Appendix A

Poisson processes

Section last updated February 4, 2009.

This section is meant to provide background on Poisson processes, which make their first appearance in Section 1.1. I have gone to greater lengths than usual not to leave out steps in the various derivations; clearly you can skip steps if they are obvious. Parts of these notes—including the effort not to skip steps—are adapted from Sections A.4 and A.5 of [Rieke et al 1997].

Photons from a conventional light source arrive at a detector as a random process, specifically a Poisson process. The defining feature of the Poisson process is that each event (photon arrival) is independent of all the others, given that we know the rate $r(t)$ at which the events occur. In these notes we'll go through the detailed consequences of this simple assumption of independence; hopefully some of the results are familiar. Note that many textbook presentations make a big deal out of the distinction between a "homogeneous" Poission process, in which the rate is a constant, $r(t) = \bar{r}$, and an "inhomogeneous" Poisson process in which it can depend on time. The general case isn't that hard, so I prefer to start there.

One should perhaps note at the outset that most light sources are not exactly Poisson, but the approximation is very good. There are many more systems for which the Poisson model is a decent if not excellent approximation, and so we'll discuss all this without further reference to photons: we are describing the statistics of arbitrary point events which occur at times $t_1, t_2, \cdots, t_N$.

The rate $r(t)$ can be thought of either as the mean rate of events that we would observe in the neighborhood of time t if we did the same experiment many times, or equivalently as the probability per unit time that we observe an event at t. Recall that there is the same dual definition for the concentration $c(\mathbf{x})$ of moelcules—either the mean number of molecule per unit volume that we find in the neighborhood of a point $\mathbf{x}$, or the probability per unit volume that we observe a single molecule at $\mathbf{x}$.

Since the events are independent, the probability density for observing events at times $t_1, t_2, \cdots, t_N$ must be proportional to a product of the rates evaluated at these times,

$$P[\{t_i\}|r(\tau)] \propto r(t_1)r(t_2)\cdots r(t_N) \equiv \prod_{i=1}^{N} r(t_i). \tag{A.1}$$

But to get the exact form of the distribution we must include a factor that measures the probability of *no* events occurring at any other times. The probability of an event occurring in a small bin of size $\Delta\tau$ surrounding time t is, by the original definition of the rate, $p(t) = r(t)\Delta\tau$, so the probability of no event must be $1 - p(t)$. Thus we need to form a product of factors $1 - p(t)$ for all times not equal to the special t_i where we observed events. Let's call this factor F,

$$F = \prod_{\{t_n\}\neq\{t_i\}} [1 - p(t_n)]; \tag{A.2}$$

as a shorthand we write

$$F = \prod_{n\neq i} [1 - p(t_n)]. \tag{A.3}$$

Then the probability of observing events in bins surrounding the t_i is

$$P[\{t_i\}|r(\tau)](\Delta\tau)^N = \frac{1}{N!} F \prod_{i=1}^{N} [r(t_i)\Delta\tau], \tag{A.4}$$

where the $N!$ corrects for all the different ways of assigning labels $1, 2, \cdots, N$ to the events we observe.

To proceed we pull out all the factors related to the t_i and isolate the terms independent of these times:

$$\begin{aligned} P[\{t_i\}|r(\tau)](\Delta\tau)^N &= \frac{1}{N!} F \prod_{i=1}^{N} [r(t_i)\Delta\tau] \\ &= \frac{1}{N!} \prod_{n\neq i} [1 - p(t_n)] \prod_{i=1}^{N} [r(t_i)\Delta\tau] \qquad \text{(A.5)} \\ &= \frac{1}{N!} \prod_{n\neq i} [1 - r(t_n)\Delta\tau] \prod_{i=1}^{N} [r(t_i)\Delta\tau] \qquad \text{(A.6)} \end{aligned}$$

$$= \frac{1}{N!}\prod_n [1 - r(t_n)\Delta\tau] \prod_{i=1}^{N} \left[\frac{r(t_i)\Delta\tau}{1 - r(t_i)\Delta\tau}\right]; \quad (A.7)$$

keep in mind that $\prod_n$ denotes a product over *all* possible times t_n.

