József Mika and Tibor Török

ANALYTICAL EMISSION SPECTROSCOPY



Akadémiai Kiadó, Budapest

J. Mika and T. Török

ANALYTICAL EMISSION SPECTROSCOPY

FUNDAMENTALS

Spectroscopic techniques have become important for determining a large number of chemical elements in a wide range of materials both solid and liquid. Optical emission spectroscopy can readily analyse most of the elements by modifying excitation procedures, optical systems and detectors. By means of special enrichment procedures very low detection limits are obtained. This capability makes emission spectroscopy a very powerful method with many of the advantages of similar techniques such as X-ray fluorescence spectroscopy. Recent developments, for example, involving the generation of emission by laser beams have added a new dimension which promises to find wide application in geochemical and metallurgical analysis.

□

This volume deals with the fundamentals of optical emission spectroscopy. The basic theory of spectral analysis is followed by a detailed account of the various methods of excitation, including arc and spark sources and their control. Resolution of the resulting radiation by both prism and grating systems as well as interference devices are described in detail with sections on the efficiency and design parameters of various systems. A vital aspect of emission spectroscopy is the detection of radiations; methods of intensity measurement for quantitative analysis are discussed in detail.

□

This volume should be indispensable to the analyst and research worker in many laboratories involved in trace element analysis. Chemists and biochemists, metallurgists and geochemists, public analysts and forensic scientists will find a great deal of information which is not readily available elsewhere.

AKADÉMIAI KIADÓ · BUDAPEST

Analytical Emission Spectroscopy

Fundamentals

Analytical Emission Spectroscopy

Fundamentals

by

József Mika and Tibor Török



Akadémiai Kiadó • Budapest 1973

This book is the enlarged English version of
Emissziós színképelemzés
published by Akadémiai Kiadó, Budapest

Translated by
László Nemes

English translation edited by
P. A. Floyd
University of Keele

© Akadémiai Kiadó, Budapest 1973

Joint edition published by
Akadémiai Kiadó, Budapest
and
Butterworths, London

Printed in Hungary at Akadémiai Nyomda

Preface

The subject of the present book is *optical* emission spectrochemical analysis. The flame-photometric method is not dealt with here since this technique nowadays constitutes the basis for a separate branch of analytical science.

The purpose of our work is to provide help for those who intend to deal with emission spectrum analysis in detail, and also to facilitate the work of those who devote their attention mainly to the practical applications of such procedures. The main emphasis of our work is a systematic and detailed treatment of the theoretical basis of excitation, light dispersion, and the determination of the wavelength and intensity of spectral lines. We have tried to limit the treatment of the fundamentals of theoretical electricity, atomic theory, optics, etc., to those aspects that are indispensable for the understanding of physical and chemical phenomena occurring in spectrochemical analysis.

The limitations in the coverage have rendered it impossible to discuss all known procedures, or to provide an exhaustive literature list for them. In excluding the obsolete methods from our scope we have aimed at a discussion of the essentials, advantages and drawbacks of the presently used methods. Also, we have tried to cover such procedures as can only be used in exceptional cases, or which have not found widespread application, if, in our opinion, they nevertheless have important or instructive features. The references listed contain those sources that were used in writing the book, and also examples of monographs and general works on the subject that represent more detailed studies.

Since the publication of the Hungarian edition the original material of the book has been revised. This work involved the correction of some

mistakes and several chapters have been rewritten to a greater or smaller extent. It is hoped that these alterations have contributed to the usefulness of the book.

Our thanks are due to Dr. P. A. Floyd for his very careful work in revising the text. We are grateful to E. Gegus, chemical engineer and Professor K. Zimmer for their great help in finding the mistakes which occurred in the Hungarian edition. Finally we should like to thank Mrs. E. K. Kállay-Tóth for her careful editorial work.

József Mika
Tibor Török

Contents

1 Introduction 13

1.1 Optical spectra 13

1.1.1 Classification of spectra 13

1.2 Structure of line spectra 14

1.2.1 Properties of line spectra 14
1.2.2 Origin of atomic spectra 15
1.2.3 Spectral line series 17
1.2.4 Term schemes 20
1.2.5 Systems having one valency electron 22
1.2.6 Structure of atomic spectra having more than one valency electron 29
1.2.7 Ion spectra 34

1.3 Absolute and relative intensity of spectral lines 37

1.3.1 Absolute intensity of spectral lines 37
1.3.2 Relative intensity of spectral lines 38
1.3.3 Relationship between line intensity and concentration 40

1.4 Band spectra and continuous spectra 42

1.4.1 Structure of band spectra 42
1.4.2 Origin of molecular spectra 43
1.4.3 Role of band spectra in emission analysis 46
1.4.4 Continuous spectra 46

2 Excitation 49

2.1 Basic principles 49

2.1.1 Requirements of excitation methods 49
2.1.2 Light sources for emission analysis 50

2.2 The arc 51

2.2.1 Fundamental properties of the arc 51
2.2.2 The arc plasma 53
2.2.3 Migration of particles in the arc plasma 57

2.2.4 Phenomena occurring at the electrodes of the arc 64
2.2.5 Light emission of the arc plasma 67
2.2.6 Relationship between the temperature of the arc plasma and the degree of ionisation 69
2.2.7 Measurement of the plasma temperature by spectra 71
2.2.8 Arc temperature and the first ionisation potential of the elements 73
2.2.9 Simultaneous element excitation 75
2.2.10 Contours of spectral lines 78
2.2.11 Self-absorption 82
2.2.12 Background radiation 85

2.3 Electric parameters of arcs 88
2.3.1 Arc voltage and arc current 88
2.3.2 Breakdown voltage 92
2.3.3 Effect of alternating current 95

2.4 Arc generators 97
2.4.1 Classification of arc generators 97
2.4.2 Power supplies for arc generators 98
2.4.3 Arc ignition 99
2.4.4 Direct current continuous-arc generators 105
2.4.5 Mains voltage a.c. arc generators 108
2.4.6 High-voltage a.c. arc generators and high-voltage rectified arc generators 108
2.4.7 Mechanically controlled, interrupted-arc generators 111
2.4.8 Electronically controlled a.c. arc generators 113
2.4.9 Arc generation by high-frequency current 114
2.4.10 Excitation by d.c. plasma jet 114
2.4.11 Excitation by high-frequency plasma jet 119

2.5 Spark excitation 122
2.5.1 Spark discharges 122
2.5.2 Determination of the temperature of the spark plasma 131
2.5.3 Basic principles of spark generators 134

2.6 Processes taking place in the discharge circuit of spark generators 135
2.6.1 Preliminary remarks 135
2.6.2 Discharge of a condenser 136
2.6.3 Periodically damped discharge of a condenser 138
2.6.4 Over-damped and critically damped discharge of a condenser 145
2.6.5 The influence of the parameters of the discharge circuit 148

2.7 Electric phenomena occurring in the charging circuit of a condenser 152
2.7.1 Preliminary remarks 152
2.7.2 Condenser charging via a transformer 153
2.7.3 Direct condenser charging via the mains current 159
2.7.4 Influence of the damping resistance 160

2.8 Principal design of high-voltage spark generators 164
2.8.1 Classification of spark generators 164
2.8.2 High-voltage, non-controlled spark generators 165
2.8.3 High-voltage spark generators having a controlling spark gap 168
2.8.4 Mechanically controlled, high-voltage spark generators 169
2.8.5 High-voltage spark generators with a rectified charging circuit 174
2.8.6 Electronically controlled, high-voltage spark generators 177
2.8.7 Pulse-spark generators 179

2.9 The principal design of low- and medium-voltage spark generators 179
2.9.1 General remarks 179
2.9.2 Medium-voltage spark generators 180
2.9.3 Mains-voltage spark generators 183

2.10 Excitation in discharge tubes 187

2.10.1 Fundamentals of the method 187
2.10.2 Excitation in Geissler tubes 192
2.10.3 Excitation by hollow-cathode tubes 195
2.10.4 Excitation in flash tubes 202
2.10.5 High-frequency excitation at reduced pressures 204

2.11 Excitation by laser 208

2.11.1 Fundamental principle of the procedure 208
2.11.2 Solid-phase lasers used for excitation 210
2.11.3 Application of lasers 215

3 Dispersion of light 221

3.1 Classification of procedures used for the dispersion of light 221

3.1.1 Different dispersion methods 221
3.1.2 Performance of dispersion devices 222

3.2 Fundamental principles of prism dispersion 224

3.2.1 Refraction of light 224
3.2.2 Prism dispersion 227

3.3 Fundamental principles of prism dispersion instruments 233

3.3.1 The construction of prism dispersion instruments 233
3.3.2 The slit 235
3.3.3 The prism 237
3.3.4 The collimating lens 245
3.3.5 The camera lens 251

3.4 Performance of prism dispersion instruments 253

3.4.1 Linear dispersion and dispersion curve of prism spectrographs 253
3.4.2 Theoretical resolving power of prism spectrographs 254
3.4.3 Practical resolution of prism spectrographs 256
3.4.4 Photographic speed of prism spectrographs 261

3.5 Error sources of prism spectrographs 262

3.5.1 Main sources of error 262
3.5.2 Consequences of light absorption 262
3.5.3 Consequences of light reflection 264
3.5.4 Effect of stray light 267
3.5.5 Effects of temperature and barometric pressure changes 271

3.6 Prism dispersion instruments 274

3.6.1 Single-prism quartz spectrographs 274
3.6.2 Autocollimating quartz spectrographs 276
3.6.3 Multiprism quartz spectrographs 279
3.6.4 Low- and medium-performance glass spectrographs 281
3.6.5 High-performance glass spectrographs 284
3.6.6 Liquid-prism spectrographs 286
3.6.7 Prism vacuum spectrographs 287
3.6.8 The Fery spectrograph 290
3.6.9 Spectroscopes 292

3.7 Illumination of prism spectrographs 296

3.7.1 General remarks 296
3.7.2 Direct illumination of spectrographs 296
3.7.3 Illumination of spectrographs by a collecting lens 298

3.7.4 Illumination of spectrographs with intermediate focusing 301
3.7.5 Illumination of spectrographs by a concave metal mirror 305
3.7.6 Time-resolved illumination of spectrographs 307
3.7.7 Attenuation of the illumination of spectrographs 312

3.8 Basic principles of grating dispersion 318

3.8.1 Gratings 318
3.8.2 Basic principles of plane-grating dispersion 318
3.8.3 Basic principles of concave-grating dispersion 326
3.8.4 Dispersion of grating spectrographs 332
3.8.5 Resolving power and optimum slit width of grating spectrographs 334
3.8.6 Photographic speed of grating spectrographs 336
3.8.7 Common diffraction gratings 336

3.9 Plane-grating spectrographs 340

3.9.1 Construction of plane-grating spectrographs 340
3.9.2 Basic principle of plane-grating spectrographs 341
3.9.3 Czerny–Turner-system plane-grating spectrographs 343
3.9.4 Ebert-system plane-grating spectrographs 344
3.9.5 High-performance Jacob–Ebert-system plane-grating spectrographs, illuminated by a predispersion unit 347

3.10 Concave-grating spectrographs 350

3.10.1 Basic principle of the instrument 350
3.10.2 Different astigmatic concave-grating spectrographs 353
3.10.3 Paschen–Runge-system spectrographs 355
3.10.4 Eagle-system spectrographs 356
3.10.5 Concave-grating spectrographs with stigmatic line formation 358
3.10.6 High-performance spectrophotometers 361
3.10.7 Concave-grating vacuum spectrographs with normal spectrum 362
3.10.8 Grating vacuum spectrographs illuminated by grazing incidence light 364

3.11 Errors of grating spectrographs 367

3.11.1 The overlap between spectra of different orders 367
3.11.2 Ruling errors of diffraction gratings 369
3.11.3 Light losses and scattered light 371
3.11.4 Errors caused by external experimental conditions 371

3.12 Light dispersion instruments used for the study of the fine structure of spectra 372

3.12.1 General remarks 372
3.12.2 The echelle spectrograph 373
3.12.3 Combination of the Fabry–Perot interferometer and the spectrograph 376

3.13 Dispersion of light by filters 380

3.13.1 General considerations 380
3.13.2 Absorption filters 382
3.13.3 Interference filters 383

4 Determination of the wavelength 385

4.1 Determination of the wavelength of optical radiation 385

4.1.1 Fundamentals 385
4.1.2 Determination of the wavelength by linear interpolation 388
4.1.3 Determination of the wavelength by the dispersion curve 389
4.1.4 Determination of the wavelength by the Hartmann formula 392
4.1.5 Determination of the wavelength using the normal spectrum of grating spectrographs 393

4.2 Optical devices used for the determination of wavelength 394
4.2.1 General considerations 394
4.2.2 Spectrum projectors 395
4.2.3 Measuring microscopes 398

5 Measurement of the intensity of spectral lines 401

5.1 Fundamentals of light-intensity measurements 401
5.1.1 Basic principles 401
5.1.2 Procedures used for the determination of the intensity of spectral lines 405

5.2 Fundamentals of photographic radiation intensity determination 409
5.2.1 Fundamentals of photography 409
5.2.2 The photographic characteristic curve 411
5.2.3 Designation of the sensitivity of photographic emulsions 417

5.3 Construction of the characteristic curve 419
5.3.1 Preliminary remarks 419
5.3.2 Construction of the characteristic curve using multiplet lines 420
5.3.3 Construction of the characteristic curve using a multi-step attenuator 423
5.3.4 Construction of the characteristic curve using a single pair of multiplet lines 424
5.3.5 Procedures based on the two-step attenuation of light 426

5.4 Photographic emulsions used for intensity measurements 430
5.4.1 Spectral plates and films 430
5.4.2 The halo effect 431
5.4.3 Spectral sensitivity of photographic emulsion 432
5.4.4 Photographic emulsions used for spectral analysis 435

5.5 Physical conditions influencing blackening or density 436
5.5.1 Deviation from the reciprocity law 436
5.5.2 The intermittency effect 438
5.5.3 Effect of the background 440
5.5.4 Callier effect 441
5.5.5 Graininess, resolving power, and emulsion micro-errors 442

5.6 Chemical conditions influencing the blackening or density 443
5.6.1 Developing 443
5.6.2 Effect of the composition of the developer 445
5.6.3 Influence of the duration and temperature of development 447
5.6.4 Agitation of the developer and emulsion macro-error 449
5.6.5 The Eberhardt effect 452

5.7 Straightening of the characteristic curve 454
5.7.1 General considerations 454
5.7.2 The *W*- or Seidel transformation 455
5.7.3 The $P_{1/2}$ transformation 456
5.7.4 The generally valid partial density transformation 458
5.7.5 The *I*-transformation 459

5.8 Background correction 462
5.8.1 General considerations 462
5.8.2 Background correction by calculation 463
5.8.3 Determination of the difference between the apparent density of the spectral lines and the background by the s.p.d. scale 464

5.9 Determination of the density of spectral lines 465
5.9.1 Fundamental principles of the procedure 465
5.9.2 Construction of spectral-line photometers 467

5.10 Fundamentals of photoelectric intensity measurements 473

5.10.1 External photoelectric effect 473
5.10.2 Photocells 476

5.11 Photomultipliers 481

5.11.1 Fundamental principles 481
5.11.2 Construction of photomultipliers 483
5.11.3 Dark current of photomultipliers 488
5.11.4 Shot effect of photomultipliers 489
5.11.5 Principal faults of photomultipliers 490
5.11.6 Power supplies for photomultipliers 492

5.12 Application of photomultipliers 495

5.12.1 Fundamental principles 495
5.12.2 Principle of integrating-condenser measurements 496
5.12.3 Principle of the graphical method 498

5.13 Visual intensity measurements 499

5.13.1 Fundamental principles of the method 499
5.13.2 Measurement by the grey-wedge method 501
5.13.3 Measurement by polarisation techniques 504
5.13.4 Auxiliary instruments used for visual intensity measurements 505

Author index 511

Subject index 517

1 Introduction

1.1 Optical spectra

1.1.1 Classification of spectra

Optical radiation is the name given to that part of the electromagnetic spectrum whose wavelength (λ) lies between 10^{-6} and 10^{-2} cm, or, given in the generally accepted unit of spectral analysis, between 10^{2} and 10^{6} Å. Radiation having a shorter wavelength belongs to the domain of X-rays and γ-rays, while radiation of longer wavelength is found in the region of microwaves and radiowaves.

