Sensors, Signals and Noise

COURSE OUTLINE

• Introduction
• Signals and Noise
• Filtering
• Sensors: PD4 - PhotoMultiplier Tubes PMT
Photo Multiplier Tubes (PMT)

- Photodetectors that overcome the circuit noise
- Secondary Electron Emission in Vacuum and Current Amplification by a Dynode Chain
- Photo Multiplier Tubes (PMT): basic device structure and current gain
- Statistical nature of the current multiplier and related effects
- Dynamic response of PMTs
- Signal-to-Noise Ratio and Minimum Measurable Signal
- Advanced PMT device structures
- APPENDIX 1: Secondary Emission Statistics and Dynode Gain Distribution
- APPENDIX 2: Understanding the PMT Dynamic Response
Circuit Noise limits the sensitivity of photodiodes...

VACUUM TUBE
PHOTODIODE

LOAD

ELECTRONICS

Detector Noise
(Dark-Current)

Circuit Noise:
very much higher than Detector Noise, sets the limit to minimum detected signal

Detector Signal (current at photocathode and anode):
just one electron per detected photon!
... but an Electron Multiplier Overcomes the Circuit Noise

Primary Detector Noise (cathode Dark-Current)

- Primary Signal (photocathode current): one electron per detected photon
- Output (anode) current: $G > 10^3$ electrons per primary electron
- Dark-current noise and/or photocurrent noise at detector output are much higher than circuit noise, which has practically negligible effect
Secondary Electron Emission in Vacuum and Current Amplification by a Dynode Chain
Secondary Electron Emission in Vacuum

- A primary electron is emitted in vacuum with very little kinetic energy $E_e < 1\text{eV}$
- Driven in vacuum by a high potential difference (some 100V), it impacts with high energy on a dynode (electrode coated with suitable material, see later)
- Energy is transferred to electrons in the dynode; some of them gain sufficient energy to be emitted in vacuum; $g > 1$ is the **yield** of secondary electrons per primary electron

![Diagram of Secondary Electron Emission](attachment:image.png)

- $V_D \approx$ from 100 to 1000 V
- $V_A > V_D$
- $I_A = g \cdot I_K$ anode current
- $I_K$ cathode current
- $V_K = 0$
- $1$ primary electron
- $g$ secondary electrons
- 1 photon hv
- Dynode
- Anode
- Cathode
Dynode materials

Secondary emitter coatings with ordinary yield:
- MgO  Magnesium Oxide
- Cs\textsubscript{3}Sb  Cesium Antimonide
- BeO  Beryllium Oxide
- Cu-Be  Copper-Beryllium alloys

Secondary emitter with high yield (due to NEA negative electron affinity, see slide 26 in PD2):
- GaP Gallium Phosphide

In the normal working range up to \(\approx500\text{V}\), the emission yield \(g\) is proportional to the accelerating voltage \(V\) (i.e. the primary electron energy)  \[ g = k_s V_D \]

At higher voltage \(g\) rises slower and tends to saturate (energy is transferred also to electrons in deeper layers, which have lower probability of escape in vacuum)

In the linear range ordinary emitters work with \(g\) values from \(\approx1.5\) to \(\approx7\) and GaP dynodes \(g\) values from \(\approx5\) to \(\approx25\)

GaP dynodes are more costly and delicate, require special care in operation and their yield tends to decrease progressively over long operation times.
Electron Multiplier

Sketch of the Principle (example with 5 dynodes)

- $V_K < V_{D1} < V_{D2} < V_{D3} < V_{D5} < V_A$
- Electron optics (i.e. potential distribution) carefully designed to lead the electrons emitted from each electrode to the next one
- $g_r > 1$ secondary electron yield of dynode $r$
- $G = g_1 \cdot g_2 \cdot g_3 \cdot g_4 \cdot g_5$ overall multiplier gain
  that is, $G=g^5$ with equal stages $g_1=g_2=\ldots=g$
Photo Multiplier Tubes (PMT): device structure and current gain
Side-on PhotoMultiplier Tubes PMTs

PMTs with side-window and opaque photocathode

The basic structure of Photomultiplier Tubes with discrete dynodes and electrostatic-focusing was first demonstrated in 1937 by the RCA Laboratories; in the following decades it was progressively improved and developed by various industrial laboratories (RCA, DuMont, EMI, Philips, Hamamatsu... )
End-on PhotoMultiplier Tubes PMT

PMTs with end-window and semitransparent photocathode

PMT Basic outline

PMT Operation sketch
PMTs with semitransparent photocathode

For comparison:
- APD Avalanche Photodiode mounted in case
- APD Avalanche Photodiode in silicon chip
PMT Gain Regulation and Stabilization

- PMTs can have high number $n$ of dynodes (from 8 to 12) and attain high gain $G$.

  With $n$ equal dynodes it is $G = g^n$; e.g. with 12 dynodes $G = g^{12}$

  - $G = 10^4$ with $g = 2,2$
  - $G = 10^5$ with $g = 2,6$
  - $G = 10^6$ with $g = 3,2$

- $G$ is controlled by the dynode bias voltage, which regulates the dynode yield $g$

- A single supply is usually employed, with high voltage $V_A$ typically from 1500 to 3000 V. The dynode voltages are obtained with a voltage-divider resistor chain; the potential difference $V_j$ between two dynodes $j$ and $(j-1)$ is a preset fraction $f_j$ of the supply $V_A$

$$V_j = f_j V_A$$
PMT Gain Regulation and Stabilization

- The supply voltage $V_A$ thus rules the yield $g_j$ of every dynode $g_j = k_S V_j = k_S f_j V_A$
  and the total gain $G = g_1 g_2 ... g_n = k_S V_1 \cdot k_S V_2 ... k_S V_n = k^n_S f_1 f_2 ... f_n \cdot V_A$
  which increases with $V_A$ much more than linearly

$$G = k^n_S f_1 f_2 ... f_n \cdot V_A^n = K_G \cdot V_A^n$$

(NB: $K_G = k^n_S f_1 f_2 ... f_n$ is constant, set by the voltage distribution and dynode characteristics)

