CHAPTER 1

OPERATING PRINCIPLES AND CONSTRUCTION

1.1	Photocathode			
1.1.1	Standard types			

- 1.1.2 Other types of photocathode
- 1.1.3 Input window

1.2 Electron-optical input system

- 1.2.1 General purpose tubes
- 1.2.2 Fast-response tubes
- 1.2.3 Cathode current saturation

1.3 Electron multiplier

- 1.3.1 Dynode materials
- 1.3.2 Dynode geometry
- **1.4** Anode collection space

1.5 Special types of photomultipliers

- 1.5.1 Tubes with GaP(Cs) dynodes
- 1.5.2 Crossed-field photomultipliers
- 1.5.3 Microchannel-plate photomultipliers
- 1.5.4 Multi-channel photomultipliers
- 1.5.5 Vacuum photodiodes and phototriodes
- 1.5.6 Smart photomultipliers
- 1.5.7 Experimental photomultiplier types

Appendix Physics of photomultipliers

- A1.1 Photoemission
- A.1.1.1 Photon absorption
- A1.1.2 Electron diffusion
- A1.1.3 Surface barrier
- A1.1.4 Photoelectron energy distribution
- A1.2 Secondary emission

OPERATING PRINCIPLES AND CONSTRUCTION

A photomultiplier converts light into an electrical signal, then amplifies that signal to a useful level by emission of secondary electrons. Figure 1.1 shows the essential elements:

- a *photocathode* which converts light flux into electron flux;
- an *electron-optical input system* which focuses and accelerates the electron flux;
- an *electron multiplier* consisting of a series of secondary-emission electrodes (dynodes); and, finally
- an *anode* which collects the electron flux from the multiplier and supplies the output signal.



Fig.1.1 Elements of a photomultiplier based on the world's first fast PMT, the 56AVP, introduced by Philips (now Photonis) in 1956

The two phenomena fundamental to the operation of a photomultiplier are *photoemission* and *secondary emission* (see Appendix to this chapter).

Photoemission is due to a fraction of the incident photons that impart all their energy to bonded electrons of the photocathode material, giving some of them sufficient energy to escape. If the number of these *photoelectrons* that strike the first dynode is n_k , and

the gain of the dynode is g_1 , the number of resulting secondary electrons is $n_k g_1$. Then, if the second dynode has a gain g_2 , it in turn emits $n_k g_1 g_2$ electrons. The process repeats from dynode to dynode up to the anode where the electrons are finally collected. If N is the number of dynodes, the number of electrons collected is

$$n_a = n_k \prod_{i=1}^N g_i$$

For example, if the gain of each dynode is 4, the current amplification M of a ten-stage multiplier is

$$M = \frac{n_a}{n_k} = \prod_{i=1}^{10} g_i = 4^{10} \approx 10^6$$

The electrons are accelerated and focused by electric fields between the dynodes, the required potential gradients usually being obtained from a voltage divider across the terminals of a high-voltage supply (Fig.1.2).



Fig.1.2 Voltage-divider high-tension supply

1.1 Photocathode

The cathodes normally used in photomultipliers are made of a deposited photoemissive semiconductor. There are two main kinds:

- semi-transparent cathodes, the most widely used, are deposited on the inside of the input window; electrons are emitted from the side opposite to the incident light. The cathode can be large (from ten to a few hundred millimetres in diameter) and the window on which it is deposited can be flat or curved.
- opaque cathodes are deposited on a metal electrode inside the tube. Electrons are emitted from the illuminated side. The area is usually limited to a few square centimetres because of the size of the focusing electrodes.

1.1.1 Standard types

The most used photocathode materials are silver-oxygen-caesium (AgOCs), antimonycaesium (SbCs), and the bi-and trialkali compounds SbKCs, SbRbCs, and SbNa₂KCs. Examples of the sensitivities of different photocathodes as functions of wavelength are shown in Fig.1.3; these are called the *spectral sensitivity characteristics*. They are limited at long wavelengths by the photoemission threshold of the material, and at short wavelengths by the transmission of the window. Table 1.1 gives the characteristics of some standard photocathodes.

The sensitivity curves of bi- and trialkali cathodes can be shifted toward the long wavelengths at the expense of blue sensitivity, by increasing their thickness; the trialkali S2OR photocathode (sometimes called S25) is an example.



Fig.1.3 Spectral sensitivity characteristics for various photocathode types. Curves of uniform quantum efficiencies, ρ, are shown for reference; see ch.2.1.2.

Type of spectral response	Composition	Type of window	Photo- emission threshold (nm)	Wavelength at maximum sensitivity (nm)	Radiant sensitivity at λ_{max} (mA/W)	Quantum efficiency at λ_{max} (%)
S1	AgOCs	1	1100	800	2.3	0.4
S4	SbCs ₃	1,2,3	680	400	50	16
S11	SbCs ₃	1	700	440	80	22
S13	SbCs ₃	2	700	440	80	22
S20	SbNa ₂ KCs	1	850	420	70	20
S20	SbNa ₂ KCs	2	850	420	70	20
S20R (ERMA*)	SbNa ₂ KCs	1	900	550	35	8
bialkali	SbKCs	1	630	400	90	28
bialkali	SbKCs	2	630	400	90	28
bialkali (GEBA**)	SbKCs	1	700	440	100	28
bialkali	SbNaK	1	700	400	50***	16***
solar blind	CsTe	2	340	235	20	10

Table 1.1 Composition and typical characteristics of photocathodes

Photocathodes The S designations (JEDEC No. 50. Oct. 1954, S curves) refer to the total spectral response, including the effect of the input window. They do not identify specific types of cathode or cathode materials, or absolute sensitivities, although they are often so used.

Windows 1. Borosilicate or lime glass or equivalent. 2. Fused silica. 3. Internal (opaque) cathode.

