IOP Publishing

Measurement, Uncertainty and Lasers

Masatoshi Kajita

Chapter 2

What is measurement uncertainty?

2.1 Introduction

To reduce measurement uncertainties, we must know their properties. In this chapter, we present various kinds of measurement uncertainties and methods to reduce them. We also introduce parameters to estimate measurement uncertainties, which depend on the use of measurements.

2.2 Statistical and systematic uncertainties

When we measure something, there is a question whether it is a real value. The reliability of measured values can be confirmed by repeating measurements many times. We will see that measurement results are distributed over a certain limited area. The uncertainty given by the finite distribution area is called 'statistical uncertainty' as shown in figure 2.1.

We cannot guarantee that the real value is in the area over which the measurement results are distributed. All measurement values can vary according to circumstances, and the real values should be defined with a certain condition. If we take measurements under another circumstance, the measured value may shift from the defined value. If we know the dependence of the measurements on the circumstances, we can make a correction of the measured values by the estimated shift. Then the uncertainty of the estimated shift becomes another uncertainty, called 'systematic uncertainty' as shown in figure 2.1.

More detailed explanations of statistical and systematic uncertainties are presented in the following subsections.

2.2.1 Statistical uncertainty

Statistical uncertainty is given by the finite broadening of the distribution area of the measurement results. This broadening can be induced by the temporal fluctuation of

doi:10.1088/2053-2563/ab0373ch2

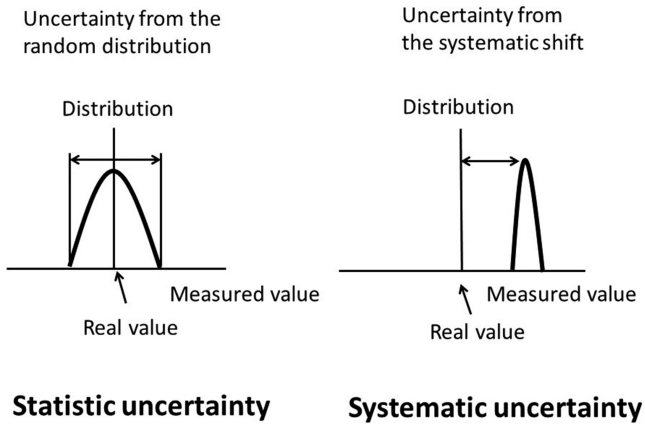

Figure 2.1. Statistic and systematic uncertainties with the distributions of measurements.

the circumstances, which can be reduced by stabilizing the circumstances. This broadening is also induced by the quantum uncertainty principle.

We consider the measurement of a physical value X. The measurement results are distributed around the real value with the measuring condition X_r. Taking measurement samples $X(i)$ $(i = 1 - N)$, we obtain the average X_{ave} and the standard deviation σ. When N is large enough, the probability distribution of X_{ave} is approximately given by (figure 2.2)

$$P(X_{\text{ave}}) = \frac{\sqrt{N}}{\sigma\sqrt{\pi}} \exp\left[-N\left(\frac{X_{\text{ave}} - X_r}{\sigma}\right)^2\right] \tag{2.1}$$

with any distribution of the measurement results (central limit theorem). Therefore, X_r is estimated to be in the range of

$$X_r = X_{\text{ave}} \pm \frac{\sigma}{\sqrt{N}}. \tag{2.2}$$

The statistic uncertainty is reduced by increasing the number of measurement samples.

