Allan deviation

Haonan Liu

May 24, 2023

1 Definition

1.1 Formal definition

The *Allan deviation* is a measure of frequency stability in atomic clocks. Due to the existence of flicker noise in atomic clocks, traditional statistical tools cannot be used since the standard deviation estimator of frequency is divergent. The Allan deviation is one way to characterize the frequency deviation. Other measures include the modified Allan deviation, the time Allan deviation, and the Hadamard deviation, etc.

Definition 1 [M-sample variance estimator]**:** *Consider a continuous-time ergodic stochastic variable $x(t)$. Starting from $t = 0$, we sample M times with a sampling interval T_c. Each sampling takes time $\tau \leq T_c$. The M-sample variance estimator is* [1]

$$\sigma_y^2(M, T_c, \tau) = \frac{M}{M-1}\left\{\frac{1}{M}\sum_{j=0}^{M-1}\left[\frac{x(jT_c+\tau)-x(jT_c)}{\tau}\right]^2 - \left[\frac{1}{M}\sum_{j=0}^{M-1}\frac{x(jT_c+\tau)-x(jT_c)}{\tau}\right]^2\right\}. \tag{1}$$

We call the time $T_c - \tau$ as the dead time between measurements. Thus there is no dead time if $T_c = \tau$.

Definition 2 [Allan variance]**:** *The Allan variance* [2] *is the ensemble average of the 2-sample variance estimator without dead time, i.e.,*

$$\sigma_y^2(\tau) \equiv \left\langle \sigma_y^2(2, \tau, \tau) \right\rangle = \frac{1}{2\tau^2}\left\langle [x_{n+1}(\tau) - 2x_n(\tau) + x_{n-1}(\tau)]^2 \right\rangle_n, \tag{2}$$

where $x_{n+k}(\tau) = x(t_n + k\tau)$, $k \in \{-1, 0, 1\}$, and t_n stands for some chosen time of the ensemble.

Here, the ensemble average over n is defined as [3]

$$\langle f_n \rangle_n = \lim_{N\to\infty}\sum_{n=1}^{N}\frac{f_n}{N}. \tag{3}$$

The *Allan deviation* is the square root of the Allan variance.

1.2 Clocks

In the context of optical atomic clocks, we can give the definition above more physical meaning. Consider a test clock A and a reference clock R that are supposed to be synchronized at $t = 0$ [4]. In reality R is usually a chosen atom with a chosen optical transition of frequency ω_0, and A is a local oscillator with frequency

[1] This is just a Bessel correction $M/(M-1)$ times a variance written in the form of $E[X^2] - E[X]^2$.

[2] Notice that we have been sloppy about the notations of estimators and the mean of the estimators. Allan variance is defined as the latter.

[3] If taken from the same time series, the ensemble average can be taken in an overlapped manner or a nonoverlapped manner.

[4] How to do this in reality?

around ω_0. The model of A is given by $e^{i\phi_A(t)}$ where $\phi_A(t) = \omega_0 t + \varphi(t)$ with $\varphi(t)$ some unwanted noise. The model of R is the perfect oscillator $e^{i\phi_R(t)}$ where $\phi_R(t) = \omega_0 t$. The time-error $x(t)$, which is also proportional to the phase error, is then given by

$$x(t) = t_A(t) - t_R(t) = \frac{\phi_A(t) - \phi_R(t)}{\omega_0} = \frac{\varphi(t)}{\omega_0}. \tag{4}$$

Notice that in this model $t_R(t) = t$. Using Eq. (4) in Eq. (2) we can calculate the Allan variance of the clock caused by the noise $\varphi(t)$.

