Appendix B

Expected Value of the *NSE* When Applied to a Process Affected by Oscillations

This appendix develops the expected value of the NSE presented in chapter 9 when the process is affected by oscillations. Therefore, we start with an analysis of the evolution of the sample-mean coming from a sampled sine-wave. Then, by using previous result, the statistic S^2 is analyzed in the case of a statistical process undergoing sine-wave oscillations. Finally, the expected value of the NSE is calculated for the same case.

Before we start, it must be noted that the expected values of statistics S^2 and *ENE* are *time-dependent* waves oscillating around an average value. This appendix focuses mainly on this average value. Even though some information about the amplitude of the waves is incidentally provided, the shape of the wave and the phases of the wave components are considered to be meaningless and obviated.

B.1 The sample mean of a sampled sine-wave

Let us consider a sine-wave $A\sin(\omega t+\delta)$ being sampled every T seconds. May $\tilde{x}_j$ be a measurement corresponding to $t = T\dot{j}$. Its value can be thus expressed as:

$$\tilde{x}_j = A\sin(\omega T j + \delta)\,; \qquad\qquad j \epsilon \mathbb{N}. \tag{B.1}$$

May $\tilde{\boldsymbol{x}}_{\boldsymbol{j}} = \{\tilde{x}_{j-n+1}, \tilde{x}_{j-n+2}, \ldots, \tilde{x}_j\}$ be a sample of size N. The definition of the sample mean is now applied:

$$\bar{\tilde{x}}_j = \frac{1}{N}\sum_{k=0}^{N-1}\tilde{x}_{j-k} \tag{B.2}$$

The value of $\tilde{x}_j$ can is substituted in previous equation and then terms depending on j are separated from the rest:

$$\begin{aligned}
\bar{\tilde{x}}_j &= \frac{A}{N}\sum_{k=0}^{N-1}\cos\left(\omega T(j-k)+\delta\right) \\
&= \frac{A}{N}Re\left(\sum_{k=0}^{N-1} e^{(\omega T(j-k)+\delta)i}\right) \\
&= \frac{A}{N}Re\left(e^{(\omega Tj+\delta)i}\sum_{k=0}^{N-1} e^{(-\omega Tk)i}\right) \\
&= \frac{A}{N}Re\left(e^{(\omega Tj+\delta)i}G(\omega T, N)\right) \qquad \text{(B.3)}
\end{aligned}$$

Term $G(\omega T, N)$ is a phasor whose module depends on the sampling frequency ratio ωT and the sample size N. Figure B.1 shows its value for some different sample sizes. By expressing phasor G as $G = |\mathrm{G}|\angle\varphi$, equation B.3 can be written as:

$$\begin{aligned}
\bar{\tilde{x}}_j &= \frac{A}{N}Re\left(|\mathrm{G}|e^{\varphi i}e^{(\omega Tj+\delta)i}\right) \\
&= \frac{A}{N}|\mathrm{G}|\cos\left(\omega Tj+\varphi_2\right)
\end{aligned}$$

Where $\varphi_2 = \varphi + \delta$.

Last equation should be more properly written like:

$$\bar{\tilde{x}}_j = \frac{A}{N}|\mathrm{G}(\omega T, N)|\cos\left(\omega Tj+\varphi_2\right) \qquad \text{(B.4)}$$

That corresponds to an out-of-phase wave whose module is attenuated with respect to the input one according to $\frac{|\mathrm{G}(\omega T,N)|}{N}$. (figure B.2)

B.2 The sample variance of a stochastic process affected by oscillations

Let us now consider a stochastic process affected by oscillations in a concrete frequency. May the population be defined as the sum of a sine-wave and an independent stochastic process. Then we can write:

$$x_j = \tilde{x}_j + \hat{x}_j; \qquad j\epsilon\mathbb{N}. \qquad \text{(B.5)}$$

Where $\tilde{x}_j = A\sin\left(\omega Tj+\delta\right)$ is a time-dependent, non-stochastic variable and $\hat{x}_j = X(\mu, \sigma_{proc})$ is a time-independent, stochastic variable.

We are now interested on calculating the expected value of the sample variance for this population. The definition of the sample variance is:

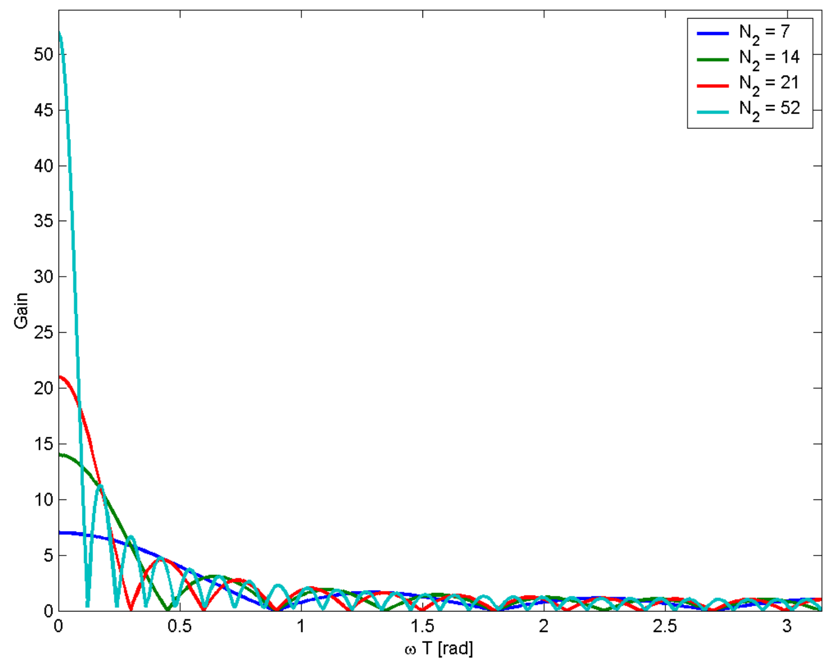

Figure B.1: Value of G

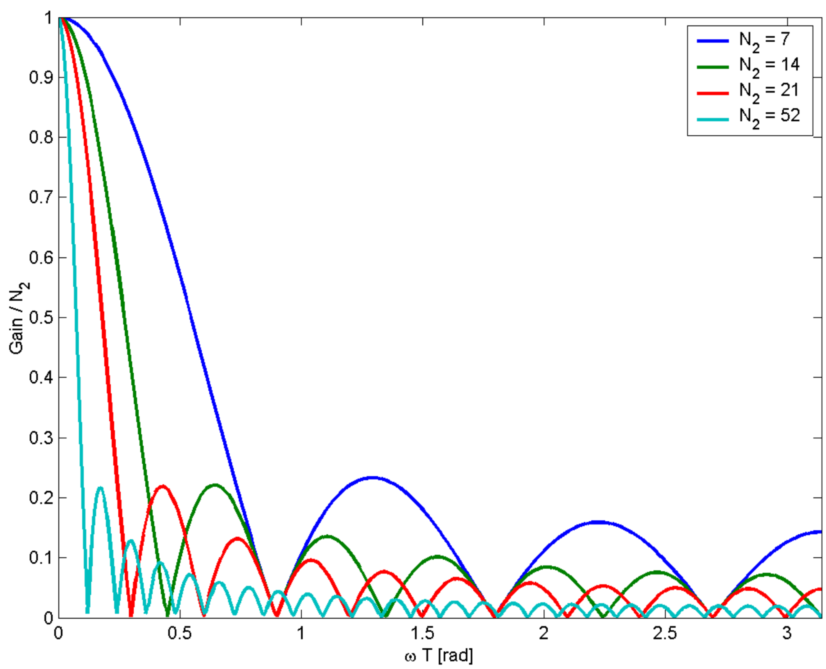

Figure B.2: Gain of the sample mean over a sine-wave input

$$S_j^2 \doteq \frac{1}{N_2-1}\sum_{k=1}^{N_2}(x_k-\bar{x}_j)^2 \tag{B.6}$$

Where N_2 is the sample size.

By substituting equation (B.5) in (B.6) and expanding the square, we get:

$$\begin{aligned} S_j^2 &= \frac{1}{N_2-1}\sum_{k=1}^{N_2}\left((\tilde{x}_k-\bar{\tilde{x}}_j)+(\hat{x}_k-\bar{\hat{x}})\right)^2 \\ &= \frac{1}{N_2-1}\sum_{k=1}^{N_2}\left((\tilde{x}_k-\bar{\tilde{x}}_j)^2+(\hat{x}_k-\bar{\hat{x}})^2+2(\tilde{x}_k-\bar{\tilde{x}}_j)(\hat{x}_k-\bar{\tilde{x}}_j)\right). \end{aligned}$$

Now the three terms in the sum are separately analyzed to finally calculate their expected values:

① $= \sum_{k=1}^{N_2}(\tilde{x}_k-\bar{\tilde{x}}_j)^2$, ② $= \sum_{k=1}^{N_2}(\hat{x}_k-\bar{\hat{x}})^2$, ③ $= \sum_{k=1}^{N_2}2\left((\tilde{x}_k-\bar{\tilde{x}}_j)(\hat{x}_k-\bar{\hat{x}})\right)$.

