INSTITUTE OF PHYSICS PUBLISHING METROLOGIA

Metrologia **41** (2004) 17–32 PII: S0026-1394(04)70012-2

Review of methods for time interval measurements with picosecond resolution

Józef Kalisz

Military University of Technology, Kaliskiego 2, 00-908 Warsaw, Poland

E-mail: jkalisz@wel.wat.edu.pl

Received 29 July 2003
Published 10 December 2003
Online at stacks.iop.org/Met/41/17 (DOI: 10.1088/0026-1394/41/1/004)

Abstract
This paper is a review of methods and techniques used for precise measurement of time intervals (TIs) or precise conversion of TIs to digital data. The following methods are described: the counter method and averaging, time stretching, time-to-amplitude conversion followed by analogue-to-digital conversion, the Vernier method, conversion utilizing tapped delay lines, and interpolation methods. Special attention has been paid to converters utilizing integrated delay lines for digital conversion of TIs, including designs with phase-locked loop and delay-locked loop circuits. This review is illustrated by design examples and contains a comprehensive list of references.

1. Introduction

Precise measurements of *time intervals* (TIs) between two or more physical events are frequently needed in many applications in science and industry. In a simple case shown in figure 1, the time interval T is measured between the leading edges of two electrical pulses applied to the inputs START and STOP of the *time-interval meter* (TIM). The pulses may be generated by the *time discriminators* used to extract the timing information from the pulses received from the detectors of some physical events, for example, light flashes. Time discrimination involves the use of advanced methods and electronic circuits to produce the pulses precisely timed relative to the related events. The definition of the 'points' on the time axis to measure the TI between them becomes a challenging issue, hard to resolve in applications demanding the highest accuracy.

It is easy to draw TIM input pulses of zero rise time. However, real pulses always have a finite slope of the leading edge, and usually the timing points are referred to the instants when the pulse edges cross an arbitrarily defined threshold level.

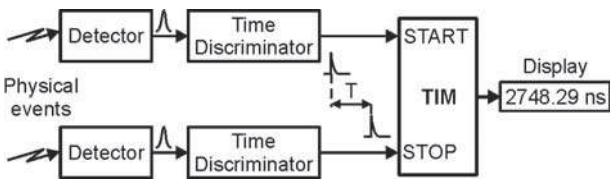

Figure 1. Principle of TI measurement.

The TIM performs conversion of a time interval T into a digital (binary) word, frequently displayed in the decimal form. Therefore a TIM is also called a *time-to-digital converter* (TDC). Originally this name referred to *non-interpolating* TIMs with a short measuring range, usually not longer than 100 ns to 200 ns. *Interpolating* TIMs, with a longer range (reaching tenths of seconds), are frequently called *time counters* (TCs).

It should be noted that the above classification with regard to TDC and TC is not obligatory in the timing community, and quite frequently a TC is called a TDC. The name 'time digitizer' is also used in both cases.

Figure 1 shows an exemplary TIM with two separate inputs, START and STOP. The TIM can also (or only) have a single COMMON input. Then the START pulse and the subsequent STOP pulse(s) should be generated on the same wire. A useful feature of that mode of operation is virtually zero offset error, but the shortest measured TI is limited. It has to be longer than the TIM dead time or the input pulse width (whichever is greater).

In the real world all repetitive 'constant' TIs have some dispersion (time jitter) caused by the measured physical phenomena and also introduced by the time discriminators used. The TIM also contributes some jitter (due to the inherent noise) and a measurement uncertainty caused by the non-linearity of conversion and the quantization process. The total statistical variation of measured TIs is usually calculated from

 © 2004 BIPM and IOP Publishing Ltd Printed in the UK

the collected data sample as an estimator, s, of the standard deviation, σ (sigma, root-mean-square (rms)), following the ISO Guide called 'standard uncertainty', and also called 'random error' or 'precision'.

What does 'precise' measurement of TI mean? We can assume that this feature can be attributed to such measurements whose standard uncertainty, s, due to the TIM is less than 1 ns (rms). In the best instruments, designed with the use of advanced methods and modern technologies, the lowest value of s is between 3 ps and 10 ps, and in interpolating TIMs it is typically about 20 ps. Values between 50 ps and 500 ps are typical for the instruments employing fully digital processing methods, i.e. without the use of any intermediate analogue processing (like time-to-amplitude (T/A) conversion or time stretching). Due to the rapid development of new methods and growth of technology, the new integrated digital TDCs can achieve $s < 50$ ps.

TIMs are used in science research (experiments in nuclear physics and astronomy), industry (dynamic testing of integrated circuits and hard drives), telecommunications (evaluation of high-speed data transfer), geodesy, and military equipment (in laser ranging systems).

The basic and most important technical parameters of TIMs are:

- measurement range (MR),
- standard measurement uncertainty or random error or precision (s),
- non-linearity of the time-to-digital conversion: differential (DNL) and integral (INL),
- quantization step (q) or least significant bit (LSB) or (incremental) resolution (r),
- dead time (T_d) or the shortest TI between the end of a measurement and the start of the next one,
- readout speed, important when the measurements are performed continuously at a high rate and with readout 'on the fly'.

In this paper I present an overview of some representative methods and techniques used for precise measurements of TIs. The main assumption is that the measured TI is defined by the specific time points on the edges of the related START and STOP pulses at the inputs of the TDC or TC.

Methods based on digital signal processing (DSP), commonly used in sampling oscilloscopes, are generally beyond the scope of this paper. DSP methods are based on precise sampling and memorizing the signal waveforms to calculate the TI between specific time points. Advanced sampling oscilloscopes and dedicated instruments are powerful (and very costly) measuring tools, which can also be used for precise TI measurements. *Real-time sampling* oscilloscopes are optimized for acquisitions of single-shot events (but not only these), while *random sampling* and *sequential sampling* oscilloscopes can be used only for visualization and processing of repetitive signals. The jitter floor can be as low as 1 ps to 3 ps. Modern sampling oscilloscopes provide high comfort of operation and comprehensive mathematical processing. Some oscilloscopes can also be used to advantage as high-quality time discriminators (figure 1). For example, after acquisition of a detector pulse, the accurate time position of the centre of gravity of such a pulse can be computed to determine the START (or STOP) instant.

The related database of publications is quite large and thus, unfortunately, this review is not complete due to the lack of space (many valuable contributions have been omitted). The classic methods used for measurement of TIs were earlier presented in a comprehensive review [1].

Many related publications may be found on the Web, for example with the aid of the search engines at www.scirus.com and odysseus.ieee.org/ieeesearch, and also in the database of the United States Patent Office (www.uspto.gov).

In the following text, in descriptions of the digital circuits the *low* (L) and *high* (H) logical levels have been assumed as equivalent to the logical states '0' and '1', respectively (positive logic).

2. Measurement methods

2.1. 'Coarse' counting

The simplest method of measuring TIs involves the use of a counter (figure 2(*a*)) which is driven by the reference clock, CP, of frequency f_0 or period $T_0 = 1/f_0$. The resolution (LSB) also equals T_0. After initial reset R the EN pulse enables the counter for the duration T to obtain the measurement result $T_p = nT_0$, where n is the decimal equivalent of the integer binary number Q read at the counter output.

When using the counter method, we assume that the measured intervals T are *asynchronous* with regard to the clock. It means that neither the origin nor the end of the measured interval is correlated in time with the clock pulses. In other words, there is a uniform probability distribution of the TI between the active edge of the clock pulse and the origin (and the end) of the measured time interval T. In such a case the maximum quantization error of a single measurement may

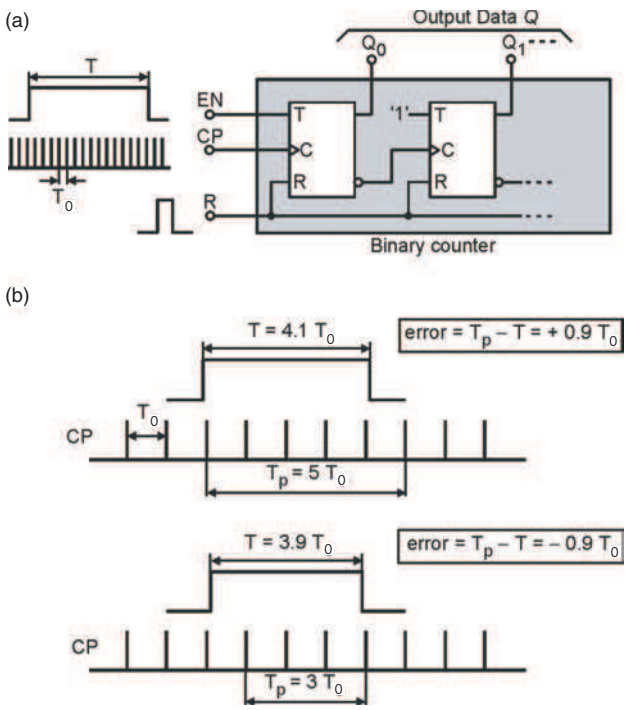

Figure 2. Counter as a simple TDC: (*a*) counting principle, (*b*) counting errors (T_p is the result of counting).

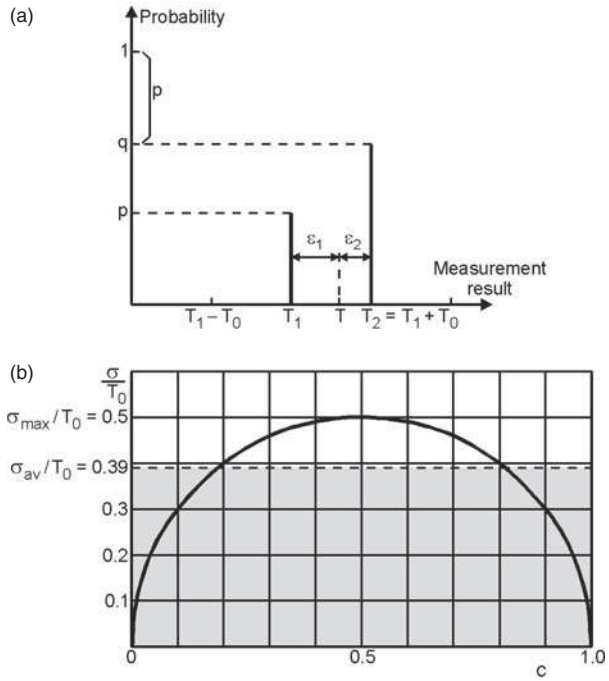

Figure 3. Quantization error inherent in the counter method: (*a*) example of a negative (ε_1) and positive (ε_2) error appearing in measurements of a constant and asynchronous time interval T, (*b*) standard deviation of measurements shown as a function of c, the fractional part of the quotient T/T_0.

reach almost $\pm T_0$, depending on the true value of the interval T and its time location with regard to the clock (figure 2(*b*)).

When measuring a series of a constant and asynchronous interval T, one obtains two results, $T_1 < T$ and $T_2 = T_1 + T_0 > T$ (figure 3(*a*)). The probability of each reading depends on the fractional part $c = \mathrm{Frc}(T/T_0)$:

$$p(T_1) = 1 - c \tag{1a}$$

and

$$q(T_2) = c. \tag{1b}$$

The measured TI is

$$T = pT_1 + qT_2 \tag{2}$$

and the quantization error is expressed by two values, $\varepsilon_1 = T_1 - T < 0$ and $\varepsilon_2 = T_2 - T > 0$.