We can also interpret Allan variance from the perspective of frequencies. Define the time-dependent frequency $\omega(t)$ of the local oscillator A as

$$\omega(t) = \frac{d\phi_A(t)}{dt} = \omega_0 + \frac{d\varphi(t)}{dt}. \tag{5}$$

Then we can define the *fractional frequency difference* as

$$y(t) = \frac{\omega(t) - \omega_0}{\omega_0} = \frac{\omega(t)}{\omega_0} - 1 = \frac{1}{\omega_0}\frac{d\varphi(t)}{dt} = \frac{dx}{dt}, \tag{6}$$

i.e., $y(t)$ is none other than the time derivative of $x(t)$.

As a summary, for the atomic clock model, $x(t)$ stands for the time reading signal in time and $y(t)$ stands for the frequency signal in time, which is the time derivative of $x(t)$. We can calculate the Allan variance using either x or y. Below we mostly choose y since it is dimensionless.

1.3 Allan variance

It is natural to define the Allan variance with $y(t)$, i.e., [5]

$$\begin{aligned}
\sigma_y^2(\tau) &= \frac{1}{2\tau^2}\left\langle \left[\int_{t_n}^{t_n+\tau} dt'\, y(t') - \int_{t_n-\tau}^{t_n} dt'\, y(t')\right]^2 \right\rangle_n, \\
&= \lim_{N\to\infty} \frac{1}{2\tau^2 N}\sum_{n=1}^{N}\left[\int_{t_n}^{t_n+\tau} dt'\, y(t') - \int_{t_n-\tau}^{t_n} dt'\, y(t')\right]^2.
\end{aligned} \tag{9}$$

Let $t_1 = -\frac{T}{2}$, $t_n = \frac{T}{2}$. We have [6]

$$\begin{aligned}
\sigma_y^2(\tau) &= \lim_{N\to\infty} \frac{1}{2\tau^2}\sum_{n=1}^{N}\frac{1}{T}\frac{T}{N}\left[\int_{t_n}^{t_n+\tau} dt'\, y(t') - \int_{t_n-\tau}^{t_n} dt'\, y(t')\right]^2 \\
&= \frac{1}{2\tau^2}\int_{-\frac{T}{2}}^{\frac{T}{2}} \frac{1}{T}\, dt \left[\int_{t}^{t+\tau} dt'\, y(t') - \int_{t-\tau}^{t} dt'\, y(t')\right]^2 \\
&= \frac{1}{2\tau^2 T}\int_{-\infty}^{\infty} dt \left[\int_{t}^{t+\tau} dt'\, y(t') - \int_{t-\tau}^{t} dt'\, y(t')\right]^2
\end{aligned} \tag{10}$$

where in the last line we have extended the upper and lower limits of the integral to infinity by assuming $y(t) = 0$ outside $[-\frac{T}{2}, \frac{T}{2}]$. Notice that we have not taken the $T \to \infty$ limit at this step. All that we have done is *zero padding.*

[5] In literature people tend to use the notion of the *average fractional frequency difference* $\bar{y}$. We define

$$\bar{y}(t,\tau) = \frac{1}{\tau}\int_0^\tau dt'\, y(t+t'). \tag{7}$$

We have the relation

$$x(t+\tau) - x(t) = \bar{y}(t,\tau)\tau. \tag{8}$$

[6] Since $y(t)$ is a stochastic variable, the integrals inside Eq. (9) are essentially Itô integrals, which means the summand itself is a stochastic variable. Therefore one should keep in mind that rigorously speaking we get a Riemann-Stieltjes integral instead of a Riemann integral by taking the $N \to \infty$ limit.

2 Stochastic differential equations

2.1 Definition

Below we calculate the Allan deviation $\sigma_y(\tau)$ given a model of $y(t)$.

