Available online at www.sciencedirect.com
ScienceDirect
Nuclear Instruments and Methods in Physics Research A 567 (2006) 48–56
NUCLEAR INSTRUMENTS & METHODS IN PHYSICS RESEARCH Section A
www.elsevier.com/locate/nima

Geiger-mode avalanche photodiodes, history, properties and problems

D. Renker*

Paul Scherrer Institute, 5232 Villigen, Switzerland

Available online 23 June 2006

Abstract

Geiger-mode avalanche photodiodes (G-APDs) have been developed during recent years and promise to be an alternative to photomultiplier tubes. They have many advantages like single photon response, high detection efficiency, high gain at low bias voltage and very good timing properties but some of their properties, the dark count rate for example, can be a problem. Several types of G-APDs are on the market and should be selected carefully for a given application.

PACS: 29.40.Pe; 29.40.Mc; 85.60.Dw

Keywords: SiPM; Photosensor; Avalanche Photodiode; Geiger-mode Avalanche Photodiode

1. Introduction

The detection of light with very low intensity became possible almost one century ago when in the year 1913 Elster and Geiter invented the photoelectric tube only few years after Einstein formulated 1905 the photoelectric workfunction [1]. It took more than 20 years until the first photomultiplier tube (PMT) was invented in the RCA laboratories and became 1936 a commercial product. Single photons were detectable from now on. Further innovations have led to highly sophisticated devices available nowadays. PMTs have two severe handicaps: They are very sensitive to magnetic fields and their price is high because the complicated mechanical structure inside the vacuum container is mostly handmade.

Many modern experiments in high-energy physics have calorimeters with high granularity inside a solenoid with strong magnetic fields. This forced the search for an alternative to PMTs.

Very successful is the PIN photodiode, which is used, in most big experiments in high-energy physics (CLEO, L3, BELLE, BABAR, GLAST). PIN describes the structure and stands for p-, intrinsic and n-silicon. Even with a state of the art charge sensitive amplifier, which is needed because the device has no internal gain, the noise is at the level of several hundred electrons and consequently the smallest detectable light flash needs to consist of even more photons.

Avalanche photodiodes (APDs) have internal gain which improves the signal to noise ratio but still some 20 photons are needed for a detectable light pulse. The excess noise factor, the fluctuation of the avalanche multiplication, limits the useful range of the gain. CMS is the first big experiment that uses avalanche photodiodes.

At the beginning of this millennium the Geiger-mode avalanche photodiode (G-APD) has been developed. This device can detect single photons like a PMT and therefore some people call it Silicon PhotoMultiplier, SiPM. The pulse height spectrum measured with a G-APD shows a resolution which is even better than what can be achieved with a hybrid photomultiplier tube (Fig. 1).

2. History

Pioneering work in the development of solid state single photon detectors was done in the 1960s in the RCA company by McIntyre [3] and by Haitz in the Shockley research laboratory [4]. Avalanche photodiodes operated in linear- and in Geiger-mode were in the sixties and early seventies a very active field of experimental and theoretical

*Tel.: +41 56 3104213; fax: +41 56 3105230.
E-mail address: dieter.renker@psi.ch.

0168-9002/$ - see front matter © 2006 Elsevier B.V. All rights reserved.
doi:10.1016/j.nima.2006.05.060

research. A model of the behavior was developed and verified with test structures (Fig. 2).

In Japan a solid-state single photon detectors with a somewhat different design with a horizontal development of the avalanche (Fig. 3) was patented in 1972 [5].

The performance of the first devices operated in Geiger-mode with a bias voltage several Volts higher than the breakdown voltage was not very good but single photons have been seen and with improving technology the development was leading to the so-called Single Photon Avalanche Diode (SPAD) and to the SLIK™ structure produced by Perkin–Elmer.

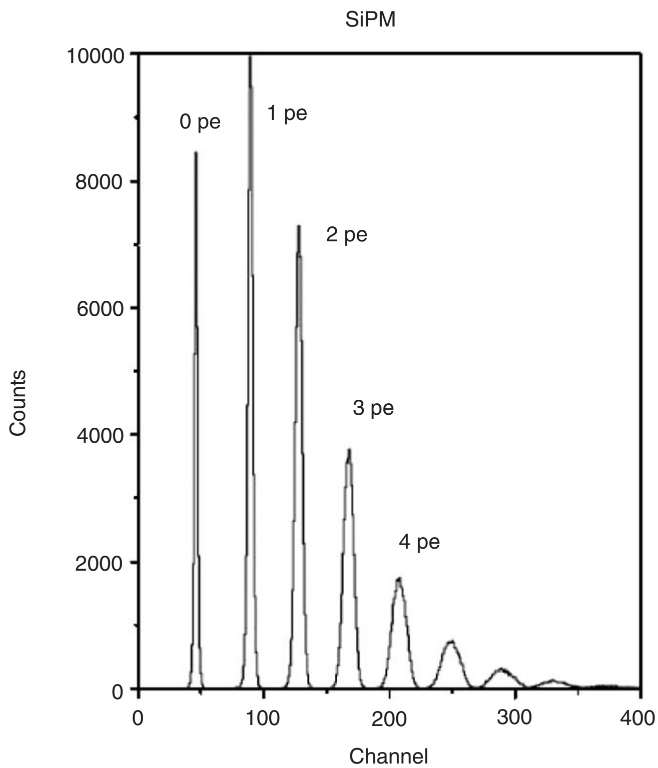

Fig. 1. Pulse height spectrum of light pulses with very low intensity recorded with a G-APD. Taken from Buzham et al. [2].

