2 Japan Standard Time and its Related Research

2-1 Fundamentals of Time and Frequency

KAJITA Masatoshi, KOYAMA Yasuhiro, and HOSOKAWA Mizuhiko

Time (frequency) is one of the fundamental physical quantities, that can be measured most accurately among them. In this paper, the principles for the measurement of the time & frequency and the estimations of its uncertainties are summarized.

Keywords
Time and frequency measurement, International system units, Frequency measurement, Allan variance

1 History and significance of frequency

Frequency is how many times a periodic phenomenon is repeated per unit time. It has been known from antiquity that using an appropriate natural phenomenon enables us to maintain this quantity extremely accurately. What human beings first recognized as such would be the periodicity and accuracy of revolutions of celestial bodies such as the sun, the moon and stars. Utilizing this, the clock time was determined and the calendar was organized. However, these are extremely large-scaled movements that are beyond human control. While one could enumerate pitch control with stringed instruments and wind instruments from such ancient times as one of what humans created by themselves and could control, it is considered that full-fledged art started from the discovery of the isochronism of the pendulum by Galileo from the 16th century that the frequency of a pendulum depends only on its length regardless of bob weights. In the 17th century, Huygens invented pendulum clocks and furthermore mechanical clocks utilizing the accurate frequency of machinery vibration caused by a hairspring and a balance wheel. In addition to these, the standard for acoustic frequency has been provided since a tuning fork was invented in the 18th century. Following this, improvements of frequency precision made two major leaps. The first was the development of crystal oscillators originating with the discovery of the piezoelectric effect of quartz crystal by Paul-Jacques Curie and Pierre Curie in 1880. The second was the development of atomic frequency standards starting with the utilization of inversion transition of ammonia molecules in 1949. The accuracy of atomic frequency standards has reached 16 digits until today since it was first developed with an exceptional accuracy of 10 digits about 60 years ago. Furthermore, it can be said that atomic frequency standards in the optical region, whose accuracy has been dramatically improved from the outset of the 21st century, is the third leap. For the history of frequency standards, refer to [1] listed in the bibliography below.

In this manner, fairly high-accuracy frequency signals are easily available today due to highly accurate oscillators. Now, how can we measure and evaluate its precision? Phases and amplitudes are the basic elements of frequency and noise is further added to real signals. As will be discussed below, since amplitude fluctuations not usually problematic with good frequency signals, it is necessary to measure and evaluate accuracy, stability, noise etc., for the phase as a function of time with appropriate measures. This paper will give an explanation of the measuring methods for frequency signals obtained from high-accuracy frequency oscillators etc. and evaluating methods for these measures. **2** provides an overview the degree of the accuracy of frequency oscillators and defines quantity as the basis for frequency evaluation. **3** discusses the principle and basis for frequency measurement and **4** explains its evaluating methods. **5** offers commentary on classifications of noise and spectral expansion and **6** presents a conclusion.

2 Accuracy of time/frequency standards

2.1 Accuracy of frequency standards (Why are atomic clocks accurate?)

Along with technological developments for oscillators discussed in **1**, the accuracy of time displayed on a clock has been improved. The frequency of the pendulum of a pendulum clock developed in the 17th century can be expressed by:

$$T = 2\pi\sqrt{l/g} \qquad (1)$$

Here, l stands for the length of the pendulum and g for gravitational acceleration. Still, this period cannot be always stable as (a) the length of the pendulum varies at a rate of about 10^{-6} with a change in temperature by 1K, and (b) gravitational acceleration varies at a rate of about 10^{-7} with a locational change on the earth, especially with a change in altitude by 1m. It holds true also with crystal clocks, whose oscillating frequencies depend on temperature etc., that clocks gain or lose as temperature changes. However, crystal clocks are more stable than pendulum clocks as the rate of vibration frequency to frequency width (called Q value) for the former is higher than that for the latter by 5–6 digits.

Soon after the advent of atomic clocks, the accuracy of clocks was improved by about 3 digits compared with the crystal clocks available up until that time. Atomic clocks take the frequency of electromagnetic waves absorbed and emitted by atoms (neutral, ions) as a standard. Originally called “atomic frequency standards,” atomic clocks do not function as clocks until they are used together with crystal clocks etc. and feedback errors. The differences between atomic clocks and conventional clocks are detailed below:

(i) Electrons ranging around a nucleus can only acquire energy intermittently. The energy conditions are unique since they can only be determined by Coulomb force between electrons and a nucleus. An atom can absorb and emit electromagnetic waves with frequency (ν) corresponding to difference in each level energy (E_i) ($\nu = E_i - E_k / h$. Here, h stands for Planck constant).

(ii) An atomic clock, unlike conventional clocks, takes a single atom (in conditions of gas) as a standard and accordingly its condition is unique (interactions between neighboring atoms and molecules etc., are not unique in a solid state).

In short, atomic clocks are the adaptation of microcosmic phenomena to macro world. As a result, even commercially available atomic clocks ensure an accuracy of more than 12 digits. This means that it will take several tens of thousands of years to generate an error of one second. The advent of atomic clocks has enabled us to experimentally demonstrate the relativistic effect that time passes more slowly by force of movements and gravity.

However, also atomic clocks have limited accuracy. What was shown in (i) above is not completely correct since there is a width of $\Delta\nu = 1/2\pi\tau$ in accordance with a limited time of τ because of uncertainty principle for energy and time (a fundamental principle of quantum mechanics). τ is sometimes limited by the interaction time between atoms and electromagnetic waves and sometimes determined by the lifespan of quantum state with a fixed phase (limited by spontaneous emission, and collision etc.). Also, electron orbits can be distorted by external electric fields and magnetic fields to cause transition frequency shifts (one caused by electric fields is called Stark shift and one caused by magnetic fields Zeeman shift). What was shown in (ii) is also not completely accurate; even gaseous atoms collide against one another or the wall of a container. Consequently, the transition frequency shift at that moment appears as collision shift. There are also frequency shifts caused by relativistic effects from the movements and gravity of atoms that freely fly around (second-order Doppler shift and gravitational shift).

Once laser cooling could decrease the temperature of atoms as far as several μK, they became sluggish and could be made to stand still at one place, which enabled us to take longer measuring time τ. In addition, with further improvements of the accuracy of clocks, the second-order Doppler shift decreased below 18 digits, cesium fountain atomic clocks have been able to measure with an accuracy of 16 digits.