We expect $x(t)$ to be a continuous time signal with a stochastic component. This means that it is reasonable to model the time evolution of $x(t)$ with a Langevin equation, i.e., for $t \in \mathbb{R}$,

$$\frac{dx}{dt} = a(x,t) + b(x,t)\xi(t), \tag{11}$$

where we call $a(t)$ the drift term and $b(x,t)\xi(t)$ the diffusion term with $b(x,t)$ the diffusion coefficient. Here $\xi(t)$ is an ergodic stochastic random variable that satisfies

$$\langle \xi(t) \rangle = 0, \tag{12}$$

$$\langle \xi(t)\xi(t') \rangle = \delta(t - t'), \tag{13}$$

where the time average $\langle \cdot \rangle$ is defined as

$$\langle f(t) \rangle = \lim_{T\to\infty} \frac{1}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} dt\, f(t). \tag{14}$$

In this sense, $dW = \xi(t)\, dt$ is a Wiener process.

For now we make some assumptions to simplify our model. We assume a and b are both real and independent of x so that $a(x,t) = a(t)$ and $b(x,t) = b(t)$. This results in the form

$$\frac{dx}{dt} = y(t) = a(t) + b(t)\xi(t). \tag{15}$$

2.1.1 Drifts

The drift term $a(t)$ can be a stochastic variable or a deterministic one. For now we suppose $a(t)$ is deterministic [7]. The simplest nontrivial case then is the linear drift

$$a(t) = a_1 t. \tag{16}$$

2.1.2 Diffusion

Physically speaking, errors in atomic clocks can be divided into two types: *statistical errors* from measurement fluctuations and other *systematic errors*. Allan deviation caused by statistical errors, which are mostly white noise, tends to decrease with increasing τ [see Eq. (42)], but Allan deviation caused by systematic errors, which are more likely a mixture of other *power-law noise* or *colored noise* types, will likely level off [Eq. (45)] or increase [Eq. (48)]. These effects can be taken into account if we consider the second order moment of the noise term $b(t)\xi(t)$, given the fact that we assume $a(t)$ is deterministic.

For convenience we define a new stochastic variable

$$B(t) = b(t)\xi(t) \tag{17}$$

with

$$\langle B(t) \rangle = 0, \tag{18}$$

$$\langle B(t')B(t+t') \rangle = A_B(t), \tag{19}$$

[7] We can also try stochastic models of $a(t)$

where $A_B(t)$ is the autocorrelation function of $B(t)$. The autocorrelation function is only a function of t due to the stationary condition of the stochastic variable $B(t)$.

It is mostly convenient to study the autocorrelation function in the frequency domain. By Wiener-Khinchin theorem, we have

$$S_B(f) = \mathcal{F}^{-1}[A_B(t)], \tag{20}$$

where $S_B(f)$ is the power spectral density of $B(t)$ that is formally defined as

$$S_B(f) = \lim_{T\to\infty} \frac{1}{T}\left|\tilde{B}(f)\right|^2. \tag{21}$$

Here we have used the convention

$$\mathcal{F}[B(t)] = \tilde{B}(f) = \int_{-\infty}^{\infty} dt\, B(t)e^{-i2\pi ft}, \tag{22a}$$

$$\mathcal{F}^{-1}\left[\tilde{B}(f)\right] = B(t) = \int_{-\infty}^{\infty} df\, \tilde{B}(f)e^{i2\pi ft}, \tag{22b}$$

We will see that the Allan variance can be calculated given a model of the power spectral density $S_B(f)$.

2.2 Analytical solution

2.2.1 No drift

We first discuss the case without drifts. From Eq. (15), assuming $a(t) = 0$, we have

$$y(t) = B(t). \tag{23}$$

Thus our $y(t)$ only includes the stochastic part $B(t)$ now. Substituting Eq. (23) into Eq. (10), we have

$$\sigma_B^2(\tau) = \frac{1}{2\tau^2 T}\int_{-\infty}^{\infty} dt \left[\int_t^{t+\tau} dt'\, B(t') - \int_{t-\tau}^{t} dt'\, B(t')\right]^2. \tag{24}$$

Writing the integrand as a convolution (or an *impulse response form*), we have

$$\sigma_B^2(\tau) = \frac{1}{2\tau^2 T}\int_{-\infty}^{\infty} dt \left[\int_{-\infty}^{\infty} dt'\, h(t-t')B(t')\right]^2, \tag{25}$$

where

$$h(t') = \begin{cases} 1, & -\tau < t' < 0 \\ -1, & 0 < t' < \tau \\ 0, & \text{elsewhere} \end{cases}. \tag{26}$$