The quenching of the breakdown was done passively. When the current fluctuations happen to go to zero the breakdown stops and needs a new triggering event to start again. The devices were therefore slow and the maximal count rate was smaller than 100 kHz. This is still the case for state of the art devices nowadays. Only the development of active quenching circuits allows high-count rates of more than 1 MHz and provides a short dead time [6].

Radiation Monitor Devices Inc. (RMD) developed an array of APDs with single photon detection capability for DIRC applications (Detection of Internally Reflected Cherenkov light). It consists of 6×14 individual APDs with a size of $150 \times 150\,\mu m^2$. It is operated in Geiger-mode and has an active quenching circuit for each APD [7].

In the Rockwell International Science Center, Stapelbroek et al. developed 1987 the Solid State PhotoMultiplier (SSPM). This is an APD with very high donor concentration which creates an impurity band 50 meV below the

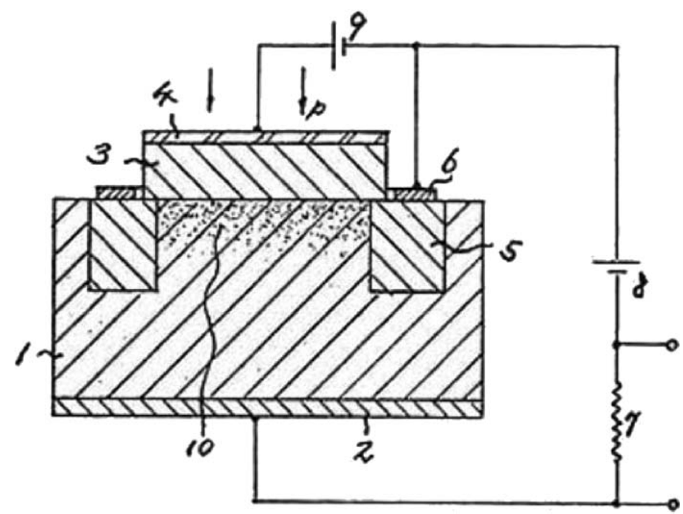

Fig. 3. Structure of the detector developed in Japan.

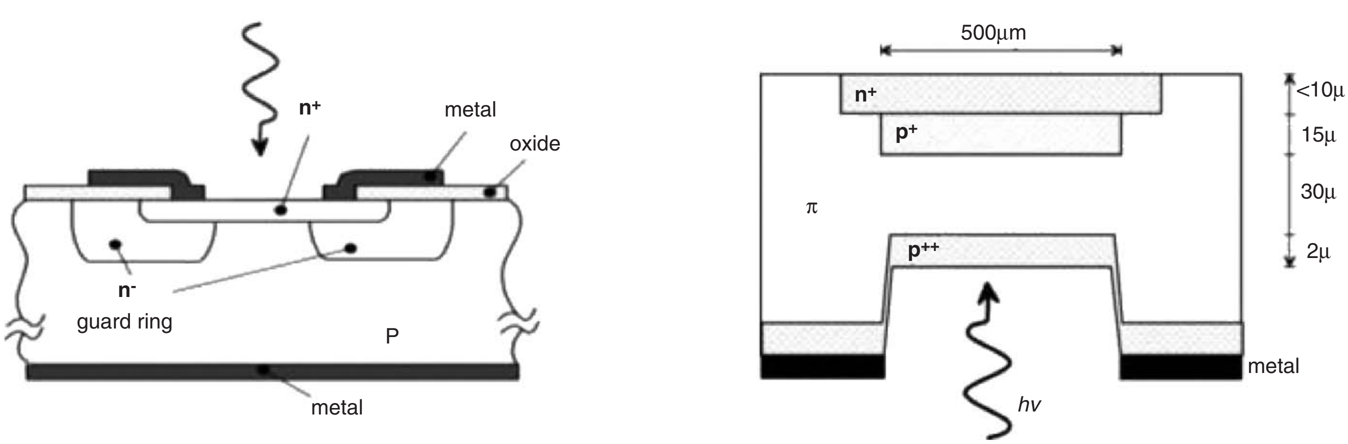

Fig. 2. The first two silicon single photon detectors. Left is the planar type from Haitz and right the reach through type made by McIntyre.

conducting band. Later this device was modified to be less sensitive to infrared light and is now called Visible Light Photon Counter (VLPC). The small band gap forces an operation at very low temperatures of few degree Kelvin (Fig. 4).

Around 1990 the MRS APDs (Metal-Resistor-Semiconductor) were invented in Russia.