It is not simple to derive the central limit theorem as a general formula [1]. Here, we derive it using the simplest model by which we get the measurement results $X_r + \sigma$ and $X_r - \sigma$ with the probability of 1/2. Repeating measurement N times, the probability $p(n)$ to measure n times $X_0 + \sigma$ and $(N - n)$ times $X_0 + \sigma$ is given by

$$p(n) = \frac{N!}{n!(N-n)!}\left(\frac{1}{2}\right)^N. \tag{2.3}$$

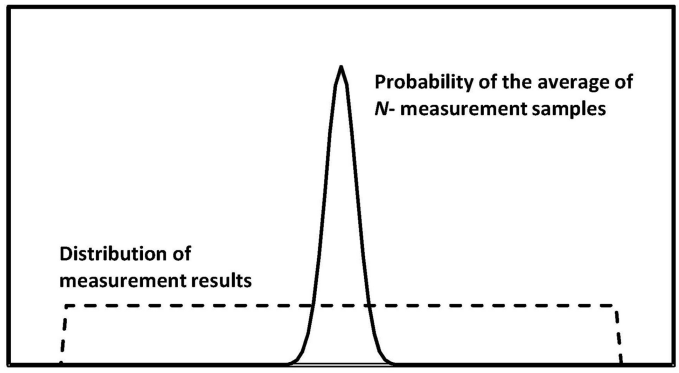

Figure 2.2. Relation between the distributions of measurement results and the probability of the average of N-measurement samples. The probability of the average of the N-measurement samples is given by a Gaussian with broadening narrower than that of measurement results with a factor of $1/\sqrt{N}$.

Here, we consider the Taylor expansion of $\ln[p(n)]$ using

$$\begin{aligned} \frac{d\ \ln[p(n)]}{dn} &= -\ln(n) + \ln(N-n), \\ \frac{d^2\ \ln[p(n)]}{dn^2} &= -\frac{1}{n} - \frac{1}{N-n}. \end{aligned} \tag{2.4}$$

Here, $\ln[p(n)]$ at n close to $N/2$ is approximately given by

$$\begin{aligned} \ln[p(n)] &= \ln p\left(\frac{N}{2}\right) - \frac{4}{N}\left(n - \frac{N}{2}\right)^2, \\ p(n) &= p(0)\exp\left[-\frac{4}{N}\left(n - \frac{N}{2}\right)^2\right]. \end{aligned} \tag{2.5}$$

The average of the measurement is given by

$$\begin{aligned} X_{\text{ave}} &= X_r + \frac{2n-N}{N}\sigma, \\ n - \frac{N}{2} &= \frac{N}{2\sigma}(X_{\text{ave}} - X_r). \end{aligned} \tag{2.6}$$

Then the probability of the average of N-sample measurements is given by

$$P(X_{\text{ave}}) = P(X_r)\exp\left[-N\left(\frac{X_{\text{ave}} - X_r}{\sigma}\right)^2\right]. \tag{2.7}$$

2.2.2 Systematic uncertainty

Statistical uncertainty is reduced by averaging many measurement samples. However, it is possible that all measurement results have a parallel shift from the real value. This is because all the measured values depend on the circumstance under which they were taken. It is not always possible to measure the values while measuring the defined value. Seeing just the measurement results, the shift of the measured value is the systematic uncertainty. Systematic uncertainty is reduced by monitoring the circumstance and giving a correction of the estimated shifts. With this correction, the systematic uncertainty is given by the uncertainty of the estimated measurement shift.

For example, thermal expansion causes a shift in the length of an object; therefore, at which temperature (T_0) the length is defined should be clarified. Repeating the measurement with another temperature T_p, the averaged length is shifted from the defined length (figure 2.3). The systematic uncertainty is reduced by monitoring T_p and giving a correction by $\alpha_p(T_0 - T_p)$, but it cannot be zero because of the uncertainties of T_p and α_p.

2.2.3 Uncertainty in atomic transition frequencies

Measurement uncertainty is particularly low with the frequency of electromagnetic waves absorbed or emitted by atoms or molecules in the gaseous state (atomic, molecular transition frequencies). Using this characteristic, atomic clocks based on the atomic transition frequencies were invented after World War II, and the uncertainty of time and frequency was reduced drastically (chapter 3). However, atomic transition frequencies also have non-zero measurement uncertainties. Here, we discuss the factors that contribute to the measurement uncertainties of atomic transition frequencies.