From the point of view of spectrochemical analysis the important part of optical radiation is found between 100 and 12 000 Å. This radiation can be approximately divided into several spectral regions, as shown in Table 1.1.

Table 1.1

Designation of optical spectral regions

Wavelength limits approx. (Å)	Name	Wavelength limits approx. (Å)	Name
100–2 000	vacuum ultra-violet	100–1 200	Schuman
		1 200–2 000	Lyman
2 000–4 000	ultra-violet		
		4 000–4 500	violet
		4 500–5 100	blue
4 000–7 600	visible	5 100–5 900	green
		5 900–6 400	yellow
		6 400–7 600	red
7 600–12 000	photographic infra-red		

Radiation from different sources is usually polychromatic, that is, it is composed of radiation of different wavelengths. The spectral density of a light beam at wavelength λ is $I(\lambda)$ and is usually called the absolute intensity of the 'spectral' line. Thus the total output of the source is given by

$$\Phi = \int_0^\infty I(\lambda)\mathrm{d}\lambda \tag{1.1}$$

while the *relative line intensity* is the following expression

$$I = \frac{I(\lambda)}{\Phi} \tag{1.2}$$

Accordingly, a diagram showing the variation of the wavelength (abscissa), and the relative line intensity (ordinate) is called the *spectrum* of the radiation. In a general sense the spectrum defines the intensity value for a given wavelength.

Depending on the distribution of the relative line intensities in the spectrum of the radiation, there may be line spectra, band spectra, and continuous spectra present.

The radiation spectrum depends on the source and on the absorption properties of the medium traversed by the radiation. In the case of emission spectra only the spectral composition of the radiation from the source is important. In emission studies it is therefore important to make sure that the absorptive medium does not alter the composition of the emitted radiation by selective absorption.

1.2 Structure of line spectra

1.2.1 Properties of line spectra[1,2]

Line spectra are radiated from relatively low-density glowing gases or vapours in the atomic or ionic state.

Line spectra, according to their broader definition, consist of narrow, bell-shaped lines at certain wavelengths (Fig. 1.1*a*) between which the radiation intensity is small (see Fig. 1.1*b*).

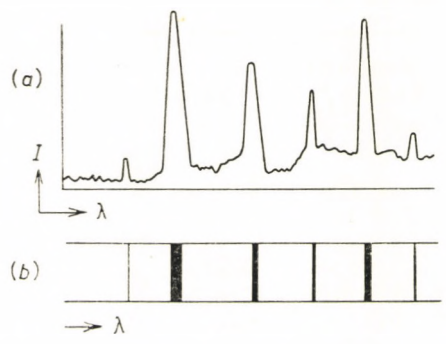


Fig. 1.1 Line spectra: (*a*) intensity distribution, (*b*) photographic spectrogram (diagrammatic)

Generally, line spectra are composed of many lines distributed in an apparently chaotic manner. The number and intensity of the spectral lines in the spectrum of a glowing gas or vapour depend on the experimental conditions and in particular

on the temperature of the gas. The higher the temperature, the greater will be the degree of ionisation of the atoms and, therefore, the greater will be the number of lines corresponding to both the ions and the atoms present. Accordingly, a distinction is made between 'atomic spectra' and 'ion spectra', although no clear-cut division can be established between these two types as the transition between them is continuous. The former are obtained generally by arc excitation, while the latter are produced mainly by spark excitation. For this reason the older and not quite precise terminology for these two types was 'arc spectrum' and 'spark spectrum', respectively.

References

1. *Grimsehs Lehrbuch der Physik*, 13th edn, Leipzig, vol. **4** (1957)
2. JOHNSON, R. C., *Atomic Spectra*, 2nd edn, London (1957)

1.2.2 Origin of atomic spectra

According to Bohr, radiation occurs when there is an atomic transition from some discrete higher energy level to a lower one. Prior to radiation, however, the atom should be *excited* by the injection of a definite amount of energy that will cause an electronic transfer from a lower energy state, E_n, into a higher energy state, E_m. The excited state is, however, not stable. The atom has a limited lifetime in the excited state and, depending on the character of this state, the lifetime is generally smaller than 10^{-8} s. An exception to this is the so-called *metastable* state, whose lifetime can exceed 10^{-3} s.

If the initial and final energy states of the atom are denoted by E_n and E_m, respectively, then the energy difference ΔE, released in the transition $E_m \rightarrow E_n$, is emitted in the form of a photon $h\nu$, obeying the Bohr *frequency rule*, which states that

$$\Delta E = E_m - E_n = h\nu \tag{1.3}$$

where h is *Planck's constant* ($6 \cdot 626.10^{-27}$ erg s). The frequency of the resultant spectral line is therefore

$$\nu = \frac{\Delta E}{h} \tag{1.4}$$

Furthermore, when it is taken into account that the velocity of light is given by $c = \lambda\nu$ (about 3.10^{10} cm s^{-1}), the wavelength of the radiation is given by

$$\lambda = \frac{hc}{\Delta E} \tag{1.5}$$

or the wave number value of a spectral line of frequency ν is

$$\nu^* = \frac{1}{\lambda} = \frac{\nu}{c} = \frac{\Delta E}{hc} \tag{1.6}$$

that is

$$\nu^* = \frac{E_m}{hc} - \frac{E_n}{hc} \tag{1.7}$$

In order to simplify the above expression we introduce the so-called *terms* T_m and T_n, where

$$-\frac{E_m}{hc} = T_m \quad \text{and} \quad -\frac{E_n}{hc} = T_n \tag{1.8}$$

from which the following expression is obtained

$$\nu^* = T_n - T_m \tag{1.9}$$

According to the Rydberg–Ritz combination principle above, the difference between two terms gives the wave number of a possible atomic line. According to the quantum theory the term values are definite quantities; thus it follows that the wave number of a spectral line can only be definite.

The intensity of the radiation originating from the emission process $E_m \rightarrow E_n$, having frequency $\nu_{mn} = (E_m - E_n)/h$, is proportional to the number of atoms transferred from the state E_m to E_n in unit time: N_{mn}, i.e.

$$I_{mn} \sim N_{mn} h\nu_{mn} \tag{1.10}$$

If at time t there are $N_m(t)$ atoms in the energy state E_m, then

$$N_{mn} = A_{mn} N_m(t) \tag{1.11}$$

where the constant A_{mn} is the *transition probability* between the states E_m and E_n. The reciprocal value of this constant is the 'average lifetime' of the state E_m, so that

$$\tau_{mn} = \frac{1}{A_{mn}} \tag{1.12}$$

In the above it has been tacitly assumed that the transition $E_m \rightarrow E_n$ is the only one possible. When there are more transitions possible from state E_m into states E_1, E_2, etc., then the mean lifetime of the state E_m is given by the expression

$$\frac{1}{\tau} = A_{m1} + A_{m2} + \ldots \tag{1.13}$$

The number of atoms having a particular energy state is referred to as the population of the electronic levels and is given by

$$N_2 = N_1 e^{-(E_2 - E_1)/kT} \tag{1.14}$$

This relationship is the Maxwell–Boltzmann distribution law, where k is the *Boltzmann constant*, having a value of $1{\cdot}38 \,.\, 10^{-16}$ erg/K and T is the absolute temperature. It follows from the above relation that the relative value of the number of atoms in the higher energy state is limited by the factor e^{-80} at room temperature. In normal circumstances, therefore, there can be only a few atoms residing on the higher energy level when the temperature is around room temperature. It also follows from the above equation that the population of the energy levels changes according to the curve shown in Fig. 1.2, where it can be seen that the population of the higher electronic levels drops very sharply.

According to Eq. (1.8) the term value is given by the negative value of the energy of the atomic state, divided by hc. The dimension of the latter is therefore cm^{-1}, although in emission analysis it is usual to use the unit eV (electronvolt). This is

the kinetic energy assumed by an electron when it is accelerated by a potential gradient of 1 V (it then has a value of $1{\cdot}555 \,.\, 10^{-12}$ erg). The advantage of using the electronvolt unit lies in the fact that the difference of term values calculated in this unit is equal to the excitation potential of the spectral line in question.

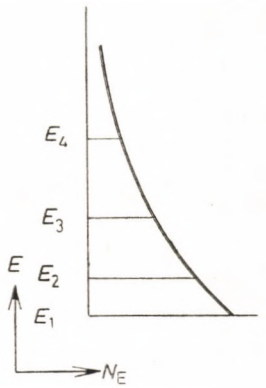


Fig. 1.2 Population of the energy levels according to the Maxwell–Boltzmann distribution law

Among the different possible energy states the lowest one is called the *ground state*. It is characteristic of this state that the atom can remain here for an indefinite time; the excited state, on the other hand, has a very short lifetime.

In considering the origin of optical line spectra in this process, it is only one or more valency electrons that are of importance out of the whole atom. The number involved is the group number in the periodic system of elements for the particular element concerned.

1.2.3 Spectral line series[1–2]

As can be seen from Eq. (1.9) there is a certain regularity concerning the possible wave number values of the spectral lines. The spectral lines can be grouped into series in such a way that for all lines belonging to the same series the value of the final term T_n is constant, while that of the initial term T_m is variable. It also follows from Eq. (1.9) that the lines of a series converge in the direction of decreasing wavelength until finally they reach the so-called series limit.

These conclusions obtained from theoretical considerations are supported by practice. The best example to illustrate this is the spectrum of hydrogen, as it is extremely simple. Upon excitation of hydrogen there appear four lines in the visible part of the spectrum, while about 20 lines can be observed in the ultra-violet region (see Fig. 1.3). These lines converge towards the smaller wavelengths up to a certain wavelength value (series limit) and are then followed by a continuum. In addition to this *Balmer series* there is a further one in the far ultra-violet part of the spectrum and several others in the infra-red region as well. The following *series formula* applies to these spectral line series

$$\nu^* = R_H \left[\frac{1}{n^2} - \frac{1}{m^2} \right] \tag{1.15}$$

The value of R_H – the Rydberg constant – is 109 677·58 cm^{-1} for hydrogen, while m and n are positive integers. The series formulae for the hydrogen line series

are therefore the following

Lyman series: $\nu^* = R_H\left(\frac{1}{1^2} - \frac{1}{m^2}\right)$, where $m = 2, 3, 4, \ldots \infty$

Balmer series: $\nu^* = R_H\left(\frac{1}{2^2} - \frac{1}{m^2}\right)$, where $m = 3, 4, 5, \ldots \infty$

Paschen series: $\nu^* = R_H\left(\frac{1}{3^2} - \frac{1}{m^2}\right)$, where $m = 4, 5, 6, \ldots \infty$

Brackett series: $\nu^* = R_H\left(\frac{1}{4^2} - \frac{1}{m^2}\right)$, where $m = 5, 6, 7, \ldots \infty$

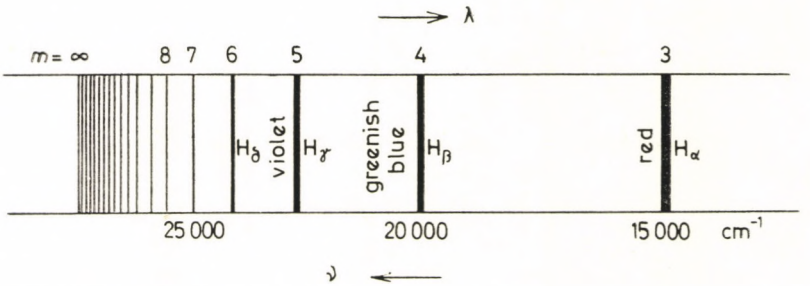


Fig. 1.3 The Balmer series of hydrogen

It has been possible to group the spectral lines of many other elements into series[1] in a similar manner. For these line series the following series formula can be used

$$\nu^* = \frac{R}{(n+a)^2} - \frac{R}{(m+b)^2} \tag{1.16}$$

Both R (the Rydberg constant) and a and b (the so-called *correction terms*) change from element to element, while n and m are integers, where the value of n is characteristic for the whole series, whereas m is the so-called *running number*. Accordingly, for a certain series $1/(n + a)^2 = G$ is constant, so the wave number of the lines belonging to this series is given by the formula

$$\nu^* = G - \frac{R}{(m+b)^2} \tag{1.17}$$

The values for m form a series of consecutive integers such that for the first member of this series the wave number ν^* is not negative and the condition is met according to which $m \geqq (n + 1)$. At the series limit $m = \infty$; thus $R/(m + b)^2$ will be equal to 0, and therefore $G = \nu^*_\infty$. Thus from the previous equation it follows that

$$\nu^* = \nu^*_\infty - \frac{1}{(m+b)^2} \tag{1.18}$$

From the above series formulae not only is it possible to establish the wave numbers belonging to the lines of any given series, but the relative positions of the different

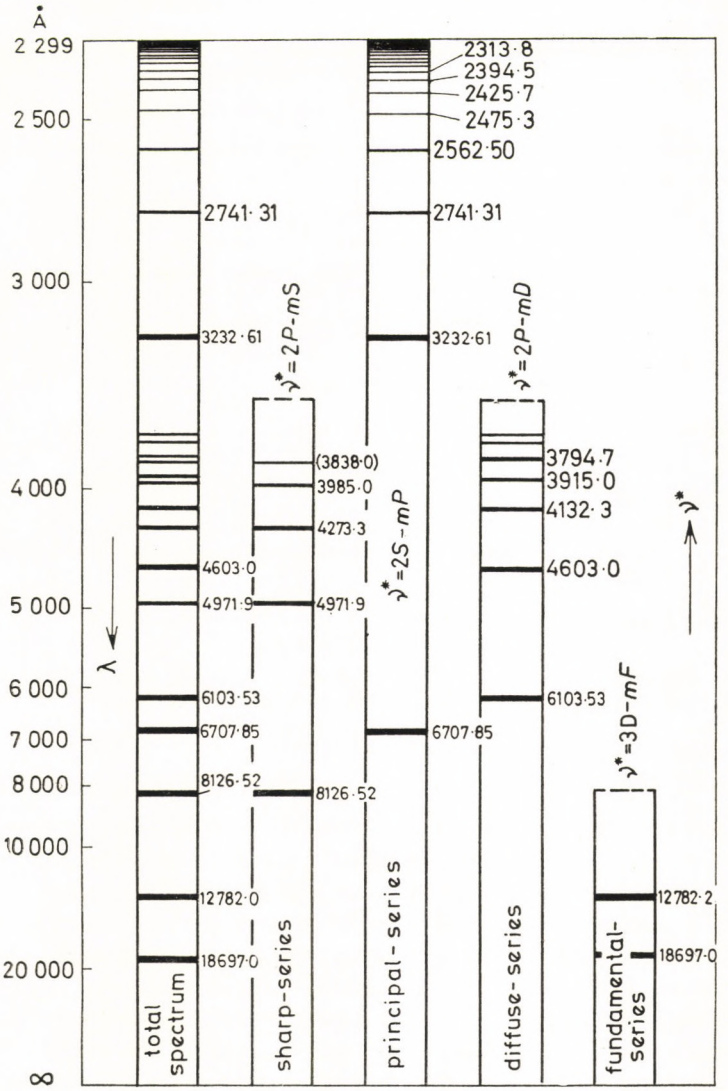


Fig. 1.4 Lithium spectrum and the line distribution of its series

series can also be determined as is shown, for example, in Fig. 1.4. This is because for certain spectral series of an element the values of constants a and b in Eq. (1.16) are identical. It is generally accepted that constants a and b are substituted by the notations s, p, d, f (sometimes g, h, etc.) in accordance with the observed character of the spectral series (see Table 1.2).

