The Averaging of Phase and Frequency Data

W.J. Riley
Hamilton Technical Services
Beaufort, SC 29907 USA
bill@wriley.com

• Introduction

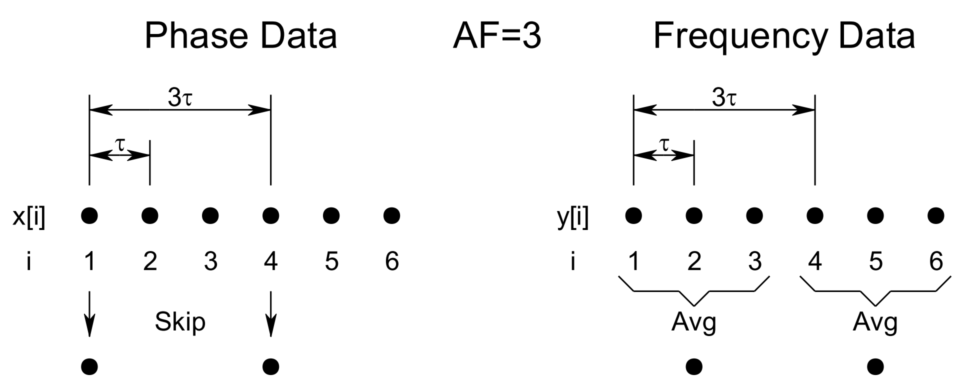
This brief note discusses the averaging of phase and frequency data in the analysis frequency stability with the objective of clarifying some of the terminology involved.

Frequency stability analysis generally applies to equally-spaced phase or frequency measurements taken at a particular *measurement interval* denoted by the lower-case Greek letter *tau* (τ). Other words used for this quantity are *sampling interval*, *measurement time*, *sampling time* or *averaging time*. The measurement and sampling terms are usually associated with the measurement process itself, while the averaging time applies to the analysis. The basic measurement interval is often denoted as τ_0 while the analysis averaging time is simply called τ.

Phase data in this context have units of seconds and are denoted by x, while frequency data are dimensionless fractional frequency denoted by y.

• Data Averaging

Data taken at a certain measurement interval τ_0 can be *averaged* to become data at an integer multiple n of the measurement tau, $\tau = n \cdot \tau_0$, and the use of the term *data averaging* can sometimes lead to confusion. Frequency data are averaged to a longer tau by ordinary *algebraic averaging*, while phase data undergo the same transformation by *decimation* (actually *downsampling*, see below). In other words, to average frequency data, one adds n adjacent frequency points and divides that sum by n, while, to average phase data, one simple uses every n^{th} point by skipping $n-1$ intermediate points, where n is called the *averaging factor*, *AF*. Thus we average frequency data by averaging and we average phase data by decimation. In both cases, we call the process averaging, but it is performed by decimation for phase data.



Adding further to the possible confusion, for a set of N phase data points there are one fewer $M=N-1$ corresponding frequency data points. This is because the frequency data are the first differences of the phase data divided by tau, and it obviously takes two phase points to form a difference. Conversely, phase data are obtained from frequency data via numerical integration by adding the product of the frequency and tau to the previous phase value.

• **Decimation vs. Downsampling**

Strictly speaking, the process used in phase averaging is downsampling because decimation usually implies low pass filtration before resampling at the lower rate in order to avoid aliasing by removing spectral components higher than the Nyquist frequency of one-half of the new sampling rate. Because simple downsampling correctly describes the phase evolution, low pass filtration alters the data and introduces additional parameters, and because the emphasis is on analyzing noise (not discrete components), downsampling rather than full decimation is used to convert phase data to a longer averaging time.

When performing spectral analysis, one should therefore be aware of the possibility of aliasing of discrete spectral components in "averaged" phase data. For example, consider a set of white phase noise data having a strong component at its original Nyquist frequency (0.5 Hz, one-half of the 1-second measurement rate), as shown in Figure 1. If these data are averaged by a factor of three by simple downsampling, the discrete component is aliased to one-third of its original frequency (≈ 0.17 Hz) in the averaged data, as shown in Figure 2. If the original data are low pass filtered at 0.5/3 Hz before being downsampled (and thus correctly decimated) as shown in Figure 3, the averaged data do not have the aliased component, but are changed, as shown in Figure 4.