Then we have from Eq. (25)

$$\sigma_B^2(\tau) = \frac{1}{2\tau^2 T}\int_{-\infty}^{\infty} dt\, [h(t) * B(t)]^2, \tag{27}$$

$$= \frac{1}{2\tau^2 T}\int_{-\infty}^{\infty} dt\, |h(t) * B(t)|^2, \tag{28}$$

$$= \frac{1}{2\tau^2 T}\int_{-\infty}^{\infty} df\, |\mathcal{F}[h * B]|^2, \tag{29}$$

$$= \frac{1}{2\tau^2}\int_{-\infty}^{\infty} df\, \left|\tilde{h}(f)\right|^2 \frac{\left|\tilde{B}(f)\right|^2}{T}, \tag{30}$$

where we have used the fact that $h(t)$ and $y(t)$ are real, the Parseval-Plancherel theorem, and the convolution theorem, respectively.

At this point, we take the $T \to \infty$, $\frac{T}{N} \to 0$ limit and use Eq. (21), yielding

$$\begin{aligned}\sigma_B^2(\tau) &= \lim_{T\to\infty} \frac{1}{2\tau^2} \int_{-\infty}^{\infty} df\, \left|\tilde{h}(f)\right|^2 \frac{\left|\tilde{B}(f)\right|^2}{T}, \\ &= \int_{-\infty}^{\infty} df\, \frac{2\sin^4(\pi f \tau)}{(\pi f \tau)^2} S_B(f). \end{aligned} \tag{31}$$

We have also explicitly calculated $\tilde{h}(f)$, which is usually called the *transfer function*. Since $B(t)$ is real, $\tilde{B}(f)$ is Hermitian and $S_B(f)$ is even. This makes the integrand of Eq. (31) even. Thus we have

$$\sigma_B^2(\tau) = \int_0^{\infty} df\, \frac{4\sin^4(\pi f \tau)}{(\pi f \tau)^2} S_B(f). \tag{32}$$

We now have a relation of the Allan deviation as a function of τ and the power spectral density of $B(t)$.

We briefly discuss the purpose of taking the $T \to$ limit. We have assumed that the support of $B(t)$ is $[-\frac{T}{2}, \frac{T}{2}]$ by using zero padding. Thus by take the $T \to \infty$ limit, we neatly recover the power spectral density. Hypothetically, we could also take the $T \to$ limit in Eq. (10), which is a matter of taste to me. Since for simulations we will need to choose a finite T and N, the theoretical result given by Eq. (32) should be the limit of the simulation results in the $N \to \infty, T \to \infty, \frac{T}{N} \to 0$ limit.

2.2.2 With drift

We now go back to Eq. (15) and calculate the Allan deviation given a linear drift term, i.e., given

$$y(t) = a_1 t + B(t). \tag{33}$$

Equation (10) then becomes

$$\begin{aligned}\sigma_y^2(\tau) &= \frac{1}{2\tau^2 T} \int_{-\infty}^{\infty} dt \left[\int_t^{t+\tau} dt'\, y(t') - \int_{t-\tau}^{t} dt'\, y(t')\right]^2 \\ &= \frac{1}{2\tau^2 T} \int_{-\frac{T}{2}}^{\frac{T}{2}} dt \left[\int_t^{t+\tau} dt_1\, [a_1 t_1 + B(t_1)] - \int_{t-\tau}^{t} dt_2\, [a_1 t_2 + B(t_2)]\right]^2 \\ &= \frac{1}{2\tau^2 T} \int_{-\frac{T}{2}}^{\frac{T}{2}} dt \left[a_1 \tau^2 + \int_{-\infty}^{\infty} dt'\, h(t-t') B(t')\right]^2 \\ &= \frac{1}{2\tau^2 T} \int_{-\frac{T}{2}}^{\frac{T}{2}} dt\, a_1^2 \tau^4 + \frac{1}{2\tau^2} \int_{-\infty}^{\infty} dt \left[\int_{-\infty}^{\infty} dt'\, h(t-t') B(t')\right]^2 \\ &= \frac{a_1^2 \tau^2}{2} + \int_0^{\infty} df\, \frac{4\sin^4(\pi f \tau)}{(\pi f \tau)^2} S_B(f) \\ &= \frac{a_1^2 \tau^2}{2} + \sigma_B^2(\tau). \end{aligned} \tag{34}$$