In the above discussion it has been tacitly assumed that the spectral lines are singlets. The series formulae given in Table 1.2 can also be used in those cases where the series are composed of multiplet lines. In this case the constants s, p, d, and f should be replaced by an adequate number of constants. For example, in the case of a triplet series by the constants p_1, p_2, and p_3. On the other hand, the Rydberg

formula is not applicable for the description of atomic spectra having more than one valency electron.

Table 1.2

Series formulae for elements

Name of series	Designation	Constants n a b	Series formula	Running No. m
principal	p	1 s p	$\nu^* = \frac{R}{(1+s)^2} - \frac{R}{(m+p)^2}$	2, 3, 4, . . .
sharp	s	2 p s	$\nu^* = \frac{R}{(2+p)^2} - \frac{R}{(m+s)^2}$	2, 3, 4, . . .
diffuse	d	2 p d	$\nu^* = \frac{R}{(2+p)^2} - \frac{R}{(m+d)^2}$	3, 4, 5, . . .
fundamental	f	3 d f	$\nu^* = \frac{R}{(3+d)^2} - \frac{R}{(m+f)^2}$	4, 5, 6, . . .

References

1. WALKER, S. and STRAW, H., *Spectroscopy*, London, vol. I (1966)
2. WHITE, H. E., *Introduction to Atomic Spectra*, New York (1939)

1.2.4 Term schemes[1]

The spectral structure of the elements can be derived from the Grotrian term schemes.† The left-hand ordinate gives electronvolt values, while the position of the horizontal lines indicates the energy values belonging to the different terms. The numbers on the left-hand side of the horizontal lines give the appropriate value for the principal quantum number. The right-hand ordinate gives the wave number values. The vertical bars connecting various horizontal lines indicate the spectral lines produced during transitions between different energy levels. The number written along these bars gives the wave number value of the spectral line in question. (Note: The wavelengths of the spectral lines can be easily calculated. When the difference between two wave numbers is $\Delta\nu^*$, the wavelength value (in ångstroms) is obtained from the following relation: $\lambda = 10^8/\Delta\nu^*$.) The thickness of the vertical bars connecting terms shows the relative intensity of the spectral line. While the wave numbers of the bars or spectral lines are obtained from the difference between the term values, the line intensities are determined by the probability of the transition between the term levels. Lines ending at the same lower term level belong to the same series, whose principal quantum number is shown at the left-hand side of the horizontal lines. Lines ending at the lowest level are called *fundamental lines.*

The term scheme shows that among the spectral lines belonging to the same series it is always the first one that has the greatest light intensity, while the intensity of the

† Fig. 1.5 illustrates the simplified term scheme of hydrogen.

other lines decreases towards the smaller wavelengths. It can also be established that among the different series the smaller the mean intensity, the higher the value of the lower term they belong to. The most intense lines of the lower series are called the *last lines*. The reason for this terminology is that these are the lines that disappear last from the spectrum when the excitation energy is gradually decreased, or when the excitation energy is not very great and the concentration of the element is gradually diminished in the light source.

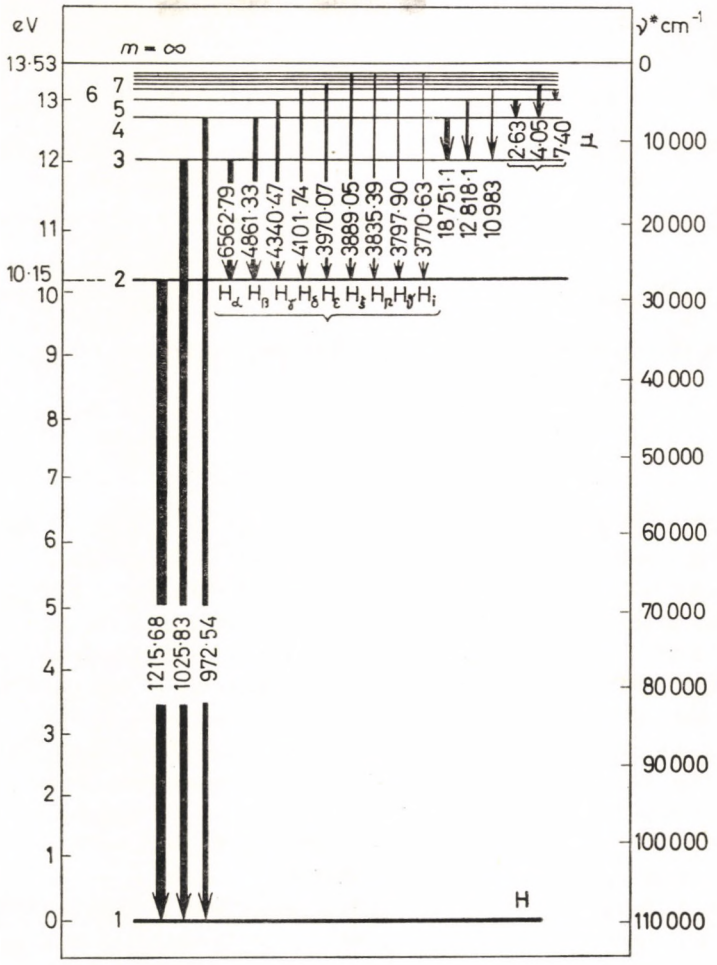


Fig. 1.5 The simplified Grotrian term scheme of hydrogen

The term scheme given in Fig. 1.5 is only a simplified drawing of the actual term scheme of hydrogen. At a higher resolution level it becomes evident that the apparent spectral lines of hydrogen are, in fact, composed of many lines (a *multiplicity*); that is, *fine structure* can be observed. Upon a further increase of the dispersion each multiplet line can be split to two components. The *hyperfine structure* of the hydrogen spectrum therefore shows a doublet structure.

In spite of the fact that both hydrogen and the alkali metals have only one valency electron, the interpretation of the spectra of alkali metals entails further difficulties.

This is because the lines belonging to the fine structure of alkali metal spectra, for example the D_1 and D_2 lines of sodium, appear already at low resolution. This is easily comprehended if we consider the correction terms in the series formula of sodium, given in Table 1.3.

Table 1.3

Values of the constants in the series formula of sodium

Name of series	Design	n	a	m	b
principal	p	1	$s = 0{\cdot}629$	2, 3, 4, . . .	$p = 0{\cdot}144$
sharp	s	2	$p = 0{\cdot}144$	3, 4, 5, . . .	$d = 0{\cdot}070$
diffuse	d	2	$p = 0{\cdot}144$	2, 3, 4, . . .	$s = 0{\cdot}629$
Bergmann	f	3	$d = 0{\cdot}070$	4, 5, 6, . . .	$f = 0{\cdot}020$

Reference

1. GROTRIAN, W., *Graphische Darstellung der Spektren von Atomen und Ionen mit ein, zwei und drei Valenzelektronen*, Berlin (1928)

1.2.5 Systems having one valency electron[1–4]

According to the *Rutherford–Bohr* theory* the number of electrons orbiting around the nucleus is given by the atomic number of the element. The arrangement of these electrons around the nucleus is such that the total energy of the atom is always at a minimum. This rule is fulfilled by the grouping of the electrons into different shells spaced at different distances from the nucleus. In the origin of optical spectra, however, only the least strongly bound electrons in the outermost shell (the *valency electrons*) are important.

As opposed to classical physics, the Rutherford–Bohr theory assumes also that the electron orbits cannot be freely chosen, and that the electrons can only move in selected orbits that conform to the quantum conditions. The Bohr theory furthermore assumes that light is only emitted if the electron is brought from a higher energy quantum state into a lower energy state.

It follows from the above that according to the Rutherford–Bohr theory the hydrogen atom is composed of a nucleus having charge e, and an electron also having charge e and orbiting with peripheral velocity v around the nucleus (Fig. 1.6). The electron has thus both kinetic and potential energy.

In calculating the kinetic energy it should be taken into consideration that the electron is kept on its circular orbit of radius r only when the centripetal force operating between the nucleus and the electron (e^2/r^2) equals the centrifugal force (mv^2/r).

* In the following we omitted more up-to-date aspects of quantum mechanics. The reason for this is that by using the Rutherford–Bohr–Sommerfeld theory not only can the fundamental principles necessary for the practical applications of spectral analysis be illustrated better, but their interpretation is also simpler.

That is, when the following condition is met

$$\frac{e^2}{r^2} = \frac{mv^2}{r} \tag{1.19}$$

The kinetic energy of the electron is therefore

$$E_{\text{kin}} = \frac{mv^2}{2} = \frac{e^2}{2r}$$

In calculating the potential energy of the electron it is advisable to start arbitrarily at the initial state of the ionised atom or, in other words, when the potential energy is zero. In this case the electron is at an infinite distance from the nucleus. Then the potential energy of the electron at a distance r from the nucleus can be given as the work performed by the attractive force (e^2/r^2) operating between the electron and the nucleus. When the electron is brought from infinity to a distance r from the nucleus, the work done is given by

$$E_{\text{pot}} = \int_{\infty}^{r} \frac{e^2}{r^2}\, dr = -\frac{e^2}{r}$$

Considering that the total energy of the electron is given by the sum of its kinetic and potential energy, we obtain from the two previous equations that

$$E = E_{\text{kin}} + E_{\text{pot}} = \frac{e^2}{2r} - \frac{e^2}{r} = -\frac{e^2}{2r} \tag{1.20}$$

The negative value of the energy occurs because we chose the zero energy state to be that of the highest energy, that is, the ionised state.

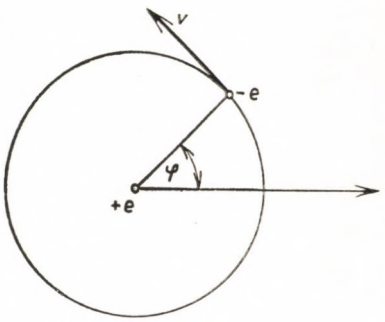


Fig. 1.6 Rutherford model of the hydrogen atom

In Eqs (1.19) and (1.20), which were derived by classical mechanical considerations, there are three unknown quantities, r, v, and e. A third equation is therefore necessary to establish the energy of the electron. Results in accordance with observations are obtained if the following quantum condition is taken into consideration

$$\int_{0}^{2\pi} p\, d\phi = nh \tag{1.21}$$

where p is the angular momentum of the electron with respect to the nucleus, ϕ is the polar angle coordinate, n is a positive integer called the principal quantum

number, and h is the Planck quantum of action. Since the angular momentum corresponding to the momentum mv is

$$p = mvr$$

it follows that

$$\int_0^{2\pi} pd\phi = mvr2\pi = nh$$

From the above two equations, therefore,

$$p = n\frac{h}{2\pi} \tag{1.22}$$

That is, according to the quantum condition the angular momentum is an integer multiple of $h/2\pi$, the unit of mechanical angular momentum. Using Eqs, (1.19), (1.20), and (1.22) it can be calculated that the total energy of the electron is given by

$$E = -\frac{2\pi^2me^4}{h^2c} \cdot \frac{1}{n^2} \tag{1.23}$$

From Eq. (1.8) the term values belonging to the energy values are obtained from

$$T = \frac{2\pi^2me^4}{h^3c} \cdot \frac{1}{n^2} \tag{1.24}$$

while the Rydberg constant is given by

$$R = \frac{2\pi^2me^4}{h^3c} \tag{1.25}$$

The calculated value of the Rydberg constant obtained from the above equation is in complete agreement with the value derived from experiments. This fact supports the practical utility of the Bohr theory. Thus the wave number of spectral lines can be calculated from Eq. (1.16).

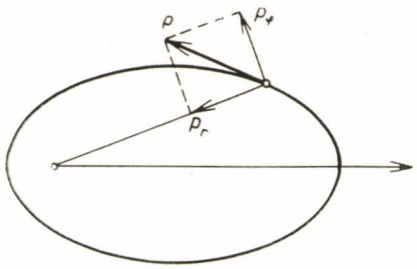


Fig. 1.7 Sommerfeld atomic model of hydrogen

The number of spectral lines arising from the combinations of the energy levels, derived in the above fashion, is much smaller than the number of the actually observed lines. This is because at higher resolution it becomes evident that the spectral lines of elements having one outer electron are composed of many closely packed lines, that is, the spectra have a *fine structure*. This is difficult to observe in the case of hydrogen, since the separation between the split components is extremely small (expressed in wavelength units) while in the case of the alkali metals the separation is conspicuous. The conclusion, therefore, is that the Bohr–Rutherford theory is unable to account

for the fine structure observed in the spectra of systems having one valency electron. This deficiency is made up by the Sommerfeld theory. According to this theory the electrons revolve in elliptical orbits that have one of their foci in the nucleus. The circular orbit is considered essentially as a particular and simple form of ellipse. Thus the quantum condition formulated by Eq. (1.19) for the orbital momentum of the electron remains valid. On the other hand, the component perpendicular to the radius vector is not identical with the orbital vector if the electron moves in elliptical orbits, rather than circular ones. Accordingly the momentum vector should be resolved into an 'azimuthal' and a 'radial' component, as shown in Fig. 1.7. The quantum conditions valid for these components are

$$p_\phi = n_\phi \frac{h}{2\pi}, \qquad p_r = n_r \frac{h}{2\pi} \tag{1.26}$$

where n_ϕ and n_r are the *azimuthal quantum number* and the *radial quantum number*, respectively.

The following relationship is valid between the two quantum numbers defined above and the principal quantum number

$$n = n_\phi + n_r \tag{1.27}$$

Consequently, two quantum numbers are enough to determine the energy of the electron. Since in the case of a circular orbit the components of the angular momentum in the direction of the orbit and perpendicular to the radius vector are identical, the *azimuthal quantum number* plus the *principal quantum number* are generally used to describe the electron energy.

The quantum number derived from the azimuthal quantum number by reducing the latter by one is called the *reduced azimuthal quantum number* (or simply azimuthal quantum number) and is denoted by l according to general usage, that is, $l = n_\phi - 1$. The possible values of the azimuthal quantum number are limited by the expression

$$0 \leqq l < n \tag{1.28}$$

The value of the azimuthal quantum number can be any integer between $l = 0$ and $l = n-1$. Thus, the following azimuthal quantum number values belong to the principal quantum number

$$n = 1 \quad \ldots \quad l = 0$$

$$n = 4 \quad \ldots \quad l = 0, 1, 2, \text{ and } 3$$

In this way the azimuthal quantum number can have several integer values and, also being the number of possible electronic levels, the number of spectral lines also increases. The different electronic levels are generally denoted by letters instead of numbers, so that the electrons having azimuthal quantum numbers, $l = 0, 1, 2, 3, 4, 5$, are denoted by the letters s, p, d, f, g, h, respectively. (Note: These letters are identical with those used for the denotation of the series (see Table 1.2). This only indicates a correspondence between the build-up of the atoms and the structure of the spectra, and it does not imply the identity of the concepts used!) As the relationships between the lengths a and b of the major and minor axes of the elliptical electron orbits and between the principal and azimuthal quantum numbers are valid, the length of the major and minor axes are proportional to n^2 and $n(l + 1)$, respec-

tively, while the eccentricity is given by the following expression $a/b = n/(l + 1)$ (Fig. 1.8).