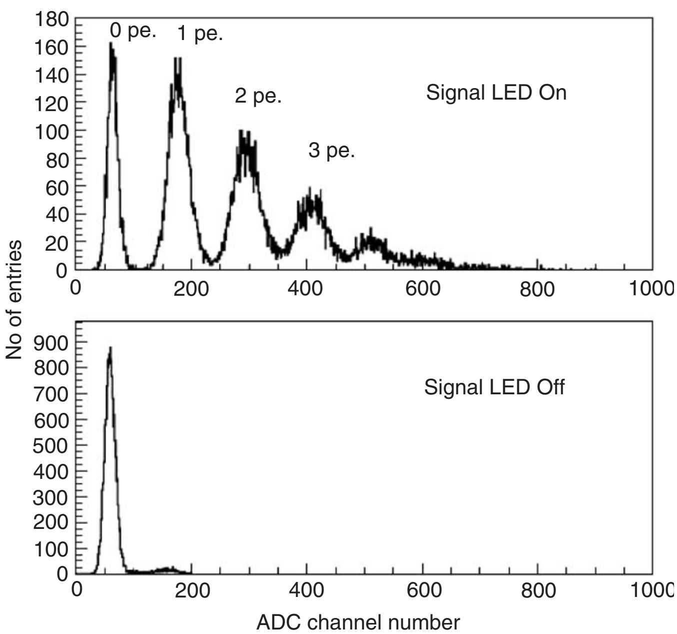

Fig. 4. Pulse height spectrum of light pulses with very low intensity recorded with a VLPC. Taken from Bross et al. [8].

A very thin metal layer (Ti, ~0.01 μm) and a layer of SiC or Si_XO_Y with a resistivity of 30–80 MΩ cm limits the Geiger breakdown by a local reduction of the electric field.

The technology is difficult because all parameters need to be controlled very precisely.

Two examples [9,10] out of a great number of different designs are shown in Fig 5.

The next step was logical: subdivide the MRS structure into many cells and connect them all in parallel via an individual limiting resistor (Fig. 6). The Geiger-mode avalanche photodiode (G-APD) is born. Key personalities in this development are Golovin [11] and Sadygov [12]. The technology is simple. The G-APD is produced in a standard MOS (Metal-Oxide-Silicon) process and promises to be cheap. An educated guess is a price of some 100$ per cm^2.

3. Properties of G-APDs

3.1. High gain

G-APDs produce a standard signal when any of the cells goes to breakdown. The amplitude A_i is proportional to the capacitance of the cell times the overvoltage.

$$A_i \sim C(V - V_b),$$

V is the operating bias voltage and V_b is the breakdown voltage.

When many cells are fired at the same time, the output is the sum of the standard pulses (Fig. 7)

$$A = \sum A_i.$$

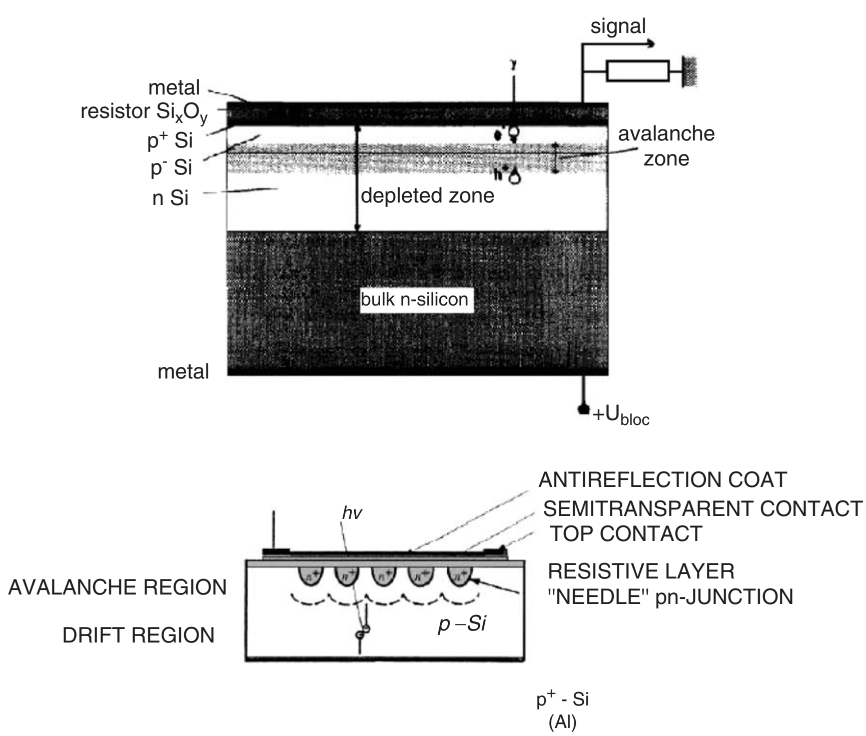

Fig. 5. Examples of MRS APDs structures. The picture on top is taken from Antich et al. [9] and the picture at the bottom is from Saveliev and Golovin [10].

The gain is in the range of 10^5–10^7. Single photons produce a signal of several millivolts on a 50 Ω load. No or at most a simple amplifier is needed and pickup noise is no more a concern and no shielding is needed. There is no nuclear counter effect—even a heavily ionizing particle produces a signal which is not bigger than that of a photon. Since there are no avalanche fluctuations as we have in normal APDs the excess noise factor is very small and it could eventually be one. Grooms theorem [13] is no more valid.

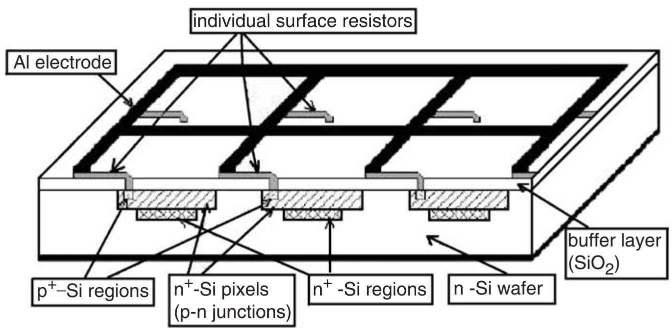

Fig. 6. Basic structure of a Geiger-mode avalanche photodiode. Taken from Sadygov [12].