- * ERMA = Extended-red multialkali, sometimes called S25 ** GEBA = Green-extended bialkali
- *** The SbNaK bialkali cathodes are intended for high-temperature operation and the sensitivity and quantum efficiencies are given for 130 °C.

1.1.2 Other types of photocathode

Semiconductors with negative electron affinity (see Appendix A1.1.3), such as GaAs, GaInAs, and GaInAsP, have sensitivities extending into the near-infrared (Fig.1.4). These materials were first limited to opaque cathodes, but technological developments have made it possible to use them in semi-transparent ones as well.



Fig.1.4 Spectral sensitivity characteristics for two negative-electron-affinity photocathode materials; quantum efficiency curves $\rho = 1\%$ and $\rho = 10\%$ shown for reference

1.1.3 Input window

The material of the input window limits the spectral sensitivity in the short wavelength region. Figure 1.5 and Table 1.2 give the characteristics of some of the glasses used. These are mainly borosilicate glass (hard glass) and lime glass (soft glass), the cut-off wavelengths of which are between 250 and 300 nm, and uv-transparent glasses and fused silica, which have cut-off wavelengths below 250 nm. In some applications, the windows used with semi-transparent cathodes are frosted to reduce reflection.



Fig.1.5 Transmission (%) as a function of wavelength λ for various glasses used in photomultiplier input windows (thickness 3 mm)

Type of window	Cut-off wavelength (decrease to 10%)	Refractive index n at λ
	(nm)	(nm)
lime glass	300	1.54 400
borosilicate	270	1.50 400
UV-glass	190	1.49 400
fused silica	160	1.47 400 1.50 250
sapphire (Al ₂ O ₃)	145	1.80 400
MgF ₂	115	1.40 400
LiF	105	1.40 400

Table 1.2 Characteristics of glasses used in photomultiplier windows

Even with LiF or MgF_2 , UV-radiation of less than 180 nm can be investigated only in vacuum because of the absorption of the air. For wavelengths less than 105 nm, there is no transparent material and windowless photomultipliers must be used in an evacuated system. (Alternatively, single-channel electron multipliers or microchannel plates may be used.)

1.2 Electron-optical input system

The purpose of the electron-optical input system is to focus all the photoelectrons onto the useful area of the first dynode. If the photocathode is an opaque one, it forms part of the internal electrode structure and the input system can be simple. However, if it is a semi-transparent one deposited on the input window, and especially if it is large, the system is more complicated and may incorporate one or more extra focusing electrodes.

The input system must satisfy two main requirements:

- It must be able to focus the maximum number of electrons onto the first dynode regardless of their initial velocity and regardless of where on the photocathode they originate. The ratio between the number of electrons reaching the useful area of the first dynode and the total number of electrons emitted by the cathode is called the *collection efficiency*; it varies with wavelength but is usually greater than 80% (see Fig.2.2).
- The transit time of electrons between the cathode and the first dynode must, as far as possible, be independent of their initial velocity and point of origin. This is particularly important in fast-response photomultipliers, and these therefore have more complex input systems.

1.2.1 General purpose tubes

Figure 1.6 shows the input system geometry of two general-purpose photomultipliers: one with focusing dynodes and one with venetian-blind dynodes. In both, the system consists of:

- the cathode
- an aluminium layer evaporated onto the inside of the tube and connected to the cathode
- an accelerating electrode at the potential of the first dynode.

Together, these three electrodes establish the electric field that determines the paths of the electrons between the cathode and the first dynode. The dotted equipotential lines show the shape of the field. The solid lines show the paths of the electrons emitted with zero initial velocity, from different parts of the cathode.



Owing to the spread of initial velocities, there is some spatial spread in the landing on the first dynode of electrons originating from the same point on the cathode. This spread is shown in Fig.1.6. Two factors that favour collection efficiency are good focusing of the photoelectrons and large area of the first dynode (e.g. a venetian-blind or foil dynode or a large box, see page 7-5). By placing the hole in the focusing electrode slightly off centre and splitting the electrode into two parts (triode input optics) it is possible to achieve almost perfect matching to the inherent asymmetry of a venetian-blind first dynode, Fig 7.2(b). For a focusing or box first dynode, however, it also becomes equally important to correctly collect the secondary electrons onto a second dynode.

1.2.2 Fast-response tubes

For fast response photomultipliers, there is an additional requirement, namely to keep transit time *variations* in the cathode/first-dynode space to a minimum. These variations have two components:

- a *chromatic* component, due to the spread of initial velocities of electrons originating from the same point. The initial velocity vector can be resolved into a velocity v_n normal to the cathode, corresponding to an energy E_n , and a velocity v_t tangential to the cathode, corresponding to an energy E_t . Compared with the transit time of an electron emitted with zero initial velocity, the transit time decrease due to a normal component of initial velocity is

$$\Delta t_{n} = \frac{1}{\mathscr{E}e} \sqrt{2m_{e}\Delta E_{n}}$$
(1.1a)

where \mathscr{E} is the electric field strength at the cathode surface, e the electron charge, and m_e the electron mass.

The tangential component of velocity also makes a contribution to the transit time fluctuations. Compared with the transit time of an electron emitted with zero initial velocity, the transit time increase due to a tangential component of initial velocity is

$$\Delta t_{t} \approx \frac{r}{v} \tag{1.1b}$$

where r is the radius of electron scatter and v the impact velocity on the first dynode.

- a *geometric* component, due to the difference in transit times between primary paths of different lengths from different points on the cathode; a contributing factor is the inhomogeneity of the electric field in the space traversed by the electrons. The difference in transit time of two electrons emitted with zero initial velocity and passing through a uniform electric field \mathcal{E} along paths of lengths L and L + Δ L is:

$$\Delta t = \Delta L \sqrt{\frac{m_e}{2e\mathscr{E}L}}$$
(1.2)

As shown by Eqs 1.1 and 1.2, transit time variations can be decreased by increasing the electric field strength at the cathode surface and by minimizing differences between the electron path lengths in the cathode/first-dynode space. In fast-response photomultipliers this is done by using a concave cathode and one or more additional focusing electrodes (Fig.1.7).