To simplify Eq. (A.7) we remember that products can be turned into sums by taking logarithms, so that

$$\prod_n [1 - r(t_n)\Delta\tau] = \prod_n \exp\left\{\ln[1 - r(t_n)\Delta\tau]\right\} \quad (A.8)$$

$$= \exp\left\{\sum_n \ln[1 - r(t_n)\Delta\tau]\right\} \quad (A.9)$$

Now when we substitute back into Eq. (A.7) we find

$$P[\{t_i\}|r(\tau)](\Delta\tau)^N = \frac{1}{N!}\prod_n [1 - r(t_n)\Delta\tau] \prod_{i=1}^{N} \left[\frac{r(t_i)\Delta\tau}{1 - r(t_i)\Delta\tau}\right]$$

$$= \frac{1}{N!}\exp\left\{\sum_n \ln[1 - r(t_n)\Delta\tau]\right\} \times \prod_{i=1}^{N} \left[\frac{r(t_i)\Delta\tau}{1 - r(t_i)\Delta\tau}\right]. \quad (A.10)$$

We are interested in the case where the time bin $\Delta\tau$ is very small (we introduced these artificially, remember), which means that we need to take the logarithm of numbers that are almost equal to one. We recall that the log of one is zero, and the Taylor series of the logarithm in the neighborhood of one is

$$\ln(1 + x) = x - \frac{1}{2}x^2 + \frac{1}{3}x^3 - \cdots. \quad (A.11)$$

In this case we apply this expansion to

$$\ln[1 - r(t_n)\Delta\tau] = -r(t_n)\Delta\tau - \frac{1}{2}[r(t_n)\Delta\tau]^2 + \cdots, \quad (A.12)$$

so our expression for the probability can be written as

$$P[\{t_i\}|r(\tau)](\Delta\tau)^N = \frac{1}{N!}\exp\left\{\sum_n \ln[1 - r(t_n)\Delta\tau]\right\} \times \prod_{i=1}^{N} \left[\frac{r(t_i)\Delta\tau}{1 - r(t_i)\Delta\tau}\right]$$

$$= \frac{1}{N!}\exp\left\{\sum_n[-r(t_n)\Delta\tau] - \frac{1}{2}\sum_n[-r(t_n)\Delta\tau]^2 + \cdots\right\}\prod_{i=1}^{N}\left[\frac{r(t_i)\Delta\tau}{1-r(t_i)\Delta\tau}\right]. \tag{A.13}$$

This expression involves a sum over bins, with factors of the bin width $\Delta\tau$. We recall that this converges, as the bins become small, to an integral:

$$\lim_{\Delta\tau\to 0}\sum_n f(t_n)\Delta\tau = \int dt\, f(t), \tag{A.14}$$

for any smooth function $f(t)$. In the present case this means that

$$\lim_{\Delta\tau\to 0}\exp\left\{\sum_n[-r(t_n)\Delta\tau] - \frac{1}{2}\sum_n[-r(t_n)\Delta\tau]^2 + \cdots\right\}$$
$$= \exp\left[-\int dt\, r(t) - \frac{1}{2}\Delta\tau\int dt\, r^2(t) + \cdots\right]. \tag{A.15}$$

Now we notice that the second integral in the exponential has an extra factor of $\Delta\tau$, which comes from the $(\Delta\tau)^2$ in the previous expression, but if we really let $\Delta\tau$ go to zero this must be negligible as long as the rate doesn't become infinite. Similarly, we have in Eq. (A.13) factors like

$$\frac{r(t_i)\Delta\tau}{1-r(t_i)\Delta\tau},$$

and again as $\Delta\tau \to 0$ we can expand this in powers of $\Delta\tau$ and drop all but the first term. This is equivalent to replacing the denominator of the fraction by 1. So, when the dust clears, the expression for the probability density of the event times becomes

$$P[\{t_i\}|r(\tau)] = \frac{1}{N!}\exp\left[-\int_0^T dt\ r(t)\right]\prod_{i=1}^{N} r(t_i), \tag{A.16}$$

where we have set the limits on the integral to refer to the whole duration of our observations, from $t = 0$ to $t = T$. Note that this is a probability density for the N arrival times t_1, t_2, $\cdots$, t_N and hence has units $(\text{time})^{-N}$.

