» Technical note

MPPC

Authors: A. Ghassemi, K. Sato, K. Kobayashi | Editors: Y. Ohashi, Y. Enomoto, Y. Adachi

Contents

Silicon as a photodetector technology

Over recent years, various optical applications have experienced a shift in their optimal choice of photodetector technology. This shift has followed a general trend away from the vacuum-based photomultiplier tube (PMT) towards solid-state silicon photodetectors. In addition to the greater cost efficiency of silicon device microfabrication and scalability of wafer processing, several technological advantages based on physical properties of silicon photodetectors have contributed to that trend:

- Considering silicon's narrow bandgap (1.14 eV) and due to higher transition probability of a photoelectron from a silicon crystal's valence band to its conduction band than emission probability of a photoelectron from an alkali-based photocathode to the vacuum level, silicon photodetectors can attain higher quantum efficiencies over a wider range of wavelengths (UV-VIS-NIR) than the PMT.

- Silicon's semiconductor properties and crystalline structure allow for formation of extrinsically-doped regions that can be depleted of charge carriers (electrons or holes) in presence of an electric field, hence diminishing carrier recombination. This, along with high (excess) carrier drift velocity in the depleted regions, provides for a higher collection efficiency of photoelectron charge (i.e. less photoelectric signal loss) than the likelihood of a photoelectron creating secondary electron emissions in bombarding a PMT's first dynode.

- Fine feature sizes attainable by modern silicon microfabrication processing allow the implementation of micron-sized PN junctions and anode/cathode regions. Combined with high electrical conductance of doped silicon, that permits the formation of strong electric field (E-field) intensities within a silicon detector's depletion layer. This enables low-voltage operation of silicon photodetectors at low supply currents, resulting in low power consumption. Furthermore, at sufficiently strong E-field intensities, a high carrier multiplication or avalanche gain within small physical dimensions of a silicon device is made possible; this internal gain mechanism is fundamental to increasing the signal level to above the noise floor of an output amplifier for signal readout. Additionally, small PN junction and anode/cathode physical dimensions result in small junction capacitance, which facilitates high-frequency response.

- When a photosensitive area larger than what monolithic chips (as restricted by wafer size) can accommodate is required, compactness of silicon photodetectors with a high fill-factor (i.e. ratio of photosensitive area to total area) permits the construction of large focal-plane arrays of tiled chips with ultrahigh (≤200 μm of dead gap) packing-efficiency.

- Silicon photodetectors are mechanically rugged and are not hampered by effects of magnetic field, hysteresis, and aging/warmup considerations unlike PMTs.

Despite the above strengths of silicon photodetectors, it should be noted that the PMT has lower dark current output per unit of photosensitive area, less capacitance per unit of area, and greater radiation hardness.

It is also noteworthy to add that silicon is the material of choice for VLSI microelectronic fabrication of mixed-signal (analog + digital) ICs, based specifically on CMOS transistors that are essential to implementation of modern signal processing schemes. Furthermore, silicon's native oxide has excellent light transmittance while silicon itself is opaque; this enables microfabrication of mirrors, waveguides, and gratings on silicon wafers. Additionally, since silicon has remarkable mechanical strength and rigidity, various three-dimensional dynamic (actuators, oscillators, etc.) or static (capacitors, inductors, bonding and anchoring pads, and other fixtures) structures can be constructed on silicon wafers via wet or dry etching processes based on particular lattice orientations. These fabrication technologies lead to integration of silicon photodetectors into MOEMS (micro-opto-electro-mechanical systems) products with tremendous potential in a wide variety of optical applications that demand large-scale cost-effective manufacturing, dimensional miniaturization, and ultralow power consumption; a similar integration of PMT is not possible due to its inherent dimensional and power-consumption limitations. Altogether, a technological revolution, known as Lab-on-Chip (LoC), of micron-scale wafer-level integration of optical measurement, photoelectric signal processing, and digital data output with enormous potential in biomedical, scientific/analytical, industrial, and consumer markets is underway as a result.

1. Operation principles & characteristics

Technological evolution from Si PN/PIN photodiodes to APD and MPPC

1-1. The PN junction & unity-gain Si photodiodes

Silicon PN and PIN photodiodes are the most basic of silicon photodetectors; a fundamental understanding of their operation principles is essential for our discussion.

Crystalline silicon at highest levels of purity has a steady equilibrium of negative (electrons) and positive (holes) charge carriers; despite possible local fluctuations, neither carrier type can attain general majority. This state of silicon crystal is referred to as *intrinsic* as almost no foreign impurities are present within the crystal lattice; intrinsic silicon's resistance is quite high, making it unsuitable for efficient charge collection. Si PN photodiodes are fabricated by forming a PN junction in crystalline silicon by doping[1] a portion of it with electron donors (i.e. atoms of group V elements like phosphorus or arsenic with five electrons in their outermost electron shell) to form an N region (where negatively-charged electrons become the majority charge carrier) and by doping an adjacent portion with electron acceptors (i.e. atoms of group III elements like boron or aluminum with one electron in their outermost shell) to form a neighboring P region (where positively-charged holes become the majority charge carrier). Since each silicon atom has four electrons in its outermost shell, a donor atom can share four of its outermost electrons with neighboring silicon atoms to form an electrovalent bond with them and participate in the silicon lattice structure by donating its fifth outermost electron to the conduction band (a lattice-wide energy level in which an electron can move from an atom to another within the crystal lattice) at room temperature. Likewise, an acceptor atom can share its three outermost electrons with neighboring Si atoms and join the crystal lattice by accepting an electron and thus introducing a hole (i.e. an absence of electron) into the valence band (a lattice-wide energy level in which a hole can move from an atom to another within the crystal lattice) at room temperature.

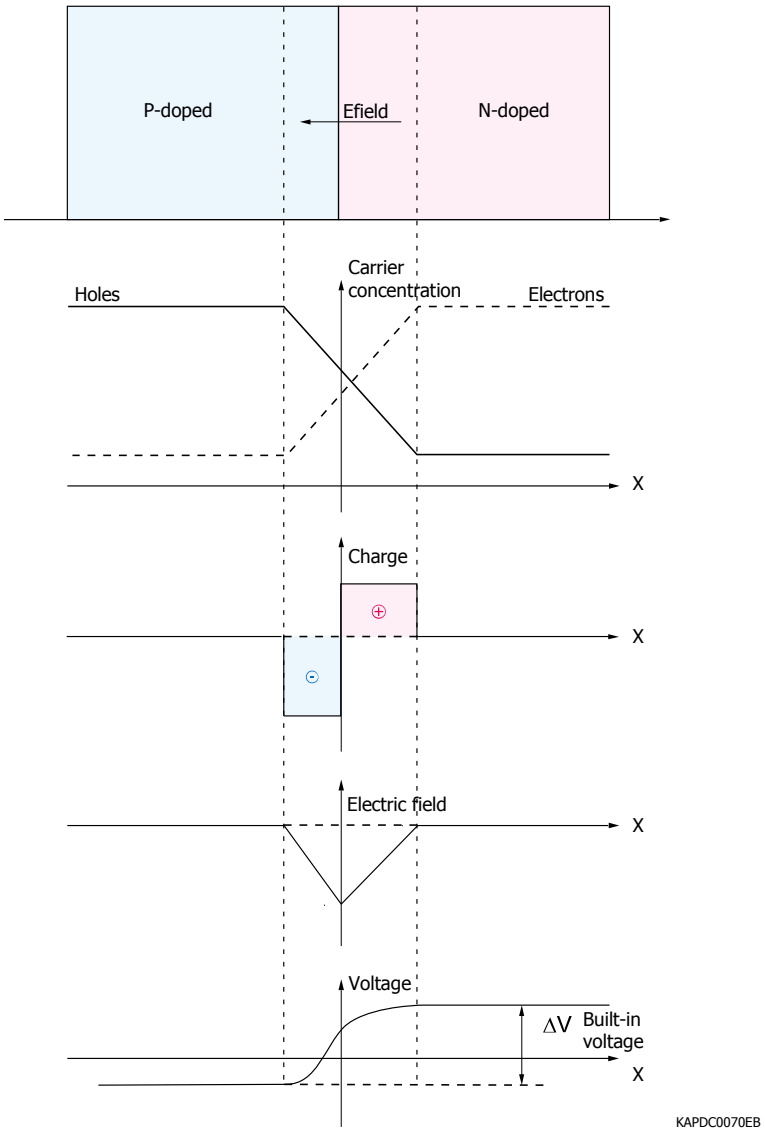
The resulting P (anode) and N (cathode) regions are commonly referred to as *extrinsic* with much lower resistance than intrinsic silicon as increases in majority carrier concentration decrease electrical resistance of a doped region. Since all participating atoms in the lattice of a monolithic semiconductor crystal share the same valence and conduction band energy levels, electrons in the conduction band and holes in the valence band can roam within the crystal's physical dimensions. Either carrier type can, however, enjoy far greater mobility under an electric field by drifting rapidly across the E-field than to diffuse through the semiconductor crystal by Brownian random walk in the absence of an electric field.

Once a PN junction is formed, considering that its N and P regions have a large concentration gradient of majority carriers between them, a diffusion current appears with electrons diffusing from the N side and holes diffusing from the P side towards the opposite side in order to recombine (i.e. a hole and an electron recombine to neutralize each other's electrical charge). The resulting recombination process depletes the layer adjacent to the junction on

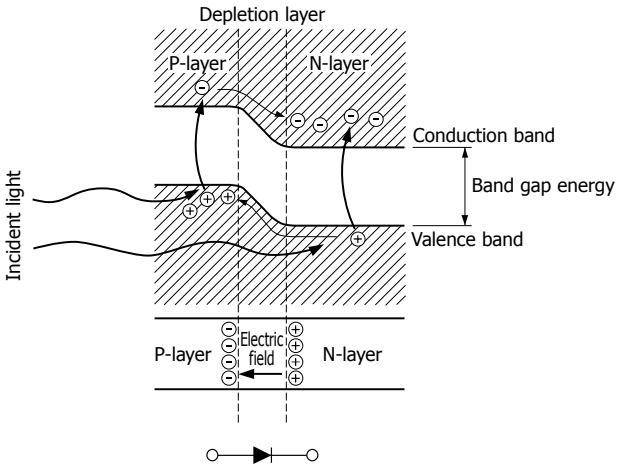
[1] Doping is a phenomenon that can be appropriately described as forming a solid-state solution through Brownian motion of the dissolved impurity atoms (i.e. dopants) into silicon (as the solvent) at elevated temperatures, which results in the dopant atoms participating in the silicon crystal lattice structure. To create this solution, the impurity atoms must be introduced into the silicon bulk by either of two means: i. Diffusion in which the Brownian motion begins by deposition of the dopant atoms at the silicon surface and progresses in depth away from the silicon surface into the silicon bulk (as a spreading Gaussian profile) by thermal agitation, and ii. Ion implantation in which the silicon bulk is bombarded by high-energy ions using an industrial particle accelerator or RF plasma oscillator; this bombardment is then followed by a phase of thermal annealing that allows the bombardment damage to the silicon lattice to be repaired while (in a similar manner to diffusion) the dopant atoms are spread away from their peak concentration (at the mean bombardment depth) by the Gaussian profile of Brownian motion and concurrently get positioned within the lattice structure. Generally speaking, diffusion is suitable for forming large but coarse doped regions at shallow depths while ion implantation is used for creating deep-buried junctions of fine size or precise profile.

each side from majority carriers (hence forming what is called the *depletion layer*). Since there is an immobile ionized donor or acceptor atom in the lattice for every majority carrier (conduction electron or valence hole) in a doped region, opposite charges form in the depletion layer from acceptor and donor ions whose hole and electron contributions have recombined and been neutralized. Consequently, the depletion layer's N-region side (with an electron majority) becomes positive due to loss of its electrons (and because of its uncompensated donor ions); the depletion layer's P-region side (with a hole majority) becomes negative due to loss of its holes (and because of its uncompensated acceptor ions). These charges cause the formation of an electric field across the depletion layer that points from the layer's positively-charged N side towards its negatively-charged P side. Pointing from the cathode (i.e. the N region) to the anode (i.e. the P region), this electric field is considered to have a negative polarity by convention; it has an absolute intensity peaking at the PN junction (where distance between opposing charges is minimal) and rolling off farther away from it into each side of the depletion layer.
While being formed, this electric field exerts a force on diffusing carriers, repelling the holes from the positively-charged N side and the electrons from the negatively-charged P side of the depletion layer. This force is in opposite direction of the carrier diffusion process that caused the E-field's formation in the first place and becomes equal to it in magnitude. Thus, net carrier migration between the P and N regions comes to an end, and charge carrier flow across the junction reaches equilibrium [Figure 1-1].

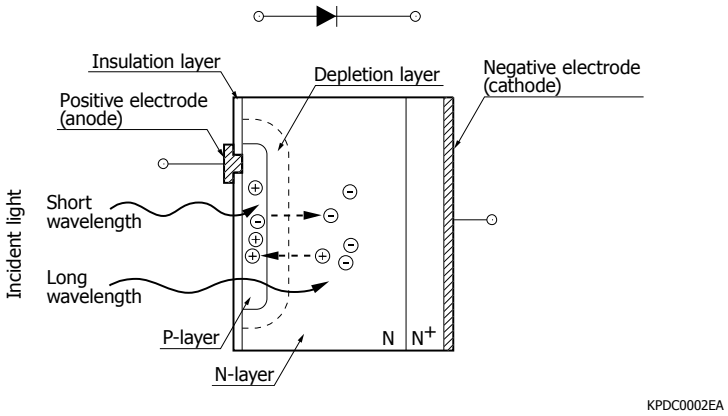
[Figure 1-1] The PN junction under equilibrium



[Figure 1-2] Visualization of silicon bandgap as the energy difference between valence and conduction bands



[Figure 1-3] Absorption of shorter wavelengths vs. longer ones in a Si PN photodiode



However, if an external source of energy can be used to generate a surplus of charge carriers in either side of the depletion region, it is possible to disturb that equilibrium and induce a net current across the junction. That external source of energy could be thermal agitation (resulting in the so-called *dark current*) or the photoelectric effect if enough energy can be provided to an electron in the valence band to transition to the conduction band and leave a hole behind in the valence band [Figure 1-2]. The resulting electron and hole are electrically compensated by each other (and not by any ion in the lattice); together, they are referred to as an *electron-hole pair* and are considered to be *excess* charge carriers.

Due to the presence of the depletion layer's native E-field, once an excess electron-hole pair is formed in the depletion layer, each carrier type (electron or hole) drifts in opposite direction towards its majority side (N or P). That is because (by the laws of electrostatics) electrons, being negatively charged, must travel against the E-field's direction (from the lower electrical potential of the depletion layer's negative P side towards the higher potential of its positive N side) and holes, being positively charged, must travel in the E-field's direction (from the higher electrical potential of the depletion layer's positive N side towards the lower potential of its negative P side). In this process, only minority carriers (electrons generated in the depletion layer's P side and holes generated in the layer's N side) drift across the PN junction to arrive at the depletion layer's outer edges; majority carriers (electrons generated in the layer's N side and holes generated in the layer's P side) do not cross the junction but drift to reach the layer's outer edges.

After the excess electrons and holes reach the respective outer edge of the depletion layer, the electrons accumulate in the N region while the holes accumulate in the P region. Hence, if a current loop is formed between the P and the N regions in order to allow for the accumulated electrons and holes to travel to the opposite regions and recombine, photocurrent can flow in that loop (if the depletion layer is exposed to light photons of energies greater than the silicon band gap, which is 1.14 eV of energy difference between silicon's conduction and valence bands). Certainly, the photoelectric effect can create electron-hole pairs outside the depletion layer as well, but most such pairs will have short lifetimes and undergo recombination before they can reach the depletion layer. That is due to poor collection efficiency in absence of an E-field outside the depletion layer since carriers would need to rely on diffusion and Brownian random walk to reach the layer (and then drift across the junction and be collected for readout).

The depletion layer's depth affects the photosensitivity and frequency response of the PN photodiode. Longer light wavelengths are absorbed at deeper depths within silicon, and thus, enlarging the depletion layer's depth increases the PN photodiode's "red" photosensitivity [Figure 1-3]. Also, analogous to increasing the distance in between the conducting plates of a parallel-plate capacitor, increasing the depletion layer's depth decreases the PN junction's capacitance, proportionally improving the PN photodiode's frequency bandwidth (considering that $Z_C(f) = \frac{1}{j2\pi fC}$ is the frequency-domain electrical impedance of a capacitor). In either case, however, an increased depletion layer depth also results in greater collection of thermally-generated carriers that form a photodetector's dark current. Nevertheless, many pulsed or red/NIR applications find a larger depletion depth to be beneficial in overall terms. For such applications, a PN photodiode's depletion layer can be deepened by biasing the PN junction in reverse, which means biasing the N region (cathode) at a higher electric potential than the P region (anode). As the reverse bias voltage is increased in magnitude, the depletion layer's depth can reach full depth of the photodiode's bulk silicon, eliminating the loss of those photo-electrically generated carriers that would otherwise need to diffuse in order to reach the depletion layer and be collected.

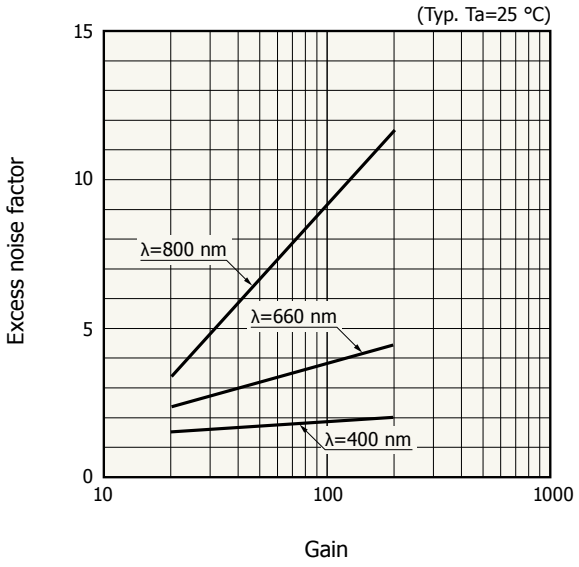
Moreover, in order to further improve the frequency response and bandwidth of PN photodiodes, a region of intrinsic silicon can be introduced in between the P and N regions, forming a PIN photodiode. This intrinsic region results in a photodiode of smaller capacitance per unit of area, and thus, increased cutoff frequency and bandwidth. The intrinsic region's role in decreasing junction capacitance can also be analogized to increasing the distance in between the conducting plates of a parallel-plate capacitor. This feature makes PIN photodiodes suitable for high-frequency pulsed applications that enjoy relatively abundant levels of light as PIN (and PN) photodiodes have no internal carrier multiplication or gain mechanism. These photodiodes operate at a gain of 1, and thus, they are suitable for detecting relatively strong light signals. With that, we will now proceed to discussing silicon photodetectors with internal gain mechanisms that are suitable for detecting lower light signal levels.

1-2. Avalanche photodiode (APD)

Aside from inducing deeper depletion, increasing the reverse bias voltage across a PN junction increases the electric force exerted on charge carriers (electrons and holes) by intensifying the E-field present in the depletion layer. Carriers accelerated by this force experience an increase in velocity and hence kinetic energy in between scattering collisions with atoms (whether silicon or dopant) in the crystal's lattice. Such collisions result in loss of carrier kinetic energy through thermal vibration of the crystal lattice (referred to as *phonon vibration*). However, if the E-field's strength is such that mean energy of carriers can exceed the silicon band gap energy during their average travel time in between consecutive collisions, sufficiently-energetic carriers emerge that are likely to ionize lattice atoms upon impact and release at least another electron-hole pair into conduction and valence bands per impact. This impact ionization effect constitutes a carrier multiplication phenomenon by which the number of drifting carriers increases rapidly, resembling an avalanche. The extent of that increase in carrier population between the original photo-carrier count and the final charge signal collected and read out is the gain of the avalanche process.

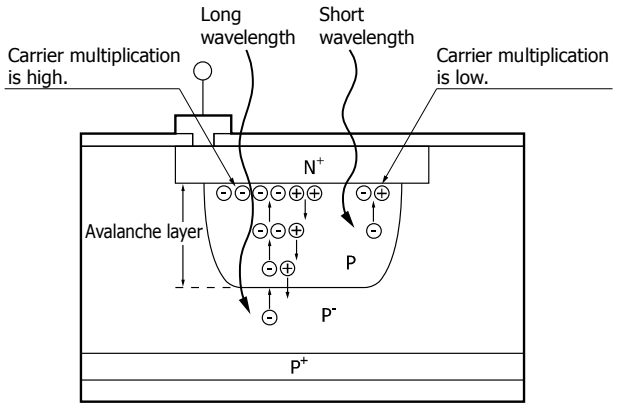
The sole benefit of avalanche gain is to increase the magnitude of the original photoelectric signal (with respect to its own inherent noise) to a level that can adequately surpass the noise floor of the readout mechanism (typically an amplifier circuit). However, since carrier multiplication in terms of E-field acceleration time/distance and energy losses to phonon collisions has an inherent randomness, the avalanching carriers develop a fluctuating spread of energies and hence multiplication gains during the avalanche process. That in turn contributes an excess noise factor [Figure 1-4] to the original signal as a random fluctuation in the APD's overall internal gain. Despite this multiplication excess noise factor, many mid-to-low light level applications nonetheless find the APD and its internal gain greatly beneficial. APD gain is dependent on the incident light wavelength [Figures 1-5, 1-6] and increases with reverse bias voltage [Figure 1-7]. Furthermore, the gain diminishes with temperature (due to increased phonon vibrations at higher temperatures and thus greater losses in kinetic energies of avalanching carriers due to increased scattering collisions) as shown in Figure 1-8.

[Figure 1-4] Plots of excess noise factor vs. gain at three different wavelengths for a Hamamatsu APD



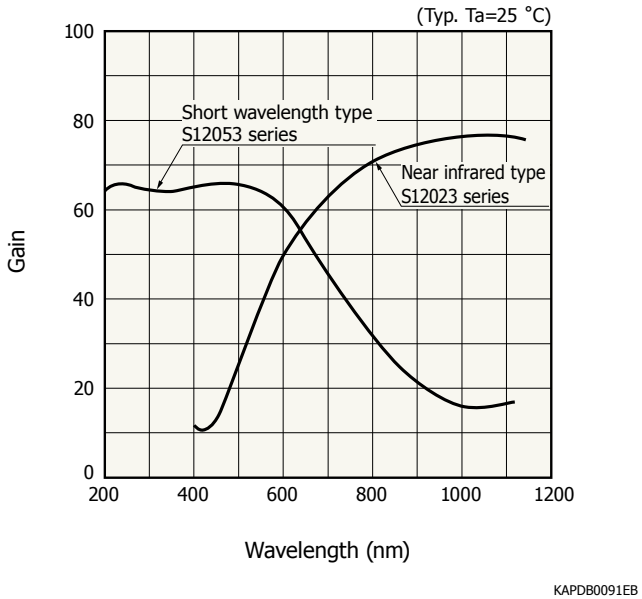
KAPDB0092EA

[Figure 1-5] Visualization of absorption of short and long wavelengths of light in a red/NIR-enhanced N-on-P APD structure



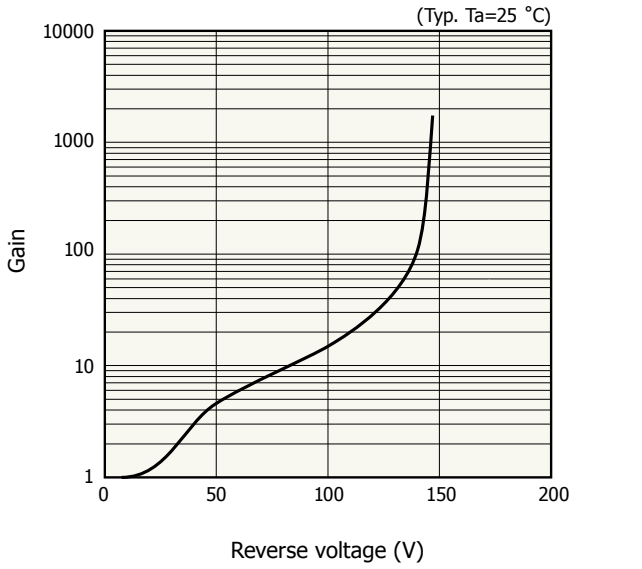
KAPDC0014EB

[Figure 1-6] Gain dependence on incident light wavelength for two Hamamatsu APDs



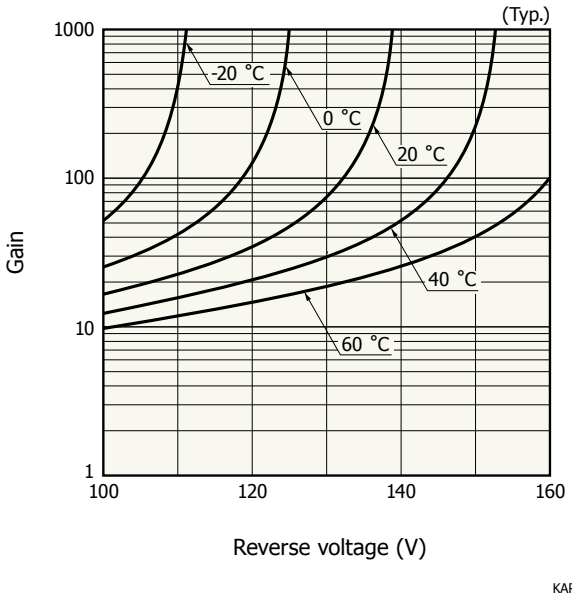
KAPDB0091EB

[Figure 1-7] Plot of APD gain vs. reverse bias voltage



KAPDB0088EA

[Figure 1-8] Effects of temperature on gain vs. reverse voltage



Amongst the above APD gain characteristics, a peculiarly curious case is the gain's dependence on wavelength; to see why, we should note that an APD can be fabricated in two alternative structures: N-on-P and P-on-N. The N-on-P structure offers greater gain and enhanced photosensitivity in response to longer wavelengths (at the expense of shorter ones) while the P-on-N structure has higher gain and photosensitivity in response to shorter wavelengths (at the expense of longer ones). This difference in structure also pertains to the MPPC as well but only affects its photosensitivity; we will revisit and explain these structures as part of our discussion of the MPPC.