Previous clocks took microwave frequency 1–50GHz as a standard that required a measurement period of more than 10 days in order to gain an accuracy of 16 digits. With the frequency shift and uncertainty on the same level, taking high frequencies as a standard is more advantageous for shortening measuring time, and many atomic quantum transitions in the optical region have narrow line width and are accordingly suitable for highly accurate frequency standards. Furthermore, because of improvements of laser technology, the possibility of determining time and frequency at a higher accuracy has been arising. However, there was a major hurdle to measure frequency in the optical region without impairing the accuracy of frequency standards in the microwave range. Frequency measurement in the optical region had not been performed until the development of femtosecond pulse laser with stable cyclic frequencies enabled us to measure optical frequencies in around 2000. At present, the accuracy of optical transition frequency of Al^+ ion is, for instance, beyond the definition by cesium and reaches 18 digits only from cumulative uncertainty factors[2].

2.2 Definition of frequency

As the accuracy of clocks improves, it becomes necessary to provide a definition for determining certain length of time. Time is one of base quantities as well as length, mass, electric current, thermodynamic temperature, amount of substance and luminous intensity that are defined in the International System of Units (SI), and its basic unit is the second (s)[3]. For expressing time units longer than a second though, the kilosecond (ks: 10^3 sec.) etc., could be used but conventional units such as minute (min: 60sec.), hour (hour: 3600sec.) etc. are used in most cases. Time units shorter than a second, millisecond (ms: 10^{-3} sec.) and microsecond (μs: 10^{-6}sec.), etc., are used as defined in SI.

Originally, one second was defined as “1/86400 of the average rotation period of the earth”. The short-term variation at that time was approximately 10^{-8}.

Consequently, following the new definition: “one second shall be 1/331556925.9747 of the revolution period of the earth” by the International Committee of Weights and Measures in 1956, the Measurement Act of Japan was also revised[4]. The accuracy of observation at this time was approximately 10^{-9}.

The world’s first atomic clock was developed with reference to the inversion transition frequency of the ammonia molecule in 1949. Subsequently in 1955, the first Cs atomic clock was developed at the National Physical Laboratory (NPL)[5]. Furthermore, after a collaborative research by NPL and the U.S. National Observatory (USNO), Cs hyperfine structure transition frequency in zero magnetic field was measured at 9 192 631 770±20 Hz by the second on the mean solar day[6]. Thereafter, Cs atomic clocks were developed at several other research institutions until 1960, and thus the utility of atomic clocks as a time standard has become to be recognized. In 1967, the plenary session of the General Conference on Weights and Measures adopted the definition: “one second shall be the time equivalent to 9192631770 times as long as the radiation period corresponding to the transition between two hyperfine structures of the ground state of cesium-113 atom" and accordingly the Measurement Act of Japan was revised in 1972[7].

However, as was discussed in **2.1**, transition frequencies can shift by force of electric fields, magnetic fields and atomic movements, etc. Therefore, the defined frequency of the cesium atom is "frequency with both the electric and magnetic fields are zero and on the condition of the atom standing still on the earth's geoid". What are known as primary frequency standards are atomic clocks made to be able to evaluate all the uncertainty factors for themselves so that the defined frequency as such can be presented without compared with other clocks. Since the cesium atom cannot be actually measured in the above-mentioned ideal condition, primary frequency standards present the defined values of time and frequency by monitoring factors causing frequency shifts (the magnitude of electric and magnetic fields etc.) and making corrections with the estimation of frequency shifts included in measured frequencies[8]-[10].

3 Basis for measurement

3.1 Frequency measurement principles

In this Chapter, we will discuss the measurement methods for frequencies or periods of alternating current signals. Since frequency is a reciprocal of period and vice versa, either can be determined by measuring the other.

3.1.1 Measurement by electric counters [11][12]

(1) Frequency Measurement Method

First, AC input signals are attenuated and amplified as far as appropriate magnitude to be modified into pulse forms. Meanwhile, output frequencies of 1, 5, 10MHz etc. that are constituted by a high stability oscillator are divided by powers of 10 and selected by the dividing ratio selecting switch to be made into clock signals of 1, 10, 100ns etc. The frequency direct count method of input signals is the method that evaluates:

$$f = \frac{N}{\tau} \tag{2}$$

as the average frequency for τ seconds when opening the gate with clock signals for τ seconds, wherein input signals of N period pass through.

(2) Period Measurement Method

There is also a method for evaluating frequency by measuring period as its reciprocal. When opening the gate only for the time of n_d period of input signals and measuring n times the clock pulse of τ_u interval,

$$T = \frac{n\tau_u}{n_d} \tag{3}$$

expresses the average period of $n\tau_u$ seconds.

(3) Time Interval Measurement Method (TIC)

The time interval measurement is performed by assigning the opening and closing operations to two input signals and then measuring the time during that span as the number of clock pulses. Suppose that the frequencies of the two signals are almost the same and their signal waveforms are sin $(2\pi\nu t + \phi_1)$ and sin $(2\pi\nu t + \phi_2)$ respectively. Then, time difference T_d for zero crossing of those two signals can be measured as:

$$T_d = n\tau_u(\phi_2 - \phi_1) = \frac{(\phi_2 - \phi_1)}{2\pi\nu} \tag{4}$$

(4) Measurement Errors

The most major factor for measurement errors is an error of ±1 count (±1 error) on the number of pulses that pass the gate time. Counting accuracy is the reciprocal of counted value. As a result, the period measurement is more advantageous for measuring low frequencies and so is the frequency direct measurement for high frequencies. Therefore, recent counters adopt the period for frequencies lower than those whose error ranges by the frequency direct measurement and the period measurement cross and frequency direct measurement for higher frequencies. Such frequency counters are called reciprocal counters.