The random error due to quantization can be expressed by the standard deviation of the related binomial probability distribution:

$$\sigma = T_0\sqrt{pq} = T_0\sqrt{c(1-c)}. \tag{3}$$

The half-circle shown in figure 3(*b*) is the plot of the normalized standard deviation $\sigma/T_0 = \sqrt{c(1-c)}$. The maximum value $\sigma_{\max} = T_0/2$ is obtained at $c = p = q = 0.5$. The average variance, $(\sigma^2)_{\mathrm{av}}$, can be calculated as the integral of the function $\sigma^2(c)$ within the bounds $0 \leqslant c \leqslant 1$. It yields $(\sigma^2)_{\mathrm{av}} = T_0^2/6 \cong 0.17T_0^2$. The *average standard deviation* can be calculated by integration of (3) or written directly by using the known formula for area of the half-circle of unity diameter

$$\sigma_{\mathrm{av}} = \frac{\pi T_0}{8} \cong 0.39T_0. \tag{4}$$

The accuracy of counter measurements can be improved by taking a series of measurements of the same interval T and *averaging* the results [2]. For a given measurement sample of size N, the results T_1 and T_2 are obtained with the numbers N_1 and N_2, respectively. Since $N_1 + N_2 = N$, the corresponding probabilities can be approximated by $p \approx N_1/N$ and $q \approx N_2/N$. The averaged result is

$$T_N = \frac{T_1N_1 + T_2N_2}{N}. \tag{5}$$

If N is large enough, then $T_N \approx T$ and the average quantization error $T_N - T$ approaches zero. The random error of T_N is lowered by $\sqrt{N}$ as compared with single-shot measurements, and the related maximum and average standard deviations are

$$\sigma_{N\,\max} = \frac{T_0}{2\sqrt{N}} \tag{6}$$

and

$$\sigma_{N\,\mathrm{av}} = \frac{\pi T_0}{8\sqrt{N}} \cong 0.39\frac{T_0}{\sqrt{N}}. \tag{7}$$

Thus at $N = 100$ the spread of the averaged result is ten times lower than that at $N = 1$. A disadvantage of the averaging method is the long time needed to take many measurements.

A substantial advantage of the counter method is the long MR (up to hundreds of seconds), which can be achieved in a relatively simple circuitry, because every additional flip-flop (FF) in the binary counter multiplies MR by two. Such *coarse* counters are also used in precise, interpolating TDCs, described in section 2.3.

A practical limitation of the counter method is the low single-shot resolution, which equals only 1 ns at a 1 GHz clock. Such a design requires a stable 1 GHz clock generator and a very fast counter, which are rather expensive devices.

The counter can be designed as a simple ripple counter (figure 2(*a*)), but usually it is synchronous, with the maximum cycle length 2^m, where m is the number of bits (FFs) of the counter. A very simple and fast synchronous counter can be obtained in the structure of the *linear feedback shift register* [3, 4], because it may require only a single XOR gate as a feedback. A disadvantage is the pseudo-random output code, which has to be converted separately to the natural binary or BCD code. In the counter design *gating of the clock pulses should be avoided* because this could cause an additional erroneous count in some cases [2].

The counter can also be used for TI measurement in the *free-running mode*. It means that in the START instant the current state of the counter is *sampled and read on the fly*, and the same operation is also performed in the STOP instant. The number of counts needed to calculate TI is determined taking into account the possible overflow of the counter or even the number of overflows (which would require a separate overflow counter). That kind of operation is preferred when a *multistop* or *multichannel* mode of measurements is needed. To achieve a low error (not greater than one LSB) during readout on the fly of fast synchronous counters, the Gray code is frequently used.

2.2. 'Fine' measurement methods

In this section the basic (non-interpolating) methods that utilize TDCs of a short MR (usually between 10 ns and 200 ns) and

have a much better accuracy than the 'coarse' counters are presented. The accuracy of those TDCs is determined mainly by the non-linearity (DNL and INL, see section 3.1) of the time-to-digital conversion, as in the commonly used analogue-to-digital converters (ADCs). As a rule, such TDCs are designed to obtain $\mathrm{INL}_{\max} < \mathrm{LSB}$. When measuring repetitively a *constant* time interval T, the observed random error, s, is mainly caused by the time jitter inherent in the electronic circuits used. The value of s may be quoted for a given T (usually $s < 10$ ps) or it may be presented as $s_{\max}$ within the whole MR.

These methods can be roughly classified as 'analogue' (A) and 'digital' (D). The most popular are

(a) TI stretching (A) followed by the counter method (D),
(b) double conversion: time-to-amplitude (A) followed by standard analogue-to-digital (A/D) conversion,
(c) the Vernier method with two startable oscillators (D),
(d) time-to-digital conversion utilizing the tapped delay line (D),
(e) the Vernier method with a 'differential' delay line comprising two tapped delay lines (D).

The methods listed above are utilized in two ways. In the first one a method is applied *without* an additional 'coarse' real-time counter and the designed TDC has a reasonably short MR. In the second one a method is applied *with* such a coarse counter, following the *interpolation* principle (section 2.3). The relevant instrument is frequently called a TC. The MR of the TC can be much longer (e.g. 40 s). Such a TC contains the coarse binary counter and a single or two TDCs of a short MR and high resolution to enhance the measurement accuracy.

An exception to this rule is the Vernier method (c), which inherently utilizes counters and in both solutions allows for a long MR.

In general, the 'digital' methods are preferred because the classic 'analogue' methods are difficult to implement in the integrated circuit technology, are more sensitive to the ambient temperature, are more susceptible to external disturbances, and have a longer conversion time.

The classic method of *time stretching* (figure 4) to obtain dual-slope conversion was already introduced in the era of vacuum tubes [5]. The time stretcher performs like a voltage amplifier, and sometimes is even called a 'time amplifier'. In the steady state the diode, D, conducts current $I_2 \ll I_1$. During the measured interval T, the capacitor, C, is charged with a constant current $(I_1 - I_2)$ and then discharged with a much smaller current, I_2. The *stretching factor* is defined as $K = (I_1 - I_2)/I_2$. The discharging time is stretched proportionally: $T_r = TK$. The total time $(T + T_r)$ is detected by a fast comparator and measured by a simple counter that provides an effective resolution $\mathrm{LSB} = T_0/(K + 1)$. Ignoring the quantization and linearity errors, when the count number is n, the measurement result is $nT_0/(K + 1)$.

It is clear that the method involves dual conversion: time/time/digital. This method was used, among others, in nuclear physics experiments [6], in precision laser ranging systems for space applications [7, 8] and for testing the dynamic parameters of CMOS digital circuits [9]. The time stretchers in these applications were built as low-cost, discrete circuits. An integrated TDC of this type was designed in BiCMOS technology [10].

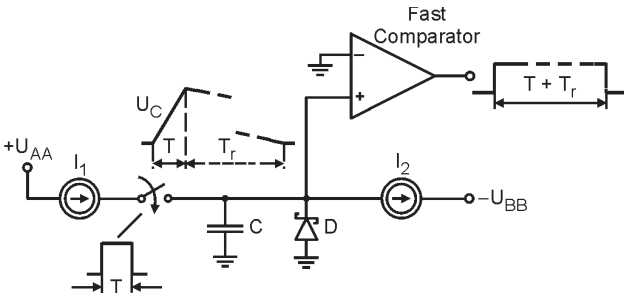

Figure 4. Linear stretching of the measured time interval T for subsequent counting.

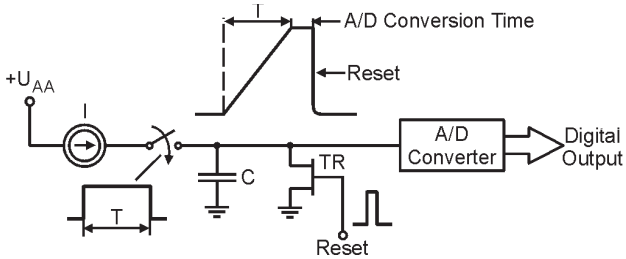

Figure 5. Conversion of TI to amplitude followed by a typical A/D conversion.

The best resolution obtainable with this method is about 10 ps. Considerable improvement became possible with the use of the *two-stage time stretching method* [11, 12]. In this approach, at $T_0 = 10$ ns ($f_0 = 100$ MHz) and $K = 10^4$ a single-shot resolution of 1 ps was obtained [11, 12]. However, the jitter level was about 5 ps and the linearity error about 10 ps. Hence the main advantage of the very high resolution (of very low value) is a small quantization error, which may be neglected.

A disadvantage of the time stretching method is the long conversion time, equal to TK, which limits the maximum frequency of measurements. A considerable shortening of this time has been possible with the use of the *two-fold interpolation* method [13, 14] and the *multiple interpolation* method [15], which can be implemented with a reasonably simple circuitry. Those methods are used only in interpolating TCs (section 2.3).

In another commonly used method the measured TI is first converted to a voltage (amplitude) by charging a capacitor with a constant current, and then the voltage is held briefly to allow its conversion to digital form by a typical, integrated A/D converter (figure 5). After conversion the capacitor is rapidly discharged to reduce dead time. Thus the conversion time in this method is equal to that of the A/D converter used. The method has been used with success in many designs [1, 16–20] and also in a commercial counter SR620 (SRS). Using modern, high-resolution, integrated A/C converters this allows us to achieve a high resolution in TI measurements. In practice an LSB value of 1 ps to 20 ps is readily achieved.

The above two methods are based on analogue processing of the measured TI. The first truly digital time conversion method has become the Vernier method (Pierre Vernier, 1584–1638, inventor of the popular Vernier caliper[1]), which

[1] http://www-history.mcs.st-andrews.ac.uk/history/Mathematicians/Vernier.html.

actually is a method of digital time stretching [21, 22]. In the basic configuration of the Vernier converter (figure 6), two *startable* oscillators (SG1 and SG2) generate signals of frequencies $f_1 = 1/T_1$ and $f_2 = 1/T_2$ differing only slightly. The incremental resolution is $r = T_1 - T_2$. The start of the waveform obtained at the output of each generator is synchronous with the active edge of the related input signal (START and STOP). The conversion is completed when coincidence of the active edges of the pulses produced by the generators is detected by the coincidence circuit (CC). Then the respective counters, CTR1 and CTR2, store the numbers n_1 and n_2. When the quantization error is ignored, the measurement result is

$$T = (n_1 - 1)T_1 - (n_2 - 1)T_2 = (n_1 - n_2)T_1 + (n_2 - 1)r. \tag{8}$$

When $T < T_1$, then $n_1 = n_2$ and $T = (n_2 - 1)r$. The use of a single counter, CTR2, is then sufficient. The longest conversion time is $n_{2\max}T_2 = T_1T_2/r$. For example, when $T_1 = 10$ ns and $T_2 = 9.9$ ns ($r = 100$ ps) that time is 990 ns.

The Vernier method described allows us to obtain a resolution below 100 ps. It was shown [23, 24] that in an improved design the resolution value can be obtained as low as 1 ps.

To achieve good accuracy of measurements utilizing the Vernier method, the startable oscillators should have high accuracy and stability, which poses a hard design challenge, especially at long TIs. Therefore the *dual interpolation* method with two Vernier converters is preferred (section 2.3).

A conceptually simple method of TI measurement is based on the use of the tapped delay line. The line is composed of a number of delay cells, each having the same (in an ideal case) propagation delay τ. The TI measurement is accomplished by sampling the state of the line during propagation of an initial (START) pulse. First conventional coaxial cables were used for this purpose, but following continued growth in semiconductor technology, new methods have been developed, based on integrated delay lines [25–35]. The first inventions in this field were filed in the early 1980s [25, 26]. The new integrated TDCs [46–49, 52–66] utilize the delay lines within the phase-locked loop (PLL) or delay-locked loop (DLL) circuits to achieve high stability and inherent calibration (section 2.3).