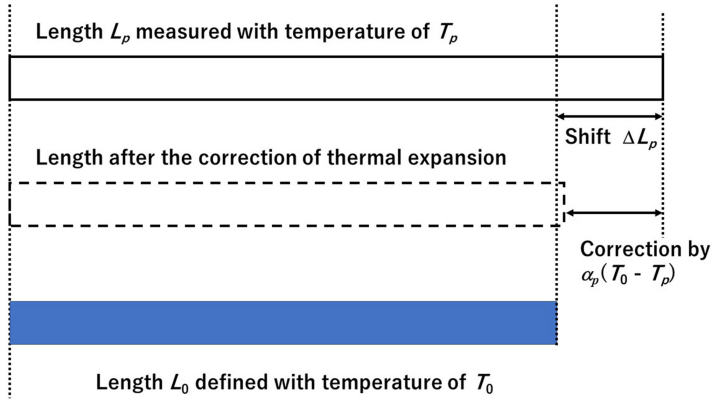

Figure 2.3. Concept of systematic uncertainty with length measurement. The length should be defined with a certain temperature T_0. Measuring with another temperature T_p, there is a shift of measurement and this gives a systematic uncertainty. This uncertainty is reduced by monitoring the temperature and giving a correction of thermal expansion $\alpha_p(T_0 - T_p)$ (α_p: thermal expansion coefficient), but it cannot be zero because of the uncertainties of T_p and α_p.

Statistical uncertainty is given by the limited interaction time τ_e between atoms and electromagnetic waves without the phase jump. The limit of τ_e is caused by several things: limited interaction time between atom and electromagnetic wave by the atomic motion, collision between atoms, or spontaneous emission transition (transition from a higher energy state to a lower energy state giving fluorescence). The transition frequency has the uncertainty of the order of $\delta f = 1/2\pi\tau_e$, because the phase procedure $2\pi f\tau_e$ must have the uncertainty of ± 1. This relation is derived from the quantum uncertainty principle between time and energy. With the real transition frequency of f_0, the distribution of the single measurement result is given by

$$I(f) = \frac{\delta f}{(f - f_0)^2 + (\delta f)^2}, \tag{2.8}$$

and δf is observed as the spectrum linewidth. The statistical uncertainty δf_{sta} is given by

$$\delta f_{\text{sta}} = \delta f \sqrt{\frac{(\tau_e + \tau_d)}{N_a T_m}} = \frac{1}{2\pi\tau_e}\sqrt{\frac{(\tau_e + \tau_d)}{N_a T_m}}, \tag{2.9}$$

where N_a is the number of atoms, τ_d is the measurement dead time (time for preparation or detection), and T_m is the measurement time.

The systematic measurement uncertainty is given by the shift of the measured atomic transition frequency caused by various things. For example, the electron orbit in the atom is distorted by the electric or magnetic fields, and the transition frequencies are shifted. The frequency shift caused by the electric field and the magnetic field are called Stark and Zeeman shifts, respectively. The Stark shift cannot be zero because the electric field is applied by the electromagnetic wave for the probe and by blackbody radiation (radiation of electromagnetic wave from objects having non-zero temperature). The Zeeman shift also exists because of the Earth's magnetic field.

Systematic uncertainty is also given by the relativistic effects as follows (chapter 6). The motion of atoms (at room temperature on the order of 200–500 m s^{-1}) gives frequency shifts called quadratic Doppler shifts. The difference in the gravitational potential also gives the frequency shift called the gravitational red shift (effect of the theory of general relativity).

The definition for the atomic transition frequency is given by the condition of zero velocity free from electric and magnetic fields on the geoid plane. Atomic clocks are based on atomic transition frequencies, where the statistical and systematic uncertainties are relatively small [2–6]. Atomic laser cooling (chapter 4) made it possible to reduce the atomic velocity to a few cm s^{-1}, and the quadratic Doppler shift was reduced significantly. Statistic uncertainty was also reduced since the development of laser cooling because of the longer interaction time between atoms and electromagnetic waves. Uncertainty on the order of 10^{-18} was obtained with several atomic transition frequencies after correction of the possible frequency shifts.