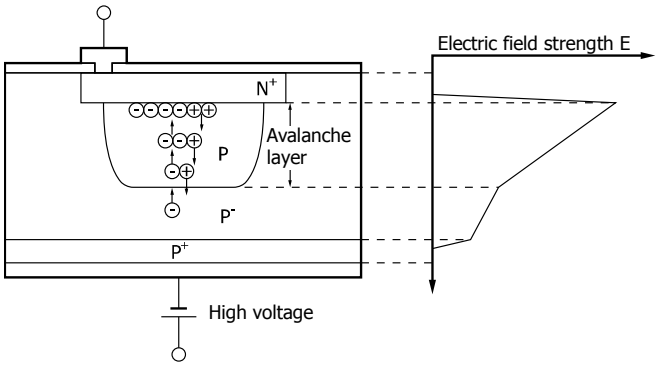
Another particularly interesting characteristic is APD gain's dependence on applied reverse voltage as shown in Figure 1-7. Based on the two knees of the plot in Figure 1-7 and the three resulting gain vs. reverse voltage regions, APD gain demonstrates three distinct behaviors with respect to reverse voltage. In the example plot shown in Figure 1-7, those distinct behaviors are (A) below about 50 V, (B) between around 50 V and about 100 V, and (C) above 100 V. For the sake of our discussion, we will address the latter two. By fitting a trend line, we observe that gain approximately has a $10^{\alpha V}$ relationship with reverse voltage under the moderate electric field intensities of region (B) but increases very steeply with reverse voltage in region (C). The steep slope of region (C) poses a concern to the use of the APD in photometric applications, since it raises the question of whether the APD's internal gain can be reliably controlled and accurately known in region (C) or not; a stable and reliably controllable gain is essential to a deterministic relationship between input and output of a detector used in a measurement. In other words, accurate determinability of APD gain is critical to its application in high-accuracy measurements. Moreover, a stable gain is a fundamental aspect of maintaining the APD's overall response linearity.

We thus proceed to demonstrate that the APD's avalanche gain response is theoretically linear. Although a comprehensive mathematical formulation of the avalanche process is quite complex, to illustrate that linearity in the simplest way possible, let's say a single photon generates an initial photoelectron and hole pair that mutually succeeds at impact ionization and creating two electron-hole pairs. In our simplified example, if electrons and holes have the same fixed ionization probability of P to generate only one electron-hole pair per ionization, we can characterize the avalanche process as forming $2 + 2P + 2P^2 + 2P^3 + \ldots + 2P^n = 2 \cdot \sum_{i=0}^{n} P^i = 2 \cdot \left(\frac{1}{1-P}\right)$ (if n $\rightarrow \infty$) avalanching carriers in average where i = 0, 1, 2, 3, … represents the sequential iterations of carrier multiplication. Now, if we increase the initial number of photoelectrons and holes triggering the avalanche to $2m$ (created by $m > 1$ detected photons), the count of avalanching carriers after $n \rightarrow \infty$ iterations of carrier

multiplication would be $2m + 2mP + 2mP^2 + 2mP^3 + \ldots + 2mP^n = 2m \cdot \sum_{i=0}^{n} P^i = 2m \cdot \left(\frac{1}{1-P}\right)$. We thus observe that the ratio between the two final avalanche populations is *m*, and therefore, there exists a linear relationship of proportionality by a factor of *m* between the two final populations (one of which was originated by 1 photon and the other by *m* photons).

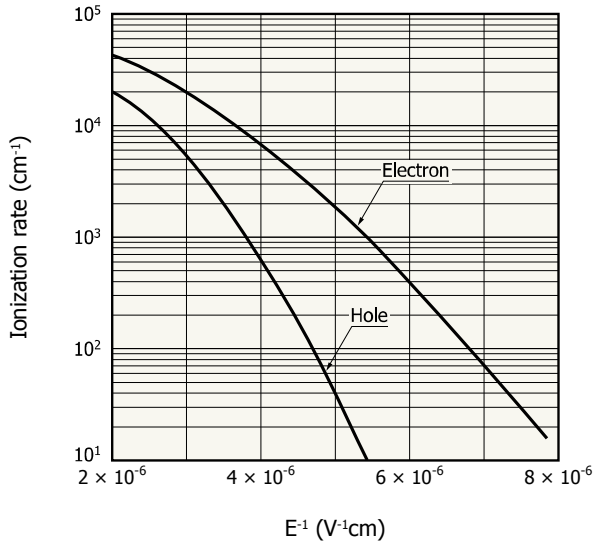
Another noteworthy conclusion from the above simplistic model is that the APD's gain (as the ratio of the avalanche's final carrier population to its initial carrier population) is $\left(\frac{1}{1-P}\right)$, which means that if $P \rightarrow 1$ under a sufficiently-intense E-field, the APD's gain approaches infinity. Nonetheless, the APD nature is actually more complex in nature than our simplistic model above: avalanching electrons and holes ionize atoms at differing probabilities under moderate E-field intensities; those differing ionization probabilities increase significantly and converge as the electric field intensifies to higher levels. Without delving into its complicated mathematical derivation, the APD's gain as a carrier multiplication factor can be formulated as $M = \frac{(\alpha_e - \alpha_h) \cdot e^{d \cdot (\alpha_e - \alpha_h)}}{\alpha_e - \alpha_h \cdot e^{d \cdot (\alpha_e - \alpha_h)}} = \frac{1-k}{e^{(k-1) \cdot \alpha_e \cdot d} - k}$ in which α_e and α_h are the ionization efficiencies (probabilities per unit of travelled distance) of avalanching electrons and holes, *k* is the ratio of ionization efficiency of a hole to that of an electron ($k = \frac{\alpha_h}{\alpha_e}$), and d is the thickness of the avalanche layer depicted in Figure 1-9; it is the distance within the APD's depletion layer over which the E-field is strong enough for the avalanche process to occur. Plots of ionization efficiencies of electron and holes and their dependencies on E-field intensity within the depletion layer have been shown for silicon in Figure 1-10.

[Figure 1-9] The avalanche region within a Si APD



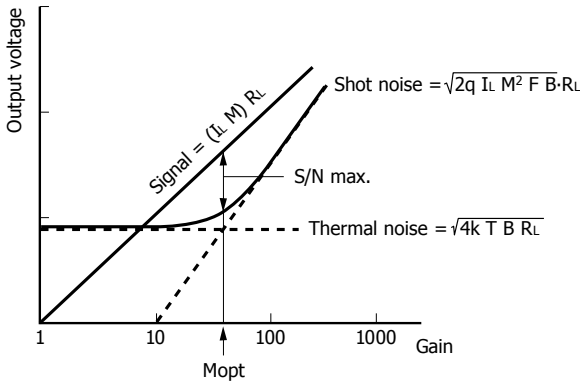
KAPDC0006EC

[Figure 1-10] Ionization efficiencies of electrons and holes drifting under an electric field in silicon at 300 K



KAPDB0382EA

[Figure 1-11] Practical determination of an APD's optimal gain



RL: load resistance
k : Boltzmann's constant
T : absolute temperature

KAPDB0033EC

In our theoretical study of APD gain, three specific scenarios are of interest:
- Optimal APD gain is the highest gain level at which the APD maintains relatively constant bandwidth and stable noise output; theoretically, that optimal gain is $M_{opt} = \frac{\alpha_e}{\alpha_h}$, and assuming $\alpha_e \gg \alpha_h$, it is attained once the condition $\alpha_e \,.\, d \approx -\ln(k) - 0.69$ [2] is met. However, since measurement of carrier ionization efficiencies in a specific APD and monitoring the establishment of the aforementioned condition are not trivial tasks, the optimal gain can instead be determined by the approach depicted in Figure 1-11. APD gain is typically on the order of few 10 up to 100 for optimal operation.

[2] Derivation: $M_{opt} = \frac{(\alpha_e - \alpha_h)\cdot e^{d\cdot(\alpha_e-\alpha_h)}}{\alpha_e - \alpha_h\cdot e^{d\cdot(\alpha_e-\alpha_h)}} = \frac{\alpha_e}{\alpha_h} \mid \alpha_e \gg \alpha_h \Longrightarrow \frac{(\alpha_e)\cdot e^{d\cdot\alpha_e}}{\alpha_e - \alpha_h\cdot e^{d\cdot\alpha_e}} = \frac{\alpha_e}{\alpha_h} \Longrightarrow \alpha_h\cdot e^{d\cdot\alpha_e} = \alpha_e - \alpha_h\cdot e^{d\cdot\alpha_e} \Longrightarrow 2\cdot e^{d\cdot\alpha_e} = \frac{1}{k} \mid k = \frac{\alpha_h}{\alpha_e} \Longrightarrow d\cdot\alpha_e = -\ln(k) - \ln(2) = -\ln(k) - 0.69$

- For $M \rightarrow \infty$, we have $(e^{\alpha_e \cdot d \cdot (k-1)} - k) \rightarrow 0$, which leads to $\alpha_e \cdot d \cdot (k-1) = \ln(k) \implies d \cdot (\alpha_h - \alpha_e) = \ln(k)$. If this theoretical condition $\left[d \cdot (\alpha_h - \alpha_e) = \ln(k)\right]$ is met between α_e and α_h over the range of d, the APD is said to be in *breakdown*; the reverse bias voltage at which this phenomenon ($M \rightarrow \infty$) occurs is referred to as the APD's breakdown voltage.

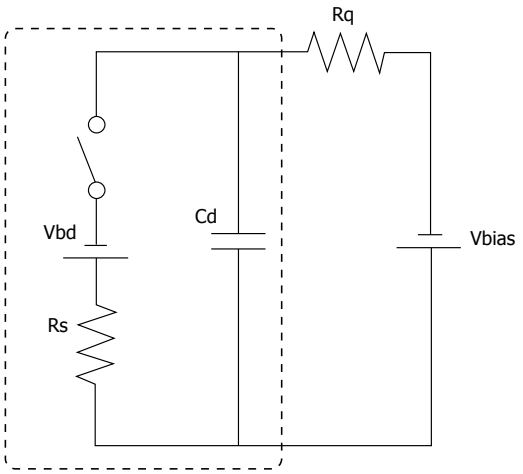
- A practical breakdown scenario occurs after the reverse bias voltage increases to a point that causes $k \rightarrow 1$ and thus results in $M \rightarrow \frac{1}{1-\alpha_e \cdot d}$. Consequently, if the bias voltage (or further increase thereof) results in $\alpha_e \cdot d \rightarrow 1$, we will have $M \rightarrow \infty$. Based on the data plot in Figure 1-10, $k \rightarrow 1$ occurs when the ionization efficiencies converge to an order of magnitude of around $10^7 m^{-1}$ at an E-field intensity of about $10^8 \frac{V}{m}$ (corresponding to a reverse bias voltage on the order of 100 V across a 10 µm-deep depletion layer).

In technical terminology, operating an APD in the breakdown state (i.e. the applied reverse bias voltage being higher than the APD's breakdown voltage) is referred to as *Geiger-mode* operation.

Now, recall our earlier discussion of Figure 1-7 (APD gain vs. reverse voltage plot) and the issue of APD gain stability and linearity in "operation region (C)" in which APD gain shoots up towards higher and higher values. That issue indeed applies to a Geiger-mode avalanche whose output charge (per avalanche) is the same for a given gain and its corresponding reverse bias voltage regardless of the initial number of photo-carriers that initiated it (contrary to the linear APD response that we described with a simplistic model earlier). This makes an APD operating in the Geiger mode to be a so-called *digital* photodetector (as opposed to being linear or *analog*), since it lacks a linear response between its input (i.e. detected amount of light) and output (i.e. amount of outputted charge); its response is indeed a binary one, corresponding to signal detection or lack thereof.
Since a single photoelectron-hole pair is all that is needed to trigger a Geiger-mode avalanche, APD gain in the Geiger mode for a given bias voltage can be obtained by dividing its output charge by the fundamental charge of an electron. Without delving into its detailed derivation, output charge of a Geiger avalanche is the product of the APD's junction capacitance and the applied operating overvoltage (i.e. difference between the applied reverse bias voltage and the APD's breakdown voltage V_{BR}). Thus, we can formulate the APD's Geiger-mode gain as: $M = \frac{C_j \times (V_{bias} - V_{BR})}{q_{electron}}$. In a typical APD, depending on its surface area, the junction capacitance C_j is on the order of several 10 to a few 100 pF while $q_{electron}$ is 1.6×10^{-19} C. Thus, it can be concluded from this equation and by the large ratio of $\frac{C_j}{q_{electron}}$ that biasing an APD in Geiger-mode requires an ultralow-ripple power supply in order to attain good gain stability.

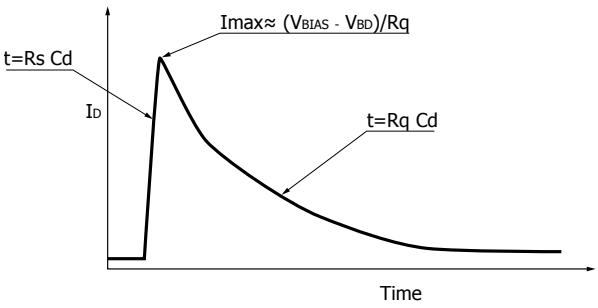
Another peculiar aspect of Geiger-mode operation is the population density of avalanching carriers in the depletion region. In Geiger mode, this density reaches such high levels that the cloud of avalanching carriers behaves like a conductor with very low resistance under a finite bias voltage. This results in a surge of current through the APD that by intense generation (through scattering collisions) and recombination continually mobilizes valence electrons from high energy states susceptible to contributing to the avalanche multiplication process, rendering the APD unable of producing discrete output pulses. It is thus necessary to quench that current surge once sufficient gain and adequate output signal amplitude have been attained so that a subsequent photoelectron event can lead to a succeeding output pulse. Figure 1-12 shows the equivalent circuit of a Geiger-mode APD (GAPD) and can be used to describe how that quenching is performed.

[Figure 1-12] Equivalent circuit of a Geiger-mode APD (GAPD)



KAPDC0073EA

[Figure 1-13] Conceptual output pulse of the equivalent circuit



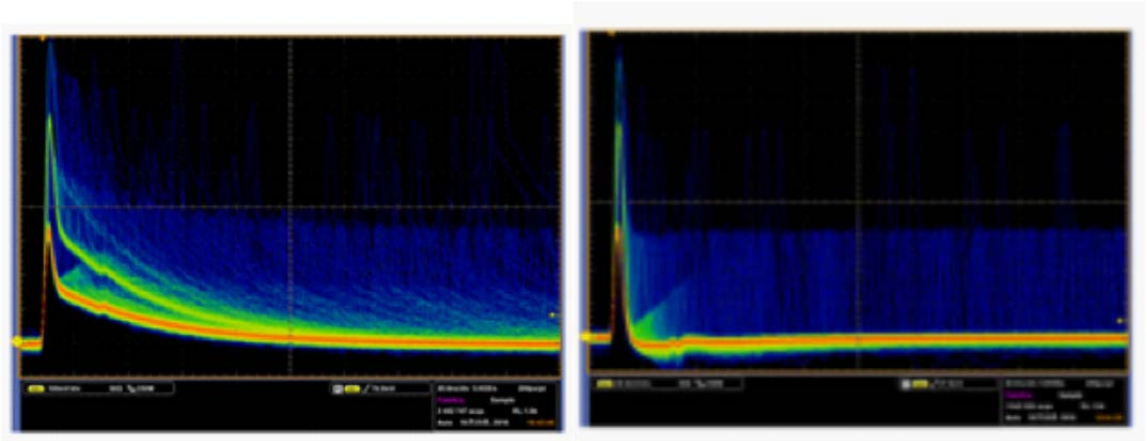
KAPDC0074EA

In this equivalent circuit, the GAPD capacitance (C_d) is initially biased at V_{bias} while the conceptual switch shown is open. Once an electron-hole pair is generated within the depletion layer (whether thermally or by the photoelectric effect), the conceptual switch closes, and C_d begins to discharge through GAPD's series resistance (R_S), whose small value (diminished by the avalanching carrier population) provides for a surge in current flow while the potential difference across C_d (let's call it V_d) exponentially decays towards the breakdown voltage (V_{bd}). This decline in V_d weakens the avalanche process, increasing R_S and decreasing C_d's discharge current flow. Based on the GAPD equivalent circuit's two current loops, the following relationship can be established to describe the net current flow through C_d: $[(V_d - V_{bd}) / R_S] + [(V_{bias} - V_d) / R_q] = I_d$. Since the current flows in the two loops would be in opposite directions (one entering C_d to recharge it and another leaving C_d to discharge it), net current I_d would flow in the direction of the stronger current flow. Thus, if at point where V_d is such that $V_d - V_{bd} = V_{bias} - V_d$, the quenching resistor R_q is so small that a recharge current continues to flow into C_d and sustains its discharge through R_S, the quenching process shall not take place. However, if R_q is large enough so that the current flow from V_{bias} cannot sustain C_d's discharge, the avalanche process would be quenched. Thus, once the discharge is exhausted and reaches its trough, the avalanche process is quenched and the conceptual switch opens. The recharge current flowing into C_d through R_q increases the voltage across C_d by $(V_{bias} - V_{bd})$ to equal V_{bias}, preparing the GAPD for the next avalanche. It should be noted that R_q values large enough to facilitate the quenching process are generally determined empirically.

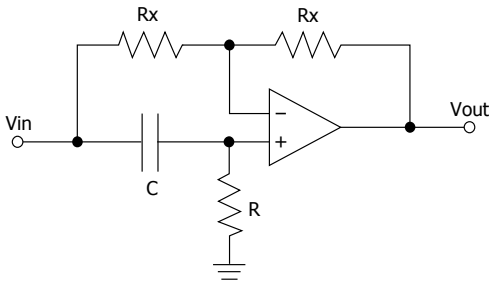
The recharge process of C_d is commonly called *recovery* and its time duration is generally characterized by the time constant $R_q \cdot C_d$. During the brief period of avalanche multiplication (corresponding to the fast time constant

of $R_S \cdot C_d$) and the subsequent recovery time, a GAPD is rather unavailable to detect a new photoelectric event; any output pulse meanwhile generated would practically have a small amplitude (depending on the remaining amount of C_d's stored charge that is available for a secondary discharge). To decrease the GAPD's recovery time [Figure 1-14], its net capacitance can be reduced by theoretically introducing a smaller capacitance in series with C_d as part of the readout scheme; the resulting decrease in recovery time comes with a tradeoff of lower overall gain due to a reduced net capacitance (since gain is proportional to GAPD's capacitance as described earlier). From a practical circuit design standpoint, this theoretical arrangement can be implemented by introducing a high-pass filter with a RC time constant desirably shorter than GAPD's recovery time at the input of the amplifier circuit used for the GAPD's readout [Figure 1-15]. The amplifier output pulse is thus shaped such that net overall gain loss (of the GAPD and the amplifier circuit combined after the introduction of the RC filter) is $\left(1 - \frac{\frac{f_{GAPD}}{f_c}}{\sqrt{1+\left(\frac{f_{GAPD}}{f_c}\right)^2}}\right)$ in which f_c is the filter's cutoff frequency $\left(\frac{1}{2\pi R_{filter} \cdot C_{filter}}\right)$ and f_{GAPD} is the frequency bandwidth of GAPD's recovery $\left(f_{GAPD} = \frac{1}{2\pi R_q \cdot C_d}\right)$ while the amplifier output pulse's fall time (90% to 10% of amplitude) can be obtained as the RMS of the amplifier's and the GAPD's fall times to be $\left(2.2 \times \sqrt{R_{filter}^2 \cdot C_{filter}^2 + R_q^2 \cdot C_d^2}\right)$ or $\left(0.35 \times \sqrt{\frac{1}{f_c^2 + f_{GAPD}^2}}\right)$ considering that $BW = \frac{0.35}{rise/fall\ time}$ and that fall time of an RC circuit is approx. 2.2 times its time constant.

[Figure 1-14] Shaping the output pulse of a GAPD in order to decrease its recovery time (MPPC: 3 × 3 mm)
(a) Before (b) After



[Figure 1-15] Schematic of a V/V high-pass amplifier circuit



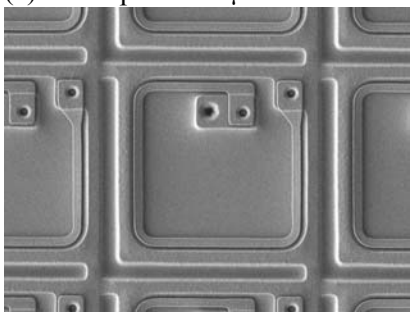
The exceedingly high levels of avalanche gain (on the order of several 10^5 to 10^6) in Geiger mode make this APD operation regime particularly interesting to applications with low light levels; specialized GAPDs have been developed for utilization in such applications. However, in order to overcome GAPD's unavailability for photoelectron detection during its recovery time and to address its lack of a linear response, a far more versatile category of silicon photodetectors has been developed based on the GAPD concept by arranging a matrix of them within the same field-of-view of the incident light signal. We will now proceed to learn about the multi-pixel photon counter or MPPC.

1-3. MPPC (multi-pixel photon counter)

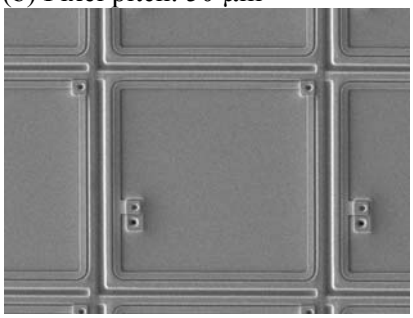
Also known as Silicon Photomultiplier or SiPM, the MPPC is a common-bias and common-output (common-cathode) matrix of GAPD elements (called pixels or microcells) connected in parallel and fabricated on a monolithic silicon crystal. Figures 1-16, 1-17, and 1-18 illustrate the layout of MPPC pixels.

[Figure 1-16] Individual MPPC pixels (microcells) with a metal-composite quenching resistor fabricated around each microcell

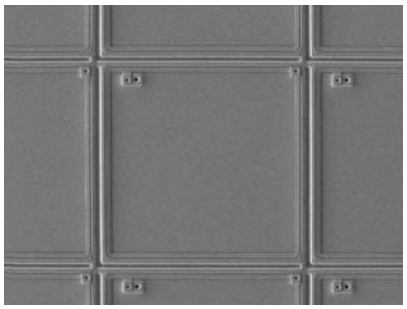
(a) Pixel pitch: 25 µm



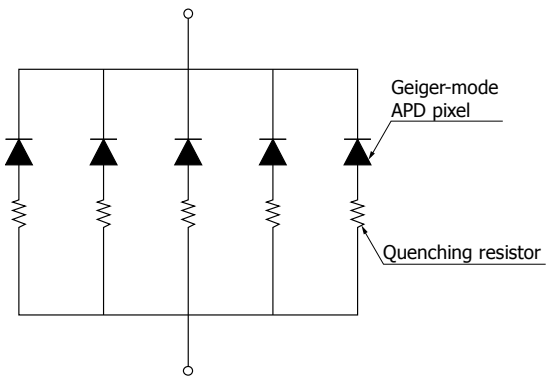
(b) Pixel pitch: 50 µm



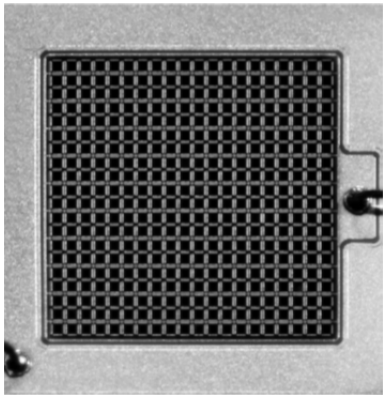
(c) Pixel pitch: 75 µm



[Figure 1-17] Conceptual illustration of the MPPC as a matrix of GAPD pixels (microcells) connected in parallel



[Figure 1-18] An actual matrix implementation of MPPC microcells



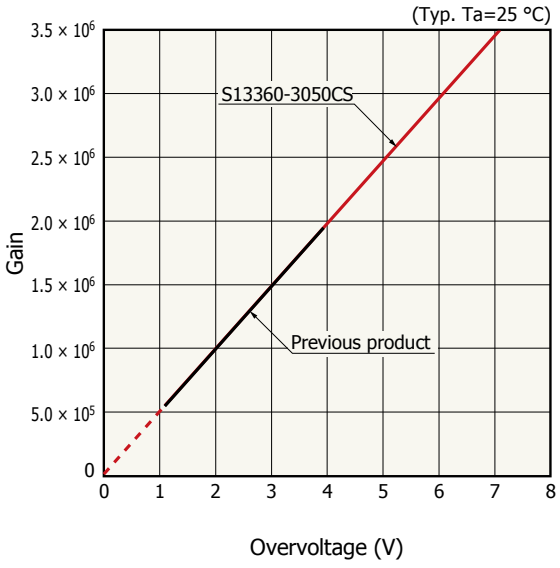
KAPDC0049EA

Similar to the APD, a MPPC pixel can be fabricated based on two distinct structures: N-on-P (for enhanced red/NIR photosensitivity) and P-on-N (for enhanced UV/blue photosensitivity).