Various ways to minimize ±1 error by mea-

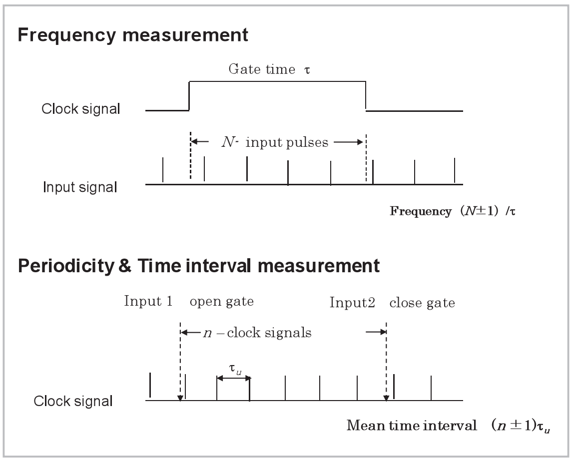

Fig.1 *The fundamental concept of the frequency, periodicity, and time interval measurements*

At the periodicity measurement, the input signals 1 and 2 are given by a common signal resorce.

suring a span shorter than pulse interval have been contrived. Major examples for that are the time expansion system that expands short time intervals immeasurable by an internal clock to measure by the same clock, the time vernier principle that synchronizes an oscillator whose frequency is slightly different from that of the reference clock with input signals to calculate the relationship between the oscillator and clock signals, the multi-phase clock system that uses multiple clocks with a slight phase shift to improve resolution, and the time-voltage conversion system that converts short time intervals less than pulse intervals into voltage to measure time from that voltage values. Among these, the time expansion system and the time vernier principle have the drawback that they need a long pause for measurement, and the multi-phase system has another that it requires a large circuit size. Therefore the time-voltage conversion system is advantageous to obtain high speed and high resolution.

Sources of error other than ±1 error are as follows[13].

(a) Trigger Errors

When input voltage signals have noise, it is converted into time error in the form of:

Trigger error = [Signal noise/Trigger voltage] × (Trigger pulse rise time).

If signal noise is smaller than trigger voltage and rise time is short, trigger error is small.

(b) Time Base Errors

Errors caused by fluctuations and drift of the frequency to be a standard for clock signals. While it is often the case that clock signals go by crystal oscillators, time base errors can be constrained if using atomic clocks as the base via an external input terminal.

(c) Trigger Level Timing Errors

These are measurement errors arising at time of measuring time intervals, which occur on the occasion that the trigger response time differs when opening and closing the gate.

For performance example, a model announced by Tektronix, Inc. in September, 2010 exerts a time resolution of 50ps and a frequency resolution of 12 digits per second[14].

3.1.2 Time difference measurement method with mixers (Dual Mixer Time Difference: DMTD) [15]

The Dual Mixer Time Difference (DMTD) is a device that measures zero crossing time intervals for two signals with the almost same frequency of ν more accurately than usual TIC. Equation (4) shows that the lower ν becomes, the smaller ±1 error becomes. This holds true for trigger errors, time base errors and trigger level timing errors. Thus, convert beat signals for the both input signals and bias signal sin $(2\pi\nu_0 t)$ into the forms of sin $[2\pi(\nu-\nu_0)t+\phi_1]$ and sin $[2\pi(\nu-\nu_0)t+\phi_2]$ to input into TIC. Here, a phase shifter makes an adjustment to hold $\phi_1-\phi_2$ constant. The principle for that is shown in Fig. 2. Time difference interval T_{dmtd} measured by this equipment is:

$$T_{dmtd}=\frac{\phi_1-\phi_2}{(\nu-\nu_0)}=\frac{\nu}{(\nu-\nu_0)}T_d \tag{5}$$

Since the error range for T_{dmtd} is as same as the case of measuring T_d directly by TIC, the error accuracy of T_d to be determined by DMTD improves by $\nu/(\nu-\nu_0)$.

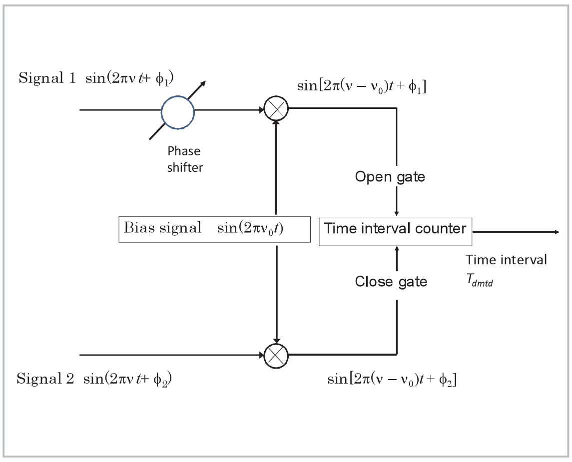

Fig.2 *The block diagram of the Dual Mixer Time Difference (DMTD)*

3.1.3 Noise measurement for frequency [16][17]

Although the frequency of an ideal frequency oscillator must be expressed by a single sine wave, an actual oscillator generates variations in oscillating output. Among these, the short-term amplitude variation is called AM noise, which is not usually brought question very much as it itself does not connect to frequency variations. On the other hand, the short-term phase variation is called phase noise, which can directly connect to frequency fluctuations and accordingly an important factor to determine characteristics of frequency sources.

The phase noise contained in an oscillator appears as the jitter in the time-domain measurement with oscilloscopes. Meanwhile, in the frequency-domain measurement with a spectrum analyzer (SA), the phase noise is observed to spread toward the bottom in the neighborhood of the carrier frequency of oscillator. This can be considered that it is spreading oscillating output energy also toward adjacent frequencies other than carrier output frequency.

It is a fundamental rule for evaluating phase noise that the intensity ratio of the offset frequency component with a gap of f_m from carrier frequency f_0 to the carrier is evaluated by dBc as follows:

$$P(f_m) = -10Log_{10}\left[\frac{I(f_0+f_m)}{I(f_0)}\right] \tag{6}$$

However, frequency components observed when phase noise is actually evaluated by SA are electric power within the finite bandwidth of Δf, which can be expressed by:

$$E(f) = \int_{f-\Delta f}^{f+\Delta f} I(f')df' \tag{7}$$

When linewidth Δf_0 for carrier frequency components is adequately narrower than Δf, they are expressed by:

$$\begin{aligned} E(f_0) &\approx I(f_0)\Delta f_0 \\ E(f_0+f_m) &\approx I(f_0+f_m)\Delta f \end{aligned} \tag{8}$$

And since values observed as:

$$\begin{aligned} L(f_m) &= -10Log_{10}\left[\frac{E(f_0+f_m)}{E(f_0)}\right] \\ &\approx -10Log_{10}\left[\frac{I(f_0+f_m)}{I(f_0)}\times\frac{\Delta f}{\Delta f_0}\right] \end{aligned} \tag{9}$$

vary according to a bandwidth of measurement, the bandwidth must be specified. Thus, since this is inconvenient for comparing measurements with different bandwidths, it is common to evaluate phase noise by using a power density of $E(f_0+f_m)/\Delta f$ per 1 Hz bandwidth. In reality, it is much more common to display what expresses:

$$\begin{aligned} P(f_m) &= -10Log_{10}\left[\frac{E(f_0+f_m)}{E(f_0)}\times\frac{1}{\Delta f}\right] \\ &\approx -10Log_{10}\left[\frac{I(f_0+f_m)}{I(f_0)}\times\frac{1}{\Delta f_0}\right] \end{aligned} \tag{10}$$

based on the unit of dBc/Hz together with offset frequencies.