The hyperfine structure of the spectra cannot be explained by the principal and azimuthal quantum numbers. These very fine differences in energy are accounted for by the Uhlenbeck–Goudsmith theory which assumes that the electron revolving about the nucleus rotates also about its own axis. The sense of this rotation can be

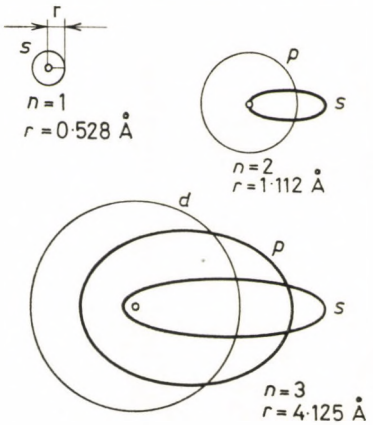


Fig. 1.8 The circular and elliptical orbitals of the hydrogen atom electrons

the same as or opposite to that of the rotation about the nucleus. According to quantum theory the angular momentum corresponding to this rotation of the electron or *spin*, cannot assume any value, except for those determined by the condition

$$p_s = s \frac{h}{2\pi}$$

where s is the *spin quantum number*. The value of the latter can be positive or negative according to the two possible spin states of the electron, and thus its numerical value can only be

$$s = \pm 1/2 \tag{1.29}$$

According to the above theory the angular momentum of the electron is built from two components. One of these is the angular momentum arising from the rotation in the orbit around the nucleus, while the other is the spin angular momentum (Fig. 1.9). For a given principal quantum number the former component is measured by the azimuthal quantum number l, while the latter is defined by the spin quantum number s, both in units of $h/2\pi$. The vectorial sum of these angular momenta is represented by the *inner quantum number* j

$$j = l + s \tag{1.30}$$

The following condition applies to this quantum number

$$j > 0 \tag{1.31}$$

The value of the inner quantum number cannot be negative and according to quantum theory it can only be an odd multiple of 1/2, that is, $j = 1/2, 3/2$, etc., provided the atom has only one electron, e.g. hydrogen. Accordingly, all energy levels are split to a doublet, except for those states having zero azimuthal quantum number.

The principal, azimuthal, and inner quantum numbers do not define the spatial orientation of the electron orbits. The plane of the orbits can be chosen at will. This 'disorder' disappears when the electrons are placed in an external magnetic field. In this case the component of the orbital angular momentum in the direction of the external magnetic field will be given by the relation $mh/2\pi$, where m is the *magnetic quantum number.* The value of the magnetic quantum number defines the direction

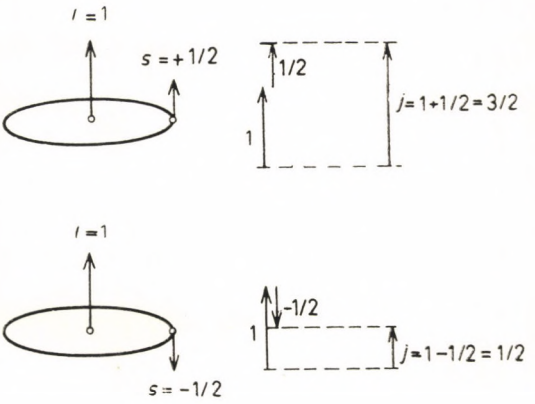


Fig. 1.9 Vector model of the one-electron atom

of the magnetic moment arising from the orbital rotation of the electron with respect to the external magnetic field. The possible number of m values and the possible energy values of the electron are limited by the fact that the number of combinations of the l and m values belonging to a given principal quantum number n is equal to n^2. It is easy to see that the possible values of m are the following

$$m = 0, \pm 1, \pm 2, \ldots, \pm l \tag{1.32}$$

When the angular momentum vector of the electron is orientated in the direction of the external field the value of m is $+l$, while in the case of opposite orientation $m = -l$. These are extreme cases, for when the angular momentum vector is perpendicular to the external field direction, then $m = 0$.

The magnetic quantum number is important when Zeeman and Stark splitting is considered in an external magnetic and electric field, respectively.

The energy and term values determined by the principal, azimuthal, and inner quantum numbers can be given by the so-called 'term symbols'. The Russell–Saunders term symbol is the following

$$n^{\mu}l_{j} \tag{1.33}$$

This symbol is composed of the following elements: n, the principal quantum number, given in arabic numbers, the exponent μ giving the multiplicity of the term, l, the azimuthal quantum number, and the subscript letter j, the inner quantum number. For example, the term symbol corresponding to $n = 2$, $l = 1$, $j = 1/2$, and $\mu = 2$ is $2^2p_{1/2}$. Exceptions are confined to terms having zero azimuthal quantum number: where $l = 0$, that is, the s-terms. These are always singlets, but in the term symbol the multiplicity is given as $\mu = 2$. Thus, for example, the notation for the case $n = 1$, $l = 0$, and $j = 1/2$ is $1^2s_{1/2}$. This can be explained by the fact that although the

'selection rules' (described later) prohibit transitions between two *s*-terms, the singlet *s*-terms can only combine with doublet *p*-terms, therefore only doublet lines can occur.

In the course of light emission from the excited atom the transitions among the energy levels, determined by the three types of quantum numbers, cannot occur in just any possible way. This is because there are *selection rules* that govern the actual possible transitions between the terms. According to these rules:

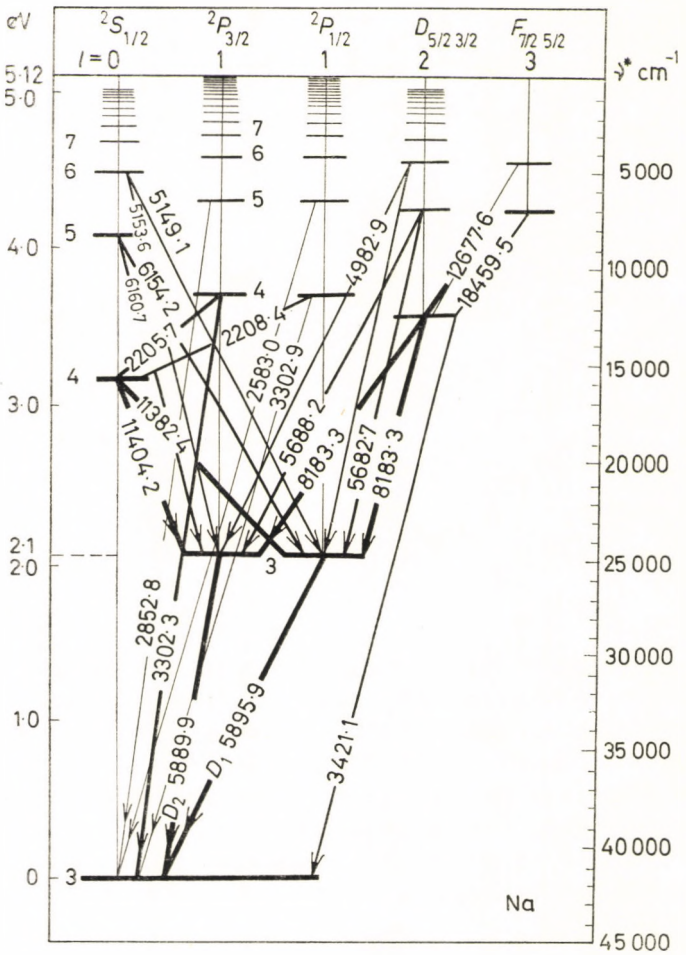


Fig. 1.10 Term scheme of sodium

(1) The change in the principal quantum number: Δn can be any integer.
(2) The change in the azimuthal quantum number is only $\Delta l = \pm 1$.
(3) The change of the inner quantum number can only be $\Delta j = 0$ or ± 1.

It follows from these rules that only those transitions are possible that occur between states of adjacent azimuthal quantum number values, or which occur between states that have the same or adjacent inner quantum number values. This rule can be illustrated by the term scheme of the sodium atom (Fig. 1.10). As illus-

trated by this scheme, the D_1 and D_2 sodium doublet at wavelengths 5889·96 and 5895·93 Å, respectively, belong to the transitions

$$3^2p_{3/2} \rightarrow 3^2s_{1/2} \qquad \text{and} \qquad 3^2p_{1/2} \rightarrow 3^2s_{1/2}$$

respectively.

References

1. BOHR, N., *Phil. Mag.*, 6th ser., **26**, 476 (1913)
2. FINKELNBURG, W., *Einführung in die Atomphysik*, Berlin–Göttingen–Heidelberg (1951)
3. WALKER, S. and STRAW, N., *Spectroscopy*, London, vol. **1** (1961)
4. WHITE, H. E., *Introduction to Atomic Spectra*, New York (1939)

1.2.6 Structure of atomic spectra having more than one valency electron

When there is more than one electron in the outermost shell of the atom, the energy state of the latter is determined by the collective energy state of all the valency electrons. The principal quantum numbers, characteristic of the main energy state of the individual electrons are, however, independent of one another. On the other hand, the azimuthal, spin, and inner quantum numbers of the electrons are related to each other by vector relations in such a way that the resultants of the combined vectors are not just any quantities, but obey new quantum conditions. The new quantum numbers belonging to these new conditions are the group quantum numbers, and are denoted by capital letters.

In the discussion of the *group azimuthal quantum number*, let us assume for the sake of simplicity that the atom has two valency electrons. When the azimuthal quantum numbers belonging to these electrons are l_1 and l_2, respectively, and l_1 is greater than l_2, then the group azimuthal quantum number has the following possible L values

$$l_1 - l_2 \leqq L \leqq l_1 + l_2 \tag{1.34}$$

If, for example, $l_1 = 3$ and $l_2 = 2$, then $l \leqq L \leqq 5$, so that the possible values of the azimuthal quantum number L are the following: $L = 1, 2, 3, 4,$ and 5. The resultant energy levels are therefore characterised not only by the principal quantum numbers of the individual electrons, but also by the defined values of the group azimuthal quantum number. Figuratively speaking, this means that the angle between the planes of the elliptical orbits of the electrons cannot be taken at will, but must have some definite value. In a manner similar to azimuthal quantum number symbols in term schemes having one outer electron, the group azimuthal quantum number is a capital letter, such that for $L = 0, 1, 2, 3, 4, 5$, etc., the letters S, P, D, F, G, H, etc., are used, respectively.

The *group spin quantum number* is equal to the arithmetical sum of the spin quantum numbers of the individual electrons. The value of S, however, cannot be negative, that is

$$S = \sum s \geqq 0 \tag{1.35}$$

The maximum value of the group spin quantum number is determined by the number, z, of valency electrons in such a way that the following relationship holds

$$S_{\max} = \frac{z}{2} \tag{1.36}$$

The values of the possible spin quantum numbers and the group spin quantum numbers in the presence of z valency electrons are summarised in Table 1.4.

Table 1.4

Relationship between the number of valency electrons (z), the possible values of the spin quantum numbers (s), and the value of the group spin quantum numbers (S)

z	s	S
1	+1/2, −1/2	—
2	+1/2, +1/2	1
	+1/2, −1/2	0
3	+1/2, +1/2, +1/2	3/2
	+1/2, +1/2, −1/2	1/2
4	+1/2, +1/2, +1/2, +1/2	2
	+1/2, +1/2, +1/2, −1/2	1
	+1/2, +1/2, −1/2, −1/2	0

As in the case of systems having only one valency electron, the quantum theory necessitates that the sum of the vectors corresponding to the group azimuthal quantum number L and the group spin quantum number S should obey new quantum conditions. The new quantum number corresponding to the vector arising this way is denoted by J and it is called the *group inner quantum number*. The quantum conditions to be met are

$$\begin{aligned} &\text{when } L > S, \text{ it is necessary that } L - S \leqq J \leqq L + S \\ &\text{when } L < S, \text{ it is necessary that } S - L \leqq J \leqq S + L \end{aligned} \tag{1.37}$$

while the number of the possible J values is given by

$$2S_{\max} + 1 \tag{1.38}$$

For example, from the above, when $L = 10$, $S = 2$, then $J = 8, 9, 10, 11$, and 12; or, when $L = 10$, $S = 5/2$, then $J = 15/2, 17/2, 19/2, 21/2, 23/2$ and $25/2$. From the previous two examples the rule emerges that, depending on whether S has an integer value or half-integer value, the values of J will be numbers that always change progressively by one unit.

The μ multiplicity of the spectra is determined by the group inner quantum number. From Eqs (1.36) and (1.38) it follows that the value of maximum multiplicity ($\mu_{\max}$) is one unit larger than the number of valency electrons, z; that is

$$\mu_{\max} = z + 1 \tag{1.39}$$

The above relation is also proved by looking at Table 1.5, which summarises the group spin quantum numbers and spectral multiplicities of atoms in successive columns of the periodic system. It becomes evident from this table that the multiplicity of the spectral lines in the odd columns of the periodic system is an even

number, while in the even columns the multiplicity is an odd number; this is known as *the Rydberg alternation rule.*

Table 1.5
Relationship between the number of valency electrons (z) and the group spin quantum number (S), and the multiplicity of the spectra

z	S	μ
1	$\pm 1/2$	2
2	0, 1	1, 3
3	1/2, 3/2	2, 4
4	0, 1, 2	1, 3, 5
5	1/2, 3/2, 5/2	2, 4, 6
6	0, 1, 2, 3	1, 3, 5, 7
7	1/2, 3/2, 5/2, 7/2	2, 4, 6, 8
8	0, 1, 2, 3, 4	1, 3, 5, 7, 9

It also follows from the above that the smaller the number of valency electrons in the element in question, the simpler will be its spectrum. Thus, the spectra of the elements in the first column of the periodic system are relatively simple, while those of the elements belonging to the other columns are so complicated that it is practically impossible to arrange them in series (Fig. 1.11).

With respect to the multiplicity of the spectral lines the following rules are observed:

It is characteristic of the spectra of elements in the first column of the periodic system (the alkalis, copper, silver and gold), that although they are composed of doublet lines, the wavelength separation of these lines is frequently large. For example, the most intense line doublet of copper is found at wavelengths 3247·5–3274·0 Å, and for gold the doublet lines occur at 2428·0–2676·0 Å. On the other

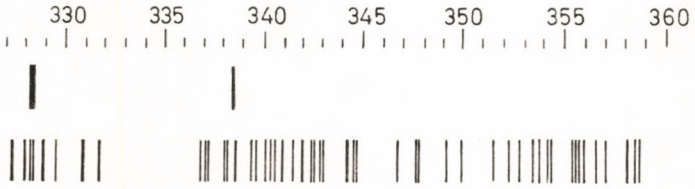


Fig. 1.11 Spectra of silver and iron

hand, the *D*-line of sodium is composed of closely associated lines at 5895·93–5889·96 Å. In the case of the above doublets it is always the line of shorter wavelength that has the greater intensity. In general, the line that has the greater intensity in the doublet is the one in which the sign of the changes of the azimuthal quantum number l, and of the inner quantum number j are identical.

In the atomic spectra of elements belonging to the second group of the periodic system (Be, Mg, Ca, etc., and Zn, Cd, Hg, etc.), there are triplet lines associated

with singlet lines (Fig. 1.12). In the triplets the highest intensity will belong to the component with the longest wavelength. The spectrum of helium is also composed of singlet and triplet lines, although when compared with the previously mentioned elements it is seen to be composed of two entirely separated term schemes. The rule holds that *there can be no interconnection* between the singlet and triplet terms (Fig. 1.13). Another important difference is that the 2^3S term of helium is a *metastable* one.