- The gain $G$ is very sensitive to even small variations of the supply $V_A$: the relative variations of supply voltage are $n$-fold amplified in the relative variations of gain

$$\frac{dG}{G} = n \frac{dV_A}{V_A}$$

- Consequently, tight requirements must be set to the stability of the high voltage $V_A$ versus ambient temperature and/or power-line voltage variations. e.g. getting $G$ stability better than 1% for a PMT with $n=12$ dynodes
  requires a high voltage supply $V_A$ better stable than 0,08%
Cautions and limits in PMT exploitation

- The parameter values in the PMT operation must be carefully selected for exploiting correctly the PMT performance. We will point out some main aspects and call the user attention on warnings reported in the manufacturer data sheets.
- For limiting self-heating of voltage divider below a few Watt, the divider current must be $< 1 \text{ mA}$, hence total divider resistance must be at least a few $\text{M} \Omega$.
- In order to avoid nonlinearity in the current amplification, variations of dynode voltages caused by the PMT current should be negligible. The PMT output current must thus be less than $1\%$ of the divider current, i.e. typically a few $\mu\text{A}$.
- This limit is acceptable for DC current, but not for pulsed optical signals. However, fast transients of dynode voltages can be limited by introducing in the last stages capacitors in low-pass filtering configurations, as sketched in the examples.

![Diagram of PMT circuit](image)
Cautions and limits in PMT exploitation

- Space-charge effects may cause nonlinearities in the amplification of fast pulsed signals. A high charge of the signal itself can significantly reduce the electric field that drives the electrons: the higher is the pulse, the slower gets the electron collection. The pulse shape is more or less distorted, depending on its size.

- Nonlinearity can occur also if the voltage signal developed on the load is high enough to reduce the driving field from last dynode to anode.

- Magnetic fields have very detrimental effect: the electrons traveling in vacuum are deviated and the operation is inhibited or badly degraded. With moderate field intensity, magnetic screens (Mu-metal shields wrapped around the vacuum tube) can limit the effects; with high intensity fields PMT operation is actually impossible.

- PMTs are fairly delicate and subject to fatigue effects and their operation is prejudiced by mechanical vibrations.
Statistical nature of the electron multiplier and related effects
Single Electron Response SER

- The PMT output is superposition of elementary current pulses that correspond to single electrons emitted by the cathode, called **Single Electron Response (SER)** pulses.
- SER current pulses are fast (a few nanosecond width) and fairly high (pulse-charge $Gq$ from $10^5$ to $10^6$ electrons). They are remarkably higher than the noise of fast circuits; with PMT weakly illuminated they are well observable on the oscilloscope screen and each of them corresponds to the detection of a single photon.
- The SER current pulses observed have all equal pulse shape, but randomly varying pulse-amplitude; i.e. $G$ is not constant, but statistical

![Graph](image.png)

- The random fluctuations of $G$ are due to the statistical nature of secondary electron emission
- Since the SER charge is much higher than the minimum measurable detector pulse*, the statistical distribution $p(G)$ of the gain $G$ (probability density of $G$ value) can be directly collected by measuring and classifying the pulse-charge of many SER pulses.

* See the OPF2 slides
**Statistical Distribution of the PMT gain**

The plot above sketches the typical appearance of the statistical distribution \( p(G) \) of the PMT gain \( G \).

- For different PMT models and different operating conditions (bias voltage distribution on dynodes; temperature of operation; etc.) remarkably different \( p(G) \) are observed. The distributions are roughly akin to gaussian, but skewed toward high \( G \) values.

- The main parameters to be considered for analyzing the PMT operation are mean gain \( \bar{G} \), gain variance \( \sigma_G^2 \) and relative variance \( \nu_G^2 = \frac{\sigma_G^2}{\bar{G}^2} \).

\[ \Delta N = \text{relative number of SER pulses observed with charge } Q \text{ in the interval } qG \pm q\Delta G/2 \]

\[ p(G) = \frac{\Delta N}{\Delta G} \text{ probability density of } G \]

\[ \text{Mean gain } \bar{G} \]

\[ \text{Gain variance } \sigma_G^2 = \frac{1}{\Delta G} \int_{-\Delta G}^{\Delta G} (G - \bar{G})^2 f(G) dG = \frac{1}{\Delta G} \int_{-\Delta G}^{\Delta G} G^2 f(G) dG - \bar{G}^2 \]

\[ \nu_G^2 = \frac{\sigma_G^2}{\bar{G}^2} \]
Excess Noise due to Gain Fluctuations

• Emission of primary electrons from cathode is a process with Poisson statistics, i.e. mean number $N_p$, variance $\sigma_p^2 = N_p$ and relative variance $\nu_p^2 = \frac{\sigma_p^2}{N_p^2} = \frac{1}{N_p}$

• Emission is followed in cascade by statistical multiplication with fluctuating $G$

• The mean of the cascade output is $N_u = N_p \cdot \bar{G}$ (two independent processes)

• The Laplace theory of probability generating functions shows that the relative variance $\nu_u^2$ of the output of a cascade is sum of the relative variance of every stage in the cascade divided by the mean value of all the previous stages. In our case:

$$\nu_u^2 = \frac{\sigma_u^2}{N_u^2} = \nu_p^2 + \nu_G^2 = \frac{1}{N_p} + \frac{\nu_G^2}{N_p} = \frac{1}{N_p} \left(1 + \nu_G^2\right)$$

