Discussion about Noise Equivalent Power and its use for photon noise calculation.

Samuel Leclercq. 2007-03-02.

Abstract.

The Noise Equivalent Power (NEP) is a concept often used to quantify the sensitivity of a detector or the power generated by a source of noise on a detector. But the literature offers different definitions and different ways to calculate it. I recall here these definitions and the results of calculations from several authors, for the particular case of photon noise from background source illuminating a detector. In the second part of the document, starting from bases of mathematical description of random processes, I show the link between the different definitions of the NEP. In the third part, starting from the fundamental properties of boson I calculate the most general expression for the photon NEP, allowing to link the various expressions found in the literature, and understand the assumptions made for each case.

Introduction.

Definitions of NEP.

Contrary of what could be thought regarding the wide use of the Noise Equivalent Power in literature to characterize the measurements limits of detectors, it is not easy to find a clear mathematical definition of the NEP in the literature!

The first definition given here comes from the Federal Standard 1037C (telecom glossary 2000) of the United State Government. "*Noise-equivalent power (NEP) is the radiant power that produces a signal-to-noise ratio of unity at the output of a given optical detector at a given data-signaling rate or modulation frequency, operating wavelength, and effective noise bandwidth. Some manufacturers and authors define NEP as the minimum detectable power per square root bandwidth [W/Hz$^{1/2}$].*" [http://www.its.bldrdoc.gov/fs-1037/fs-1037c.htm]. Although is widely used on internet or articles with different phrasing ([Benford] [Léna] [Goodrich] [Atis] [Anghel et al] [Das et al] [Besse] and others), it is surprisingly confusing by its lack of mathematical precision and generality. Indeed, it is not clear whether it should be expressed in W or W/Hz$^{1/2}$, it does not give the relation between each component of the formula, and it is defined only for radiation detectors (for instance NEP produced by an electronic device can't be defined that way).

There is another definition coming from the theory of signal and random processes, which is much more general and mathematically precise, though rarely used. "*Given a property X which may fluctuate with a finite amplitude in a finite frequency band Δf and which correspond to a random ergodic process. Given a system able to measure this property with a defined power conversion factor called sensitivity. The Noise Equivalent Power of the system measuring the fluctuating property is the ratio of the ergodic process monolateral spectral density over the sensitivity of the system, it is expressed in units of W/Hz$^{1/2}$.*" [Leclercq, mais viens d'une autre source à retrouver !!!]. Although more rigorous than the previous definition, this one is tougher and uses terms that also need to be defined (spectral density, ergodic process, sensitivity). For that reason it is rarely used in the optical detector literature (or it is used in a more simple but less rigorous way [Dutoit] [Zweiacker]).

One can ask whether these two definitions are equivalent or at least compatible with each other ? I'll show that the first definition is actually a direct consequence of the second with some restrictions on the spectral density of the noise source process.

Some calculations of photon noise NEP.

There are several references in the literature that give formulas for the calculation of NEP from photon noise on detectors. I give here three references that are particularly interesting because they give different formulations, thus representative of the wide variety of formulations that can be found in the literature, and in addition they present a special interest for the specification of background limited bolometers for the IRAM 30m telescope.

1) In the book "Observational Astrophysics" [Léna], the author introduces the NEP in the chapter dealing with the conversion of n photons into electrons in a photoelectric detector. He define first the detector as a filter with a given quantum efficiency η and an idealized transfer function being a rectangle function Π_f cutting off at a given frequency f_c. Then he calculates the variance of the photocurrent:

$$\sigma_i^2 = \overline{i^2} - \bar{i}^2 = \int_{-\infty}^{+\infty} \left[(\eta \bar{n})^2 \, \delta(f) + \eta \bar{n} \right] \Pi_{f_c}(f) \, df - (\eta \bar{n})^2 = 2 \eta \bar{n} \, f_c$$

where the bar above a letter is a mean value, $\delta(t)$ is the Dirac function, and the term in the integral is the filter spectral density for a poissonian stochastic process. The signal-to-noise ratio for an incident radiation power $P(\nu) = \bar{n}\, h\nu$ is:

$$\frac{i^2}{\sigma_i^2} = \eta \frac{P(\nu)}{2 h\nu \, f_c}$$

Then he applies this calculation to the situation where the detector simultaneously receives photons from a faint source and from a dominating background. The minimum detectable signal ($S/N=1$) is then limited by the background fluctuations so that $i_S = \sigma_B$, which is an incident power at the detector:

$$P_S = h\nu \frac{i_S}{\eta} = h\nu \frac{\sigma_B}{\eta} = h\nu \left(\frac{2 P_B}{\eta h\nu} f_c \right)^{1/2}$$

This power, referred to a unit band-pass $f_c = 1Hz$ is measured in $W/Hz^{1/2}$; it is the **Noise Equivalent Power** from the background on the detector.

2) In the article "Noise Equivalent Power of background limited thermal detectors at submillimeter wavelengths" [Benford et al], the authors start from the mean square fluctuation in the number of photons detected per mode for a telescope of main beam efficiency $\eta_{MB}(\nu)$, a blackbody emissivity $\varepsilon(\nu)$, and an optical efficiency (product of optics transmission and detector absorptivity) $\alpha(\nu)$ [Fellgett et al]:

$$\sigma_n^2 = n(1 + \varepsilon \eta_{MB} \alpha n)$$

Using the Bose-Einstein statistical expression for the number of photons, and multiplying the above expression by the energy per photon $h\nu$ and the number of modes N, then yields the spectral density of the mean square fluctuations in the radiation power detected:

$$\sigma_{P_\nu}^2 = 2N (h\nu)^2 \frac{1}{e^{\frac{h\nu}{kT}} - 1} \left[1 + \frac{\varepsilon \, \eta_{MB} \, \alpha}{e^{\frac{h\nu}{kT}} - 1} \right]$$

where the factor of 2 derives from the fact that a thermal detector is a square law detector and consequently doubles the mean square fluctuations [Robinson].