2.3 Accuracy and stability

The measurement uncertainty is discussed with regard to accuracy and stability. Figure 2.4 shows the distributions of measurements with regard to 'high' and 'low' levels of 'accuracy' and 'stability'. Accuracy indicates the reliability of the final results after the averaging of many experimental results and correction of measurement shifts induced by various causes. Accuracy can also be high if the difference between each measurement result is large. To achieve high accuracy, the systematic uncertainty should be small because the statistical uncertainty can be reduced by repeating a measurement many times. The accuracy is estimated from the possible shift of measurements by theoretical calculation or experimental measurement with different circumstances. Comparison of measurements obtained using different devices is also performed.

Stability refers to the constancy of each measurement result. The stability can be high if the measurement results are equally shifted from the real value. During the repetition of a measurement over a long period, there may be temporal change of the measurement values induced by changes in the circumstance. Therefore, we distinguish 'short-term stability' and 'long-term stability' from the averaging time. Short-term stability is determined from the statistical uncertainty covering short measurement times, with which the change of the circumstance is negligibly small. To achieve high short-term stability with the atomic transition frequency, a narrow spectrum linewidth (δf in equation (2.9)) and a large number of atoms (N_a in equation (2.9)) are required. With long measurement times, the statistical uncertainty is suppressed, but the influence of the change in conditions (e.g. electric field and magnetic field) becomes significant. Therefore, long-term stability is determined by the systematic uncertainty. The standard deviation is generally used for the

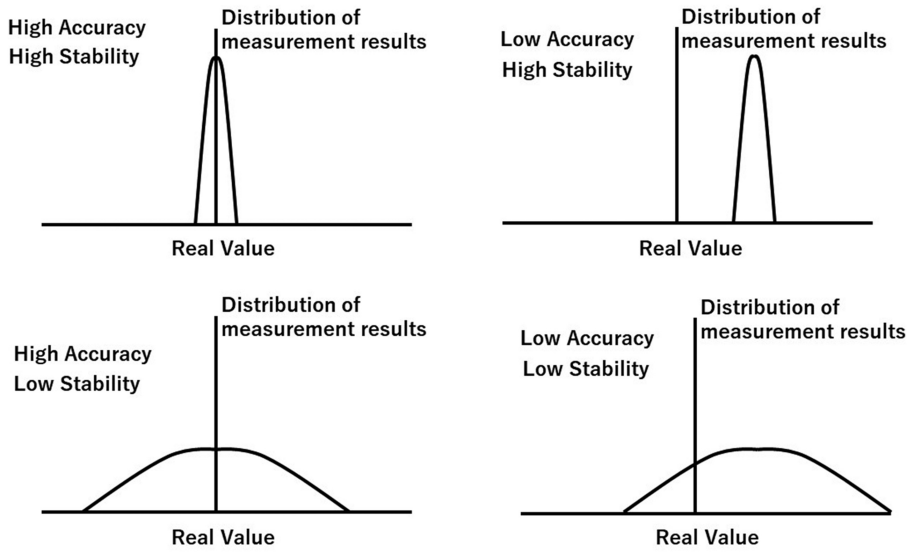

Figure 2.4. Schematic of the distribution of measurements with high and low 'accuracy' and 'stability'.

estimation of statistical uncertainty, but it is not useful to include measurements with a temporal change of circumstance. With the linear drift of measurement value, the standard deviation is divergent.