The tapped delay lines can be used in different configurations (figure 7). In the simplest one (*a*), the delay line is created by a train of N cells containing latch FFs, which are initially transparent (STOP = H) and reset (because START = L). The rising edge of the START pulse propagates through consecutive latches having the propagation delay τ until the falling edge of the STOP pulse appears, which latches the state of all FFs (samples the current state of the line) and stops the propagation. The measured TI is the sum of propagation times of all FFs that store the state H, or $T = k\tau$, where k is the highest position of the FF storing the state Q = H. The output data are obtained in the thermometer code, which should be converted to the natural or BCD binary code as needed by the application.

The delay line can also be created as a train of buffers, each having the delay τ. In the scheme shown in figure 7(*b*) the state of the line is sampled (by the rising edge of the STOP pulse) and held (in the edge-triggered D flip-flops FF1, . . . , FFN). The measurement result is determined by the highest position of the FF storing the H state. That method has been used in the commercial frequency and time interval analyser HP5371A [29] to obtain a 200 ps resolution.

If in that configuration the FF inputs of the clock (C) and data (D) are interchanged, we obtain the circuit shown in figure 7(*c*). Here the line operates like a *multiphase clock*

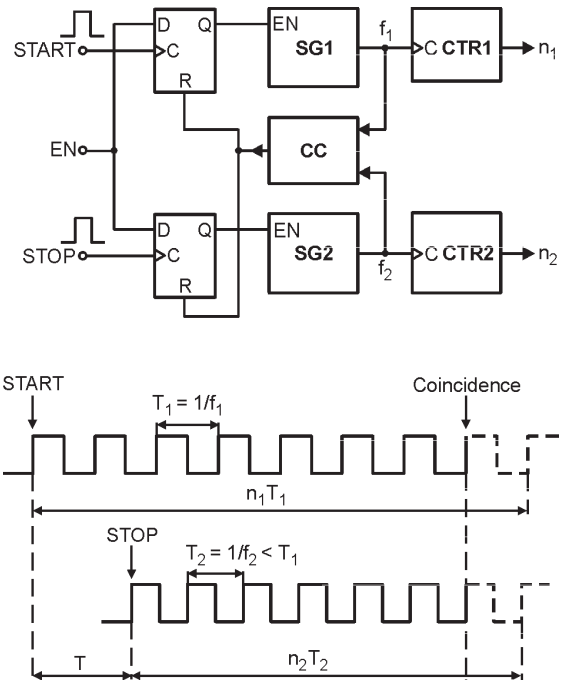

Figure 6. TDC based on the Vernier method: (*a*) circuit example, (*b*) example of conversion process at EN = H.

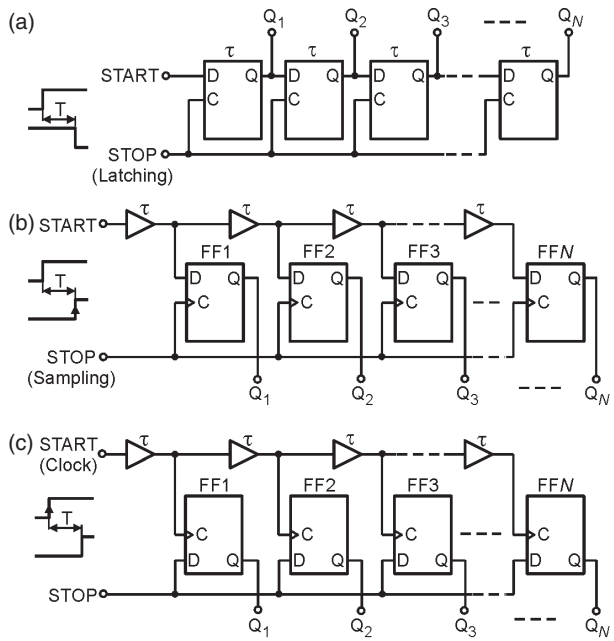

Figure 7. TDC utilizing the tapped delay line: (*a*) line comprising latches, (*b*) line comprising buffers with simultaneous sampling of its state, (*c*) line comprising buffers with successive sampling of the state of the STOP input.

sampling the state of the STOP input. When the STOP pulse appears, the nearest clock edge changes a FF output to H. If this does not disable triggering the next FF (by an additional logic), its output will also be set in the H state after the delay τ, and so on. Then the measurement result is given by the lowest position of the FF that stores the H state.

The above-described 'delay line' techniques represent direct time/digital conversion, that is without any intermediate processing. The sampling operation results in a negligible conversion time and therefore such converters are also called *flash* TDCs. If the readout time is ignored, the dead time of the circuit (*a*) is equal to the time needed to reset all latches in the line. When the line is reset serially (by setting START = L), the dead time is $N\tau$, but when using the parallel reset (using a separate reset input of all latches) the dead time also becomes negligibly short. Using separate reset inputs of the FFs in the circuits (*b*) and (*c*) also results in a negligible dead time.

One may point out that the use of the tapped delay line for measurement of TIs is equivalent to the use of a fast counter driven by a startable clock. For example, the line composed of latches with $\tau = 2$ ns is equivalent to the counter driven by a clock of frequency equal to 500 MHz. The number, N, of latches in the line is, however, much greater than the number, n, of FFs in the equivalent counter ($N = 2^n$). The measuring range can also be enlarged much more easily using the counter approach. To double that range using the counter, only one FF should be added ($n' = n + 1$), while the length of the line should be increased two-fold ($N' = 2N$).

An improvement of the basic tapped delay line has been the 'pulse-shrinking' delay line [30, 31], which offers better resolution. It was utilized in a TDC designed for space instrumentation [32].

A fine resolution of the TDC can also be obtained using two lines of slightly different cell delays creating the *differential delay line*, usually fabricated in an application-specific integrated circuit (ASIC) [25, 26, 65].

Such a TDC was also designed in a more cost-effective CMOS FPGA technology [33, 34]. The basic time-coding circuit of this converter is shown in figure 8. It contains two delay lines with 63 delay cells and the output decoder. Each delay cell contains the latch L having a delay τ_1 (between the input, D, and the output, Q) and being a part of the first delay line, and the non-inverting buffer, B, having a delay $\tau_2 < \tau_1$ and being a part of the second delay line. The input TI is defined between the rising edges of the pulses START and STOP and coded in the first delay line by setting the H level at the Q output of the last cell whose C input change (L $\rightarrow$ H) is ahead of a similar change at the input D. An average resolution ($\tau_1 - \tau_2$) of about 200 ps was obtained, covering the 10 ns range with 50 cells. The maximum conversion time is $63\tau_1$.

Each cell set to the H level generates the reset input signal to the previous cell in a local feedback loop. In this way, all the cells preceding the last set cell are cleared, and the output from the line is obtained in the '1-out-of-63' code. To convert it into 6 bit natural binary code, an array of multi-input OR gates of the FPGA device was used. No separate reset input is needed because in the initial state the line consisting of open latches (when STOP = L) is transparent to the input START = L.

Figure 9 shows the logic structure of the single delay cell created within the logic block of the FPGA device (QuickLogic). The latch (τ_1) is built with the multiplexer N and the gates D and E. The non-inverting buffer (τ_2) is realized by the gate F. One or more of the free inputs shown can be connected to the node D and/or the buffer input to increase the respective delays as needed to obtain minimum linearity error of conversion.

The design of a precise TDC with FPGA technology is not easy, because to obtain low linearity error many trial-and-error designs have to be tested before sufficient experience is gained and a satisfactory result is obtained. Software simulators are not accurate enough in such applications. However, when the design is accepted it can be used for mass production without further modifications.

It may be noted that the above method of TI measurement is similar to the Vernier method described earlier with two startable oscillators (figure 6). The delays τ_1 and τ_2 may be regarded as equivalent to the periods T_1 and T_2. Therefore the differential line is also called a *Vernier delay line* [65], and

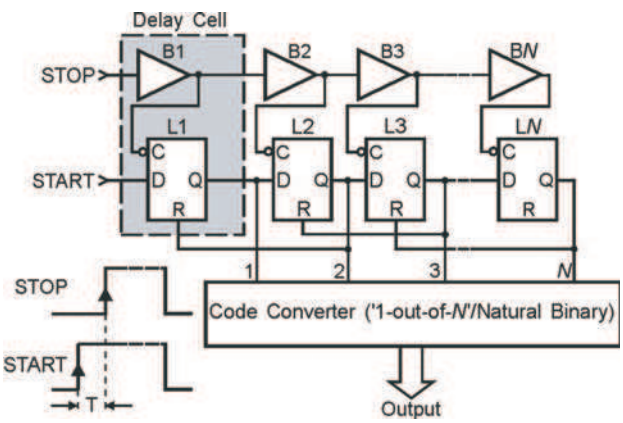

Figure 8. Example of the differential (Vernier) tapped delay line [33].

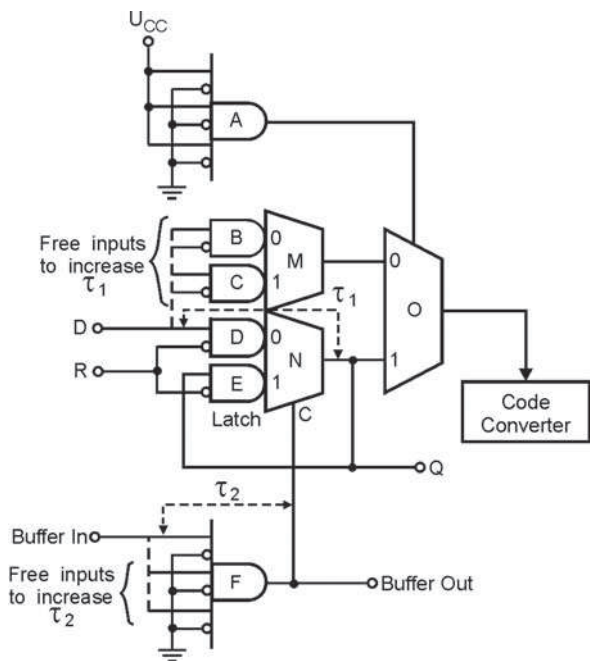

Figure 9. Delay cell of the differential line shown in figure 8, created in the FPGA logic block (pASIC1, QuickLogic) [33].

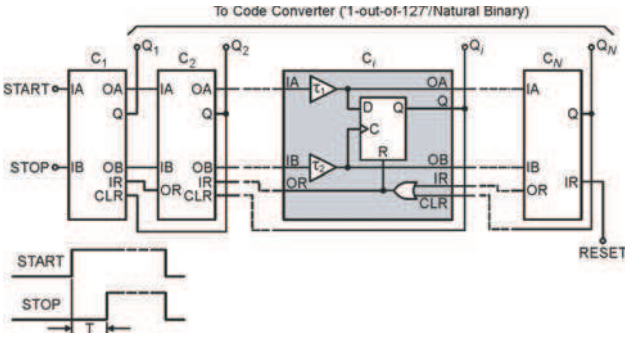

Figure 10. Simplified logic diagram of the differential delay line providing a 100 ps resolution in an FPGA device (pASIC2, QuickLogic) [35].

the respective converter is called a *Vernier TDC with delay lines*.

In the improved converter with a Vernier delay line, also designed in the CMOS FPGA device, a 100 ps resolution was obtained [35]. The logic of the delay line is shown in figure 10. Here the time difference is created by two buffers of delays τ_1 and τ_2, and the coincidence is detected by the D FF whose output Q is set at the H level.

The use of a more economical, reprogrammable FPGA device (Virtex XCV300) resulted in the design of TDCs with 100 ps and 500 ps resolution [81]. To obtain a 100 ps resolution the intrinsic carry delay between the logic slices within the FPGA configurable logic blocks has been utilized to create a single, 32-tap delay line of the type shown in figure 7(*b*). The TDC with 500 ps resolution utilizes two 16-tap delay lines with 1 ns delay per tap. An additional shift of 0.5 ns between the lines creates a virtual 32-tap delay line with a 0.5 ns resolution.