To calculate the total power mean square variations the authors integrate the expression over the frequency band of detection, and use three assumptions to simplify the expression: (1) the detector performance is limited by the atmosphere which temperature is high enough so that the Rayleigh-Jeans approximation ($n \approx kT/h\nu$) can be use in the photons frequency band of the detector; (2) the detector is in the diffraction limited case so that the number of modes is 2 when including both polarization of light ; (3) the bandwidth of the instrument is narrow and the observed sources are smaller than the beam size so that the number of modes is independent of frequency. Using the first definition for the NEP given in the introduction, assuming that the observed source is a perfect blackbody and the atmospheric background contribute in both the main beam and error beam, the authors establish the NEP expression in the following manner:

$$NEP \propto Noise / Signal$$

$$Signal \propto \eta_{MB}(\nu)\ \alpha(\nu)\ (1-\varepsilon(\nu))$$

$$Noise \propto \sqrt{\varepsilon(\nu)\alpha(\nu)}$$

$$NEP^2 = \int \frac{4\varepsilon(\nu)}{\eta_{MB}^2\ \alpha(1-\varepsilon(\nu))^2} kT\, h\nu \left[1+\varepsilon(\nu)\alpha\frac{kT}{h\nu}\right] d\nu$$

The authors call this last expression the **Background Radiation Equivalent Noise Equivalent Power**, hence defined as the source noise power yielding a signal to noise of 1 when observing a source through the atmosphere.

3) In the article “Photon noise in photometric instrument at far-infrared and submillimeter wavelengths” [Lamarre] the author starts also from the mean square fluctuations of the number of photons, but for the more general case of g cells (or modes or states) of the phase space (g>>1):

$$\sigma_n^2 = \bar{n}\left(1+\frac{\bar{n}}{g}\right)$$

For an incident radiation characterized by a mean optical spectral power P_ν measured by a detector with a quantum efficiency η in the frequency band $\Delta\nu$ during the time t, the mean number of photons detected is:

$$\bar{n} = \frac{Q_\nu\ \Delta\nu\ t}{h\nu}$$

where $Q_\nu = \eta \cdot P_\nu$ is the mean optical spectral power of the effectively detected radiation. The number of modes can be written as the product of the inverse of space coherence Δ_s by the inverse of time coherence $\Delta\nu \cdot t$ [Kastler ?][Mandel ?]. The fluctuations of the absorbed energy are then:

$$\sigma_W^2 = (h\nu)^2\sigma_n^2 = \left(Q_\nu h\nu + \frac{Q_\nu^2}{m\Delta_s}\right) t\,\Delta\nu$$

where the term m has been introduced to take into account the polarization of light in the number of cells available in the phase space. These photons fluctuations induce fluctuations on the detector signal expressed in terms of electrical NEP as:

$$NEP_{ph}^2\ B = \frac{\sigma_W^2}{t^2}$$

where B is the equivalent bandwidth of a perfect integrator with an integration time t and is equal to $1/(2t)$ [Bracewell]. As in Benford et al, the NEP has 2 terms. The first one is purely poissonian (signal grows as $\bar{n}$ and noise as $\sqrt{\bar{n}}$) and can be interpreted as a quantum noise also called shot noise when it occurs in devices. The second term is the boson factor of the radiation and can be interpreted as a phenomenon of photon bunching or interferences of waves at the origin of the diffraction phenomena. The rest of the article is dedicated to a discussion about the space coherent factor. After giving the explicit case $\Delta_S=c^2/(\nu^2 U)$ valid only when the incident radiation is produced by an incoherent source and is uniformly distributed over a detector which beam throughput (U) is large with respect to the etendue of coherence (ν^2/c^2), the author presents the semi classical theory of Hanbury Brown and Twiss (HBT) describing intensity interferometry in visible light [Hanbury Brown and Twiss] and derives a general expression for $\Delta_s(\nu)$, called the partial coherence factor by HBT. Considering the dependencies of Δ_s over the system dimensions and considering the effect of state polarization p the author gives the general expression of the NEP:

$$NEP_{ph}^2 = 2\int h\nu Q_\nu \, d\nu + \left(1+p^2\right)\int \Delta_s(\nu) Q_\nu^2 \, d\nu$$

$$\Delta_s(\nu) = f\left(\frac{r_s r_d \nu}{cZ}\right) \qquad \frac{c^2}{\nu^2 U} \le \Delta_s(\nu) \le 1 \qquad 0 \le p \le 1$$

This is the general expression of the **electric Noise Equivalent Power** created by incident radiations on a detector. In this expression P is the degree of polarization and the partial coherence factor depends on the source dimension r_s, the detector dimension r_d, the distance between the source and the detector Z, and the radiation wavelength $\lambda=c/\nu$.

One can ask whether these three different results for the photon NEP created on a detector are compatible with each other or not ? Do they describe the same characteristics or is there at least some links between them ?
In the following part I will focus on the definition of the Noise Equivalent Power, then based on this definition I will show the links between the different NEP calculations and establish the expression(s) I'll use from now for the specifications of bolometric detectors.

Defining NEP using random process formalism.

The phenomenon of *noise* forms the limit to the measure of everything in nature. Such a central problem of science has obviously been studied and formalized mathematically. Many books and papers deal with the theory of random process, so I recall here only the basics definitions that will allow to study the specific case of the NEP notion. These definitions are mostly from [leclercq], based on the lectures from [Diu][Pouvil][Puech][Borg], and [Léna].