The difference in photosensitivity of these structures is caused by how much farther an electron, which has greater ionization efficiency and hence avalanche probability in silicon than a hole, must travel within the depletion layer before being collected. In the N-on-P structure, longer light wavelengths create an electron-hole pair in the P region due to such wavelengths' greater absorption depths in silicon. On the other hand, in the P-on-N structure, shorter wavelengths create an electron-hole pair also in the P region due to such wavelengths' shallower absorption depths in silicon. As explained in our discussion of a PN photodiode's operation, a photoelectron generated in the depletion layer's P side is swept across the PN junction in the opposite direction of the depletion layer's E-field. Prior to collection, that electron would consequently face a longer pathlength, which, combined with its higher ionization efficiency and avalanche probability, leads to a greater likelihood of avalanche initiation and hence detection.

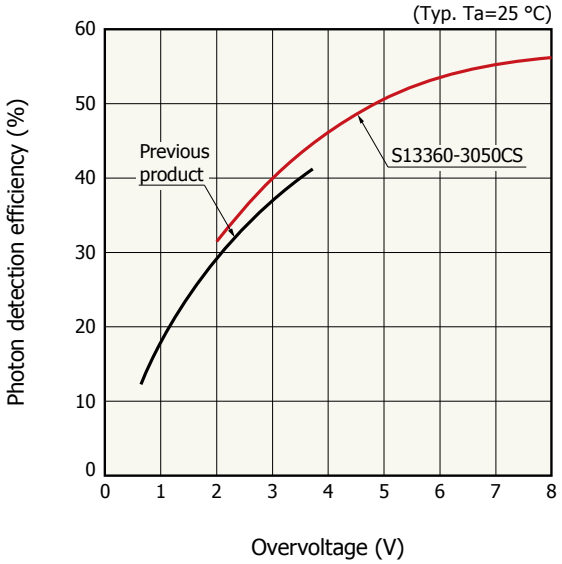
The MPPC has a set of fundamental characteristics that are shown in the plots of Figures 1-19 to 1-22.

[Figure 1-19] Gain vs. overvoltage (S13360-3050CS)



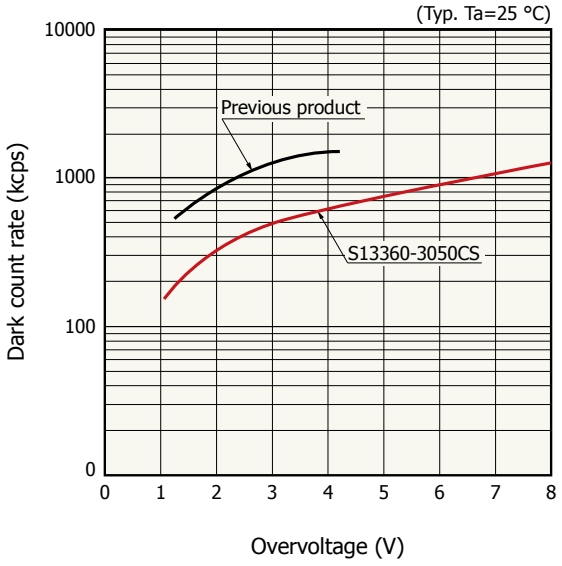
KAPDB0307EB

[Figure 1-20] PDE vs. overvoltage (S13360-3050CS)



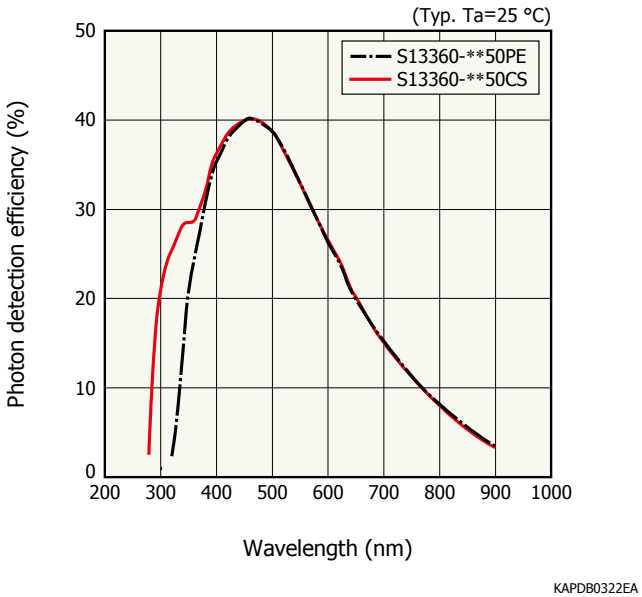
KAPDB0308EB

[Figure 1-21] Dark count rate vs. overvoltage (S13360-3050CS)



KAPDB0310EB

[Figure 1-22] PDE vs. wavelength (S13360 series)



Some MPPC characteristics are dependent on temperature. Of particular note are the temperature dependencies of photon detection efficiency (PDE), gain, and dark count rate (DCR). MPPC's PDE is the product of silicon's quantum efficiency (QE) times MPPC pixel fill-factor (ratio of pixel photosensitive area to total area) times Geiger-mode avalanche probability. Silicon's QE increases at elevated temperatures due to increased phonon vibrations: those vibrations facilitate the transition of electrons to the conduction band and the formation of holes in the valence band and thus improve silicon's photosensitivity (although, that facilitation pertains not only to photo-carriers but also to thermal carriers that constitute a silicon photodetector's dark current). However, avalanche probability has a decreasing relationship with increased temperature, countering QE's increasing relationship with temperature. As a result, MPPC's PDE is affected by the more dominant of these opposing factors as temperature changes and could even remain stable over some ranges of temperature.

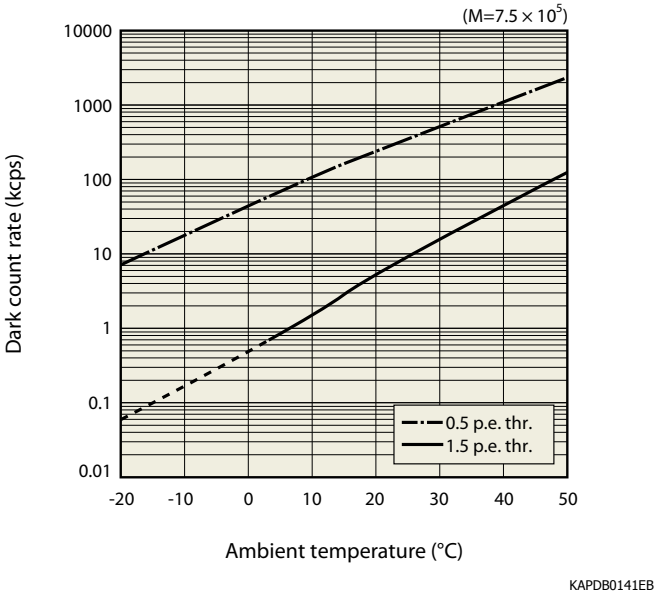
Like in case of the APD, in absence of temperature control, the MPPC's biasing scheme requires a temperature-compensation circuit that would adjust the applied bias voltage with changes in temperature in order to maintain a constant MPPC gain. Also, like in the case of the APD, MPPC gain declines with increases in temperature due to increased phonon vibrations at higher temperatures and thus greater scattering collisions and losses in kinetic energies of avalanching carriers; based on our earlier definition of GAPD gain, this decline in gain stems from an increasing breakdown voltage as temperature increases.

Being a noise that occurs randomly and uncorrelated with photon-initiated output pulses, DCR is detrimental to utilizing the MPPC for single-photon counting applications. However, if the time window during which single photon events are most likely to occur is sufficiently narrow and can be known to the detector system by availability of an external trigger, DCR is no longer detrimental to single-photon counting with the MPPC. For example, if the time window in question is 10 ns in duration, the likelihood of a single dark count at DCR = 1000 kcps (cps: counts per second) occurring during that time window is fairly slim, and it therefore becomes possible to detect single photons using the MPPC with a high degree of statistical confidence. Nevertheless, if the above timing constraints on single-photon events cannot be met, DCR can be reduced by thermoelectric cooling of the MPPC at a rate of about $\left(\frac{1}{2}\right)$ for every approx. 10 ºC reduction in temperature as shown in Figure 1-23.

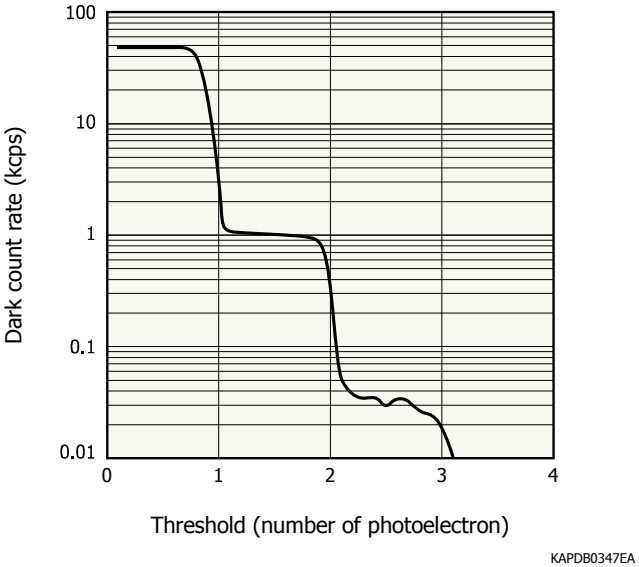
On the other hand, in applications with several photons or more per light event, DCR can be excluded from readout by increasing the discriminator level. As a general rule of thumb, for every one single-photoelectron (1

p.e.) output pulse height increase in the discriminator's threshold level, DCR readout diminishes by about an order of magnitude as shown in Figure 1-24.

[Figure 1-23] A typical MPPC's DCR vs. temperature



[Figure 1-24] DCR vs. counting discriminator threshold level

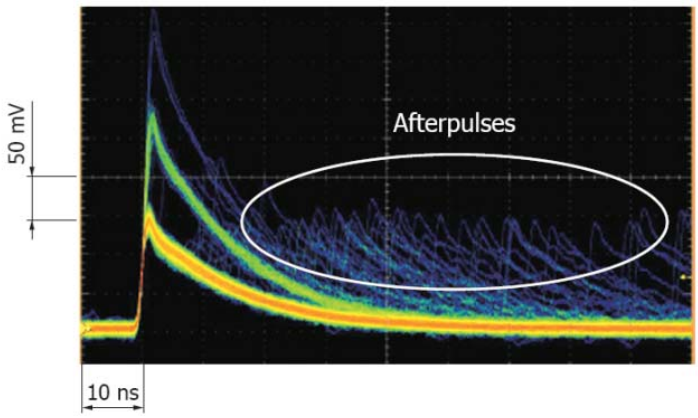


In addition to DCR, MPPC operation has two peculiar types of noise that are correlated in occurrence with detection of photon signal and deserve individual discussion:

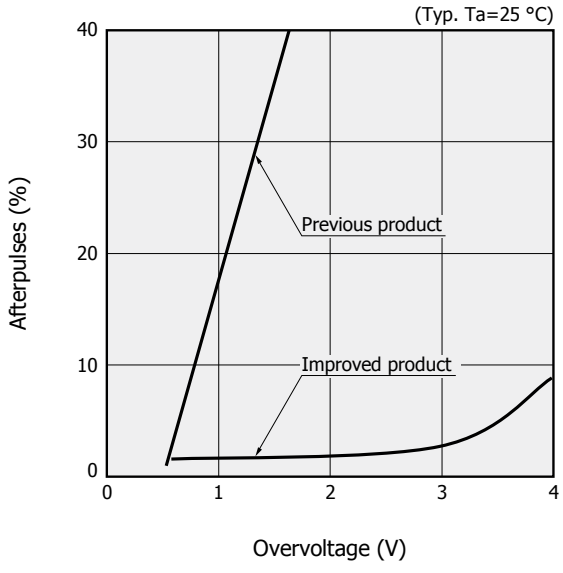
i. After-pulses: During an avalanche process (whether triggered by a photoelectron or a thermal carrier), a small portion of avalanching carriers get trapped in impurity energy levels but are released after short delays (typically on the order of few 1-10 ns) upon receiving the required energy (albeit small) to reenter the conduction or valence band. Upon their release, these carriers initiate new avalanche pulses, which appear with delays after the genuine parent pulse and are hence referred to as *afterpulses*. Please see Figure 1-27 for a waveform showing strong after-pulsing. If an afterpulse is released during the MPPC pixel's recovery time, its amplitude will be shorter than that of a regular (1 p.e.) pulse and can thus be discriminated against by pulse-height analysis (PHA) and excluded

from data processing. However, if the afterpulse is released after the MPPC pixel's recovery, it will have a full (1 p.e.) pulse amplitude and will be indistinguishable in amplitude or shape from a genuine output pulse. In ultralow DC light level applications where consecutive single-photon events in close timing proximity are unlikely or in ultralow-rate pulsed applications (where time differences between succeeding pulses are large) or when the timing of a pulse can be known by availability of a trigger (regardless of what the pulse rate might be), a time-delay filtering algorithm could be used to exclude full-height afterpulses. Nevertheless, Hamamatsu MPPCs have greatly diminished levels of after-pulsing as shown in Figure 1-26.

[Figuer 1-25] Waveform showing the occurrence of afterpulses



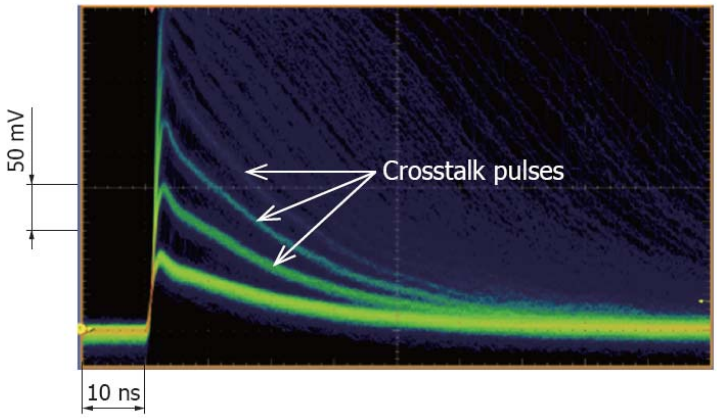
[Figuer 1-26] Much reduced afterpulsing of Hamamatsu MPPC



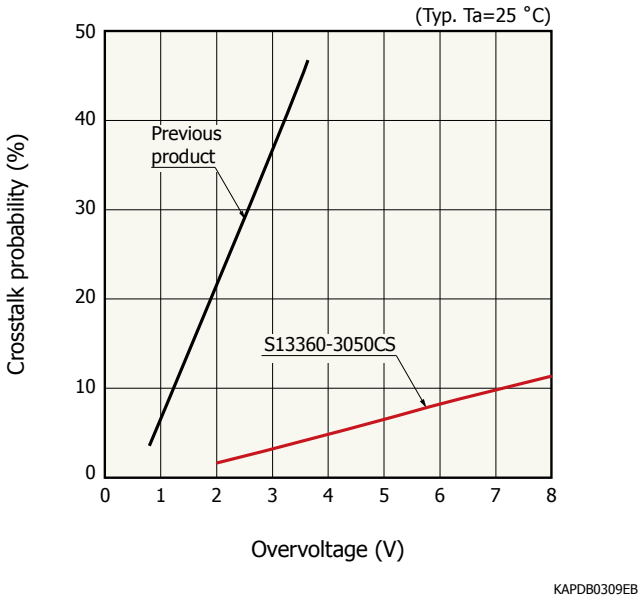
ii. Optical crosstalk: During the avalanche process, kinetic energy of avalanching carriers (even if larger than silicon's band gap energy) need not necessarily contribute to carrier multiplication. As discussed earlier, through scattering collisions, a portion of that energy is lost as heat to phonon vibrations. In a less likely phenomenon, it is also possible for that energy to be emitted as photons. When that occurs, the resulting photons can travel to the neighboring MPPC pixels and initiate avalanches in them. This undesirable phenomenon is referred to as *optical crosstalk*. If a crosstalk photon produces an electron-hole pair in the depletion layer of a neighboring pixel and triggers an avalanche process as a result, this situation is called *prompt* crosstalk, whose output pulse appears simultaneously with the genuine original pulse. Please see Figure 1-27 for a waveform showing strong prompt optical crosstalk. With the exception of using an upper-level discriminator in single-photon counting, there is

unfortunately no way to discriminate against prompt crosstalk pulses in order to exclude them from data processing. However, in the alternative scenario that the crosstalk photon creates an electron-hole pair outside an MPPC pixel's depletion layer, the resulting electron-hole pair can only trigger an avalanche process after successfully (albeit improbably) reaching the depletion layer by diffusion. This introduces a delay between the original pulse and its crosstalk noise, which is called *delayed* crosstalk. Delayed crosstalk can be excluded by the same approach as for excluding full-height afterpulses.

[Figure 1-27] Waveform showing the occurrence of prompt crosstalk



[Figure 1-28] Much reduced crosstalk of Hamamatsu MPPC



Note: Dark pulse overlap (pileup effect) is eliminated.

Consequently, (prompt) crosstalk is a particularly detrimental source of noise to applications with ultralow-level light signals (such as multi-photon counting). As shown in Figure 1-28, however, Hamamatsu MPPCs enjoy substantially diminished levels of crosstalk, thanks to optical trenches that are implemented around every MPPC pixel in order to block crosstalk photons from reaching neighboring pixels. Furthermore, considering that optical crosstalk is correlated with signal detection and increases with gain (as an indicator of the intensity of the avalanche process and its byproduct photons), operation of the MPPC at a lower gain while maintaining a higher overvoltage and hence PDE is essential to overcoming the detrimental impact of crosstalk in low-signal applications. Since gain is directly proportional to both overvoltage and also MPPC pixel capacitance while PDE

takes no effect from pixel capacitance, it becomes of paramount importance to select MPPCs with low pixel capacitance for low-signal applications.

Our discussion of characteristics of silicon photodetectors, including the MPPC, has been qualitative so far. However, in order to select suitable photodetectors for practical real-world applications, it is necessary to develop a quantitative framework to assess photodetector performance under an application's conditions. In the following sections 2 and 3, we will overview generalized quantitative methods for selection of silicon APD and MPPC products for a broad range of applications. A primary objective has been to describe methods that could be used for any application as long as certain basic pieces of information are available about the application; these methods could be used for photodetector selection consistently, regardless of the particular application involved. However, it is important to note that the methods described are meant to serve as preliminary and general guidelines (so-called *back-of-the-envelope* calculations); they are intended to serve as early indicators of what product(s) to begin considering. As the design and development works advance, these methods would not substitute extensive simulation and comprehensive evaluation under the application's specific conditions. With this understanding, let's now proceed to discussing preliminary calculation methods for assessing APD and MPPC performance.

2. APD & MPPC performance parameters

Each and every application that involves detection of light (regardless of its source: laser, LED, lamp, scintillation, different luminescence effects, etc.) can be defined and characterized by the following parameters:

2-1. Signal

An optical application's input signal level is simply the amount of light that is to be detected. The input light signal has a spectral distribution, which is typically represented in simplistic calculations (such as the methods to be described here) by the peak wavelength.
In quantifying the amount of input signal, the dimension or unit of measurement is essential for making proper calculations. The two dimensions that could be used by the methods described in this manual are Watts [W] and number of photons. Note that the former is normalized to time while the latter is not (although the latter could be, and both can be normalized to illumination area). A signal level S expressed in either of these units can be converted to the other by the following formula:

$$S\,[photons] = S\,[W] \cdot \lambda \cdot T / (h \cdot c) \qquad \textbf{(2-1)}$$

in which λ is the light wavelength of interest, T is illumination time for one measurement reading,
h is Planck's constant, and c is the speed of light. There are other units (lm, lx, etc.) used to express the amount of light present, but those units require tabulated reference data or complex calculations for their proper use, and hence, they are not discussed here as they exceed the intended scope of our discussion. Input light signal can be converted to output signal of a photodetector by the following relationship:

$$S_{output}[e-] = \left(S_{input}[photons] \cdot QE \cdot CE \cdot M\right) + S_{dark}[e-] \qquad \textbf{(2-2)}$$

in which S_{dark} is the detector output charge that is not generated as a result of the photoelectric effect during the measurement, M is gain, and CE is collection efficiency.
Often, Photon Detection Efficiency (as defined by $PDE = QE \times CE$) is used as part of equation 2-2; in the case of MPPC, CE = *avalanche probability* × *fill factor* (in which fill-factor is the ratio of a MPPC pixel's photosensitive area to the pixel's total area). In the case of a fully-depleted APD, CE can be assumed to 1.

2-2. Noise

The intrinsic uncertainty or random fluctuation in a measured signal is noise. For an optical signal, noise is characterized through a histogram (a so-called pulse height distribution or PHD) of incidents of detected light pulses with varying heights (corresponding to differing counts of photons); the resulting histogram could be closely fitted into a profile similar to that of a Poisson probability distribution. That leads us to conclude that noise characteristics of a photon signal can be modeled by the following Poisson probability distribution function and its corresponding mean (μ) and standard deviation (σ) with $\boldsymbol{m}$ being the expected number of detected photons on average:

$$\boldsymbol{P_X(k) = \frac{m^k \cdot e^{-m}}{k!}} \ \boldsymbol{, \mu = m \ , \sigma = \sqrt{m}} \qquad \textbf{(2-3)}$$

To model the detection of light using the above probability model, the standard deviation is considered a measure of randomness or uncertainty (i.e. noise) of the Poisson random variable (i.e. photon signal), and the mean is the expected value of the signal. In other words, intrinsic noise of a light signal is described by the square root of its mean. A photodetector's dark output also has a Poisson probability distribution. A noise whose random behavior can be characterized by a Poisson distribution is often referred to as *shot* noise.

2-3. Signal-to-noise ratio (S/N)

As its name suggests, it is the ratio of signal to noise as calculated for a detector's output. In light detection applications with unity gain, it is fundamentally defined by:

$$\boldsymbol{S/N = \frac{S_{input} \times PDE}{\sqrt{n^2_{photon\ shot} + n^2_{dark\ shot} + n^2_{readout}}}} \qquad \textbf{(2-4)}$$

in which $\boldsymbol{n_{photon\ shot}}$ is the photon shot noise, $\boldsymbol{n_{dark\ shot}}$ is the dark shot noise, and $\boldsymbol{n_{readout}}$ is the readout noise generated by the output amplifier circuit (at its bandwidth frequency). More practically, considering that APDs are typically read out in analog (linear) mode, there are two S/N equations for APDs:

i. when read out by resistive trans-impedance amplifiers for relative measurements:

$$\boldsymbol{S/N = \frac{\Phi^2 \cdot S^2_{input} \cdot M^2}{2 \cdot q \cdot \Delta f \cdot M^2 \cdot F \cdot [(\Phi \cdot S_{input}) + S_{dark}] + \left(2 \cdot q \cdot \Delta f \cdot \frac{S_{dark}}{M}\right) + JTN + n_{readout}}} \qquad \textbf{(2-5)}$$

in which $\boldsymbol{q}$ is the fundamental electron charge, $\boldsymbol{\Delta f}$ is readout amplifier bandwidth, $\boldsymbol{JTN}$ is Johnson thermal noise, and $\boldsymbol{\Phi = QE \cdot \lambda \,/\, 124000}$ is APD unity-gain photosensitivity [A/W] in whose calculation $\boldsymbol{QE}$ is a percentage and $\boldsymbol{\lambda}$ is wavelength [nm]. $\boldsymbol{JTN = \frac{4kT \cdot \Delta f}{R_L}}$ is a Gaussian noise component that originates from thermal generation of current in the APD's load resistor (even without any external voltage applied across it). In $\boldsymbol{JTN}$'s calculation, $\boldsymbol{R_L}$ = 50 Ω is typically assumed, and thus, it can be simplified to $\boldsymbol{JTN} \approx 3.3 \times 10^{-22}$ C × $\boldsymbol{\Delta f}$ or $2 \times 10^{-3} e^- \times \boldsymbol{\Delta f}$ for room temperature and could hence be ignored for relatively low frequencies or if other noise factors or the signal are

comparatively large. Additionally, considering that $\left(2 \cdot q \cdot \Delta f \cdot \frac{S_{dark}}{M}\right)$ tends to be small compared to other components of the denominator, equation (2-5) can be further simplified to:

$$S/N = \frac{\Phi^2 \cdot S_{input}^2 \cdot M^2}{2 \cdot q \cdot \Delta f \cdot M^2 \cdot F \cdot [(\Phi \cdot S_{input}) + S_{dark}] + n_{readout}} \qquad \textbf{(2-6)}$$

in equations (2-5) (2-6), S_{input} is in the unit of W, and S_{dark} is in the unit of A and includes the multiplication effect of APD gain.

ii. when read out by capacitive trans-impedance amplifiers (a.k.a. charge amplifiers) for photometric or absolute measurements:

$$S/N = \frac{QE \cdot M \cdot S_{input}}{\sqrt{F \cdot [(QE \cdot M^2 \cdot S_{input}) + (M \cdot S_{dark})] + n_{readout}^2}} \qquad \textbf{(2-7)}$$

in which M is gain and $F = M^x$ is APD excess noise factor for whose calculation x is provided as excess noise index in Hamamatsu APD datasheets (for a certain illumination wavelength but yet a reasonable generalized estimation for our purpose). In equation 2-7, S_{input} is in the unit of photons, and S_{dark} is in the unit of electrons.

It is noteworthy that both equation (2-5) and equation (2-7) become applicable to PN and PIN photodiodes by setting M = 1 and F = 1.