It is a useful exercise to check the normalization of the probability distribution in Eq. (A.16). We want to calculate the total probability, which involves taking the term with N events and integrating over all N arrival times, then summing on N:

$$\sum_{N=0}^{\infty}\int_0^T dt_1\int_0^T dt_2\cdots\int_0^T dt_N P[\{t_i\}|r(t)]$$
$$= \sum_{N=0}^{\infty}\int_0^T dt_1\int_0^T dt_2\cdots\int_0^T dt_N \frac{1}{N!}\exp\left[-\int_0^T dt\, r(t)\right]\prod_{i=1}^{N} r(t_i). \tag{A.17}$$

Notice that the exponential does not depend on the $\{t_i\}$ or on N, so we can take it outside the sum and integral. Furthermore, although we have to integrate over all the N different t_i together (an N dimensional integral), we see that the integrand is just a product of terms that depend on each individual t_i. This means that really we have a product of N one dimensional integrals:

$$\sum_{N=0}^{\infty}\int_0^T dt_1\int_0^T dt_2\cdots\int_0^T dt_N P[\{t_i\}|r(t)]$$
$$= \exp\left[-\int_0^T dt\, r(t)\right]\sum_{N=0}^{\infty}\frac{1}{N!}\int_0^T dt_1\cdots\int_0^T dt_N\, r(t_1)\cdots r(t_N) \tag{A.18}$$
$$= \exp\left[-\int_0^T dt\, r(t)\right] \times \sum_{N=0}^{\infty}\frac{1}{N!}\int_0^T dt_1\, r(t_1)\int_0^T dt_2\, r(t_2)\cdots\int_0^T dt_N\, r(t_N) \tag{A.19}$$
$$= \exp\left[-\int_0^T dt\, r(t)\right]\sum_{N=0}^{\infty}\frac{1}{N!}\left[\int_0^T dt\, r(t)\right]^N. \tag{A.20}$$

Recall that the series expansion of the exponential function is

$$\exp(x) = \sum_{N=0}^{\infty}\frac{1}{N!}x^N, \tag{A.21}$$

so we can actually do the sum in Eq. (A.20):

$$\exp\left[-\int_0^T dt\, r(t)\right] \times \sum_{N=0}^{\infty}\frac{1}{N!}\left[\int_0^T dt\, r(t)\right]^N$$

$$= \exp\left[-\int_0^T dt\, r(t)\right] \times \exp\left[+\int_0^T dt\, r(t)\right] \tag{A.22}$$

$$= 1, \tag{A.23}$$

which completes our check on the normalization of the distribution.

Next we would like to derive an expression for the distribution of counts, which we write as $P(N|\langle N\rangle)$ to remind us that the shape of the distribution depends (as we will see) only on its mean. To do this we take the full probability distribution $P[\{t_i\}|r(\tau)]$, pick out the term involving N events, and then integrate over all the possible arrival times of these events:

$$P(N|\langle N\rangle) = \int_0^T dt_1 \cdots \int_0^T dt_N P[\{t_i\}|r(\tau)] \tag{A.24}$$

$$= \int_0^T dt_1 \cdots \int_0^T dt_N \frac{1}{N!} \exp\left[-\int_0^T dt\, r(t)\right] \prod_{i=1}^N r(t_i). \tag{A.25}$$