The Vernier delay line was also used in the TDC fabricated in the CMOS 0.7 µm technology [65] and containing the cells of a structure similar to that described in [35]. In this design a high stability and a 30 ps resolution were obtained.

2.3. Interpolation methods: coarse and fine measurements together

The interpolation methods are used when both a long measuring range and a high resolution are required. The long MR is provided by the coarse counter driven by the reference clock (LSB $= T_0$), while the high resolution is obtained by the fine interpolators.

By definition, interpolation is a method for determination of an approximate value of a function within a range bounded by two function values. With regard to the TI measurement, when a timing event, T_x, occurs between two succeeding states of the coarse counter, say between n and $(n+1)$, then $T_x/T_0 = n_x + c_x$, where the coarse counter content $n_x = \mathrm{Int}(T_x/T_0)$ is an integer part of the ratio T_x/T_0 and $c_x = \mathrm{Frc}(T_x/T_0)$ represents the respective fractional part, measured by the interpolator.

The interval T measured with the use of the interpolation method is decomposed into three intervals. One interval (which may be quite long) is measured in real time by the coarse counter. The remaining two short intervals are defined at the beginning and at the end of the interval T (the first when the START pulse appears and the next at the STOP pulse) and are measured by a single or two interpolators. Each interpolator contains the synchronizer, which produces a short TI (usually between T_0 and $2T_0$), which is measured by a fine TDC of a short range (usually $2T_0$). The TDC utilizes one of the fine conversion methods described in the preceding section and provides high resolution (LSB $= T_0/K$, where $K = 10$–10^4).

If a startable oscillator is used for coarse counting, then the START interpolator is not needed because in such a case $c_{\mathrm{START}} = 0$.

If the interpolation is performed at the beginning and at the end of TI, then such an interpolation is called *dual interpolation*. Sometimes that term is referred to a single interpolator, which performs the interpolation in two succeeding steps and using two separate electronic circuits, which create a *tandem interpolator* [13, 14]. To avoid ambiguity, in this paper the latter case will be called a *two-stage interpolation*.

Two-stage interpolation can be realized in two ways: consecutively by two circuits connected in series [13, 14] or simultaneously by two parallel circuits [61]. The corresponding two-stage interpolator may be called, respectively, a *serial* or *parallel* (flash) interpolator.

The interpolation can also be realized in more than two stages. Recently a three-stage parallel interpolator was designed [63], and earlier a *multiple interpolation method* was described [15], where a single interpolator with a feedback is repeatedly used in a few consecutive steps at each input event.

The dual interpolation method was first introduced with the use of the Vernier or digital time stretching method [21] and later with the use of the analogue processing [36, 37]. The Baron method [21] was greatly improved, called 'dual Vernier' and used with success in a commercial TC [38]. The main invention [39] has been the use of two free-running oscillators stabilized by the PLL. The oscillators can be momentarily stopped (PLLs switched out) and then started exactly in-phase with the beginning of the interpolated TIs. The PLLs are switched on automatically when their phase detectors discover time coincidences. In this way a high accuracy of measurement of even very long TIs was achieved (20 ps resolution, 10 s range).

A detailed analysis of the Nutt method [36, 37] has been presented in [12]. The method has been used with time stretchers [6–15, 17, 36] and with T/A + A/D converters [16, 18–20, 37]. The latter method has been used in the commercial counter SR620 (Stanford Research Systems). A common capacitor for both interpolators has also been used in some designs. 'Pulse-shrinking delay lines' were used in the specialized TDCs [31, 32]. CMOS FPGA technology has also been utilized to design a single-chip interpolation TC [35, 40, 41]. The leading designs in the CMOS ASIC technology are presented in the next section.

Figure 11(*a*) shows an example illustrating the Nutt method. It has become very popular due to the relative simplicity of design and low cost while providing high resolution and large MR. The measured time interval, T, is decomposed into three parts:

$$T = n_{\mathrm{C}} T_0 + T_{\mathrm{A}} - T_{\mathrm{B}}, \tag{9}$$

where n_{C} is the content of the coarse counter operating in real time. The TIs T_{A} and T_{B} are measured between the leading edge of the input pulse (START, STOP) and the second nearest

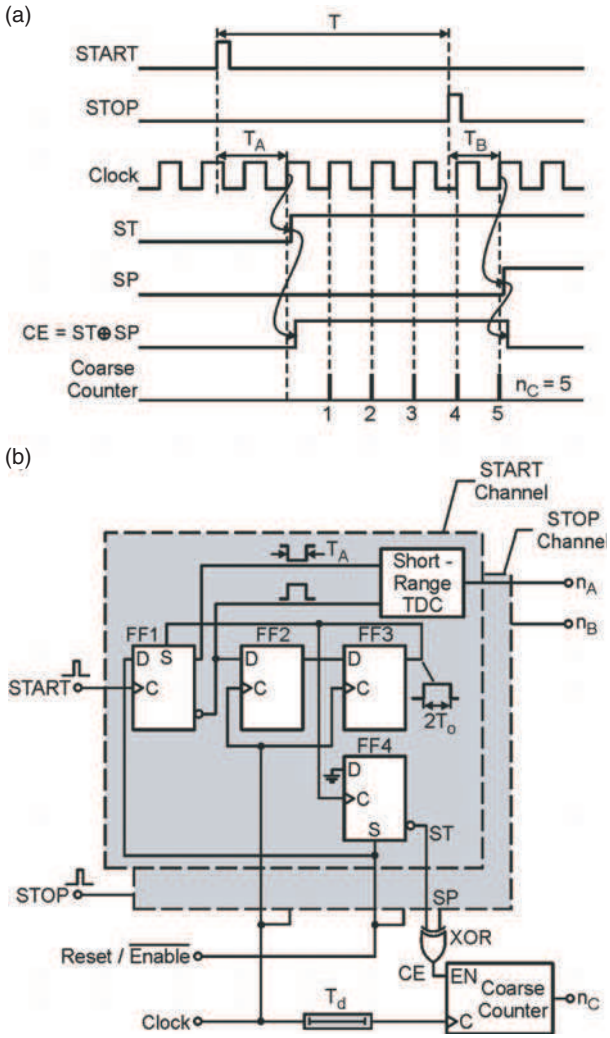

Figure 11. The Nutt interpolation method: (*a*) example of waveforms, (*b*) example of the relevant circuit diagram.

clock pulse. The internal signals, ST and SP, are generated to create the signal enabling the coarse counter. When using time stretchers, the intervals T_A and T_B are stretched by the respective factors, K_A and K_B, and then counted to obtain the counter contents n_A and n_B. Denoting the respective resolutions as $\tau_A = T_0/K_A$ and $\tau_B = T_0/K_B$, we get

$$T_A \cong n_A \tau_A \qquad \text{and} \qquad T_B \cong n_B \tau_B. \tag{10}$$

These TIs can be expressed similarly when the faster and more precise conversion method of T/A followed by A/D is used (section 2.2).

An example of the interpolating TC is illustrated by the simplified logic circuit shown in figure 11(*b*). Similar designs are commonly used [6–20]. The circuit contains two interpolators with short-range TDCs and the coarse counter. In each interpolator the flip-flop FF1 sets the H level at the negated output when the leading edge of the asynchronous input pulse appears. The 2 bit shift register (FF2 and FF3) is a two-stage synchronizer detecting the second nearest clock pulse. The leading edge of the pulse appearing at the FF3 output triggers FF4, sets ST = H, and completes the complementary pulses of width T_A at the FF1 outputs. The coarse counter is enabled by the XOR gate when CE = ST ⊕ SP = H. The optional delay, T_d, compensates for the propagation time of FF4 and XOR gate if timing is critical (at a high clock frequency).

Traditionally, the main reason for detecting the 'second nearest' clock pulse instead of the 'first nearest' has been the strong non-linearity of the initial part of the transfer characteristic of the popular TDCs employing the intermediate time stretching or T/A conversion methods. The 'third nearest' approach was also used [12].

The synchronizers with two or more stages also help to reduce or virtually eliminate the adverse influence of the metastability effect on the accuracy of the interpolating counters [6, 16, 24, 60]. The well-known effect of *metastability* in FFs [75–77] can be observed when the signal at the FF data input (D, T, J, K) changes state within a very short time window around the active edge of the clock signal applied to the control input (C) of an edge-triggered FF. This results in random stretching of the FF propagation time, and in some cases even the final logical state of FF cannot be predicted. The effect of such a stretching in the single-stage synchronizer (comprising only a single FF), though appearing very seldom, may be observed [35].

There are three main sources of errors contributing to the combined standard uncertainty, s, of the interpolating counters: non-linearity of both embedded interpolators, quantization error, and jitter.

In typical applications, the input START and STOP pulses are asynchronous or are not correlated in time with the reference clock. Then the linearity error is a function of the measured interval, T. That function is repetitive modulo clock period T_0 when T is varied, mainly due to the non-linearity of the short-range TDCs in the interpolators [12]. Thus we may expect that the behaviour $s(T)$ is the same when T changes to $T \pm iT_0$, where i is an integer. The accurate plot of the function $s(T)$ within at least one clock period, T_0, is the most representative measure of the standard uncertainty for a given interpolating counter. To avoid excessive averaging in channels of the plot $s(T)$, a sufficiently narrow channel width should be chosen. An example of a measured function $s(T)$ with a channel width $0.1T_0$ is shown in figure 17. Clearly a width of $0.05T_0$ would result in a better accuracy.

It also means that the popular plots showing the statistical dispersion of measurements of a constant interval T are actually of minor value because many different plots may be generated within the 'window' T_0. The designer might want to show the best plot (with $s_{\min}$), but for correct evaluation of the design, the plot with $s_{\max}$ should rather be shown.

The quantization error appearing in the interpolation method is the difference between the quantization errors produced by the START and STOP interpolators. For a given interval T measured asynchronously, the error induced by the START interpolator conforms to the uniform distribution, but the STOP events are strongly correlated in time with the START events (9). In a simplified theoretical case, when both interpolators have the same values of ideally linear conversion factor K being an integer, the value of the quantization error may be negative or positive, with the respective, normalized probabilities [12]

$$P_1(\eta_x \geqslant \eta_c) = 1 - \eta_c \tag{11a}$$

and

$$P_2(\eta_x < \eta_c) = \eta_c, \tag{11b}$$

where $\eta_x = \mathrm{Frc}(Kx)$ at $0 \leqslant x \leqslant 1$, $\eta_c = \mathrm{Frc}(Kc)$, $c = \mathrm{Frc}(T/T_0)$, and $K = T_0/\mathrm{LSB}$. This means that the fraction c is decomposed into Kc quantization steps, each one of width LSB. The quantization error appears only within the last step.

In the sample of asynchronous measurements of a time interval T the probabilities (11) correspond to the normalized numbers of two-valued hits differing by a single LSB. Then the quantization error can be represented by the binomial distribution, as in the simple TCs (section 2.1). The error has a zero mean value but its standard deviation strongly depends on η_c or the measured interval T (cf (3) and figure 3(*b*)):

$$\sigma = \mathrm{LSB}\sqrt{(1-\eta_c)\eta_c}. \tag{12}$$

The maximum standard deviation $\sigma = 0.5$ LSB is obtained at $\eta_c = 0.5$ and the average standard deviation is (cf (4))

$$\sigma_{\mathrm{av}} = \frac{\pi \mathrm{LSB}}{8} \cong 0.39\ \mathrm{LSB}. \tag{13}$$

In real counters the conversion factors K in the interpolators may be not equal, not exactly integers, and not strictly linear. Then the quantization steps may not be identical (influenced by non-linearity) and the relevant probability distributions may be distorted. Some authors [46, 47, 56, 58] assume that the overall error contribution due to quantization in the real interpolating counter can be approximated by the rms error of a simple uniform quantizer, or $\sigma \cong \mathrm{LSB}/\sqrt{12} \cong 0.29\,\mathrm{LSB}$. That measure seems too optimistic in this application.