Let's call **x** a random (or stochastic) variable which can take the values x_n or x, depending on they are discreet or continue, with the normalized probabilities P_n or density of probabilities $p(x)$. The mathematical expectancy of any function f of the random variable is:

$$E\{f(\mathbf{x})\} = \sum_n f(x_n) P_n = \int_{-\infty}^{+\infty} f(x) p(x) dx$$

$$\text{where} \quad \sum_n P_n = 1 \quad \text{and} \quad \int_{-\infty}^{+\infty} p(x) dx = 1$$

The mean and the variance of **x** are defined as:

$$m = E\{\mathbf{x}\}$$
$$\sigma_x^2 = E\left\{(\mathbf{x} - E\{\mathbf{x}\})^2\right\} = E\left\{\mathbf{x}^2\right\} - E\{\mathbf{x}\}^2$$

where the last equality is obtain thanks to the distributive property of the integration and the normalization of the probabilities giving $E\{a\,E\{b\} + c\} = E\{a\}E\{b\} + E\{c\}$.

Now let's defined an experiment which possible results ζ have probabilities $P(\zeta)$ and are assigned with a function of time $\mathbf{x}(t,\zeta)$. The family of functions created are called stochastic process, and the notation $\mathbf{x}(t)$ is used to represent this process. The mean, autocorrelation and autocovariance are respectively defined as:

$$m_x(t) = E\{\mathbf{x}(t)\}$$
$$R_{xx}(t_1, t_2) = E\left\{\mathbf{x}(t_1)\ \mathbf{x}^*(t_2)\right\}$$
$$C_{xx}(t_1, t_2) = E\left\{[\mathbf{x}(t_1) - m(t_1)]\,[\mathbf{x}(t_2) - m(t_2)]^*\right\}$$
$$\Rightarrow \sigma_x^2(t) = C(t,t) = R(t,t) - m^2(t)$$

where the asterisk denotes the complex conjugate.

The process is stationary in the simple case it is conserved in time in the probabilistic point of view. In other word a translation in time doesn't change its statistics (mean, variance an so on). Redefining the time variables as $t_1=t$, $t_2=t+\tau$ allow to write the autocorrelation and the autocovariance as depending on a unique time variable (e.i. $R(\tau)$ and $C(\tau)$).

A stochastic process is ergodic when the temporal means are asymptotically equal to the means on all possible realizations of the process, that is to say equal to the mathematical expectancies:

$$\lim_{T\to\infty} x_T \equiv \lim_{T\to\infty} \frac{1}{T}\int_0^T x(t)\,dt = m_x$$

$$\lim_{T\to\infty} R_T(\tau) \equiv \lim_{T\to\infty} \frac{1}{T}\int_0^T x(t)\,x(t+\tau)\,dt = R_{xx}(\tau)$$

The ergodicity, which implies stationarity, is obviously linked to the question of knowing whether the measure during a finite time of a random process is representative of its real physical properties. Associated to the problem of measuring is the inevitable process of filtering the signal. Filtering can be described in the frequency domain and in the time domain, and the Fourier Transform is the useful tool linking the two domains.

The spectral density of a signal is the square of the magnitude of the continuous Fourier transform of the signal, for an ergodic process one has:

$$S_{xx}(f) \equiv \lim_{T\to\infty} \frac{1}{T}\,\tilde{x}(f)\,\tilde{x}^*(f) \qquad \left[x\_unit^2 / Hz\right]$$

$$S_{xx}(f) = FT\left[R_{xx}(\tau)\right] \qquad\qquad R_{xx}(0) = \int_{-\infty}^{+\infty} S_{xx}(f)\,df$$

The equality with the Fourier transform of the autocorrelation constitutes the Weiner-Khintchin theorem. The restriction of the spectral density to the positive frequency domain is called monolateral spectral density: $S_x(|f|) = 2S_{xx}(f)$.

Let consider a system measuring a property x likely to fluctuate with a finite amplitude as an ergodic process in a finite frequency band. The function linking the value of x to the real power dissipated into the system is the sensitivity (or power response) of the system to the measured property: $pr_x\,[x\_unit/W]$. The **Noise Equivalent Power** of the process measured in the detection system is defined as the ratio of the square root of the monolateral spectral density in the finite frequency band over the system sensitivity:

$$NEP \equiv \frac{\sqrt{S_x(f,\Delta f)}}{pr_x} \qquad \left[W/\sqrt{Hz}\right]$$

This is the power dissipated by the ergodic process into the measuring system, expressed in the frequency domain taking into account the system bandwidth. This definition is general in the sense that no assumption is done on the physical process making x fluctuate, and no assumption is done on the physical process linking the value of x to the dissipated power. Though it implies two important restraining points. First, the fluctuating process has to be ergodic. Second the measurement is done in a finite time and in a finite frequency band, therefore the system is a filter giving an estimation of the random process characteristics. This estimation procedure appears in the spectral density expression through the bandwidth variable (Δf). How can we interpret this NEP, in terms of measured signal and signal over noise characteristic of the system ?

The measuring system acts as an operator transforming the input ergodic process $\mathbf{x}(t)$ to a new ergodic process $\mathbf{y}(t)$ at the system output. If the system is linear (obeying associative and distributive rules), then the operator is the convolution product of the input process with the system response $h(t)$. The Fourier transform allow to write the operation in the frequency space as a simple product.

$$y(t) = x(t) * h(t) = \int_{-\infty}^{+\infty} x(t-\tau)\; h(\tau)\; d\tau$$

$$\tilde{y}(f) = \tilde{x}(f)\; \tilde{h}(f)$$

Consider the system acts as an ideal low-pass filter (most measuring systems are low pass filters), which transfer function $\tilde{h}(f)$ can be expressed as a ideal rectangle function, and suppose the input signal is centered (null mean) and its spectral density can be considered as constant (white noise) in the frequency band of the filter, then:

$$\tilde{h}(f) = \Pi\left(\frac{f}{2f_c}\right) \qquad \text{where } f_c \text{ is the filter cutoff frequency}$$

$$\sigma_y^2 = R_{yy}(0) = \int_{-\infty}^{+\infty} S_{xx}(f)\; \tilde{h}(f)\; df \approx S_{xx}\, 2f_c = S_x\; f_c = pr_x\; NEP^2\; \Delta f$$

So for a centered ergodic process with spectral density considered as frequency independent in the measuring system bandwidth Δf=2f_c, the NEP is the power dissipated into the system by unit bandwidth due to the process variance.

