<u>Dynamic Resolution for Optical Encoders</u>

By Robert M. Setbacken

An incremental optical rotary encoder consists of an illumination source, a static grating or mask, a rotating grating or code wheel, and a set of optical detectors. (Figure 1.) The measurement resolution is derived from the grating dimensions. The grating structures are created by deposition of an opaque material on the glass surfaces in a controlled and repeatable manner. The number of alternating clear and opaque patterns placed upon the perimeter of the rotating disk, and on the stationary mask, define the number of "cycles" the encoder will generate per revolution. When light is projected through the mask and disk, the intensity of light passing through the two structures will vary as the rotating disk moves with respect to the fixed mask. When the opaque portion of the disk is in line with the clear portion of the mask, a minimum amount of light will pass through the assembly. Conversely, when the clear sections of both pieces are aligned, a maximum amount of light will pass. This type of grating is referred to as an Amplitude grating, because it changes the amount of light transmitted as the disk rotates. This is different from a Diffraction grating, which modulates the light using interference principles, and which are usually constructed to use reflected light rather than transmitted light. In either case, the resulting light signal is modulated at the base resolution of the device, or Cycles per Revolution (CPR). Encoders using transmitted light detection can be readily manufactured with up to 10,000 CPR. Because motion of the disk simply generates output cycles, and nothing is known about the true absolute position, these encoders are termed incremental encoders.

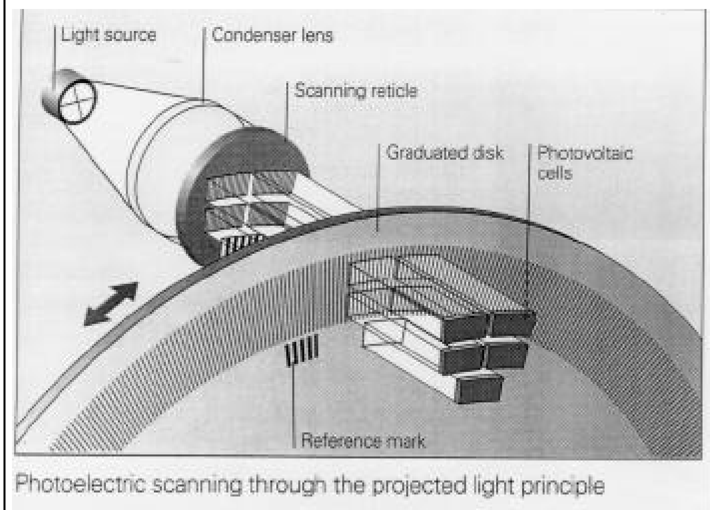

Figure 1

The light transmitted through the disk/grating in this fashion will have an intensity that is roughly sinusoidal. Some manufacturers amplify these sinusoidal signals and transmit them directly to the controller. These encoders are aptly termed sinusoidal output encoders. An alternative to the sinusoidal encoder is one that has TTL style outputs. Encoders of this type compare the light detector output with a threshold level via a comparator circuit. This allows a digital signal to be generated with a period equal to the cyclic fluctuation of the incident light. For either type of encoder, the output will usually consist of two channels "in quadrature". This means that two output signals in the base resolution are

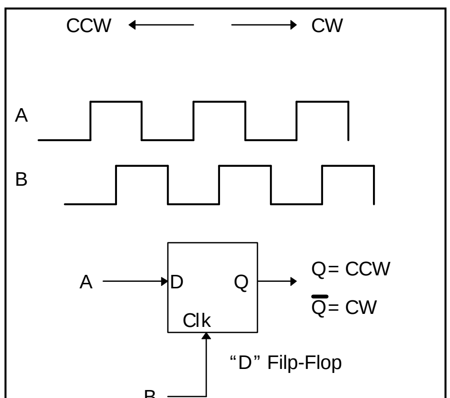

Figure 2

generated, but one is shifted 90° with respect to the other. By evaluating which signal is more positive with respect to the other, direction of rotation can be derived. And example of these outputs and the method for decoding direction is provided in Figure 2.