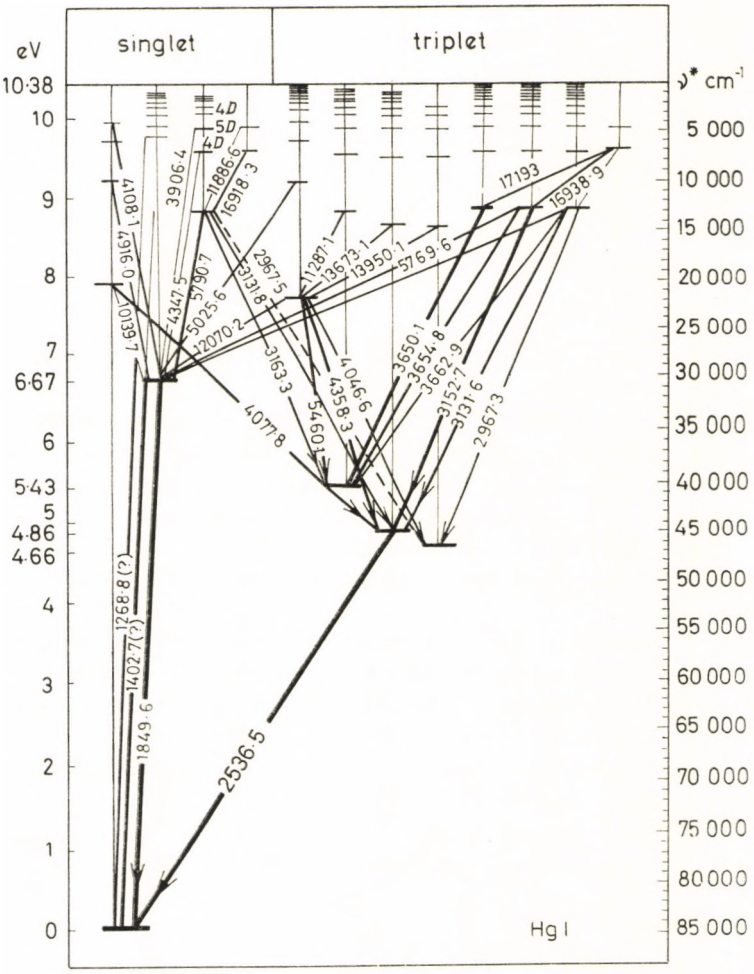


Fig. 1.12 Term scheme of the atomic spectra of mercury

From the point of view of the practical spectral analysis, a property of the multiplets is very important; namely, that the relative intensity of the multiplet lines is largely independent of the method used for the excitation of the spectrum.

The *selection rules* of systems having more than one valency electron are similar to those for one valency electron systems. There is a difference, however, since the value of the azimuthal quantum number can change to zero as well. These selection rules are as follows:

(1) There can be any change in the principal quantum number; Δn can be any integer.
(2) The allowed changes of the group azimuthal quantum number are the following: $\Delta L = 0$ and ± 1.
(3) The change of the group inner quantum number can be either $\Delta J = 0$ or ± 1.
(4) According to the Pauli exclusion principle[1] the atom cannot have two or more such electrons that have the same values for all four quantum numbers.

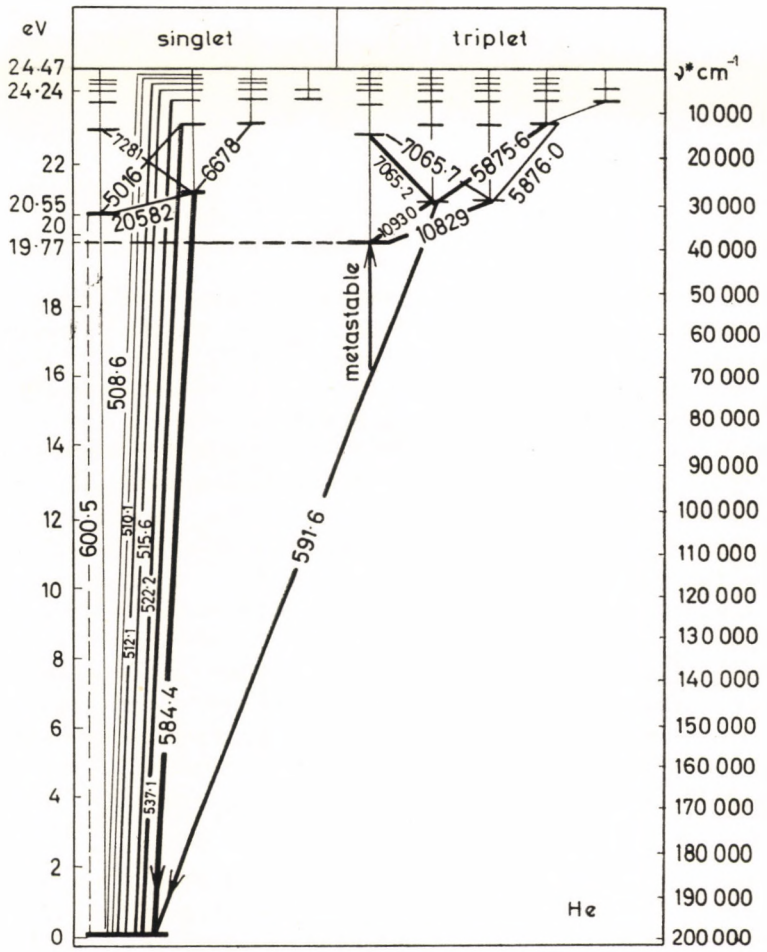


Fig. 1.13 Term scheme of helium

The term symbols of systems having more than one outer electron can be written in a way similar to that for one-valency electron systems. The only difference is that the corresponding upper case symbol is used. The Russell–Saunders term symbols are summarised in Table 1.6.

Table 1.6

Russell–Saunders term symbols

Singlet term	Doublet term		Triplet term		
n^1S_0	n^2S_1		n^3S_1		
n^1P_1	n^2P_2	n^2P_1	n^3P_2	n^3P_1	n^3P_0
n^1D_2	n^2D_3	n^2D_2	n^3D_3	n^3D_2	n^3D_1
n^1F_3	n^2F_4	n^2F_3	n^3F_4	n^3F_3	n^3F_2

Reference

1. PAULI, N., *Z. Phys.*, **31**, 765 (1925)

1.2.7 Ion spectra

In the previous section it has been tacitly assumed that during excitation the electrons remain within the atom. When, however, the energy transferred to the atom during excitation is increased, a point can be reached when one electron is completely removed from the atom to produce a positively charged ion. The work necessary for this effect is called the ionisation energy. This energy is generally measured in electronvolts and its value is called the *ionisation potential* of the atom.

When the energy is increased above the ionisation potential during the excitation of elements having more than one valency electron, the place of the electron removed is taken by another outer electron that is then raised to the excited state. In this way *ion lines* are produced. Provided enough energy is transferred to the atom, more electrons in the outer shell can be removed, thereby allowing further ionisation states to be achieved. Each ionisation state has its characteristic wavelength in the emission spectrum. Thus *atomic lines and ion lines* can be distinguished. It is generally accepted that the atomic lines are denoted by the chemical symbol with the roman numeral I placed after it, while higher roman numerals denote the ion lines of the element in question. This is the *spectroscopic notation* for the lines. For example, for aluminium the following notation is used:

Al I: atomic line; chemical symbol Al
Al II: ion line of a singly ionised atom; chemical symbol Al^+
Al III: ion line of a doubly ionised atom; chemical symbol Al^{2+}
Al IV: ion line of a triply ionised atom; chemical symbol Al^{3+}

Naturally, the spectra of singly, doubly, and multiply ionised states also have the same sort of regularity that is exhibited by atomic spectra. The lines of ion spectra can also be grouped into series, like atomic spectral lines (Fig. 1.14). Since in successively higher ionised states the number of outer electrons is decreased one by one, it follows from the Sommerfeld–Kossel displacement law that the spectrum of a singly ionised atom is similar to the atomic spectrum of the element having an atomic number of one less. Also the spectrum of a doubly ionised atom is similar to that of the element having an atomic number of two less or to the ion spectrum of the element of atomic number of one less. For example, the structure of the spec-

trum of aluminium belonging to group III of the periodic system in the singly ionised state is similar to the structure of the atomic spectrum of magnesium, whereas the spectral structure of the doubly ionised aluminium atom is similar to that of the atomic spectrum of sodium (see Figs 1.15, 1.14, and 1.10, respectively).

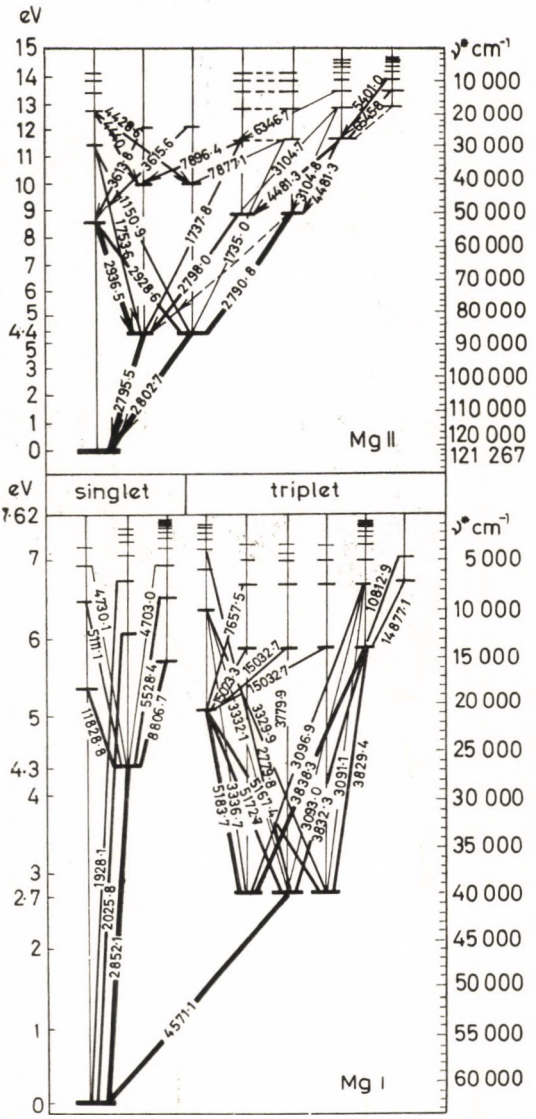


Fig. 1.14 Term schemes of the atomic and ion spectra of magnesium

It follows from the above that by increasing the degree of ionisation, the number of spectral lines becomes greater. When, however, the atomic and ion spectra of a given element are compared, it will be seen that *although the ion spectra contain many more lines than the atomic spectra, the intensity of the former is much less than those of the latter* (Fig. 1.16). This finding apparently contradicts the fact that for the excitation of ionic spectra significantly greater energy is necessary than for the excitation of atomic spectra. The contradiction disappears when it is taken into

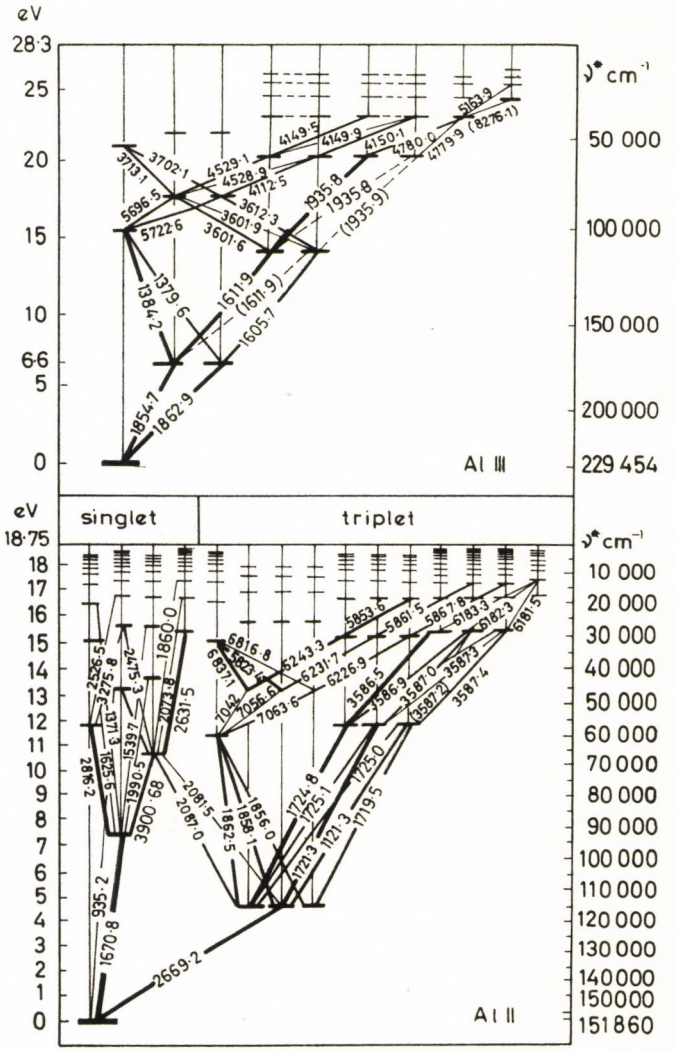


Fig. 1.15. Term schemes of the spectra Al II and Al III of aluminium

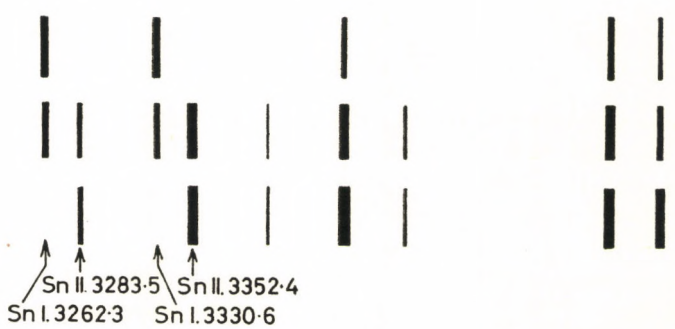


Fig. 1.16 Atomic and ion spectra of tin

able 1.7

The periodic system, the first ionisation potential (V_I) of the elements, and the structure of the characteristic electron shells. (The heavy line separates the easily ionisable elements from those which are difficult to ionise)

Period				Group I				Group II				Group III				Group IV				Group V				Group VI				Group VII				Group VIII						Group O	
				a		b		a		b		a		b		a		b		a		b		a		b		a		b									
	s	*p*	*d*	*p*	*s*	*d*	*s*	*p*	*s*	*d*	*s*	*d*	*s*	*s*	*p*	*d*	*s*	*s*	*p*	*d*	*s*	*s*	*p*	*p*	*d*	*s*	*s*	*p*	*d*	*s*	*s*	*p*	*d*	*s*	*d*	*s*	*d*	*s*	*s*	*p*
1.	1																																							
																																							2
																														1										2. He
																														H										24·50
					1				2																					13·60										
2.	2				3. Li				4. Be				2 1			2 2																								
	2	2			5·39				9·32				5. B			6. C																								
													8·30			11·26																								

account that the energy needed for the excitation of ion spectra is distributed over the increased number of lines.

From Table 1.7 it is seen that the alkali metals have the lowest ionisation energy, while the highest values belong to the noble gases. Within each group of the periodic system, the lower the ionisation potential the greater is the period number to which the element belongs. It is more difficult, therefore, to ionise the lithium atom than the cesium atom. Furthermore, taking into account that the lowest excitation energy of the elements has more or less the same trend as that of the first ionisation potential, it is seen that the first ionisation energy of the metallic elements is significantly lower than that of the metalloids. In fact, so much so that the technically important elements require from 6 to 9 eV for ionisation to occur, barring a few exceptions. This state of affairs has the advantageous consequence that by using appropriate excitation the spectra of all the metals can be obtained simultaneously. On the other hand, the excitation of non-metallic elements, like the noble gases, hydrogen, halogens, carbon, nitrogen, oxygen, phosphorus and sulphur, necessitates the application of special techniques.

1.3 Absolute and relative intensity of spectral lines

1.3.1 Absolute intensity of spectral lines[1]

The *emission coefficient* relating to a spectral line of given photon energy $h\nu$ in a certain plasma is the term used for the amount of energy emitted by a unit volume of the plasma in unit solid angle $(1/4\pi)$ in unit time. The value of the emission coefficient i is given by

$$i = \frac{1}{4\pi} A_{nm} \eta_m h\nu \text{ (erg/s cm}^3 \text{ sr)} \tag{1.40}$$

where A_{nm} is the probability of a transition between quantum states $m \rightarrow n$ belonging to the spectral line in issue and η_m is the number of particles in unit plasma-volume excited to quantum state m (i.e. the density of particles).