As in the discussion leading to Eq. (A.20) we notice that the exponential factor can be taken outside the integral, and that really we have a product of N one dimensional integrals rather than a full N dimensional integral:

$$P(N|\langle N\rangle) = \int_0^T dt_1 \cdots \int_0^T dt_N \frac{1}{N!} \exp\left[-\int_0^T dt\, r(t)\right] \prod_{i=1}^N r(t_i)$$

$$= \frac{1}{N!} \exp\left[-\int_0^T dt\, r(t)\right] \int_0^T dt_1 \cdots \int_0^T dt_N \prod_{i=1}^N r(t_i)$$

$$= \frac{1}{N!} \exp\left[-\int_0^T dt\, r(t)\right] \left[\int_0^T dt\, r(t)\right]^N \tag{A.26}$$

$$\equiv \frac{1}{N!} \exp(-Q) Q^N, \tag{A.27}$$

where we have defined

$$Q = \int_0^T dt\, r(t). \tag{A.28}$$

In particular, the probability that no events occur in the time from $t = 0$ to $t = T$ is $P(0) = \exp(-Q)$, or

$$P(0|\langle N\rangle) = \exp\left[-\int_0^T dt\, r(t)\right]. \tag{A.29}$$

With the probability distribution of counts from Eq. (A.27), we can compute the mean and the variance of the count. To obtain the mean we compute

$$\begin{aligned} \langle N \rangle &\equiv \sum_{N=0}^{\infty} P(N) N && \text{(A.30)} \\ &= \sum_{N=0}^{\infty} \frac{1}{N!} \exp(-Q) Q^N N && \text{(A.31)} \\ &= \exp(-Q) \sum_{N=0}^{\infty} \frac{1}{N!} Q^N N. && \text{(A.32)} \end{aligned}$$

Now we have already made use of the series expansion for the exponential, Eq. (A.21), and to sum this last series we notice that

$$Q^N N = Q \frac{\partial}{\partial Q} Q^N, \tag{A.33}$$

so that

$$\begin{aligned} \langle N \rangle &= \exp(-Q) \sum_{N=0}^{\infty} \frac{1}{N!} Q^N N \\ &= \exp(-Q) \sum_{N=0}^{\infty} \frac{1}{N!} Q \frac{\partial}{\partial Q} Q^N && \text{(A.34)} \\ &= \exp(-Q) Q \frac{\partial}{\partial Q} \sum_{N=0}^{\infty} \frac{1}{N!} Q^N && \text{(A.35)} \\ &= \exp(-Q) Q \frac{\partial}{\partial Q} \exp(+Q), && \text{(A.36)} \end{aligned}$$

where in the last step we recognize the series for the exponential. Now the derivative of the exponential is just the exponential itself,

$$\frac{\partial}{\partial Q} \exp(+Q) = \exp(+Q), \tag{A.37}$$

so that

$$\begin{aligned} \langle N \rangle &= \exp(-Q) Q \frac{\partial}{\partial Q} \exp(+Q) \\ &= \exp(-Q) Q \exp(+Q) = Q. && \text{(A.38)} \end{aligned}$$

We see that the mean count is what we have called Q, the integral of the rate.

Now we can write the count distribution directly in terms of its mean:

$$P(N|\langle N\rangle) = \exp(-\langle N\rangle)\frac{\langle N\rangle^N}{N!}, \tag{A.39}$$

which is what we need to start the discussion of photon counting in vision.

We can do a very similar calculation to find the variance of the count distribution. We start by computing the average of N^2,

$$\langle N^2\rangle = \sum_{N=0}^{\infty} N^2 P(N). \tag{A.40}$$

Substituting for $P(N)$ from Eq. (A.27) and rearranging, we have

$$\begin{aligned}
\langle N^2\rangle &= \sum_{N=0}^{\infty} N^2 P(N) \\
&= \sum_{N=0}^{\infty} N^2 \exp(-Q)\frac{1}{N!}Q^N && \text{(A.41)} \\
&= \exp(-Q)\sum_{N=0}^{\infty}\frac{1}{N!}N^2 Q^N. && \text{(A.42)}
\end{aligned}$$

