Harnessing Nature's Randomness: Physical Random Number Generator

G. A. Barbosa

Abstract— **Random number generators are indispensable for a multitude of tasks; from electronic games to secure communications. Most generators have been made either in software or determinist hardwired devices such as the Linear-Feedback-Shift-Registers; while gaining in costs or speed, the "random" sequences generated are actually deterministic, obeying clear generating algorithms, despite all randomness appearance of their outputs. From the other side, Nature presents a multitude of sources of true randomness that can be explored. Commercial random generators exist based on physical processes as the source of randomness. Difficulties are always present to extract Nature's randomness. This paper presents guidelines for construction of a fast (telecommunication speed) Physical Random Number Generator. It discusses the fundamental physical elements involved, technicalities of signal recording and its limitations, and the final bit extraction. The need for randomness tests is emphasized and the impossibility of guaranteeing true randomness of a finite sequence is discussed.**

Keywords— **Random, Physical processes, Cryptography.**

I. Introduction

RANDOM processes are intrinsic to Nature. This statement is usually accepted as a fundamental truth in Physics, together with the assumption that this randomness pervades the whole Universe. Quantum Mechanics itself is based on this randomness assumption. However, these beliefs are not unanimous [1].

A reason for this non-unanimity –even nowadays– is that probing Nature's randomness is a daunting task. Man's interaction with a random process interferes with the process itself or forcefully introduce filters during observation or detection of the involved phenomenon. The obtained outcome is always some biased picture of the fundamental process. This biasing is mostly created by the detecting instrumentation. Another fundamental problem is that the finite time window necessary to acquire data leads to samples of finite length. Being finite, they cannot fully characterize the random-like phenomenon: momentum powers of all orders necessary to a full characterization of a generic probability distribution cannot be obtained. One has to be satisfied with approximate results, not with the un-achievable idealized goal. Nevertheless, it is assumed that it is possible to obtain records of this "filtered" and finite set of data that passes many or all available statistical tests for a random phenomenon. One should be satisfied if no deterministic patterns are seen –the pragmatic approach normally used.

In a distinct way, man-made devices –consisting of electronic circuitry or software based– designed to produce the most possible randomness are *deterministic* devices by principle, regardless the complexity level that could be associated to them [2] such as the use of nonlinearities or superpositions of complex processes. These Pseudo Random Number Generators (PRNG) encompass the large majority of random generators in use nowadays.

Besides the randomness necessary for security applications other predicates are usually considered for a random number generator, including speed and cost. Speed is the second most desired feature, necessary for telecommunications.

Physical Random Number Generators (PhRG), by its turn, are devices trying to harness the random characteristics inherent to some physical phenomena. A PhRG designed to last in the existing fast advancing technological scenario should operate in *principles* that are untouched by the technology itself. As such, technological improvements can be incorporated in the system without modifications of the physical source being probed.

PRNGs have been widely described in the literature while PhRG implementations are not so common. In principle, PhRGs are free of the "deterministic" tag. Recent PhRG implementations include devices recording single-photon events [3] (detectors placed at the two ports of a beamspliter (BS)), Nyquist electrical noise [4] and chaotic lasers [5]. These PhRG implementations are not free of problems such as: the existence of bounds on speed due to the need of weak laser intensities [3] to avoid appearance of photons in both BS ports, slowness of electrical noise based processes [4], instabilities [5]. Nevertheless, they are a step forward in achieving true randomness, when compared to PRNGs.

This work shows steps necessary to construction of a Physical Random Generator (PhRG) based on the observation (fast detection and recording) of an elementary random *physical* phenomenon: photon number fluctuations at very short sampling times. The discussed device is aimed to extract intrinsic short-time intensity fluctuations of a coherent field (laser light) to produce random streams in a rate adequate for telecommunications.

II. Basic conditions for light sampling

A lasing device, fed by a current with a random stream of a large number of electrons can be used as the light source for a PhRG. A normal laser, gaseous or semiconductor, would fulfil this condition. The light state generated by a laser is well described by a coherent state where $\langle n\rangle = |\alpha|^2$ is the average number of photons in one coherence time τ_c, α is the complex laser amplitude, and the photon number variance is $\sigma^2 = \langle(\Delta n)^2\rangle = \langle n\rangle$; σ is the standard deviation. The probability for occurrence of n photons in a coherent state within sampling times $\Delta t \ll \tau_c$ is Poissonian distributed

$$p(n) = \frac{e^{-\langle n\rangle}\langle n\rangle^n}{n!}\,. \tag{1}$$

The probabilistic occurrence of these photon numbers reflects existing quantum fluctuations inherent to Nature and, in principle, they exist at all frequencies. The Poissonian occurrence of photon numbers has been called light's "shot-noise", like bal-

G. A. Barbosa, PhD, CEO of QuantaSec – Consulting and Projects in Physical Cryptography Ltd., Av. Portugal 1558, Belo Horizonte MG 31550-000 Brazil. Phone: +11 55 (31) 3441–4121, e-mail: GeraldoABarbosa@gmail.com

Invited Paper

listic occurrences of independent events (e.g., $\sim$ rain drops on a roof).

A single-mode laser will be discussed for this work. In the photon shot-noise limit (where the light noise predominates over other noise sources), intensity measurements can be performed to observe short time fluctuations ΔI, that deviate from the mean intensity $\langle I\rangle$ according to the Poissonian statistics (1):

$$\frac{\sqrt{\langle(\Delta I)^2\rangle}}{\langle I\rangle}=\frac{\sqrt{\langle(\Delta n)^2\rangle}}{\langle n\rangle}\rightarrow\frac{1}{\sqrt{\langle n\rangle}}\,. \tag{2}$$

Eq. (2) shows that the relative noise decreases as $\langle n\rangle$ increases. This makes deviations from the average intensity of an intense laser very difficult to be detectable.

A crucial characteristics associated to the statistical distribution given by (1) is that successive photon numbers, n_1 and n_2, present *no* correlation: $\langle n_1 n_2\rangle=\langle n_1\rangle\langle n_2\rangle$. This is the main property that guarantees that *if* one is able to extract these inherent fluctuations to generate bits, no correlations will appear among them.

As the physical phenomenon itself presents no bandwidth limitation, the PhRG can be made to follow any advances in optoelectronic technology. Usually the main speed restrictions arise from the light detector itself and the amplification circuitry bandwidth.

The device to be discussed [6] relies on the properties of a coherent light state, such as the one produced by a laser working well above threshold but with a damped intensity to increase the relative light fluctuations. This damping should be made by gray light filters (absorbers) without decreasing the laser current itself. This way, the coherent properties are preserved while the relative fluctuations are increased (See Eq. (2)). It may appear that decreasing the laser current could be a simpler way to get the desired low intensity. However, decreasing the current to obtain the desired light levels could put the laser close to the lasing threshold. Close to this threshold, the photon statistics are similar to the statistics of thermal fields. Differently from coherent state statistics, thermal fields present photon number correlations. These correlations are reflections of photon bunching that, given the occurrence of a photon, it is quite probable that a second photon will occur [7]. Therefore, whenever low intensities are desired the coherent state intensity shall not be diminished by drastically reducing the laser current. Instead, it should be damped with neutral filters. This avoids the mixing with thermal field statistics. This mixing produces detrimental features that will show up when statistical tests for randomness become more stringent.

Eventual integration of light source and detector on the same chip should produce the best light source and detector combination. However, this integration is not trivial nowadays but one should expect it to become more accessible in the near future. At the moment, coupling a laser source with a on-chip electronics circuitry is the way to go.

The detection circuitry, following a fast optical detector and a fast analog to digital conversor (AtoD) should discretize the input analog signals and discriminate for signals above and below the average intensity value. This "above" or "below" signals will be converted in fixed amplitude signals + or −, constituting the bit sequence. It is frequent that detection electronics are not perfectly symmetric in the charging and discharging of its circuitry. This may lead to asymmetric amplitude distributions and procedures are usually taken to minimize this problem. One way is to work with the time derivatives of the fluctuating signals [5], or else, a differential phase shift keying scheme (DPSK) can be used, where the difference of two successive modulations defines the bit, either 0 if no change occurs or 1 if a change has occurred.

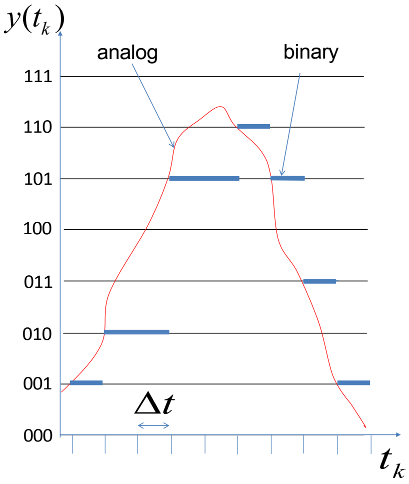

Figure 1. Binary levels (blue) representing an analog signal (red) varying in time t, with a 3-bit ADC. The number of available levels, above zero, is 2^3-1. With an n-bit ADC and digital time units Δt, the digital output values are given by $y(t_k)=b_n(t_k)2^{n-1}+b_{n-1}(t_k)2^{n-2}+\ldots+b_2(t_k)2^1+b_1(t_k)2^0$ $(t_k=k\Delta t)$. Following this rule, the notation at the plot ordinate represents the sequence of three bits b_3, b_2, b_1.