It is also noteworthy to mention that since the inverse of the integration time of a charge amplifier would represent its max. sampling rate, and considering that measurement bandwidth would be half of the sampling rate based on Nyquist's theorem, the denominator of equation (2-7) lacks the factor 2 that exists in the denominator of equations (2-5) (2-6) as both S_{input} and S_{dark} would be accumulated over the charge amplifier's integration time.

In case of MPPC, considering the binary nature of the readout scheme in photon counting, readout noise is forgone, and the following equation[3] would be used per measurement reading:

$$SNR = \frac{N_{photon} \times PDE}{\sqrt{(N_{photon} \times PDE) + N_{dark}}} = \frac{N_{total}}{\sqrt{N_{total} + 2 \cdot N_{dark}}} \qquad \textbf{(2-8)}$$

in which N_{photon} is the number of photons incident onto the MPPC, $(N_{photon} \times PDE)$ is the number of photoelectrons detected and N_{dark} is the MPPC dark (to 1 p.e. in height under no illumination) output pulse count during a measurement. The latter portion of equation (2-8) is particularly useful in experimental determination of photon-counting S/N with N_{total} being obtained from dividing the MPPC's total output charge

[3] The derivation of the denominator of equation (2-8)'s second half is often a point of curiosity. If we define (T – D) as a random variable to be the photo-signal (i.e. Dark-subtracted Total output signal measured under illumination), then its variance would be: $\sigma_{T-D}^2 = \sigma_T^2 + \sigma_D^2 - \sigma_{TD}$. Since total output signal and dark signal are uncorrelated, we have: $\sigma_{TD} = 0$. Thus, RMS combination of shot noises of dark signal and photo-signal becomes: $\sqrt{\sigma_{T-D}^2 + \sigma_D^2} = \sqrt{\sigma_T^2 + \sigma_D^2 + \sigma_D^2} = \sqrt{\sigma_T^2 + 2\sigma_D^2}$

by the amount of output charge corresponding to 1 p.e. pulse height. As we will discuss in section 4, these counts would be calculated by performing PHD analysis on integrated MPPC output pulse data. Keep in mind that equation (2-8) is simplistic as it assumes no correlated noise, but considering the greatly diminished levels of crosstalk and after-pulsing in Hamamatsu's MPPCs (down to few %) as mentioned in the previous section, it is a practical approximation. Please bear in mind that equations (2-4) to (2-8) assume that the signal to be measured is the only light flux incident on the detector (i.e. no background light).

2-4. Linearity

The extent to which the output of a photodetector has a linear relationship with its input (as defined by $f(x) = b.x + c$ in which b and c are real constants) is the measure of a photodetector's linearity and is fundamentally defined by:

$$Linearity = \frac{\frac{A_{output}(t_2) - A_{output}(t_1)}{A_{output}(t_1)}}{\frac{A_{input}(t_2) - A_{input}(t_1)}{A_{input}(t_1)}} \mid t_2 > t_1 \qquad \textbf{(2-9)}$$

in which A is signal amplitude and t represents passage of time. If the ratio of the 2 relative changes is < 1, nonlinearity exists. For practical purposes, nonlinearity is typically of greater interest than linearity itself:

$$Nonlinearity\ [\%] = 100 - Linearity\ [\%] \qquad \textbf{(2-10)}$$

Since ideal photodetector response is theoretically linear (within the constraints of S/N > 1 and up to saturation), linearity can be practically obtained by:

$$Linearity\ [\%] = |Real\ Output / Ideal\ Output| \times 100 \qquad \textbf{(2-11)}$$

In terms of detector response, there are 3 particular definitions of linearity:

- DC linearity: This is a measure of how mean output of a detector changes linearly with respect to changes in average light input over a period of time. Understanding the concept of '*DC*' linearity within the context of an average over a single measurement's time duration is important: linearity can be feasibly assessed by relying on data points of mean response (integrated during each measurement reading and thus averaged per its time duration) for signals with randomly fluctuating amplitudes or nonharmonic repetitions.

- Pulse height (or amplitude) linearity: This is the relative extent by which a change in an input light pulse's amplitude results in a change in the photodetector's output. For APDs, this linearity is limited by saturation effect of the junction capacitance's rate of charge-up and discharge as influenced or further limited by the readout circuitry.

In the case of MPPC, upper linearity limit is characterized by availability of pixels to detect succeeding photons while a portion of the pixel population is recovering from the detection of the preceding photons. While ideal MPPC response is theoretically linear as defined by:

$$N_{fired} = N_{photon} \times PDE \qquad \textbf{(2-12)}$$

MPPC's real response can be predicted with high accuracy by:

$$N_{fired} = N_{pixel} \cdot \left(1 - e^{\frac{-N_{photon} \times PDE}{N_{pixel}}}\right) \mid PW < T_{recovery} \qquad \textbf{(2-13)}$$

$$N_{fired} = \frac{N_{pixel} \times PW}{T_{recovery}} \cdot \left(1 - e^{\frac{-N_{photon} \times PDE}{\frac{N_{pixel} \times PW}{T_{recovery}}}} \right) \mid PW > T_{recovery} \qquad \textbf{(2-14)}$$

in which N_{fired} is the number of MPPC pixels fired (i.e. undergoing Geiger-mode avalanche) by the incident photons, N_{pixel} is the number of MPPC's pixels, N_{photon} is the number of incident photons per light pulse, PW is the width of the incident light pulse, and $T_{recovery}$ is the recovery time of a MPPC pixel.

- Pulse rate linearity (Detector bandwidth): This is a measure of a photodetector's pulse height linearity as a function of input signal pulse rate. Ideally, there must be no dependence; however, in practice, an undesirable effect known as *pulse pileup* occurs with output pulse heights (i.e. output signal level differences between peaks and valleys of pulses) decreasing as frequency increases to exceedingly higher levels.

For a fixed readout impedance, pulse rate linearity is limited by a detector's internal capacitance. For a pulsed input signal of fixed amplitude but increasing frequency, bandwidth is defined as the frequency [Hz] at which amplitudes of output pulses decline by a certain amount compared to a DC input signal of the same amplitude. In APD's case, response bandwidth is considered to be limited by a cutoff response frequency defined by:

$$f_c = \frac{1}{2\pi \cdot C_t \cdot R_L} \qquad \textbf{(2-15)}$$

for which R_L = 50 Ω is typically assumed.

Bandwidth is a concern in the design of output amplifier circuits and other readout electronics, considering that a larger amplifier bandwidth allows the passage of a wider spectrum of noise frequency components to the output while bandwidth must be larger than the highest-frequency component of the input light signal in order to allow its proper detection. Thus, the designer of a detector system seeks to select a detector with a cutoff frequency that is by a conservative margin above the highest-frequency component of the signal and then design a readout amplifier circuit whose bandwidth is also by a conservative margin above the highest-frequency component of the signal.
In a pulsed application, if the study of individual light pulses is intended for instance, the signal's highest-frequency component is the product of the constant 0.35 and the inverse of the shortest pulse rise or fall time (10% to 90% or vice versa of amplitude) that is to be measured.

Furthermore, the readout amplifier circuit's cutoff frequency at -3 dB should be designed to be at least twice that of the signal's highest-frequency component whose measurement is desirable (but as a general rule of thumb, 4 times is an advisable design target).

Assessing MPPC pulse-rate linearity is a complex computational effort that is beyond the scope of this manual; however, if the expected time interval between 2 consecutive light pulses will be longer than $T_{recovery}$, the application is within the pulse-rate linearity range of the MPPC. If not, that does not necessarily mean that the application exceeds the MPPC's pulse-rate linearity, but complex simulation or actual experimentation would be needed in order to assess that.

2-5. Dynamic range

DR is typically expressed as a ratio between two levels of the input signal. One level (as numerator of the ratio) is the highest amount of input signal at which the detector maintains its response linearity (i.e. nonlinearity < application's requirement). The other (denominator of the ratio) is the lowest amount of input signal at which the detector behaves linearly.

Over a detector's ***DR***, nonlinearity at the lower limit is typically limited by noise (whether dark or readout noise or a combination thereof), or in other words, the amount of input signal that yields S/N = 1 (often measured and divided by square root of the bandwidth and then specified as *noise-equivalent power* $[\frac{W}{\sqrt{Hz}}]$). On the other hand, nonlinearity at ***DR***'s upper limit is typically caused by saturation effects.

2-6. Time response

As represented by a photodetector's rise and fall times, this is an indicator of how closely the output of a photodetector temporally resembles the shape of its input. That is particularly important for applications in which maintaining the pulse shape integrity of the input signal is desirable for pulse shape discrimination (PSD); in those cases, the detector's rise and fall times must be shorter that rise/fall times of input light pulses. A fully-depleted silicon photodetector's rise time is dominated by carrier drift time within its depletion layer while its fall time is proportional to its capacitance (for a fixed readout impedance).

2-7. Time resolution

The uncertainty that exists in determining the timing of a detected event with respect to a reference point in time (which could be another detected event) is called time resolution. In optical applications, that is the overall uncertainty in timing the detection of an input light pulse and is fundamentally defined by:

$$\sigma_{Overall\ Timing} = \sqrt{\sigma^2_{Pulse\ Timing} + \sigma^2_{Time\ Stamping} + \sigma^2_{Detector\ Jitter}} \qquad \textbf{(2-16)}$$

In this formula, pulse timing is that aspect of the detector's output pulse that is used for determining its detection time. For example, if the time measurement system is *edge-triggered*, pulse timing would be a portion of the rise time of the pulse (depending on the trigger's set threshold). On the other hand, if *level-triggered*, pulse timing would be the time duration of that portion of the pulse shape that defines level. In how the photodetector output signal is amplified and used for triggering, fluctuations in the formation of this timing parameter (called *amplitude time walk*) cause a variance in measuring time.

On the other hand, time stamping is the recording of the signal's timing by the measurement system once the trigger requirement has been met; this parameter also experiences a variance (typically due to digitization noise of the measurement system).

However, since caused by the photodetector alone and as the limiting factor of the overall time resolution of a light detection system, detector jitter is the parameter of interest to our discussion. Considering the Poisson nature of photoelectron signal and noise, we can conclude that the jitter is proportional to the inverse of the square root of the number of photoelectrons:

$$\sigma_{Detector\ Jitter} \propto \frac{1}{\sqrt{N_{photon} \times PDE}} \qquad \textbf{(2-17)}$$

This relationship provides a highly effective tool in *preliminary* technical considerations, but one must understand its limitation: it is only applicable to timing at point of charge generation and excludes any variability in delays that might be present in charge collection and readout.

3. Case studies of APD/MPPC performance calculations

In this section, we will conduct studies of APD/MPPC performance calculations for product selection. As part of this review, we will learn a number of key techniques to perform such calculations. As discussed in the previous section, please keep in mind that optical power [W], whose dimension is normalized to time, can be converted to units of photons (and vice versa) using equation (2-1).

Let's suppose that an application demands the following conditions:

Peak wavelength of approx. 450 nm,
10 to 10^6 photons per pulse,
Pulse rates of 10 kMHz to 1 MHz,
Typical pulses having widths of approx. 8 ns, rise times of approx. 3 ns, and decay (fall) times of approx. 5 ns.

And with the following requirements: absolute intensity of each pulse must be measured with nonlinearity below 10% while the optical design allows for a detector photosensitive area or field-of-view (FOV) of 3 mm per channel.
Note that the application conditions are fairly generic; many seemingly important details, such as what source is producing the light signal, do not affect these back-of-the-envelope calculations.

3-1. MPPC S/N

We focus on studying photodetector S/N at the signal's low end. We use the typical specs of a regular Hamamatsu MPPC like the S13360-3050 (3 × 3 mm, 50 μm pixels):

- Total pixel count of 3600
- Typical ***PDE*** @ 450 nm = 40% at an overvoltage of 3 V
- Typical Gain of 1.7×10^6 at an overvoltage of 3 V
- Typical dark count rate (***DCR***) = 500 kcps at an overvoltage of 3 V
- Terminal capacitance ($\boldsymbol{C_t}$) = 320 pF

Using equation (2-2), we calculate charge output from a 10-photon input pulse: $10 \times 0.4 \times (1.7 \times 10^6) \times (1.6 \times 10^{-19})$ = 1090 fC. The node sensitivity of a typical digitizing readout (like a QDC, which is the charge-digitizing equivalent of an oscilloscope with similar node sensitivity) is on the order of low 10 s of fC/LSB (like 25 fC/LSB in the case of CAEN V965), and since 1090 fC > 25 fC, we conclude that S13360-3050 will be suitable for applications such as particle or nuclear physics in which use of flash digitizers (a QCD is simply a digitizer with on-board charge-to-voltage conversion) is prevalent; that is commonly the case if pulse shape information is required. However, such readout schemes are highly costly and power-intensive, making them unsuitable for common scientific, industrial, or consumer applications.

A cost-effective alternative readout scheme is use of photon-counting circuity such as a multi-channel analyzer (MCA), consisting of discriminator/scaler, counter, and other signal processing functions. Towards using equation (2-8) to calculate S/N for this scheme, the desired counting integration time must be determined. For the sake of our discussion, let's assume that is 1ms for which S13360-3050's ***DCR*** would yield 500 k × 1 m = 500 dark counts.

Using equation (2-8), we set S/N = 1 and solve $\frac{N_{photon} \times 0.4}{\sqrt{(N_{photon} \times 0.4) + (500)}} = 1$ to obtain $N_{photon} \approx 57$. At a min.

pulse rate of 10 kHz, 10 photons per pulse would yield 100 photons in 1ms. Alternatively, 100-photon S/N can be calculated using equation (2-8) to be $\frac{\mathbf{100 \times 0.4}}{\sqrt{\mathbf{(100 \times 0.4) + (500)}}} \approx 1.7$.

We repeat the same for S13360-3025 (3 × 3 mm, 25 μm pixels) with ***PDE*** @ 450 nm = 25% and typical ***DCR*** = 400 kcps to obtain an incident light level of $\boldsymbol{N_{photon}}$ = 82 for S/N = 1 (by solving $\frac{\boldsymbol{N_{photon} \times 0.25}}{\sqrt{\boldsymbol{(N_{photon} \times 0.25) + (400)}}} = 1$) or alternatively calculating S/N for 100 photons as $\frac{\mathbf{100 \times 0.25}}{\sqrt{\mathbf{(100 \times 0.25) + (400)}}} = 1.2$.

It is noteworthy that *"excellent"* S/N is generally considered to be ≥ 10. Most instrument designers typically have a target performance of S/N > ***X*** in mind for which ***X*** is greater than 1 (even if less than 10). Thus, it is important to use the value of ***X*** that represents the instrument designer's target S/N in the above calculations. For example, if ***X*** = 5, the above calculations would yield $\boldsymbol{N_{photon}}$ = 312 for S13360-3050 and $\boldsymbol{N_{photon}}$ = 453 for S13360-3025.

These results would mean that a portion of the lower range of the expected signal levels could not be detected with the target S/N (= 5) performance in this case.

3-2. MPPC linearity

We now assess how linear a MPPC's response would be under this application's conditions:

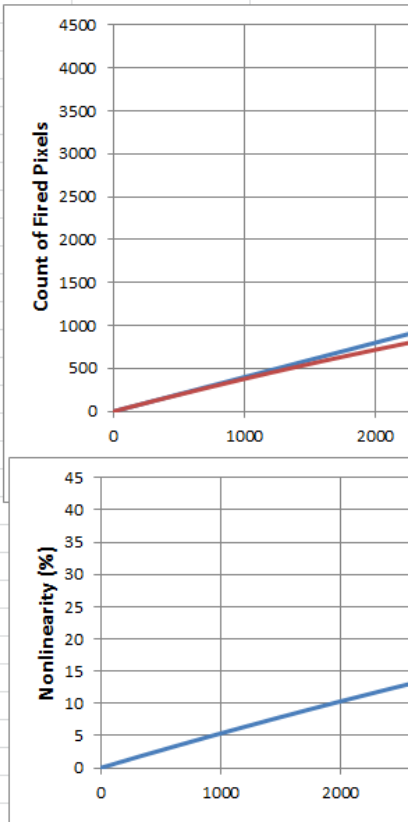
- **MPPC pulse-height linearity:** The question we face here is: up to how many 450 nm photons would S13360-3050 or S13360-3025 be able to detect with 10% max. nonlinearity ?
To answer this, we first obtain the pixel capacitance[4] by utilizing the nominal value of the MPPC's typical gain, assuming a unit of electrons for it and converting it to charge in coulombs, and then dividing the resulting charge by the specified over-voltage corresponding to that gain value. We thus have: $\frac{(1.7E6\ e-) \times (1.6E{-}19 \frac{C}{e-})}{3} \approx 91$fF. Using the quenching resistor value[5] of a 50 μm MPPC pixel, we then calculate S13360-3050's pixel recovery time to be 63 ns [approx. 4.6 × 91f × 150k where 4.6 = –ln(0.01) corresponds to a 99% MPPC recovery] and compare to the application's light pulse width. Since the condition $\boldsymbol{PW < T_{recovery}}$ is met (8 ns < 63 ns), we use equation (2-13) to plot the MPPC's expected response and compare that to its ideal response as obtained from equation (2-12) in Microsoft® Excel® and look for the point at which the 2 plots diverge by 10%. We perform the comparison by plotting the resulting nonlinearity using the combination of equation (2-10) and equation (2-11).

[4] Another approach could also be used to determine pixel capacitance; it consists of dividing the specified MPPC terminal capacitance by pixel count, which yields 88 fF (= 320 pF / 3600 pixels) in the case of S13360-3050. Please note, however, that measurement of MPPC gain is affected by the quenching resistor's parasitic capacitance while measurement of terminal capacitance is affected by the parasitic capacitances of the quenching resistor and also MPPC package and traces, and thus, either method overestimates the MPPC pixel's junction capacitance (considering that both parasitic capacitances are in parallel to the junction capacitance). This overestimation becomes particularly significant for MPPCs with smaller pixel sizes (10 μm and 15 μm in Hamamatsu's lineup) whose junction capacitances are relatively quite small.

[5] MPPC Rq values: 1 MΩ for 10 μm pixels, 1 MΩ for 15 μm pixels, 300 kΩ for 25 μm pixels, 150 kΩ for 50 μm pixels

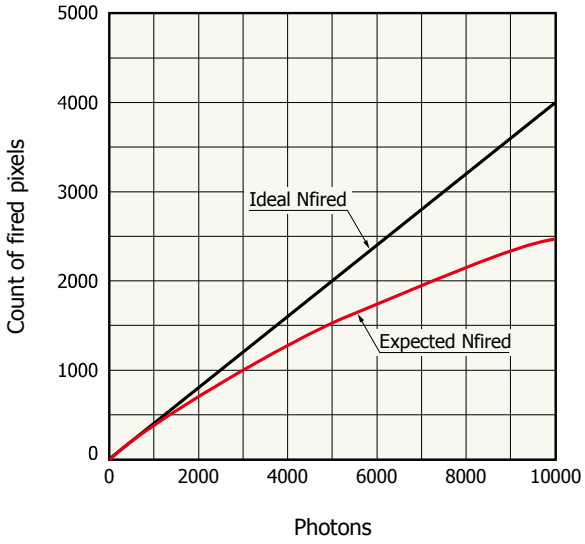
[Figure 3-1] Linearity calculation with Microsoft® Excel®

	A	B	C	D
1	PDE	0.4		S13360-3050
2	Npixel	3600		
3	Nphoton	Ideal Nfired	Expected Nfired	Nonlinearity (%)
4	5	=A4*0.4	=3600*(1-EXP((-1*A4*0.4/3600)))	=(100*ABS(C4-B4)/B4)
5	10	=A5*0.4	=3600*(1-EXP((-1*A5*0.4/3600)))	=(100*ABS(C5-B5)/B5)
6	15	=A6*0.4	=3600*(1-EXP((-1*A6*0.4/3600)))	=(100*ABS(C6-B6)/B6)
7	20	=A7*0.4	=3600*(1-EXP((-1*A7*0.4/3600)))	=(100*ABS(C7-B7)/B7)
8	25	=A8*0.4	=3600*(1-EXP((-1*A8*0.4/3600)))	=(100*ABS(C8-B8)/B8)
9	30	=A9*0.4	=3600*(1-EXP((-1*A9*0.4/3600)))	=(100*ABS(C9-B9)/B9)
10	35	=A10*0.4	=3600*(1-EXP((-1*A10*0.4/3600)))	=(100*ABS(C10-B10)/B10)
11	40	=A11*0.4	=3600*(1-EXP((-1*A11*0.4/3600)))	=(100*ABS(C11-B11)/B11)
12	45	=A12*0.4	=3600*(1-EXP((-1*A12*0.4/3600)))	=(100*ABS(C12-B12)/B12)
13	50	=A13*0.4	=3600*(1-EXP((-1*A13*0.4/3600)))	=(100*ABS(C13-B13)/B13)
14	55	=A14*0.4	=3600*(1-EXP((-1*A14*0.4/3600)))	=(100*ABS(C14-B14)/B14)
15	60	=A15*0.4	=3600*(1-EXP((-1*A15*0.4/3600)))	=(100*ABS(C15-B15)/B15)
16	65	=A16*0.4	=3600*(1-EXP((-1*A16*0.4/3600)))	=(100*ABS(C16-B16)/B16)
17	70	=A17*0.4	=3600*(1-EXP((-1*A17*0.4/3600)))	=(100*ABS(C17-B17)/B17)
18	75	=A18*0.4	=3600*(1-EXP((-1*A18*0.4/3600)))	=(100*ABS(C18-B18)/B18)
19	80	=A19*0.4	=3600*(1-EXP((-1*A19*0.4/3600)))	=(100*ABS(C19-B19)/B19)
20	85	=A20*0.4	=3600*(1-EXP((-1*A20*0.4/3600)))	=(100*ABS(C20-B20)/B20)
21	90	=A21*0.4	=3600*(1-EXP((-1*A21*0.4/3600)))	=(100*ABS(C21-B21)/B21)
22	95	=A22*0.4	=3600*(1-EXP((-1*A22*0.4/3600)))	=(100*ABS(C22-B22)/B22)
23	100	=A23*0.4	=3600*(1-EXP((-1*A23*0.4/3600)))	=(100*ABS(C23-B23)/B23)
24	200	=A24*0.4	=3600*(1-EXP((-1*A24*0.4/3600)))	=(100*ABS(C24-B24)/B24)
25	300	=A25*0.4	=3600*(1-EXP((-1*A25*0.4/3600)))	=(100*ABS(C25-B25)/B25)
26	400	=A26*0.4	=3600*(1-EXP((-1*A26*0.4/3600)))	=(100*ABS(C26-B26)/B26)
27	500	=A27*0.4	=3600*(1-EXP((-1*A27*0.4/3600)))	=(100*ABS(C27-B27)/B27)
28	600	=A28*0.4	=3600*(1-EXP((-1*A28*0.4/3600)))	=(100*ABS(C28-B28)/B28)
29	700	=A29*0.4	=3600*(1-EXP((-1*A29*0.4/3600)))	=(100*ABS(C29-B29)/B29)
30	800	=A30*0.4	=3600*(1-EXP((-1*A30*0.4/3600)))	=(100*ABS(C30-B30)/B30)



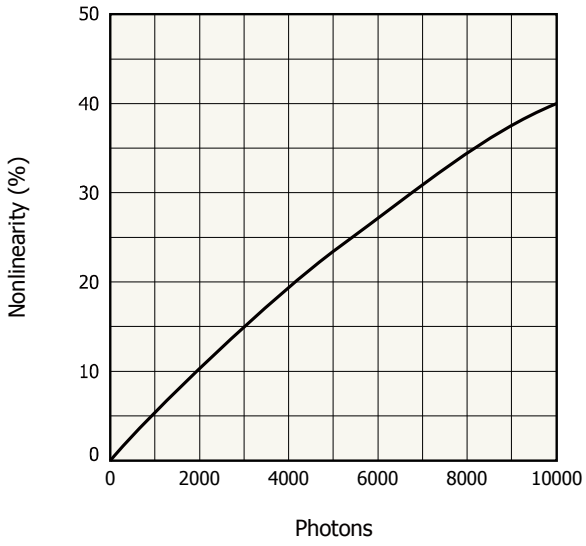
For S13360-050, the point of 10% nonlinearity is at about 2000 photons:

[Figure 3-2] Count of fired pixels vs. photons



KAPDB0348EA

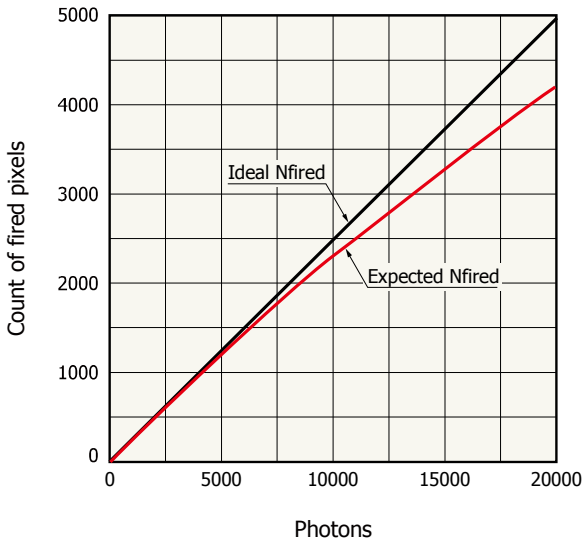
[Figure 3-3] Nonlinearity vs. photons



KAPDB0349EA

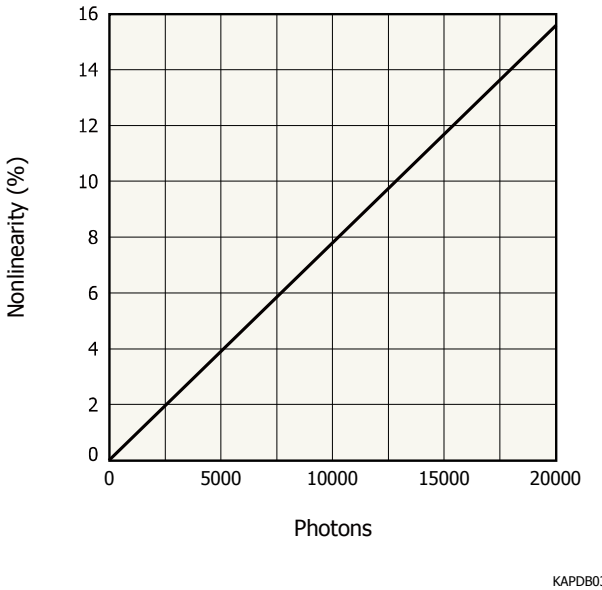
For S13360-025, by repeating the same calculations and plots, we find the point of 10% nonlinearity to be at about 12000 photons:

[Figure 3-4] Count of fired pixels vs. photons



KAPDB0350EA

[Figure 3-5] Nonlinearity vs. photons



- MPPC pulse-rate linearity: From the application conditions, we see that the shortest inter-pulse time is longer than S13360-3050's pixel recovery time (92 ns > 63 ns), so this application is within the pulse-rate linearity range of the S13360-3050.