It should be noted that in a general case the quantization error is a repetitive (modulo LSB) function of T, like the linearity error (modulo T_0), and the two error sources combine to create a non-linear function $s(T)$ within a 'window' T_0, or $s(c)$.

The jitter error is caused by the noise inherent in the components used, jitter contributed by the external signals (including the clock), and the noise induced by the environment (including the power supply). The jitter contribution generally does not depend on c and creates a 'floor level', which usually is below 10 ps (rms).

Thus, when measuring the characteristic $\sigma(c)$ at $0 < c < 1$, we can distinguish the almost constant jitter floor and the variable error caused by non-linearity and quantization.

2.4. Interpolating TDCs in CMOS ASIC technology

When ASICs became generally available for custom design and their manufacturing became economically feasible, the trend in design of precise time converters shifted towards 'pure' digital conversion methods, based on the use of custom designed, integrated delay lines, and synchronous counters needed to obtain greater dynamic range. TI measurements are accomplished by sampling (actually reading) the current states of the line and the counter, and storing them 'on the fly', without interruption of the counting process. Those solutions have the following distinctive features:

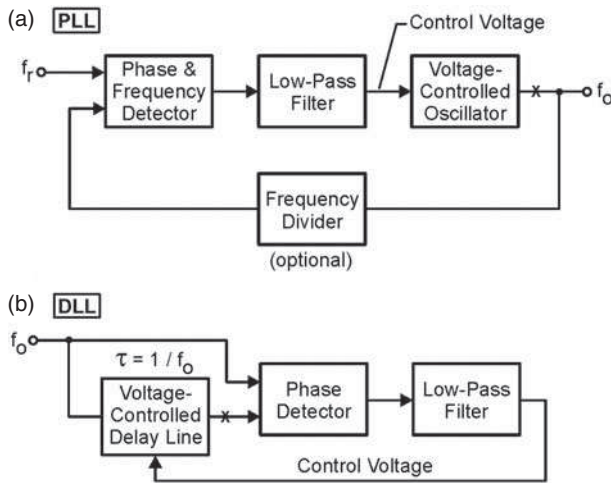

Figure 12. Basic block diagrams of the PLL (*a*) and DLL (*b*).

- the delay lines and the complete chips can be designed specifically for a required application containing a dedicated control logic and offering, for example, a multistop (multisampling) operation;
- the use of an internal PLL or DLL provides easy, automatic stabilization of the measuring range and quantization step (resolution) against ageing and changes of the ambient temperature and supply voltage;
- the conversion time is virtually zero and the dead time can be minimized by the use of additional registers and first-in-first-out (FIFO) memory;
- lower power dissipation, lower chip count, and better reliability can be obtained than in older technologies.

Precision, integrated TDCs with delay lines are grouped in two categories, depending on the use of a PLL or DLL circuit [30]. Both techniques are comparatively described in the textbook [42]. Figure 12 shows the relevant basic circuits.

The PLL (figure 12(*a*)) contains a voltage-controlled oscillator (VCO), whose frequency f_0, after optional dividing, is compared with the reference frequency, f_r. The difference is detected, filtered, amplified, and used to adjust the frequency of the VCO to minimize the difference.

The PLL is a much older idea than the DLL, and was invented already in the era of vacuum tubes. The design and analysis of PLL circuits have been described in many articles and books (e.g. [43, 44]). PLL circuits are commonly used for frequency synthesis. The frequency f_0 can be much higher than f_r and can be controlled easily by changing the dividing ratio in the frequency divider. In a simplified model, the jitter introduced by the reference source is reduced by virtue of the low-pass behaviour of the PLL. That is why PLLs were first developed to recover data and timing from noisy communication channels; the VCO acts much like a flywheel in a mechanical system. However, the inherent jitter of the VCO is present directly at the output. The output jitter is also influenced by the low-frequency input jitter (from the reference source), which exists within the PLL band.

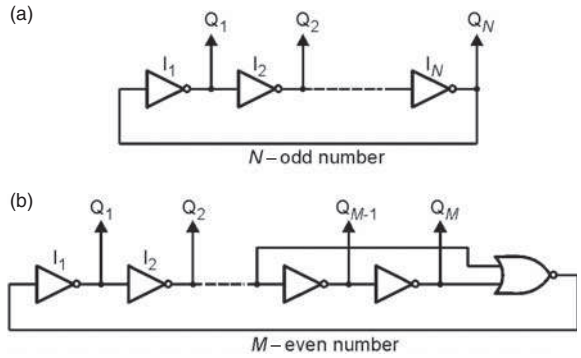

Figure 13. Ring oscillators used in PLL circuits: (*a*) with odd number of inverters, (*b*) with even number of inverters.

In the DLL approach (figure 12(*b*)) the loop contains the voltage-controlled delay line (VCDL). The delay, $N\tau$, of the line is varied to align the phases at the inputs of the phase detector. In the ideal case, $N\tau = 1/f_0$. A DLL provides superior jitter performance when a clean reference clock is available.

Both PLL and DLL circuits have also been successfully used for aligning the clock in complex digital devices and systems. For this purpose the circuit elements inducing the clock skew are inserted in places marked by 'X' in figures 12(*a*) (without frequency divider) and (*b*).

For TI measurements, the free-running VCO in the integrated PLL loop is usually designed as a ring oscillator. In the basic configuration it is created by the delay line that contains an odd number, N, of inverters. The output of the line is connected to its input, as shown in figure 13(*a*). The oscillation period is

$$T_0 = N(t_{p\mathrm{LH}} + t_{p\mathrm{HL}}), \tag{14}$$

where $t_{p\mathrm{LH}}$ and $t_{p\mathrm{HL}}$ are the respective propagation times of the inverters $I_1, \ldots, I_N$. In the modified ring counter [45] containing additional gates, the period (14) has been almost halved.

If $t_{p\mathrm{LH}} = t_{p\mathrm{HL}} = \tau$, then $T_0 = 2N\tau$, and the signals at the outputs $Q_1, \ldots, Q_N$ represent a multiphase clock with delay step $\tau = 1/(2Nf_0)$. The number of timing signals delayed by τ is equal to $2N$ (both rising and falling edges). The ring counter can also be designed with the asynchronous SR flip-flops instead of inverters. The invention [80] shows that idea yet with vacuum tubes and quartz stabilization without using a DLL or PLL. However, it can be implemented in a modern integrated circuit.

In the CMOS time digitizer [46] a four-stage ring oscillator with frequency 125 MHz ($T_0 = 8$ ns) was used to obtain LSB = 1 ns. In this design the control voltage of the VCO was used also to stabilize the separate, startable ring oscillator [31, 32].

Another approach is based on the use of the PLL ring oscillator as a multiphase clock driving the clock inputs of the edge-triggered D FFs of the associated register, as shown previously in figure 7(*c*). Such a principle was used in the design of ASIC TDCs of designations TMC-TEG3 [47] and F1 [48].

The four-channel TMC-TEG3 was developed using a 0.5 µm CMOS sea-of-gates technology. It contains an asymmetric, 32-stage ring oscillator (figure 13(*b*)), which delivers an even number of timing signals ($M = 32$). The oscillation period is also given by equation (14), where N should be replaced by M. The number of equally spaced timing signals (rising edges only), delayed by the resolution ($t_{p\mathrm{LH}} + t_{p\mathrm{HL}}$), is also equal to M (in this case, not $2M$). A detailed analysis of this oscillator is presented in the text [47] where the main circuit blocks of the TDC have also been described.

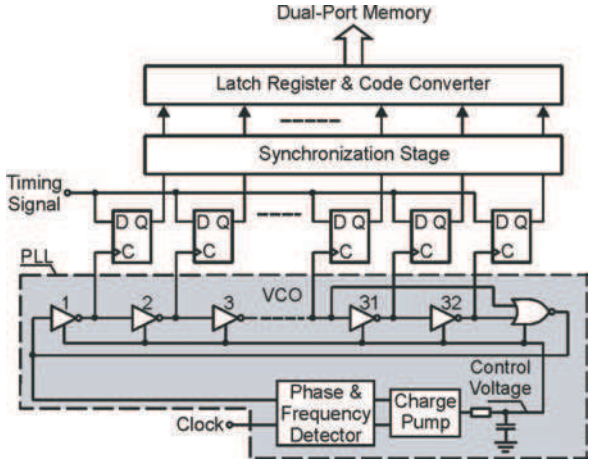

Figure 14. Example of a TDC with the PLL circuit [47].

The time-digitizing circuit of a single channel contained in the TMC-TEG3 chip is shown in figure 14. It is also representative of other PLL designs. The PLL comprises a phase/frequency detector, a charge pump, a low-pass filter, and a VCO (ring oscillator). An external capacitor, C, is used in the filter. The time-coding register is the same as previously shown in figure 7(*c*). At a typical VCO frequency of 40 MHz, the resolution is 25 ns/32 = 781 ps.

The eight-channel F1 was fabricated with 0.6 µm CMOS sea-of-gates technology and contains 19 inverters of 150 ps typical delay to create a VCO of the same structure as shown in figure 14, utilizing both edges of the clock. Thus the typical frequency, f_0, of the VCO is $1/(38 \times 150\,\mathrm{ps}) \approx 175$ MHz. The 19 bit register stores the data representing the 'fine' part of the measured TI, which is measured with 150 ps resolution. After conversion to the natural binary code this part represents an interval equal to a number n times 150 ps. The 'coarse' counter is driven directly by the VCO and counts the number, m, of periods $T_0 = 5.7$ ns. The dynamic range is determined by the 16 bit data word or $2^{16} \times 150\,\mathrm{ps} \approx 9.8\,\mu\mathrm{s}$.

In general, when using a VCO to count its periods and store its state 'on the fly', the result of a TI measurement is obtained as a difference of the data sampled first at the START (n_1, m_1) and then at the STOP (n_2, m_2) event. This is the *dual sampling* principle:

$$T = (m_2 - m_1)T_0 + (n_2 - n_1)r, \tag{15}$$

where r is the TDC resolution (LSB). It should be noted again that the conversion time is virtually zero or the dual sampling method allows for design of a 'flash' TDC. The multistop (multisampling) operation can be performed by reading the numbers n and m at succeeding events.

In this way continuous measurements of succeeding TIs between pulses in a train can be performed. For example, a typical reference time clock generates pulses of 1 s period, which is closely related to the standard time. Measurements of the succeeding periods to a picosecond accuracy by a TDC utilizing a better clock reference (or at least of a known performance) can be utilized for a detailed statistical analysis of the evaluated source.

A four-channel AMS110 device [49] has been manufactured in a 0.8 µm BiCMOS technology and has an adjustable resolution in the range 125 ps to 175 ps. The free-running ring oscillator utilizes the delay line containing eight pseudo-ECL differential buffers (one of them is inverting) and generates clock pulses of adjustable frequency (500 MHz maximum), synchronized by the PLL to the external reference of 31.25 MHz. When an external hit occurs, the status of the delay line is sampled, stored in the register, and converted to a 4 bit natural binary code. At the maximum VCO frequency of 500 MHz, the minimum LSB is 2 ns/16 = 125 ps. The clock also feeds a 10 bit coarse counter (read on the fly), giving a total dynamic range of $2^{14} \times$ LSB or $16\,384 \times 125$ ps $= 2.048$ µs minimum. This TDC (and the improved model, AMS111) has been designed with a focus on high readout speed and reduction of the pile-up of data obtained during physical experiments.

