

PHOTOMULTIPLIER TUBES

principles & applications

Re-edited September 2002 by S-O Flyckt and Carole Marmonier**,
Photonis, Brive, France*

**Email: so.flyckt@photonis.com*

***Email: c.marmonier@photonis.com*

FOREWORD

For more than sixty years, photomultipliers have been used to detect low-energy photons in the UV to visible range, high-energy photons (X-rays and gamma rays) and ionizing particles using scintillators. Today, the photomultiplier tube remains unequalled in light detection in all but a few specialized areas.

The photomultiplier's continuing superiority stems from three main features:

- large sensing area
- ultra-fast response and excellent timing performance
- high gain and low noise

The last two give the photomultiplier an exceptionally high gain x bandwidth product.

For detecting light from UV to visible wavelengths, the photomultiplier has so far successfully met the challenges of solid-state light detectors such as the silicon photodiode and the silicon avalanche photodiode. For detecting high-energy photons or ionizing particles, the photomultiplier remains widely preferred. And in large-area detectors, the availability of scintillating fibres is again favouring the use of the photomultiplier as an alternative to the slower multi-wire proportional counter.

To meet today's increasingly stringent demands in nuclear imaging, existing photomultiplier designs are constantly being refined. Moreover, for the analytical instruments and physics markets, completely new technologies have been developed such as the foil dynode (plus its derivative the metal dynode) that is the key to the low-crosstalk of modern multi-channel photomultipliers. And for large detectors for physics research, the mesh dynode has been developed for operation in multi-tesla axial fields. Recent developments include very large hemispherical photomultipliers with excellent time response for cosmic ray experiments, and ultra-fast tubes with a time jitter of less half a nanosecond.

This book describes the operating principles of the photomultiplier tube and surveys its many diverse applications, such as medical imaging, nuclear and high-energy physics including the latest cosmic-ray research aimed at opening new windows on the universe.

CONTENTS

CHAPTER 1	Operating principles and construction
CHAPTER 2	Fundamental characteristics
CHAPTER 3	Dark current and noise
CHAPTER 4	Operating characteristics and considerations
CHAPTER 5	Supply and operating advices
CHAPTER 6	Scintillation counting
CHAPTER 7	Applications of scintillation counting and analysis
CHAPTER 8	Non-scintillator applications of photomultipliers

INDEX

- absolute spectral sensitivity 2-2
- absorbed photon flux 1-25
- absorption coefficient 1-27
- activation analysis 7-17
- active dividers 5-10
- aerogel 8-14
- afterpulses 4-40
 - afterpulse ratio 4-40
 - luminous reactions 4-40, 4-53
 - residual gases 4-40
- AGASA 8-19
- air-shower experiments 8-18
- aluminium layer 1-8, 1-9
- AMANDA 8-17
- Altuglas 6-5
- ambient light 3-6
- ambient pressure 4-54
- ambient temperature effects 4-42
- analytical applications 9-11
 - liquid scintillation counting 7-12
 - radio-immuno assay (RIA) 7-15
- angle of incidence 4-9
- Animal PET 7-11
- annihilation peak 6-25
- anode
 - a current generator 1-17
 - blue sensitivity 2-7
 - collection space 1-16, 4-25
 - collection grid 1-16, 4-25
 - load 5-21, 6-9
 - luminous sensitivity 2-7
 - mosaic 1-21
 - radiant sensitivity 2-7
 - resistor 5-17
 - response pulse width 2-9, 4-58, 6-9
 - rise time 2-10, 4-6, 4-58
 - spectral sensitivity 2-2
 - typical anode pulse 2-9, 4-58
- anode load 5-21, 6-9
 - continuous operation 5-21
 - pulse operation 5-22, 6-8
- anode pulse rise time 2-10, 4-6, 4-58
 - delta pulse response 2-9, 4-56, 4-58
 - step response 2-10, 4-56, 4-58
- anode response pulse width 2-9, 4-10
- Animal PET 7-11
- ANTARES 8-17
- applying voltage 5-2
- attenuation length 6-6
- AUGER 8-19
- Auger electrons 1-33
- autocorrelation function 3-11, 5-46
- azimuthal response 7-5
- back scattering 1-34, 2-9
- back scatter peak 6-25
- background radiation 3-5
- BaF₂ cameras 7-8
- bandgap 1-27, 1-30
- bandwidth effects 3-25, 3-27
- bar code scanners 8-8
- BBQ 6-6
- Benthos sphere 1-23
- beta sources 7-9
- BGO 7-10
- binomial distribution 3-10
- bioluminescence 8-2
- bipolar layer 1-30
- blue sensitivity 2-3
- body scanners 7-6
- Boltzmann constant 3-3
- BooNE 8-15
- BOREXINO 1-15
- borosilicate 1-7, 4-54
- box dynodes 1-14
- calorimeters
 - electromagnetic 7-25
 - fibre 7-27
 - hadronic 7-26
 - lead glass 7-27
 - VPD and VPT 7-27
- cardiac cameras 7-6
- cathode current
 - cut-off voltage 1-11
 - fluctuations 3-11
 - proportionality 1-11
 - saturation 1-11
- cathode resistivity 2-16, 4-26
- cathode sensitivity 2-2

CdWO₄ 6-35
 CdZnTe 7-8
 centre-edge difference 2-11, 4-14
 chance coincidence rate 7-14
 channel PMT 1-23
 chemiluminescence 8-2
 Cherenkov effect 8-12

- aerogels 8-14
- detectors 8-12, 8-16, 8-18
- formula 8-12
- gases 8-14
- in PMT window 3-5, 4-53

 Child-Langmuir equations 4-25
 circular cage dynodes 1-15
 CLIA 7-16, 8-2
 collider detectors 7-23
 cold emission 3-4
 collection efficiency 1-8, 1-12, 2-5, 4-7, 4-8

- relative 2-5

 collimators 7-2, 7-7
 Compton

- backscattering 2-9, 7-22
- edge 6-24
- effect 6-24

 conduction band 1-26, 6-29
 conductive coating 5-3
 construction 1-2
 continuous mode 6-14
 cooling 4-49, 5-33
 cosmic ankle 8-19
 Corning blue sensitivity 2-2
 cosmic background 3-5, 4-53
 cosmic-ray detectors 7-30, 8-14
 counting accuracy 6-16
 counting plateau 6-15
 CPM 1-23
 cross talk 1-21
 CsF 7-10
 current amplification 1-3
 CZT 7-8

 dark current 2-12, 3-2, 4-52, 5-3, 5-33

- exposure to light 3-6
- field emission 3-4
- leakage current 2-17, 3-2, 4-49, 4-54
- measurements 2-17
- piezoelectric effects 2-18
- pulse rate 2-18, 4-52, 5-3
- pulse spectra 2-14
- reduction 5-33
- stabilization time 2-17, 5-3
- thermionic emission 3-3, 4-48, 5-33

 dark pulses 2-12, 4-52, 5-3
 decoupling capacitors 5-11
 dead time 6-16
 delta function 2-9, 2-12, 4-56, 4-58, 4-61

- centre-edge difference 2-11, 4-14
- jitter 2-11, 4-15
- response 2-9, 4-19, 4-56, 4-58
- standard deviation 2-11
- time resolution 2-11, 4-60
- transit-time difference 2-11
- transit-time spread 2-11, 4-15, 4-17

 density measurement 7-20
 depth-of-interaction 7-8
 detectivity 3-24
 distance curve 7-4
 drift compensation 5-38
 DUMAND-Hawaii 8-16
 DUMAND-Lake Baikal 8-16, 8-18
 dynode

- box dynodes 1-14
- circular cage 1-15
- crossed-field 1-18
- emission coefficient 1-13
- foil 1-16
- GaP(Cs) 1-17
- geometry 1-14, 1-15
- iterative 1-12
- linear focusing 1-9, 1-14
- material 1-13
- mesh 1-15
- microchannel plates 1-18
- signal 5-19
- venetian blind 1-9, 1-14

 electrolysis of glass 3-8
 electron diffusion 1-28
 electron multiplier 1-12
 electron-optical input 1-8, 7-5

- triode 1-9

 emission spectrum 6-38
 energy bands 1-26
 energy resolution 2-14, 3-14, 3-19, 6-20

- energy dependence 2-14
- radiation sources 2-15, 6-21, 6-25, 7-9, 7-12, 7-14, 7-18, 7-24, 7-25
- statistical fluctuations 2-14

with scintillators 2-15, 3-14, 3-19
 energy spectrum 6-21
 composition 6-24
 energy spectrometry 6-19
 environmental considerations 4-42
 atmosphere 4-54
 magnetic fields 4-51
 mechanical stress 4-55
 radiation 4-52
 temperature 4-42
 expectation value 6-17
 exposure
 ambient light 3-6
 humidity 4-54, 5-3
 light flash 3-6
 radiation 4-52
 temperature 4-43
 equipotential lines 1-8
 equivalent anode dark current input 3-23
 equivalent noise bandwidth 3-24, 3-32
 equivalent noise input 3-23
 escape depth 1-28
 escape peak 6-25

 fall-off curve 7-4
 FDG 7-10
 feedback resistance 4-2
 Fermi level 1-26
 field emission 3-4
 filtered sensitivity 2-2
 flame spectrometry 8-4
 Fluor-18 7-10
 flying spot scanners 8-8
 Fly's Eye 8-19
 foil dynodes 1-16
 Fourier transform 3-12
 fused silica, quartz 1-7, 4-54
 FWHM 6-28

 GaAs, GaInAs, GaInAsP 1-5, 1-27, 1-33
 gain 1-2, 1-12, 4-6, 5-30
 adjustment 5-30
 curves 5-24
 fluctuations 1-12
 gain-voltage characteristics 2-13
 gamma cameras 7-2
 azimuthal response 7-4
 collimators 7-2
 distance curve 7-4
 energy window 7-3
 fall off curves 7-4
 PHR 7-6
 spatial resolution 7-6
 SPECT 7-7
 stability 7-4
 summation signals 7-3
 tomography 7-7
 uniformity 7-4
 weighting factors 7-3
 whole body 7-6
 gamma sources 7-22
 GaP(Cs) dynodes 1-17, 1-19
 gating 5-37
 Gaussian distribution 3-14
 Gaussian light pulse 4-60
 general purpose tubes 1-8
 GRB 8-18
 green sensitivity 2-5
 GSO 6-35, 7-11
 GZK theory 8-19

 He penetration 4-39, 4-40, 4-54
 HEGRA 8-18
 hexagonal array 7-6
 high energy physics 7-23
 high pressure 4-54
 high-voltage polarity 3-8, 5-2
 HiRes 8-19
 hodoscopes 7-30
 hot electrons 1-28
 HPD 1-23
 humidity 4-54, 5-3
 hybride PMTs

 ICECUBE 8-17
 industrial gamma sources 7-22
 initial velocities 1-8, 4-16
 inorganic scintillators 6-35
 input optics 7-5
 insulators 1-26
 integral quantum efficiency 6-33
 interference filters 2-16

 jitter 2-11, 4-15, 4-17
 Johnson noise 3-17

 K27 6-6, 6-7, 6-44
 KAMIOKANDE 8-14, 8-15

KamLAND 8-16
 laser scanners 8-9
 laser telemetry 8-10
 leakage current 3-2
 LIA 8-4
 life of a PMT 4-33, 4-54
 light guides 6-4, 6-6
 lime glass 1-7, 4-40, 4-54
 linear-focusing dynodes 1-14
 linearity 4-20, 4-56, 5-25
 cathode resistivity 4-26, 4-48
 composite radiation method 4-30
 damping resistors 4-25
 differential linearity 4-30
 divider current 4-20
 dual pulse method 4-30
 integral linearity 4-30
 measurements 4-28
 overlinearity 4-23
 power supply 4-20
 progressive voltage dividers 4-26, 5-7
 reservoir capacitors 4-24, 5-11
 space charge 4-25
 three pulse burst method 4-31
 variation with high-voltage 4-26
 XY method 4-29
 zener diodes 4-24, 5-9
 liquid scintillation counting 7-12
 lithium fluoride 1-7
 low pressure 4-54
 LSND 8-15
 LSO 6-35,
 LSO 7-11
 Lucite 6-4
 luminous sensitivity 2-2, 2-15

 magnesium fluoride 1-7
 magnetic monopoles 8-16
 magnetic fields 1-20, 4-51
 magnetic shielding 5-3, 5-35
 Malargüe 8-19
 mamography scanners 7-11
 mean gain deviation 4-35
 medical applications 7-2
 mesh dynodes 1-15
 metal photocathode 1-26
 metal-can multiplier 1-21
 metal secondary emitters 1-36

 microchannel-plate PMTs 1-18
 good spatial resolution 1-20
 limited life time 1-21
 limited pulse charge capacity 1-21
 limited rate and linearity 1-21
 magnetic field immunity 1-20, 4-51
 very fast response 1-20
 microphony 4-55
 MiniBooNE 8-15
 MINOS 8-15
 moisture 2-17
 monochromatic sensitivity 2-2, 2-16, 4-42
 multichannel analyser 6-19
 multi-channel PMTs 1-21, 8-15
 multi-electron response 4-13
 mu-metal 5-36
 muon trackers 7-29

 Nautilus sphere 1-23
 negative electron affinity 1-5, 1-13, 1-28, 1-36
 negative polarity 5-2
 NEMO 8-17
 NESTOR of Pylos 8-17
 neutrino oscillation 8-15
 neutrino point sources 8-16
 neutron activation analysis 7-20
 noise
 equivalent power 3-23, 3-27, 3-29
 Johnson noise 3-17
 multiplier contribution 3-15
 photon noise 3-9
 shot noise 3-18
 spectrum 3-12
 statistical nature 3-8
 with scintillators 3-14, 3-18
 non-destructive analysis 7-17
 non-proportional light yield 6-20
 non-scintillator applications 8-2
 nuclear medicine 7-2
 number of stages 4-6

 oil-well logging 7-23
 OPERA 8-15
 operating advice 5-2
 operating modes 6-14
 continuous mode 6-14
 counting plateau 6-15
 pulse mode 6-14
 pulse counting 6-15

- operating range 5-23
 - choice of working point 5-25
 - gain and dark current 5-24
 - gain adjustment 5-30, 5-38
 - large dynamic range 5-26
 - linearity compromises 5-29
 - working diagram 5-28
- operational amplifiers 5-42
- optical grease 6-4
- organic scintillators 6-41, 6-45
- output connections 5-17
 - anode resistor 5-17
 - anode load 5-21
 - dynode signal 5-19
 - output cable 5-18
- overlinearity 4-22

- pair production 6-25
- PbWO₄ 6-35, 6-36
- PET scanners 7-8
- photocathode
 - absolute spectral sensitivity 2-2
 - diameter 4-6
 - luminous sensitivity 2-2
 - materials 1-4
 - measurements 2-15
 - monochromatic sensitivity 2-2, 4-42
 - opaque 1-4
 - radiant sensitivity 1-5, 2-2
 - resistivity 2-16, 4-26, 4-48
 - semi-transparent 1-4
 - spectral sensitivity 1-5, 4-5, 4-42
 - survey table 1-5
 - temperature 4-42
 - windows 1-5
- photodiode versus PMT 4-2
- photoelectric effect 6-24
- photoelectron energy 1-31
- photoemission 1-2, 1-24
 - angle of incidence 4-9
 - apparent electron affinity 1-30
 - apparent negative electron affinity 1-31, 1-36
 - bipolar layer 1-30
 - coefficient 1-13, 1-35
 - electron diffusion 1-28
 - energy distribution 1-31
 - photon absorption 1-24
 - quantum efficiency 1-24, 6-33
 - Schottky effect 1-31
 - temperature dependence 4-43
 - the surface barrier 1-28
 - theory 1-24, 1-33
 - threshold 1-5, 1-29, 1-31, 3-3
 - wavelength dependence 4-7
- photometric units 2-2
- photon absorption 1-24
- photon counting 5-48
- photo peak 7-3
- PHR 2-14, 7-4
- pilot-U 6-45
- pin connections 5-41
- piezoelectric effect 2-18
- Planck's constant 2-4
- PMT input optics 7-5
- Poisson distribution 3-9, 3-10, 3-16
- polarity 5-2
- POPOP 7-13
- positron
 - emitters 7-9
 - scanners 7-9
 - sources 7-9
- potential barrier 1-23, 1-35
- proton decay experiments 8-14
- proton life 8-14
- pulse amplitude distributions 6-22
- pulse counting 6-15
- pulse height spectrum 6-21
 - composition 6-24
 - FWHM 6-21
- pulse mode 5-22, 6-14
- pulse response 4-10, 4-56
 - FWHM 4-11
 - sigma 4-11
- pulse sampling 5-50
- PWO₄ calorimeters 7-26

- quadratic detection 5-47
- quantum efficiency 1-5, 1-24, 2-4, 6-33
- quartz, fused silica 1-7, 4-54

- radiant sensitivity 1-5, 2-2, 2-16, 4-42
- radiation absorption 6-42
- radiation effects 4-52
- radiometric units 2-2
- Raman spectrometry 8-7
- reducing photocathode area 5-33
- reflection coating 6-3
- reflection coefficient 1-26

- refractive index
 - scintillators 7-10
 - windows 1-7
- resolution time 6-15
- RIA counters 7-15
- Richardson's law 3-3, 4-48
- rise time 2-10

- safety 5-41
- sapphire 1-7
- Schottky effect 1-31, 3-13
- scintillation counting 6-2
- scintillator coupling 6-3
 - direct 6-4
 - light guides 6-4, 6-6
- scintillator pulses 6-8
 - current pulse 6-8
 - decay time constant 6-8, 6-27
 - RC constant 6-9
 - voltage pulse 6-8
- scintillator
 - characteristics 6-32
 - coating 6-3
 - coupling 6-3
 - effectiveness 6-32
 - efficiency 6-32
 - emission spectra 6-32, 6-38, 6-43
 - finish 6-3
 - fundamentals 6-29
 - hygroscopy 6-41
 - inorganic 6-29, 6-35
 - integral quantum efficiency 6-33
 - interaction 6-20
 - light guides 6-4
 - luminescence centres 6-30
 - matching factor 6-33
 - organic 6-30, 6-40, 6-44
 - properties 6-35, 6-45
 - quenching centres 6-30
 - radiation absorption 6-36
 - response time 6-39
 - shape 6-2
 - survey table 6-35, 6-45
 - temperature effects 6-40
 - transparency 6-43
 - traps 6-30
- secondary emission 1-2, 1-13, 1-33
 - coefficient 1-35, 2-9
 - current gain 2-6, 2-17
 - exponent of voltage 1-12, 2-6
 - fluctuations 1-12
 - secondary paths 1-8
- semiconductor photocathode 1-26
- sensitivity contours 4-7
- separating signal from noise 5-45
 - autocorrelation 5-46
 - photon counting 5-48
 - pulse shape sampling 5-50
 - quadratic detection 5-47
 - synchronous detection 5-45
- shot noise 3-18
- signal processing 5-41
 - operational amplifiers 5-42
- signal-to-noise ratio 3-13, 3-22, 3-24, 3-27, 4-2
- signal transfer in linear systems 4-56
- Si hybrid PMTs 1-23
- Silastene 5-4
- single-electron response 4-12
- single-electron spectrum 2-8, 3-16
 - peak-to-valley ratio 2-8
 - resolution 2-9
- smart PMTs 1-22, 8-16, 8-17, 8-20
 - DUMAND experiments 1-22, 8-16
- SNO experiment 8-15
- socket leakage current 2-17
- solar neutrino experiments 8-14
- space charge 4-25
- spatial resolution 1-20
- SPECT 7-7
- spectral density 3-12
- spectral sensitivity
 - characteristics 1-4, 4-5
 - survey table 1-5
 - temperature influence 4-42
- Spectrosil™ 1-7, 4-54
- stability 4-32, 7-5
 - compensation 5-38
 - high current drift 4-33
 - hysteresis 4-39
 - long term drift 4-33, 5-38
 - low current drift 4-34, 5-38
 - mean gain deviation 4-35
 - operating life 4-33, 4-54
 - short term shift 4-36
 - temperature 4-42, 4-49
- standard deviation 3-10
- statistical
 - distributions 3-9
 - fluctuations 2-14, 3-2, 3-8
 - mean value 3-10

step response 2-10, 4-56, 4-58
 sum peaks 6-25
 SUPER KAMIOKANDE 8-15
 superposition principle 4-57
 supplementary illumination 5-39
 surface barrier 1-28
 synchronous detection 5-45

 temperature compensation 5-40
 temperature cycling 4-50
 temperature effects 4-42
 thermionic emission 3-3, 4-48, 5-33
 thermionic work function 1-28
 thermopile 2-16
 thickness measurement 7-20
 timing definitions 2-4
 time jitter 2-11, 4-15, 4-17
 time-of-flight detectors 7-28
 time resolution
 FWHM 2-11, 4-58
 open cathode 2-11
 sigma 2-11
 time spectrometry 6-26
 TMAE 7-8
 tracking detectors 7-24
 transit time
 difference 2-11, 4-14
 jitter 2-11, 4-15, 4-17
 spread 2-11, 4-15, 4-17
 tungsten filament lamp 2-15

 unit step 2-10, 4-56, 4-58
 useful life 1-21
 UV-glass 1-7

 vacuum level 1-28

 vacuum photodiodes (VPDs) 1-21, 7-27
 vacuum phototriodes (VPTs) 1-21, 7-27
 valence band 1-26
 venetian-blind dynodes 1-14, 2-14
 vertex detector 7-24
 Veto counters 8-14
 voltage dividers 1-13, 5-6
 decoupling capacitors 5-11, 5-14
 equal step 5-6
 intermediate 5-7
 multiple tubes 5-31
 multiple supplies 5-10
 progressive 5-7
 resistive 5-8
 wiring precautions 5-16
 zener diodes 5-9

 wave length shifters (WLS) 6-7, 6-43, 8-13
 white light sensitivity 2-2, 2-15
 Wiener-Khintchine theorem 3-12
 windowless photomultipliers 1-8
 windows 1-5, 1-6
 cut-off 1-7
 materials 1-7
 refractive index
 transmission curves 1-7
 work functions 1-29
 working point 5-28

 X-escape peak 7-19
 X-fluorescent analysis 7-19

 Y7 6-6, 6-7

 zener diodes 4-24, 5-9
 ZnWO₄ 6-35

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.

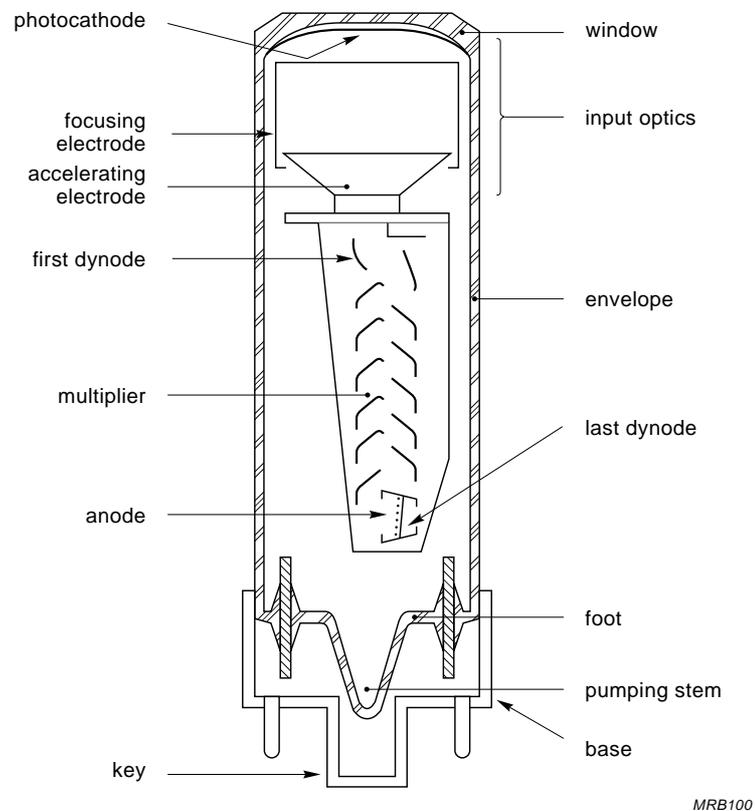


Fig.1.1 Elements of a photomultiplier based on the world's first fast PMT, the 56AVP, introduced by Philips (now Photonis) in 1956

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

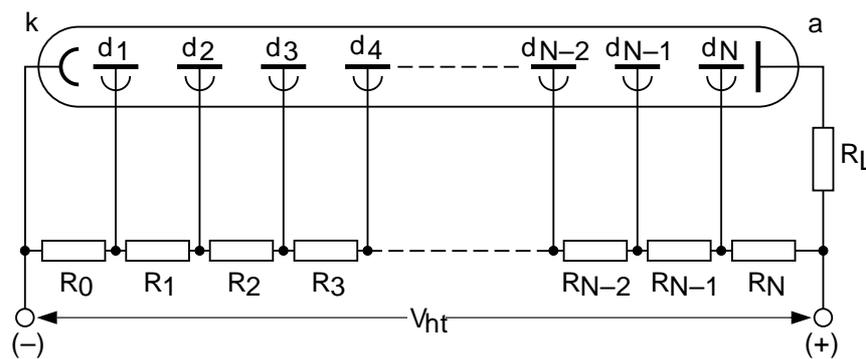
the gain of the dynode is g_1 , the number of resulting secondary electrons is $n_k g_1$. Then, if the second dynode has a gain g_2 , it in turn emits $n_k g_1 g_2$ electrons. The process repeats from dynode to dynode up to the anode where the electrons are finally collected. If N is the number of dynodes,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



MRB101

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), antimony-caesium (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.

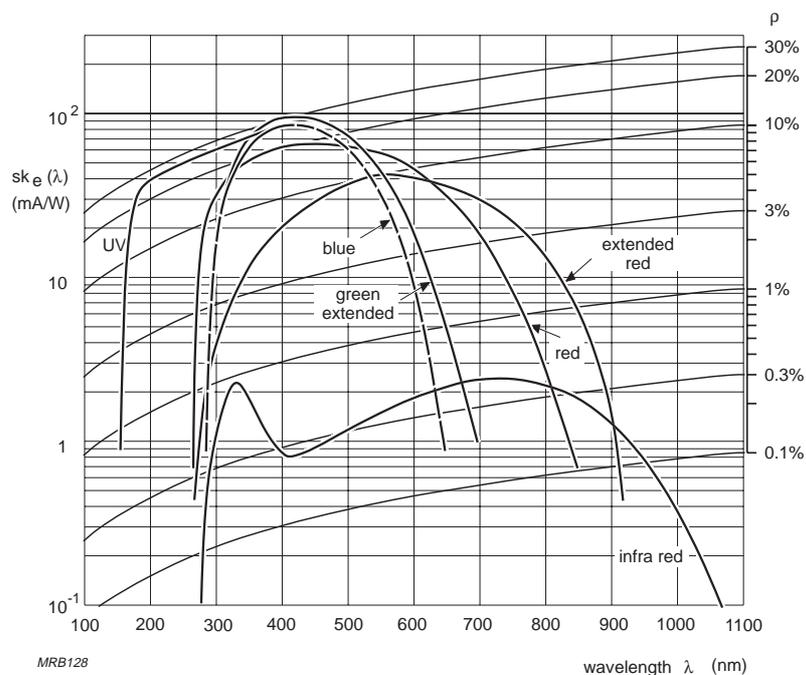


Fig.1.3 Spectral sensitivity characteristics for various photocathode types. Curves of uniform quantum efficiencies, ρ , are shown for reference; see ch.2.1.2.

Table 1.1 Composition and typical characteristics of photocathodes

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

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.

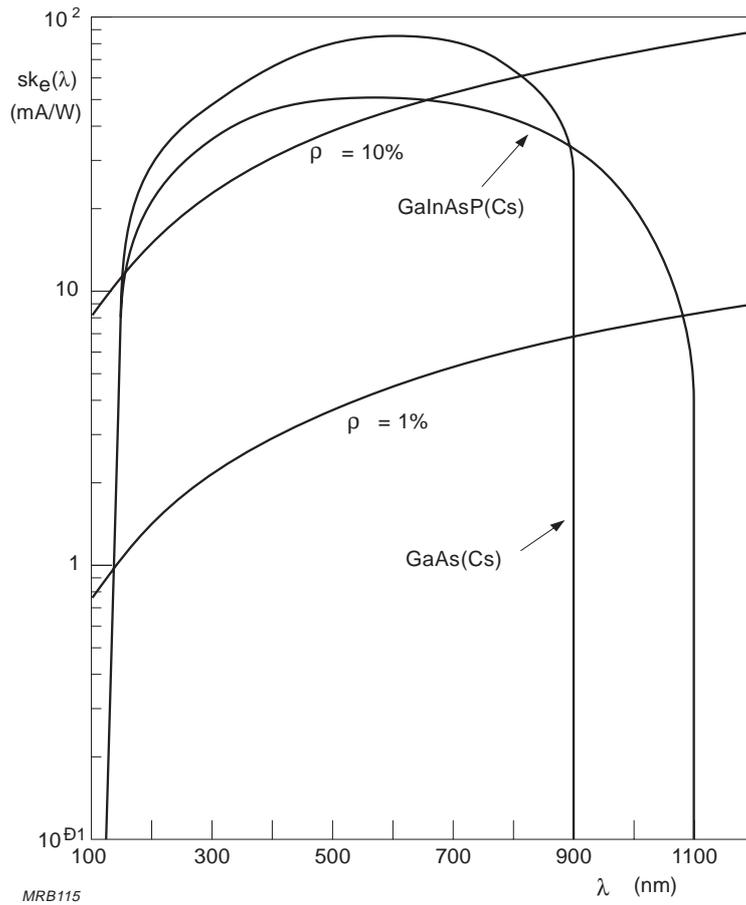


Fig.1.4 Spectral sensitivity characteristics for two negative-electron-affinity photocathode materials; quantum efficiency curves $\rho = 1\%$ and $\rho = 10\%$ shown for reference

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.

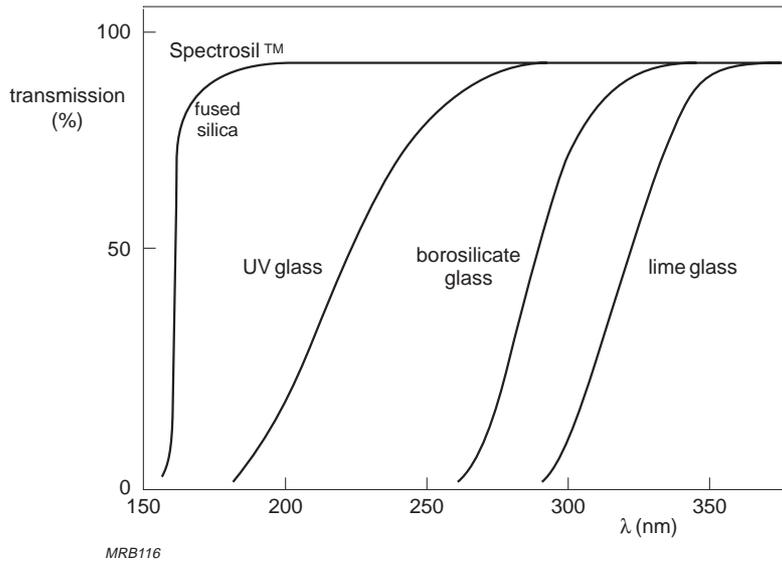


Fig.1.5 Transmission (%) as a function of wavelength λ for various glasses used in photomultiplier input windows (thickness 3 mm)

Table 1.2 Characteristics of glasses used in photomultiplier windows

Type of window	Cut-off wavelength (decrease to 10%) (nm)	Refractive index n at λ (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_2O_3)	145	1.80 400
MgF_2	115	1.40 400
LiF	105	1.40 400

Even with LiF or MgF₂, 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:

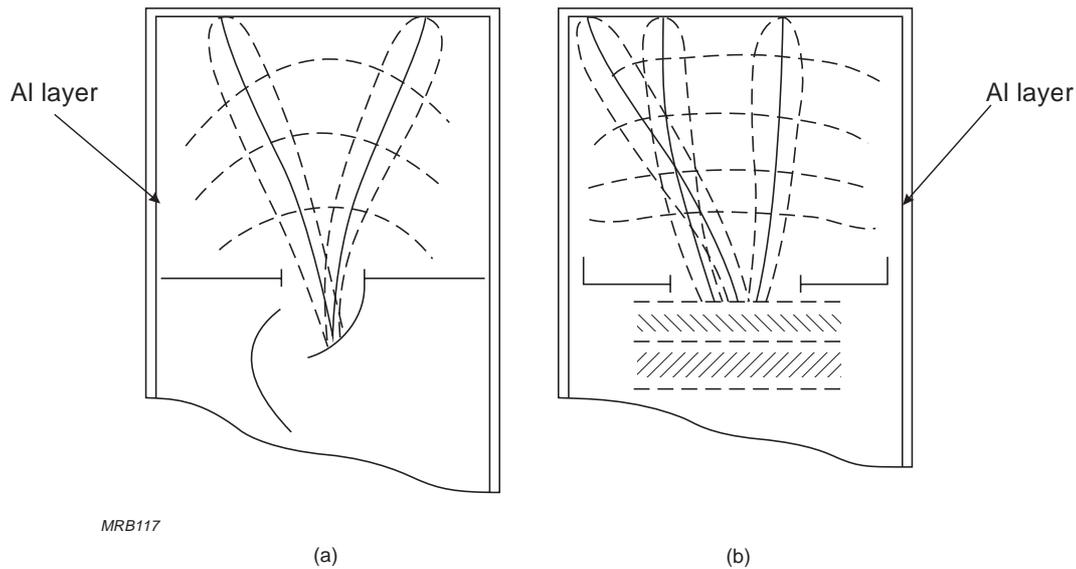
- It must be able to focus the maximum number of electrons onto the first dynode regardless of their initial velocity and regardless of where on the photocathode they originate. The ratio between the number of electrons reaching the useful area of the first dynode and the total number of electrons emitted 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.



MRB117
 (a) (b)
 Fig.1.6 Input electron-optics of a photomultiplier with (a) focusing dynodes and (b) venetian-blind dynodes, showing equipotential lines and electron trajectories

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

$$\Delta t_n = \frac{1}{\mathcal{E}e} \sqrt{2m_e \Delta E_n} \quad (1.1a)$$

where \mathcal{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

$$\Delta t_t \approx \frac{r}{v} \quad (1.1b)$$

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 $L + \Delta L$ is:

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

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).

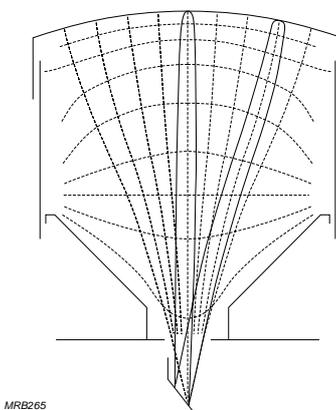


Fig.1.7 Example of a large fast photomultiplier input optics

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.

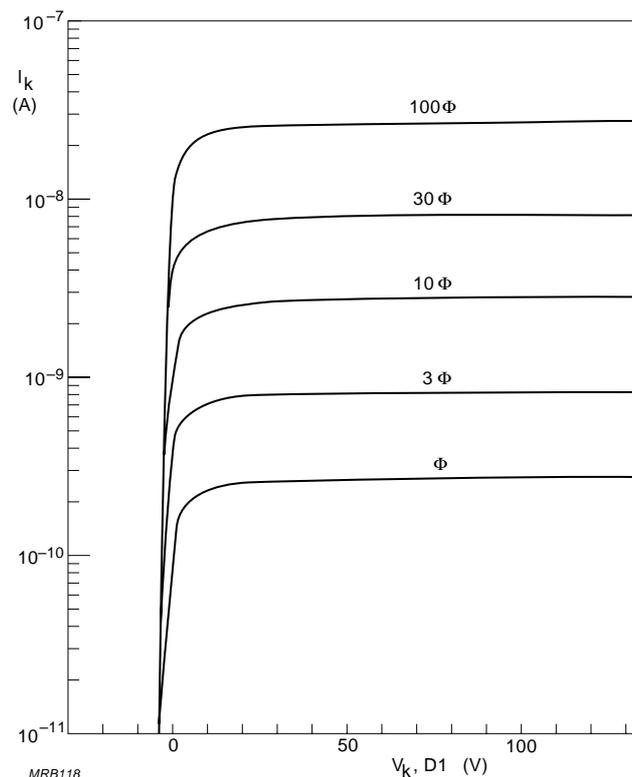


Fig.1.8 Photocurrent variation as a function of cathode-to-first-dynode voltage, with relative incident flux as parameter

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 :

$$n_{i-1} \delta_i = k_i V_i^\alpha \quad (1.3)$$

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

$$g_i = \delta_i n_i \quad (1.4)$$

and the total gain M of the multiplier is

$$M = \prod_{i=1}^N g_i \quad (1.5)$$

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 \bar{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:

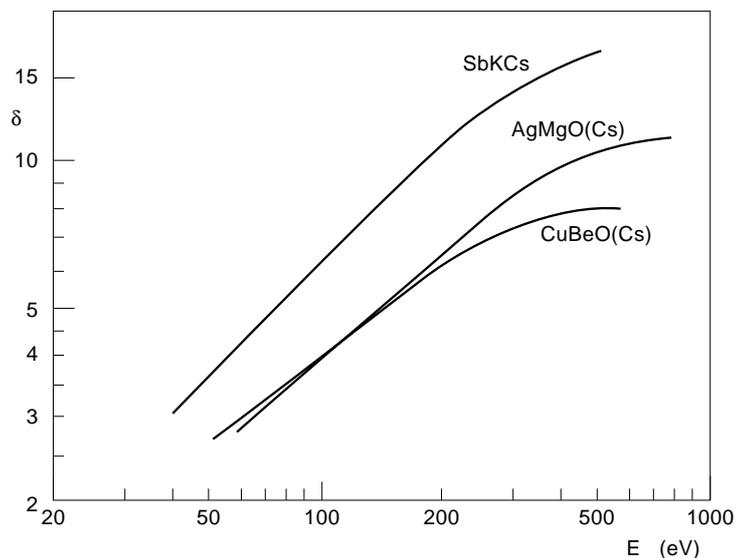
- variations in initial electron velocity
- variations in electron trajectories.

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 Al₂O₃, 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.



MRB119

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

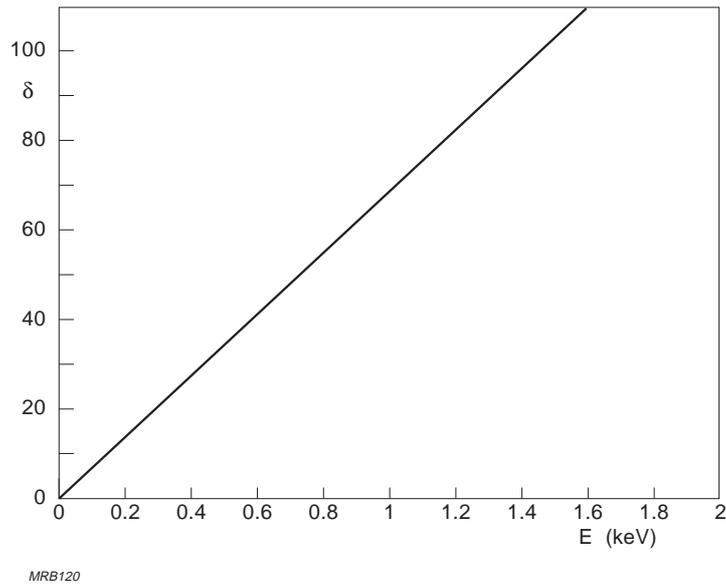


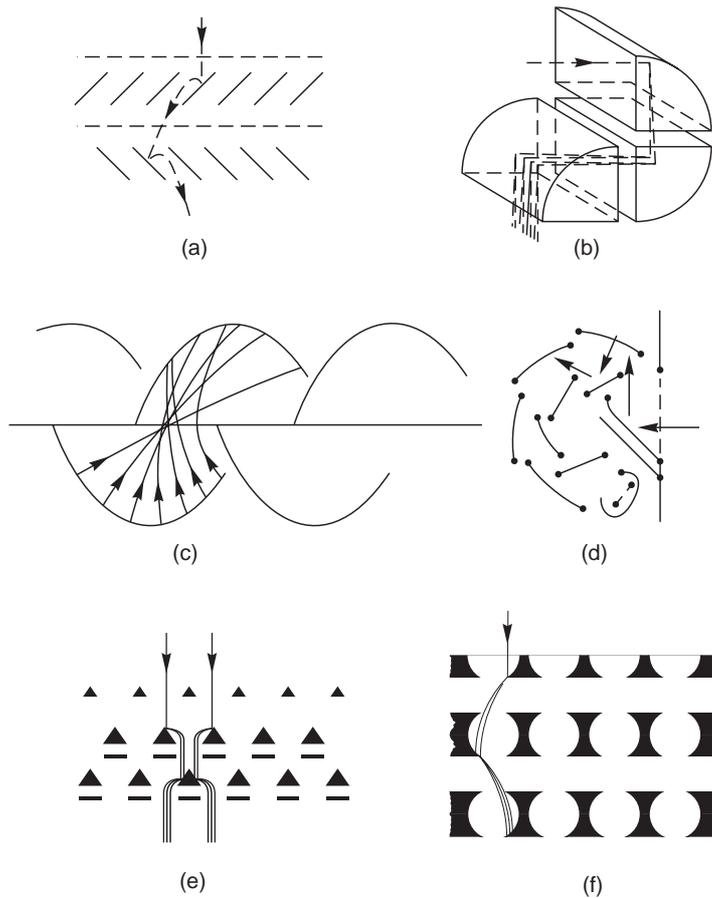
Fig.1.10 Secondary emission coefficient of GaP(Cs) as a function of incident primary-electron energy

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.



MRB121

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.

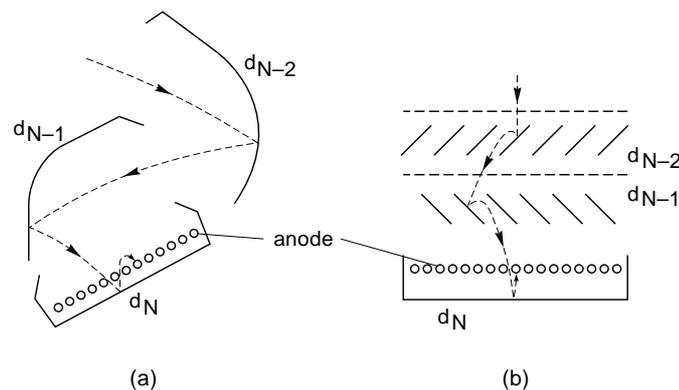
Foil dynodes (Fig.1.11(f)). These consist of perforated metal foils with well defined apertures. Their collection area is high and their collection efficiency from dynode to dynode is about the same as that of venetian blind dynodes. They can work in magnetic fields up to some tens of mT and their low 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.

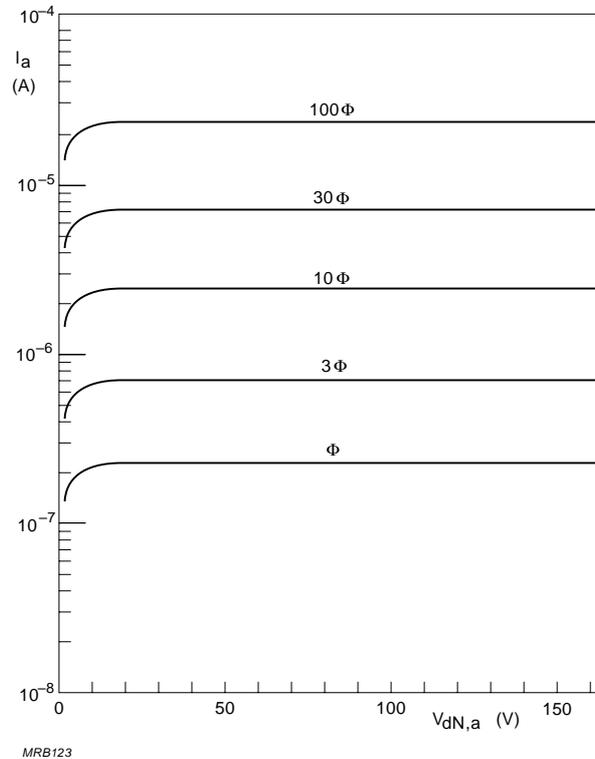


Fig.1.13 Anode current as a function of last-dynode-to-anode voltage, with relative incident flux ϕ as parameter

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 μm). Figure 1.15 shows the electron multiplication in a single microchannel; one electron at the input can give rise to as many as 10^8 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 10^4 ; higher gain can be obtained by arranging two microchannel plates in cascade.

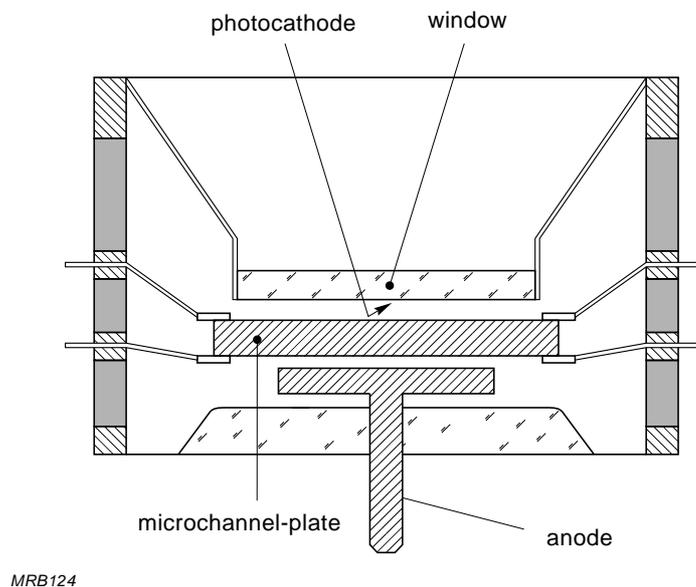
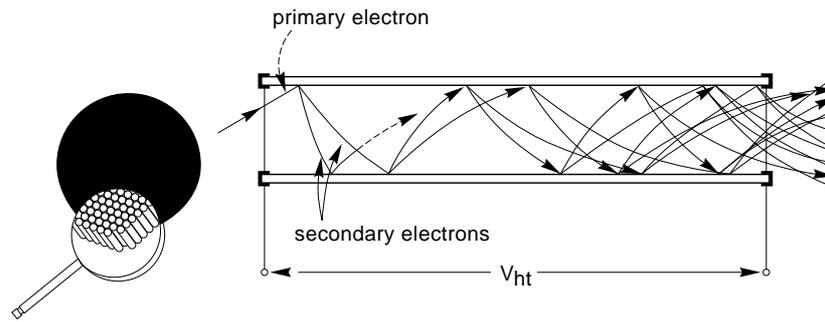


Fig.1.14 Section through a microchannel-plate photomultiplier



MRB125

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.

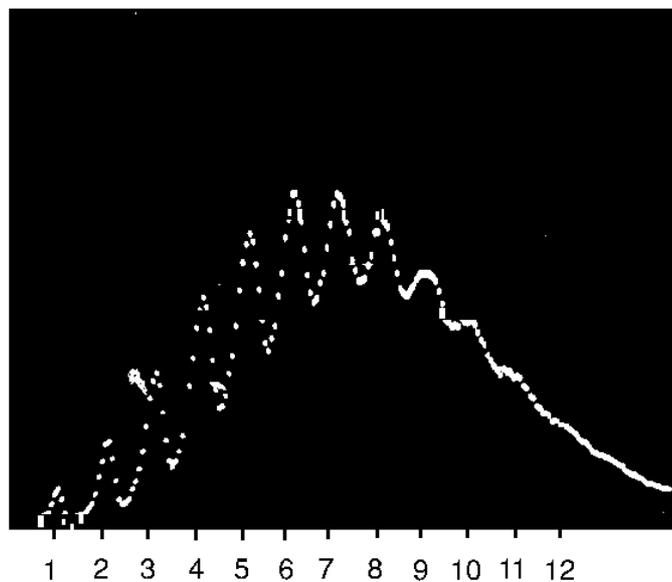
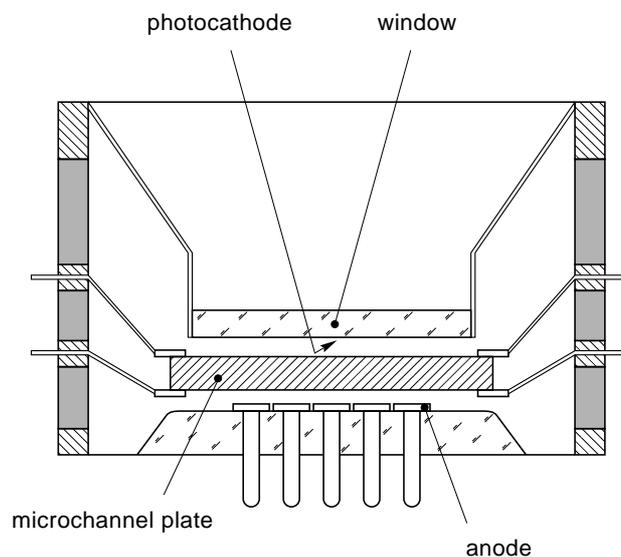


Fig.1.16 Distribution of anode charge per pulse of a microchannel-plate photomultiplier working in the saturation region: vertical axis, relative pulse frequency; horizontal axis, number of photoelectrons per pulse

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 $\mu\text{A}/\text{cm}^2$ and the long recovery time per channel.
- *Limitations of pulse charge capacity per channel* (5×10^{-14} to 5×10^{-13} 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 (5×10^{-10} to 5×10^{-9} C/cm²).
- *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, beam-profile 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 PbWO₄ 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.

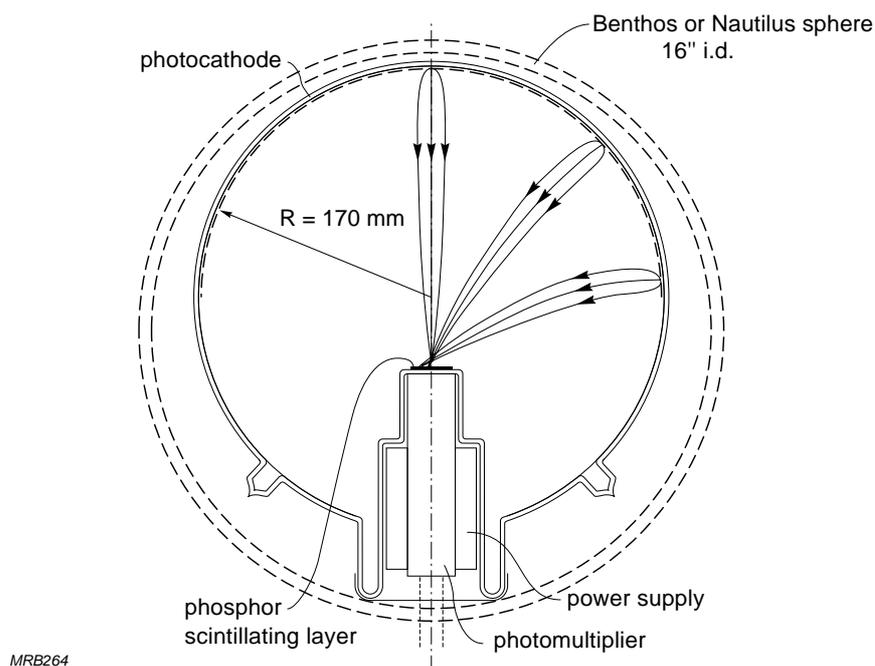


Fig.1.18 Smart photomultiplier combination in a glass pressure sphere for deep underwater muon and neutrino detectors (DUMAND)

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 photomultiplier* (CPM) a photocathode is evaporated onto a small entrance window (≈ 10 mm diameter) that is indium-sealed onto a spiralised single-channel 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.

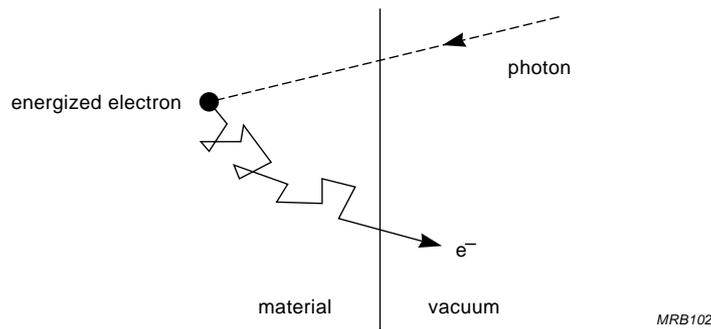


Fig.A1.1 Photoemission

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.

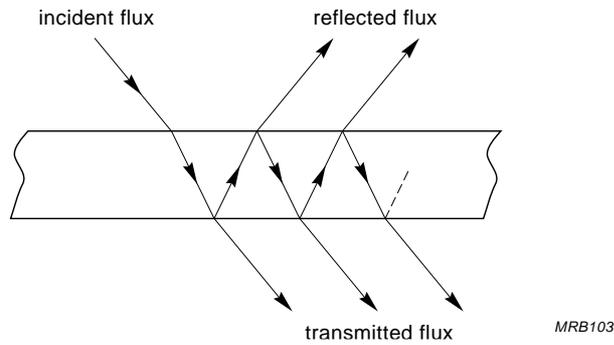


Fig.A1.2 Light transmission and reflection in a semi-transparent layer

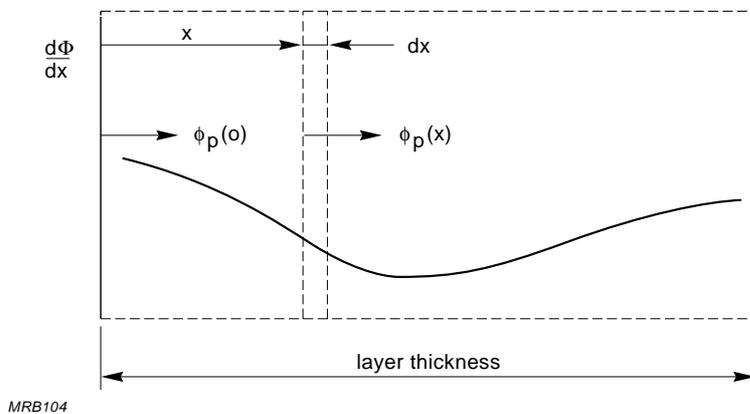


Fig.A1.3 Relative photon absorption $d\Phi/dx$ in a photoemissive layer

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

$$\frac{d\Phi_p(\nu, x)}{dx} = -\alpha(\nu) \Gamma(x) \Phi_p(\nu, x) \quad (\text{A1.1})$$

where $\alpha(\nu)$ 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:

$$\Phi_p(\nu, x) = \Phi_p(\nu, 0) [1 - r(\nu)] \exp -\alpha(\nu)x \quad (\text{A1.2})$$

where $\Phi_p(\nu, 0)$ is the incident flux and $r(\nu)$ the *reflection coefficient* at the input surface.

The absorption coefficient $\alpha(\nu)$ depends on the energy $h\nu$ 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 E_0 . 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.

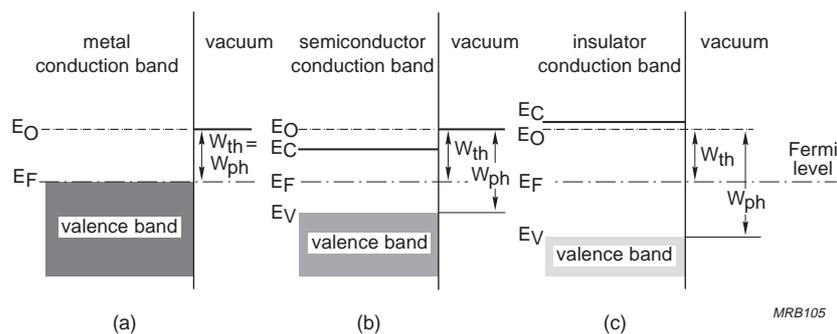


Fig.A1.4 Energy bands in (a) a metal, (b) a semiconductor, (c) an insulator

Figure A1.5 shows the variation in the absorption coefficient $\alpha(\nu)$ as a function of the energy $h\nu$ of the incident photons, for a few types of semiconductor. For most of them the absorption coefficient is greater than $10^4/\text{cm}$ for $h\nu > E_g$ (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.

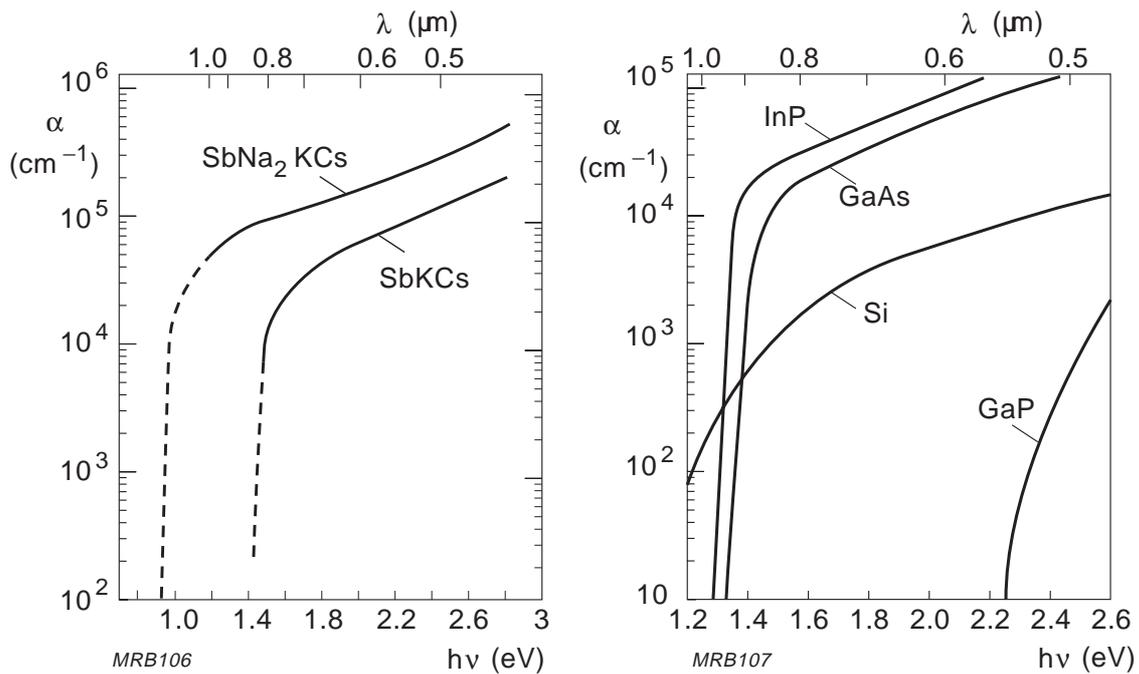


Fig.A1.5 Absorption coefficients α as functions of photon energy $h\nu$ for various photoemissive layers

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)

$$\frac{dn}{dx} = -\alpha_e \frac{d\Phi_p(\nu, x)}{dx} \quad (\text{A1.3})$$

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 $h\nu \geq W_{th}$. 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).

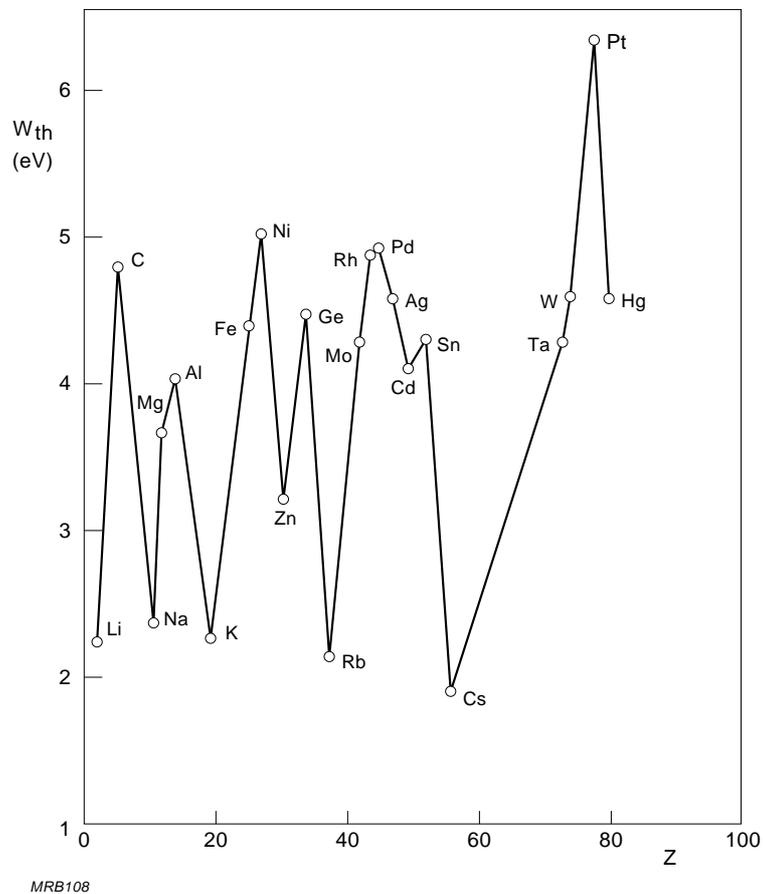


Fig.A1.6 Work functions of pure metals, in order of atomic number

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:

$$E_g = E_C - E_V \quad \text{and} \quad E_A = E_0 - E_C$$

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 $W_{th} = E_0 - E_F$.

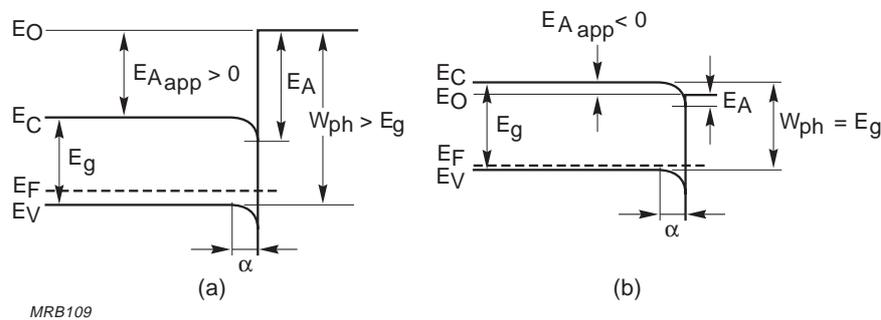


Fig.A1.7 Energy-band diagrams showing (a) positive and (b) negative electron affinity

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:

$$W_{\text{ph}} = E_{\text{g}} \quad (\text{A1.6})$$

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

$$\Delta W_{\text{ph}} = \sqrt{\frac{e\mathcal{E}}{4\pi\epsilon_0}} \quad (\text{A1.7})$$

where ϵ_0 is the permittivity of free space, e is the electron charge, and \mathcal{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.

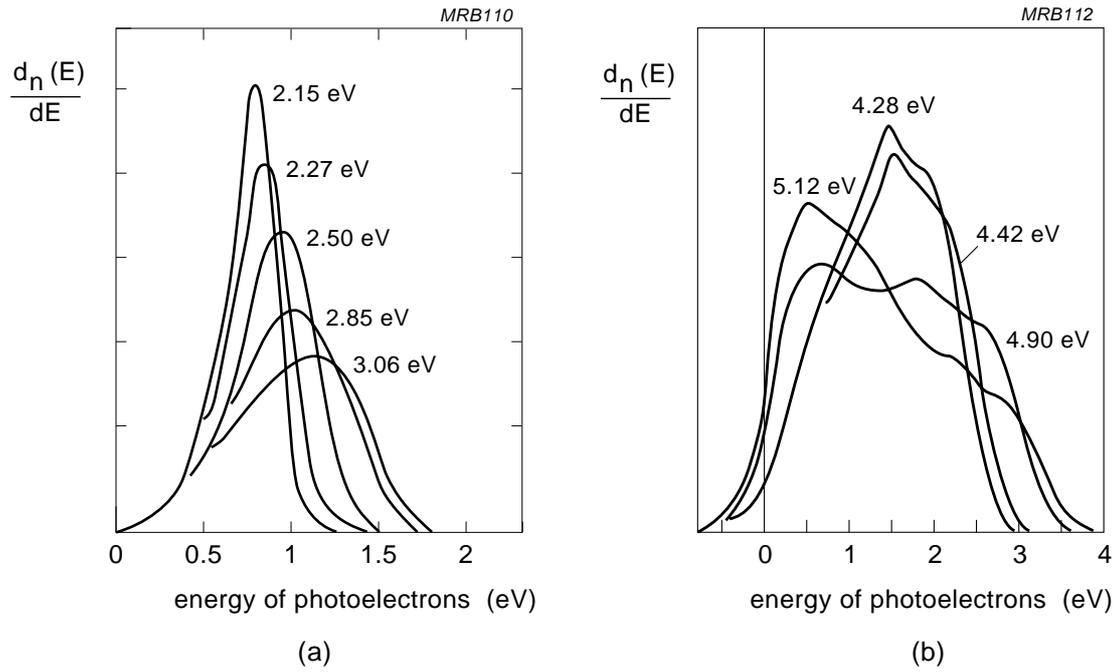


Fig.A1.8 Relative distribution of photoelectron energies, E_{ph} , from a layer of SbKCs at 290 K, for incident photon energies (a) from 2.15 eV to 3.06 eV, and (b) from 4.28 eV to 5.12 eV

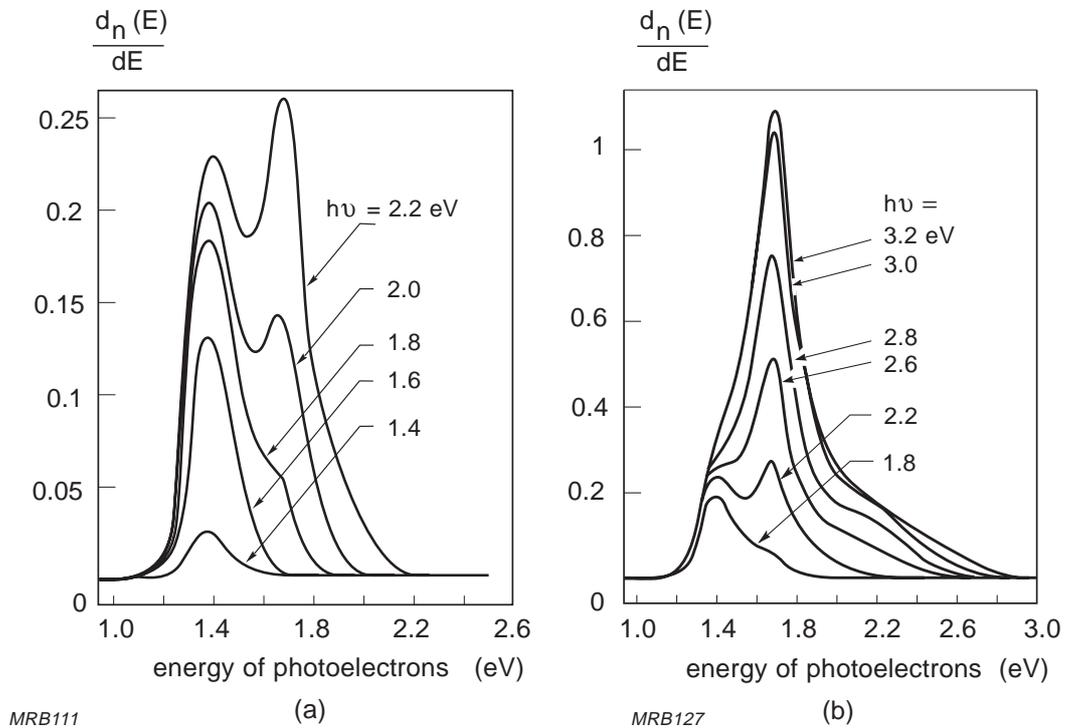


Fig.A1.9 Photoelectron energy distribution (in electrons per photon per eV) from a layer of GaAs(Cs) for incident-photon energies (a) from 1.4 eV to 2.2 eV, and (b) from 1.8 eV to 3.2 eV

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

$$\frac{dE}{dx} = \frac{K}{E_p} \ln \left(\frac{aE_p}{E_e} \right) \quad (\text{A1.8})$$

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

$$R = \int_0^{E_p} \frac{dE}{|dE/dx|}$$

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 ($E_p < 10$ keV), introduces the concept of penetration depth R defined by the relation:

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

Within a limited energy range the penetration depth varies as E_p^α , where α is slightly greater than unity but depends on the atomic number of the material. It has been shown that for secondary emission materials such as Al_2O_3 , MgO and BeO

$$R = \frac{1.15 \times 10^{-5}}{\rho} E_p^{1.35} \quad (\text{A1.10})$$

where R is in cm, ρ in g/cm^3 , 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,

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

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,

$$\frac{dn}{dx} = \frac{1}{\epsilon} \frac{E_p}{R} \quad (\text{A1.11})$$

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).

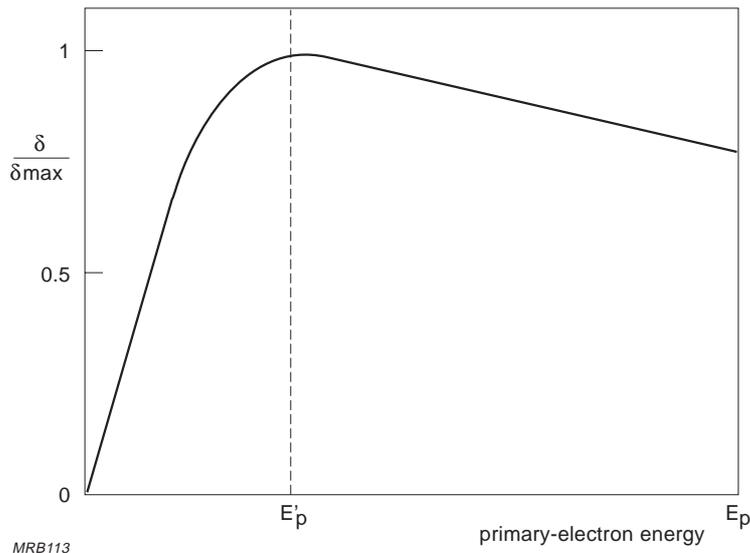


Fig.A1.10 Relative variation of secondary-emission coefficient δ as a function of primary-electron energy E_p . For energies above E'_p , δ falls off proportional to $E_p^{1-\alpha}$

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.

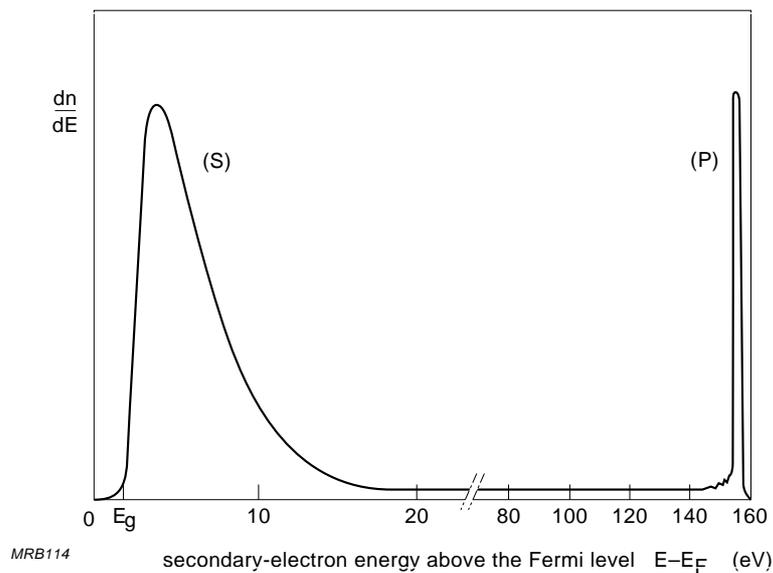


Fig.A1.11 Distribution of secondary-electron energies E above the Fermi level E_F .

CHAPTER 2

FUNDAMENTAL CHARACTERISTICS

2.1 Sensitivity characteristics

- 2.1.1 Photocathode sensitivity
- 2.1.2 Quantum efficiency
- 2.1.3 Input system collection efficiency
- 2.1.4 Gain
- 2.1.5 Anode sensitivity
- 2.1.6 Single-electron spectrum

2.2 Time characteristics

- 2.2.1 Response pulse width
- 2.2.2 Rise time
- 2.2.3 Transit time differences
- 2.2.4 Transit time spread, time resolution

2.3 Dark current

2.4 Energy resolution

2.5 Measurement of the characteristics

- 2.5.1 Cathode sensitivity
- 2.5.2 Gain
- 2.5.3 Dark current
- 2.5.4 Dark pulse rate

2.6 Interpretation of data

FUNDAMENTAL CHARACTERISTICS

The fundamental characteristics of any given type of photomultiplier are specified in its data sheets. This chapter describes what they are, what they signify, how some of them are measured and, finally, how they are commonly presented in the data sheets. Some of the characteristics relate to the tube as a whole, others to only one of the three functional parts

- photocathode
- electron-optical input system
- electron multiplier and anode.

Collectively, they can be divided into four classes relating to sensitivity, time, dark current, and energy resolution. (Stability and linearity characteristics are dealt with separately in Chapter 4.)

2.1 Sensitivity characteristics

2.1.1 Photocathode sensitivity S_k

This is the ratio of the cathode current I_k (less the dark current, see §2.3) to the incident flux Φ . Expressed in radiometric units it is called *cathode radiant sensitivity*:

$$S_k(\text{A/W}) = \frac{I_k (\text{A})}{\Phi_e (\text{W})} \quad (2.1)$$

Expressed in photometric units it is called *cathode luminous sensitivity*:

$$S_k(\text{A/lm}) = \frac{I_k(\text{A})}{\Phi_v(\text{lm})} \quad (2.2)$$

In most applications the radiation is not monochromatic but has a spectral composition to which the cathode is not uniformly sensitive (see Fig.1.3 and 1.4). To specify the photocathode sensitivity completely, therefore, it is also necessary to specify the spectral composition at which the stated sensitivity was measured.

Cathode radiant sensitivity

The radiant sensitivity is customarily specified for a specific wavelength λ . If $d\Phi_e$ is the incident flux increment in a wavelength increment $d\lambda$ centred on λ , and dI_k is the corresponding cathode current increment, then

$$S_{k,\lambda} = \lim_{d\lambda \rightarrow 0} \frac{dI_k}{d\Phi_e}$$

is the *monochromatic sensitivity* or *absolute spectral sensitivity* at the wavelength λ .

Relative spectral sensitivity is the ratio between the sensitivity at a given wavelength λ and at a reference wavelength, usually the one at which sensitivity is greatest. It can differ appreciably from one tube to another of the same type, particularly in the neighbourhood of the photoemission threshold and especially in tubes with S20, S20R or S1 cathodes. The radiant sensitivity of S20 cathodes, for instance, may vary by a factor of 10 at 800 nm but by a factor of only 2 at 450 nm.

Cathode luminous sensitivity

The luminous sensitivity is usually specified for a spectral composition that is typical for a particular application. A customary distinction is:

- luminous sensitivity to white light, which is used to characterize tubes intended for photometry; the reference illuminant for specifying this is a tungsten-filament lamp with a colour temperature of 2856 K.
- luminous sensitivity to blue light, which is used to characterize tubes intended for scintillation counting. (The emission of many scintillators peaks at about 430 nm.) This sensitivity, also called CB (Corning Blue) sensitivity, is based on the light of a 2856 K tungsten-filament lamp transmitted through a Corning C.S. No. 5-58 filter ground and polished to half stock thickness (Fig.2.1). In the data sheets this is specified as ‘blue sensitivity’ expressed in $\mu\text{A}/\text{lmF}$ (where F stands for filtered, indicating that the same unit can be used also for other specified filters, e.g for ‘red sensitivity’).

For a bialkali cathode (see Fig.1.3) the ratio of CB to white-light luminous sensitivity is

$$\frac{S_k \text{ (C.B.)}}{S_k \text{ (white)}} \approx 0.15$$

Radiant sensitivity is often specified at 400 nm; the relation between CB luminous sensitivity and the radiant sensitivity at 400 nm is

$$\frac{S_k \text{ (CB)}}{S_{k,400\text{nm}}} \approx 0.125 \times 10^{-3} \text{ (W/lm)}$$

Both these relations are subject to some variation from tube to tube.

Cathode sensitivity is never uniform over the whole surface of the cathode; moreover, the non-uniformity varies with wavelength (see Fig.4.3). For some applications it may be important to measure this non-uniformity and its wavelength dependence.

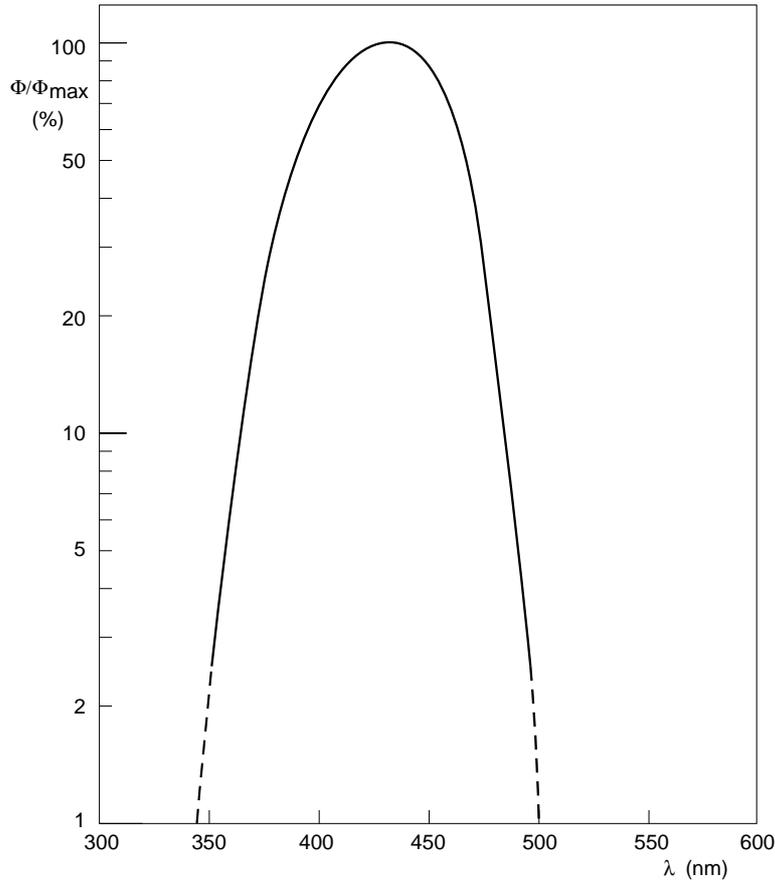


Fig.2.1 Relative spectral sensitivity characteristic of 2856 K tungsten light transmitted by a Corning filter CS No. 5-58 ground to half stock thickness

2.1.2 Quantum efficiency, ρ

This characteristic, which is another way of expressing cathode sensitivity, is the ratio of the number of photoelectrons emitted, n_k , to the number of incident photons, n_p . It is usually specified for *monochromatic* light and is related to the absolute spectral sensitivity by

$$\rho = \frac{n_k}{n_p} = S_{k,\lambda} \frac{h\nu}{e} = S_{k,\lambda} \frac{hc}{\lambda e}$$

where e is electron charge, h is Planck's constant, and c is the speed of light in vacuum. With $hc/e = 1.24 \times 10^{-6} \text{ Wm/A}$

$$\rho(\%) = 124 \frac{S_{k,\lambda}(\text{mA/W})}{\lambda (\text{nm})} \quad (2.3)$$

The curves of uniform quantum efficiencies drawn in Fig.1.3 illustrate the extent to which the quantum efficiencies of standard cathodes vary with wavelength.

For light that is not monochromatic, the ratio of the number of photoelectrons emitted to the total number of incident photons (in the region where the incident light spectrum and the cathode sensitivity spectrum overlap) is called the *integral quantum efficiency*. This is often an important specification in high-energy physics applications involving scintillators and wavelength shifters (§7.4).

2.1.3 *Input system collection efficiency, η*

This is the ratio between the number of photoelectrons reaching the first dynode and the number leaving the cathode, usually expressed in per cent. It is a function of the initial velocities of the electrons and therefore varies with wavelength, decreasing at shorter wavelengths as the photon energy increases (Fig.2.2); a slight recovery is observable at wavelengths below about 200 nm for which so far no satisfactory explanation has been put forward. Like cathode sensitivity, collection efficiency is not uniform with respect to the whole cathode surface; it also varies according to the geometry of the input system. The non-uniformity is a function of wavelength and is a measurable quantity. No universally accepted methods exist for measuring absolute collection efficiency, and it is often more useful to speak about *relative collection efficiency*, i.e. efficiency relative to a well-defined situation, for example an energy-resolution measurement.

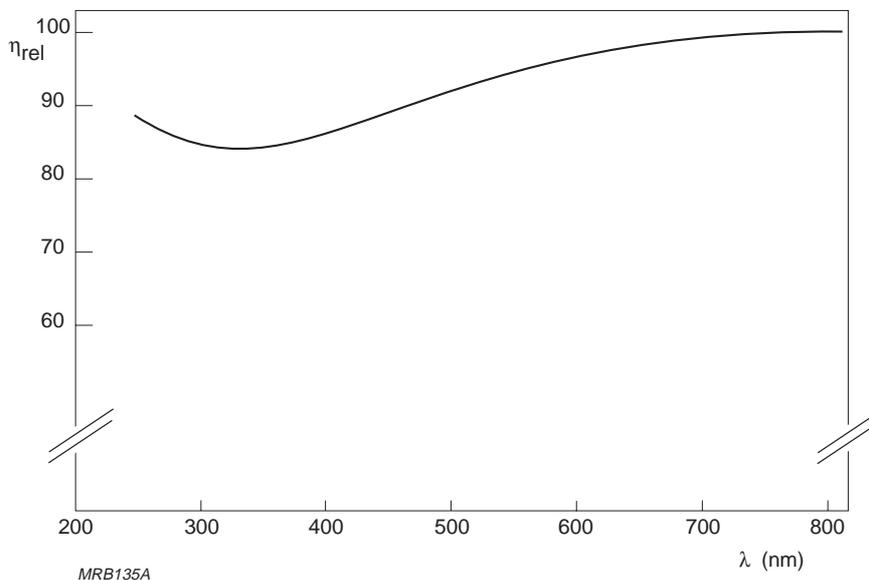


Fig.2.2 Example of relative input system collection efficiency as a function of wavelength

Collection efficiency depends on the voltage applied between the cathode and first dynode. Under optimum conditions it is possible to obtain a mean collection efficiency of more than 80%, averaged over the whole cathode surface, for wavelengths longer than 400 nm.

2.1.4 Gain, G

The gain of a photomultiplier is the ratio I_a/I_k , where I_a is the anode current due to a cathode photocurrent I_k :

$$G = \frac{I_a}{I_k} \quad (2.4)$$

For an N -stage tube in which η is the collection efficiency of the first dynode, δ_i the secondary emission coefficient of the i th dynode, and η_i the collection efficiency of the i th multiplier stage,

$$G = \eta(\delta_1\eta_1)(\delta_2\eta_2) \dots (\delta_N\eta_N) = \eta \prod_{i=1}^N \delta_i\eta_i = \eta \prod_{i=1}^N g_i \quad (2.5)$$

where g_i is the gain of the i th stage.

The gain increases rapidly as a function of the applied voltage. If the collection efficiency of all stages approaches 100%, from Eqs 1.3 and 2.5

$$G = \prod_{i=1}^N k_i V_i^\alpha \quad (2.6)$$

where k_i is a proportionality constant, V_i the interdynode voltage per stage, and α is between 0.6 and 0.8. As V_i is a fraction k'_i of the supply voltage V_{ht} , this can be written

$$G = \prod_{i=1}^N k_i (k'_i)^\alpha V_{ht}^\alpha = K V_{ht}^{N\alpha} \quad (2.7)$$

where the constant K depends on the material of the dynodes and the voltage division between them. Thus, for a 10-stage tube the gain increases as about the 7th power of the supply voltage (see Fig.2.7), doubling for each 10% voltage increase. With present-day tubes, gains of 10^6 are often obtained at supply voltage between 800 V and 1200 V.

The ratio M of the gain to the input system collection efficiency η ,

$$M = \frac{G}{\eta} = \prod_{i=1}^N g_i \quad (2.8)$$

represents the gain of the electron multiplier alone; that is, the number of electrons delivered to the anode for each electron received at the first dynode.

Gain varies somewhat with temperature, incident-light wavelength, and mean anode current.

2.1.5 Anode sensitivity, S_a

This is the ratio of the anode current I_a to the incident flux Φ . Like photocathode sensitivity, it can be specified in radiometric or photometric terms.

Anode radiant sensitivity

$$S_a(\text{A/W}) = \frac{I_a(\text{A})}{\Phi_e(\text{W})} \quad (2.9a)$$

where Φ_e is the total radiant flux measured over the whole spectrum.

Anode spectral sensitivity

$$S_{a,\lambda}(\text{A/W}) = \lim_{d\lambda \rightarrow 0} \frac{dI_a(\text{A})}{d\Phi_e(\text{W})} \quad (2.9b)$$

where $d\Phi_e$ is the flux increment in a wavelength increment $d\lambda$.

Anode luminous sensitivity (e.g. to white or blue light),

$$S_a(\text{A/lm}) = \frac{I_a(\text{A})}{\Phi_v(\text{lm})} \quad (2.10)$$

where Φ_v is the luminous flux measured in the spectrum of interest. In the data sheets anode blue sensitivity is expressed in A/lmF (where F stands for filtered).

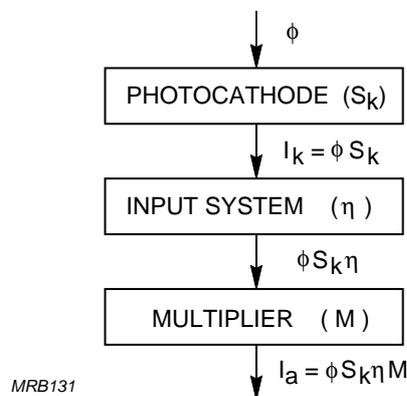


Fig.2.3 Photomultiplier sensitivity characteristics

Anode sensitivity can also be expressed in terms of cathode sensitivity, input system collection efficiency, and multiplier gain (Fig.2.3):

$$S_a = \frac{I_a}{\Phi} = S_k \eta M \quad (2.11)$$

or, from Eq.2.8,

$$S_a = GS_k \quad (2.12)$$

The sensitivity characteristic measured at the anode does not correspond exactly to that of the cathode. The differences are due mainly to the variations of η as a function of wavelength, especially in the ultraviolet.

2.1.6 Single-electron spectrum (SES)

When a photomultiplier is used to detect very weak signals such that the pulses from individual photoelectrons are well separated in time, it is often advantageous to count those pulses. Because of the nature of the secondary-emission process (§ 3.2.5), the single-electron pulses show very large amplitude fluctuations.