Fig.1.7 Example of a large fast photomultiplier input optics

1.2.3 Cathode current saturation

Figure 1.8 shows an example of photocurrent variation as a function of the voltage applied between cathode and first dynode, for several values of incident flux. Because the initial velocity of the photoelectrons is usually not zero, some of them can be collected by the first dynode (or one of the auxiliary electrodes connected to it) even if the potential difference between the cathode and the first dynode is zero. This can be stopped by making the first dynode negative with respect to the cathode. The voltage at which the photocurrent becomes zero is called the *cut-off voltage*. It depends on wavelength, photocathode material, and input system design, and is never more than a few volts.

As the first dynode is made more positive with respect to the photocathode, the photocurrent tends towards a *saturation value*. The voltage at which this is reached depends on photocathode composition and structure, the variation in the initial photoelectron velocities, and the layout of the collection electrodes. Depending on the type of photomultiplier, the *saturation voltage* is between a few volts and a few tens of volts. The *saturation current* is **proportional** to the incident flux, as shown in Fig.1.8.



Fig.1.8 Photocurrent variation as a function of cathode-to-firstdynode voltage, with relative incident flux as parameter

1.3 Electron multiplier

The dynodes of which the electron multiplier consists are so arranged that the electric fields between them cause the electrons emitted by each dynode to strike the next with an energy of a few hundred electron-volts. As a result of secondary emission, the number of electrons increases from dynode to dynode, giving the required multiplication. However, not all the electrons emitted by one dynode reach the next one. Each interdynode space of the multiplier is characterised by a *collection efficiency* n. The secondary emission coefficient δ_i of dynode i and the collection efficiency n_{i-1} of the space that precedes it are *both* increasing functions of the voltage V_i between the dynodes d_i and d_{i-1} (the rate of increase being faster for δ_i than for n_{i-1}). The product of n_{i-1} and δ_i varies as a power of the applied voltage V_i :

$$\mathbf{n}_{i-1} \, \boldsymbol{\delta}_i = \mathbf{k}_i \, \mathbf{V}_i^{\boldsymbol{\alpha}} \tag{1.3}$$

where k_i is a constant and the exponent α is usually between 0.65 and 0.75.

For expressing the gain, it is more convenient to associate the secondary emission factor δ_i of dynode i with the collection efficiency n_i of the interdynode space that immediately follows it; the gain g_i of dynode i is then

$$g_i = \delta_i n_i \tag{1.4}$$

and the total gain M of the multiplier is

$$\mathbf{M} = \prod_{i=1}^{N} \mathbf{g}_i \tag{1.5}$$

where N is the total number of stages.

The structure of the multiplier is mostly iterative, all dynodes (except, sometimes, the first one or two and the last) being identical. Because of the high curvature of the electron trajectories entering it, the first interdynode space forms a critical coupling between the electron-optical input system and the iterative part of the multiplier. For that reason, a higher voltage is often applied between the first two or three dynodes than between the others.

The gain of each stage fluctuates about a statistical mean. The same holds for the overall gain M of the multiplier, which is therefore specified in terms of a mean value \overline{M} and a probability distribution. Gain fluctuations can be reduced by

- increasing the secondary emission coefficient
- improving the uniformity of the secondary emission coefficient
- equalizing the collection efficiencies of the stages.

The electron transit time between first dynode and anode also fluctuates, for the same reasons as in the input system:

- variations in initial electron velocity
- variations in electron trajectories.

Important factors governing the gain and time response of the multipliers are the materials and geometry of the dynodes.

1.3.1 Dynode materials

Materials with satisfactory secondary emission coefficients are either insulators or semiconductors. For the structure of the dynodes, materials such as AgMg, CuBe and NiAl are used; these do not in themselves have useful secondary emission coefficients, but the oxides, such as MgO, BeO and Al_2O_3 , that can be formed on their surfaces do. The alkali metals used in forming the photocathode also serve to reduce the electron affinity of the dynode surfaces. Some photoemissive materials (particularly SbCs₃) are also good secondary emitters and can be deposited and processed in the same way for both photocathodes and dynodes. Figure 1.9 gives examples of the secondary emission of some typical surfaces.

Semiconductor materials with negative electron affinity have been developed which make it possible to obtain higher secondary emission coefficients. One such material is GaP(Cs), the secondary emission characteristic of which is plotted in Fig.1.10. With primary-electron energies between 400 eV and 800 eV, this material gives secondary emission coefficients between 20 and 50.



Fig.1.9 Secondary emission coefficients of three commonly used dynode materials, as functions of incident primary-electron energy



Fig.1.10 Secondary emission coefficient of GaP(Cs) as a function of incident primary-electron energy

1.3.2 Dynode geometry

Venetian-blind dynodes (Fig.1.11(a)). These are parallel strips slanted with respect to the axis of the tube and stacked parallel to the photocathode. The large surface of the first dynode makes it possible to use a fairly simple electron-optical input system. The arrangement has high collection efficiency and good gain stability but mediocre time characteristics, response being slow owing to the low electric field at the surface of the dynodes.

Box dynodes (Fig.1.11(b)) also known as *box-and-grid dynodes*. These too have a large collection area at the first dynode, and hence good collection efficiency, but the low electric field at the internal surface of the boxes does not contribute to good time characteristics and the collection of the secondary electrons onto the second dynode becomes important.

Linear focusing dynodes (Fig.1.11(c). The dynodes are designed to ensure progressive focusing of the electron paths through the multiplier. This reduces the variation in transit times between stages and makes for very fast response. Because of the need to achieve the best coupling between the input electron-optics and the multiplier, the first dynodes are shaped and arranged differently from the rest.