Here we have used the fact that our integral for Allan deviation was only carried out during the time interval $[-\frac{T}{2}, \frac{T}{2}]$. This result does not change when we take the $T \to \infty$ limit. Therefore we see that a linear drift term $a_1 t$ will add a τ^2 behavior to the Allan deviation. Generally, for a drift term of the form

$$y_D(t) = a_n t^n, \tag{35}$$

the modification to the Allan deviation is

$$\begin{aligned}\sigma_D^2(\tau) &= \frac{1}{2\tau^2}\int_{-\frac{T}{2}}^{\frac{T}{2}} dt\left(\int_t^{t+\tau} dt_1\, a_n t_1^n - \int_{t-\tau}^{t} dt_2\, a_n t_2^n\right)^2 \\ &= \frac{1}{2\tau^2}\frac{a_n^2}{(n+1)^2}\int_{-\frac{T}{2}}^{\frac{T}{2}} dt\left[(t+\tau)^{n+1}+(t-\tau)^{n+1}-2\tau^{n+1}\right]^2. \end{aligned} \tag{36}$$

1. For $n = 1$, we have

$$\sigma_D^2(\tau) = \frac{1}{2}a_1^2\tau^2 \tag{37}$$

   which does not depend on T.

2. For $n > 1$, the modification terms will diverge as we increase T. Allan variance does not seem to be a good estimator in this case.

Thus nonzero monomiol drifts can cause a power law modification to the Allan variance.

2.3 Colored noise

In this section we consider the situation where the stochastic process $B(t)$ can be described by various types of colored noise and calculate the Allan variance $\sigma_B^2(\tau)$. In literature, it is convenient to use the monomial power spectral densities, i.e.,

$$S_B(f) = c_n|f|^n, \tag{38}$$

where we have used the absolute value sign since $S_B(f)$ is even. Specifically, for $n = -2, -1, 0, 1, 2$, the noise types are named RW FM, F FM, W FM, F PM, and W PM [8], although I am not a fan of using these names.

For each n value, we can calculate the autocorrelation functions of these noise types. **Mathematica** tells us that, given the power spectral density in Eq. (38), the corresponding autocorrelation function $A_B(t) = \mathcal{F}^{-1}[S_B(f)]$ is given by

$$A_B(t) = -\frac{c_n\Gamma(n+1)\sin\left(\frac{n\pi}{2}\right)}{2^n\pi^{n+1}}\frac{1}{|t|^{n+1}} = \begin{cases} -\dfrac{c_n}{\pi^{n+\frac{1}{2}}}\dfrac{\Gamma(\frac{n+1}{2})}{\Gamma(-\frac{n}{2})}\dfrac{1}{|t|^{n+1}}, & n\,\text{even} \\ -\dfrac{c_n\Gamma(n+1)}{2^n(i\pi)^{n+1}}\dfrac{1}{t^{n+1}}, & n\,\text{odd}\end{cases}. \tag{39}$$

Notice that use of Eq. (39) requires care when n takes values of nonpositive integers. With Eq. (38), it is also straightforward to calculate the corresponding behavior of the Allan variance $\sigma_B^2(\tau)$ using Eq. (32).

Below we explicitly calculate $A_B(t)$ and $\sigma_B^2(\tau)$ given different values of n.