The value of the transition probability A_{nm} can only be calculated by means of complicated theoretical quantum considerations. Values for certain lines can be found in tabulated form from various references.[2-6]

The particle density η_m in the above expression can be calculated by assuming a Boltzmann-type, statistical distribution of the particles. The corresponding value for the unit plasma volume in thermal equilibrium is

$$\eta_m = \eta_0 \frac{g_m}{U_0} e^{-E_m/kT} \tag{1.41}$$

where η_0 is the density of particles in the same ionisation state and having all possible excited configurations, while U_0 is the so-called partition function occurring in the zero-ionisation state; that is, the atomic state is also included in the concept of ionised states. The quantity g_m is the statistical weight of the m excited state, E_m is the excitation energy belonging to the m excited state, k is the Boltzmann constant, and T is the absolute temperature of the radiating plasma.

The statistical weight of the excited state is also established by complicated theoretical quantum considerations. Its value for certain terms can be found in special tables.[2-6]

The value of the *state sum* is obtained by summation

$$U_0 = \sum g_m e^{-E_m/kT} \tag{1.42}$$

In the above calculation, however, all possible excited states for the given ionised state of the particles should be taken into consideration; thus m can vary from 1 to ∞.

On the basis of the above considerations the absolute intensity of the spectral lines can be calculated, if the dimension l of the homogeneous plasma having layer thickness l is regarded as radiant energy in the direction perpendicular to the layer. In this case the emission coefficient i in Eq. (1.40) is modified to l_i. Thus, according to Eqs (1.40) and (1.41), the absolute intensity of the spectral line can be given in the form

$$I = \frac{1}{4\pi} A_{nm} \eta_0 h\nu \frac{g_m}{U_0} l_i e^{-E_m/kT} \text{ (erg/s cm}^2 \text{ sr)} \tag{1.43}$$

References

1. WALKER, S. and STRAW, H., *Spectroscopy*, London, vol. **I** (1961)
2. SOMMERFELD, A., *Atombau und Spektrallinien*, Braunschweig (1931)
3. WIESE, W. L., SMITH, M. W. and GLENNON, B. M., *Atomic Transition Probabilities*, Washington, vol. I, 'Hydrogen through Neon' (1966)
4. WIEN, W. and HARMS, F., *Handbuch der Experimentalphysik*, vol. **22**, Leipzig (1927)
5. FINKELNBURG, W. and MAECKER, H., *Elektrische Bögen und thermisches Plasma*, in *Handbuch der Expenmentalphysik*, vol. **22**, Gasentladungen **2**, Berlin–Göttingen-Heidelberg (1956)
6. CORLICH, C. H. and BOZMANN, W. R., *Experimental Transition Probabilities for Spectral Lines of Seven Elements*, Washington (1966)

1.3.2 Relative intensity of spectral lines

Using Eq. (1.43) given in the previous sub-section, the relative intensities of two spectral lines that belong to two very different term combinations can be calculated.

The calculation for multiplet lines becomes very simple as one of the energy levels belonging to these lines is identical and the principal quantum number is also the same. In this case the values of l_i, U_0, η_0, and T are the same for both lines and what is more, E_m and A_{nm} can also be regarded as practically the same. Thus, the relative intensity can be written as

$$I_1/I_2 = g_{m,1}/g_{m,2} \tag{1.44}$$

That is, the intensity relationship of two multiplet lines is identical to the ratio of the statistical weights of the two different energy levels.

According to the qualitative intensity rule[1] the intensity relationship of multiplet lines can be approximated in more simple cases by the application of the following conditions:

(1) Among the multiplet lines the one of *highest intensity* will be the one for which the condition

$$j + 1/2 = l + 1 \tag{1.45}$$

is met for both terms involved. That is, the inner quantum number j plus one-half gives the azimuthal quantum number l plus one.

(2) The next most intense line will be the one for which the above given condition does not apply; that is

$$j + 1/2 \neq l + 1 \tag{1.46a}$$

but for the term transition it is valid that

$$\Delta j = \Delta l \tag{1.46b}$$

(3) The least intense multiplet lines will be those for which none of the above conditions are fulfilled; that is

$$j + 1/2 \neq l + 1 \qquad \text{and} \qquad \Delta j \neq \Delta l \tag{1.47}$$

The lines with the lowest intensity will be those with the greatest difference between Δj and Δl.

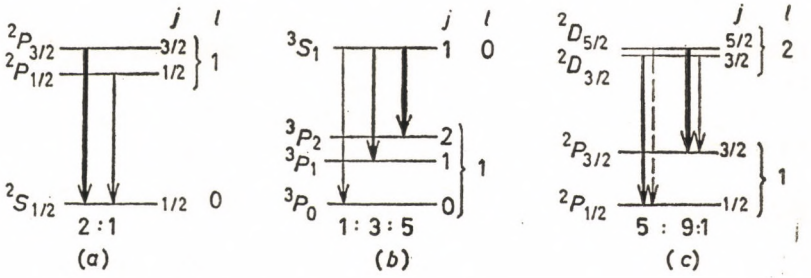


Fig. 1.17 Relative intensity of multiplet lines

The relationships given above are directly applicable only to doublet spectral lines; that is, when j has the value of a half. In *the case of triplet* lines j is an integral number, thus the value of j should be used instead of $j + 1/2$.

The *quantitative intensity rules*[2-4] are based on the determination of the statistical weights of the multiplet levels. According to these rules the statistical weights of the levels – in simple cases – are given by the number of possible magnetic splitting components $(2j + 1)$. If in the case of the $P \rightarrow S$ multiplet the S-term is common, then the intensity relationships of the members of the multiplet are given by the relation of the $(2j + 1)$ values belonging to the P-terms.

According to the above rule the intensity relation of a doublet, for example, the $P \rightarrow S$ doublet of an atom having one valency electron, like the sodium doublet, Na D_1 5889·96/Na D_2 5895·83, or the copper doublet, Cu 3247·55/Cu 3273·97, is 2 : 1 (Fig. 1.17*a*). The intensity relations of triplets of two-valency-electron elements, for example, Zn 4680·1/Zn 4722·2/Zn 4810·5, are as follows: 1 : 3 : 5 (Fig. 1.17*b*).

In more complicated cases, when both terms are doubled, as in $D \rightarrow P$ and $P \rightarrow D$ transitions, the *sum rule* should be applied. According to this rule, the sum of the intensities of two lines belonging to transitions having either a common initial level or a common final level, for example $(I_a + I_c)$, is such that the intensity of the third line (I_b) is the statistical weight of the common level for the first two lines $(2j + 1)$, which is related to the statistical weight of the third line term, which differs from that of the first two lines.

According to the above treatment, the intensity relations of multiplet lines corresponding to the transition between two doublet terms can be calculated. When the

azimuthal quantum number of the two upper levels is l, then this quantum number will be $l - 1$ for the two lower terms (Fig. 1.17*c*). Thus the appropriate j-values are directly obtained from the relation $j = l \pm 1/2$. Among the terms there are three kinds of possible transitions according to the selection rule $\Delta j = 0 \pm 1$. On the basis of the sum rule the following two equations can be written for the intensity values I_a, I_b and I_c of the developed triplet lines *a*, *b* and *c*

$$(I_a + I_c) : I_b = 2l : (2l + 2)$$

$$(I_b + I_c) : I_a = 2l : (2l - 2)$$

Using these equations the following can be derived

$$I_a : I_b : I_c = (l - 1).(2l + 1) : (l + 1).(2l - 1) : 1 \qquad (1.48)$$

Accordingly, the intensity relationship of the lines can be calculated from the azimuthal quantum numbers of the upper levels. For example, the calculated intensity relationship of multiplet lines belonging to a transitional type $^2D \rightarrow {}^2P$ is 9:5:1; as in the case of In 2560·23/In 2710·27/In 2713·94.

Using this procedure the relative intensities of multiplet lines can be calculated with acceptable precision[5].

References

1. GROTRIAN, W., *Graphische Darstellung der Spektren von Atomen und Ionen mit ein, zwei und drei Valenzelektronen*, Berlin (1928)
2. DORGELO, H. B., *Z. Phys.*, **13,** 206 (1923); **22,** 170 (1924); **26,** 756 (1925)
3. BURGER, H. C. and DORGELO, H. B., *Z. Phys.* **23,** 258 (1924)
4. ORNSTEIN, L. S. and BURGER, H. C., *Z. Phys.*, **24,** 41, 410 (1924); **28,** 135 (1925); **31,** 340 (1925)
5. KEREKESNÉ (CSÉTI), S., *Magyar Kém. Foly.* **70,** 536 (1964)

1.3.3 Relationship between line intensity and concentration

According to the generally accepted view, the fundamental relationship between the concentration c of an element in the plasma and the intensity of its spectral line is given by the Scheibe–Lomakin equation. According to this equation

$$I = \sqrt[b]{\frac{c}{a}} \qquad (1.49)$$

where a and b are constants.[1,2] These quantities, however, will only be constants over narrow concentration ranges. In addition, the value of constant a depends on many factors, such as various physical constants of the atom and its spectral line, the chemical composition of the bulk material at the radiation source and the enthalpy change of the process, etc.[3] The value of b also changes from element to element, generally $b \approx 1$, exceptionally (e.g. in the case of self-reversed line) $b > 1$[4], and $b < 1$ never occurs.

The *Malpica equation* also expresses a similar mathematical relation[5]

$$I = Ac^n \qquad (1.50)$$

In this relation the value of the constant A depends on the composition of the material

examined and the temperature of the plasma, while the value of n changes, according to the factors influencing the emission process, in the range from 0·1 to 1·5.

The drawback of the previous two relations is that the constants can only be used as such when the relative concentration interval is narrow. This disadvantage can be traced back to the fact that the intensity of the spectral lines does not depend on the concentration of the element in the plasma alone, but is also influenced by self-absorption. This latter influence is exhibited when the concentration is increased; the intensity does not increase linearly, but decreases when a certain concentration level is reached. The function describing this behaviour has therefore a maximum (see Section 2.2.11). This maximum, however, is not taken into account in the previously given intensity relationships.

To eliminate this drawback, any derived expression must contain a term that takes into account the concentration-dependent behaviour of self-absorption.[6,7]

The above condition is met by the approximate relation, given by Kerekesné–Ág,[7,8] which can be expressed as follows

$$I = Ac^n e^{-\beta c} \tag{1.51}$$

This expression contains an exponential self-absorption term: exp $(-\beta c)$, that converges to unity at low concentrations. This relation is therefore equivalent to the Malpica equation for small concentrations.

The great advantage of the above relation (Eq. 1.51) is that without changing the values of the constants A, n, and β, the calculated intensity values are in very good agreement with the experimentally established value over very broad concentration intervals (Table 1.8). Furthermore, it has the advantage that it can be easily transformed into logarithmic form.

Table 1.8

The calculated and measured intensity of the Cd 3252·53 line.
(The 'measured'-I values are obtained from arc-generated spectra, when 0·1–25 mg CdO was placed in the crater of the supporting carbon electrode)[7,8]

measured I	0·15	0·44	0·51	0·60	0·74	1·67	2·36	3·54	5·28
calculated I	0·17	0·32	0 52	0 68	0·82	1·51	2·34	3·50	5·29
deviation of I	+0·02	−0·12	+0·01	+0·08	+0·08	−0·16	−0·02	−0 04	+0·01

The values of the constants in Eq. (1.51) derived from approximate calculations, can be obtained by the Gauss least-squares equation:

$$\sqrt{\sum_i (y_x - y_i)^2} = \text{a minimum.}$$

These values based on few measurements are the following

$$\begin{aligned} \beta &= 0{\cdot}1016 \sum_i \log c_i \log I_i + 0{\cdot}04164 \sum_i \log I_i - 0{\cdot}01244 \sum_i c_i \log I_i \\ n &= 0{\cdot}1892 \sum_i \log c_i I_i - 0{\cdot}0443 \sum_i \log I_i + 3{\cdot}447\,\beta \\ A &= \text{num log}\,(0{\cdot}1 \sum_i \log I_i - 0{\cdot}234\,n + 2{\cdot}284\,\beta) \end{aligned} \tag{1.52}$$

References

1. SCHEIBE, G., LINSTRÖM, C. F., and SCHNETTLER, O., *Z. angew. Chem.*, **44**, 134 (1931)
2. LOMAKIN, B. A., *Z. anorg. allg. Chem.*, **187**, 75 (1930)
3. FRISQUE, J., *Analyt. Chem.*, **32**, 1484 (1960)
4. ZENTAI, P., *Magyar Kém. Foly.*, **70**, 538 (1964)
5. MALPICA, J. T. M., *Gen. elect. Rev.*, **43**, 288 (1940)
6. PITTWELL, L. R., *Appl. Spectrosc.*, **10**, 77 (1962)
7. KEREKESNÉ (CSÉTI), S. and ÁG, A., *Acta Chim. Hung.*, **41**, 281 (1964)
8. KEREKESNÉ (CSÉTI), S., *Acta Chim. Hung.*, **53**, 335 (1967)

1.4. Band spectra and continuous spectra

1.4.1 Structure of band spectra[1-10]

Band spectra or *molecular spectra* are emitted by excited gaseous molecules. On photographic spectrograms these bands appear as more or less broad strips, with one edge relatively sharp, while the other end is diffuse, indicating a fairly rapid drop in

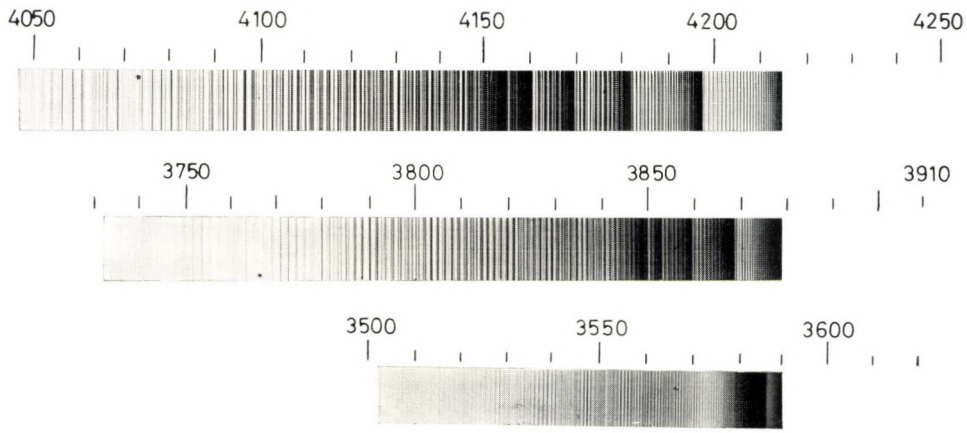


Fig. 1.18 Cyanogen bands

intensity (Fig. 1.18). In fact, these bands are composed of many narrow lines which converge towards a maximum and at a certain wavelength abruptly disappear. This position in the spectrum is called the *band head.* The difference between series formed by lines in bands and in line series is that, in the case of the important band the intensity increases toward the band head, while in line series it decreases towards the series limit.

There is a certain regularity in the appearance of the bands in the spectrum, such that *band series* are formed. These can be made up of a few or many hundreds of bands, that not only appear in different regions of the optical spectrum (infra-red, visible, and ultra-violet regions), but may in fact overlap each other. Generally, the more complicated the band spectra, the more complicated is the structure of the excited molecule. Therefore, the simplest band spectra are found in the case of diatomic molecules (N_2, H_2, CN, etc.).