Optical Encoders are used in a wide variety of applications, from printers to machine tools. Generally, they are selected as the feedback device for servo systems requiring high levels of accuracy, but also can be the sensor of choice for very low cost applications. More and more often, they are also being chosen because of their ability to provide a wider dynamic range than available from competitive technologies. For example, the z-axis of a milling machine will perform high-speed bulk material removal, and it will also perform thread tapping operations. The difference in velocities for these two operations can exceed four orders of magnitude, (80db). Since most digital controllers would like to obtain much more than a single "count" per measurement period, low speed operation requires very high resolution from the feedback device. The higher the resolution, the more cycles can be detected during a measurement cycle, which gives the numerical algorithms more "bits" to work with. For example, assume a milling machine needs to operate between 1 rpm and 10,000 rpm It uses a controller with a 200 microsecond measurement cycle, and at least 10 counts per measurement period are desired. This would require a measurement resolution of 3,000,000 counts. If the encoder industry can only readily provide 10,000 count devices, how can this be realized?

The solution lies is a process called interpolation. There are many methods of interpolation, some are virtually free, and others are complex and expensive. The cost, of course, is directly related to the degree of interpolation produced. Methods of interpolation can be provided to some extent "for free" in a TTL encoder. In these encoders, the two quadrature signals generate four distinct quadrants in the cyclic output, and by monitoring the rising edges of these signals, four positions per cycle are distinctly identified. In this way, the user realizes a 4 times multiplication of the base device resolution. (Refer to Figure 2, there are two vertical transitions in each of the A and B signals every cycle, for a total of four "counts") So, an encoder with 10,000 transparent/opaque line-pairs spaced around the code wheel, would develop 40,000 "counts" per revolution if edge counting was used. Because virtually every TTL encoder ever made provides this feature, it is widely used and in fact leads to confusion. A 2000 cycle encoder is sometime referred to as a 2000 count encoder, which could be provided by a 500 cycle encoder if edge counting was assumed. The buyer is warned to be careful when ordering!

Real interpolation can not really be done with a TTL encoder, it must use a sinusoidal device. With sinusoidal outputs in quadrature, each output cycle can be resolved into finer and finer positions using a number of methods. The most obvious is to use simple trigonometry. Each electrical cycle is a sinusoidal signal, and each signal passes through 360° electrical degrees per cycle. Since a sin/cos output is available because of the quadrature output relationship, the $\tan^{-1}$ function can be used to determine the angular position *within a cycle* of the output. Depending upon the accuracy of the system implementation, each cycle can be subdivided substantially. There are numerous examples of applications where theses types of signals have been successfully interpolated by a factor of 1000. If the basic encoder had a 3000 CPR capability, a 1000-fold interpolation can produce the 3,000,000 counts per revolution required by the previous example. Additionally, when the encoder is operated at high speed, the interpolation can be bypassed, and the natural encoder output utilized. This allows the signal transmission frequencies to be kept low, an minimizes transmission difficulties when long cables are required.

With DSP solutions, the costs for production of a system demonstrating this level of performance are certainly coming down. However, they are still not insignificant. In addition, there may be times when only 5 or 10-fold interpolation would be desired, and a low cost solution is mandatory. For these types of applications, the sinusoidal signals can be divided using simple resistor bridges and combined using exclusive-OR logic. Interpolation of this type can be realized for only a few dollars, but allows a 10,000 cycle encoder to provide 400,000 counts per revolution! (The output of an interpolation circuit is TTL Quadrature.)

One of the characteristics of all of these methods is that they employ a fixed interpolation rate once in operation. In most cases, this is accomplished by a dedicated circuit design. In some cases, the interpolation factor can be "programmed" by the user at time of manufacture by selection of a set of fixed connections and components when the device is incorporated into the encoder circuitry. In this second case, the multiplication factor is usually fixed at power-up of the device, and is not changed during operation.

Another characteristic is that, even if the interpolation rate is allowed to change, this change can not be accomplished without loss of information, or "missing a few counts". For most cases, the interpolated outputs are generally used simply for speed control, and so missing a few position counts can easily be filtered out. But what if it mattered? What if the system needed to keep track of true position, and to achieve extended resolution? For these applications Dynamic Resolution Enhancement is available. This method provides with system with the ability to change the interpolation factor during operation without loss of information.

An example circuit has been designed to provide 1X, 2X, 5X and 10X interpolation factors. See Figure 3. Two control inputs allow the user to select any of the four possible interpolation rates, (RES_1 and RES_2). These inputs can change freely during operation of the device, but must be held constant at a point approximately 45° prior to the allowed switching point in order to assure stable values to be supplied to the control electronics. In order to assure this setup time is provided, an output signal HOLD is provided. It is assumed this will be used to latch the resolution values into a buffer external to the device, so that they will not change during the setup period. A second output pin, ACK is used to acknowledge that a change in the commanded resolution has been detected by the device. ACK will be held in the acknowledge state until the resolution change has been accomplished. At this time, the acknowledge output will return to the clear state. The ACK signal allows the user to synchronize his position information with the new output scale factor so that position information is not corrupted. The input and output signal relationships are shown in Figure 4. Examples of various resolutions with respect to the base resolution are shown in Figure 5.