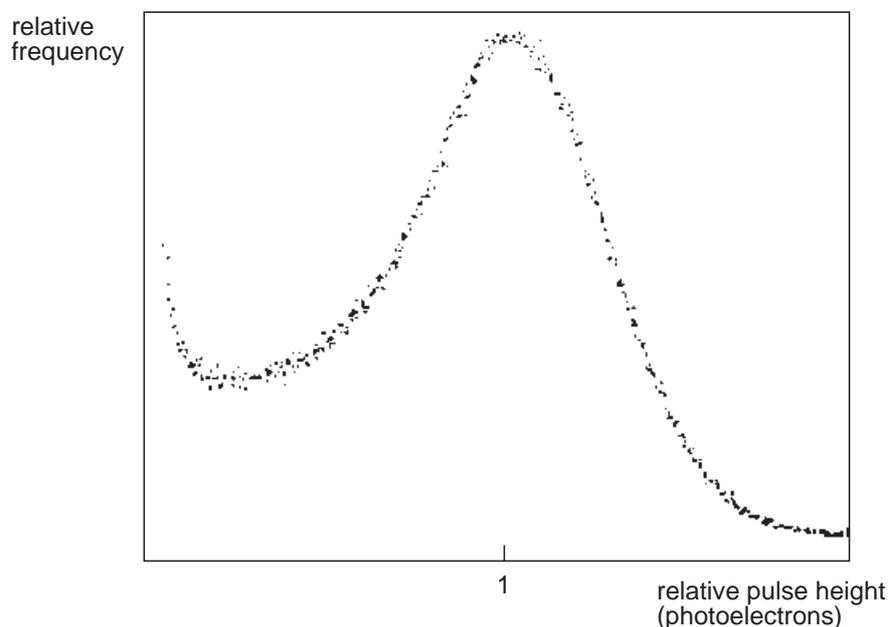


Fig.2.4 Typical single-electron spectrum. Resolution 67% FWHM.
Peak-to-valley ratio 2.8:1

The corresponding amplitude distribution (or *single-electron spectrum*) can be observed with a multi-channel pulse-height analyser (Fig.2.4) and can be described by several parameters:

- the amplitude corresponding to the centroid of the spectrum. This is, of course, the *mean* amplitude; if the secondary emission were free from any fluctuation, all pulses would have that amplitude
- the peak-to-valley (P/V) ratio. With a secondary-emission coefficient for the first dynode of at least 6 to 8, the SES will show a peak and a P/V ratio can be estimated. This can be used for monitoring the actual gain of the photomultiplier. The P/V ratio is often given in the data sheets

- single-electron resolution. If the first dynode has a high secondary-emission coefficient (above 12), the P/V ratio may exceed 2 and it is possible to estimate a single-electron resolution. This is defined as the FWHM of the single-electron peak divided by the position of the peak on the multichannel analyser (expressed as a percentage). Events producing 2 simultaneous photoelectrons will give a second peak with twice the pulse height
- relative variance for the distribution.

Note: there is always a large proportion (10 to 20%) of very small pulses (below 1/3 of the single-electron peak position). They are real *signal* pulses caused by photoelectrons being inelastically back-scattered by the first dynode.

2.2 Time characteristics

2.2.1 Response pulse width, t_w

This is defined as the full width at half maximum (FWHM) of the anode current pulse delivered in response to a delta-function light pulse (Fig.2.5). Although it is not practicable to generate true delta-function light pulses, it is practicable to generate light pulses whose FWHM is much less than that to be measured. If $t_{w,l}$ is the FWHM of such a light pulse, and $t_{w,a}$ the FWHM of the corresponding anode current pulse, the response pulse width t_w is given by

$$t_w = \sqrt{t_{w,a}^2 - t_{w,l}^2} \quad (2.13)$$

provided both $t_{w,a}$ and $t_{w,l}$ are approximately gaussian. If t_w is more than a few nanoseconds, it can be satisfactorily measured using light pulses of 1 ns FWHM.

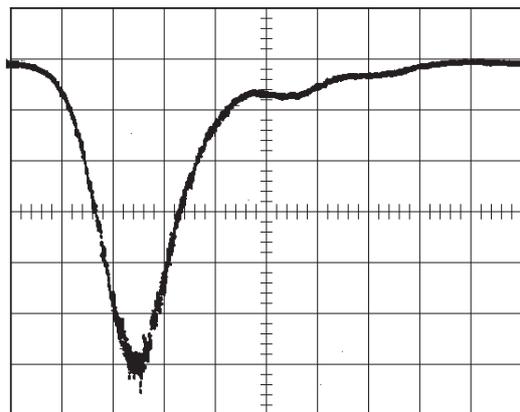


Fig.2.5 Anode current pulse response of a fast-response photomultiplier: vertical scale 20 mA/div., horizontal scale 2 ns/div

Modern lasers are capable of generating light pulses with FWHM < 100 ps. Such lasers are expensive, however, and are usually only found in advanced measuring systems. Today, new blue-emitting fast (1 - 2 ns) LEDs are commonly used.

The response pulse width depends on the illumination level and is minimum when this level is so low that the probability of more than one photoelectron being emitted per light pulse is very small (§4.4.2); this is called the single-electron response (SER). In fast photomultipliers, t_w increases by a few tenths of a nanosecond when the amplitude of the anode pulse increases from a few milliamperes (single-electron operation) to, say, a hundred milliamperes (multi-electron operation).

2.2.2 Rise time, t_r

Step-response rise time is properly defined as the time required for the anode current to increase from 10% to 90% of its final value in response to a unit step input $\epsilon(t)$. Measured under these conditions, the rise time $t_{r,\epsilon}$ (Fig.2.6(b)) approximately equals the response pulse width t_w .

However, due to the difficulty of producing unit steps of light, the rise time is by convention defined as the 10% to 90% rise time of the anode current pulse in response to a light pulse that approximates a delta function (Fig.2.6(c)). This is designated $t_{r,\delta}$, or more simply t_r . It varies from about 1 ns for photomultipliers with linear-focusing dynodes to about 20 ns for those with venetian-blind dynodes or box first dynodes.

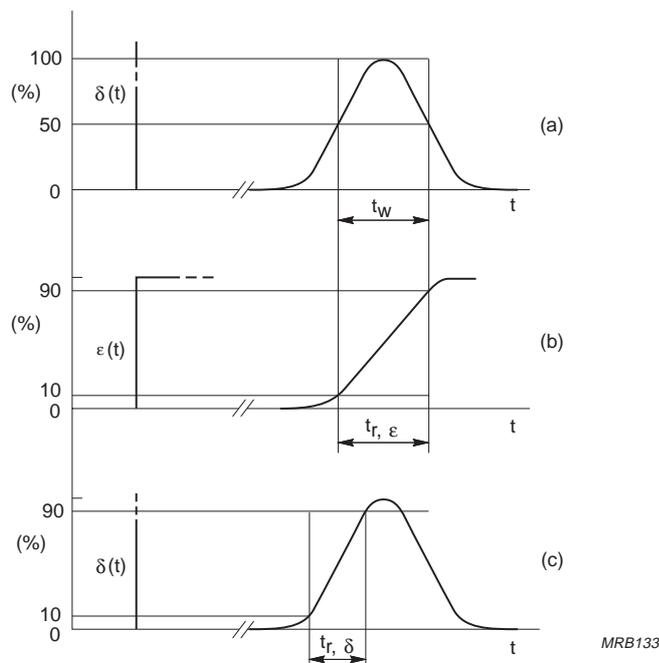


Fig.2.6 Delta-function input response-pulse width t_w , step-function input response rise time $t_{r,\epsilon}$, and delta-function input response rise time $t_{r,\delta}$

2.2.3 Transit-time differences

The interval between the arrival of a light pulse at the cathode and that of the corresponding current pulse at the anode is called the transit time. Its mean value \bar{t}_t , evaluated over a statistically large number of pulses, varies as $1/\sqrt{V_{ht}}$ and is usually of the order of several tens of nanoseconds.

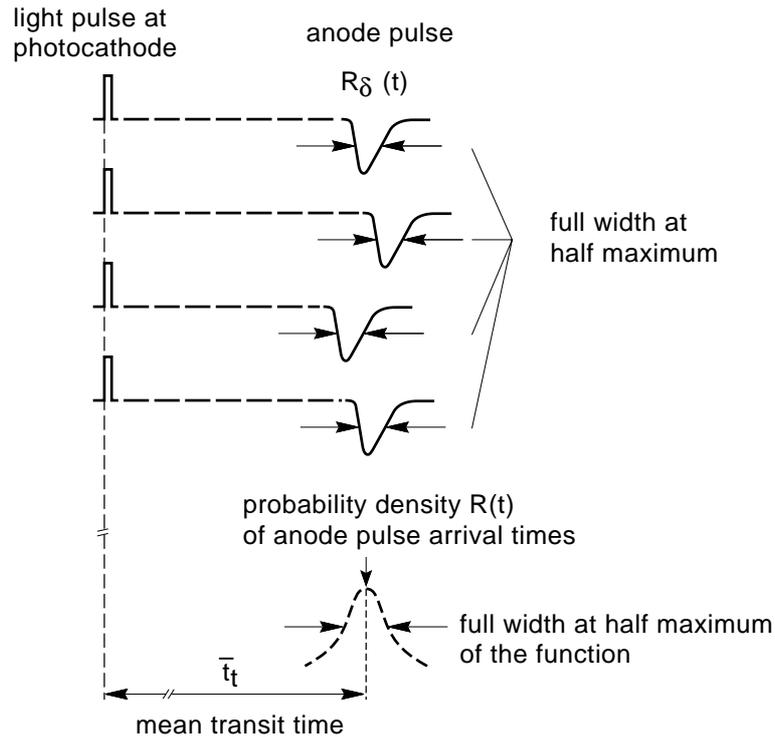
In general, the mean transit time differs according to where on its surface the cathode is illuminated. When measured with reference to one point of illumination at the centre of the cathode and another at the edge, the corresponding transit-time difference is called the *centre-edge difference* (often referred to as Δt_{CE}).

2.2.4 Transit-time spread, time resolution

Transit time spread ('jitter') is the transit-time fluctuation observed when identical light pulses strike the same part of the cathode (Fig.2.7). The time resolution of a tube, R_t , is defined as the FWHM of the probability distribution of the fluctuations. It is practically proportional to $1/\sqrt{n_{k,i}}$, where $n_{k,i}$ is the number of photoelectrons per pulse. Like the transit-time differences, the time resolution depends on the size and location of the illuminated part of the cathode; it also depends on the voltage applied to the electron-optical input system and on the spectral character of the illumination.

In the data sheets, transit-time spread is specified in terms of the standard deviation σ of the probability distribution of the transit-time fluctuations. It is a worst-case value (as defined by IEC) based on single-photoelectron pulses originating from points distributed over the whole surface of the cathode ('open cathode'), i.e. including the centre-edge difference.

How the time characteristics of a photomultiplier depend on the different parts of the tube, such as the input system, the multiplier, and the anode collection space, is dealt with in Chapter 4.



MRB132

Fig.2.7 Response-pulse jitter due to transit time fluctuations

2.3 Dark current

Even in total darkness, a current can still be measured at the anode of a photomultiplier; its causes include thermionic emission, field effects, and leakage currents. It depends particularly on the composition of the cathode, and, throughout the usual range of supply voltages, varies ideally as the gain (Fig.2.8).

Observation of the dark current using a wide-band oscilloscope reveals pulses of widely varying amplitudes (Fig.2.9). These are called dark current pulses, or dark noise pulses. Integrating the dark pulses yields one component of the dark current.

The amplitude distribution, called the dark current pulse spectrum, varies according to the type of dynode (Fig.2.10), but it is often still broader than the SES, and the shape may vary considerably between different samples of the same tube.

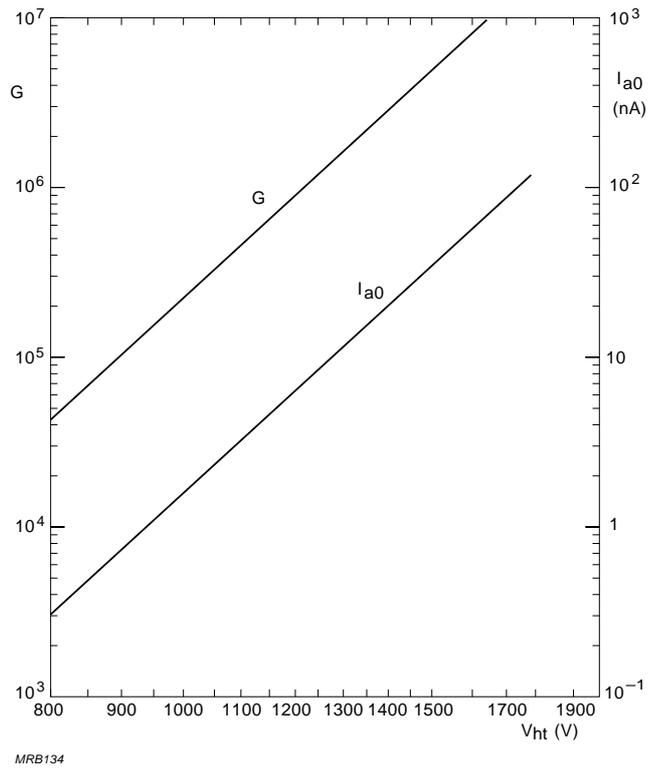


Fig.2.8 Examples of photomultiplier gain and anode dark-current variation as functions of applied voltage

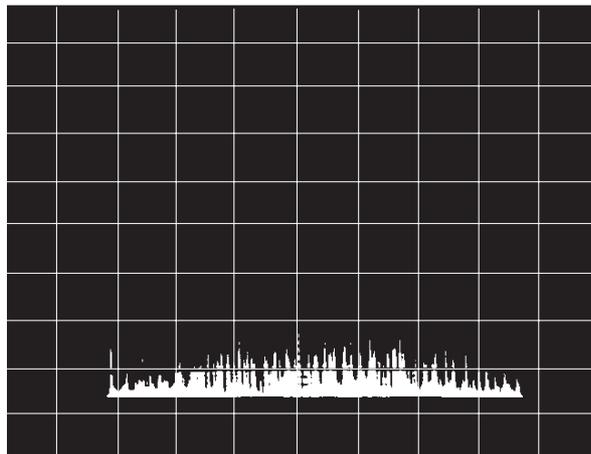


Fig.2.9 Example of dark current pulses on a wideband oscilloscope. Vertical axis, amplitude; horizontal axis, time

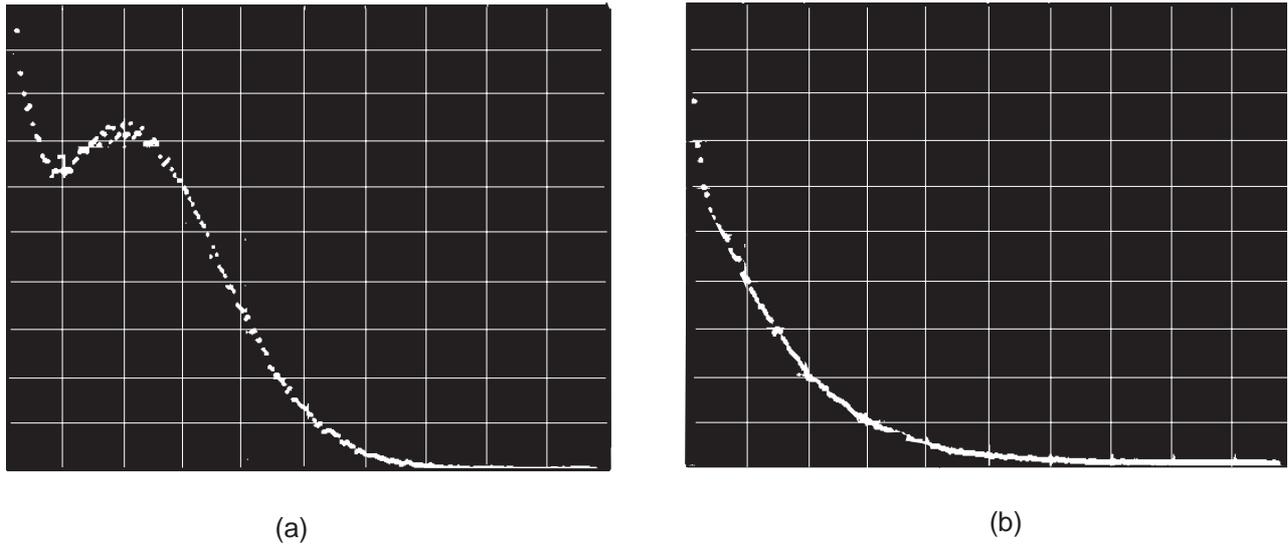


Fig.2.10 Examples of dark-current pulse spectra (a) with linear-focusing CuBe dynodes, and (b) with venetian blind dynodes. Vertical axes, pulse frequency; horizontal axes, pulse amplitude

If the dark pulses were mainly due to single-electron thermionic emission from the cathode, their spectrum would be typical a single-photoelectron spectrum (§3.2.5). In fact, this is never the case so thresholds are always used to suppress unwanted pulses.

A commonly specified photomultiplier characteristic is the mean dark pulse rate, measured at a specific gain and temperature and with reference to a specific threshold. A threshold of about 1/4 of the mean peak amplitude is often used.

2.4 Energy resolution

An important application of photomultipliers is scintillation counting (Chapter 6), one objective of which is to measure the energy of nuclear radiations. The photomultiplier is coupled to a scintillator which emits light pulses in response to α -, β - or γ - radiation, the average quantity of light per pulse being proportional to the radiation energy dissipated in the scintillator. The mean anode charge per pulse is proportional to the quantity of light and, hence, also to the energy dissipated.

This proportionality holds only on average, however. For equal quantities of energy dissipated in the scintillator, the quantity of light that actually reaches the photomultiplier cathode fluctuates from pulse to pulse, as does the generation of photoelectrons, the collection efficiency, and the gain of the photomultiplier. Taking all this into account, the energy resolution of the scintillator/photomultiplier combination is given by

$$R_e = \frac{\Delta q_{a,s}}{\bar{q}_{a,s}} \quad (2.14)$$

where $\Delta q_{a,s}$ is the FWHM of the probability distribution of the anode charge fluctuations and $\bar{q}_{a,s}$ is the mean anode charge per pulse. The energy resolution (also called *pulse height resolution* or PHR) is a measure of a scintillation counter's ability to discriminate between closely similar pulses; it varies approximately as $1/\sqrt{E}$, where E is the radiation energy dissipated in the scintillator. (As is shown in Chapter 3, energy resolution is closely related to the noise which sets an ultimate limit to the accuracy of measurements made with scintillation counters.)

Energy resolution is a characteristic of a scintillation counter as a whole, not of a photomultiplier by itself; the contributions of the scintillator and photomultiplier are not independent of each other and cannot be treated separately. In photomultiplier data sheets, therefore, when energy resolution values are given, the type and dimensions of the scintillator with which they were measured are also specified, as is the type of radiation. Sometimes the energy resolution is measured against more than one type of radiation; for example, the 662 keV γ -radiation of ^{137}Cs , the 122 keV γ -radiation of ^{57}Co , and the 5.9 keV X-radiation of ^{55}Fe . Note that the lower the gamma energy, the better is the test for judging the photomultiplier energy resolution characteristics. At higher energies (662 keV), the test is dominated by the scintillator quality and gives little information about the photomultiplier contribution to the energy resolution.

2.5 Measurement of the characteristics

2.5.1 Cathode sensitivity

The cathode luminous sensitivity for white light is measured using a 2856 K tungsten-filament lamp calibrated for luminous intensity in a specific direction. With the cathode at a distance d from the lamp (Fig.2.11), the incident flux Φ_v is calculated from

$$\Phi_v = \frac{I_v A}{d^2}$$

where I_v is the calibrated intensity of the lamp and A the area of the cathode. After subtracting the cathode dark current from the measured cathode current, the luminous sensitivity S_k is calculated from Eq.2.2.

To ensure that the electric field is the same as in normal operation, the input system electrodes and at least the first two dynodes are operated at their normal voltages. The other electrodes are strapped so that the gain is low and a measurable cathode current can flow without causing excessive anode current. To ensure that the measured current

is stabilized in the saturation region, the cathode-to-first-dynode voltage is set at about 100 V.

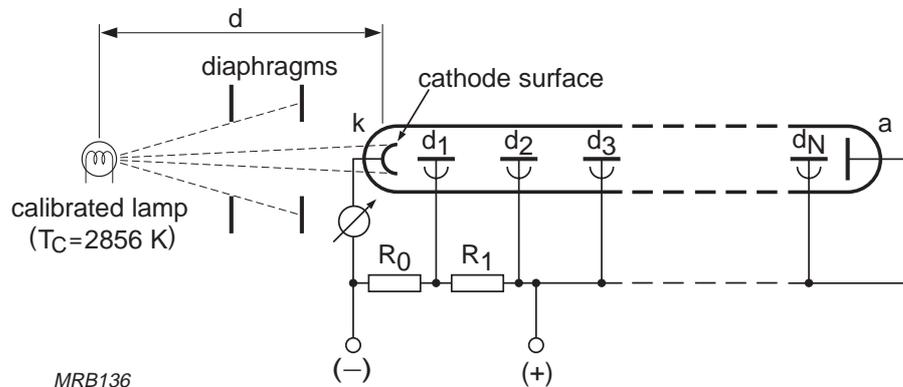


Fig.2.11 Set-up for measuring cathode sensitivity

To measure the cathode radiant sensitivity at a specific wavelength the calibrated lamp is replaced by a monochromatic source; for instance, a tungsten-filament or mercury vapour lamp combined with interference filters, or a monochromator. The radiant flux at the cathode is usually measured with a thermopile, precautions being taken to eliminate spectral leaks due to imperfections of the interference filters. The effect of such leaks can be significant because of the wide bandwidth of the thermopile and possible fast variations in flux density of the light source. To avoid problems, the interference filters may be augmented by high-pass or low-pass filters.

The measurement of radiant sensitivity is always difficult, and it is usually necessary to compare the results with those obtained by other means. For instance, the measured radiant and luminous sensitivities can be compared via conversion factors calculated from the relative spectral sensitivity characteristics of the cathode. The known emissivity characteristics of tungsten make it possible to calculate the flux transmitted through a given filter by a calibrated 2856 K lamp. Another way of cross-checking the measurement is with other photomultipliers with the same type of cathode calibrated as secondary standards.

Precautions must also be taken against the effects of cathode resistivity (§4.5.2). If the resistivity is high, the incident flux must be attenuated by calibrated neutral-density filters.

The uncertainty of radiant sensitivity measurements is seldom less than 5%.

2.5.2 Gain

Direct measurement of gain is usually practical only for low values ($G < 10^4$) because the anode current must not be allowed to exceed a certain maximum. Higher gains must be measured indirectly or in stages. Several acceptable methods are described in IEC Publication 306-4; the one generally used is described below.

First, as with the measurement of cathode sensitivity, the tube is connected with only the input system electrodes and the first two dynodes at their normal voltages, and the cathode dark current is measured. Then the tube is exposed to an accurately reproducible flux and the corresponding cathode current I_k is determined by subtracting the dark current from the measured current. Next, the tube is reconnected normally and the anode dark current is measured. The tube is then exposed again to the same flux as before, but with neutral density filters in the light path to attenuate it by a known factor F , and the corresponding anode current I_a is determined. The gain is then

$$G = \frac{I_a}{I_k} F \quad (2.16)$$

due to excitation of energy levels in the glass window and the photocathode layer. If the incident flux is accurately known, the anode sensitivity can be evaluated directly, using Eq.2.9 or 2.10.

2.5.3 *Dark current*

To measure the dark current the tube is enclosed in a chamber (e.g. as described in IEC Publication 306-1) that excludes all radiation, visible and invisible, in the sensitivity range of the cathode. The current is measured with a galvanometer in the anode circuit (Fig.2.12). If the tube were connected in normal polarity, the galvanometer would be at a high voltage with respect to ground and leakage currents to ground in the instrument could cause significant error. Therefore the tube is connected in negative polarity (anode grounded, cathode at high negative potential), although this necessitates special precautions:

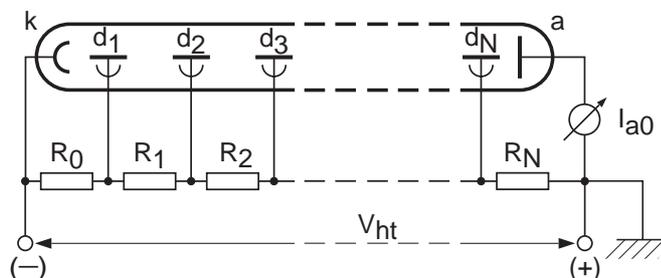
- the tube envelope must be perfectly insulated from its surroundings, particularly surrounding metal at ground potential, to prevent any leakage that might provoke dark current instability.
- considerable time must be allowed for the dark current to stabilize after the high voltage is applied (much more time than when the tube is connected in normal polarity, see Fig.3.6). The more recently the tube has been exposed to light, the more important – and the longer – the stabilization time (§3.1.5 and 3.1.6); tubes with exceptionally low dark current may require several hours (up to 24 hours). If the high voltage is altered during the measurement procedure, time must again be allowed for the dark current to stabilize at the new voltage.

Other precautions must also be taken to reduce leakage currents which may be significant in comparison with very low dark currents (less than a nanoampere):

- the socket and base of the tube must be perfectly clean and free of any trace of moisture; leakage currents on the socket can be measured by applying the high

voltage with the socket empty. Likewise, all connections must be perfectly clean, with no trace of grease or moisture.

- the insulation of the anode lead must be of very high quality (e.g. teflon, diallylphthalate). To minimize piezoelectric effects which could superimpose spurious currents on the dark current, a low capacitance lead should be used.

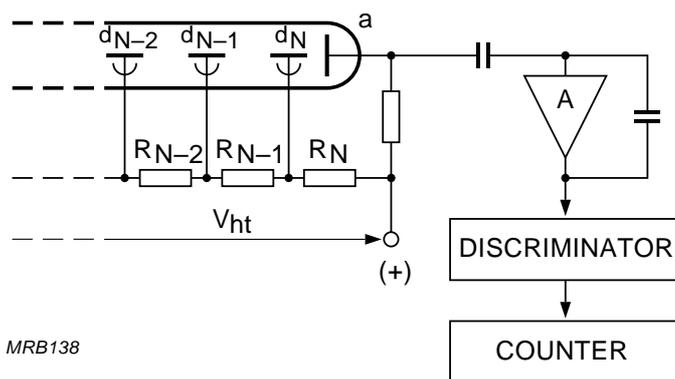


MRB137

Fig.2.12 Dark-current measurement; the tube is connected in negative polarity with (+) supply terminal grounded.

2.5.4 Dark pulse rate

Figure 2.13 shows a set-up for measuring the dark pulse rate; the tube can be connected in positive polarity, because the leakage currents, being DC, are blocked by the capacitor. In positive polarity the dark current stabilizes quickly. It is still advisable, however, to keep the tube in total darkness for several hours (for example overnight) before the measurement, especially if it has recently been exposed to light.



MRB138

Fig.2.13 Dark-pulse-rate measurement

The dark pulses are integrated in a charge-sensitive amplifier and converted into voltage pulses that are fed to a discriminator, the threshold of which is set so that only pulses that correspond to a charge exceeding a certain level will be counted. In the spectra of Fig.2.10, for example, the chosen threshold might be a quarter of the amplitude

corresponding to the centre of gravity of the single-electron distribution; that is, a quarter of the mean anode-pulse amplitude that would result from a single photoelectron (popularly referred to as ‘a ¼ photoelectron’).

By applying a voltage pulse $U(t)$ to a capacitor C at the input, the discriminator can be calibrated from the relation

$$q(t) = C \int_0^t U(t) dt$$

2.6 Interpretation of data

The information presented in photomultiplier data sheets is ranged under the following subdivisions and headings.

General description. As an aid to making a preliminary choice, this gives general information about size, construction, photocathode, and intended applications.

General characteristics. Details of window, photocathode and electron optical input multiplier system.

Recommended voltage dividers. Divider networks for alternative voltage distributions (§5.2).

Output characteristics. Gain or anode sensitivity, dark current, time characteristics, energy resolution, behaviour in a magnetic field, and linearity and stability parameters. In some cases these are given for alternative voltage distributions to enable the user to optimize specific performance aspects for an intended application.

The qualification *typ.* (typical) preceding a value means that it is derived from measurements on a representative sample and corresponds to the 50% point (median value) of the cumulative relative frequency curve for the sample. Values marked *typ.* may vary slightly from sample to sample; such values are always accompanied by an upper or lower limit defining the acceptability limit for the tube. Measured actual values for each tube are entered in an individual characteristics sheet which accompanies the tube and usually gives:

- the cathode luminous, CB filtered or radiant sensitivity
- the supply voltage required for obtaining a specified gain or anode sensitivity
- the dark current at that value of gain or anode sensitivity. The sign \approx or the qualification *approx.* preceding a value means that the value is determined by regular sampling from production; for example, certain time characteristics that vary very little from tube to tube of a given type.

Limiting values (Absolute maximum rating system). These are limits for voltages, gains currents, and temperatures that the user must observe to avoid damaging the tube. Where stated limits for different operating parameters cannot be applied simultaneously, the most restrictive should be observed. For example, if the maximum voltage were applied to each stage of a photomultiplier, the maximum gain for the tube as a whole would be exceeded; thus the latter takes precedence.

Notes. These refer to the output characteristics and specify either the definitions of certain terms or the conditions for measuring certain characteristics.

Performance curves. These include spectral sensitivity characteristics and such other characteristics as gain, anode sensitivity and dark current, as functions of supply voltage.

Mechanical data. Dimensions, mass, base designations, and base-pin identification.

Accessories. Type numbers of sockets etc.

The Photonis photomultiplier catalogue contains an introductory section which further defines the terms used in the data and describes generally applicable methods and conditions for measuring photomultiplier characteristics. It also discusses considerations relating to such factors as noise, linearity and stability, and includes operating notes regarding power supply, pulsed and continuous operation, and precautions to be taken to ensure accuracy, minimize dark current and maximize useful life. Individual data sheets are also available for each type.

CHAPTER 3

DARK CURRENT AND NOISE

3.1 Dark current

- 3.1.1 Leakage currents
- 3.1.2 Thermionic emission
- 3.1.3 Field emission
- 3.1.4 Background radiation
- 3.1.5 Exposure
- 3.1.6 High-voltage polarity

3.2 Statistical nature of noise

- 3.2.1 Photon noise
- 3.2.2 Cathode current fluctuations
- 3.2.3 Noise spectrum
- 3.2.4 Noise in scintillation detectors
- 3.2.5 Noise contribution of the electron multiplier
- 3.2.6 Johnson noise
- 3.2.7 Scintillation detection

3.3 Equivalent noise input and minimum discernible signal

- 3.3.1 Definitions
- 3.3.2 Minimum value of noise equivalent power
- 3.3.3 Effect of bandwidth
- 3.3.4 Measurement of noise equivalent power

Appendix Noise statistics and bandwidth

- A3.1 Practical scintillation spectra
- A3.2 Noise equivalent bandwidth

DARK CURRENT AND NOISE

This chapter discusses disturbances that interfere with the current or voltage to be measured and limits the accuracy of the measurement. This includes principally dark current and noise.

3.1 Dark current

The dark current is not, strictly speaking, a noise; however, noise that is associated with it does impose a limitation on the detection of very low energy radiation.

The current that flows in the anode circuit when voltage is applied to a photomultiplier in total darkness has two components:

- a continuous one due to leakage on glass and insulation surfaces,
- an intermittent one, consisting of pulses of a few nanoseconds duration.

The effect of the various causes of dark current varies according to the operating and environmental conditions (applied voltage, gain, temperature, humidity etc.), and also according to the tube's history (past storage and illumination conditions, etc.). Some of the causes are merely temporary in their effect, in which case the dark current eventually settles down to a stable level. Others are permanent.

The permanent causes of dark current (i.e. those that are independent of the history of the tube) are mainly:

- leakage currents
- thermionic emission
- field emission
- background radiation.

3.1.1 Leakage currents

These are the sole cause of the continuous component of the dark current and are due to the surface conductivity of the electrode supports, envelope, base, and socket. Surface conductivity on the inside of the tube is affected by the alkali metals used, and on the outside by agents such as dust, moisture and grease.

The dark current component due to leakage currents varies roughly linearly with the high voltage applied (that is, much less markedly than the gain, which varies exponentially at a high power of the voltage). It is therefore the predominant component when the tube operates at low gain. It is also the predominant component at low temperatures, where thermionic emission is less significant.

3.1.2 Thermionic emission

One of the main causes of the pulse component of the dark current is *thermionic emission* from the photocathode. This emission obeys Richardson's law

$$J = AT^2 \exp(-W_{th}/kT) \quad (3.1)$$

in which J represents current density, A is a constant, T is absolute temperature, and k is the Boltzmann constant. The term W_{th} is the thermionic work function of the photocathode material; for semiconductors, it is less than the photoemission threshold (§A1.1.3), being about 1 to 2 eV for cathodes sensitive to visible light.

Although the thermionic work function and the photoemission threshold are separate quantities, within a given family of tubes, there appears to be a statistical correlation that makes high red and infrared sensitivity incompatible with low dark current.

At room temperature, the thermionic emission of photocathodes with maximum sensitivity in the range 300 to 500 nm is between 10 and 1000 electrons/cm²s. It increases, however, as the sensitivity extends towards the long wavelengths (lower electron affinity) and with an S1(C) photocathode can be as high as a few million electrons/cm²s. Thermionic emission also occurs at the dynodes of the multiplier.

As Eq.3.1 shows, thermionic emission decreases rapidly as temperature decreases. Figure 3.1 shows the variation in the number of dark current pulses per second as a function of temperature for bialkline (SbKCs) cathodes and trialkaline cathodes with extended red sensitivity (SbNa₂KCs). At normal temperatures, thermionic emission is the predominating cause of the dark current, at least at normal supply voltages. At low temperatures, it becomes negligible compared with other causes, and the dark pulse rate tends towards a plateau as the temperature decreases.

Dark pulses due to thermionic emission are mainly of the single-electron type (§2.3), and those originating in the electron multiplier are amplified less than those from the cathode. The amplitude distribution of the pulses depends largely on the multiplier design (Fig.2.9).

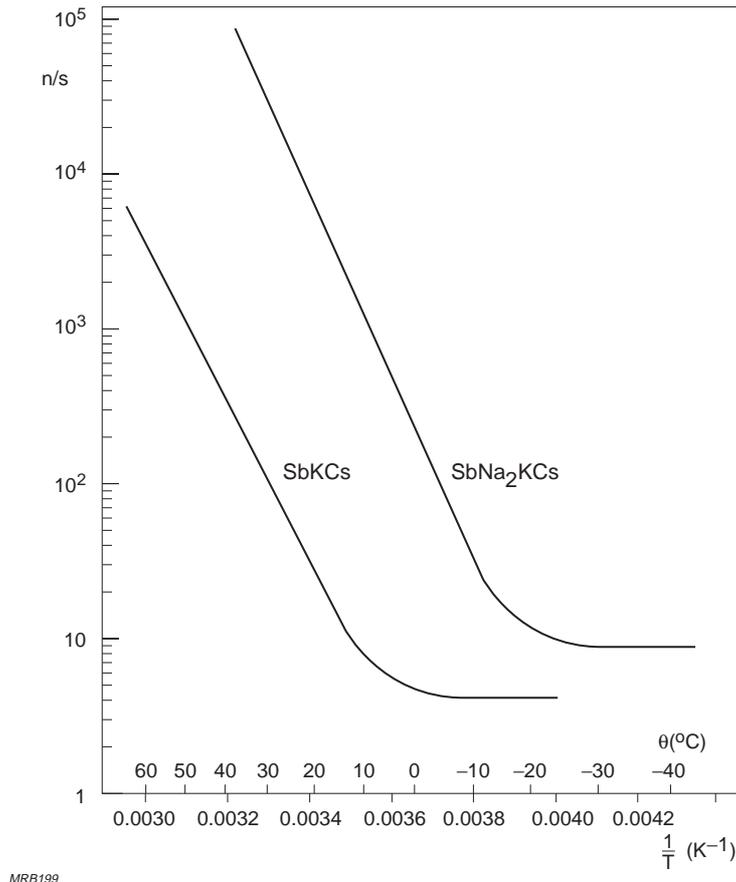


Fig.3.1 Number of dark pulses per second as a function of temperature, for SbKCs and SbNa₂KCs photocathodes

3.1.3 Field emission

Although the electric fields in a photomultiplier are fairly low, there is some electron emission due to field effect (cold emission) because of inevitable roughness of the electrodes; this is aggravated by the adsorption of alkali metals (mainly caesium) at the electrode surfaces, which considerably reduces their electron affinity. Electrons emitted by field effect bombard the envelope glass and other surfaces causing emission of photons which can reach the photocathode.

The dark pulse rate due to field emission does not depend much on temperature. It does depend on the applied voltage, however, and increases faster than the gain, which is one of the principal factors that sets a practical limit to gain. Figure 3.2 shows the three ranges of supply voltage in which each of the three causes of dark current, so far discussed, predominates.

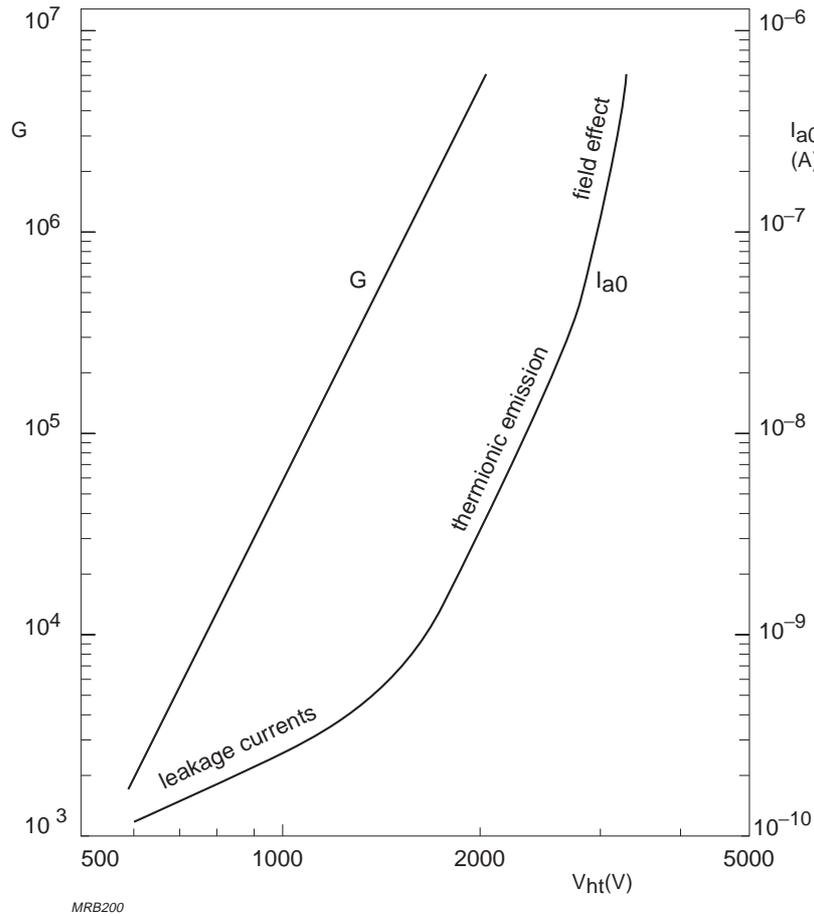


Fig.3.2 Major causes of dark current versus supply voltage

3.1.4 Background radiation

Background radiation, including that due to the materials of the tube (e.g. ^{40}K), is another cause of dark pulses. High energy charged particles (e.g. cosmic rays) can give rise to Cherenkov radiation in the tube window, which in turn causes photoemission. Cherenkov radiation can generate several photons at a time, so the dark pulses it causes (multi-electron noise) are often of high amplitude.

3.1.5 Exposure

Of the many *temporary* causes of dark current, the two most often encountered are previous exposure to light and negative-polarity connection of the photocathode.

Exposure to normal light, even when no voltage is applied, considerably increases the subsequent dark current, owing to excitation of the photocathode itself and the glass of the envelope. The dark current increase depends on the wavelength, the incident flux, and the duration of exposure. Figures 3.3 to 3.5 show how the dark current of an S13 photocathode (SbCs_3 on fused silica) stabilizes after different conditions of exposure. (In the graphs, the number of electrons per second constituting the dark current during the stabilization period has been normalized with respect to the steady-state dark current.) After prolonged exposure to sunlight, the time required to stabilize the dark current may be as long as 48 hours. Hence the necessity of guarding photomultipliers against exposure to ambient light and, if possible, storing them in the dark when they are not in use.

A similar increase of dark current occurs if a photomultiplier in operation is accidentally subjected to a brief, intense flash of 'UV' light.

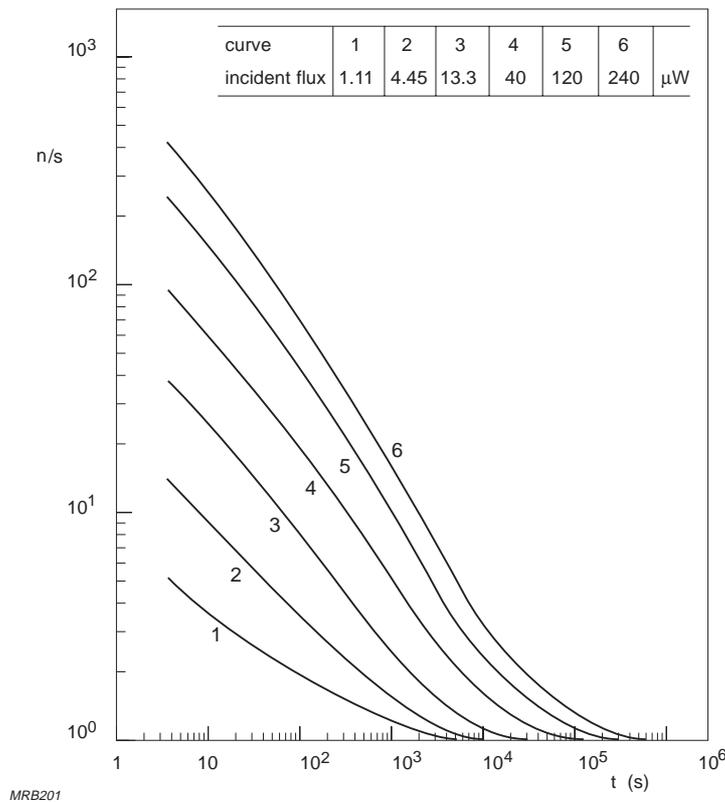


Fig.3.3 Dark-current decay of an S13 cathode at 0 °C following exposure for 100 s at $\lambda = 366$ nm, with incident flux as parameter. Vertical scale, relative number of electrons per second

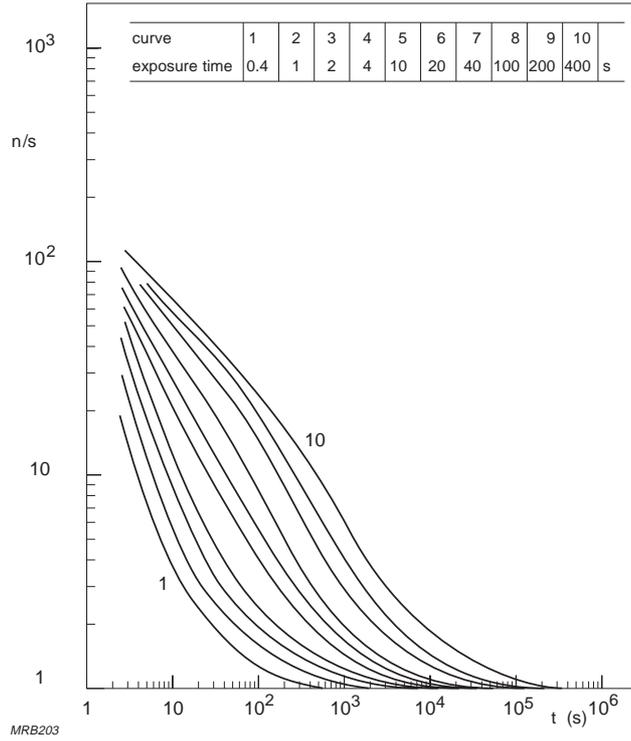


Fig.3.4 Dark-current decay of an S13 cathode at 0 °C following exposure to an incident flux of 120 μW at $\lambda = 366 \text{ nm}$, with exposure time as parameter. Vertical scale, relative number of electrons per second

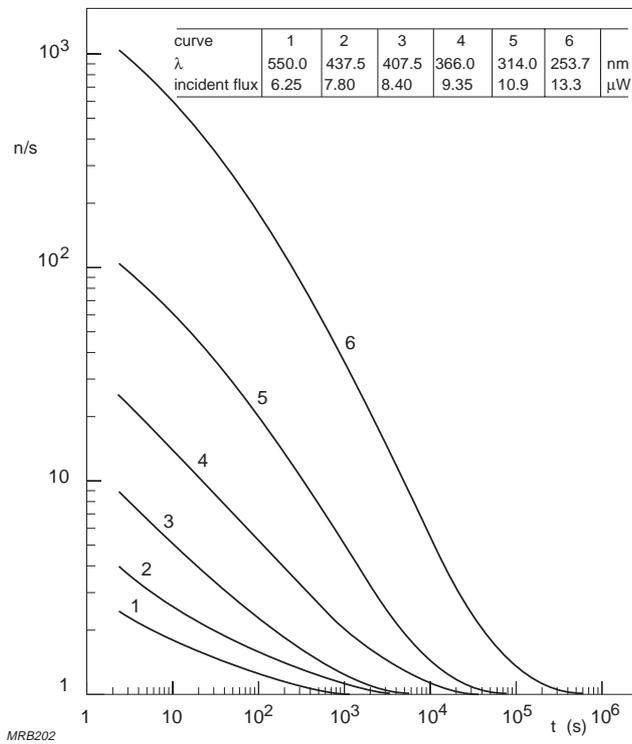


Fig.3.5 Dark-current decay of an S13 cathode at 0 °C following exposure for 100 s to equal numbers of photons, with wavelength as parameter ($\phi \propto 1/\lambda$). Vertical scale, relative number of electrons per second

3.1.6 High-voltage polarity

In some applications it is necessary to ground the anode and apply a high negative potential to the cathode (§5.1). When this is done the dark current immediately assumes a value much higher than normal (curve 1, Fig.3.6) and may take more than half an hour to settle back. The lower the normal dark current is, the more pronounced the rise and the longer the settling time.

If a tube connected in negative polarity is not properly insulated from its surroundings, insulation breakdown may occur between the envelope and earth; this generates a high, unstable, dark current and can quickly destroy the tube by electrolysis of the glass. The risk can be guarded against by coating the wall of the tube with conductive paint, taken to photocathode potential through a protective resistor of about 10 M Ω , and/or enclosing it in an adequate thickness of insulation (§5.1).

When the photocathode is grounded and the anode positive, the dark current stabilizes quickly. This polarity should therefore be used whenever possible.

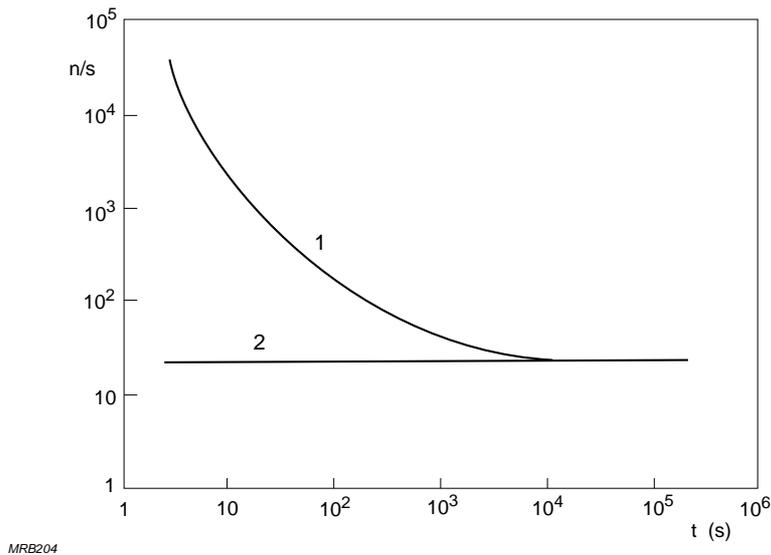


Fig.3.6 Dark-current behaviour following application of high voltage in 1) negative, 2) positive polarity. Vertical scale, number of dark pulses per second

3.2 Statistical nature of noise

It is important to recognize the irreducible nature of noise, which always accompanies the signals to be measured and cannot be cancelled or compensated. All the causes of noise encountered in photomultipliers have a common physical origin, namely the spontaneous fluctuation of currents and voltages due mainly to the discontinuous nature of radiation, electricity and matter. Noise is, therefore, closely related to the

statistical nature of photoemission and secondary emission and is inherent in the signal. Consider the following example.

A photomultiplier with a cathode sensitivity $S_k = 10 \text{ mA/W}$ and operated at a gain $G = 10^6$ is exposed to a continuous flux $\Phi_e = 10^{-9} \text{ W}$. The resulting anode voltage across a load resistance $R_L = 5 \text{ k}\Omega$ is $V_a = \Phi_e S_k G R_L = 50 \text{ mV}$. However, when this is applied to an oscilloscope with a stray capacitance of 50 pF (equivalent noise band $B_N = 1/4R_L C = 1 \text{ MHz}$) a peak-to-peak fluctuation of about $\pm 20 \text{ mV}$ is observed, corresponding to an anode current fluctuation of $\pm 4 \text{ }\mu\text{A}$. That this fluctuation is noise inherent in the signal is evidenced by the fact that it disappears completely when the incident flux is removed. In this example, the fluctuation is relatively large because the number of photoelectrons emitted in a single period, corresponding to the reciprocal of the noise band ($1/B_N = 1 \text{ }\mu\text{s}$), is small and fluctuates considerably from one period to the next.

If the incident flux is not continuous but pulsed, the resulting anode current fluctuates from pulse to pulse. The amplitude of the fluctuation determines the energy resolution of the photomultiplier; it is of the same statistical nature as the noise and yields to the same analytical procedures.

3.2.1 Photon noise

Photon emission is a random process, the number of photons emitted during like intervals being subject to a statistical distribution. To begin with we shall assume that fluctuations in the number of photons striking the cathode and the number of photoelectrons emitted both follow a *Poisson distribution*.

Consider a photocathode constantly illuminated by a source (e.g. a tungsten filament lamp) from which photons are emitted independently of each other. Assume that N_p photons are received by the photocathode during a fairly long period T (Fig.3.7). We can divide this period into a large number of intervals τ . A photon emitted during the period T has a probability $p = \tau/T$ of being received during an interval τ and a probability $(1 - p)$ of being received during the complementary interval $(T - \tau)$.

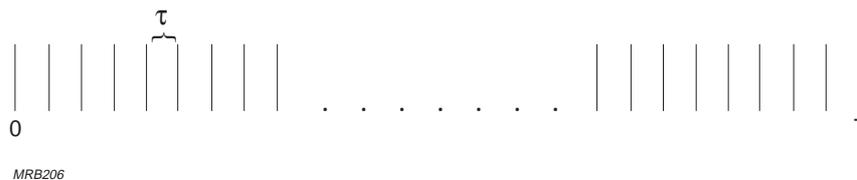


Fig.3.7 The probability of a photon arriving during the interval τ is $p = \tau/T$

All photons emitted during the period T have the same probability p of being received during an interval τ . If the number of photons reaching the cathode during the period T is N_p , and the number during the interval τ is $n_{p,\tau}$, then the number during the complementary interval $(T - \tau)$ is $(N_p - n_{p,\tau})$. Thus, the possible number of combinations of N_p photons taken $n_{p,\tau}$ at a time during the period T is

$$\binom{N_p}{n_{p,\tau}} = \frac{N_p!}{n_{p,\tau}!(N_p - n_{p,\tau})!}$$

and the probability $P(n_{p,\tau})$ of obtaining $n_{p,\tau}$ photons during the interval τ is

$$P(n_{p,\tau}) = \binom{N_p}{n_{p,\tau}} \left(\frac{\tau}{T}\right)^{n_{p,\tau}} \left(1 - \frac{\tau}{T}\right)^{N_p - n_{p,\tau}} \quad (3.2)$$

This is the *binomial distribution*, with mean value

$$\bar{n}_{p,\tau} = N_p \frac{\tau}{T} = \bar{n}_p \tau$$

and standard deviation

$$\sigma = \sqrt{N_p \frac{\tau}{T} \left(1 - \frac{\tau}{T}\right)} = \sqrt{\bar{n}_p \tau \left(1 - \frac{\tau}{T}\right)}$$

where $\bar{n}_p = N_p/T$ is the average number of photons received per unit time.

If T is taken fairly large or, which amounts to the same, if the probability $p = \tau/T$ is very small, the binomial distribution tends towards the *Poisson distribution*:

$$P(n_{p,\tau}) = \frac{(\bar{n}_{p,\tau})^{n_{p,\tau}} \exp(-\bar{n}_{p,\tau})}{n_{p,\tau}!} \quad (3.3)$$

In this case, the mean value and the standard deviation are, respectively

$$\begin{aligned} \bar{n}_{p,\tau} &= \bar{n}_p \tau \\ \sigma_{n_{p,\tau}} &= \sqrt{\bar{n}_{p,\tau}} = \sqrt{\bar{n}_p \tau} \end{aligned} \quad (3.4)$$

Hence, the relative variance is

$$V_{n_{p,\tau}} = \frac{\sigma_{n_{p,\tau}}^2}{\bar{n}_{p,\tau}^2} = \frac{1}{\bar{n}_{p,\tau}} = \frac{1}{\bar{n}_p \tau} \quad (3.5)$$

from which it is evident that the larger the mean number of photons received during the interval τ , the smaller are the fluctuations in the number.

3.2.2 Cathode current fluctuations

Photoemission is a random process which can usually be treated on its own. Each photon that strikes the cathode has a probability ρ (quantum efficiency) of liberating an electron and a probability $(1 - \rho)$ of not liberating one. When the randomness of photon arrival is also to be taken into account, the mean value $\bar{n}_{k,\tau}$ and standard deviation σ of the number of photoelectrons emitted during an interval τ are, respectively

$$\begin{aligned} \bar{n}_{k,\tau} &= \bar{n}_p \rho \tau \quad \text{and} \\ \sigma_{n_{k,\tau}} &= \sqrt{\bar{n}_p \rho \tau} \end{aligned} \quad (3.6)$$

Hence the relative variance is

$$V_{n_{k,\tau}} = \frac{\sigma_{n_{k,\tau}}^2}{\bar{n}_{k,\tau}^2} = \frac{1}{\bar{n}_{k,\tau}} = \frac{1}{\bar{n}_p \rho \tau} \quad (3.7)$$

For the interval τ the value of the cathode current $I_{k,\tau}$ is given by

$$I_{k,\tau} = \frac{e n_{k,\tau}}{\tau}$$

where e is the electron charge. $I_{k,\tau}$ differs from the mean current over a large number of intervals, I_k , by an amount

$$i_{k,\tau} = I_{k,\tau} - I_k = \frac{e}{\tau} (n_{k,\tau} - \bar{n}_{k,\tau}) \quad (3.8)$$

the mean square value of which is

$$\overline{i_{k,\tau}^2} = \frac{e^2}{\tau^2} \sigma_{n_{k,\tau}}^2 = \frac{e^2 \bar{n}_p \rho}{\tau} = \frac{e I_k}{\tau} \quad (3.9)$$

3.2.3 Noise spectrum

When the photocathode is illuminated by a constant flux, the photocurrent $I_k(t)$ consists of a constant component I_k and a fluctuating component $i_k(t)$:

$$i_k(t) = I_k(t) - I_k$$

Here, $i_k(t)$ is a true random quantity whose mean value $\bar{i}_k(t)$ over a long period T is zero. However, there may be a certain correlation between values of $i_k(t)$ measured at different times; this is expressed by the *autocorrelation function*:

$$\gamma(s) = \overline{i_k(t) \cdot i_k(t + s)} \quad (3.10)$$

For $s = 0$, the autocorrelation function assumes a maximum equal to the mean square value $\overline{i_k^2(t)}$. As s increases, $\gamma(s)$ decreases and tends towards zero. Thus, we can characterize $i_k(t)$ by a time constant τ_0 such that $\gamma(s)$ becomes less than a specified very low value ($\gamma(s) \leq \epsilon$) for $|s| \gg \tau_0$. If τ_0 is very small compared with the observation time of the signal $i_k(t)$, which it usually is in practice, the correlation can be disregarded.

We can express the autocorrelation function $\gamma(s)$ another way by using the Fourier transform

$$\gamma(s) = \int_0^{\infty} w(f) \cos(2\pi fs) df \quad (3.11)$$

where $w(f)$ represents the spectral density or *noise spectrum* of $i_k(t)$. The inverse Fourier transform expresses the noise spectrum as a function of $\gamma(s)$:

$$w(f) = 4 \int_0^{\infty} \overline{i_k(t) \cdot i_k(t + s)} \cos(2\pi fs) ds \quad (3.12)$$

This is the Wiener-Khintchine theorem which shows that the spectral density is independent of the frequency as long as $2\pi f\tau_0 \ll 1$.

If we set s equal to zero in Eq.3.10 and 3.11, we obtain

$$\gamma(0) = \overline{i_k^2(t)} = \int_0^{\infty} w(f) df \quad (3.13)$$

which relates the spectral density to the mean square value of $i_k(t)$. The term $\overline{i_k^2(t)}$ represents the total noise power throughout the frequency spectrum, as a function of the parameters that characterize the random nature of the photon emission and the photon-electron conversion.

Consider the mean value of the fluctuating component $i_k(t)$ of the photocurrent over an interval τ that is very small compared with the observation time:

$$i_{k,\tau} = \frac{1}{\tau} \int_t^{t+\tau} i_k(t) dt$$

The quantity $i_{k,\tau}$ fluctuates randomly from one interval to another; its mean square value

$$\overline{i_{k,\tau}^2} = \frac{1}{\tau^2} \int_0^\tau \int_0^\tau i_k(t) i_k(t') dt dt'$$

is independent of time because of the stationary nature of the random variable $i_k(t)$. However, if $\tau \gg \tau_0$,

$$\overline{i_{k,\tau}^2} = \frac{w(f)}{2\tau} \quad (3.14)$$

so, by substitution from Eq.3.9,

$$w(f) = 2eI_k \quad (3.15)$$

which shows that, in the frequency range where $\tau \gg \tau_0$, the spectral density $w(f)$ is constant.

Finally, by combining Eq.3.13 and 3.15 we can write:

$$\overline{i_k^2}(t) = 2eI_k \int_0^\infty df$$

or, for a frequency interval Δf ,

$$\overline{i_k^2}(t) = 2eI_k \Delta f \quad (3.16)$$

which is the well known *Schottky* formula. Dividing by I_k^2 and substituting from Eq.3.9, gives

$$\frac{\overline{i_k^2}(t)}{I_k^2} = \frac{2e\Delta f}{I_k} = \frac{2\Delta f}{\bar{n}_p \rho} \quad (3.17)$$

the square root of which is the reciprocal of the signal-to-noise ratio:

$$\frac{N}{S} = \sqrt{\frac{2\Delta f}{\bar{n}_p \rho}} \quad (3.18a)$$

So the signal-to-noise ratio due to cathode current fluctuation under conditions of constant illumination, is

$$\frac{S}{N} = \sqrt{\frac{\bar{n}_p \rho}{2\Delta f}} \quad (3.18b)$$

in the frequency interval Δf .

3.2.4 Noise in scintillation detectors

So far we have considered only the fluctuations of a continuous photocurrent due to constant illumination of the cathode. However, we can extend the same reasoning to the photocurrent of a scintillation detector, even though the scintillation photon pulses decay exponentially, provided we choose sampling intervals long enough to include effectively the whole of a scintillation. As a first approximation we can assume that the number of photons per scintillation, $n_{p,s}$, follows a Poisson distribution; if $n_{p,s}$ is large, the Poisson distribution tends towards a gaussian distribution. The number of photoelectrons per scintillation, $n_{k,s}$, then also follows a gaussian distribution, with relative variance

$$v_{n_{k,s}} = \frac{\sigma_{n_{k,s}}^2}{\bar{n}_{k,s}^2} = \frac{1}{\bar{n}_{k,s}} \quad (3.19)$$

The probability distribution of the number of photoelectrons is usually determined on the basis of an anode pulse histogram generated by a multichannel pulse-height analyser. When the scintillations are due to monoenergetic radiation (e.g. X-or γ -rays), the histogram has a more or less well-defined peak corresponding to photoelectric absorption of the radiation in the scintillator. If the FWHM of the peak corresponds to $\Delta n_{k,s}$ electrons emitted by the cathode, the energy resolution of the photomultiplier-scintillator combination is

$$R_e = \frac{\Delta n_{k,s}}{\bar{n}_{k,s}}$$

which, for a gaussian distribution, reduces to

$$R_e = 2.36\sqrt{v_{n_{k,s}}}$$

or, from Eq.3.19,

$$R_e = \frac{2.36}{\sqrt{\bar{n}_{k,s}}} \quad (3.20)$$

Comparison of Eq.3.20 and 3.18 illustrates the close relation between energy resolution and signal-to-noise ratio.

3.2.5 Noise contribution of the electron multiplier

To take account of fluctuations in the secondary emission of the dynodes, a more extended treatment is required. Fluctuations in electron multiplication have been treated statistically by Lombard and Martin, using the method of generating functions; here, we merely summarize the main results.

First, assume that the random processes at all stages of the electron multiplier obey a Poisson distribution and occur independently of each other. The number of electrons reaching the anode during an interval τ is

$$n_{a,\tau} = n_{k,\tau} G = n_{p,\tau} \rho G \quad (3.21)$$

where G is the photomultiplier gain. The relative variance of $n_{a,\tau}$ is

$$v_{n_{a,\tau}} = \frac{\sigma_{n_{a,\tau}}^2}{\bar{n}_{a,\tau}^2} = \frac{1}{\bar{n}_p \rho \tau} (1 + v_G) \quad (3.22)$$

where v_G is the relative variance of the gain. It has been shown that when all electron-multiplier stages except the first have the same gain g ,

$$v_G = v_\eta + \frac{1}{\eta} \cdot \frac{1}{g_1} \cdot \frac{g}{g-1} \quad (3.23)$$

where g_1 is the (usually higher) gain of the first stage, and v_η is the relative variance of the collection efficiency η of the input optics. In the ideal case, when η is independent of the point on the photocathode from which the electrons originate,

$$v_\eta = \frac{1 - \eta}{\eta} \quad (3.24)$$

e.g. for $\eta = 0.95$, $v_\eta \approx 0.05$.

If we introduce the fluctuating component $i_a(t)$ of the anode current,

$$i_a(t) = I_a(t) - I_a = i_k(t) G$$

Eq.3.17 and 3.20 can be replaced by

$$\frac{\overline{i_a^2(t)}}{I_a^2} = \frac{2\Delta f}{\bar{n}_p \rho} (1 + v_G) = \frac{2e\Delta f}{I_k} (1 + v_G) \quad \text{and} \quad (3.25)$$

and

$$R_e = 2.36 \sqrt{\frac{1 + v_G}{\bar{n}_{k,s}}} \quad (3.26)$$

The Poisson distribution assumption for secondary emission is only a rough approximation. The single-electron spectrum, which reveals the probability distribution of the electron multiplier gain, provides an accurate means to test it experimentally.

Equation 3.23 can be restated in the form

$$v_G = v_\eta + \frac{v_M}{\eta} \quad (3.27)$$

where

$$v_M = \frac{1}{g_1} \cdot \frac{g}{g - 1} \quad (3.28)$$

is the relative variance of the electron multiplier gain. This can be calculated from the single-electron spectrum obtained as follows.

The cathode is illuminated by a constant flux sufficient to cause the emission of fewer than 10^4 electrons per second, so that the mean interval between successive electrons is at least 100 μ s. The resulting anode charges are then integrated with a time constant of less than 1 μ s so that each voltage pulse has a very low probability (< 1%) of being due to the emission of more than one electron at the cathode. The resulting anode pulse histogram obtained from a multichannel analyser constitutes the single-electron spectrum.

Figures 2.4 and 3.8 show examples of single-electron spectra obtained in this way from photomultipliers with focusing and with venetian-blind dynodes. In neither case is the Poisson-distribution assumption confirmed: the variance is larger than that given by Eq.3.28, the discrepancy being greater for venetian-blind than for focusing dynodes. Anything that interferes with the input system focusing or impairs the collection efficiency increases the variance.

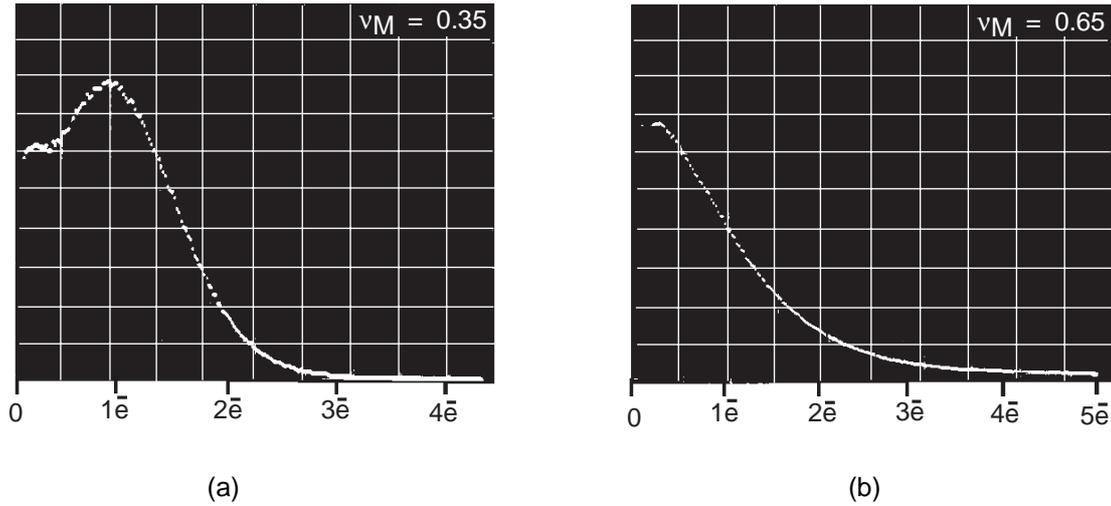


Fig.3.8 Single-electron spectra of photomultipliers with CuBe dynodes a) linear-focusing, b) venetian-blind. Vertical axes, pulse frequency (arbitrary scale; horizontal axes, anode pulse amplitude referred to number of electrons per pulse at the cathode

Considering that the relative variance v_G only appears as a corrective factor in Eq.3.25 and 3.26, it is simpler and sufficient to use Eq.3.27 and to consider v_η as an experimental factor. For tubes with focusing dynodes, v_η is between 0.1 and 0.2, depending on structural details and the voltages used. For venetian-blind dynode structures, v_η is between 0.2 and 0.4. At very low voltage (for example, 100 V between cathode and first dynode), v_η may be as high as 0.5.

Equation 3.28 has the merit of highlighting the predominant effect of the first-stage gain g_1 . For example, if $g_1 = 7$, $g = 4$, and $v_\eta = 0.1$, v_G can be about 0.3; whereas if $g_1 = g = 3$, and $v_\eta = 0.4$, v_G can be close to unity.

3.2.6 Johnson noise

Equation 3.25 can be rewritten as follows to give the reciprocal of the signal-to-noise ratio at the anode

$$\frac{N}{S} = \sqrt{\frac{\overline{i_a^2}}{I_a^2}} = \sqrt{\frac{2e\Delta f (1 + v_G)}{I_k}} = \sqrt{\frac{2e\Delta f (1 + v_G)}{\Phi S_k}} \quad (3.29)$$

where Φ is the incident flux and S_k the cathode sensitivity. *Comparison of this with Eqs 3.17 and 3.18 illustrate one of the fundamental advantages of amplification by secondary emission; the signal-to-noise ratio is only slightly degraded (20% in the worst case) and, to a first approximation, is nearly independent of the gain used.* Hence, there is no objection to using the maximum gain, provided the mean anode current stays within the permissible limits and no feedback phenomena interfere with the operation (§5.5.2).

In any case, there is a minimum gain for which the signal noise (or shot noise) becomes predominant compared with the Johnson noise in the load resistance R_L . (For simplicity, intrinsic amplifier noise is here disregarded). The RMS value of the signal noise is given by

$$R_L \sqrt{i_a^2} = GR_L \sqrt{2e\Delta f (1 + v_G) I_k} \quad (3.30)$$

and the RMS value of Johnson noise by

$$\sqrt{4kTR_L\Delta f} \quad (3.31)$$

from which it is possible to calculate that the minimum gain for which the Johnson noise is negligible compared with the signal noise,

$$G \gg \sqrt{\frac{2kT}{eR_L I_k (1 + v_G)}} \quad (3.32)$$

or, with $T = 300 \text{ K}$, $k = 1.38 \times 10^{-23} \text{ J/K}$, $v_G = 0.5$

$$G^2 \gg \frac{3.5 \times 10^{-2} (\text{V})}{R_L (\Omega) I_k (\text{A})}$$

An approximation giving a reasonable margin of safety is

$$G_{\min}^2 = \frac{1 (\text{V})}{R_L (\Omega) I_k (\text{A})} \quad \text{or} \quad G_{\min} = \frac{1 (\text{V})}{R_L (\Omega) I_a (\text{A})} \quad (3.33)$$

where the mean anode current $I_a = GI_k$

3.2.7 Scintillation detection

Pulse mode. This is the mode used in nuclear spectrometry. Equation 3.26 shows that if the number of light photons per scintillation is proportional to the energy E_{ph} of the X- or γ -photons absorbed in the scintillator, the energy resolution should vary inversely as $\sqrt{E_{ph}}$. However, this is not observed experimentally, especially at high energy levels. When the γ -radiation of ^{137}Cs (662 keV) is absorbed in a NaI(Tl) scintillator for example, the number of photons emitted should give a resolution of about 5%; however, the value observed experimentally is closer to 7%. Reasons for the discrepancy are:

- scintillation efficiency is not uniform throughout the bulk of the scintillator;
- cathode sensitivity is not uniform – scintillations from different parts of the scintillator give rise to different numbers of electrons;
- variations in the energy conversion process in the scintillator.

Faulty optical coupling, incomplete collection at the first dynode, or anything tending to impair cathode uniformity also impairs the energy resolution.

A theoretical analysis which has served as the basis for all subsequent statistical treatments of energy spectrometry has shown that the energy resolution can be expressed by the general equation

$$R_e = 2.36 \sqrt{\left(v_{n\ p,s} + \frac{1 - \rho + v_G}{\bar{n}_{p,s} \rho} \right)} \quad (3.34)$$

where ρ is the photocathode quantum efficiency, v_G the relative variance of the photomultiplier gain (Eq.3.23 or 3.27) and $v_{n\ p,s}$ the relative variance of the photon distribution, which must be determined empirically.

Letting $k = E_{ph}/\bar{n}_{p,s}$ represent the ratio of the initial energy of the X- or γ -photons completely absorbed in the scintillator to the mean number of photons received by the photocathode, Eq.3.34 can be rewritten in the form

$$R_e^2 = \alpha + \frac{\beta}{E_{ph}} \quad (3.35)$$

where

$$\alpha = 2.36^2 v_{n\ p,s}$$

and

$$\beta = 2.36^2 (1 - \rho + v_G) \frac{k}{\rho}$$

In Fig.3.9, a straight line corresponding to Eq.3.35 is superimposed on values obtained by measurement. At energies below about 300 keV Eq.3.35 is in good agreement. From the intersection of the line with the ordinate it is possible to determine α and $v_{n\ p,s}$ empirically. At higher energies the resolution is better than predicted by Eq.3.35, as shown by the measured values being below the straight line. Various explanations for this have been proposed.

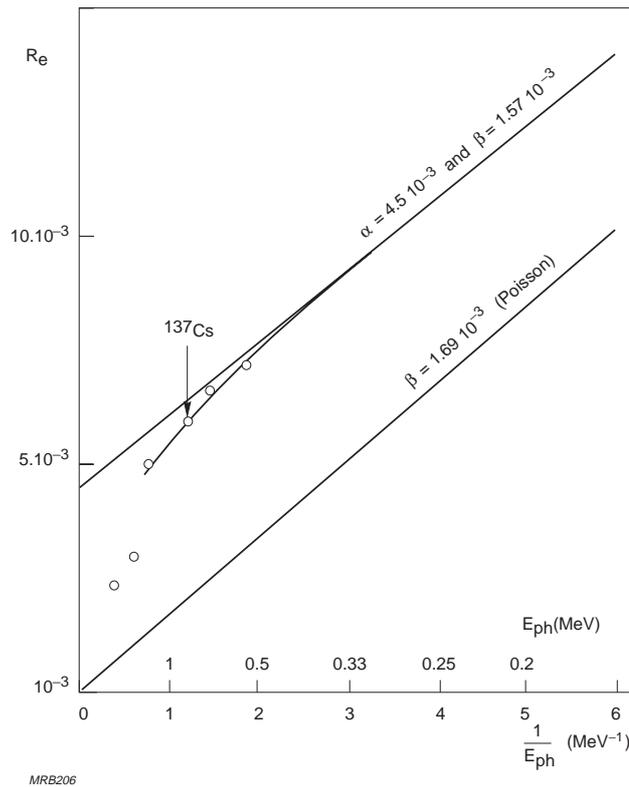


Fig.3.9 Energy resolution as a function of the X or γ photon energy E absorbed in the scintillator

The lower line in Fig.3.9 is based on an assumed Poisson distribution, $v_{n p,s} = 1/\bar{n}_{p,s}$, for which Eq.3.35 simplifies to

$$R_e^2 = \frac{\beta'}{E_{ph}}$$

where $\beta' = 2.36^2 (1 + v_G) k/\rho$. This is identical to Eq.3.26 and represents a limiting case for energy resolution; the slope of the line depends on the quantum efficiency and the variances of the gain.

Continuous mode. In some applications, scintillation detectors are used in a continuous mode. In scanning electron microscopy, for example, electrons reflected from a specimen are accelerated onto a scintillator and the resulting output of a photomultiplier provides the video signal for a television monitor. Continuous thickness measurements of sheet metal, paper, etc. by the absorption of β - or γ -radiation are also based on this mode.

The quantity to be measured is the arrival rate of the quanta (X-, γ -photons, α -, β -particles, etc.). However, because some give rise to no scintillation, there is an

uncertainty in the number of quanta, n_q , which depends on the probability p of their interaction with the scintillator. Thus, if Q is the ratio of the number of electrons received at the anode, n_a , to the number of interactions in the scintillator per unit of time,

$$Q = \frac{n_a}{n_q p} \quad \text{and}$$

$$n_a = n_q p Q \quad (3.36)$$

which is equivalent to Eq.3.21, so the same analysis applies. Replacing the number of photons $n_{p,\tau}$, the quantum efficiency ρ and the gain G by the n_q , p and Q respectively, we can write Eq 3.17 in the form

$$\frac{\overline{i_k^2(t)}}{I_k^2} = \frac{2\Delta f}{\overline{n_q p}} \quad (3.37)$$

This allows only for fluctuations in the number of interactions in the scintillator; however, the number of photons created per interaction, the number of photoelectrons emitted, and the number of secondary electrons are also subject to fluctuation, independently of each other and of the number of scintillator interactions. Fortunately, the total of these fluctuations can be evaluated on the basis of the anode pulse amplitude distribution corresponding to the *single-quantum spectrum* (SQS). This can be determined with a multichannel analyser by irradiating the scintillator with quanta of the type in question (X- or γ -photons, α -, β -particles, etc.). The spectrum is characterized by a relative variance v_Q . With an NaI(Tl) scintillator, for example, this ranges from 0.002 for the γ -radiation of ^{57}Co ($E_\gamma = 122$ keV) to 0.450 for the γ -radiation of ^{137}Cs ($E_\gamma = 662$ keV). Some spectra with an exponential shape or with widely separated peaks (§A.3) can have a variance close to unity. By analogy with Eq.3.25, v_G can be replaced by v_Q and Eq.3.37 rewritten in the form

$$\frac{\overline{i_a^2(t)}}{I_a(t)^2} = \frac{2\Delta f}{\overline{n_q p}} (1 + v_Q) \quad (3.38)$$

v_Q , like v_G , is a corrective term that lowers the S/N ratio by a factor $1/(1+v_Q)^{1/2}$, which is usually between 0.7 and 0.9; the number of interactions in the scintillator remains the predominant factor. Table 3.1 compares the parameters used in the pulse-mode and analogue-mode analysis and illustrates their close parallelism.

Table 3.1

Statistical parameters in the detection of continuous light and scintillation pulses.

continuous light	scintillation pulses
<div style="text-align: center;"> <div style="border: 1px solid black; padding: 5px; width: fit-content; margin: 0 auto;">photon flux</div> <p>⇓</p> <p><i>photon noise</i></p> <p>relative variance, $\frac{1}{\bar{n}_p}$</p> </div>	<div style="text-align: center;"> <div style="border: 1px solid black; padding: 5px; width: fit-content; margin: 0 auto;">X, γ, α, β quanta, etc.</div> <p>⇓</p> <p><i>quantum noise</i></p> <p>relative variance, $\frac{1}{\bar{n}_q}$</p> </div>
<div style="text-align: center;"> <div style="border: 1px solid black; padding: 5px; width: fit-content; margin: 0 auto;">photoemission quantum efficiency, ρ</div> <p>⇓</p> <p><i>random number of photoelectrons</i></p> <p>relative variance, $\frac{1}{\bar{n}_p \rho}$</p> </div>	<div style="text-align: center;"> <div style="border: 1px solid black; padding: 5px; width: fit-content; margin: 0 auto;">scintillator interactions interaction probability, p</div> <p>⇓</p> <p><i>random number of scintillations</i></p> <p>relative variance, $\frac{1}{\bar{n}_q p}$</p> </div>
<div style="text-align: center;"> <div style="border: 1px solid black; padding: 5px; width: fit-content; margin: 0 auto;">multiplication by secondary emission</div> <p>⇓</p> <p><i>single-electron spectrum (SES)</i></p> <p>relative variance, $v_M \rightarrow v_G$</p> </div>	<div style="text-align: center;"> <div style="border: 1px solid black; padding: 5px; width: fit-content; margin: 0 auto;">photon-current conversion</div> <p>⇓</p> <p><i>single-quantum spectrum (SQS)</i></p> <p>relative variance, v_Q</p> </div>
<div style="text-align: center;"> <p><i>noise-to-signal ratio</i></p> $\frac{N}{S} = \sqrt{\frac{2\Delta f(1 + v_G)}{\bar{n}_p \rho}} \quad (\text{Eq.3.25})$ </div>	<div style="text-align: center;"> <p><i>noise-to-signal ratio</i></p> $\frac{N}{S} = \sqrt{\frac{2\Delta f(1 + v_Q)}{\bar{n}_q p}} \quad (\text{Eq.3.38})$ </div>

3.3 Equivalent noise input and minimum discernible signal

Because the noise sources dealt with so far cause uncertainty in the detection of small signals or inaccuracy in their measurement, we have pursued an analysis that emphasizes the relation between noise and energy resolution and leads to a mathematical expression for signal-to-noise ratio. Now we shall give a more general definition of signal-to-noise ratio and introduce the new quantities, *noise equivalent power* and *equivalent noise input*. These can be measured experimentally and used for determining the *minimum discernible signal*. In practice, however, the minimum discernible signal also depends on the method of detection, the observer, and the probability that can be accepted of obtaining a spurious signal.

3.3.1 Definitions

Several definitions relating to maximum sensitivity are sanctioned by usage or recommended by the IEC. However there are still no full and precise standards on the subject. Here we shall give the most useful definitions, together with the abbreviations under which they are often known.

Equivalent anode dark current input, EADCI, of an individual photomultiplier is the flux that must be applied to the photocathode to produce an anode current equal to the dark current I_{ao} . Hence, it is the ratio of the anode dark current to the anode sensitivity

$$\text{EADCI} = \frac{I_{ao}}{S_a} \quad (3.39)$$

expressed in watts or lumens, depending on the unit of S_a . It varies greatly with the experimental conditions (temperature, humidity, stabilizing time), and is of practical interest only when the tube is operating continuously.

Signal-to-noise ratio, S/N, defined at the system output is the ratio between the RMS values of the output signal and the overall noise (signal noise plus dark current noise) within the system bandwidth.

Noise equivalent power, NEP, (symbol P_N) is the incident flux that produces an RMS output signal equal to the RMS noise measured at the output under specific operating conditions. It may be expressed in watts or lumens depending on the nature of the application and the incident flux, and is significant only in relation to the specified set of operating conditions; (for instance: incident flux spectrum, modulation frequency, measuring equipment bandwidth, illuminated cathode area, operating temperature). The lower the noise equivalent power, the better the photomultiplier is able to detect low flux inputs.

Some publications specify P_N for a bandwidth of 1 Hz based on the assumption of constant spectral density throughout the frequency range concerned. It is then expressed in $W/Hz^{1/2}$ or $lm/Hz^{1/2}$.

Equivalent noise input, ENI, (symbol E_N) is the ratio of the noise equivalent power to the area of the cathode, assuming the whole cathode to be uniformly illuminated. It is expressed in W/m^2 or lm/m^2 . For a bandwidth of 1 Hz it is expressed in $W/m^2 Hz^{1/2}$ or $lm/m^2 Hz^{1/2}$. Like noise equivalent power, a specified equivalent noise input is significant only in relation to a specified set of operating conditions.

Detectivity, D , is the reciprocal of the noise equivalent power:

$$D = \frac{1}{P_N} \quad (3.40)$$

It is expressed in W^{-1} or lm^{-1} .

3.3.2 Minimum value of noise equivalent power

The photomultiplier and its load resistance account for the greater contribution to the total noise measured, the electronics for a much smaller contribution. Two of the main sources of noise are:

- the shot noise associated with the photomultiplier signal current, Eq.3.30
- the Johnson noise due to the load resistance, Eq.3.31.

Both apply over a frequency interval of Δf . All the other causes of noise can be taken into account by assuming an additional random component of the anode current with a mean square value $\overline{i_{a,n}^2}$ over the interval Δf . The signal-to-noise ratio of the photomultiplier is then

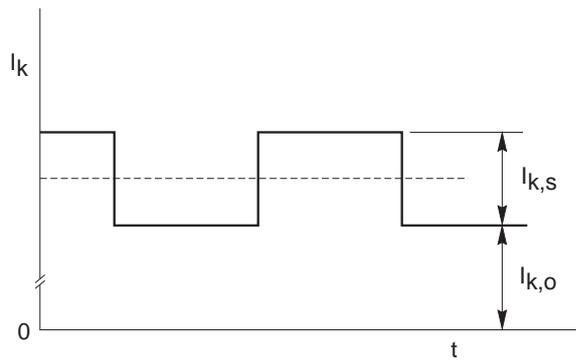
$$\frac{S}{N} = \frac{I_a}{\sqrt{\frac{4kT\Delta f}{R_L} + 2e G^2 I_k \Delta f (1 + v_G) + \overline{i_{a,n}^2}}} \quad (3.41)$$

The term $\overline{i_{a,n}^2}$ mainly represents the noise associated with the different components of the dark current (thermionic noise, field emission noise, etc.) plus such additional noise as the input noise of the electronics. It does not necessarily obey the Schottky formula, nor should it be regarded as wideband white noise; it is sufficient to bear in mind that it applies to a frequency interval Δf equal to the energy bandwidth B_N , or *equivalent noise bandwidth*, of the measuring circuit.

It is more convenient to express the signal-to-noise ratio as a function of the cathode current; Eq.3.41 then becomes

$$\frac{S}{N} = \frac{I_k}{\sqrt{B_N \left(\frac{4kT}{G^2 R_L} + 2eaI_k + \frac{\overline{i_{k,n}^2}}{B_N} \right)}} \quad (3.42)$$

where $a = 1 + v_G$ and $\overline{i_{k,n}^2} = \overline{i_{a,n}^2} / G^2$ is the calculated additional anode noise referred to the cathode. In Eq.3.42 this term is referred to the unit of bandwidth ($\overline{i_{k,n}^2} / B_N$).



MRB207

Fig.3.10 Cathode current when the input flux is chopped

Equation 3.42 shows that *the most effective way to improve the signal-to-noise ratio is to reduce the bandwidth B_N of the measuring circuit*, at least in the frequency range where the $\overline{i_{k,n}^2} / B_N$ is still small. A light modulator combined with a very narrow band AC amplifier is often used for this (Fig.3.13) and has the advantage of automatically cancelling the DC component of the dark current, although the noise associated with the dark current is amplified normally. If the amplifier bandwidth is very narrow (a few hertz), only the fundamental component of the modulated signal is amplified.

If the light is modulated by a symmetrical square wave (Fig.3.10), the RMS value of the fundamental component of the cathode current due to the signal is

$$\frac{I_{k,s}}{\pi} \sqrt{2} \quad (3.43)$$

and the mean noise current associated with the signal and measured in the equivalent noise bandwidth B_N is

$$\overline{i_{k,s}^2} = \frac{2ea I_{k,s} B_N}{2}$$

Therefore, assuming that G^2RL is large enough to allow the Johnson noise to be neglected, the ratio of the RMS signal and noise currents is

$$\frac{S}{N} = \frac{I_{k,s}}{\pi} \sqrt{\frac{2}{\left(eaI_{k,s}B_N + \overline{i_{k,n}^2} \right)}} \quad (3.44)$$

The noise equivalent power, P_N , is found by setting the S/N ratio equal to unity:

$$P_N = \frac{I_{k,s}}{S_k} \text{ at } \frac{S}{N} = 1 \quad (3.45)$$

If the equivalent noise bandwidth B_N tends towards zero,

$$\lim_{B_N \rightarrow 0} \frac{P_N}{\sqrt{B_N}} = \frac{\pi}{S_k} \sqrt{\frac{\overline{i_{k,n}^2}}{2B_N}} \quad (3.46)$$

provided the spectral density $\overline{i_{k,n}^2}/B_N$ tends towards a finite limit as B_N tends towards zero. This shows that P_N has significance only when the bandwidth is defined in which $\overline{i_{k,n}^2}$ is specified (width and centre frequency). The value of $\sqrt{\overline{i_{k,n}^2}/B_N}$ can be of the order of 10^{-16} A/Hz^{1/2} for a photomultiplier having a photoemission threshold in the visible spectrum ($W_{ph} > 1.5$ eV).

A limiting case to consider is that in which the noise component $\overline{i_{k,n}^2}$ represents only the noise power associated with the dark current, that is to say

$$\overline{i_{k,n}^2} = 2ea I_{ko} B_N \quad (3.47)$$

where I_{ko} is the equivalent dark current of the photocathode, corresponding to the anode dark current divided by the gain, $I_{ko} = I_{ao}/G$. Equation 3.47 represents a minimum value for the noise component $\overline{i_{k,n}^2}$, disregarding the DC component of the dark current (i.e. the leakage current, which is usually negligible) and the noise contribution of the measuring circuits. The minimum noise equivalent power P_N referred to the bandwidth is then

$$\lim_{B_N \rightarrow 0} \left(\frac{P_N}{\sqrt{B_N}} \right) = \frac{\pi}{S_k} \sqrt{eaI_{ko}} \quad (3.48)$$

3.3.3 Effect of bandwidth

Equation 3.48 is valid only when the bandwidth is very small (a few hertz); we shall now consider how the noise equivalent power varies as bandwidth increases.