Depending on the origin of the bands, different types can be produced, such as rotational, vibrational–rotational, and electronic bands. As the first two types appear in the far and middle infra-red regions, respectively, only the electronic bands in the visible and ultra-violet regions have any practical importance in emission analysis.

The electronic bands are normally composed of three superimposed series of lines known as the *P*, *Q*, and *R branches*. These can then be further divided into sub-branches, although it often happens that one branch is missing from the spectrum. The lines belonging to the branches can be approximated by the Deslandres branch formula

$$\nu^* = A + B \pm 2Bn - Cn^2 \tag{1.53}$$

where A, B, and C are constants and n is a running number having the values: 1/2, 3/2, 5/2, etc. When the $h\nu^*$ values calculated from the above empirical formula are plotted as a function of n, three parabolic curves are obtained. Such a plot is called a *Fortrat diagram*.

R eferences

1–3. HERZBERG, G., *Molecular Spectra and Molecular Structure*, vol. **I**, Spectra of Diatomic Molecules (1939, reprinted in 1951)
vol. **II**, Infrared and Raman Spectra of Polyatomic Molecules (1945, reprinted in 1947)
vol. **III**, Electronic Spectra and Electronic Structure of Polyatomic Spectra (1966)

4. KOHLRAUSCH, K. W. F., *Aufbau der Materie*, Wien (1949)
5. BARROW, G. M., *Molecular Spectroscopy*, New York–San Francisco–Toronto–London (1962)
6. WALKER, S. and STRAW, H., *Spectroscopy*, vol. **I**, London (1961)
7. KOVÁCS, J., *Rotational Structure of the Spectra in Diatomic Molecules*. In print
8. BRAND, J. C. D. and SPEAKMAN, J. C., *Molecular Structure*, Edward Arnold, London (1961)
9. JOHNSON, R. C., *An Introduction to Molecular Spectra*, Methuen, London (1949)
10. PENNER, S. S., *Quantitative Molecular Spectroscopy and Gas Emissivities*, Pergamon, London–Paris (1959)

1.4.2 Origin of molecular spectra

It was seen in the previous chapter that during atomic excitation only the kinetic and potential energies of the valency electrons are changed. It has therefore been assumed that the nucleus is in a stationary position. In considering the origin of molecular spectra it should also be taken into consideration that the molecule is composed of nuclei, in addition to the electrons, that are not stationary but move periodically with respect to one another. Assuming the molecule is composed of two atoms (this is the simplest and most frequent case in emission spectroscopy) we have to take into account that in addition to the movement of the electrons, the nuclei rotate about their mass centre and also vibrate along the line connecting them (Fig. 1.19). Thus the total energy of the molecule E is not given exclusively by the electron energy E_e, but the vibrational energy E_v and the rotational energy E_r also contribute. Therefore the total energy is given by

$$E = E_e + E_v + E_r \tag{1.54}$$

During the excitation of the molecule the vibrational and rotational energies also

change, as well as the electronic energy, thus

$$\Delta E = \Delta E_e + \Delta E_v + \Delta E_r \tag{1.55}$$

According to quantum theory, not only are the atomic energy states governed by quantum mechanics, but the electronic, vibrational, and rotational energies are also bound by these conditions. Each of these energies can therefore be expressed by quantum numbers. Accordingly, the relation expressed by Eq. (1.3), that is, the Bohr frequency postulate, $\Delta E = h\nu$, is also valid in the molecular case. Also, it follows that not only is the relation $\lambda = hc/\Delta E$ valid, but also the selection rules with respect to the corresponding possible changes of quantum number apply as well.

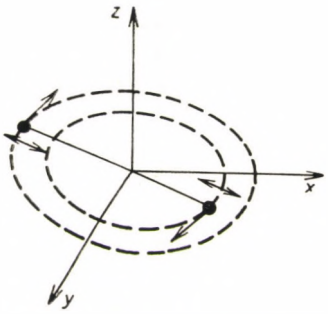


Fig. 1.19 Kohlrausch model of a diatomic molecule

The electron quantum numbers of the molecule are denoted by Greek capital letters (Σ, Π, Δ, Φ, etc.) in a similar fashion to the term notations for atoms (S, P, D, F, etc.). On the other hand, the vibrational and rotational quantum numbers are denoted by v and J, respectively. The quantum number belonging to the higher energy level is superscripted by a single prime (for example v'), while that of the lower level is denoted by a double prime (for example v''), so that in emission a transition is written as follows: $J' \rightarrow J''$.

It can be proved theoretically that the electron energies in a molecule, or changes in their energy, are greater than the vibrational and rotational energies or their respective energy changes. In accord with experimental observations, the electronic vibrational, and rotational spectra appear at subsequently longer-wavelength spectral intervals. If it is also taken into consideration that the energy changes occurring in the course of molecular electronic transitions are – in terms of order of magnitude – identical to those of the atomic spectra, it is evident that the wavelengths of lines in electronic band systems are of the same order of magnitude as the wavelengths of atomic spectral lines.

The regularity of the band systems can be illustrated by the term scheme given in Fig. 1.20. It can be seen that the coarsest structure of the spectrum is determined by the two singlet levels of the electronic state, denoted by $^1\Sigma$ and $^1\Pi$. The fine details of the two main energy levels are determined by the vibrational quantum numbers $v = 0, 1, 2$, etc. Furthermore, each electron-vibrational quantum state is subdivided according to the rotational states quantised by $J = 0, 1, 2$, etc. The fine structure of the spectrum is therefore determined by the changes of the rotational quantum number J for the unchanged vibrational state v belonging to a given $\Sigma \rightarrow \Pi$ transition, in such a manner that according to the selection rule governing the changes of J, $\Delta J = J' - J'' = 0, \pm 1$ are the only possibilities. Therefore, according

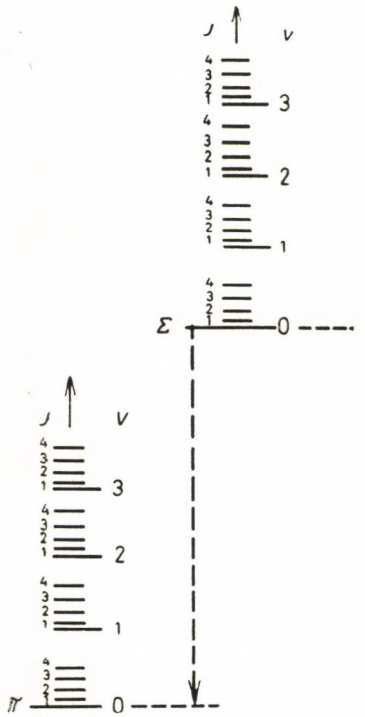


Fig. 1.20 Details of the theoretical term scheme for the CO molecular spectrum

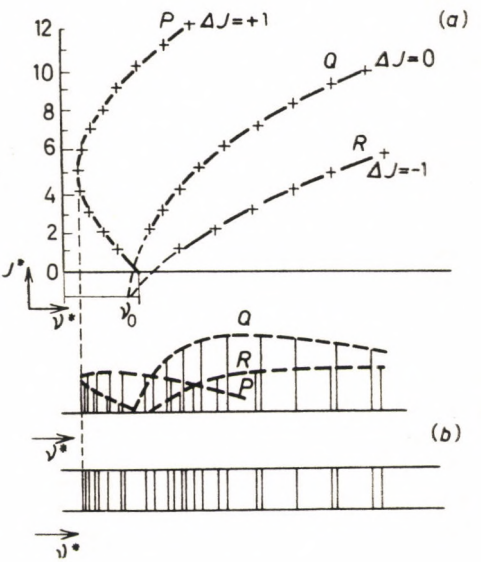


Fig. 1.21 (*a*) Fortrat diagram, (*b*) origin of band heads

to the parameters $\Delta J = +1$, 0, and -1, three band series belong to the values $J'' = 0$, 1, 2, etc.

The distribution of the lines in a band composed of the previously described three band series (denoted by *P*, *Q*, and *R*) is well illustrated by the Fortrat diagram mentioned earlier. In Figure 1.21 the origin of the band head can be seen quite clearly and is brought about by the crowding of the spectral lines. In contrast to

the series limit of atomic spectra, however, the number of converging lines is always finite. The top figure also shows that the molecular spectra are more complicated than the atomic or ion spectra.

References

See Sub-section 1.4.1

1.4.3 Role of band spectra in emission analysis

In the practical application of emission analysis it is always important to consider the origin of molecular spectra, even though their role is generally of secondary importance. The reason for this is that the band heads are only rarely used for analytical purposes (an exception is the CaF_2 band head in the determination of fluorine). More often than not the presence of band heads is disturbing, for when the difference between the wavelengths of the analysis line and a spectral line are small, they cannot be separated at the given resolution level.

In excitation in the presence of air, it is the 'air bands' (band lines belonging to N_2, O_2, and NO molecular spectra) that interfere with the analysis. When metals are analysed metal-oxide bands can arise that appear in the ultra-violet spectral region. When carbon electrodes are used, it should be taken into account that in the upper ultra-violet spectral region and in the low wavelength part of the visible spectrum cyanide bands will occur. Upon excitation of aqueous solutions extensive OH bands appear in the middle part of the ultra-violet spectrum, thereby increasing the analyst's problems.

1.4.4 Continuous spectra[2,3]

In emission spectral analysis the continuous spectrum, that is, the spectrum in which there is a continuous wavelength distribution, appears as *background radiation* that interferes with the line spectra. The interference stems partly from the fact that the intensity of the background radiation adds to the intensity of the spectral line and

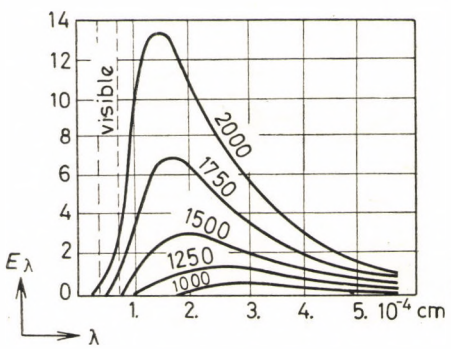


Fig. 1.22 Energy distribution of black body radiation at different temperatures

in this way wrong results can be obtained. On the other hand, when the background radiation is very strong, certain lines in the spectrum that have a low intensity cannot be identified. There are a number of physical processes that give rise to the appearance of continuous spectra that forms this background radiation.

It is well known that all bodies whose temperature is above absolute zero emit radiation. The wavelength distribution of the emitted radiation is given by the Planck radiation law. According to this law the value of the emitted energy in units of erg cm^{-2} s^{-1} is as follows

$$S = \frac{2c_1}{\lambda^5} \frac{1}{e^{c_2/\lambda T} - 1} \tag{1.56}$$

where $c_1 = 5{\cdot}955.10^{-6}$ erg cm^{-2} s^{-1} and $c_2 = 1{\cdot}432$ cm grad, λ is the wavelength in cm, and T is the absolute temperature. According to the above equation the radiant energy changes continuously as a function of wavelength. The numerical value of the wavelength belonging to the maximum energy density is given by the Wien displacement law

$$\lambda_m = 0{\cdot}288 \frac{1}{T} \tag{1.57}$$

It can be seen from the above expression that λ_m drops when the value of T increases. That is, when the absolute temperature of the radiating body increases the radiation maximum is displaced towards the shorter wavelengths (Fig. 1.22). The relative energy distribution with respect to the temperature of the radiating body and the radiation wavelength is summarised in Table 1.9.

Table 1.9

The relative distribution of radiant energy according to the temperature of the radiating body and the wavelength of the radiation, when the radiation value at wavelength 5 600 Å is normalised to 100[1]

Å	2 000 °C	2 200 °C	2 400 °C	2 800 °C	3 750 °C	6 000 °C
4 000	2·194	5·096	7·512	13·82	34·90	96·37
4 200	5·878	8·676	12·00	19·94	43·14	100·5
4 400	10·14	13·94	18·16	27·52	51·72	103·3
4 600	16·49	21·26	26·25	36·55	60·45	104·8
4 800	25·52	31·03	36·45	47·03	69·14	105·5
5 000	37·85	43·55	48·95	58·78	77·37	105·0
5 200	54·62	59·13	63·67	71·64	85·56	103·9
5 400	74·49	77·94	80·72	85·45	93·01	102·2
5 600	*100·0*	*100·0*	*100·0*	*100·0*	*100·0*	*100·0*
5 800	130·5	125·4	121·2	115·1	106·2	97·51
6 000	166·3	153·9	144·2	130·4	111·7	94·62
6 200	207·7	185·5	168·9	145·8	116·5	91·63
6 400	254·3	220·0	194·7	161·1	120·7	88·40
6 600	306·0	256·9	212·8	176·1	124·0	85·09
6 800	362·9	295·7	249·3	190·5	126·7	81·83
7 000	424·7	336·8	277·2	204·4	128·8	78·47
7 200	490·8	378·6	305·2	217·7	130·4	75·20

Continuous spectra can also be radiated by excited atoms or molecules of gases under suitable conditions and not only by solid and liquid substances. The range and intensity distribution of these continua, however, greatly depend on the experi-

mental conditions. Thus, in the case of excitation under appropriate conditions in gas discharge tubes, a continuum is joined to the series limit of the spectral lines of certain elements (hydrogen, noble gases, alkali metals, alkaline earths) whose intensity drops steadily towards the shorter wavelengths. When the gas pressure and the current intensity flowing through the tube are increased, this *series-limiting continuum* is extended towards the longer wavelengths and is accompanied by broadening of the spectral lines. Continua can also be observed in the molecular spectra of oxygen and the halogens, etc.

References

1. MOHLER, H., *Chemische Optik*, Aarau (1950)
2. KOHLRAUSCH, F., *Praktische Physik*, 19th edn, Leipzig (1955)
3. *Grimsehls Lehrbuch der Physik*, 13th edn, vol. **3**, Leipzig (1955)

2 Excitation

2.1 Basic principles

2.1.1 Requirements of excitation methods

The light source for emission analysis is obtained by the excitation of the substance to be examined. In principle, therefore, the ideal light source is the one whose spectrum contains the spectral lines of *all* the elements constituting the material under study. This requirement can be rarely attained in practice, however. For example, difficulties arise when in addition to the metallic elements that are easily excited there are non-metallic elements present that are difficult to excite, or when the easily excited elements to be determined are present in trace quantities. The excitation requirements depend on the specific task, such that there are different problems to be solved in qualitative and quantitative spectral analysis, respectively.

In the case of *qualitative spectral analysis* it is generally sufficient to determine which elements are present by visual inspection of the excited spectrum. In theory, therefore, it is sufficient to have an excitation method just sensitive enough so that the most sensitive lines of the elements in question can still be observed in the excited spectrum.

One further requirement in practical analysis is that the spectrum obtained by excitation should be easily surveyed, that is, it should be an atomic spectrum composed of a few lines.

In quantitative analysis, however, it is not enough (even from a theoretical aspect) to be content with a sufficiently sensitive excitation method. It is essential to have spectra that can be reproduced under the same excitation conditions.

In quantitative spectral analysis an excitation is regarded as sufficiently sensitive when the intensity of selected analytical lines in the spectra can be appropriately changed to produce a readable intensity. This latter requirement is achieved in different circumstances that generally depend on whether the concentration of the element to be determined in the sample is low or high. When the elements determined are present in high concentration, reproducible analytical results can only be achieved

by using ion spectra. On the other hand, elements in low concentrations can be determined with sufficient accuracy only in spectra composed mainly of atomic lines. In order to perform quantitative spectral analysis, therefore, excitation methods are required that provide the possibility of altering the character of the spectra within wide limits.