Fig.1.11 Dynode configurations: (a) venetian blind, (b) box, (c) linear focusing, (d) circular cage, (e) mesh and (f) foil

Circular cage dynodes (Fig.1.11(d)). In contrast to linear focusing dynodes, this arrangement leads to highly compact multipliers.

Mesh dynodes (Fig.1.11(e)). These consist of parallel mesh planes of thin wire. Their collection efficiency from dynode to dynode is low, but well designed mesh dynodes can operate in magnetic fields of about a tesla. Though gain is reduced in fields of this magnitude, it is still around 10^4 for say a 16-stage tube.

They are used in position-sensitive multi-channel tubes but because of the high crosstalk introduced, the spatial position of an event must be found by centroiding over many channels, and this makes them too slow for many applications.

Foil dynodes (Fig.1.11(f)). These consist of perforated metal foils with well defined apertures. Their collection area is high and their collection efficiency from dynode to dynode is about the same as that of venetian blind dynodes. They can work in magnetic fields up to some tens of mT and their low crosstalk allows them to be used in position-sensitive multi-channel tubes in which all signals are read in parallel.

1.4 Anode collection space

The shape and arrangement of the electrodes at the output end of the multiplier again differ from those used in the iterative part. The geometry must be suitable for:

- collecting all the secondary electrons emitted by the last dynode
- minimizing space charge effects to ensure linear response in pulse-mode operation (§4.5)
- matching the anode impedance to the characteristic impedance of the output connection.

Figure 1.12 shows the collection spaces of two general purpose photomultipliers. The collector, or *anode*, is a grid positioned close to the last dynode so that secondary electrons emitted by the next-to-last dynode pass through it but those emitted by the last dynode are collected. Such a configuration makes it possible to obtain a high electric field between the last dynode and anode and so reduce the space charge effect in the last stage (§4.5.2). Furthermore, the last dynode forms a partial electrostatic screen around the collector.



MRB122

Fig.1.12 Collection-space configurations: (a) with focusing dynodes, (b) with venetian-blind dynodes

Figure 1.13 is an example of anode collection space current-voltage characteristics. The current becomes practically constant when the last-dynode-to-anode voltage is more than about ten volts. The photomultiplier then behaves as an almost perfect current generator.



Fig.1.13 Anode current as a function of last-dynode-to-anode voltage, with relative incident flux ϕ as parameter

1.5 Special types of photomultiplier

1.5.1 Tubes with GaP(Cs) dynodes

The high secondary emission coefficient (Fig.1.10) of GaP(Cs) makes it possible to reduce random fluctuations of gain and greatly improve energy resolution, especially for light pulses that liberate only a few electrons from the cathode. With a first dynode of GaP(Cs), it is possible to discriminate between light pulses that liberate 1, 2, 3, 4 or 5 photoelectrons. The time characteristics are also improved because of the more favourable multiplication statistics and the reduction of the total number of dynodes required.

1.5.2 Crossed-field photomultipliers

In a crossed-field tube, electrons are focused along their whole travel from cathode to anode by crossed magnetic and electric fields. Provided the fields are very uniform, tight electron trajectories and low electron transit times with little fluctuation can be obtained at each stage of the tube. Such tubes have pulse rise times of less than 0.5 ns.

1.5.3 Microchannel-plate photomultipliers

Study of the secondary emission properties of certain glasses led to the development of channel electron multipliers and, from there, to microchannel plates and microchannel-plate photomultipliers. Detailed treatment of these is beyond the scope of this book.

A microchannel-plate photomultiplier (Fig.1.14) consists of a microchannel-plate electron multiplier with a proximity focused cathode at one side and an anode at the other. The microchannel plate is a compact aggregation of parallel glass tubes of microscopic diameter (6 – 25 μ m). Figure 1.15 shows the electron multiplication in a single microchannel; one electron at the input can give rise to as many as 10⁸ at the output, depending upon the voltage across the tube and its length/diameter ratio. Incorporation in a microchannel plate normally limits the gain to 10⁴; higher gain can be obtained by arranging two microchannel plates in cascade.



Fig.1.14 Section through a microchannel-plate photomultiplier



Fig.1.15 Electron multiplication in a microchannel

When the electron content of a pulse in a channel reaches a certain level the channel saturates. It is possible to turn this to advantage by operating under saturation conditions to reduce gain fluctuations appreciably. This is especially worthwhile when detecting very low current pulses (photon counting, for example). Figure 1.16 shows a pulse-height distribution for a very small number of photoelectrons.



Fig.1.16 Distribution of anode charge per pulse of a microchannel-plate photomultiplier working in the saturation region: vertical axis, relative pulse frequency; horizontal axis, number of photoelectrons per pulse

Advantages of microchannel-plate photomultipliers are:

- *Fast response*. Because of the short distance the electrons travel (1 to 2 mm) and the high electric field (1 20 kV/cm), the transit time is extremely short. Transit-time fluctuations are also small and transit-time differences from different points on the photocathode are negligible. The response pulse duration is therefore short (FWHM ≈ 0.2 ns) and the time resolution δ_t (< 100 ps) is better than that of photomultipliers with discrete dynodes.
- *Insensitivity to magnetic fields*. Because of the short travel, the high electric field, and the constraint imposed by the microchannels, the electron paths are scarcely affected by magnetic fields up to 0.1 T.
- Spatial resolution. The honeycomb structure of a microchannel plate makes it possible to convert an optical image on the cathode into a corresponding electron image on the anode by simple proximity focusing. Spatial information can be recovered by using a mosaic anode (Fig.1.17); numerous methods for improving the spatial resolution electronically are possible.