Assume that the bandwidth of the amplifier in the measuring circuit is equivalent to that of an RC circuit, the DC component of the signal being blocked (Fig.3.13). When the incident flux is modulated by a symmetrical square wave (Fig.3.10), the RMS value of the periodic component of the cathode current due to the signal is

$$\sqrt{(I_k - I_{ko})^2} = m \frac{I_{k,s}}{2} \quad (3.49)$$

where the modulation coefficient $m \leq 1$ and depends on the bandwidth and modulation frequency. If the bandwidth is taken as the 3 dB cut-off frequency of the RC circuit.

$$B_{3dB} = f_c = \frac{1}{2\pi RC}$$

and the modulation frequency f_m as 100 Hz, Fourier analysis shows that the value of m is

$$m = \frac{2}{\pi} \sqrt{2 \sum_{n=1}^{\infty} \frac{1}{(2n-1)^2} \cdot \frac{1}{1 + \left(\frac{100(2n-1)}{B_{3dB}}\right)^2}}$$

Figure 3.11 shows the variation of m as a function of B_{3dB} . The signal-to-noise ratio is thus

$$\frac{S}{N} = \frac{m I_{k,s}}{2 \sqrt{\left\{ e a I_{k,s} B_N + \overline{i_{k,n}^2} \right\}}} \quad (3.50)$$

and, from Eq.3.45, the noise equivalent power is

$$P_N = \frac{2eaB_N}{m^2 S_K} \left\{ 1 + \sqrt{1 + \frac{m^2 \overline{i_{k,n}^2}}{e^2 a^2 B_N^2}} \right\} \quad (3.51)$$

When the bandwidth is high (in practice, $B_{3dB} > 10^3$ Hz), m tends towards unity and the expression for P_N/NB_N tends towards

$$\lim_{B_N \rightarrow \infty} \left(\frac{P_N}{\sqrt{B_N}} \right) = \frac{4ea}{S_k} \sqrt{B_N} \quad (3.52)$$

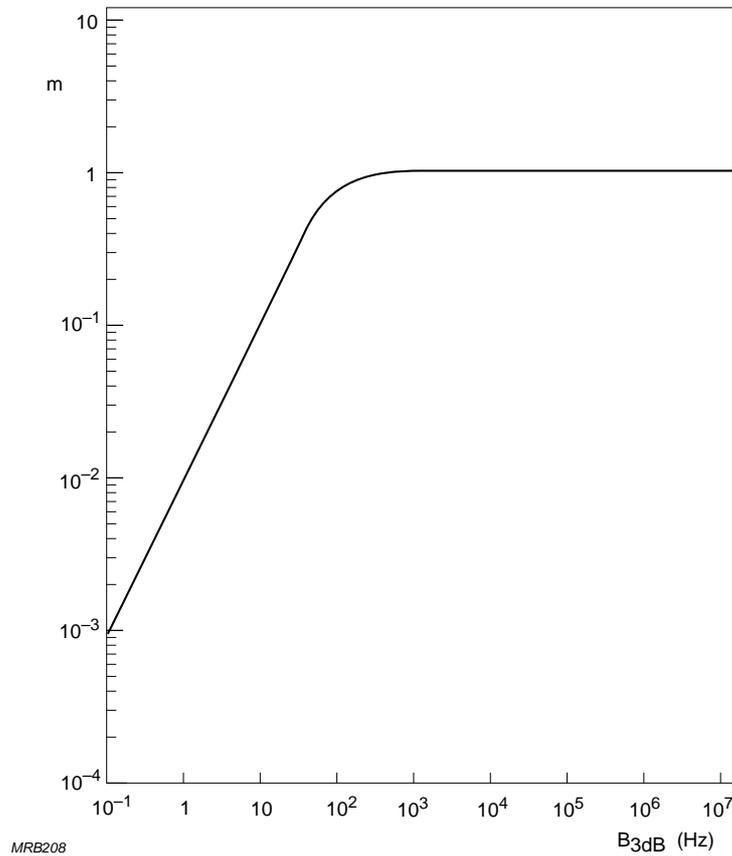


Fig.3.11 Modulation coefficient m as a function of the 3 dB bandwidth of the measuring circuit. ($f_m = 100$ Hz)

Figure 3.12 shows the variation in the noise equivalent power P_N as a function of the energy bandwidth B_N for two values of $i_{k,n}^2$. The solid line on the left correspond to very low values of bandwidth ($B_N < 10$ Hz) centred about the fundamental frequency of the modulated signal. This is where the smallest value of the noise equivalent power is obtained; the practical minimum is determined by the noise associated with the dark current, hence the interest in keeping the dark current as low as possible. The solid lines on the right correspond to high values of bandwidth ($B_N > 10^3$ Hz, for a modulation frequency of 100 Hz). This is where the noise equivalent power P_N becomes proportional to the energy bandwidth of the measuring circuit.

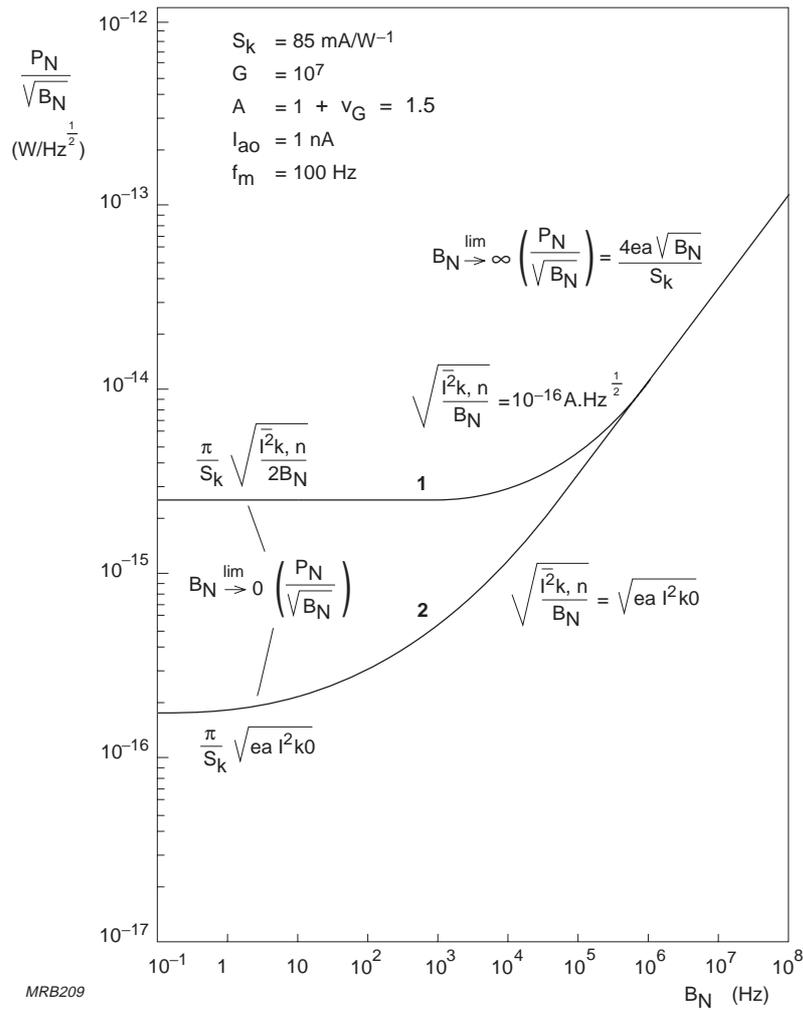


Fig.3.12 Noise equivalent power (flux as a function of the measuring circuit passband, for two value of additional noise: 1) $10^{-16} \text{ A/Hz}^{1/2}$, 2) $7 \times 10^{-18} \text{ A/Hz}^{1/2}$

3.3.4 Measurement of noise equivalent power

Figure 3.13 shows the set-up used for determining the minimum value of the noise equivalent power. The incident flux is mechanically chopped to obtain a symmetrical square-wave voltage at the photomultiplier anode; the usual chopping frequency is 100 Hz. The photomultiplier signal is applied to a filter with a 3 dB bandwidth of 1000 Hz (the preceding amplifier blocks the DC component). A voltmeter measures the RMS value of the amplified voltage.

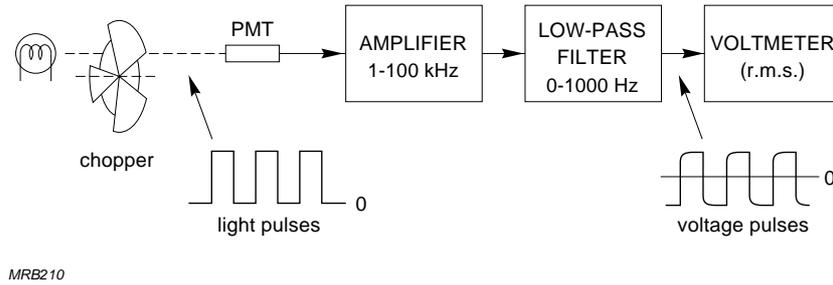


Fig.3.13 Set-up for measuring noise equivalent power

The minimum noise equivalent power is determined by measuring the signal-to-noise ratio. The effective bandwidth of the measuring set-up being 1000 Hz, the signal-to-noise ratio at a bandwidth of 1 Hz is obtained by calculation. Two measurements are carried out in succession:

- first, with no light, the RMS noise voltage V_N is measured for an energy bandwidth $B_N = (\pi/2)1000$ Hz. Assuming the spectral density of the noise to be constant throughout the frequency range under consideration, the RMS noise voltage for a bandwidth of 1 Hz is then

$$V_N (1 \text{ Hz}) = \frac{V_N}{\sqrt{\frac{\pi}{2} 1000}} \quad \text{V/Hz}^{1/2} \quad (3.53)$$

- next, with chopped light, the RMS signal voltage V_S is measured for a 3 dB bandwidth of 1000 Hz. The RMS value calculated for a bandwidth of 1 Hz is the RMS value of the fundamental component of the modulated voltage. Thus, from Eq.3.43 and 3.49,

$$V_S (1 \text{ Hz}) = \frac{2V_s\sqrt{2}}{m\pi}$$

where m takes account of the finite bandwidth (Fig.3.11). For a bandwidth of 1000 Hz and a modulation frequency of 100 Hz, $m = 0.968$; hence,

$$V_s (1 \text{ Hz}) = 0.930 V_s \quad (3.54)$$

Thus, from Eqs 3.53 and 3.54, for an energy bandwidth of 1 Hz,

$$\frac{V_S (1 \text{ Hz})}{V_N (1 \text{ Hz})} = 0.930 \frac{V_S}{V_N} \sqrt{1000 \frac{\pi}{2}}$$

Finally, the minimum noise equivalent power is given by

$$P_N = \Phi \frac{V_N (1 \text{ Hz})}{V_S (1 \text{ Hz})} \quad (3.55)$$

or

$$P_N = \frac{V_N \Phi}{0.930 V_S \sqrt{1000 \frac{\pi}{2}}} \quad (3.56)$$

expressed in $\text{W/Hz}^{1/2}$ or in $\text{lm/Hz}^{1/2}$ depending on the unit of incident flux, Φ .

APPENDIX 3

A3.1 Practical scintillation spectra

β-spectrum The β-spectrum is intermediate between the uniform and the triangular distributions; its relative variance v is close to 0.4.

γ-spectrum consisting of a single gaussian peak (^{55}Fe , ^{57}Co), $R_e = 2.36\sqrt{v}$.

$$^{55}\text{Fe}: E_\gamma = 5.9 \text{ keV}, R_e = 0.40, v = 2.9 \times 10^{-2}$$

$$^{57}\text{Co}: E_\gamma = 122 \text{ keV}, R_e = 0.10, v = 1.8 \times 10^{-3}$$

γ-spectrum with Compton distribution For a typical case, ^{137}Cs ($E_\gamma = 662 \text{ keV}$), $v \approx 0.45$.

Complex spectrum with several peaks approximating delta functions When the peaks are narrow and entirely separate v is between 0.5 and 1.

A3.2 Noise equivalent bandwidth

Consider a linear transmission system having a transfer function $G(f)$ for a signal of frequency f and unit amplitude. If a randomly varying signal $i(t)$ – for example, the noise component of photomultiplier anode current – is applied to such a system, its response $i_s(t)$ is given by

$$\overline{i_s^2(t)} = \int_0^\infty w(f) |G(f)|^2 df$$

where $w(f)$ is the energy density spectrum of $i(t)$.

If $w(f)$ can be assumed to be independent of frequency so that it can be replaced by a constant w_0 , then the original transmission system can be replaced by a notional one which fulfils the following conditions:

- it transmits a constant power density equal to $w_0 G_0^2$ throughout a frequency interval $B_N = f_2 - f_1$;
- the total noise power it transmits throughout this interval is equal to that of the original system.

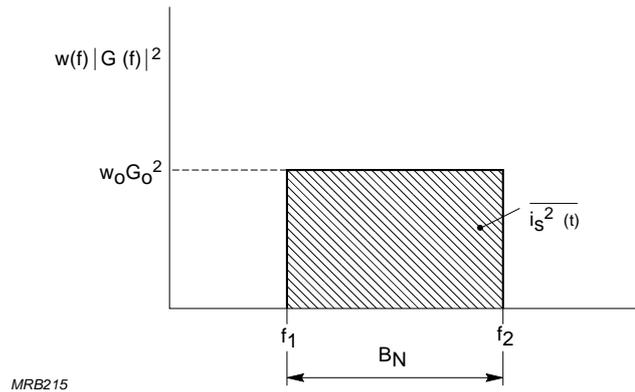


Fig.A3.1 Definition of noise pass-band B_N .

For such a system

$$w_0 G_0^2 B_N = w_0 \int_0^{\infty} |G(f)|^2 df$$

whence

$$B_N = \frac{1}{G_0^2} \int_0^{\infty} |G(f)|^2 df$$

where G_0 may be either the maximum value of $G(f)$ or its value at the centre frequency between f_1 and f_2 .

B_N defines the energy bandwidth, or *noise equivalent bandwidth*, of the actual transmission system. The transfer function of such a system may take a variety of forms; a commonly encountered one is equivalent to that of an *RC* circuit,

$$|G(f)| = \frac{G_0}{1 + (2\pi f RC)^2}$$

for which $B_{3dB} = 1/2\pi RC$ and the noise equivalent bandwidth is

$$B_N = \frac{1}{4RC} = \frac{\pi}{2} B_{3dB}$$

This treatment applies only to linear systems; non-linear systems require a more general interpretation of noise equivalent bandwidth.

CHAPTER 4

OPERATING CHARACTERISTICS AND CONSIDERATIONS

4.1	Signal-to-noise ratio comparison	4.7	Afterpulses
		4.7.1	Luminous reactions
		4.7.2	Ionization of residual gases
4.2	Photomultiplier selection criteria	4.7.3	Afterpulse factor
4.3	Factors affecting sensitivity	4.8	Environmental considerations
4.3.1	Wavelength	4.8.1	Temperature
4.3.2	Collection efficiency	4.8.2	Magnetic fields
4.3.3	Angle of incidence	4.8.3	Radiation
4.4	Time characteristics	4.8.4	Atmosphere
4.4.1	Pulse response: determining factors	4.8.5	Mechanical stress
4.4.2	Pulse response: measurement	Appendix	Signal transfer in linear systems
4.4.3	Transit time differences	A4.1	Pulse and step response
4.4.4	Transit time spread	A4.1.1	Superposition principle
4.4.5	Frequency response	A4.1.2	Rise time and FWHM
4.5	Linearity	A4.2	Time resolution
4.5.1	External factors affecting linearity	A4.2.1	Delta-function light pulse
4.5.2	Internal factors affecting linearity	A4.2.2	Arbitrary light pulse
4.5.3	Linearity measurement		
4.6	Stability		
4.6.1	Long-term drift		
4.6.2	Short-term drift		

OPERATING CHARACTERISTICS AND CONSIDERATIONS

Photomultiplier characteristics that need to be considered in most applications include sensitivity, time and frequency response, stability, linearity, and possible environmental effects at a certain gain.

Before choosing a photomultiplier for a given application, however, it is well to establish that a photomultiplier is in fact the best type of detector for that application. Alternatives, such as a vacuum or semiconductor photodiode plus a high-gain, low-noise amplifier, may offer advantages in size, power supply, or cost. Other criteria on which the choice may depend include spectral sensitivity, frequency response, and output current range. Leaving all these out of consideration, however, the area where a photomultiplier clearly excels is in its ability to detect very low-level light; and in particular, below a certain threshold, to do so with a better signal-to-noise ratio than any alternative detector. In many applications this is decisive. For the light level to be detected and the required signal-to-noise ratio, a photomultiplier may be the only choice.

4.1 Signal-to-noise ratio comparison: photomultiplier vs photodiode

Whether the detector is a photomultiplier, or a photoemissive cell, or a photodiode plus amplifier, the critical factor governing signal-to-noise ratio is the quantity of light received, which often depends on the size of the sensitive surface. To make the following comparison independent of size, therefore, it will be based on the assumption of equal photocurrents: I_k for the photomultiplier, and I_d for the photodiode.

Equation 3.42 gave the signal-to-noise ratio of a photomultiplier with cathode current I_k , gain G , load resistance R_L , and anode dark-current noise (referred to the cathode) $i_{k,n}$. The equivalent expression for a photodiode, plus low-noise, high-gain amplifier connected as a current-voltage converter with feedback resistance R_f is

$$\frac{S}{N} = \frac{I_d}{\sqrt{B_N \left(\frac{4kT}{R_f} + 2eI_d \right) + \overline{i_{d,n}^2} + \overline{i_{o,n}^2}}} \quad (4.1)$$

where $i_{d,n}$ is the intrinsic noise current of the detector, and $i_{o,n}$ the input equivalent noise current specified by the amplifier manufacturer. (The amplifier input equivalent noise voltage, $e_{o,n}$, also specified by the manufacturer, is usually negligible and is here disregarded.) Equating Eq.3.42 with Eq.4.1 gives the photocurrent at which both

detectors have like signal-to-noise ratios,

$$R_f I_k = R_f I_d = \frac{2kT}{e(a-1)} \left(1 - \frac{R_f}{G^2 R_L} \right) + R_f \frac{\overline{i_{d,n}^2} + \overline{i_{o,n}^2} - \overline{i_{k,n}^2}}{2e(a-1)B_N} \quad (4.2)$$

If $G^2 R_L \gg R_f$, this simplifies to

$$R_f I_k = R_f I_d = \frac{2kT}{e(a-1)} + R_f \frac{\overline{i_{d,n}^2} + \overline{i_{o,n}^2} - \overline{i_{k,n}^2}}{2e(a-1)B_N} \quad (4.3)$$

The two terms of Eq.4.3 correspond to two limiting cases.

If R_f is low (say, less than $10^5 \Omega$), the second term is negligible compared with the first. For $a = 1.5$ (a being $1 + V_G$, where V_G is the photomultiplier gain variance), the signal *voltage* below which the photomultiplier has the better signal-to-noise ratio is

$$R_f I_k = R_f I_d \leq \frac{2kT}{e(a-1)} \approx 100 \text{ mV}$$

The photocurrent or incident flux to which this corresponds can be lowered only by increasing the feedback resistance R_f . But that can be done only at the expense of bandwidth.

If R_f is high (say, more than $10^7 \Omega$), the first term becomes negligible compared with the second. Then, since $i_{d,n}$ and $i_{k,n}$ are also negligible compared with $i_{o,n}$, the voltage threshold below which the photomultiplier has the better signal-to-noise ratio tends toward

$$R_f I_k = R_f I_d \leq R_f \frac{\overline{i_{o,n}^2}}{2e(a-1)B_N}$$

which, for $\sqrt{\overline{i_{o,n}^2}/B_N}$, corresponds to a *photocurrent*

$$I_k = I_d \approx 10^{-7} \text{ A}$$

Thus, at very low light levels requiring a high feedback resistance R_f , the signal-to-noise ratio of a photodiode is limited by the amplifier input equivalent noise current $i_{o,n}$.

In Eqs 4.2 and 4.3 the following practical values apply.

For a photomultiplier

$$\begin{aligned}\sqrt{i_{k,n}^2/B_N} &< 10^{-14} \text{ A/Hz}^{1/2} \text{ with S1 cathode} \\ &< 10^{-16} \text{ A/Hz}^{1/2} \text{ with other cathodes}\end{aligned}$$

for a vacuum photodiode

$$\begin{aligned}\sqrt{i_{d,n}^2/B_N} &< 10^{-14} \text{ A/Hz}^{1/2} \text{ with S1 cathode} \\ &< 10^{-16} \text{ A/Hz}^{1/2} \text{ with other cathodes}\end{aligned}$$

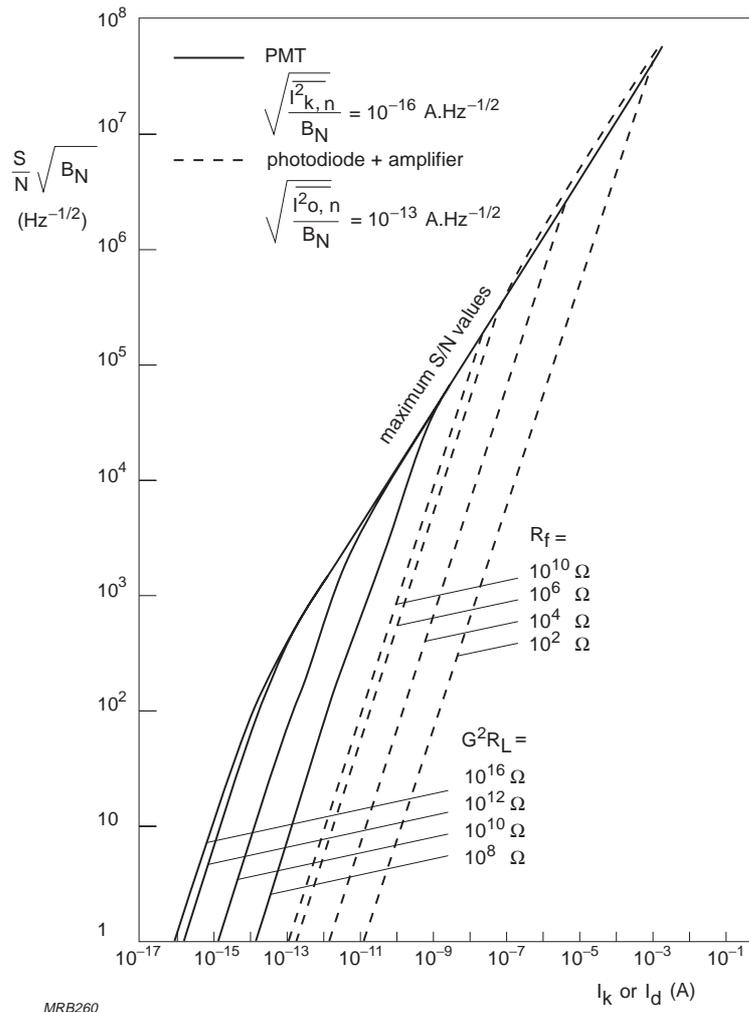
for a low-noise silicon photodiode

$$\sqrt{i_{d,n}^2/B_N} \leq 5 \times 10^{-15} \text{ A/Hz}^{1/2}$$

and for a low-noise operational amplifier with FET input

$$\sqrt{i_{o,n}^2/B_N} > 10^{-13} \text{ A/Hz}^{1/2}$$

Figure 4.1 shows the signal-to-noise ratio, based on these values and at 1 Hz bandwidth, as a function of photocurrent for a photomultiplier and for a photodiode plus low-noise amplifier. For the photomultiplier, G^2R_L is taken as parameter, and for the photodiode R_f . The intersections of the curves with the horizontal axis ($S/N = 1$) indicate the photocurrents at which the signal is no longer distinguishable from the noise. The intersections of the dashed curves with the solid ones indicate the photocurrents at which the signal-to-noise ratio of a photodiode becomes competitive with that of a photomultiplier.



MRB260
 Fig.4.1 Signal-to-noise ratio as a function of photocurrent for a photomultiplier (solid lines) and a photodiode plus low-noise amplifier (dashed line)

4.2 Photomultiplier selection criteria

Points to consider in choosing a tube for a given application are photocathode characteristics, response speed, and number of stages.

Photocathode spectral sensitivity and size. If the light to be detected is monochromatic, choose a tube whose maximum sensitivity is as close as possible to the same wavelength. If it is not monochromatic, look for the best match between sensitivity and spectral distribution, using the matching factor described in A6.3. Bear in mind, though, that the greater the sensitivity in the red, the greater the thermionic emission. And that relative variations of sensitivity versus temperature and from tube to tube, are largest close to the photoemission threshold.

The choice of photocathode diameter depends partly, but not entirely, on the size of the incident light beam or source. Beam or source size determines the minimum practical diameter but not the maximum. If the equipment or installation does not impose strict constraints on size or weight, tubes with medium-diameter cathodes are often preferable to those with small ones. They are generally more stable and have higher permissible anode currents. Dark current does not vary in strict proportion with cathode diameter; and, if the dark current of a tube with a large-diameter cathode is inconveniently high, it can be reduced by reducing the *effective* cathode diameter in the way described in §5.8.2.

Response speed. If the rise time required is about one nanosecond or less, or the required bandwidth more than about 100 MHz, choose a fast-response tube. Such tubes also have the best time resolution.

If the rise time required is a few nanoseconds, or the required bandwidth between 50 and 100 MHz, a standard tube with linear-focusing dynodes is a good choice. Tubes with venetian-blind dynodes are comparatively slow and not suitable for bandwidths of more than about 10 MHz.

Number of stages. General-purpose tubes usually have eight or ten stages and a gain of 10^3 to 10^7 at an applied voltage of 600 to 1800 V. Lowering the voltage impairs fast response and linearity, so if lower gain is required, choose a tube with fewer stages. For gain higher than 10^7 , choose one with more stages. The additional stages enable interdynode voltages to be kept moderate and so prevent dark current from becoming excessive. A tube with twelve stages will safely give a gain of 10^8 ; setting the gain higher than that limits the output current pulse range (§5.5.3) and is seldom justified. A fast-response tube operating at a gain of 10^9 with type A voltage distribution (§5.2.1) approaches its linearity limit even with single-electron pulses.

4.3 Factors affecting sensitivity

The anode sensitivity of a photomultiplier varies according to the part of the cathode surface from which photoemission originates. The variation can be mapped by scanning the cathode with a narrow beam of light and plotting the resulting anode current variation. Figure 4.2 is a plot of the anode sensitivity variation measured relative to one diameter of a 32 mm SbKCs cathode; the dashed line shows the corresponding variation of cathode sensitivity measured along the same diameter with the tube connected as a diode. Comparison of the two curves illustrates the relative variation of the two components of anode sensitivity given in Eq.2.11: namely, the

photocathode sensitivity S_k and the collection efficiency η of the cathode/first-dynode space.

Figure 4.2 was plotted with the scanning beam normal to the cathode; a different angle of incidence would give a different sensitivity contour.

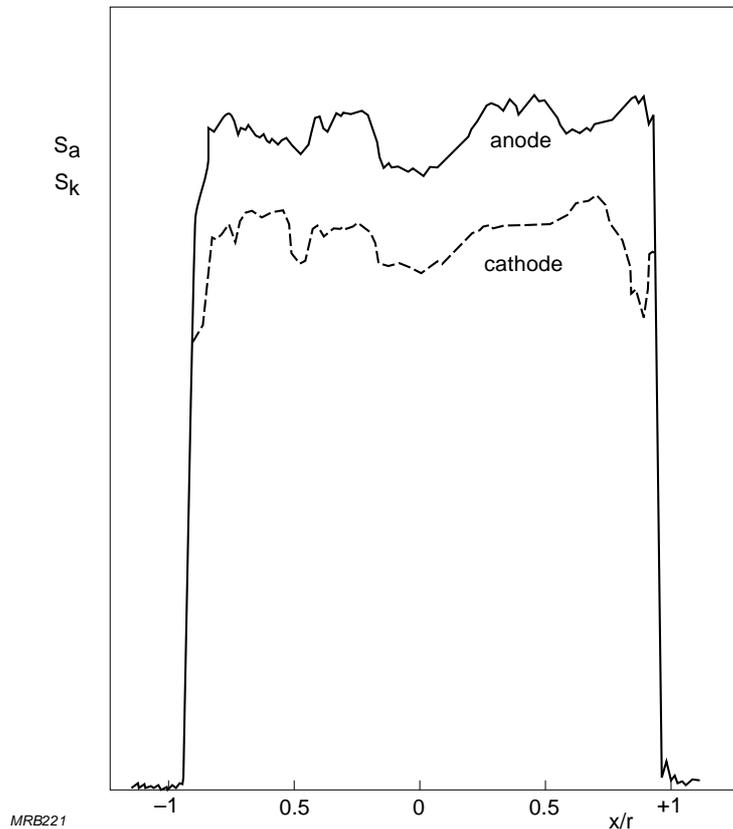


Fig.4.2 Example anode and cathode sensitivity variation measured at $\lambda = 424 \text{ nm}$ along one cathode diameter of a photomultiplier with a 32 mm SbKCs cathode. Vertical axis, relative sensitivity (arbitrary scale); horizontal axis, distance from cathode centre, normalized with respect to cathode radius. Curves offset for clarity

4.3.1 Wavelength

Cathode sensitivity variations depend very much on wavelength. They are usually low in the spectral range corresponding to maximum sensitivity and increase towards the ends of that range. SbCs (S11) and bialkali cathodes are commonly used near their maximum sensitivity wavelengths, where uniformity of sensitivity is generally good. Trialkali (S20, S20R) cathodes, however, are often used near their threshold wavelengths, where uniformity is less good. Figure 4.3 shows the sensitivity variation of an S20 cathode measured at wavelengths of 424 nm, 629 nm, and 800 nm along a diameter aligned with the dynodes; note how the good uniformity at 424 nm deteriorates at the longer wavelengths. For this reason, the measurement of uniformity

of response for tubes with alkali cathodes with, for example, green (560 nm) LEDs can give misleading results.

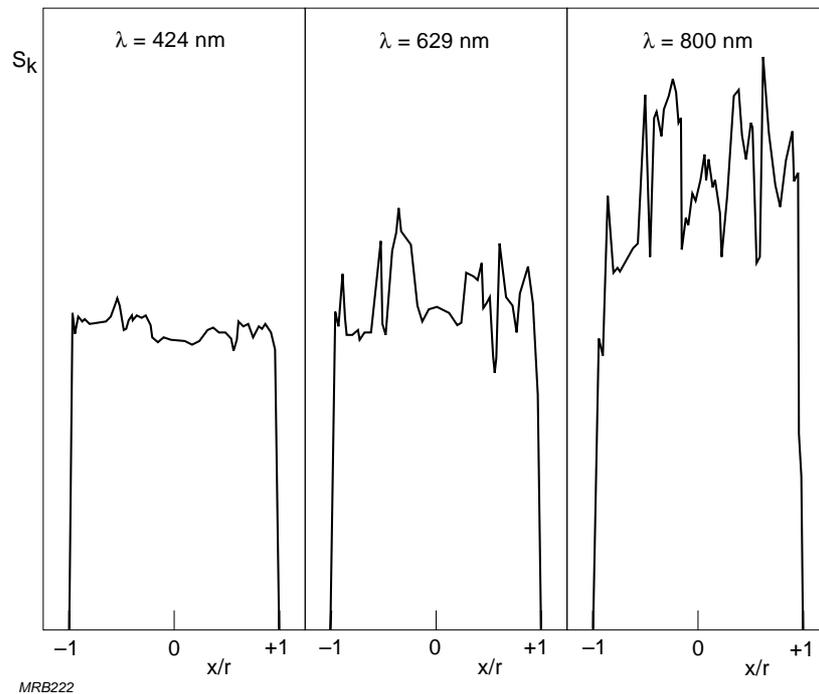


Fig.4.3 Examples of cathode sensitivity variation measured along one diameter of a SbNa_2KCs cathode at three wavelengths

4.3.2 Collection efficiency

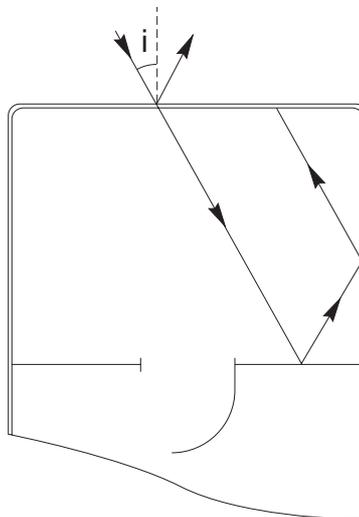
The electron-optical input system is meant to direct all photoelectrons, regardless of their points of origin or initial velocities, onto the useful area of the first dynode. Design features that can contribute to this include: a spherically curved photocathode, a large-area first dynode (e.g. venetian blind), and an electrode in the input system whose potential can be adjusted to compensate for assembly tolerances.

In tubes with focusing dynodes the uniformity of collection is generally best along a diameter perpendicular to the plane of symmetry. The ratio of the first-to-second dynode voltages is also important, for it influences the effective area of the second dynode. To optimize performance, the second dynode potential, and that of any adjustable focusing electrode in the input system, must be carefully adjusted with the tube biased for maximum gain and preferably with the cathode fully illuminated.

Magnetic fields, including the earth's, to which a photomultiplier is exposed have an adverse effect on collection efficiency.

4.3.3 Angle of incidence

Depending on its angle of incidence, light that passes through a semitransparent photocathode may either be reflected back to it from the interior of the tube (Fig. 4.4) or impinge on some internal photoemissive surface. In either case, the resulting photoelectrons may be collected by the multiplier and contribute to the anode signal. This is an important cause of the variation of apparent anode sensitivity with angle of incidence. Figure 4.5(a) shows examples of anode sensitivity contours measured on the same tube at three angles of incidence. As the transmission coefficient of the photoemissive layer varies with wavelength, so does the amount of light that penetrates to the interior of the tube. Thus, the variation of apparent sensitivity with angle of incidence is also a function of wavelength.



MRB223

Fig.4.4 Light that passes through a semitransparent cathode may be reflected back to it by internal surfaces

Increasing the thickness of the photoemissive layer reduces the amount of light it transmits and, hence, the amount of internal reflection. This effect is clearly evident in a comparison of the angle-of-incidence related sensitivity variations of the thin S20 and the thick S20R photocathode.

In some cases nearly all internal reflection effects can be eliminated by frosting (sandblasting or etching) the outside of the input window, without any loss of absolute sensitivity (Fig.4.5(b)). This is not effective, however, when the tube is optically coupled to a scintillator or light guide by means of a matching compound such as silicone grease.

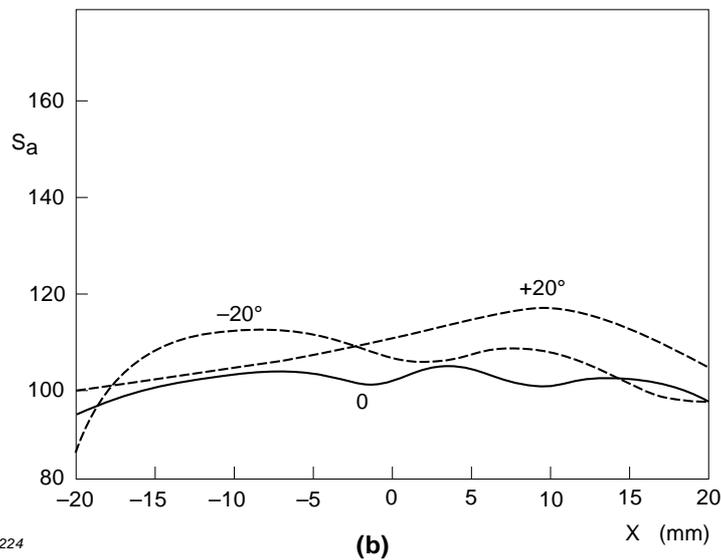
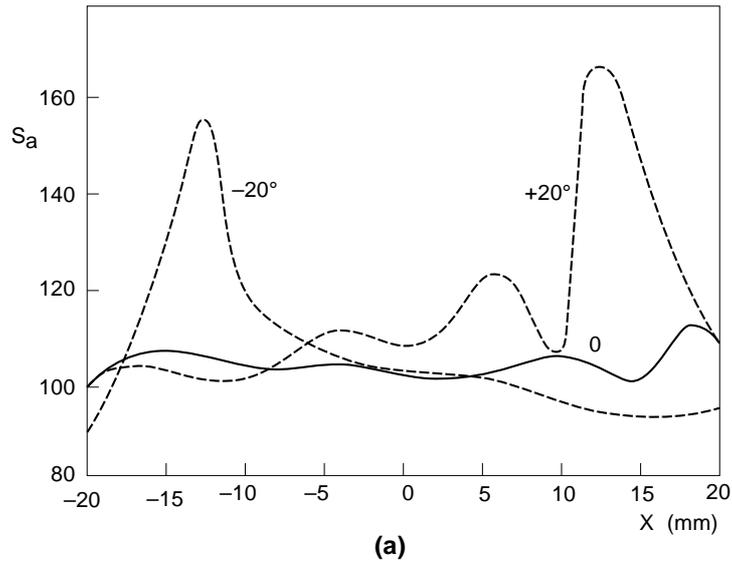


Fig.4.5 Examples of relative anode sensitivity variation at three angles of incidence, measured at $\lambda = 629 \text{ nm}$ along one diameter of a tube with SbNa_2KCs cathode; with (a) polished and (b) frosted input window. x is the distance from the cathode centre

4.4 Time characteristics

4.4.1 Pulse response: determining factors

When a light pulse of infinitesimal duration (delta-function pulse) excites the photomultiplier, the duration of the resulting anode pulse is not infinitesimal. The lengthening is due to electron transit time variations in the cathode/first-dynode space and in successive stages of the electron multiplier. The probability distribution of the variations in each stage is called the *specific response* of the stage. The individual

stages are not, however, statistically independent; there is some correlation between the transit time spreads in successive stages, and this makes it difficult to calculate the pulse response of the tube from the response of the respective stages. However, if the probability distribution of electrons arriving at the anode is assumed to be approximately gaussian with a standard deviation σ_R , then the response $R_\delta(t)$ to a delta-function light pulse is given by

$$R_\delta(t) = \frac{1}{\sigma_R \sqrt{2\pi}} \exp \left\{ - \frac{(t - t_t)^2}{2\sigma_R^2} \right\} \quad (4.4)$$

where t_t is the mean transit time.

To a good approximation, the variance σ_R^2 of the pulse response can be represented by the sum of the variances of the response of each stage:

$$\sigma_R^2 = \sigma_{k,d1}^2 + \sigma_{d1,d2}^2 + \dots + \sigma_{dN-1,dN}^2 + \sigma_{dN,a}^2 \quad (4.5)$$

Generally, it can be assumed that all stages beyond the second have equal response; hence,

$$\sigma_R^2 = \sigma_{k,d1}^2 + \sigma_{d1,d2}^2 + (N - 1) \sigma_{d,d}^2 \quad (4.6)$$

and the response pulse width (FWHM) is

$$t_w = 2.36 \sigma_R \quad (4.7)$$

This is least when the response is due to a single photoelectron, for then there are no transit time differences in the cathode/first-dynode space and the term $\sigma_{k,d1}^2$ in Eq.4.6 vanishes.

Electrodynamic effects in the anode collection space can also affect the pulse response, altering the anode pulse shape and increasing t_w . Among these are the electromagnetic effect of electron movement close to the anode collector grid and the generation of high-frequency currents due to oscillation of electrons about this grid. The often imperfect matching of the anode to the output transmission line can also give rise to oscillations at the pulse trailing edge ('ringing').

Effect of applied voltage. The transit-time fluctuations that affect pulse response have two main causes:

- the initial velocity spread of electrons emitted by different electrodes; the contribution of this cause varies as $1/V_{d,d}$ (where $V_{d,d}$ is the interdynode voltage).

- the difference in transit time due to different points of emission from the same dynode; this contribution varies as $1/\sqrt{V_{d,d}}$.

Thus, the lengthening of the response pulse at each stage depends on a factor between $V_{d,d}^{-1}$ and $V_{d,d}^{-1/2}$. Provided the permissible limits are observed, increasing the voltage per stage is an effective way to improve the pulse response.

Effect of wavelength. The photoelectron energy distribution depends on the incident light wavelength. This, therefore, also affects transit time fluctuations, but only in the electron-optical input system and the first multiplier stage. Moreover, as all stages of the tube contribute about equally to the lengthening of the response pulse, the overall effect of wavelength is only moderate. The existence of other causes of response pulse lengthening, apart from transit time fluctuations, further diminishes the overall significance of wavelength.

4.4.2 *Pulse response: measurement*

Single-electron response. There are two ways to measure the response to emission of a single photoelectron:

- with continuous light so attenuated that the average interval between successive photoelectrons is much greater than the least interval the measuring set-up is able to resolve;
- with light pulses so attenuated that the probability of each pulse giving rise to only one photoelectron is much greater than the probability of its giving rise to more than one.

If fluctuations in the number of photons per light pulse follows a Poisson distribution, so will the number, $n_{k,i}$, of photoelectrons emitted in response to them:

$$P(n_{k,i}) = \frac{(\bar{n}_{k,i})^{n_{k,i}}}{n_{k,i}!} \exp(-\bar{n}_{k,i})$$

Thus the probability of no photoelectron being emitted is $P(0) = \exp(-\bar{n}_{k,i})$; the probability of only one being emitted is $P(1) = \bar{n}_{k,i} \exp(-\bar{n}_{k,i})$; and the probability of more than one being emitted is $P(n_{k,i} > 1) = 1 - P(0) - P(1)$. Since the ratio $P(n_{k,i} > 1)/P(1)$ tends toward $\bar{n}_{k,i}/2$ as $\bar{n}_{k,i}$ tends toward zero, it is evident that it is possible to so attenuate the light pulses that the probability of more than one photoelectron being emitted per pulse is negligible compared with that of only one being emitted. In practice single-electron operation is obtained by so attenuating the light that less than one anode pulse occurs per hundred light pulses. The ratio $P(n_{k,i} > 1)/P(1)$ is then less than 5×10^{-3} .

Multi-electron response is measured with subnanosecond pulses – as from fast LEDs, semiconductor lasers, spark sources, or Cherenkov sources – which are short compared with the several nanosecond pulse-widths t_w of present-day photomultipliers.

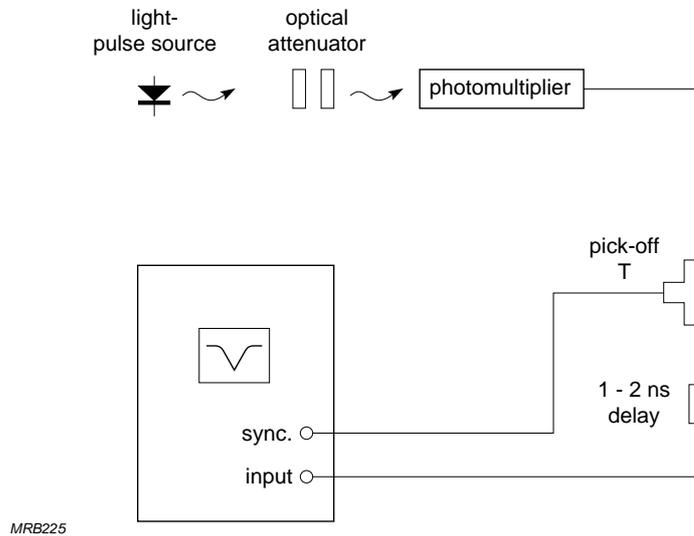


Fig.4.6 Set-up for measuring pulse response

Figure 4.6 shows the usual set-up for measuring pulse response. The oscilloscope may be synchronized by a signal with 1 or 2 ns lead time taken from one of the last dynodes, or by one taken from a pick-off T that precedes a 1 or 2 ns delay as shown. The response $R^*(t)$ measured on the oscilloscope is the convolution of the illumination function $L(t)$, the pulse response $R_\delta(t)$ of the photomultiplier, and the pulse response $S(t)$ of the measuring set-up (transmission line, delay line, oscilloscope):

$$R^*(t) = L(t) * R_\delta(t) * S(t) \quad (4.8)$$

When the measurement is made under single-electron conditions, $L(t)$ approximates a delta function and Eq.4.8 simplifies to

$$R^*(t) = R_\delta(t) * S(t)$$

The pulse response of the measuring set-up must be accurately known; estimates may introduce significant error. It is preferable to use a set-up whose response is known to approximate a delta function. Then, in the single-electron case

$$R^*(t) \approx R_\delta(t)$$

and in the multi-electron case

$$R^*(t) \approx L(t) * R_{\delta}(t) \quad (4.9)$$

$R^*(t)$ then has a standard deviation

$$\sigma_R^* = \sqrt{\sigma_L^2 + \sigma_R^2} \quad (4.10)$$

where σ_L and σ_R are the standard deviations of the illumination function $L(t)$ and the photomultiplier response $R_{\delta}(t)$.

4.4.3 Transit time differences

Transit time differences are due mainly to differences in electron path lengths but also, in part, to differences in electric field strength. The largest originate in the cathode/first-dynode space, where path-length differences are greatest; the electron multiplier, where path lengths are more nearly equal, does not make so important a contribution.

Effect of applied voltage. Like transit-time fluctuations, transit-time differences vary inversely as the square root of inter-electrode voltage. As it is the cathode/first-dynode space that contributes most, it is here that the most can be gained by applying the maximum permissible voltage.

Effect of wavelength. Incident light wavelength does not significantly affect transit time differences. In fact, even its effect on overall transit time is small. At wavelengths from 250 nm to 900 nm the initial energy of the photoelectrons is only a fraction of an electron-volt, which is negligible compared with the several hundred electron-volts they may gain from the electric field between the cathode and first dynode. The observed variation in transit time as a function of wavelength amounts to only about 1 ps/nm.

Measurement of transit-time differences. Transit time is determined by measuring the interval between signals known to be synchronous with light pulses at the cathode and the resulting anode pulses. Transit-time differences can be measured, and mapped, by focusing the light pulses on different parts of the cathode and noting the corresponding transit times. Figure 4.7 is an example of the transit-time differences measured with reference to perpendicular cathode diameters of a fast-response photomultiplier. Figure 4.8 shows a set-up for measuring both transit-time differences and transit-time spread.

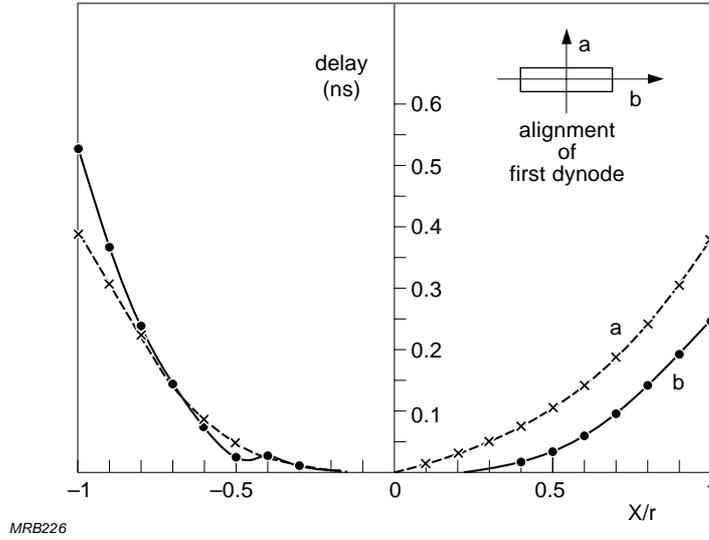


Fig.4.7 Transit-time differences of a fast-response photomultiplier as functions of distance from the cathode centre, measured along perpendicular diameters

4.4.4 Transit-time spread

The pulse-to-pulse fluctuation of transit time (jitter) limits the accuracy of photomultiplier-based time measurements. For single-electron operation it has been suggested that the variance σ_{tt}^2 of the total transit-time fluctuation is

$$\sigma_{tt}^2 = \sigma_{k,d1}^2 + \sigma_m^2 \quad (4.11)$$

where $\sigma_{k,d1}^2$ and σ_m^2 are the variances of the transit-time fluctuations in the cathode/first-dynode space and the electron multiplier. If all stages of the multiplier except the first are identical,

$$\sigma_m^2 = \frac{\sigma_{d1}^2}{g_1} (1 + v_{g1}) + \frac{\sigma_{dd}^2}{g_1(g - 1)} (1 + v_g) \quad (4.12)$$

where g_1 is the gain of the first stage, g the average gain of subsequent stages, v_{g1} and v_g are the relative variances of those gains, and σ_{d1}^2 and σ_{dd}^2 are respectively the variances of the transit-time fluctuations in the first and subsequent stages. Equations 4.11 and 4.12 show that the stages that predominate in determining the transit-time spread are the cathode/first-dynode space and the first multiplier stage.

Transit-time fluctuations in the cathode/first-dynode space have two components: a *chromatic* one due to the spread of photoelectron initial velocities, and a *geometric* one due to path-length differences.

The *chromatic component* has a variance σ_{vi}^2 which can be resolved into components σ_{vn}^2 and σ_{vt}^2 corresponding to the normal and tangential components of initial velocity:

$$\sigma_{vi}^2 = \sigma_{vn}^2 + \sigma_{vt}^2 \quad (4.13)$$

If the probability distribution of the initial velocities is known, σ_{vn} and σ_{vt} can be found from Eq.1.1. For fast-response tubes σ_{vi} is between 50 and 500 picoseconds, depending to some extent on incident-light wavelength and applied voltage.

The *geometric component* has a variance σ_{cb}^2 which can be calculated if the relation between photoelectron points of origin and transit-time differences can be formulated. If, for instance, the transit time varies roughly as the square of the distance of the point of origin from the cathode centre, then $\sigma_{cb} \approx 0.3 \Delta t_{tmax}$, where Δt_{tmax} is the transit-time increment due to origination at the maximum distance from the centre.

For the total contribution of the electron-optical input system

$$\sigma_{k,d1}^2 = \sigma_{vn}^2 + \sigma_{vt}^2 + \sigma_{cb}^2 \quad (4.14)$$

For fast-response photomultipliers, $\sigma_{k,d1}$ is usually between 150 and 350 ps, depending on the cathode/first-dynode voltage and the incident-light wavelength.

The contribution of the electron multiplier, σ_m , is mainly from the first stage, for two reasons:

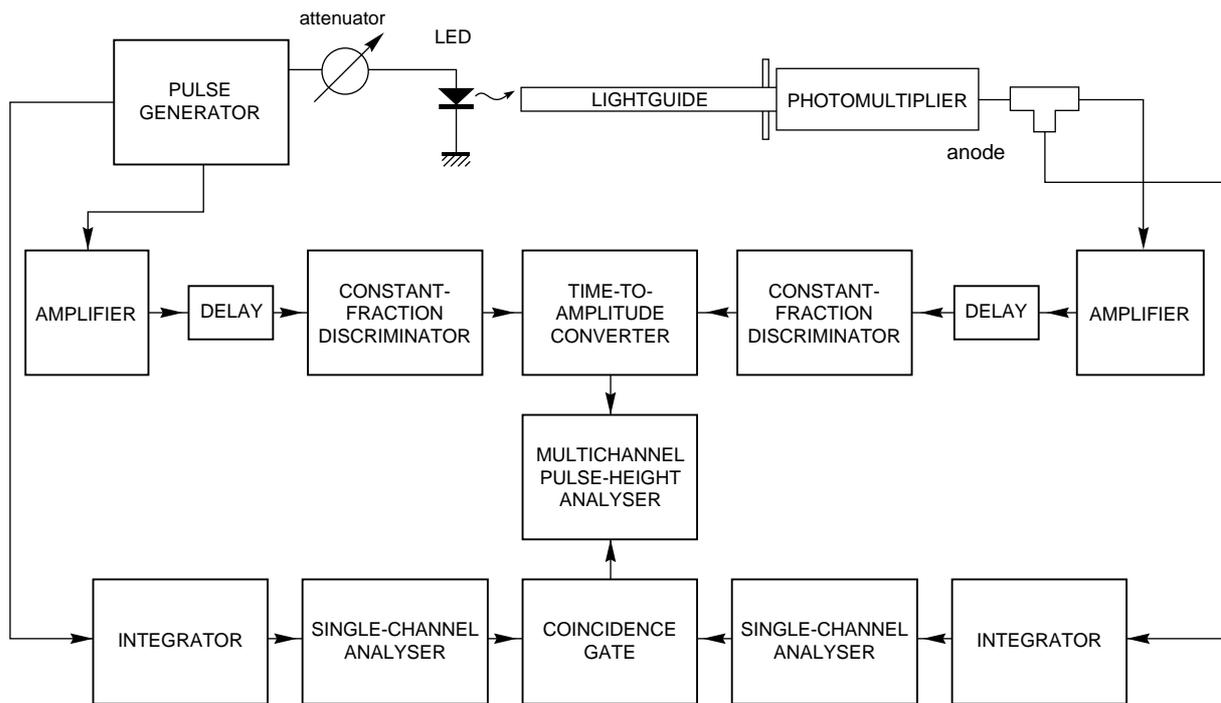
- the number of secondary electrons there is smaller than in subsequent stages, so the standard deviation tends to be larger;
- the role of the first stage as a coupling zone between the electron-optical input system and the iterative part of the multiplier also tends to increase the standard deviation.

Again, there is a chromatic component and a geometric one. The chromatic component is due to the spread of secondary electron initial velocities. The geometric one, which may predominate and can be an important factor in the effect of wavelength on overall transit-time spread, is due to the scatter of electrons on the first dynode. It thus depends on the primary-electron velocity spread. In fast-response tubes σ_m is between about 150 ps and 250 ps, depending on voltage and wavelength.

Effect of applied voltage. The spread of initial velocities varies as $1/V_{d,d}$, and the spread of transit-time differences as $1/\sqrt{V_{d,d}}$. Furthermore, the stage gain g in Eq.4.12 varies as a power of $V_{d,d}$ between 0.65 and 0.75. Hence the overall transit time spread varies as $V_{d,d}^{-n}$, where n is between 0.5 and 1.

Effect of wavelength. Wavelength affects the energy distribution of the photoelectrons and, hence, their initial velocity spread, which increases as wavelength decreases. Although wavelength changes have little effect on the geometric component in the electron-optical input system, they do affect the chromatic component there, σ_{vj} , and the geometric component σ_{cb} in the first multiplier stage. The overall effect depends on the relative importance of these contributory effects. With a bialkali cathode, a wavelength increase from 400 nm to 560 nm decreases the transit-time spread about 40%.

Measurement of transit-time spread. Transit-time spread is measured by recording the intervals between a clocked series of light pulses and the corresponding series of anode pulses (Fig.4.8). The transit time probability distribution depends on the mean number of photoelectrons, $\bar{n}_{k,i}$, emitted per light pulse, the variance being greatest for single-electron operation. The measured probability distribution can also depend to some extent on the statistics of photon emission because the timing reference chosen is a light pulse.



MRB220

Fig.4.8 Set-up for measuring transit-time differences and transit-time spread

Single-electron operation. Let $L(t)$ be the probability distribution of the instants of photon emission, such that

$$\int_0^{\infty} L(t) dt = 1$$

and $R(t)$ the probability distribution of the occurrence of the corresponding anode pulses referred to the instants of photoelectron emission. Then the measured probability distribution of the transit time will be

$$R^*(t) = L(t) * R(t) \tag{4.15}$$

The single-electron time resolution, designated R_t^* , is the FWHM of the $R^*(t)$ curve. Provided the light pulses with which it is measured are very short in comparison with $R(t)$, R_t^* is an accurate measure of the transit time spread. Figure 4.9 is a single-electron time resolution curve of a fast-response photomultiplier.

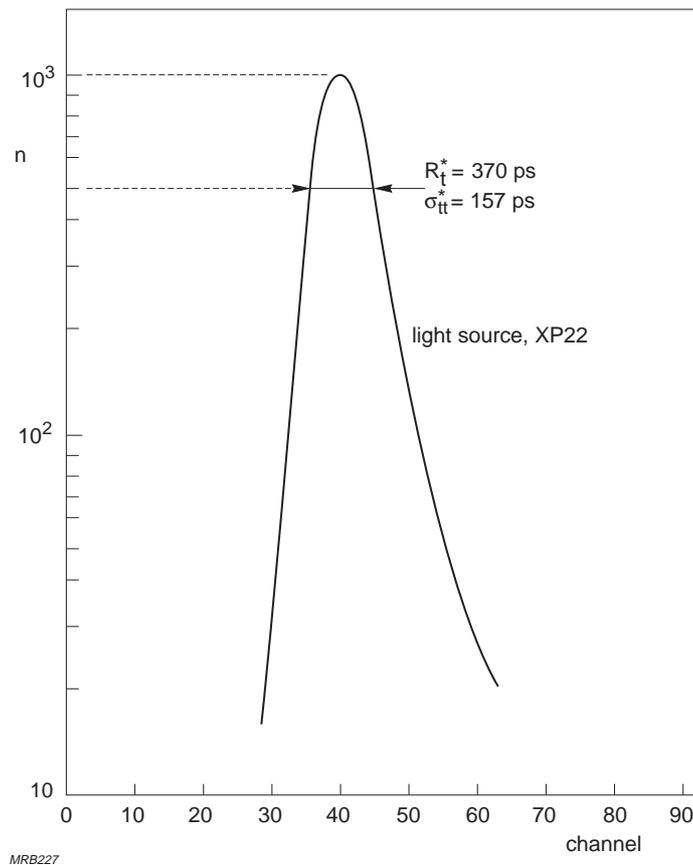


Fig.4.9 Time resolution of a fast-response photomultiplier, for single-electron pulses originating from a single point on the cathode at an illumination wavelength of 560 nm, measured according to the method of Fig.4.8. Vertical axis, number of pulses per channel (arbitrary scale); horizontal axis, channel number (about 40 ps per channel)

If $L(t)$ and $R(t)$ are approximately gaussian, with variance σ_L^2 and σ_{tt}^2 , then $R^*(t)$ will also be gaussian, with a variance

$$\sigma_{tt}^{*2} = \sigma_L^2 + \sigma_{tt}^2 \tag{4.16}$$

and FWHM

$$R_t^* = 2.36 \sigma_{tt}^* \quad (4.17)$$

Multi-electron operation. Provided L(t) and R(t) are gaussian,

$$\sigma_{tt}^{*2} = \frac{\sigma_L^2 + \sigma_{tt}^2}{\bar{n}_{k,i}} \quad (4.18)$$

For fast-response tubes in single-electron operation at an illumination wavelength of 400 nm, R_t^* is usually less than 1 ns; in operation with an average of 10 photoelectrons per light pulse this is divided by $\sqrt{10}$ and becomes less than 320 ps.

4.4.5 Frequency response

The frequency response G(f) is important in applications involving modulated light. Its upper limit is mainly due to statistical effects and imperfect matching of the output to the external circuit. The frequency response can be derived from the pulse response $R_\delta(t)$ via the Fourier transform

$$G(f) = \int_{-\infty}^{\infty} R_\delta(t) \exp(-j2\pi ft) dt \quad (4.19)$$

The narrower the pulse response, the higher the cut-off frequency. If the pulse response is gaussian with a variance σ_R^2 ,

$$R_\delta(t) = \frac{1}{\sigma_R \sqrt{2\pi}} \exp \left\{ - \frac{(t - t_t)^2}{2 \sigma_R^2} \right\} \quad (4.20)$$

and, from Eq.4.19,

$$G(f) = \exp \left[- 2(\pi f \sigma_R)^2 - j2\pi f t_t \right] \quad (4.21)$$

the absolute value of which is

$$| G(f) | = \exp \left[- 2(\pi f \sigma_R)^2 \right] \quad (4.22)$$

This corresponds to a 3 dB bandwidth

$$B_{3dB} = \frac{0.133}{\sigma_R} \quad (4.23)$$

The bandwidth can also be expressed as a function of the step-response rise time $t_{r,\epsilon}$. If the pulse response is gaussian, with a variance σ_R^2 (and FWHM $t_w = 2.36 \sigma_R$), the rise time in response to a unit step $\epsilon(t)$ is (§A4.1.2)

$$t_{r,\varepsilon} \approx 1.11 t_w = 1.11 \times 2.36 \sigma_R \quad (4.24)$$

whence, from Eq.4.23,

$$B_{3dB} \approx \frac{0.35}{t_{r,\varepsilon}} \quad (4.25)$$

For a photomultiplier with $t_w \approx 3$ ns, $B_{3dB} \approx 105$ MHz.

(Some authors have suggested other ways of describing the step response of a photomultiplier yielding a comparable 3 dB bandwidth.)

Note. In accordance with IEC standard, Publications 306-4 (1971) and 462 (1974), the pulse response is specified in terms of delta-function rise time (t_r) and FWHM (t_w); this is not to be confused with the step-response rise time $t_{r,\varepsilon}$ used for calculating the bandwidth.

4.5 Linearity

The degree of proportionality between the number of electrons collected at the anode and the number of incident photons is called *charge linearity*. The degree of proportionality between incident flux and anode current is called *current linearity*; in this relationship therefore, *time* is an additional parameter. Limits on both charge and current linearity are set by internal and external factors.

4.5.1 External factors affecting linearity

Power supply. Changes in interelectrode voltages affect gain (Fig.4.10) by influencing the dynode secondary emission factors and the electron trajectories.

Divider current. When the electrode voltages are derived from a resistive divider across a stabilized power supply, the anode current I_a tends to lessen the potential between the last dynode and the anode. This upsets the voltage distribution throughout the divider and causes an increase of gain comparable to what would be caused by increasing the high voltage by the same amount.

The current through an iterative divider ($R = R_1 = R_2 \dots = R_N$) when there is no anode current is

$$I_p = \frac{V_{ht}}{(N + 1)R} \quad (4.27)$$

where V_{ht} is the high voltage and R is the common value of the resistors. When there is an anode current, the currents in the resistors are altered as shown in Fig.4.11. To keep V_{ht} constant, I_p must assume a new value

$$I_p' = I_p + \Delta I_p = I_p + \frac{1}{N + 1} \sum_{i=0}^N I_i \quad (4.28)$$

From Eq.2.7 the gain at divider current I_p' is

$$G = KV_{ht}^{N\alpha} = KR^{N\alpha} \prod_{i=0}^{N-1} (I_p + \Delta I_p - I_i)^\alpha \quad (4.29)$$

and the gain when the anode current is zero is

$$G_m = KR^{N\alpha} I_p^{N\alpha}$$

Thus, the ratio of the gain at divider current I_p to that when the anode current is zero is

$$\frac{G}{G_m} = \prod_{i=0}^{N-1} \left(1 + \frac{\Delta I_p - I_i}{I_p} \right)^\alpha \quad (4.30)$$

or, if we neglect terms higher than the first order

$$\frac{G}{G_m} = \left[1 + \sum_{i=0}^{N-1} \left(\frac{\Delta I_p}{I_p} - \frac{I_i}{I_p} \right) \right]^\alpha$$

Substituting the value of ΔI_p from Eq.4.28 gives

$$\frac{G}{G_m} = \left[1 + \frac{I_N}{I_p} \left(\frac{N}{N+1} - \frac{1}{N+1} \sum_{i=0}^{N-1} \frac{I_i}{I_N} \right) \right]^\alpha \quad (4.31)$$

As the quantity under the summation is always less than unity, the quantity between the inner parentheses is positive and the gain G is an *increasing* function of the ratio $I_N/I_p = I_a/I_p$.

Throughout the range of variation $I_N \leq I_p$ the ratio I_N/I_p is nearly independent of the gain of each stage; so, by setting

$$\frac{I_i}{I_{i-1}} = g_i = g$$

and noting that $g^N \gg 1$, Eq.1.31 can be simplified to

$$\frac{G}{G_m} = \left\{ 1 + \frac{I_a}{I_p} \left(\frac{N}{N+1} - \frac{1}{(N+1)(1-g)} \right) \right\}^\alpha$$

Or, for large values of N,

$$\frac{\Delta G}{G} = \frac{G - G_m}{G} \approx \alpha \frac{N}{N+1} \frac{I_a}{I_p} \quad (4.32)$$

Equation 4.32 expresses the relative variation of the gain as a function of the ratio of the anode current I_a to the nominal divider current at $I_a = 0$, provided the decrease of voltage across the terminals of the last stage does not impair collection efficiency. Under these conditions (and provided no internal factors work against it), the ratio $\Delta G/G$ has the same sign as I_a/I_p : an increase of I_a results in an increase of gain (Fig.4.12). This increase (or overlinearity) as a function of I_a is largely independent of N. For $\alpha = 0.7$, $N = 10$, and $I_a/I_p = 0.1$, it amounts to about 7%.

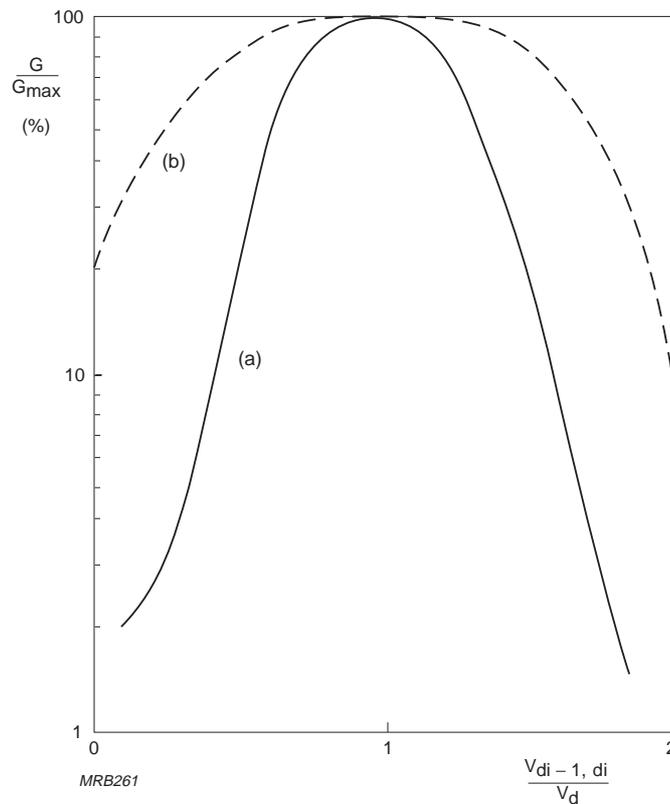


Fig.4.10 Gain variation as a function of the voltage between successive dynodes, for a tube with (a) linear focusing dynodes, (b) venetian-blind dynodes. V_d = nominal interdynode voltage

When the ratio I_a/I_p approaches unity, Eq.4.32 no longer holds. The voltage drop in the last stage, which increases with I_a , becomes too great and collection efficiency declines rapidly, leading to an abrupt decrease of gain. Internal factors may accentuate this effect (§4.5.2). To maintain linear operation in resistive voltage dividers, a good rule is to ensure that the ratio $I_a/I_p \leq 0.01$ (see Fig.4.12 and §5.2.2).

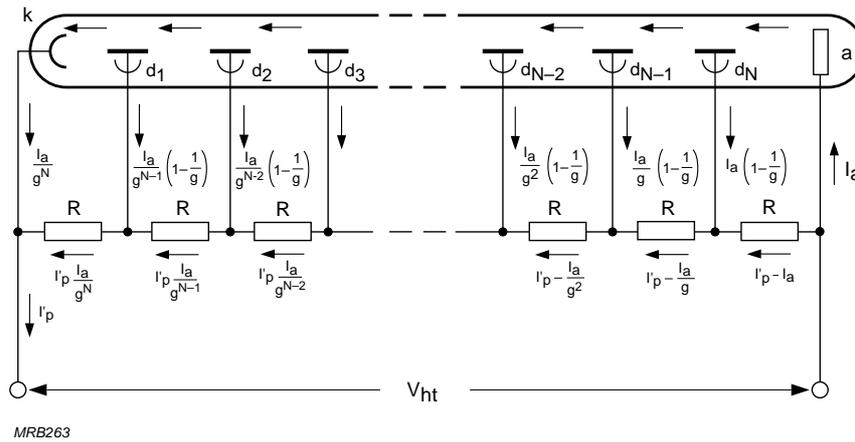


Fig.4.11 Current distribution in an iterative voltage divider

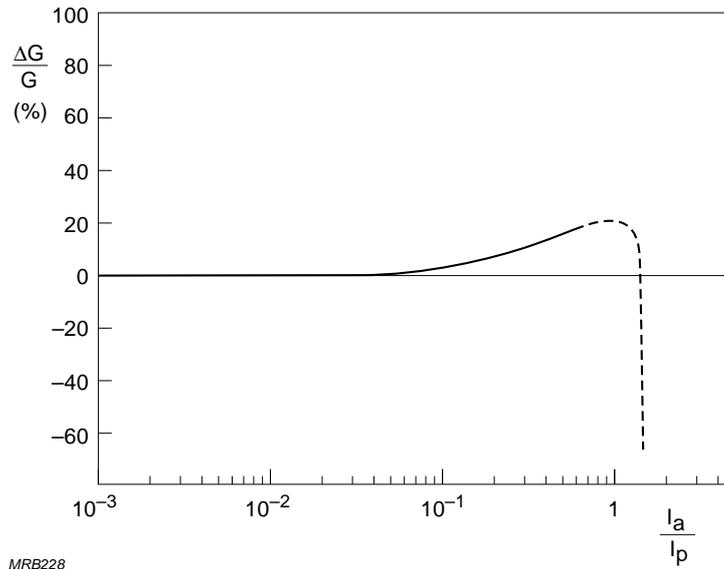


Fig.4.12 Gain variation (overlinearity) as a function of the ratio I_a/I_p

The maximum value of the ratio $\Delta G/G$ depends mainly on the voltage across the tube and how the drop between anode and last dynode affects the voltage distribution among the first stages.

To ensure that the ratio $I_a/I_p \leq 0.01$, I_p may be increased, but this solution is limited by power dissipation which may not exceed a few watts.

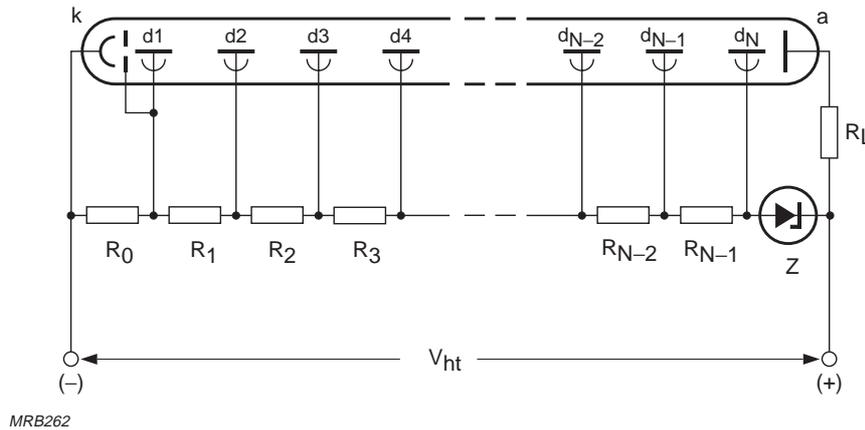


Fig.4.13 Voltage divider with zener diode for improved linearity

The dependence of gain on anode current can be lessened by substituting a zener diode for the resistance R_N and possible for R_{N-1} (Fig.4.13); the overlinearity is then eliminated and the anode current at which the rapid decrease of gain occurs becomes about ten times higher.

A voltage divider composed solely of zener diodes is never used: firstly, because it makes it impossible to adjust the gain by adjusting the supply voltage; and secondly, because it provides no current-limiting action to guard against the consequences of accidental overexposure of the cathode. In high-current applications, therefore, active dividers (employing transistors) are often preferred.

Reservoir capacitors. When the anode current can reach high values for only a small fraction of the time (short-pulse operation), it is preferable to connect reservoir (or decoupling) capacitors to the dynodes (§5.2.5). The charge stored by the capacitors must be sufficiently large compared with that supplied by each dynode when pulses pass through the tube so that the dynode potentials will not vary by more than one or two volts. Calculation of the required capacitance values differs according to whether the decoupling is parallel or series; in the latter case voltage variations are

cumulative. When space or insulation considerations limit the size of the capacitors that can be used, their values must be calculated with especial care, as inadequate decoupling affects charge linearity rather than current linearity, which could cause misleading results.

Damping resistors. When a photomultiplier is operating in the pulse mode, a high-frequency spurious oscillation superimposed on the anode pulses may be observed, even with pulses as wide as a few hundred nanoseconds. This oscillation, which affects the linearity characteristic of the tube, usually producing an overlinearity, may appear abruptly when the anode current exceeds a certain level. One way of overcoming this effect is to connect a 50 Ω non-inductive resistor in series with each of the last two or three dynodes. Fast-response photomultipliers with plastic bases have such resistors built-in. For other types they must be wired into the socket, between the base and the decoupling capacitors.

Anode load. The voltage developed across the anode load subtracts from the last-dynode to anode voltage and, if it is not negligible compared with that, may affect linearity. As the load voltage rarely exceeds a few volts, however, this is seldom the case.

4.5.2 *Internal factors affecting linearity*

Space charge. At high currents, space charge can influence the electron trajectories, causing collection losses; at still higher currents it can cause some electrons to return to the surfaces from which they originate. The condition resembles that of a space-charge limited diode with parallel-plane electrodes, for which the relation between current density J_s (in A/cm²) and electric field is given by the Child-Langmuir equation

$$J_s = 2.2 \times 10^{-6} \frac{V^{3/2}}{d^2} \quad (4.33)$$

where V is the interelectrode voltage in volts, and d the interelectrode distance in centimetres.

The current density is normally highest between the last dynode and the anode. To ensure a high field there, the anode is positioned close to the surface of the last dynode and made in the form of a grid through which the electrons pass on their way from the next-to-last dynode. Then, it is the field between the next-to-last dynode and the anode, which is 3 to 5 times lower, that sets the limit for current linearity in most photomultipliers.

That limit can be raised by using a progressive instead of an equal voltage distribution in the last stage (§5.2.1), so as to raise the voltage between the last two stages to as much as 300 V or more. To maintain correct focusing between dynodes without unduly increasing the gain, the interelectrode voltages are progressively decreased in the anode to cathode direction so that the nominal value applies at the terminals of the first stages. For tubes with focusing dynodes, the data sheets give, in addition to the conventional voltage distribution, one or two examples of recommended progressive distributions. Using these, the maximum pulse current in linear operation can be increased from 10 – 50 mA to 100 – 300 mA. For some special tubes, linear pulse operation can be obtained at currents of more than 5 A.

For tubes with venetian-blind or box-and-grid dynodes, the maximum pulse currents for linear operation are smaller (10 – 50 mA) because of the very low electric fields between all dynodes other than the last.

For most tubes, the current linearity limit due to space charge varies as V_{ht}^n , where n is between 2 and 3. This is merely approximate, but when the limit at one voltage is known from the published data it gives a practical indication of the limit at another voltage, especially if the onset of saturation is progressive. If linearity is not important, the maximum anode current that can be obtained before saturation is several times greater than the maximum for linear operation.

The space charge phenomena that limit current linearity last for times comparable to the transit times between dynodes, that is, 1 to 2 ns. Even when linearity errors are severe, there is no charge accumulation and the errors are strictly related to the electron current passing between the last dynodes.

Current linearity is important when pulses are wide compared with the pulse response of the tube; when they are of the same order as the pulse response, it is no longer relevant. The significant parameter then is charge linearity. Depending on the shape of the pulses, higher peak anode currents, can be obtained under short-pulse conditions while still maintaining good charge linearity.

Data sheets specify only the current linearity limit, not the charge linearity limit, and for a worst-case situation with anode pulses about 100 ns wide.

Cathode resistivity. The electron-optical input system is designed on the assumption that the cathode is an equipotential surface. Any departure from that condition is likely to alter the electron trajectories and affect the collection efficiency of the first dynode. This is what happens, at least in the case of semitransparent cathodes having

no underlying conductive layer, when the cathode current is too large in relation to the surface resistivity.

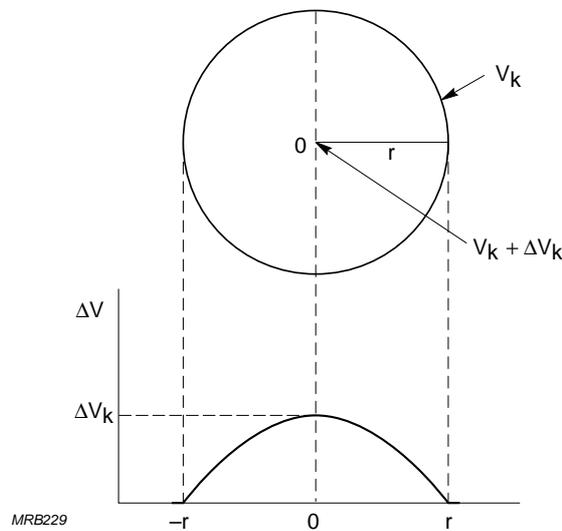


Fig.4.14 Potential distribution due to cathode resistivity; r is the cathode radius

Consider a circular cathode of uniform sensitivity, uniformly illuminated and emitting a total current I_k (Fig.4.14). Let R_{\square} be its surface resistivity (the bulk resistivity divided by the thickness); the potential difference between the centre and the edge is then

$$\Delta V_k = \frac{R_{\square} I_k}{4 \pi} \quad (4.34)$$

If it exceeds a few volts, this potential difference increases the input-system convergence and causes loss of electrons emitted from the cathode edge. More complicated effects occur when only small areas of the cathode are illuminated. These lead to a dynamic variation of gain as a function of cathode current; in other words, to linearity errors. However, for the cathode current normally encountered, such phenomena are practically significant only with bialkali cathodes. For these, the surface resistivity is about $10^{10} \Omega$ at ambient temperature, decreasing slightly with illumination but increasing rapidly as temperature decreases. A tube with a 45 mm diameter bialkali cathode exhibits a non-linearity of a few percent at a mean cathode current of about 10 nA at ambient temperature; at -30°C the same non-linearity occurs at a current of only 0.1 nA. In tubes with larger cathode diameters, the currents at which comparable non-linearity occurs are even lower. In tubes with S11 and SbRbCs cathodes, comparable non-linearity at ambient temperature occurs at currents respectively about 100 times and 10 times higher.

Fortunately, the distributed capacitance of the cathode (about 1 pF) is sufficient to store a charge of about 10^{-12} C. At a gain of 10^5 , this corresponds to an anode pulse of 100 mA amplitude and 1 μ s duration; cases in which cathode resistivity actually presents a problem are therefore fairly uncommon.

Gain drift. Gain may undergo more or less reversible variations when the mean anode current varies. Although this too constitutes a linearity error, by convention it is treated as an instability (§4.6).

4.5.3 Linearity measurement

Many methods of linearity measurement have been developed but all are limited to an accuracy no better than 2%. Two types of gain drift may interfere with the measurement:

- long-term, time-dependent drift (§4.6.1)
- short-term shift due to changes of illumination (§4.6.2).

To avoid these, the measurement must be made quickly and with a mean current not exceeding a few microamperes. The measurement should result in determining the anode current at which space charge limiting starts to become evident, avoiding all other causes of linearity limiting. The methods described below are for measuring either current or charge linearity.

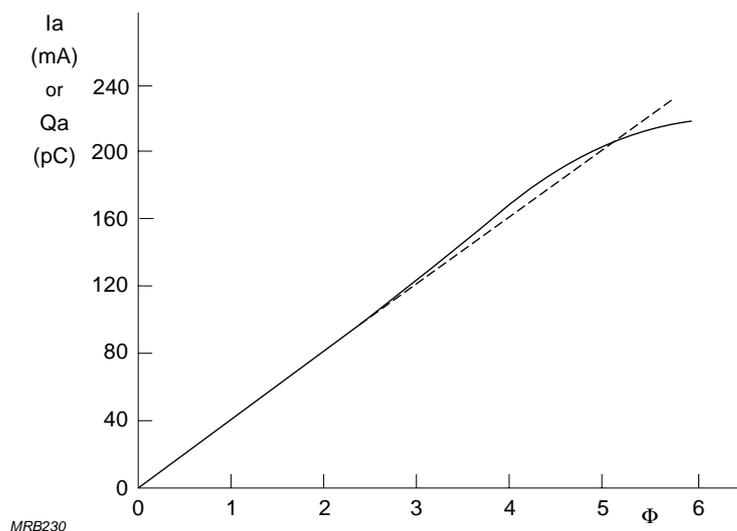
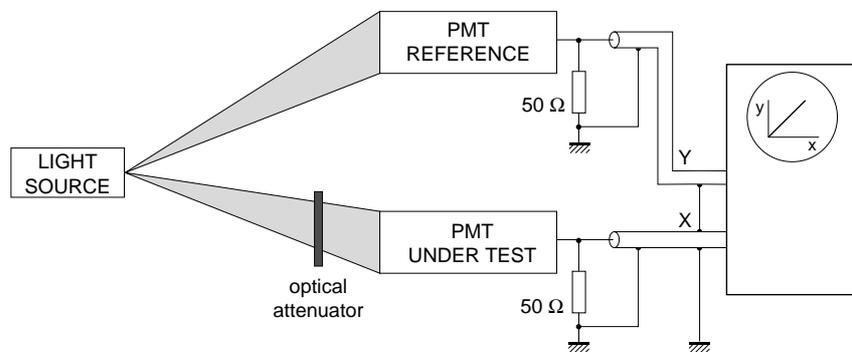


Fig.4.15 Typical current or charge linearity characteristics of a photomultiplier operating from a supply with type B voltage division (flux Φ in arbitrary units)

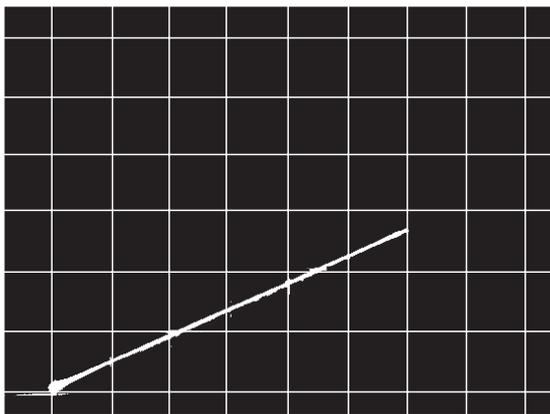
Figure 4.15 shows a typical linearity curve, in which a slight overlinearity appears before saturation. Such overlinearity is often observed with voltage dividers designed for delaying the onset of saturation at high current levels (§5.2.1). It can be corrected by adjusting the voltages of the stages immediately preceding the last, but at the cost of lowering the current threshold beyond which saturation occurs (§5.2.3).

XY method. This method makes use of an oscilloscope having identical X and Y deflection factors (Fig.4.16). The anode pulse of the photomultiplier under test deflects the beam parallel to one axis, and the anode pulse of a reference photomultiplier operating in its known linear region deflects it parallel to the other. The measurement is therefore one of instantaneous current linearity. Both tubes are excited simultaneously by light pulses of a few hundred nanoseconds.

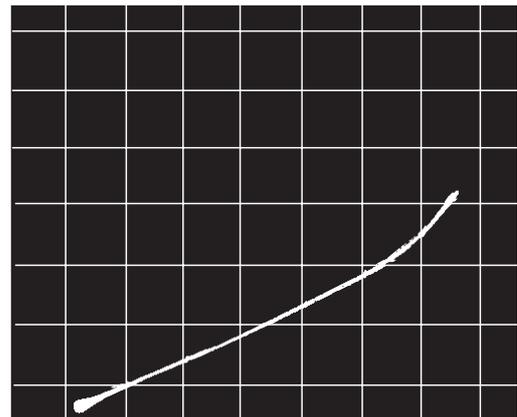


MRB231

Fig.4.16 Set-up for the XY-method of measuring linearity



(a)



(b)

Fig.4.17 Oscillogram obtained by the method of Fig.4.16 showing (a) linear and (b) non-linear response. Scales, 50 mA/div

At each level of pulse current corresponding to the setting of an optical attenuator, an oscilloscope trace is obtained. This may have, for example, the shape of Fig.4.17(b), with a linear part and a curved part; the linearity error is the percentage by which the curved part of the trace departs from the line of the straight part. This is a measurement of *integral linearity*. In Fig.4.17(b), the tube under test is linear within 5% up to 200 mA and within 10% up to 300 mA.

Dual pulse method. The flux from a practically monochromatic source (for example a LED) can be calibrated with great accuracy. Two such light sources controlled by separate generators send light pulses to the photomultiplier under test. The width of the pulses is about 100 ns and their amplitudes are in a fixed ratio, for example 2:1. Provided the timing of the two generators is known, the pulses from each can be recognized by the processing electronics. The pulse repetition frequency must decrease with increasing pulse height to ensure that the mean anode current of the photomultiplier remains constant (at a value $< 1 \mu\text{A}$) throughout the whole range of pulse amplitude variation.

The method consists in increasing the light-pulse amplitudes, while maintaining their ratio (for example, by the use of neutral filters), and monitoring the height or charge ratio of the current pulses supplied by the photomultiplier. In this way, a pulse-current value can be determined beyond which the height or charge ratio differs by a given amount from the initial ratio. This is a *differential linearity* measurement. An advantage of it is that it does not require the use of a reference photomultiplier.

The ratio of the photomultiplier pulses can be measured with a multichannel pulse-height analyser, or with an oscilloscope, but not so accurately.

Composite radiation method. This method consists in exposing a scintillator-photomultiplier combination to radiation from a composite γ -source emitting several known energies between a few hundred keV and a few MeV (Fig.4.18). The relative heights of the integrated current pulses of the photomultiplier are then measured with a multichannel pulse-height analyser. When the mean amplitudes of the pulses corresponding to each γ line are plotted against energy, the resulting curve has an initial linear part followed by a curved part indicative of either overlinearity or saturation. The anode pulse height beyond which the curve departs from straightness by a given amount is a measure of the integral linearity of the scintillator-photomultiplier combination, and from this it is possible to determine the linearity of the photomultiplier alone. (As inorganic scintillators are not linear at low energies, pulses corresponding to γ -energies of less than a few hundred keV should be disregarded.)

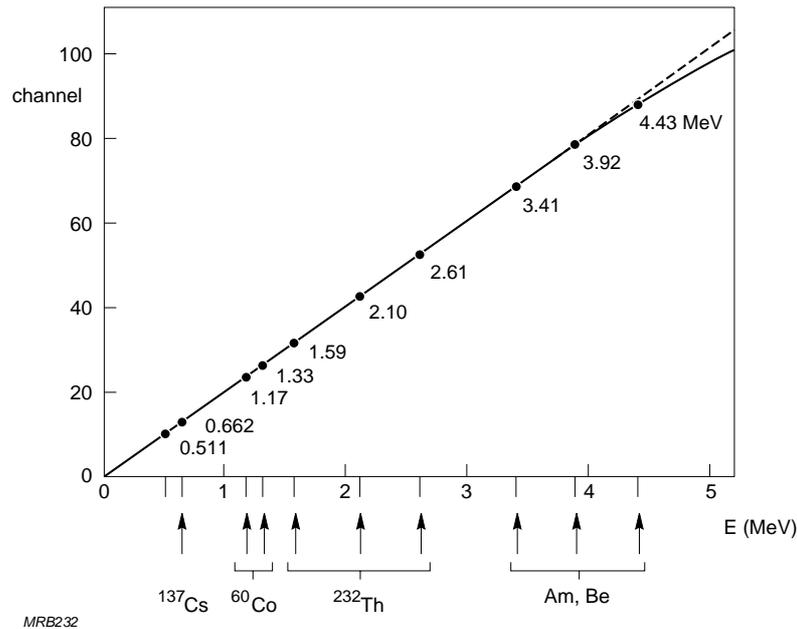


Fig.4.18 Distribution of γ -ray energies emitted by a source consisting of ^{60}Co , ^{137}Cs , ^{232}Th and AmBe as used in measuring photomultiplier linearity. Vertical scale, channel number of multichannel pulse-height analyser

The composite radiation method is especially applicable to high-gain tubes: at a gain of 10^7 an energy of 1 MeV absorbed in a NaI(Tl) scintillator gives an anode pulse of about 30 mA. A major drawback of the method is the near-impossibility of adjusting the amplitude of the light pulses. The amplitude of the anode pulses can therefore be adjusted only by varying the gain. Since this means varying the high voltage, it can itself affect the linearity to be measured.