This theorem states that the resolution of an assembly of a scintillator and a semiconductor photodetector is independent of the area of the detector because the photon statistic improves with the square root of the detector area but becomes worse at the same time with the square root of the detector capacitance which is proportional to the area.

The output signal is proportional to the number of fired cells as long as the number of photons in a pulse (N_{photon}) times the photodetection efficiency PDE is significant smaller than the number of cells N_{total} (Fig. 8)

$$A \approx N_{\text{firedcells}} = N_{\text{total}}\left(1 - e^{\frac{N_{\text{photon}}\ \text{PDE}}{N_{\text{total}}}}\right)$$

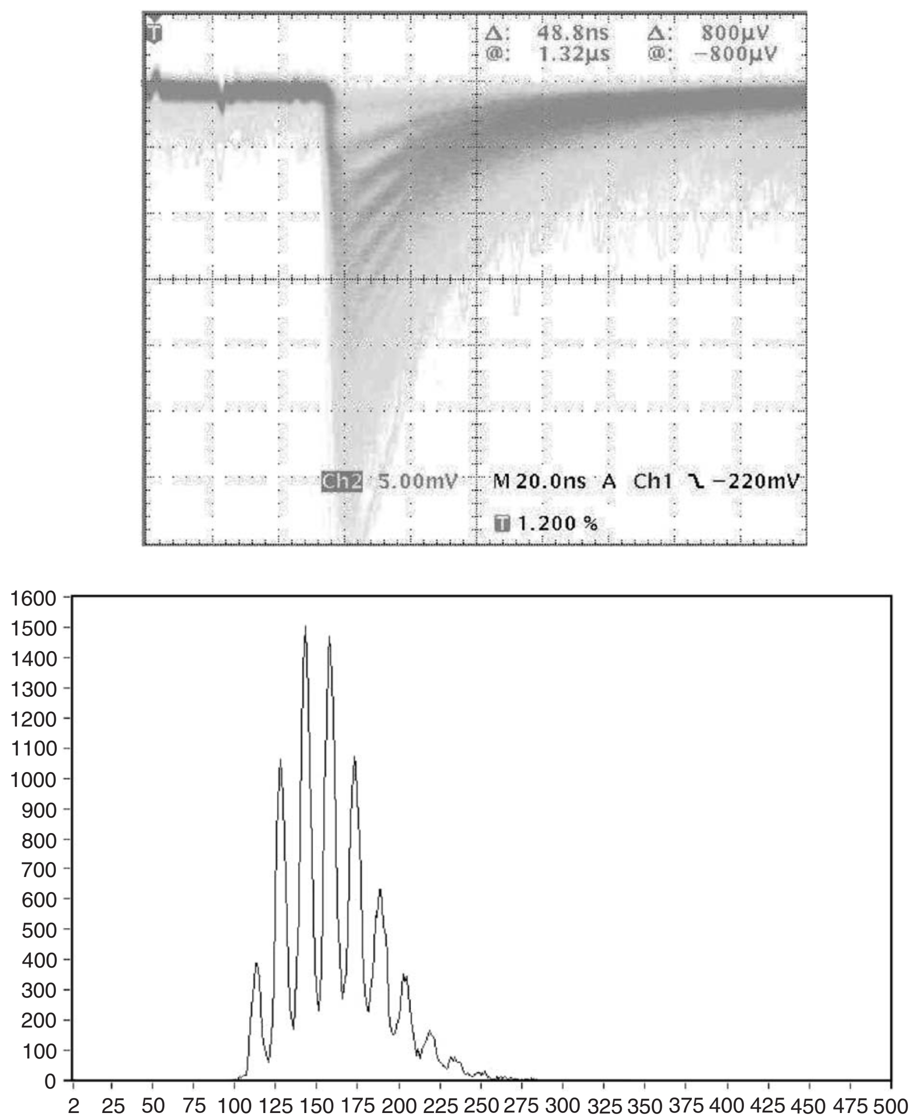

Fig. 7. Oscilloscope picture of the signal from a G-APD (Hamamatsu 1-53-1A-1) recorded without amplifier (a) and the corresponding pulse height spectrum (b).

two or more photons which hit one cell look exactly like only one single photon.

3.2. Dark counts

A breakdown can be triggered by an incoming photon or by any generation of free carriers (Fig. 9). The latter produces dark counts with a rate of 100 kHz to several MHz per mm^2 at 25 °C and with a threshold at half of the one photon amplitude. Thermally generated free carriers can be reduced by cooling. There is a factor 2 reduction of the dark counts every 8 °C. Another possibility is to operate the G-APDs at lower bias resulting in a smaller electric field and thereby lower gain. Field-assisted generation (tunneling) can only be reduced by a smaller electric field.

The dark counts can be reduced in the G-APD production process by minimizing the number of generation-recombination centers, the impurities and crystal defects.

One G-APD was irradiated by the author with a dose of 5 kGy from a ^{60}Co source. No measurable effect was found. An irradiation with protons or neutrons which generate defects by displacement of the silicon atoms in the lattice is not yet done. An increase of the dark count rate can be predicted.

3.3. Optical crosstalk

In an avalanche breakdown there are in average 3 photons emitted per 10^5 carriers with a photon energy higher than 1.14 eV, the band gap of silicon [15]. When these photons travel to a neighboring cell they can trigger a breakdown there (Fig. 10).