III. Mesoscopic states and bit recording

As Eq. (2) indicates, for large intensity, the relative number fluctuation goes to zero. Some constraints already discussed are here repeated to emphasize general requirements to achieve the operational level:
1) A low current, lasers deviate from the coherent state operation. Therefore, to obtain coherent states in the *mesoscopic* regime (above strictly quantum but below very large intensities where fluctuations are negligible), the laser intensity is decreased through use of neutral filters and *not* by decreasing the electronic current.
2) Usual communication detectors do not have single photon sensitivities. Their minimum detection threshold are usually of the order of a few hundred of photons. This indicates that one has to utilize photon numbers, in Δt, above $\langle n\rangle_{\Delta t}\sim 10^3$.
3) A third constraint relates to the use of fast (linear) analog-to-digital (AtoD) recorders. An analog signal s is digitally recorded using b bits that produce 2^b levels equally spaced. See Fig. 1. Analog signal levels occurring within bit levels are rounded by the digital technique. The spacing between the binary levels define the available ADC's resolution. An ADC with

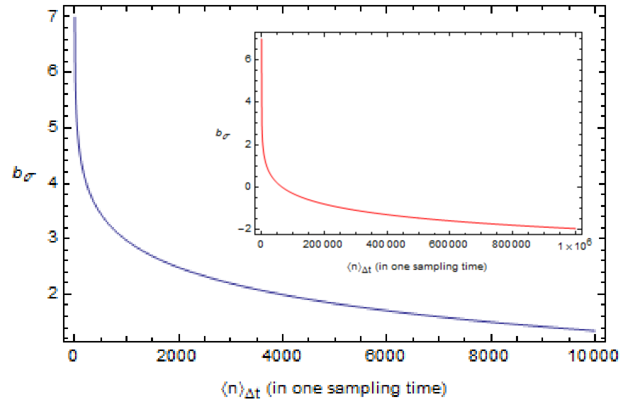

Figure 2. b_σ versus $\langle n\rangle_{\Delta t}$, for $b = 8$ ($2^b = 256$ levels), in the mesoscopic range. Inset for higher intensities. $\lambda = 1.55\mu$m, $\tau_c = 0.3\mu$s, or $\tau_c/\Delta t = 300$. It is seen that for higher intensities, the AtoD recorder saturates ($b_\sigma < 1$).

a high number of bits has an increased resolution ($1/2^n$ of the full signal range) but usually has a decreased speed. Small bit-number ADC (e.g., 8 bit) are usually faster and preferable, say, for telecommunication uses.

4) Use of a single detector reduce costs and eliminates the need for intensity balance in homodyne setups. Homodyne detection, while presenting a simple way to eliminate the average intense signal and extracting the desired noise, demands a constant precise balance, both optically as well as electronically, of the two detectors. The single detector use, when optimized for extraction of signals in the desired intensity range, offers a lower cost system without compromising speed.

A. Bits for average signal and noise

The digital recording of a laser light intensity signal as a function of time should reveal both the average signal level (or $\langle n\rangle$) and the fluctuations ($\pm\sigma$) around the average. For high intensity, the σ contribution for the signal gets smaller than the ADC's resolution and only the average signal is detected. Working in such conditions would rule out the possibility to have a record of σ. An ADC's with b bits has to accommodate both average and signals around average: $\langle n\rangle_{\Delta t} \pm \sigma$. Moreover, $\pm\sigma$ should be detectable by the ADC. Assuming that the laser intensity $I \propto \langle n\rangle_{\Delta t}/\Delta t$ and that the optical detector operates in a linear regime, it is expected that the relative proportion holds:

$$\frac{\langle n\rangle_{\Delta t}+\sigma}{\sigma} = \frac{2^b}{2^{b_\sigma}} \rightarrow b_\sigma = b + \log_2\frac{\sigma}{\langle n\rangle_{\Delta t}+\sigma}, \qquad (3)$$

where b_σ is the number of bits "reserved" for σ (within the set of bits b).

Fig. 2 illustrates the dependence of b_σ with $\langle n\rangle_{\Delta t}$ for an analog-to-digital recorder of $b = 8$. It is seen that for a few thousands of photons a few bits b_σ are available to record the fluctuation σ. However, as the number of photons increase, the inset shows that the analog-to-digital recorder saturates and no bits b_σ are available to record the fluctuation σ.

The operation regime for the laser should be "shot-noise" limited, where the Poissonian statistics of light predominates well above thermal radiation residues and electronic noises. The establishment of conditions to guarantee shot-noise operation is not trivial. Section "Methods" sketches general guidance lines to achieve shot-noise limited signals.

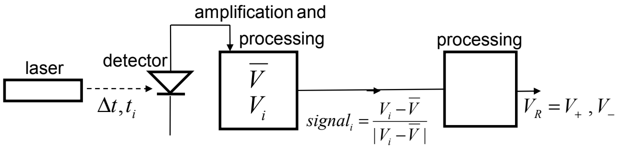

Figure 3. A shot-noise limited laser illuminates a fast light detector. After amplification, the signals V_i are recorded within a time interval Δt around time t_i. A processing circuit average the signals and classify each of them above or below the average $\overline{V}$. Random signals above or below the average value are converted in constant voltage amplitudes V_+ or V_- representing the random bits.

IV. PhRG's block diagram and commercial components

The PhRG's block diagram is sketched on Fig. 3. A laser with sufficient intensity to produce shot-noise limited light – where the light noise predominates over all electronic noises– illuminates a fast light detector. The processing units may contain an ADC for fast processing. Just to consider some concrete examples, using commercially available components, a fiber-optic connected continuous wave (CW) laser operating in single mode can be used, at $\lambda = 1.550\mu$m, bandwidth $\Delta\nu < 477$ kHz, and with a controlled temperature of 25^0C; its coherence time is $t_C \simeq 0.3\mu$s. An InGaAs PIN detector could be used, with a bandwidth $\Delta\nu \simeq$ 2GHz, with a transimpedance amplifier operating under a battery power source supplying 5V and a photovoltage bias of 10V; the detector responsivity is 0.8A/W. The amplified signals can be recorded by a 1GHz analog-to-digital (AtoD) circuitry with 1Gb of memory and 8 bit resolution. The signals are to be acquired within time windows $\Delta t \simeq 10^{-9}$s, much shorter than the laser coherence time t_C and, therefore, representing the true statistics of the light fluctuations, as given by Eq. (1). A signal processor average the signals in the AtoD and classify the recorded data as being above or below their average with signal$_i$ = $(V_i - \overline{V}/|V_i - \overline{V}|)$. Further processing converts the signal$_i$ sequences into constant amplitude signals V_+ or V_- that represent the sequence of random bits. Any formatting can be applied to this output stream.

A. Signal simulations

The described PhRG is part of an effort [8] to develop new cryptographic schemes. As it is not yet ready for operation, some computer simulations will be presented for pedagogical purposes. This way the reader can better understand the comments already made.

For the moment, the ADC operation will be ignored and analog signals will be treated for simplicity (for the ADC operation just think of discretized levels). Fig. 4 shows a sample of an analog CW signal (simulated as a signal taken at the amplification output stage). Comparisons, say, at each Δt (e.g., 2×10^{-9}s), between the instantaneous obtained value for the intensity (or an output voltage V) with respect the average intensity produce the bit values bit_i according to the rule

$$\text{signal}_i = \frac{I_i - \bar{I}}{|I_i - \bar{I}|}, \text{ and } \text{bit}_i = \frac{1 + \text{signal}_i}{2}. \tag{4}$$

For the fluctuations within the inset in Fig. 4, rule (4) gives the binary sequence 0,1,0,1,1,1,0,1,1,1,0,1,0,1, 0,1,0,0,1,0,1,1,1,0,1,0,1,1,1,0,1,1,0,0,1,0,0,0,1,1,1,0,0,0,1,0, 1,0,1,1.