3-3. APD S/N

Now, let's assess APD's S/N at those signal levels above which MPPC linearity falls short of the application's linearity requirement. We choose a blue-enhanced APD of suitable size, like Hamamatsu's S8664-30K, and utilize its characteristics ($\boldsymbol{QE}$ @ 450 nm = 75%, $\boldsymbol{I_d}$ = 1 nA, $\boldsymbol{M_{opt}}$ = 50, $\boldsymbol{F} = 50^{0.2} \approx 2.2$) in equation (2-7) to calculate S/N. Considering that the application is photometric (i.e. measuring the amount of incident light in absolute terms for which a charge amplifier is required), we also use the readout noise spec of a sufficiently fast charge amplifier (one that can resolve a single pulse at the max. pulse rate); we note 993e- in case of Analog Devices AD8488. Note that the charge amplifier's bandwidth (as inverse of its integration time) must be at least twice that of the application's max. expected pulse rate, and hence, we will use 2 MHz as the min. amplifier bandwidth limit. We proceed to calculate $\boldsymbol{S_{dark}} = \frac{\boldsymbol{I_d}}{\boldsymbol{q} \times \Delta \boldsymbol{f}}$ = 1n / (2M × 1.6 × 10^{-19}) = 3125e-. Using equation (2-7), we thus have

$$\boldsymbol{SNR} = \frac{\mathbf{0.75 \times 50 \times 2000}}{\sqrt{\mathbf{2.2 \times [(0.75 \times 50^2 \times 2000) + (50 \times 3125)] + 993^2}}.} \approx \mathbf{24}$$ for 2000 photons and

$$\boldsymbol{SNR} = \frac{\mathbf{0.75 \times 50 \times 12000}}{\sqrt{\mathbf{2.2 \times [(0.75 \times 50^2 \times 12000) + (50 \times 3125)] + 993^2}}.} \approx \mathbf{63}$$ for 12000 photons.

As mentioned before, *excellent* S/N is typically considered to be ≥ 10. Therefore, in such a photometric application using a charge amplifier, we conclude that S8664-30K can perform well at those signal levels at which S13360-3050 and S13360-3025 exhibit excessive nonlinearity. Thus, it is imperative to use S13360-3050 (instead of S13360-3025) in photon-counting mode to detect the smaller pulses in his application but then consider using S8664-30K for pulses > 2000 ph.

As a side exercise, we calculate the APD's S/N for making a relative measurement (requiring a resistive trans-impedance amplifier for readout) under the same conditions. First, let's explore the case that the APD and its

output amplifier circuit are intended for use as a pulse counter, which would make the signal's highest-frequency component to be the max. expected pulse rate of 1 MHz. Since the output amplifier's bandwidth must be at least twice that of the measurement frequency (which would be the frequency of the signal's highest-frequency component that is to be detected), we adopt 2 MHz as our min. amplifier bandwidth limit (or cutoff frequency in other words). Using equation (2-6) along with S8664-30K's characteristics [$\boldsymbol{\Phi}$ @ 450 nm ≈ 0.3 A/W [derived from QE by equation (2-5)], $\boldsymbol{I_d}$ = 1 nA, $\boldsymbol{M_{opt}}$ = 50, $\boldsymbol{F}$ = $50^{0.2}$ ≈ 2.2] and Texas Instruments amplifier OPA380 characteristics (readout noise spec of 10$\frac{\text{fA}}{\sqrt{\text{Hz}}}$ resulting in an estimated readout noise of approx. 14 pA at 2 MHz), we have:

$\boldsymbol{S/N} = \frac{\mathbf{0.3^2 \times (8.8 \times 10^{-16} \times 2E6)^2 \times 50^2}}{\mathbf{2 \times 1.6E{-}19 \times 2E6 \times 50^2 \times 2.2 \times [(16 \times 8.8E{-}10) + 10^{-9}] + (1.4 \times 10^{-11})}} < \mathbf{1}$ for detecting pulses of 2000 photons or 0.88 fJ of incident 450 nm light per pulse at a max. rate of 1 MHz, and

$\boldsymbol{S/N} = \frac{\mathbf{0.3^2 \times (5.28 \times 10^{-15} \times 2E6)^2 \times 50^2}}{\mathbf{2 \times 1.6E{-}19 \times 2E6 \times 50^2 \times 2.2 \times [(16 \times 5.28E{-}9) + 10^{-9}] + (1.4 \times 10^{-11})}} < \mathbf{1}$ for detecting pulses of 12000 photons or 5.28 fJ of incident 450 nm light per pulse at a max. rate of 1 MHz.

Now, let's study the case that the APD and its output amplifier circuit will be used to perform pulse-shape discrimination (PSD) on signal pulses, which would require the signal's highest-frequency component to be obtained from its rise time (since shorter than the fall time) by the following approximation $\frac{0.35}{3\ ns}$ ≈ 117 MHz. Like before, since the output amplifier's cutoff frequency must be at least twice that of the measurement frequency, we adopt 234 MHz as our min. amplifier bandwidth limit. Using equation (2-6) along with S8664-30K's characteristics ($\boldsymbol{\Phi}$ @ 450 nm ≈ 0.3 A/W, $\boldsymbol{I_d}$ = 1 nA, $\boldsymbol{M_{opt}}$ = 50, $\boldsymbol{F}$ = $50^{0.2}$ ≈ 2.2) and Analog Devices amplifier AD8015 specifications (indicating a readout noise spec of 3$\frac{\text{pA}}{\sqrt{\text{Hz}}}$ resulting in an estimated readout noise of approx. 46 nA at 234 MHz), we have:

$\boldsymbol{S/N} = \frac{\mathbf{0.3^2 \times (8.8 \times 10^{-16} \times 2.34 \times 10^{8})^2 \times 50^2}}{\mathbf{2 \times 1.6E{-}19 \times 234E6 \times 50^2 \times 2.2 \times [(16 \times 8.8E{-}10) + 10^{-9}] + (4.6 \times 10^{-8})}} < \mathbf{1}$ for 2000 photons or 0.88 fJ of incident 450 nm light per pulse with a rise time of 3 ns, and

$\boldsymbol{S/N} = \frac{\mathbf{0.3^2 \times (5.28 \times 10^{-15} \times 2.34 \times 10^{8})^2 \times 50^2}}{\mathbf{2 \times 1.6E{-}19 \times 234E6 \times 50^2 \times 2.2 \times [(16 \times 5.28E{-}9) + 10^{-9}] + (4.6 \times 10^{-8})}} < \mathbf{1}$ for 12000 photons or 5.28 fJ of incident 450 nm light per pulse with a rise time of 3 ns.

These results show that the above signal amplitudes are too low for relative detection at such high bandwidths using S8664-30K and the aforementioned amplifiers. So, let's calculate at what signal levels we could attain S/N = 1 and S/N = 10 using S8664-30K and the above amplifiers at the same bandwidths.

For the earlier scenario of detecting and counting 450 nm light pulses at a rate of 1 MHz using S8664-30K, we have:

$\frac{\mathbf{0.3^2 \times (S_{input} \times 2E6)^2 \times 50^2}}{\mathbf{2 \times 1.6E{-}19 \times 2E6 \times 50^2 \times 2.2 \times [(16 \times 8.8E{-}10) + 10^{-9}] + (1.4 \times 10^{-11})}} = \mathbf{1}$ which yields $\boldsymbol{S_{input}}$ ≈ 124 fJ or ≈ 282 thousand 450 nm photons per pulse.

$$\frac{0.3^2 \times (S_{input} \times 2E6)^2 \times 50^2}{2 \times 1.6E{-}19 \times 2E6 \times 50^2 \times 2.2 \times [(16 \times 8.8E{-}10) + 10^{-9}] + (1.4 \times 10^{-11})} = \mathbf{10}$$ which yields S_{input} ≈ 39.4 pJ or ≈ 896 thousand 450 nm photons per pulse.

For the latter scenario of performing PSD on 450 nm light pulses with a rise time of 3 ns (and thus, amplifier bandwidth of 234 MHz) using S8664-30K, we have:

$$\frac{0.3^2 \times (S_{input} \times 2.34 \times 10^8)^2 \times 50^2}{2 \times 1.6E{-}19 \times 234E6 \times 50^2 \times 2.2 \times [(16 \times 8.8E{-}10) + 10^{-9}] + (4.6 \times 10^{-8})} = \mathbf{1}$$ which yields S_{input} ≈ 61 fJ or ≈ 139 thousand 450 nm photons per pulse.

$$\frac{0.3^2 \times (S_{input} \times 2.34 \times 10^8)^2 \times 50^2}{2 \times 1.6E{-}19 \times 234E6 \times 50^2 \times 2.2 \times [(16 \times 8.8E{-}10) + 10^{-9}] + (4.6 \times 10^{-8})} = \mathbf{10}$$ which yields S_{input} ≈ 19.3 pJ or ≈ 439 thousand 450 nm photons per pulse.

3-4. APD linearity

We now discuss how linear S8664-30K's response would be under this application's conditions:

- Pulse-rate linearity: For this, a calculation of response cutoff frequency using equation (2-15) would be made based on the terminal capacitance values specified in Hamamatsu's APD datasheets for a load resistance of 50 Ω. However, APD cutoff frequency values (calculated in the same way) are provided in Hamamatsu's APD datasheets (so no need to calculate!). In the case of S8664-30K, the specified cutoff frequency is 140 MHz, which far exceeds this application's max. pulse rate of 1 MHz.

- Pulse-height linearity: For calculating the upper limit of pulse-height linearity, one would calculate the APD's charge storage capacity by using $Q = C_t \times V_{bias}$ in which V_{bias} is the APD's reverse bias voltage for the desired gain; please note that plots of APD gain and terminal capacitance vs. reverse voltage are provided in Hamamatsu APD datasheets. For our case study, we will use S8664-30K's terminal capacitance of 22 pF and the reverse bias voltage of 360 V for the optimal gain of M_{opt} = 50 to obtain a charge storage capacity of 7.9 nC. Back-calculating from that charge storage capacity by taking S8664-30K's QE (0.75 @ 450 nm) and optimal gain (M_{opt} = 50) into account, we arrive at a photon count of 1.3 × 10^9, which is larger than the max. photon count of 10^6 per pulse in this application. Thus, this application is within S8664-30K's pulse-height linearity.

Furthermore, one also needs to take the readout circuity into account. For example, Hamamatsu's H4083 charge amplifier has a 2 pF storage capacitor biased up to the rail voltage of 12 V, which by $Q = C \times V$ yields a max. charge storage capacity of 24 pC or 1.5 × 10^8 electrons (or alternatively obtained by max. output voltage of 12 V divided by gain of 0.5 V/pC). Like the previous step, back-calculating from that amount of charge by taking S8664-30K's QE (0.75 @ 450 nm) and optimal gain (M_{opt} = 50) into account, we arrive at a photon count of 4 × 10^6, which is larger than the max. photon count of 10^6 per pulse in this application. Thus, pulse-height linearity would not be limited by the charge amplifier circuitry if H4083 is utilized in this case.

3-5. Conclusion

In conclusion, please note that multiple products could turn out to be suitable for a given set of application conditions. By taking price information into consideration, those options can be trimmed down to one or more

candidates for characterization and evaluation. With that, considering their peculiar complexities, we will dedicate the next section to describing methods of measuring MPPC characteristics and discussing their specifics.

4. MPPC characterization measurements

Sometimes referred to by other terms such as a data-point, a sampling, a reading, or an event, a measurement is a ***single*** experimental attempt [temporally (per a duration of time) and spatially (per a detector channel and within its physical size and field-of-view)] to quantify an optical signal.
The result of a measurement is thus a piece of quantitative data that represents essential information about the true signal. Measurement is a very fundamental and seemingly simple concept, so why is it explored here? There are 3 important points to keep in mind:

- **Measurement Frequency:** Understanding the required bandwidth for a measurement is of paramount importance to proper identification and definition of an application's requirements and in order to compare those requirements with characteristics of candidate photodetectors. Fundamentally, based on Nyquist's sampling theorem, the sampling frequency must be at least twice the measurement frequency, which in turn would be equal to the frequency of the signal's highest-frequency component that is intended to be detected. Furthermore, as discussed before, a readout amplifier circuit's cutoff frequency at -3 dB should be designed to be at least twice that of the measurement frequency (but as a general rule of thumb, 4 times is an advisable design target).

For better illustration, the following examples portray cases in which a detrimental bandwidth mismatch is present:

- Using an amplifier with a cutoff frequency of 1 MHz to measure pulses with frequency of 900 kHz and rise/fall times of 1 μs.

- Using an imager with a spatial resolution of 10 lp/mm to contact-image a min. feature size of 20 μm. [Reason: 10 lp/mm is theoretically able to observe a min. feature size of 1 mm /(2 × 10) = 50 μm]

- Using an oscilloscope with a sampling rate of 150 MHz to look for multiple randomly-occurring events with durations of 10 ns each (Reason: 150 M < 2 × 1/10 n = 200 M).

- **Measurement resolution and range:** Besides measurement frequency, the resolution of the measurement instrumentation in its ability to accurately resolve the smallest expected quantity of or change in the parameter (that is to be measured) over the full range of its expected values is also of great importance. The following are examples in which there is a detrimental mismatch between signal characteristics and measurement resolution and range:

- Using an 8-bit digitizer with a conversion factor of 500 ke-/LSB to plot pulse height distribution for up to 150 p.e. pulses of a MPPC at gain = 10^6 (Reason: 150 ×10^6 / 5 × 10^5 = 300 > 2^8).

- Using a detector with a rise time of 10 ns to discriminate the pulse shapes of 2 signals with rise times of 1 ns and 8.5 ns.

- Using a spectrometer with a resolution of 30 nm to distinguish emission or excitation peaks 10 nm apart.

- Using a 16-bit/sample ADC with a bit rate of 50 Mbps to record the output waveform of an MPPC with a dark count rate of 4 Mcps for under dark conditions. (Reason: 50 M / 16 < 4 M)

- Normalizations: While some signal characteristics (like amplitude, timing or frequency) can be appropriately obtained from measuring a single parameter (whether once or a multitude of times), others (like flux or power) would by definition include normalizations to time and/or spatial information and hence require measurements of 2 or more parameters in order to be quantified.

These differences are important, since a typical instrument designer is naturally concerned with the performance of her overall instrumentation: she could be designing for an application condition that combines a series of measurements and contains one or more normalizations instead of consisting of a single parameter alone. In order to assess a photodetector's suitability for a given application, one must bear that in mind and see how the way detector characteristics have been specified compares with the designer's intended application conditions and target requirements.

Towards that, one would begin by finding out to what temporal or spatial parameters a stated application condition might have been normalized; when in doubt, one should make sure about the dimension or unit of the application condition in question. Furthermore, it is also important to understand the scope of those normalizations.

For example, one needs to examine whether the incident light "power" is applied to the entirety of a detector's photosensitive area or normalized to pixel count or unit of area. As another example, it is important to verify whether *peak* power or *average* power is meant when a reference to incident light "power" is made; peak and average power have different temporal normalizations in that the former is the ratio of energy content of a single pulse to the time duration of that pulse while the latter takes the time in between consecutive pulses into account. As a further example, if various spectral band components of a broad spectrum of light are to be detected by different detector channels separately (applicable to 1D or 2D serial- or parallel-readout multichannel detectors), performance of each detector channel must be assessed independently in accordance to its particular expected input light signal conditions.

With the above caveats in mind, we proceed to introducing methods of measuring MPPC characteristics. In what follows, we will focus on the characterization of Hamamatsu's S13360-3050CS as an example, but please be assured that these measurement methods can be adapted to other MPPC models with little or no adjustments.

Before we begin, however, let's briefly discuss the choice of a suitable power supply to bias the MPPC. At Hamamatsu Photonics, we typically utilize a source meter to bias the MPPC, since it allows us to control the applied voltage via a PC connection and also to set a max. voltage of 80 V for safe biasing of the MPPC. Furthermore, a source meter enables automated I-V curve characterization and can thus be used in QC screening of mass-produced parts. Therefore, despite its relatively high cost, a source meter is recommendable as a suitable biasing solution in performing the measurements described hereinafter.

4-1. Gain and breakdown voltage (V_{BR}) measurement

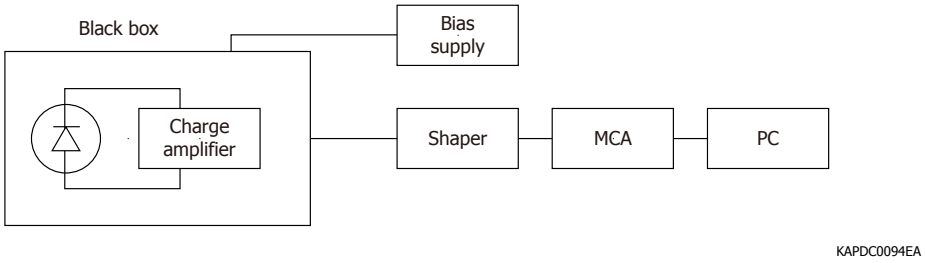
- Measurement Principle:
MPPC gain is the factor by which Geiger-mode avalanche (whether initiated by the photoelectric effect or thermal carrier excitation) multiplies the initiating electron to form the MPPC's output charge per avalanche. Following the mechanism that we explained in Section 1, that gain is proportional to overvoltage (V_{over}), which is the difference between the bias voltage applied to the MPPC and the MPPC's inherent breakdown voltage (V_{BR}).

The simplest measurement of V_{BR} can be derived from the x-intercept of a MPPC's gain vs. bias voltage plot. Assuming a proportionality factor of $\boldsymbol{A}$ between gain and overvoltage, we can say Gain $= \boldsymbol{A} \times V_{over} = \boldsymbol{A} \times (V_{bias} - V_{BR})$, and at the gain-voltage plot's x-intercept (at which Gain = 0), V_{bias} thus corresponds to the breakdown voltage.

Determining the V_{BR} of MPPC before beginning other measurements is practically beneficial and recommendable. Most MPPC characteristics whose measurement we are going to introduce in this section, namely photon detection efficiency, dark count rate, prompt or delayed crosstalk probability, and afterpulse probability, are dependent on V_{over}. Additionally, V_{BR} differs between each and every pair of MPPCs even if they are the same product – for example, amongst a small batch of S13360-3050CS, V_{BR} of each is different from the rest ; operating them at the same applied bias voltage would likely results in differing characteristics. It is thus recommended to compare or calibrate the characteristics of MPPCs at the same overvoltage (or the same gain if practical).

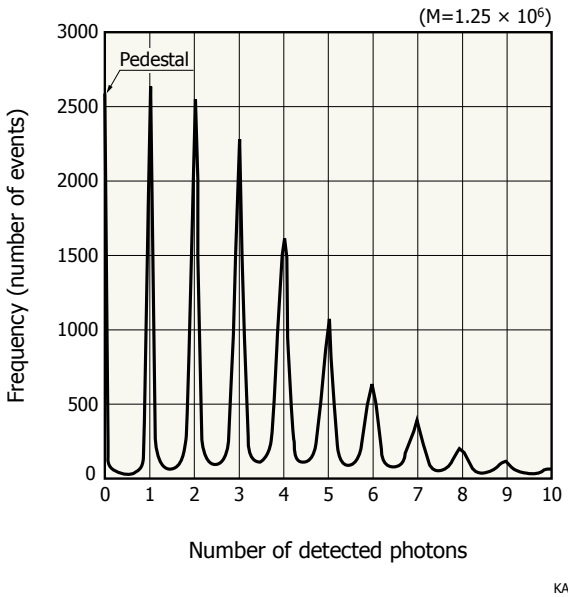
- Measurement setup: Figure 4-1 shows the gain measurement setup.
In this setup, the MPPC is placed inside a dark box and electrically connected to it (if metallic) via a common ground in order to decrease the parasitic impedance that forms between them. Output of the charge amplifier goes to a shaping amplifier, which is followed by a MCA. Finally, a dedicated FPGA circuit sends the MCA output to a PC through a USB connection. We analyze the data coming from the MCA with a LabVIEW software program in this setup.

[Figure 4-1] Gain measurement setup



- Measurement procedure:
In our setup, the MCA output is used to produce a histogram of the shaping amplifier's output pulses; this histogram would consist of pulse count and digitized pulse height (proportional to the MPPC's output charge) as its y and x axes, respectively. Figure 4-2 shows an example of this histogram. As shown, the histogram has several peaks, which correspond to (from left to right) the noise pedestal (population of readout noise pulses) and populations of 1 p.e., 2 p.e., … pulses of the MPPC. Based on the properties of these peaks, we can compute the net output charge corresponding to a 1 p.e. pulse by calculating the peak-to-peak interval between any two consecutive peaks (excluding the pedestal) along the x-axis after which we convert the interval value (in digital counts [LSB]) to equivalent charge amount based on the charge amplifier's gain [V/C] and the MCA's A/D conversion factor [LSB/V]. This 1 p.e. charge is, by definition, equal to the gain of the evaluated MPPC at the applied bias voltage. We repeat this measurement process for various applied voltages to obtain the gain-voltage plot. We then perform linear fitting on the resulting plot; the x-intercept of the fitted line represents V_{BR} of the MPPC under evaluation.
It is noteworthy to point out a potential pitfall in performing this measurement: the MPPC output pulse populations must be sufficiently large to be statistically significant in order to create a meaningful histogram. This condition could be difficult to attain for low overvoltage levels or if the MPPC's dark count rate is too small. In such scenarios, low-intensity illumination of MPPC photosensitive area should be utilized to increase the output pulse counts.

[Figure 4-2] Histogram example of shaping amplifier's output pulse

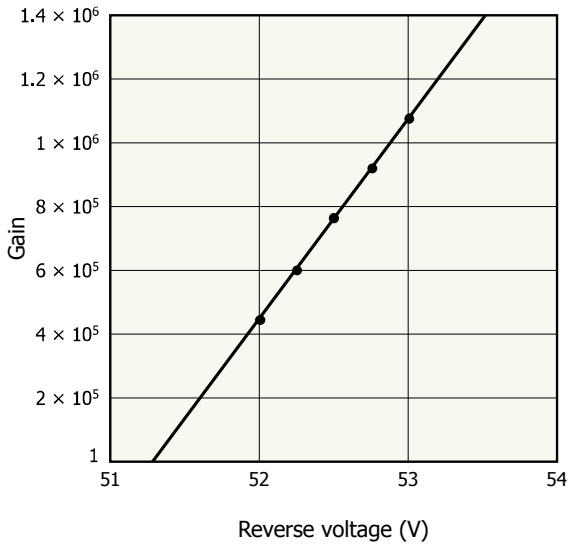


KAPDB0133EA

- Measurement result:

Figure 4-3 shows the gain-voltage plot and its linear fit for S13360-3050CS. From this plot, we conclude that V_{BR} of this MPPC is 51.29 V.

[Figure 4-3] Gain vs. voltage plot example (S13360-3050CS)



KAPDB0361EA

4-2. Breakdown voltage measurement by obtaining Vpeak from I-V curve

- Measurement Principle:

There is another way of obtaining MPPC's V_{BR} that is useful when evaluating many MPPC units of the same type. We use this measurement method in the mass inspection of our MPPC products.

V_{peak} is defined as the inflection point of *log(I)* vs. *V* curve. The critical property of V_{peak} is that the subtraction $V_{peak}-V_{BR}$ depends only on the type of MPPC and not on the actual value of V_{BR}. Combined with the fact that

measuring the I-V curve only around V_{peak} is less cumbersome than measuring the gain-voltage plot, this measurement method is highly efficient for evaluating V_{BR} for a large batch of MPPCs.

To illustrate with an example, let us assume that we want to evaluate V_{BR} values of 100 units of S13360-3050CS. Measuring the gain-voltage plot 100 times for all units would be very demanding, so we instead obtain V_{BR} of just one unit, considering that measuring gain-voltage plot only one time is not too much of a burden. For the same unit, we then obtain its V_{peak} by the method to be described here shortly, so that we can calculate the value of (V_{peak}−V_{BR}) from these two measurement results. This value is common amongst all 100 units of S13360-3050CS in our hypothetical batch, and we can calculate V_{BR} of the remaining 99 units by measuring their individual I-V curves and obtaining their particular V_{peak} values.