Particularly for the estimation of frequency stability, Allan variance is often used for any averaging time [7]. The Allan variance is given by

$$\sigma_y(M, T_a, \tau) = \sqrt{\frac{1}{M-1}\left\{\sum_{i=0}^{M-1}\left[\frac{x(iT_a+\tau)-x(iT_a)}{\tau}\right]^2 - \frac{1}{M}\left[\sum_{i=0}^{M-1}\frac{x(iT_a+\tau)-x(iT_a)}{\tau}\right]^2\right\}}, \tag{2.10}$$

where $x(t)$ denotes the phase at the time t. Here, $|x(iT_a + \tau) - x(iT_a)|/\tau$ shows the average frequency between iT_a and $iT_a + \tau$. The Allan variance is the variation of

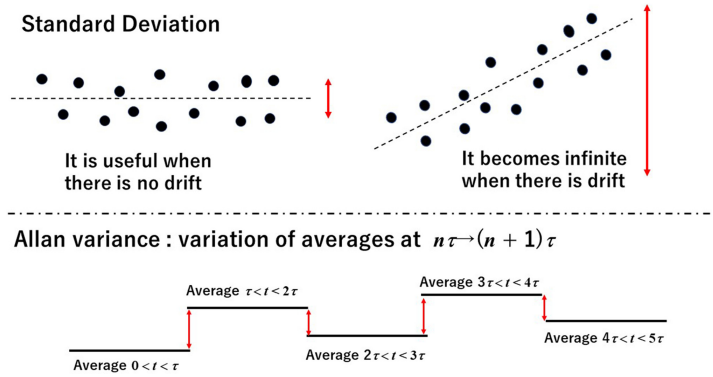

Figure 2.5. Concepts of standard deviation and Allan variance with $M = 2$ and $T_a = \tau$ (equation (2.7)). Standard deviation is not useful when there is a drift. Allan variance is the variation of the average at various measurement time periods, which is also useful when there is a drift.

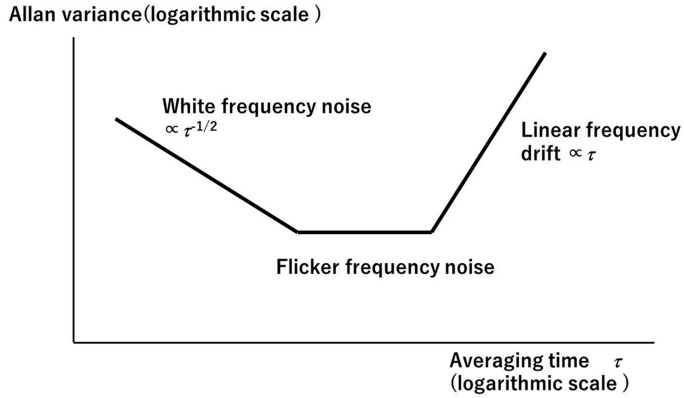

Figure 2.6. Dependence of the Allan variance on the averaging time τ considering the white frequency noise, flicker frequency noise, and linear frequency drift.

the averaged frequency for the period of τ at various time periods. Figure 2.5 shows the case with $M = 2$ and $T_a = \tau$. Figure 2.6 shows the dependence of the Allan variance on τ.

The short-term stability is generally given by the white frequency noise, and the Allan variance is proportional to $1/\sqrt{\tau}$, as shown with the standard deviation. With longer τ, the flicker frequency noise (called $1/f$ noise) is more significant than the white frequency noise, and the Allan variance is constant with τ. When there is a linear frequency drift, the Allan variance is proportional to τ with large τ.

References

[1] https://en.wikipedia.org/wiki/Central_limit_theorem
[2] Heavner T P *et al* 2005 *Metrologia* **42** 411
[3] Clairon A, Salomon C, Guellati S and Phillips W 1991 *Europhys. Lett.* **16** 165
[4] Chou C W *et al* 2010 *Phys. Rev. Lett.* **104** 070802
[5] Ushijima I *et al* 2015 *Nat. Photon.* **9** 185
[6] Nicholson T L *et al* 2015 *Nat. Commun.* **6** 6896
[7] https://en.wikipedia.org/wiki/Allan_variance