The optical crosstalk acts like shower fluctuations in an APD. It is a stochastic process and introduces an excess noise factor like in a normal APD or like in a PMT.

With a dedicated design with grooves between the cells which act as an optical isolation the crosstalk can be reduced. Operation at relatively low gain again is advantageous.

3.4. Afterpulsing

Carrier trapping and delayed release causes afterpulses during a period of several microseconds after a breakdown (Fig. 11).

Afterpulses with short delay contribute little because the cells are not fully recharged but have an effect on the recovery time. Operation at low temperatures elongate the delayed release by a factor of 3 when the temperature is reduced by 25 °C.

3.5. Photon detection efficiency

The photon detection efficiency (PDE) is the product of quantum efficiency of the active area (QE), a geometric

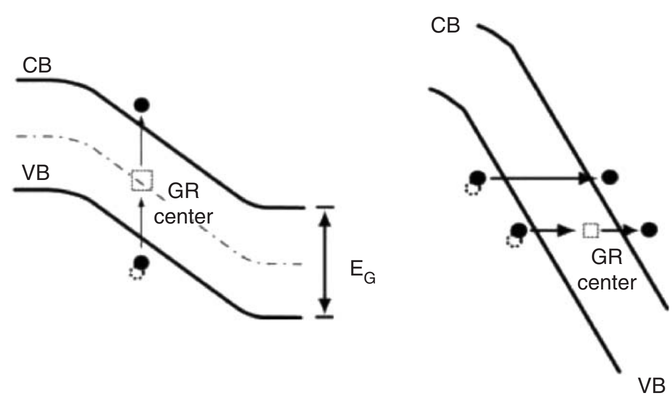

Fig. 9. Thermally (left) and field-assisted generation (right) of free carriers which can trigger a breakdown.

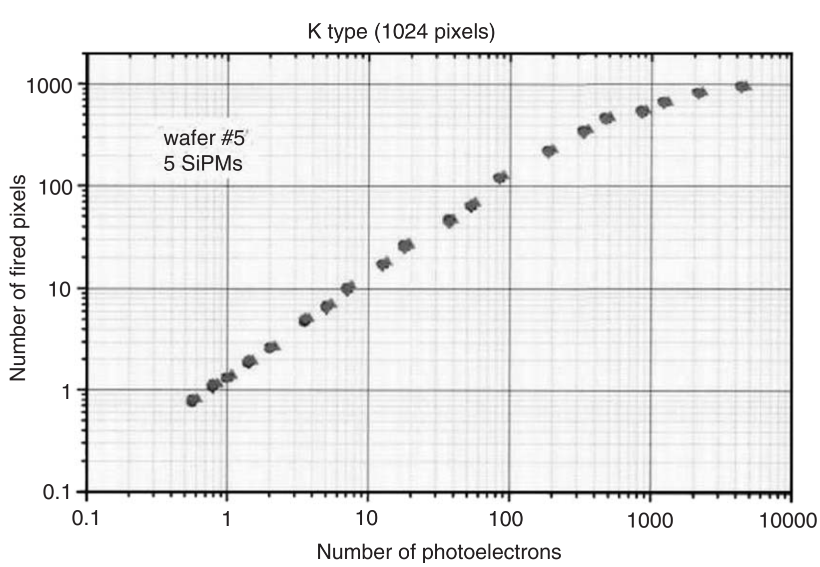

Fig. 8. The number of fired cells which is proportional to the signal amplitude as function of the primary photoelectrons. From Andreev et al. [14].

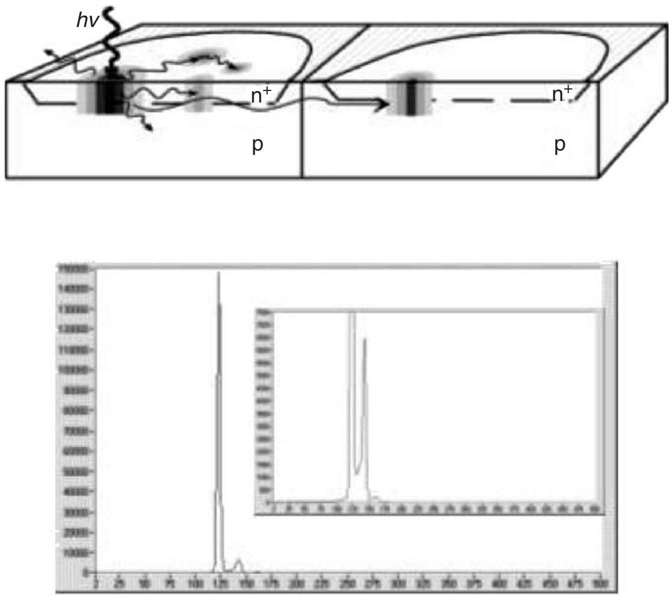

Fig. 10. Optical crosstalk in a G-APD (top) and a pulse height spectrum of events triggered by one single carrier (bottom). Events where a second and even a third cell was fired are visible as satellite peaks. The insert is the same histogram but with the vertical scale is reduced by a factor of 20.