Since the measure is done in a finite bandwidth and a finite time t_m (rectangle function in the time domain) it gives only an estimation of the signal characteristics. To know how good is the estimation, it is necessary to calculate the variances of the measured characteristics. For instance the variances of the signal mean is:

$$\sigma_{y_T}^2 = \frac{1}{t_0}\int_{-\infty}^{+\infty}\left[R_y(\tau) - m_y^2\right]\left[\Pi\left(\frac{\tau}{t_m}\right) * \Pi\left(\frac{\tau}{t_m}\right)\right] d\tau \approx \frac{\sigma_y^2}{2f_c\; t_m}$$

As expected, the estimated means tends toward the real mean as t_m increases. This is also true for the signal autocorrelation and autocovariance. Thus the precision on the measured power will increase as the square root of time. This characteristic shows the very interest of the NEP expressed in W/Hz$^{1/2}$; it is an information about the system performances to measure a given process independently of the integration time.

Generally the signal delivered by a receptor can be written as the sum of the useful signal and a background noise term $x_m(t)=x_s(t)+x_b(t)$. The *noise* is the fluctuations causing a difference between the estimators for the signal (measurements made over a finite time t_m) and the true average values of the signal; it is given by the signal variance. The ratio signal to noise is an estimation of the relative differences between the estimator and the wanted quantity [Léna]. In the case of additive independent noise sources, the estimated signal to noise is:

$$\frac{S}{N} = \frac{\overline{x_m} - \overline{x_b}}{\sqrt{\sigma_s^2 + \sigma_b^2}}$$

where the overbar is used to mark the mean over the measurement time.

Let suppose the background noise dominates the measurement, and is ergodic with a spectral density approximately constant in the frequency band Δf and null elsewhere (either due to the noise structure itself or due to the limited detector bandwidth). In terms of dissipated power into the detector, a signal over noise ratio of 1 will lead to:

$$P_s = \frac{\overline{x_s}}{pr} = \frac{\sigma_{\bar{b}}}{pr} = NEP\sqrt{\Delta f}$$

So the **Noise Equivalent Power** can be interpreted *the input signal power that produces a signal-to-noise ratio of unity at the output of a given detector at a given data-signaling rate or modulation frequency, and effective noise bandwidth; it is the minimum detectable power per square root bandwidth.* This "practical" definition is the one given in the introduction, and in the limit of the special conditions given above for the noise and the detector, it is indeed equivalent to the "statistical" definition using the monolateral spectral density. As it is shown in the next section, this is particularly true for background limited optical detectors.

Calculating the photon noise NEP for background limited detectors.

Let start this study from the basics of statistical physics which establish for a system of free particles in a thermostat (box fixing the mean energy of the particles) the probability that a state l is occupied [Diu]:

$$P_l = \frac{e^{-\beta(E_l - \mu N_l)}}{Z} \quad \text{with the partition function} \quad Z = \sum_l e^{-\beta(E_l - \mu N_l)} \quad \text{and}\ \beta = \frac{1}{kT}$$

where E is the system energy, N the number of particles, μ the chemical potential, k the Boltzmann constant, and T the system temperature.

Because they are bosons, an unlimited number of photons can occupy a cell l of the phase space. Noticing that $(1-x)\sum_0^{n-1} x^n = 1 - x^n$, the partition function ζ_l of a state l can be written as a simple expression without the sum sign, and since the occupation probability P_l does not depend on the other states, the mean number of bosons in the state l (or mode occupancy number) can also be express as a simple expression called the Bose-Einstein statistics:

$$\zeta_l = \sum_{n_l=0}^{\infty} X_l^{n_l} \quad \text{where} \quad X_l \equiv e^{-\beta(E_l-\mu)} \Rightarrow P_l = \frac{e^{-\beta(E_l-\mu)n_l}}{\zeta_l} = X_l^{n_l}\left(1-X_l\right)$$

$$\bar{n}_l = \sum_{n_l} n_l P_l = (1-X_l)X_l \frac{d}{dX_l}\sum_{n_l} X_l^{n_l} = \frac{(1-X_l)X_l}{(1-X_l)^2} = \frac{1}{e^{-\beta(E_l-\mu)}-1}$$

Using the same reasoning, one can deduce the expression of the variance of the number of photons:

$$\overline{n_l^2} = \sum_{n_l} n_l^2 P_l = (1-X_l)X_l \frac{d}{dX_l}\sum_{n_l} n_l X_l^{n_l} = X_l\left(\frac{1}{1-X_l} + \frac{2X_l}{(1-X_l)^2}\right) = \bar{n} + 2\bar{n}^2$$

$$\Rightarrow \quad \sigma_l^2 = \overline{\delta n_l^{\,2}} = \overline{n_l^{\,2}} - \bar{n}_l^{\,2} = \bar{n}_l\left(1+\bar{n}_l\right)$$

Let consider a system with g cells of phase space and let's call n the total number of photons. One has:

$$\bar{n}_l = \frac{\bar{n}}{g} \Rightarrow \sigma^2 = g\sigma_l^2 = \left(\bar{n} + \frac{\bar{n}^2}{g}\right)\left(\frac{g-1}{g+1}\right)$$

The term (g-1)/(g+1) take into account the fact that photons can "travel" from one cell to another [Kastler], when g is big enough this term can be neglected and the expression is identical to the one presented in the introduction, with poissonian component plus a photon bunching component.

The problematic for a real detector receiving photons is to have a good estimation of the number of states (or cells of the phase space) available. The rigorous demonstration giving a correct and general expression of the number of available states would require a longer and more technical paper, so I choose to give a feeling of the estimation of this number thanks to the basics of electromagnetism and physical reasoning on the extremes asymptotical cases. The more general and correct formulation is based on the semiclassical HBT theory [Hanbury Brown and Twiss], and can be found in other papers [Lamarre].