• The variance $\sigma_u^2$ thus is

$$\sigma_u^2 = N_p \bar{G}^2 \nu_u^2 = N_p \bar{G}^2 \left(1 + \nu_G^2\right) = \sigma_p^2 \bar{G}^2 \left(1 + \nu_G^2\right)$$

In conclusion, the PMT:
1) amplifies the input variance by the square gain $\bar{G}^2$, like an amplifier and
2) further enhances it by the **Excess Noise Factor** $F$ due to the gain fluctuations

$$\sigma_u^2 = \sigma_p^2 \cdot \bar{G}^2 \cdot F$$

with

$$F = 1 + \nu_G^2 > 1$$
PMT Noise versus Amplifier Noise

• A PMT amplifies by $\bar{G}^2$ the input noise like an amplifier and further increases it by the **Excess Noise Factor** $F$: $\sigma_u^2 = \sigma_p^2 \cdot \bar{G}^2 \cdot F$

• We will see that it is $F \leq 2$ for most PMT types and $F$ is close to unity for high quality **PMT types**. The factor of increase of rms noise is always moderate $\sqrt{F} \leq 1.4$ and often near to unity. Reasonably approximated evaluations can be obtained by neglecting the excess noise, i.e. with $F=1$.

• Further explanations and comments on the gain fluctuation are given in Appendix 1

• As modern alternative to a PMT, one could propose a vacuum tube photodiode coupled to a high-gain and low-noise amplifier chip, possibly with amplifier chip inside the vacuum tube. It would offer practical advantages: more simple, rugged and compact structure, lower operating voltage, etc..

• In fact, a PMT outperforms such «photodiode-with-amplifier-inside» by detecting optical signals smaller by orders of magnitude. We can better understand the matter by gaining a better insight about how these devices work.
PMT Noise versus Amplifier Noise

• In the amplifier a signal gains energy from the power supply by modulating the bias current in transistors, which must be active all the time. The amplifier noise sources are always active (shot noise of transistor bias current; Johnson noise of resistors)

• In a PD-amplifier combination it is the amplifier noise that sets the limit to the minimum measurable signal, since it is much higher than the photocathode dark-current noise

• In a PMT, the electrons of the signal gain energy directly from the voltage supply: the bias voltage accelerates them and the kinetic energy gained is exploited in the impact to generate other free electrons. There is no bias current in the multiplier chain, the current flows only when electrons are injected from the cathode.

• In a PMT there are no noise sources in the dynode chain; the minimum signal is limited by the dark-current noise and/or the photon-current noise at the cathode.

• The cathode noise is indeed slightly increased by the gain fluctuations in the dynode chain, but in practice this is always a minor effect and often it is negligible.
Dynamic response of PMTs
PMT Dynamic Response: SER pulse

PMT response to a single photon

\[ i_s(t) \text{ SER current pulse} \]

\[ \text{Photon arrival} \]

\[ \text{Transit Time } t_b \]

\[ T_w \text{ pulse width} \]

\[ \text{pulse barycenter} \]

\[ T_j \text{ transit time jitter} \]

\[ T_b = t_b \]

\[ p_b(t_b) \text{ Probability Density} \]

\[ \text{Transit Time distribution} \]
PMT Dynamic Response: SER pulses

- Differently from vacuum tube photodiodes, in PMT the rise of a SER current pulse is delayed (from ≈10ns to some 10ns dependent on PMT type and bias voltage) with respect to the photon arrival. The dynodes electrostatically screen the anode, so that only electrons traveling from last dynode to anode induce current (Shockley-Ramo theorem).

  The PMT transit time $t_b$ is defined as the delay of the pulse barycenter.

- The transit time $t_b$ randomly fluctuates from pulse to pulse, with a transit time jitter $T_j$ (full-width at half maximum FWHM of the $t_b$ distribution) from a few 100ps to a few ns depending on PMT type and bias voltage. $T_j$ is due to the statistical dispersion of the electron trajectories in the first stages of the multiplier.

- The SER pulse width $T_w$ (FWHM from a few ns to various ns, depending on PMT type and bias voltage) is always wider than the transit time jitter: $T_w \approx 5$ to 10 times $T_j$. It is due to the statistical dispersion of the electron trajectories in all the multiplier.

- $T_w$ has very small fluctuations, practically negligible

- Further explanations and comments on these parameters are given in Appendix 2
PMT Dynamic Response: Multi-photon pulses

PMT response $i_p(t)$ to a multi-photon $\delta$-like light pulse:
derived from 1) SER pulse waveform and 2) transit time distribution

$$i_p(t) = \int_0^\infty p_b(t_b) i_s(t-t_b) dt_b = p_b * i_s$$

**Single-Photon arrival**

**Multi-Photon arrival**

**$p_b(t_b)$**

**SER (normalized)**

**Transit Time distribution**

**$\delta$-response (normalized)**

**$0$**

**$t_b$**

**$t$**

**$T_w$**

**$T_j << T_w$**

**$T_p = \sqrt{T_w^2 + T_j^2} \approx T_w$**
PMT Dynamic Response: Multi-photon pulses

- The δ-response is a convolution $i_p = p_b * i_s$, hence its FWHM $T_p$ is quadratic sum of FWHMs $T_w$ and $T_j$ of the components

  $$T_p = \sqrt{T_w^2 + T_j^2}$$

- since $T_j / T_w$ is small (from 0.1 to 0.2) the width of the δ-response is practically equal to the SER current pulse width

  $$T_p \approx T_w \left[1 + \frac{1}{2} \left(\frac{T_j}{T_w}\right)^2\right] \approx T_w$$