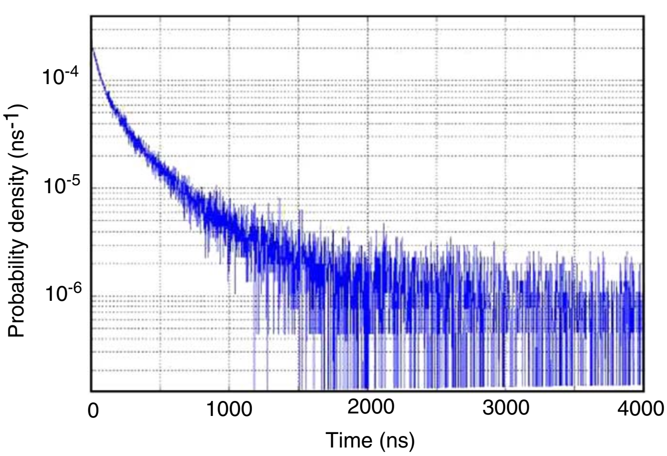

Fig. 11. The probability for a delayed release of carriers as function of time after a breakdown event. From Cova et al. [16].

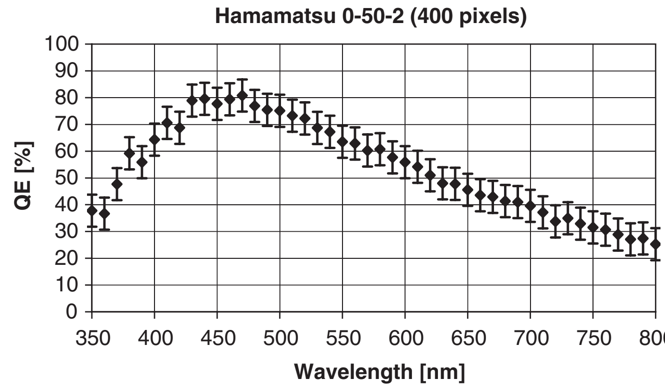

Fig. 12. Quantum efficiency as function of the wavelength.

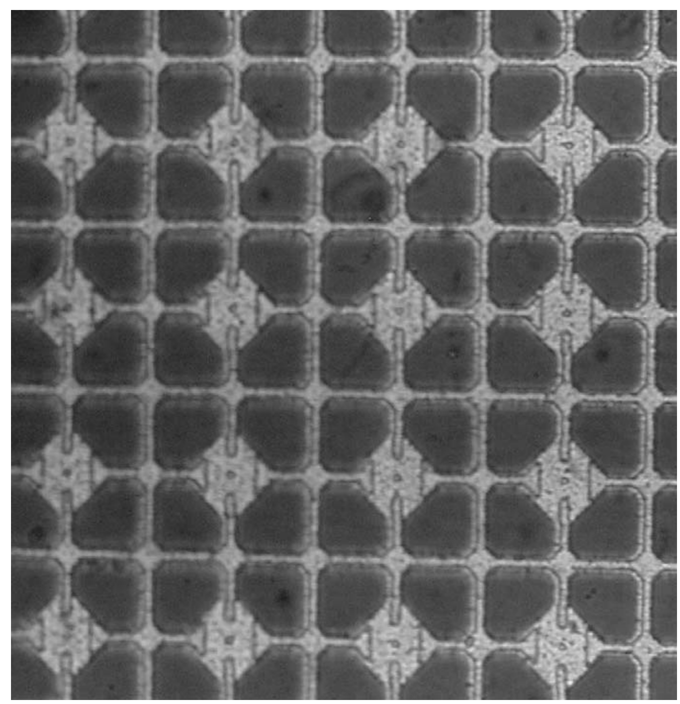

Fig. 13. Microscopic view of a G-APD produced at JINR in Dubna, Russia.

factor (ε, ratio of sensitive to total area) and the probability that an incoming photon triggers a breakdown (P_{trigger})

$$\text{PDE} = \text{QE} \times \varepsilon \times P_{\text{trigger}}.$$

The QE is maximal 80–90% depending on the wavelength. It peaks in a relative narrow range of wavelengths because the sensitive layer of silicon is very thin. In the case shown in Fig. 12 the G-APD structure is p-silicon on a n-substrate and the p-layer is only 0.8 μm thick. Devices with inverse structure, n-silicon on a p-substrate are more sensitive for green and red light and less for blue light because only the photons with longer wavelengths penetrate deep enough into the silicon.

The geometric factor ε needs to be optimized depending on the application. Since some space is needed between the cells for the individual resistors and is needed to reduce the optical crosstalk the best filling can be achieved with a small number of big cells.

For example in a camera for air Cherenkov telescopes the best possible PDE is wanted. Since the number of photons is small big cells are suitable and a geometric factor of 50% and more is possible. On the other hand LSO crystals for PET produce many photons and 1000 or more can be collected at the end face of the crystals. In order to avoid a saturation effect the number of cells needs to be big and the cells small. The geometric factor will be in the range of 20–30% (Fig. 13).