Method using bursts of three pulses. In this method, a special pulse generator, giving burst of three calibrated pulses of increasing amplitude is used to drive one LED. The light reaches the cathode, passing through a neutral optical filter giving a five-fold attenuation. The anode pulses feed a multichannel analyzer (Fig.4.19 (a)). and the registered channel numbers p_1 , p_2 , and p_3 become p'_1 , p'_2 and p'_3 when the filter is withdrawn and an five-fold *electrical* attenuator is placed before the multichannel analyser. An on-line calculator can be used to calculate the ratios $a = p_2/p_1$ and $a' = p'_2/p'_1$ and to calculate the linearity deviation $(a' - a)/a$. In the same way, the ratios $b = p_3/p_1$ and $b' = p'_3/p'_1$ give the linearity deviation $(b' - b)/b$.

A typical test setup could, for example, comprise the following:

- burst of three pulses of 50 ns giving, when the filter is ON, three anode pulses of 6, 20 and 30 mA peak value at the operating gain
- time between each pulse of 100 μsec
- burst frequency of a few kilohertz.

When the filter is OFF, the anode pulses should be 30, 100 and 150 mA.

Fig.4.19(b) shows the results of this test. The linearity deviation is then calculated between the channel numbers corresponding to the ideal values 100 – 30 mA on the one hand, and 150 – 30 mA on the other hand, with respect to the tube linearity at very low level. One advantage of this method (which measures the differential linearity) is that the result is not affected by the possible gain shift due to the change of mean anode current that occurs when the filter is withdrawn.

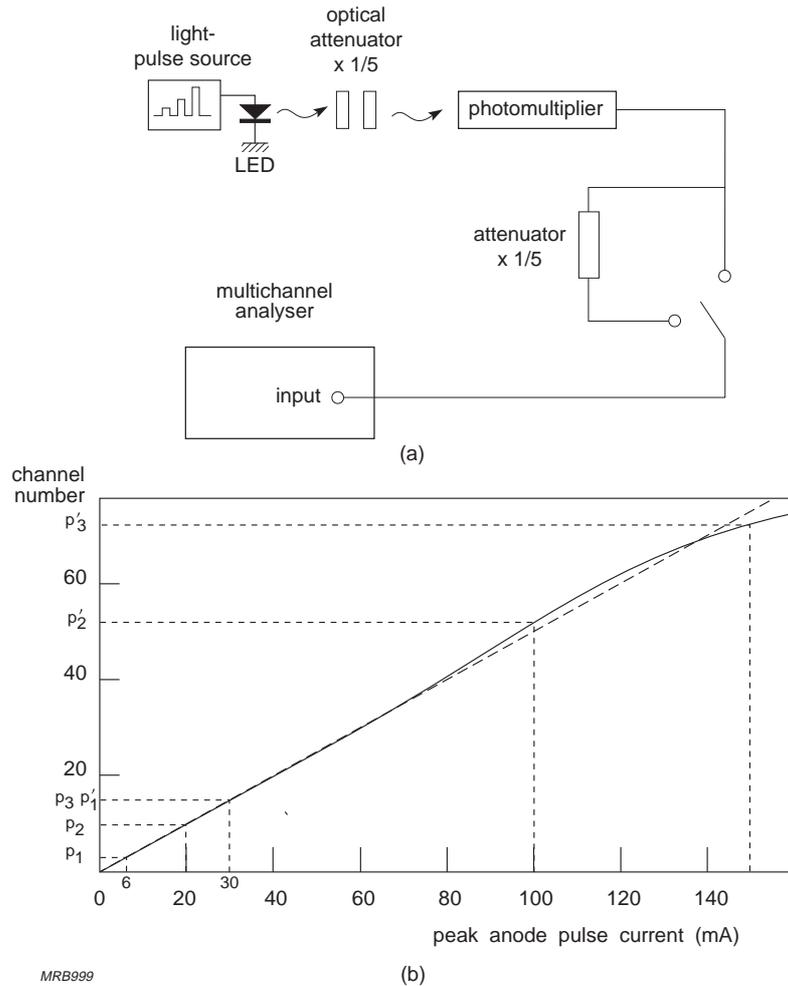


Fig.4.19 Pulse-linearity test; (a) block diagram of the 'three-pulses method'; (b) example of an experimental linearity curve obtained with the three-pulses method (channel number in arbitrary units)

4.6 Stability

The term 'stability' is used to describe the relative constancy of anode sensitivity with time, temperature, mean current, etc. The most important departures from constancy are:

- **long-term drift**, which is a time-dependent variation of gain under conditions of constant illumination
- **short-term shift**, which is a variation of gain following a change in mean current.

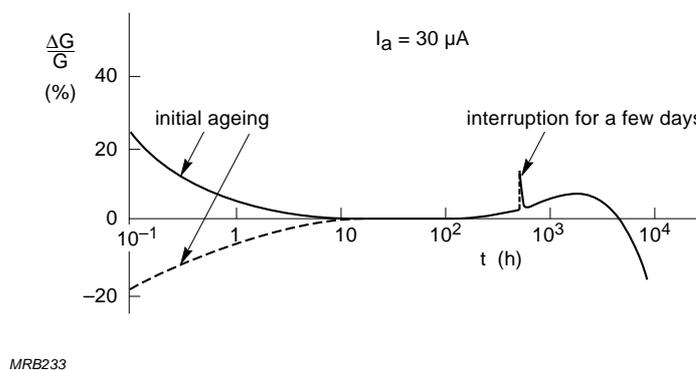
4.6.1 Long-term drift

Two modes of long-term drift can be distinguished, according to whether the mean anode current is high or low.

High-current drift; operating life. Certain more or less irreversible effects are observable at mean anode currents larger than about $10\ \mu\text{A}$. After long storage (e.g. a few months), a photomultiplier exhibits a large drift of gain for the first one or two days of operation. For some thousands of hours after that the gain is relatively stable, then it slowly decreases as a function of the total charge handled. The rate of these variations varies roughly as the anode current of the tube.

Operating life, defined as the time required for anode sensitivity to be halved, appears to be a function of the total charge delivered. Values of 300 to 1000 coulombs are typical. If the incident flux is reduced (by, say, 90%) or cut off completely, or if the supply voltage is switched off for several days, the following sequence can be observed when the original operating conditions are restored: first, a certain recovery of sensitivity accompanied by a renewed initial drift; then, a tendency to catch up fairly quickly with the slow decline of sensitivity at the point at which it was interrupted.

Figure 4.20 illustrates the relative gain variation of a photomultiplier operating at a mean anode current of $30\ \mu\text{A}$. The initial drift, which can be considered an ageing period, is between 20% and 40%. The duration of the ageing period depends on the anode current; at $10\ \mu\text{A}$ it is about 24 hours. As long as the mean current does not fall below about $100\ \text{nA}$, ageing is still observable though very slow.



MRB233
Fig.4.20 Relative gain variation of a photomultiplier operating at high average current

In most cases, if the gain is high and the cathode current low, the variations of anode sensitivity reflect variations of gain due to changes in the surface state of the dynodes. This is commonly attributed to the mobility under electron bombardment of the alkali metal molecules (mainly potassium and caesium) with which they are coated, though the exact mechanism is probably more complicated than the literature suggests.

When mean anode current is only a few microamperes, total charge delivered is no longer the decisive factor for operating life. Other effects, such as helium migration through the glass or internal migration and diffusion balances, determine the end of useful life, which is then measured in years and is independent of the mode of operation. The experience of many users would even seem to indicate that continuous, uninterrupted operation results in better long-term stability of performance characteristics than storage.

Photomultipliers with S1 cathodes deserve separate mention. Even at anode currents of only a few microamperes they exhibit large short-term drift which is independent of the gain adjustment. This drift is reversible, and the process can be speeded up by heating the tube for a few hours at the maximum permissible temperature, which suggests that there is some exchange of molecules between the dynodes and surfaces not subject to electron bombardment.

Low-current drift. When a photomultiplier is switched on and subjected to more or less constant illumination, its gain changes over the first few hours or days. The amount of change differs from type to type and even from one specimen to another of the same type. In most cases, though, the rate of change quickly decreases to as low as one per cent a month (Fig.4.21), and the higher the current the quicker the gain stabilizes. It is sometimes worthwhile to speed the process by operating the tube initially at a current up to ten times higher than that expected in the intended application. It is also advisable to leave the tube switched on even when it is idle. If the tube is stored for a time comparable with its former operating period, the gain change reverses and is repeated when the tube is again put into service.

The ANSI¹⁾ test, which is used to characterize this type of drift, employs a scintillator and a ¹³⁷Cs source positioned so as to produce a fixed count rate between 10³ and 10⁴ per second. After a stabilization period of 30 to 60 minutes, the height

¹⁾ ANSI: American National Standards Institute N42.9-1972

of the ^{137}Cs peak (662 keV) is recorded every hour for the next 16 hours and the mean gain deviation (MGD) calculated according to the formula:

$$\text{MGD} = \frac{\sum_{i=1}^{17} |p - p_i|}{17} \cdot \frac{100}{p} \quad (4.35)$$

where p is the mean height of the peak averaged over the 17 readings and p_i the height corresponding to the i th measurement.

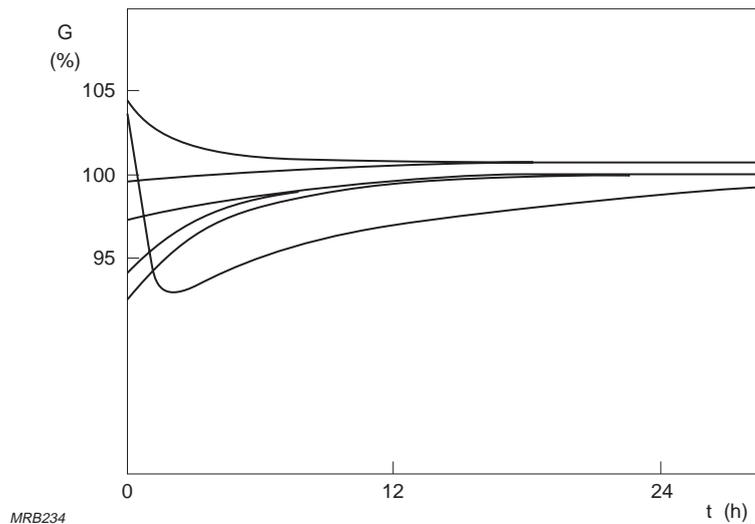


Fig.4.21 Examples of initial low-current drift

This type of drift is not related to the high-current long-term drift previously described. Though its major cause is also related to change in the structure of the emissive surfaces, other factors, such as the charge distribution at insulator surfaces (e.g. dynode spacers), may also play an important part. The drift is much less ($\text{MGD} < 1\%$, typically) in tubes with multialkali cathodes and CuBe or SbCs coated dynodes than in those with S11 cathodes or AgMg dynodes.

The ANSI test specification does not mention the anode sensitivity at which the test is to be performed. However, when a figure for long term stability is given, the mean anode current during the test must be specified. Values of about a microampere are generally used because they are broadly representative of most applications. For convenience, the scintillator and source used in the ANSI test may be replaced by a LED.

Figure 4.22 gives some examples of anode sensitivity variation curves having the same maximum deviation but different MGD values. For some applications, one may want to know the stability over a long period (for example, a month). It is possible to determine an MGD over such a period, but the measurement is more difficult because of the likelihood of drift in the measuring system itself. For such measurements, a radioactive source in combination with a scintillator is preferable to a LED because its long-term stability is much better.

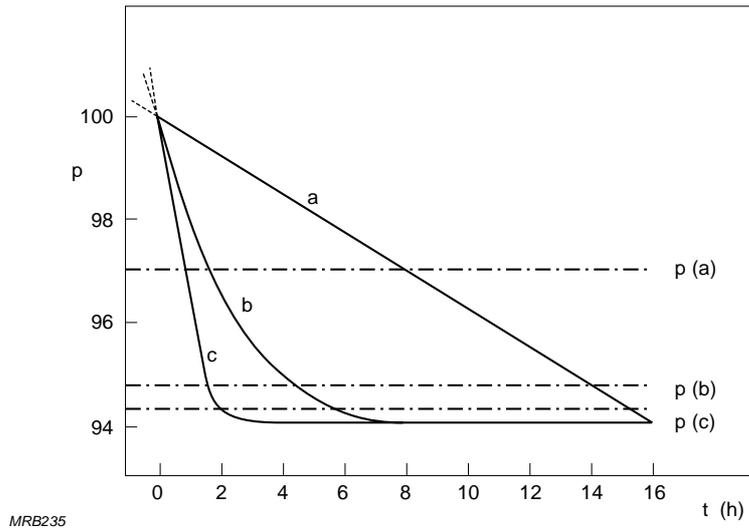


Fig.4.22 Anode sensitivity curves showing the same absolute change over 16 hours but different values of mean gain deviation (MGD) according to the ANSI method. Curve a: MGD = 1.6%; curve b: MGD = 1.1%; curve c: MGD = 0.75%; p(a), p(b) and p(c) are the corresponding mean heights of the pulse peak averaged over 17 readings

4.6.2 Short-term shift

When the flux to which a photomultiplier is exposed gives rise to a mean anode current of less than 10 μA , the gain is usually sufficiently stabilized after about 10 or 15 minutes for its long-term drift to be disregarded. If the flux is then changed abruptly, the anode current, instead of assuming a new value abruptly, starts a new drift phase before stabilizing again (Fig.4.23). Thus, the gain becomes a function (often an increasing one) of the mean value of the anode current reckoned over an interval of a second or longer.

For most photomultipliers, the time required to stabilize the gain after changing the average flux is around a second. But in some cases, and especially for tubes with S11 cathodes, this fast shift is augmented by one with a much longer time constant (about an hour). Figure 4.24 gives an example of shift with a single, short time constant; and Fig.4.25 an example with two time constants, one short and one long.

Two methods are used to measure the gain shift due to a change of average flux. The ANSI test uses a ^{137}Cs radioactive source and NaI(Tl) scintillator. After a stabilization time of at least 15 minutes, the position of the ^{137}Cs absorption peak is recorded at a count rate of 10^4 per second. The source is then moved to reduce the rate to 10^3 counts per second and the new position of the peak is recorded. The shift is characterized by the relative shift of the peak, $\Delta p/p$.

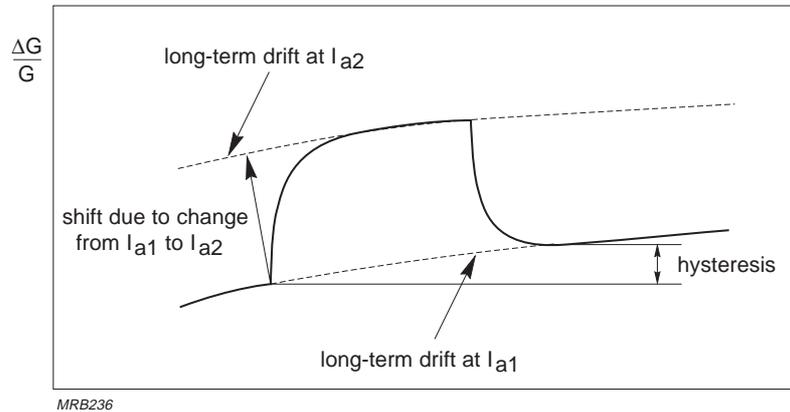


Fig.4.23 Long-term gain drift and short-term shift due to change of operating conditions

The ANSI test specification does not mention the anode sensitivity at which the test is to be performed. For the stability figures to have meaning, the extreme values of the anode mean current must be quoted. To take account of typical photomultiplier applications, the test is usually performed between 300 nA and 30 nA, 1 μA and 100 nA or even between 10 μA and 100 nA.

Another method, easier to set up, uses two independent LEDs that illuminate the photomultiplier simultaneously. One emits pulses of adjustable intensity and frequency, or simply a continuous flux of adjustable intensity, for setting the mean anode current to any desired level throughout the applicable range. The other emits pulses of fixed intensity and frequency. The mean height of the anode pulses due to these is a measure of the relative gain at the set level.

One cause of shift may be that charges on internal insulators (dynode spacers) are affected by the passage of electrons (scatter, for example), and that this in turn modifies the focusing between stages. In present-day photomultipliers, particularly those with venetian-blind dynodes, careful design of the electrode structure practically eliminates this effect. But even so, variations in gain due to variations in secondary

emission can still be observed, which suggests the influence of phenomena at the level of the emissive layer itself.

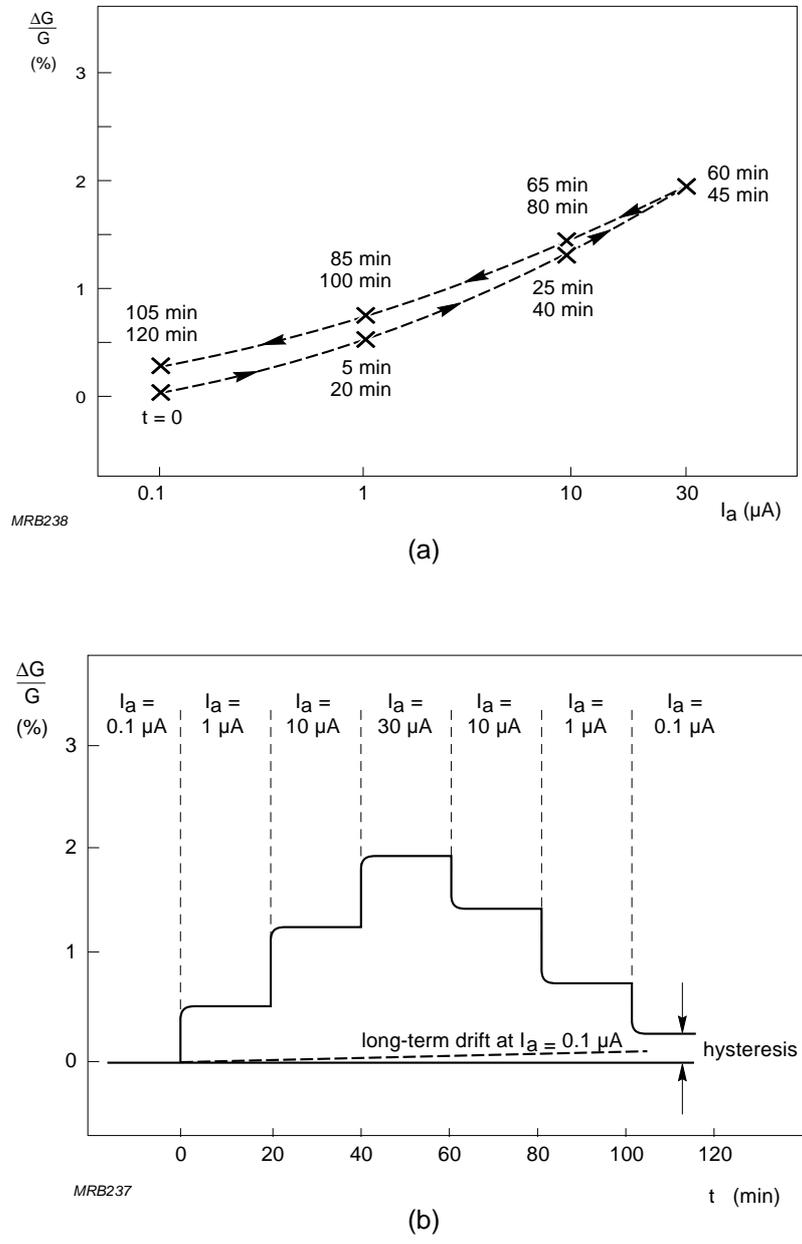


Fig.4.24 Gain shift of a photomultiplier with a single, short stabilization time constant: (a) as a function of mean anode current, (b) as a function of time. Measured 5 minutes and 15 minutes after illumination changes made every 20 minutes

Tubes with bialkali cathodes and CuBe venetian-blind or SbCs coated dynodes are usually considered the most stable in respect of shift, gain variations of less than 1% being common for anode current variations of ten to one (from 100 nA to 1 μA).

Figures 4.24 and 4.25 show that the gain does not return exactly to its original value when the flux does. This hysteresis reflects an interaction between long-term and short-term stability parameters. During prolonged operation, the higher the current the quicker the hysteresis tends to disappear. Here again, accelerated ageing at medium to high current has a useful effect.

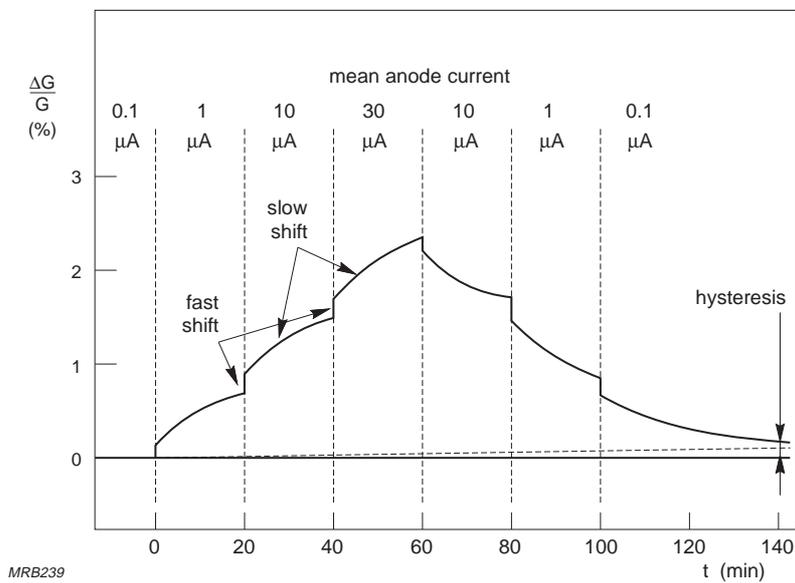
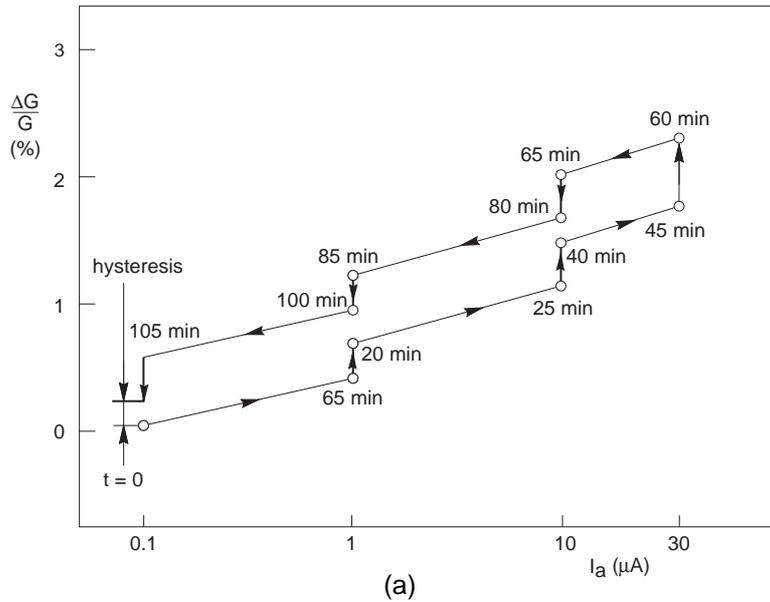


Fig.4.25 Gain change of a photomultiplier with one short and one long stabilization time constant: (a) as a function of mean anode current, (b) as a function of time. Measuring conditions as for Fig.4.24

4.7 Afterpulses

Afterpulses are spurious pulses that appear in the wake of true pulses. They can be observed on an oscilloscope during detection of very short flashes such as scintillation and laser pulses. As they are time-correlated with the true pulses, they are particularly inconvenient in chronometry and time spectrometry applications using coincidence techniques; in counting applications they often spuriously limit the number of true pulses that can be registered. Afterpulses have two main causes – luminous reactions and ionization of residual gases – which can be distinguished by the interval that separates the afterpulse from the true pulse.

4.7.1 *Luminous reactions*

When the electrodes are bombarded by electrons they emit photons. Although the luminous efficiency of these impacts is very low, in some tubes it is possible for light emitted even by the last stages to make its way back to the photocathode and so give rise to afterpulses. With respect to the true pulses from which they originate, these are delayed by the sum of the light and electron pulse transit times, typically 20 to 100 ns.

4.7.2 *Ionization of residual gases*

Ionization of gas traces also gives rise to afterpulses. The traces may be of residual gases left in the tube after evacuation or desorbed by materials of its structure, or of helium that has migrated through the glass. Ionization afterpulses are delayed by a few hundred nanoseconds to as much as several microseconds with respect to the true pulses from which they originate.

Ionization afterpulses can be subdivided according to whether they originate in the electron-optical input system or the electron multiplier.

Primary photoelectrons in the electron-optical input system can generate positive ions that are accelerated towards the cathode and there give rise to emission of one or more secondary electrons. The transit time of the ions depends more on the input-system electric field and the mass of the ions than on the distance from the cathode at which they originate. The usual ions are H_2^+ , and He^+ , and CH_4^+ ; in a photo-multiplier with a 50 mm cathode, these have transit times of about 0.3 μs , 0.4 μs and 1 μs , respectively, at a cathode to first-dynode voltage of 300 V. The amplitude of the resulting afterpulses, relative to the true pulses, increases very rapidly with increasing cathode to first-dynode voltage.

The amount of helium entering the tube depends on the kind of glass used, its surface area, and the ambient helium partial pressure (typically $0.7 \text{ Pa} \approx 5 \times 10^{-3} \text{ torr}$). For applications in which higher helium partial pressures are encountered, tubes with soft (lime) glass envelopes are preferred.

Ionization afterpulses originating in the electron multiplier come mainly from the last stages, where the electron current is largest. The relative amplitude of these pulses is independent of the cathode to first-dynode voltage but it does increase rapidly with gain. The delays observed for afterpulses of this type amount to a microsecond or more (e.g. about $3 \mu\text{s}$ for Cs^+ ions).

4.7.3 Afterpulse ratio

Each true pulse may be followed by one or more afterpulses (Fig.4.26). The relation between true pulses and afterpulses is quantified by an *afterpulse ratio* which may be stated as

- the ratio of the *number* of afterpulses to the number of true pulses, or
 - the ratio of *charge* transferred by the afterpulses to charge transferred by the true pulses,
- in either case expressed as a percentage.

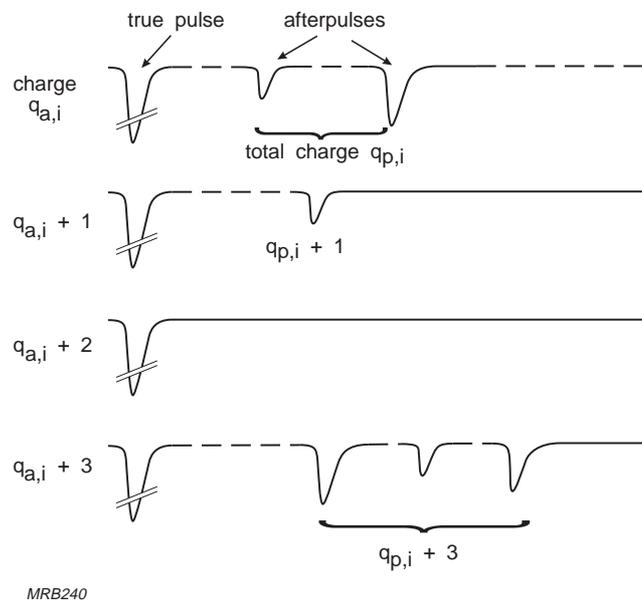


Fig.4.26 Examples of the development in number and charge of afterpulses

In specifying the afterpulse ratio, the interval in which afterpulses are counted following each true pulse must also be stated. The ratio can vary greatly depending on the measuring conditions and can be interpreted differently according to the application.

Afterpulse ratio stated as a charge ratio is given by the expression

$$F_{\text{apq}} = \frac{\sum q_{\text{p},i}}{\sum q_{\text{a},i}} \times 100 \quad (4.36)$$

where $q_{\text{a},i}$ is the charge transferred by true pulses and $q_{\text{p},i}$ the charge transferred by afterpulses. The ratio is usually less than 1% and, as long as the gain is not too high, does not vary much with the number of true pulses or the amount of charge they contain. When the charge transferred by each true pulse is very small (e.g. in single-electron operation), that transferred by each afterpulse may be as large or even larger. However, as proportionally fewer true pulses are then followed by afterpulses, the charge ratio remains about the same.

The afterpulse ratio can be reduced by decreasing the cathode to first-dynode voltage as far as the application will permit, or by working with minimum gain, or both.

The effects of afterpulses can be minimized by using coincidence techniques, by blanking the photomultiplier for a set interval after each true pulse (§5.10), or simply by using electronics with sufficiently long dead time.

4.8 Environmental considerations

Environmental factors – chiefly temperature, magnetic fields, background radiation, and atmosphere – can affect the operation of a photomultiplier in varying degrees, temporarily or permanently. To a large extent the effects can be guarded against or compensated.

4.8.1 Temperature

By the nature of their photoemissive and secondary emissive materials, photo-multipliers are also sensitive to temperature variations. These affect three of the main characteristics:

- spectral response (the shape of the curve)
- dark current (the thermionic component)
- anode sensitivity and gain (secondary emission coefficients).

Changes in characteristics due to temperature variations within the permissible limits are usually reversible, though there may be some hysteresis that disappears only gradually.

Effects of temperature on the photoemission and secondary-emission surfaces are complex. They depend not only on the composition of the surfaces but also, to some extent, on the type of tube; and even between tubes of the same type there are appreciable differences. However, tendencies and average values can be identified.

In storage as well as use, photomultipliers must be kept within the temperature limits specified in their data sheets, usually $-30\text{ }^{\circ}\text{C}$ to $80\text{ }^{\circ}\text{C}$ ($50\text{ }^{\circ}\text{C}$ for types with S1 cathodes). Beyond those limits effects such as sublimation of the cathode or stresses in the glass may occur. **Always consult the manufacturer before considering operation of a photomultiplier outside its published temperature limits.**

Effect on spectral sensitivity. The spectral sensitivity characteristic does not vary much with temperature. The greatest relative variation is usually observed close to the photoemission threshold. For a given application, therefore, it is advisable to choose a tube with a type of cathode that makes it possible to operate far from the threshold.

The dependence of monochromatic sensitivity on temperature differs from one type of cathode to another. For a given type it is defined (in percent per degree at $20\text{ }^{\circ}\text{C}$) by the ratio

$$\alpha_k = \frac{\Delta S_k}{S_{k(20^{\circ}\text{C})}\Delta\theta} \times 100 \quad (4.37)$$

where ΔS_k is the change of cathode sensitivity observed over a temperature interval $\Delta\theta$. For commonly used photocathodes α_k usually varies continuously with wavelength, in some cases passing through the zero and changing sign.

The value of α_k and its variation with wavelength depend not only on the composition of the cathode but also on the structure of the tube, and therefore differ considerably from one type to another. The figures given below are merely indicative.

S11 cathode. Of all photocathodes, this type exhibits the widest variation of α_k as a function of wavelength; see Figs 4.27 and 4.28. The sign change of two of the specimens in Fig.4.28 at about 580 nm wavelength is supported by other observations but is not the rule; the temperature coefficient of many S11 cathodes keeps the same sign at least throughout the range of practical wavelengths (400 nm to 620 nm).

Bialkali SbKCs cathode. The temperature coefficient is very low in wavelength range 400 nm to 500 nm (Fig.4.29) where sensitivity is maximum, and may go to zero there or change sign. In Fig.4.30, where the coefficient is plotted for temperature intervals $-20\text{ }^{\circ}\text{C}$ to $20\text{ }^{\circ}\text{C}$, and $20\text{ }^{\circ}\text{C}$ to $60\text{ }^{\circ}\text{C}$, it is nowhere greater than 0.15% in the range 400 nm to 500 nm. Furthermore, its variation with temperature is very small at short wavelengths.

S20 and S20R trialkali cathodes. These are characterized by a negative temperature coefficient throughout most of the useful spectrum. Figure 4.31 shows the variation of relative sensitivity as a function of temperature, of both types for several wavelengths; and Fig.4.32, the variation of temperature coefficient as a function of wavelength. Note that type S20R, whose response extends farther into the red, has a larger temperature coefficient than type S20.

Whatever the type of cathode, the residual change of sensitivity (hysteresis) following a temperature cycle is usually much less than 1% and can be disregarded.

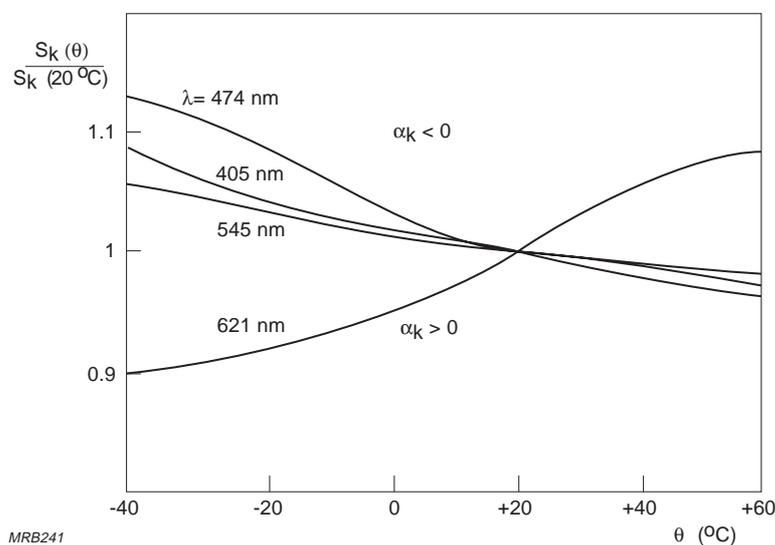
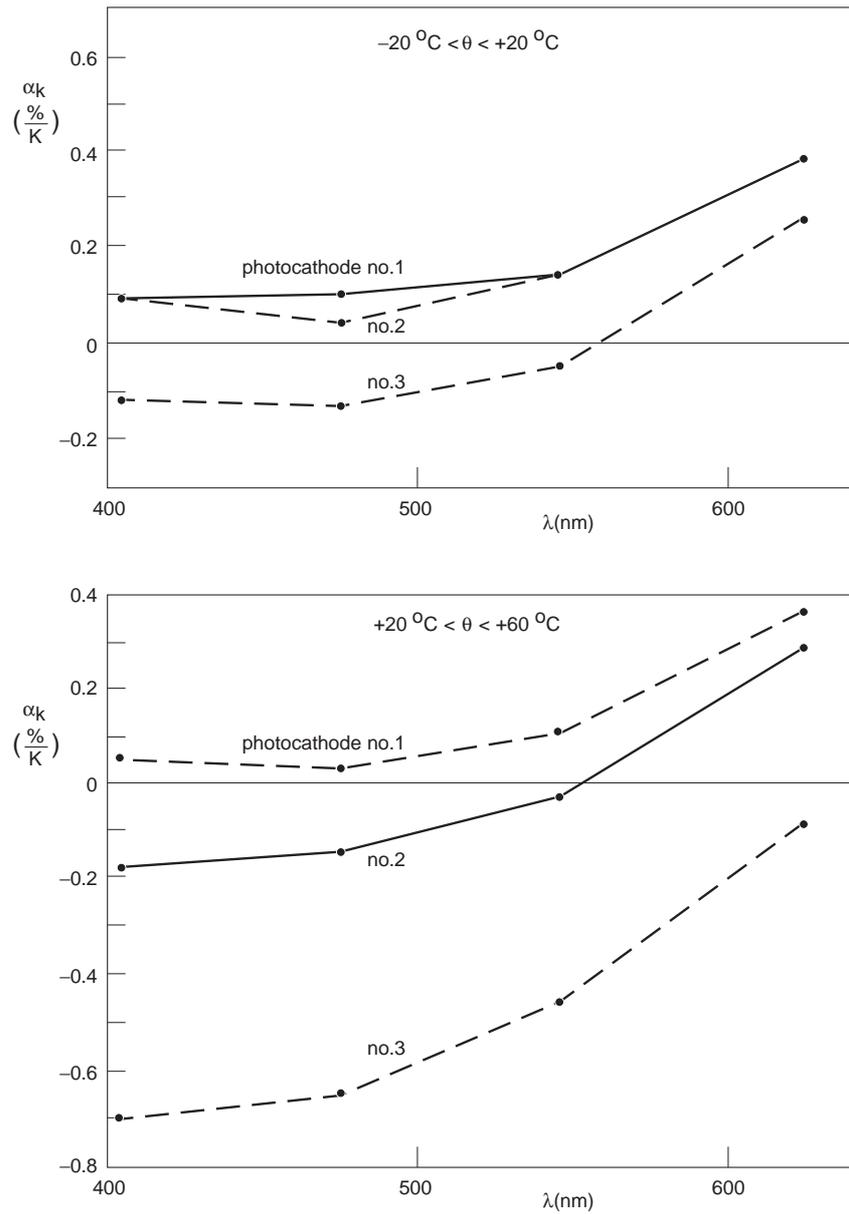


Fig.4.27 Relative variation of sensitivity of a type S11 cathode as a function of temperature, with wavelength as parameter



MRB242

Fig.4.28 Examples of the variation of the temperature sensitivity coefficient α_k of type S11 cathodes as a function wavelength

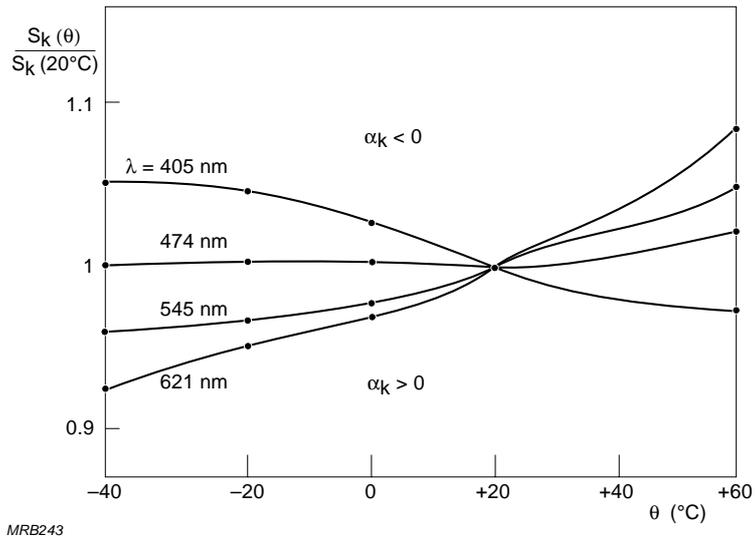


Fig.4.29 Relative variation of sensitivity of a SbKCs bialkali cathode as a function of temperature, with wavelength as parameter

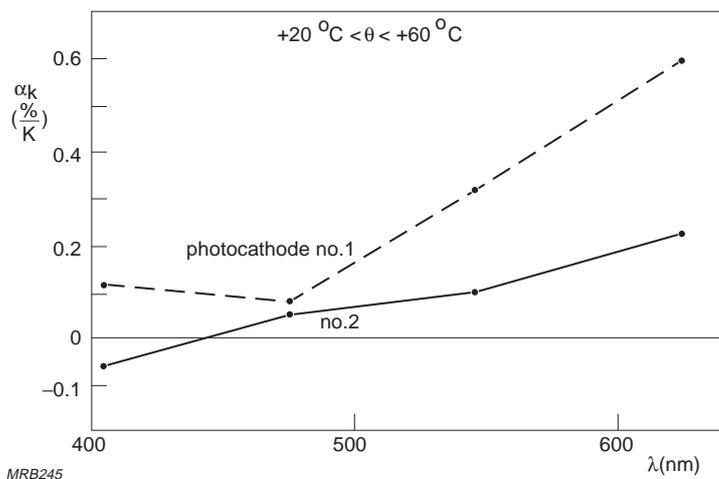
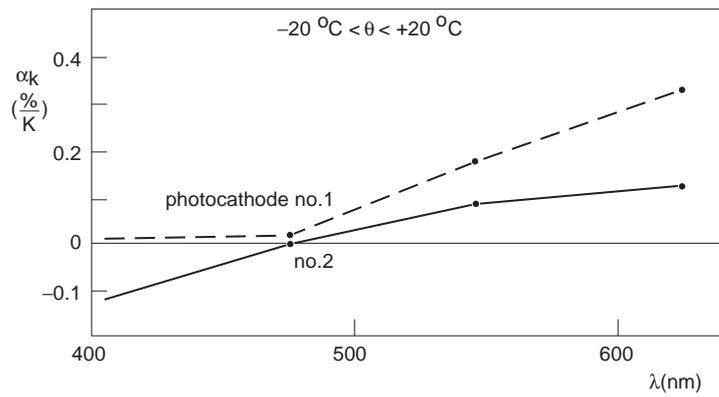


Fig.4.30 Examples of the variation of the temperature sensitivity coefficient α_k of SbKCs bialkali cathodes as a function of wavelength

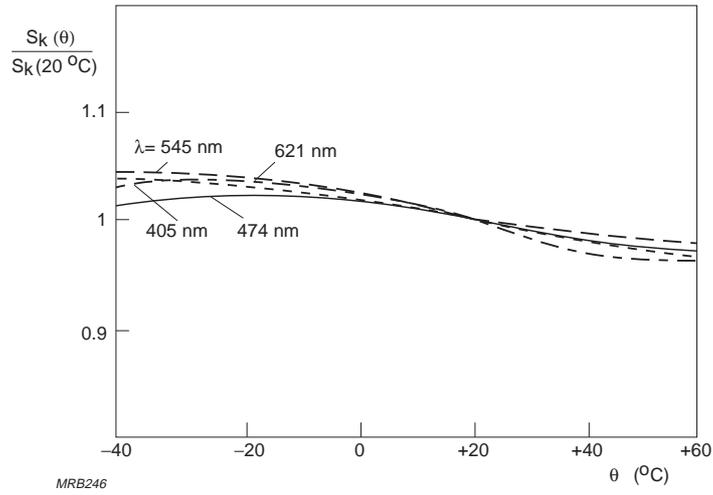


Fig.4.31 Relative variation of sensitivity of type S20 cathode as a function of temperature, with wavelength as parameter

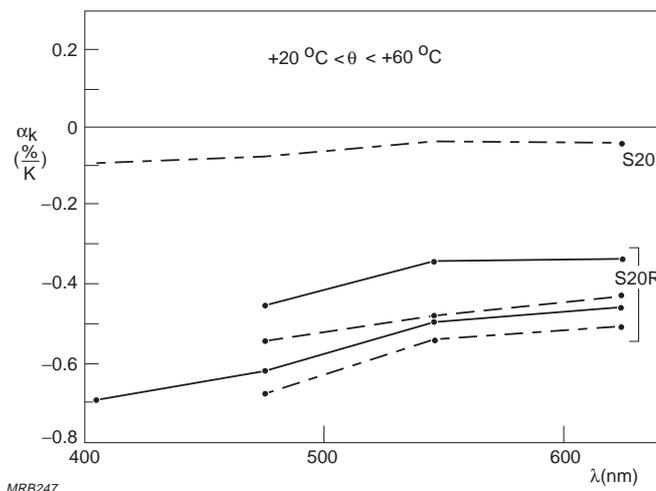
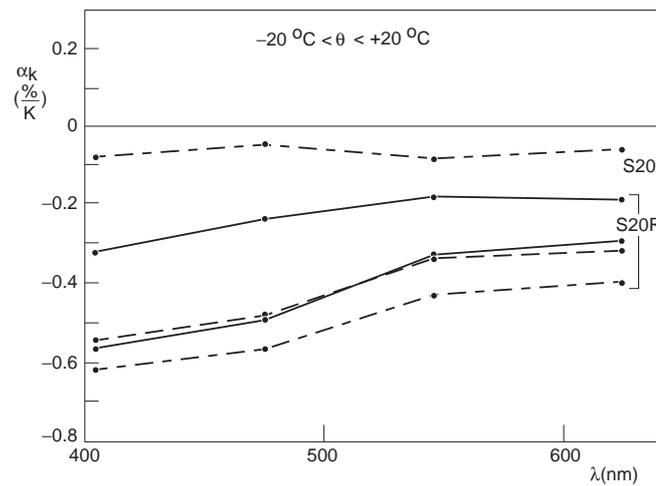


Fig.4.32 Examples of the variation of the temperature sensitivity coefficient α_k of type S20 and S20R cathodes as a function of temperature

Effect on cathode resistivity. The resistivity of photocathodes varies inversely with temperature. This can limit the minimum operating temperature, especially of bialkali SbKCs cathodes which, at room temperature, have a resistivity a hundred to a thousand times greater than that of S11 and S20 cathodes (Fig.4.33). The practical minimum for bialkali cathodes is $-30\text{ }^{\circ}\text{C}$ (if cathode current is more than 0.1 nA); and for all other types, about $-100\text{ }^{\circ}\text{C}$.

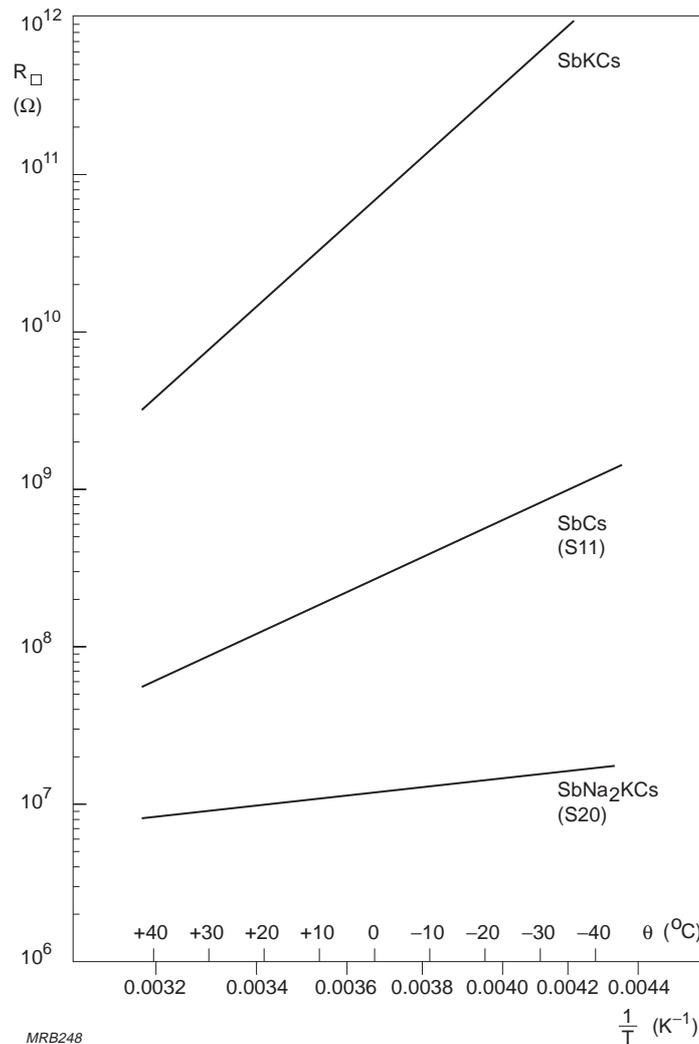


Fig.4.33 Surface resistivities of three photoemissive materials as functions of temperature

Effect on the dark current. The thermionic component of the dark current obeys Richardson's law, Eq.3.1; therefore, both dark current and dark pulse rate increase strongly with temperature. The rate of increase depends mainly on the cathode material but may differ considerably from one tube to another of the same type. This

is because some components of the dark current, such as thermionic emission, field emission, leakage current, etc., have different temperature coefficients and differ in relative importance from tube to tube. For the same reason, the dark current and dark pulse rate do not usually continue to decrease below a certain temperature, and in some cases may even increase, particularly in tubes with multialkali cathodes.

Table 4.1 Thermionic emission characteristics of common photocathodes

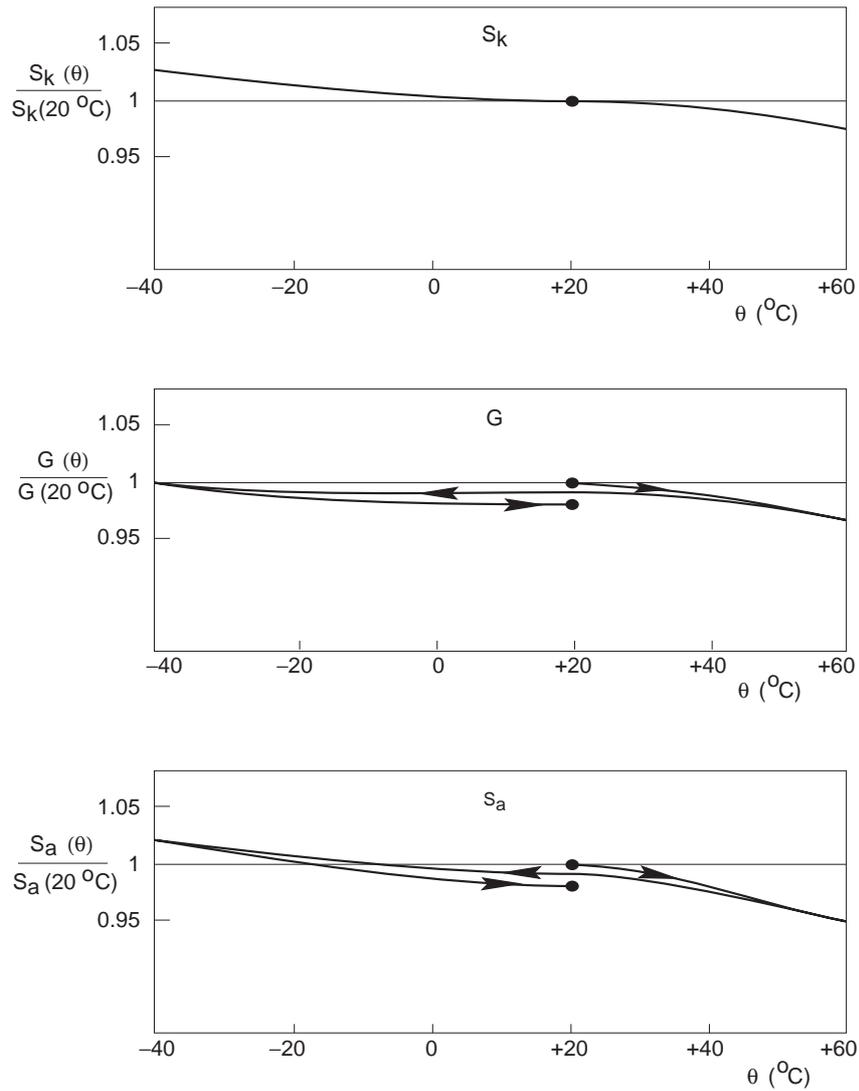
Type of cathode	Thermionic emission at 20 °C (A/cm ²)	Temperature rise for which thermionic current doubles (K)	Minimum operating temperature (°C)
AgOCs (S1)	10 ⁻¹³ – 10 ⁻¹¹	5 – 7	-100
SbCs (S11)	10 ⁻¹⁶ – 10 ⁻¹⁵	6 – 15	-20
SbKCs	10 ⁻¹⁹ – 10 ⁻¹⁷	4 – 5	-20
SbNa ₂ KCs (S20)	10 ⁻¹⁹ – 10 ⁻¹⁵	4	-40
SbNa ₂ KCs (S20R)	10 ⁻¹⁷ – 10 ⁻¹⁵	4	-40

Effect on gain and anode sensitivity. Dynode secondary emission also varies with temperature (though less so than cathode thermionic emission) and correspondingly affects gain. The temperature coefficient of gain is usually negative and depends not only upon the composition of the dynodes but also upon that of the cathode and, to some extent, the structure of the multiplier. For CuBe and AgMg dynodes the coefficient is about -0.1% per degree with bialkali and trialkali (S20) cathodes, and between -0.5% and -1% per degree with S11 cathodes. It is smaller in tubes with venetian-blind dynodes than in those with focusing dynodes.

Variations in anode sensitivity reflect variations in both cathode sensitivity and multiplier gain. At certain temperatures and wavelengths these may be equal and opposite, cancelling each other.

After a temperature cycle, anode sensitivity usually does not return exactly to its previous value. The hysteresis is mainly in the multiplier gain (cathode sensitivity hysteresis being negligible) and tends to disappear after long storage. Figure 4.34 shows the relative cathode sensitivity, gain, and anode sensitivity hysteresis of a tube with bialkali cathode and CuBe venetian-blind dynodes.

Gain hysteresis hardly varies from tube to tube. It is not cumulative and, in fact, tends to diminish after repeated temperature cycles. As an example, a tube with venetian-blind dynodes showed a -2% gain change after a first $-40\text{ }^{\circ}\text{C}$ to $60\text{ }^{\circ}\text{C}$ to $-40\text{ }^{\circ}\text{C}$ cycle, but only -0.8% after a second cycle.



MRB249

Fig.4.34 Temperature variation of cathode and anode sensitivity and gain of a tube with a bialkali cathode at a wavelength of 405 nm, showing hysteresis.

4.8.2 Magnetic fields

Magnetic fields even as weak as the earth's affect photomultiplier performances. This can be demonstrated by rotating a horizontally mounted tube about its main axis. The resulting variation of anode sensitivity is due to the varying effect of the earth's field on the electron trajectories, and the corresponding variation of collection efficiency in all stages. Highly focused tubes, in which the electron impact areas on the dynodes are small, are the most sensitive to magnetic effects; a transverse flux density of a few tenths of a millitesla can reduce gain by 50%. In a tube with venetian-blind dynodes, the field required to produce the same effect would be up to three times as large.

Magnetic influence is greatest in the electron-optical input system, where electron trajectories are longest. Increasing the voltage across the input system increases the energy of the electrons and decreases the sensitivity to magnetic fields. A photomultiplier tube is least sensitive to magnetic fields parallel to its axis.

Tubes with linear focusing dynodes are most sensitive to magnetic influence when the field is parallel to the dynodes (axis y in Fig.4.35). (For precautions against magnetic effects see §5.9.)

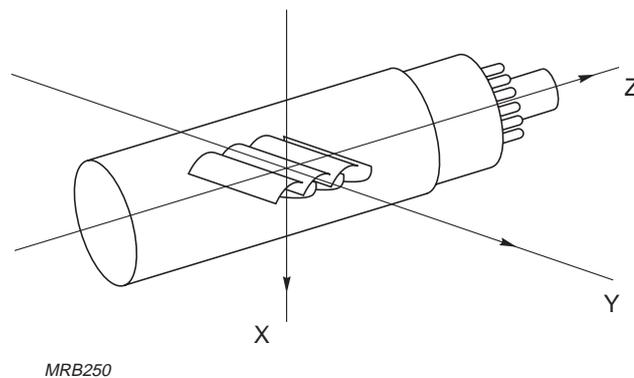


Fig.4.35 Axes used in measuring magnetic sensitivity

Magnetic sensitivity is measured relative to three perpendicular axes (Figs 4.35 and 4.36). Data sheets give either the measured sensitivity curves or the values of magnetic flux density parallel to each axis at which gain is halved. The data are for optimum operating conditions. Magnetic sensitivity is greater when electrode potentials are not optimum, as is the case when gain is deliberately decreased by defocusing a dynode or the accelerating electrode.

Strong fields may permanently magnetize some parts of a photomultiplier, lastingly affecting its performance. If that happens, the tube can be demagnetized with a coil producing a flux density of about 10 mT, at 50 Hz.

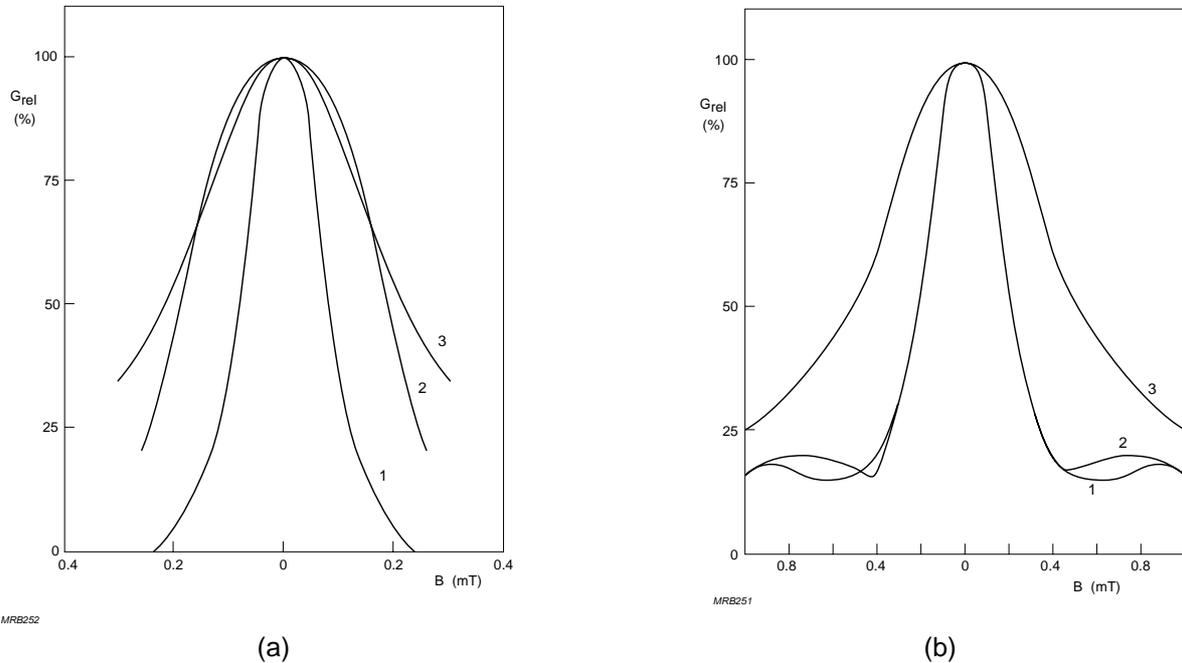


Fig.4.36 Relative gain variation as a function of magnetic field: (a) for a tube with linear focusing dynodes, (b) for a tube with venetian-blind dynodes.
 curve 1: field aligned with y-axis (Fig.4.35)
 curve 2: field aligned with x-axis
 curve 3: field aligned with z-axis

4.8.3 Radiation

Ionizing radiation can give rise to secondary effects such as dark current increase or, if the radiation is sufficiently intense, permanent loss of sensitivity and gain.

Dark current increase. Background radiation interacts with parts of the tube, mainly the glass, to produce light pulses that increase the dark current or dark pulse rate. The effect can be demonstrated by bringing a radioactive source close to the tube; for high-energy γ -radiation (more than 1 MeV) the dark current will increase considerably, owing to pair production in the glass.

Light pulses due to background radiation have two causes: scintillation and the Cherenkov effect.

Scintillation is due to interaction of low-energy α - and β -radiation with the glass of the envelope; such radiation may come from the surroundings or from the glass itself (e.g. from traces of ^{40}K). The scintillation efficiency is very low – about 10 photons

per MeV, not all of which reach the cathode – so the resulting pulses are mostly single-electron.

Cherenkov effect occurs when charged particles traverse a medium – e.g. the glass of the envelope – at more than the speed of light in that medium. Such particles may be due to cosmic radiation or may be generated in the medium itself by high-energy γ -radiation. The resulting anode pulses are of large amplitude, corresponding to 10 to 20 photoelectrons; moreover, they are followed by low-amplitude afterpulses due to phosphorescence of the glass excited by the ultraviolet content of the Cherenkov emission.

Cherenkov effect due to cosmic radiation can be a troublesome source of dark pulses in many applications; however, there are several ways of keeping it, or its influence on the measurement, to a minimum.

- *Thin input window.* The amplitude of light pulses due to the Cherenkov effect is proportional to the thickness of the glass: about 500 photons/cm in the wavelength range 300 nm to 600 nm. Input windows should therefore be as thin as possible to help keep the Cherenkov effect low.
- *Inverting the tube.* The direction of cosmic rays is usually more or less vertical. Positioning the tube with the input window down reduces their effect; Cherenkov radiation then tends to be emitted outwards, and only the fraction reflected back by the glass-air interface reaches the cathode. If the window is up, all of it reaches the cathode. (Mounting the tube horizontally gives an intermediate result. The glass cross-section presented to cosmic rays is much smaller but the distance the rays travel in the glass is much larger, so the count rate is lower but the pulse amplitude higher).
- *Window material.* Glass that limits transmission of ultraviolet light attenuates that component of the Cherenkov radiation. Glass windows give less Cherenkov-effect dark current than fused silica windows.
- *Anti-coincidence.* With an anti-coincidence system using two detectors it is possible to eliminate a large part of the dark current due to cosmic radiation. An anti-coincidence detector above the photomultiplier, or better still, completely surrounding it, makes it possible to cancel all pulses registered simultaneously.
- *High discrimination.* As pulses due to cosmic rays are of high amplitude, this characteristic can often be used as a basis for discrimination. When low-energy radiation is to be detected, for instance, the counting window can be centred on the relevant energy band so as to exclude high-amplitude pulses.

Materials in the surroundings may have a significant effect. Cosmic rays interact with them to produce showers of secondary particles that give rise to trains of noise pulses in the photomultiplier. A concrete roof a few centimetres thick is not enough to block cosmic radiation but, on the contrary, may be a troublesome source of noise due to secondary particles (spallation noise).

Gain and sensitivity loss. Permanent loss of gain and sensitivity is a serious risk only in very high radiation environments, such as parts of extraterrestrial space where electron flux can reach 10^{10} electrons per square centimetre per second. Under those conditions the emissive properties of the dynodes change and the input window darkens, affecting the transmission at shorter wavelengths. Lithium fluoride and lime glass windows are more sensitive than others to prolonged radiation. The photoemissive layer is relatively unaffected, probably because its absorption coefficient for ionizing radiation is low.

Permanent alteration of gain and sensitivity becomes noticeable only after exposure doses of about 10^4 rad.

4.8.4 Atmosphere

Humidity. Because of the high voltages used, operation in a damp atmosphere can lead to insulation problems. Condensation gives rise to leakage currents which increase the dark current. Local insulation breakdowns may also occur. Take particular care to avoid condensation on the glass, at the pins, and especially inside the plastic base. If moisture does get into the base, it will be necessary to drill a hole in the base key to enable it to escape.

Ambient pressure. Photomultipliers can operate satisfactorily at low ambient pressure, but precautions against flashover at the pins are necessary at pressures below 10 kPa (≈ 75 torr). *For operation or storage at high ambient pressure, consult the manufacturer; permissible pressures differ from type to type and are not usually given in the data sheets.*

Helium partial pressure. Glass is permeable to helium, the rate of penetration being proportional to the helium partial pressure. Of the glasses used in photomultipliers, lime glass (soft glass) is the least permeable; borosilicate glass and fused silica (hard glasses) are, respectively, about 100 and 1000 times more permeable. Helium intrusion increases the afterpulse factor and shortens the life: a tube with a fused silica window in a helium partial pressure of 100 kPa has a useful life of only a few days; this is a hazard to be guarded against in helium-cooled high-energy physics

experiments. The partial pressure of atmospheric helium is normally about 0.7 Pa, which is low enough to allow an average useful life of some ten years. Finally, note that helium penetration increases with temperature.

4.8.5 *Mechanical stress*

Like all electron tubes, photomultipliers should be protected against mechanical and temperature stress. Vibration or shock transmitted to the dynodes can modulate the gain (microphony). Especially robust types are available for use in hostile environments.

APPENDIX 4. SIGNAL TRANSFER IN LINEAR SYSTEMS

A4.1 Pulse and step responses

A transmission system having an input signal $U_1(t)$ and an output signal $U_2(t)$, is called linear if its response to the sum of two input signals acting together is equal to the sum of its responses to the two acting separately.

When a unit-area input pulse $U_1(t)$ of width Δt and amplitude $1/\Delta t$ is applied, the output pulse $U_2(t)$ is wider than Δt (Fig.A4.1). As Δt tends toward zero, $U_1(t)$ tends toward the delta function $\delta(t)$, such that

$$\lim_{\Delta t \rightarrow 0} \int_0^{\Delta t} \delta(t) dt = 1$$

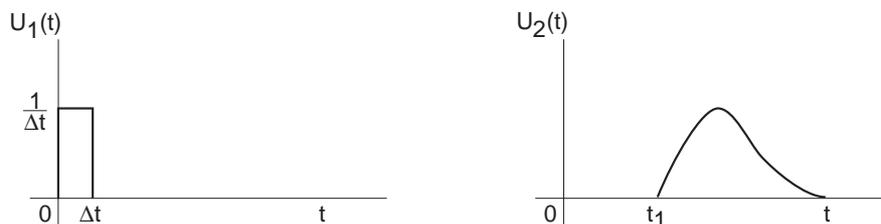
The corresponding output pulse $U_2(t)$ then represents the *pulse response*, $R_\delta(t)$.

The *step response* $R_\varepsilon(t)$ is the response to a unit step $\varepsilon(t)$ such that

$$\begin{aligned} \varepsilon(t) &= 0, \text{ for } t < 0 \\ \varepsilon(t) &= 1, \text{ for } t \geq 0 \end{aligned}$$

The unit step is related to the delta function by the expression:

$$\delta(t) = \frac{d}{dt} \varepsilon(t) \tag{A4.1}$$



MRB253A

Fig.A4.1 Response to a rectangular pulse

A4.1.1 Superposition principle

An arbitrary input signal $U_1(t)$ can be represented by the superposition of very narrow adjacent pulses of width ds (Fig.A4.2). The response of the system to one such pulse occurring at an instant δ is (to within the second order) equal to the product of the pulse area and the pulse response of the system, shifted by the time s :

$$dU_2(t) = U_1(s) ds \cdot R_\delta(t - s)$$

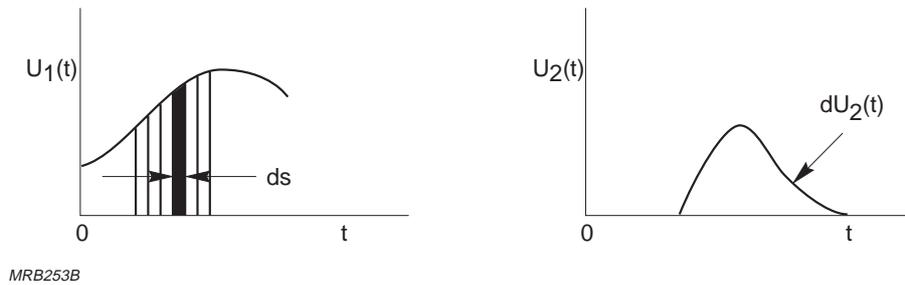


Fig.A4.2 Superposition of elemental pulses

Provided the system is linear, its response to the input signal $U_1(t)$ is the summation of the individual responses to the elementary pulses; thus,

$$U_2(t) = \sum_s dU_2(t)$$

or, letting ds tend toward zero,

$$U_2(t) = \int_{-\infty}^{+\infty} U_1(s) R_\delta(t - s) ds \quad (\text{A4.2})$$

This integral is known as the convolution product and is written

$$U_2(t) = U_1(t) * R_\delta(t) \quad (\text{A4.3})$$

Interchanging the variables s and $t-s$ gives an equivalent form of Eq.A4.2,

$$U_2(t) = \int_{-\infty}^{+\infty} U_1(t - s) R_\delta(s) ds \quad (\text{A4.4})$$

i.e., the convolution product is commutative.

From Eqs A4.2, A4.3 and A4.4 it is possible to calculate the step response $R_\epsilon(t)$ if the pulse response $R_\delta(t)$ is known:

$$R_\epsilon(t) = \epsilon(t) * R_\delta(t) \quad (\text{A4.5})$$

This equation is useful when it is difficult to measure the step response directly. With photomultipliers, for example, it is easier to simulate repetitive delta-function pulse inputs than it is to simulate repetitive unit-step inputs.

Using the superposition principle and similar reasoning to the above, it can be shown that the output of a linear system can also be expressed as a function of the step response $R_\epsilon(t)$ by the relation:

$$U_2(t) = \frac{d}{dt} [U_1(t) * R_\epsilon(t)] \quad (\text{A4.6})$$

This is equivalent to Eq.A4.3 and can be derived from it via Eqs A4.1 and A4.5.

A4.1.2 Rise time and FWHM

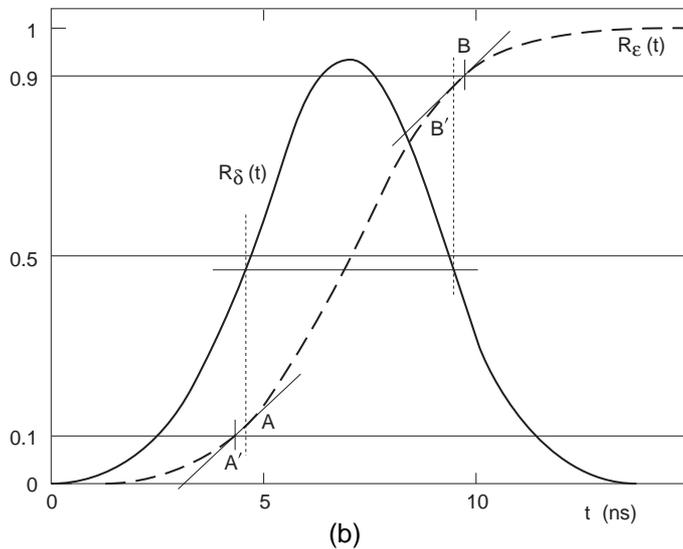
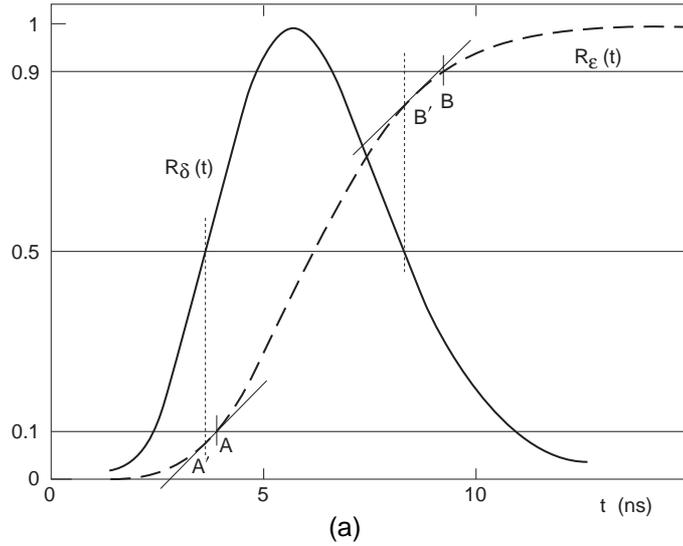
Equation A4.5 can be simplified by noting that the pulse response $R_\delta(t)$ and the step response $R_\epsilon(t)$ are zero for $t \leq 0$; thus,

$$R_\epsilon(t) = \int_0^t R_\delta(s) ds \quad (\text{A4.7})$$

Several workers have proposed that the pulse response $R_\delta(t)$ of a photomultiplier can be accurately represented by a function of the type

$$R_\delta(t) = \epsilon(t) \frac{\sqrt{m+1}}{m! \sigma_R} \left(\frac{\sqrt{m+1}}{\sigma_R} t \right)^m \exp \left(- \frac{\sqrt{m+1}}{\sigma_R} t \right) \quad (\text{A4.8})$$

where $m = 8$ (solid line in Fig.A4.3(a)). The dashed line in Fig.A4.3(a) represents the step response $R_\epsilon(t)$ derived from the pulse response $R_\delta(t)$ via Eq.A4.5 or A4.7. The step response is usually characterized in terms of the rise time $t_{r,\epsilon}$ between 10% and 90% of maximum (points A and B). Figure A4.3(a) illustrates an interesting consequence of Eq.A4.7: the rise time between points A' and B', where the slope of the step response is half of its maximum, is equal to the full width at half maximum t_w of the pulse response. Points A', B' are located close to points A, B respectively.



MRB254

Fig.A4.3 Pulse response (a) of the type $at^m \exp(-bt)$; (b) Gaussian pulse response

Another case to consider is the gaussian function (Fig.A4.3(b)) with standard deviation σ_R

$$R_\delta(t) = \epsilon(t) \frac{1}{\sigma_R \sqrt{2\pi}} \exp \left\{ -\frac{(t - t_l)^2}{2\sigma_R^2} \right\} \quad (\text{A4.9})$$

which is sometimes assumed for the pulse response of a photomultiplier, even though it is symmetrical; it is easier to manipulate than Eq.A4.8 and facilitates approximation. For the gaussian function points A' and B' are very close to points A and B , and the

step response rise time is nearly equal to the pulse response full width at half maximum:

$$t_{r,\varepsilon} \approx 1.11 t_w \quad (\text{A4.10})$$

Gaussian light pulse. The photomultiplier pulse response will be as given by Eq.A4.9, and the convolution of this with a gaussian light pulse is:

$$R^*(t) = \frac{1}{2\pi\sigma_L\sigma_R} \int_0^t \exp\left\{-\frac{(s-t_L)^2}{2\sigma_L^2}\right\} \exp\left\{-\frac{(t-t_t-s)^2}{2\sigma_R^2}\right\} ds \quad (\text{A4.11})$$

or:

$$R^*(t) = \frac{f(t)}{2\pi\sigma_L\sigma_R} \exp\left\{-\frac{(t-t_t-t_L)^2}{2(\sigma_L^2 + \sigma_R^2)}\right\} \quad (\text{A4.12})$$

It can be shown that if $(t_L + t_t)$ is greater than about twice the quantity $\sqrt{\sigma_L^2 + \sigma_R^2}$, the function $f(t)$, which increases monotonically, quickly approaches the asymptote

$$\sqrt{\left(2\pi \frac{\sigma_L^2 \sigma_R^2}{\sigma_L^2 + \sigma_R^2}\right)} \quad (\text{A4.13})$$

When t is in the range $t_L + t_t \pm 2\sqrt{\sigma_L^2 + \sigma_R^2}$, the function $f(t)$ approaches this asymptote to within less than 1%; $R^*(t)$ can then be written in the form

$$R^*(t) = \frac{1}{\sqrt{2\pi(\sigma_L^2 + \sigma_R^2)}} \exp\left\{-\frac{(t-t_t-t_L)^2}{2(\sigma_L^2 + \sigma_R^2)}\right\} \quad (\text{A4.14})$$

which is a gaussian function with variance $\sigma_L^2 + \sigma_R^2$.

A4.2 Time resolution

Pulse transit time in a photomultiplier (Fig.A4.4) fluctuates from pulse to pulse. As the instant to which it is referred is the instant of illumination, however, its probability density distribution cannot be considered independently of the statistics of photon emission. Illumination that takes the form of, say, a delta-function pulse gives a different density distribution than illumination by a pulse of arbitrary width.

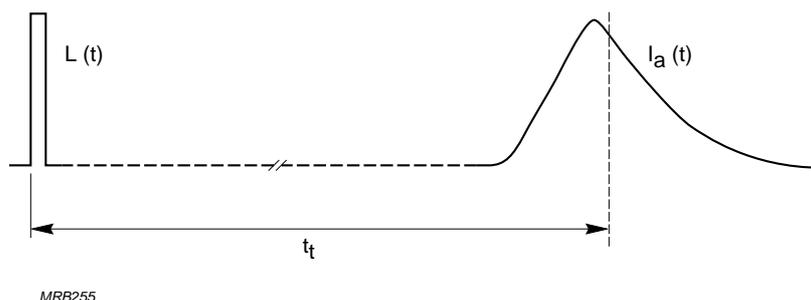


Fig.A4.4 Pulse transit time t_t .

A4.2.1 Delta-function light pulse

It is convenient to normalize the illumination function with respect to the mean number of photoelectrons per pulse, $\bar{n}_{k,i}$, thus

$$L(t) = \bar{n}_{k,i} \delta(t)$$

The pulse transit time can be defined as the interval separating the occurrence of the light pulse and a definable reference point on the ensuing current pulse – say, its centre of gravity. Let $R(t)$ represent the probability density of the transit time t_t when the tube operates repeatedly under single-electron conditions (one photoelectron per pulse, Fig.A4.5); the mean transit time is \bar{t}_t and the variance $\sigma_{tt(1)}^2$. Now assume that the cathode emits $n_{k,i} \gg 1$ electrons per pulse, and that each of these gives rise to an identical anode pulse; the sum of the $n_{k,i}$ anode pulses is then the total response of the tube. The transit times of the individual pulses obey the $R(t)$ probability density; however, they conform to it exactly only if their number is infinite. In actuality, therefore, the multi-electron pulse transit time fluctuates about a mean value that must be determined with reference to a large number of light pulses.

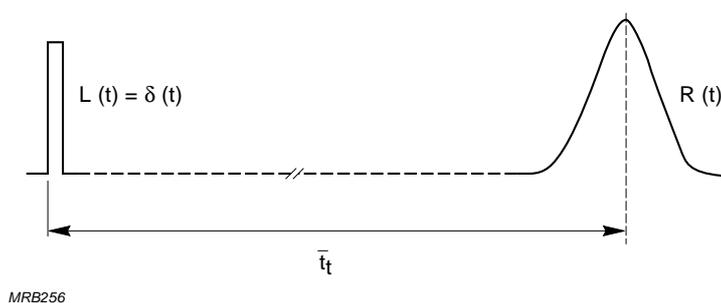


Fig.A4.5 Pulse transit-time probability distribution

To determine the variance of the fluctuations when each light pulse releases $n_{k,i}$ photoelectrons, consider the period during which the $n_{k,i}$ elementary responses arrive at the anode to be divided into infinitesimal intervals dt (Fig.A4.6). The fraction $dn_{k,i}$

of current pulses arriving at the anode during a single interval dt fluctuates from one interval to the next because the $n_{k,i}$ elementary pulses are randomly divided amongst the intervals; and also because the number $n_{k,i}$ itself fluctuates from one light pulse to the next. The first of these factors predominates; if the number $n_{k,i}$ is large enough, its fluctuations are negligible compared with those of the fraction $dn_{k,i}$ arriving during a single interval dt . To simplify calculation without affecting the generality of the results, we can therefore replace $n_{k,i}$ by its mean value $\bar{n}_{k,i}$ and $dn_{k,i}$ by $n_k(t)$. A variation $\Delta n_k(t)$ in this number shifts the centre of gravity of the transit time distribution by an amount

$$\Delta t_t = \frac{\Delta n_k(t)}{\bar{n}_{k,i}} (t - \bar{t}_t) \text{ if } \Delta n_k(t) \ll \bar{n}_{k,i}$$

or,

$$\sigma_{tt}^2 = \frac{\sigma_{nk}^2(t)}{\bar{n}_{k,i}^2} (t - \bar{t}_t)^2$$

where σ_{tt}^2 and $\sigma_{nk}^2(t)$ are the variances of t_t and $n_k(t)$.

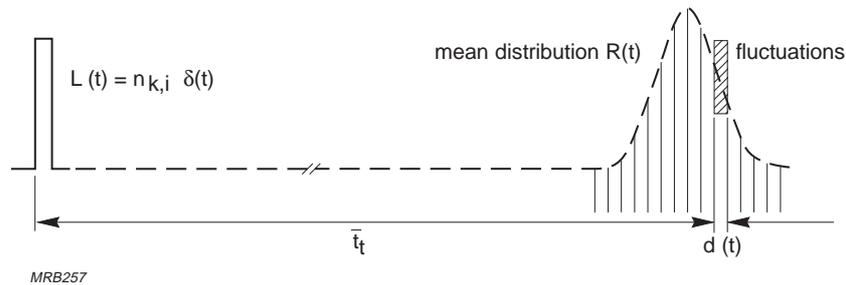


Fig.A4.6 Transit-time distribution of $n_{k,i}$ elementary response pulses

If we assume that the fluctuations in the number $n_k(t)$ during the intervals dt are absolutely independent of each other, then $n_k(t)$ obeys a Poisson law and we can write

$$\sigma_{nk}^2(t) = \bar{n}_{k,i} R(t) dt$$

whence

$$\sigma_{tt}^2(t) = \frac{R(t) dt}{\bar{n}_{k,i}} (t - \bar{t}_t)^2$$

or, by integrating over all the intervals dt :

If we assume $R(t)$ to be gaussian with a mean value \bar{t}_t and a variance $\sigma_{tt(1)}^2$

$$\sigma_{tt}^2(n_{k,i}) = \frac{1}{\bar{n}_{k,i}} \int_{-\infty}^{\infty} (t - \bar{t}_t)^2 R(t) dt \quad (\text{A4.15})$$

$$R(t) = \frac{1}{\sigma_{tt(1)}\sqrt{2\pi}} \exp \left\{ -\frac{(t - \bar{t}_t)^2}{2\sigma_{tt(1)}^2} \right\} \quad (\text{A4.16})$$

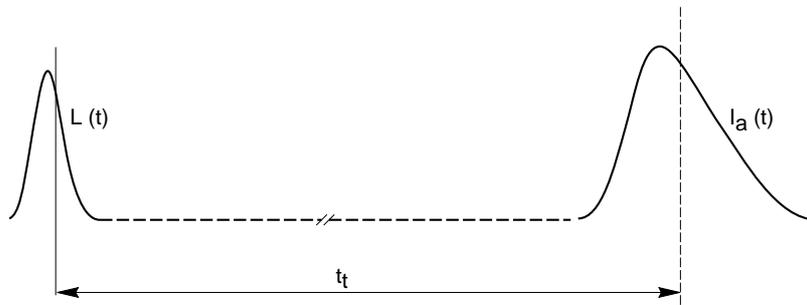
and Eq.A4.15 becomes:

$$\sigma_{tt}^2(n_{k,i}) = \frac{\sigma_{tt(1)}^2}{\bar{n}_{k,i}} \quad (\text{A4.17})$$

That is: when $n_{k,i}$ photoelectrons are emitted in response to a delta-function light pulse, the transit-time variance is equal to the single-electron pulse transit time variance divided by $n_{k,i}$.

A4.2.2 Arbitrary light pulse

For a light pulse of measurable width, transit time must be referred to definable points on both the light pulse and the ensuing anode pulse (Fig.A4.7).



MRB259

Fig.A4.7 Reference points for transit-time measurement must be clearly definable

Consider first the single-electron case. Relative to the chosen reference point, the instant of photoemission fluctuates from pulse to pulse with a density distribution corresponding to the illumination function L(t) which, in this case, may be normalized to unity.

$$\int_0^{\infty} L(t) dt = 1 \quad (\text{A4.18})$$

As photoemission and electron multiplication are separate events in cascade, the density distribution of the times after the reference point on the light pulse when anode pulses occur is given by the convolution product

$$R^*(t) = L(t)*R(t) \quad (\text{A4.19})$$

If $L(t)$ and $R(t)$ are approximately gaussian, with standard deviations σ_L and $\sigma_{tt(1)}$, then $R^*(t)$ is also gaussian and its variance is

$$\sigma_{tt(1)}^{*2} = \sigma_L^2 + \sigma_{tt(1)}^2 \quad (\text{A4.20})$$

Now consider the case when $n_{k,i} \gg 1$ photoelectrons are emitted per light pulse. As before, these can be assumed to give rise to $n_{k,i}$ elementary anode pulses the transit times of which will obey the density distribution of Eq.A4.19. The reasoning applied in the case of the delta-function light pulse can then be extended to the arbitrary light pulse by substituting $R^*(t)$ for $R(t)$, and Eq.A4.15 can be rewritten

$$\sigma_{tt}^{*2}(n_{k,i}) = \frac{1}{\bar{n}_{k,i}} \int_{-\infty}^{\infty} (t - \bar{t}_t)^2 R^*(t) dt \quad (\text{A4.21})$$

and the transit time variance is

$$\sigma_{tt}^{*2}(n_{k,i}) = \frac{\sigma_L^2 + \sigma_{tt(1)}^2}{\bar{n}_{k,i}} \quad (\text{A4.22})$$

The foregoing reasoning disregards the effect of gain fluctuations on the shape of the anode pulse and, hence, the position of a reference point such as its centre of gravity. It has been shown that, for large values of $\bar{n}_{k,i}$ gain fluctuations with a relative variance v_G increase both the delta-pulse and the arbitrary-pulse transit time variances by a factor $(1 + v_G)$.

CHAPTER 5

SUPPLY AND OPERATING ADVICES

5.1	Applying voltage	5.7	Supply for several tubes
5.1.1	Polarity		
5.1.2	Rate	5.8	Dark current reduction
5.2	Voltage dividers	5.8.1	Cooling
5.2.1	Types of voltage distribution	5.8.2	Reducing effective cathode area
5.2.2	Resistive dividers		
5.2.3	Zener-diode dividers	5.9	Magnetic shielding
5.2.4	Multiple power supplies		
5.2.5	Decoupling	5.10	Gating
5.2.6	Wiring precautions		
5.3	Output connections	5.11	Drift compensation
5.3.1	Anode resistor	5.11.1	Anode sensitivity drift
5.3.2	Output cable high-voltage connection	5.11.2	Temperature-dependent dark-current drift
5.3.3	Signal taken from dynode	5.12	Pin connections and safety precautions
5.4	Anode load		
5.4.1	Continuous operation	5.13	Signal processing
5.4.2	Pulse operation	5.13.1	Operational amplifiers
		5.13.2	Separating the signal from the noise
5.5	Operating range		
5.5.1	Gain and dark current characteristics	5.13.3	Detection at very low light levels
5.5.2	Choice of operating conditions: continuous operation		
5.5.3	Choice of operating conditions: pulse operation		
5.6	Gain adjustment		

SUPPLY AND OPERATING ADVICES

Correct use of a photomultiplier calls for observance of certain rules and circuit techniques. Those described here are indicative of present-day practice and sufficient to serve as a working guide.

5.1 Applying voltage

5.1.1 Polarity

There are two ways of applying the high voltage to a photomultiplier:

- *positive polarity*, with the cathode earthed and the anode at high positive potential (Fig.5.1)
- *negative polarity*, with the anode earthed and the cathode at high negative potential (Fig.5.2).

The choice depends on the application.