- The finite SER pulse width establishes a finite bandwidth $f_p$ for the PMT employed as analog current amplifier

  $$f_p = 1/k_a T_w$$

  (the coefficient $k_a$ is from $\approx3$ to $\approx10$, depending on the SER pulse waveform)
Signal-to-Noise Ratio and Minimum Measurable Signal
Noise sources and filtering with PMTs

- $n_p$ photoelectron rate $\rightarrow I_p = n_p q$ photocurrent
- $n_D$ dark electron rate $\rightarrow I_D = n_D q$ cathode dark current
- $n_b$ electron rate due to photon background $\rightarrow I_b = n_b q$ photon background current
- $n_B = n_D + n_b$ total background electron rate $\rightarrow I_B = n_B q$ total background current

Noise sources:
- at cathode: $S_{ip} = 2qI_p = 2q^2 n_p$ photocurrent noise, increases with the signal
- at cathode: $S_{iB} = 2qI_B = 2q^2 n_B$ background noise, independent from the signal
- at anode: resistor load noise $S_{iR}$ and preamplifier noise $S_{iA}$ and $S_{iR}$

Let’s deal with S/N and minimum measurable signal in the basic case: constant signal current $I_p$ and low-pass filtering (typically by Gated Integration)
We consider cases with wide-band load, i.e. with \( \frac{1}{4R_L C_L} \ll f_F \), such that

a) the filtering effect of \( C_L \) is negligible

b) the circuit noise can modeled simply by a current generator

\[
S_{ie} = S_{iA} + S_{iR} + \frac{S_{vA}}{R_L^2}
\]

which can be referred back to the input (at the photocathode) as \( S_{ie}/G^2 F \)
Role of the Circuit Noise

- The circuit noise $S_{ie}$ can be modeled by a shot current at the anode: $I_e = S_{ie}/2q$ with electron rate $n_e = I_e/q = S_{ie}/2q^2$
- With wide band preamplifier and low resistance $R_L \approx$ few kΩ the circuit noise typically is $\sqrt{S_{ie}} \approx 2 \text{pA}/\sqrt{\text{Hz}}$ or more. The equivalent shot electron rate is $n_e \approx 10^{14}$ el/s or more
- Referred to input (cathode), the circuit noise is modeled by a shot current with reduced electron rate $n_e/G^2$. For instance, with $G = 10^6$ it is $n_e/G^2 \approx 100$ el/s
- The circuit noise referred to the input added to the background noise $S_{iB} = 2qI_B = 2q^2n_B$ gives the constant noise component (i.e. NOT dependent on the signal)

$$S_{iB} + \frac{S_{ie}}{G^2 F} = 2qI_B + \frac{2qI_e}{G^2 F} = 2q^2 \left( n_B + \frac{n_e}{G^2 F} \right)$$
The role of the circuit noise is assessed by comparing it to the constant noise source of the PMT, the background noise $S_{ib} = 2qI_B = 2q^2n_B$.

The background electron rate at the cathode $n_B$ may vary from a few $\text{el/s}$ to a few $10^6 \text{el/s}$, depending on the photocathode type and operating temperature and on the background light level (see Slides PD2).

In most cases of PMT application it is $n_B >> n_e/G^2F$ : the equivalent electron rate $n_e/G^2F$ is totally negligible with respect to $n_B$, the circuit noise plays no role.

In cases with moderate gain $G$ and/or very low dark current the circuit noise contribution may be significant and is very simply taken into account, by employing the resulting density of constant noise component in the evaluation.
Minimum Measurable Signal

For the sake of simplicity in the following computations we consider:

a) **negligible circuit noise.** Anyway, we know when it must be taken into account and how to do it, by considering an increased constant component of noise.

b) negligible excess noise, i.e. $F = 1$. Anyway, cases with non-negligible $F > 1$ can be taken into account simply by introducing the factor $\sqrt{F}$ to decrease the S/N and increase the noise variance and the minimum signal computed with $F=1$.

The minimum signal $I_{p,\text{min}}$ is reached when $S/N = 1$ : we will see that the result markedly depends on the **relative size of constant noise vs photocurrent noise**
Minimum Signal limited by Photocurrent-Noise

- The simplest **extreme case is with negligible background noise**: only photocurrent noise matters. With noise band-limit \( f_F = 1/2T_F \) (GI filtering)

\[
\frac{S}{N} = \frac{I_p}{\sqrt{2qI_p f_F}} = \frac{I_p T_F}{\sqrt{qI_p T_F}} = \sqrt{\frac{I_p T_F}{q}} = \sqrt{n_p T_F} = \sqrt{N_p}
\]

\( N_p = n_p T_F \) is the **number of photoelectrons** in the filtering time \( T_F \).