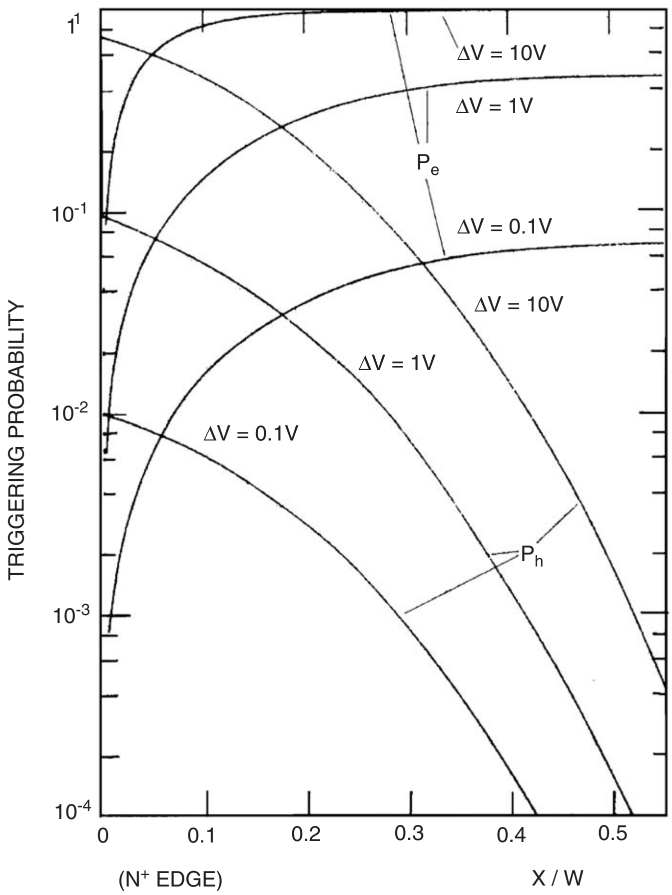

Fig. 14. Triggering probability for different positions of carrier generation (see text).

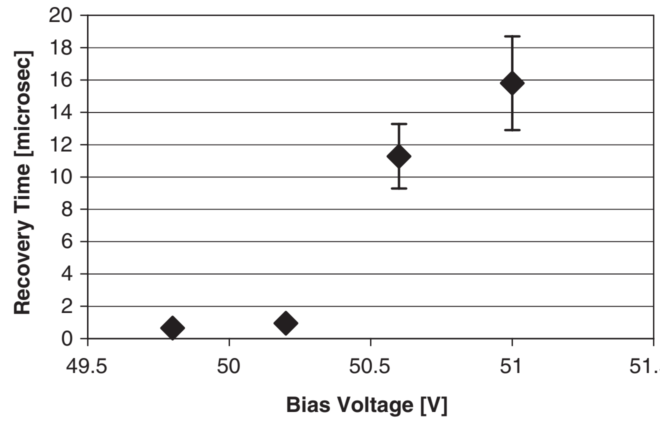

Fig. 15. Recovery time for 4 different settings of the bias voltage.

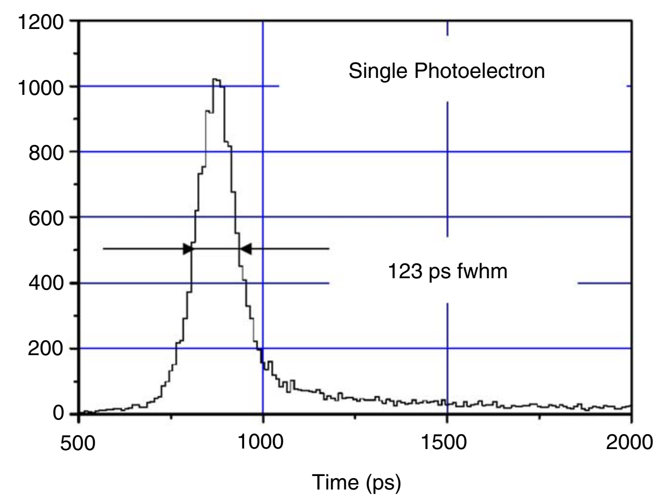

Fig. 16. Time resolution for single photons. See text.

The triggering probability depends on the position where the primary electron–hole pair is generated and it depends on the overvoltage. A high gain operation is favoured. Electrons have in silicon a better chance to trigger a breakdown than holes. Therefore a conversion in the p-layer has the highest probability. This has been calculated by Oldham et al. [17]. They define an avalanche region with width W and the position X which runs from 0 to *W* starting at the *n*-edge (Fig. 14).

A semiconductor material other than silicon, in which the holes have a higher mobility and higher ionization coefficient, like GaAs could have a very high trigger probability.

3.6. Recovery time

The time needed to recharge a cell after a breakdown has been quenched depends mostly on the cell size due to its capacity and the individual resistor (RC). Afterpulses can prolong the recovery time because the recharging starts anew. This can be reduced only by operation at low gain (Fig. 15). Some G-APDs need hundreds of microseconds after a breakdown until the amplitude of a second signal reaches 95% of the first signal. The smallest values of the recovery time have G-APDs with small cells and small resistors.

Polysilicon resistors are used up to now for the quenching of the breakdown. They change their value with temperature. Therefore, there is a strong dependence of the recovery time on the temperature. It would be advantageous to use instead a metal alloy with high resistivity like FeCr.

3.7. Timing

The active layers of silicon are very thin (2–4 μm), the avalanche breakdown process is fast and the signal amplitude is big. Therefore, very good timing properties even for single photons can be expected. Fluctuations in the avalanche development are mainly due to a lateral spreading by diffusion and by the photons emitted in the avalanche [18,19]. The vertical build-up contributes only little to the timing. Fig. 16 shows a measurement of the time response of a G-APD for single photons [2]. The authors state a 40 ps contribution from the used laser and the electronics each. The result then is a time resolution with a standard deviation of 42 ps Operation at high overvoltage (high gain) improves the time resolution.

The tail to the right visible in Fig. 16 can be explained by carriers created in field free regions which have to travel by diffusion. It can take several tens of nanoseconds until they reach a region with field and trigger a breakdown (Fig. 17). At low gain the lateral spreading of the depleted volume can be incomplete and can enhance the diffusion tail.

