reverse polarity protection	Yes
Housing	Stainless steel
External dimensions	M12xl (thread) X 60 mm length

$$V_H = k_H \frac{1}{\delta} I_c B \sin\theta \tag{5.2}$$

where I_c is the applied current, $B\sin\theta$ represents the component of the magnetic flux density vector perpendicular to the current path, θ is the angle between the magnetic flux density vector and Hall element surface, δ is the thickness of Hall element and k_H is Hall constant (m^3/C).

Specifications of a typical Hall effect sensor are listed in Table 5.7. This sensor is used to scan moving electromagnetic objects, preferably toothed ferromagnetic racks [15].

Encoder systems with Hall effect devices are arranged differently than those comprising MR elements. Hall sensors are typically placed between a moving, magnetically salient target, e.g., ferromagnetic ruler with teeth, and a bias PM which excites the magnetic field. In the case of low resolution of positioning systems a long *flexible magnetic strip* distributed along the motion track serves as the magnetically salient target. This strip is made out of ferrite material or low energy NdFeB PMs mixed with rubber, and is usually alternatively magnetized, i.e., N, S,...N, S with pole pitch of a few millimeters. The alternatively magnetized flexible strip permits achieving the repeatability up to ±5 μm (1.22 μm resolution) with 4096 x multiplier (electronic circuit). The relative position is-determined by counting the number of poles or target saliencies (steel teeth) moving through the sensor, while the speed is obtained from the frequency at which they pass. Meanwhile, the movement

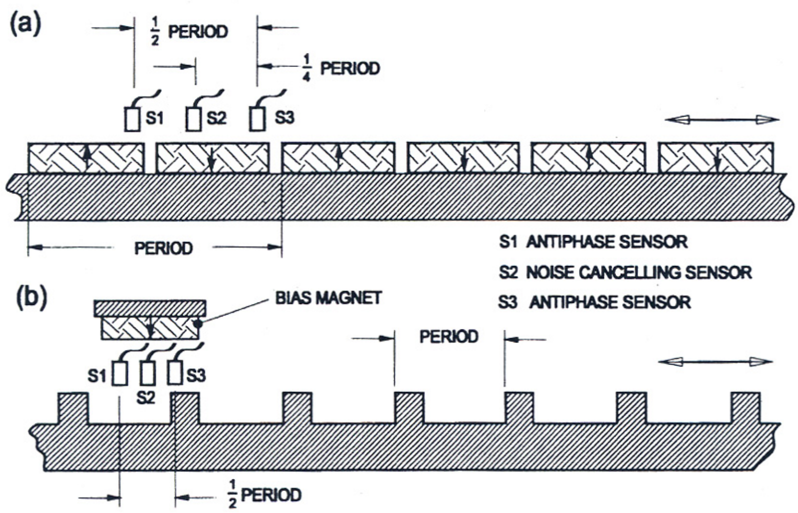


Figure 5.14 Noise canceling magnetic sensor array.

direction is obtained from the relative timing of two sensors in the quadrature with the target saliency. This flexible-strip-based linear encoder is used in LBMs with inner air-cored armature winding and moving PMs (Lighting Series) manufactured by *Anorad* [8].

In linear motors utilized for propulsion with an array of magnetic poles N, S,...N, S and short pole pitch, the installation of an additional magnetic strip for the encoder is not necessary. Encoder sensors are located near the surface of the guide away and the field produced by PMs is used as the magnetic target.

5.2.2 Noise Cancellation

One of disadvantages of linear magnetic encoders is their sensitivity to external magnetic fields and temperature changes. Sometimes, the *magnetic noise* can exceed the sensor generated signal up to one order of magnitude. Therefore, noise cancellation techniques aimed at suppression of unwanted disturbance signals are required.

One of simple noise cancellation methods is based on an array of three magnetic sensors [110]. Two of them are situated half of the magnetic period apart (in anti-phase relationship, i.e., 180° out-of-phase), while the third sensor is placed between the two remaining. The three-sensor array is shown in Fig. 5.14.

The individual sensor output signal is a function of the magnetic flux density created by passing magnetic targets and the ambient noise of magnetic origin. For magnetic poles of alternative polarity (Fig. 5.14), the three sensor output signals can be expressed as

$$s_1 = \sin\frac{\pi}{\tau}x + N \tag{5.3}$$

$$s_2 = \sin\left(\frac{\pi}{\tau}x + \frac{1}{2}\pi\right) + N \tag{5.4}$$

$$s_3 = \sin\left(\frac{\pi}{\tau}x + \pi\right) + N \tag{5.5}$$

where x is the pole or saliency position in the direction of motion and N is the noise signal. Quadrature positions s_{1c} and s_{2c} are derived from the three sensor signals s_1, s_2 and s_3 as follows

$$s_1 = s_1 - s_2 = \sqrt{2}\cos\left(\frac{\pi}{\tau}x + \frac{1}{4}\pi\right) \tag{5.6}$$

and

$$s_{2c} = s_3 - s_2 = \sqrt{2}\cos\left(\frac{\pi}{\tau}x + \frac{3}{4}\pi\right) \tag{5.7}$$