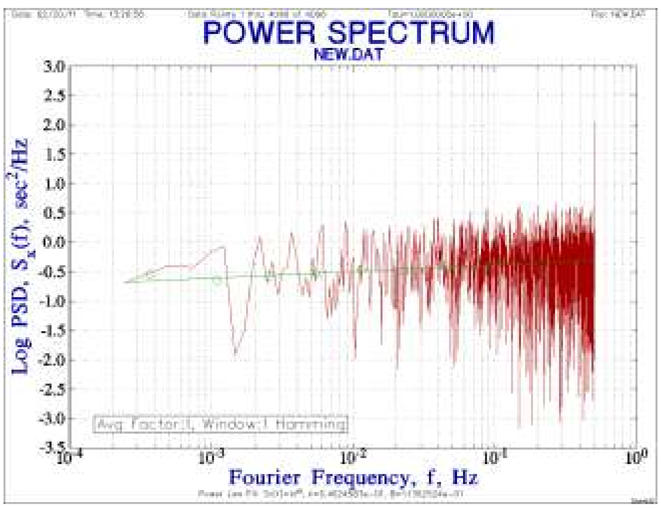


Figure 1. PSD of Original Data

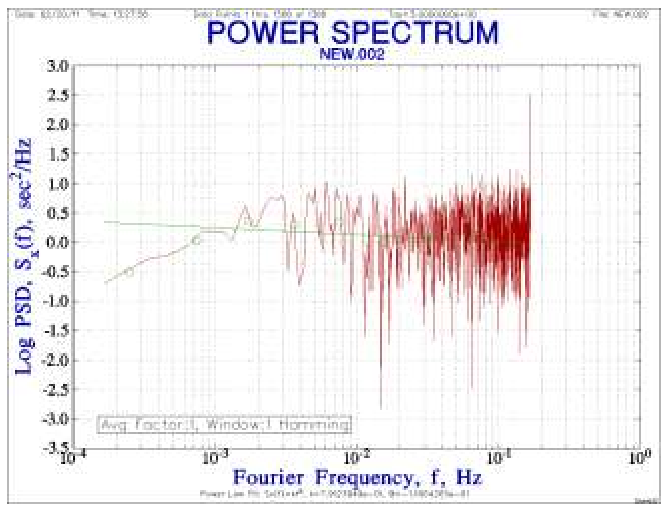


Figure 2. PSD of Averaged Data

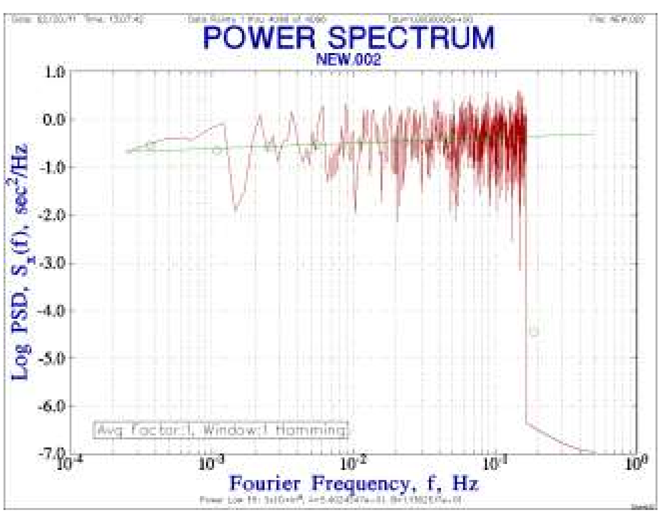


Figure 3. PSD of Filtered Original Data

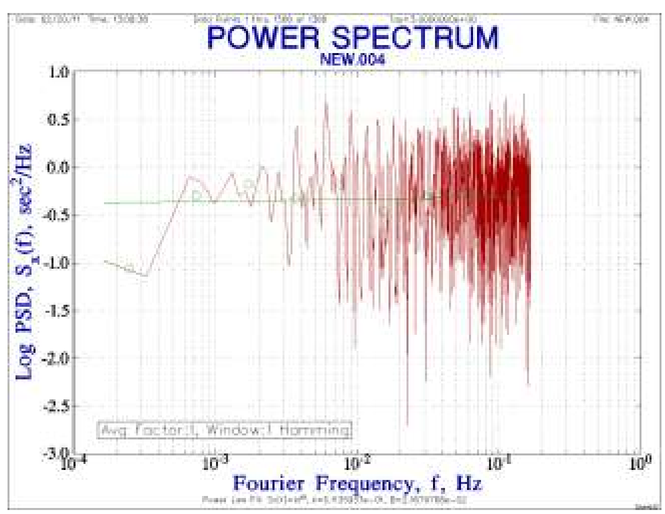


Figure 4. PSD of Averaged Filtered Data

• **Allan Deviation**