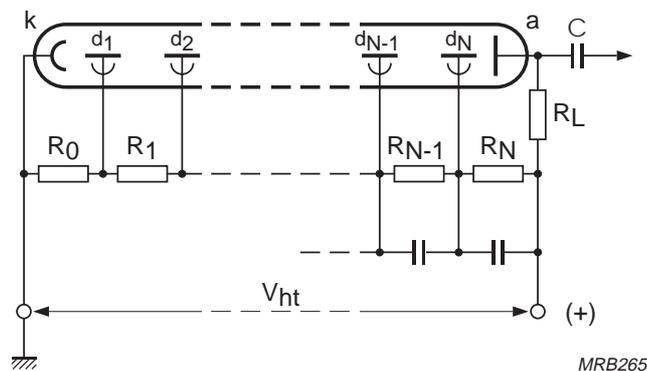


Fig.5.1 Positive-polarity voltage supply

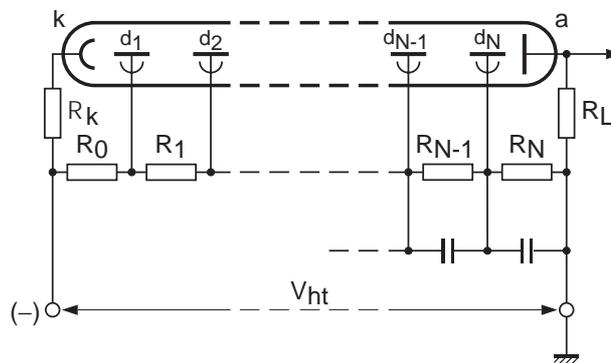


Fig.5.2 Negative-polarity voltage supply

For pulse counting, positive polarity is usually preferred; the capacitor C isolates the measuring circuits from the high voltage. Positive polarity is convenient because it does not necessitate insulating the tube from its surroundings. Magnetic screens and the coatings of scintillators should, however, be kept at photocathode (earth) potential.

For applications involving detection of continuous flux or very brief pulses, where use of a coupling capacitor would be unpractical, negative polarity has to be used. This necessitates special precautions to minimize its effect on dark current and to guard the tube against the potentially destructive effect of voltage gradients across the glass.

With tubes whose dark current is low (those with alkali cathodes, for example), the time required for the dark current to settle after switch-on may delay measurements for half an hour or more. With all tubes, insulation defects or capacitive currents between the cathode and nearby earthed parts may make the dark current high and unstable. There are two ways to guard against this:

- mount the tube free-standing, supported only by its socket and out of contact with surrounding parts. At normal atmospheric pressure a few millimetres of air space is enough to afford adequate insulation.
- If the tube must be in contact with surrounding parts (e.g. for cooling), insulate it from them (Fig.5.4). Coat the envelope with a conductive layer (e.g. metallic paint) to neutralize any potential gradients on the glass that might give rise to leakage currents, and connect the conductive coating to cathode potential via a safety resistor of, say, 10 M Ω . Enclose the conductive coating in an insulating coating with a high insulation resistance. If good heat transfer through the coating is required a material with high thermal conductivity should be used.

If the environmental humidity is very low and the resistance of the insulation (available) is high enough, the conductive coating can be dispensed with. This requires an insulation resistance of at least $10^{15}\Omega$ (e.g. a few tenths of a millimetre of Teflon or a few millimetres of Silastene, curve 4 in Fig.5.3).

Figure 5.3 illustrates the dark-current behaviour of a magnetically shielded tube (Fig.5.4) under various conditions of insulation. Polarity was negative but the voltage was increased gradually to its nominal value to limit the initial dark-current transient. It is clear that the better the insulation, the better the dark current stability.

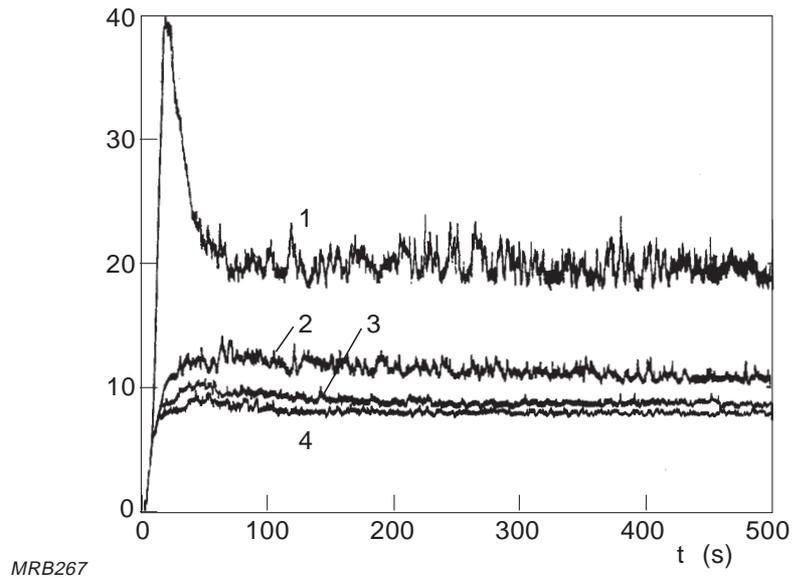


Fig.5.3 Stabilization of anode dark current I_{a0} (arbitrary scale) following gradual application of voltage in negative polarity, showing the effect of different ways of insulating the tube from its magnetic shield.

- 1: tube without conductive coating, wrapped with insulating tape
- 2: tube with conductive coating connected to cathode potential, wrapped with insulating tape
- 3: tube with conductive coating connected to cathode potential, insulated with 2 mm of Silastene
- 4: tube insulated with 4 mm of Silastene (with or without conductive coating)

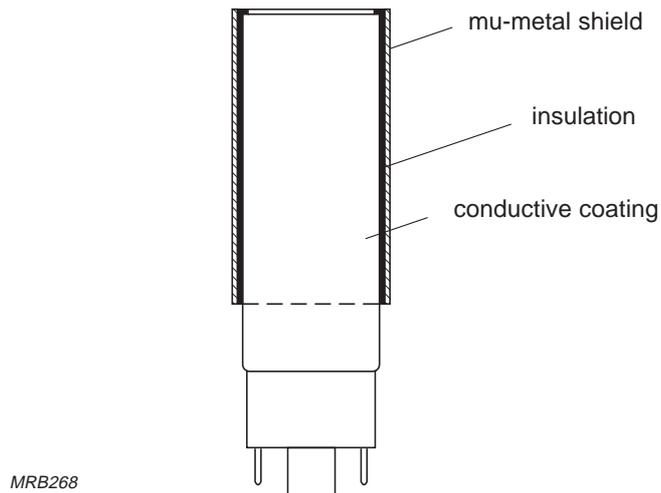


Fig.5.4 Magnetic shielding, insulation, and conductive coating

Figure 5.5 shows the dark current behaviour of a negatively-connected tube with differing provisions for heat transfer to the magnetic shield. Heat transfer coefficients in both cases were approximately equal.

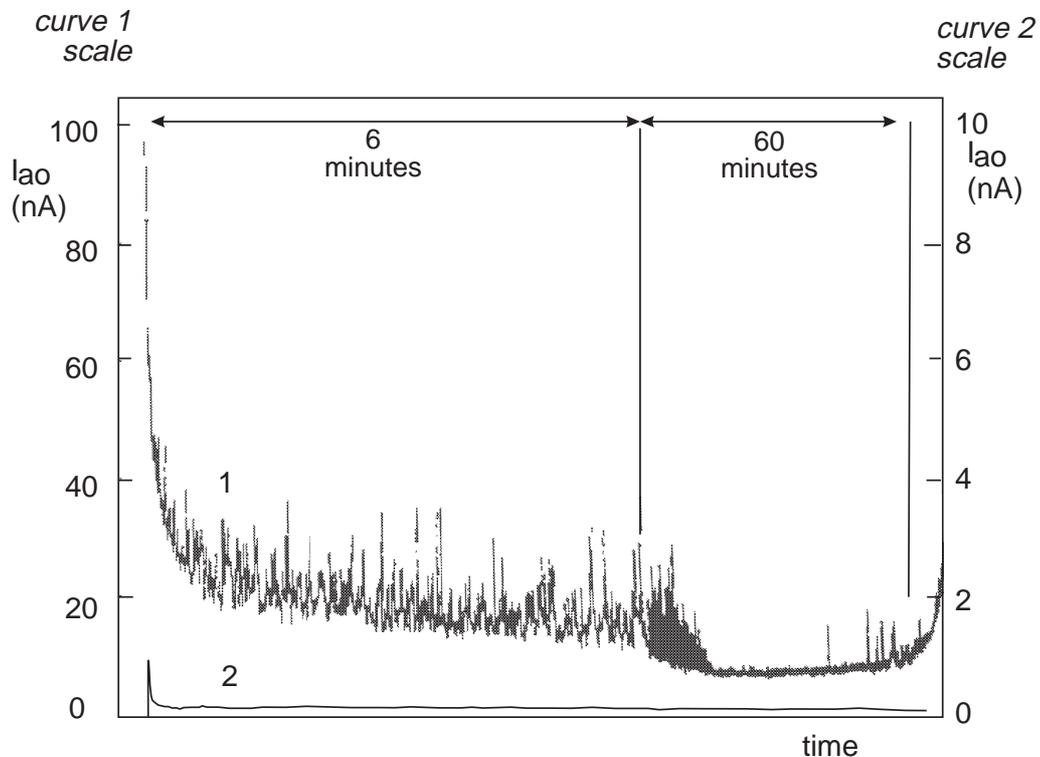
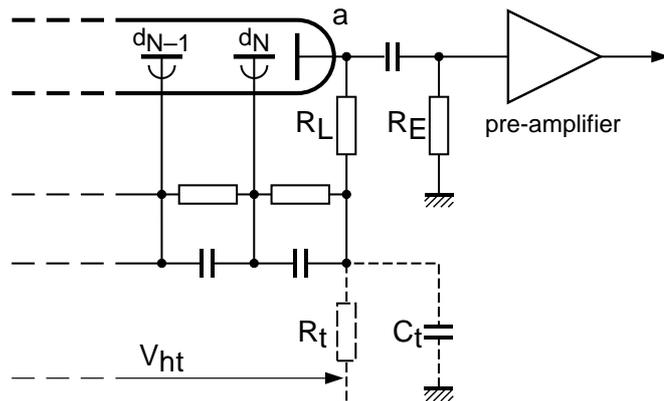


Fig.5.5 Dark-current stabilization of a negative-polarity connected tube in thermal contact with its magnetic shield, showing the effect of different ways of achieving thermal contact and electrical insulation.

- 1: tube without conductive coating, thermal contact between tube and shield via crumpled aluminium foil
- 2: tube with conductive coating connected to cathode potential, thermal contact between tube and shield via 3 mm of thermally conductive insulation

5.1.2 Rate of voltage application

Positive polarity If the high voltage is applied abruptly to a photomultiplier connected in positive polarity (Fig.5.6), an initial pulse of amplitude $V_{ht}R_E/(R_L + R_E)$ is coupled through the capacitor to the preamplifier input and causes damage. To avoid this it is advisable to use a circuit decoupling network (shown dotted in Fig.5.6) with a time constant R_tC_t of at least one second. An alternative is to shunt a protection diode across the resistor R_E to eliminate all positive-going pulses at the amplifier input.



MRB270

Fig.5.6 Network for decoupling the high-voltage switch-on transient from the amplifier input

Negative polarity If voltage is applied abruptly to a tube connected in negative polarity, the amplitude of the initial dark-current transient (Fig.5.5, trace 1) may be high enough to damage sensitive measuring apparatus. Applying the voltage gradually reduces the transient or may even eliminate it. The RC time constant should be a few seconds.

5.2 Voltage dividers

The choice of voltage divider, including the type of voltage distribution, depends on:

- the application of the tube; e.g. continuous, pulse, or high-current operation, etc.
- the performance required; e.g. gain, linearity, timing, stability, etc.

5.2.1 Types of voltage distribution

Recommended voltage distributions are given in the data sheets for each tube. There are three main types, designated A, B and C.

Type A (*equal steps*, Fig.5.7(a)) Voltages between all the iterative dynodes are equal. This distribution gives maximum gain for a given supply voltage and is particularly suitable for photometry and nuclear spectrometry applications.

Type B (*progressive*, Fig.5.7(b)) Interdynode voltages increase progressively in the anode direction, becoming 8 to 10 times as high in the last stages as in the first. This distribution makes it possible to obtain anode pulses of several hundred milliamperes peak with good linearity. Gain, however, is much lower than with type A distribution for the same total high voltage.

Type C (*intermediate*, Fig.5.7(c)) Interdynode voltages increase only in the last stages. Time characteristics are optimized; gain and pulse linearity are also satisfactory. So, type C distribution is particularly suitable for fast response photomultipliers.

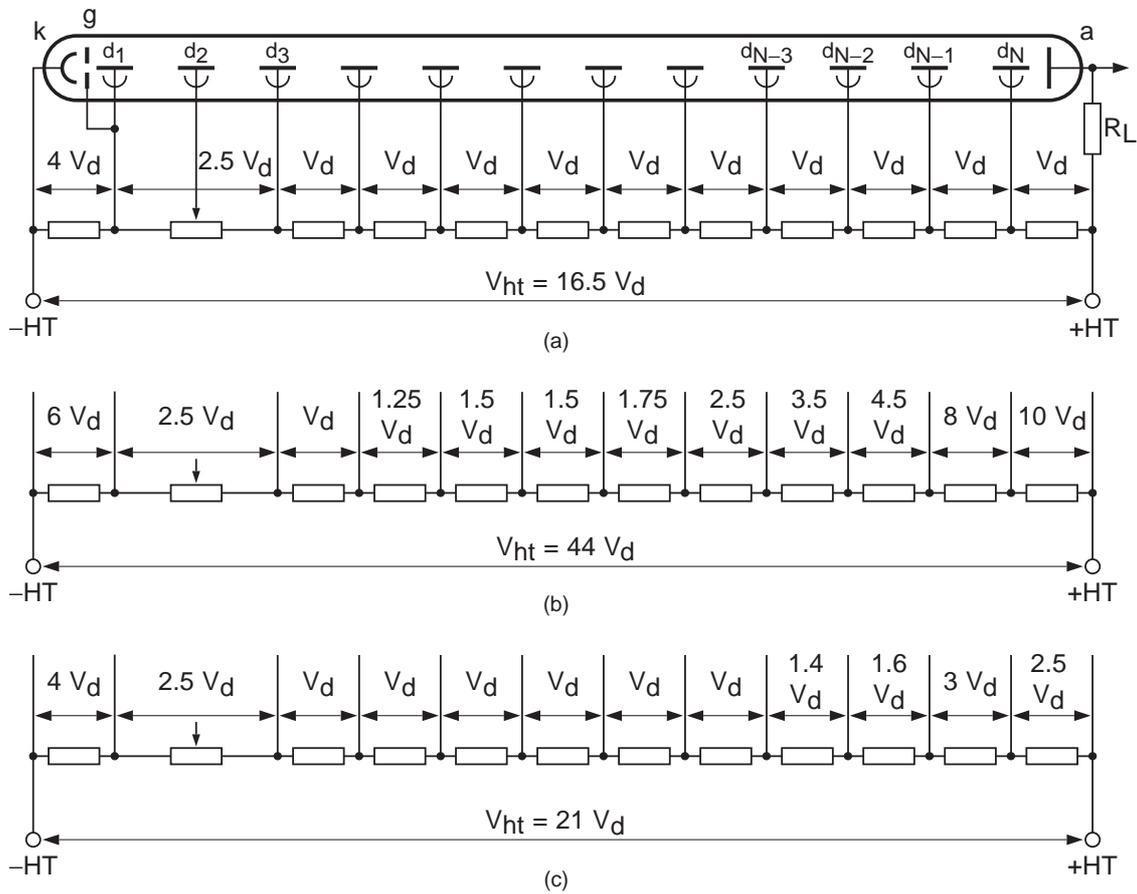


Fig.5.7 Voltage dividers: (a) type A equal step, (b) type B progressive, (c) type C intermediate

Table 5.1 compares the gain and pulse linearity of a fast response photomultiplier for each type of distribution.

Table 5.1
Gain and pulse linearity of a fast response photomultiplier,
with three types of voltage distribution

Type of voltage distribution	Gain $V_{ht} = 2500 \text{ V}$	Linear within 2% for current pulse amplitudes up to
A	1.2×10^8	40 mA
B	0.7×10^6	250 mA
C	2×10^7	100 mA

Other types of distribution are sometimes offered. The type A1 distribution, for example, has a higher cathode to first-dynode voltage than the type A to ensure good collection efficiency even when the tube operates at low voltage. The gain, though, is lower than with type A distribution.

5.2.2 Resistive dividers

Design of a resistive voltage divider depends on the supply voltage, the voltage distribution, and the anticipated mean anode current I_a . To ensure that voltage variations due to anode current variations are negligible, the nominal divider current I_p must be much larger than I_a ; a good rule is

$$\frac{I_p}{I_a} \geq 100 \quad (5.1)$$

Consider a type C voltage distribution with $V_{ht} = 2500 \text{ V}$, (Fig.5.7(c)) an interdynode voltage increment V_d , and a division ratio such that there are precisely 21 such increments. Let the maximum anticipated mean anode current I_a be $10 \mu\text{A}$. The divider current I_p must then be at least 1 mA, which means a total divider resistance of $2.5 \text{ M}\Omega$ and an incremental resistor value of $2.5 \times 10^6 / 21 \approx 120 \text{ k}\Omega$.

The resistors must be properly rated for power and voltage. The latter is important because some of them must withstand several hundred volts continuously. Tolerances should be not greater than 5%.

The ratio specified by Eq.5.1 represents a minimum for maintaining good linearity. Two other considerations limit the maximum value of the ratio.

- Heat due to dissipation in the divider ($I_p^2 R$) can cause an increase in the dark current, especially if the divider and tube are housed close together or with the tube vertical, cathode uppermost.
- Low divider current gives some protection against accidental overexposure of the cathode; as soon as anode current rises proportionately, gain drops abruptly and prevents the anode current from becoming excessive.

5.2.3 Zener-diode dividers

Zener diodes can be substituted for some of the resistors in the divider to keep certain inter-electrode voltages constant.

They can be used between cathode and first dynode to keep input collection efficiency constant regardless of supply voltage and gain setting; and between the dynodes of the first two multiplier stages to keep the gains of those stages constant (Fig.5.8(a)). This is useful where certain minimum voltages are necessary in the first stages because S/N or PHR is critical but the overall gain need not be high. It cannot be done, however, in those fast-response tubes that incorporate an accelerating electrode internally connected to one of the higher ranking dynodes; the potential of such an electrode must be in constant proportion to the voltage between cathode and first dynode.

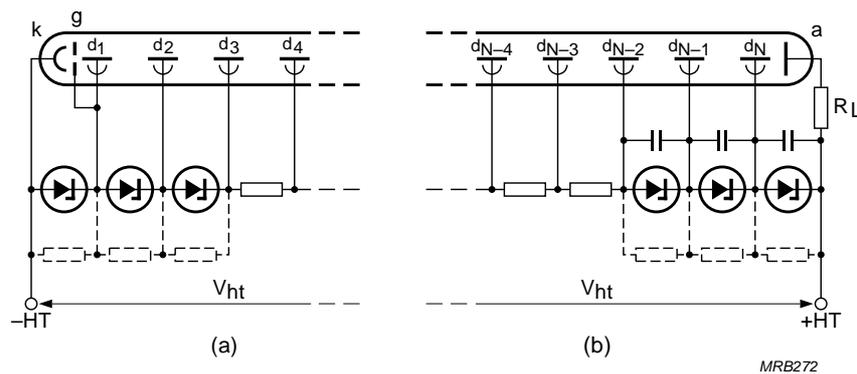


Fig.5.8 Inter-electrode voltage stabilization with zener diodes: (a) in first stages, (b) in last stages (dashed lines, protection resistors)

Another place to use zener diodes is in the last stages of the divider (Fig.5.8(b)), to stabilize the voltages there throughout a wider range of anode current variations. This

also makes it possible to accommodate smaller ratios of I_p/I_a than with a purely resistive divider (§4.5.1).

In certain applications a drawback of using zener diodes is that they limit the freedom of gain adjustment. Altering the supply voltage to adjust the gain would also alter the overall voltage distribution, for it would affect the voltages across the resistor stages but not the zener-stabilized ones. As linearity is very dependent on the overall voltage distribution, a divider with zener diodes should be designed for a specific supply voltage and that voltage should be adhered to as closely as possible. Departure from it invites the risk of either overlinearity or premature saturation. The risk is considerably less if only the very last stage is zener stabilized.

Whether zener diodes are used in the higher or lower stages of the divider, they should be shunted by resistors to protect those stages from receiving the full supply voltage in the event of a diode going open-circuit. The values of the resistors should be 2 to 3 times what they would be in a purely resistive divider.

The temperature coefficients of the zener diodes is an important consideration. Variation of zener voltage with temperature can cause variation in gain.

Note, though zener-diode dividers offer some advantages in special situations, these advantages are sometimes over-stressed and often the best solution is a specially tailored resistive divider.

More elaborate voltage dividers including active components such as transistors are also used to cope with high mean anode currents (details can be found in the Photonis Photomultiplier catalogue).

5.2.4 Multiple power supplies

If negative-polarity connection is not objectionable, using more than one supply can ease some of the constraints on a resistive divider mentioned earlier. It enables good linearity to be obtained over a wide anode current range without excessive dissipation in the resistors. The supplies may be either in series (Fig.5.9(a)) or parallel (Fig.5.9(b)). In either case, the one connected to the cathode is a high-voltage low-current (<1 mA) supply. The other, feeding the anode and last three or four dynodes, is a low-voltage high-current supply. The terminal common to both is usually earthed and, in fact, has to be if the supplies themselves are not isolated from earth.

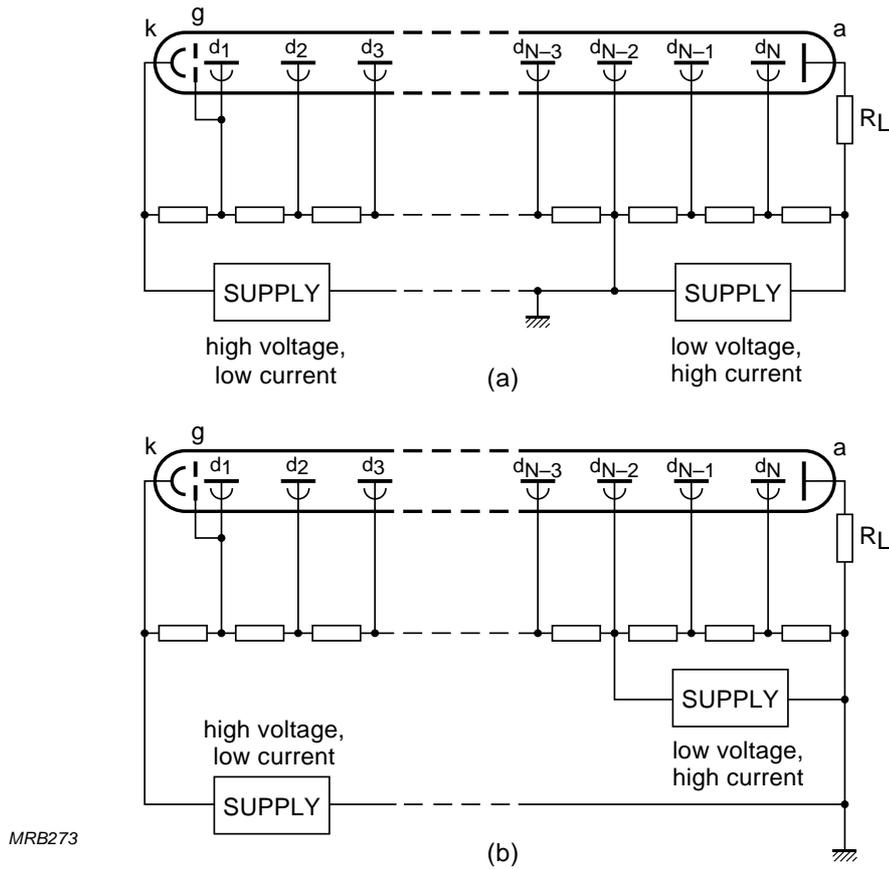


Fig.5.9 Dual voltage supplies: (a) series, (b) parallel

A parallel supply as in Fig.5.9(b) can also be used to provide additional current through zener diodes fitted to the last stages of a voltage divider. In this way the operating point of those stages can be set independently of the current through the resistive part of the divider, which can be kept low. In high-energy physics each of the last four or five stages is often provided with its own parallel supply, making it possible for the photomultiplier to accept high count rates without loss of linearity.

5.2.5 Decoupling

Provided the dynodes are adequately decoupled, instantaneous values of current in pulse operation may greatly exceed the mean value of the divider current. In observing the rule of Eq.5.1, the value taken for I_a should be the mean anode current based on the anticipated pulse amplitude and duty factor. Decoupling may be either parallel (Fig.5.10(a)) or series (Fig.5.10(b)); the former requires high-voltage capacitors which are larger and more expensive.

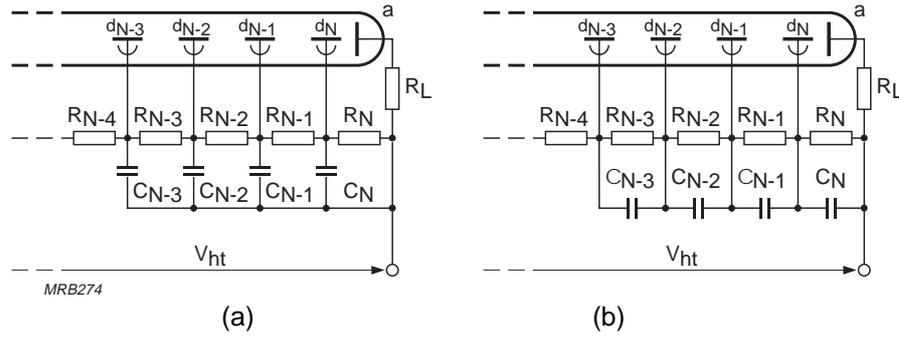


Fig.5.10 Dynode decoupling: (a) parallel, (b) series.

The decoupling capacitors act as reservoirs to restore the charge transferred by pulses passing through the tube. Let Δq_N be the pulse charge supplied by the last dynode, and ΔV_N the maximum voltage change that can be tolerated on that dynode. The required value of capacitor C_N is then given by

$$C_N \geq \frac{\Delta q_N}{\Delta V_N} \quad (5.2)$$

The charge added per stage increases in direct proportion to the gain; therefore, the required value of C_{N-1} is

$$C_{N-1} = \frac{C_N}{g_N}$$

where g_N is the gain due to dynode d_N . If all stages have equal gain, the general expression for the required capacitance in the i -th stage is simply

$$C_i = \frac{C_N}{g^{N-i}} \quad (5.3)$$

Working backwards from the last stage to the first, this rule should be applied until the value it gives for C_i is comparable to the stray capacitance of the stage (usually about 20 pF).

The foregoing is based on the assumption that the voltage changes ΔV_i on the dynodes occur independently of each other, but that is not always so. Consider the case where the pulse duration is large compared with the overall transit time, and the dynodes are series decoupled. The voltage changes can then be fed back down the decoupling network in such a way that

$$\Delta V_i = \sum_{j=i}^N \Delta V_j$$

For adequate decoupling the capacitance in each stage except the last must then be increased by a safety factor about equal to the stage gain. Equation 5.3 then becomes

$$C_i' = g \frac{C_N}{g^{N-i}} = \frac{C_N}{g^{N-i-1}} \quad (5.4)$$

Example A photomultiplier adjusted for an overall gain of 3×10^5 detects NaI(Tl) scintillation pulses due to γ -radiation from a ^{60}Co source. The full width at half maximum of the resulting anode pulses is about $0.3 \mu\text{s}$, and their amplitude about 1 mA; the pulse charge is therefore

$$\Delta q_N \approx I_a t_w = 0.3 \times 10^{-9} \text{ C} \quad (5.5)$$

If the maximum tolerable change in last-dynode voltage is $\Delta V_N = 1 \text{ V}$ (i.e. about 1% of the interdynode voltage), the minimum decoupling capacitance required at that dynode is

$$C_N \geq \frac{\Delta q_N}{\Delta V_N} = 0.3 \text{ nF}$$

The decoupling capacitors recharge with a time constant determined by the divider resistances. It is not necessary for each capacitor to recharge fully before every pulse but only for it to recover the charge lost during the preceding pulse. What is important is that the amplitude of the capacitor voltage variation in each stage should not exceed the tolerable value of ΔV_i in that stage. This condition is satisfied when the rule given in Eq.5.1 is observed; that is, when

$$I_a \leq \frac{I_p}{100} = \frac{V_{ht}}{100 \sum_{i=1}^N R_i} \quad (5.6)$$

where V_{ht} is supply voltage and R_i the voltage divider resistance at each stage. Together, Eq.5.2 and (5.3) define the minimum capacitance needed to decouple each stage, and Eq.5.6 the minimum divider current to restore the capacitor charges. Apart from considerations of bulk and cost, there is no strict upper limit to the decoupling capacitances.

The examples that follow illustrate the calculation of decoupling capacitance. For the sake of comparison the following conditions are common to all three:

– peak anode pulse amplitude, \hat{I}_a	1 mA
– full width at half maximum, t_w	0.3 μ s
– charge per pulse, $\Delta q_N = \hat{I}_a t_w$	0.3×10^{-9} C
– number of pulses per second, n	10^4
– multiplier gain per stage, g	3
– tolerable voltage variation at last dynode, ΔV_N	1 V

The mean anode current $I_a = n\Delta q_N$ and the conditions to be satisfied are those of Eqs 5.1 and 5.2.

Fixed-frequency pulses From Eq.5.1

$$I_p \geq 100I_a = 100n\Delta q_N = 100 \times 10^4 \times 0.3 \times 10^{-9} = 0.3 \text{ mA}$$

From Eq.5.2

$$C_N = \frac{\Delta q_N}{\Delta V_N} = 0.3 \times 10^{-9} = 0.3 \text{ nF}$$

whence, from Eq.5.4

$$C_{N-1} = 0.3 \text{ nF} \quad C_{N-2} = 0.1 \text{ nF} \quad C_{N-3} = 33 \text{ pF}$$

and $C_{N-4} = 11 \text{ pF}$, which is comparable with the stray capacitance and can be neglected, as can all lower stages. For C_N to C_{N-3} four 1 nF capacitors can be used for simplicity.

Random-frequency pulses. Let $\bar{n} = 10^4/\text{s}$ be the mean about which the frequency fluctuates, and assume that the maximum instantaneous frequency is $3\bar{n}$. Two methods of approach are possible.

1. As the charge per pulse is the same as before, the same decoupling capacitances will suffice. If the calculated values (0.3 nF, 0.3 nF, 0.1 nF, 33 pF) are chosen, however, the divider current will have to be tripled to allow for the instantaneous excursions of pulse repetition frequency to $3\bar{n}$
2. The mean anode current is the same as before, so it should be practical to work with the same divider current. It is, but the capacitors will have to supply three times as much charge during the instantaneous frequency excursions. This leads, as before, to the choice of 1 nF capacitors and is preferable to tripling the divider current; in this case five are required, one each for C_N to C_{N-3} and another for C_{N-4} (calculated value 0.33 nF).

Pulse bursts Consider bursts of 10 ms containing 10^4 pulses and recurring once a second. For simplicity, each burst can be regarded as a single 10 ms pulse conveying a total charge

$$\Delta q_N' = 10^4 \Delta q_N = 3 \times 10^{-6} \text{ C}$$

and the anode current is, as before, 3 μA . The required divider current is therefore 0.3 mA. However, the decoupling capacitance C_N becomes

$$C_N = \frac{\Delta q_N'}{\Delta V_N} = 3 \mu\text{F}$$

Applying Eq(5.4), this leads to a requirement for capacitors in twelve stages before the indicated decoupling capacitance becomes comparable with the stray capacitance of the stage. Fortunately, though, it is not necessary to decouple stages in which the dynode current is less than 1% of the anode current. In the present example (gain per stage, $g = 3$), this is the case for all stages lower than $N - 4$. Using standard values, the actual capacitor requirement is therefore

$$C_N = 3.3 \mu\text{F} \quad C_{N-1} = 3.3 \mu\text{F} \quad C_{N-2} = 1 \mu\text{F} \quad C_{N-3} = 330 \text{ nF}$$

$$C_{N-4} = 100 \text{ nF}$$

Even so, some of these are inconveniently large for capacitors that must have a high working voltage.

One practical alternative is to use a higher current (e.g. $I_p \approx 3 \text{ mA}$) in a divider with zener diodes in the last four stages. Five 10 nF capacitors would then suffice.

Another alternative, which makes it possible to accept even large pulse burst, is to use a divider with separate supplies for the last four stages.

The networks of Fig.5.10 are not the only practical ones. Figure 5.11 shows a 2-and-2 series arrangement that is often used when many dynodes have to be decoupled and capacitor lead lengths must be kept to an absolute minimum (alternate dynode pinning).

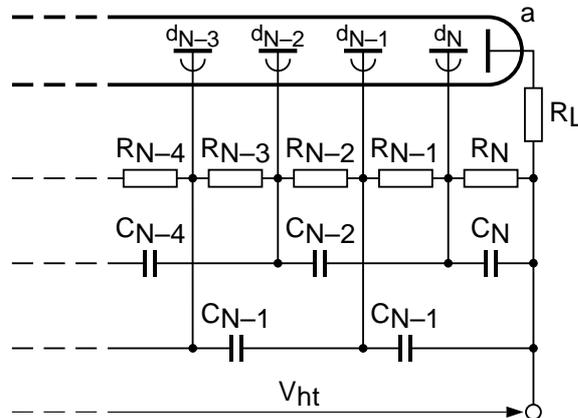
5.2.6 Wiring precautions

When the tube is connected in positive polarity, observe the usual wiring rules for high-voltage/low-current; take especial care over insulation of the output stage. Keep decoupling capacitor leads short to minimize stray inductance. (Voltage divider resistor leads are not so critical; a printed wiring board may be used).

When the tube is connected in negative polarity take especial care over insulation of the cathode connection and, if the best possible time characteristics are required, observe the following additional rules:

- use the 2-and-2 decoupling network of Fig.5.11.
- connect the last two decoupling capacitors to the braided sheath of the coaxial cable and earth the sheath there as well as the output end. If other coaxial cables are used (e.g. for double anode output, dynode output), earth them at the same point. Terminate coaxial cables in their characteristic impedance.
- some fast response tubes incorporate integral damping resistors, in the base (see Fig.5.12). If the type used does not, connect such resistors externally between the last two dynodes and their decoupling capacitors. Use non-inductive 50 Ω resistors.

If the socket used has pin contacts, it can be mounted direct to a printed wiring board. This considerably simplifies wiring of the voltage divider.



MRB275

Fig.5.11 Series decoupling of alternating dynode pairs

5.3 Output connections

5.3.1 Anode resistor

Whether the tube is connected in positive or negative polarity, the anode potential must be fixed.

If the tube is connected in negative polarity and direct coupled to the measuring apparatus (Fig.5.13(a)), the anode potential is clamped by the internal resistance of the apparatus. However, if the output is disconnected even briefly while the high voltage is still applied, the anode will acquire a negative charge which may damage the apparatus when connection is restored. Therefore it is advisable to fit a protection resistor (dotted in Fig.5.13(a)) between the anode and earth. As it is shunted across the high internal resistance of the photomultiplier, the protection resistance must also be high. The value chosen depends mainly on the load circuit and is typically $\geq 10 \text{ k}\Omega$.

If the tube is connected in positive polarity and capacitively coupled to the measuring circuitry (Fig.5.13(b)), a resistor between the anode and the positive terminal of the high voltage supply is essential; this resistor can also constitute the anode load. Once again, the resistance must be reasonably high and depends mainly on the input impedance of the measuring circuitry.

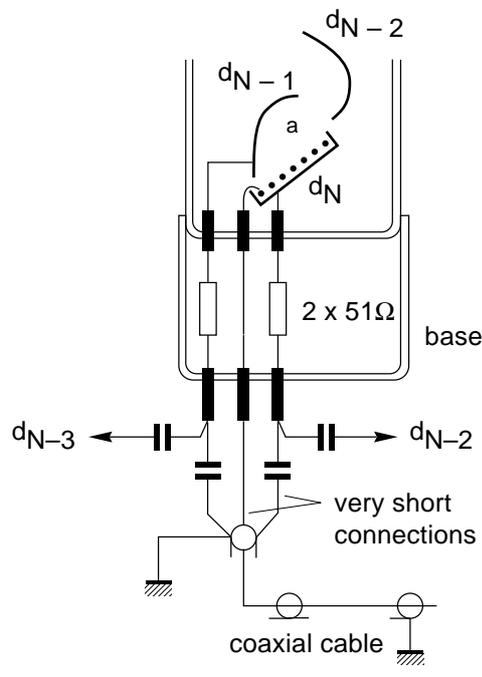


Fig.5.12 Connection of damping resistors and decoupling capacitors to the last two dynodes

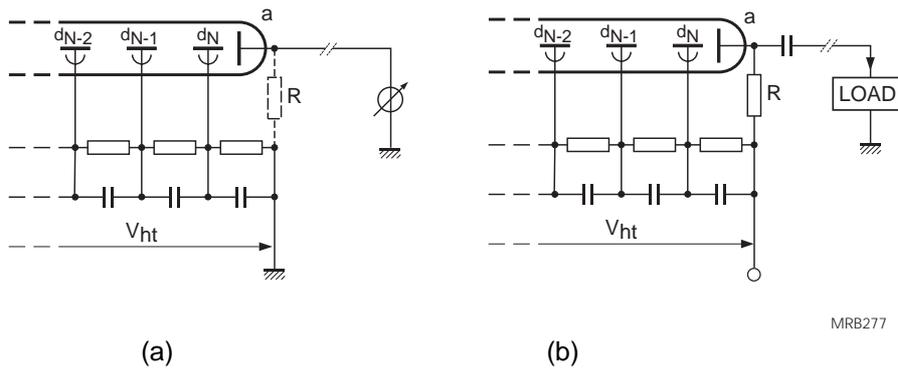


Fig.5.13 Fixing the anode potential with the tube connected in (a) negative, (b) positive polarity

5.3.2 Output cable high-voltage connection

Figure 5.14 shows two ways of using the output cable for connection of the high voltage supply. In Fig.5.14(a) resistors R_1 and R_2 , together with the capacitance of the cable, form the anode load; if the signal is to be integrated, this is acceptable. If the cable is long, the alternative shown in Fig.5.14(b) can be used, with the step-down transformer matched to the characteristic impedance of the cable.

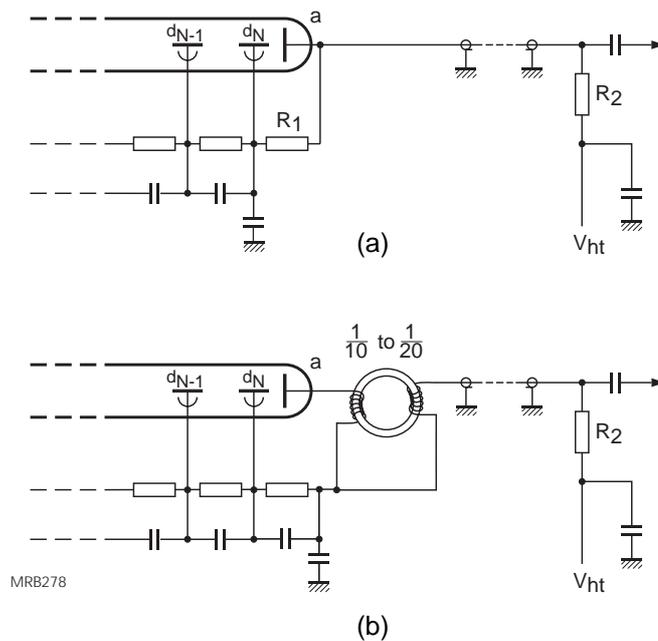


Fig.5.14 Voltage supply via the output cable

5.3.3 Signal taken from dynode

There are two cases in which a signal may have to be taken from a dynode:

- when the measurement to be made requires a signal synchronous with the anode signal
- when it is desirable to limit gain by using fewer stages.

Procedures applicable to the two cases are different.

Synchronous dynode signal. The requirement for a synchronous signal from a dynode usually arises in connection with the detection of very short pulses. The signal may be required either for synchronizing an instrument or for supplying additional charge or amplitude information. The problem is to obtain the required signal without disturbing the anode signal. Figure 5.15 shows two solutions. To obtain an amplitude comparable with that at the anode, the signal is usually taken from the last dynode but need not be if amplitude is not a governing consideration.

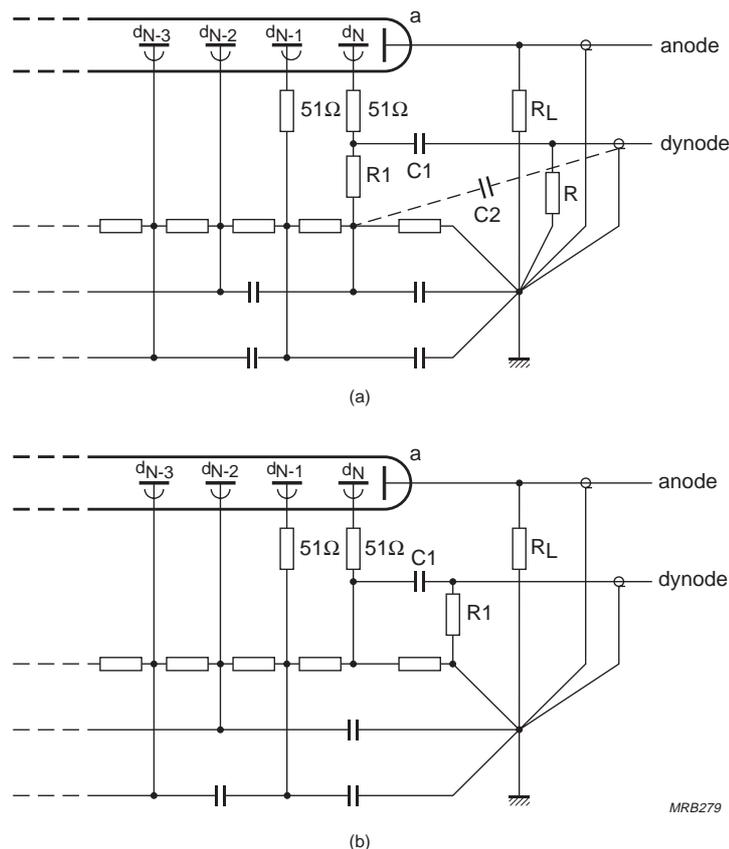


Fig.5.15 Taking an auxiliary signal from the last dynode

The coupling capacitance C_1 must be such that the product $C_1 Z_c$ (where Z_c is the characteristic impedance of the coaxial cable) is much greater than the expected pulse duration, otherwise the pulses will be differentiated. Using a single earth point, as indicated, minimizes inductive effects. If, in spite of that, oscillation does occur, connecting a second capacitor, $C_2 = C_1$, as shown in Fig.5.15(a), will minimize it.

Resistance R_1 must be large compared with Z_c but not too large. If the voltage across it exceeds a few volts there is a risk of disturbing the anode signal. For $Z_c = 50 \Omega$, the usual value of R_1 is 200Ω to 300Ω . When the dynode output is not in use, terminate the cable with its characteristic impedance.

Pulses taken from a dynode are positive-going. If negative-going pulses are required, as is standard in fast nuclear instrumentation, a coaxial cable transformer can be used.

Dynode output. If the incident flux is large and high gain is not required, fewer multiplier stages can be used. The voltages at the still active dynodes need not be changed but the overall high voltage can be reduced. The problem is to obtain a linear output signal from dynodes that have not been designed for that. Figure 5.16 shows two solutions.

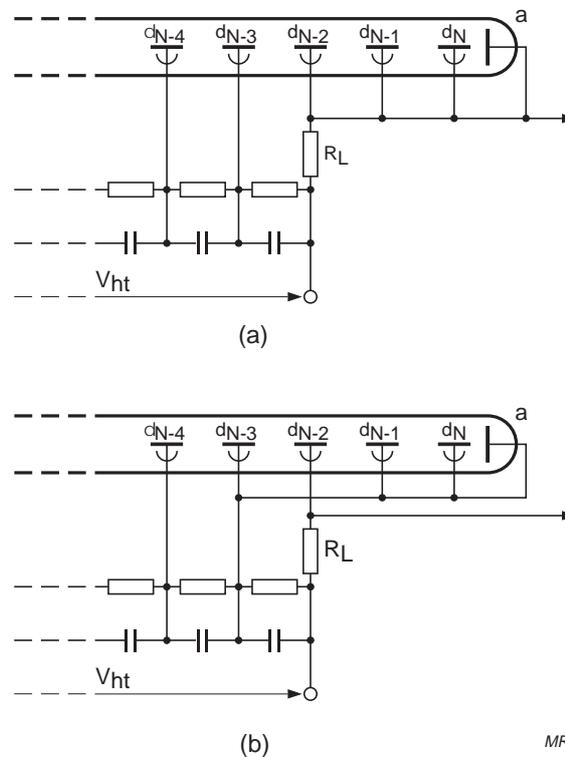


Fig.5.16 Reducing the number of active stages

In Fig.5.16(a) the inactive electrodes and the last active dynode are wired together for collecting the signal. In Fig.5.16(b) the signal is taken from one dynode, and the inactive higher-ranking dynodes and anode are connected together and taken to the potential of the next lower ranking dynode. This is an attempt to simulate the normal output geometry, in which the anode is situated between the last two dynodes. The (b) circuit is preferable to the (a) circuit for two reasons:

- it puts less stray capacitance in parallel with the load; the larger the load and the greater the bandwidth required, the more important this is.
- it gives somewhat better pulse linearity. Even so, the linearity scarcely extends to more than a few milliamperes, for dynodes, unlike the anode, are not designed for linear operation at high peak currents.

In neither circuit should unused electrodes be left floating or simply earthed; that would disturb electron trajectories elsewhere in the multiplier.

5.4 Anode load

5.4.1 Continuous operation

As long as the operating point is in the saturated current region, (Fig.5.17) a photomultiplier behaves as an almost perfect current generator; the anode current depends only on the incident flux and is completely independent of the load. Nevertheless, even in the ideal case the current through the load does lessen the voltage between the last dynode and anode. It is therefore important to check that this voltage remains high enough throughout the intended operating range to ensure complete electron collection at the anode. Figure 5.17 shows two load lines superimposed on a set of current/voltage characteristics. The extent of the linear operating range depends on the load line slope.

In practice the characteristics in the saturation region do have a slight slope (which may be either positive or negative). This accentuates the effect of the load on linearity and must be taken into account if the tube is operated in such a way as to cause excursions of more than, say, 10 V in the last-dynode-to-anode voltage.

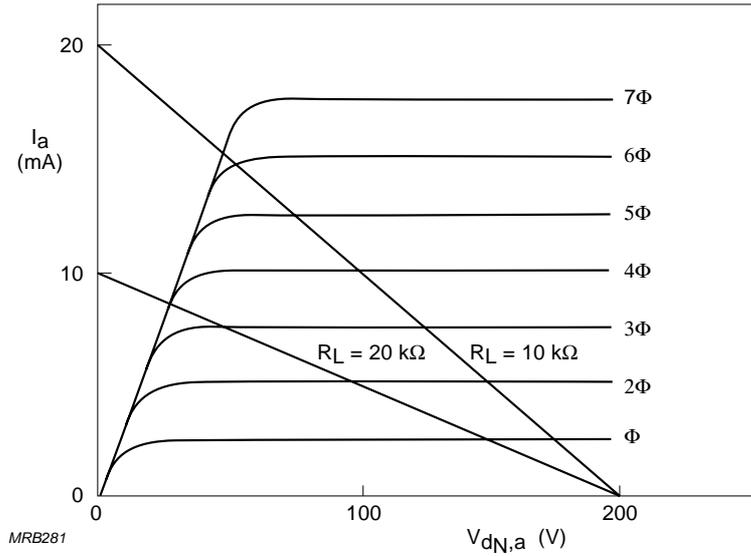


Fig.5.17 Effect of anode charge on linearity (idealized current/voltage curves)

5.4.2 Pulse operation

There are three modes of response, depending on the RC time constant of the load.

Charge mode. When the load time constant is much greater than the anode current pulse width, the pulse charge is integrated. The voltage across the load resistor is then proportional to the total charge in each pulse.

If the mean number of photons per light pulse is $\bar{n}_{p,i}$, the mean charge supplied to the anode is

$$\bar{q}_{a,i} = \bar{n}_{p,i} \rho e G \quad (5.7)$$

where ρ is the quantum efficiency of the photocathode, e the electron charge, and G the overall gain of the tube. The amplitude of the resulting voltage pulse at the anode is

$$V_a = \frac{\bar{q}_{a,i}}{C} \quad (5.8)$$

Thus, if maximum output is required at a specified value of RC, it is desirable for C to be as small as possible.

Current mode. When the time constant is much less than the anode current pulse width, the voltage across the load resistance varies as the anode current. A typical application of this mode is in the detection of very short pulses in time spectrometry (§6.3.2); the anode load is usually the 50 Ω impedance of a coaxial cable. The amplitude of the voltage pulse is

$$V_a = \frac{\bar{q}_{a,i} R}{t_w} = \frac{\bar{n}_{p,i} \rho e G R}{t_w} \quad (5.9)$$

where t_w is the full width at half maximum of the anode current pulse. Thus, for a tube operating under single-electron conditions with $\bar{n}_{p,i} \rho = 1$, $t_w = 2.4$ ns, $G = 3 \times 10^7$ and $R = 50$ Ω,

$$V_a = 100 \text{ mV}$$

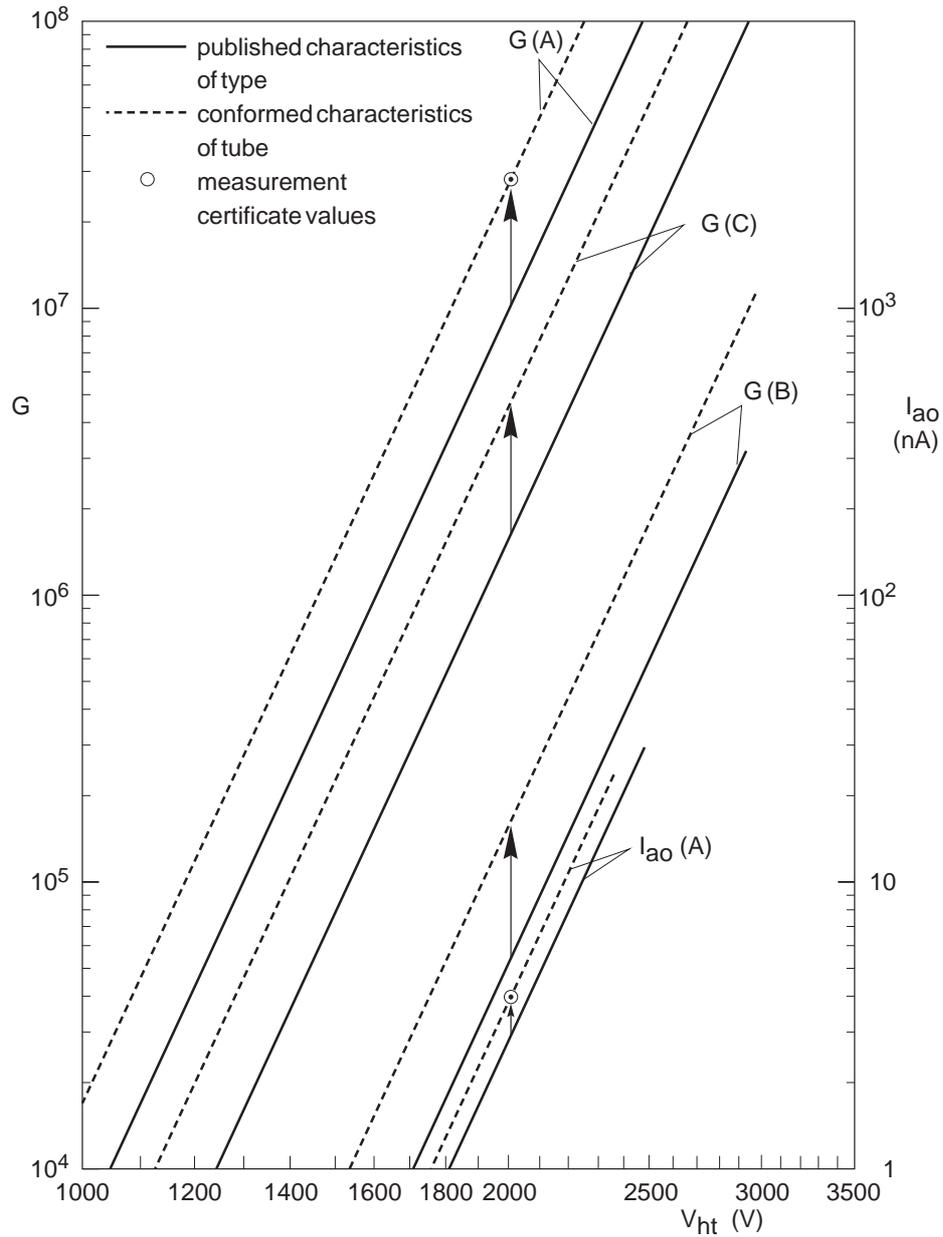
That is, emission of a single photoelectron gives rise to a mean anode pulse of 100 mV into 50 Ω.

Intermediate mode. When the time constant is comparable with the anode current pulse width, the response is intermediate between charge mode and current mode. A typical instance is in nuclear spectrometry, where the scintillator light pulse usually has an exponential decay. The shape of the resulting voltage pulse depends on the ratio of the time constant RC and the scintillator pulse decay time τ . When $RC \gg \tau$, the pulse amplitude is as given by Eq.5.8; and when $\tau \gg RC$, it is as given by Eq.5.9. Between those extremes, the response has to be determined with the aid of graphs (see Fig.6.9 to 6.13).

5.5 Operating range

5.5.1 Gain and dark current characteristics

The gain and dark current characteristics plotted in the data sheets for each type are merely typical; there is some variation from tube to tube. However, each tube is accompanied by a measurement certificate stating the gain and dark current measured at a specific applied voltage and with a specific voltage distribution. To obtain the actual characteristics of that tube it is sufficient to transfer the measured points to the corresponding data sheet graph and translate the graphed characteristics so that they pass through the transferred points. Figure 5.18 shows an example. Note that in the example the measured gain is given only for type A voltage distribution. To obtain the corrected gain characteristics for type B and C distributions the corresponding lines are shifted by the same amount as the line for type A.



MRB281

Fig.5.18 Conforming the measured gain and dark current characteristics of an individual tube to the published characteristics of the type. G(A), G(B) and G(C) are the gain characteristics with the recommended type A, B and C voltage dividers; $I_{ao}(A)$, the dark current with the recommended type A voltage divider

This procedure yields a good approximation; in reality the slopes of the characteristics also differ from tube to tube, but only by a few percent. The procedure is equally applicable to anode sensitivity characteristics when those, instead of gain, are given in the data sheets and measuring certificate.

Because it is a factory measurement taken after only a short stabilization period, the value of the dark current given in the measurement certificate tends to be pessimistic.

5.5.2 Choice of operating conditions: continuous operation

If the illumination level can be set at will, there is a temptation to set it high to obtain a good signal-to-noise ratio. How high it actually can be set depends on the cathode current level at which the effects of cathode resistivity become significant; for bialkali SbKCs cathodes that level is about 1 nA.

With the cathode current determined by the working conditions, the required gain and anode current depend on the characteristics of the measuring or signal processing circuits used. The minimum practical gain is that which corresponds to the minimum electrode voltages specified in the data sheets of the tube; these are based on considerations of linearity and minimum gain fluctuation. The maximum practical gain is usually determined by dark-current and signal-to-noise ratio considerations or pulse linearity limits.

For good gain stability with time, the mean anode current should be kept as low as possible.

5.5.3 Choice of operating conditions: pulse operation

In pulse operation the following factors affect the choice of operating range:

- detection efficiency
- energy and time resolution
- pulse linearity
- instrument triggering threshold
- maximum count rate.

Their significance varies according to whether the illumination is about constant (small dynamic range) or widely varying (large dynamic range).

Small dynamic range. This is often encountered in spectrometry applications, where optimum time or energy resolution is required. The important thing is to minimize response fluctuations by:

- optimizing the voltages at the electron-optical input system and the first two multiplier stages to minimize transit-time and gain fluctuations
- optimizing the collection of light at the photocathode.

Beyond the threshold set by the first of these recommendations, the practical minimum for the gain depends on the sensitivity or detection threshold of the signal processing circuits. The practical maximum is set either by dark current considerations or by linearity limits that come into play at pulse peaks. Beyond a certain value of applied voltage, the dark pulse rate increases faster than the gain (§3.1.3). The linearity limit may be due to either the photomultiplier or the circuits.

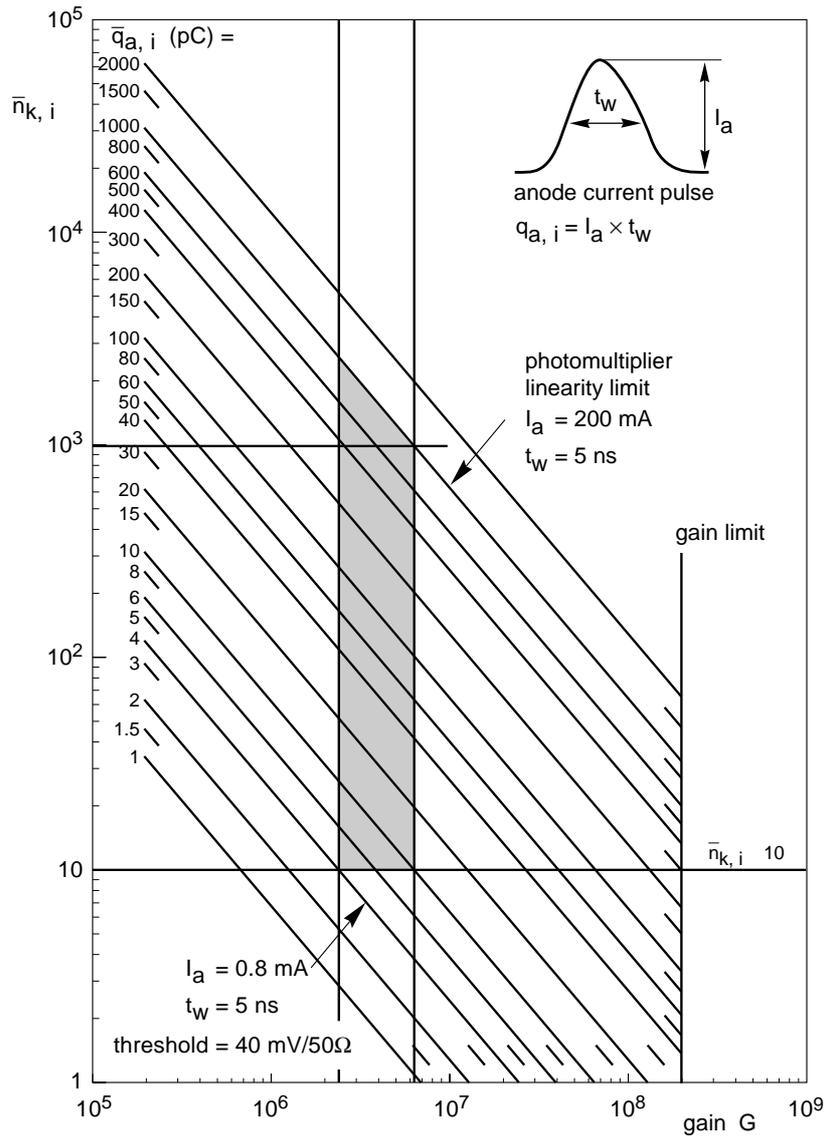
In pulse counting applications, the upper and lower gain limits can be found experimentally by plotting the variation of count rate as a function of applied voltage. This reveals a distinct counting plateau (see Fig.6.16) within which to set the operating point.

In other low dynamic-range applications, the choice of operating point may depend on other criteria, such as the bandwidth and gain of the signal processing circuits.

Large dynamic range This commonly applies in high-energy physics, where light pulses often vary over a wide range. The same fundamental rules apply as for low dynamic range, but four additional criteria outlined below are now decisive.

- The minimum number of photoelectrons per pulse that has to be detected.
- The sensitivity threshold of the electronics, which determines the amplitude or charge of the minimum detectable anode pulse.
- The maximum allowable anode pulse charge, which depends on the linearity limit of the tube at the chosen operating voltage.
- The ratio of maximum to minimum pulse amplitude or number of photoelectrons per pulse (dynamic range). If this is referred to the anode, and the minimum number of photoelectrons is very small, it should be expressed in terms of pulse charge rather than amplitude. So long as the interval between successive photoelectrons is shorter than the response pulse width of the tube, the anode pulse resembles a unique multi-electron pulse. But when the interval is longer than the response pulse width, the tube resolves the individual electrons into discrete single-electron pulses. Under these conditions it makes no sense to speak of amplitude.

Figure 5.19 illustrates how these criteria determine the practical boundaries of operation. The figure shows the log-log relation between the mean number of photoelectrons per pulse, $\bar{n}_{k,i}$, and the gain G



MRB281b

Fig.5.19 Photomultiplier working diagram in terms of photoelectrons per pulse, $n_{k,i}$, gain, and mean anode charge per pulse

In Fig.5.19 the first of these conditions corresponds to the bold horizontal line at $n_{k,i} = 10$; the second corresponds to the isocharge line $q_{a,i} = 4$ pC (third from bottom). Their intersection defines the minimum acceptable gain, about 2.5×10^6 . Assume that, with the type of tube and voltage divider used, this is obtained at $V_{ht} = 2200$ V. At that voltage the pulse-peak linearity limit, $\hat{I}_{a(max)}$, given above is reduced by the ratio

$$\frac{\hat{I}_{a(\max)}}{250} = \left(\frac{2200}{2500} \right)^n \quad (5.13)$$

where the exponent n is between 2 and 3, depending on the type of tube. For the present case, assume it is 2; then $\hat{I}_{a(\max)} \approx 200$ mA and, from Eq.5.11 and $t_w = 5$ ns,

$$q_{a,i(\max)} = 1000 \text{ pC}$$

which is indicated by the second isocharge line from the top. This is the third boundary of the practical operating region.

The fourth boundary is set by the required dynamic range which implies a maximum of 1000 photoelectrons per pulse. The $q_{a,i} = 1000$ pC isocharge line crosses the $n_{k,i} = 1000$ ordinate at a point corresponding to $G = 6.25 \times 10^6$. Thus, the practical boundaries of operation are:

$$\begin{array}{l} n_{k,i} = 10 \\ q_{a,i} = 4 \text{ pC} \end{array} \quad \Rightarrow \quad G_{\min} = 2.5 \times 10^6 \quad \begin{array}{l} n_{k,i} = 1000 \\ q_{a,i} = 1000 \text{ pC} \end{array} \quad \Rightarrow \quad G_{\max} = 6.25 \times 10^6$$

If $G \leq 2.5 \times 10^6$, pulses containing less than 10 photoelectrons do not exceed the sensitivity threshold of the electronics. And if $G > 6.25 \times 10^6$, the required dynamic range of 100 cannot be accommodated.

Defining these boundaries gives a first approximation of the required operating point. It is then necessary to check that a gain in the indicated range can be obtained at the assumed $V_{ht} = 2200$ V. If not, a new value of V_{ht} will have to be assumed and the boundaries redetermined.

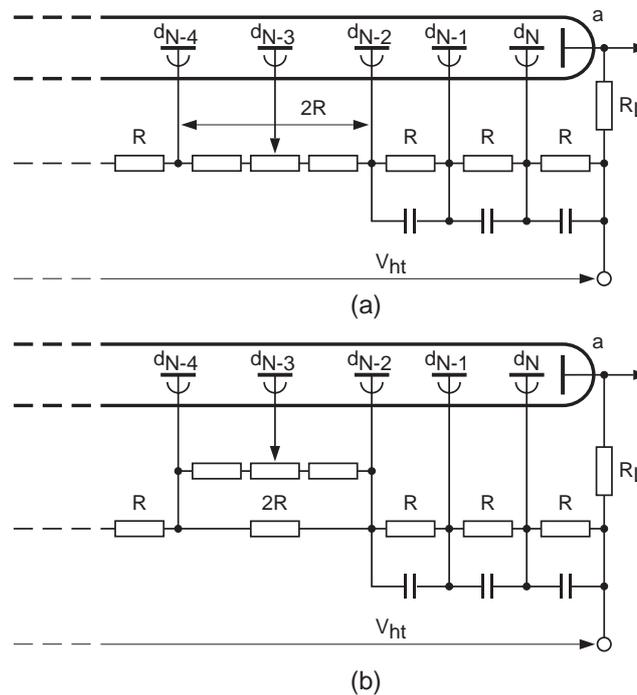
Uncertainty about the value of n in Eq.5.13 reflects on the accuracy with which the upper gain limit can be determined; however, this is seldom significant except when the gain operating range is narrow. Also, the current linearity limit given in the data sheets is a nominal value from which individual tubes may deviate; allowance should be made for this in the calculations. The pulse peak linearity mostly departs very slowly from the 2% linearity limit and higher pulses may still be linear to within say 3 to 4%.

In some cases of large dynamic range operation, other criteria have to be taken into account; especially, the high mean anode current due to a high count rate may affect stability (§4.6.1), necessitating reconsideration of the initial parameters by, for example, reducing the tube gain and compensating for this with additional gain in the preamplifier.

5.6 Gain adjustment

Gain characteristics differ from tube to tube. Sometimes, though, it is necessary to ensure that a number of tubes working together operate at equal gain. There are two ways to do this.

Supply voltage adjustment. Gain can be adjusted by adjusting the high voltage supplied to each tube. If the tubes do not have separately adjustable supplies but are fed from a common supply, their voltages can still be adjusted by ballast resistors connected in series with their respective voltage dividers. Even though the current from the supply is practically constant, the ballast resistors should be decoupled.



MRB282

Fig.5.20 Alternative circuits for adjusting gain by adjusting the voltage of one dynode

Dynode voltage adjustment. This is often used when it is not practical to adjust the high voltage supply to each tube. Gain can be altered by altering the voltage of any dynode (see Fig.4.10), but an intermediate one is always chosen to avoid interfering with the collection efficiency of the electron-optical input system or the output stage. Of the two adjustment circuits shown in Fig.5.20, the (b) version is preferable if the divider current is high; it makes it possible to use a high-value potentiometer (about $1\text{ M}\Omega$) with a low power rating ($\leq 0.75\text{ W}$). In both the (a) and (b) versions, resistors should be connected on both sides of the potentiometer to limit its working voltage; in practice, the range of control variation required is usually far less than the

maximum possible. As all terminals of the potentiometer are at a fairly high voltage, the potentiometer must be well insulated.

Dynode voltage adjustment is more effective with focusing than with venetian-blind dynodes. A disadvantage of it is that it can impair stability and increase susceptibility to magnetic fields; on the other hand, time characteristics are relatively unaffected.

5.7 Supply for multiple tubes

When many photomultipliers are used together (as in hodoscopy, scintigraphy, tomography) the high voltage can be supplied either separately to each or by a single supply common to all.

Separate supplies are preferable. They prevent any reaction between tubes and, if they are adjustable, facilitate individual gain adjustment. Compact, adjustable and non-adjustable, individual supply modules are marketed, as well as supplies with several, separately adjustable output channels. The latter, though, are usually bulky and expensive.

Common supply to a number of tubes is often used when the current required from each is low. Each then has its own voltage divider to minimize reaction between tubes and its own provision for gain adjustment, which may be a potentiometer either in series with the divider (Fig.5.21) or controlling the voltage of one of the dynodes (Fig.5.22).

If the application requires the voltages at the electron-optical input system or the output stages to be zener stabilized, it may be advantageous to use a single zener diode (or series-connected group of diodes) for all the parallel-connected tubes. This ensures sufficient current through the diodes even when the current through the individual voltage-divider chains is low.

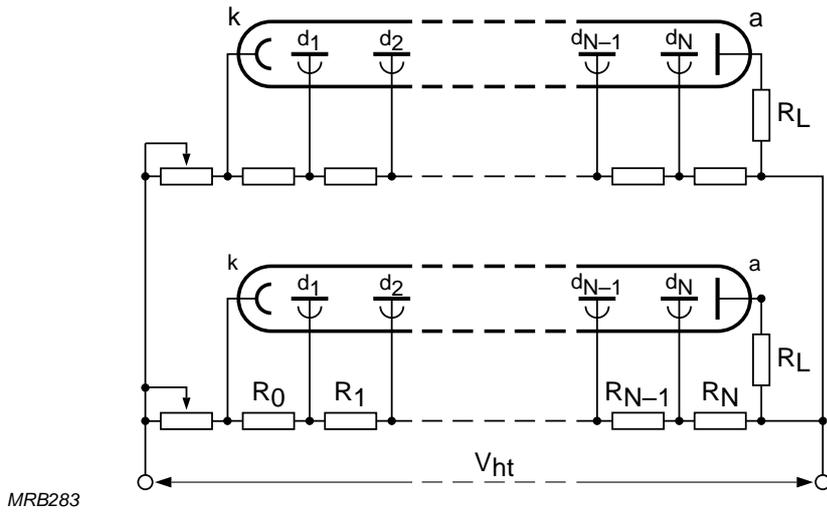


Fig.5.21 High-voltage supply to photomultipliers in parallel

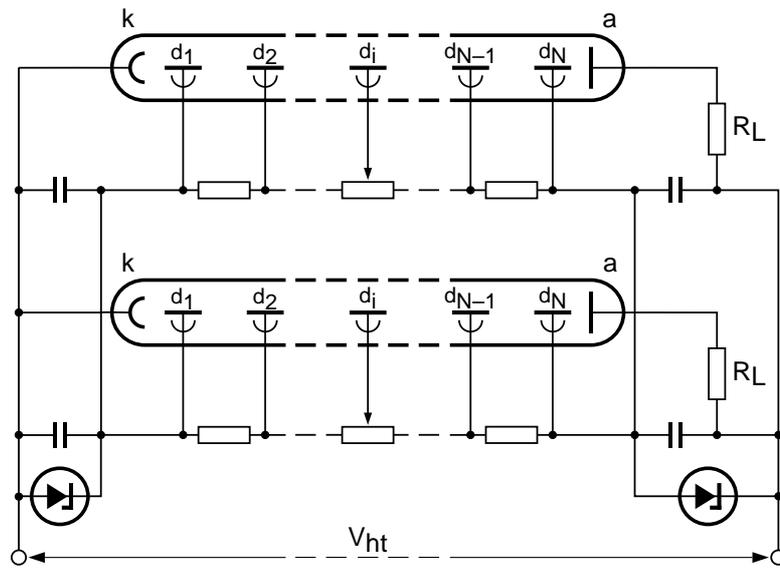


Fig.5.22 High-voltage supply to photomultipliers in parallel, with common zener-diode stabilization

5.8 Dark current reduction

5.8.1 Cooling

As the dark current is partly thermionic it can be reduced by cooling the tube, but only down to a certain temperature depending on the photocathode composition. Below that temperature leakage current, field emission, and other temperature-independent components of the dark current predominate. Depending on the requirements of the application, cooling can be in a cold chamber, by refrigerating coils, or by Peltier coolers. Whichever is used, precautions against condensation should be taken by enclosing the tube either in a partial vacuum or a dehumidified atmosphere. Moisture condensed on the tube impairs insulation, increases leakage current, and can affect the transmission of the input window.

It is preferable to connect an artificially cooled tube in positive polarity so that no special precautions need be taken as regards insulation. If the tube does have to be connected in negative polarity, a thermally conductive electrical insulator must be interposed between it and the cooling provision.

5.8.2 Reducing effective photocathode area

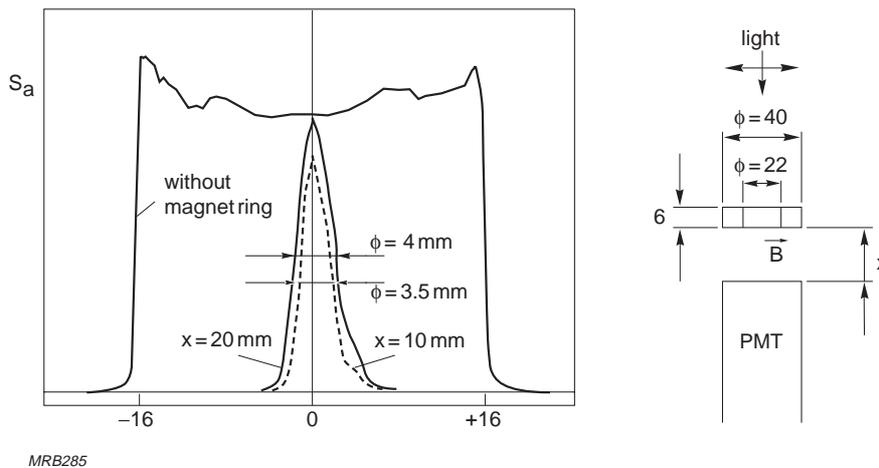
If less than the whole area of the photocathode is to be used (as when working with collimated light, for example), dark current can be reduced by rendering the unused area inactive. This eliminates the thermionic component of the dark current from the inactive area and is therefore the more rewarding the larger that component is. It is particularly advantageous with S20, S20R, and S1 cathodes which, being more sensitive toward the long wavelengths, have relatively high thermionic dark current components.

An outer zone of the cathode can be made inactive by defocusing the electrons from it so that they are not collected by the first dynode. This can best be done magnetically, either

- by means of an axially-magnetized toroidal permanent magnet concentric with the cathode and slightly in front of it. Figure 5.23 illustrates the effective-area reduction, plotted along one cathode diameter, that can be obtained in this way.
- or by means of a solenoid surrounding the cathode. Figure 5.24 illustrates the area-reduction effect, and Fig.5.25 compares this with the effect on dark current. The dark current decreases less rapidly than the effective area because not all of it originates from the cathode. A potential disadvantage of the electromagnetic method is that heat dissipated in the solenoid may raise the cathode temperature.

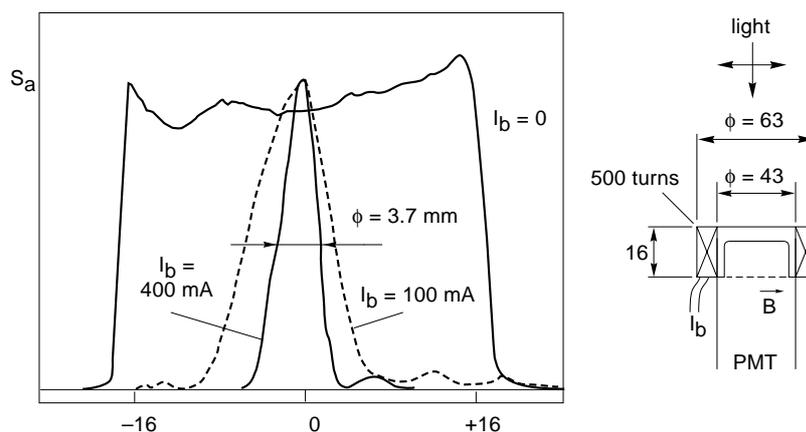
All these methods are more difficult to use with small-diameter tubes. Such tubes seldom have an accessible focusing electrode, and the input to the electron multiplier is often off-centre; a centred magnet or solenoid therefore upsets the electron-optical input system.

Dark current can also be cancelled or compensated electronically (§5.11.2), and dark pulses can be excluded by coincidence techniques. However, neither these nor any of the methods described above has any effect on signal noise, which is fundamentally irreducible (Chapter 3).



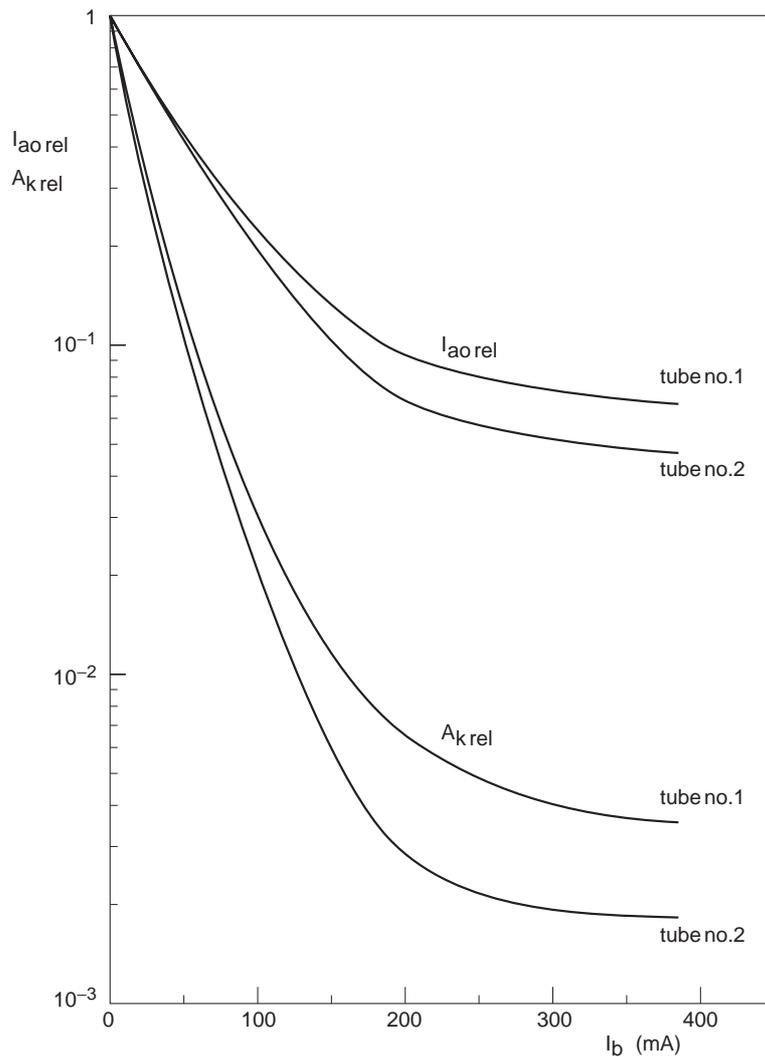
MRB285

Fig.5.23 Effective-cathode-area reduction due to a permanent-magnet ring in front of the cathode (32 mm diameter S1 cathode)



MRB286

Fig.5.24 Effective-cathode-area reduction due to a solenoid surrounding the cathode (32 mm diameter S1 cathode)



MRB287

Fig.5.25 Relative dark-current (I_{a0}) and cathode-area (A_k) reduction as functions of solenoid current I_b in two tubes with 32 mm diameter S1 cathodes, using the method of Fig.5.24. Note that the dark-current reduction is much less than the effective-cathode-area reduction

5.9 Magnetic shielding

Since fields as weak as the earth's can affect sensitivity (§4.8.2), a mu-metal shield is always desirable. At flux densities of more than a few milliteslas, however, such a shield saturates and becomes ineffective. It must then be surrounded by a supplementary shield, usually of soft iron. Shielding effectiveness, based on material and dimensions, can be determined from Fig.5.26. For example, a mild-steel shield with an inner diameter of 70 mm and thickness 5 mm ($r_{in}/r_{ext} = 0.875$) in a flux

density of 100 mT gives an attenuation of about 30. An inner mu-metal shield with an inner diameter of 57 mm and thickness of 0.8 mm ($r_{in}/r_{ext} = 0.973$) gives an additional attenuation of about 300, leaving a residual flux density that is well below the 0.05 mT to 0.1 mT sensitivity threshold of most tubes.

Magnetic shields should extend about one diameter beyond the cathode plane, especially if the magnetic flux density is high. If the tube is operated in negative polarity, the shielding must be completely insulated from the glass (§5.1.1) or connected to the cathode potential via a protective resistor.

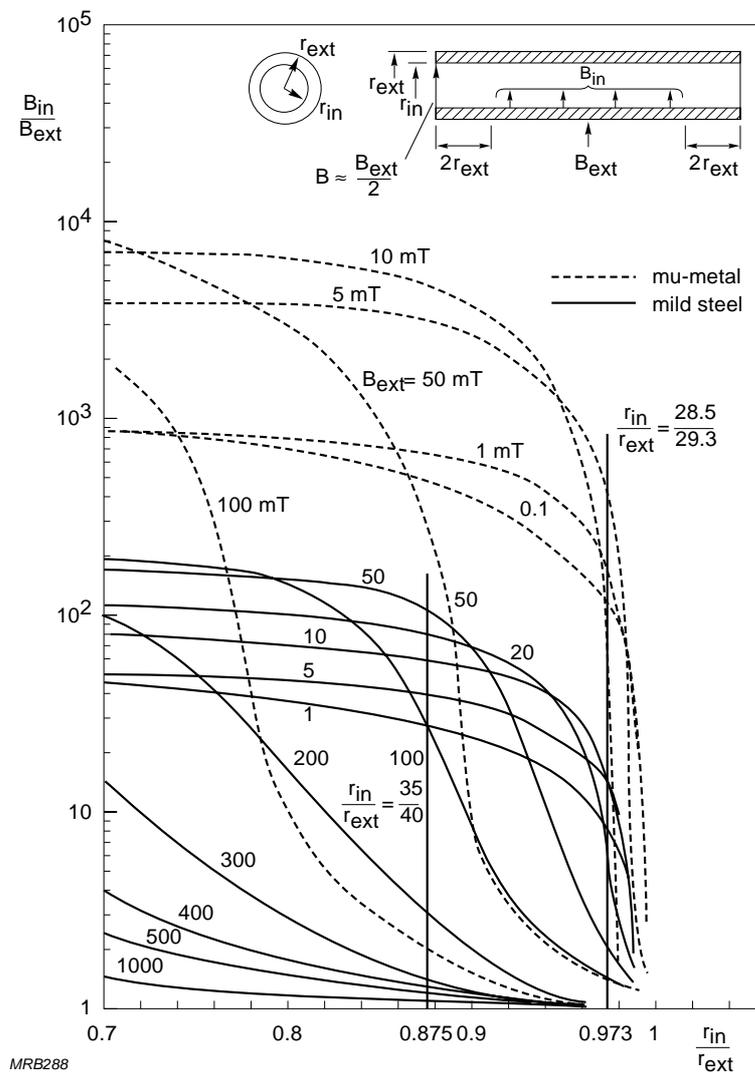


Fig.5.26 Comparative effectiveness of mu-metal and mild-steel magnetic shields

5.10 Gating

In applications in which there is intermittent exposure to high illumination, it may be desirable to guard the tube or its circuits against overloading. This can be done by gating the tube via the high voltage supply, the cathode, one or more of the dynodes, or the cathodes and dynodes together.

High-voltage supply gating. The tube can be gated at any desired frequency by chopping the high-voltage supply but, as this requires a low-impedance source of high-voltage pulses, it is usually not practical. Moreover, the anode sensitivity only reaches its maximum some 10 μs to 50 μs after the high voltage does. Finally, high-voltage supply chopping aggravates dark-current and gain instabilities.

Positive cathode pulse gating. Applying a positive pulse to the cathode has similar disadvantages to high-voltage supply chopping. Moreover, the photo-sensitivity of the first dynode limits the effectiveness of the gating.

Negative focusing-electrode gating. Provided the focusing electrode is independently accessible, applying a 50 V to 100 V negative pulse to it is the easiest method of gating. However, the effectiveness of this too tends to be limited by first-dynode photoemission.

Negative-dynode gating. Applying a negative pulse via a capacitor or coaxial-cable transformer to one or more of the dynodes (usually the 3rd to the 5th) can reduce gain several hundredfold, depending on the amplitude of the pulse.

Combined cathode and dynode gating. During negative-dynode gating, photoemission continues. If the illumination is intense the recovery time is long and there is a risk of cathode fatigue. There is also a risk of impaired gain stability due to scattered charges built up on the glass and insulators by electron bombardment. A way out of these difficulties is to gate a dynode and the cathode together, simultaneously lowering gain and suppressing photoemission. This can be done by applying a pulse between the cathode and the third dynode, using a separate voltage divider earthed at the third dynode to supply the first two stages. Although the amplitude of the pulse has to be several hundred volts, gain can be reduced by a factor of 10^{-4} in this way.

The tube can be permanently blocked and then unblocked whenever an event to be counted or measured is expected. This is difficult, however, the unblocking pulse amplitude, on which the gain of the tube depends, has to be very precisely controlled.

To protect the tube against excess illumination an external coincidence circuit is preferable, such a circuit might make use of a gate in the anode circuit.

Another gating technique described in the literature makes use of a cross-bar grid fitted to the input window to localize the area of photoemission.

5.11 Drift compensation

5.11.1 Anode sensitivity drift

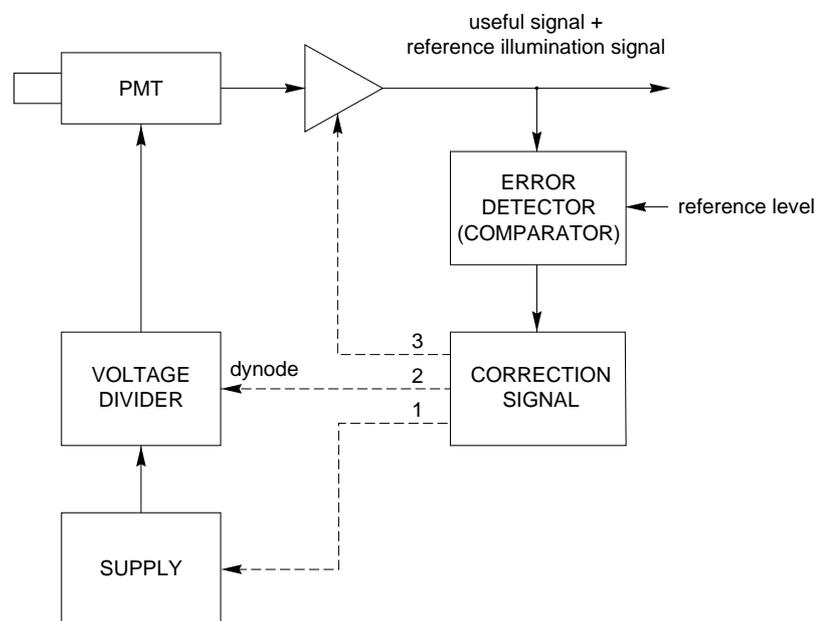
Anode sensitivity drift can be compensated by

- varying the gain of the tube or the preamplifier of the measuring circuit
- varying the illumination.

Varying the gain. Figure 5.27 illustrates the principle. A signal due to a reference illumination of the tube is compared with a fixed reference level to derive an error signal. The error signal is fed to a control circuit which regulates

- the high-voltage supply
- the potential of one dynode
- or the gain of an amplifier in the anode circuit

to equalize the illumination signal and the reference signal so as to make the error signal zero.



MRB289

Fig.5.27 Correction of anode-sensitivity drift by varying the gain. The signal due to a reference illumination of the cathode is compared with a fixed reference level to derive an error signal. The gain can be altered by any one of the three routes shown

Depending on the type of application, the reference illumination may be a calibrated radioactive (α or γ) source/scintillator combination or a LED. A calibrated radioactive source integral with the scintillator normally used with the tube makes it possible to compensate drift in both the anode sensitivity of the tube and the luminous efficiency of the scintillator at the same time.

Systems incorporating microprocessors for compensating the drift of several photomultipliers at once are described in the literature.

Varying the illumination. Gain shift due to wide variation of illumination can be reduced with the aid of a constant source of supplementary illumination, or compensated by means of a variable one.