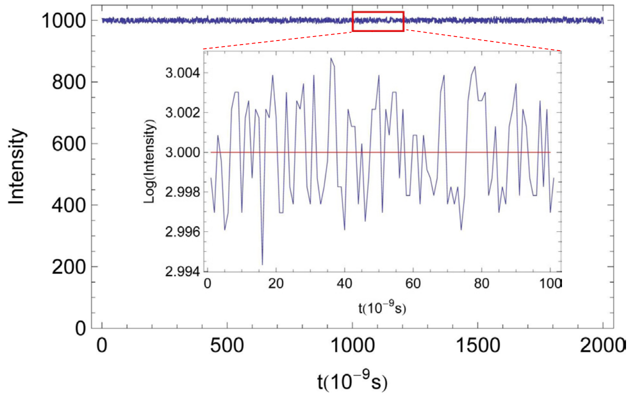

Figure 4. Simulation of a low laser intensity with Poissonian fluctuations as a function of time (the Intensity units are arbitrary). Inset: Detail of the laser intensity (log) within the small time window (in red). The red line indicates the mean intensity value.

A PhRG should operate continuously generating bits in a very high rate. The single laser and single detector scheme are overly superior with respect to stability than homodyne systems or comparison systems where two detectors are used.

V. Random number sequences: generalities

Given a finite sequence of supposedly random numbers, say the bit sequence above, one may ask if this sequence is truly random or not.

Quite generally, a physical source of entropy or numbers sequences n can be characterized by a probability distribution $p(n)$. Equivalently, $p(n)$ can be characterized by its moments of all orders (e.g., $\langle n \rangle, \langle (n - \langle n \rangle)^2 \rangle, \ldots$). However, being finite a sequence could not reveal all moments of the statistical source. Similarly, the occurrence of 0s and 1s, and distinct groups of 0 and 1 have to occur randomly. Observation of a finite sequence of 0s and 1s may reveal group patterns in the sequence. Patterns can be generated deterministically. That means that the given finite sequence could be compressed. Differently, a true random source should generate a sequence that, as its length increases, the associated entropy will increase linearly producing a sequence that could not be compressed.

The idea of randomness is perhaps better appreciated through the concept of "complexity of a string" [9]: The complexity of a string s is the length of the string's shortest description in some fixed universal description language. In other words, the complexity of a string is defined by the length of the *program* that describe that string. For example, a sequence of 10^6 consecutive "1"s, followed by another sequence of 10^6 consecutive "0"s produce a sequence of 2×10^6 bits. However, a short program such as "From $i = 1$ to 10^6, Print 1. From $i = 10^6 + 1$ to 10^6, Print 0 " produces the same sequence – with a short program. In other words, the sequence can be highly compressed and, therefore, is not random.

A description of s of minimal length, $d(s)$, uses the fewest number of characters and it is called a *minimal* description of s. The length of $d(s)$, i.e. the number of characters in the description, is the Kolmogorov complexity of s, written $K(s) = |d(s)|$. Unfortunately, K is not a computable function [10].

Nevertheless, it is clear that, under this definition, a perfect random sequence will need a program at least as long as the string itself to define the string, that is to say, the sequence cannot be compressed. Although these definitions may clarify the difficulties involved, they do not help much in the practical sense.

A. Statistical tests

The best one can do evaluate randomness is to apply a variety of tests. A satisfactory sequence that passes a given test will be said "random for that particular test".

There are known statistical test suites developed for this purpose. An example is the "A Statistical Test Suite for Random and Pseudo-random Number Generators for Cryptographic Applications", described in NIST's Special Publication 800 - 22/Revision 1:
http://csrc.nist.gov/groups/ST/toolkit/rng/
documentation-software.html .

Another one is the "DieHard" battery of tests:
http://www.stat.fsu.edu/pub/diehard/ .

VI. From signal detection to signal-to-noise ratio: METHODS

This section assumes that the readers have some familiarity with basic concepts of quantum mechanics and that some of the cited references are to be consulted. It also supposes familiarity with basic thermodynamics. Those not interested in these formalisms should skip the derivations. However, the resulting equations can be used; they are the end-products of the section.

The laser source characteristics adequate as an optical noise source for a PhRG were briefly discussed. Fast detectors are another essential part of the device. Detectors are also sources of noise (purely shot and thermal noises) and the understanding of the mixtures of light noise and other noises arising from detectors have to be understood to allow one to control or balance these sources and to make possible extraction of uncorrelated bits.

Above all, detectors are an important part of our tool set to understand the Universe. As a short comment, at the instant of their actions, detectors define our interface between the past and the future, in the classical view of time as a constantly moving arrow. What they record (=past) can be checked against our predictions (=future) of this same event and contribute to our primary sketch of the Universe. These logs are classical, in the sense that they can be faithfully copied. Interpretation of these records one-by-one or, in a correlated form, gives support or not to our immediate expectations or even to more broad concepts as our views of a classical or quantum world. Therefore, our understanding of the detector's construction not only show

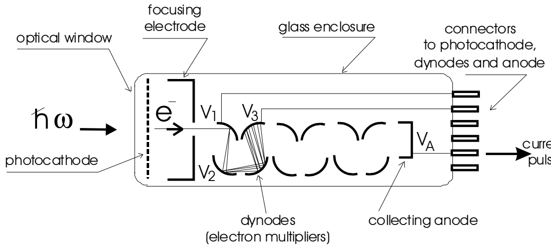

Figure 5. *PMT elements and gain mechanism* – A photon may transfer energy to a photoelectron in the photocathode. This electron hits the first dynode after acceleration by a voltage difference, ejecting electrons from the material. Successive accelerations and collisions result in a charge pulse at the collecting anode.

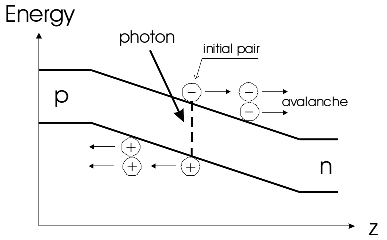

Figure 6. *Gain mechanism for an APD* – A photon creates an electron-hole pair in a semiconductor *p-n* junction. A strong field is maintained at the junction such that these charges are accelerated and, whenever they gain energy ($\geq E_g$, the gap energy) a secondary pair will be created at the collisions. Each new pair may contribute to create an "avalanche" of charges at the collecting outputs. The physical structure of an APD may contain a volume where light absorption creates electron-hole pairs and an electric field separates the two kinds of charges and sweeps one of the carriers towards a multiplication region where a strong electric field accelerates this charge causing impact ionizations as a gain mechanism.

their limitations but also allow us to improve them, widening our perceptions about the Universe.

This section discusses optical detectors and, in particular, detectors that operate by photo-electron absorption processes. The optically sensitive materials used in their construction limit their wavelength bandwidth and their amplification electronics usually impose their frequency bandwidths. Noise sources will also be discussed as well as mathematical tools to understand their inner-workings. Only *direct* detection, used for the presented PhRG, is treated here; homodyne and heterodyne detection will not be discussed.

Among commercial detectors are the single photon sensitivity detectors, known as photon-counting detectors (SPCDM, or single photon counting detection modules), *photon multipliers* (PMT), and integrated semiconductor detectors known as *avalanche photo-diodes* (APD). Silicon PIN detectors are also used for low-light detection but their sensitivity is much lower than APD's and they are not adequate for single photon counting. However, an APD can be built with a faster electronics and have a wide use in telecommunication with a corresponding lower cost. Less common are the cryogenic detectors operating by photon absorption and using photon-to-thermal energy conversion. They can have very high sensitivities and resolution but they are slower. For a review on single-photon detectors see [11]. This note is not concerned with single photon detectors but with detectors for telecommunication that, apart of being low-cost compared with single photon detectors, could be able to detect mesoscopic number of photons. The expected photoelectron signals at the detector output are well represented by analog signals.

Usually, silicon APDs are optimized to work between 300 and 1100 nm, germanium between 800 and 1600 and InGaAs from 900 to 1700nm. Only silicon APDs present dark current low enough for commercial use in *non-gated* single photon detectors. Non-silicon APDs can also be used in gated operation for single-photon counting; the gate operation avoids the excessive thermal noise that builds up if they were used in a continuous operation and help diminishing after-pulses when operated to be sensitive to weak avalanches [12], [13].

The detectors based on the photoelectric effect can be classified as *annihilation* detectors. Figs. 5 and 6 sketch the gain mechanisms in these devices.

A. Symmetric and asymmetric devices

Annihilation detectors differ from field detectors (antennas) in fundamental aspects. One of them is that annihilation detectors are not sensitive to the field polarization, whereas an antenna is. Another difference resides in the action of the photon annihilation and creation operators $\widehat{a}$ and $\widehat{a}^+$ (These are the fundamental quantum operators that annihilate and create photons). In the classical limit of optics $\hbar\omega \to 0$ and, therefore, a detector in contact with the heat bath at temperature T_K may, with high probability, gain sufficient energy to emit a photon with energy $\hbar\omega \ll k_B T_K$ (k_B is the Boltzmann's constant). In other words, $\widehat{a}$ and $\widehat{a}^+$ will have similar contributions in the interaction process. In the microwave range, $\hbar\omega \sim k_B T_K$. In the optical range $k_B T_K \ll \hbar\omega < 2m_0c^2$ (the upper limit taken as the energy of an electron-positron pair creation); therefore, the heat bath will have a small probability to create a photon. One may also say that in the optical range annihilation processes dominate over creation ones—an asymmetry between these processes. Single-photon sensitivity detectors are usually submitted to cooling processes to further reduce the probability of *dark* noise or electron emission in the detector when no desired light is present. In the optical range, the interaction between detector and field then proceeds mainly through the electric field operator $\widehat{E}^{(+)}$ (that involves only annihilation field operators).

