Optics Communications 354 (2015) 197–201

Contents lists available at ScienceDirect

Optics Communications

journal homepage: www.elsevier.com/locate/optcom

Can an "impulse response" really be defined for a photoreceiver?

F. Javier Fraile-Pelaez

Dept. de Teoría de la Señal y Comunicaciones, Universidad de Vigo, ETS Ingenieros de Telecomunicación, Campus Universitario, E-36310 Vigo, Spain

ARTICLE INFO

Article history:
Received 24 March 2015
Received in revised form
22 May 2015
Accepted 25 May 2015
Available online 28 May 2015

Keywords:
Photodetector
Impulse response
Linear system
Single photon
Quantum noise

ABSTRACT

In this paper we examine the validity of the concept of impulse response employed to characterize the time response and the signal-to-noise ratio of p-i-n and similar photodetecting devices. We analyze critically the way in which the formalism of analog linear systems has been extrapolated, by employing results from macroscopic electromagnetic theory such as the Shockley–Ramo theorem or any equivalent approach, to the extreme case of a single-photon detection. We argue that the concept of "response to an optical impulse" is ill-defined in the customary terms it is envisioned in the literature, this is, as an output current pulse having a certain predictable, calculated temporal shape, in response to the detection of an optical "Dirac delta" impulse, conceived in turn as the absorption of a single photon.

1. Introduction

It is well known that the ultimate sensitivity of any photodetector is determined by the quantum noise of the radiation itself. Specifically for a photodiode receiver, this means that the noise present at the output current of the diode should be an exact reproduction of the intrinsic noise of the impinging radiation, with no any other excess noise contributions. Obviously, the current–voltage amplification at the electronic stage of the receiver should also be noiseless, consistent with the " ideal receiver" assumption. In other words, once all additional (electronic) noise sources have been removed, the process of noiseless photodetection amounts to photon-counting through ideally equivalent temporal electron-counting.

On the other hand, the functional modelling of a photoreceiver systematically makes use of an essential concept taken from linear systems theory: the "impulse response of the receiver," [1] which is required for the analysis of both signal and noise performance of any linear, invariant system. For an analog system, the impulse response is defined as the output time signal when the input is an instantaneous impulse of unit area, i.e. a Dirac delta, $\delta(t)$, which contains all frequencies homogeneously, from 0 to ∞. Such an impulse is unrealizable (and surely unphysical), but its mathematical usefulness makes it convenient to assume its existence, at least in the approximate form of a physical impulse having a duration much shorter than any characteristic time of the system. Thus, in the case of an electrical circuit, one can think of a delta-like impulse of voltage, for example. In the case of incoherent optical reception, the input "signal" is the time-varying optical power $P(t)$, so the input impulse is to be described mathematically as $P(t) = \delta(t)$.

It should be kept in mind that all signals are inherently analog in this formalism. Actually, to a great extent, the Dirac delta works as an *ad hoc* artifact intended to allow hypothetical point-like objects (masses, charges, etc.) to "live" in continuous spatial or temporal domains, which would otherwise be unconceivable; if space–time is thought continuous, at least differential intervals are needed to contain a non-null amount of any magnitude, since a discrete *point* is, in mathematical terms, a zero measure set, thus meaningless. Only if one accepts that a space or time point can accommodate a "Dirac delta" (of charge, say), can the problem be skipped.

The above considerations lead us to the following point. Consider the optical signal to be a narrow-band modulated optical flux $\bar{Q}(t) = P(t)/(h\nu)$ (photons/s), where the overbar denotes statistical averaging and ν is the central optical frequency. This corresponds well to the archetypical case of a laser (or even LED) beam modulated in intensity by a low frequency (baseband, RF, microwave) signal varying like $P(t)$. Contrary to what is frequently implied in the literature, the "unit" impulse at the input of the detector is not *necessarily* "one photon"—in spite of the cardinal number. This confusion, detected in many textbook presentations, arises surely from the fact that the electromagnetic field, roughly speaking, happens to be quantized *in amplitude*, whereas the Dirac delta formalism was never intended to deal with "quantized analog" signals—incidentally, a concept which does not exist in linear systems theory.

As far as the signal part of the signal and noise calculations is

E-mail address: fj_fraile@com.uvigo.es

http://dx.doi.org/10.1016/j.optcom.2015.05.063
0030-4018/

concerned, the problem can be surmounted easily for two related reasons. First, the unit amplitude of the Dirac delta is purely conventional and without consequences in a linear system; obviously, if the impulse $A\delta(t)$ is employed at the system input, the system response to the unit impulse will merely be the actual output divided by A. In other words, it is the temporal condensation that matters, not the amplitude. Second, in view of the previous consideration, any sufficiently short optical pulse, yet simultaneously intense enough to clear up any concern on signal level quantization, will be a perfectly valid approximation to the unit input impulse.

Things change when the focus is put on the noise. Particularly in the case of the photonic signal noise, the inherent amplitude quantization cannot be disguised anymore and the solution described above is unfeasible. One thus has to confront a frequently overlooked issue which threatens the typical automatic extension of the linear system formalism to handle noise of quantum origin. In Sections 2 and 3 we review, very briefly, the standard theory of the signal noise as routinely applied to the linear system model of photoreceiver. The problems carried out by the accepted formalism are discussed in Section 4. Section 5 contains the conclusions.