MRB126

Fig.1.17 Microchannel-plate photomultiplier with mosaic anode

However, microchannel-plate photomultipliers also have limitations that exclude them from most applications. Amongst these are:

- Limitations on count rate and linearity in the pulse mode. These are due to the inherent limitation on the average current at the output to a few μ A/cm² and the long recovery time per channel.
- Limitations of pulse charge capacity per channel $(5 \times 10^{-14} \text{ to } 5 \times 10^{-13} \text{ C})$. First, this limits the gain to about 10^6 ; and second, it limits the current that can be delivered per pulse, thereby changing the shape of the pulse and lengthening it. The second effect is due to the limitation on the maximum anode charge that can be obtained in the linear region $(5 \times 10^{-10} \text{ to } 5 \times 10^{-9} \text{ C/cm}^2)$.
- Limitations of useful life. The large total surface of a microchannel plate makes it difficult to outgas the channels completely, even at high temperature. Internal electron bombardment therefore generates ions that bombard the cathode and quickly reduce its sensitivity. There are ways of countering this, but they are not without drawbacks. For example, a thin film of aluminium deposited over the input end of the plate will stop the ions from reaching the cathode, but at the expense of collection efficiency. Furthermore, as the cumulative charge handled by the plate increases, the gain decreases (but can be restored to some extent by increasing the voltage). Together, these effects limit the useful life to a cumulative charge of about 0.2 to 0.5 C/cm²; that is, about a thousandth of what a conventional photomultiplier can deliver.

Because of the potential possibilities of microchannel-plate photomultipliers there has been wide interest in them in high-energy physics (time-of-flight measurement, beamprofile plotting), nuclear physics (measuring positron life), photon-counting photometry, the bio-sciences and laser physics. Their disadvantages, however, plus the expense of microchannel-plate technology, have until now impeded their industrial development.

1.5.4 Multi-channel photomultipliers

In order to overcome the life disadvantage of microchannel-plate photomultipliers, Photonis invented the foil multiplier, see Fig.1.11(f), in which a package of 10 foil dynodes of CuBe replaces the microchannel plate to make multi-channel photomultipliers with many fully parallel channels. This concept has resulted in a whole family of such tubes with 64 and 96 channels (pixels) with a crosstalk between channels below 5%. The concept has been partly copied by Hamamatsu in their metal-can multi-channel photomultiplier family with 4, 16 or 64 parallel channels and a better packing ratio between useful and dead area.

Both families still suffer from relatively low quantum efficiency and mediocre gain uniformity between pixels and mediocre linearity, but have nevertheless proved to be very useful in reading out scintillating fibres in many physics experiments and animal PET scanners as their lifetime is now comparable to normal photomultipliers.

1.5.5 Vacuum photodiodes and phototriodes for strong magnetic fields

In response to the demands set by high-energy physics experiments, vacuum photodiodes have been developed that can work in magnetic fields up to 0.3 T at angles up to 70° to axial. In conjunction with special low-noise preamplifiers these give excellent results at the light levels available from NaI(Tl) scintillators.

For the much lower light levels available from lead glass or $PbWO_4$ scintillators, vacuum phototriodes have been developed with gains of 4–10 at a few tesla axial field, which is sufficient to lift the signal above the preamplifier noise level. These can work in fields up to 1 T, also at angles up to 70° to axial. Variants with segmented dynodes have been developed in which the signal is taken off separately between each dynode segment and the common anode; with these, the strong axial magnetic field helps considerably to reduce crosstalk between dynode channels.

1.5.6 Smart photomultipliers

These are a combination of a large electro-optical preamplifier tube and a small photomultiplier tube (Fig.1.18). Photo-electrons are accelerated in the preamplifier by a voltage of 25 to 40 kV towards a fast scintillator optically coupled to a conventional linear focused small photomultiplier tube. A first-stage gain of 30 to 50 offers the 'smartness', i.e. the facility to discriminate between one, two three and more photoelectrons. These tubes have high immunity to ambient magnetic fields (the earth's, for example), a transit time difference of a few nanoseconds and they are able to detect a low level multi-electron signal in a high single photoelectron background environment, for example, in deep underwater muon and neutrino detectors (DUMAND) see §8.4.



Fig.1.18 Smart photomultiplier combination in a glass pressure sphere for deep underwater muon and neutrino detectors (DUMAND)

1.5.7 Experimental photomultiplier types

There are also some types of photomultiplier that have so far made little or no progress beyond the laboratory. Amongst these are photomultipliers with silicon dynodes operating by transmission. Current efforts are concentrated on photomultipliers in which the photoelectrons impinge on a silicon diode, silicon avalanche diode, silicon diode array or a charge-coupled device (CCD). The major problem here is to avoid chemical attack on the Si diode structure by aggressive alkalis used for the photocathode processing. This technology has now been realized by DEP for the CERN LHC high-energy physics CMS experiment on an industrial scale in the *hybride photon detector* (HPD) with an electron back-bombarded silicon-photodiode array as target. The gain is limited to a few thousand at an electron-accelerating voltage of 7 - 10 kV. However, these HPDs can work in an axial magnetic field of up to a few tesla.

In the *channel photomultipler* (CPM) a photocathode is evaporated onto a small entrance window (≈ 10 mm diameter) that is indium-sealed onto a spiralised singlechannel electron multiplier (SCEM). Its cathode sensitivity is lower than normal PMTs but its noise is also lower, even at a very high gain of up to 10^8 . As its pulse-height linearity is limited at high gain it is suitable mainly as a broad-dynamic-range counter for analytic applications. Compared with traditional photomultiplier technologies, however, costs remain the limiting factor.

APPENDIX 1. PHYSICS OF PHOTOMULTIPLIERS

A1.1 Photoemission

For descriptive purposes, photoemission can be divided into three phases (Fig.A1.1):

- absorbed photons impart energy to electrons in the material
- energized electrons diffuse through the material, losing some of their energy
- electrons reaching the surface with sufficient excess energy escape from it.