The sensor output quadrature signals s_{1c} and s_{2c} are digitized for the complete noise cancellation enhancement. In some applications, the noise N depends on the position of the magnetic pole within the strip along the reaction rail. This results in an incomplete noise cancellation. However, the digitization with zero-crossing detection occurring at points of geometrical symmetry will fully cancel the noise signals. Fig. 5.15 depicts the sensor output signal with resulting zero-crossing digitization.

5.2.3 Signal Interpolation Process

The interpolation is a process of an encoder signal subdivision into phase shifted copies. It can be applied to sinusoidal outputs in the quadrature only. For example, the TTL signals cannot be interpolated. To enhance the resolution effectiveness, i.e., the overall accuracy, the interpolated signals are recombinated in electronic circuitry.

The sinusoidal signals formed by the incremental encoders are processed by the digitizing electronic units. These are often incorporated into a numerical motion controller and enclosed in a separate housing. Three of the commonly used interpolation methods, i.e. (a) analog digital interpolation using a resistor networks, (b) digital interpolation with look-up and tracking counter, and (c) digital interpolation with, arc-tangent calculator, have successfully been applied to the *Heidenhain, GmbH* encoders.

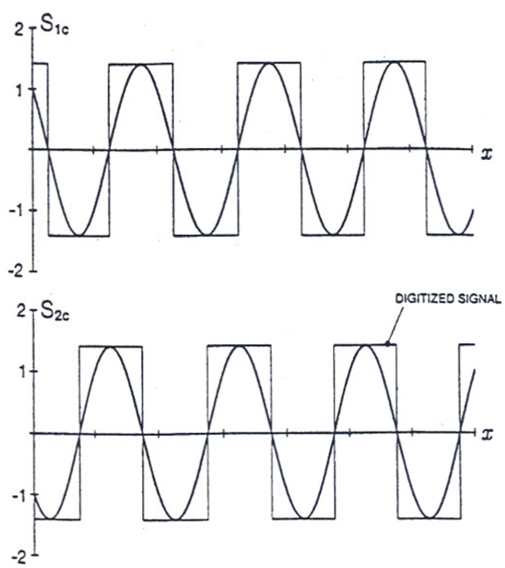


Figure 5.15 Quadrature sensor output signals S_{1c} and S_{2c} digitized by sensing zero-crossings.

The first method makes use of the trigonometric identity $\sin(\alpha+\beta)=\sin\alpha\cos\beta+\cos\alpha\sin\beta$ to develop the phase-shifted copies of the original signals.

The encoder LS 774/LS 774C manufactured by *Heidenhain, GmbH* is based on the analog-digital conversion with 5-fold interpolation (the so called $5\times$ interpolator). The scanning signals s_{1c} and s_{2c} are amplified and interpolated in the resistor network that generates collateral phase-shifted signals by using vector algebra. The 5-fold interpolation process is shown in Fig. 5.16. Ten signals are produced with a phase shift ranging from 0 to 162° electrical. After conversion to the quadrature, these signals are combined into two square-wave trains by exclusive-OR (XOR) gates. The trains of impulses have the frequency five times greater than that of the scanning input signals and are phase-shifted by the quarter of period. Each edge of the signals S_1 and S_2 can be used as a counting pulse within one period. The reference pulse S_0 is gated between the two successive edges of S_1 and S_2. The 20 μm grating period of the encoder, which combines 5-fold interpolation with 4-fold electronic evaluation is capable of achieving 1 μm measuring step. A similar process can be used for 10 or 25-fold interpolation that results in 1/40 or 1/100 measuring step of the grating period.

In interpolation processes utilizing higher subdivisions (50-fold and above), digital methods are required. Two scanning signals are first