B. Quantum Efficiency

The intrinsic bandwidth $\Delta\omega$ of a detector is basically determined by the material employed to make the photocathode and other elements such as the optical material of the collecting window. The electronic circuitry after the detecting elements will also contribute to the effective bandwidth of the detecting system. In general, one is interested in a detection bandwidth $\delta\omega$ quite narrow compared with the optical frequency ω_0 of the incident photons ($\delta\omega \ll \omega_0$.)

The photoelectric material is made with a very low value of the *work function* (the energy necessary to extract an electron from the material), which defines the lower limit for the detectable optical frequency. The material employed in the optical window of a cooled detector frequently defines the maximum optical frequency detectable.

The emission probability associated with the photoelectric effect has a conversion or *quantum efficiency* η_{pe} smaller than unity ($\eta_{pe} < 1$). The *quantum efficiency* is measured in the averaging process of converting photons to photoelectrons (electrons emitted in the photoelectric effect), and is given by

$$\langle n_{pe} \rangle = \eta_{pe} \langle n \rangle \, , \tag{5}$$

where $\langle n_{pe} \rangle$ is the average number of photoelectrons emitted and $\langle n \rangle$ is the average number of incident photons reaching the detector area A. This implies that there is *no* assurance that a photoelectron will appear after a given photon hits the photoelectric material – (5) only express *average* quantities.

This parameter η_{pe} has a stochastic origin related to atomic processes. One may interpret this uncertainty in the emission time of an excited atom as due to stochastic fluctuations of the *vacuum* electric field (existing field in the absence of photons). $\eta_{pe} < 1$ also impose limitations on experiments involving photon pair detection with two detectors (*coincidence* detection), because each undetected photon results in a loss of coincidence between the detectors.

A photon detector is usually made to multiply the initial charge ejected from the photocathode to result in a charge pulse easily treated by conventional electronics. This amplification process is known as the detector *gain G*. After this gain process an electric current $i(t) = Ge(dn_{pe}/dt)$ appears at the anode. The ratio σ between the photoelectric current density and the incident photon intensity is

$$\sigma = \frac{e \frac{dn_{pe}}{dt}}{\hbar\omega \frac{dn}{dt}} = \frac{e}{\hbar\omega} \frac{dn_{pe}}{dn} = \frac{e}{\hbar\omega} \eta_{pe} \, . \tag{6}$$

The ratio σ, known as “radiant sensitivity” of the photocathode, is usually furnished by the detector’s maker. In a photomultiplier, it can be measured extracting the charge pulse directly from the first dynode (See Fig. 5), thus avoiding the gain mechanism.

C. Temporal Response; Amplification and Discrimination; Formatting

The gain involve processes occurring in a time interval τ_d. In APDs, within this time, a newly arriving photon cannot produce a distinct amplification pulse and thus produces no count in the external electronics. τ_d defines the detector’s dead time. Some new detectors, including cryogenic ones, aim to identify the arrival of two or more photoelectrons within τ_d.

For some applications, such as coincidence counts between two detectors, in order to shorten the time resolution below τ_d, electronic techniques utilize the rate of increase (or decrease) of charge variation (time derivatives during a pulse formation) from one detector to trigger a time counter. A time is measured when the second detector gives a signal. This way, current commercial detectors may present a time resolution of $\sim 10^{-10}$s between two events (shorter than the dead time τ_d).

Some noise sources contribute to degradation of the photodetection, among them thermionic emission (proportional to temperature and dependent on specific materials) and cosmic rays. Photodetector engineering tries to optimizes their signal to noise ratio. Cosmic rays can be eliminated in coincidence detection, due to the negligible probability of both detectors being excited simultaneously. They cannot be eliminated in a single detector but are minimized by a small detection area.

In a general way, an electronic circuit *amplifies* the signal appearing after the gain process and chops some of them in a *discrimination* process to reduce events that come from thermal noise. Electrons emitted due to thermal emission produce charge pulses of less intensity in the gain stage, because they usually have much less initial kinetic energy than those produced in the photoelectric effect. This fact can be used to set a discrimination level to result in one charge pulse for each photoelectron emitted. *Formatting* electronics are now usually built into many detecting systems giving approximately a standard digital output (TTL, ECL, etc.) for each analog charge pulse generated.

D. The Quantum Process of Photodetection (basics)

The theory of photo-detection is an area of study by itself [14]. Some outstanding landmarks were established by Glauber with his work on optical coherence and by Mandel on the theory of photon statistics [15]. The reader is strongly suggested to consult these references. In this section, a simplified approach to the generation of photoelectrons from single photon streams, using a phenomenological response function [16], is utilized to introduce some readers in this subject.

An electric field quantum operator $\widehat{E}(z,t) = \widehat{E}^{+}(z,t) + \widehat{E}^{-}(z,t)$, where $\widehat{E}^{-}(z,t) = (\widehat{E}^{+}(z,t))^{+}$, describes light propagation along the z-axis in an isotropic medium with dielectric susceptibility ϵ, where $\omega = kv = kc/n$. For example, one could write, for a x-polarized field

$$\widehat{\mathbf{E}}^{-}(z,t) = \hat{\mathbf{x}} \sum_{\omega} \sqrt{\frac{\hbar\omega}{2\epsilon V}} \, \widehat{a}_{\omega}^{+} \exp\left[-i\omega\left(z/v - t\right)\right] \, , \tag{7}$$

where V is the quantization volume. In the classical limit, quantum operators $\widehat{a}_{\omega}$ and $\widehat{a}_{\omega}^{+}$ become the field amplitudes a and a^{*}.

If one considers ω in a continuum, it may be convenient to write the Hamiltonian $\widehat{H}$ for a free mode and the number operator $\widehat{N}$ as

$$\widehat{H} = \int_0^{\infty} \hbar\omega \, \widehat{a}^{+}(\omega)\widehat{a}(\omega) \, d\omega \, , \; \widehat{N} = \int_0^{\infty} \widehat{a}^{+}(\omega)\widehat{a}(\omega) \, d\omega, \tag{8}$$

where $[\widehat{a}(\omega), \widehat{a}^{+}(\omega')]_{-} = \delta(\omega - \omega')$. The practical transition from a discrete to a continuum is made by substituting $\sum_{k_z} \rightarrow (L_z/2\pi) \int dk_z$, where L_z is the quantization length of the field and writing the quantization volume V as the product of the mode area A_c times the length $L_z = v\delta t$ ($v = c/n$), where $\delta t = 1/\delta\nu$ is the separation time interval between modes. Also $\widehat{a}(\omega) \rightarrow \widehat{a}_{\omega}/\sqrt{\delta\omega}$, giving

$$\widehat{E}^{-}(z,t) = \frac{i}{\sqrt{2\pi}} \int_0^{\infty} d\omega \sqrt{\frac{\hbar\omega}{2\epsilon A_c v}} \widehat{a}^{+}(\omega) \exp\left[-i\omega\left(z/v - t\right)\right]. \tag{9}$$

As discussed previously, one is usually interested in a narrow frequency range around the average field frequency ω_0. Writing $\omega = \omega_0 + \omega'$, this condition is $\omega'/\omega_0 \ll 1$ and in this case,

$$\widehat{E}^{-}(z,t) \simeq i \frac{1}{\sqrt{2\pi}} \sqrt{\frac{\hbar\omega}{2\epsilon A_c v}} \exp\left[i\omega_0\left(z/v - t\right)\right]$$

$$\times\int_{-\infty}^{\infty} d\omega'\,\widehat{a}^{+}(\omega')\exp\left[-i\omega'(z/v-t)\right] \ . \qquad (10)$$

The field intensity operator, $\widehat{I}(t)=\widehat{E}^{-}(t)\widehat{E}^{+}(t)$, is connected to the photoelectric current operator $\widehat{I}_e$ through a response function $D(t-t')$ of the photodetector by

$$\widehat{I}_e(z,t)\equiv\frac{\widehat{N}_e(t)}{dt}=e\int_{-\infty}^{\infty} dt'\,D(t-t')\widehat{E}^{-}(z,t')\widehat{E}^{+}(z,t') \ . \qquad (11)$$