- In fact, the S/N can be obtained directly from the Poisson statistics of photoelectrons: with mean number \( N_p \), the variance is \( \sigma_p^2 = N_p \) and

\[
\frac{S}{N} = \frac{N_p}{\sigma_p} = \frac{N_p}{\sqrt{N_p}} = \sqrt{N_p}
\]

- **Remark** that in this case the noise is **NOT constant**, independent from the signal: as the signal goes down, **also the noise goes down!!**
Minimum Signal limited by Photocurrent-Noise

• By making lower and lower $I_p$, when $S/N = 1$ the minimum signal $I_{p,min-p}$ is reached

$$\left( \frac{S}{N} \right)_{min} = 1 = \sqrt{\frac{I_{p,min-p}T_F}{q}} = \sqrt{n_{p,min-p}T_F} = \sqrt{N_{p,min-p}}$$

• The minimum measurable photocurrent signal $I_{p,min-p}$ corresponds to just one photoelectron in $T_F$, the filter weighting time:

$$I_{p,min-p} = \frac{q}{T_F}, \quad n_{p,min-p} = \frac{1}{T_F}, \quad N_{p,min-p} = 1$$

• Observing the complete S/N equation

$$\frac{S}{N} = \frac{I_p}{\sqrt{2qI_pT_F + 2qI_Bf_F}} = \frac{I_pT_F}{\sqrt{qI_pT_F + qI_BT_F}} = \frac{n_pT_F}{\sqrt{n_pT_F + n_BT_F}} = \frac{N_p}{\sqrt{N_p + N_B}}$$

we see that the background noise is truly negligible only if $I_B \ll I_p$ for any $I_p$ down to the minimum $I_{p,min-p}$, i.e. only if

$$I_B \ll \frac{q}{T_F}, \quad n_B \ll \frac{1}{T_F}, \quad N_B \ll 1$$
Minimum Signal limited by Photocurrent-Noise

\[ \frac{S}{N} = \frac{N_p}{\sigma_p} = \frac{N_p}{\sqrt{N_p}} \]

\[ N_{p,\text{min}-p} = 1 \]

Signal measured by charge, in terms of number of photoelectrons \( N_p = n_p T_F \)
Minimum Signal limited by Background Noise

• The **opposite extreme case is with negligible photocurrent noise**: only background noise matters. More precisely, it’s the case where the limit current \( I_p = I_{p,min-p} \) computed with only the photocurrent noise is much lower than the background current \( I_B \)

\[
I_B \gg \frac{q}{T_F} \quad n_B \gg \frac{1}{T_F} \quad N_B \gg 1
\]

• There is now a different **minimum signal** \( I_{p,min-B} \) **limited by the background noise**

\[
I_{p,min-B} = \sqrt{\frac{qI_B}{T_F}} \quad n_{p,min-B} = \sqrt{n_B} \quad N_{p,min-B} = \sqrt{N_B}
\]

• In **intermediate cases both noise components** contribute to limit the minimum signal, which is computed from

\[
\frac{S}{N} = \frac{N_{p,min}}{\sqrt{N_{p,min} + N_B}} = 1 \quad 2^{nd} \text{ order equation that leads to } N_{p,min} = \frac{1}{2} \left(1 + \sqrt{1 + 4N_B}\right)
\]

\(NB: \text{ the other solution is devoid of physical meaning}\)
Minimum Signal limited by Noise

Signal charge, in terms of number of photoelectrons $N_p = n_p T_F$
Effect of Background Subtraction

The background can be evaluated in absence of optical signal by a separate measurement with weighting time $T_{FS}$, which gives

\[
\text{Mean } N_{BS} = n_B T_{FS} = N_B \frac{T_{FS}}{T_F} \quad \text{variance } \sigma_{BS}^2 = N_{BS} \quad \text{relative variance } \nu_{BS}^2 = \frac{1}{N_{BS}}
\]

Rescaled to $T_F$ it gives

\[
\text{Mean } N_B = N_{BS} \frac{T_F}{T_{FS}} \quad \text{variance } \sigma_{B2}^2 = N_B^2 \nu_{BS}^2 = N_B \frac{T_F}{T_{FS}}
\]

Subtracting it from the total, we get the net signal with increased noise

\[
\text{Signal } \quad N_{tot} - N_B = N_p + N_B - N_B = N_p
\]

\[
\text{Noise } \quad \sigma_m^2 = N_p + N_B + N_B \frac{T_F}{T_{FS}} = N_p + N_B \left(1 + \frac{T_F}{T_{FS}}\right)
\]

With background subtraction, the equation for evaluating the minimum signal is

\[
\frac{S}{N} = \frac{N_{p,\min}}{\sqrt{N_{p,\min} + N_B \left(1 + \frac{T_F}{T_{FS}}\right)}} = 1
\]
Effect of Background Subtraction

\[ \frac{S}{N} = \frac{N_{p,\text{min}}}{\sqrt{N_{p,\text{min}} + N_B \left(1 + \frac{T_F}{T_{FS}}\right)}} = 1 \]

- When the minimum signal is set by the photocurrent noise and the background noise has negligible effect, the conclusions previously drawn are still valid.
- When the background noise has a significant role in the S/N, the conclusions previously drawn must be revised.
- The revision simply consists in increasing by the factor \((1+T_F/T_{FS})\) the spectral density of the background noise and related quantities (equivalent electron repetition rate, variance etc.).
- In case of background measurement with equal filter weighting time \(T_{FS}=T_F\) the correction factor is simply 2.
- In case of background measurement with much longer filter weighting time \(T_{FS}>>T_F\) the correction is practically negligible, the background subtraction has negligible influence on the S/N.
Minimum Signal vs Filter Weighting Time

- Let now turn our attention from the signal charge (or electron number) to the current (or electron rate $n_p$) and highlight how the minimum signal varies with the filter weighting time $T_F$ (always in the basic case of constant signal and low-pass filtering).

- With $T_F$ long enough to have $n_B >> 1/T_F$, the minimum signal (limited by the background noise) is inversely proportional to the square root of $T_F$.

  $$n_{p,\text{min-B}} = \sqrt{n_B / T_F}$$

For comparison, recall that the decrease of the minimum signal with $\sqrt{1/T_F}$ is customary in cases with DC signal, constant white noise and low-pass filtering.