Constant supplementary illumination. When the mean anode current is only 20-30 nA but the relative variation is large, gain shift due to the variation can be reduced by using a constant source of supplementary illumination (e.g. a LED) to increase the mean anode current. A capacitor preceding the signal processing circuits eliminates the resulting DC component.

This method has the disadvantage that operating the tube at a higher mean anode current may introduce another causes of gain instability – long-term drift.

Variable supplementary illumination. In pulse applications the gain shift due to variations of mean anode current can be compensated by a supplementary light source (LED) which can be varied in opposition to the mean anode current variation. Figure 5.28 illustrates the principle. Integrated anode pulses are compared with a fixed reference to generate an error signal which controls the current to the LED; the greater the difference between the integrated pulses and the reference, the more light the LED emits. Once again, a coupling capacitor preceding the signal processing circuits eliminates the DC component due to the LED.

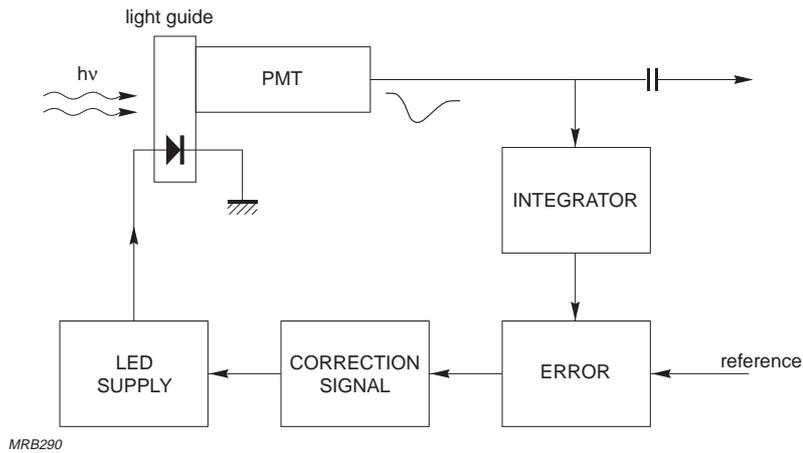
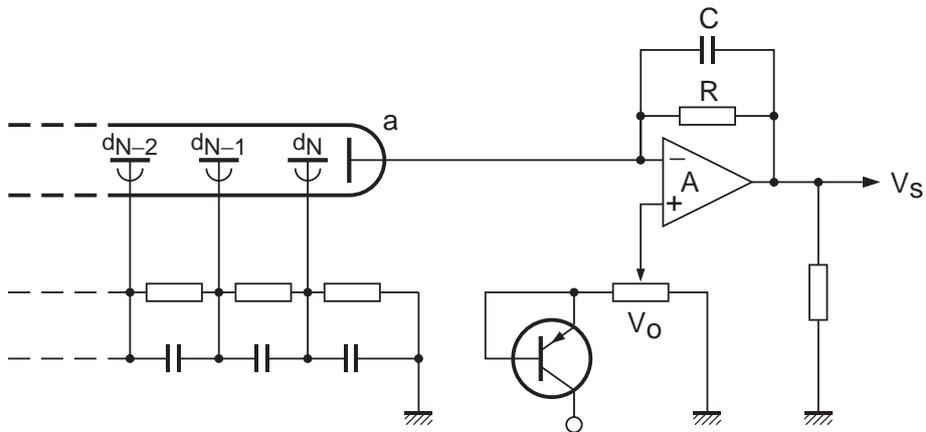


Fig.5.28 Correction of gain drift by varying a supplementary light source

5.11.2 Temperature-dependent dark-current drift

Temperature-dependent dark-current drift can be compensated by means of a current source (e.g. a selected diode or transistor) obeying the same temperature law as the dark current. Figure 5.29 shows a typical circuit. When there is no signal from the photomultiplier the output of the differential amplifier is $V_s = V_0 - RI_{a0}$, where I_{a0} is the dark current. V_0 is adjusted to make $V_s = 0$ at the prevailing temperature. Provided the transistor from which V_0 is derived follows the same temperature law as the dark current, the voltage V_s due to the dark current remains zero at other temperatures. Diodes or transistors that follow exactly the same law as the dark current of a given photomultiplier are not easy to find, but a selection can usually be made that gives no more than a small discrepancy over a practical temperature range.



MRB291

Fig.5.29 Circuit for compensating temperature-dependent dark-current drift

5.12 Pin connections and safety precautions

To save space it is sometimes necessary to make connections direct to the pins of a tube with glass base, instead of using a socket. Do so with contact clips if possible; if it is not possible, spot-weld the connections. *Do not solder them*; soldering heat can crack the glass.

For such applications it is also possible to order tubes with flying leads which can be safely soldered.

Internally connected pins (marked 'i.c.' in the data sheet) must never be connected. They are used only during manufacture of the tube. If used, the applied potential may distort the field distribution within the tube.

Non-connected pins (marked 'n.c.' in the data sheet) should not be used either, even as intermediate wiring points. They are provided only to facilitate interchangeability of similar tubes (i.e. an electrode that is externally connected in one type may be internally connected in a similar type using the same socket and operating voltages). Connections made to an n.c. pin could cause leakage currents or local insulation breakdown.

Ensure that there are adequate safeguards for those working near high voltage. Voltages as high as 3 kV are used with some tubes, and some power supplies can deliver more than 10 mA.

Handle photomultipliers with due regard for the fact that they contain a high vacuum, glass is fragile, and the flat input window is highly stressed. Take particular care with tube of 100 mm diameter or more, because of risk of implosion.

5.13 Signal processing

In a few applications involving steady-state operation the anode current of a photomultiplier can be read from a galvanometer. In most applications, however, and especially those in which the tube forms part of a control system, amplification is necessary, and the amplifier must have stability and noise characteristics comparable with those of the tube. Integrated-circuit operational amplifiers that meet this requirement are available nowadays fairly inexpensively. Such an amplifier is usually a direct-coupled differential amplifier with very high input impedance ($>10\text{ M}\Omega$) and high open-loop gain ($>10^5$) and near-zero output impedance. The transfer function of the amplifier is determined solely by the input and feedback networks associated with it.

5.13.1 Operational amplifiers

Figure 5.30 shows an operational amplifier connected as a voltage amplifier. If A is the open-loop gain of the amplifier, $V_s = A(V_2 - V_1) = -AV_1$. Because of the very high input impedance of the amplifier, the current I_3 is negligible; therefore,

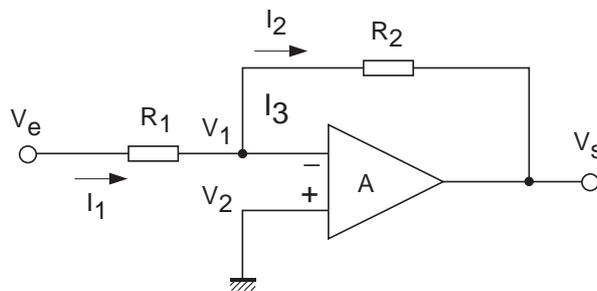
$$I_1 = \frac{V_e - V_1}{R_1} = I_2 = \frac{V_1 - V_s}{R_2}$$

and the gain with feedback is

$$G = \frac{V_s}{V_e} = - \frac{\frac{R_2}{R_1}}{1 + \frac{1}{A} + \frac{R_2}{AR_1}}$$

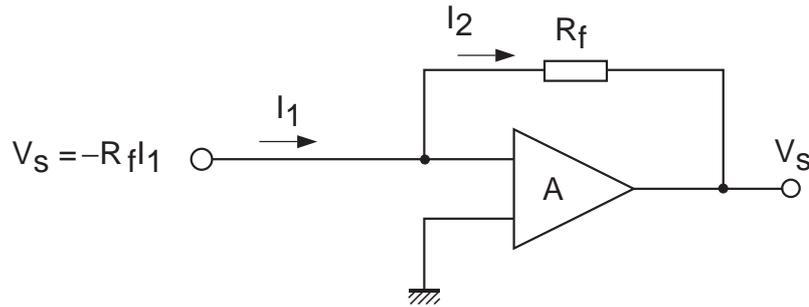
which, since A is very large, reduces to $G = -R_2/R_1$.

To ensure stability the resistances R_1 and R_2 must be very stable. The same holds for the feedback resistance R_f in the current/voltage converter of Fig.5.31, a circuit which finds wide application in view of the fact that a photomultiplier is an almost ideal current generator.



MRB292

Fig.5.30 Voltage amplifier (inverting)



MRB293

Fig.5.31 Current-voltage converter

Operational amplifiers can also be used for such mathematical operations as addition, subtraction, and integration; see Fig.5.32 to 5.34. A practical example of subtraction, for instance, is the arrangement for compensating temperature-dependent dark-current drift shown in Fig.5.29. Integration is required in some scintillation counting applications, where it is necessary to measure the quantity of light contained in individual scintillations, whatever their pulse shape or duration. The anode charge q_a due to a single scintillation is proportional to the quantity of light it contains. Integrating the anode current over the duration of the scintillation gives a voltage V_s proportional to q_a and, hence, to the quantity of light. An electronic switch (represented by S in Fig.5.34) discharges the integrating capacitor between scintillations.

In applications in which the photomultiplier output can vary by several orders of magnitude it is often desirable to convert it to logarithmic form. This can be done by incorporating a diode or transistor in the feedback loop (Fig.5.35) so as to exploit the logarithmic characteristic of a forward-biased p-n junction. A transistor has an advantage over a diode in that earthing its base establishes an exact zero reference for the output voltage.

When the photomultiplier output is a very low-level, varying direct current, the zero-drift of a DC amplifier can be a significant source of error. An alternative (Fig.5.36) is then to chop the signal at a frequency must higher than its own range of variation, amplify the resulting modulated squarewave signal in a stable AC amplifier, demodulate the amplifier output, and restore the original but now amplified DC signal by means of a low-pass filter.

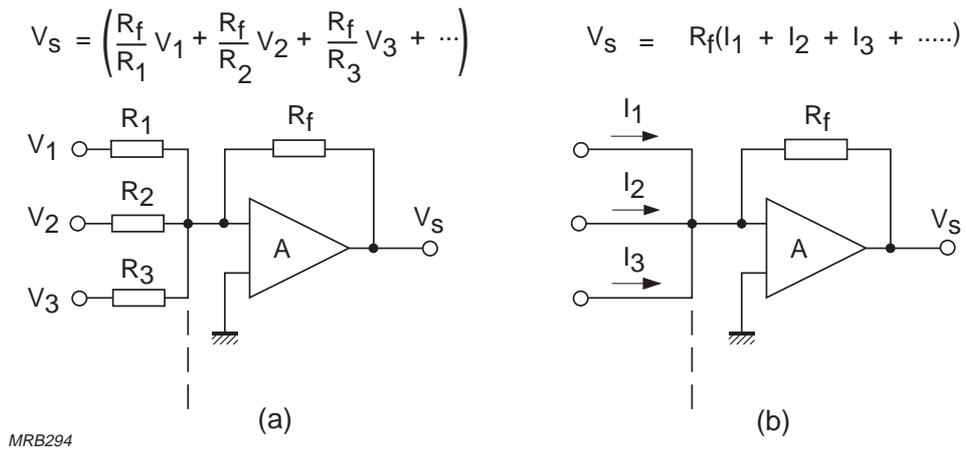


Fig.5.32 (a) Voltage adder, (b) current adder

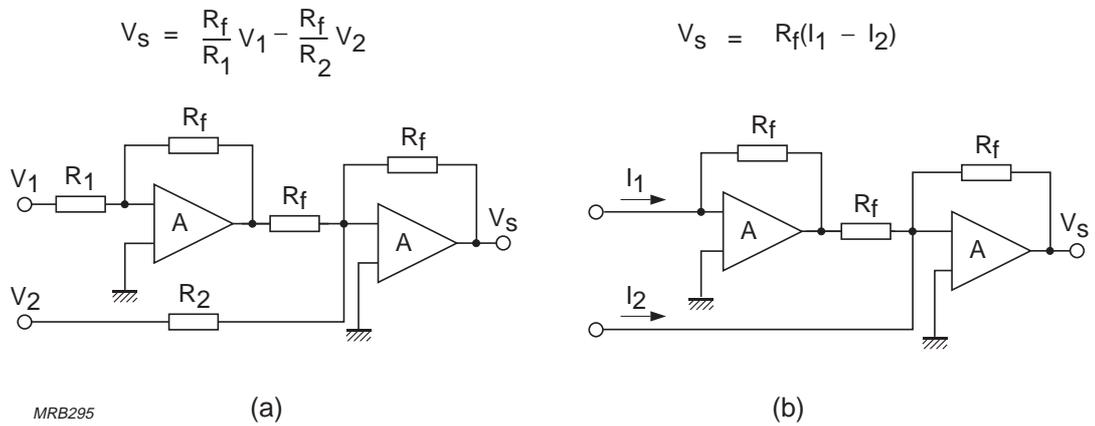


Fig.5.33 (a) Voltage subtractor, (b) current subtractor

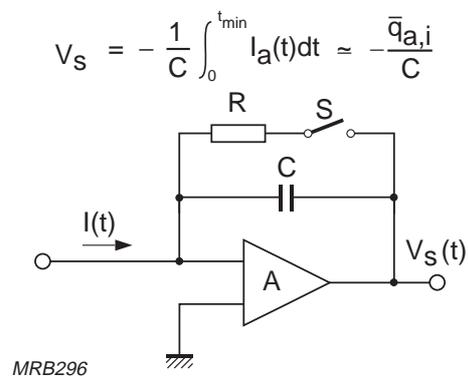


Fig.5.34 Integrator. For integrating the charge due to light pulses, t_{\min} must exceed the pulse duration. Switch S discharges the capacitor between pulses

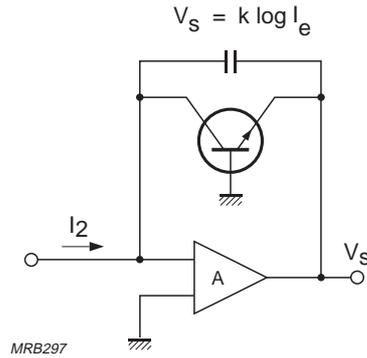
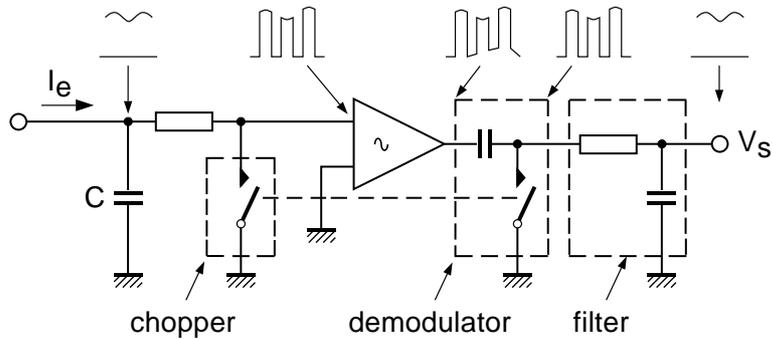


Fig.5.35 Logarithmic amplifier

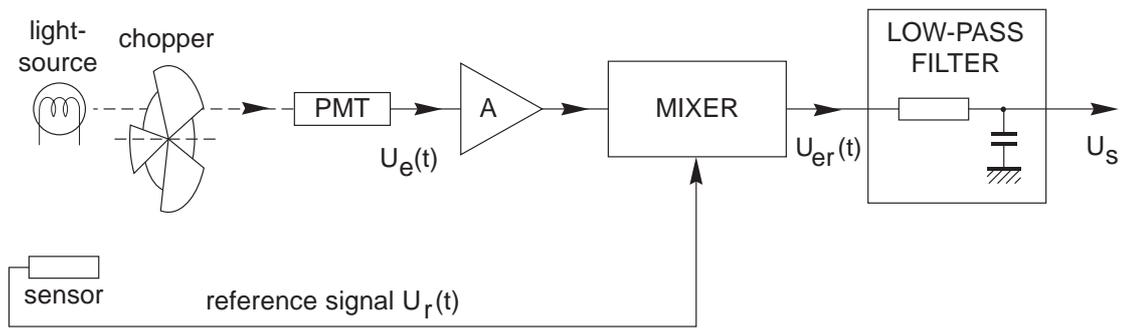


MRB298

Fig.5.36 Chopper-stabilized amplifier

5.13.2 Separating the signal from the noise

Synchronous detection. A principle similar to that of the chopper stabilized amplifier is often used in photometric instruments when the signal information is known to occupy only a narrow bandwidth; it is based on reducing the bandwidth of the measurement to increase the signal-to-noise ratio (see §3.3.3). The light flux to be measured is first chopped at a frequency at least twice its own highest frequency of variation (Fig.5.37) so that the resulting photomultiplier output is itself a squarewave carrier modulated by the signal information. The modulated carrier is amplified and then demodulated with the aid of a reference signal synchronous with the chopper that switches the gain of a mixer between +1 and -1. A low-pass filter at the mixer output eliminates carrier frequency components and leaves an amplified continuous signal proportional to the input.



MRB299

Fig.5.37 Synchronous detection

Autocorrelation. Another method for separating the signal from the noise is based on the autocorrelation function

$$\gamma(\tau) = \frac{1}{T} \int_0^T U_1(t)U_2(t - \tau) dt$$

where U_1 and U_2 represent signals to be discriminated. When $U_1 = U_2$ the autocorrelation function $\gamma(\tau)$ attains a maximum at $\tau = 0$. If the signal is periodic, $\gamma(\tau)$ is also periodic and has the same period. If the signal is aperiodic, $\gamma(\tau) = 0$ except when $\tau = 0$, in which case $\gamma(\tau)$ corresponds to an impulse function. Therefore, in a signal-processing system incorporating a delay channel for discriminating signals from noise by autocorrelation, aperiodic noise is quickly suppressed and periodic signals become clearly distinguishable as soon as the delay τ is made significantly greater than zero.

Figure 5.38 shows an autocorrelation system for extracting a low-level periodic signal, $U(t) = U_0 (1 + \cos \omega t)$ from a photomultiplier output which is corrupted by a noise $u(t)$. After being bandpass filtered around the frequency ω , the photomultiplier output is applied to two equal-gain amplifier channels, one incorporating a delay τ and the other undelayed. The outputs of the two channels are multiplied together and integrated over a time T , giving

$$\gamma(\tau) = \frac{1}{T} \int_0^T U(t)U(t - \tau) dt = \frac{A^2}{T} \int_0^T \{U_o(1 + \cos \omega t) + u(t)\} [U_o \{1 + \cos \omega(t - \tau)\} + u(t - \tau)] dt$$

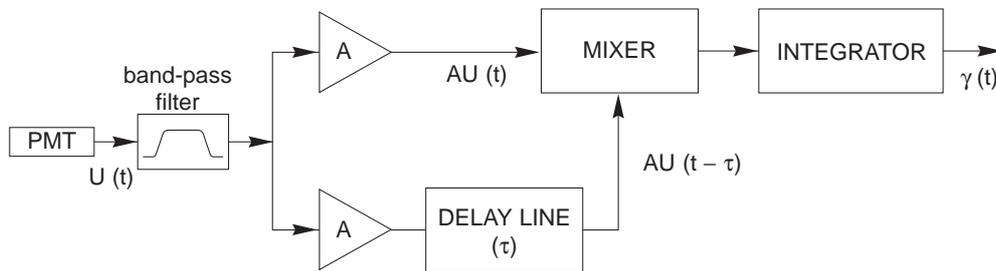
The discrimination principle rests on finding the least value of τ which is an integral multiple of the signal period and for which the fluctuating component, corresponding to terms containing the noise $u(t)$, becomes negligible. The equation then simplifies to

$$\gamma(\tau) \approx \frac{A^2 U_o^2}{T} \int_0^T (1 + \cos \omega t) \{1 + \cos \omega(t - \tau)\} dt \approx \frac{A^2 U_o^2}{2} (2 + \cos \omega\tau)$$

which, since τ is chosen as an integral multiple of the signal period, reduces to

$$\gamma(\tau) \approx \frac{3}{2} A^2 U_o^2$$

Thus, the output signal is proportional to the square of the input signal.



MRB300

Fig.5.38 Autocorrelation detector

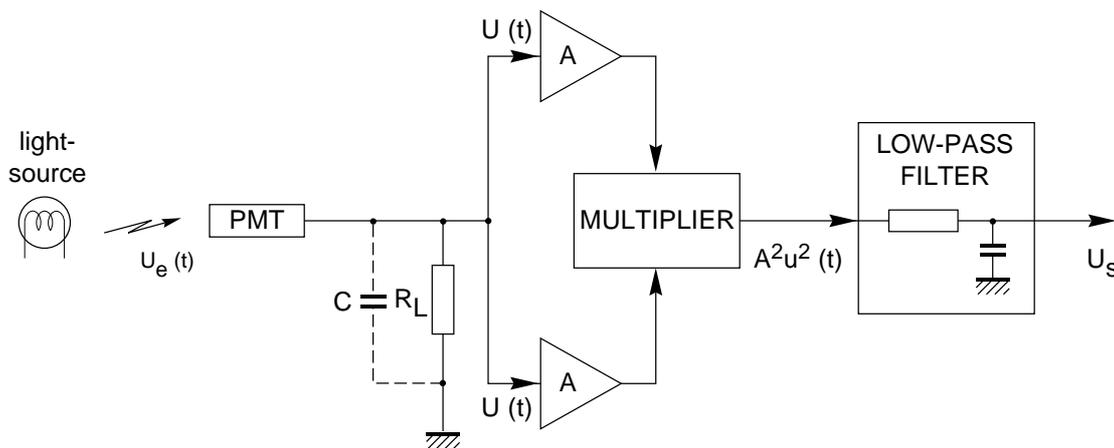
5.13.3 Detection at very low light levels

Quadratic detection. When a photomultiplier is exposed to a constant or slowly varying flux the resulting anode current consists of a DC component, which varies as the flux, and a fluctuating ‘shot-noise’ component which varies as the square root of the DC component. If the flux is very low the energy transferred by the fluctuating component can exceed that transferred by the DC component, and in that case it may be more expedient to measure the fluctuating than the DC component.

An advantageous way to do this is by quadratic detection, which gives an output proportional to the square of the fluctuation – and, hence, to the incident flux.

In the quadratic detector shown in Fig.5.39 the DC component of the voltage across the load resistor R_L is too low to be useful and the fluctuating component $u(t)$ is amplified by a matched pair of wideband amplifiers. The amplifier outputs are multiplied by each other to produce a voltage $A^2u^2(t)$ which is then low-pass filtered to obtain its mean value, the output voltage U_s , which is thus proportional to the incident flux.

As the noise contributions of the two amplifiers are not correlated with each other, they do not significantly contribute to the output voltage U_s . However, the photomultiplier dark-current noise and the load-resistor thermal noise, which are applied equally to both channels, do contribute to the output voltage and are thus a factor limiting the sensitivity of quadratic detection.

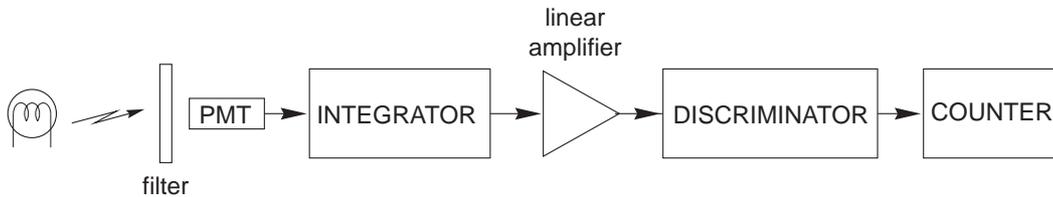


MRB301

Fig.5.39 Quadratic detector

Photon counting. In photon counting the photomultiplier is operated under single-electron conditions. Flux levels as low as a few tens of photons per second can be measured – and the method has the advantage of making it possible to eliminate several otherwise disturbing factors from the measurement. Among these are the DC component of the dark current, low-amplitude pulses originating in the electron multiplier, and high-amplitude pulses of other than photoelectric origin. Photon counting can also be used for determining the shape of fast, low-intensity light pulses, as in certain applications of scintillation counting.

Low flux levels. In the photon-counting system of Fig.5.40, the level of the incident flux is such that the cathode emits only single electrons. The individual anode charges due to these are integrated to produce proportional voltage pulses which are passed by a discriminator to a pulse counter whose output over a set time is a measure of the incident flux.



MRB302

Fig.5.40 Photon-counting detector

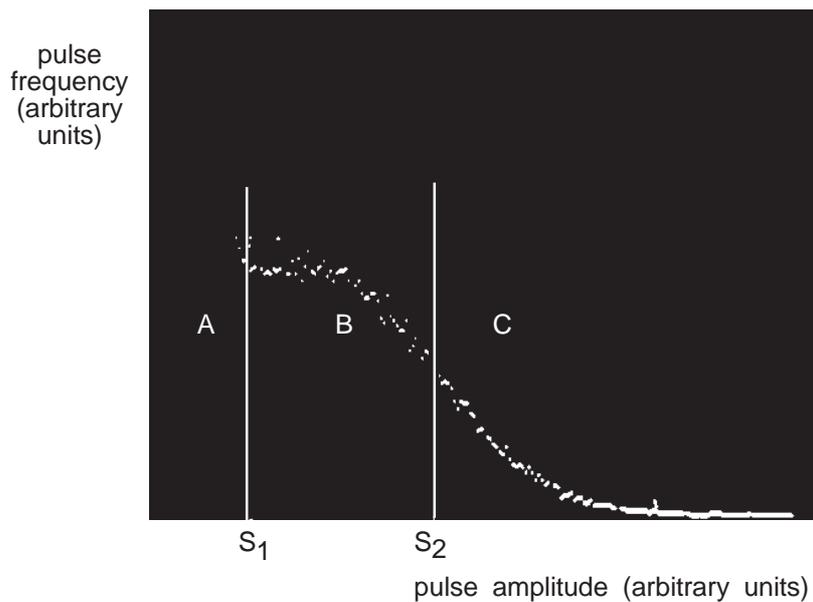


Fig.5.41 Pulse-amplitude spectrum of photomultiplier operating under single-electron conditions, showing lower and upper thresholds for photon counting (defining counting window B)

Because of statistical fluctuations in the electron multiplication, the amplitudes of the single-electron pulses are distributed according to the SES (§2.1.6). The dark noise pulses are distributed according to a spectrum whose general shape is often quite

different (Fig.5.41) with a large quantity of very small pulses but also some very large pulses (cosmic rays, afterpulses). To optimise the S/N (signal/dark-noise) ratio, two thresholds S_1 and S_2 are adjusted to give the best pulse-amplitude ‘window’.

However, there are also some dark pulses within the discriminator window, and the count must be corrected for them. If \bar{n}_p is the mean number of incident photons and \bar{n}_d the mean number of non-excluded dark pulses per unit time, the total count during a period τ will be $\tau(\bar{n}_p\rho\eta + \bar{n}_d)$, where ρ is the quantum efficiency of the cathode and η the collection efficiency of the electron-optical input system. Subtracting the number of dark pulses counted during a like period without illumination gives the correct count

$$n_e = \tau(\bar{n}_p\rho\eta + \bar{n}_d) - \bar{n}_d\tau$$

which (from §3.2) has a variance

$$\sigma_n^2 = \bar{n}_p\rho\eta\tau + 2\bar{n}_d\tau$$

At very low flux levels n_d is usually much larger than n_p and the signal-to-noise ratio becomes

$$\frac{S}{N} = \frac{\bar{n}_p\rho\eta\tau}{\sqrt{2\bar{n}_d\tau}}$$

which increases as the square root of τ . Thus, improving the signal-to-noise ratio by increasing the counting time is analogous to the improvement that can be obtained in charge-integration methods by narrowing the bandwidth.

The signal-to-noise ratio can also be improved by cooling the cathode (§5.8.1) to reduce the number of single-electron pulses of thermionic origin, and by magnetically reducing the cathode area from which thermionic electrons can be collected (§5.8.2).

Determining the shape of fast, low-intensity pulses. Repetitive pulses detected by a photomultiplier can be displayed in real time on an oscilloscope, provided their duration is more than a few tens of nanoseconds and their intensity corresponds to more than a few thousand photons. However, if their duration is short compared with the pulse response of the photomultiplier (a few nanoseconds), this is not possible, even if the number of photons per pulse is small enough for them all to be resolved. Then, a sampling method must be used and the pulse shape reconstructed statistically. For this it is essential for the transit-time fluctuations of the tube to be small.

Figure 5.42 shows a set-up for reconstructing (delineating) fast, low-intensity pulses by sampling. (Note the similarity to the transit-time spread measuring set-up of Fig.4.8). The pulses received by photomultiplier PMT1 are so attenuated that the tube operates under single-electron conditions. Those received by PMT2 are not attenuated but are delayed sufficiently to ensure a measurable time difference between the outputs derived from the two tubes. Discriminators 1 and 2 produce standardized timing pulses coinciding with chosen reference points on their respective input pulses. From the time differences between the discriminator pulses the time-to-amplitude converter and the multichannel analyser construct a histogram that reflects the photon distribution in the original light pulses.

The auxiliary channel at the bottom of the diagram minimizes statistical errors due to the effect of pulse-shape fluctuations and walk errors in the discriminators. The auxiliary channel gates the multichannel analyser only when the pulses from both photomultipliers satisfy predetermined amplitude criteria.

Let $L(t)$ be the probability density of photon-emission instants during the light pulse, and $R(t)$ the probability density of the single-electron current-pulse transit time of PMT1. The probability density of the corresponding anode-current pulse is then given by the convolution

$$L^*(t) = L(t) * R(t)$$

If $L(t)$ and $R(t)$ are both gaussian with variances σ_L^2 and σ_{tt}^2 , then $L^*(t)$ is also gaussian and has a variance

$$\sigma_{L^*}^2 = \sigma_L^2 + \sigma_{tt}^2$$

For a fast-response tube $\sigma_{tt} \leq 0.4$ ns. If this is small in comparison with σ_L , then $L^*(t)$ is practically identical to $L(t)$.

The time reference in the set-up of Fig.5.42 is derived from a second photomultiplier PMT2 which also has transit-time variations. To take this into account, let $\bar{n}_{k,i}$ be the mean number of photoelectrons per pulse at the PMT2 cathode. Then, if $R(t)$ for PMT2 is gaussian, the probability density $(L')^*(t)$ of the corresponding anode current pulse is also gaussian and has a variance

$$[(\sigma_{L'})^*]^2 = \frac{\sigma_L^2 + \sigma_{tt}^2}{\bar{n}_{k,i}}$$

which is $\bar{n}_{k,i}$ times smaller than the corresponding variance for PMT1.

The probability density distribution $L^{**}(t)$ of the histogram constructed by the multichannel analyser is given by the convolution.

$$L^{**} = (L')^*(t) * L(t) * R(t)$$

which, provided its standard deviation is large compared with $(\sigma_L)^*$ and σ_{tt} , is practically identical to the illumination function $L(t)$.

Thus, this method is a good way for reconstructing brief, low-intensity light pulses provided the transit time fluctuations in the two photomultipliers are low compared with the light pulse duration.

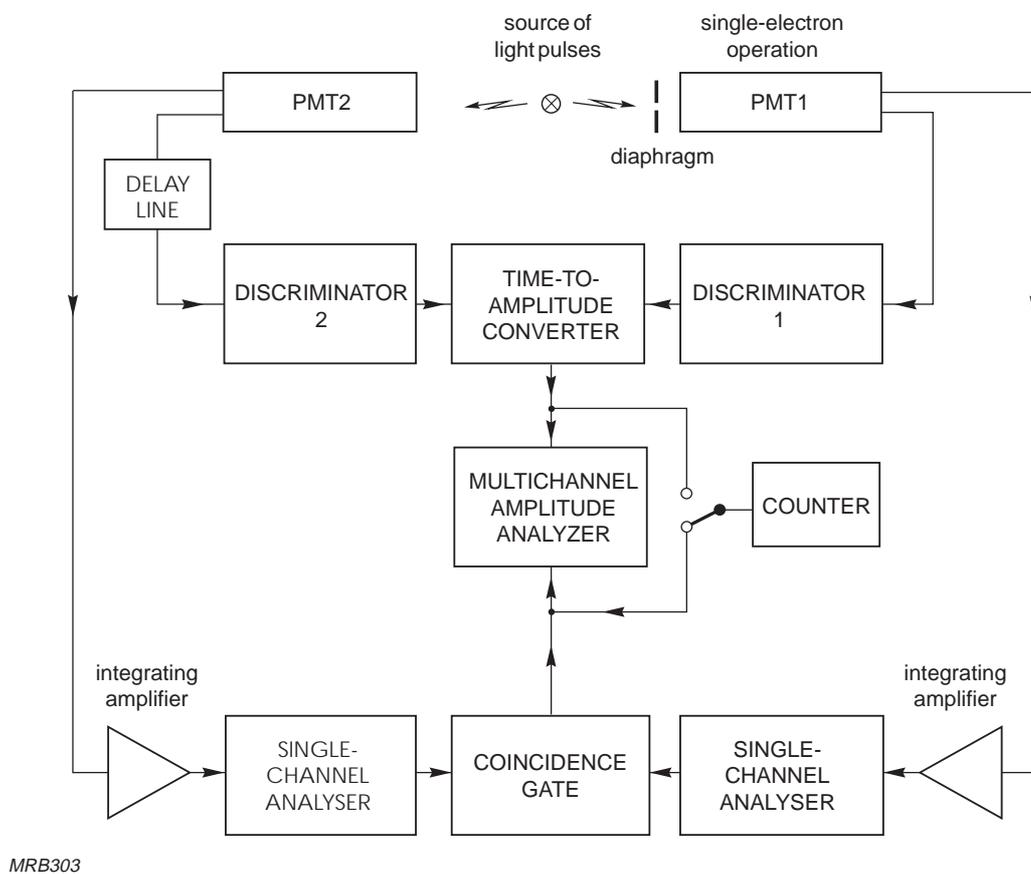


Fig.5.42 Set-up for reconstructing fast, low-intensity pulses by sampling

CHAPTER 6

SCINTILLATION COUNTING

6.1 Scintillator-photomultiplier combination

- 6.1.1 Scintillator shape
- 6.1.2 Scintillator finish
- 6.1.3 Scintillator-photomultiplier coupling

6.2 Electrical signal characteristics

- 6.2.1 Current pulse
- 6.2.2 Voltage pulse

6.3 Operating mode

- 6.3.1 Continuous mode
- 6.3.2 Pulse mode

Appendix Scintillator fundamentals

- A6.1 Inorganic scintillators
- A6.2 Organic scintillators
- A6.3 Scintillator characteristics
- A6.4 Scintillator properties
 - A6.4.1 Inorganic-scintillator properties
 - A6.4.2 Organic-scintillator properties

SCINTILLATION COUNTING

Nuclear radiation absorbed in certain materials excites flashes of light – scintillations. The first laboratory application of this phenomenon was the spintharoscope, in which scintillations of a zinc sulphide screen are observed through a microscope. Early twentieth century investigators who patiently counted the scintillations excited by different radioactive substances, thereby helping to lay the foundations for today's understanding of radioactivity, were the first scintillation counters.

Depending critically as it did on a trained observer's eye and powers of concentration, the spintharoscope was an instrument of inherently limited usefulness. What put scintillation counting on a practical footing was the invention of the photomultiplier, the development of electronic circuits to exploit the photomultiplier's capabilities, and the formulation of more efficient scintillators. Together, these freed it from the constraints of the research laboratory and made it accessible to a varied range of scientific and industrial applications.

Today, scintillation counting is routine and widely used: not only in high-energy and radiation physics research but also in medicine, for diagnostic imaging and analysis, and in industry for thickness and density measurement, non-destructive analysis, and oil-well logging.

6.1 Scintillator-photomultiplier combination

Unlike the zinc sulphide screen, nearly all present-day scintillators are transparent. How much of the light from such a scintillator reaches the photomultiplier cathode depends on the shape, absorption, refractive index and surface finish of the scintillator, and how it is coupled to the cathode window. (Liquid scintillators, a special case, are considered in §7.2.1)

6.1.1 Scintillator shape

Scintillation photons are emitted in all directions; some reach the cathode directly, others by internal reflection. The best scintillator shape is one in which the total number of photons reaching the cathode is maximum and the number of internal reflections minimum; efficiencies of various scintillator shapes are dealt with extensively in the literature and in suppliers' catalogues. For many applications, truncated-cone and parabolic shapes give the most satisfactory results, but simple shapes like rods and blocks, which are less expensive to fabricate, are often good enough. Lately, scintillating fibres have become commercially available.

6.1.2 Scintillator finish

The surface finish and symmetry of the scintillator determine how much light escapes from it or is internally absorbed before reaching the photomultiplier cathode. Losses tend to be higher in highly symmetrical scintillators. They can be lessened by spoiling the symmetry or altering the finish of some surfaces, for instance by roughening or applying a reflective coating to them.

In long scintillators multiple internal reflections play an important role in transporting light from one end to the other. Scintillation plates are an example; the parallel surfaces should be polished to promote reflection and, in some cases, reflectively coated. The most suitable coatings are MgO, TiO₂, and Al. Figure 6.1 shows the reflection coefficients of these materials as a function of wavelength.

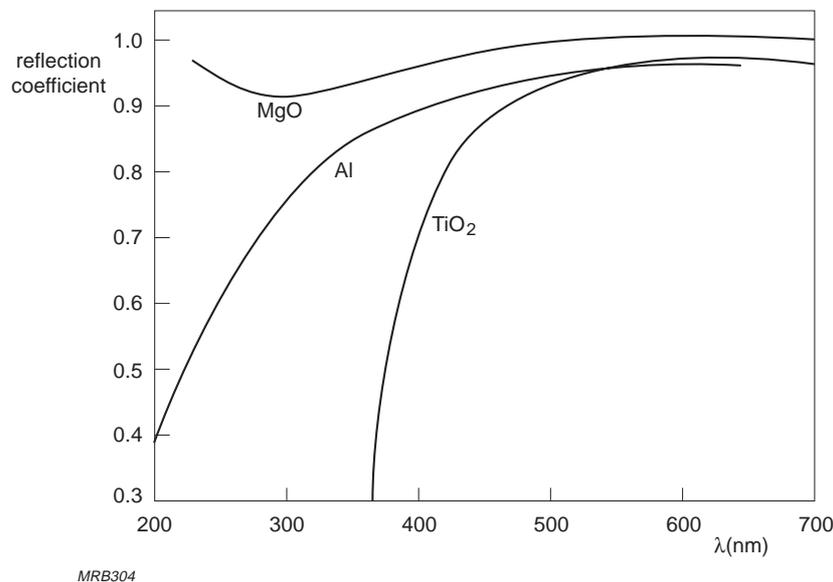
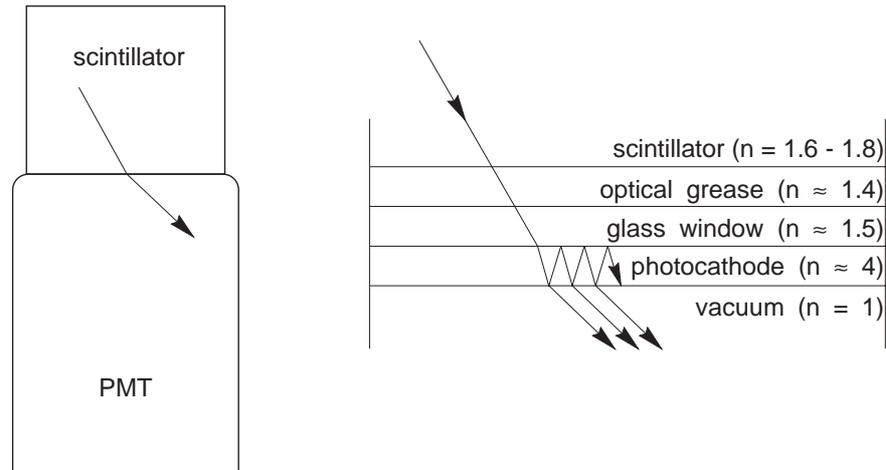


Fig.6.1 Reflection coefficients of aluminium, titanium dioxide and manganese oxide as functions of wavelength

6.1.3 Scintillator-photomultiplier coupling

The way the scintillator is coupled to the photomultiplier affects the proportion of photons that reach the cathode. Coupling efficiency is especially important in spectrometry, where maximum photon collection is essential. The fact that a semi-transparent cathode has a higher refractive index than the glass and vacuum on either side of it is fortunate; it encourages multiple internal reflections inside the cathode and so increases the probability of photoemission (Fig.6.2).



MRB305

Fig.6.2 Light transmission through the window and through the photocathode

Direct coupling. If the scintillator output surface is no larger than the cathode window, the two can simply be butted together. Use of a mating compound such as silicone grease, with a refractive index close to the refractive indices of the scintillator and the glass, is necessary to minimize interface losses.

Light-guide coupling. When it is not practical to couple the scintillator direct to the cathode window, either because its output surface is too large or too small or inconveniently shaped, or because operating circumstances necessitate keeping the two apart, a light guide has to be used. The usual materials are glass, fused silica, polystyrene, polyvinyltoluene, and especially polymethyl methacrylate (Perspex*, Lucite*, Plexiglass*, Altuglas*). The last named is less UV transparent than fused silica, but it has the advantage of being much easier to fabricate (Fig.6.3).

* Registered trademark

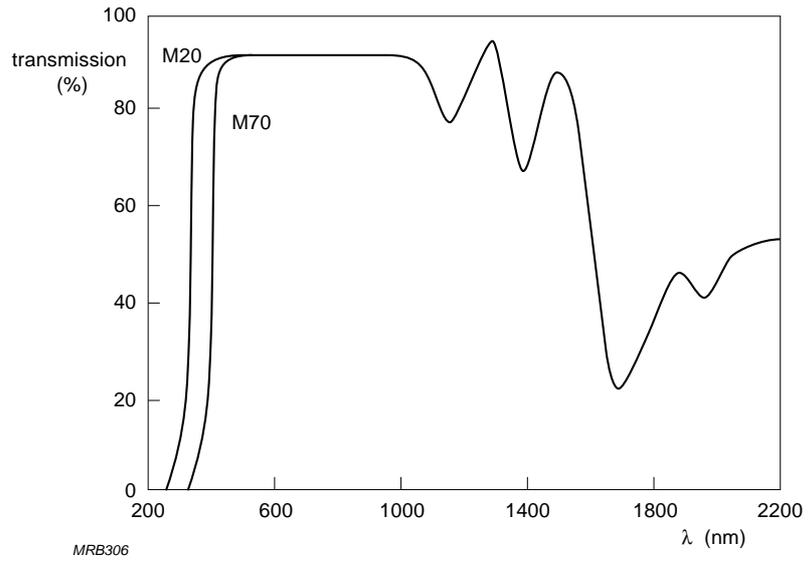
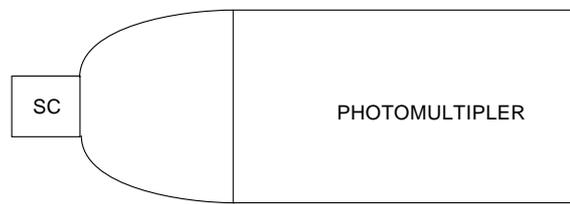
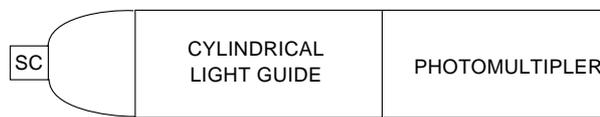


Fig.6.3 Transmission coefficient of Altuglas (composition M20 and M70) as a function of wavelength



(a)



(b)

MRB307

Fig.6.4 Logarithmic-spiral light guide (a) between scintillator and PMT, (b) with an intermediate cylindrical light guide

When the scintillator output surface is smaller than the cathode window, light guides with logarithmic-spiral (Fig.6.4), parabolic, or other special shapes can theoretically provide coupling efficiencies of 100%. In practice, however, losses due to internal reflection and absorption usually reduce that to something closer to 70%.

If the size and shape of the scintillator output surface are comparable to those of the cathode window, the light guide is cylindrical. Losses in such a light guide are characterized by an *attenuation length* in which the light intensity is halved; for a 25 mm diameter cylinder of polished methacrylate the attenuation length for visible light is more than a metre. Attenuation length is wavelength dependent and can be increased by using wavelength shifters (e.g. BBQ, K27, Y7 etc.) to shift the scintillation light to longer wavelengths, see below.

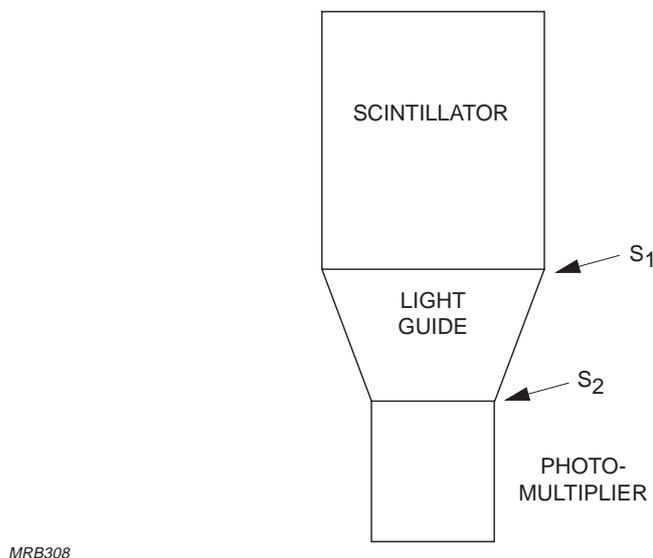
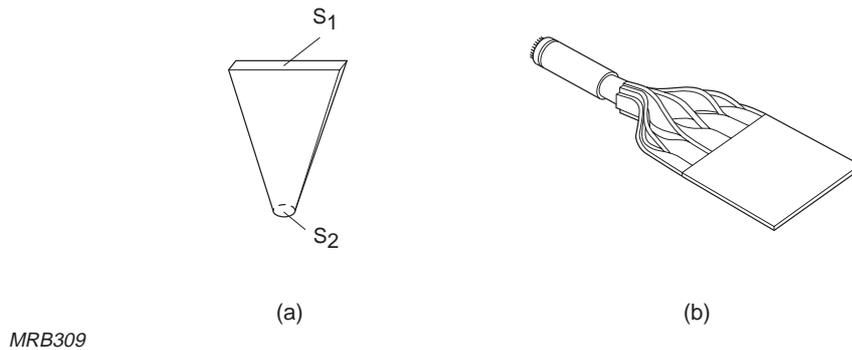


Fig.6.5 Truncated cone light guide

When the scintillator output surface is larger than the cathode window no light-guide shape is even theoretically 100% efficient. The best and most popular compromise is a truncated cone (Fig.6.5). Optimum design calls for a trade-off between the included angle of the cone and its length. The larger the angle, the sooner the internal reflections reach the critical angle of incidence and escape through the lateral surface of the light guide. On the other hand, the smaller the angle, the longer the light guide and the larger the absorption losses. The coupling efficiency never exceeds the ratio of the cathode area to the scintillator output area (S_2/S_1 , Fig.6.5). Therefore, the only practical application for such a light guide is when operating circumstances

necessitate having the scintillator remote from the cathode window; when that is not the case, direct coupling works just as well.

The fishtail and strip-light illustrated in Fig.6.6 are two examples of light guides designed to adapt specially shaped scintillators to circular cathode windows. Many others are described in the literature.



MRB309

Fig.6.6 The fishtail (a) and strip-light (b) are two examples of light guides designed to adapt specially shaped scintillators to circular cathode windows

Wavelength shifters. Particle physics experiments often use large scintillator plates that cannot be coupled direct to a photomultiplier; moreover, the light they emit is toward the blue end of the spectrum, where the transmission of most light guides is far from good. To overcome this difficulty, transparent media incorporating fluorescent substances are used that shift the colour toward the green, yellow, or red, where transmission is better. The well-known wavelength shifters BBQ, Y7 and K27 shift blue scintillation light into the green part of the spectrum where tubes with extended green sensitivity bialkali cathodes have good sensitivity. Tubes with S20 cathodes could also be used, but price considerations normally limit them to use only with yellow or red wavelength shifters. Use of different fluorescent substances in different parts of a wavelength shifter can make particle discrimination possible.

Wavelength shifters usually take the form of a bar or a sheet air-coupled to the scintillator. It may, in fact, integrate the light from many scintillators and may be coupled to a separate light guide or itself serve as one. Wavelength-shifting fibres are gaining in popularity despite their low efficiency when used, for example, together with flat scintillators.

Wavelength-shifting layers can also be deposited on the photomultiplier window. The main disadvantages of this are low efficiency, loss of time resolution and degraded linearity.

6.2 Electrical signal characteristics

6.2.1 Current pulse

The light emission $L(t)$ of a scintillator excited at time $t = 0$ can be approximated by

$$L(t) = \frac{\bar{n}_{p,s}}{\tau} \exp\left(-\frac{t}{\tau}\right)$$

where $\bar{n}_{p,s}$ is the mean number of photons per scintillation and τ is the decay time constant of the scintillator. The resulting number of electrons at the photomultiplier anode is given by the convolution product

$$S(t) = L(t) * R_{\delta}(t)$$

where $R_{\delta}(t)$ is the photomultiplier pulse response. If the photomultiplier response-pulse width t_w is negligible compared with τ , this reduces to

$$S(t) \approx \frac{\bar{n}_{a,s}}{\tau} \exp\left(-\frac{t}{\tau}\right)$$

where $\bar{n}_{a,s}$ is the mean number of anode electrons per scintillation; or, in terms of the anode current I_a and anode charge $q_{a,s}$

$$I_a(t) = \frac{\bar{q}_{a,s}}{\tau} \exp\left(-\frac{t}{\tau}\right) \quad (6.1)$$

6.2.2. Voltage pulse

In the circuit of Fig.6.7 the current pulse of Eq.6.1 gives rise to a voltage pulse

$$V_a(t) = \frac{\bar{q}_{a,s}}{C} \frac{\theta}{\tau - \theta} \left\{ \exp\left(-\frac{t}{\tau}\right) - \exp\left(-\frac{t}{\theta}\right) \right\} \quad (6.2)$$

where $\theta = R_L C$. Figure 6.8 shows the pulse shape; the amplitude is given by

$$V_a = \frac{\bar{q}_{a,s}}{C} \left(\frac{\theta}{\tau} \right)^{\frac{1}{1 - \frac{\theta}{\tau}}}$$

As θ/τ tends toward infinity, V_a tends toward the limit $\bar{q}_{a,s}/C$. Figure 6.9 shows the effect of θ/τ on relative amplitude, and Fig.6.10 its effect on pulse shape; in Fig.6.11 the curves of Fig.6.10 are normalized to unit amplitude. Figure 6.12 shows the

variation of rise time with θ/τ , and Fig.6.13 the variation of full width at half maximum, in both cases normalized with respect to τ .

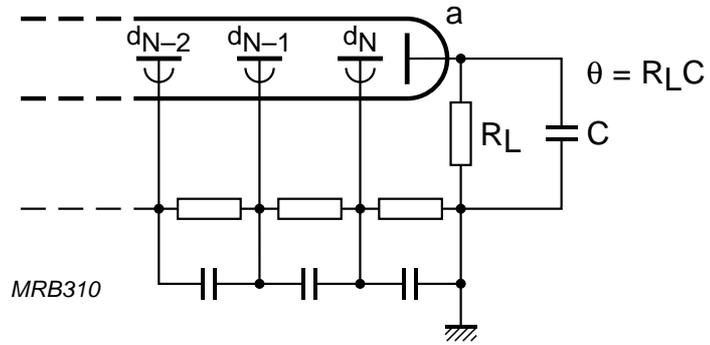


Fig.6.7 Anode circuit with the anode load $R_L C$

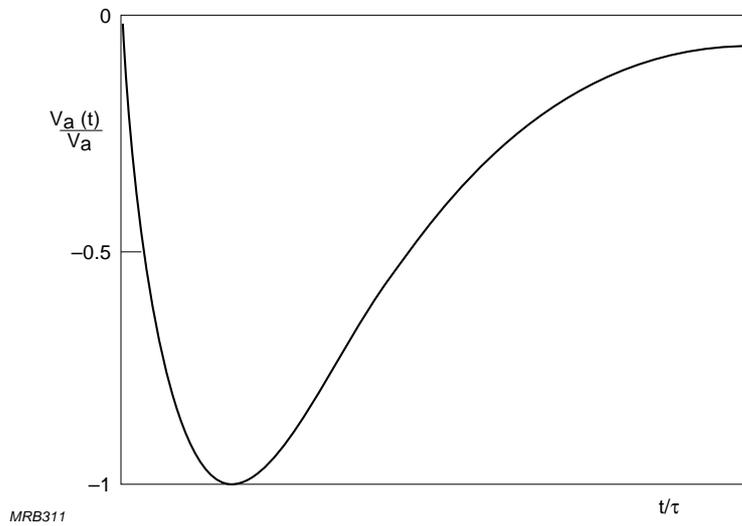


Fig.6.8 Voltage pulse shape over the anode load $R_L C$

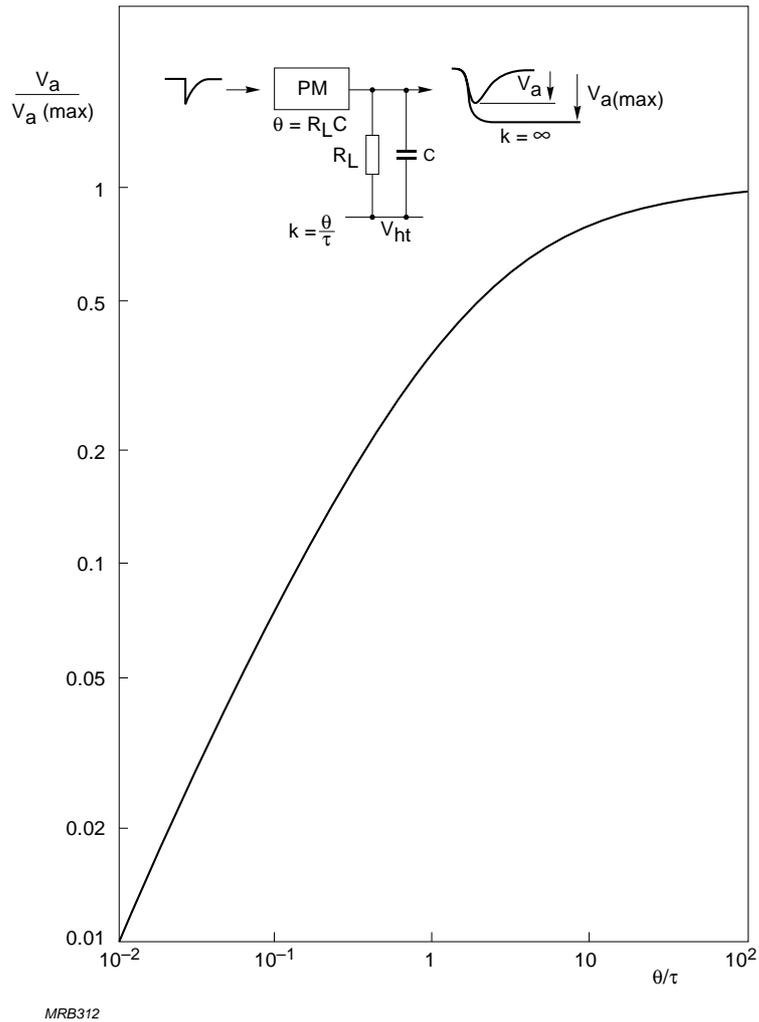


Fig. 6.9 Relative pulse amplitude variations as a function of the ratio θ/τ , τ being the scintillator decay time constant

Special cases

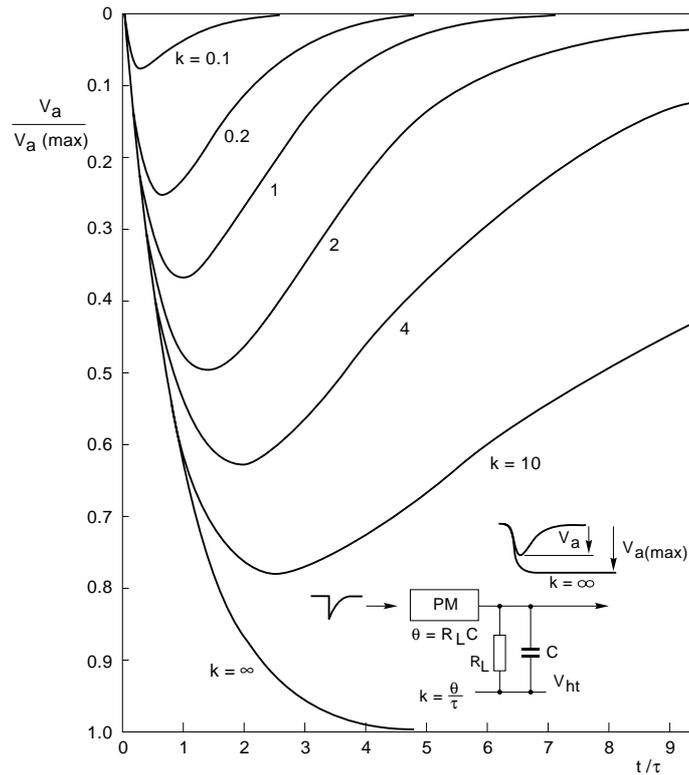
When $\theta \ll \tau$, Eq.6.2 approaches the limit

$$V_a(t) = \frac{\bar{q}_{a,s}}{C} \frac{\theta}{\tau} \left\{ 1 - \exp\left(-\frac{t}{\theta}\right) \right\} \exp\left(-\frac{t}{\tau}\right)$$

corresponding to an amplitude

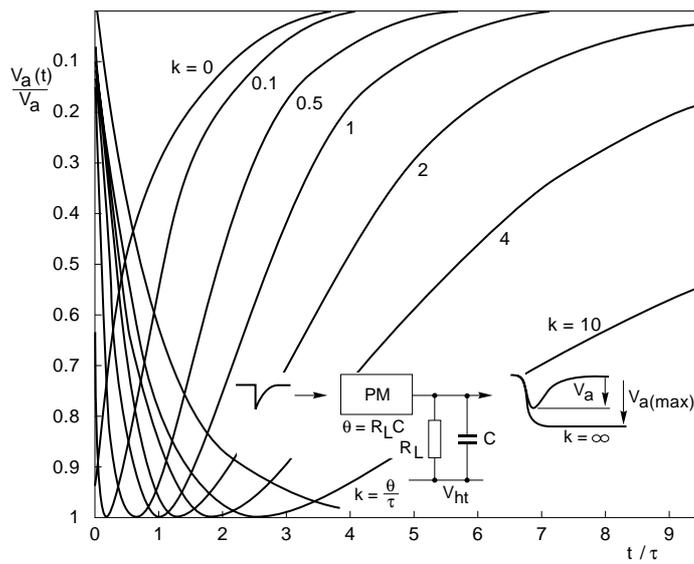
$$V_a = \frac{\bar{q}_{a,s}}{C} \frac{\theta}{\tau}$$

attained at a time $t_m = \tau \ln(\tau/\theta)$, which tends toward zero as θ/τ does.



MRB313

Fig. 6.10 Relative pulse amplitude shapes for an anode load $R_L C$ with $k = \theta/\tau$ as parameter



MRB314

Fig. 6.11 Relative pulse amplitude shapes for an anode load $R_L C$ with $k = \theta/\tau$ as parameter and the amplitudes normalized to 1

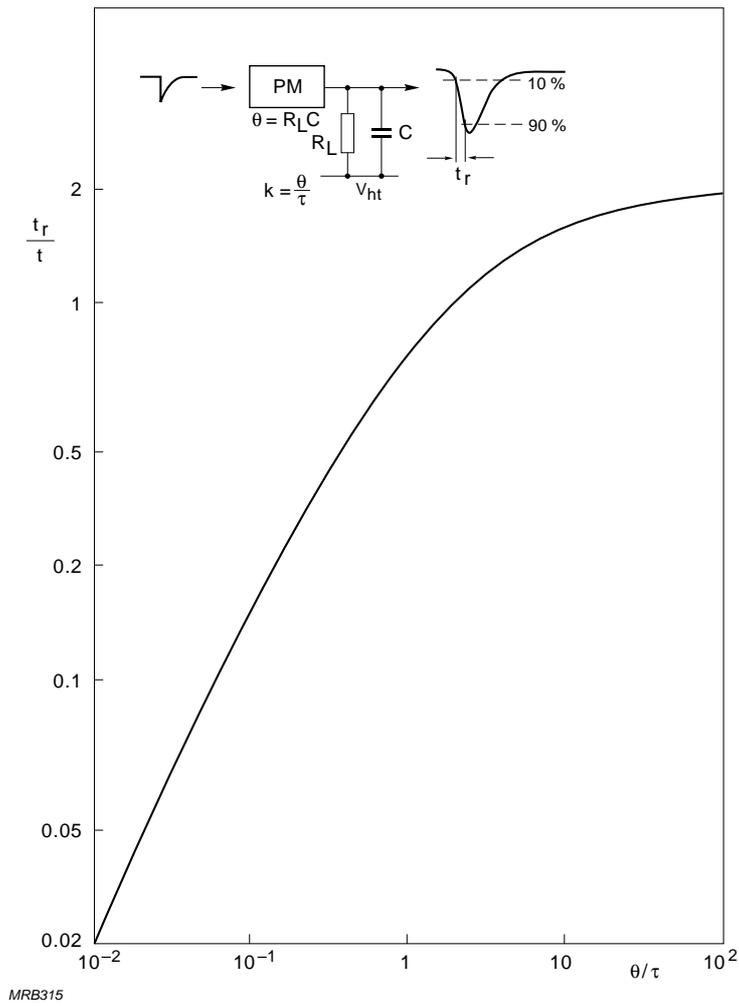


Fig.6.12 Variation of pulse rise time with θ/τ

When $\theta = \tau$, Eq.6.2 reduces to

$$V_a(t) = \frac{\bar{q}_{a,s}}{C} \frac{t}{\tau} \exp\left(-\frac{t}{\tau}\right)$$

corresponding to an amplitude

$$V_a = \frac{\bar{q}_{a,s}}{C} \exp(-1)$$

attained at a time $t_m = \tau$.

When $\theta \gg \tau$, Eq.6.2 approaches the limit

$$V_a(t) = \frac{\bar{q}_{a,s}}{C} \left\{ 1 - \exp\left(-\frac{t}{\tau}\right) \right\} \exp\left(-\frac{t}{\theta}\right)$$

corresponding to an amplitude

$$V_a = \frac{\bar{q}_{a,s}}{C} = V_{a \text{ max}}$$

attained at a time $t_m = \tau \ln(\theta/\tau)$, which tends toward infinity as θ/τ does.

Low values of θ/τ allow high count rates and lessen the risk of pulse pile-up, but, because of the larger bandwidth of the circuit, they worsen the signal-to-noise ratio.

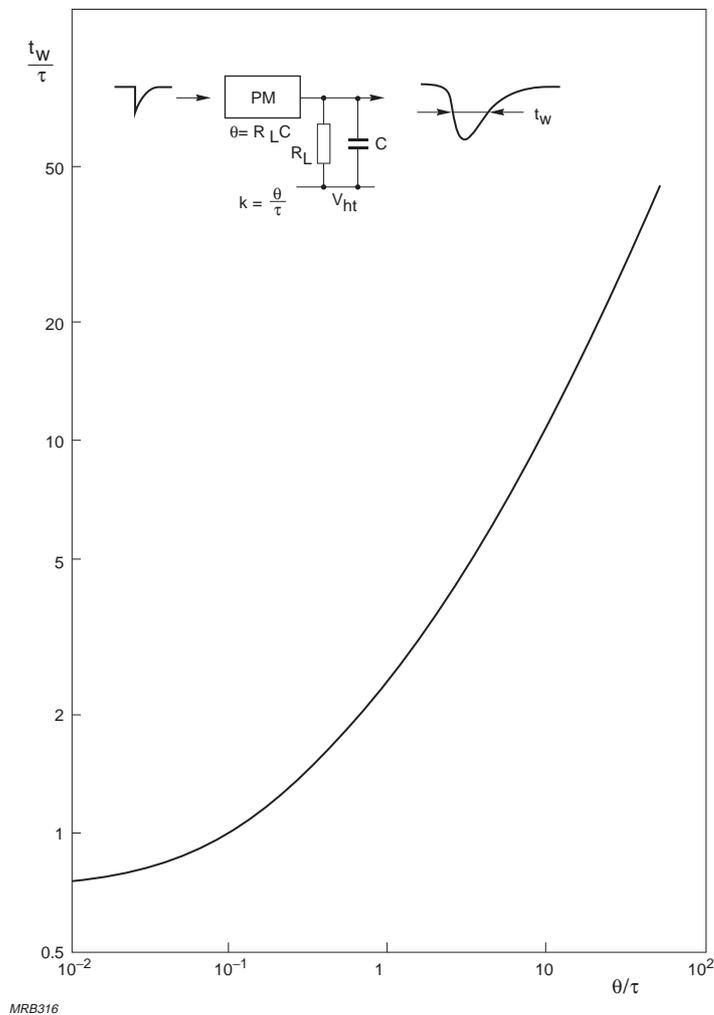


Fig.6.13 Variation of full width at half maximum t_w with θ/τ

6.3 Operating modes

6.3.1 Continuous mode

When a scintillator is used in the continuous mode (as in thickness monitoring, for instance) the photomultiplier anode charges are integrated rather than counted (Fig.6.14). Owing to the quantum nature of radiation and the statistics of energy conversion in the scintillator and photomultiplier, there is therefore some fluctuation of the resulting signal (§3.2.7). The amplitude of this fluctuation depends on the time constant $R_L C$ and the bandwidth of the output electronics or instrument.

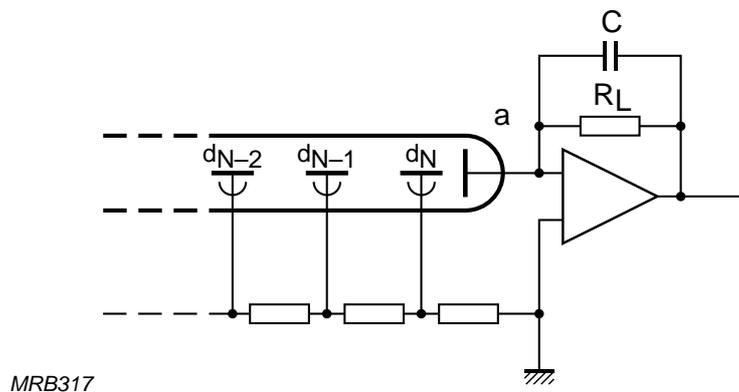


Fig.6.14 Circuit for a photomultiplier in continuous mode

6.3.2 Pulse mode

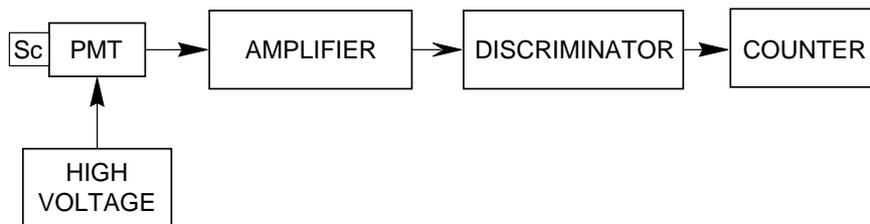
In pulse-mode operation the integration time constant has to be short; how short depends on:

- the pulse characteristics to be measured (amplitude, frequency)
- the pulse characteristics required by the electronics (rise time, duration).

For pulse counting, the time constant is usually made as short possible to be compatible with the high resolution of the detector system. For energy spectrometry it must be long enough to integrate all the charges due to a single event; however, if is too long there is a risk of pulse pile-up. For time spectrometry it is made as short as possible (§5.4.2), often of the same order of magnitude as the transit-time spread of the tube.

Pulse counting

In the set-up illustrated in Fig.6.15 the discriminator selects pulses that exceed a set threshold, or are between a set minimum and maximum (the amplitude window), and passes them to the counter. The pulse frequency registered by the counter is a measure of the radiation intensity to which the scintillator is exposed.



MRB318

Fig.6.15 Block diagram of a scintillator pulse counter

Together, the discriminator threshold and the gain characteristic of the tube define a counting plateau (Fig.6.16) within which the count rate due to a constant radiation flux varies very little as a function of the high voltage applied to the tube. Such a plateau is usually described in terms of its slope (in per cent change of count rate per volt) and its length in volts. Afterpulses, whose rate increases with the applied voltage (§4.7), and dark noise, which is related to field emission, are important factors affecting the slope.

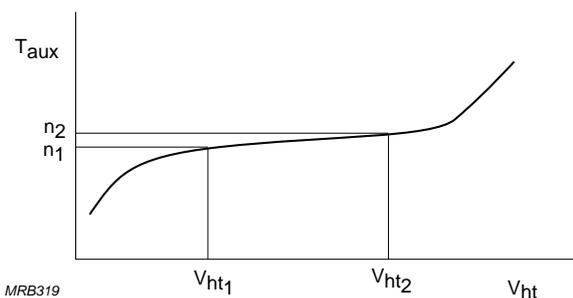


Fig.6.16 The discriminator threshold and gain characteristics of the tube define a counting plateau

Resolution time is the least interval between successive scintillations that the counter can distinguish. It is determined either by the decay time constant of the scintillator or the time characteristics of the electronics. With present-day scintillators resolution times of the order of 1 ns can be obtained.

If the dead time of the counter, τ_d , is comparable with the mean interval separating successive scintillations, the counting error may be appreciable. If n_e is the observed count rate, the true count rate corrected for the dead time is

$$n = \frac{n_e}{1 - n_e \tau_d}$$

Environmental effects may increase or decrease the count rate. Background radioactivity increases it. Absorption between the source and the scintillator decreases it. Absorption can be a significant cause of counting error for highly ionizing radiations such as α - and soft β -rays, especially if their free path in the ambient medium is short compared with the distance between source and scintillator. Scatter may increase or decrease the count rate, depending on whether the net deflection of radiation is into or away from the detector.

Statistical considerations also affect counting accuracy. Even when the dead time of the counter is small compared with the mean time interval between scintillations (and environmental effects are negligible), the accuracy of the count is limited by the stochastic nature of the phenomena observed. For example, when measuring radiation intensity (nuclear particles or photons), the measured value of the count rate will have no fixed level but will vary from observation to observation. An estimate of the accuracy of the count rate measured can be made on a statistical basis.

The value of the radiation intensity under observation can be defined as a true mean count rate m taken over an infinite number of observations. The measured count rate is then a variable X whose value x is distributed about m according to a probability distribution appropriate to the phenomena under observation. This distribution will be characterized by the true mean value m and an ideal standard deviation σ_X . Although, when a single observation of X is made, there is no way of knowing whether this is close to m , when N observations x_i of X are made the results can be analysed statistically and an indication of the *probable* accuracy obtained.

First, calculate the experimental mean

$$\bar{x} = \frac{1}{N} \sum_{i=1}^N x_i$$

and the variance about the experimental mean

$$s^2 = \frac{1}{N} \sum_{i=1}^N (x_i - \bar{x})^2$$

The calculated values of both \bar{x} and s^2 will vary from one set of observations to another, and are therefore themselves random variables. They are, however, unbiased estimators of the values m and σ_X required.

When N is large it can be shown that:

- the experimental mean \bar{x} approaches the true mean m
- the value $(\bar{x} - m)$ has a gaussian distribution of variance σ_X^2/N .

Let the mean experimental value be the expectation E of that value, so that the mean value of the experimental variance is $E(s^2)$. Then, with m as the origin of the distribution

$$Ns^2 = \sum_{i=1}^N x_i^2 - N\bar{x}^2$$

The average of N observations is

$$NE(s^2) = NE(x_i^2) - NE(\bar{x}^2) = N\sigma_X^2 - \frac{N\sigma_X^2}{N} = (N - 1)\sigma_X^2$$

and

$$E(s^2) = (N - 1)\frac{\sigma_X^2}{N}$$

Thus, s^2 approaches σ_X^2 as N increases.

Using the gaussian distribution, it is possible to determine the minimum number of observations, N , that must be made so that, with a probability P , \bar{x} is within a given percentage of m . This probability is termed the confidence level of the measurement, Fig.6.17, and is given by $P = (1 - 2\alpha)$, where α is the probability that \bar{x} is further from m by the required amount in any one direction.

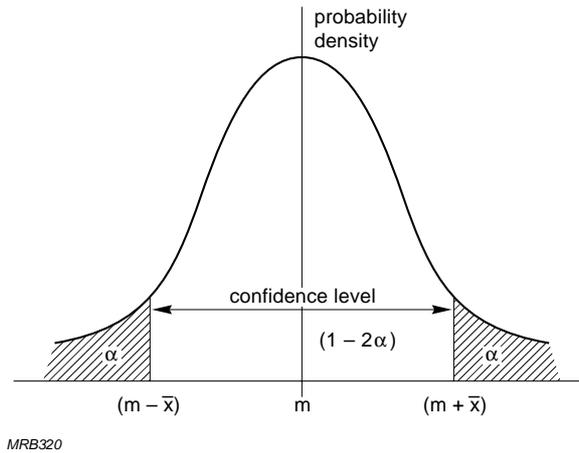


Fig.6.17 The confidence level is the probability that an observation \bar{x} is within a given percentage of the origin m of the distribution

Values of 2α are given in Table 6.1 and plotted in Fig.6.18 for $\kappa\sqrt{N}/\sigma_X$, where the modulo $\kappa = m/(m - \bar{x})$. Thus, to find the number of observations N such that, with a confidence level of 95%, \bar{x} is within 10% of m , read the value of $\kappa\sqrt{N}/\sigma_X$ for $2\alpha = 1 - 0,95 = 0,05$. In Fig.6.18 this is about 2, so, since $\kappa = m/10$

$$\frac{m\sqrt{N}}{10\sigma_X} > 2$$

and

$$N > 400 \left(\frac{\sigma_X}{m} \right)^2$$

Table 6.1 Spot value of 2α

$\kappa\sqrt{N}/\sigma_x$	0.5	0.6745	1	1.349	2	2.024	2.698	3
2α	0.617	0.5	0.317	0.178	0.0455	0.0341	0.00706	0.00272

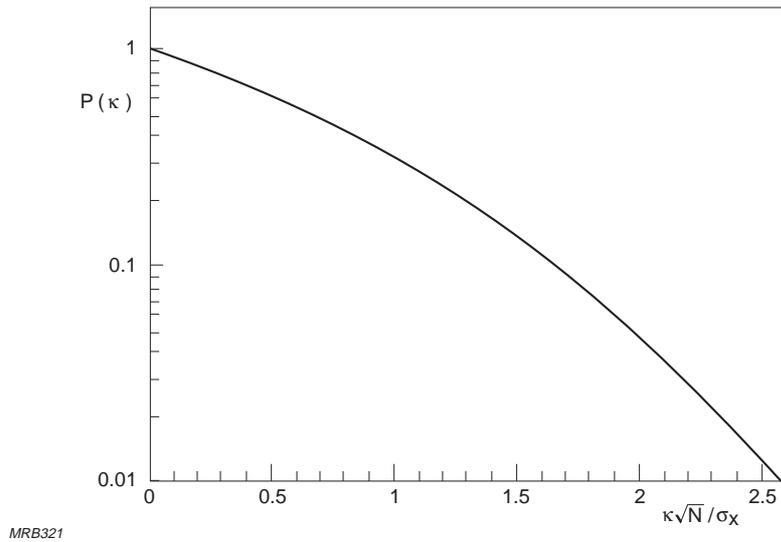
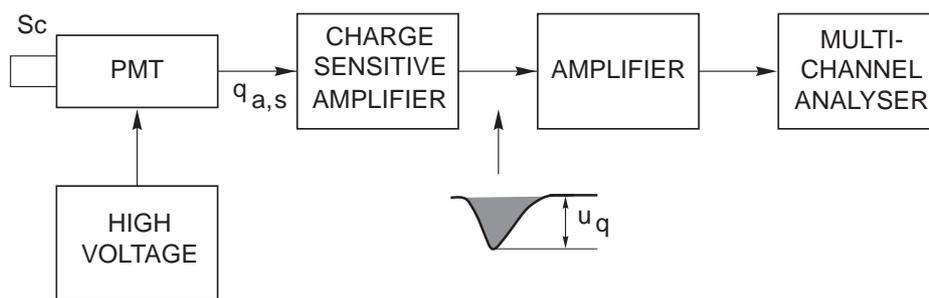


Fig.6.18 Probability $P(\kappa)$ that a variable $(\bar{x} - m)$ is greater than the modulo κ

Energy spectrometry

Figure 6.19 illustrates an energy spectrometer using a multichannel pulse-height analyser. Each interaction in the scintillator gives rise to a photomultiplier anode charge $q_{a,s}$ which is proportional to the energy spent in the interaction. The charge amplifier converts the charge into a proportional voltage pulse u_q which is then amplified and applied to the multichannel analyser. As successive pulses occur the analyser constructs a histogram of their amplitudes.



MRB322

Fig.6.19 Energy spectrometer using a multichannel pulse-height analyser

Energy resolution. For homogeneous (monoenergetic) radiation completely absorbed in the scintillator, the histogram should theoretically consist of a single narrow peak

whose abscissa corresponds to the characteristic energy of the radiation. For non-homogeneous radiation the histogram should consist of a number of such peaks, each with an abscissa corresponding to one energy component of the radiation and a height proportional to the relative frequency of occurrence of that component.

The narrowness of the peaks is a measure of the energy resolution of the scintillator-photomultiplier combination. The following effects always cause some widening of the peaks, and hence loss of resolution:

- Photon emission in the scintillator fluctuates from one monoenergetic interaction to another, mainly because of inhomogeneities in the scintillator and an inherent resolution-degrading component originating from the non-proportionality of the light yield, typical for each scintillator*.
- Photon collection at the cathode fluctuates from one scintillation to another because of variation of the points of origin of successive scintillations, consequent variations of photon path lengths and angles of incidence, and local variations of scintillator window, coupling compound, and cathode window transparency and refractive index.
- Photon-electron conversion efficiency fluctuates because of local variations of cathode sensitivity.
- Electron collection and multiplication fluctuate, mainly because of variations in photoelectron points of origin.

All these cause the voltage pulse u_q to fluctuate from one scintillation to another. If the number of photons emitted per scintillation is large, the pulse height distribution is approximately gaussian, with mean value \bar{u}_q and standard deviation σ_u . From Eq.2.14 the energy resolution R_e is then $\Delta u_q / \bar{u}_q$, where $\Delta u_q = 2.36 \sigma_u$. Or, in terms of the photomultiplier parameters (Eq.3.34),

$$R_e = 2.36 \left(v_{np,s} + \frac{1 - \rho + v_G}{\bar{n}_{p,s} \rho} \right)^{3/2} \quad (6.3)$$

where $\bar{n}_{p,s}$ is proportional to the primary energy E_p absorbed in the scintillator. Equation 6.3 shows that R_e varies inversely as $\sqrt{E_p}$. (The Poisson distribution assumption, $v_{np,s} = 1/\bar{n}_{p,s}$, is, however, not entirely valid, so the resolution is in fact not quite so good as Eq.6.3 would suggest.)

* This is treated in detail in "Inorganic Scintillation Detectors in γ -ray Spectrometry" by M. Moszynski, Elsevier Science, NIM 2002, in print.

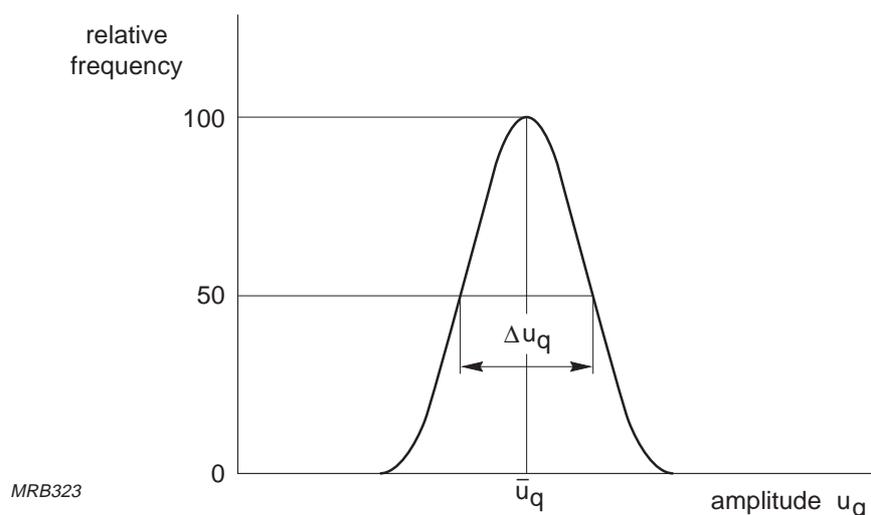


Fig.6.20 Energy resolution $\Delta u_q/u_q$

Energy resolution is measured by taking the FWHM of the full-energy peak (Fig.6.20), dividing it by the energy corresponding to the peak maximum, and multiplying by 100. Commercial diameter 50 mm × 50 mm and diameter 38 mm × 38 mm aluminium encased NaI(Tl) scintillators used in combination with present-day photomultipliers give resolutions of 6–7% for the 662 keV γ -radiation of ^{137}Cs , and 8–9% for the 122 keV γ -radiation of ^{57}Co (Fig.6.21). These are standard sources for such measurements.

A standard source for low-energy measurements is the 5.9 keV X-radiation of ^{55}Fe , and the scintillators usually have a window of beryllium instead of aluminium, for minimum attenuation of the radiation. In the example shown in Fig.6.22 the energy resolution is 30%; the measurement was made using a 32 mm diameter, 2 mm thick NaI(Tl) scintillator with a 36.8 mg/cm² beryllium window.

Another parameter of interest is the peak-to-valley ratio P/V defined as the ratio of the peak value of the amplitude distribution curve to the minimum value of the valley to the left of the peak. The P/V ratio in Fig.6.22 could be estimated as 40:1.

Because electron collection and multiplication in the photomultiplier input stages is a dominant statistical factor in the conversion of incident photon energy to electrical energy, the energy resolution also depends strongly on the voltages applied to those stages. Figure 6.23 shows examples of energy resolution variation as a function of cathode to first-dynode voltage obtained with three radiation sources: ^{55}Fe , ^{57}Co , and ^{137}Cs . From these it is apparent that energy resolution is best when the cathode to first dynode voltage is between about 200 V and 300 V.

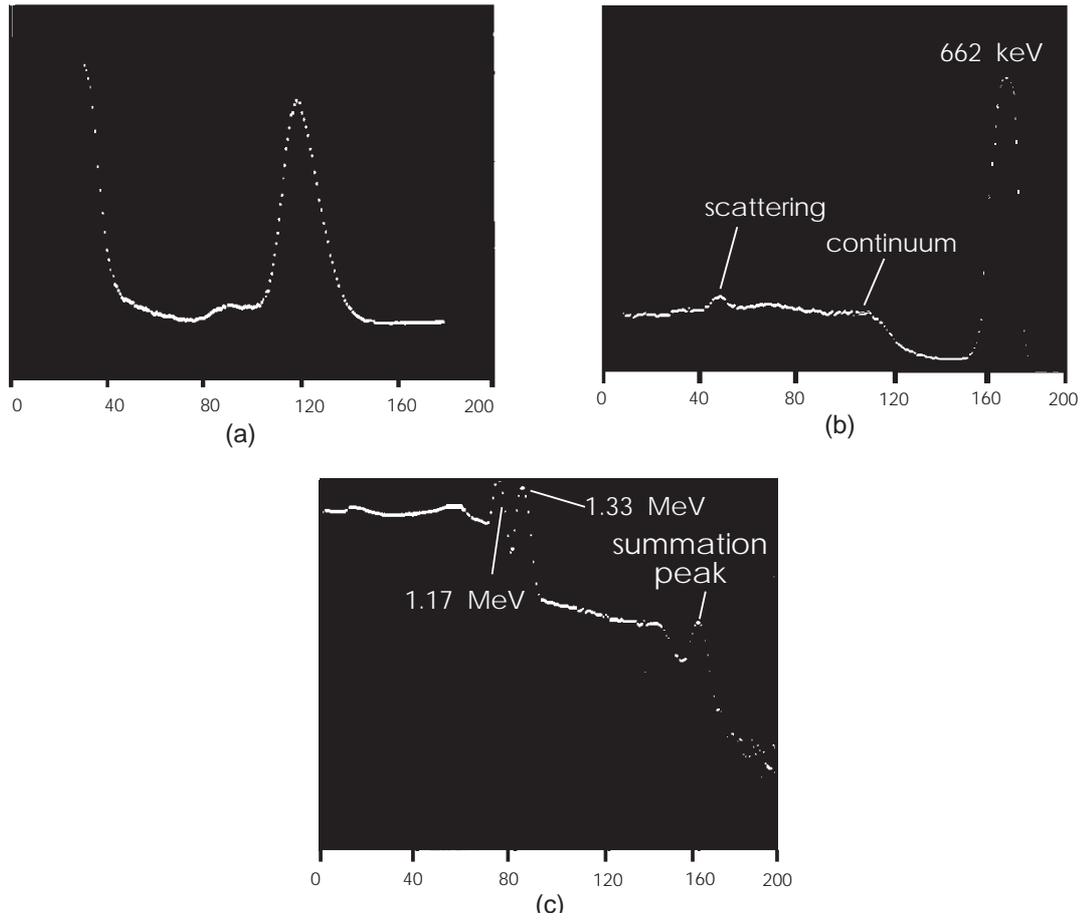


Fig.6.21 Scintillator pulse amplitude distribution using sources of (a) ^{57}Co , (b) ^{137}Cs and (c) ^{60}Co

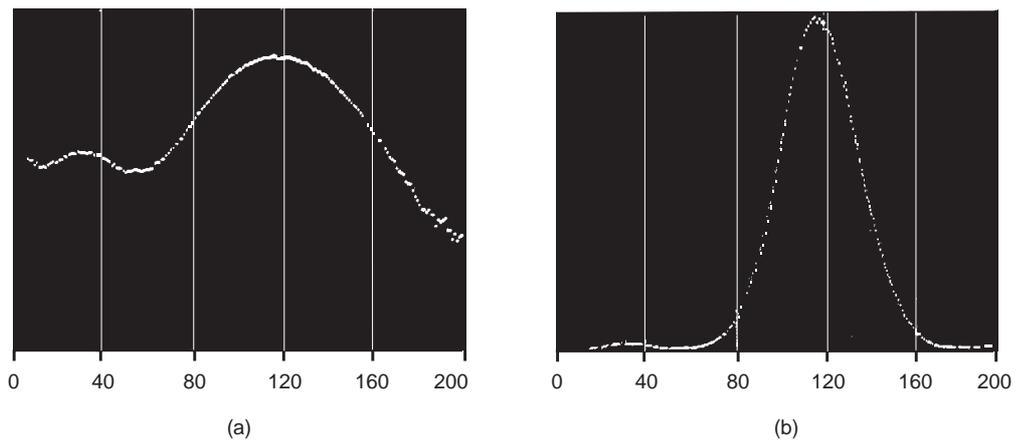


Fig.6.22 Scintillation pulse amplitude distribution for a ^{55}Fe source with (a) a logarithmic and (b) a linear scale on the y axis

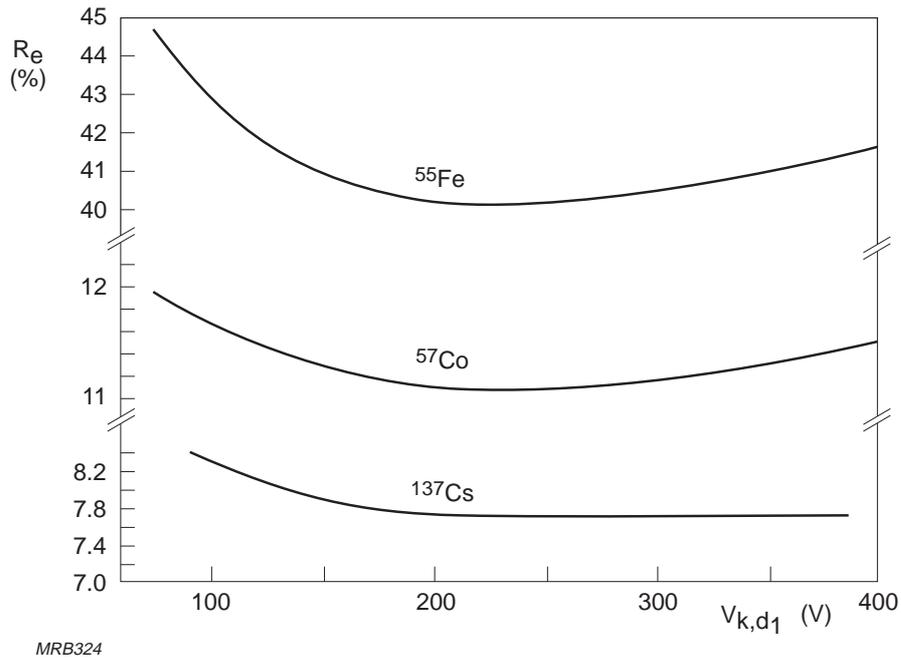


Fig.6.23 Energy resolution variation as a function of cathode to first-dynode voltage. Sources ^{55}Fe , ^{57}Co and ^{137}Cs

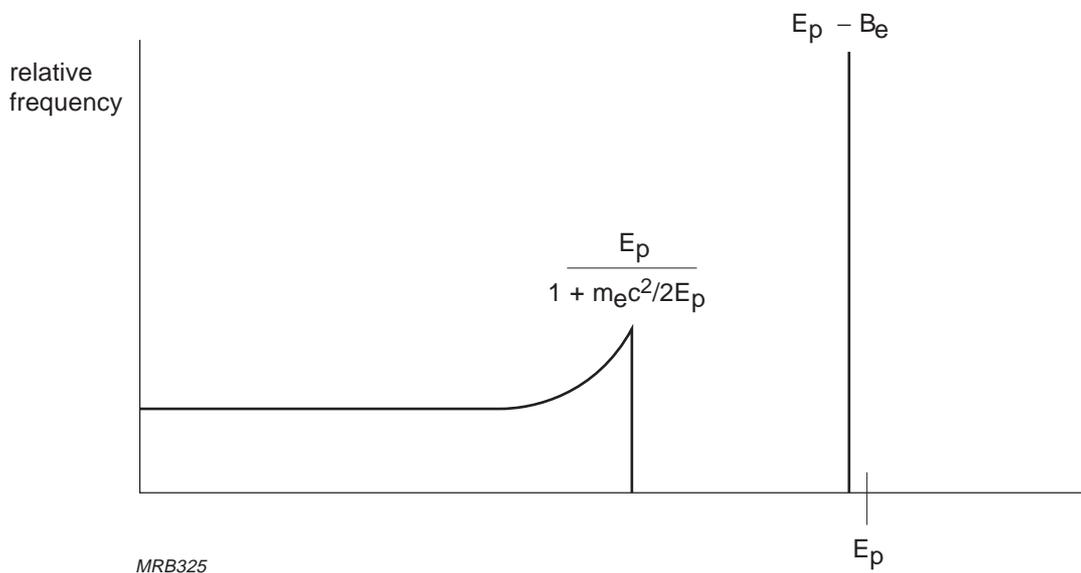


Fig.6.24 Spectrum of primary electron energies produced in a scintillator by monoenergetic γ radiation of energy E_p , showing the Compton continuum, the Compton edge and the photoelectric line

Composition of the pulse-height spectrum. The proportionality between the primary energy E_p absorbed in the scintillator and the mean number of photons received at the cathode, $\bar{n}_{p,s}$, depends on the interaction processes by which radiation absorbed in the scintillator liberates electrons and thereby excites fluorescence (see Appendix to this chapter). If there were but one such process, the pulse-height spectrum due to a monoenergetic source of X- or γ -radiation would consist of a single peak – the ‘full energy peak’ – corresponding to the energy E_p .