2. Optical shot noise in the photoreceiver model

Quantum noise is almost synonymous of shot noise as far as a photoreceiver is concerned. Only two or three simple statistical concepts are needed to describe the photodetection process in the simple fashion in which it is usually modelled, and a straightforward correspondence can be apparently established between the mathematical route and the physical route. Thus, assuming a coherent light source, the random arrival times of the photons are governed by a Poisson distribution characterized by its average $\bar{N}$, related in turn to the average rate of the photon flux through $\bar{N} = \bar{Q}T$, with T being the "photocounting" period. One could anticipate at this point that T will be roughly equal to the inverse of the bandwidth, which is basically true.

Next, as an optical–electrical transducer, the photodetector transmutes the photon absorptions into charge carriers—the mathematical consequence being a mere multiplication of the actual instantaneous photon flux, $Q(t) = \sum_k \delta(t - t_k)$, by the electron charge q to arrive at the same delta train function, but this time as an electrical current rather than a photon flux: $i_\delta(t) = q\sum_k \delta(t - t_k)$. Certainly, this impossible current is only a conceptual intermediate step toward the "real" current, which in general is described by the expression

$$i(t) = q\sum_k M_k h_k(t - t_k), \tag{1}$$

where $h_k(t)$ is the shape of the *current pulse* generated across the terminals of the photodiode by the k-th absorbed photon. The functional form (1) has its grounds on the classical or semiclassical formalism of photodetection, both of which predict the photocurrent to be generated by a Poisson process at a rate given by the light intensity (see for example [2]). Expression (1) specifically describes a filtered Poisson process, the (low-pass) filtering being due to the finite duration of $h_k(t)$. In the presence of additional intensity noise, expression (1) would still be valid, then describing a doubly stochastic, filtered Poisson process [3]. The prefactor M_k accounts for the possibility of the detector being an avalanche photodiode (APD) with average gain $\bar{M}$, while the subindex k of h_k reflects the fact that the actual pulse shape may vary due to several circumstances. For example, even in a p-i-n photodiode, the shape of current pulse will vary depending on the specific location within the photodiode where the photon has been absorbed [4]. Other different situations can also be modelled by

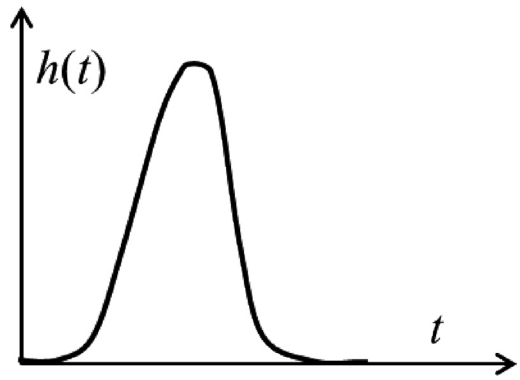

Fig. 1. When applied to a photodetector, the formalism of linear systems seeks to calculate the "impulse response" as the photocurrent pulse at the output of the device which corresponds to the detection of just one photon. In the absence of internal gain, such elementary current pulse is predicted to have the form $i(t) = qh(t)$, with q being the electronic charge and $h(t)$ the pulse shape, satisfying $\int_{-\infty}^{\infty} h(t)\, dt = 1$.

a variable impulse response (see for example [5] and references therein). Expression (1) is most often oversimplified by ignoring the random character of h_k and writing a fixed $h(t)$, sketched in Fig. 1, which is then identified with the "impulse response" of the linear system, its Fourier transform $H(\omega)$ being the photodetector transfer function.

Considering the specificities of an APD is unnecessary for the purpose of the present discussion, so we will take $M_k = 1$ and focus on a p-i-n photodiode. The shape of $h_k(t)$ is determined by the geometry and structure of the diode, mainly the width of the intrinsic layer. Fig. 1 sketches the form of the current pulse, which—always within the frame of the described approach—arises from the transit of *one* electron–hole pair photogenerated (typically) somewhere in the space charge region, toward the positive and negative, respectively, electrodes of the structure. These transit times determine the ultimate bandwidth of the photodetector.

3. Impulse response and sub-electron charge

The area of any elementary current pulse as described above is given by

$$\int_{-\infty}^{\infty} i(t)\, dt = q\int_{-\infty}^{\infty} h_k(t)\, dt = q, \tag{2}$$

which manifests the transfer of one electron charge during the duration of the pulse, or, expressed more accurately, the passage of a total charge q across an imaginary plane located at any point along the electrical circuit. Thus, at the end of the "flight time" of the electron and the hole, assuming that they do not recombine before being collected at the electrodes, one can safely say that a total charge of one electron has moved, as a conduction current, along the whole circuit. However, expression (1) has a very discomforting feature. If $h_k(t)$ is truly a current shape and the actual pulse duration spreads, say, from $t=0$ to $t = T_p$, one should be able to observe a *fractional* charge q_f given by

$$q_f = q\int_{t_1}^{t_2} h_k(t)\, dt \tag{3}$$

during any finite interval $[t_1, t_2]$, with $0 \le t_1 < t_2 \le T_p$. However striking this consequence of the formalism may look, seemingly it has never deserved a remark in any textbook or article, passing completely unnoticed in the literature to the author's knowledge.

It is necessary to recall the origin of the theory leading to this somewhat stunning result (3). Essentially, this is the Ramo or