The concept of DLL (figure 12(*b*)) when implemented in MOS technology was first described under the name of 'synchronous delay line' [50]. The delay line was controlled by a feedback loop containing the switched-capacitor low-pass filter. The 'variable delay line PLL' was utilized for CPU-coprocessor synchronization [51]. In this design the low-pass filter was designed as a commonly used charge pump. In a TDC design [52], voltage-controlled delay elements (time memory cells (TMCs)) with a feedback loop were introduced and an LSB of 0.77 ns was obtained. A simple feedback loop with the delay line was also used in a design of a TDC, with a 0.75 ns LSB [53].

The design issues related to both PLL and DLL applications have been presented in [30], where the asynchronous 'pulse-shrinking delay line' was also introduced. The latter approach and a DLL were used to design a CMOS TDC [27] based on the interpolation method and having an LSB of 0.78 ns.

A general block diagram of the TDC utilizing DLL is shown in figure 15. Following the basic definition of DLL (figure 12(*b*)) the delay of the line is adjusted by the loop to be equal to the clock period or $N\tau = T_0$. To take a measurement when an event (START or STOP) appears, the states of the line and the coarse counter are sampled and stored in related registers on the fly. The sampling register is controlled by the N-phase clock generated at the taps of the delay line and can be regarded as an array of N single interpolators.

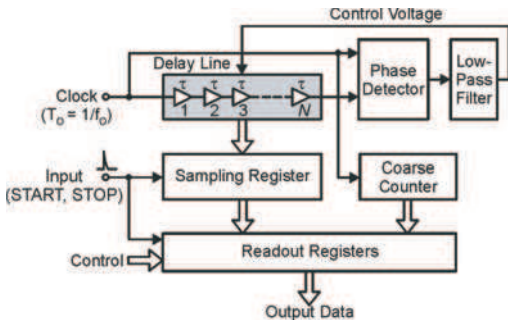

Figure 15. Block diagram of a typical TDC with DLL.

A 16-channel TDC with DLL was developed in a 1 µm CMOS technology, and an LSB of 1.56 ns was obtained [54]. To achieve better resolution, an array of DLLs was proposed [55], and using the same technology, a 150 ps resolution was measured in a tested array. This concept was later used to design a four-channel TDC with an LSB of 89 ps and a dynamic range of 3.2 µs [56].

A TDC with a DLL working in 'true' Vernier mode [57] allows us to obtain a resolution, LSB, of a smaller value than the delay, τ, of a single cell in the line. The main assumption is that the line delay, $N\tau$, can be a multiple of T_0, or $N\tau = HT_0$, where H is such a number that the greatest common divisor of N and H is 1. Then the resolution can be calculated as

$$\mathrm{LSB} = \frac{T_0}{N} = \frac{\tau}{H}. \tag{16}$$

In a test chip, MTD144 [57], the lowest value of LSB = 46.9 ps was obtained at $T_0 = 6$ ns ($f_0 = 166.6$ MHz), $H = 5.5$, and $N = 64$. Both edges of the clock were utilized, which effectively doubles the length of the delay line ($N = 128$, $H = 11$). The delay of a single cell is $\tau = 516$ ps.

Almost the same resolution (48.8 ps) was obtained in the TDC with a simple DLL and the tapped RC delay line [58]. The linearity of that line was corrected using the commonly used but time-consuming statistical method (code density test) [12, 18, 33–35].

A 16-channel TDC chip with a 0.5 ns resolution [59] was developed in a 0.8 µm CMOS technology for physics research. The TDC integrates one 60 MHz counter, 16 typical DLL-controlled delay lines with 32 taps of 500 ps delay each, and one calibration channel.

In a nine-channel TDC design fabricated in a 0.8 µm CMOS technology [60], a single delay line with 32 cells, a 6 bit coarse counter, and a 50 MHz clock were used to get a resolution of 625 ps within a 960 ns range. The metastability effects occurring in FFs used for synchronization of the input pulses were analysed and a *two-phase synchronizer* was used. In this circuit the START and STOP pulses are synchronized to both rising and falling edges of the reference clock.

The following TDC from the same laboratory [61, 62] was designed in a 0.8 µm CMOS technology for application in a precise laser rangefinder and featured a resolution of 92 ps and a range of 3 µs. In this design a *parallel, two-stage interpolation method* has been introduced.

Figure 16 shows a simplified block diagram of this TDC. The coarse counter is fed by the reference clock of 85 MHz frequency or $T_0 = 11.765$ ns. The first interpolation stage (DLL1, Register 1) was designed as a typical DLL configuration. The delay line controlled by DLL1 contains 16 cells, C1, ..., C16, of delay $\tau_1 = T_0/16 \approx 735$ ps each. The main invention is the second stage of interpolation. The time location of the synchronized signal is found by a 16-input wired-OR gate, whose output is a sampling signal for the second interpolator. It contains an array of 16 delay lines, DL1, ..., DL16, fed in parallel by the asynchronous START (or STOP) input pulse. The delay of the consecutive lines is

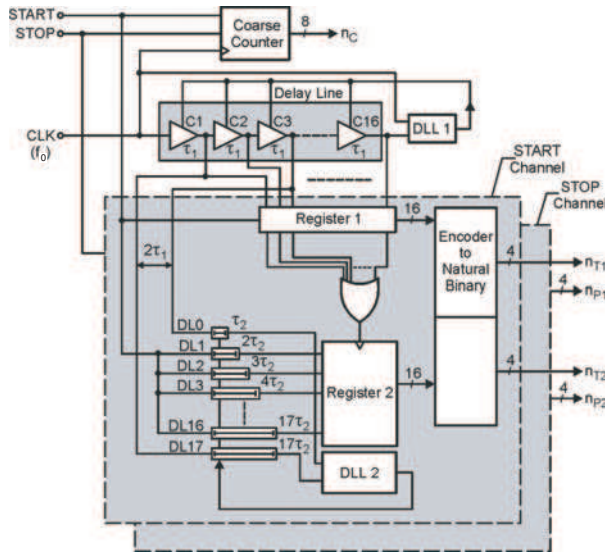

Figure 16. Block diagram of the TDC with two-stage interpolation [61, 62].

incremented by τ_2, the fine resolution determined by the DLL2. This loop is referenced by the delay $2\tau_1$ and the delay difference of the loop lines DL17 and DL0, equal to $16\tau_2$. Thus $\tau_2 = \tau_1/8 \approx 92$ ps. The array, DL1, . . . , DL16 is divided into three segments: the measuring segment containing eight middle lines, DL5, . . . , DL12, and two side segments, DL1, . . . , DL4 and DL13, . . . , DL16, which allow some timing mismatches between the synchronized and unsynchronized input signal paths caused by manufacturing process and temperature variations.

The array of parallel delay lines used is functionally equivalent to the single tapped delay line shown in figure 7(*b*). However, the delays of the parallel lines can be independently fine tuned in the design process to minimize linearity error and the resulting random error. The cumulative delay error inherent in the tapped lines is eliminated. This means more freedom for the designer and results in a better TDC accuracy.

A *three-stage interpolation* has been used in the improved, nine-channel TDC designed in the same laboratory and fabricated in a 0.6 µm CMOS technology [63, 64]. The clock period is first divided by 16, next by 4, and finally by 8. At a clock speed of 66 MHz, this gives the resolution of 151.5 ns/512 = 29.6 ps, and the 15 bit coarse counter covers a range of 496 µs. To lower the linearity error and the related random error, a linearity correction was utilized, similar to that described in [18, 33–35]. In this way the random error was lowered from about 30 ps to below 20 ps over an ambient temperature of −40 °C to +60 °C.

A better stability and lower random error were also obtained in the FPGA counter [40, 41] as a result of non-linearity correction and introduction of an external DLL circuit controlling the supply voltage of the FPGA device [66].

3. Some design issues

3.1. Correction of non-linearity

The non-linearity of conversion is the main cause of measurement uncertainty in precise TDCs. When all possible options in the optimization of circuit design have been utilized, there remains a possibility of correction of rough digital data by suitable processing. The idea is the compensation of the non-linearity error in all bins within the measurement range of the TDC used. When using the interpolation methods it refers to the correction of fine interpolators only.

To make a correction, first the linearity error has to be identified. The data vector obtained is then used to correct every measurement result on the fly. A commonly used method for non-linearity identification is the statistical method, called also the 'statistical code density test' [12, 18, 33–35, 58, 64, 66]. Using a test generator that produces an approximately Poissonian train of pulses (randomly appearing in time), one has to take a large number, N, of test measurements to obtain a discrete histogram consisting of a number, M, of channels (bins). In practice, a common RC generator may be used for this purpose, but not a stabilized one to avoid timing correlations with the reference clock driving the TDC under test. The number N should be sufficiently large to obtain a sufficiently small random uncertainty of the content n in every channel. That uncertainty is approximately equal to $1/\sqrt{n}$. In an ideal case the content in each channel should be the same ($n_s = N/M$), but in a real case in every (ith) channel there is a differential non-linearity (DNL):

$$l_i = \frac{n_i - n_s}{n_s}. \tag{17}$$

The integral non-linearity INL, referred to the jth channel, is obtained by summation:

$$L_j = \sum_{i=1}^{j} \frac{l_i}{M}. \tag{18}$$

To describe the linearity error by a single value, usually the maximum value of L_j or $L_{j\,\max}$ ($1 \leqslant j \leqslant M$) is selected, which represents the worst case.

The correction vector, containing M values of L_j, allows one to perform a suitable correction of the measurement data. That method is mostly effective in interpolating counters, where correction is performed on data obtained from the interpolators. The correction is usually performed by the microprocessor or PC used to control the measurements. The correction vectors can be stored in the EEPROM memory or in a file used by the processing software.

A dramatic lowering of the random error can be obtained using that correction (figure 17, [40]). However, one can

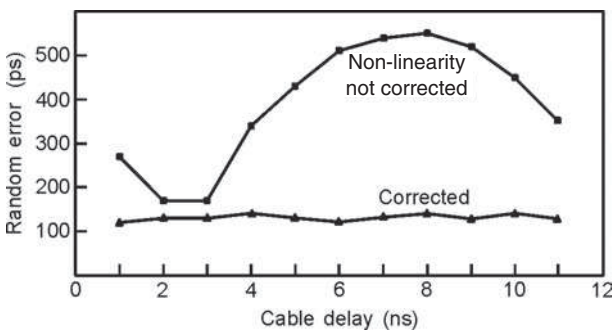

Figure 17. Effect of the non-linearity correction [40].

expect such a behaviour only when the ambient temperature is constant and equal to that when the correction vectors were determined. If it is not, introduction of the stabilizing DLL circuit helps considerably [64, 66].

3.2. Offset error

The offset error (non-zero readout at $T = 0$) in TCs can be easily compensated by making a series of a few hundred measurements at $T = 0$ or when START and STOP inputs are shorted. The mean value calculated on the basis of that sample is the offset error which can be subtracted automatically from every other measurement result performed later. The test generator used must produce asynchronous pulses, not correlated in time with the reference clock of the TC. It may be the same generator as that used to determine the non-linearity correction vectors. However, the sample size needed here is much smaller.

3.3. LSB versus standard uncertainty

In the review presented of TDCs and the conversion methods used, the obtainable resolution (LSB expressed in picoseconds) was assumed as a distinctive feature, which makes a comparison feasible. This has been assumed as an analogy to commonly used A/D converters. However, with reference to TDCs, the LSB value allows for only partial evaluation of a given TDC. A more representative feature of *non-interpolating* TDCs is the maximum value of the INL, $\mathrm{INL}_{\max}$. In *interpolating* TDCs such a feature is the maximum value of the standard uncertainty, s (random error, precision, sigma value). The value of s is calculated as an estimator of the standard deviation in a sample of measurements of a constant time interval T, and $s_{\max}$ is found in a set of s values obtained when T is varied within a single clock period (section 2.3).