In reality, however, the composition of the spectrum is not that simple (Fig.6.24). There are not one but three primary processes by which radiation is absorbed, each of which leaves its own signature on the spectrum. Moreover, there are various secondary phenomena which also play a part. The three primary processes are: photoelectric effect, Compton effect, and pair production. Each of these liberates electrons of a characteristic energy E_s which excite fluorescence of a corresponding intensity.

– *Photoelectric effect:* $E_{s,p} = E_p - B_e$

Here, B_e is the binding energy of the liberated electron, which is usually ejected from a K or L shell. When the vacancy left by its ejection is again filled, the energy B_e reappears as characteristic X-rays and Auger electrons. If these occur within the response time for the primary photoelectric event, the fluorescence they excite also contributes to the photo-peak corresponding to that event. In that case, the photo-peak becomes equivalent to a full-energy peak E_p .

– *Compton effect:* $0 \leq E_{s,c} \leq E_p/(1 + m_e c^2/2E_p)$

where m_e is the rest mass of the electron and c the free-space velocity of light ($m_e c^2 = 0.511$ MeV). Besides the primary electron of energy $E_{s,c}$, a Compton interaction yields a secondary photon of energy

$$E'_p = \frac{E_p}{1 + (1 - \cos\theta)\frac{E_p}{m_e c^2}}$$

where θ is the angle between the path of the secondary photon and the original X- or γ -photon. If the secondary photon is absorbed within the scintillator, the Compton interaction contributes to the full-energy peak. Otherwise, it yields a continuous spectrum (the ‘Compton continuum’) extending from zero up to the energy $E_p/(1 + m_e c^2/2E_p)$, which corresponds to a spectral feature known as the ‘Compton edge’.

- *Pair production*: $E_{s,pp} = E_p - 2m_e c^2$
 where $2m_e c^2$ represents the rest-mass energy of an electron-positron pair created by the absorption of a γ -photon. The contribution of pair-production events to the spectrum depends on whether both, one, or neither of the photons due to subsequent annihilation of the positron escapes from the scintillator without further interaction. If both escape, the pair-production event contributes to a spectral peak at $E_p - 1.02$ MeV, which is usually superimposed on the Compton continuum. If only one escapes, the event contributes to a peak at $E_p - 0.511$ MeV. And if neither escapes, it contributes to the full-energy peak E_p . All three of these possibilities may appear in a single spectrum.

Secondary effects that may influence the shape of the pulse-height spectrum include X-escape, back-scatter, annihilation, and sum peaks:

- *X-escape peak*. If an X-photon due to a photoelectric interaction escapes from the scintillator, the scintillation energy for that interaction will be correspondingly diminished. In an NaI(Tl) scintillator the loss amounts to some 28 keV. In the spectra of low-energy radiation sources this gives a secondary peak 28 keV to the left of the photo-peak. In higher-energy spectra the X-escape peak is so close to the photo-peak that it is indistinguishable.
- *Back-scatter peak*. Compton effect interactions can scatter low-energy photons back into the scintillator from the surroundings. These may produce an observable photo-peak in the Compton continuum, usually centred at an energy slightly greater than $E_p / (1 + 2E_p / m_e c^2)$.
- *Annihilation peak*. Pair-production interactions in the immediate surroundings (or in the radiation source itself) may scatter annihilation photons into the scintillator. These produce a spectral peak at 0.511 MeV.
- *Sum peaks*. Sources that emit γ -radiation of two distinct energies, E_{p1} and E_{p2} , produce a spectrum with two corresponding full-energy peaks. However, when the two radiations are emitted simultaneously within the response time of the scintillator-photomultiplier combination, they produce a third spectral peak corresponding to $E_{p1} + E_{p2}$. Examples of sources that produce such peaks are ^{60}Co (see Fig.6.21(c)) and ^{24}Na . The latter produces a sum peak corresponding to its 1.274 MeV γ -radiation plus the 0.511 MeV due to annihilation of a pair-production positron.

In general, the larger the scintillator, the more prominent is the full-energy peak relative to such features as the Compton continuum and pair-production, back-scatter and X-escape peaks. This is illustrated by Fig.6.25 in which comparable spectra obtained with three sizes of scintillator have been normalized with respect to the area under the full-energy peak. The fraction of the total recorded pulses that appears under the full-energy peak is called the photoelectric efficiency, ϵ_{ph} , of the scintillator.

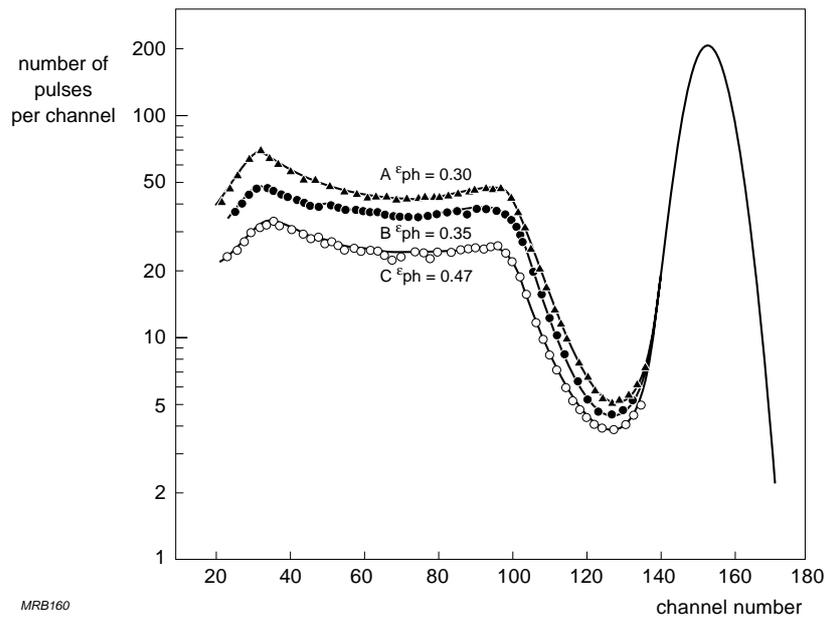


Fig.6.25 Scintillation pulse amplitude distributions with ^{137}Cs sources and scintillators of dimensions $\phi 38 \text{ mm} \times 25 \text{ mm}$ (A), $\phi 50 \text{ mm} \times 50 \text{ mm}$ (B) and $\phi 75 \text{ mm} \times 75 \text{ mm}$ (C). The distributions are normalized to give the same area under their full energy peak curves

Time spectrometry

Time spectrometry is used to measure time relations between events such as the emission of particles or photons in cascade or the de-excitation of nuclear states. It requires fast-response photomultipliers and fast-response scintillators with high scintillation effectiveness (Appendix A6.3).

The time spectrometry principle can also be used to measure scintillator response time, as in the set-up shown in Fig.6.26. Here, the two γ -photons emitted in cascade by a ^{60}Co source provide the basis for measurement. Pulses from the two photomultipliers are amplified, shaped, and applied via discriminators to a time-to-pulse-height converter. A multichannel analyser at its output generates a distribution plot (Fig.6.27) of the time differences measured for each pulse pair.

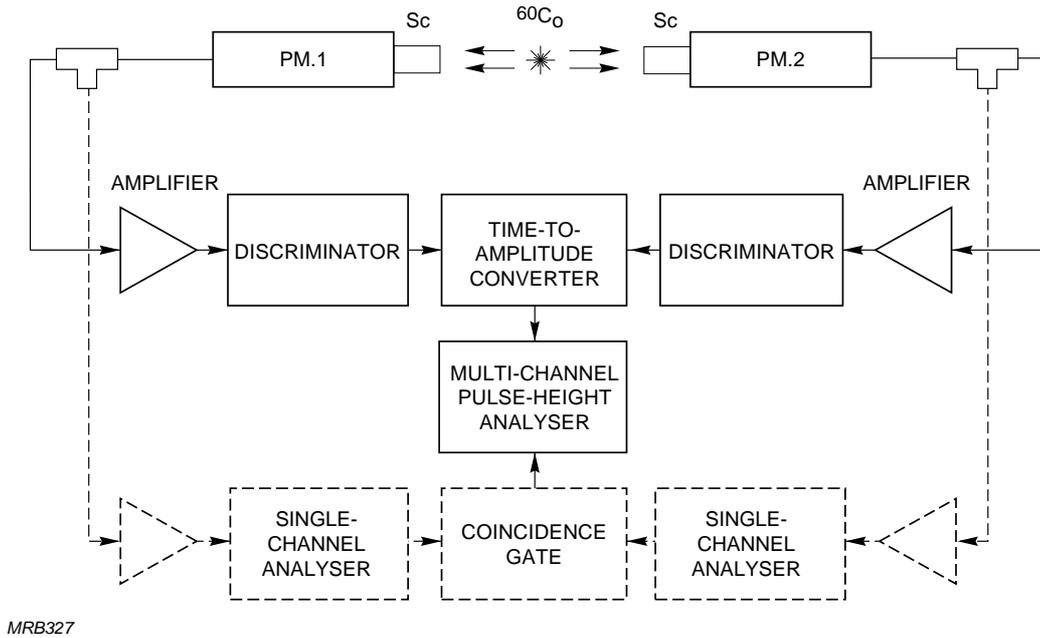


Fig.6.26 Basic setup for time spectrometry

To minimize errors due to photomultiplier pulse amplitude fluctuations (‘walk errors’), the parallel channel shown dotted in Fig.6.26 gates the multichannel analyser to accept time-proportional pulses only when the corresponding photomultiplier pulses occupy a set amplitude window. Walk errors can be further reduced by using discriminators in the main channel that trigger at an optimum point on the pulse leading edges regardless of peak pulse amplitude (‘constant fraction triggering’).

With respect to the instant of γ -photon emission, the response-time variance of either of the two detectors is, from Eq.4.18

$$\sigma_t^{*2} = \frac{\tau^2 + \sigma_t^2}{\bar{n}_{k,s}}$$

where τ is the scintillation decay time constant, σ_t the standard deviation of the photomultiplier transit time fluctuations, and $\bar{n}_{k,s}$ the mean number of photoelectrons emitted at the photocathode per scintillation. If the statistics are gaussian, the variance of the two detectors together can be obtained by summing the individual variances; thus,

$$\sigma_{t,t}^{*2} = 2 \frac{\tau^2 + \sigma_t^2}{\bar{n}_{k,s}}$$

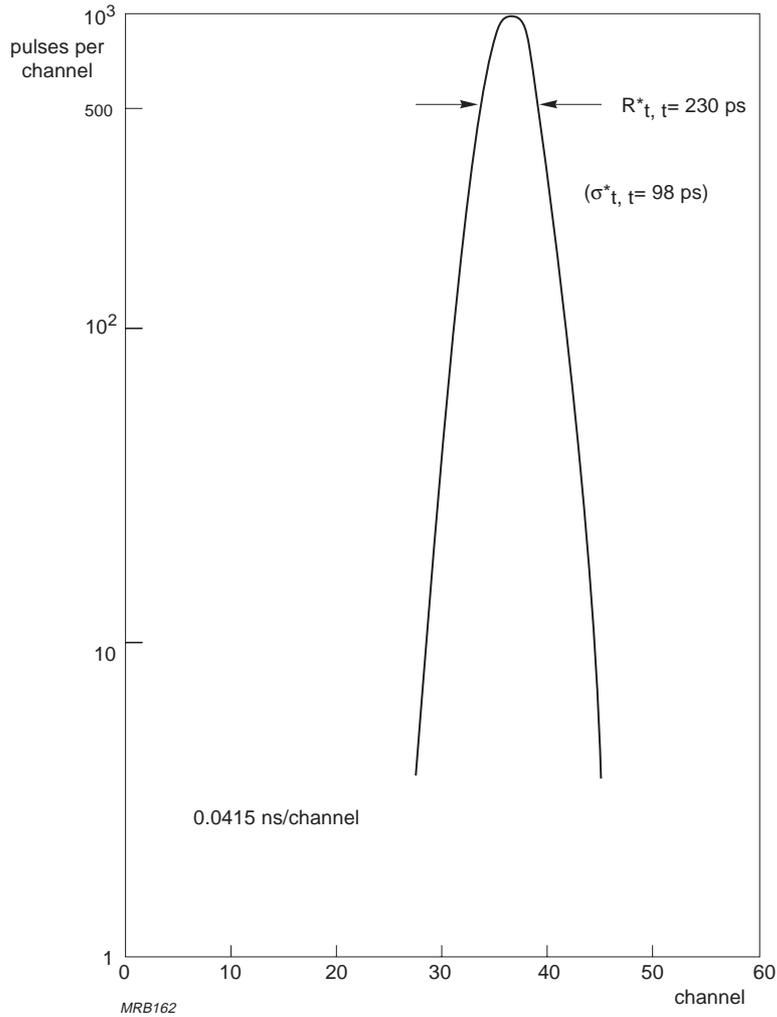


Fig.6.27 Coincidence time resolution distribution from two fast scintillation counters for ^{60}Co coincidence γ - γ radiation

The lowest values of τ for fast-decay plastic scintillators are a few nanoseconds, whereas the lowest values of σ_t for fast-response photomultipliers are only a few tenths of a nanosecond. The greater part of the response time variance is therefore due to the scintillator.

The spectrometer time resolution $R_{t,t}^*$ is defined as the FWHM of the observed response-time distribution,

$$R_{t,t}^* = 2.36 \sigma_{t,t}^*$$

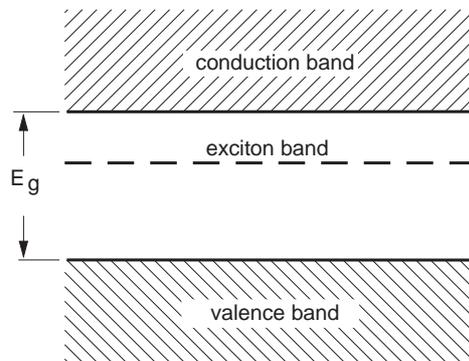
Since $\sigma_{t,t}^*$ varies inversely as the square root of the mean number of photoelectrons excited per scintillation, $\sqrt{\bar{n}_{k,s}}$, it is obvious that the choice of an efficient scintillator is very important in time spectrometry. Using very fast response photomultipliers and the fastest plastic scintillators (or BaF_2) for which $\tau < 3$ ns, values of $R_{t,t}^* < 200$ ps can be obtained.

APPENDIX 6 SCINTILLATOR FUNDAMENTALS

Scintillators are substances which convert some of the energy of incident radiation into detectable photons. They fall into two main classes, inorganic and organic, which differ not only chemically but also in the processes by which they release photons.

A6.1 Inorganic scintillators

Luminescence in inorganic scintillators is mainly a crystal property; it does not occur in liquids or gases, except in some inert gases in which it is an atomic or molecular property. Inorganic scintillators are crystalline compounds in which luminescence originates from emission centres in the compounds themselves or formed by activating agents introduced in controlled amounts (§A6.4).



MRB329

Fig.A6.1 Energy band diagram in an inorganic crystal

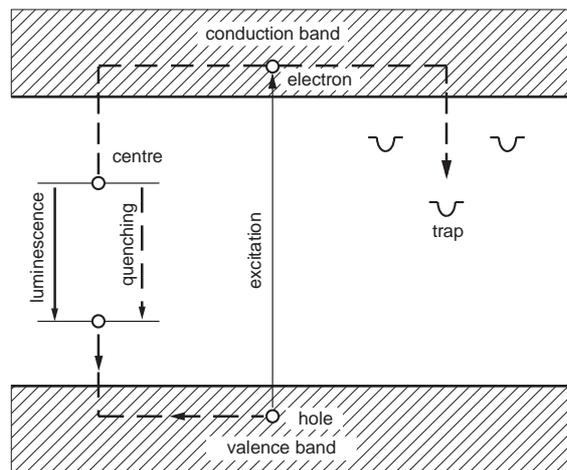
Scintillator crystals are usually insulators or semiconductors having an energy gap of a few electron volts between the valence and conduction bands (Fig.A6.1). Under the influence of ionizing radiation, electrons in the valence band can acquire sufficient energy to reach the conduction band, leaving behind charge deficiencies (positive 'holes') in the valence band. The crystal is then ionized and photoconductive. If the energy imparted to an electron is insufficient to raise it to the conduction band, it remains bound to a hole, at an energy level just below the conduction band. The electron-hole pair is called an 'exciton'.

This description applies only to a perfect crystal. Imperfections due to lattice dislocations or impurities in the crystal create additional energy levels in the energy gap into which excitons in the gap or electrons in the conduction band can move (Fig.A6.2). These levels correspond to activation centres of three types:

Luminescence centres, at which recombination of an electron-hole pair brings the centre to an excited state from which it returns to the ground state by photon emission; this is called fluorescence. The electron-hole recombination that excites the centre can be due either to capture or an exciton or to simultaneous capture of an electron from the conduction band and a hole from the valence band.

Quenching centres, which are like luminescence centres except that the excitation energy is dissipated as heat (phonons) instead of light.

Traps, which are metastable levels at which electrons and holes, or excitons, can remain for a long time before acquiring sufficient thermal energy to return to the conduction and valence bands or to move to a luminescence or quenching centre. When they do move to a luminescence centre the delayed emission that results is called phosphorescence.



MRB330

Fig.A6.2 Discrete levels within the forbidden band caused by crystal impurities

A6.2 Organic scintillators

Luminescence in organic substances is a molecular phenomenon which can occur in the solid, liquid or vapour phase as well as in liquid or solid solutions and plastic and glassy states.

Organic compounds form molecular crystals in which the occurrence of luminescence is associated with transitions between different energy levels of the electron systems: the singlet levels S_0, S_1, \dots, S_n ; and the triplet levels T_1, T_2, \dots, T_n (Fig.A6.3).

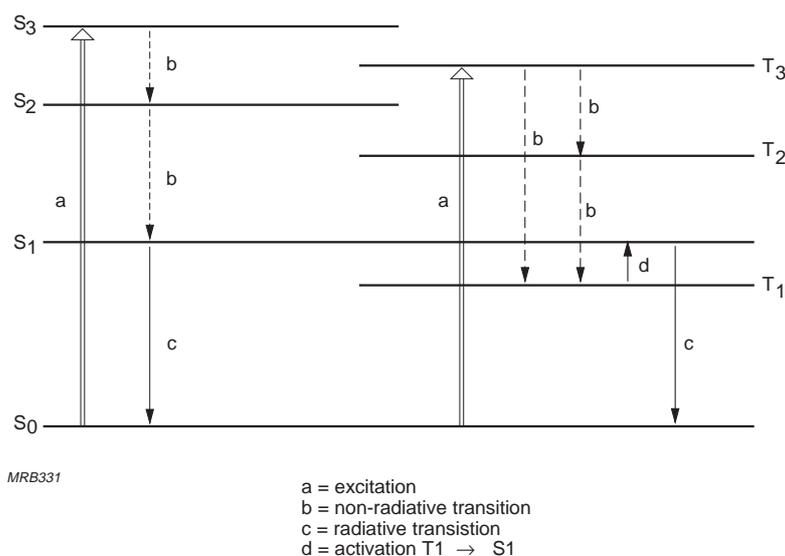


Fig.A6.3 In organic compounds, luminescence is associated with electron transitions between different energy levels (singlet and triplet states)

In the ground state the electrons are at the S_0 level. Incident radiation can either excite a molecule, raising its system of π electrons to a high singlet level (S_n), or ionize it, forming a free electron and a positive ion. Fast recombination of the electron and ion produces an excited molecule at a triplet level T_n or, less probably, a singlet level S_n . Non-radiative internal conversion phenomena then quickly return the excited levels to the lowest triplet or singlet levels, T_1 or S_1 ; these phenomena take about 10 to 100 ps.

The S_1 excited state, which lasts for about a nanosecond, can return to the ground state S_0 by either of two processes: thermal degradation (inhibition) or photon emission. The latter is called prompt fluorescence.

The T_1 excited state lasts much longer, and the molecule may return from it to the ground state by either of two routes. A direct return from T_1 to S_0 with photon emission is called phosphorescence. The emission is at a longer wavelength than fluorescence and the decay, which is exponential, is also longer – sometimes approaching a millisecond. Alternatively, interaction between two molecules in excited triplet states may raise one of them from the T_1 to the S_1 state, from which it can then return to the ground state; as before either by thermal degradation or by photon emission. In this case the emission is of the same wavelength as fluorescence but the decay is non-exponential.

A6.3 Scintillator characteristics

Scintillators are usually characterized in terms of scintillation effectiveness, scintillation efficiency, emission spectrum, integral quantum efficiency and response time.

Scintillation effectiveness ϵ_s is the ratio of the number of photons emitted per unit time, n_p , to the energy of the radiation absorbed in the scintillator, E_a ,

$$\epsilon_s = \frac{n_p}{E_a}$$

usually expressed in photons per MeV.

Scintillation efficiency η_s is the product of scintillation effectiveness ϵ_s and the energy $h\nu$ of the emitted photons,

$$\eta_s = \epsilon_s h\nu = \frac{n_p h\nu}{E_a}$$

usually expressed as a percentage.

An important consideration in energy spectrometry and a phenomenon that is often forgotten is scintillator linearity. This is a measure of the energy range over which scintillation effectiveness and scintillation efficiency are constant; this range is usually not too wide.

Emission spectrum is characterized by two parameters:

λ_m , the wavelength of maximum emission

$\Delta\lambda$, the FWHM of the emission band.

A useful measure of the spectral matching between a specific scintillator and a specific photomultiplier is the matching factor F_m :

$$F_m = \frac{\int_0^{\infty} R_{e,\lambda} \Phi_{e,\lambda} d\lambda}{\int_0^{\infty} \Phi_{e,\lambda} d\lambda}$$

where $\Phi_{e,\lambda}$ is the spectral distribution of the scintillator emission and $R_{e,\lambda}$ is the relative spectral sensitivity characteristic of the photomultiplier,

$$R_{e,\lambda} = \frac{S_{a,\lambda}}{S_m}$$

where $S_{a,\lambda}$ is as defined by Eq.2.9b and S_m is the sensitivity corresponding to the maximum of $S_{a,\lambda}$.

The matching factor F_m corresponds to the area ratio A_1/A_2 in Fig.A6.4. Table A6.1 lists the matching factors between some photocathodes and 2856 K tungsten light.

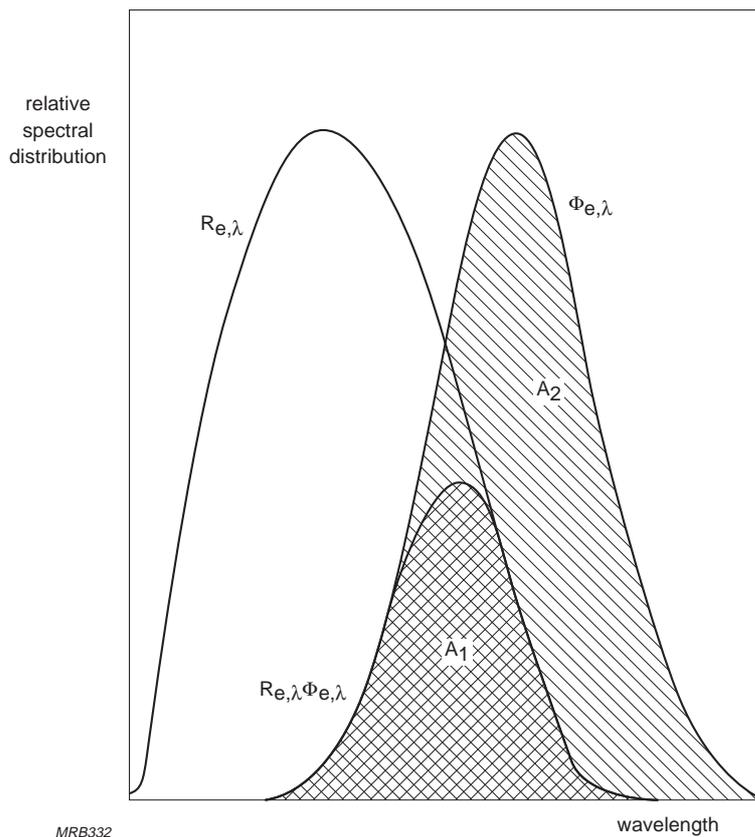


Fig.A6.4 Definition of matching factor $F_M = A_1/A_2$

Table A6.1 Matching factors between several photocathodes and 2856 K tungsten light

type of photocathode	S11 S13	bialkaline	S4	S20	S20R	S1
matching factor F_m	0.028	0.018	0.021	0.056	0.126	0.251

Integral Quantum Efficiency, IQE is defined for a non-monochromatic radiation as the ratio of the average number of emitted photoelectrons to the average number of received photons whatever their energy is.

$$\text{IQE} = \frac{hc}{e} \frac{\int_0^{\infty} S_{a,\lambda} \Phi_{e,\lambda} d\lambda}{\int_0^{\infty} \Phi_{e,\lambda} \lambda d\lambda}$$

with $e = 1.6 \times 10^{-19} \text{ C}$
 $h = \text{Planck's constant}$
 $c = \text{speed of light in vacuum.}$

For the wavelength shifters BBQ, Y7 and K27, typical IQE values are:

Type of photocathode	IQE
monoalkaline S11 SbKCs ₃	10-15%
bialkaline SbKCs	10-15%
green-extended bialkaline (GEBA) SbKCs	12-18%
trialkaline S20 SbNa ₂ KCs	15-22%

Response time is a measure of the width of the scintillator pulse, which is usually much greater than the stopping time of the ionizing particle or photon that causes it. After rising abruptly, the pulse falls according to a composite law in which at least two components can usually be distinguished:

- a fast exponential decay (fluorescence) with a time constant τ
- a slower decay (phosphorescence, delayed fluorescence) following a more complex law.

The time constant τ is called the response of the scintillator.

A6.4 Scintillator properties

A6.4.1 Inorganic-scintillator properties

Except for the inert gases and some specially formulated glasses, inorganic scintillators are either monocrystals or polycrystalline powders. The monocrystals are transparent to their own luminescence, the polycrystalline powers are less so.

The principle monocrystalline scintillators are alkali halides, mostly doped with heavy metals such as thallium or europium: NaI(Tl), CsI(Tl), CsI(Na), KI(Tl), CaI₂(Eu), and CaF₂(Eu). Crystals of pure (undoped) alkali halides like NaI and CsI are also very useful.

Other, more recently adopted monocrystalline scintillators are CaWO_4 , CdWO_4 , PbWO_4 , ZnWO_4 , BaF_2 , CsF , CeF_3 , $\text{Bi}_4\text{Ge}_3\text{O}_{12}$ (colloquially, BGO), and the Ce-doped $\text{Gd}_2\text{SiO}_5:\text{Ce}$ (gadolinium orthosilicate or GSO), Lu_2SiO_5 (lutetium orthosilicate or LSO), $\text{Y}_2\text{SiO}_5:\text{Ce}$ (yttrium orthosilicate or YSO) and $\text{YAlO}_3:\text{Ce}$ (yttrium aluminium perovskite or YAP).

Table A6.2 Properties of some inorganic scintillators

scintillator composition	density (g/cm^3)	index of refraction	wavelength of maximum emission (nm)	decay time constant 1/e (ns)	scintillation pulse height ¹⁾	notes
NaI	3.67	1.78	303	60	190	2)
NaI(Tl)	3.67	1.85	410	250	100	3)
CsI	4.51	1.95	310	10	6	3)
CsI(Tl)	4.51	1.79	565	1000	45	3)
CsI(Na)	4.51	1.84	420	630	85	3)
⁶ LiI(Eu)	4.06	1.96	470-485	1400	35	3)
$\text{CaF}_2(\text{Eu})$	3.19	1.44	435	900	50	
BaF_2	4.88	1.49	190/220 310	0.6 630	5 15	
BGO	7.13	2.15	480	300	10	
ZnWO_4	7.87	2.2	480	5000	26	
CdWO_4	7.90	2.3	540	5000	40	
PbWO_4	8.28	2.16	480	2/7/26	0.8	
CsF	4.65	1.48	390	5	5	3)
CeF_3	6.16	1.68	300 340	5 20	5	
LSO	7.40	1.81	420	42	75	
GSO	6.71	1.9	440	60	20	
YSO	4.45	1.8	420	35	50	
YAP	5.50	1.9	370	30	40	
$\text{ZnS}(\text{Ag})$	4.09	2.35	450	200	150	4)
$\text{ZnO}(\text{Ga})$	5.61	2.02	385	0.4	40	4)

¹⁾ relative to NaI(Tl) ²⁾ at 80 K ³⁾ hygroscopic ⁴⁾ polycrystalline

Polycrystalline powders include the doped sulphides $\text{ZnS}(\text{Cu})$, $\text{ZnS}(\text{Ag})$ and $\text{CdS}(\text{Ag})$, $\text{ZnO}(\text{Ga})$, and the oxide ZnO . NaI(Tl) is also sometimes used as a polycrystalline powder. Being partly opaque to their own luminescence, polycrystalline powders are

mostly used as thin screens which detect only heavy and highly ionizing particles with good efficiency.

Table A6.2 summarizes the properties of some principal inorganic scintillators. In general they have the following properties in common:

- high absorption for X- and γ -photons and nuclear particles such as α and β
- good matching with the spectral sensitivity of general-purpose photomultipliers
- practically linear scintillation effectiveness over a wide energy range
- decay time constants mostly in the region of a microsecond or less.

The list of inorganic scintillators grows yearly, and investigators worldwide are currently devoting a lot of effort to the search for new heavy, non-hygroscopic scintillator materials that are also inexpensive to produce. Interest is presently centred on the fluorides (e.g. ThF_4 , BaLiF_3 , LiYbF_4 , BaYb_2F_4 and PbF_2). Applications for new materials will increase in the coming years, notably in high-energy physics experiments. For the CERN CMS experiment at the Large Hadron Collider (LHC), considerable effort in cooperation with Russian and Chinese centres have gone into the development of the new PbWO_4 scintillator. In nuclear medical imaging, especially for the strongly growing market of PET scanners (Chapter 7), considerable effort is being devoted to finding heavier and faster inorganic scintillators than the fairly new and currently preferred LSO which is gradually replacing the traditional BGO.

Radiation absorption. Because of their high density (3 to 4 g/cm^3) and the high atomic number of iodine ($Z = 53$), monocrystals of the alkali iodides are particularly well suited for detecting X- and γ -photons; see Fig.A6.5. They are also highly absorbent for electrons and heavy charged particles (α , protons, etc.); see Fig.A6.6 and A6.7.

The large neutron cross-section of ${}^6\text{Li}$ makes ${}^6\text{LiI}(\text{Eu})$ an efficient neutron detector. Mixtures of ZnS with ${}^6\text{Li}$, ${}^{10}\text{B}$, ${}^{235}\text{U}$, or even methyl methacrylate, make excellent detectors for thermal neutrons.

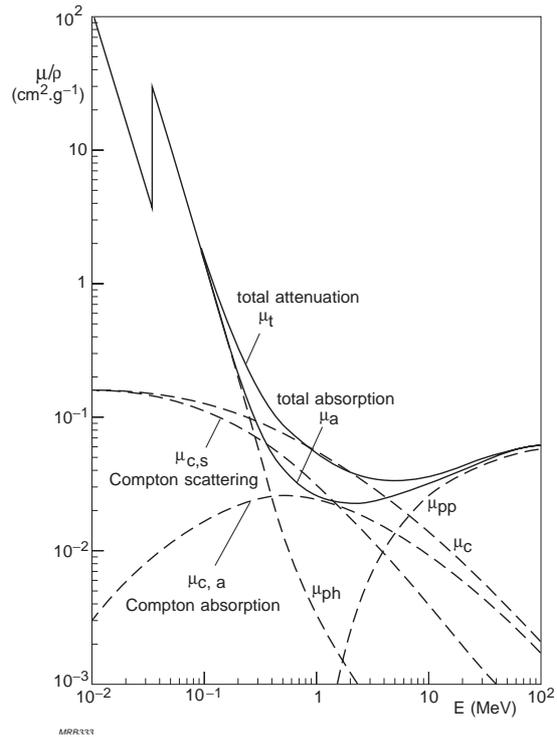


Fig.A6.5 Mass attenuation and mass absorption coefficients for X- and γ -radiation in NaI(Tl), $\rho = 3,67 \text{ g/cm}^3$

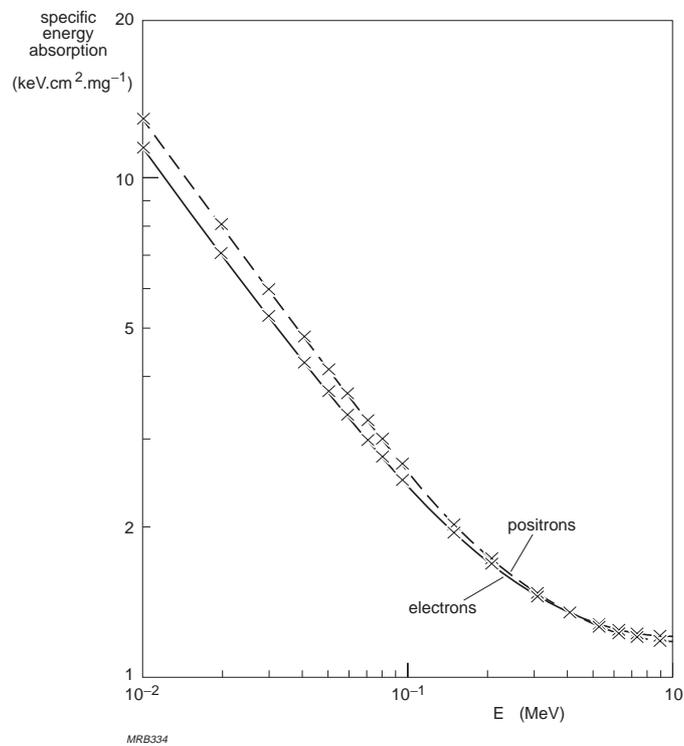


Fig.A6.6 Specific energy absorption for electrons and positrons in pure NaI as a function of their energy (*U.S. National of Bureau Standards, circular number 577, 1956*)

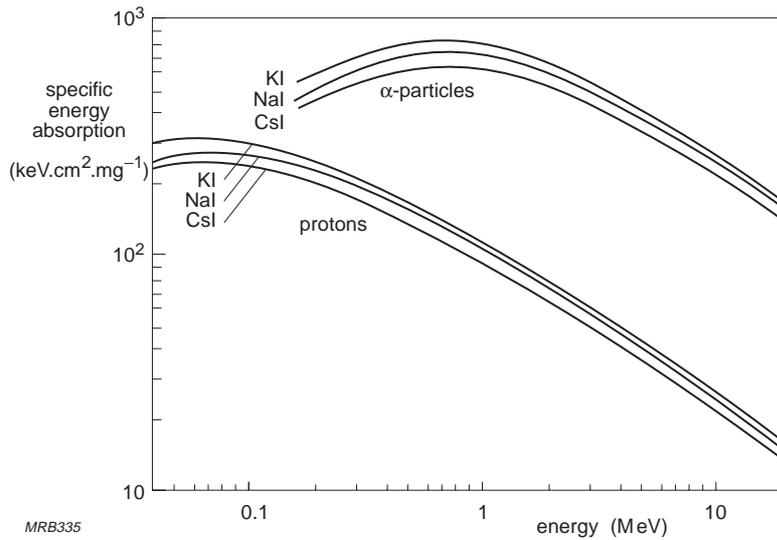


Fig.A6.7 Specific energy absorption for ionizing α -particles and protons in pure KI, NaI and CsI scintillators

Emission spectrum. Inorganic scintillators usually have two emission bands: one due to the activator and another, at shorter wavelength, due to the crystal lattice (see entries for doped and undoped NaI and CsI in Table A6.2). The relative intensity of the emission due to the crystal lattice decreases as the concentration of the activator increases; at high enough concentrations (e.g. 0.1% to 0.5% Tl in NaI) only the emission due to the activator is practically significant. Figure A6.8 compares the emission spectra of Tl-doped NaI ($\lambda_m = 410$ nm) and the undoped scintillator $\text{Bi}_4\text{Ge}_3\text{O}_{12}$ ($\lambda_m = 480$ nm).

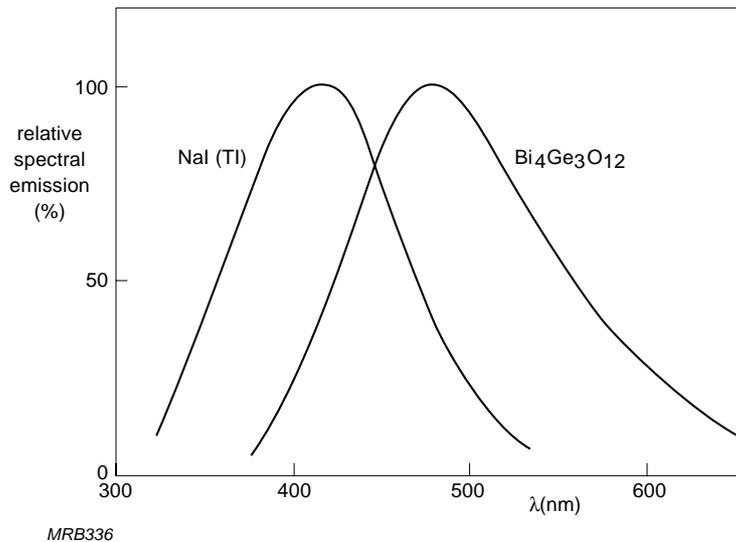


Fig.A6.8 Normalized emission spectra of NaI(Tl) and $\text{Bi}_4\text{Ge}_3\text{O}_{12}$

Scintillation efficiency. The efficiency of most scintillators used in radiation spectrometry can be taken as constant over a wide range; however, a 15% to 20% increase of efficiency at the low-energy end of the range is often observed.

Of the monocrystalline scintillators, the efficiencies of NaI and CsI are among the highest known. NaI(Tl) has an efficiency of about 13%, emitting four or five photons or 3 eV average energy per 100 eV of absorbed energy.

ZnS has higher efficiency than any monocrystalline scintillator. ZnS(Ag) and ZnS(Cu) have efficiencies of about 25% and 30% for α -particles, regardless of the energy of the particles, and about half that for β -particles.

The efficiency of most scintillators as particle detectors varies according to the type of particle being detected. NaI(Tl) is most efficient for protons and deuterons and progressively less so for electrons, α -particles, and heavy ions.

Response time. The shape of the scintillation pulse depends on factors such as the type of crystal, the nature and concentration of the activator, the type of radiation detected, and the temperature. In general, the rise time is very fast: a few tenths of a nanosecond, and the fall time can be resolved into two components (Fig.A6.9):

- a prompt, exponential component with a time constant from a few hundred nanoseconds to a few microseconds
- a delayed component lasting several microseconds.

In NaI(Tl) the prompt component has a decay time constant of about 350 ns at very low concentrations of Tl, decreasing to 230 ns at normal concentrations (0.1% to 0.5%).

The fact that the delayed component may account for as much as 20% to 40% of the total emission of inorganic scintillators sets a limit to their usefulness in high-count-rate applications.

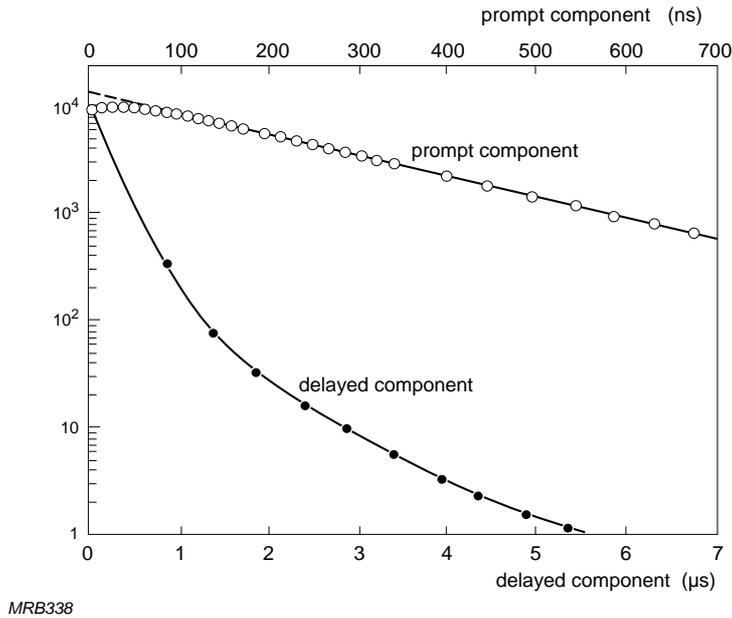


Fig.A6.9 Decay time components of a NaI(Tl) scintillator for γ -radiation
(*Rev. Sci. Instr.* 32 (1952) 1044)

Temperature effects. Temperature influences scintillation effectiveness, pulse decay time and emission spectrum. Figure A6.10 shows the scintillation effectiveness of CsI(Tl), CsI(Na) and NaI(Tl) as functions of temperature; note that the effectiveness of NaI(Tl) has a conveniently broad maximum around 20 °C. For use at very low temperatures special fast photomultipliers are now feasible despite the increasing resistance of the photocathode layer.

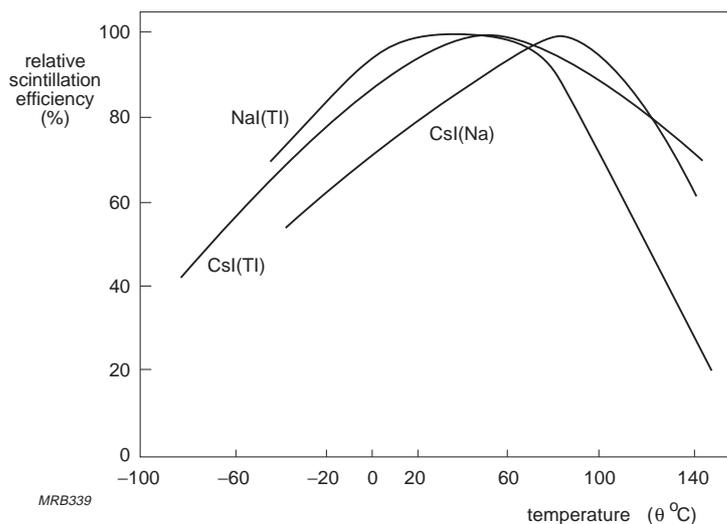


Fig.A6.10 Scintillation efficiency as a function of temperature for several inorganic scintillators (*Harshaw Chemical Company*)

Hygroscopy. NaI(Tl), CsI(Tl), CsI(Na), CsF, KI(Tl) and $^6\text{LiI}(\text{Eu})$ are all hygroscopic and have to be hermetically encased. BGO and BaF_2 , which are not hygroscopic, have high absorption coefficients for X- and γ -photons and very low delayed fluorescence. In some applications, such as medical PET scanners (Chapter 7), these properties compensate for the relatively low scintillation effectiveness of these materials. The new non-hygroscopic LSO scintillator, with a much shorter decay time constant than BGO and with similar stopping power, shows considerable improvements and therefore is the recent scintillator of choice for PET scanners. Other lutetium-based scintillators are under investigation worldwide.

A6.4.2 Organic-scintillator properties

Most organic scintillators are either monocrystalline compounds or liquid or plastic solutions.

Among the most efficient monocrystalline compounds are naphthalene, anthracene, fluoranthane, p-terphenyl (PT), p-quaterphenyl (PQ), and trans-stilbene.

Liquid scintillators are usually binary or ternary systems consisting of a solvent and small concentrations of one or two solutes that alter the emission wavelength. The most widely used solvents are alkyl benzenes such as xylene, benzene, toluene, etc. Common primary solutes are p-terphenyl, DPO, and PBD; common secondary solutes are POPOP, BBO, and BPO.

Chemically, plastic scintillators are very similar to liquid ones. Just as the alkyl benzenes are the most effective liquid solvents, their polymers, such as polyvinylbenzene, polyvinyltoluene, etc., are the most effective plastic ones. PT, DPO, and PBD are useful primary solutes for both liquid and plastic scintillators; common secondary solutes include BBO, TBP, POPOP, and DPS. An advantage of plastic scintillators is the ease with which they can be formed into a variety of shapes and sizes: the large scintillator sheets used in high-energy physics experiments are an example.

Table A6.3 compares the properties of some organic scintillators; in general, they have in common:

- high absorption for electrons and fast neutrons, but low γ -detection efficiency
- practically linear scintillation effectiveness over a wide energy range
- linear scintillation response for relativistic particles
- faster response than most inorganic scintillators.

Radiation absorption. Because of their comparatively low density (1 to 1.25 g/cm³) and the low atomic numbers of their constituent elements, organic scintillators are less absorbent for charged particles and X- and γ -photons than inorganic ones (Fig.A6.11). X- and γ -absorption, in particular, is nearly all due to Compton effect and can be increased by introducing a small concentration of a heavy-metal (e.g. Pb, Sn) organic salt.

Figure A6.12 shows the calculated path lengths of four types of particle in an organic scintillator. As the governing parameters, density and H/C ratio, are fairly similar in most organic scintillators, the path lengths do not differ much from one to another.

Because of their hydrogen content, organic scintillators are considerably more absorbent than inorganic ones for fast neutrons and are thus also used as neutron moderators.

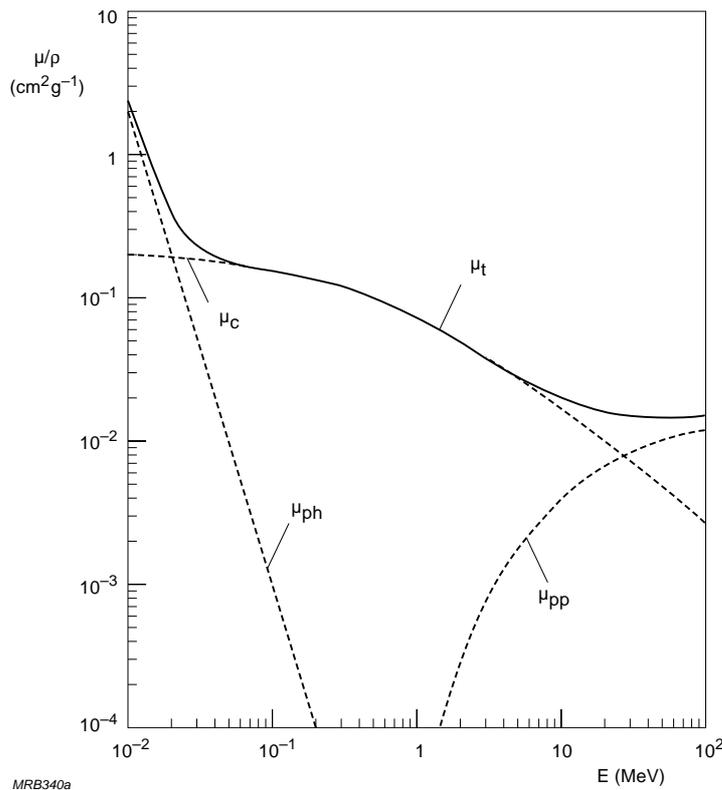


Fig.A6.11 Mass absorption coefficient for X- and γ -radiation in an organic anthracene C₁₄H₁₀ scintillator

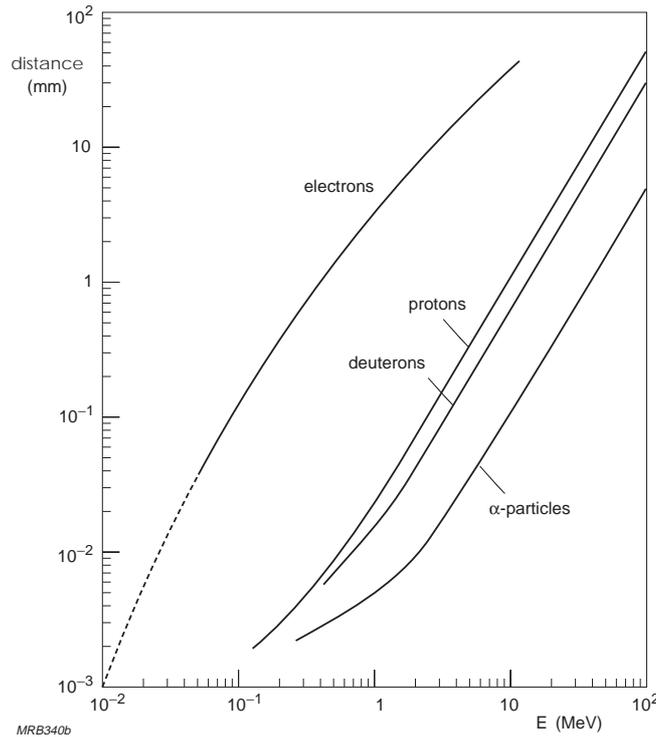


Fig.A6.12 Calculated pathlength of four types of charged particle in an organic scintillator (*Nuclear Enterprises Ltd.*)

Emission spectrum. The useful emission of most organic scintillators peaks at about 400 nm. Emission at shorter wavelengths is largely suppressed by auto-absorption due to overlap of the emission and absorption spectra (a wavelength shifting process with low Stoke's shift).

The emission of binary (single solute) liquid and plastic scintillators peaks at wavelengths below 400 nm. A small concentration of a secondary solute will shift the peak to between 400 nm and 500 nm where it is better matched to the spectral sensitivity of most photomultipliers.

The observable emission spectrum of plastic scintillators depends not only on the properties of the constituents but also on where the emission originates. Emission excited within about a micron of the output surface has a different spectral distribution from that originating in the interior of a (long) scintillator.

Transparency. Especially in large scintillators, transparency is of critical importance. Two factors that play an important part in it are selective absorption and multiple internal reflections; light intensity does not usually decrease exponentially with distance as Lambert's law would suggest. Therefore it is customary to characterize

the transparency of an organic scintillator by its half-thickness, the distance in which light intensity is halved. General information of organic scintillators can be found in manufacturers' catalogues.

Wavelength shifters (WLS) like BBQ, Y7, K27 etc. are plastics in which small concentrations of a secondary solute absorbs short-wavelength radiation, e.g. radiation peaking at around 400 nm, and re-emits it at a peak of around 500 to 550 nm. These materials are characterised by:

- a large Stoke's shift
- minimal overlap of absorption and emission spectra.

Here too (as with inorganic scintillators), an intensive search for new, superior materials is currently going on, driven by the needs of the physics community.

Sometimes, wavelength shifters are formed by evaporating organic materials onto the photomultiplier window. In this way, limeglass- or borosilicate-window tubes can be used to detect UV radiation (§6.1.3 and §8.3.1).

Table A6.3 Properties of some organic scintillators

scintillator	density (g/cm ³)	index of refraction	wavelength of maximum emission (nm)	decay time constant 1/e (ns)	scintillation pulse height ¹⁾	H/C ratio ²⁾
Monocrystals						
naphthalene	1.15	1.58	348	11	11	0.800
anthracene	1.25	1.59	448	30-32	100	0.714
trans-stilbene	1.16	1.58	384	3-8	46	0.857
p-terphenyl	1.23		391	6-12	30	0.778
Plastics ³⁾						
BC-400	1.032	1.581	423	2.4	65	1.103
BC-404	1.032	1.58	408	1.8	68	1.107
BC-408	1.032	1.58	425	2.1	64	1.104
BC-412	1.032	1.58	434	3.3	60	1.104
BC-414	1.032	1.58	392	1.8	68	1.110
BC-416	1.032	1.58	434	4.0	38	1.110
BC-418 (pilot-U)	1.032	1.58	391	1.4	67	1.100
BC-420	1.032	1.58	391	1.5	64	1.100
BC-422	1.032	1.58	370	1.6	55	1.102
BC-422Q	1.032	1.58	370	0.7	11	1.102
BC-428	1.032	1.58	480	12.5	36	1.103
BC-430	1.032	1.58	580	16.8	45	1.108
BC-434	1.049	1.58	425	2.2	60	0.995
BC444	1.032	1.58	428	285	41	1.109
BC470	1.037	1.58	423	2.4	46	1.098

¹⁾ relative to anthracene

²⁾ ratio of hydrogen to carbon atoms

³⁾ Saint Gobain Crystals & Detectors (Bicron), Newbury, Ohio, USA

CHAPTER 7

APPLICATIONS OF SCINTILLATION COUNTING AND ANALYSIS

7.1 Nuclear medicine

7.2 Analytical applications

7.2.1 Liquid scintillation counting

7.2.2 Radio-immuno assay

7.3 Industrial applications

7.3.1 Non-destructive analysis

7.3.2 Thickness and density measurement

7.3.3 Oil-well logging

7.4 High-energy physics applications

7.4.1 Collider detectors

7.4.2 Fixed-target detectors

7.5 Cosmic-ray detectors

APPLICATIONS OF SCINTILLATION COUNTING AND ANALYSIS

7.1 Nuclear medicine

The information sought in most medical applications of scintillation detection is a clinically significant internal image of the body. Here we will treat only the imaging modality using photomultiplier tubes, i.e. nuclear medicine.

The radioactively labelled substances administered are chosen for their affinity to the specific organs or tissues to be studied. From the local radiation intensities registered by scintillation detectors it is then possible to construct images of the organs or tissues under investigation. In cardiac studies, for example, labelled substances injected into a vein enable dynamic images of the heart action to be obtained. A growing pharmacopoeia of radioactive tracers is steadily increasing the number of organs, tissues and functions that are accessible to this type of investigation.

An imaging system for nuclear medicine consists mainly of

- a collimator to select radiation emitted at a specific angle
- either one large scintillator covered by many large photomultipliers or many small scintillators read out by several smaller photomultipliers
- preamplifiers, pulse-shaping and analysing electronics, and discriminators
- a computer for constructing an image based on the information derived from the detected radiation.

Within this general framework there are numerous variations.

γ -cameras. The γ -camera invented by Anger offers dynamic information and has a fields of view up to 45 cm in diameter or 40 cm \times 60 cm in rectangular format. In its usual form (Fig.7.1) it consists of a hexagonally or square close-packed array of photomultipliers coupled via a light guide (Fig.7.3) to a single, large scintillator. Radiation reaches the scintillator through a collimator of radiopaque material (e.g. lead, tungsten) with a dense, sieve-like array of parallel holes.

Apart from parallel-hole collimators, several other types exist, e.g. diverging and converging hole collimators to enlarge the field of view or increase the resolution at a specific depths within the body, slant-hole collimators to enable oblique imaging, and line-focused collimators for whole-body scanning.

Since all the photomultipliers view the scintillator, they all respond more or less to each scintillation, their signals being a function of their distance from the scintillation point and the total amount of light emitted by the scintillator, which is a function of

the energy of the γ -quanta. The coordinates of the scintillation point are calculated from summation signals obtained by passing the photomultiplier signals through a summation network incorporating weighting factors (resistors or capacitors). These weighting factors are chosen such that the summation signals either increase or decrease linearly with the x or y coordinates. As the photomultiplier signals are energy dependent, so too are the summation signals and hence the scintillation coordinates. To remove this energy dependence of the calculated coordinates the following algorithm is often used:

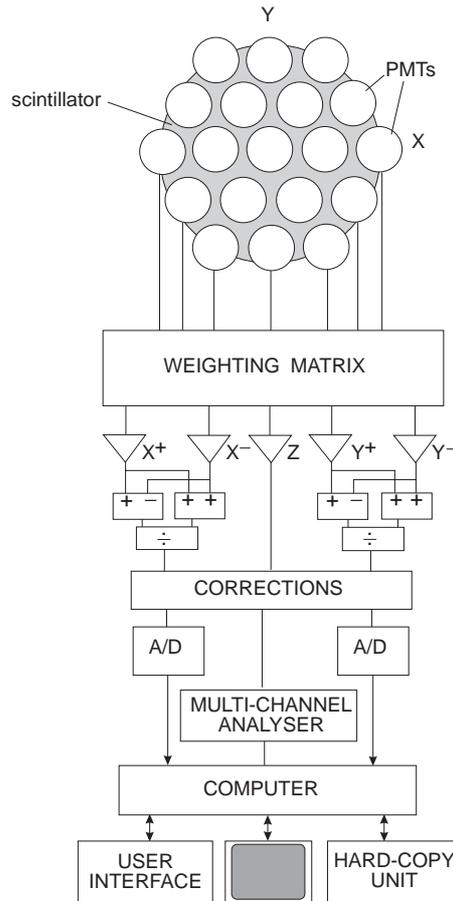
$$x = \frac{X^+ - X^-}{X^+ + X^-}$$

in which X = summation signals. The y coordinate is similarly calculated from the Y⁺ and Y⁻ summation signals.

Besides the summation signals defining position, there is also a fifth summation signal E or Z – the energy summation signal. Its function is to exclude (discriminate) undefined coordinate information caused by scattered radiation. This is necessary because the scattering process not only affects the direction but also lowers the energy of the γ -quanta.

The weighting factors for the energy summation signal are chosen such that the signal is virtually independent of the position of the scintillation. The energy signal is analysed by a multichannel analyser (MCA) in which, by appropriate choice of energy window, only so called *photo-peak events* are accepted while signals with other energies are excluded. This results in better contrast. The image is obtained by acquiring tens of thousands to millions of accepted events (counts).

Originally, the image was displayed on a cathode ray tube (CRT) with a long-persistence phosphor for direct viewing or integrated with a polaroid film cassette for recording. Nowadays, all γ -cameras are integrated with a computer. The position signals are digitized and the images are stored in memory and displayed via a monitor or sent to a hardcopy unit. The computer can be used for several forms of image processing and reconstruction (whole body, tomography). Its use is increasingly being extended toward providing total system control and user interfacing (Fig.7.1).



MRB342

Fig.7.1 Operating principle of a γ -camera

The most important demands on photomultipliers for use in γ -cameras are

- uniform azimuthal response
- excellent pulse-amplitude resolution over a large dynamic range. γ -energies vary between 50 keV and 400 keV but the dynamic range is much larger due to the signal variation caused by the variation in distance to the scintillation point (by factors of between 30 and 100).
- very stable gain as a function of time, count rate and magnetic field.

The energy resolution of a photomultiplier varies inversely as the square root of the signal amplitude, which decreases as the distance of the scintillation from the photomultiplier axis increases. This decrease is commonly described by fall-off curves in which the photomultiplier response is plotted as a function of the off-axis distance (in any direction) of the scintillation. By designing the front end of the tube so as to favour the collection of distant light, it is possible to broaden the fall-off curve somewhat, and thereby improve the energy resolution of the tube for distant scintillations, improve the uniformity and linearity of the picture and, most important,

improve the camera's spatial resolution. Also, the use of hexagonal or square tubes, which can be close-packed with less dead space between them or with the dead space more uniformly distributed, makes it possible to obtain a larger sum signal and thus improve the energy resolution of the entire camera.

The interrelated aspects of azimuth response and tube energy resolution with picture uniformity, linearity and spatial resolution are the most important parameters for characterizing photomultipliers for γ -camera applications. The closer the scintillations are to the photocathode, the more critical these parameters are. They are less critical in cameras with thick scintillators and thick light guides between the scintillator and the photomultipliers, so here the use of hexagonal tubes may bring only minor advantages. Hence, different camera designs call for differently optimized tubes.

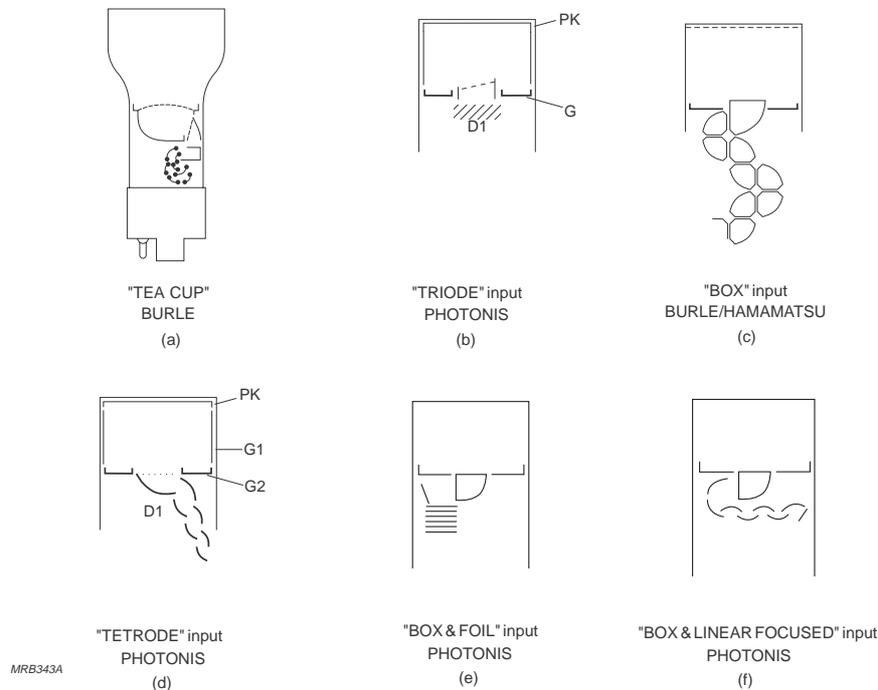


Fig.7.2 Examples of γ -camera tube input systems

The quest for improved azimuthal response and tube resolution has led to several developments in electron-optical input systems over the years. Amongst these are the 'teacup' first dynode of Burle, the asymmetric 'triode' input to venetian-blind dynodes and the 'tetrode' input to linear focusing dynodes of Photonis. Others include the 'box & grid' multiplier of Burle and Hamamatsu and the 'box & linear focusing' and 'box & foil' multiplier systems of Photonis (Fig.7.2). These newer multipliers have also enabled tube manufacturers to make much shorter tubes, resulting in thinner and, for the patient, less threatening γ -camera heads whilst simplifying the mechanical counterweights needed for rotating cameras. At the same

time, the quest for better energy resolution has led to improvements in cathode sensitivity, collection efficiency, and first-dynode multiplication statistics. And with present-day tubes, gain drifts can be kept well below 1% per month under the conditions prevailing in γ -cameras.

In modern γ -cameras, a whole palette of tricks and corrections is used to optimise the images. Important parameters are:

- spatial resolution
- uniformity
- energy resolution
- linearity (distortion)
- speed/temporal resolution
- stability (time, temperature, position).

Some of these properties result in conflicting parameter settings for the camera (e.g. resolution versus uniformity). Most of them, however, are more or less interdependent.

The shape of the response curve (fall-off versus distance) plays an important role in spatial resolution, linearity and uniformity. The shape is influenced by several factors (e.g. crystal and light-guide thickness and refractive index, shape and size of photomultipliers). Furthermore, it can be influenced by optical tricks (redistribution of light) and electronically (non-linear amplifier characteristics, threshold). It is also possible to introduce special processing for the signals contributing to the energy summation signals (to reduce noise).

During the last fifteen years, the γ -camera has seen many impressive improvements including the introduction of digital correction systems for controlling linearity, uniformity, energy uniformity and photomultiplier stabilization (off-line and on-line).

Traditionally it was common for cameras to contain 37, 61 or 91 circular tubes in a hexagonal array. Recently, the most popular and lowest-price concepts that still offer good spatial resolution make use of 55 circular tubes of 76 mm diameter with 6 elongated-hexagonal 40 mm tubes along the sides of the rectangular array, or simply 48 square 76 mm tubes forming a rectangular array. Large rectangular cameras have replaced rectilinear and hybrid scanners for whole-body scanning. Small-field-of-view cameras with thirty-seven 1.5 inch or 2 inch tubes are used for cardiac studies (sometimes as mobile cameras), an application in which gain stability as a function of anode current is critically important: count rates commonly go from background to as much as half a million for a few seconds and then down to background again.

For the tubes nearest the patient's heart, the variation in anode current is considerable. Here, a low gain hysteresis is also important.

A major application of γ -cameras is single-photon emission computed tomography (SPECT, Fig.7.3), in which cross-sectional images of the body are generated by one, two or even three rotating camera heads around the patient. This imposes a further demand on the photomultipliers used: insensitivity to variations of ambient magnetic field. The constructions shown in Fig.7.2 have proved themselves able to meet this requirement fully and at the same time provide good azimuthal response.

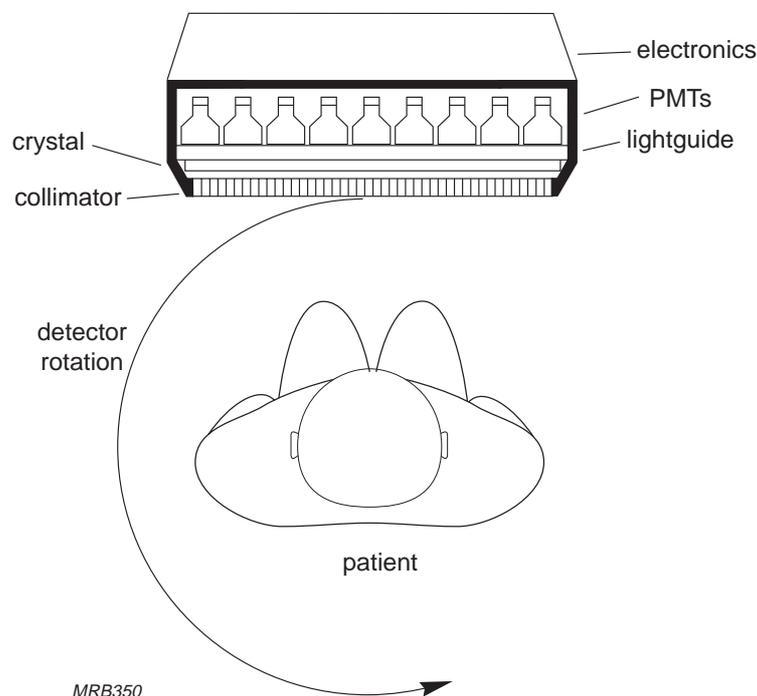


Fig.7.3 Principle of single-photon emission computed tomography (SPECT). The γ -camera head (sometimes 2 or 3) is moved around the patient and acquires projections from several angles. These are back-projected by computer to give transversal slices

Because of their inherently superior energy resolution, semiconductor radiation detectors appear to promise advantages over the scintillator and photomultiplier combination; theoretically at least, a γ -camera using high-purity germanium detectors should provide superior contrast and spatial resolution. So far, however, fulfilment of this promise has not been economically feasible. The low efficiencies of prototype systems, together with the technological difficulty of manufacturing large arrays of sufficiently uniform detectors, has discouraged serious ventures into production.

A further drawback of germanium is that it has to be cooled to liquid-nitrogen temperature. Two candidates for room-temperature detectors, CdTe and HgI₂, have been under investigation for many years but have yet to yield commercially usable results. Again, the stumbling blocks have been material quality and detector size. Detector polarization under radiation has also presented problems that have not yet been fully solved. At the moment, it looks as if semiconductor detectors have little chance of supplanting scintillators and photomultipliers in large field-of-view γ -cameras within the next ten years. However, new CdZnTe (CZT) room-temperature solid-state detectors may slowly become feasible for small, dedicated cardiac and thyroid γ -cameras with better spatial resolution. This will, however, require a significant reduction in the cost of growing and selecting useful (high-quality) single-crystal pieces out of a multi-crystal ingot.

Another idea that has been explored (and abandoned) is to combine a scintillator with a large image intensifier observed by a number of photomultipliers or other photo-sensors. Here, problems with the image intensifier (especially its size) have been the principal barrier to continued progress.

Experiments with gas-multiplication wire chambers have led to useful prototype cameras only for low γ -energies, where the efficiency of such a system is acceptable. Trials with the fast UV-emitting BaF₂ scintillator have recently raised the efficiency of such systems to useful levels. However, the UV-readout process involving a 'liquid' TMAE photocathode (not very stable over time) has so far been tried only on a laboratory scale, and new liquid photocathodes with greater overlap with the 190/225 nm fast BaF₂ peak have yet to be found.

Many ideas for replacing the photomultiplier by silicon or avalanche photo-diodes fall down on the basis of surface area, performance, price of the diode and the high number of electronic channels needed. Recently, however, a small camera has been commercialised based on an array of small CsI(Tl) scintillators plus silicon photodiodes for the niche market of local cardiac doctors.

Nevertheless, since large field-of-view γ -camera applications form a backbone of the world photomultiplier market, it seems likely that the scintillator-photomultiplier partnership will survive for many years yet, and that γ -camera developments, such as adding depth-of-interaction (DOI) corrections, will continue to motivate photomultiplier improvements as they have in the past.

Positron scanners. In positron emission tomography (PET), Fig.7.4, a positron-emitting isotope is administered to the patient and the body region under investigation

is surrounded by rings of small scintillators and photomultipliers. A computer constructs cross-sectional metabolic images based on detection of the two coincident 511 keV γ -photons which are emitted 180° apart when a positron annihilates with an electron in the patient's body. Owing to the very short half-lives of the four commonly used isotopes: ^{11}C (20 min), ^{13}N (10 min), ^{15}O (2 min) and ^{18}F (110 min), they must be produced and processed into labelled compounds on site. This calls for a dedicated cyclotron and radiochemistry facilities and for many years this has limited the use of PET to large, well-funded research hospitals. Scanners for measuring the blood flow in the coronary arteries were therefore the first to go into clinical use in hospitals. These make use of the long-lived isotope ^{82}Rb milked from ^{82}Sr (produced commercially by several accelerator centres around the world), eliminating the need for a cyclotron on site.

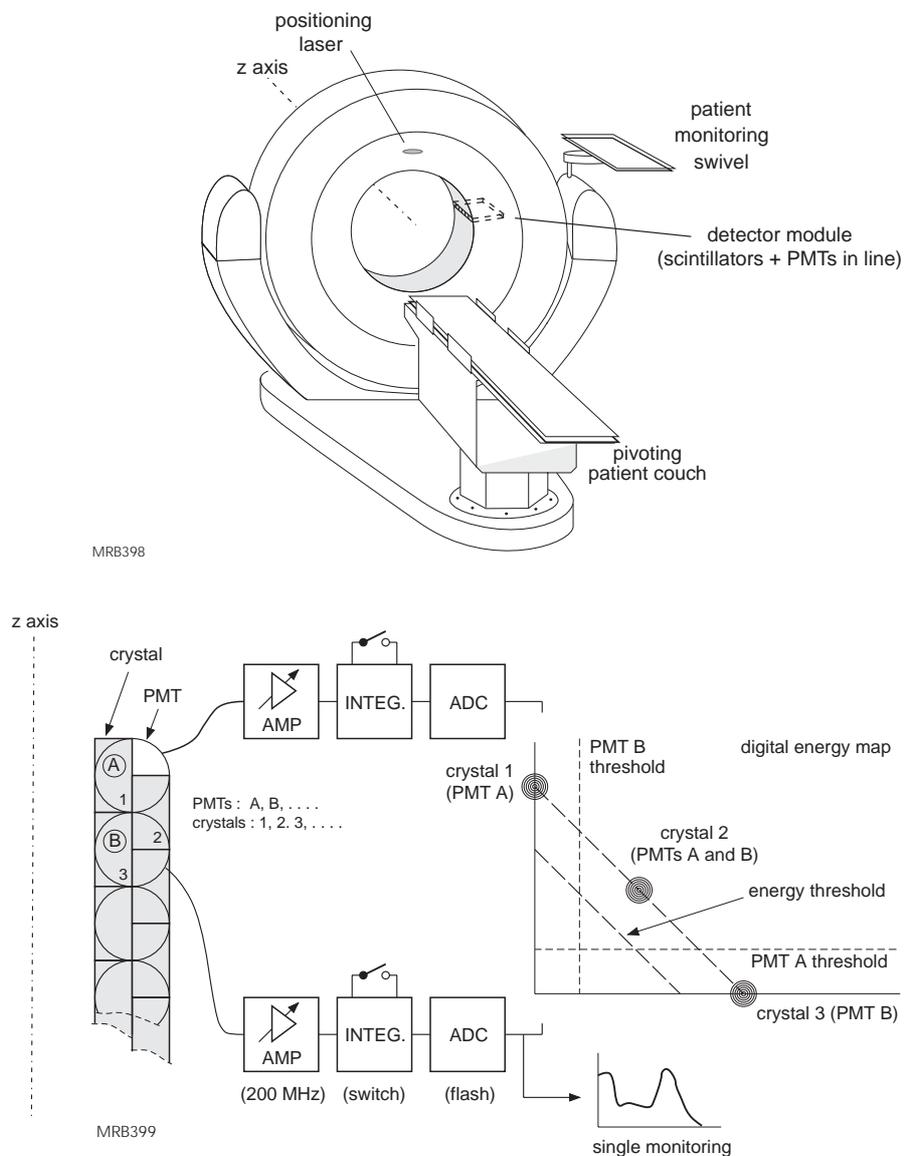


Fig.7.4 Principle of the PET scanner (courtesy of Positron Corporation, Houston, Texas). Light output from crystal 1 is primarily detected by PMT A, while light output from crystal 2 is detected by PMTs A and B equally

For some years, however, a few major medical-imaging companies have been investing heavily in PET scanner designs for clinical use, as the PET modality has been proven to be far superior to other modalities in clinical oncology. This has also been further promoted by the introduction of fluorodeoxyglucose (FDG) marked with the positron emitter ^{18}F , a type of sugar that concentrates in malignant tumours because of their faster growth than surrounding tissue. Owing to the 110 minute half life of the ^{18}F isotope, FDG can be transported from a cyclotron centre to several PET hospitals in densely populated areas offering better economy to the PET modality. Another approach that has recently become popular is to put the cyclotron in a van and visit several hospitals to a schedule of need. This market boom has resulted in PET applications becoming another backbone of the world photomultiplier market, further driving the development of new, dedicated PMT types.

The first PET scanners used NaI(Tl) scintillators and slow coincidence detection; the principal demands on the photomultipliers were good energy resolution and moderate time resolution. Later scanners use multiple detector rings with the green-emitting scintillator $\text{Bi}_4\text{Ge}_3\text{O}_{12}$ (colloquially, BGO). These scanners therefore require photomultipliers with good green sensitivity (which is provided by the introduction of green-extended bialkali cathodes), and improved time resolution, which calls for good collection efficiency in multi-electron mode. Note, however, that the time resolution is, in any case, limited by the slow decay time constant (300 ns) and dimensions of the BGO crystals.

A disadvantage of BGO as a scintillator is that it has a refractive index of 2.15, which makes it difficult to couple to the normal glass windows of photomultipliers (refractive index about 1.5), so the light gets trapped in the scintillator. Trials of other glasses have not yet led to a practical solution to this problem. Design effort has concentrated instead on trying to match the scintillators to the windows of the photomultipliers in such a way as to achieve the required dense rectangular packing of detectors without too much 'geometrical' light loss.

PET scanners with NaI(Tl) or BGO scintillators use slow coincidence detection. By using fast coincidence detection it becomes possible to 'unfold' a lot of noise from the picture. With this in view a number of scanners were developed around the fast scintillator compound CsF. This emits in the blue and ultraviolet and requires fast photomultipliers with UV-transparent windows. To attain the required efficiency for 511 keV γ -photons, CsF scintillators have to be several centimetres long. This sets a limit to the time resolution of the scintillator itself (as well as introducing parallax problems) and therefore puts severe demands on the time resolution of the photo-

multiplier. For multi-electron events, coincidence times below 400 ps can be resolved between two opposite channels.

The discovery of the fast 190/225 nm peak (decay time constant 0.6 ns) of BaF₂ has practically killed all interest in the highly hygroscopic CsF and directed most of the development effort in fast coincidence scanners toward BaF₂ and fast quartz window photomultipliers. Though PET scanners with two scintillators working with one photomultiplier (GSO as the fast channel and BGO as the slow channel) have been produced, GSO prices have so far limited further interest. Nevertheless, with a decay time constant of 35 ns and high stopping power, GSO now forms the basis of one of the new scanners. Others use blocks of many small BGO 'needles', a few millimetres in cross-section and about 25 mm long coupled to several 19 mm diameter photomultipliers or to two 24 mm square, dual photomultipliers. Future developments are expected to see the blocks read out with one single, quadrupole tube (four tubes in one envelop) developed by Photonis for decreasing the readout cost.

The recent introduction of the lutetium orthosilicate (LSO) scintillator by one major company, offering a slightly higher density than BGO, a decay time constant of only 42 ns combined with a refractive index of only 1.81 and a light output seven times that of BGO has contributed to a major leap in PET scanner performance. Several other new inorganic scintillators are under investigation to compete or further improve on the parameters of LSO.

Animal PET scanners. As a spin-off from PET scanners there has in recent years appeared a new market for photomultipliers for so called animal PET scanners. These are miniaturised PET scanners for studying the influence of new pharmaceuticals on rodents such as mice. Most profit from the new LSO scintillator forming narrow scintillator needles that are read out by multi-channel photomultipliers, sometimes via fibre light guides. The market, however, is still limited to pharmacological research.

Mamography scanners. The introduction of multi-channel photomultipliers has also stimulated a new market dedicated to mammography scanners based on different versions of the new scintillators, and some small companies have recently been created to promote such systems.

7.2 Analytical applications

Scintillation counting has opened the way to numerous techniques for measuring specific biochemical levels in body tissues and fluids with the aid of radioactive

tracers. Two of the most widely used are liquid scintillation counting and radio-immuno assay.

7.2.1 Liquid scintillation counting

The aim of liquid scintillation counting (LSC) is generally to trace the pathways of specific substances in the body or to measure their tendency to concentrate in specific tissues or fluids. For example a quantity of the substance in question is labelled with a radioactive tracer; after it has been assimilated, a sample is taken of the body tissue or fluid under investigation and the concentration of the labelled substance is determined by measuring its radioactivity.

Because of the organic nature of the samples to be measured, the radioactive tracer is usually an isotope of hydrogen or carbon. The most commonly used are the β -emitting isotopes ^3H and ^{14}C , with half-lives of 12.26 and 5500 years respectively, and maximum radiation energies of 18 keV and 156 keV (Fig.7.5). To detect such soft radiation it is necessary to incorporate the sample in the scintillation medium itself, hence the choice of a liquid. This not only prevents the β -particles from being stopped prematurely by air or detector windows, it also makes the response of the scintillator substantially isotropic (' 4π geometry'), which greatly improves counting efficiency at the very low energy levels involved.

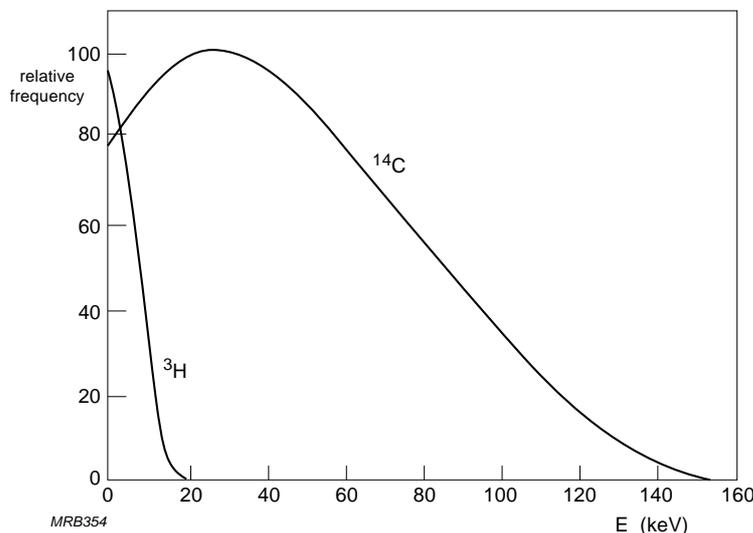


Fig.7.5 Relative spectral energy E of β emission from tritium (^3H) and carbon (^{14}C)

The liquid scintillator is a 'cocktail' consisting of at least one fluorescent aromatic solute in an aromatic solvent. One effective combination is PPO (2,5 diphenyl

oxazole) in toluene, which has a broad spectral emission peak at 370 nm. To avoid the use of a photomultiplier with a UV-glass or quartz window, a secondary solute that acts as a wavelength shifter may be added. A popular secondary solute is POPOP (1,4-bis-2-(5-phenyloxazolyl)-benzene) which shifts the spectral emission peak to 420 nm.

In a typical liquid scintillator, β -particles of 5 keV (mean energy for ^3H) will produce only about 50 photons, not all of which will reach the photomultiplier cathode. The resulting current pulses are therefore of the same order of magnitude as the dark pulses, so coincidence techniques must be used to distinguish them (Fig.7.6).

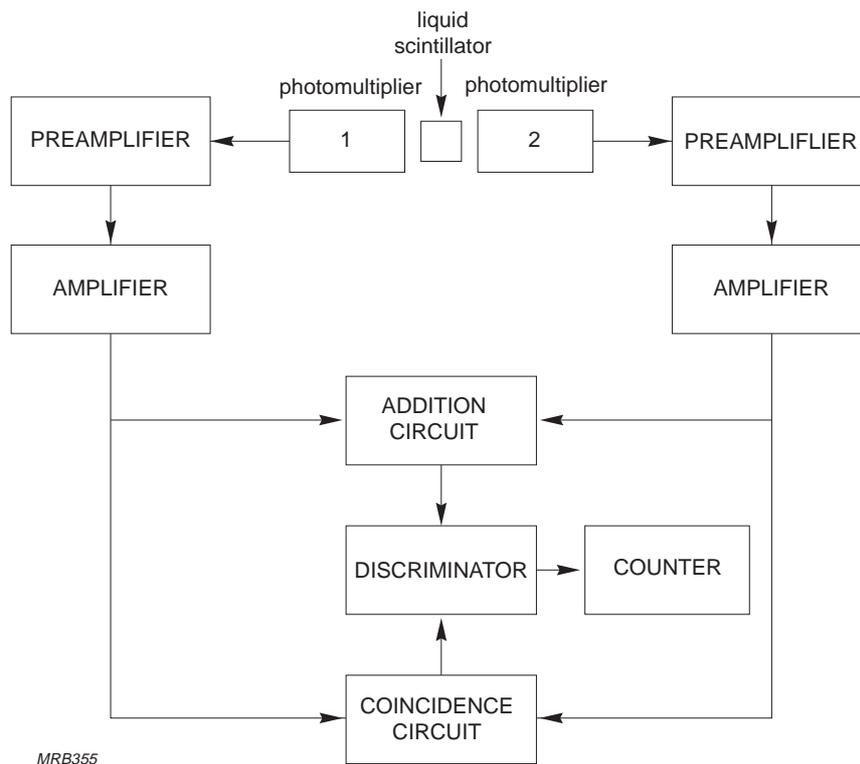


Fig.7.6 Schematic diagram of a liquid scintillation counter

The scintillator is observed by two photomultipliers 180° apart whose outputs are added together. Whenever a scintillation is observed by both tubes in coincidence, the added outputs are gated to a discriminator; provided their combined amplitude falls within a set window, the pulse is then passed to a counter. Except for chance coincidences, which are rare, random dark-current pulses from both tubes occur independently of each other and are rejected.