1. $n = 0$.

   We have

$$S_B(f) = c_0, \tag{40}$$
$$A_B(t) = c_0\delta(t), \tag{41}$$
$$\sigma_B^2(\tau) = c_0\frac{1}{\tau}. \tag{42}$$

   Typical white noise. The Allan variance is proportional to τ^{-1}. For $n \neq 0$ the time correlation tells us that we have nonmarkovian noise.

[8] Abbreviations stand for Random Walk (RM), Frequency Modulation (FM), Flicker (F), White (FM), Phase Modulation (PM).

2. $n < 0$.

   (a) $n = -1$.
   We have

$$S_B(f) = c_{-1}|f|^{-1}, \tag{43}$$
$$A_B(t) = -2c_{-1}(\ln 2\pi|t| + \gamma), \tag{44}$$
$$\sigma_B^2(\tau) = c_{-1}\ln 16. \tag{45}$$

   Here γ is the Euler's constant or "Euler gamma". The Allan variance is constant in τ.

   (b) $n = -2$.
   We have

$$S_B(f) = c_{-2}|f|^{-2}, \tag{46}$$
$$A_B(t) = -2c_{-2}\pi^2\frac{t^2}{|t|}, \tag{47}$$
$$\sigma_B^2(\tau) = \frac{4\pi^2 c_{-2}}{3}\tau. \tag{48}$$

   The Allan variance is proportional to τ.

   (c) $n = -3$. We have

$$S_B(f) = c_{-3}|f|^{-3}, \tag{49}$$
$$A_B(t) = 2c_{-3}\pi^2 t^2\big(-3 + 2\gamma + \ln 4\pi^2 t^2\big). \tag{50}$$

   The Allan variance as defined in Eq. (2) diverges. It can still be calculated given a lower frequency cutoff though.

3. $n > 0$.

   (a) $n = 1$.
   We have

$$S_B(f) = c_1|f|, \tag{51}$$
$$A_B(t) = -\frac{c_1}{2\pi^2}\frac{1}{t^2}, \tag{52}$$
$$\begin{aligned}\sigma_B^2(\tau) &= \lim_{f_h\to\infty}\frac{c_1}{2\pi^2}\frac{3\gamma + \ln 4 - 4\mathrm{Ci}[2\pi f_h\tau] + \mathrm{Ci}[4\pi f_h\tau] + 3\ln\pi f_h\tau}{\tau^2}\\ &= \lim_{f_h\to\infty}\frac{c_1}{2\pi^2}\frac{3\gamma + \ln 4 + 3\ln\pi f_h\tau}{\tau^2},\end{aligned} \tag{53}$$

   where Ci is the cosintegral function and f_h is the higher frequency cut off that corresponds to the minimum time resolution of the measurements.

   (b) $n = 2$.
   We have

$$S_B(f) = c_2|f|^2, \tag{54}$$
$$A_B(t) = -\frac{c_2}{4\pi^2}\delta^{(2)}(t), \tag{55}$$
$$\begin{aligned}\sigma_B^2(\tau) &= \lim_{f_h\to\infty}\frac{c_2}{8\pi^3}\frac{12\pi f_h\tau - 8\sin 2\pi f_h\tau + \sin 4\pi f_h\tau}{\tau^3}\\ &= \lim_{f_h\to\infty}\frac{3c_2}{2\pi^2}\frac{f_h}{\tau^2}.\end{aligned} \tag{56}$$

(c) $n > 2$.

So on.

Therefore, the Allan deviation defined in Eq. (2) is only a good measure of the $n = 0, -1, -2$ cases, which correspond to the RW FM, F FM, and W FM noise types. The τ dependence is τ^{-1}, τ^{0}, and τ^{1}, respectively. For $n \leq -3$ or $n \geq 1$, either a lower or higher frequency cutoff needs to be considered. In reality these components can be seen due to practical constraints on the frequency cutoffs. We can simulate these effects by manually introducing the cutoffs.