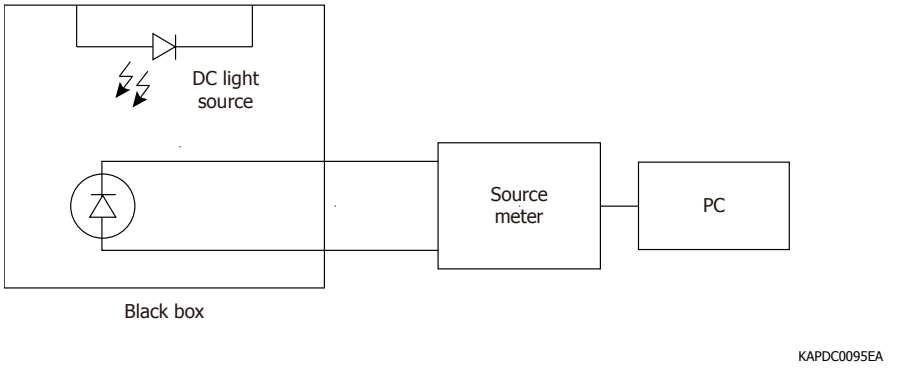
- Measurement Setup:

Figure 4-4 shows the V_{peak} measurement setup.

The setup is the same as that of gain measurement but with the key differences being utilization of a stable light source (such as an LED) and a different output readout scheme. In this setup, we directly connect the MPPC to a source meter and simply read out the MPPC's output current. We repeat this readout for various applied voltages to obtain the I-V curve.

We transfer this I-V curve data from the source meter to a PC in order to perform I-V curve analysis and obtain the V_{peak} of each MPPC under evaluation.

[Figure 4-4] V_{peak} measurement setup

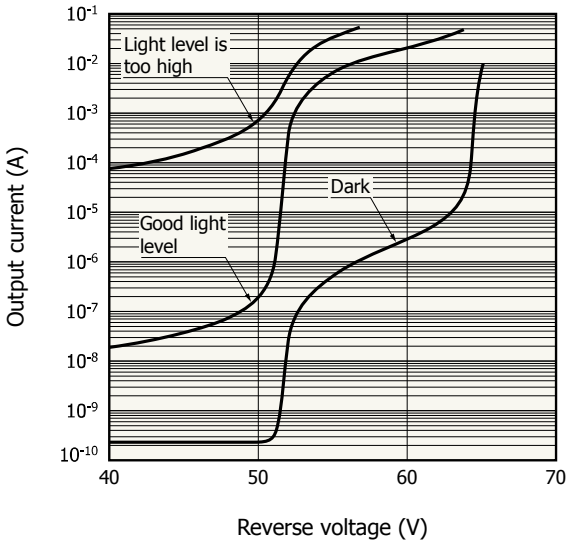


- Measurement procedure:

Figure 4-5 shows a measured I-V curve of S13360-3050CS under various luminance levels coming from the light source. Under dark condition, the I-V curve comprises bulk current and surface leakage current. The latter strongly affects the I-V curve's baseline under dark condition and the resulting V_{peak}. By illuminating MPPC with the light source, however, we can ignore the baseline and calculate the V_{peak} reliably.

In practice, one should experimentally determine the suitable LED luminance by repeating the I-V curve measurement for various light levels while avoiding excessive illumination of the MPPC.

[Figure 4-5] I-V curve example of V_{peak} measurement



- Measurement result:
Table 4-1 shows the V_{peak} measurement results for three units of S13360-3050CS. For reference, measured V_{BR} of each of those MPPCs (via the x-axis intercept method) has also been listed. As you can see, the difference $V_{peak}-V_{BR}$ has the same value amongst all three.

[Table 4-1] V_{peak} (and V_{BR}) measurement result for 3 pcs of S13360-3050CS

Sample no.	1	2	3	Unit
Vpeak	51.47	51.57	51.87	V
V_{BR}	51.29	51.41	51.70	V
Difference	0.18	0.16	0.17	V

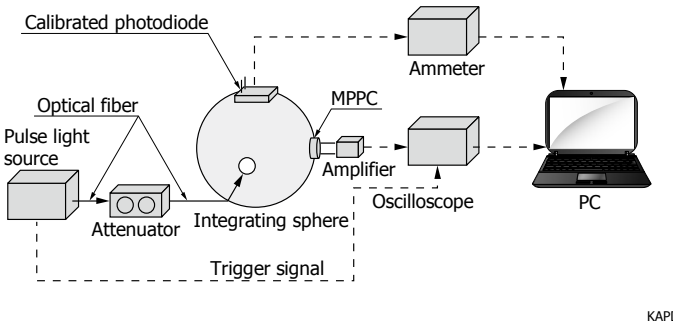
4-3. Photon detection efficiency (PDE) vs. bias voltage measurement

- Measurement principle: A key aspect of measuring MPPC's PDE is the exclusion of correlated noise (optical crosstalk and afterpulses) from the measured data. While several such techniques have been devised over the recent years, the best-known MPPC PDE measurement technique (also utilized by Hamamatsu) has been described extensively in [11]. We will provide an overview of that technique as performed by Hamamatsu in this subsection.

In simple terms, the aforementioned technique relies on measuring the amplitudes of MPPC output pulses (instead of output current), plotting a pulse height histogram, and then calculating the ratio of non-photon (< 1 p.e. in height) event count (also known as the *pedestal* event count and noted as N_{ped} in [11]) to total event count (= N_{ped} + $N_{\geq 1p.e.}$ and noted as N_{tot} in [11]). This ratio is not affected by crosstalk and afterpulses because correlated noise does not occur in absence of signal detection; it only appears in correlation with genuine (thermal or photoelectric) events of 1 p.e., 2 p.e., … in height.

- Measurement setup: Figure 4-6 shows the PDE measurement setup.

[Figure 4-6] Hamamatsu's PDE measurement setup



The following are typical examples of instruments that can be used in this measurement setup:

Pulse light source: PLP-10 (Hamamatsu)
Wavelength: 408 nm
(several other wavelengths such as 655 nm, 851 nm, etc. are also available)
Pulse width: 88 ps FWHM
(wavelength dependant)
Optical Attenuator[6]: DA-100-3U-850-50/125-M-35 (OZ optics)
Integrating sphere: 3P-GPS-033-SL (Labsphere)
Power meter: 2936-R (Newport)
Bias supply: GS610 (Yokogawa) or 2636B (Keithley) and the like
Amplifier: Linear amplifier (internal product and not for sale), bandwidth: 450 MHz
Oscilloscope: SDA 760Zi (LeCroy)
Measurement software program is written in LabVIEW 2010.

The pulsed light source used in this measurement should emit monochromatic light and must have a pulse width shorter than the MPPC rise time (which is on the order of a few nanoseconds).

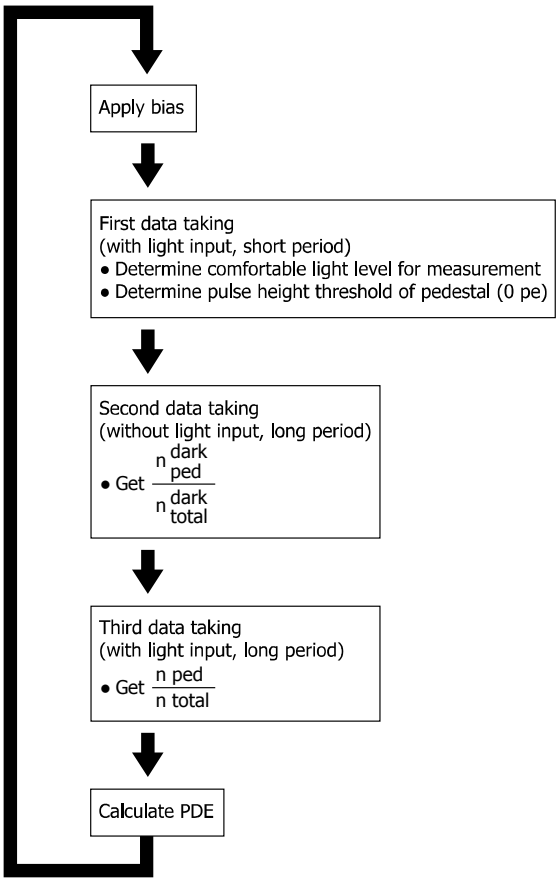
In utilizing the integrating sphere, we mount the MPPC and the power meter's photodiode head on two output ports of the sphere. To hold it on the output port, the MPPC is mounted using a fixture with a small aperture radius (0.5 mm or 0.3 mm depending on the MPPC's photosensitive area in order to ensure the illumination of the MPPC photosensitive area only).
The ratio of output light levels of the sphere's two ports should be measured in advance by mounting the same photodiode head on each port and dividing the resulting power meter outputs. It should be emphasized that light intensity in determining the output power ratio of the sphere's ports should be high enough to allow for the power meter's photodiode head to detect it with good accuracy through the small aperture of MPPC's fixture. When using a pulsed light source as in our case, higher light intensities can be achieved by increasing the light pulse frequency.

[6] The listed optical attenuator is designed for an attenuation wavelength of 850 nm. When the optical signal wavelength is different from the attenuator's design wavelength, the actual attenuation level would differ from what the user sets the attenuator to operate at. Nevertheless, for the purpose of this measurement, we do not need to know the attenuation level since we can measure the attenuator's output (as the number of photons that are illuminated onto the MPPC) by the power meter. In that sense, we only use the attenuator to adjust the amount of light incident onto the MPPC; how much light is lost (between the source and the attenuator's output) is irrelevant to this measurement.

To ensure high S/N for good measurement accuracy, the output amplifier used to read the MPPC must have relatively low levels of noise. Since an amplifier's noise performance and its bandwidth are competing factors, it is acceptable to compromise on the bandwidth even if that would result in the distortion of the MPPC output pulse shape.
For digitization and analysis of the linear amplifier's output pulses, we use the oscilloscope SDA 760Zi (LeCroy), which has the capability to perform PHA/PHD (i.e. create pulse height analysis/distribution histograms) at a high data throughput.

- **Measurement procedure:** Figure 4-7 shows a flowchart of our PDE measurement program's underlying algorithm. We wrote this program to automatically measure MPPC's PDE vs. bias voltage.

[Figure 4-7] Flow-chart of our PDE measurement program's algorithm



The flowchart loop begins with setting an initial bias voltage, which must be controlled by a temperature-compensation scheme if a temperature-controlled enclosure or ambient environment is unavailable. Then, we proceed to the main measurement procedure that comprises three data taking steps: the first measures the light intensity incident onto the MPPC and determines the N_{ped} threshold while the following two steps measure N_{ped}.

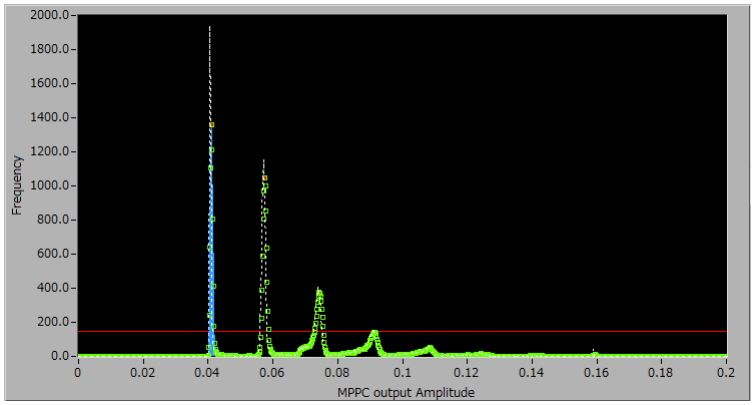
For determining N_{ped} in the following two steps, we use simpler criteria (than what has been described in [11]) by counting the number of events below the threshold set by Step A. This approach is suitable when there is good separation between discrete p.e. peaks and a reasonably high (≥ 2) peak-to-valley (P/V) ratio exists between 0 p.e. and 1 p.e. peaks and the valley between them. If such distinct separation between peaks is not attained, the 0 p.e. peak can be fitted to a Gaussian distribution curve with the area under it calculated as described in [11]. For the case of S13360-3050CS, as Figure 4-8 shows, our simple method can be reliably used. This measurement is not

intended to take N_{ped}/N_{tot} precisely, and a short measurement time period is acceptable. In this case, we set it to 10 seconds. If the light level is too low or too high to measure the 0 p.e. ratio effectively for the following measurement, we change the light intensity by changing the optical attenuation level and repeat collecting the data until the light intensity becomes suitable for the following steps.

Step B is for measuring $N_{ped}^{dark}/N_{tot}^{dark}$ in dark condition. Now, we set the light attenuation level to the maximum (40 dB in this example) to ensure that the MPPC is practically under dark condition. The time period of data taking should be long enough to exclude the statistical fluctuation of 0 p.e. events, and we thus set the time to 2 minutes in this step.

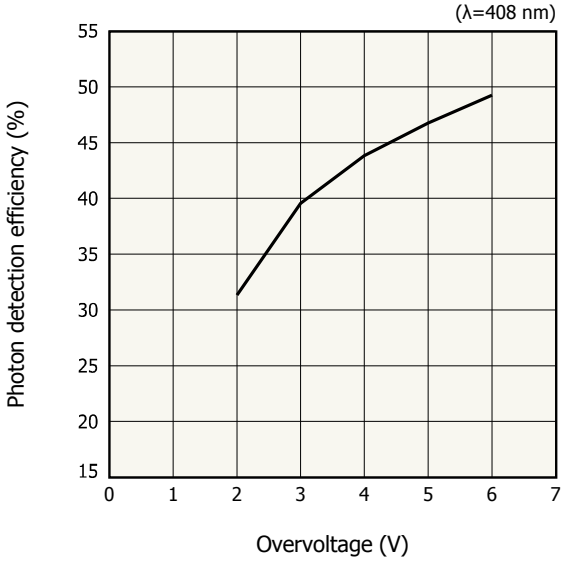
Step C is for measuring N_{ped}/N_{tot}. We set the light attenuation to the level utilized in Step A but then take the data with a comparatively long period (2 minutes – same as the Step B). At the same time,
we obtain the number of input photons (incident onto the MPPC) per light pulse (n_{in}) from the average value of powermeter output during data collection as divided by the light pulse rate.
After these data collection steps, we finally calculate the PDE value for this bias voltage using $N_{ped}^{dark}/N_{tot}^{dark}$, N_{ped}/N_{tot} and n_{in}. Subsequently, we return to the beginning of the flowchart loop and repeat this procedure for the next bias voltage.

[Figure 4-8] MPPC characterization and measurement program



- Measurement result: Figure 4-9 shows the result of measuring S13360-3050CS's PDE with our setup.

[Figure 4-9] A PDE measurement example (S13360-3050CS)



4-4. Dark count rate (DCR) and prompt crosstalk measurement using counter and CR filter

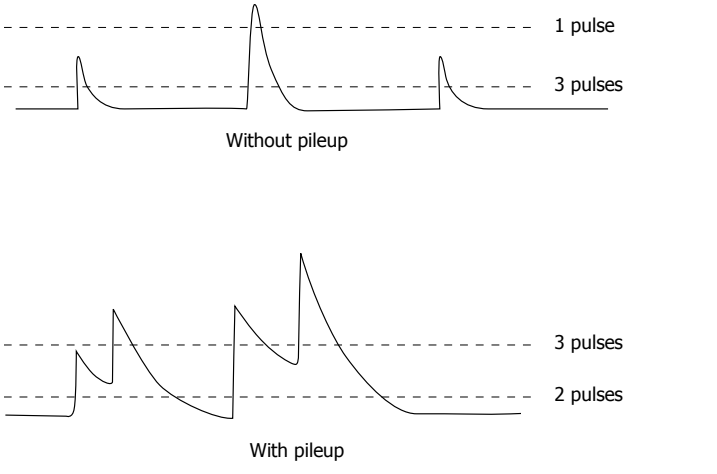
- Measurement principle: In this measurement, our goal is to collect the data for the step plot of DCR vs. discriminator threshold that is shown in Figure 1-26 of Section 1-3. From that plot, DCR can be obtained as the value of the first plateau, and the prompt crosstalk probability can be calculated by the ratio [(second plateau) / (first plateau)].

It is important to note, however, that pulse pileup, which is the overlapping of output pulses described in Section 2, can have a degrading effect on the accuracy of this measurement.

Figure 4-10 shows the effect of pulse pileup. In ideal cases, MPPC output pulses are completely separated, and from the view of counter, these pulses have completely discrete levels of pulse height with little randomness derived from gain fluctuation and white noise. However, if an MPPC has high DCR and/or long fall time because of large terminal capacitance and/or high probability of delayed crosstalk and afterpulses, pulse pileup under dark conditions can take place as shown in the lower part of Figure 4-10.

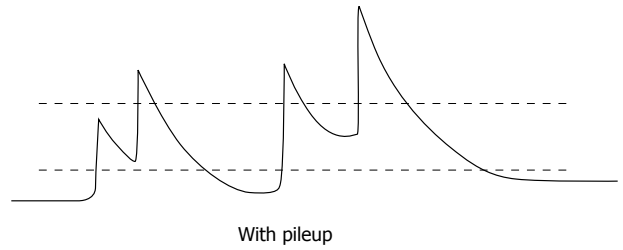
Due to pulse pileup, a lower than actual number of MPPC output pulses would be counted, and the resulting steps of DCR vs. discriminator threshold plot become less steep, hindering the accurate determination of the number of counts of each plateau. Furthermore, groupings of pulses can be counted as “one big pulse”, especially in the case of < 1 p.e. pulses whose population would be greater, and thus, the count of pulses at low threshold becomes smaller than the actual value. Generally speaking, these effects make prompt crosstalk probability larger than the actual value because the count of pulses exceeding the 0.5 p.e. threshold in height is diminished while the count of pulses whose heights exceed the 1.5 p.e. threshold is inflated.

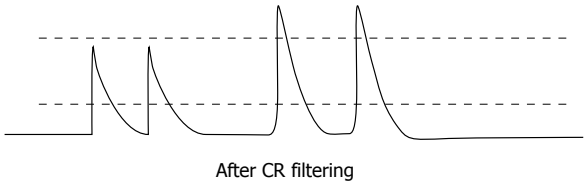
[Figure 4-10] Pulse pileup effect on measurements using counter



To exclude this degradation of measurement by pulse pileup, we can use a high-pass filter for pulse shaping after the linear amplifier’s output. As a result, output pulse widths become narrower and pileup is decreased as shown in Figure 4-11.

[Figure 4-11] Pulse pileup reduction using CR filter

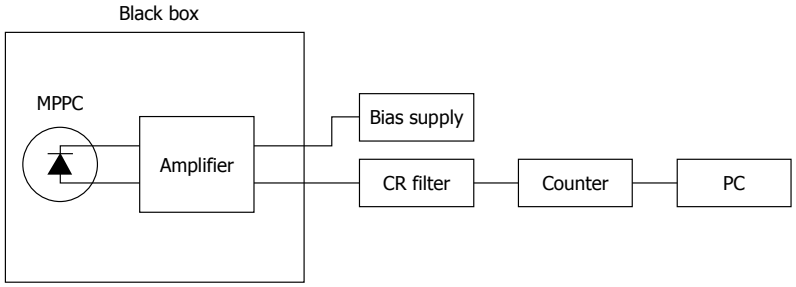




KAPDC0081EA

- **Measurement setup:** Figure 4-12 shows the measurement setup.

[Figure 4-12] DCR measurement setup using counter and band-pass filter



KAPDC0082EA

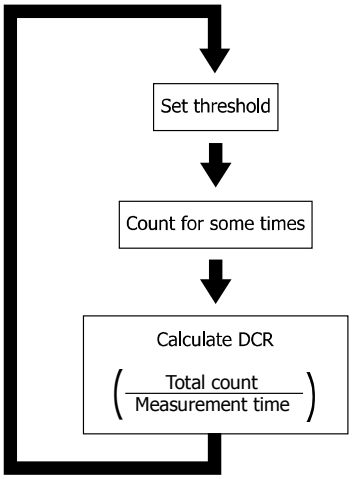
Amplifier: Linear amplifier (internal product and not for sale), bandwidth: 1 GHz
High-pass filter: internally made and caged in test box.
Counter: 53131A (Agilent)

In this setup, the linear amplifier should have sufficient bandwidth to replicate the MPPC rise time (as the highest-frequency component of the MPPC output) in order to effectively reject the pulse pileup effect. For a detailed explanation on how to determine the readout amplifer's optimal bandwidth, please refer to Section 2.

The time constant of high-pass filter should be set according to the pulse shape of the MPPC that you want to evaluate. For the purpose of characterizing S13360-3050CS, we use the high-pass filter with 100 pF capacitor and 51 Ω resistor.

As explained in Section 1, the high-pass filter decreases the output pulse height and so extra amplifier gain might be required for effective pulse output measurement. In that case, one solution is implementing another linear amplifier between the first stage and the high-pass filter. We use Mini-Circuit ZX60-14012L-S+ for this purpose.
- **Measurement Procedure:** Figure 4-13 shows flowchart of this measurement program's algorithm.

[Figure 4-13] Flow-chart of DCR and prompt crosstalk measurement program using a counter

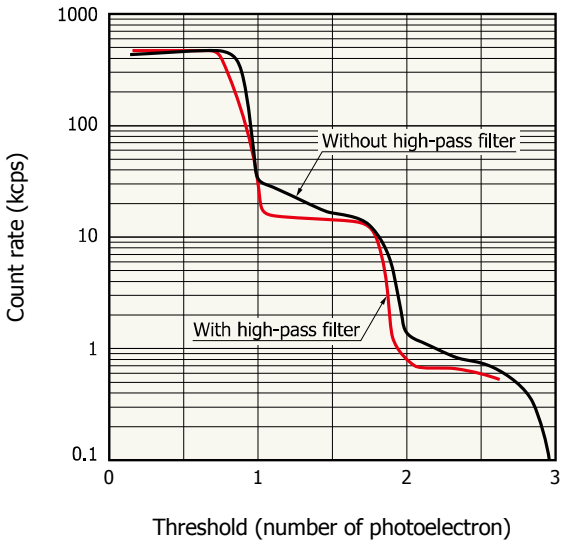


KAPDC0092EA

The counting integration time for each threshold value should be reasonably large enough to exclude statistical fluctuation of the measured count. For the case of S13360-3050CS, we set the integration time to 2 seconds. If DCR of the measured MPPC is small because of its small photosensitive area or low temperature or other factors, integration time should be set longer.

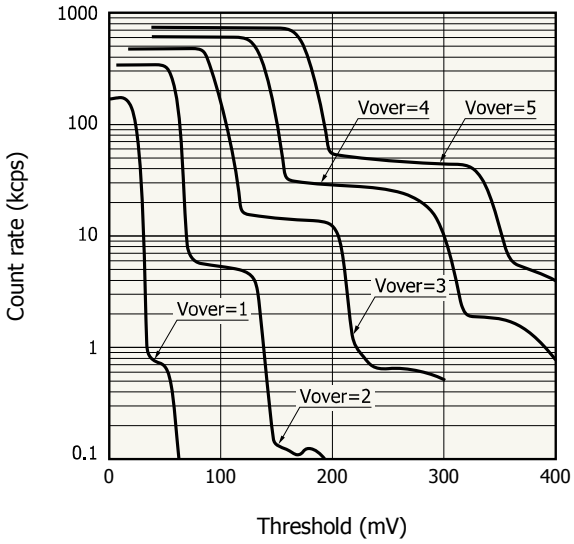
- **Measurement Result:** Figure 4-14 shows the measurement result for S13360-3050CS with and without high-pass filter. The high-pass filter's effect is clearly appreciable. Figure 4-15 shows the step plot for various overvoltages (with high-pass filter); we can see the increases in gain, DCR and prompt crosstalk probability. Figure 4-16 shows the DCR and prompt crosstalk probability obtained by this measurement.

[Figure 4-14] Comparison of the data with and without high-pass filter (S13360-3050CS)

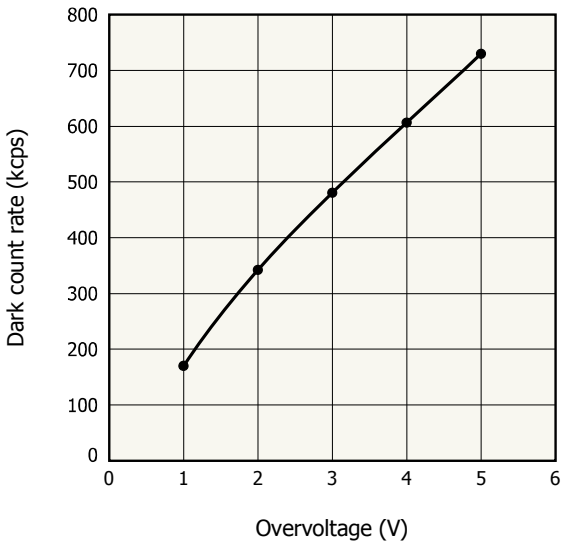


KAPDB0363EA

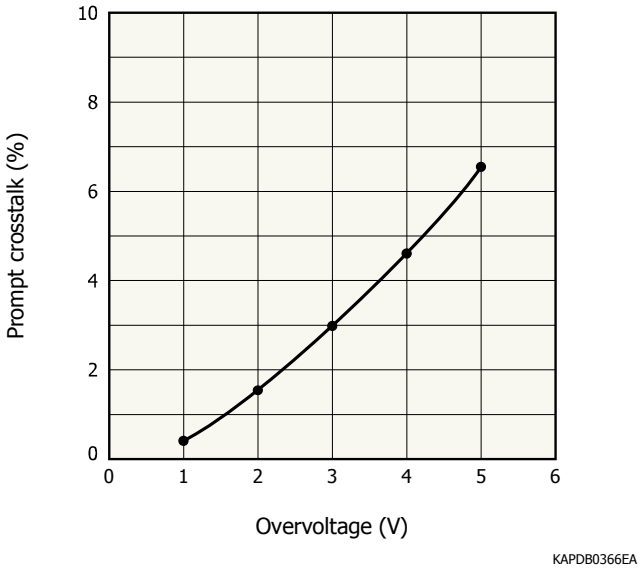
[Figure 4-15] Measurement results using counter (S13360-3050CS)



[Figure 4-16] DCR and prompt crosstalk measurement results (S13360-3050CS)
(a) Dark count rate vs. overvoltage



(b) Prompt crosstalk vs. overvoltage



4-5. Measurement of various MPPC characteristics using digitizer and digital pulse processing

- **Measurement principle:** In this measurement, we measure various MPPC charasteristics by detecting MPPC output events with digitizer and software processing of the resulting data.
This approach is commonly referred to as digital pulse processing (DPP).