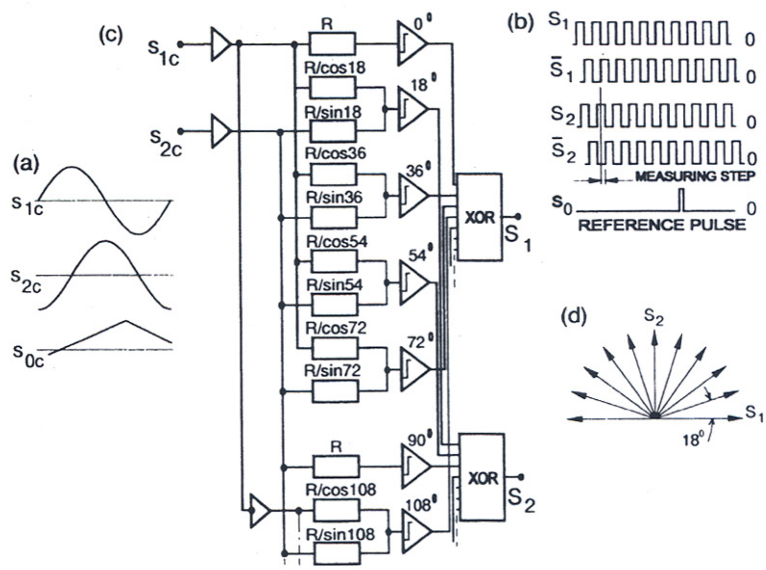


Figure 5.16 Interpolation process with resistor network: (a) scanning signals, (b) measuring signals after 5-fold interpolation and digitization, (c) electronic circuit, (d) phasor diagram.

amplified, then quantified in the sample-and-hold circuitry and finally digitized into regular intervals in the A/D converter. These digitized voltages define a single address (row and column) in a look-up table describing an instantaneous position (Fig. 5.l7a). The actual position is compared with the value determined in the previous cycle which is stored in a tracking counter. The tracking counter produces incremental square-wave signals (0° and 90°) from the differences between previous and current positions. The look-up table interpolation method is used in the EXE 650 (50-fold interpolation) and EXE 660 (100-fold interpolation) encoders manufactured by *Heidenhain, GmbH.*

The most advanced interpolators use microprocessor technology. Fig. 5.l7b illustrates the digital interpolation process employing an *arc-tangent* calculator. The microprocessor calculates the tangent S_1 / S_2 from two digitized input voltages. The corresponding angle value (*arc-tangent*) which indicates the position within one signal period is derived from the table stored in EPROM. The analog input signals s_{1c} and s_{2c} are simultaneously converted into quadrature waveforms and signal periods are determined. The actual position is derived from the evaluated period and the calculated angle. Finally, to compensate for the system errors, appropriate correction values are read from the table stored in the RAM. After the error correction, the digital signal is transmitted to the motion control unit.

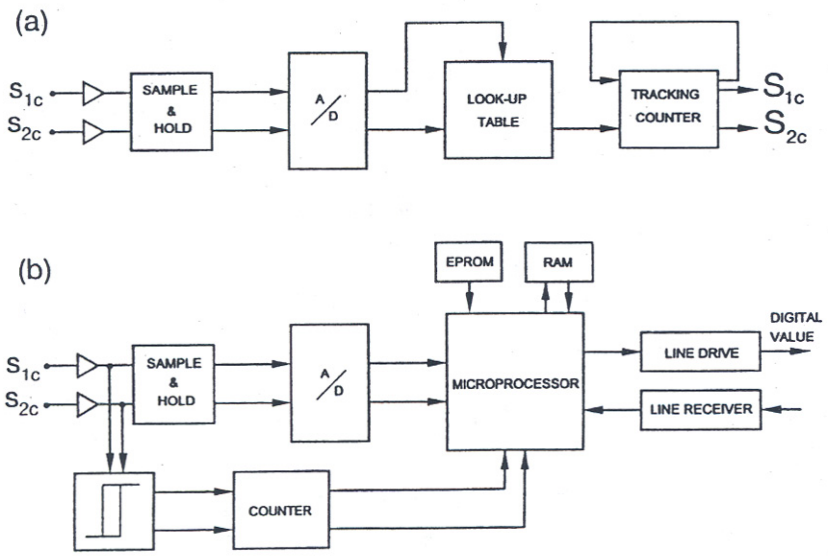


Figure 5.17 Digital interpolation methods: (a) using look-up table and tracking counter, (b) using microprocessor to compute arctan.

5.2.4 Transmission of Speed and Position Signals

The speed and position control systems are limited by the *pulse per meter* counts and the maximum linear *speed/frequency* response rate which depends on

- mechanically permissible traversing speed,
- minimum possible edge separation of the square-wave output signals S_1 and S_2,
- maximum input frequency of the interpolating and the digitizing electronics.

The maximum traversing speed

$$v_{tr} = \tau_p f \tag{5.8}$$

depends on the maximum input frequency f of the interpolation and digitizing electronics and the scale grating period τ_p .If f is in kHz, and τ_p is in μm, the speed v_{tr} is in mm/s. An example of the relationship between maximum traversing speed and the grating period at various maximum permissible input frequencies is illustrated in Fig. 5.18.