The drawbacks at time of the measurement of phase noise with SA are that phase noise lower than SA's phase noise cannot be measured and that phase noise and amplitude noise are indistinguishable. The Phase Lock Loop (PLL) method is superior to measurements with SA in that since it constrains carrier signals by phase-locking of the reference oscillator for a signal source to be measured by a phase difference of $\pi/2$, it can measure pure phase noise with amplitude noise removed in high dynamic ranges. One can find a correlation between measured results of two signal sources to reduce phase noise of measuring instrument itself. Still, there is a drawback for PLL method that it can measure finite offset frequencies.

3.2 Uncertainty of frequency measurement

3.2.1 Average and variation for measurement samples

If one wants to obtain high-accuracy measured values, they should repeat measurements as often as possible and apply averaging operations to obtained data to raise accuracy. In this case, not only average values but also variations in measured data are important factors.

The ideas of average and variation will be demonstrated with a simple example below. Let probabilities in the binomial distribution taking conditions A and B be p and $1-p$ respectively. Assuming that probability for the number of samples taking A condition among N items of samples to be n is $P(n)$, the average value of n is expressed as:

$$\bar{n} = \sum_{n=0}^{N} nP(n) \tag{11}$$

Here, when using:

$$P(n) = \sum_{n=0}^{N} \frac{N!}{n!(N-n)!} p^n (1-p)^{N-n} \tag{12}$$

the following can be obtained:

$$\begin{aligned}\bar{n} &= Np \sum_{n=0} \frac{(N-1)!}{(n-1)!(N-n)!} p^{n-1} (1-p)^{N-n} \\ &= Np[p + (1-p)]^{N-1} = Np\end{aligned} \tag{13}$$

Next, Standard Deviation σ that is often used as the parameter for variations in samples is expressed as:

$$\begin{aligned}\sigma^2 &= \sum_{n=0}^{N} P(n)(n-\bar{n})^2 \\ &= \sum_{n=0}^{N} P(n)[n(n-1)+n] - (Np)^2 \\ &= N(N-1)p^2 + Np - (Np)^2 = Np(1-p)\end{aligned} \tag{14}$$

Thus, the ratio of Standard Deviation to average value becomes:

$$\frac{\sigma}{\bar{n}} = \sqrt{\frac{1-p}{Np}} \tag{15}$$

which proves to vary inversely with the square root of N (the number of samples).

The least square method is often used to estimate a most probable value from the results of repeated measurements. Consider the case that there is parameter set λ to be determined and most probable value X_j to jth measurement is given by function $X_j(\lambda)$ of λ. Then, evaluate parameter λ that makes the following least:

$$\Phi(\lambda) = \frac{1}{2} \sum_{j=1}^{N} \left[x_j - X_j(\lambda)\right]^2 \tag{16}$$

3.2.2 Uncertainty of time/frequency

In general, there are factors that cause gaps of measured values from true values: ones that cause statistical dispersion in spit of zero average and ones that cause systematic gaps from true values for certain reasons. Uncertainty when estimating true values from measured values can be estimated from uncertainty u_A resulted from statistical factors (Type A) and systematic uncertainty u_B(Type B) as:

$$u^2 = u_A^2 + u_B^2 \tag{17}$$

In the above, u_A denotes the ratio of fluctuation in measured values to average values. Suppose that fluctuation haven't been estimated by Standard Deviation; then the ratio becomes small if the number of measured data is increased as shown in Equation (15). In order to make u_B small, it is necessary to identify factors to be causes for gaps from true values and add corrections per measurement condition. The measured value distribution for the case that u_A and u_B are large and small will be shown in Fig. 3.

An important factor for estimation of u_A when really comparing time/frequency between two atomic clocks is the dispersion of the difference in time/frequency between the two clocks and it does not matter whether or not their average values agree.

Meanwhile, what determines u_B is whether the gaps from true values caused by electric fields, magnetic fields, relativistic effects etc., can be made small or precisely estimated to correct.

In case of u_B, how much the average of measured values of two clocks each agrees serves as a standard for estimating gaps from true values. In order to minimize u_B, it is necessary to perform frequency measurements under various conditions and determine parameters to be shift factors by means of the least square method. Specifically, it can be pointed out that one measures variations in measured frequencies by changing magnetic field sizes in order to identify the influence of Zeeman effects and estimates frequency in zero magnetic field, and so on.

4 Basis for evaluation of frequency stability [11]

In this Chapter, we will discuss what parameters one should use to estimate the frequency changes necessary to estimate statistical uncertainty u_A for time/frequency, which is useful when comparing and discussing characteristics of clocks measured under different conditions. While frequency changes to be discussed below mainly cover frequency fluctuations that randomly change over time, the frequency drift that frequency changes in a single direction over time is also partly discussed.

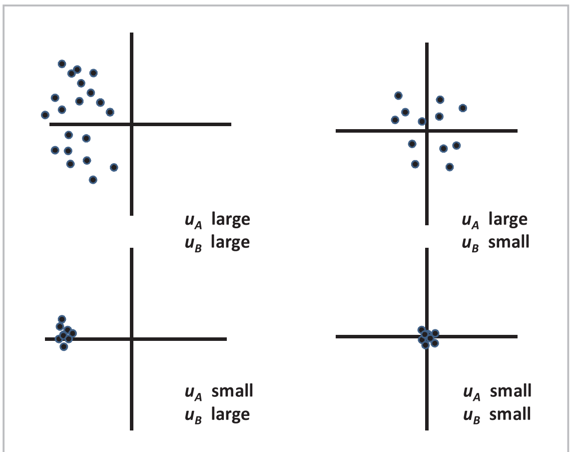

Fig.3 The distribution of measurements with large (or small) statistical uncertainty u_A and systematic uncertainty u_B