- With $T_F$ short enough to have $n_B << 1/T_F$, the minimum signal (limited by the photocurrent noise) is proportional to $1/T_F$.

  $$n_{p,\text{min-p}} = 1/T_F$$

Why the variation with $T_F$ is now much steeper??

- It is because now the noise density is no more constant, but increases with the photocurrent !! Making $T_F$ shorter the bandwidth is increased, hence the rms noise; the minimum photocurrent thus is increased, but this in turn increases the noise density, which further increases the rms noise.
Minimum Signal vs Filter Weighting Time

Signal current, in terms of **photoelectron rate** $n_p [el/s]$ (plotted for the case $n_B = 10000 \text{ el/s}$)

- **Limit by photocurrent noise**
  \[ n_{p,min-p} = \frac{1}{T_F} \]

- **Limit by background noise**
  \[ n_{p,min-B} = \frac{n_B}{T_F} \]

**Combined limit**

\[ n_{p,min} = \min \{ n_{p,min-p}, n_{p,min-B} \} \]
Minimum Signal vs Filter Weighting Time

a) With very short $T_F \approx 10$ ns ($\approx$ SER width):
the photocurrent noise sets a very high limit $n_{p,\text{min}} = 1/T_F \approx 10^8 \text{ el/s}$, such that any $n_B$ met in real cases is much lower and plays no role

b) With fairly short $T_F$, such that $T_F \ll 1/n_B$:
it is still $n_{p,\text{min}-p} = 1/T_F > n_B$, the effect of the background rate remains negligible and the minimum signal decreases as $1/T_F$

c) With definitely longer $T_F$, i.e. long enough to have $T_F >> 1/n_B$:
the limit $1/T_F$ set by the photocurrent noise now does not play any role, because it is much lower than the background rate $1/T_F \ll n_B$;
the minimum signal is limited by the background noise and decreases as $\sqrt{1/T_F}$.

In summary:
if one starts from very short $T_F$ (i.e. very wide bandwidth) and gradually increases $T_F$:
• in a first phase, the minimum signal $n_{p,\text{min}}$ decreases as $1/T_F$
• when $n_{p,\text{min}}$ arrives to the level of $n_B$, the decrease slows down to $\sqrt{1/T_F}$
Progress in PMT device structures
In order to get PMTs more simple, compact, robust and less sensitive to mechanical vibrations, minitubular electron multipliers were introduced (in the late years 60’s)

- A special glass capillary tube with $D < 1\text{mm}$, called Continuous Channel Multiplier CCM or Channeltron, is at once voltage divider and electron multiplier; the inner surface is chemically treated and converted in a semiconductor layer with high resistivity and secondary electron emission yield $g \approx$ from 1,2 to 3.

- For a given applied voltage the gain depends on the ratio $L/D$. As $L/D$ increases the number of impacts increases, but the yield decreases because the impacting electron energy decreases. Maximum gain is attained with $L/D \approx 50$

- Gain $G$ from $10^5$ to $10^6$ is attained with applied voltage in the range 2 to 3kV

- No need to focus electrons within the multiplier, but the electron optics from cathode to multiplier input must be carefully designed to get good collection efficiency.
Neutralizing the Ion Feedback in Channel

In order to exploit CCMs it is necessary to neutralize the effect of Ion Feedback.

- In the last part of the channel the density of energetic electrons is high and creation of free heavy ions (ionized atoms) by collision with residual gas molecules (or with the wall material) becomes probable.
- The free ions drift in the field and by impacting on the wall cause a strong emission of electrons. If the impact occurs near the channel input the emitted electrons undergo all the channel multiplication.
- This is a positive feedback effect, which enhances the current amplification in uncontrolled way and may even cause a self-sustaining breakdown current in the multiplier.
- The effect is avoided by bending the axis of the multiplier tube. Due to the large mass and small charge, a free ion has small acceleration in the electric field and its trajectory is almost straight; the ions thus impact in the last part of the channel, hence the emitted electrons undergo a much lower amplification.
CCM Performance

- Dark current is lower in CCM-PMTs than in dynode-PMTs, which collect also electrons from auxiliary input electrodes contaminated in the photocathode fabrication.
- The excess noise factor $F>2$ is significantly higher than dynode-PMTs, because the statistical dispersion in the electron multiplication is clearly greater.
- The inner layer resistance is in $G\Omega$ range, the current in this voltage divider is low $<1\mu A$, hence for avoiding nonlinearity the mean output current must not exceed a few nA. This sets a strict limit to the product of mean photon rate and PMT gain.
- The amplification of a pulse signal leaves a charge on the multiplier surface near to the output. A high charge modifies the electric field, impairing the amplification of the following pulses during a long recovery transient (discharge through the inner layer resistance, with time constant of milliseconds or more). To avoid this, the product of multiplier gain and input pulse charge and/or repetition rate must be limited.
- Strong nonlinearities due to space charge may occur for high pulses and high gain.