The Allan deviation is correctly processed even if data containing a strong component at its original Nyquist frequency is averaged by a factor of three, as shown in Figures 5 and 6.

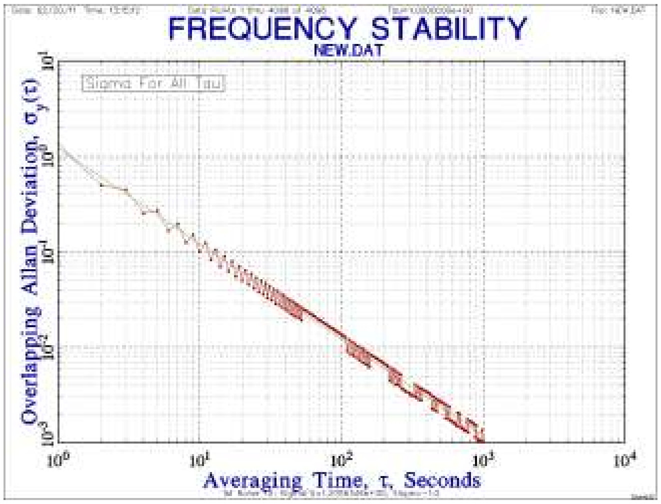


Figure 5. ADEV of Original Data

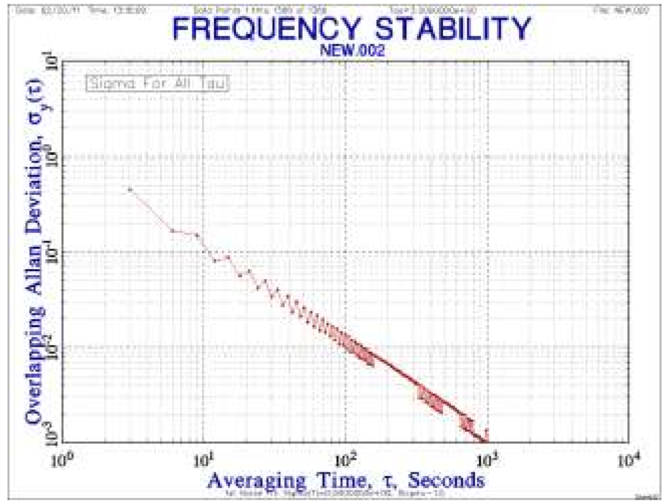


Figure 6. ADEV of Averaged Data

According to Dr. David Howe of NIST:

The frequency response of AVAR looks like a ½-octave-wide band pass filter, as shown in Figure 7. The peak in the response is at $f \cdot \tau = 0.5$, or at $\tau = 0.5/f = 0.5T$, where f is a Fourier component of fractional frequency deviation y and T is simply the period of that same component. Now suppose we only have one Fourier signal at $f = 1/T$ and no noise. At a period twice the sampling interval, we have $T = 2 \cdot \tau_0$, or $\tau_0 = 0.5T$, hence that signal is precisely at the peak or center of AVAR's filter. If we decimate by 3, this is the equivalent of averaging of y such that $\tau = 3 \cdot \tau_0$, or $T' = 2 \cdot \tau = 2 \cdot 3\tau_0) = 6 \cdot \tau_0$ at the center. Even though AVAR is maximally responsive to a signal at a long period of T' due to the "decimation" of the phase data, that response originates from period T, the undecimated phase data. The response to frequencies above the center peak *appears* as an aliasing effect, but it's more closely tied to a digital-filtering effect that is inherent to AVAR's sampling of y, shown in Figure 8. An analog ½-octave wide band pass filter built from an RC network would respond the same way to a frequency higher than the band pass center, but would be smooth, without the zeroes, of course. The picket-fence response makes undesirable zeroes appear in the response. Thêo1, and similarly ThêoH, are significantly more efficient for determining noise type at long-term tau than AVAR because of its smooth, more-ideal band pass response.

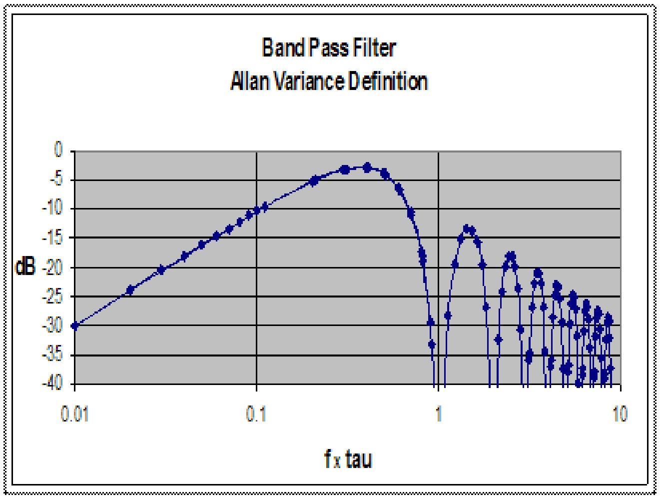


Figure 7 AVAR Bandpass Filter Characteristic

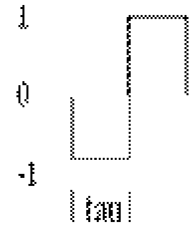


Figure 8. AVAR Sampling Function

• **Missing Points**

Missing phase or frequency points must be included in the data set to maintain the correct time spacing. Those gaps are often represented by a value of zero, and special means must be taken when averaging such data. For phase data averaging (decimation), intermediate gaps are simply ignored. For frequency data averaging, one forms the algebraic average of the non-gap points. In both cases, the averaging process will have the side effect of removing the gap(s) unless all of the associated points are missing, in which case the averaged point will also become a gap.

• **Phase Averaging in the Modified Allan Deviation**

Still more semantic confusion can arise regarding the phase averaging that is part of the calculation of the modified Allan deviation (MDEV), a stability measure most often applied to distinguish between white and flicker PM noise processes, or as the basis of the time deviation statistic. In that context, actual algebraic averaging is applied to the phase data as part of the MDEV calculation.

• **Averaged Phase Measurements**

Clock measuring systems occasionally perform algebraic averaging as phase data are collected as a way to reduce the measurement noise. While that does lower the noise by a factor on the order of the square root of the averaging factor, it also changes the statistical properties of the data in a fashion similar to the modified Allan deviation.

• **Frequency Averaging in the Allan Deviation**

If frequency values are sampled directly, as with a frequency discriminator or tight PLL and an analog-to-digital converter, they must be averaged before using them to calculate the Allan deviation. Oversampling can be used to support that averaging process. An oversampling factor of ten is generally sufficient.

File: The Averaging of Phase and Frequency Data.doc
W.J. Riley
Hamilton Technical Services
November 10, 2010
Revision C, November 3, 2011