Output signal $V(t)$ when discussing frequency stability is expressed as follows:

$$V(t) = V_0[1 + \varepsilon(t)]\sin[2\pi\nu_0 t + \phi(t)] \quad (18)$$

Here, V_0 and ν_0 are amplitudes of output signals and averages of frequencies, and ε and ϕ denote amplitude fluctuations and phase fluctuations. Since the actually obtained amplitude can be mostly regarded as a fairly small fluctuation for estimating frequency changes, V_0 can be usually deemed as invariable. Suppose instantaneous phase $\Phi(t)$ as:

$$\Phi(t) = 2\pi\nu_0 t + \phi(t) \quad (19)$$

then instantaneous frequency $\nu(t)$ is evaluated by:

$$\nu(t) = \frac{1}{2\pi}\frac{d\Phi(t)}{dt} = \nu_0 + \frac{1}{2\pi}\frac{d\phi(t)}{dt} \quad (20)$$

Here, assume:

$$|\varepsilon| << 1 \qquad \frac{1}{2\pi\nu_0}\left|\frac{d\phi(t)}{dt}\right| << 1 \quad (21)$$

Then, define dimensionless parameter $y(t)$ standing for frequency changes as:

$$y(t) = \frac{\nu(t)}{\nu_0} - 1 = \frac{1}{2\pi\nu_0}\frac{d\phi(t)}{dt} \quad (22)$$

Also, define time change $x(t)$ as:

$$x(t) = \int y(t)dt = \frac{\phi(t)}{2\pi\nu_0} \quad (23)$$

As shown in **2**, $y(t)$ cannot actually measure a value for that moment but one averaged by frequency counter's gate time τ:

$$\bar{y}_k(\tau) = \frac{x(t_k + \tau) - x(t_k)}{\tau} = \frac{1}{\tau}\int_{t_k}^{t_k+\tau} y(t)dt \quad (24)$$

In order to estimate the size of frequency changes, it would be natural to first evaluate the dispersion:

$$\left\langle \bar{y}_k(\tau)^2 \right\rangle = \lim_{K\to\infty}\frac{1}{K}\sum_{k=1}^{K}\bar{y}_k(\tau)^2 \quad (25)$$

of frequency changes:

$$\left\langle \overline{y}_k(\tau)^2 \right\rangle = \left\langle \frac{[x(t_k+\tau)-x(t_k)]^2}{\tau^2} \right\rangle$$ (26)

$$= \frac{1}{\tau^2}\left[\left\langle x(t_k+\tau)^2\right\rangle + \left\langle x(t_k)^2\right\rangle - 2\left\langle x(t_k)x(t_k+\tau)\right\rangle\right]$$

Here, define the autocorrelation function for phase and frequency:

$$R_x(\tau) = \lim_{T\to\infty}\frac{1}{T}\int_0^T x(t)x(t+\tau)dt$$

$$R_y(\tau) = \lim_{T\to\infty}\frac{1}{T}\int_0^T y(t)y(t+\tau)dt$$ (27)

$R_x(\tau)$ is based on the unit of s^2 and $R_y(\tau)$ is non-dimensional. Then, one can readily understand:

$$R_y(0) = \left\langle y(t)^2 \right\rangle$$ (28)

Autocorrelation function is a parameter used to indicate how much effect remains after gaps from average values of phase and frequency at a certain time change only by time τ. By using this parameter, Equation (26) can be modified into:

$$\left\langle \overline{y}_k(\tau)^2 \right\rangle = \frac{2[R_x(0)-R_x(\tau)]}{\tau^2}$$ (29)

As will be discussed below, speed for the time change of phase should be considered with frequency components to be more understandable. For that, consider power spectral density $S_{x,y}(f)$ that is the application of Fourier transform to the autocorrelation function of phase change:

$$S_{x,y}(f) = 4\int_0^\infty R_{x,y}(\tau)\cos(2\pi f\tau)d\tau$$

$$R_{x,y}(\tau) = \int_0^\infty S_{x,y}(f)\cos(2\pi f\tau)df$$ (30)

$S_x(f)$ is based on the unit of s^2/Hz and $S_y(f)$ on the unit of /Hz. The case of $\tau = 0$ leads to:

$$\left\langle y(t)^2 \right\rangle = \int_0^\infty S_y(f)df$$ (31)

which indicates that the power spectral density represents the frequency components of frequency fluctuations. Using the power spectral density leads to:

$$\left\langle \overline{y}_k(\tau)^2 \right\rangle = \frac{2}{\tau^2}\int_0^\infty S_x(f)[1-\cos(2\pi f\tau)]df$$

$$= \frac{4}{\tau^2}\int_0^\infty S_x(f)\sin^2(\pi f\tau)df$$ (32)

Furthermore, consider that temporal differentiation in the time domain corresponds to multiplication by $2\pi f$ in the frequency domain, which leads to:

$$S_y(f) = (2\pi f)^2 S_x(f)$$ (33)

then Equation (32) can be expressed by:

$$\left\langle \overline{y}_k(\tau)^2 \right\rangle = \int_0^\infty S_y(f)\frac{\sin^2(\pi f\tau)}{(\pi f\tau)^2}df$$ (34)

In reality, the power spectral density of frequency changes can be expressed by the polynomial in f:

$$S_y(f) = \sum h_a f^a$$ (35)

It is often the case from the past measured data that discussions go on in the range of $a = -2, -1, 0, 1, 2$. Here, each term denotes as follows:

- $a = -2$ Random walk FM noise
- -1 Flicker FM noise
- 0 White FM noise
- 1 Flicker PM noise
- 2 White PM noise

White PM and FM noise and flicker noise will be discussed in **5** below.

Now, when viewing Equation (31), it turns out that this integration diverges with $a = -2, -1$ in $f \to 0$ range. It does not actually reach an infinite value as the residual of frequency changes is evaluated from finite measured data following Equation (22). Still, since acquired values depend on how to draw samples, this integration is not preferable as a stability scale. This divergence of integral values represents that frequency is endlessly changed while sampling is repeated with random walk and flicker FM noises over an extended time period. When evaluating the residual of frequency changes, the average frequency to be a standard cannot be obtained until the entire measurement is completed, which is also problematic.