The trick is once again to write the extra factors of N (here N^2) in terms of derivatives with respect to Q. Now we know that

$$\frac{\partial^2}{\partial Q^2}Q^N = N(N-1)Q^{N-2}, \tag{A.43}$$

so we can write

$$Q^2\frac{\partial^2}{\partial Q^2}Q^N = (N^2 - N)Q^N, \tag{A.44}$$

which is almost what we want. But we can use the formula in Eq. (A.33) to finish the job, obtaining

$$N^2 Q^N = Q^2\frac{\partial^2}{\partial Q^2}Q^N + Q\frac{\partial}{\partial Q}Q^N. \tag{A.45}$$

Now we can substitute into Eq. (A.42) and follow the steps corresponding to Eq's. (A.34) through (A.38):

$$\langle N^2\rangle = \exp(-Q)\sum_{N=0}^{\infty}\frac{1}{N!}N^2 Q^N$$

$$= \exp(-Q) \sum_{N=0}^{\infty} \frac{1}{N!} \left[Q^2 \frac{\partial^2}{\partial Q^2} Q^N + Q \frac{\partial}{\partial Q} Q^N \right] \quad \text{(A.46)}$$
$$= \exp(-Q) Q^2 \frac{\partial^2}{\partial Q^2} \sum_{N=0}^{\infty} \frac{1}{N!} Q^N + \exp(-Q) Q \frac{\partial}{\partial Q} \sum_{N=0}^{\infty} \frac{1}{N!} Q^N$$
$$= \exp(-Q) Q^2 \frac{\partial^2}{\partial Q^2} \exp(+Q) + \exp(-Q) Q \frac{\partial}{\partial Q} \exp(+Q) \quad \text{(A.47)}$$
$$= \exp(-Q) Q^2 \exp(+Q) + \exp(-Q) Q \exp(+Q) \quad \text{(A.48)}$$
$$= Q^2 + Q. \quad \text{(A.49)}$$

Now since we have already identified Q as equal to the mean count, this means that the mean square count can be written as

$$\langle N^2 \rangle = \langle N \rangle^2 + \langle N \rangle. \quad \text{(A.50)}$$

But the variance of the count is defined by

$$\langle (\delta N)^2 \rangle \equiv \langle N^2 \rangle - \langle N \rangle^2 \quad \text{(A.51)}$$
$$= [\langle N \rangle^2 + \langle N \rangle] - \langle N \rangle^2 = \langle N \rangle. \quad \text{(A.52)}$$

Thus the variance of the count for a Poisson process is equal to the mean count.

The next characteristic of the Poisson process is the interval between events. The probability per unit time that we observe an event at time t is given by the rate, $r(t)$. The probability that we observe no events in the interval $[t, t + \tau)$ is given by

$$P(0) = \exp\left[-\int_t^{t+\tau} dt' \, r(t') \right]. \quad \text{(A.53)}$$

The probability per unit time that this interval is closed by an event is again the rate, now at time $t + \tau$. Thus the probability per unit time that we see events at t and $t+\tau$, with no events (an empty internal) in between is given by

$$P(t, t + \tau) = r(t) \exp\left[-\int_t^{t+\tau} dt' \, r(t') \right] r(t + \tau). \quad \text{(A.54)}$$

In the simple case that the rate is constant, this is just $P(t, t+\tau) = r^2 e^{-r\tau}$. On the other hand, if the rate varies, the average probability for observing two events separated by an empty interval of duration τ is

$$P_2(\tau) = \left\langle r(t) \exp\left[-\int_t^{t+\tau} dt' \, r(t') \right] r(t + \tau) \right\rangle, \quad \text{(A.55)}$$

where $\langle \cdots \rangle$ is an average over these variations in rate.

If we ask for the probability density of intervals, this is really the conditional probability that the next event will be at $t + \tau$ given that there was an event at t. To form this conditional probability we need to divide by the probability of an event at t, but this is just the average rate. Again, in the simple case of constant rate, this yields the probability density of inter–event intervals,

$$p(\tau) = re^{-r\tau}. \tag{A.56}$$

This exponential form is one of the classic signatures of a Poission process. We can think of it as arising because the moment at which the interval closes has no memory of the moment at which it opened, and so the probability that there has not ben an event must be a product of terms for the absence of an event in each small time slice $\Delta\tau$, as in the derivation above, and this product becomes an exponential.