Not all photons incident on a photoemissive material cause electron emission. The ratio of the number of emitted electrons to the number of incident photons is called the *quantum efficiency*. Because of the conditions affecting photon absorption and the energy losses that occur at each stage of the photoemission process, the quantum efficiency is always less than unity. Factors affecting it are the wavelength of the incident light and the composition and thickness of the photoemissive material.



Fig.A1.1 Photoemission

A.1.1.1 Photon absorption

When light strikes a layer of photoemissive material, only a fraction of it imparts energy to the electrons in the material. The rest is either reflected from the front or back of the layer, or passes through it and, therefore, does not contribute to the photoemission process (Fig.A1.2). Metallic layers are highly reflective in the visible part of the spectrum. Semiconductor layers are less so, usually having a reflection coefficient of less than 0.5, depending on wavelength and the thickness of the layer.

The absorption in a layer can seldom be described simply. Because of the thickness of the photocathode layer (a few hundredths to a few tenths of a micron), multiple internal reflections occur and interference effects increase the total absorption, especially close to the emissive surface. Figure A1.3 shows an example of variation in the photons flux absorbed, $d\Phi_p/dx$, at a distance x from the surface of a semiconductor layer.



Fig.A1.2 Light transmission and reflection in a semi-transparent layer



Fig.A1.3 Relative photon absorption $d\Phi/dx$ in a photoemissive layer

Flux absorption as a function of the distance from the surface can be represented by the general equation

$$\frac{d\Phi_{p}(v,x)}{dx} = -\alpha(v) \Gamma(x) \Phi_{p}(v,x)$$
(A1.1)

where $\alpha(v)$ is the *absorption coefficient* of the material, and the complex function $\Gamma(x)$ takes account of the effect of reflections from both surfaces of the layer. If the layer is thick enough (theoretically, semi-infinite), the term $\Gamma(x)$ tends toward unity and the absorption follows an exponential law:

$$\Phi_{p}(v,x) = \Phi_{p}(v,0) \ [1-r(v)] \exp(-\alpha(v)x)$$
(A1.2)

where $\Phi_{p}(v,0)$ is the incident flux and r(v) the *reflection coefficient* at the input surface.

The absorption coefficient $\alpha(v)$ depends on the energy hv of the photons and on the energy band structure of the material forming the layer. Figure A1.4 illustrates the energy band structures typical of metals, semiconductors, and insulators. In metals, the conduction band, at least at low temperature, is filled up to the Fermi level, while all the higher energy levels are almost empty. Hence, it is possible for electrons in the conduction band to be excited by photons and to be emitted into the vacuum if the photon energy is greater than the difference w_{th} separating the Fermi level E_F and vacuum level E₀. In a semiconductor or insulator, the Fermi level is in the forbidden band and the occupancy of the conduction band, even at room temperature, is usually much too low to support an appreciable photoelectric effect. Electrons can therefore be emitted only from the top of the valence band.



Fig.A1.4 Energy bands in (a) a metal, (b) a semiconductor, (c) an insulator

Figure A1.5 shows the variation in the absorption coefficient $\alpha(v)$ as a function of the energy hv of the incident photons, for a few types of semiconductor. For most of them the absorption coefficient is greater than 10⁴/cm for hv>E_g (where E_g is the band-gap energy of the semiconductor). Owing to interference effects, thicknesses of a few tens of nanometres (as in multi-alkali cathodes) are enough to absorb the greater part of the incident radiation. For photon energies less than the band-gap energy the absorption coefficient decreases rapidly.



Fig.A1.5 Absorption coefficients α as functions of photon energy hv for various photoemissive layers

In a slice of thickness dx (Fig.A1.3), the number of electrons, n, excited by transfer of photon energy is proportional to the flux absorption (e.g. A.1.1)

$$\frac{\mathrm{dn}}{\mathrm{dx}} = -\alpha_{\mathrm{e}} \frac{\mathrm{d}\Phi_{\mathrm{p}}(\mathbf{v},\mathbf{x})}{\mathrm{dx}} \tag{A1.3}$$

where α_e represents the fraction of absorbed photons which excite electrons. For visible light, α_e is generally equal to unity.

A1.1.2 Electron diffusion

The probability that an excited electron will be emitted depends on the energy-loss process governing its passage to the surface of the photocathode. This process differs according to whether the photocathode is a metal or a semiconductor.

In metals the conduction band is partially occupied by free electrons that are in thermal equilibrium with the crystal lattice. An electron excited in the conduction band will lose its excess energy mainly through multiple collisions with free electrons in the material and will regain its state of thermal equilibrium after travelling a fairly short distance. As electrons require an excess of energy to escape from the surface, only those excited very near to the surface can contribute to photoemission. The *escape depth* is of the order of a few atomic radii.

In semiconductors, the conduction band is almost empty and the probability that excited electrons will collide with other free electrons while diffusing through the conduction band is very low. Energy loss is therefore mainly due to interaction with the crystal lattice (creation of phonons) and is very low. Thus excited electrons (*hot* electrons) can travel relatively long distances through the material before reaching thermal equilibrium. After they do reach thermal equilibrium, they diffuse down to the bottom of the conduction band over ever greater distances (up to several microns) before recombining with holes and returning to the valence band. As in the case of metals, the existence of an energy barrier at the surface of the semiconductor makes the emission of electrons in thermal equilibrium impossible. Only electrons that have retained an excess of energy can escape into the vacuum. As the average energy loss per electron-phonon collision is about 0.05 eV, and as the mean free path between two collisions is between 2.5 nm and 5.0 nm, the escape depth can be some tens of nanometres.

It is possible to modify the surface barrier of certain types of semiconductors to achieve *negative electron affinity* so that even electrons in thermal equilibrium which retain an energy of only a few kT above the bottom of the conduction band can be emitted into the vacuum. The *escape depth* then becomes equal to the diffusion length of the electrons in thermal equilibrium; that is, it increases to several microns.