Then, samples for frequency changes of a finite time period are taken by N sample variance that can be defined below:

$$\sigma_y^2(N,T,\tau)=\left\langle \frac{1}{N-1}\sum_{n=1}^{N}\left(\bar{y}_n-\frac{1}{N}\sum_{k=1}^{N}\bar{y}_k\right)^2\right\rangle$$
$$=\int_0^\infty S_y(f)\frac{\sin^2(\pi f\tau)}{(\pi f\tau)^2}\left\{1-\frac{\sin^2(N\pi fT)}{N^2\sin^2(\pi fT)}\right\}df \quad (36)$$

Here, T denotes repeated frequencies for measurement of τ seconds and $t_{k+1}=t_k+T$ ($k=0, 1, 2, ,,,,$) is effected. It means that each measurement involves a blank time of $T-\tau$. Of these, N sample variance under conditions of $T=\tau$ and $N=2$ is especially called Allan variance σ_y^2:

$$\sigma_y^2(\tau)=\frac{1}{2}\left\langle (y_{k+1}-y_k)^2\right\rangle$$
$$=\frac{1}{2}\left(\left[x(t_k+2\tau)-2x(t_k+\tau)+x(t_k)\right]^2\right) \quad (37)$$
$$=2\int_0^\infty S_y(f)H(\pi f\tau)df$$
$$H(\pi f\tau)=\frac{\sin^4(\pi f\tau)}{(\pi f\tau)^2}$$

$H(f)$ will be shown in Fig. 4. It implies that there is no divergence in $f\rightarrow 0$ range for the Allan variance even with $a=-2, -1$. This is also the measured result of frequency changes for finite measuring time τ, which corresponds to the fact that there is no problem with its sampling for over an infinite time period.

Allan variance's τ-dependence will be discussed in the following way:

$$-3<a<1$$
$$\sigma_y^2(\tau)=2h_a\int_0^\infty f^a\frac{\sin^4(\pi f\tau)}{(\pi f\tau)^2}df$$
$$=2h_a(\pi\tau)^{-(a+1)}\int_0^\infty q^{a-2}\sin^4(q)dq\propto\tau^{-(a+1)}$$
$$q=\pi f\tau$$

$$a=-2\quad \sigma_y^2(\tau)=\frac{2\pi^2h_{-2}\tau}{3}$$
$$a=-1\quad \sigma_y^2(\tau)=2h_{-1}\ln 2 \quad (38)$$
$$a=0\quad \sigma_y^2(\tau)=\frac{h_0}{2\tau}$$

This result can be easily understood from that $f_{max}^a\times\Delta f$ is proportional to $\tau^{-(a+1)}$ as $H(f)$ peaks at around $f_{max}=3/(2\pi\tau)$ and integral width Δf is approximately $1/(\pi\tau)$. To express this further in different words, it corresponds to dividing the magnitude of fluctuations with frequency component f_{max} by the sample size proportional to τ. The above-mentioned infinite integral will be infinite with $a=1, 2$. Then, consider the case of using a low-pass filter that shields frequency components higher than maximum frequency component f_h. In this case, Equation (37) must be replaced with the finite integral of:

$$\sigma_y^2(\tau)=2h_a\int_0^{f_h}f^aH(\pi f\tau)df$$
$$=2h_a(\pi\tau)^{-(a+1)}\int_0^{q_h}q^aH(q)dq \qquad q_h=\pi f_h\tau \quad (39)$$

It results in:

$a=1$

$$\sigma_y^2(\tau)=\frac{3h_1}{(2\pi\tau)^2}\ln(2\pi f_h\tau) \quad (40)$$

$a=2$

$$\sigma_y^2(\tau)=\frac{3h_2f_h}{(2\pi\tau)^2} \quad (41)$$

Thus, since the case of either $a=1$ (flicker PM noise) or 2 (white PM noise) approximately shows τ^{-2}-dependence, the two kinds of noise will be indistinguishable. This result is caused by the fact that since $H(f)$ reaches a local maximum in all frequencies expressed by $f=3/(2\pi\tau)\times(2m+1)$ (m is an integer) (ref. Fig. 4) and all the frequency components for $a=2$ get involved evenly with the Allan variance, the interpretation that the Allan variance stands for the frequency components for $3/(2\pi\tau)$ of frequency fluctuations turns out to be invalid.

What was contrived to differentiate between the cases of $a = 1$ (flicker PM noise) and $a = 2$ (white PM noise) was the Modified Allan variance $Mod\sigma_y^2(\tau)$. Let τ be the integral multiple of least-time unit τ_0 that is $\tau = n\tau_0$, then $Mod\sigma_y^2(\tau)$ is defined:

$$Mod\sigma_y^2(\tau)=\frac{1}{2}\left\langle\left[\frac{1}{n}\sum_{j=1}^{n}\left(\bar{y}_{j+n}-\bar{y}_j\right)\right]^2\right\rangle$$

$$\bar{y}_{j+n}(\tau)=\frac{1}{\tau}\int_{t_k+(j+n)\tau_0}^{t_k+(j+2n)\tau_0} y(t)dt \qquad (42)$$

$$\bar{y}_j(\tau)=\frac{1}{\tau}\int_{t_k+j\tau_0}^{t_k+(j+n)\tau_0} y(t)dt$$

Calculate Equation (42) to lead to[18]:

$$Mod\sigma_y^2(\tau)=2\int_0^\infty S_y(f)H_{n-\mathrm{mod}}(f)df$$

$$H_{n-\mathrm{mod}}(f)=H(f)M_n(f) \qquad (43)$$

$$M_n(f)=\frac{\sin^2(\pi f\tau)}{(\pi f\tau)^2}\frac{(\pi f\tau/n)^2}{\sin^2(\pi f\tau/n)}$$

Looking at the above, it turns out that $M(f)$ attenuates as $(\pi f\tau)$ grows while it is approximate to 1 in $(\pi f\tau)<1$ range. Figure 4 shows $H_{10\text{–mod}}(f)$ ($H_{n\text{–mod}}(f)$ of $n = 10$) to compare with $H(f)$. Since contribution to the integral of Equation (43) is limited within $(\pi f\tau)<1$ range with $a<1$, $Mod\sigma_y^2(\tau)$ is about the same as $\sigma_y^2(\tau)$. However, a large difference arises when the part of $(\pi f\tau) \gg 1$ largely contributes and $Mod\sigma_y^2(\tau)/\sigma_y^2(\tau)$ is inversely proportional to n in case of $a = 2$. $Mod\sigma_y^2(\tau)$ becomes proportional to τ_0/τ^3 to show τ-dependence that is largely different from the case of $a = 1$. This can be interpreted in a way that while all frequency components for $f = 3/(2\pi\tau)\times(2m+1)$ make an even contribution in the Allan variance, $Mod\sigma_y^2(\tau)$ becomes proportional to $\tau^{-(a+1)}$ also in case of $a = 2$ as in case of $a<1$ as only frequency components for $f = 3/(2\pi\tau)$ contribute in the Modified Allan variance.