For the sake of our discussion, we denote an “event” as data obtained within a time window that would contain MPPC output pulses (one or more) of any height. Specifically speaking, in the case of S13360-3050CS, the full width of a single output pulse is about 200 ns; we digitize the output waveform during a time window on the order of microseconds for DCR measurement and on the order of hundreds of nanoseconds for other measurements. You should choose the time window according to your measurement objective. A smaller time window leads to the reduction of valid events and consequently to the diminishment of afterpulse output whereas a larger time window results in a larger data size for one event and thus a decrease in measurement throughput .

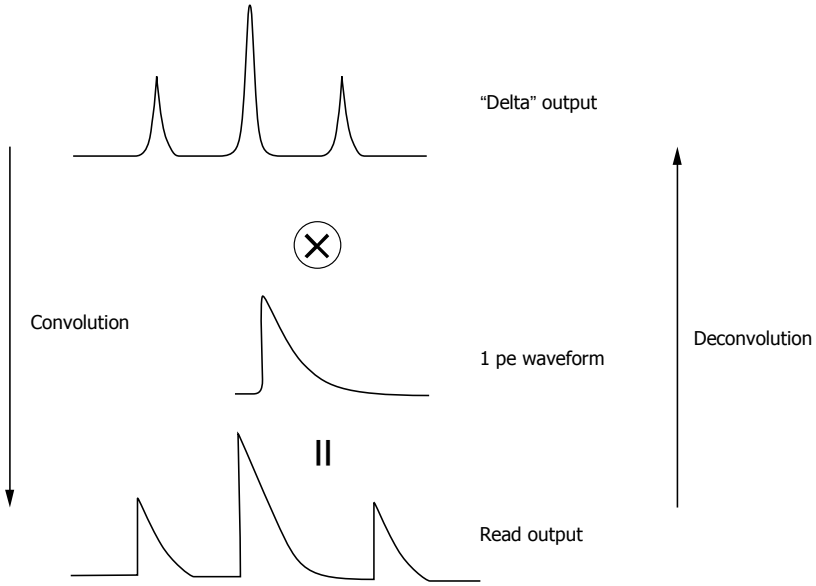
In a manner similar to the hardware solution (high-pass CR filter) described earlier, we resolve the issue of pulse pileup through software processing of an MPPC event using a deconvolution filter.
The basic concept of this method is well-known in the digital image processing field and described in references such as [12] while its practical application to the photosensor signal processing is described in [13] (in Japanese). There are other solutions for pulse pileup rejection such as simply differentiating the output and suppressing baseline discrepancy as described in [14].

A conceptual illustration of the deconvolution filter is shown in Figure 4-17. We assume that the MPPC output can be represented by the convolution of Dirac delta-function output events and the 1 p.e. MPPC waveform. The delta-function event is free from pileup since each such pulse would have a very sharp and distinct shape. If we define the specific shape of the delta waveform and measure the 1 p.e. waveform of the MPPC under measurement, we can mathematically compute the deconvolution filter that converts the real MPPC output events to delta-function events and thus reject the pulse pileup effect. Detailed mathematical explanation of how the deconvolution filter is computed has been provided in [13] and [14].

In addition to the deconvolution filter, we use another filter to reduce white noise , which disturbs the detection of delta-function peaks. Hamamatsu’s choice of such filters is the Wiener filter that determines the reduction in pulse height of a 1 p.e. MPPC output pulse as a function of frequency by computing the fourier transform of the 1 p.e.

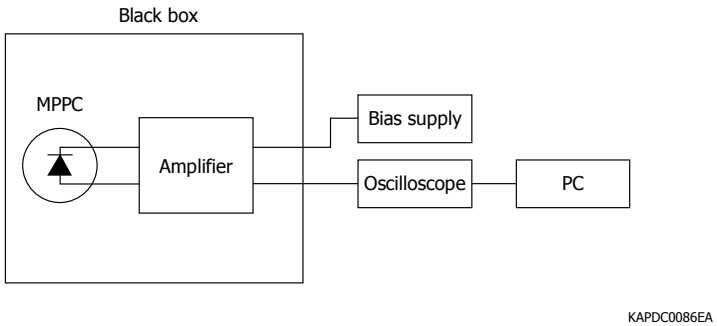
waveform in frequency domain after white noise filtration. Other methods are also effective; one of the simplest is to apply a low-pass filter to the event or equivalently set the oscilloscope bandwidth to a small value compared to the highest-frequency component of a 1 p.e. MPPC so that white noise can be filtered out while the MPPC output pulse maintains its amplitude.

[Figure 4-17] Basic concept of waveform deconvolution



- **Measurement Setup:** Figure 4-18 shows the measurement setup for this method. The MPPC is placed in a black box, and we proceed to recording the dark output events.
We use the oscilloscope as a digitizer in this measurement; board-level digitizers such as Flash ADCs are also suitable if they have sufficient node sensitivity, bandwidth, and sampling rate for precisely recording MPPC output pulse shapes. For our purpose, the 1 GHz bandwidth and 10 GS/s sampling rate of oscilloscope DPO7104 (Tektronix) is adequate.

[Figure 4-18] Measurement setup using a digitizer (oscilloscope)



- **Measurement procedure:** Before starting the data collection, we first form the deconvolution and wiener filters and then combine them.

Figure 4-19 shows how to make the deconvolution filter.
First, we obtain the MPPC's 1 p.e. waveform (the method of obtaining it is explained in the next subsection), and we then define the delta-function pulse shape based on the obtained 1 p.e. waveform. We afterwards use the

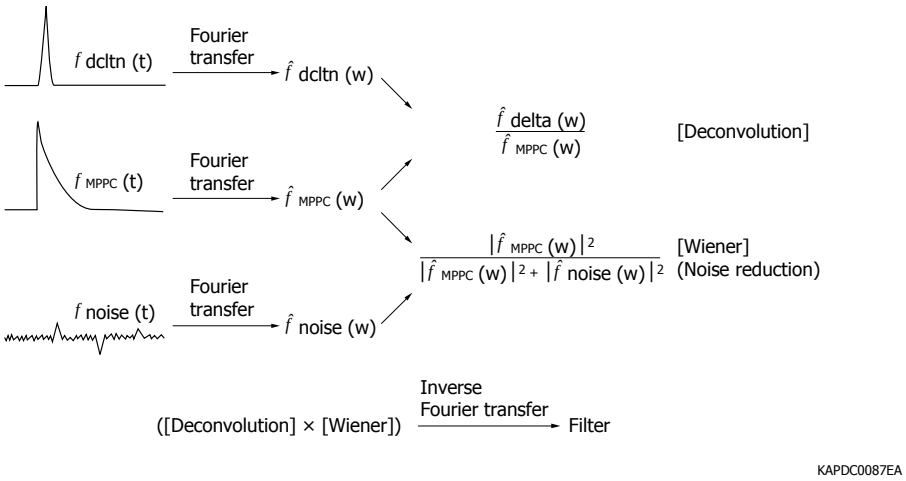
Blackman window function[7] as the delta function's shape and determine the window's optimal width by setting various widths and checking the deconvolution result for each of them. We then perform the Fourier transform operation on each waveform and calculate the ratio of its intensity for every frequency and obtain the deconvolution component of the filter.

We then proceed to building the wiener component of the filter by using the Fourier form of 1 p.e. pulse. We first obtain the white noise by recording the output waveform while applying a bias voltage slightly below the MPPC's breakdown voltage and calculating the Fourier transform of the resulting noise waveform. Using the two Fourier functions, we then calculate their power ratio as described in Figure 4-19 to obtain the Wiener component of the filter.

Finally, we form the product of deconvolution component and Wiener component and perform inverse Fourier transform on the product. The resulting waveform is the pulse processing filter we want.

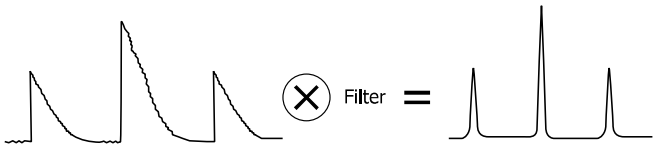
In this example, we use LabVIEW's library functions for the Fourier and inverse Fourier transform operations. You can use other math library functions prepared by other programming languages (such as MATLAB); there should be no differences in the results.

[Figure 19] Mathematical basis for creating a filter to cancel the effects of MPPC output pulse pileup and white noise



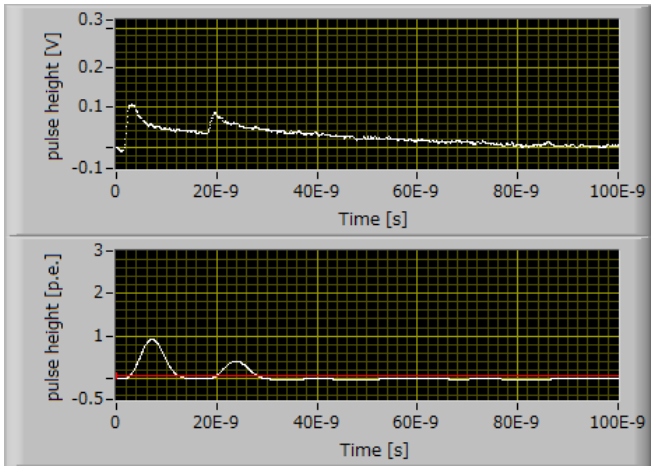
After creating the filter, we take a MPPC output event and convolute the entire event with the filter to obtain the delta pulse output. The effect of deconvolution is shown in Figure 4-20, and an example of it for S13360-3060CS is shown in Figure 4-21.

[Figure 4-20] "Delta" output with white noise reduced in a single step using the filter we made



[7] $w(t) = 0.42 - 0.5\cos 2\pi t + 0.08\cos 4\pi t,\ 0 \le t \le 1$

[Figure 4-21] Deconvolution example (S13360-3050CS)



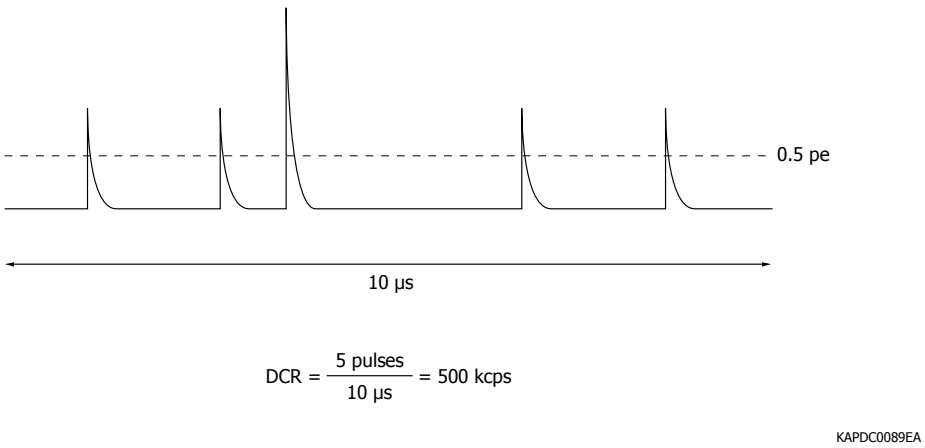
Processing the filtration's output can be performed by two approaches, which we now proceed to explain.

DCR and prompt crosstalk measurement:

In the first approach, the digitizer is triggered arbitrarily; an example of the trigger source would be 417 NIM pocket pulser (made by Phillips Scientific).

Figure 4-22 shows the concept of how to obtain DCR and prompt afterpulse probability in this measurement. To each MPPC output event recorded, we apply the aforementioned filtration in order to obtain a delta-function output event and then count the number of pulses in the event with a 0.5 p.e. discriminator threshold. The DCR can thus be obtained as the following simple ratio:
[(Number of pulses)/(Event time window)].
The prompt crosstalk can also be obtained easily by counting the number of pulses with a 1.5 p.e. threshold during the same time window and calculating the following ratio: [(number of >1.5 p.e. pulses)/(number of >0.5 p.e. pulses)].

[Figure 4-22] The principle of DCR measurement method using a digitizer



- Measurement Result: Figure 4-23 shows the prompt crosstalk probability obtained by this measurement. We also plotted these counting results with and without CR filter for comparison; the effect of pileup rejection can be observed.

[Figure 4-23] Prompt crosstalk vs. overvoltage

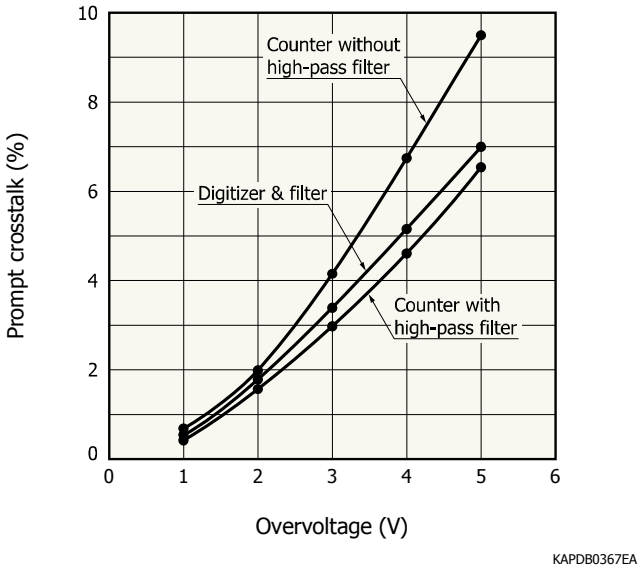


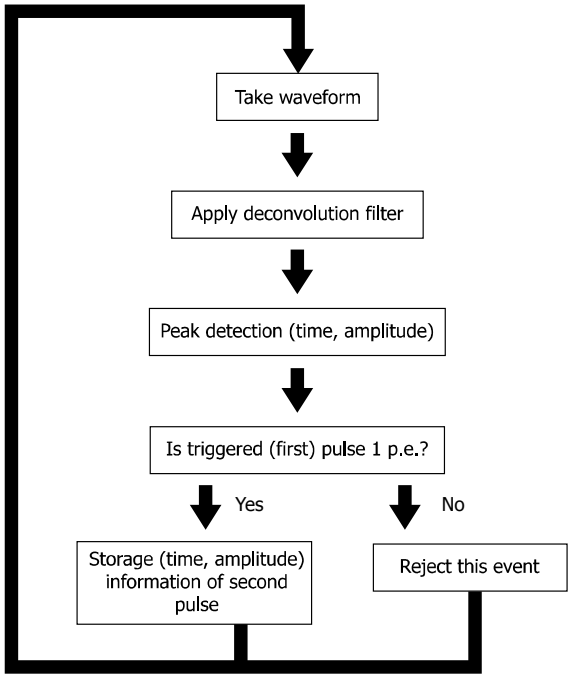
Figure 4-23: Prompt crosstalk comparison as measured by A. pulse counter with & without high-pass filter and B. digitizer. Use of a digitizer can cancel the effect of pulse pileup (as deduced from the linear slope) as would the using of a counter with high-pass filter.

Recovery time, afterpulse, and delayed crosstalk measurement:

Figure 4-24 shows flowchart of this measurement. In this case, we trigger the digitizer with MPPC output pulses. We then apply the filter to the output event, check whether the first pulse (used as trigger) has 1 p.e. pulse height, and reject the > 1 p.e. pulses because such pulses have differing probabilities of correlated noise pulses and hence disturb the following analysis. For the 1 p.e. pulse event, we then check arrival time and amplitude of next pulse and store the information into a memory array as part of this repetition loop.

The peak detection threshold after filtration should be as low as possible unless the remaining white noise reaches the threshold. In our example, we take the pulse height histogram of white noise output before beginning this loop, fit the histogram with Gaussian function, then take the 6σ value of the fitted function and use that value as the peak detection threshold.

[Figure 4-24] Flowchart of data taking routine for our measurement program using digitizer and deconvolution



Repeat until statistical error becomes sufficiently small for the intended purpose

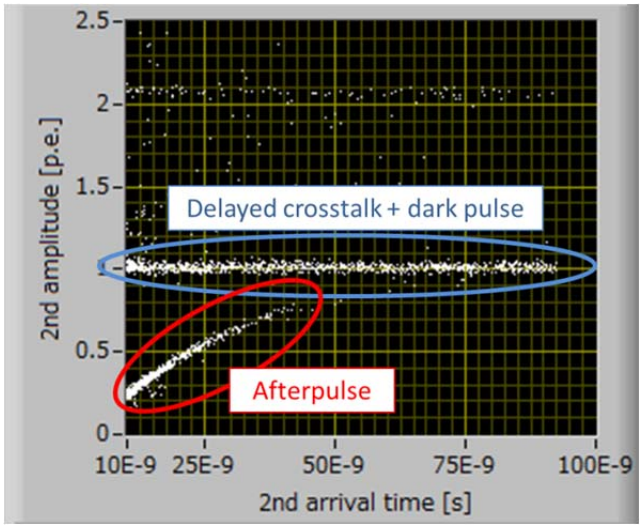
KAPDC0093EA

After data collection, we make a scatter plot of secondary pulses (arriving after the trigger pulse) in the manner shown in Figure 4-25 whose x-axis is the arrival time of a secondary pulse and y-axis is its amplitude. In practice, for verifying the measurement process's progress, we have created a measurement software program to show this scatter plot whenever needed.

We now study the two event groupings circled by red and blue circles in Figure 4-25. The red group most likely consists of afterpulse events of the first (trigger) pulse considering that their pulse heights are less than 1 p.e. whereas the blue group consists of a combination of delayed crosstalk and accidental dark pulses that are not related to the first pulse considering that their heights are about 1 p.e.

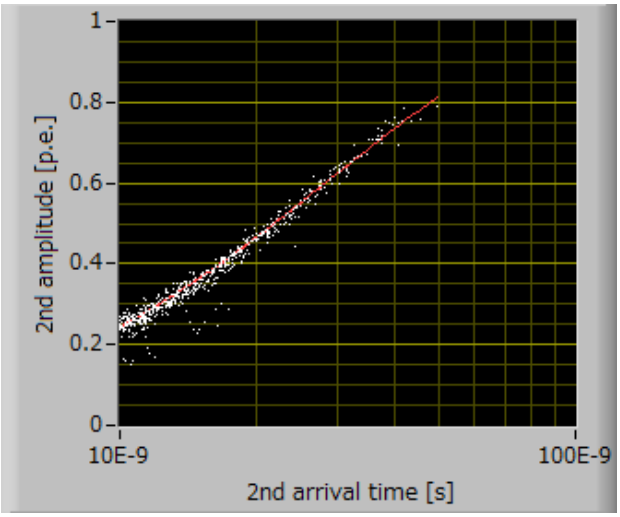
We now isolate these groups from the scatter plot and analyze them separately to obtain the MPPC characteristics.

[Figure 4-25] Scatter plot of timing and amplitude of secondary pulses



- Recovery Time Analysis: Figure 4-26 shows the scatter plot of afterpulse events. From this, we can obtain the recovery time of this MPPC by fitting this plot to an exponential function and calculate its time constant. By this analysis, we obtain the recovery time of S13360-3050CS to be 30ns.

[Figure 4-26] Trend line of afterpulse events of Figure 4-25 (circled in red)

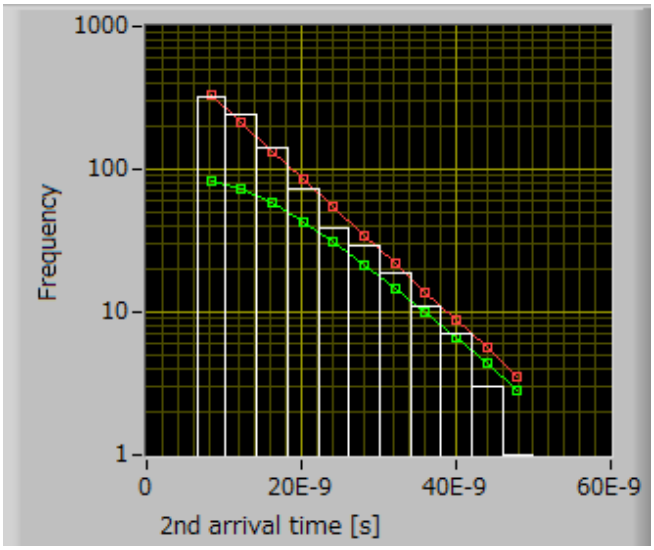


The red line shows the exponential fitting of those events.

- Afterpulse Analysis: Because of its nature, definition of the afterpulse probability is rather subjective due to the wide range of pulse heights and time delays that could constitute its definition. Thus, the concrete definition of MPPC afterpulse is dependent on the end-use application conditions and requirements. If the MPPC output current is read out continuously, all afterpulses and their charge contents will be integrated and included in the MPPC output current. On the other hand, when the MPPC output is read out as individual pulses, afterpulses whose pulse heights are lower than the counting discriminator threshold are ignored and would not be taken into account for assessing afterpulse probability; this means that measurement and quantification of afterpulse probability to a large extent depend on what lower-level discriminator threshold is used in the counting setup.

In the following, two definitions and analysis methods of afterpulse are showcased. The first definition simply sums the afterpulse events and calculates the percentage ratio of that sum to the total number of events. By this definition, afterpulse probability only counts the number of afterpulses whose pulse height exceeds the threshold level of the counter's discriminator for pulse detection.
Hence, in specifying afterpulse probability based on this definition, the discriminator threshold level should be referenced.
An alternative but more complex definition utilizes a weighted integral of afterpulse events based on their pulse heights. Towards that, we first make a histogram of afterpulse events according to their arrival times (x-axis of Figure 4-26) as shown in Figure 4-27. Assuming that afterpulse probability has an exponential distribution, we fit this histogram to an exponential function to obtain the afterpulse probability distribution curve vs. arrival time. Afterwards, we calculate the product of this curve with the exponential function whose time constant corresponds to the recovery time (explained in preceding subsection) and proceed to compute the area under the resulting curve. Finally, we obtain the afterpulse probability as the ratio of this area to the total number of triggered events.

[Figure 4-27] Histogram of afterpulse events

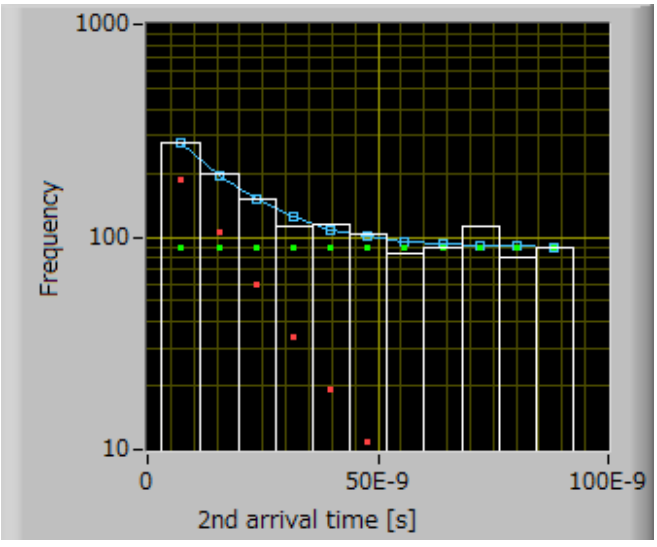


Note that compared to Figure 4-26, the y-axis now changes to frequency. Red line represents the exponential fitting of the histogram while the green line shows the product of afterpulse fitting and recovery time curve.

- **Delayed crosstalk analysis:** Figure 4-28 shows the histogram of the mixed events of delayed crosstalk and accidental dark pulses consisting of the events circled in blue in Figure 4-25. Different from afterpulse analysis, the resulting histogram does not appear as an exponential curve because it has two exponential components with different time constants.

To differentiate these, we have to fit this histogram to the sum of two exponential functions with different parameters. After a proper fitting is carried out, we can obtain the total number of delayed crosstalk events by calculating the area under the fitted exponential that corresponds to the delayed crosstalk component. Then, we can calculate the delayed crosstalk probability as the ratio of that area to the total number of triggered events.

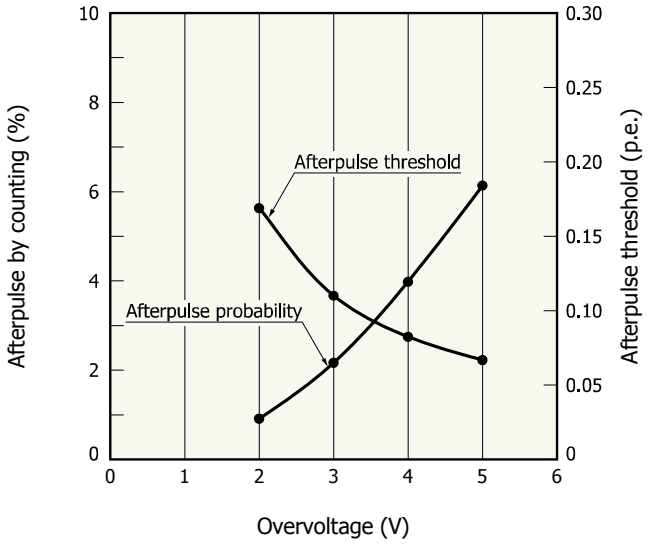
[Figure 4-28] Histogram of mixed (delayed crosstalk and dark pulse) events



Red and green dots represent two exponential components.