Exactly like the demonstration of the Planck law of Black Body radiation [Diu], we start the reasoning from the Maxwell equations allowing to formulate the electromagnetic field as progressive waves into a closed box: $\vec{E} \propto \exp\left[i\left(\vec{K}\vec{r} - \omega t\right)\right]$. The wave vector verifies $\left|\vec{K}\right| = \omega/c$ and $K_i = n_i\, 2\pi/L$, where ω is the wave pulsation, c is the speed of light, L is the typical length of one of the box side, n_i is a positive integer. The total number of modes $\vec{K}$ available between K and $K+dK$ is the volume of the sphere skin with a thickness dK divided by the minimal size of one mode:

$$g = \frac{4\pi K^2\, dK}{(2\pi/L)^3} = 4\pi V \frac{\nu^2}{c^3} d\nu$$

where V is the box volume and $\nu = \omega/2\pi$ is the wave frequency.

For a detector with a surface A_d receiving photons from a solid angle Ω, during a time t_m, the box of available modes is the cylinder of surface A_d and height $c{\cdot}t_m$. Because the number of available directions for the vectors $\vec{K}$ is limited by the incoming solid angle, the volume of available modes between K and $K+dK$ is not a sphere anymore but a cap. Thus the number of modes available is:

$$g = \frac{\Omega K^2\, dK}{(2\pi)^3/V} = \left(\frac{A_d \Omega}{\lambda^2}\right)\left(t_m\, d\nu\right)$$

where $\lambda = c/\nu$ is the photons wavelength. Since $\Omega \approx A_s / Z^2$, where A_s is the surface of the source and Z is the distance between the source and the detector, one can define a quantity conserved all along the light path called the beam throughput $U \equiv A\Omega = A_d A_s / Z^2$ (this is the variable U used in the introduction). The term $\lambda^2/A\Omega$ can be interpreted as the spatial coherence factor of the beam, and the term $1/t_m d\nu$ as the time coherence factor. From quantum mechanics principle, the minimum number of states available is an integer and can't be less than one. This reasoning actually holds for each one of the coherence factors so that the minimum value for the number of modes available is $g=1$. One can feel that the minimum size for a photon state (phase cell) will be limited by the diffraction pattern on the detector plane. This result appears in the demonstration of the Zernicke-van Cittert theorem which can be stated as: *if the linear size of a quasi-monochromatic radiation source and the distance between two points of its image on a screen are both small compared with the distance between source and screen, the modulus of the complex degree of coherence is equal to the modulus of the spatial Fourier transform of the source intensity, normalized by the total intensity of the source* [Léna] [Born and Wolf]. One could reach the same results following Heisenberg's uncertainty principle. These reasoning are also valid in the time domain for the time coherence factor [Mandel] [Kastler], but we will assume that we always verify $t_m \cdot d\nu > 1$, which is the case for all instrument measuring higher frequencies than the mid range radio domain. Taking into account the degree of polarization ($p=0$ for an unpolarized ray, and $p=1$ for a ray polarized in one direction), the generalized expression of the number of modes available and the spatial coherence factor $\Delta_s(\nu)$ can be written as:

$$g = \frac{t_m \, d\nu}{\Delta_s(\nu)\left(1+p^2\right)/2} \qquad 1 \le \Delta_s(\nu) \le \frac{\lambda^2}{A\Omega}$$

$$\Delta_s(\nu) = \frac{\iint\limits_{\vec{\rho}_1} \iint\limits_{\vec{\rho}_2} \iint\limits_{\vec{r}_1} \iint\limits_{\vec{r}_2} \sqrt{\prod_{i,j} I_\nu(\vec{\rho}_i, \vec{r}_j)} \, \cos\left[\frac{2\pi\nu}{cZ} \Delta\vec{\rho}\,\Delta\vec{r}\right] d^2\vec{\rho}_1 \; d^2\vec{\rho}_2 \; d^2\vec{r}_1 \; d^2\vec{r}_2}{\left(\iint\limits_{\vec{\rho}} \iint\limits_{\vec{r}} I_\nu(\vec{\rho},\vec{r}) \; d^2\vec{\rho} \; d^2\vec{r}\right)^2}$$

where I_ν [W/(m^2 sr Hz)] is the specific intensity of the light received on the infinitesimal surface element d^2r of the image from the infinitesimal element $d^2\rho$ of the source, and Z is the distance between the source and the image. Checking that the asymptotical values of coherence factor are indeed the one we gave with the previous simple reasoning, is curiously easier for the case $\Delta_s(\nu)=1$ than for $\Delta_s(\nu)=\lambda^2/A\Omega$. Indeed, when the beam throughput is much smaller than the coherence throughput $\left((\Delta\rho\,\Delta r/Z)^2 << \lambda^2\right)$, the cosine term tends toward 1 and therefore $\Delta_s(\nu)=1$. In the case of a large uniform beam such that the specific intensity can be considered as constant, the mathematical trick to perform the integration consist of using the Fourier transform of the intensity to make the cosine term disappear and use the Parseval-Plancherel theorem (if, $F(\omega) = FT\left[f(x)\right]$, then $\int_{-\infty}^{+\infty} \left|f(x)\right|^2 dx = (1/2\pi)\int_{-\infty}^{+\infty} \left|F(\omega)\right|^2 d\omega$) to come back in the real space so that the numerator of $\Delta_s(\nu)$ is equal to $I_0^2 \, A_s \, A_d \, (cZ/\nu)^2$. As can be seen the general formulation is not really practical, hopefully in the case the detector efficiency is uniform on its surface, the spatial coherence factor can be written in a more simple way allowing to calculate it in practice for any beam size:

$$\Delta_s(\nu) = \frac{1}{A_d^2} \iint_{A_d} \iint_{A_d} \frac{\left| L\left((x-x')\nu/cZ\,,(y-y')\nu/cZ\right)\right|^2 dx\,dy\ \ dx'\,dy'}{\left| L(0,0)\right|^2}$$

where L is the Fourier transform of the source brightness ($B_\nu(\vec{\rho}\,)$) at frequency ν. In the case of a punctual source viewed through a circular pupil, it will be Airy pattern. The value of the coherence factor depends only on the value of $\mu = (r_s r_d \nu)/(cZ)$, where r_s and r_d are the characteristic dimensions of the source and the detector.