Dynamic Resolution Enhancement provides system designers with a flexible and low cost means of providing significant performance improvements at a very low installed cost. Resolution can be increased an order of magnitude, and the overhead can require as little as one wire to implement the control input. For system designers who are trying to squeeze a bit more performance out of an existing system, or who do not have the resources to develop DSP or ASIC based interpolation interfaces, this can be a quick, and effective solution to providing performance improvements.

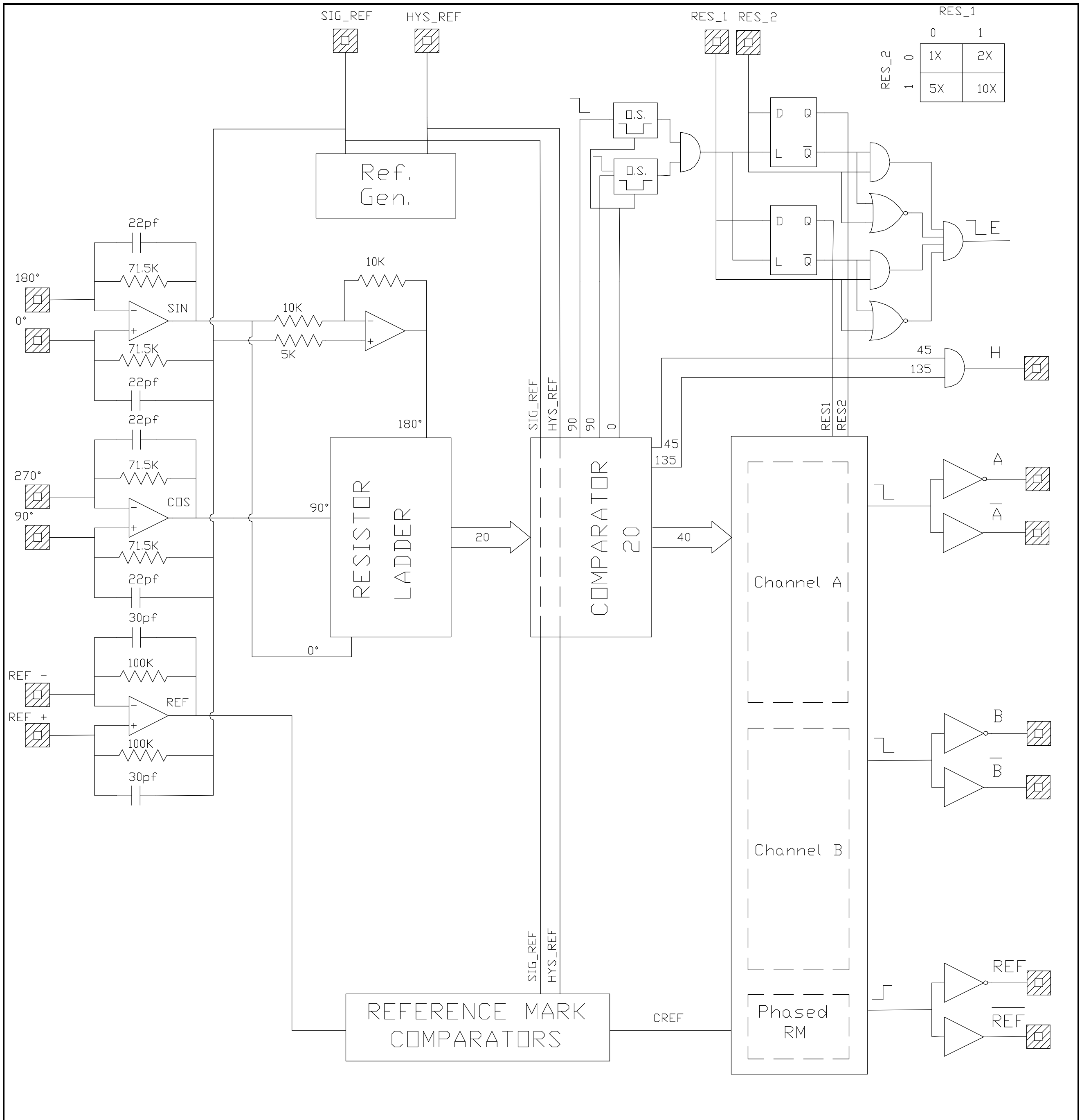

Figure 3 - Interpolation Circuitry

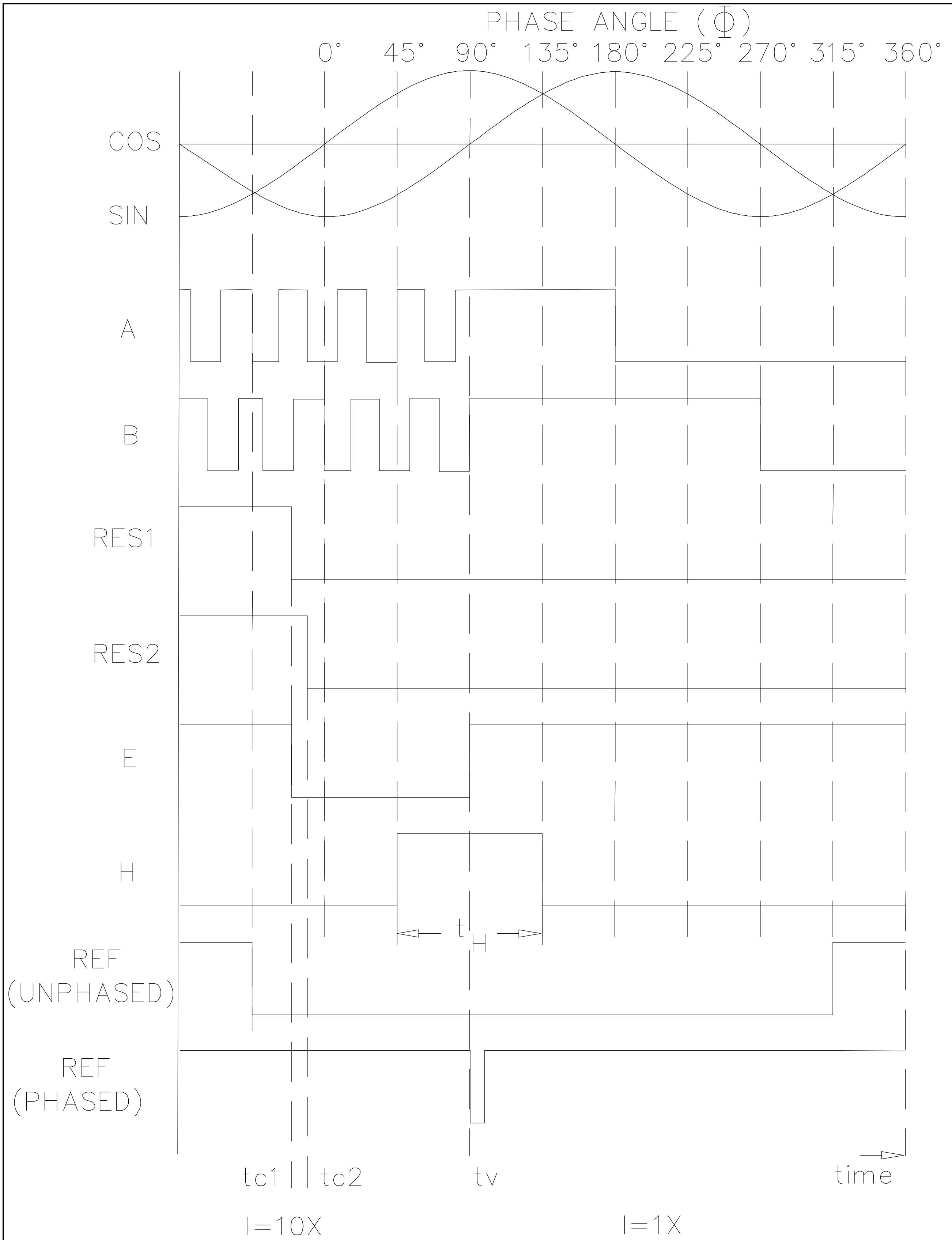
PHASE ANGLE (Φ)
0° 45° 90° 135° 180° 225° 270° 315° 360°
COS
SIN
A
B
RES1
RES2
E
H
t H
REF (UNPHASED)
REF (PHASED)
tc1
tc2
tv
time
I=10X
I=1X