There is a commonly accepted rule that in a 'good' converter the condition $\mathrm{INL}_{\max} \leqslant \mathrm{LSB}$ or $s_{\max} \leqslant \mathrm{LSB}$ should be met, and in many designs it is met. In some designs it is not met, however. For example, in the TDC described in [11, 12] a value of LSB of 1 ps was obtained while $s_{\max}$ was 23 ps (mainly due to non-linearity). Indeed, it is easier to lower LSB than s and the only advantage resulting from the LSB being lower than s may be a negligible error contribution due to quantization.

3.4. Uncertainty of measurement of long TIs

The standard uncertainty observed in measurements of long TIs (measured or generated) greatly depends on the quality of the reference clock used. In a specific case, when the Allan variance, $\sigma_y^2(\tau)$, of the reference generator is known and the frequency instability is mainly caused by white noise, the sigma value of the jitter can be calculated as [69]

$$\sigma_T \approx \sqrt{T\tau}\,\sigma_y(\tau), \tag{19}$$

where τ is the averaging period. When $\tau = 1\,\mathrm{s}$, we get a simple formula $\sigma_T \approx \sqrt{T}\,\sigma_y(1)$.

3.5. Robust estimation

When the input timing signals are disturbed by some external noise sources, in the statistical spectra of measurement data some 'outliers' or results whose values differ much from the 'right' value may appear and their frequency of occurrence is much greater than would follow from the default probability distribution (figure 18). To eliminate such disturbed results during data processing, some robust estimation methods may be applied [18, 70]. They also appear effective in laser rangefinders using TDCs, for robust measurement of time of flight of laser pulses. A detailed analysis [71, 82] shows how significant improvement of accuracy can be obtained using new methods of robust estimation and optimized hardware.

3.6. Application example

An FPGA counter [40, 41] has been used to design a versatile, virtual time/frequency counter in the form of a PC card with a PCI or PXI interface (figure 19) [72]. The accompanying software provides control, display (figure 20), statistical processing of data, and diagnostics. The resolution is 200 ps at single-shot measurements within the range 0 s to 43 s. The selectable sample size, N, can lower the resolution according to the rule $(200\,\mathrm{ps})/\sqrt{N}$. The START and STOP inputs can be preset to have a 50 Ω or 1 MΩ input impedance and AC/DC coupling and to accept pulses of positive/negative polarity. The input threshold levels can be set manually or found automatically by the software. The frequency can be measured up to 1.1 GHz using the built-in frequency divider.

The new model of the counter card can be driven by an external clock signal of 10 MHz obtained, for example, from an atomic reference generator.

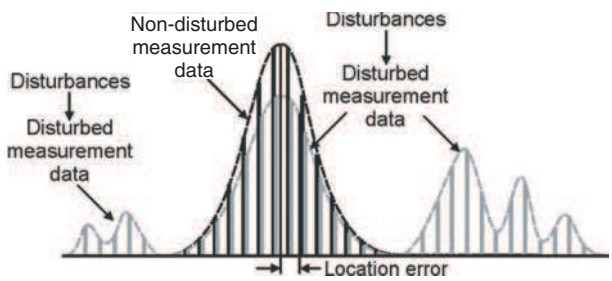

Figure 18. Deterioration of the measurement results caused by external disturbances.

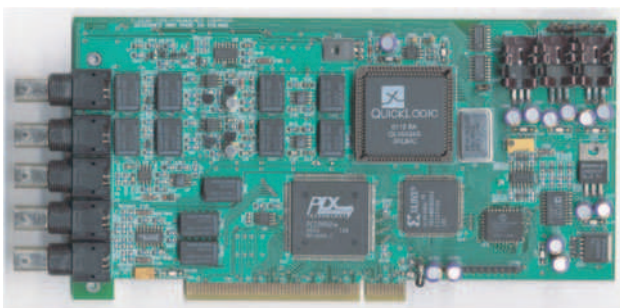

Figure 19. Time/frequency counter on a PC board with PCI interface.

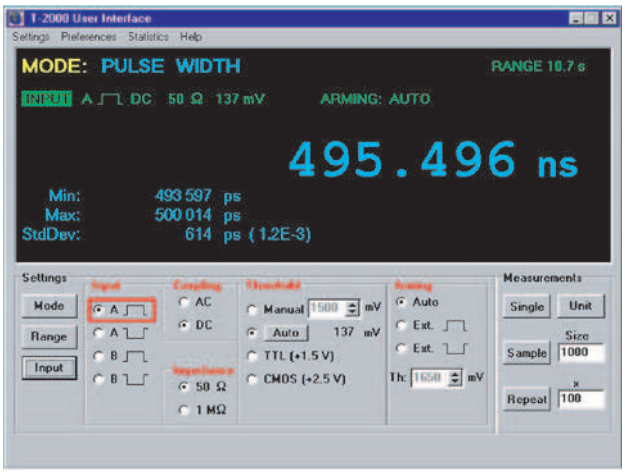

Figure 20. Example of a virtual front panel of the counter shown in figure 19.

3.7. Delay generators

When performing tests of TDCs a crucially needed instrument is a generator of precise time delays. A simple solution is a fast pulse generator driving a 'cable box' to produce pulses of different delays. A highly precise delay generator can also be designed using a programmable integrated circuit and a typical signal generator [73].

The measurements of time jitter produced by delay generators are frequently performed with the aid of high-frequency digital oscilloscopes with real-time sampling. There are some methods and hardware/software solutions offered commercially for this purpose. However, accurate measurement standards are still not provided and one may expect different results to be obtained when using different oscilloscopes and methods. In particular, when testing the jitter of the signal generated by a highly stable source, it should be noted that the generator of the sampling pulses is also based on a highly stable reference generator inside the oscilloscope. This implies a possibility of timing correlation of both generators, which may result in lowering of the displayed jitter value.

3.8. Analogue versus digital conversion methods

The 'analogue' TDCs, i.e. based on analogue processing (conversion of T to a voltage by charging a capacitor) still provide better resolution than 'digital' ones. A resolution of 1 ps to 5 ps can be obtained easily [9, 11, 12, 16–20]. However, the standard uncertainty (precision) obtainable, s, is not so good, usually between 15 ps and 25 ps.

The low cost and simple technology of this approach are advantageous. On the other hand, to lower the errors caused by the temperature sensitivity and time drift inherent to analogue instruments, a repetitive calibration must be performed. Although it may be highly precise [11] and automated using advanced adaptive algorithms [17, 67] analogue TDCs today already seem a bit old-fashioned.

Integrated digital TDCs are inherently stabilized by the embedded PLL or DLL circuits and thus their calibration is not required. The best single-chip TDC known at the time of writing this text (June 2003) features a resolution of 30 ps and precision below 20 ps [64]. It should be noted, however, that similar results were also obtained 'digitally' much earlier [37, 38], though in a multichip and rather complex TDC. Probably, new designs will appear in the future with still better parameters resulting from an advanced IC technology (SiGe, CMOS with design rules of 0.3 µm and below) and still more precise conversion methods.

4. Final remarks

It seems that the following methods will further be utilized and developed:

- Measurements in *parallel channels* integrated on the same chip and processing the output data to obtain better accuracy. For example, in one TDC [48], a 'high resolution' mode is provided when two channels operate in parallel, with an input time shift of LSB/2. This old idea [79] can be extended by increasing the number of parallel channels.
- An *averaging* approach can be utilized in different ways. An interesting method [68] involves the use of a group of simultaneously started and stopped integrated ring oscillators (of arbitrary frequencies) to calculate an accurate measurement result.
- *Correction of non-linearity* of the interpolators [18, 33–35, 64].
- *Improved methods based on the Vernier principle* [57] and with a DLL stabilizing delay difference [65].
- *DSP methods* used in advanced oscilloscopes and in dedicated instruments [74, 78].

Many other ideas and techniques may be utilized. They can be even the very old ones. Reviewing the old database of the United States Patent Office (www.uspto.gov) can result in discovering some exciting inventions, which can be used in a modern technology (for example [79, 80]).

Acknowledgments

The author wishes to thank the leading researchers in the field and collaborators in Digital Systems Laboratory at MUT for many helpful comments and suggestions which contributed to improvement of the original text. In particular I deeply appreciate the comments received from B Turko (Lawrence Berkeley Laboratory), J Kostamovaara and A Mäntyniemi (University of Oulu), C Hervé (European Synchrotron Radiation Facility), and J Christiansen (CERN). Many thanks to K Różyc who prepared and many times re-edited the figures.

References

[1] Porat D I 1973 Review of subnanosecond time-interval measurements *IEEE Trans. Nucl. Sci.* **20** 35–51
[2] 1970 Time interval averaging *Hewlett-Packard Application Note* 162-1
[3] Wakerly J F 2000 *Digital Design, Principles and Practices* 3rd edn (Englewood Cliffs, NJ: Prentice Hall)
[4] Alfke P 1996 Efficient shift registers, LFSR counters, and long pseudo-random sequence generators *Application Note* XAPP 052, Xilinx Corp.

[5] Moody N F 1952 *Electron. Eng.* **24** 289–93

[6] Wiedwald J D 1973 A CAMAC high resolution time interval meter *IEEE Trans. Nucl. Sci.* **20** 242–5

[7] Leskovar B and Turko B 1977 Optical timing receiver for the NASA laser ranging system *Lawrence Berkeley Laboratory Report* LBL 6133

[8] Leskovar B and Turko B 1978 Optical timing receiver for the NASA spaceborne ranging system *Lawrence Berkeley Laboratory Report* LBL 8129

[9] Kalisz J, Pawłowski M and Pełka R 1988 Präzisions–Zeitintervall–Mess-system *Elektronik* **14** 65–8

[10] Räisänen-Ruotsalainen E, Rahkonen T and Kostamovaara J 1996 A BiCMOS time-to digital converter with time stretching interpolators *Proc. European Solid-State Circuit Conf. ESSCIRC'96 (Neuchatel, 17–18 September 1996)* p 4

[11] Kalisz J, Pawłowski M and Pełka R 1985 A method for autocalibration of the interpolation time interval digitiser with picosecond resolution *J. Phys. E: Sci. Instrum.* **18** 444–52

[12] Kalisz J, Pawłowski M and Pełka R 1987 Error analysis and design of the Nutt time-interval digitiser with picosecond resolution *J. Phys. E: Sci. Instrum.* **20** 1330–41

[13] Turko B 1979 A modular 125 ps resolution time interval digitizer for 10 MHz stop burst rate and 33 ms range *IEEE Trans. Nucl. Sci.* **26** 737–45

[14] Turko B 1980 Space borne event timer *IEEE Trans. Nucl. Sci.* **27** 399–404

[15] Kalisz J, Pawłowski M and Pełka R 1986 A multiple-interpolation method for fast and precise time digitizing *IEEE Trans. Instrum. Meas.* **35** 163–9

[16] Kostamovaara J and Myllylä R 1986 Time-to-digital converter with an analog interpolation circuit *Rev. Sci. Instrum.* **57** 2880–5

[17] Kalisz J, Pawłowski M and Pełka R 1993 Improved time-interval counting techniques for laser ranging systems *IEEE Trans. Instrum. Meas.* **42** 301–3
Kalisz J, Pawłowski M and Pełka R 1992 *Proc. Conf. on Precision Electromagnetic Measurements CPEM'92 (Paris, 9–12 June 1992)* pp 412–13