A1.1.3 Surface barrier

To understand the mechanism by which electrons escape into the vacuum, first consider the case of *metals* (Fig.A1.4(a)). In the conduction band all the energy levels higher than the Fermi level (E_F) are almost empty. The potential energy E_0 of an electron *in vacuum* exceeds the Fermi level by a quantity W_{th} , called the *thermionic work function*:

$$W_{th} = E_0 - E_F \tag{A1.4}$$

To be emitted into the vacuum, an electron that occupies the highest energy level in the metal must gain sufficient kinetic energy to overcome this *potential barrier*. The additional energy can be provided either by an increase in temperature, or by a photon with sufficient energy. In this case, the thermionic work function W_{th} and the *photoemission threshold* W_{ph} (Fig.A1.4(a)) are the same. Therefore photoemission from a metal can occur only if the incident light has an energy $hv \ge W_{th}$. For most metals W_{th} is greater than 3 eV. Only in the alkali metals (Li, Na, K, Rb, Cs) is it low enough for photoemission to occur in response to visible light (Fig.A1.6).



Fig.A1.6 Work functions of pure metals, in order of atomic number

Now consider a *semiconductor* (Fig.A1.4(b)). The Fermi level is in the forbidden band. The valence band (with upper limit E_v) is almost full, whereas the conduction band (with lower limit E_c) is almost empty. Therefore, high photoemission efficiency can be

obtained only by imparting to electrons in the valence band sufficient energy to enable them to overcome the band-gap energy E_g and the electron affinity E_A of the conduction band:

$$E_g = E_C - E_V$$
 and $E_A = E_0 - E_C$

Thus, the photoemission threshold of a semiconductor is

$$W_{ph} = E_g + E_A \tag{A1.5}$$

which is always greater than the thermionic work function $W_{th} = E_0 - E_F$.



Fig.A1.7 Energy-band diagrams showing (a) positive and (b) negative electron affinity

Semiconductors generally have work functions of 5 - 6 eV, and so can be sensitive only to wavelengths less than about 300 nm. However, the photoemission threshold can be shifted towards longer wavelengths by heavy p-doping in combination with appropriate surface treatment. Because of the ionic nature of the bonds involved, the adsorption of alkali metals (such as caesium) at the surface of a semiconductor results in the appearance of a bipolar layer, the main effect of which is to reduce the work function. Furthermore, as a fraction of the electron charge of the adsorbed atoms is given up to the semiconductor, a local change in the occupancy of the energy levels at the surface results in a *curvature* of the energy bands towards the lower levels. Figure A1.7(a) shows the energy band diagram of such a semiconductor. Note that the electrons from the valence band, which are excited at a greater distance than α from the surface of the semiconductor, only have to overcome an *apparent electron affinity* E_{Aapp} < E_A to be emitted into the vacuum. With caesium, the most efficient metal, the work function can be reduced to about 1.4 eV. And with co-adsorption of oxygen this can be brought down to less than 1 eV.

For certain widths of the forbidden band and doping of the semiconductor, the minimum level of the conduction band can be raised above the vacuum level, as shown in Fig.A1.7(b); this is called *apparent negative electron affinity*. It is then possible for excited electrons that have reached thermal equilibrium by interaction with the crystal lattice to be emitted into the vacuum and the photoemission threshold becomes equal to the band-gap:

$$W_{ph} = E_g \tag{A1.6}$$

By appropriate choice of semiconductor and surface treatment it is possible to produce photo-emitters that are sensitive up to longer wavelengths and have a higher photoemission efficiency close to the threshold.

Under certain conditions, an external electric field can have a measurable effect on photoemission efficiency. Schottky has shown that an *accelerating* electric field established near the surface of the material reduces the potential barrier there by an amount

$$\Delta W_{\rm ph} = \sqrt{\frac{e \mathscr{E}}{4\pi\varepsilon_0}} \tag{A1.7}$$

where ε_0 is the permittivity of free space, e is the electron charge, and \mathscr{E} is the electric field strength at the surface. This effect, known as the *Schottky effect*, shifts the photoemission threshold towards the longer wavelength and thereby improves the sensitivity of the layer, mainly close to the threshold. However, as can be shown by substituting practical values in Eq.A1.7, the Schottky effect is appreciable only at electric field strengths greater than several kilovolts per centimetre. Fields of that strength normally occur only in microchannel-plate photomultipliers.

A1.1.4 Photoelectron energy distribution

In travelling toward the surface, excited electrons lose energy due to multiple inelastic collisions. This is the main factor contributing to the more or less complicated angular and energy distributions with which they are emitted. These distributions depend on the crystalline structure of the material and vary with the energy of the incident flux.



Fig.A1.8 Relative distribution of photoelectron energies, E_{ph}, from a layer of SbKCs at 290 K, for incident photon energies (a) from 2.15 eV to 3.06 eV, and (b) from 4.28 eV to 5.12 eV



Fig.A1.9 Photoelectron energy distribution (in electrons per photon per eV) from a layer of GaAs(Cs) for incident-photon energies (a) from 1.4 eV to 2.2 eV, and (b) from 1.8 eV to 3.2 eV

Studies of semiconductor photo-emitters of the alkali-antimony type have shown their photoelectron energy distributions to have certain common characteristics (see Fig.A1.8). For low-energy incident photons (less than about 3 eV, Fig.A1.8(a)) the distribution curve has a single peak corresponding to an average energy, with respect to the vacuum level, of several tenths of an electron-volt; this average increases with the energy of the incident photons. For higher-energy incident photons the curve widens (Fig.A1.8(b)) and has two poorly-defined peaks: one for slow electrons whose average energy is independent of that of the incident photons, and one for fast electrons whose energy increases with that of the photons.