Figure 4

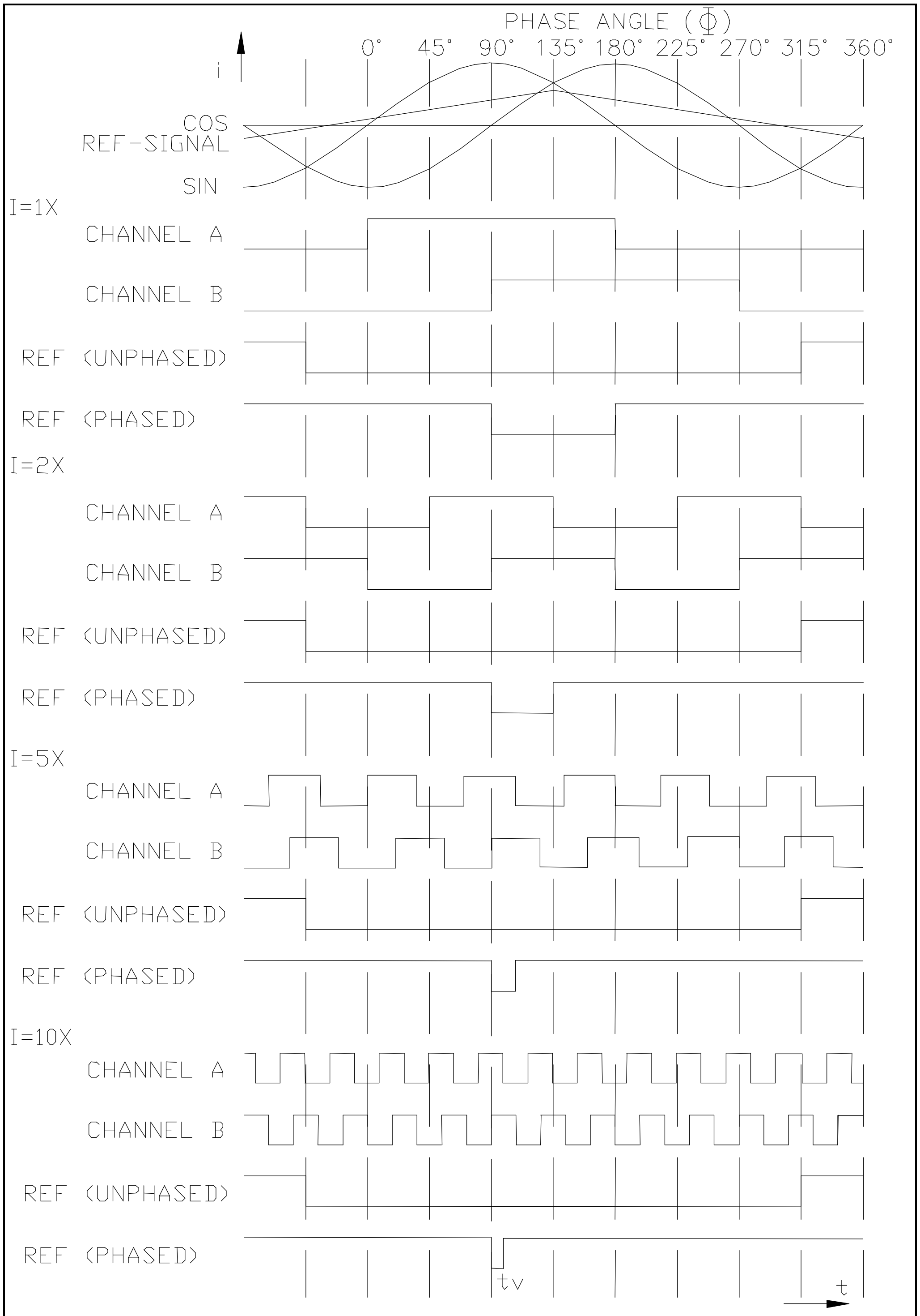
PHASE ANGLE (Φ)
0° 45° 90° 135° 180° 225° 270° 315° 360°
i
COS
REF-SIGNAL
SIN
I=1X
CHANNEL A
CHANNEL B
REF (UNPHASED)
REF (PHASED)
I=2X
CHANNEL A
CHANNEL B
REF (UNPHASED)
REF (PHASED)
I=5X
CHANNEL A
CHANNEL B
REF (UNPHASED)
REF (PHASED)
I=10X
CHANNEL A
CHANNEL B
REF (UNPHASED)
REF (PHASED)
tv
t

Figure 5