Now consider a source emitting a mean spectral power P_ν. Since each photon caries and energy $h\nu$, the mean spectral power is related to the mean number of photons emitted by the source during a time t_m by the simple expression $P_\nu d\nu = \overline{n}_e\, h\nu / t_m$. Consider that between the source and the detector output the signal is attenuated by a factor $\gamma(\nu)$, which can gather many different things that we will discuss latter (emissivity, efficiency, transmission, absortivity). If Q_ν is the spectral power dissipated into the detector then $Q_\nu=\gamma(\nu)P_\nu$. Let’s define n_d as number of photons detected so that $Q_\nu d\nu = \overline{n}_d h\nu / t_m$. The fluctuations of the numbers of photons detected during the time t_m is:

$$\overline{\delta n_d^2} = \overline{n}_d + \frac{\overline{n}_d^{\,2}}{g} = \frac{Q_\nu t_m\, d\nu}{h\nu} + \frac{\Delta_s(\nu)}{t_m\, d\nu}\left(\frac{Q_\nu t_m\, d\nu}{h\nu}\right)^2$$

The trick that will allow to link together all the definitions and results shown in the introduction resides in the definition of the measuring system power response (or sensitivity). If one includes some attenuation factor γ_s into the power response of the detector, then the power response is the conversion factor between the number of photons before the objects responsible for the attenuation(s) and the power dissipated into the detector. So the general expression of the power response is:

$$pr = \frac{\gamma_s(\nu)}{h\nu / t_m}$$

This possibility to include an attenuation factor into the detector power response will lead to two different interpretation of the definition of the NEP.

- If the power response has no attenuation factor (or γ_s=1) then the NEP is the power dissipated by the noise source into the detector, that is to say the noise <u>*power at the output of the detector*</u>. As in Lamarre’s article I call this NEP the **electrical NEP**, it is general in the sense that it is generally used for other noise sources than incoming radiations; for example it is applied for the Johnson noise (or thermodynamic noise) created in a resistor or for any other electrical noise such as shot noise in a transistor, or phonon noise in devices sensitive to electron-phonon coupling, or many other noise processes.
- If the power response includes an attenuation factor smaller than unity, then the NEP is the <u>*power at the input of the detector*</u> that would produce the same signal level as the noise source. This is equivalent to the first definition given in the introduction. I call it the **optical NEP** as a reminder of what seems to be its most common use (according to the ease one can find the Federal Standard 1037C definition of NEP on the internet, and Lamarre uses also this term to distinguish the two meaning of the NEP). Though this definition does not specify what kind of power is at the detector input; it is actually not

necessarily an optical power ! Moreover this definition does not specify what is considered as the input of the detector; as it will appear in the applications below it is sometime useful to include part of the detector's environment after the input. It is important to stress that in these conditions the attenuation factor $\gamma_s(\nu)$ in the power response is not necessarily the same as the attenuation factor $\gamma_b(\nu)$ giving the relation between the number photons emitted by the noise source and the number actually detected.

Using the last equation of the previous section, the general expression of the noise equivalent power due to the noise source on the detector will be:

$$NEP\sqrt{df} = \frac{\sigma_{\bar{b}}}{pr} = \frac{h\nu\sqrt{\overline{\delta n_{bd}^2}}}{\gamma_s(\nu)\, t_m} \propto \frac{\sqrt{\gamma_b(\nu)}}{\gamma_s(\nu)}$$

Integrating in a band Δν of the frequency domain and considering that the system is a perfect integrator (rectangle function) with a bandwidth Δf, and taking care of the polarisation, will give the most general expression for the **photon NEP**:

$$NEP^2 = \frac{1}{t_m \Delta f}\left(\int_{\Delta\nu} h\nu \frac{\gamma_b(\nu)}{(\gamma_s(\nu))^2} P_\nu(\nu)\, d\nu + \int_{\Delta\nu} \frac{(1+p^2)\Delta_s(\nu)}{2}\left(\frac{\gamma_b(\nu)}{\gamma_s(\nu)} P_\nu(\nu)\right)^2 d\nu\right)$$

The minimum measurement time is usually chosen to satisfy the Shanon-Nyquist criteria so that all the information filtered by the detector is sampled. So if the system is a perfect integrator one has $1/t_m=2\Delta f$. If the system is not a perfect integrator, then one has to consider Δf as a the equivalent noise bandwidth (ENBW) which is defined as the bandwidth of an ideal low-pass filter which could pass the same power of white noise than a real filter [Gualtieri]. So calling f_c the cutting frequency of a first order low pass filter, and $\tau=RC$ its time constant, the calculation gives (see [Leclercq]) $\Delta f = f_c \pi/2 = 1/(4\tau)$.

Now let check that with the previous expression one can find the three formulations given in the introduction when using different assumptions about the various components of the equation.

1) Suppose the measurement time satisfies the Shanon-Nyquist criteria ($1/t_m=2\Delta f$), the system power response does not include any attenuation factor ($\gamma_s=1$), and the attenuation factor between incoming photons of the noise source and the detected photons is the quantum efficiency of the detector ($\gamma_b=\eta$). These conditions give:

$$NEP^2 = 2\int h\nu\,\eta P_\nu\, d\nu + (1+p^2)\int \Delta_s(\nu)\,(\eta P_\nu)^2\, d\nu$$

This is exactly the same formula than Lamarre's NEP_{ph}; the (electrical) Noise Equivalent Power for *detected* photons.