Our last task is to evaluate averages over Poisson processes, such as the one in Eq (1.23),

$$\left\langle \sum_{\mathrm{i}} V_0(t - t_{\mathrm{i}}) \right\rangle = \sum_{N=0}^{\infty} \int_0^T dt_1 \cdots \int_0^T dt_N P[\{t_{\mathrm{i}}\}|r(t)] \sum_{\mathrm{i}} V_0(t - t_{\mathrm{i}}). \tag{A.57}$$

The key idea is to proceed simply and systematically, looking at one term in our sum and doing the integrals one at a time.

One term in the sum means that we choose, for example $\mathrm{i} = 1$ *and* one particular value of N. Then what we are trying to do is the following:

$$\begin{aligned} &\int_0^T dt_1 \cdots \int_0^T dt_N P[\{t_{\mathrm{i}}\}|r(t)] V_0(t - t_1) \\ &= \int_0^T dt_1 \cdots \int_0^T dt_N \exp\left[-\int_0^T d\tau\, r(\tau) \right] \\ &\qquad \times \frac{1}{N!} r(t_1) r(t_2) \cdots r(t_N) V_0(t - t_1). \end{aligned} \tag{A.58}$$

Notice that the exponential factor (along the the $1/N!$) is constant and comes outside the integral. Now we rearrange the order of the integrals:

$$\begin{aligned} &\int_0^T dt_1 \int_0^T dt_2 \cdots \int_0^T dt_N r(t_1) r(t_2) \cdots r(t_N) V_0(t - t_1) \\ &\qquad = \int_0^T dt_1\, r(t_1) V_0(t - t_1) \int_0^T dt_2\, r(t_2) \cdots \int_0^T dt_N\, r(t_N) \end{aligned}$$

$$(A.59)$$

$$= \left[\int_0^T dt_1\, r(t_1) V_0(t - t_1)\right] \left[\int_0^T d\tau r(\tau)\right]^{N-1}. \tag{A.60}$$

But the fact that we chose i = 1 was arbitrary; we would have gotten the same answer for any i $= 1, 2, \cdots, N$. Thus summing over i is the same as multiplying by N. This leaves us with the sum on N, so we put everything back to together to find

$$\begin{aligned}
\left\langle \sum_{\mathrm{i}} V_0(t - t_{\mathrm{i}}) \right\rangle &= \exp\left[-\int_0^T d\tau\, r(\tau)\right] \left[\int_0^T dt_1\, r(t_1) V_0(t - t_1)\right] \\
&\quad \times \sum_{N=0}^{\infty} \frac{N}{N!} \left[\int_0^T d\tau r(\tau)\right]^{N-1} &\text{(A.61)} \\
&= \exp\left[-\int_0^T d\tau\, r(\tau)\right] \int_0^T dt_1\, r(t_1) V_0(t - t_1) \\
&\quad \times \sum_{N=0}^{\infty} \frac{1}{N!} \left[\int_0^T d\tau r(\tau)\right]^{N} &\text{(A.62)} \\
&= \exp\left[-\int_0^T d\tau\, r(\tau)\right] \int_0^T dt_1\, r(t_1) V_0(t - t_1) \\
&\quad \times \exp\left[+\int_0^T d\tau\, r(\tau)\right] &\text{(A.63)} \\
&= \int_0^T dt_1\, r(t_1) V_0(t - t_1). &\text{(A.64)}
\end{aligned}$$

Thus what we have shown is that our simple model of summing pulses from single photons generates a voltage that responds linearly to the light intensity,

$$\langle V(t) \rangle = V_{\mathrm{DC}} + \int dt' V_0(t - t') r(t'), \tag{A.65}$$

which is Eq (1.24) in the main text.

In the interests of posting quickly, I'll stop here. Will come back add one more section about the power spectrum of fluctuations.