In Fig. 5, τ-dependence of the Allan variance and the Modified Allan variance is shown.

Other than frequency fluctuations of the kinds that were considered in the above, there is frequency drift. Frequency drift has properties in proportion to τ^2 both in the Allan variance and the Modified Allan variance. Measured drift is usually considered separately from random noise, and yet it corresponds to f^{-3} FM noise as long as it is not removed. Both the Allan variance and the Modified Allan variance that are evaluated respectively from Equations (37) and (43) will diverge for this noise.

Frequency fluctuations also do not necessarily take forms that can be expressed by f^a. An example of this is in the case that the phase is modulated with the sinusoidal wave:

$$\phi=\phi_m\sin(2\pi f_m t) \qquad (44)$$

In this case, from:

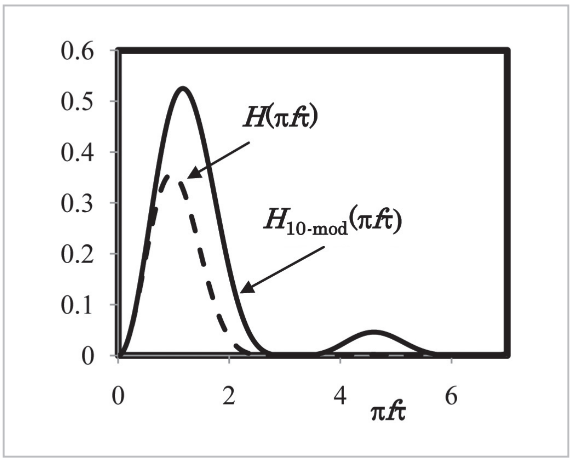

Fig.4 $H(\pi f\tau)$ and $H_{10\text{-}mod}(\pi f\tau)$ as a function of $\pi f\tau$

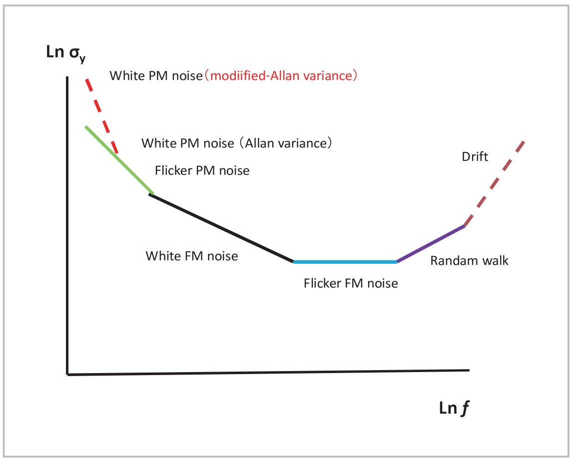

Fig.5 The dependence of Allan variance and modified-Allan variance on the measurement time τ

$$y(t) = (f_m / \nu_0)\phi_m \cos(2\pi f_m t) \quad (45)$$

$$S_y(f) = (1/2)(f_m / \nu_0)^2 \phi_m^2 \delta(f - f_m) \quad (46)$$

the following is derived:

$$\sigma_y^2(\tau) = \left(\frac{f_m}{\nu_0}\right)^2 \phi_m^2 \frac{\sin^4(\pi f_m \tau)}{(\pi f_m \tau)^2} \quad (47)$$

This becomes with $\tau \ll 1/(2 f_m)$:

$$\sigma_y^2(\tau) \approx \left(\frac{f_m}{\nu_0}\right)^2 \phi_m^2 (\pi f_m \tau)^2 \quad (48)$$

and is consequently indistinguishable from frequency drift in this range. However, the Allan variance peaks at around $\tau = 1/2 f_m$ and decreases when τ exceeds it. This holds with the Modified Allan variance as well.

In order to really measure the stability of time and frequency, it is necessary to compare more than two frequencies and measure frequency differences between them. Variance σ_{AB} of frequency difference between frequencies A and B can be evaluated from variances σ_A and σ_B of frequencies A and B each by:

$$\sigma_{AB}^2 = \sigma_A^2 + \sigma_B^2 \quad (49)$$

In case of $\sigma_A \gg \sigma_B$, measured σ_{AB} can be safely regards as σ_A. However, when the stabilities of each frequency are comparable (when measuring the highest level of stability in particular), it is necessary to estimate stabilities of each frequency by using three frequencies in the form of:

$$\sigma_A^2 = \frac{\sigma_{AB}^2 + \sigma_{AC}^2 - \sigma_{BC}^2}{2} \quad (50)$$

5 Spectrum for oscillating frequency changes [11]

The dependence of noise that determines uppermost limits of frequency stability on frequency source properties will be discussed below.

FM noise occurs due to noise within the oscillating loop and PM noise due to that outside of the loop. This corresponds to the fact that phase changes are observed as phase changes as they are since there is no feedback outside of the oscillating loop, but they are observed as frequency changes to offset phase changes that can occur within the loop. The higher speed phase changes within the loop, the smaller oscillating frequency changes become. Thus, the dependence of PM noise on f differs from that of FM noise. Suppose that the maximum width of phase noise $\Delta\theta$ is 2π and the maximum frequency change is spectral width $\Delta\nu$ of the oscillator; then the correlation between frequency fluctuations ($y = \delta\nu/\nu_0$) and phase changes ($x = \Delta\theta/\nu_0$) is estimated as:

$$\frac{\delta\nu}{\Delta\nu} = \frac{\Delta\theta}{2\pi}$$
$$y = \frac{\delta\nu}{\nu_0} = \frac{\Delta\theta}{2\pi Q} = \frac{x\nu_0}{2\pi Q} \qquad Q = \frac{\nu_0}{\Delta\nu} \quad (51)$$

Consequently, the following is derived:

$$S_y(f) = \left(\frac{\nu_0}{2\pi Q}\right)^2 S_x(f) \quad (52)$$

Meanwhile, phase noise (PM noise) outside of the loop becomes:

$$\delta\nu = (2\pi f)\Delta\theta \quad y = (2\pi f)x$$
$$S_y(f) = (2\pi f)^2 S_x(f) \quad (53)$$