Term ①:

$$E\,(①) = (N_2-1)\sigma^2. \tag{B.7}$$

Term ②:

$$\begin{aligned} ② &= \sum_{k=1}^{N_2}\left(\tilde{x}_k^2+\bar{\tilde{x}}_j^2-2\tilde{x}_k\bar{\tilde{x}}_j\right) = \sum_{k=1}^{N_2}\tilde{x}_k^2+N_2\bar{\tilde{x}}_j^2-2\bar{\tilde{x}}\sum_{k=1}^{N_2}\tilde{x}_k \\ &= \sum_{k=1}^{N_2}\tilde{x}_k^2-N_2\bar{\tilde{x}}_j^2 = \sum_{k=1}^{N_2}A^2\cos^2(\omega Tk+\delta)+N_2\bar{\tilde{x}}_j^2 \end{aligned}$$

Now, using $\cos^2(\alpha) = \frac{1}{2}(1+\cos(2\alpha))$ we can express last equation in terms of $|\mathrm{G}|$ defined in previous section:

$$\begin{aligned} ② &= \frac{A^2}{2}\sum_{k=1}^{N_2}(1+\cos(2\omega Tk+2\delta))-N_2\bar{\tilde{x}}_j^2 \qquad \text{(B.8)} \\ &= \frac{A^2N_2}{2}+\frac{A^2}{2}|\mathrm{G}(2\omega T,N_2)|\cos(2\omega Tj+\varphi_3)+N_2|\mathrm{G}(\omega T,N_1)|\cos(\omega Tj+\varphi_2). \end{aligned}$$

Where N_1 is the size of the sample used to calculate the sample mean $\bar{x}$ (not necessarily the same as N_2). This result corresponds to its self expected value since there is no stochastic term. Besides it is a time-dependent value.

Term ③: Since $\tilde{x}_j$ and $\bar{\tilde{x}}$ are independent from $\hat{x}_j$ and $\bar{\hat{x}}$, the expected value can be written as

$$\begin{aligned} E\,(③) &= 2N_2E\,(\tilde{x}_k-\bar{\tilde{x}}_j)\,E\left(\hat{x}_k-\bar{\hat{x}}\right) \\ &= 2N_2E\,(\tilde{x}_k-\bar{\tilde{x}}_j)\,(\mu-\mu) = 0 \qquad \text{(B.9)} \end{aligned}$$

Finally we get:

$$\begin{aligned} E\left(S_j^2\right) &= \sigma^2 + \frac{1}{2}\frac{A^2 N_2}{(N_2-1)} + \frac{1}{2}\frac{A^2}{(N_2-1)}|\mathrm{G}(2\omega T, N_2)|\cos(2\omega Tj + \varphi_3) + \\ & \frac{N_2}{(N_2-1)}|\mathrm{G}(\omega T, N_1)|\cos(\omega Tj + \varphi_2). \end{aligned} \tag{B.10}$$

Equation (B.10) corresponds to a time-dependent wave around a constant value. This constant value can be considered the average expected value we are looking for:

$$\overline{E}\left(S^2\right) = \sigma^2 + \frac{1}{2}\frac{A^2 N_2}{(N_2-1)} \tag{B.11}$$

B.3 Expected value of the *NSE* under induced oscillations

The NSE presented in chapter 9 is defined as:

$$NSE = S^2 - S_{\bar{x}}^2 N_1 \tag{B.12}$$

Where $S_{\bar{x}}^2$ is the sample variance calculated over a data set containing sample-mean values and N_1 is the sample size *used to calculate the sample mean.* (Do not confuse with the sample size used to calculate the sample variance. This one is N_2).

Its expected value is:

$$E(NSE) = E\left(S^2\right) - E\left(S_{\bar{x}}^2\right) N_1. \tag{B.13}$$

The first term in last equation was calculated in (B.11), the second one can be easily calculated as follows:

From equation (B.5), the sample mean value of x_i can be written as:

$$\bar{x}_j = \bar{\bar{x}}_j + \tilde{\bar{x}}_j. \tag{B.14}$$

Where $\bar{\bar{x}}_j = X(\mu, \sigma_{mean})$ (with $\sigma_{mean} = \frac{\sigma_{proc}}{\sqrt{N_1}}$).

Term $\tilde{\bar{x}}_j$ was calculated in section B.1 and can be expressed as:

$$\tilde{\bar{x}}_j = B\cos(\omega Tj + \delta). \tag{B.15}$$

Where:

$$B = \frac{A}{N_1}|G(\omega T, N_1)|. \tag{B.16}$$

Now, if we compare (B.5) and (B.14), the average expected value of the sample variance over a sample coming from the population $\bar{\bar{x}}_j$ can be calculated according to (B.11):

$$
\begin{aligned}
E\left(S_{\bar{x}}^2\right) &= \sigma_{mean}^2 + \frac{1}{2}\frac{B^2 N_2}{(N_2-1)} \\
&= \frac{\sigma^2}{N_1} + \frac{1}{2}\frac{A^2|\mathrm{G}(\omega T, N_1)|^2}{N_1^2}\frac{N_2}{(N_2-1)}
\end{aligned}
\tag{B.17}
$$

Where B has been substituted according to (B.16).

Finally, substitution of (B.11) and (B.17) in (B.13) gives:

$$
E\left(NSE\right) = \frac{1}{2}\frac{A^2 N_2}{(N_2-1)}\left(1 - |\mathrm{G}(\omega T, N_1)|\right)
$$