The phenomenological function $D(t-t')$ defines the causal process generating the electric current in time t. If one considers the time response of the photodetector to be much shorter than the frequency bandwidth considered, the response time can be approximated by $D(t-t')\simeq D\delta(t-t')$ and, therefore, substituting the above definitions in Eq. (11) results in

$$\begin{aligned}\widehat{I}_e(z,t)=eD\frac{\hbar\omega_0}{4\pi vA}&\int_{-\infty}^{\infty}d\omega'\int_{-\infty}^{\infty}d\omega''\,\widehat{a}^{+}(\omega')\,\widehat{a}(\omega'')\\ &\times\exp\left[-i(\omega'-\omega'')(z/v-t)\right] \ . \end{aligned}\qquad (12)$$

The operator number for the photoelectrons and the current operator for these same electrons, in this "instantaneous" response approximation, are related by

$$\begin{aligned}\widehat{N}_e(z)&\equiv\int_{-\infty}^{\infty}\widehat{I}_e(z,t)dt\\ &=eD\frac{\hbar\omega_0}{4\pi\epsilon vA}\int_{-\infty}^{\infty}d\omega'\int_{-\infty}^{\infty}d\omega''\widehat{a}^{+}(\omega')\,\widehat{a}(\omega')\\ &\quad\times\exp\left[-i(\omega'-\omega'')z/v\right]\\ &\quad\times\int_{-\infty}^{\infty}dt\exp\left[-i(\omega'-\omega'')t\right]\\ &=eD\frac{\hbar\omega_0}{2\epsilon vA}\int_{-\infty}^{\infty}d\omega'\widehat{a}^{+}(\omega')\widehat{a}(\omega')=eD\frac{\hbar\omega_0}{2\epsilon vA}\widehat{N}. \end{aligned}\qquad (13)$$

An average can be taken over the number operators for the photoelectrons and for the photons, giving

$$D\langle\widehat{N}\rangle=\frac{2\epsilon vA}{e\hbar\omega_0}\langle\widehat{N}_e\rangle \ . \qquad (14)$$

Using the definition of the detector efficiency,

$$D=\frac{2\epsilon vA}{e\hbar\omega_0}\frac{\langle\widehat{N}_e\rangle}{\langle\widehat{N}\rangle}\equiv\frac{2\epsilon vA}{e\hbar\omega_0}\eta_{pe} \ . \qquad (15)$$

The photoelectron current operator can now be written

$$\begin{aligned}\widehat{I}_e(z,t)=e\frac{\eta_{pe}}{2\pi}&\int_{-\infty}^{\infty}d\omega'\int_{-\infty}^{\infty}d\omega''\widehat{a}^{+}(\omega')\,\widehat{a}(\omega'')\\ \times\exp&\left[-i(\omega'-\omega'')(z/v-t)\right] \ , \end{aligned}\qquad (16)$$

and from the definition of a Fourier transform $f(t)=(1/\sqrt{2\pi})\int_{-\infty}^{\infty}d\omega f(\omega)\exp i\omega t$, one arrives at the *instantaneous* photoelectron current operator

$$\widehat{I}_e(z,t)=\frac{d\widehat{N}_e}{dt}=e\eta_{pe}\widehat{a}^{+}(t-z/v)\widehat{a}(t-z/v) \ . \qquad (17)$$

From now on z will be taken as $z=0$. The number operator $\widehat{a}^{+}(t)\widehat{a}(t)$ is the photon number *intensity* at time t (Units of $\widehat{a}(t)$ are $t^{-1/2}$; see Eqs. (8)). Taking averages of the operators one arrives at

$$\langle d\widehat{N}_e\rangle=e\eta_{pe}\langle\widehat{a}^{+}(t)\widehat{a}(t)\rangle dt\equiv R_1(t)dt=dP_1(t) \ , \qquad (18)$$

which defines the differential photodetection probability $dP_1=R_1(t)dt$ for *one* photoelectron in t within dt. The rate of photodetection—for single events—is then

$$\frac{dP_1}{dt}=R_1(t)=e\eta_{pe}\langle\widehat{a}^{+}(t)\widehat{a}(t)\rangle \ . \qquad (19)$$

Consider a photon state $|\psi\rangle$ in the number representation, describing single photons at instant times $t_1,t_2,\dots t_n$: $|\psi\rangle=|1_{t_1},1_{t_2},\dots 1_{t_n}\rangle$. Applying the photoelectron current operator $\widehat{I}_e$ to $|\psi\rangle$, one has

$$\begin{aligned}\widehat{I}_e|\psi\rangle&=e\eta_{pe}\widehat{a}^{+}(t)\widehat{a}(t)|1_{t_1},1_{t_2},\dots 1_{t_n}\rangle\\ &=e\eta_{pe}\widehat{a}^{+}(t)\widehat{a}(t)\left[\widehat{a}^{+}(t_1)\widehat{a}^{+}(t_2)\dots\widehat{a}^{+}(t_n)|0\rangle\right]. \end{aligned}\qquad (20)$$

Successive applications of $[\widehat{a}(t),\widehat{a}^{+}(t_j)]=\delta(t-t_j)$ to the products of operators in this equation gives

$$\begin{aligned}\widehat{I}_e|\psi\rangle&=e\eta_{pe}\widehat{a}^{+}(t)\left[\delta(t-t_1)\widehat{a}^{+}(t_2)\dots\widehat{a}^{+}(t_n)\right.\\ &\qquad\left.+\widehat{a}^{+}(t_1)\widehat{a}(t)\widehat{a}^{+}(t_2)\dots\widehat{a}^{+}(t_n)\right]|0\rangle=\dots\\ &=e\eta_{pe}\left(\sum_{i=1}^{n}\delta(t-t_i)\right)|\psi\rangle, \end{aligned}\qquad (21)$$

showing that the eigenvalue of the photoelectron current operator is a succession of sharp charge pulses at instants t_i (this result is qualitatively intuitive). Of course, different models for the response function D, instead of $D(t-t')=D\delta(t-t')$, give different distributions for the resulting current. With these, the delta pulses in Eq. (21) will be modified to pulses with less sharp shapes. In fact, good descriptions of practical current pulses can be achieved with simple models for D. For example, functions that depend only on a small number of "moments" (or Fourier components) are particularly useful; such as

$$D(t-t')\simeq\frac{\mu}{\sqrt{\pi}}e^{-\mu^2(t-t')^2} \qquad (22)$$

where μ is adjusted to fit charge pulses that usually depend on the particular detection system in use.

However, Eq. (21) gives a good pictorial view of the "shot-noise" process.

E. Photon Detection Probability and Field Distributions

A useful connection between photon number and field distributions can be derived using the coherent basis representation $|\alpha\rangle$ defined [14] as (k_s designate modes) $\widehat{a}_{k_s}|\alpha_{k_s}\rangle=\alpha_{k_s}|\alpha_{k_s}\rangle$. The coherent state $|\alpha\rangle$ introduced by Glauber is

$$|\alpha\rangle=e^{-|\alpha|^2/2}\sum_{n=0}^{\infty}\frac{\alpha^n}{\sqrt{n!}}|n\rangle \ . \qquad (23)$$

In the coherent basis the average photon number will be $\langle\alpha_{k_s}|\widehat{a}^{\dagger}_{k_s}\widehat{a}_{k_s}|\alpha_{k_s}\rangle = |\alpha_{k_s}|^2$. A field distribution in the diagonal representation [14] is written $P(\{\alpha_{k_s}\})$. $P(\{\alpha_{k_s}\})$ is associated to each specific field (laser, thermal etc). A density operator ρ for this field can be written

$$\rho = \int P(\{\alpha_{k_s}\})|\{\alpha_{k_s}\}\rangle\langle\{\alpha_{k_s}\}|d^2\{\alpha_{k_s}\} \ , \tag{24}$$

and normalized with $Tr[\rho] = \int P(\{\alpha_{k_s}\})d^2\{\alpha_{k_s}\} = 1$. A photon number probability distribution can be defined as

$$p(\{n_{k_s}\}) = Tr\left[\rho|\{n_{k_s}\}\rangle\langle\{n_{k_s}\}|\right] \ . \tag{25}$$

Working with these equations and summing the total number of photons counted $n = \sum_{\{n_{k_s}\}} n_{k_s}$, the Mandel's relationship connecting $p(n)$ and $P(\{\alpha_{k_s}\})$ is obtained:

$$p(n) = \int P(\{\alpha_{k_s}\})\frac{\left(\sum_{k_s}|\alpha_{k_s}|^2\right)^n}{n!}e^{-\sum_{k_s}|\alpha_{k_s}|^2}d^2\{\alpha_{k_s}\} \tag{26}$$

and for a single mode case $\{\alpha_{k_s}\} \to \alpha$

$$p(n) = \int P(\alpha)\frac{|\alpha|^{2n}}{n!}e^{-|\alpha|^2}d^2\alpha \ . \tag{27}$$

Several probability distributions can be calculated as, for example,

1. *Laser with amplitude and phase constant or uniform phase*

$$p(n) = e^{-\langle n\rangle}\frac{\langle n\rangle^n}{n!} \ , \langle n\rangle = |\alpha|^2 \ . \tag{28}$$

2. *Thermal field (random phases)* (See Section VIII, Eq. 8.8. in Ref. [14])

$$p(n) = \frac{1}{1+\langle n\rangle}\left(\frac{\langle n\rangle}{1+\langle n\rangle}\right)^n \ . \tag{29}$$

3. *Superposition of laser and thermal light*

$$p(m) = \frac{\langle n_T\rangle^m}{(1+\langle n_T\rangle)^{m+1}} \times L_m\left[-\frac{\langle n_L\rangle}{\langle n_T\rangle(1+\langle n_T\rangle)}\right]e^{-\frac{\langle n_L\rangle}{1+\langle n_T\rangle}} \ , \tag{30}$$

where L_m is the Laguerre function, $\langle n_L\rangle$ and $\langle n_T\rangle$ are the average numbers for the laser photons and thermal photons, respectively.