[18] Kalisz J, Pawłowski M and Pełka R 1994 Precision time counter for laser ranging to satellites *Rev. Sci. Instrum.* **65** 736–41

[19] Räisänen-Ruotsalainen E, Rahkonen T and Kostamovaara J 1997 A high resolution time-to digital converter based on time-to-voltage interpolation *Proc. ESSCIRC'97 (Southampton, 16–18 September 1997)* pp 332–5

[20] Määttä K and Kostamovaara J 1998 High-precision time-to-digital converter for pulsed time-of-flight laser radar applications *IEEE Trans. Instrum. Meas.* **47** 521–36

[21] Baron R G 1957 The Vernier time-measuring technique *Proc. IRE* pp 21–30

[22] Barton R D and King M E 1971 Two Vernier time-interval digitizers *Nucl. Instrum. Methods* **97** 359–70

[23] Aveynier J and Van Zurk R 1970 Vernier chronotron reflex *Nucl. Instrum. Methods* **78** 161–70

[24] Otsuji T 1993 A picosecond-accuracy, 700-MHz range, Si-bipolar time interval counter LSI *IEEE J. Solid State Circuits* **28** 941–7

[25] Hoppe D R 1982 Differential time interpolator *US Patent* 4,433,919, priority: 7 September 1982
Hoppe D R 1982 Time interpolator *US Patent* 4,439,046 (the same inventor and priority)

[26] Genat J F and Rossel F 1984 Ultra high-speed time-to-digital converter *French Patent* 84 07344 *US Patent* 4 719 608, priority 1984

[27] Davis R M 1984 Apparatus for determining interval between two events *US Patent* 4,468,746, 26 August 1984

[28] Dalzell D T 1989 Electronic pulse time measurement apparatus *US Patent* 4,875,201, 17 October 1989

[29] Stephenson P S 1989 Frequency and time interval analyzer measurement hardware *Hewlett-Packard J.* **40** 35–41

[30] Rahkonen T and Kostamovaara J 1991 The use of stabilized CMOS delay lines in the digitization of short time intervals *Proc. IEEE Symp. on Circuits and Systems (Singapore, 1991)* vol 4, pp 2252–3
Rahkonen T and Kostamovaara J 1993 *IEEE J. Solid-State Circuits* **28** 887–94

[31] Räisänen-Ruotsalainen E, Rahkonen T and Kostamovaara J 1995 A low-power CMOS time-to-digital converter *IEEE J. Solid State Circuits* **30** 984–90

[32] Paschalidis N, Karadamoglou K, Stamatopoulos N, Paschalidis V, Kottaras G, Sarris E, Keath E and McEntire R 1998 An integrated time to digital converter for space instrumentation *7th NASA Symp. on VLSI Design (Albuquerque, October 1998)* University of New Mexico

[33] Kalisz J, Szplet R, Pasierbiński J and Poniecki A 1997 Field-Programmable-Gate-Array-based time-to-digital converter with 200-ps resolution *IEEE Trans. Instrum. Meas.* **46** 51–5

[34] Pełka R, Kalisz J and Szplet R 1997 Nonlinearity correction of the integrated time-to-digital converter with direct coding *IEEE Trans. Instrum. Meas.* **46** 449–52

[35] Szplet R, Kalisz J and Szymanowski R 2000 Interpolating time counter with 100 ps resolution on a single FPGA device *IEEE Trans. Instrum. Meas.* **49** 879–83

[36] Nutt R 1968 Digital time intervalometer *Rev. Sci. Instrum.* **39** 1342–5

[37] Nutt R 1970 Digital time intervalometer with analogue vernier timing *US Patent* 3,541,448, 17 November 1970

[38] Chu D C, Allen M S and Foster A S 1980 Universal counter resolves picoseconds in time interval measurements *Hewlett-Packard J.* **29** 2–10

[39] Chu D C 1979 Double Vernier time interval measurement using triggered phase-locked oscillators *US Patent* 4,164,648, 14 August 1979

[40] Kalisz J, Szplet R, Pełka R and Poniecki A 1997 Single-chip interpolating time counter with 200-ps resolution and 43-s range *IEEE Trans. Instrum. Meas.* **46** 851–6

[41] Kalisz J, Szplet R, Pełka R and Poniecki A 1998 Single-chip low-cost time counter for distance measurements with 3 cm resolution *J. Opt.* **29** 199–205

[42] Dally W J and Poulton J W 1998 *Digital System Engineering* (Cambridge: Cambridge University Press)

[43] Razavi B (ed) 1996 *Monolithic Phase-Locked Loops and Clock Recovery Circuits—Theory and Design* (New York: IEEE)

[44] Stensby J 1997 *Phase-Locked Loops—Theory and Applications* (Boca Raton, FL: CRC Press)

[45] Rothermel A and Dell'ova F 1992 Analog phase measuring circuit for digital CMOS ICs *Proc. ESSCIRC'92 (Copenhagen, 21–23 September 1992)* pp 331–3
Rothermel A and Dell'ova F 1993 *IEEE J. Solid State Circuits* **28** 853–6

[46] Loinaz M J and Wooley B A 1995 A CMOS multichannel IC for pulse timing measurements with 1-mV sensitivity *IEEE J. Solid State Circuits* **30** 1339–48

[47] Arai Y and Ikeno M 1996 A time digitizer CMOS gate-array with a 250 ps time resolution *IEEE J. Solid State Circuits* **31** 212–20

[48] Braun G *et al* 1999 F1—an eight channel time-to-digital converter chip for high rate experiments *Proc. 5th Workshop on Electronics for LHC Experiments* Snowmass

[49] Hervé C and Torki K 2002 A 75 ps rms time resolution BiCMOS time to digital converter optimised for high rate imaging detectors *Nucl. Instrum. Methods Phys. Res.* A **481** 566–74

[50] Bazes M 1985 A novel precision MOS synchronous delay line *IEEE J. Solid State Circuits* **20** 1265–71

[51] Johnson M G and Hudson E L 1988 A variable delay line PLL for CPU-coprocessor synchronization *IEEE J. Solid State Circuits* **23** 1218–23

[52] Arai Y and Ohsugi T 1989 TMC—a CMOS time-to-digital converter VLSI *IEEE Trans. Nucl. Sci.* **36** 528–31

[53] Kleinfelder S *et al* 1991 MTD132—a new sub-nanosecond multi-hit CMOS time-to-digital converter *IEEE Trans. Nucl. Sci.* **38** 97–101
[54] Ljuslin C, Christiansen J, Marchioro A and Klingsheim O 1994 An integrated 16-channel CMOS time-to-digital converter *IEEE Trans. Nucl. Sci.* **41** 1104–8
[55] Christiansen J 1996 An integrated high resolution CMOS timing generator based on an array of delay locked loops *IEEE J. Solid State Circuits* **31** 952–7
[56] Mota M and Christiansen J 1998 A four channel, self-calibrating, high-resolution, time-to-digital converter *Proc. 1998 Int. Conf. Electronics, Circuits and Systems (Lisboa, September 1998)* vol 1, pp 409–12
[57] Gorbics M S, Kelly J, Roberts K M and Sumner R L 1997 A high resolution multihit time-to-digital converter integrated circuit *IEEE Trans. Nucl. Sci.* **44** 379–84
[58] Mota M and Christiansen J 1999 A high-resolution time interpolator based on a delay locked loop and an RC delay line *IEEE J. Solid State Circuits* **34** 1360–6
[59] Bailly P *et al* 1999 A 16 channel digital TDC chip *Nucl. Instrum. Methods Phys. Res.* A **433** 432–7
[60] Mäntyniemi A, Rahkonen T and Kostamovaara J 1997 A 9-channel integrated time-to-digital converter with subnanosecond resolution *Proc. 40th Midwest Symp. Circuits and Systems, August 1997)* vol 1, pp 189–92
[61] Mäntyniemi A, Rahkonen T and Kostamovaara J 1998 An integrated CMOS time-to-digital converter with 92 ps LSB *Proc. Midwest Symp. Circuits and Systems (August 1998)* pp 180–3
[62] Mäntyniemi A, Rahkonen T and Kostamovaara J 1999 A high resolution digital CMOS time-to-digital converter based on nested delay locked loops *Proc. IEEE Int. Symp. Circuits and Systems ISCAS'99 (July 1999)* vol 2, pp 537–40
[63] Mäntyniemi A, Rahkonen T and Kostamovaara J 2002 An integrated 9-channel time digitizer with 30 ps resolution *Digest Techn. Papers, 2002 IEEE Int. Solid-State Circuit Conf. (February 2002)* vol 1, pp 266–465
[64] Mäntyniemi A, Rahkonen T and Kostamovaara J 2002 A nonlinearity corrected CMOS time digitizer IC with 20 ps precision *Proc. IEEE Int. Symp. Circuits and Systems ISCAS'2002 (May 2002)* vol 1, pp 513–16
[65] Dudek P, Szczepański S and Hatfield J V 2000 A high-resolution CMOS time-to-digital converter utilizing a Vernier delay line *IEEE J. Solid State Circuits* **35** 240–7
[66] Kalisz J, Orżanowski T and Szplet R 2000 Delay-locked loop technique for temperature stabilisation of internal delays of CMOS FPGA devices *Electron. Lett.* **36** 1184–5
[67] Pełka R 1991 Adaptive calibration of time interval digitizer with picosecond resolution *IEEE Trans. Instrum. Meas.* **40** 315–16
[68] Chmielewski K 2000 Precise digitising of time intervals by Vernier method *Doctoral Thesis* Military University of Technology, Warsaw (in Polish)
[69] Kalisz J 1988 Determination of short-term error caused by the reference clock in precision time-interval measurement and generation *IEEE Trans. Instrum. Meas.* **37** 315–16
[70] Poniecki A 1999 Robust estimation methods in metrology of time intervals with picosecond resolution *Doctoral Thesis* Military University of Technology, Warsaw (in Polish)
[71] Sondej T 2003 Efficient methods of data processing in precision laser rangefinders with high-speed microcontrollers *Doctoral Thesis* Military University of Technology, Warsaw (in Polish)
[72] http://www.vigo.com.pl
[73] Kalisz J, Poniecki A and Różyc K 2003 A simple, precise, and low-jitter delay/gate generator *Rev. Sci. Instrum.* **74** 3507–9
[74] Artyukh Y, Vedin V and Rybakov A 1998 A new approach to high performance continuous time interval counting *Proc. XI Conf. Application of Microprocessors in Automatic Control and Measurement (Warsaw, 13–14 October 1998)*
[75] Veendrick H J M 1980 The behavior of flip-flops used as synchronizers and prediction of their failure rate *IEEE J. Solid-State Circuits* **15** 169–76
[76] Horstmann J U, Eichel H W and Coates R L 1989 Metastability behavior of CMOS ASIC flip-flops in theory and test *IEEE J. Solid-State Circuits* **24** 146–57
[77] Foley C 1993 Characterizing metastability *Tektronix Report* Foley C 1993 *Proc. 2nd Int. Symp. Advanced Research in Asynchronous Circuits and Systems (Fukushima, 18–21 March 1996)* pp 175–84
[78] http://www.wavecrest.com
[79] Gross J E 1958 Time-interval measuring device *US Patent* 2,831,162, 15 April 1958
[80] Tarczy-Hornoch Z 1966 Free running multi-stable state circuit for time interval measurement *US Patent* 3,235,796, 15 February 1966
[81] Zielinski M, Chaberski D and Grzelak S 2003 Time-interval measuring module with short dead-time *Metrol. Meas. Syst.* **10** at press
[82] Sondej T, Pełka R and Poniecki A 2003 Optimized data processing in precision laser rangefinder with embedded microcontroller *Metrol Meas. Syst.* **10** at press