2) Suppose the measurement time satisfies the Shanon-Nyquist criteria ($1/t_m=2\Delta f$), the system power response includes the quantum efficiency of the detector ($\gamma_s=\eta$), which is also the only attenuation factor applied to the noise source ($\gamma_b=\eta$), the number of modes is so high that the photon noise can be considered as purely poissonian ($A\Omega>>\lambda^2$ so that $\Delta_s\approx 0$), and the radiation is monochromatic ($P_\nu(\nu)=\delta(\nu)P_b$). These conditions give:

$$NEP^2\Delta f = 2\Delta f\,\frac{h\nu\, P_b}{\eta}$$

This is exactly the same formula than the expression of P_S in Léna's book; the (optical) Noise Equivalent Power of *incident* photons in a totally incoherent beam.

3) Suppose the measurement time satisfies the Shanon-Nyquist criteria ($1/t_m=2\Delta f$), the light has two polarization (not polarized p=0), the system power response includes the main beam efficiency ($\eta_{MB}(\nu)$ applied for punctual sources), the atmosphere transmission ($t_a(\nu)=1-\varepsilon(\nu)$), and the optical efficiency (product of the optics transmission and the detector absorptivity $\alpha(\nu)=t_o(\nu)\cdot a(\nu)$) so that its attenuation factor can be written $\gamma_s=\eta_{MB}\alpha(1-\varepsilon)$, the noise source is the atmosphere which is extended (not attenuated by the main beam efficiency) and its radiation is attenuated by the optical efficiency $\alpha(\nu)$ so that $\gamma_b=\alpha$, and finally the integration window of the wave frequencies $\Delta\nu$ is small enough to consider optical efficiency and main beam efficiency as constant. These conditions give:

$$NEP^2=\int_{\Delta\nu}\frac{2\alpha}{\left(\eta_{MB}(1-\varepsilon(\nu))\alpha\right)^2}\;P_\nu\left[h\nu+\Delta_\nu(\nu)\alpha\frac{P_\nu}{2}\right]d\nu$$

We want to explicit the spectral power using macroscopic measurable parameters. To do that we simply use the equation concept that allowed us to introduce the spectral power and we explicit the number of emitted photons a function of the atmosphere temperature thanks to the mode occupancy number (n_l) calculated previously. Using $n_e=gn_l$ would give the blackbody spectral density (for a rigorous demonstration using all the "ingredients" presented previously of the Planck law of the blackbody radiation see a statistical physics book [Diu]), but the atmosphere is not a perfect blackbody and the number is attenuated by the emissivity $\varepsilon(\nu)$ (see references dealing with radiation transfer for a generalization of thermal radiation to "greybody" and introduction to the concepts of opacity, emissivity and transmission [Leclercq] [Born and Wolf ?]) so that:

$$P_\nu d\nu=\bar{n}_e\,h\nu/t_m=\varepsilon(\nu)g\bar{n}_l\,h\nu/t_m$$

$$\Rightarrow\;P_\nu=\varepsilon(\nu)\frac{2h\nu}{\Delta_s(\nu)}n_l$$

where the factor 2 comes from the two polarizations of light ($p=0$). Assuming the atmosphere thermal energy is much higher than the detected photons energy ($kT>>h\nu$), one can use the Rayleigh-Jean approximation for the mode occupancy number and deduce from the previous expression of the NEP a new expression:

$$\bar{n}_l=\frac{1}{e^{\frac{h\nu}{kT}}-1}\xrightarrow{kT>>h\nu}\frac{kT}{h\nu}\qquad\Rightarrow\qquad P_\nu(kT>>h\nu)=\frac{2\varepsilon(\nu)}{\Delta_s(\nu)}kT$$

$$NEP^2=\int\frac{1}{\Delta_s(\nu)}\frac{4\varepsilon(\nu)}{\eta_{MB}^2\,\alpha(1-\varepsilon(\nu))^2}kT\,h\nu\left[1+\varepsilon(\nu)\alpha\frac{kT}{h\nu}\right]d\nu$$

This would be exactly the same formula as the one giving the *Background Radiation Equivalent NEP* in Benford's article at the condition that detector is smaller than the Airy disk created by the diffraction through the telescope pupil, that is to say $A\Omega<\lambda^2$ so that $\Delta_s\approx1$. In his article, Benford uses the number of modes N instead of the inverse of the spatial coherence factor, and says $N=A\Omega/\lambda^2=1$ at the diffraction limit. If the size of the detector is equal to the FWHM (full width half maximum) of the Airy disk, then one has indeed $A\Omega/\lambda^2=1$, but as we saw previously, $1/\Delta_s$ is actually underestimated, which could explain why he talks about excess noise afterward in his article. Benford call this NEP the Background Radiation Equivalent NEP to stress that it gives the incident power that a source outside the atmosphere should have to create a signal over noise of one when the noise is dominated by the atmosphere radiation.

Conclusion about NEP.

Most of the noises phenomenon occurring in devices are due to thermodynamic fluctuations of some quantities that are statistically described by the realizations of stationary and ergodic random processes with finite power. The mean square fluctuation, called variance, of a random process is at the origin of the noise signals in the devices. The square modulus of the ergodic process Fourier Transform is called the spectral density (or power spectral density, even though the word power refers to the square of the signal and not to the Watt unit !). The spectral density is shown to be the Fourrier Transform of the variance and describes fully the noise process, its unit is the square of the process unit per hertz. In quadratic detectors the signal measured is the square of the process amplitude (for example electromagnetic waves), or more specifically the energy or power of the process. The sensitivity or power response is the quantity giving the correspondence between the process unit and the power dissipated into (or detected by) the detector, its unit is the process unit per Watt. The noise Equivalent Power is equal to the square root of the spectral density divided by the power response of the device; it has the units of Watt per square root hertz. The concept of NEP is particularly useful for the characterization of quadratic detectors; it quantifies the noise level in the signal measured by the detector in terms of dissipated (detected) power or incident power. When the noise can be considered as white in the detector bandwidth, the NEP is independent of any time variable so that the noise power varies only as the square root of integration time. Some defines the NEP as the incident radiation power that would produce a signal over noise of unity. In that case the power response is corrected with the detector efficiency and the result is the correspondence between the process unit and the power incident to the detector, not the power dissipated into it. Because of this ambiguity about the power at the input or the output of the detector one can define several NEPs. From the various usages found in the literature one can establish four major definitions which are not equivalent but are linked to each other:

The **statistical NEP** is the most general definition: square root of spectral density versus power response, giving the power dissipated into the detector due to the noise.