In conclusion, CCM-PMTs are

a) **well** suitable and provide very good performance for detecting **pulses with moderate repetition rate and small size** (down to single photons).

b) **NOT** well suitable for **many-photon-pulses** (e.g. for scintillation detectors of ionizing radiation) and for **stationary light intensity**.
Micro-Channel Plate Multiplier MCP

- For overcoming the CCM limitations, the multiplier concept evolved (in early years 70’s) to the MicroChannel Plate MCP, implemented with sophisticated glass technology.
- An array of many thousands of multiplier microtubes is embedded in beehive structure into a plate. All channels are biased in parallel with the same high voltage $V_D$, applied via metal electrodes deposited on the two faces of the plate.
- The MCP has a **planar geometry**, well matched to a planar end-window photocathode; focusing of photoelectrons on the multiplier is simply provided by a high voltage from cathode to multiplier input (**proximity-focusing** geometry)
Micro-Channel Plate Multiplier MCP

- MCPs are implemented with small diameter $D$ from 50μm to 5μm and the useful area (sum of the channel input sections) is $\approx$50 to 60% of the total plate area.
- Each channel operates as an individual miniaturized CCM: the gain is optimized still with $L/D \approx 50$.
- To avoid ion feedback by bending the channel axis is not convenient for MCPs; the same principle is exploited by two MCPs with inclined channel axis, mounted in series with channel axis of the first and second MCP forming an angle.
MCP Performance

Most of the limitations that plague CCMs are relaxed for MCPs with illumination distributed on the cathode because:

a) Electrons emitted from the same position of the cathode do not enter all in the same microchannel; they are distributed over a group of facing channels in the MCP.

b) The perturbation of the voltage distribution in a channel affects the multiplication and collection of electrons just in that channel and closest neighbors, not farther.

It follows that:

1) the limit to the output mean signal current is much higher; it is a small percentage of the total bias current of the MCP, not of a single microchannel

2) also many-photon optical pulses are correctly linearly processed, since the pulse photoelectrons are multiplied in parallel in different microchannels

• The statistical gain distribution of MCPs is similar to CCMs, significantly wider than for dynode-PMTs, with excess noise factor significantly higher $F>2$

• The dynamic response of MCPs is remarkably superior to that of dynode PMTs. The transit time $T_b$ and its jitter $T_j$ are remarkably shorter; in fast MCP types they are reduced down to $T_b \approx 1\text{ns}$ and $T_j \approx \text{a few }10\text{ps}$.

Also the SER pulse-width $T_w$ is shorter, down to $T_w \approx \text{a few }100\text{ps}$. 
Miniaturized Metal-Channel PMTs

Various applications of wide interest in Medical Care, Life Sciences, Environmental Monitoring, etc. require photodetectors with PMT performances and miniature size, ruggedness, low sensitivity to vibrations and reduced sensitivity to magnetic fields.

In late years 90’s mechanical micromachining techniques were employed to implement PMTs with compact size (≈1cm) in metal case, with miniaturized metal dynodes in array structure that defines multiplier channels.

Figures by courtesy Hamamatsu Photonics K.K.
Micro PMTs

In recent years micromachining techniques have been developed within the silicon microelectronic fabrication technology for producing Micro-Electro-Mechanical Systems MEMS and Micro-Opto-Electro-Mechanical Systems MOEMS. Exploiting these techniques new miniaturized PMT structures have been developed. The multiplier is fabricated in a silicon wafer; the chip has a few millimeter size and is mounted in a flat package, in sandwich between two glass layers as here outlined.
APPENDIX 1:
Secondary Emission Statistics and Distribution of Dynode Gains
Secondary Emission Statistics and Gain Distribution

- PMT is a cascade of statistical multiplication stages (dynodes)
- Let’s consider the statistics of electrons emitted by a dynode $j$ and collected at the next dynode $j+1$ in response to one electron that impacts on dynode $j$, which is characterized by mean number $g_j$, variance $\sigma_j^2$ and relative variance $v_j^2 = \frac{\sigma_j^2}{g_j^2}$
- From the Laplace theory of probability generating functions we get

$$v_G^2 = \frac{\sigma_G^2}{G} = v_1^2 + \frac{v_2^2}{g_1} + \frac{v_3^2}{g_1g_2} + ... + \frac{v_n^2}{g_1g_2...g_{n-1}}$$

- The key importance of the first dynode for reducing the gain fluctuations is intuitive and quantitatively defined by the equation: $v_1^2$ has full weight in $v_G^2$ and the yield $g_1$ reduces the weight of all the following $v_j^2$. Therefore, the cathode-to-first dynode stage deserves a higher voltage than the dynode-to-dynode stages
Secondary Emission Statistics and Gain Distribution

- Secondary emission has Poisson statistics if all electrons in a stage are focused and impact in the same position of the dynode. Otherwise, non uniformity of emission over the dynode produces distributions with greater variance (composed Poissonian or Polya distributions). Anyway, Poisson statistics is at least a fair approximation.

- For a dynode with Poisson statistics it is

\[ \sigma_j^2 = g_j \quad \text{and} \quad v_j^2 = \sigma_j^2 / g_j^2 = 1 / g_j \]

- Hence for the complete multiplier we get

\[
v_G^2 = \frac{\sigma_G^2}{G^2} = \frac{1}{g_1} + \frac{1}{g_1 g_2} + \ldots + \frac{1}{g_1 g_2 \cdots g_n}
\]

with all dynodes with equal yield \( g_d \) this gives

\[
v_G^2 = \frac{1}{g_d} \left( 1 + \frac{1}{g_d} + \frac{1}{g_d^2} \ldots + \frac{1}{g_d^{n-1}} \right) = \frac{1}{g_d} \cdot \frac{1 - \left(1/g_d\right)^n}{1 - 1/g_d}
\]

Note that \( (1/g_d)^n \ll 1 \); e.g. with \( g_d = 2 \) and \( n=10 \) it is \( (1/g_d)^n \approx 0.001 \)
Secondary Emission Statistics and Gain Distribution

• Neglecting \((1/g_d)^n \ll 1\) we get

\[
v_G^2 \approx \frac{1}{g_d - 1}
\]

which confirms that the excess noise factor \(F\) is not high

<table>
<thead>
<tr>
<th>(g_d)</th>
<th>(v_G^2)</th>
<th>(F = 1 + v_G^2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>3</td>
<td>0,5</td>
<td>1,5</td>
</tr>
</tbody>
</table>