The encoder signal is subdivided in the subsequent electronics. The subdivision factor should remain in the reasonable proportion to the

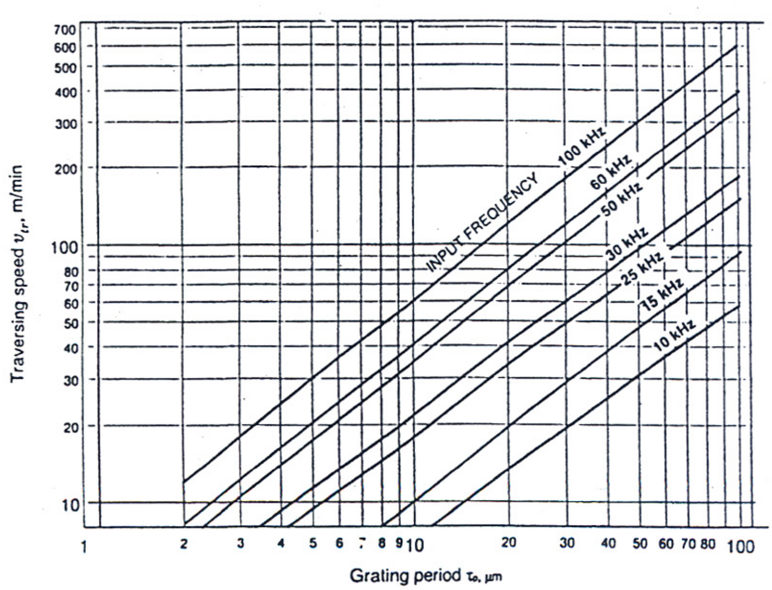


Figure 5.18 Maximum permissible input frequencies for EXE interpolation unit. Courtesy of *Heidenhain, GmbH*, Traunreut, Germany.

accuracy of the encoder. For example, the subdivision factor of 1024 applied to the 10 μm or 40 μm signal period gives the resolution of approximately 10 nm and 40 nm, respectively.

Feed drives of machine tools can reach linear speeds over 2 m/s, while the handling equipment - over 5 m/s. It can be calculated that the velocity of 1 m/s and measuring step of 0.1 μm (after a 4-fold evaluation) result in the input frequency of 2.5 MHz. Owing to large distances separating the encoder and the processing electronics (up to 50 m), the interpolating and digitizing circuit is often connected as a separate unit between them. For signals with frequencies above 1 MHz, short cables need to be employed, in order to preserve good transmission quality. Therefore, high speed motion systems utilize encoders containing interpolation and digitizing circuits. If the high frequency transmission signal is unavoidable, e.g., in the system with high traversing speed and small measuring steps, a linear encoder with sinusoidal output signals should be used. This sinusoidal signal should be 1 V (peak-to-peak) at the cutoff frequency of 200 kHz with amplitude of -3 dB. In this case, the cable length can reach 150 m.

Low traversing speed with high uniformity of motion requirement sets another limit for the measurement system. To maintain adequately uniform speed a resolution of 0.1 μm and higher may be required.

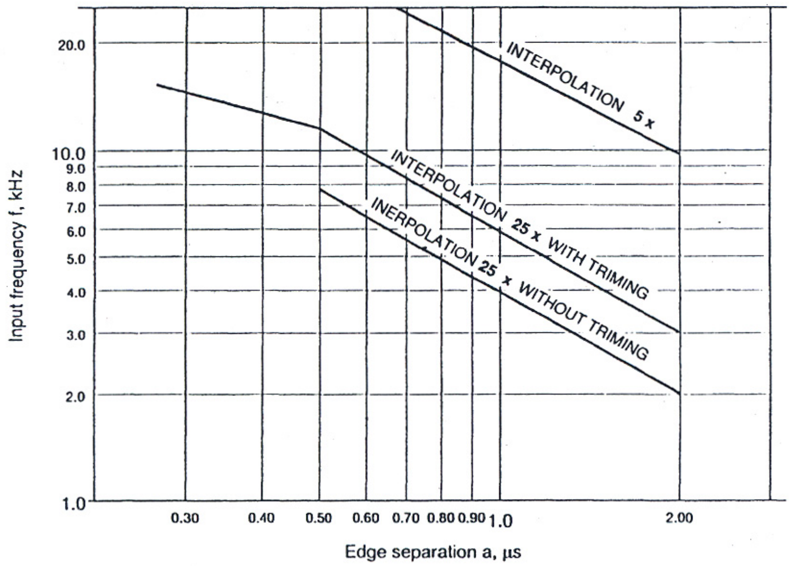


Figure 5.19 Edge separation diagram applied to non-clock EXE interpolation unit. Courtesy of *Heidenhain, GmbH,* Traunreut, Germany.

In general, the control electronics limits the minimum edge separation for square wave output signals. The relationship between input frequency f and the edge separation a for a given interpolation factor, is shown in Fig. 5.19. The input frequency f can be found from eqn (5.8).