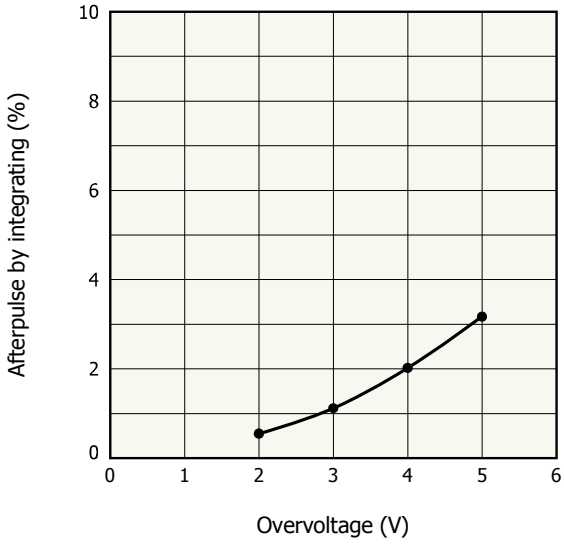
- **Measurement Results:** Figure 4-29 shows the probabilities of delayed crosstalk and afterpulsing of S13360-3050CS for various over-voltages.

[Figure 4-29] Afterpulse and delayed crosstalk probabilities (S13360-3050CS)
(a) Afterpulse by counting vs. overvoltage



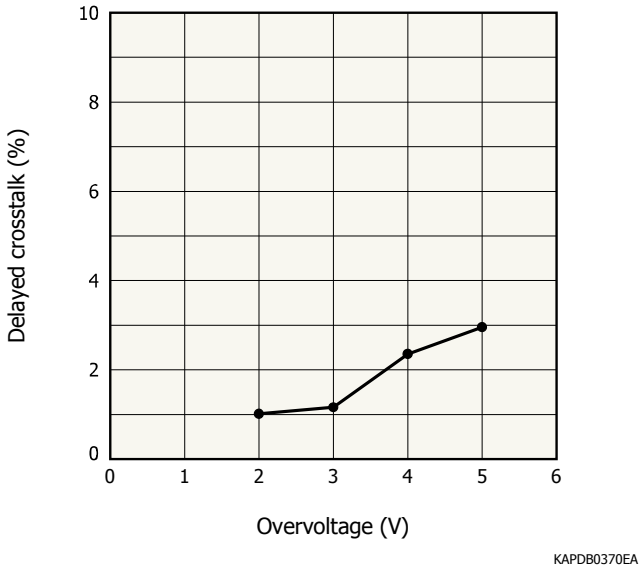
KAPDB0368EA

(b) Afterpulse by integrating vs. overvoltage



KAPDB0369EA

(c) Delayed crosstalk vs. overvoltage



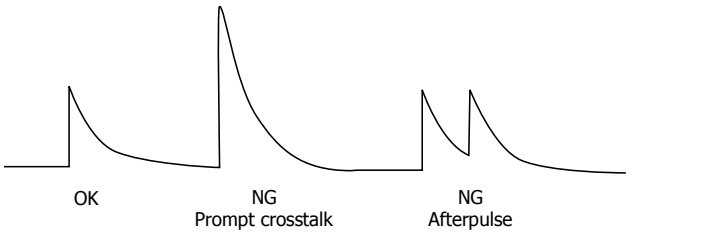
4-6. Single-photon pulse shape measurement

- **Measurement principle:** Here we will explain how to take the single-photon/photoelectron (1 p.e.) waveform of the MPPC. The 1 p.e. waveform is essential to making the filter required for digital pulse processing and utilization of digitizer in MPPC characterization.

- **Measurement setup:** The setup for this measurement would be basically similar to that of measurements using a digitizer. One difference is that the bandwidth of each measurement setup component must be high enough to avoid degrading the high frequency component of MPPC output waveform. In this case, the maximum bandwidth of 1 GHz is sufficient for taking the 1 p.e. waveform of S13360-3050CS. Another difference is that the min. data collection time window would be set by the full width of the 1 p.e. waveform. In this case of S13360-3050CS, a time window of 200 ns would be adequate.

- **Measurement procedure:** We obtain the MPPC's output waveform. Then, after checking the number of pulses in the waveform and their heights, we decide whether the triggered pulse is a genuine 1 p.e. waveform or not (i.e. prompt/delayed crosstalk or afterpulse following a dark pulse) and store an output pulse only when it is confirmed to have a 1 p.e. waveform. Afterwards, this step must be repeated a sufficient number of times to ensure the elimination of white noise by averaging all stored 1 p.e. waveforms.

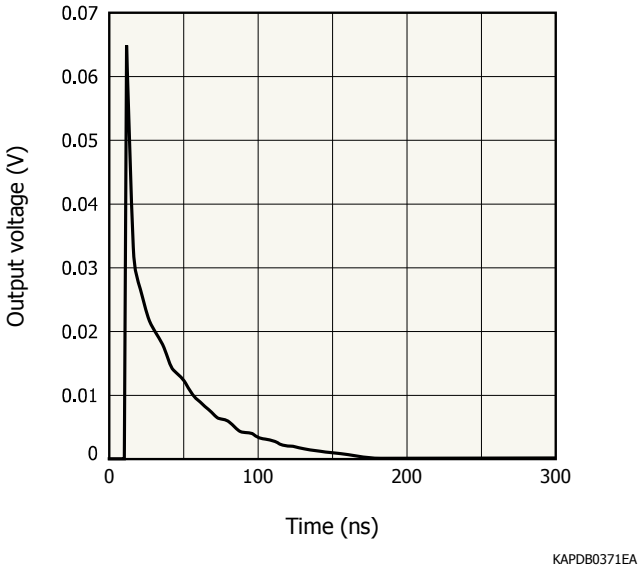
[Figure 4-30] "Genuine" 1 p.e. waveform of dark pulse



Others should be excluded from 1 p.e. waveform storage and accumulation.

- **Measurement result:** Figure 4-31 shows the result of measuring the 1 p.e. waveform of S13360-3050CS.

[Figure 4-31] 1 p.e. waveform (S13360-3050CS)



4-7. Time resolution ability

Photodetector time resolution characteristics are critical to attaining proper performance in direct-TOF (time-of-flight) applications (such as TOF-PET [positron emission tomography] in medical imaging or LiDAR [light (imaging), detection, and ranging] in 3D scanning to determine the arrival times of signal photons or radiation. In this subsection, we will describe measuring two particularly-relevant time resolution characteristics: coincidence time resolution (CTR) and single photon time resolution (SPTR).

- CTR:

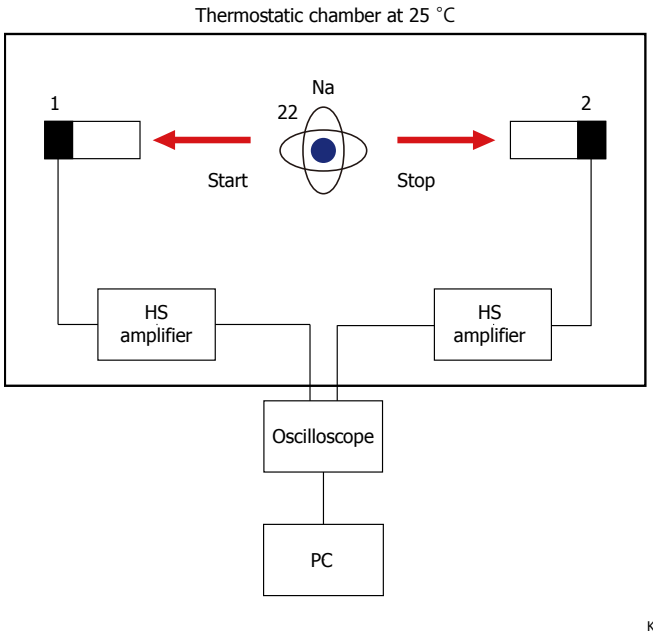
➢ Measurement principle

In TOF-PET, CTR is measured by using two detector channels, each consisting of a pair of scintillator crystal and photodetector, to detect gamma rays, having 511 keV in energy, that are emitted simultaneously from a positron decay and the resulting electron-positron annihilation.
The measurement setup is shown in Figure 4-32. The detection times obtained from detector channels 1 and 2 should ideally show the same value if each detector has the same distance from the radiation source as the other. However, variations in underlying factors such as the scintillator's light emission yield, the photodetector's internal charge collection efficiency, photon detection efficiency, and output pulse shape, and the output amplifier's characteristics cause some timing variation between the two detector channels. The probability distribution curve of the timing difference between detector channels 1 and 2 has a bell-shaped Gaussian-like profile. CTR is theoretically defined by the FWHM of this distribution profile.

The result of this measurement does not represent the time resolution of the photodetector (i.e. MPPC) itself but the overall system time resolution, including the scintillator, amplifier, trigger circuitry, and other instruments utilized besides the photodetector.

➢ Measurement setup

[Figure 4-32] CTR (coincidence resolving time) measurement setup



➢ Equipment

Radiation source: Na-22, etc.
Scintillator: LFS, LSO, LYSO, etc.
Amplifier: linear amplifier (internal product and not for sale), bandwidth: 1 GHz
Oscilloscope: SDA 760Zi (LeCroy)

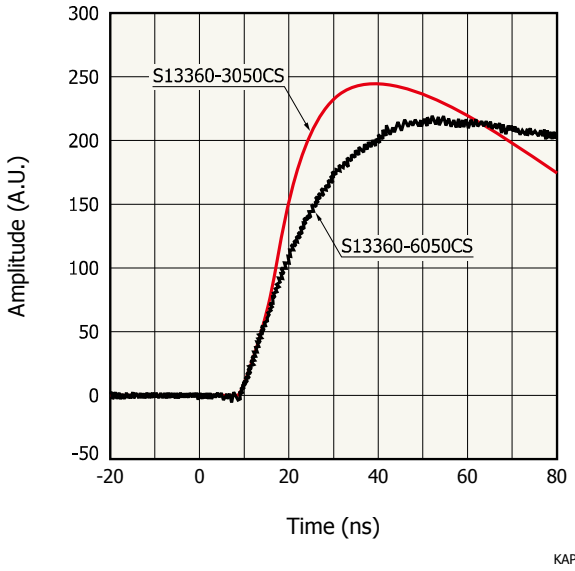
➢ Procedure

The radioisotope should be positioned in the middle of detectors to equalize its distance from them. Originating from the annihilation of the positron byproduct of Na-22's radioactive decay, gamma rays (if emitted in the proper opposite directions) reach detector channels 1 and 2 at the same time. The gamma rays undergo scintillation once detected with the emitted scintillation light in turn detected by each photodetector (i.e. MPPCs). Generally speaking, CTR measurement is limited by the system's timing jitter. To measure CTR with a specific target time jitter, the utilized output amplifier circuitry must have sufficient bandwidth to allow the jitter's proper detection.

The amplified output pulses are then read out by a high-speed oscilloscope. The oscilloscope can calculate the time differences between coincident pairs of pulses outputted by the two detector channels, can profile the distribution of those differences (as obtained from a sufficiently-large population of such coincident pulses), and then produce the FWHM of that distribution in order to determine the measured CTR. The utilized oscilloscope should also have adequately wide input bandwidth and high sampling rate in order to resolve picosecond timing. One can also use a combination of TAC and MCA instead of a wide-bandwidth high-speed oscilloscope.

Figure 4-33 shows the linear output amplifier's output pulse shape coming from two Hamamatsu S13360 series MPPCs. The key to high-accuracy measurement results is to decrease baseline fluctuations as much as possible. The optimal way to do so would depend on the measurement setup and the target application, but would consist of controlling the trigger threshold level, the bias voltage applied to the MPPC, and the measurement's optical conditions amongst other factors.

[Figure 4-33] Output pulse form linear output amplifier



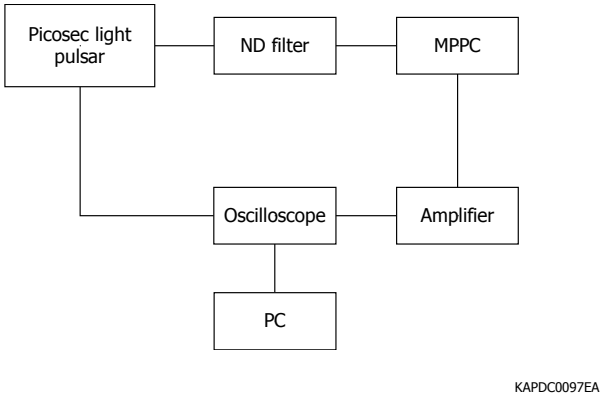
- SPTR:

➢ Measurement principle

SPTR is a parameter that characterizes the MPPC's time resolution alone; this is in contrast to CTR, which evaluates a system's overall time resolution performance. The basic idea of this measurement is to illuminate the MPPC with a single-photon input signal light level and to then determine the detection times of single photons by recording the MPPC's output pulses (using a 1 p.e. trigger threshold level). The measurement setup is shown in Figure 4-34.

The suitable light source for this measurement would be able to produce light pulses of picosecond width; Hamamatsu's PLP-10 would be one such light source.

[Figure 4-34] SPTR measurement setup



➢ Equipment

Light source: PLP-10 (Hamamatsu Photonics K.K.)
ND filters
Amplifier: linear amplifier (internal product and not for sale), bandwidth: 1 GHz
Oscilloscope: SDA 760Zi (LeCroy)

➢ Procedure

The basic measurement method is the same as CTR except that SPTR measurement relies on an output pulse height of only 1 p.e. Once a light intensity that results in majority 1p.e. output pulses is attained, pulse heights greater than 2 p.e. and lower than 1pe, which would correspond to crosstalk and afterpulsing, must be excluded by using upper- and lower-level discriminator thresholds set at 1.5 p.e. and 0.5 p.e. levels. This measurement requires careful cabling to minimize parasitic noise as SPTR is easily affected by baseline deteriorations.

The variation of the timing information obtained by this method includes not only MPPC's pulse time jitter but also the light source's pulse emitting jitter. Thus, MPPC's SPTR is derived as:

$$SPTR(MPPC) = \sqrt{(measured\ SPTR)^2 - (PLP\ jitter)^2} \qquad \textbf{(4-1)}$$

4-7. Dynamic range and linearity measurement:

As discussed in Section 2, MPPC's linearity is limited by its PDE, pixel count per unit of area, and recovery time. In this subsection, we will describe a generalized measurement method to estimate the MPPC's pulse-rate linearity.

➢ Measurement Principle

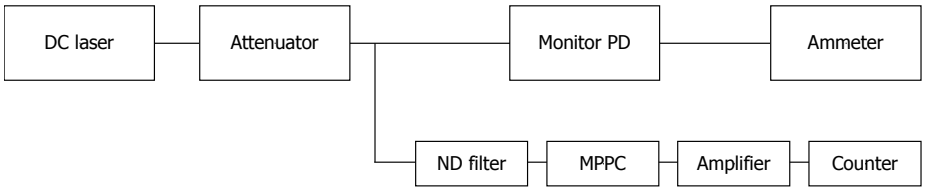
Before we discuss the measurement procedure, let's recall two key aspects of MPPC linearity and dynamic range from Section 1: First, a MPPC pixel cannot distinguish the number of simultaneously-incident photons upon it; the pixel response is limited to outputting a 1 p.e. pulse whose height does not depend on the number of incident photons. Second, each pixel requires a certain duration of recovery time. A photon's incidence onto a pixel after a preceding photon's detection cannot cause a full-height secondary output signal within the recovery time.

The first aspect (assuming we're restricted to a fixed pixel size to maintain PDE) can be alleviated by dispersing the input light signal (using a diverging lens or by other means) to decrease the number of incident photons per unit area onto the MPPC. However, both aspects can be addressed by using a MPPC of smaller pixel size; such a MPPC would have an increased count of pixels per unit of area (i.e. more pixels within the same optical field-of-view) available to detect incident photons and a decreased pixel recovery time due to lower pixel capacitance. However, these advantages have a tradeoff of loss in PDE due to the lower fill-factor of a smaller pixel; thus, the lower limit of the linearity of such a MPPC is shifted upwards.

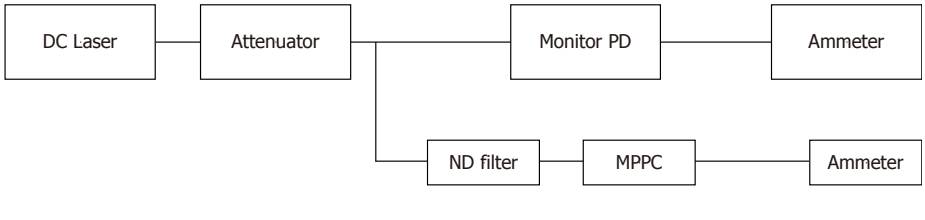
In order to measure MPPC linearity, we use the photon counting technique for the lower levels of the MPPC's linearity where the signal level intensities are low and the current measurement technique (similar to measuring I/V curve) for the upper levels of MPPC's linearity. In this measurement, we use a DC light source for implementing both techniques; we control the light level using an optical attenuator and ND filters. A key issue is accurate determination of the light intensity that the MPPC has detected at each obtained data point. This is accomplished by precise monitoring of the incident light intensity using a calibrated monitor photodiode. A peculiar aspect of doing so is in the photon-counting regime: the light intensity incident on the MPPC is adjusted to reach that regime using a precise ND filter while the calibrated photodiode can monitor the total (pre-attenuation) light intensity present with high accuracy.

➢ Measurement setup

[Figure 4-35] Measurement setup (Low intensity light level)



[Figure 4-36] Measurement setup (High intensity light level)



➢ Equipment

Light source: CW Laser
Optical Attenuator
ND filter
Amplifier: linear amplifier (internal product and not for sale), Bandwidth: 1GHz
Counter: 53131A (Agilent)

➢ Procedure

We first measure and calibrate the ND filter's attenuation rate using the calibrated photodiode.
We recommend an ND filter that has 2-3 orders of magnitude attenuation capability. In the following measurement, the attenuation rate obtained here is used to estimate the light intensity incident onto the MPPC. One can adjust that light intensity over the MPPC's full dynamic range (from photon counting to saturation) by controlling the optical attenuator.

✓ Low intensity light level

While the light intensity on the MPPC is at the photon-counting level, one can measure the MPPC's output response linearity by gradually increasing the light level. This task is similar to measuring dark counts but with the lower-level discriminator's threshold set to be 0.5 p.e. Here, we define "photon counting level" to be a light intensity that results in MPPC output pulse heights of only 1 p.e. with enough inter-pulse time interval for continual detection of individual incident photons. Figure 4-35 shows the corresponding setup.

✓ High intensity light level

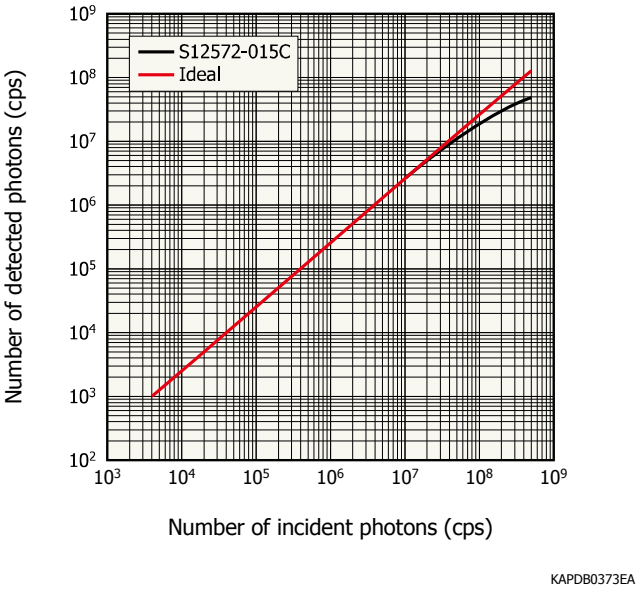
We switch to the current measurement technique once the light intensity becomes high enough to result in the occurrence of ≥ 2 p.e. output pulse heights in order to avoid counting errors due to pulse pileup. Figure 4-36 shows the corresponding setup. In this measurement, the MPPC's output current is measured using an ammeter, which is then followed by subtraction of dark current in order to obtain the net photon-correlated current. As increases in incident light intensity continue, one will eventually see the output's saturation.

The combined linearity plot (of low and high input light intensities) is obtained by using the pulse count rate of the low-intensity measurement (by dividing the integrated count of detected 1 p.e. pulses at each data point by the

integration time used for measuring each data point) and of the high-intensity measurement (by dividing the measured net (dark-subtracted) current by the charge output corresponding to a 1 p.e. pulse).

Figure 4-37 is the measured linearity plot of a Hamamatsu MPPC with a 15 µm pixel pitch. In this plot, the light level is indicated as number of incident photons per unit of time (in counts per second or cps).

[Figure 4-37] Measurement result of dynamic range (S12572-015C)



References

Section 1

[1] Donald A. Neamen, *Semiconductor Physics and Devices*, 3rd ed. New York, NY: McGraw-Hill, 2003.

[2] Ben G. Streetman, *Solid State Electronic Devices*, 3rd ed. Englewood Cliffs, NJ: Prentice Hall, 1990.

[3] John Gowar, *Optical Communication Systems*, London, UK: Prentice Hall Intl. Ltd., 1984.

[4] Silvano Donati, *Photodetectors: Devices, Circuits, and Applications*, Upper Saddle River, NJ: Prentice Hall, 2000.

[5] S. M. Sze, *Physics of Semiconductor Devices*, 2nd ed. New York, NY: John Wiley & Sons, 1981.

[6] D. Renker and E. Lorenz, "Advances in Solid State Photon Detectors" *Journal of Instrumentation*, volume 04, issue 04, pp. 04004, April, 2009.

Section 2

[7] James W. Nilsson, Susan A. Reidel, *Electric Circuits*, 6th ed. Englewood Cliffs, NJ: Prentice Hall, 2000.

[8] Tudor E. Jenkins, *Optical Sensing Techniques and Signal Processing*, London, UK: Prentice Hall Intl. Ltd., 1987.

[9] M. Grodzicka, et al. "MPPC Array in the Readout of CsI:Tl, LSO:Ce:Ca, LaBr :Ce, and BGO Scintillators" *IEEE Trans. Nucl. Sci.*, vol. 59, no. 6, Dec. 2012.

[10] Helmuth Spieler, "Fast Timing Methods for Semiconductor Detector" *IEEE Trans. Nucl. Sci.*, vol. NS-29, no. 3, pp. 1143-1158, June, 1982.

Section 4

[11] P. Eckert, et al. (2010, April 1). *Characterization Studies of Silicon Photomultipliers* (2nd ed.) [Online]. Available: https://arxiv.org/abs/1003.6071

[12] Steven W. Smith, *The Scientist & Engineer's Guide to Digital Signal Processing*, San Diego, CA: California Technical Publishing, 1997.

[13] H. Oide, "半導体光検出器 PPD の基本特性の解明と，実践的開発に向けた研究," M.S. thesis in Japanese, 2009. [Online]. Available: https://www.icepp.s.u-tokyo.ac.jp/papers/ps/thesis/master/oide_mthesis.pdf

[14] Adam N. Otte, et al. (2016, June 16). *Characterization of Three High Efficiency and Blue Sensitive Silicon Photomultipliers* [Online]. Available: https://arxiv.org/abs/1606.05186

[15] J. Rosado, S. Hidalgo. (2015, Oct. 21). *Characterization and Modeling of Crosstalk and Afterpulsing in Hamamatsu Silicon Photomultipliers* (2nd ed.) [Online]. Available: https://arxiv.org/abs/1509.02286

HAMAMATSU

www.hamamatsu.com

HAMAMATSU PHOTONICS K.K., Solid State Division
1126-1 Ichino-cho, Higashi-ku, Hamamatsu City, 435-8558 Japan, Telephone: (81) 53-434-3311, Fax: (81) 53-434-5184
U.S.A.: Hamamatsu Corporation: 360 Foothill Road, Bridgewater, N.J. 08807, U.S.A., Telephone: (1) 908-231-0960, Fax: (1) 908-231-1218
Germany: Hamamatsu Photonics Deutschland GmbH: Arzbergerstr. 10, D-82211 Herrsching am Ammersee, Germany, Telephone: (49) 8152-375-0, Fax: (49) 8152-265-8
France: Hamamatsu Photonics France S.A.R.L.: 19, Rue du Saule Trapu, Parc du Moulin de Massy, 91882 Massy Cedex, France, Telephone: 33-(1) 69 53 71 00, Fax: 33-(1) 69 53 71 10
United Kingdom: Hamamatsu Photonics UK Limited: 2 Howard Court, 10 Tewin Road, Welwyn Garden City, Hertfordshire AL7 1BW, United Kingdom, Telephone: (44) 1707-294888, Fax: (44) 1707-325777
North Europe: Hamamatsu Photonics Norden AB: Torshamnsgatan 35 16440 Kista, Sweden, Telephone: (46) 8-509-031-00, Fax: (46) 8-509-031-01
Italy: Hamamatsu Photonics Italia S.r.l.: Strada della Moia, 1 int. 6, 20020 Arese (Milano), Italy, Telephone: (39) 02-93581733, Fax: (39) 02-93581741
China: Hamamatsu Photonics (China) Co., Ltd.: B1201, Jiaming Center, No.27 Dongsanhuan Beilu, Chaoyang District, Beijing 100020, China, Telephone: (86) 10-6586-6006, Fax: (86) 10-6586-2866

Cat. No. KAPD9005E01 Mar. 2017