• If the first dynode has yield \(g_1\) different from all others with equal \(g_d\), we get

\[
v_G^2 = \frac{1}{g_1} \left(1 + \frac{1}{g_d} + \frac{1}{g_d^2} + \cdots + \frac{1}{g_d^{n-1}}\right) = \frac{1}{g_1} \cdot \frac{1}{1 - 1/g_d} \approx \frac{1}{g_1} \cdot \frac{g_d}{g_d - 1}
\]

and it is confirmed that a more efficient first dynode can remarkably reduce \(F\)

<table>
<thead>
<tr>
<th>(g_1)</th>
<th>(g_d)</th>
<th>(v_G^2)</th>
<th>(F = 1 + v_G^2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>2</td>
<td>0,5</td>
<td>1,5</td>
</tr>
<tr>
<td>10</td>
<td>2</td>
<td>0,2</td>
<td>1,2</td>
</tr>
<tr>
<td>20</td>
<td>2</td>
<td>0,1</td>
<td>1,1</td>
</tr>
</tbody>
</table>
APPENDIX 2:
Understanding the PMT Dynamic Response
Causes of Time Dispersions in the SER

- The finite FWHM jitter $T_j$ of the transit time and the finite pulse FWHM $T_w$ of the SER originate from the dispersion of the electron transit times from cathode-to-first dynode, from first-to-second dynode, etc.

- The electron transit time from each electrode to the following electrode in the chain statistically fluctuates because
  
  - a) electrons are emitted from different positions on the electrode $\rightarrow$ different trajectories $\rightarrow$ different transit times
  
  - b) electrons have random initial velocity $\rightarrow$ different trajectories even starting from same position $\rightarrow$ different transit times.

The size of the effects increases with the initial velocity values, which are fairly small for photoelectrons at the cathode (kinetic energy around 1eV or lower) and quite higher for secondary electrons from dynodes (kinetic energy up to tens of eV, increasing with dynode-to-dynode voltage)

- The transit time differences can be minimized (but not eliminated) by careful design of the geometry and potentials of electrodes (i.e. of the electron-optics system)
Causes of Time Dispersions in the SER

• For characterizing the width of a distribution the FWHM is quite practical, but for mathematical analysis the variance $\sigma$ is more appropriate. For a gaussian distribution $\frac{FWHM}{\sigma} = 2.36$ and comparable ratios are found for other distributions.

• We will denote by $\sigma_{tj}$ the SER transit time variance, $\sigma_{tw}$ the SER rms width ($2^\circ$ central moment of the waveform), $\sigma_{tk}$ the cathode-to-dynode 1 transit time variance, $\sigma_{t1}$ the dynode1-to-dynode 2 transit time variance, $\sigma_{tr}$ the dynode r-to-dynode(r+1) transit time variance.

• The statistical processes in the various stages are independent, hence the overall variance is a quadratic sum of weighted contributions of the various stages.

• The weights can be estimated (at least approximately) by carefully considering the physical processes involved.
SER Transit Time Variance $\sigma_{tj}$

- Cathode to first dynode: only one electron travels. Any transit time shift is passed to all the descendence of secondary electrons generated in the multiplier chain. $\sigma_{tk}^2$ thus contributes in full to $\sigma_{tj}^2$, its weight is unity.

- Dynode 1 to dynode 2: $g_1$ electrons travel there (on average). The barycenter of the electron set is shifted by the average of individual shifts of the $g_1$ electrons. $\sigma_{t1}^2$ reduced by $g_1$ contributes to $\sigma_{tj}^2$, i.e. its weight is $1/g_1$.

- Dynode 2 to dynode 3: $g_1 \cdot g_2$ electrons travel there (on average), etc. $\sigma_{t2}^2$ reduced by $g_1 \cdot g_2$ contributes to $\sigma_{tj}^2$, i.e. its weight is $1/g_1 \cdot g_2$ and so on for the following dynodes.

$$\sigma_{tj}^2 \approx \sigma_{tk}^2 + \frac{1}{g_1} \sigma_{t1}^2 + \frac{1}{g_1 g_2} \sigma_{t2}^2 + ....$$

We see that the cathode-first dynode stage deserves higher voltage also for minimizing the SER transit time jitter.

NB a thorough mathematical analysis shows that the fluctuations of the number of traveling electrons increase the weight of the dynode stages, but with a moderate factor of increase (about the excess noise factor F) so that the equation above is a good approximation.
SER Pulse Width $\sigma_{tw}$

- The SER waveform is the result of cascaded statistical distributions in time in the various stages of charge multiplication and collection.
- It is given by the convolution of the distributions in time of the various stages and finally also of the current waveform induced on the PMT anode by a single electron (given by the Shockley-Ramo theorem, as seen for vacuum phototubes)
- In the **cathode-to-first-dynode** stage there is **just one** electron: a fluctuation of its transit time is rigidly passed to all the descendant electrons and does not contribute to their dispersion in time. In conclusion $\sigma_{tk}^2$ does **NOT** contribute to $\sigma_{tw}^2$, whereas it fully contributes to $\sigma_{tj}^2$
- Everyone of the following dynode stages fully contributes to the carrier dispersion: all $\sigma_{tr}^2$ have full weight in $\sigma_{tw}^2$
- The mean-square-width $\sigma_{ta}^2$ of the anode current induced by an electron (see SR theorem) also has full weight in $\sigma_{tw}^2$
- In conclusion all stages contribute in full, except the first one that does not contribute at all to the SER pulse width

$$\sigma_{tw}^2 \approx \sigma_{t1}^2 + \sigma_{t2}^2 + \ldots + \sigma_{tn}^2 + \sigma_{ta}^2$$