The photoelectron energy distribution curves of semiconductors with negative electron affinity (Fig.A1.9) also have identifiable peaks. With these materials the emitted electrons are mostly in thermal equilibrium; so, for incident photon energies less than 1.6 eV, the energy distribution curve of GaAs(Cs), for example, has a single peak centred around 1.4 eV, corresponding to electrons from the top of the valence band. The average energy of these electrons with respect to the vacuum level is about 0.15 eV – the value of the negative electron affinity. As the energy of the incident photons increases, a second peak due to higher-energy electrons gradually appears; this is better defined than the peak for alkali-antimony semiconductors and corresponds to electrons in thermal equilibrium at a higher level of the conduction band.

A1.2 Secondary emission

As in the case of photoemission, three stages can be distinguished:

- absorbed primary electrons impart energy to electrons in the material
- energized electrons diffuse through the material
- electrons reaching the surface with sufficient excess energy escape into the vacuum.

Within the energy range of interest (up to a few hundred electron-volts), there are four fundamental processes by which the primary electrons lose energy:

- ionization of atoms at deep energy levels; this process comprises a series of events during which increasing numbers of electrons share energy by elastic and inelastic interactions, part of the energy being spent in the creation of phonons (heat)
- generation of X-rays due to rearrangement of the electron shells; the X-rays can either be emitted from the material or can excite other electrons (Auger electrons)
- excitation of electrons between valence band levels
- excitation of volume and surface plasmons.

The first process is the most important when the electrons have sufficient energy for the probability of wide-angle elastic scattering to be high. This has the effect of causing a

back-diffusion (back scatter) of 10% to 30% of the incident electrons and *dispersion* of the electron beam.

The energy loss of a primary electron of energy E_p is given by

$$\frac{dE}{dx} = \frac{K}{E_{p}} \ln \left(\frac{aE_{p}}{E_{e}} \right)$$
(A1.8)

where K is a constant that allows for the characteristics of the material, a is almost unity, and E_e is the average excitation energy. The depth of penetration, R, of the electron beam into the material is given by

$$\mathbf{R} = \int_{0}^{\mathbf{E}_{p}} \frac{\mathrm{d}\mathbf{E}}{|\mathrm{d}\mathbf{E}/\mathrm{d}\mathbf{x}|}$$

Several approximations of Eq.A1.8 have been proposed to account for the transparency of thin layers to electrons. One of them, valid only for low energy electrons ($E_p < 10 \text{ keV}$), introduces the concept of penetration depth R defined by the relation:

$$\frac{dE}{dx} = -\frac{E_p}{R}$$
(A1.9)

Within a limited energy range the penetration depth varies as E_p^{α} , where α is slighly greater than unity but depends on the atomic number of the material. It has been shown that for secondary emission materials such as Al₂O₃, MgO and BeO

$$R = \frac{1.15 \times 10^{-5}}{\rho} E_{p}^{1.35}$$
(A1.10)

where R is in cm, ρ in g/cm³, and E_p in keV; for initial energies less than 1 keV, R is therefore only a few nanometres. Thus secondary emission is clearly a surface phenomenon.

Because of the diversity of energy exchange processes involved, it is difficult to derive an exact expression for the number of free carriers as a function of penetration depth. It is reasonable to assume, however, that the density of free electrons is proportional to the energy loss of the incident beam; that is,

$$\frac{\mathrm{dn}}{\mathrm{dx}} = \frac{1}{\varepsilon} \left| \frac{\mathrm{dE}}{\mathrm{dx}} \right|$$

where n is the number of free electrons, and ε the average energy required to generate a single electron-hole pair. Then, from Eq.A1.9,

$$\frac{\mathrm{dn}}{\mathrm{dx}} = \frac{1}{\varepsilon} \frac{\mathrm{E}_{\mathrm{p}}}{\mathrm{R}} \tag{A1.11}$$

The electrons freed in the material migrate toward the surface. In doing so, they lose a fraction of their excess energy by collision with other electrons or by interaction with the crystal lattice (creation of phonons). Only the electrons arriving at the surface with sufficient excess energy to overcome the potential barrier are released and so contribute to the *secondary emission*. The closer to the surface the electrons are excited, the greater their probability of escape; as in the case of photoemission, the probability decreases exponentially as the depth increases. On the other hand, the total number of electrons excited increases in proportion to the initial energy E_p of the primary electrons.

The ratio of the number of secondary electrons emitted to the number of primary electrons is called the *secondary emission coefficient* (δ). This ratio first increases with the initial energy of the primary electrons to a maximum (when the initial energy of the primary electrons is E'_p), then decreases gradually as the proportion of electrons excited at greater depth in the material increases (Fig.A1.10).



Fig.A1.10 Relative variation of secondary-emission coefficient δ as a function of primary-electron energy E_p . For energies above E'_p , δ falls off proportional to $E_p^{1-\alpha}$

Just as metals are mediocre photo-emitters, they are also poor secondary emitters. They are characterized by low escape depths (about 3 nanometres) and low secondary emission coefficients (between 1 and 2) corresponding to low primary energies (about 500 eV). On the other hand, some insulators and those semiconductors that have low

potential barriers are good secondary emitters. Such materials may have escape depths of a few hundredths of a micron.

As with photoemission, and for the same reasons, semiconductors with *apparent negative electron affinity* make it possible to obtain much higher secondary emission coefficients. In heavily p-doped semiconductors on which an electropositive metal such as caesium is deposited, the electrons are in thermal equilibrium with the bottom of the conduction band and can more easily escape into the vacuum (Fig.A1.7b). The average escape depth is considerably increased and can equal the diffusion length of the electrons in thermal equilibrium.

The energy distribution of secondary electrons is quite wide. Figure A1.11 shows an example for a material with positive electron affinity. The part marked (S), asymmetrically distributed about a mean energy of a few electron-volts, is the spectrum of real secondary electrons. The narrow peak (P) represents primary electrons that are reflected almost without energy loss. And the long, low-level, intervening line represents the few primary electrons that undergo multiple inelastic collisions.



Fig.A1.11 Distribution of secondary-electron energies E above the Fermi level ${\rm E_{F}}.$