The **electrical NEP** is the same as the statistical NEP but in the particular case that the ergodic process creating the noise is white in the system bandwidth.

The **optical NEP** is the electrical NEP divided by detector quantum efficiency. It is the power at the entrance of the detector due to the noise source.

The **Background Radiation Equivalent NEP** is a particular case of optical NEP for ground observatories where the atmosphere is the dominant noise source and is included into the detector definition so that the NEP is the power that a punctual source should have *before the atmosphere* to obtain a signal over noise of unity.

Appendix: NEFD, NET and integration time.

To know the performance of a detector to observe a given source the BRE NEP can be very interesting. But in astrophysics sources are rarely described in terms of power, which implies to know the detector throughput and is therefore not universal. The concept of *flux* (or flux density) is independent of any detector geometry so it is often used to describe a source. The flux is defined as the power per unit surface per signal frequency. The usual flux unit in radioastronomy is the Jansky: $1Jy=10^{-26}W/m^2/Hz$.

The **Noise Equivalent Flux Density** is defined as the level of flux density required to obtain a unity signal to noise ratio in one second of integration with the detector. The flux required is at the input of the detector, but again there are some ambiguities in this definition. Some do not include the atmosphere, some do. Some includes the observing mode, some don't. To write a general expression for the NEFD derived from the NEP, let define two correction factors: o_m the observing mode factor and γ_x the attenuation factor multiplied to γ_s in the NEP expression to specify the location where the equivalent flux is calculated. The general expression for the NEFD of a detector with a spectral width $\Delta\nu << \nu$ and a collecting surface A_d is then:

$$NEFD = o_m \frac{NEP(\gamma_x)}{A_d\ \Delta\nu} \approx o_m \frac{NEP}{A_d \int_{\Delta\nu} \gamma_x d\nu} \qquad \left[\frac{\text{Jy}}{\sqrt{\text{Hz}}}\right]$$

Where the notation $NEP(\gamma_x)$ is used to indicate that in the NEP expression one has to use $\gamma_x\gamma_s$. If the NEP used is the optical NEP and γ_x=t_o the transmission of the optics bringing the signal to the detector, then the NEFD is the flux at the entrance of the optics that would create a unity signal to noise ratio at the detector output per second. Often in the literature people define the NEFD out of the atmosphere. If the sky is observed continuously, the observing mode would be o_m=1. But in radioastronomy a very common observing mode, called ON-OFF, consist of subtracting the atmosphere power using the difference of too nearby fields. The noise of the two fields images will add quadratically, increasing the NEFD by a factor $\sqrt{2}$. In addition, in the basic application of this mode only half of the time is spent on the source, and since the noise power increases as the square root of time, the NEFD will be increased by another $\sqrt{2}$ factor. To avoid several calculation steps, some authors include these correction factors in their calculation and give directly the NEFD for ON-OFF observations, hence in that case o_m=2.

In the radioastronomy community the unit of temperature (Kelvin) is sometimes even more used than the Jansky to characterize a source flux. The conversion is done with the Rayleigh-Jeans approximation of the Planck formula of the blackbody radiation (explained in the previous section). Naturally another concept is used to express the noise in terms of temperature; the **Noise Equivalent Temperature**. The NET is defined similarly to the NEP for systems measuring a property with temperature variations, like thermometers. Thus the NET definition uses the system thermal response $tr_x\left[x\_unit/K\right]$ instead of the power response:

$$NET \equiv \frac{\sqrt{S_x(f,\Delta f)}}{tr_x} \qquad \left[K/\sqrt{Hz}\right]$$

The calculation of the power emitted by a source at 1K Rayleigh-Jeans, gives the conversion factor to calculated a NET knowing a NEP:

$$NET = NEP \frac{1[K]}{P_{RJ}(1K)[W]}$$

Another information very often used in literature to characterize instruments is the integration time t_σ necessary to detect a given source with a signal to noise ratio σ. Since $NEP\sqrt{\Delta f} = P_{noise}$ and Δf=1/2t_m, the integration time is calculated from the NEP, or NEFD or NET as:

$$t_\sigma(P_s) = \frac{1}{2}\left(\sigma\frac{NEP}{P_s}\right)^2 \qquad t_\sigma(F_s) = \frac{1}{2}\left(\sigma\frac{NEFD}{F_s}\right)^2 \qquad t_\sigma(T_s) = \frac{1}{2}\left(\sigma\frac{NET}{T_s}\right)^2$$

To get rid of ½ factor in the time expressions, some include it in the Noise Equivalent parameters definitions, and indicate this choice using seconds in the parameters:

$$NEP' = \frac{NEP}{\sqrt{2}} \ \left[W \cdot \sqrt{s}\right] \qquad NEFD' = \frac{NEFD}{\sqrt{2}} \ \left[Jy \cdot \sqrt{s}\right] \qquad NET' = \frac{NET}{\sqrt{2}} \ \left[K \cdot \sqrt{s}\right]$$

The time dimension is sometimes used differently; the system performances are given as the detection limit giving signal to noise ratio σ after a given amount of time (for example the flux giving σ=3 after 2 hours).

References.

[Benford et al]
[Lamarre]
[Leclercq]
[Léna]

[Goodrich]
[Atis]
[Anghel et al]
[Das et al]
[Besse]
[Dutoit]
[Zweiacker]
[Fellgett et al]
[Robinson]
[Kastler ?][Mandel ?]
[Bracewell]
[Hanbury Brown and Twiss]
[Diu]
[Pouvil][Puech][Borg]
[Born and Wolf]
[Gualtieri]