If n_1 and n_2 are the average dark pulse rates of the two tubes and τ is the resolving time of the coincidence circuit, the chance coincidence rate $n_f = 2n_1n_2\tau$. This usually amounts to less than one count per minute.

Besides the chance coincidences due to the dark pulse rates of the two tubes, however, other spurious coincidences also occur. One source is background radiation, for instance high-energy cosmic rays that penetrate to the scintillation sample. Another is crosstalk between tubes, which occurs when a light flash in one tube – due to the background radioactivity of its own glass, internal ionization, or a cosmic ray – is also seen by the other tube. The background count due to these causes can be minimized, but not wholly eliminated, by good shielding and careful setting of the discriminator window.

The counting efficiency E , in per cent, is

$$E = \frac{n_c - n_b}{N} 100$$

where n_c is the count rate, n_b the background count rate, and N the actual disintegration rate. A figure of merit combining the counting efficiency and background effects is E^2/n_b , where n_b is the background count rate measured under normal operating conditions but with no sample in the liquid scintillator. Though coincidence counting does reduce the counting efficiency E , it considerably increases the figure of merit by excluding nearly all the dark pulses due to thermionic emission and background radiation.

The sample chamber and vial must be designed to ensure maximum photon collection by both photomultipliers and equal division of photons between them. In present-day equipment the energy loss per photon in the scintillator, between point of origin and photocathode, averages 400 eV. Knowing this value and the energy spectrum of the radioisotope to be measured, one can calculate the efficiency for ^3H as a function of photocathode sensitivity, with photon energy loss as parameter (Fig.7.7).

The counting efficiency of present-day liquid scintillation counters is generally better than 60% for ^3H . Provided the sample chamber is well designed and well shielded, the background count rate n_b may be as low as 15 to 18 per minute, giving a figure of merit greater than 200. Modern counters with automatic sample changers can measure and register hundreds of samples in a single loading.

Photomultipliers for liquid scintillation counting require high quantum efficiency at the scintillation wavelength and low dark-count rate. Bialkali cathodes (SbKCs),

whose sensitivity is best toward the short wavelength end of the visible spectrum are generally preferred, sometimes with quartz input windows. The tube envelope should be of low ^{40}K glass to minimize background radiation, and the input window should be as thin as possible to minimize its volume.

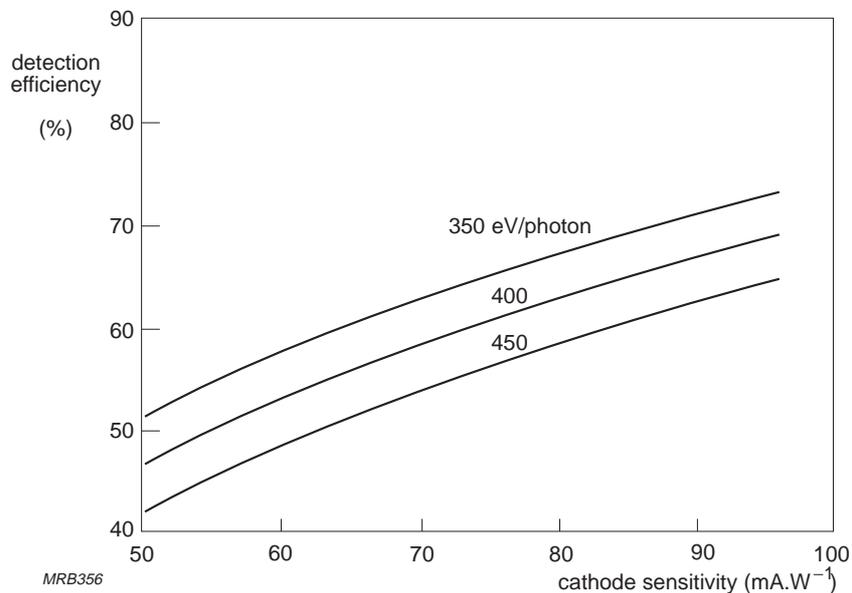


Fig.7.7 Detection efficiency as a function of photocathode sensitivity for ^3H , with energy loss as parameter

7.2.2 Radio-immuno assay

Radio-immuno assay (RIA) is a technique for measuring minute concentrations of specific substances, such as hormones, in biological fluids. A known quantity of the substance sought is labelled with a radioactive tracer and mixed with a specimen of the biological fluid under investigation. A comparatively small quantity of a binding agent (e.g. an antibody) with a specific affinity for the substance is then added to the specimen and allowed to react with both the labelled and unlabelled molecules of the substance. When the reaction has reached equilibrium the bound and the remaining free molecules of the substance are separated from each other, for instance by centrifuging. From the concentration of the radioactively labelled molecules in either fraction, as measured by scintillation counting, it is then possible to determine the concentration of the unlabelled molecules in the original specimen. In this way it has proved possible to detect and measure quantities as minute as picogrammes (10^{-12} g). Because it does not expose the subject to any radioactivity at all, RIA has become the method of choice for a large number of diagnostic tests – in fact the only practical method for some – and is now practised on a very wide scale as a routine procedure.

The tracers used in RIA are γ -emitters, and the detectors are well-type NaI(Tl) scintillators coupled to individual photomultipliers. Commonly used tracers are ^{125}I (35 keV γ -radiation) and, to a less extent (because it requires greater detection efficiency and has a shorter half-life) ^{131}I (360 keV γ -radiation). For determining the concentration of substances that contain cobalt but not iodine (e.g. vitamin B₁₂), one of the radioisotopes ^{57}Co , ^{58}Co or ^{60}Co may be used.

Counters for use in RIA are available as automatic sample changers, having one or more detectors, or as simple table-top equipment with up to 20 detectors in which the sample tray is changed by hand. Automatic sample changers, for which high capacity and therefore high efficiency are important, commonly use 2 inch scintillators coupled to similar-sized photomultipliers. Table-top models use 19 mm or 29 mm scintillators and photomultipliers.

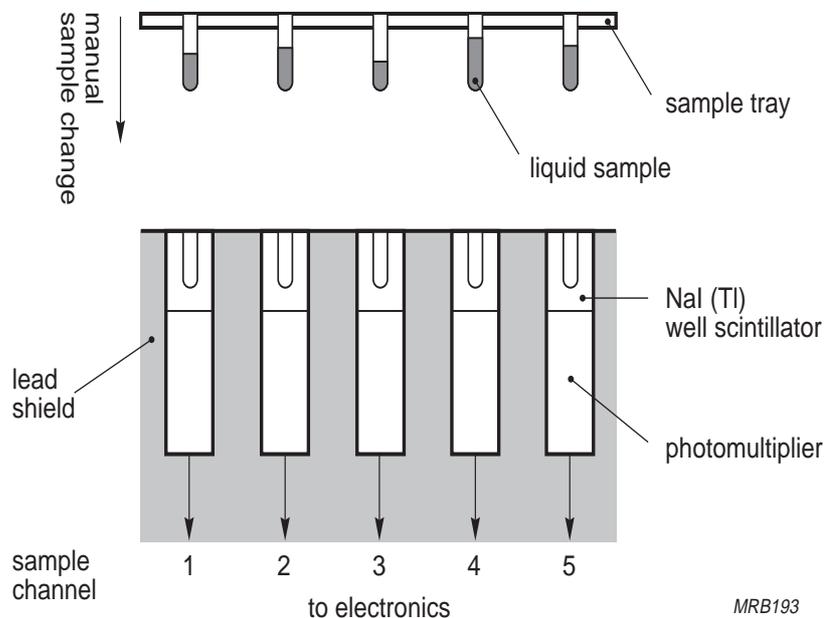


Fig.7.8 RIA counter with 5 channels

Because of the well in the scintillator the γ -interaction efficiency is high but the light transport to the photomultiplier is less than ideal. This puts demands on the photomultiplier energy resolution at the low γ -energies of some tracer isotopes (e.g. ^{125}I); it also imposes a requirement for low noise at those energy levels, plus good stability, to ensure that the detected energy peak remains within the discriminator window. Normally, 8- to 10-stage standard tubes afford sufficient gain and energy resolution, and are reasonably priced. The RIA technique is, nevertheless, currently experiencing serious competition from another technique known as chemiluminescence immuno assay (CLIA), see §8.1.2.

7.3 Industrial applications

Of the many industrial applications of scintillation counting, the three treated briefly here – non-destructive analysis of materials, thickness and density measurement, and oil-well logging – illustrate common principles and can be regarded as typical.

7.3.1 Non-destructive analysis

Two widely used methods of non-destructive analysis based on scintillation counting are:

- *Activation analysis*, in which the specimen is subjected to nuclear radiation so that the constituent elements become radioactive and can be identified and quantified by their own radiation
- *X-fluorescent analysis*, in which the specimen is exposed to γ , X or charged-particle radiation to stimulate emission of X-ray spectra characteristic of the constituent elements.

Activation analysis. Neutron irradiation is the most common means of activation; however, irradiation with high-energy γ -rays or charged particles, such as protons or deuterons, is also possible. The characteristics of the induced radioactivity that identify the constituent elements are its type (α , β , γ , X), its decay half-life, and its energy spectrum. As a rule these are sufficient for good analytical discrimination without supplementary chemical separation. The advantages of activation analysis are: speed; independence of the form in which elements are combined; preservation of the specimen; and great sensitivity, especially if the irradiation flux is large.

Short-lived radioisotopes can be identified by plotting their decay and determining the half-life. If the half-lives of constituent isotopes differ sufficiently they can be distinguished by a graphical breakdown; Fig.7.9 shows the decay of a mixture of ^{165}Dy and $^{152\text{m}}\text{Eu}$ as the sum of their respective half-lives. The greater the number of radioisotopes in the specimen and the more similar their half-lives, the less accurate the graphical method is.

For radioisotopes with very long half-lives it is more convenient to analyse the energy spectrum of the emitted radiation. Analysis by γ -spectrometry after irradiation with thermal neutrons can be done simply and quickly. There are several methods for distinguishing the characteristic energy peaks in a pulse-height spectrum; Fig.7.10 shows a subtractive analysis of the spectrum obtained from activated impurities in a specimen of natural calcium fluoride.

Neutron-activation analysis systems have recently been installed in baggage-handling departments of several major airports in an attempt to discover hidden explosives by detecting the nitrogen they contain. These systems are large, complex and expensive and, so far, they have met with only limited success. This is because they operate with thermal neutrons from a ^{252}Cf source. Only with fast neutrons from a dedicated particle accelerator would it be technically feasible to reliably detect the small amount of nitrogen present in explosives in a reasonable time (whilst the baggage is conveyed on a moving belt for instance). Cost and false alarms then become the limiting factors.

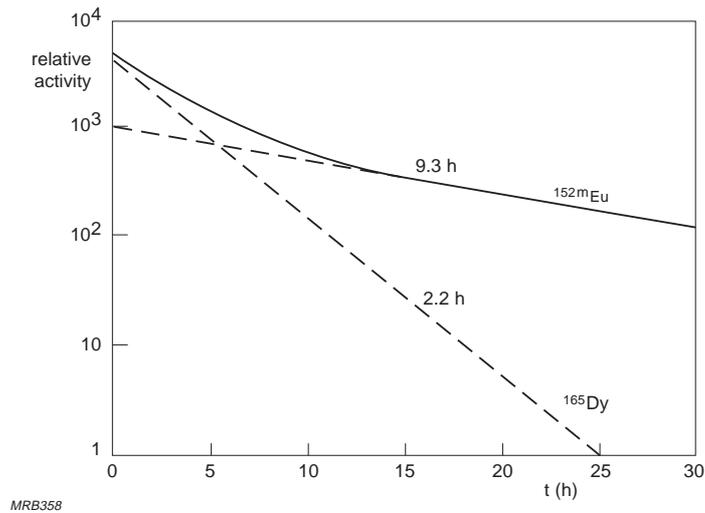


Fig.7.9 Illustration of the graphical separation of two half lives

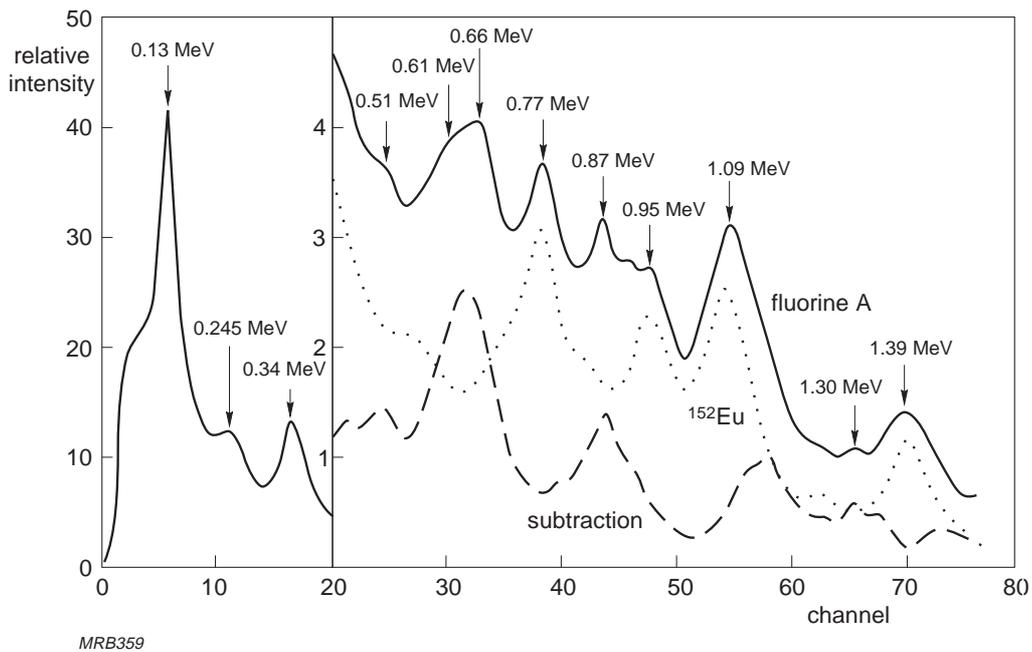


Fig.7.10 Subtractive analysis of a natural calcium fluoride spectrum

X-fluorescent analysis is based on the characteristic X-ray emission spectra of the elements constituting the specimen. The emission may be stimulated by bombardment with charged particles or irradiation with X- or γ -rays. Electron bombardment gives low emission efficiency (about 1%) and affects only the surface of the specimen. Proton bombardment gives good analytical sensitivity but requires the use of a particle accelerator; for most analyses X- or γ -irradiation is preferred.

High-power X-ray generators are available which yield an intense, highly penetrating flux capable of stimulating X-ray emission from far below the surface of the specimen. Radioisotope sources of X- or γ -radiation are also practical. They are highly stable, can supply a wide range of energies, and have the advantage of being small enough for use in portable instruments.

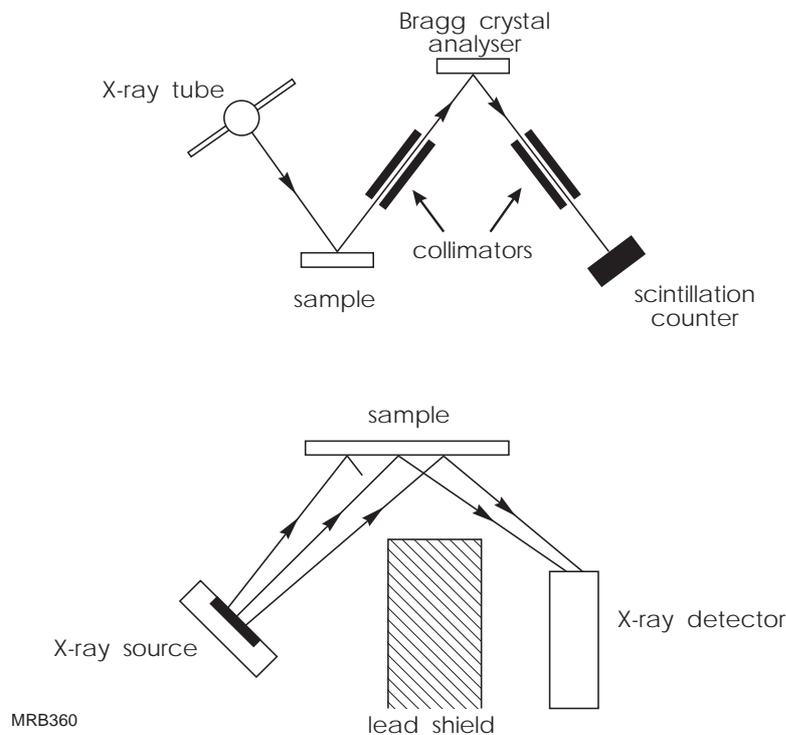


Fig.7.11 Schematic diagram of two X-ray spectrometers (a) dispersive, (b) non-dispersive

Two types of spectrometer are used, dispersive and non-dispersive (Fig.7.11). The efficiency of dispersive spectrometers is low, but this is compensated by the high intensity of the radiation; a multichannel dispersive spectrometer can identify a large number of elements simultaneously. In non-dispersive spectrometers the analysis is

based solely on pulse-height discrimination; owing to the low resolution of scintillation counters toward the low-energy end of the spectrum, this type of spectrometer is limited to analysis of spectra with comparatively few lines. The alternative is to use a high-purity germanium or Si(Li) detector.

7.3.2 Thickness and density measurement

Scintillation counters can also be used for measuring thickness and density by means of ionizing radiation. Though they require more complicated electronics than Geiger-Müller counters, they have the following advantages which are in many applications decisive: short resolution time, output pulse proportional to the radiation absorbed in the scintillator, and very high efficiency for X- and γ -radiation. The photomultipliers are usually general-purpose types with good shock and vibration resistance.

Thickness measurement can be either by transmission or backscatter (Fig.7.12), the choice being usually one of convenience. The radiation may be α , β or γ – each has advantages that suit it to some applications and limitations that exclude it from others.

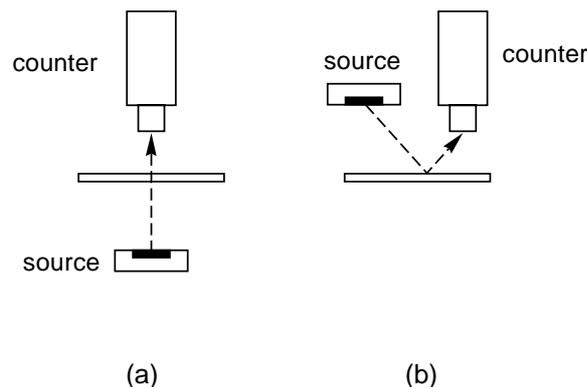


Fig.7.12 Principle of thickness measurement (a) by transmission, (b) by reflection

α -radiation has low penetration even in air and is therefore suitable for measuring only very small thicknesses, and then only by transmission, not backscatter. The detector should be spaced no more than a few millimetres from the source. Results are subject to variation due to environmental factors such as atmospheric pressure and humidity.

β -radiation has much greater penetration than α -radiation and can be used for backscatter as well as transmission measurements. Table 7.1 lists characteristics of some β -sources; those of low energy have the disadvantage that their radiation is too

quickly scattered away from the detector. On the other hand, some of the higher-energy ones emit radiation of more than one energy and type (e.g. β and γ), which complicates the interpretation of the results. Half-life is also a practical consideration: the longer it is, the less often the measuring set-up needs to be recalibrated.

Measurement by β -transmission is useful in the paper, rubber, plastics, metal and tobacco industries. The measurement may be continuous and is often integrated into an automatic process-control system. β -backscatter is useful for measuring the thickness of coatings.

Table 7.1 Some β -radiation sources

source	half-life	max. energy (MeV)	measurable thickness (mg/cm ²)
³ H	12.26 yr	0.018	small range
⁶³ Ni	125 yr	0.067	small range
¹⁴ C	5568 yr	0.155	2 – 10
¹⁴⁷ Pm	2.26 yr	0.223	4 – 15
⁸⁵ Kr	10 yr	0.695	25 – 100
²⁰⁴ Tl	2.7 yr	0.770	20 – 150
⁹⁰ Sr(+ ⁹⁰ Y)	28 yr (64.4 h)	0.54 (2.26)	60 – 550
¹⁴⁴ Ce(+ ¹⁴⁴ Pr)	285 d (2.6 yr)	0.31 (3.10)	150 – 1250
¹⁰⁶ Ru(+ ¹⁰⁶ Rh)	1 yr (24 h)	0.039 (3.50)	200 – 2000

γ -radiation has sufficient penetration for measuring the thickness of dense materials. As the best results are obtained by Compton backscattering the γ -rays in the material, only sources with γ -ray energies from a few tens of keV to a few MeV are normally used; characteristics of some γ -sources are listed in Table 7.2.

Thickness measurement by γ -transmission has applications in metallurgy. With appropriate precautions against the high level of radioactivity required, it can be used to check the uniformity of, for example, aluminium ingots.

Surfaces that are conveniently accessible from only one side – for instance, tank walls, ships' hulls, pipes, and road paving – can be measured by γ -backscatter. With a 0.1 mCi ⁶⁰Co source it is possible to measure up to 15 mm of steel or 30 mm of aluminium by backscatter.

Table 7.2 Industrial γ -sources

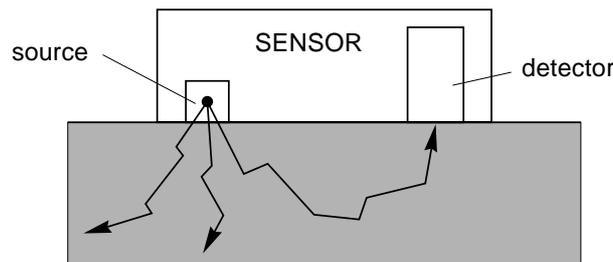
source	half-life	energy (MeV)	half-thickness in steel* (mm)
^{55}Fe	2.94 yr	0.006	0.01
^{170}Tm	0.35 yr	0.084	1.0
^{85}Kr	10.6 yr	0.520	12.3
^{137}Cs	30 yr	0.662	16.5
^{60}Co	5.25 yr	1.17 & 1.33	33.8

* The thickness in which the radiation flux (photons/cm²s) is halved

Density measurement involves the same principles as thickness measurement: the measured material absorbs or scatters radiation in proportion to its density.

An important application of density measurement by transmission is flow monitoring, with a radiation source and a scintillation counter disposed on opposite sides of the conduit. This is especially useful for monitoring high-pressure, high-temperature, or hazardous fluids. Detection of flow conditions such as the onset of turbulence is also possible.

Soil density can be measured in situ by Compton scattering. Portable instruments based on this principle are used to obtain continuous density profiles of road surfaces (Fig.7.13).



MRB362

Fig.7.13 Soil density measurement by means of Compton scattering

7.3.3 Oil-well logging

In exploratory drilling, a bore-hole probe containing a γ -ray or neutron source and a scintillation detector can aid in assessing the probability of hydrocarbon deposits. The detector is shielded from the direct radiation of the source and responds only to

Compton scattering from the walls of the bore hole. By analysing the variations in count rate as the probe descends, an experienced operator can draw valuable conclusions about the structure and composition of the successive geological strata.

By the nature of the application, it is obvious that the photomultipliers used have to be extremely rugged, especially in the now popular activity of ‘logging while drilling’ (LWD). They also have to be able to withstand high temperatures without unacceptable loss of anode sensitivity. At the 10 km depths to which a bore-hole probe may be lowered, temperatures often reach 150 – 200 °C, far above the temperature (80 – 100 °C) at which the performance of most photocathodes seriously declines. The best choice of cathode material is bialkali SbNaK, without Cs. The cathode, as well as other materials in the tube, has to be specially processed for this application.

7.4 High-energy physics

7.4.1 Collider detectors

Scintillation detectors find extensive application in the very large installations used in high-energy physics to analyse the fragments generated in particle-particle collision experiments.

Designed to cover 4π steradians around the collision point so as to account for all the energy involved, such an installation (Fig.7.14) commonly includes:

- a vertex detector closely surrounding the beam pipe
- a tracking detector
- an electromagnetic calorimeter surrounding the tracker
- a hadron calorimeter surrounding the electromagnetic calorimeter
- a time-of-flight detector
- muon chambers.

A solenoid (usually many metres long and several metres in diameter) generates a strong (multi-tesla) magnetic field coaxial with the beam pipe to facilitate identification of charged fragments.

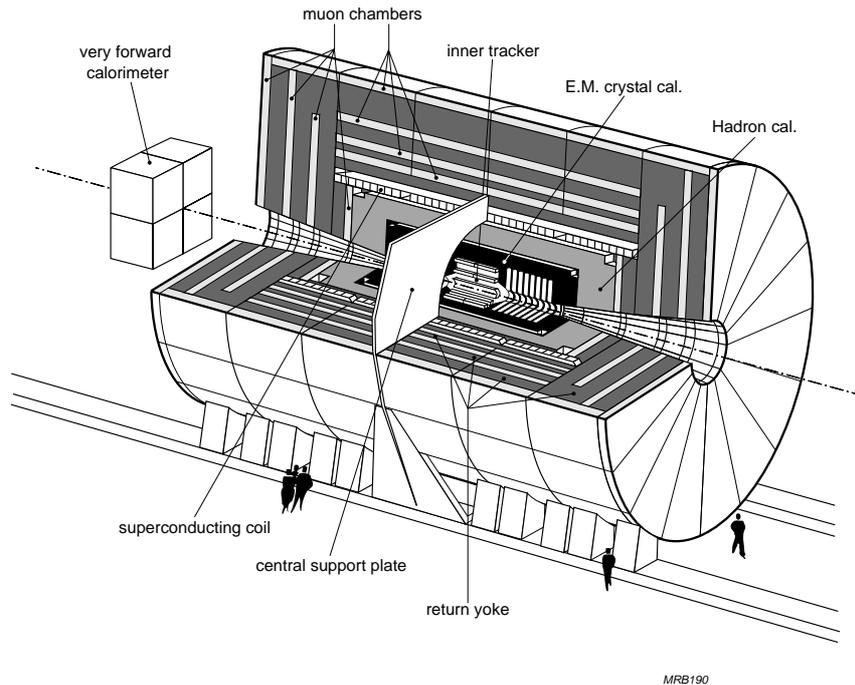


Fig.7.14 CMS 4π detector proposal (courtesy of CERN, Geneva)

Vertex detector. This may consist of thin organic scintillating fibres read out at one or both ends by photomultipliers. The energy deposited in the fibres is small and the attenuation of the light before it reaches the photomultipliers (located outside the magnetic field) is large; most events therefore give rise to only a few photoelectrons.

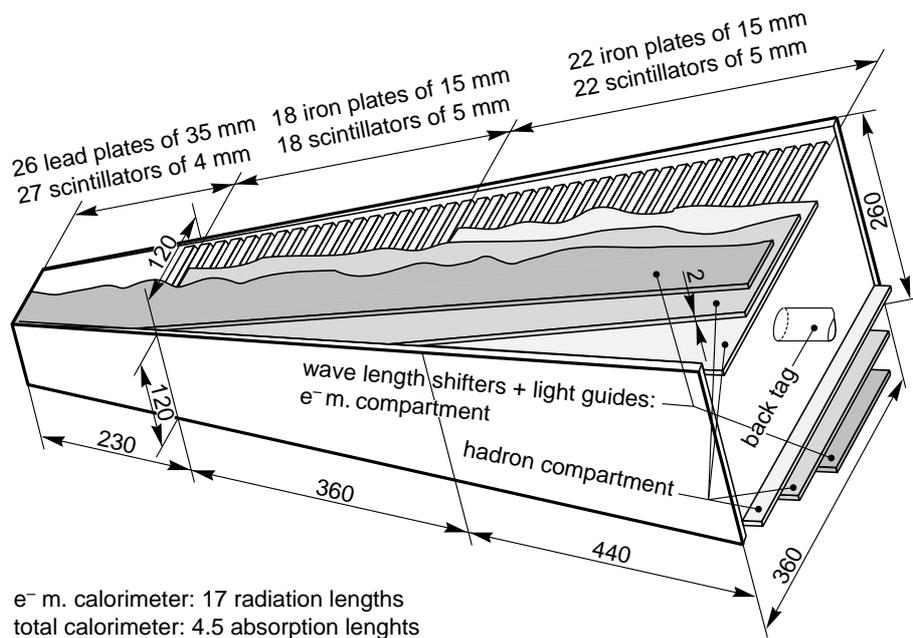
Today's designs are based on silicon-strip detector arrays or even CCDs with low-noise preamplifiers that eliminate the use of photomultipliers in vertex detectors.

Trackers. A central tracking detector traditionally uses many thousands of wires and gas multiplication. It distinguishes between the tracks of charged and neutral fragments by means of the magnetic field generated by a solenoid. With the collision rates in proton-proton colliders increasing to 60 – 70 MHz, more advanced tracking techniques will be needed. In this respect, new scintillating fibres combined with multi-channel photomultipliers may be one solution. Since the light attenuation in these fibres is less for green wavelengths (especially if the fibres have suffered radiation damage), tubes with extended green response photocathodes should be used. As multi-channel photomultipliers cannot operate in magnetic fields of several tesla, the fibres have to convey the light out of the field. The resulting attenuation means that the tubes have to work in single photo-electron mode. Competing detectors include cooled silicon pixel arrays and avalanche diode arrays that can operate inside the field, thin-walled, small-diameter gas multiplication tubes ('straw tubes'), microstrip gas counters and even silicon-strip detectors.

Electromagnetic calorimeter. This is located outside the tracker and consists of a barrel detector with end caps to stop and analyse electromagnetically interacting fragments. The barrel and end caps often consist of plastic scintillators interleaved with some material with high atomic number (high stopping power) such as lead, iron, or uranium.

As the photons generated in the scintillators are in the ultraviolet, the scintillators are doped to shift the wavelength to around 400 nm. Wavelength shifting rods or plates that also act as light guides are often used to shift the light further towards the green-yellow part of the spectrum (Fig.7.15) where the attenuation length is very low (so that the photomultipliers can be located outside the magnetic field).

Such calorimeters often use thousands of 1 to 2 inch photomultipliers. Important requirements are that their gain must be very stable and, in view of the large dynamic range called for (1000 – 2000), their pulse amplitude linearity must extend at least to 100 – 150 mA. In addition, because of the large quantity involved, they must be reliable and reasonably priced.



MRB192

Fig.7.15 Calorimeter tower based on the wavelength shifter principle (courtesy of CERN, experiment UA 2)

The best choice of photocathode is extended green sensitive bialkali. The integral quantum efficiency with BBQ-, Y7- or K27-doped light guides is typically 12 – 20%.

Because of the large dimensions involved, it is never possible to derive accurate time information using calorimeter towers, so response time is not an important factor and most types of photomultiplier will be fast enough. Linearity and stability, however, are important. The severe demands on stability, in particular, have been a major stimulus to improving photomultiplier performance, leading to benefits in several other areas (gamma cameras for example).

Finally, mesh-dynode photomultipliers able to sustain reasonable gain in an axial magnetic field have been developed for situations where the electromagnetic calorimeter must remain inside the field. New HPDs (see §1.5.7) can work in an axial magnetic field of up to a few tesla.

Hadron calorimeter. The construction is similar to that of the electromagnetic calorimeter: a barrel with end caps in which e.g. organic scintillators (scintillating fibres) interleaved with high-atomic-number material are coupled to photomultipliers by BBQ-, Y7- or K27-doped wavelength-shifters/light-guides. Here though, the dynamic range required is usually lower, so tubes with pulse amplitude linearity up to 100 mA will be sufficient. The stability requirements are just the same. Other calorimeter designs also exist based on, for example, liquid argon without photomultipliers.

PbWO₄ calorimeters

As the next challenge in high-energy physics is to find the signs of the Higgs particle(s) that is supposed to define the mass of the other particles in the standard model, future Large Hadron Collider (LHC) experiments will be mainly directed at very-high energy particles. Therefore, with much efforts, a new inorganic scintillator, the PbWO₄ has been developed in cooperation with Chinese and Russian institutes and industry. Its high stopping power (density 8.28 g/cm³) and short decay time constants of 2, 7 and 26 ns make it ideal for the huge CMS electromagnetic calorimeter. Its very low light output of only 0.8% of NaI(Tl) will easily be compensated by the very high energy of the particles of interest. Moreover, its peak emission wavelength of 480 nm makes possible to read it out with silicon avalanche photodiodes (APDs) that can operate in the perpendicular magnetic field of the calorimeter barrel. For the calorimeter end caps, where the radiation is very high close to the beam line, low-priced Russian vacuum phototriodes (VPTs) will be used.

This development will also increase the availability of PbWO₄ for lower-energy detectors in medium-energy physics experiments, where, however, owing to the lower energy of the particles, green-extended-response photomultipliers will be needed.

Lead-glass calorimeters. Alternative constructions for electromagnetic calorimeters placed outside of the magnetic field may use long lead-glass blocks instead of interleaved plastic scintillators. The photomultipliers, coupled direct to the blocks without light guides or wavelength shifters, are usually 2 or 3 inch tubes. As the light generated is due to Cherenkov radiation (§8.3) and is strongest toward the blue and UV end of the spectrum, normal SbKCs-cathode tubes are the best choice, for sensitivity as well as stability and price. The light output is sufficient for a 6- or 8-stage tube; in fact, even a gain of $10^2 - 10^3$ gives a signal large enough to be easily handled.

Fibre calorimeters. Instead of interleaving plastic sheets and high-Z material, calorimeters can also be produced from scintillating fibres uniformly stacked through the high-Z material. The signal from many fibres is then coupled to one photomultiplier. Other designs comprise scintillating tiles sandwiched between lead tiles, with the scintillating tiles read out via bundled wavelength-shifting fibres coupled to clear fibres that convey the light from all the tiles to a single photomultiplier tube. As the signal is composed of thousands of photoelectrons, only moderate gain is needed in the photomultiplier, but good stability and wide dynamic range are very important. For the best linearity, linear-focusing tubes are used, and since the light signal is in the green part of the spectrum to minimize light attenuation in the fibres, the best result is achieved with a green extended bialkali SbKCs photocathode.

Magnetic-field sensitivity: photodiodes, phototriodes. Because of the strong magnetic field surrounding the solenoid, the photomultipliers used require heavy shielding. This complicates design and limits the freedom with which the tubes can be placed.

As an alternative, detectors have been constructed with Photonis vacuum photodiodes (VPDs) that could be rendered reasonably immune to strong magnetic fields within 30° of axial. In conjunction with NaI(Tl) scintillators and special low-noise preamplifiers to compensate for their own unity gain, these have given good results. For use with lead-glass blocks (whose light output is very much lower than that of NaI(Tl)), vacuum phototriodes (VPTs) able to work in strong magnetic fields within 30° of axial with a gain of about 10 have also been used in end caps. Results are, again, excellent, when sufficiently low-noise preamplifiers are used. New glasses that give higher light output by scintillation may further advance this line of development.

Experiments have shown that BGO and CsI(Tl) scintillators in combination with large-area (1 cm^2) silicon photodiodes or avalanche photodiodes (APDs) and low-

noise preamplifiers can also achieve low enough noise levels to be feasible at high collision energies and moderate counting rates. Their advantages over photomultipliers are short length and immunity to magnetic fields.

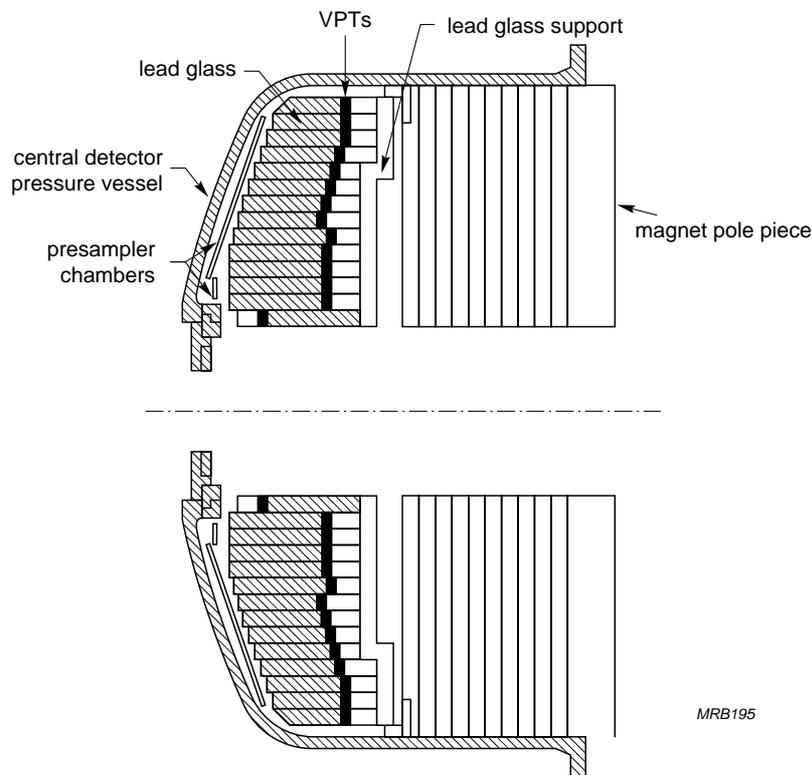


Fig.7.16 Lead-glass calorimeter end caps and Photonis VPTs (courtesy of OPAL collaboration, CERN, Geneva)

Disadvantages are high sensitivity to temperature variations, and high capacitance which induces preamplifier noise and makes the read-out channel slow to integrate all the charge in the pulse. They are fragile and their long term reliability is still not fully proven. This means placing several photodiodes per scintillator for redundancy reasons and/or to cover the required area. For these reasons, low-capacitance VPTs and VPDs could again become attractive alternatives for the next generation of ‘fast’ collider experiments.

Time-of-flight (TOF) detectors. In high-energy physics installations it is also important to obtain time-triggering information about the fragments generated. This is usually done with the aid of long rods of very fast plastic scintillators (e.g. Pilot-U, now replaced by BC418), with decay time constants of only 1 – 2 ns, observed from both ends by fast photomultipliers whose outputs are compared in time as well as

amplitude. TOF detectors often consist of a barrel placed inside or between the calorimeters.

When the photon emission of the scintillator and the electron emission of the photomultiplier cathode both obey laws of the type

$$f(t) = \left(\frac{\bar{n}}{\tau}\right) \exp\left(-\frac{t}{\tau}\right)$$

the variance of the time registered by the photomultiplier is

$$\sigma_t^{*2} = \frac{\tau^2 + \sigma_t^2}{\bar{n}_{k,s}}$$

where τ is the decay time constant of the scintillator, σ_t^2 is the transit-time variance of the photomultiplier, and $\bar{n}_{k,s}$ the mean number of photoelectrons per scintillation pulse. For optimum timing accuracy, therefore, the scintillator should be chosen for high scintillation efficiency and short decay time; and the photomultipliers, for high cathode sensitivity at the scintillation wavelength (to maximize $\bar{n}_{k,s}$) and small transit-time variance. Since the transit-time variance of fast photomultipliers is usually negligible compared with the decay time constant of even very fast scintillators, it is the latter and the length of the rod (different light paths) that mainly limit the time resolution in such detectors. It is therefore common to trigger already on the very first (directly arriving) few photons which demands a fast photomultiplier with high gain ($\approx 10^7$). Some years ago, to improve timing, Photonis developed a special tube with a 'screening' anode grid as proposed by Moszynski and this improvement is now being further explored to get the very best timing. Making the rods from scintillating fibre bundles can improve performance.

In some experiments, Cherenkov radiation in a gas-filled volume is used instead of scintillator rods to measure time-of-flight. Then, it is the photomultipliers that mainly limit the time resolution, since the Cherenkov radiation arrives at the photomultiplier practically instantaneously (§8.3).

Muon trackers. Until recently, muon detectors used wire/gas multiplication principles to determine the tracks of the muons penetrating the calorimeters. With the development of scintillating fibres, however, discussions now centre around the development of muon detectors comprising many layers of fibres read out by multi-channel photomultipliers incorporating extended-green bialkali photocathodes.

7.4.2. Fixed-target detectors

Hodoscopes. In experiments in which a beam of accelerated particles is focused onto a target, the resulting fragments scatter in various (forward) directions and are subsequently separated by a magnetic field. To trace their paths, hodoscopes are used consisting of crossed X-Y arrays of many flat, parallel, plastic scintillator strips or fibres coupled to photomultipliers. Apart from small size (19, 25 or 29 mm), the main photomultiplier requirement is high gain, for the typical scintillation excites only a few photoelectrons. In experiments in which pulse-height analysis is required this imposes demands on the linearity of the tubes. Usually, though, this is not the case; the usual information required from each scintillator is a simple yes/no.

Note: the hodoscope is also often used as a time trigger, in which case the principal demand on the tubes is narrow transit-time spread. Since most plastic scintillators emit at around 400 – 420 nm, a high-sensitivity bialkali SbKCs photocathode is necessary.

In the future, good spatial resolution is expected to become important, and here multi-channel tubes in combination with scintillating fibres are likely to be the best option. And since the attenuation of the fibres is lowest in the green part of the spectrum, these tubes should have extended-green photocathodes.

Veto and RICH counters

As modern veto counters and ring-imaging Cherenkov counters (RICH) do not use scintillators they are treated in Chapter 8.

7.5 Cosmic-ray detectors

Although some cosmic-ray detectors use liquid scintillators and some use the scintillation of the earth's atmospheric volume, these detectors are all treated in Chapter 8.

CHAPTER 8

NON-SCINTILLATOR APPLICATIONS OF PHOTOMULTIPLIERS

8.1 Analytical applications

- 8.1.1 Bioluminescence
- 8.1.2 Chemiluminescence
- 8.1.3 Flame spectrometry
- 8.1.4 Raman spectrometry

8.2 Non-analytical applications

- 8.2.1 Image conversion
- 8.2.2 Laser scanners
- 8.2.3 Laser telemetry

8.3 Cherenkov radiation

8.4 Cherenkov experiments

- 8.4.1 High-energy physics experiments
- 8.4.2 Proton decay experiments
- 8.4.3 Solar neutrino experiments
- 8.4.4 Neutrino oscillation experiments
- 8.4.5 DUMAND experiments
- 8.4.6 Air-shower experiments
- 8.4.7 Towards a common goal
- 8.4.8 A vital tool for physics

NON-SCINTILLATOR APPLICATIONS OF PHOTOMULTIPLIERS

Scintillation counting is by no means the only application of photomultipliers. Many applications exist in which the photomultiplier tube itself is used as a low light-level detector. The brief survey given here is meant only to convey an impression of the diversity of applications in which, because of their large-area sensitivity, gain, and time characteristics, photomultipliers continue to be the detectors of choice – or are but slowly yielding to new solid-state rivals.

8.1 Analytical applications

8.1.1 *Bioluminescence*

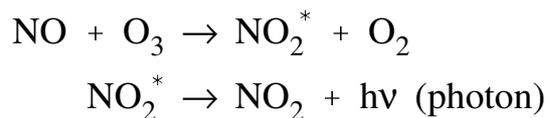
The availability of suitable reagents has stimulated interest in bioluminescence as an analytical technique. Bioluminescence assay of adenosine triphosphate (ATP), for example, is possible using the luciferin-luciferase ('firefly') reaction.

Because of the extremely low light levels produced in such reactions, single photon counting is required and signals have to be integrated over intervals of some seconds; the reaction can be followed by comparing the luminous activity during successive intervals. Tolerable noise levels are so low that photomultipliers often have to be specially selected to satisfy specifications calling for less than a few hundred counts per second above a threshold corresponding to a fraction of a photoelectron. Moreover, since the spectrum of bioluminescence is between 460 nm and 580 nm, maximum green sensitivity combined with low noise is required. Since these are conflicting requirements, a fine balance has to be found between them in the preparation of the bialkali cathodes to assure optimum performance. Finally, to provide sufficient gain, tubes with 11 or 12 stages are needed.

8.1.2 *Chemiluminescence*

Although the light produced by a chemiluminescent reaction is characteristic of the reaction, the spectrum is wide (Fig.8.1) and not easy to analyse to identify the reacting substances. Analysis is therefore based on prior choice of a reagent known to chemiluminescence with the substance sought and optical filtering to limit the detector response to an analytically significant part of the spectrum.

Chemiluminescence has been widely used for qualitative and quantitative analysis of air pollution. The concentration of ozone, for example, can be measured by its reaction with ethylene or with rhodamine B; the latter gives high luminous efficiency but depends on the surface state of the rhodamine. Nitric oxide can be measured by its reaction with ozone,



As with the NO_2^* , molecules can also lose their excitation energy by collision with other molecules, without radiation; the reaction is carried out at a pressure of only a few hundred pascals (Fig.8.2). The intensity of the luminescence is proportional to the concentration of NO; using photomultipliers it is possible to measure concentrations as small as a few parts per billion (ppb = 10^{-9}). In air pollution assays, the total concentration of NO plus NO_2 can be measured by first reducing the NO_2 to NO.

As the reaction chamber excludes background light, it is possible to reduce noise and increase sensitivity by synchronous detection (§5.13.2); however, it is more usual merely to minimize dark count by cooling the photocathode.

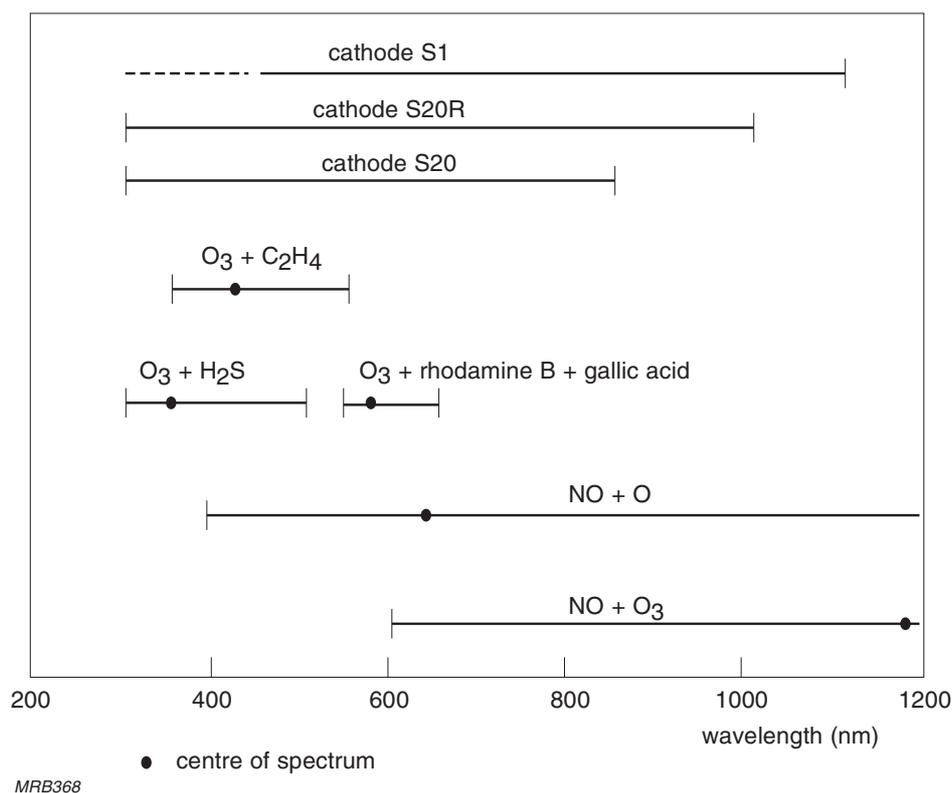


Fig.8.1 Spectra over which various chemiluminescent substances radiate compared with the spectra over which S1, S20 and S20R photocathodes respond

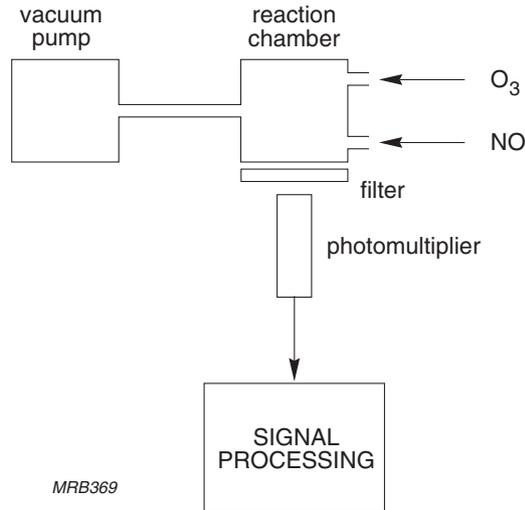


Fig.8.2 Measurement of chemiluminescence

Recently, chemiluminescence light immuno-assay (CLIA or LIA) has become a strong competitor of RIA (§7.2.2) as a sensitive analytical tool. Instead of using ^{125}I isotopes (as in RIA), CLIA uses acridinium esters which emit photons at around 430 nm when they oxidize. This wavelength is ideal for detection by standard SbKCs low-noise photomultipliers.

8.1.3 Flame spectrometry

Flame spectrometry can be based on either emission (bright-line) or absorption (dark-line) spectra. Emission is commonly used for identifying the constituent elements of a specimen and measuring their concentrations; absorption, for measuring the concentrations of elements already known, or suspected, to be present in a specimen.

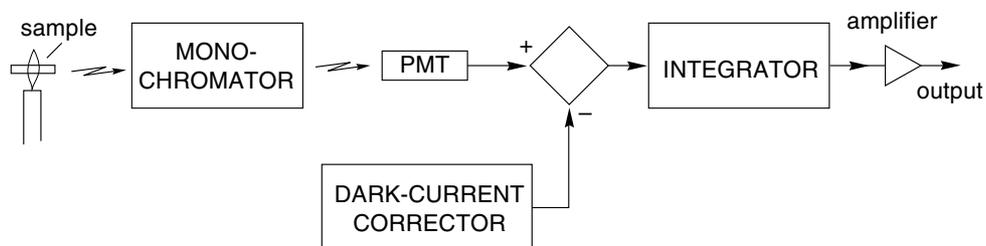


Fig.8.3 Principle of emission spectrometry

Figure 8.3 shows an emission spectrometer in which a monochromator isolates spectral lines corresponding to the constituent elements of the specimen and a photomultiplier measures the line intensities to determine the element concentrations. The spectrum due to the flame alone can be excluded by moving the specimen into and out of the flame for equal intervals and synchronously switching the gain between +1 and -1 (Fig.8.4); this also cancels the effect of the photomultiplier dark current.

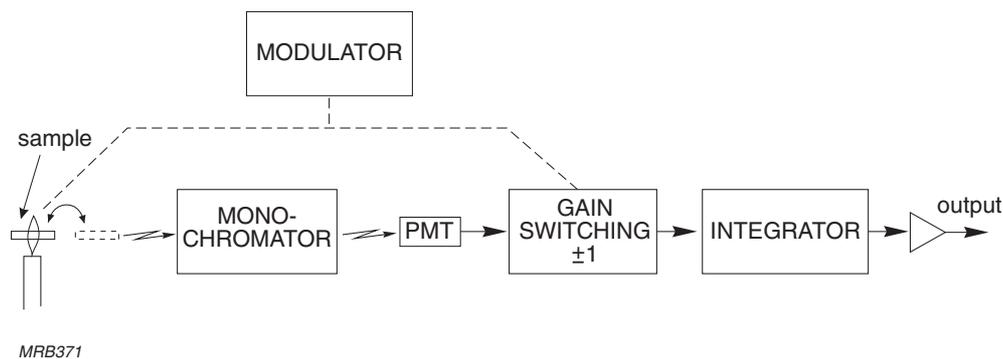
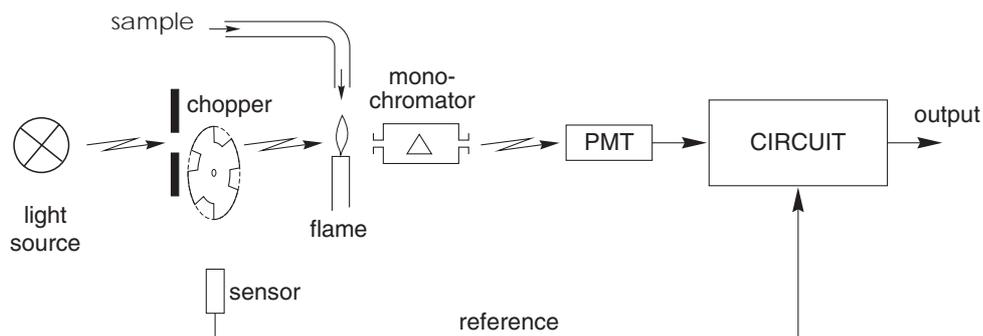


Fig.8.4 Measures for excluding dark current effects and the spectrum of the flame

Figure 8.5 shows an absorption spectrometer as used for measuring small concentrations, mainly of metallic elements, in aqueous or organic solutions (see Table 8.1). The vaporized specimen, reduced partly to the atomic state by the flame, is traversed by effectively monochromatic light of a wavelength chosen to reveal absorption by the element sought; hollow-cathode gas-discharge lamps are commonly used as light sources. By measuring the amount of absorption at each wavelength investigated, the photomultiplier determines the concentration of the corresponding elements. Reference solutions of known composition and concentration are used to calibrate the instrument.

As in the emission spectrometer, the spectrum of the flame itself is a disturbing factor. Its influence can be cancelled either by chopping the light from the monochromatic source and using synchronous detection, as in the arrangement shown, or by comparing measurements made with and without the specimen.



MRB378

Fig.8.5 Principle of absorption spectrometry

Table 8.1 Absorption spectrometry excitation wavelengths and detection thresholds of some elements in aqueous solution

element	wavelength (nm)	detection threshold ($\mu\text{g/ml}$)	element	wavelength (nm)	detection threshold ($\mu\text{g/ml}$)
Ag	328.1	0.002	Mo	313.3	0.04
Al	309.3	0.02	Na	589.0	0.0005
As	193.7	0.1	Nb	334.9	4
Au	242.8	0.03	Ni	232.0	0.01
B	249.8	3.0	Pb	217.0	0.02
Ba	553.6	0.03	Pd	244.8	0.03
Be	234.9	0.002	Rb	780.0	0.001
Bi	223.1	0.04	Sb	217.6	0.05
Ca	422.7	0.002	Se	196.0	0.3
Cd	228.8	0.002	Si	251.6	0.2
Co	240.7	0.01	Sn	224.6	0.05
Cr	357.9	0.005	Sr	460.7	0.006
Cs	852.1	0.01	Ta	271.5	4
Cu	324.8	0.003	Te	214.3	0.5
Fe	248.3	0.018	Ti	364.3	0.1
Hg	253.7	0.2	Tl	276.8	0.03
In	303.9	0.05	V	318.5	0.05
K	766.5	0.003	W	255.1	2
Li	670.8	0.002	Y	410.2	0.3
Mg	285.2	0.0002	Zn	213.9	0.001
Mn	279.5	0.003			

8.1.4 Raman spectrometry

Raman spectrometry is based on wavelength shifts due to inelastic scattering of photons when light passes through a transparent medium; it is a molecular phenomenon with no atomic counterpart. The observed shifts are characteristic of the molecular structure of the medium.

Because of the very low relative intensity of the scattered light, Raman spectrometry is practical only with a high-intensity light source for excitation. The appearance of industrial lasers has in recent years led to its widespread adoption as a means of measuring air pollution (Fig.8.6); Table 8.2 lists some of the lasers used, and Table 8.3 the Raman-effect wavelength shifts observable for several atmospheric gases excited by a nitrogen laser.

The amount of wavelength shift due to the Raman effect depends solely on the molecular structure of the medium and not on the wavelength of the excitation light. However, as the intensity of the light at the shifted wavelength varies inversely as the fourth power of the excitation wavelength, it is best to choose the shortest wavelength consistent with the spectral response of the detector used. To obtain a sufficiently high signal-to-noise ratio, the detector chosen is usually a high-gain photomultiplier which, for maximum sensitivity, may be operated as a photon counter (§5.13.3).

Table 8.2 Wavelengths and mean output powers of lasers used in Raman spectrometry

laser	λ (nm)	P_o mean (mW)
helium-cadmium	325	15
helium-cadmium	441.6	50
nitrogen	337.1	
ruby with frequency doubler	347.1	
argon	488	50 – 2000
argon	514.5	50 – 2000
neodymium ((YAG) with frequency doubler	530	
krypton	530.8	100 – 400
krypton	568.2	100 – 400
krypton	647.1	100 – 400
helium-neon	632.8	10 – 200
ruby (quasi continuous)	694.3	1000

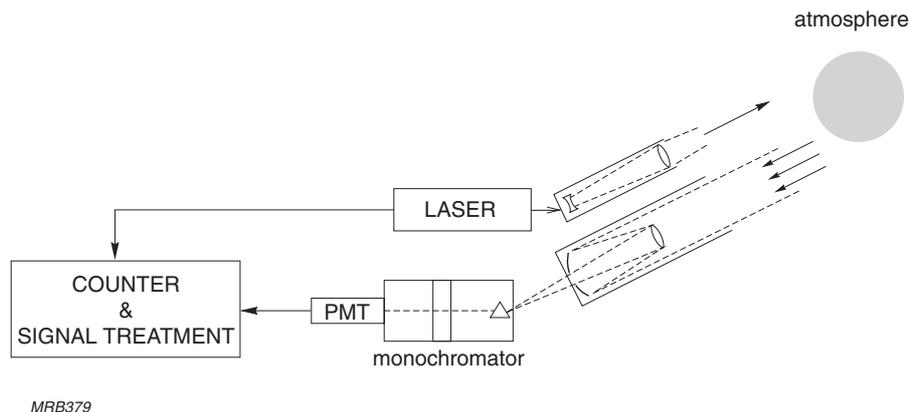


Fig.8.6 Principle of Raman spectrometry for measuring atmospheric pollution using a laser as the excitation light source

Table 8.3 Raman-effect wavelengths λ_r and wave-number shifts $\Delta\sigma$ ($\sigma = 1/\lambda$) in some molecular gases excited at a wavelength of 337.1 nm

	CO ₂ (2v ₂)	CO ₂ (v ₁)	O ₂	CO	N ₂	CH ₄ (v ₁)	CH ₄ (v ₂)	H ₂ O (v ₁)	H ₂	
λ_r	352.4	353.7	355.7	363.4	365.8	373.8	375.3	384.4	392.1	nm
$\Delta\sigma$	1285.5	1388.3	1554.7	2145.0	2330.7	2914.2	3020.3	3651.7	4160.2	cm ⁻¹

8.2 Non-analytical applications

8.2.1 Image conversion

Photomultipliers have been used for many years in the flying-spot scanners that convert photo-transparencies into video signals for television transmission (Fig.8.7). A television raster traced on the screen of a very-short-persistence cathode-ray 'flying-spot' tube is focused on the transparency to be televised. On the other side of the transparency a photomultiplier converts the transmitted light into a signal synchronous with the raster scan. Good uniformity of anode sensitivity is an important requirement, and in a three-tube scanner for colour television each tube must also have high monochromatic cathode sensitivity; in the red channel, tubes with S20 cathodes give the best signal-to-noise ratio. Another requirement is that, for adequate bandwidth (about 5 MHz), the anode load must be small and the tube must be able to operate stably at the high anode currents used. This is therefore one of the few applications where the life of the photomultiplier tube is critical and where it is quite common for the tubes to need replacing every year.

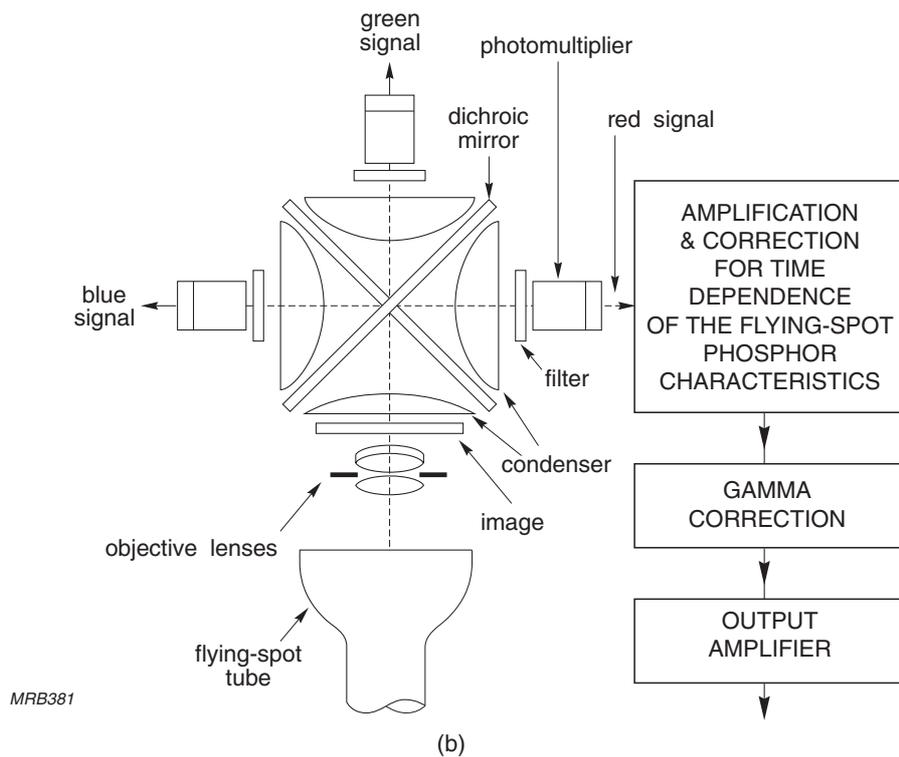
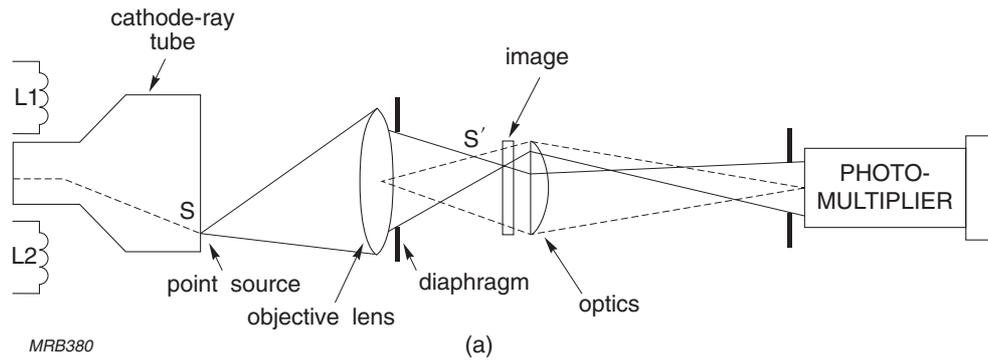


Fig.8.7 Conversion of photo-transparencies into a video image using a flying-spot scanner and photomultiplier tubes (a) monochrome system, (b) three-colour system

8.2.2 Laser scanners

Helium-neon lasers are frequently used in scanners for facsimile machines and bar-code readers, sometimes with photomultipliers as detectors.

In facsimile machines an important requirement is that the photomultiplier gain should show little hysteresis after several milliseconds' exposure to a high light level. As the facsimile scanning system often includes rapidly moving mechanical and

optical parts, high immunity to microphony is also important. These days, linear silicon diode arrays are often used in place of photomultipliers in facsimile scanners.

Large detection area, good time characteristics, and low noise made photomultipliers a natural first choice for use in scanners for reading the Universal Product Code (UPC), the bar identification code nowadays found on many packages. However, in such a scanner a photomultiplier has to draw a continuous anode current of 50 to 80 μA , which causes large gain changes (e.g. 5% to 20%) during the first thousand hours of life and a gradual degradation of gain over the next few thousand hours. Because of this the tubes had to be regularly replaced. Improvements in scanner optics have since made it possible to use large-area silicon photodiodes as detectors. However, in fast drum scanners for reproducing high-quality colour posters or scanning colour photographs for 'catalogue' presentations, the photomultiplier remains the first choice as detector.

In certain industrial scanning applications, such as the detection of surface defects in paper and rolled steel, where optical design is made difficult by the large areas to be scanned, large-diameter, red-sensitive photomultipliers are still preferred.

8.2.3 *Laser telemetry*

Distances to remote objects can be measured by timing the reflection of laser pulses from them. With electronics capable of resolving 5 ns intervals, accuracies within 75 cm can be obtained over ranges limited only by atmospheric absorption or the earth's curvature. With more precise electronics, accuracy is proportionately greater. The estimated 150 cm inaccuracy in the distance from the earth to the moon measured by laser-pulse reflection in 1969 was due largely to the experimental error, which becomes significant over such a distance, in the then accepted value of the velocity of light.

In the telemetry set-up illustrated in Fig.8.8 the transmitted laser pulse opens the timing gate and the reflected pulse closes it. The interference filter limits the photomultiplier response to the same wavelength as the laser pulse. Nevertheless, background light of that wavelength and photomultiplier dark pulses can both give rise to spurious responses. To a large extent these can be excluded by adjustment of the detection threshold. They can be still further reduced by using a coincidence detector as shown in Fig.8.9. If the pulses from both photomultipliers are shaped to a width τ corresponding to that of the laser pulses, the mean frequency of random coincidences is $n = 2n_1n_2\tau$, where n_1 and n_2 are the mean frequencies of the dark pulses from photomultipliers 1 and 2.

The frequency of random coincidences can itself be minimized by setting a range window that brackets the estimated distance to be measured. This not only improves sensitivity but also makes it possible in many cases to use low-power lasers that operate at or near the peak of the photomultiplier sensitivity curve (i.e. below 700 nm).

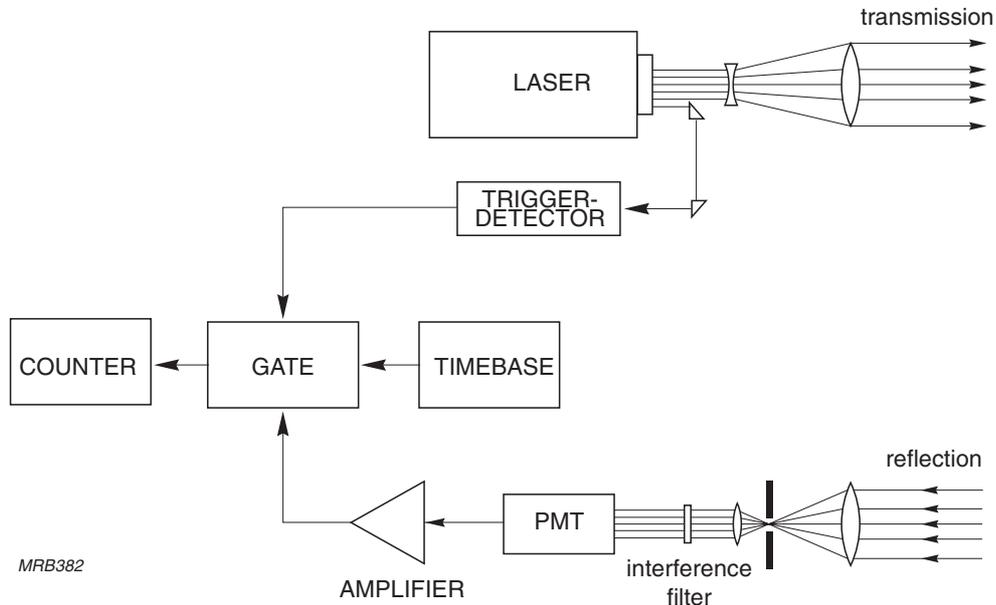


Fig.8.8 Principle of laser telemetry

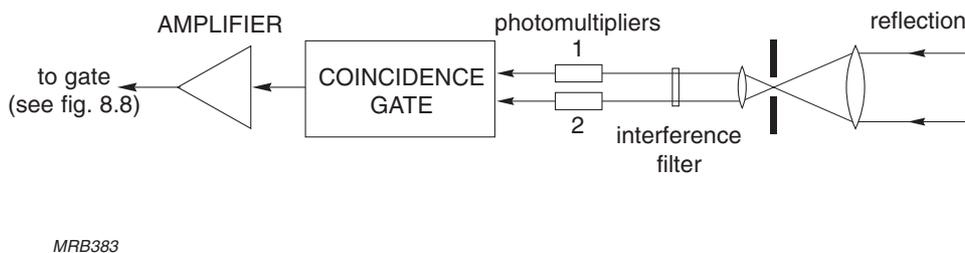


Fig.8.9 The use of a coincidence detector to reduce spurious responses in laser telemetry

Solid state lasers suitable for long-range telemetry can generate 10 ns pulses with a peak power of about a megawatt at a repetition rate of up to 50 per second. Neodymium-doped YAG and glass lasers emit at a wavelength of 1060 nm, ruby lasers at 694 nm. With the aid of frequency-doubling crystals these wavelengths can be halved to 530 nm (green) and 347 nm (near-ultraviolet) respectively, which are closer to the peak sensitivity wavelengths of most photomultipliers.

Semiconductor lasers suitable for short-range telemetry generate similar pulse widths but much lower peak powers (about a watt); their pulse repetition rates can be as high as 1000 per second.

For best resolution, the photomultipliers should be fast-response types. However, if compactness or cost is more important than resolution, as in some industrial applications, small-diameter general-purpose types will do.

An example of laser telemetry is local laser communication between buildings where the photomultiplier's good signal-to-noise ratio, large area and speed makes it the preferred detector over silicon or avalanche photodiodes.

8.3 Cherenkov radiation

Cherenkov radiation is polarized light generated when a charged particle moves through a medium faster than the speed of light in the medium. At an angle to the path of the particle the light is reinforced by constructive interference so that a conical wavefront is formed with the particle at its apex; the phenomenon is an electromagnetic counterpart of the shock wave generated by a supersonic aircraft or the bow wave of a ship. The half-angle of the conical wavefront is $\theta = \arccos(c/nv)$, where c is the speed of light in vacuum, n the refractive index of the medium, and v the speed of the particle. With the aid of photomultipliers it is possible to measure the angle of the wavefront by measuring the radius of the projected 'Cherenkov ring' (Fig.8.10) and thereby to determine the speed of the particle; if the type of particle is known, its energy can also be calculated. By suitable configuration of the optics it is also possible to make the photomultiplier respond only to particles of a specific speed. As the light pulses are practically instantaneous, Cherenkov detectors are often used as triggers in high-energy physics experiments.

The number of detected photoelectrons is given by the so-called Cherenkov radiation formula:

$$n_e = N_0 L \sin^2\theta \quad (8.1)$$

with

$$N_0 = \frac{2\pi}{137} \int_{\lambda_1}^{\lambda_2} \rho(\lambda)t(\lambda)r(\lambda) \frac{d\lambda}{\lambda^2} \quad (8.2)$$

and in which $\rho(\lambda)$ is the quantum efficiency of the counter and $t(\lambda)$ the transparency of the Cherenkov medium and $r(\lambda)$ the reflectivity of the mirrors used. If no mirrors

are used $r = 1$. L is the particle track length in the Cherenkov medium and θ is the angle of the Cherenkov radiation relative to the particle trajectory.

In practice, N_0 is around 50/cm for glass-window tubes and around 100/cm for quartz-window tubes. As $\sin^2\theta$ is small, these values for N_0 mean that gas Cherenkov counters always tend to be rather long. Also, since N_0 increases with the proportion of the UV spectrum that can be used (from Eq.8.2), Cherenkov counters tend to use fast-response quartz-window tubes, often as large as 5 inches in diameter to catch as much of the UV light as possible. Since such tubes are expensive, recent practice has been to use standard tubes optimized for visible light and to shift the UV component into the visible spectrum by covering the tube window with a wavelength shifting layer of, for example TPB (tetraphenyl-butadiene) or PTP (1,4 diphenylbenzol). Disadvantages here, however, are that the layers must be deposited under vacuum and that they are somewhat unstable.

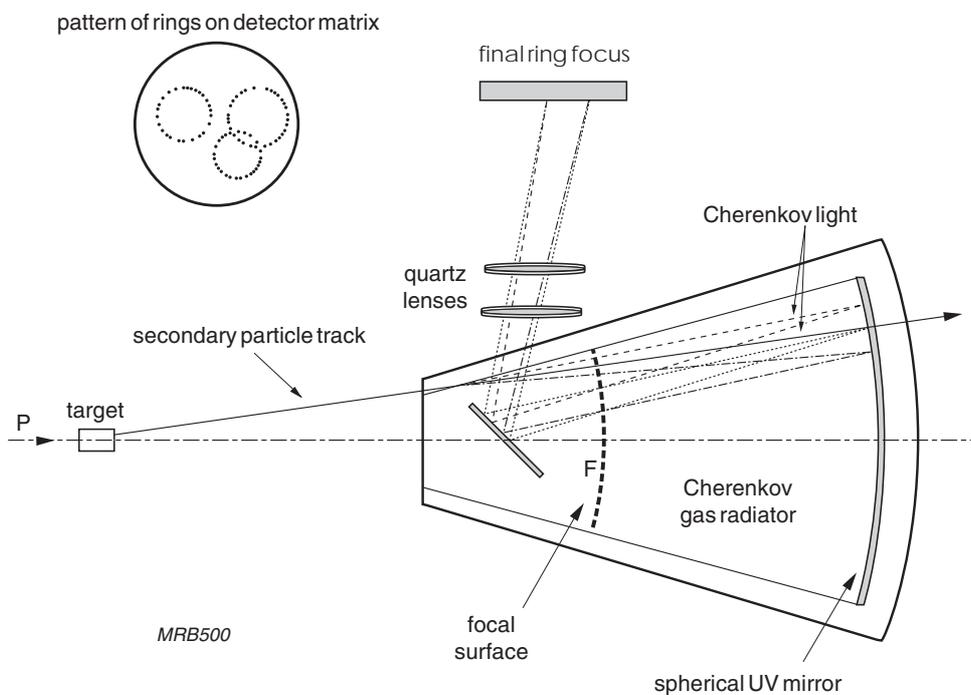


Fig.8.10 Cherenkov detector optical system in ring-focusing mode (courtesy of CERN, Geneva)

8.4 Cherenkov experiments

8.4.1 High-energy physics experiments

In high-energy physics experiments it is often of interest to discriminate particles that exceed a certain minimum speed. This can be done by choosing a Cherenkov medium, often a gas, with an index of refraction corresponding to that minimum speed. Indices as low as 1.03 can be obtained with recently developed solid materials

called *aerogels* which can be shaped into blocks. This has led to the development of large veto-counter boxes where the Cherenkov light produced by electrons, pions or kaons in about 5 cm thick aerogel blocks is reflected by the box walls to between 5 and 15 5-inch photomultipliers connected together. Such veto counters with specially developed tubes by Photonis that sacrifice speed for improved photoelectron collection efficiency have been shown to give excellent particle separation signals, enabling a sharper veto than earlier gas counter designs.

Another use of the new aerogel is in the ring imaging Cherenkov (RICH) counters used in the HERMES experiment at DESY, Hamburg, Germany, which employs four thousand 19 mm diameter UV-glass Photonis photomultipliers.

8.4.2 Proton decay experiments

In ‘swimming-pool’ experiments (like the IMB detector, USA), in which the Cherenkov medium is a large volume of clean water, thousands of 5-inch ETL and later 8-inch hemispherical Hamamatsu photomultipliers were used to look for radiation due to proton decay. This decay has been predicted theoretically but as yet not observed. The IMB experiment was able to put the limit of proton life to at least 10^{32} years, thereby dismissing the first grand unification theory uniting all the known forces (the theory that set off the first race for the discovery of proton decay).

The KAMIOKANDE experiment in Japan used a thousand 20-inch hemispherical photomultipliers in an attempt to detect proton decay. This experiment (together with the IMB experiment) recorded the majority of the neutrinos detected on earth from the supernova 1987A*. The follow-up on the water-Cherenkov experiment, the SUPER KAMIOKANDE, used twelve thousand 20 inch Hamamatsu photomultipliers and had moved the limit of proton decay to more than 10^{33} years before an unfortunate accident caused a large number of the tubes to implode. The plan is to restore the experiment within the coming years. This, however, is not enough since one way to confirm modern physical theories would be to find proton decay existing (or not) in nature with a lifetime below 10^{35} years. Considerably larger experiments are therefore being proposed (HYPER KAMIOKANDE in Japan and UNO in the US) with more than 50 000 large photomultipliers each.

8.4.3 Solar neutrino experiments

The solar-neutrino flux has been found, by a massive thirty-year long deep-mine chlorine-to-argon solar neutrino-exchange-induced experiment, to be lower than

*Shared Nobel Prize in physics 2002

theoretically predicted* and several experiments have been designed to investigate the reason for this, which is supposed to be caused by the solar electron neutrino being changed into another ‘neutrino’ flavour, a phenomenon called ‘neutrino oscillation’.

The Solar Neutrino Observatory (SNO) in the Sudbury mine, Canada, makes use of around ten thousand 8-inch hemispherical Hamamatsu photomultipliers using heavy water (D₂O) as the Cherenkov medium. Recently published results indicate that the observed solar neutrino deficit indeed has its cause in neutrino oscillations.

The results are supposed to be confirmed by the boron-doped liquid-scintillator BOREXINO experiment in the Gran Sasso tunnel complex in Italy. This makes use of about fifteen hundred 8-inch ETL photomultipliers that start to collect data in 2003.

8.4.4 Neutrino oscillation experiments

To confirm that the neutrino flavour change (oscillations) is genuine, several large experiments have been built offering conflicting results. To confirm the debated results of a former liquid scintillator experiment in Los Alamos (LSND), a larger experiment, known as MiniBooNE at Fermilab, is just now starting to collect data. This makes use of about fifteen hundred 8-inch Hamamatsu photomultipliers. If the LSND results are confirmed, it will mean that there exists more than the three neutrino flavours of the standard model and will stimulate a need for an even larger experiment, i.e. BooNE, with more photomultipliers.

Two large experiments are already in construction, one in a mine in Minnesota (MINOS) as the target of a neutrino beam from Fermilab, and the other experiment in Gran Sasso, Italy, (OPERA), as the target of a neutrino beam from CERN, Geneva. Both will make use of emulsion detectors and plastic scintillators coupled via wavelength-shifting fibres to multichannel photomultipliers from Hamamatsu.

SUPER KAMIOKANDE, before its collapse, was already used, and will be used again, as the target of a neutrino beam from the Japanese KEK accelerator. It also has given convincing results in the study of atmospheric neutrinos, confirming that neutrino oscillations really occur in nature.

*Shared Nobel Prize in physics 2002

Since nuclear reactors are also a free source of neutrinos, several experiments have attempted to find neutrino oscillations by placing detectors at different distances from reactors, although with less convincing results so far. The hope is now that the KamLAND experiment, the old Kamioka detector filled with a liquid scintillator and using thirteen hundred 17-inch and seven hundred 20-inch Hamamatsu photomultipliers, will give convincing oscillation results. Because of its location, the experiment is sensitive to the reactor-generated neutrinos from 16 reactor sites (51 reactors) in Japan and 4 sites (18 reactors) in South Korea.

8.4.5 DUMAND* experiments

In the DUMAND Lake Baikal experiment, several hundred ‘smart’ 16-inch Russian photomultipliers are used on the bottom of the lake at a depth of 1400 m. This spectacular Cherenkov experiment will search for a theoretically predicted particle called the *magnetic monopole* (which has been proposed as the mechanism that triggers proton decay).

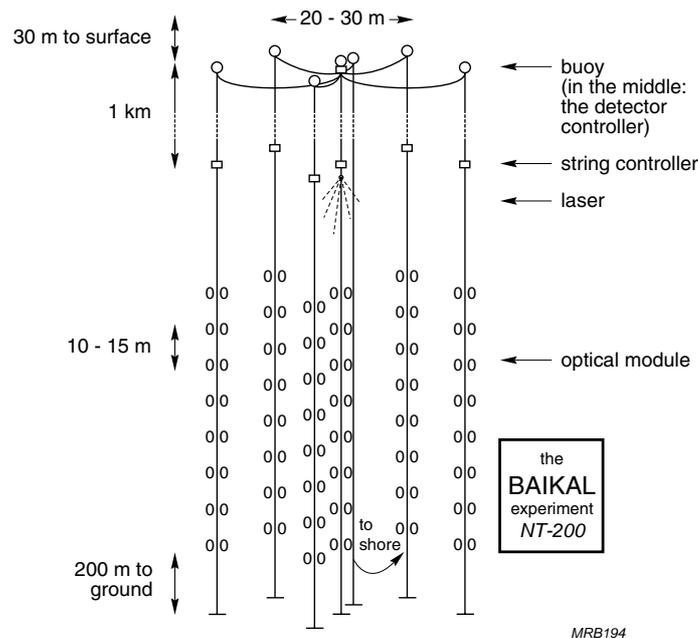


Fig.8.11 DUMAND Lake Baikal preliminary experiment (courtesy of INR, Moscow)

A similar arrangement on an even larger scale was planned for the DUMAND Hawaii experiment, where a cubic kilometre of ocean water was to be observed at 5 km depth to study possible galactic neutrino point sources. Although this DUMAND pioneer experiment was unfortunately cancelled, the extensive experience gained of

* Deep Underwater Muon And Neutrino Detector

The Photonis 15-inch ‘smart’ photomultiplier tube (see §1.5.6), which is capable of detecting low level multi-electron signals in a high single-photoelectron background environment, has been designed specifically for the Lake Baikal and Hawaii experiments. A Russian copy of the Photonis 15-inch ‘smart’ tube has also been developed for the Lake Baikal experiment. Although these ‘smart’ tubes are no longer in production, there is renewed interest in such tubes for extending the Baikal, NESTOR and NEMO experiments. Several new ideas are even now on the drawing boards of the photomultiplier suppliers to optimise the tubes for the new demands and for reducing costs.

8.4.6 Air-shower experiments

Cosmic-ray air showers, created when cosmic particles interact with the earth’s atmosphere at high altitude, are another source of Cherenkov radiation. In an extension to the HEGRA experiment on La Palma island, the Cherenkov radiation has been observed using a one square kilometre array of large hemispherical ETL photomultipliers combined with mirrors directed at the night sky (so called Aerobic detectors interest in which is also growing for an installation in India).

The EGRET experiment on board the Compton Gamma Ray Observatory satellite has demonstrated the existence of many γ -point sources with energies up to around 1 GeV. γ -ray telescopes are currently only able to detect γ -sources with energies above 1 TeV, and the present challenge for users of γ -ray telescopes is to extend their sensitivity to be able to detect γ -energies down to 1 GeV. There are therefore currently three new γ -ray telescope clusters in preparation:

- the HESS cluster of 16 telescopes on a high-plateau in Namibia, the first four using 960 29-mm-diameter low-afterpulse Photonis photomultipliers each
- the MAGIC planned cluster of three very large telescopes, the first one using 25 mm diameter hemispherical photomultipliers from ETL. This first telescope is now under construction at the HEGRA site at the La Palma island
- the VERITAS cluster of 7 telescopes. This is an extension of the Whipple site currently being prepared in the Arizona desert in which each telescope will use five hundred 29 mm diameter low-afterpulse UV-sensitive photomultipliers.

There exist spectacular cosmic phenomena seen by satellite and γ -ray telescopes, called gamma-ray bursts (GRB) – very-energetic, very-short-lived observations indicating an enormous cosmic energy “burst” of unknown origin. Further experiments are already being proposed to help investigators understand these GRBs.

Outside Salt Lake City, Utah, the Fly's Eye experiment has for the past few years been running the two so called HiRes eyes, providing a stereoscopic view from two mountain tops using 22 and 42 mirrors respectively, each focusing onto an array of two hundred and fifty six 39 mm hexagonal Photonis photomultiplier tubes. The experiment makes use of the nitrogen fluorescence UV lines stimulated by the secondary particles of the cosmic-ray air showers. Together with the older Japanese air-shower experiment, AGASA, the HiRes has recorded cosmic-ray events with an energy above 10^{20} eV, the highest cosmic ray energy ever recorded. This is in conflict with the GZK theory stating that events above about 10^{19} eV should not exist, the particle's energy supposedly being limited by interaction with the microwave cosmic ray background. To get more statistics to determine if such events are genuine, a third HiRes stereo-viewing eye on another mountain top with another 5 to 8000 tubes is planned.

This puzzle of the unexpected 'ankle' in the cosmic-ray spectrum has initiated the building of a very large air-shower detector at Malargüe, Argentina (see inside of the back cover), named AUGER after the first cosmic-ray air-shower experimentalist Pierre Auger who studied these showers using Geiger-Mueller tube coincidence signals in 1938. It consists of sixteen hundred 3 meter-diameter energy-autonomous water tanks, spread out over a very large area on the high-elevation pampas plateau. Each tank makes use of three Photonis 9-inch photomultipliers to detect the water-Cherenkov radiation originating from the air-showering particles. Complementing these are four fly's eye fluorescence detectors with more than 10 000 Photonis 38 mm hexagonal photomultipliers (as with the HiRes) to penetrate the atmosphere. If this "AUGER-South" detector is successful in detecting some high-energy ($>10^{20}$ eV) cosmic-ray events, the plan is to build a similar "AUGER-North" detector close to the HiRes in Utah, possibly making use of the HiRes as a fluorescent detector.

8.4.7 Towards a common goal

The DUMAND experiments are generally aimed at detecting very-high-energy neutrino cosmic point sources and the γ -ray experiments are looking for very-high-energy γ -ray bursts and γ -ray point sources with energies down to below 10 GeV. The air-shower experiments are looking for very-high-energy cosmic air-shower events. Although these are all different detection techniques, the common goal is to determine the nature of the gravitational energy-exchange processes believed to occur on a range of scales from binary-star systems and supernovae through to galactic nuclei, or to detect manifestations of cosmic forces and phenomena as yet unknown to mankind.

Finally, the apparent confirmation of neutrino oscillations implies that at least two of the traditional three neutrinos have mass. It also provides insight into the related challenge of understanding the possible charge-parity (CP) violation in the neutrino sector. This opens up a whole new aspect of neutrino physics – and the future need for new, extremely large experiments with very many photomultipliers. Owing to their large volume, these can also serve as proton-decay detectors.

8.4.8 A vital tool for physics

Cosmic-ray studies are not only opening up new and exciting areas in physics, they also provide a strongly growing market for photomultipliers whose combined traditional features of:

- large area coverage
- excellent time resolution
- low noise
- attractive price per square metre of detection

make them the only real choice for cosmic-ray research.

These combined features are, moreover, the principal reason why photomultipliers rapidly became the most widely used detector for early nuclear physics experiments, and it is also why they are likely to be the detectors of choice for the future generations of large astro-particle physics experiments in our new century. What is more, there is little doubt that all these new, very large experiments on the horizon will stimulate the development of ever larger tubes with lower noise (through ‘smartness’), higher quantum efficiency, pixelisation ‘smartness’ and, most important, low price per square metre of sensitive area.

The success of the photomultiplier over the past fifty years in meeting the challenges of the scientific community will doubtless continue into the future with completely new hybrid and ‘smart’ designs as yet only at the embryonic stage. No other detector known to mankind can match the photomultiplier’s sensitivity in detecting the fundamental particles that make up our physical world. No other detector is capable of penetrating so deeply into nature, and of providing us with such valuable insight into the nature of our universe.

Acknowledgements

We wish to thank the following persons for checking the accuracy of the details in this chapter and for all their enthusiastic and constructive comments:

John Learned, University of Hawaii, Honolulu, Hawaii, USA

John Matthews, University of Utah, Salt Lake City, Utah, USA

Peter Bosetti, RWTH Aachen, Germany and the Foundation Vijlen Institute for Physics, The Netherlands