F. Noise considerations

The detection process has some fundamental random contributions: 1) The photon absorption is statistical in nature. 2) The immersion of detector and associated electronics in the environment at temperature T_K produces electronic thermal excitations or thermal noise that are also recorded [4]. 3) The electronic avalanche in the gain process is statistical.

Data recording at high speeds, such as done by AtoD converters, introduce their particular error sources. Traditional technical sources of error include nonlinearities in the conversion processes, electro-magnetic interferences, gain error produced by amplifier distortions, offset error, and AtoD conversion errors. Some of these may become a significant source of error. Clock jitter, for example, introduces uncertainty in the collection time of the signal. Fast AtoD converters may present cross-talk between the analog and digital components. Many of these error sources can be technically reduced.

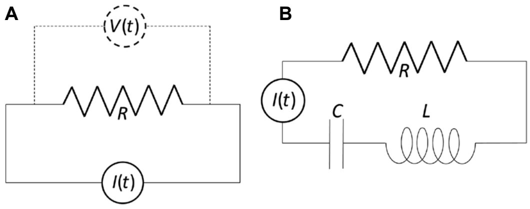

Figure 7. A: A real resistance R at temperature T_K can be represented as a lossless resistance R where a current $I(t)$ appears due to the thermally excited electrons. A voltage $V(t)$ appears at the resistance ends. B: RLC components in series.

A parameter that gives a good estimate of the signal that could be obtained under presence of fundamental noises is the *signal-to-noise* ratio *SNR*. This ratio could be defined, in the *number basis*, as the ratio of the average signal square to the variance

$$SNR = \frac{\langle n\rangle^2}{\langle(n-\langle n\rangle)^2\rangle} \ , \tag{31}$$

where $\langle n\rangle$ is the average detected number of photons. SNR parameters can be written for any quantity of interest such as voltages, currents, phase and so on. As a simple warning, another common use is writing SNR as the square root of (31).

G. Fluctuation spectra

Detectors are usually connected to an impedance that could be, in the simplest case, a resistor or the effective resistance of a pre-amplifier. Understanding the effects of the thermal noise in this resistor by itself is important to derive the effective noise of a detector coupled to an external circuit. Even neglecting microscopic aspects describing the behavior of electrons in the resistor, thermodynamic arguments and macroscopic reasonings are of great help to understand this noise source.

A resistor R coupled to an ideal amplifier tuned at a frequency $\overline{\omega}$ with a bandwidth $\Delta\omega$ will produce a fluctuating signal in the amplifier. This signal can be traced by the current $I(t)$ generated by electrons set in motion by thermal energy. A corresponding fluctuating voltage emf $V(t) = RI(t)$ will be detected across the resistor. An equivalent circuit is shown in Fig. 7-A. A more general circuit to represent a real resistance connected to LC components in series is in Fig. 7-B. All components are assumed to be at thermal equilibrium at temperature T_K under ergodicity conditions. A voltage $V_j(t)$ will appear at each j-component ends. The equipartition theorem for the energy establishes that for each degree of freedom the average energy is $k_BT_K/2$. Therefore,

$$\langle\frac{1}{2}LI(t)^2\rangle = \frac{1}{2}k_BT_K \text{ and } \langle\frac{1}{2}CV_C^2\rangle = \frac{1}{2}k_BT_K \ , \tag{32}$$

where V_C is the potential difference across the capacitor. Thus, for example

$$\langle I(t)^2\rangle = \frac{k_B}{L}T_K \tag{33}$$

Parseval's theorem gives $\int_{-\infty}^{\infty} I(t)^2 dt = \int_{-\infty}^{\infty} |\widetilde{I}(\nu)|^2 d\nu$. The average $\langle I(t)^2\rangle$ can be expressed by

$$\begin{aligned}\langle I(t)^2\rangle &= \lim_{\tau\to\infty}\frac{1}{\tau}\int_{-\tau}^{\tau} I(t)^2 dt = \lim_{\tau\to\infty}\frac{1}{\tau}\int_{-\tau}^{\tau}|\widetilde{I}(\nu)|^2 d\nu \\ &\equiv \int_{-\infty}^{\infty} S_I(\nu)d\nu\,,\end{aligned} \tag{34}$$

where $S_I(\nu)$ is the spectral density of the photo-current $I(t)$. Therefore, for the current $I_n(t)$ caused by the thermal noise

$$\int_{-\infty}^{\infty} S_{I_n}(\nu)d\nu = \frac{k_B}{L}T_K\,. \tag{35}$$

Common responses from light detectors are voltage outputs. One may want to write the SNR ratio as a function of the voltage

$$SNR = \frac{\langle V\rangle^2}{\langle (V-\langle V\rangle)^2\rangle} = \frac{\langle V\rangle^2}{\langle V^2\rangle - \langle V\rangle^2} = \frac{\langle V\rangle^2}{\langle (\Delta V)^2\rangle}\,. \tag{36}$$

One then need to obtain the average and fluctuation of V to calculate Eq. (36). For example, the current in the circuit shown in Fig. 7-B is given by

$$L\frac{d}{dt}I(t) + RI(t) + \frac{1}{C}\int_{-\infty}^{t} I(t')dt' = V\,. \tag{37}$$

Looking at $e^{i2\pi\nu t}$ $(= e^{i\omega t})$ fluctuations, one obtains the circuit impedance $Z(\omega) = V(\omega)/I(\omega) = R + i\left(\omega L - \frac{1}{\omega C}\right)$. For a circuit where the energy is mostly stored in the inductance field, one may neglect the stored charge given by $\int_{-\infty}^{t} I(t')dt' \to 0$, that would otherwise reside in the capacitor. This gives a LR circuit whose response extends to very high frequencies. This gives $Z = Z(\omega) = R + i\omega L$. From $Z(\omega)I(\omega) = V(\omega)$ one may infer the relationship between S_I and the corresponding voltage spectrum S_V:

$$S_V = |Z|^2 S_I = |R + i\omega L|^2 S_I\,. \tag{38}$$

Therefore, for the noise

$$\int_{-\infty}^{\infty} S_{I_n}(\nu)d\nu = \frac{k_B}{L}T_K = \int_{-\infty}^{\infty}\frac{S_{V_n}(\nu)}{|Z|^2}d\nu\,. \tag{39}$$

Considering that the frequency response associated with $S_{V_n}(\nu)$ is uniform up to very high frequencies, one may write $S_{V_n}(\nu) \to S_{V_n}(0)$. This gives

$$\begin{aligned}\int_{-\infty}^{\infty}\frac{S_{V_n}(\nu)}{|Z|^2}d\nu &\simeq S_{V_n}(0)\int_{-\infty}^{\infty}\frac{1}{R^2+(2\pi\nu L)^2}d\nu \\ &= \frac{S_{V_n}(0)}{2}\frac{1}{LR} = \frac{k_B}{L}T_K\,,\end{aligned} \tag{40}$$

and therefore

$$S_{V_n}(\nu) = S_{V_n}(0) = 2k_BT_KR\,, \tag{41}$$

$$S_{I_n}(\nu) = \frac{S_{V_n}(\nu)}{R^2} = \frac{2k_BT_K}{R}\,. \tag{42}$$

The detector output usually goes to a bandwidth limited pre-amplification stage, that will set the overall bandwidth limit in frequency $\Delta\nu_B$. Similarly to Eq. (34), the connection between the average $\langle V(t)^2\rangle$ and $S_{V_n}(\nu)$ is

$$\begin{aligned}\langle V(t)^2\rangle &= \int_{-\Delta\nu_B}^{\Delta\nu_B} S_V(\nu)d\nu = 2k_BT_KR\int_{-\Delta\nu_B}^{\Delta\nu_B}d\nu \\ &= 4k_BT_KR\Delta\nu_B\,.\end{aligned} \tag{43}$$

This treatment exemplifies the use of fluctuations and laws of energy equipartition to derive connections between frequency spectra and thermodynamic quantities. Similar treatment can be applied to distinct circuits. Eq. (43) was investigated by J. B. Johnson in [17].