Now, as was shown in **4**, phase noise is of two kinds: one called white noise (without dependence on f) and the other flicker noise (in proportion to f^{-1}). White noise is thermal noise when ν_0 is in the microwave region and quantum noise (fluctuations in photon number) when it is in the optical region. This difference owing to the regions of ν_0 is caused by different impacts that the discontinuity of light energy makes on fluctuations according to whether one-photon energy ($h\nu_0$) is larger or smaller than $k_B T_0$ (T_0 is the absolute temperature of the oscillator). Flicker noise occurs through phase modulation with the nonlinearity of circuit elements and such. Express the power spectral density of phase modulation within and outside the loop as:

$$S_{x-in}(f) = \alpha_{in} f^{-1} + \beta_{in}$$

$$S_{x-out}(f) = \alpha_{out} f^{-1} + \beta_{out}$$

Then, the following is derived:

$$S_y(f) = h_{-1} f^{-1} + h_0 f^0 + h_1 f^1 + h_2 f^2$$

$$h_{-1} = \left(\frac{\nu_0}{2\pi Q}\right)^2 \alpha_{in} \qquad h_0 = \left(\frac{\nu_0}{2\pi Q}\right)^2 \beta_{in} \quad (54)$$

$$h_1 = 4\pi^2 \alpha_{out} \qquad h_2 = 4\pi^2 \beta_{out}$$

Furthermore, consider random walk noise ($\propto f^{-2}$) and apparent noise due to linear frequency drift ($\propto f^{-3}$); then they can be expressed by:

$$S_y(f) = h_{-3} f^{-3} + h_{-2} f^{-2} + h_{-1} f^{-1} + h_0 f^0 + h_1 f^1 + h_2 f^2 \quad (55)$$

The dependence of the Allan variance on measuring time τ was shown in Fig. 5. The Allan variance decreases in short-range τ as taking longer τ. It is often the case that white FM noise is dominant for Cs atomic clocks etc. In such a case, the Allan variance can be kept down by raising Q-values. For hydrogen masers etc., mostly PM noise is more dominant than white FM noise as (ν_0/Q) is low. If taking τ longer than a certain time, flicker FM noise ($\propto f^{-1}$) becomes conspicuous. Thus, the Allan variance becomes constant to τ and accordingly sets a limit on frequency stability. As was shown in Equation (54), limit values of the Allan variance due to flicker FM noise become small when Q-values of the oscillator rise. When Q-values become high, the Allan variance noise can be made small even in the region where white FM noise is dominant.

If taking measuring time τ even longer, the Allan variance adversely becomes large. This holds either due to random walk or frequency drifts.

The dependence of the Allan variance on τ does not necessarily take as simple a form as described above. For instance, when the phase is sinusoidally modulated, as shown in Equation (47), the Allan variance becomes an increasing function of τ in a short τ-range and a decreasing function in a long τ-range.

6 Conclusion

Time/frequency are physical quantities that are measurable with the highest degree of accuracy and play an important role for contemporary science and technology. Their measurement and evaluation has a close connection with the passage of time and averaging time, and thus has aspects that cannot be evaluated only by average and Standard Deviation as the case of other qualities. The stability scale of the Allan variance is a typical instance; it forms an important basis for measurement and evaluation of frequencies to understand how it differs from normal Standard Deviation and relates to the power spectral density and types of noise. For understanding of this, the present paper offered a general consideration of measurement techniques for time/frequency and provided an outline of crucial ideas for estimating its uncertainty. It will be the author's pleasure if it helps relevant parties work on measurement and evaluation of frequency.

References

1 M. Hosokawa, "Physics Education at University (Phys. Soc. of Japan)," Vol. 14, No. 3, p. 125, 2008. (in Japanese)

2 C.W. Chou, D. B. Hume, J. C. Koelmeij, D. J. Wineland, and T. Rosenband, Phys. Rev. Lett., Vol. 104, pp. 070802(1–4), 2010.

3 http://www-lab.ee.uec.ac.jp/text/misc/si_units.html

4 Japanese law 61st , 3-3 (April 15th 1958). (in Japanese)

5 L. Essen and J. V. L. Parry, Nature, Vol. 176, pp. 280–282, 1955.

6 W. Markovitz, R. G. Hall, L. Essen, and J. V. L. Parry, Phys. Rev. Lett., Vol. 1, pp. 105 –107, 1958.

7 Japanese law 27th , 3-3 (May 9th, 1972). (in Japanese)

8 M. Kumagai et al., "Caesium Atomic Fountain Primary Frequency Standard NICT-CsF1," Special issue of this NICT Journal, 2-3, 2010.

9 K.Matsubara et al., "$^{40}Ca^{+}$ Ion Optical Frequency Standard," Special issue of this NICT Journal, 3-2, 2010.

10 A. Yamaguchi et al., "A Strontium Optical Lattice Clock," Special issue of this NICT Journal, 3-3, 2010.

11 K. Yoshimura, Y. Koga, and N. Oura, "Frequency and Time – Fundamental of Atomic clock & System of Atomic Time," The Inst. of Electro., Info. and Comm. Eng., 1989. (in Japanese)

12 http://www.yokogawa.co.jp/tm/TI/keimame/torjikan/kouzou.htm

13 http://www.yokogawa.co.jp/tm/TI/keimame/torjikan/kakudo.htm

14 http://ednjapan.cancom-j.com/news/2010/9/7224

15 F. Nakagawa, M. Imae, Y. Hanado, and M. Aida, IEEE Trans. Inst. Meas. Vol. 54, pp. 829–832, 2005.

16 http://em-field.jp/reduce-pn/l.html

17 http://gate.ruru.ne.jp/rfdn/TechNote/SpaCN.asp

18 J. Vanier and C. Audoin, "The Quantum Physics of Atomic Frequency Standard," p. 251, Adam Hilger, Bristol and Phyladelphia.

(Accepted Oct. 28, 2010)

KAJITA Masatoshi, *Ph.D.*

Senior Researcher, Space-Time Standards Group, New Generation Network Research Center

Atomic Molecular Physics, Frequency Standard

KOYAMA Yasuhiro, *Ph.D.*

Group Leader, Space-Time Standards Group, New Generation Network Research Center

Space Geodesy, Radio Science

HOSOKAWA Mizuhiko, *Ph.D.*

Executive Director, New Generation Network Research Center

Atomic Frequency Standards, Space-Time Measurements