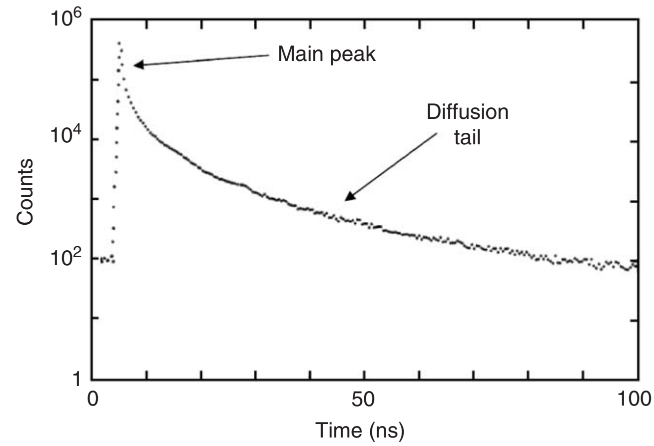

Fig. 17. Logarithmic representation of the timing distribution.

3.8. More features

There are more features which are not mentioned yet:

- G-APDs's work at low bias voltage (~50 V),
- have low power consumption ($<50\,\mu W/mm^2$),
- are insensitive to magnetic fields up to 15 T,
- are compact and rugged,
- have a very small nuclear counter effect (small sensitivity to charged particles traversing the device) and
- tolerate accidental illumination

These attributes make G-APDs superior to PMTs in many applications. The main drawback is the small area available presently. The largest device has an area of $9\,mm^2$.

Unless there is a significant improvement of the design and the production process which results in a reduction of the dark count rate the area cannot exceed few cm^2.

4. Choice of parameters

There are many different designs possible. In the following list a number of design choices are given together with the consequences on some operating parameter:

- Semiconductor material—has influence on the PDE and the range of wavelengths
- p-silicon on a n-substrate—highest detection efficiency for blue light
- n-silicon on a p-substrate—highest detection efficiency for green light
- Thickness of the layers—range of wavelengths, optical crosstalk
- Doping concentrations—operating voltage and its range
- Impurities and crystal defects—dark counts, afterpulses
- Area of the cells—gain, geometric factor, dynamic range, recovery time
- Value of the resistors—recovery time, count rate/cell
- Type of resistors—temperature dependence
- Optical cell isolation (groove)—crosstalk

Many applications need the highest possible photon detection efficiency but don't need high dynamic range (RICH, DIRC, IACT, EUSO, photon correlation studies, fluorescence spectroscopy, single electron LIDAR, neutrino detectors). Here the best is a G-APD with p- on n-silicon structure, large cells (50–100 μm^2), small value of the individual resistors and optical isolation between the cells.

Other applications need large dynamic range (HEP calorimeters, PET, SPECT, scintillator readout, Smart PMT, radiation monitors). The best is p- on n-silicon structure again, small cells (5–30 μm^2) and relative thick p-layer. An optical isolation is not needed.

Some applications like a tile calorimeter are better off with a n- on p-silicon structure.

5. Conclusions

Multi-cell APDs operated in Geiger-mode are now an alternative to PM's. They are the better choice for the detection of light with very low intensity when there is a magnetic field and when space and power consumption are limited. Most of the devices are still small ($1 \times 1\,mm^2$) but areas of $3 \times 3\,mm^2$ are available and a G-APD with $10 \times 10\,mm^2$ is planed. Also planed is a monolithic array of 4 diodes with $1.8 \times 1.8\,mm^2$ each.

The development started some 10 years ago but still there is a broad room for improvements. Many parameters can be adjusted to optimize the devices and to tailor them for special needs.

References

[1] A. Einstein, Ann. Phys. 17 (1905) 132.
[2] P. Buzhan, et al., Nucl. Instr. and Meth. A 504 (2003) 48.
[3] R.J. McIntyre, J. Appl. Phys. 32 (6) (1961) 983.
[4] R.H. Haitz., J. Appl. Phys. 35 (5) (1964) 1370.
[5] K. Yamamoto, Patent No. JP S49-59585, 1972.
[6] F. Zappa, et al., ESSCIRC (2002).
[7] S. Vasile, et al., IEEE Trans. Nucl. Sci. NS-46 (4 Part 1) (1999) 848.
[8] A. Bross, et al., Nucl. Instr. and Meth. A 477 (2002) 172.
[9] P.P. Antich, et al., Nucl. Instr. and Meth. A 389 (1997) 491.
[10] V. Saveliev, V. Golovin, Nucl. Instr. and Meth. A 442 (2000) 223.

[11] V. Golovin, Patent No. RU 2142175, 1998.
[12] Z. Sadygov, Patent No. RU 2102820, 1998.
[13] D.E. Groom, Nucl. Instr. and Meth. 219 (1984) 141.
[14] V. Andreev, et al., Nucl. Instr. and Meth. A 540 (2005) 368.
[15] A. Lacaita, et al., IEEE TED 40 (3) (1993) 577.
[16] S. Cova, et al., NIST Workshop on Single Photon Detectors (2003).
[17] W.G. Oldham, et al., IEEE TED 19 (9) (1972) 1056.
[18] A. Lacaita, et al., Appl. Phys. Lett. 57 (5) (1990) 489.
[19] A. Lacaita, et al., Appl. Phys. Lett. 62 (6) (1993) 606.