For an electric current originated from laser excitation the instants t_i will be randomly (Poissonian) distributed, and for a given light power P the average value of the excited photo-electron current is (see Eq. 6)

$$\langle I_{pe}(t)\rangle = \sigma P = \eta_{pe}\frac{e}{\hbar\omega}P\,. \tag{44}$$

The instantaneous value $I_G(t)$ of the amplified $I_{pe}(t)$ by a circuitry with a time constant t_c and gain G is

$$I_G(t) = \int_0^{\infty}\frac{e^{-t'/t_c}}{t_c}GI_{pe}(t-t')dt' \tag{45}$$

The current $I_G(t)$ will show asymmetric amplified spikes instead of the point-like Dirac's deltas and with a decaying time given by t_c.

H. Signal to noise ratio

Let us consider that the output current $I(t)$ after an amplifier stage of gain G, and time constant t_c (e.g., $t_c = RC$), is constituted by the sum of the current contributions given by:
1) light of average frequency ω_0 with power $P(t)$ giving the average current $G\eta_{ep}e\langle P(t)\rangle/(\hbar\omega_0)$,
2) electronic thermal excitations (Johnson's noise) and,
3) dark current I_{dk} generated by crystallographic defects within the depletion region of the semiconductor being used as the photo-sensitive material (dark currents in PIN photodiodes could be of order ~ 100pA or less).

In avalanche photodetectors (APD), intermediate energy levels can also be populated by the electronic avalanche. These energies would decay shortly after causing "after pulses" that may modify substantially the photo-electron statistics. For each detecting system used, one should understand and consider the causes of deviations from the direct photo-current caused by the primary photo-excitation.

Collecting the above contributions,

$$\langle I_e(t)\rangle \simeq G\eta_{ep}\,e\frac{\langle P(t)\rangle}{\hbar\omega_0} + \langle I_{e,s}(t)\rangle + \langle I_{dk}\rangle\,. \tag{46}$$

The shot noise current will fluctuate around this mean value.

The distinction between Johnson's noise and the electric shot noise is not always clear. While some forms of shot noise occurs even at a temperature of $0\,K$ (e.g., if originated by light's incidence), Johnson's noise is caused by thermal excitations and

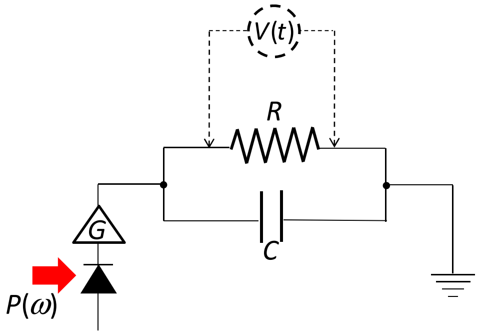

Figure 8. Voltage output. The detector current is amplified with gain G; the voltage is measured at the ends of the parallel RC circuit being probed.

do not exist at $0\,K$. Shot noise in matter is also called ballistic noise, and it is connected with processes where the mean free path of a particle (e.g., electron) is long compared with the atomic positions in the medium. Thermal equilibrium noise appear even with no net current present. Sometimes the physical process is such that a clear distinction between shot and Johnson noise cannot be made. However, for cases where this distinction can be made more apparent, a simplified derivation can be seen in [18].

Eq. (34) shows the connection between $\langle I^2\rangle$ and the current spectral density S_I. For current fluctuations due to the shot noise, this connection holds and for a narrow band S_I, one obtains the spectral density proportional to $\langle I\rangle B$:

$$S_I(\nu_0)B = \langle I^2\rangle = 2e\langle I\rangle B\,. \tag{47}$$

Actually, the total current density *fluctuation* spectrum is given by all contributions, and considering that a gain G also exists, combining Eqs. (47) and (46), one obtains

$$S_I(\nu) = G^2\eta_{ep}\,e^2\frac{P(\nu_0)}{\hbar\omega_0} + \frac{2k_BT_K}{R} + S_{I_{dk}}(\nu)\,. \tag{48}$$

Considering that samplings are taken at a specific frequency such that the t_c cutoff is very low compared to it, using Eq. (38) one may write the mean-square voltage *fluctuation*

$$\langle\Delta V(t)^2\rangle = \int_{-\infty}^{\infty} S_I(\nu)\,|Z(\nu)|^2\,d\nu \simeq S_I\int_{-\infty}^{\infty}|Z(\nu)|^2\,d\nu, \tag{49}$$

where S_I, from Eq. (48), includes the main contribution from the dark noise. In general, the variance of the filtered current *fluctuation* is

$$\langle\Delta I(t)^2\rangle = \int_{-\infty}^{\infty} F(\omega,\tau_c)S_I(\omega)\frac{d\omega}{2\pi}\,, \tag{50}$$

where $F(\omega,\tau_c)$ is the applied linear filter.

For a detector system ending in a parallel RC combination, where the voltage $V(t)$ is probed between the capacitor or resistor ends (see Fig. 8), the impedance Z to be inserted in Eq. (49) is given by

$$\frac{1}{Z} = \frac{1}{Z_R} + \frac{1}{Z_c} = \frac{1}{R} + \frac{1}{\frac{-i}{\omega C}}\,. \tag{51}$$

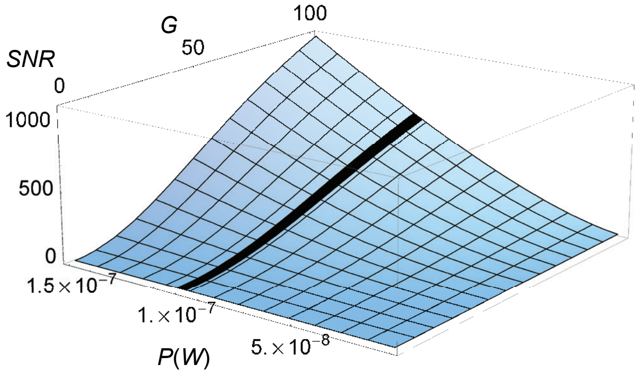

Figure 9. SNR as a function of the optical power P and the gain G. Used parameters are $T_K = 300K$, $\hbar = 1.055\times10^{-34}$Js, $k_B = 1.38\times10^{-23}$J/K, $e = 1.60\times10^{-19}$C, $\eta_{ep} = 0.8$, $R = 50\Omega, C = 20p$F, $\lambda = 1.55\mu$m, $P_{dk} = 1\times10^{-10}$W. The solid line indicates SNR values at $P = 100n$W and for variable gain.

Thus

$$\int_{-\infty}^{\infty}|Z(\nu)|^2d\nu = \frac{1}{2}\frac{R}{C} \rightarrow \langle\Delta V(t)^2\rangle \simeq S_I(\nu_0)\frac{1}{2}\frac{R}{C}\,, \tag{52}$$

and therefore

$$\langle\Delta V(t)^2\rangle = \frac{1}{2}\frac{R}{C}\left[G^2\eta_{ep}\,e^2\frac{P(\nu_0)}{\hbar\omega_0} + \frac{2k_BT_K}{R} + S_{I_{dk}}(\nu_0)\right] \tag{53}$$

$$\langle V(t)\rangle = G\eta_{ep}\,e\frac{\langle P(\nu_0)\rangle}{\hbar\omega_0}R\,. \tag{54}$$

$S_{I_{dk}}$ can be written using an equivalent power P_{dk} for the dark noise

$$S_{I_{dk}} = eG\left(\eta_{ep}\,e\frac{P_{dk}}{\hbar\omega}\right)\,. \tag{55}$$

Using the obtained relationships, the SNR ratio with respect to voltage measurements is

$$SNR = \frac{\langle V(t)\rangle^2}{\langle(V(t)-\langle V(t)\rangle)^2\rangle} = \\ \frac{\langle V(t)\rangle^2}{\langle\Delta V(t)^2\rangle} = \frac{\eta_{ep}\,(P(\omega)/(\hbar\omega))2RC}{\left[1 + \frac{2k_BT_K}{RG^2e^2\eta_{ep}(P(\omega)/(\hbar\omega))} + \frac{P_{dk}(\omega)}{GP(\omega)}\right]}\,. \tag{56}$$

Eq. (56) is one of the main results in this section. It incorporates the leading aspects of the detection process and noise from fundamental sources. It can be used as a guidance tool in the optimization process to obtain a good signal to noise ratio, with light effects predominating over thermal sources and others. Fig. 9 shows SNR as a function of the optical power P and the gain G.

As a warning, the idealized voltmeter in Fig. 8 is, in practice, an instrument with particular noise sources. Although it is usually assumed that the voltage probes have a negligible effect on the measurement, their influence may be detected. In particular, for AtoD converters, one should examine the instrument's noise to understand its influence on the obtained data. For example,

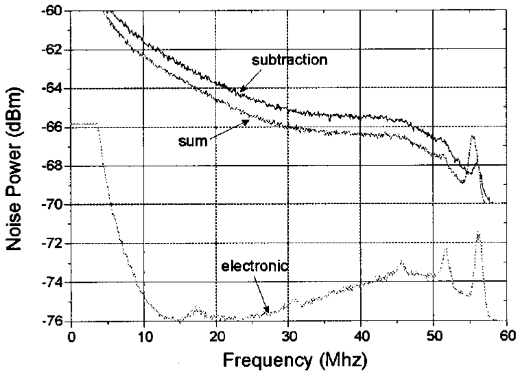

Figure 10. Example of background electronic noise compared with optical shot-noise signals for a diode laser with an external cavity. The lowest line is the electronic level (peaks are resonances in the detecting system) and upper lines are optical signals.

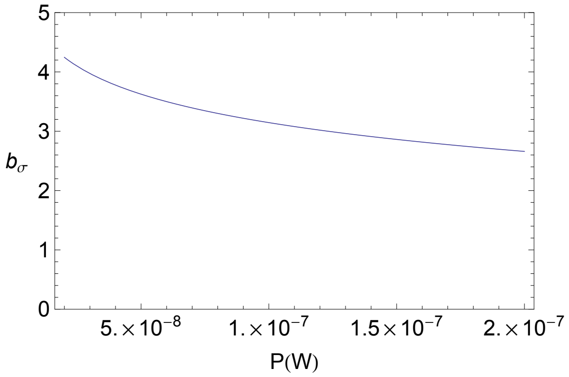

Figure 11. Number of bits available to record the fluctuations around the average optical signal. Laser power $P \sim 100 n$W, $\Delta t = 1\ ns \rightarrow 3$ bits.

turning a light source off, one can measure the background electronic noise. Fig. 6 in [19], reproduced in Fig. 10 shows a measurement of the electronic noise for a particular laser, using a spectrum analyzer.

To record fluctuations of the optical field around the average optical signal itself, a couple of conditions have to be obeyed: Firstly, the average intensity reaching the detector in Δt has to excite it. This requires a minimum number of photons (detector dependent) and, for telecommunication detectors, here estimated at ~ 600 photons in Δt. Assuming a detecting system with parameters as given in Fig. 9, an attenuated laser power of $P = 100\, n$W, probed at intervals of 1 ns, produce an average of ~ 780 photons. Secondly, the background noises (noises associated to the detecting system) should give a smaller contribution than the signal corresponding to the fluctuations of the optical signal around its average. The value $\langle n \rangle$ (in 1 sec) is given by the laser power (in MKS units), as $P = \langle n \rangle_{1s} \hbar\omega_0$. Within a sampling time Δt, the average number of photons is $\langle n \rangle_{\Delta t} = \langle n \rangle \Delta t$. Assuming $\Delta t \ll \tau_c$, the standard deviation σ from $\langle n \rangle$, or average fluctuating number of photons, is $\sigma_{\Delta t} = \sqrt{\langle (n - \langle n \rangle)^2 \rangle_{\Delta t}} = \sqrt{\langle n \rangle_{\Delta t}}$. With an applied gain $G \sim 100$, Fig. 9 shows a $SNR \sim 600$. Assuming that the total signal fulfills all of 2^b levels in the b-bits recording system, Eq. (3) gives the available bits for the signal fluctuation around its average for an 8-bits recording system (256 levels). Fig. 11 shows that for $P \sim 100\, n$W, there are ~ 3 bits available to record the fluctuating signal (± 8 in 256 levels).

Details related to a specific laser as well as to the ADC system used may heavily influence the final results due to the order of magnitude variations for some parameters; case-by-case have to be studied.

VII. Conclusions

Guidelines for construction of a basic fast multi-purpose PhRG were described as a initial contribution for construction of PhRGs for practical use in a variety of situations. Understanding the principles in these guidelines will also help the development of PhRG miniaturizations; speed gain and reduction in cost are to be expected.

Acknowledgments

This work is inserted in the project "key**BITS**" being implemented under support of RENASIC (Rede Nacional de Excelência em Segurança da Informação e Criptografia)–Ministério da Defesa/Ministério da Ciência e Tecnologia/Financiadora de Estudos e Projetos–BRAZIL.

References

[1] A. Einstein: "...I do not approve of the purely statistical way of thinking on which the new theories are founded...". From a letter to H. A. Lorentz in June 17, 1927, about Quantum Statistics.

[2] http://en.wikipedia.org/wiki/
Pseudorandom_number_generator

[3] 16 MHz Quantis-Quantum Random Number Generator, id Quantique SA (http://www.idquantique.com).

[4] For example: 250k bit/sec, True Random Number Generation IC "RPG100B" (http://www.fdk.com/whatsnew-e/release050930-e.html); VIA PadLock RNG. See also a cryptographic system in R. Mingesz, Z. Gingl, L. B. Kish, "Johnson(-like)-Noise-Kirchhoff–loop based secure classical communicator characteristics, for ranges of two to two thousand kilometers, via model-line", Phys. Letters A **372**, 978-984 (2008).

[5] I. Reidler, Y. Aviad, M. Rosenbluh, I. Kanter, "Ultrahigh-Speed Random Number Generation Based on a Chaotic Semiconductor Laser", Phys. Rev. Letters **103**, 024102/1-5 (2009).

[6] G.A. Barbosa, United States Patent #US 7,831,050 B2 (Filed on Dec. 1, 2004; Prior Publication Data Ju. 14, 2005). Fig. 3, descriptions and claims. See also INPI-Brazil, PI0405814-3 (2004).

[7] B. Qi, Y-M Chi, H-K Lo, and L. Qian, "High-speed quantum random number generation by measuring phase noise of a single-mode laser", Optics Letters **35**, 312-314 (2010).

[8] G. A. Barbosa and UFMG Team, *Telecomm. Platform*, See www.renasic.org.br/comsic/bin/view/LAPROJ/WebHome.

[9] A. Kolmogorov, "Logical Basis for Information Theory and Probability Theory", IEEE Transactions on Information Theory **14**, 662664 (1968).

[10] G. Chaitin, *Meta Math!:The Quest for Omega*. New York: Pantheon Books, 2005.

[11] R. H. Hadfield, "Single-photon detectors for optical quantum information applications", Nature Photonics **3**, 696-705 (2009).

[12] Z. L. Yan, A. W. Sharpe, J. F. Dynes, A. R. Dixon, and a. J. Shields, "Multi-gigahertz operation of photon counting InGaAs avalanche photodiodes", Appl. Phys. Letters **96**, 071101/1-3 (2010).

[13] T. Mueller, F. Xia and P. Avouris, "Graphene photodetectors for high-speed optical communications", Nature Photonics **4**, 297 (2010).

[14] R.J. Glauber, "Coherent and Incoherent States of the Radiation field", Physical Review **131**, 2766-2788 (1963). iptions and claims.

[15] L. Mandel and E. Wolf, *Optical Coherence and Quantum Optics* (Cambridge University Press, New York, 1995).

[16] B. Yurke, *Squeezed Light* (Course notes, University of Rochester, 1989).

[17] J. B. Johnson, "Thermal agitation of electricity in conductors", Phys. Rev. **32**, 97-109 (1928).

[18] L. Callegaro, "Unified derivation of Johnson and shot noise expressions", Am. J. Phys. **74**, 438-440 (2006).

[19] C. L. Garrido-Alzar, S. M. de Paula, M. Martinelli, R. J. Horowicz, A. Z. Khoury, G. A. Barbosa, “Transverse Fourier analysis of squeezed light in diode lasers”, J. Opt. Soc. Am. B **18**, 1189-1195 (2001).

G. A. Barbosa was born in Brazil, in 1943. PhD (Physics)/University of Southern California, 1974. Areas of work: Quantum Optics and Condensed Matter (Theory and Experiment). Full Professor, Universidade Federal de Minas Gerais/MG/Brazil (up to 1995); Northwestern University (2000/2012) and CEO, QuantaSec Consultoria e Projetos em Criptografia Física Ltda /Brazil. Member of American Physical Society and Sociedade Brasileira de Física. Established several research laboratories including the first Quantum Optics laboratory in Brazil, where a quantum image was first demonstrated worldwide. Has patents in cryptography, including the very first Brazilian patent in quantum cryptography: INPI-PI9806314. This system is similar to a BB84 protocol but works in time coincidences with two independently polarized photons to eliminate background noises. Another patent covers the alpha eta *encryption* system (inventors: H. P. Yuen, P. Kumar, and G. A. Barbosa), developed under support from Defense Advanced Research Projects Agency (DARPA)- Department of Defense. A patent being implemented through support from Renasic (Rede Nacional de Segurança da Informação e Criptografia) is entitled “Fast multi-photon key distribution scheme secured by quantum noise” (US-2005-0152540-A1 and Brazil-INPI 002872/G. A. Barbosa).