The other excitation requirement for quantitative spectral analysis is that the excitation be *reproducible*. This is indicated when the *relative* intensities of certain lines of the same wavelength are in sufficient agreement in many spectra developed under identical experimental conditions. The reproducibility of spectra depends, however, not only on the experimental conditions influencing the excitation, but also on the nature of the analytical task to be solved. In different quantitative analyses the experimental conditions that influence the excitation and secure the reproducibility of the spectra can themselves be different. In the case of quantitative spectral analysis it is also required that the excitation parameters influencing the character of spectra are reproducible themselves. This is the only possible way the intensity relationship of lines of the same wavelength can be identical in spectra obtained under identical conditions.

In summary, the optimal excitation requirements, from the point of view of quantitative spectral analysis, are as follows:

(1) It is applicable to varied analytical tasks.
(2) It secures the required sensitivity.
(3) It provides for the possibility of changing the character of the spectra.
(4) It gives line intensity ratios that are reproducible.

The commonly used excitation methods are more or less successful in satisfying the above requirements.

2.1.2 Light sources for emission analysis[1]

In emission spectral analysis only those light sources that give line spectra can be used. This requirement can be met in different ways and depends on the state of the material to be examined. In the analysis of gaseous samples the atoms can be directly excited by injecting a sufficiently large influx of electrical energy. In the analysis of solids and liquids (the usual sample state for spectral analysis) the material should be first vaporised such that the atoms arising in this way, sometimes due to dissociation, can be brought into the excited state. The excitation proper can be produced by either thermal or electrical energy. The vaporisation of the material can only be brought about *directly* by the application of thermal energy.

As the above shows, the whole excitation process can be realised by either thermal or electrical energy alone. However, the *'purely thermal'* or *'purely electrical'* excitation can only be of limited use in emission spectral analysis. More often than not excitation methods that apply both thermal and electrical energy are used. Such are the arc and the spark excitation methods. In practice, however, the *excitation itself* is effected by thermal energy during arc excitation, while in the case of spark excitation the electrical energy also plays an important role in addition to the thermal energy.

Accordingly, the radiation or light sources for emission analysis can be divided into four groups. These are:

(1) Purely thermal excitation
(2) Arc excitation
(3) Spark excitation
(4) Purely electrical excitation

The above classification is not rigorous, as strictly speaking spark excitation – like a.c. arc excitation – is only a special case of d.c. arc excitation. Thus d.c. arc excitation is the principal basis for the above methods and theoretically has a particular importance in this type of analysis.

Finally, it should be noted that the *flame photometric* procedures are nowadays not regarded as part of emission spectral analysis, although they apparently belong to it. This separate treatment is justified by the observation that flame excitation gives rise to band spectra in general, such that spectral lines have a secondary importance.

References

1. WIEN, W. and HARMS., F., *Handbuch der Experimentalphysik*, vol. **21**, Leipzig (1927)

2.2 The arc

2.2.1 Fundamental properties of the arc[2]

An arc is developed when two conductor electrodes, set at a distance corresponding to the breakdown voltage, are connected to a d.c. power supply that allows a continuous electrical discharge of relatively high intensity to pass between the two electrodes.[1,3]

The arc is characterised by the extremely high temperature of the discharge between the electrodes. The gas molecules and atoms in the discharge are split into ions and electrons (thermal ionisation) to form the thermal or *arc plasma.* The arc plasma is that part of the gas between the electrodes which radiates light and in which there are migrating, electrically charged particles (electrons and ions) among the neutral particles.[4] As a consequence of the potential difference between the electrodes, the small-mass electrons migrate quickly along the field lines towards the anode, while the relatively heavy, positively charged ions migrate slowly towards the cathode. This migration produces the current conductance between the electrodes. The current is conducted mainly by the highly mobile electrons, whereas the ions carry less than 1% of the current.

The current passing through the arc maintains the temperature of the plasma in such a way that a stationary state is established with respect to the thermal energy produced and dissipated. This thermal equilibrium is reached because energy taken up by the charged particles is not used *directly* for ionisation, but set free during collisions to heat up the gas. When the temperature increases, however, the degree of ionisation of the gas is also increased. In this way the current intensity is increased until an equilibrium state is established between the temperature increase brought about by the current and the increase of energy dissipated in form of heat losses. This loss of heat is caused by several factors, such as conductance of heat, radiation

of heat, light emission, convections, etc.[5] The development of the equilibrium state is also influenced by convection, the magnetic field, and the properties of the gas in the electrode gap as well. This is because the physical properties of the gases, such as heat and electrical conductance, specific gravity, specific heat, etc., change with temperature.

A further characteristic of the arc is that the discharge column is contracted in front of the cathode and the anode, so that the '*burning spot*' emerging from the electrodes has a much smaller area than the average cross-section of the arc plasma. This phenomenon is not observed in glow discharges or in gas discharges. The burning spot of the arc on the frontal part of the electrodes is mobile, although the diameter of the circle of movement is roughly the same as the width of the arc.

Furthermore the arc is characterised by the fact that no arc can be formed without the incandescent cathode. The strongly incandescent plasma is surrounded by a weakly radiating mantle called the *aureola*. It is also characteristic that upon interrupting the electric current the intensity of the light emitted by the plasma drops simultaneously with the temperature. On the other hand, light emitted from just in front of the cathode quickly disappears, although the thermal radiation here does not cease abruptly, but drops according to the thermal cooling curve.

An additional characteristic of the arc is that the voltage drop between the electrodes is not a linear one but is relatively large in the immediate vicinity of the electrodes. The voltage drop in front of the cathode is generally about 10–20 V in the case of metal electrodes, while it is around 10 V for carbon electrodes. The anode voltage drop is somewhat smaller for metal electrodes, whereas it is slightly larger for carbon electrodes.

Since the arc is produced by the simultaneous action of so many different physical processes, a general account of the individual phenomena and the actual role of the physical processes in the formation of the arc have yet to be adequately clarified. A further difficulty in the analysis of the phenomena is that the physical and chemical properties of high-temperature gases are not yet completely understood. At the moment we have to be content with the fact that the theoretical relationships derived for the arc are more or less in agreement with the empirical findings. At present it is not possible to give even an approximate theoretical picture of the influence of the experimental conditions on the properties of the arc. We are thus limited to observations and hypotheses derived therefrom. A further difficulty in the way of an adequate explanation is that there are only a few external experimental factors that can be changed independently – the terminal voltage of the supply, the series resistance, the distance of the electrodes, and the external temperature and pressure of the gas.

The shape of the arc is governed by the field lines of the electric field between the electrodes for a *free burning arc* of the type used in spectrochemical analysis. It is generally elliptical in shape, provided there is not an order of magnitude difference between the distance of the electrodes and the diameter of the arc. Such an arc is called *electrode-stabilised*, since the shape and location of the arc are determined by the relative position of the electrodes. In spectrochemical analysis the usual form of the latter is an arc that is vertically orientated; the horizontal version is only rarely used.

References

1. WEIZEL, W. and ROMPE, R., *Theorie elektrischer Lichtbögen und Funken*, Leipzig (1949)
2. ROMPE, R. and STEENBECK, M., *Ergebnisse der Plasmaphysik und der Gaselektronik*, Akad. Verlag, Berlin, 1967, Band **I**
3. FINKELNBURG, W. and MAECKER, H., *Elektrische Bögen und thermisches Plasma*. Handbuch der Experimentalphysik. vol. **XXII.** Gasentladungen **2,** Berlin–Göttingen–Heidelberg (1956)
4. WEISS, R., *Z. Physik*, **138,** 170 (1954)
5. PETERS, R., *Naturwissenschaften*, **41,** 571 (1954)

2.2.2 The arc plasma[1,2,11,12]

The temperature of the arc plasma varies generally between 6000 and 10 000 K. At such temperatures metals and other elements are in the atomic form, whereas molecules originally composed of many atoms are completely dissociated. Simple molecules are dissociated into atoms to such an extent that they have only a small, temperature-dependent portion, remaining in the original form (Fig. 2.1). In the arc plasma the atoms present are ionised to varying degrees depending on the arc temperature.

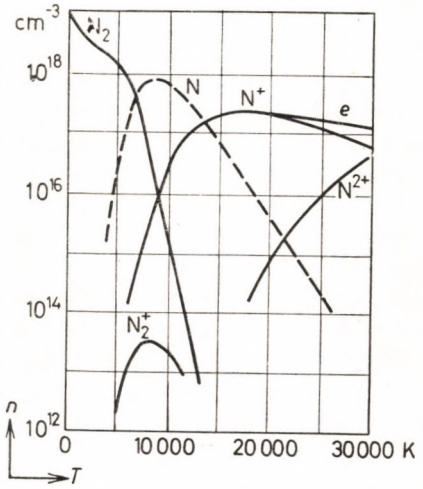


Fig. 2.1 Variation of the density (n cm^{-3}) of nitrogen particles in the plasma with temperature, at a pressure of 1 atm [3]

According to the above, the arc plasma is composed of electrons and a corresponding number of positive ions, partly derived from the electrodes and partly from the surrounding gas. Furthermore, various neutral atoms and simple molecules may also be present. The arc plasma, therefore, contains not only charged particles, but also neutral particles. The abundance of free, positive and negative charges brought about during ionisation are of course of equivalent number. The arc plasma is therefore quasi-neutral, that is, the ion and electron density values are generally equivalent. This latter statement is not valid when applied to the part of the arc plasma located immediately in front of the electrodes.

The thermal conductance of the plasma is also different from that of the gases. Provided thermodiffusion is disregarded, the slightly warm gases lose their heat energy by heat conduction to an environment that has the same pressure. On the other hand, the heat losses of the plasma do not occur only via the previously men-

tioned mechanism, but are connected with extensive local dissociation and ionisation processes taking place at the very-high-temperature, central part of the plasma. The temperature does, however, drop towards the external parts of the plasma (Fig. 2.2).

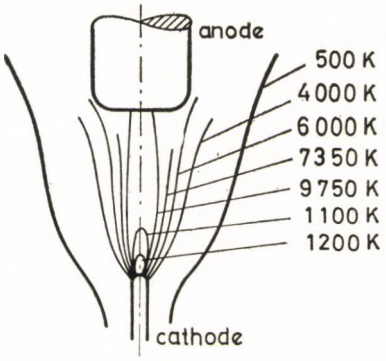


Fig. 2.2 Isotherms in the cylindrically symmetrical, 200 A carbon arc[10]

If, for the sake of simplicity, it is assumed that the plasma is spherically shaped, then the following (approximately valid) expression can be derived *for the general case*

$$\frac{T - T_\infty}{T_p - T_\infty} = \frac{R_p}{r} \tag{2.1}$$

where R_p is the radius of the spherical plasma, T_p is its temperature, and T_∞ is the temperature of the surroundings.[4] The temperature of the plasma therefore drops in proportion to $1/r$ from the value of the ambient temperature, when the surroundings are considered to be an infinite distance away. The drop is thus continuous as in principle the plasma has no geometrical limits. The volume of the plasma is, however, uncertain, even if we accept the assumption that the plasma is formed by the luminous part of the gas between the electrodes. This is because atoms having lower excitation energies are excited at lower temperatures than those of higher excitation energy. Thus the volume of the luminous part of the gas (or the volume of the plasma) is influenced by the presence of the easily excited elements.

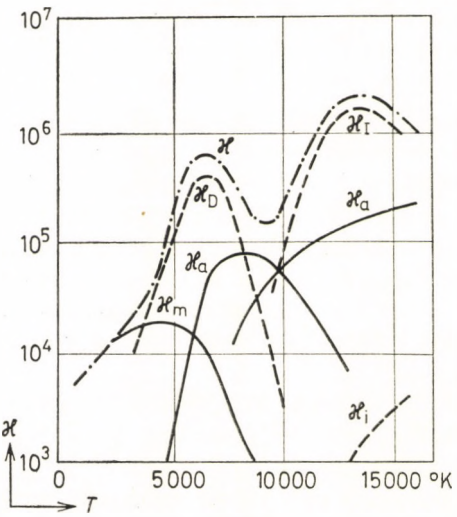


Fig. 2.3 Variation of the thermal conductivity of the nitrogen plasma with temperature. κ is the coefficient of thermal conductivity, κ_m, κ_a, κ_i are the thermal conductivities of molecules, atoms, and ions, respectively, κ_D is the conductivity due to dissociation, while κ_I is the ionisation thermal conductivity[3]

Along with the spatial drop of the plasma temperature, the dissociation or degree of ionisation of the neutral particles present in the plasma is reduced radially. This reduction is accompanied by a decrease in the concentration of the ions, and electrons as well. This has the consequence that the atoms, ions, and electrons diffuse out from the plasma, while the gas molecules of the surroundings diffuse into it. In the course of these migrations, the energy used for the dissociation and ionisation of the molecules is lost from the plasma. In the plasma, therefore, not only does the heat conduction of the plasma bring about a continuous loss of heat, but so does the *reaction heat conduction* and the *ionic heat conduction*. For the dissociation and ionisation of the gas molecules penetrating the plasma a further amount of energy is required. The temperature of the plasma cannot be maintained without a continuous supply of energy. This energy can be transferred as either thermal or electrical energy. The arc and the arc plasma are thus sustained by the electric current so that the different forms of heat loss are made up by the electrical energy supplied. As can be seen in Fig. 2.3, for arc plasma temperatures below 10 000 K, the value of the heat conductance coefficient κ is primarily influenced by the dissociation heat conductance coefficient κ_D. At such temperatures the ionisation heat conductance coefficient, κ_I has a relatively minor role.

It follows from the above that a further characteristic of the plasma is the existence of *local thermal equilibrium*.[5,6] The existence of such an equilibrium can be experimentally proved. Without recourse to this equilibrium no theoretical explanation can be given for the relationship between the mobility of the charged particles, the diffusion coefficients, the velocity of the charged particles, and the field strength.

In the arc plasma there are repulsive forces operating between particles of like charge and attractive forces between oppositely charged ones, in spite of the fact that the plasma is quasi-neutral. This results in the occurrence of local electrical fields within the plasma. The repulsive and attractive forces drop inversely with the square of the distance, but they cannot be neglected even at greater distances. On the other hand, the neutral particles attract one another only when they are brought together at collision distances. The attractive force arising in this way operates only for a very short time, about 10^{-13} seconds, thus the neutral particles do not interrupt one another's movement for most of the time.

All particles in the plasma have kinetic energy corresponding to their temperature. The kinetic energy of the free electrons and ions is also influenced by the electrical field between the electrodes. As a result of all these effects the following phenomena occur:

(1) Particles of like charge repel one another. The collision of electrons, however, does not influence the current conduction since the total impulse of colliding electrons is conserved.
(2) The collision of the neutral particles is an elastic one.
(3) Elastic collision occurs also when an electron of relatively low energy collides with a ground-state atom. In this case the electron retains its kinetic energy, but the direction of its motion will change.
(4) The ground-state atoms become excited or will be ionised when they collide with such an electron whose kinetic energy is great enough for the excitation.
(5) A *secondary collision* can occur when an electron of relatively low energy collides with an excited atom. In this case the electron takes up the excitation energy

of the atom in such a way that the electron kinetic energy is increased. Simultaneously the atom is brought back to its ground state.

(6) *Recombination* can occur when there is a collision between an electron and an ion that leads to the formation of an atom; the ionisation energy being transformed into light energy.

(7) The collision of ions and atoms is generally an *elastic* one. It can happen, however, that during the collision the charge of the ion is transferred to the atom. Thus the ion is recombined while the atom is ionised.

(8) In the plasma *triple collisions* frequently occur. In this case three particles collide in such a manner that the electron forms an atom with the ion, that is, the latter suffers recombination. The energy released is taken up by the third particle.*

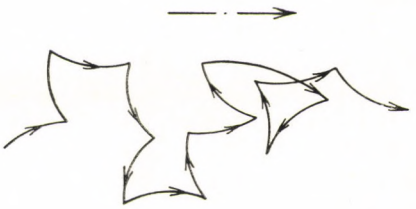


Fig. 2.4 Electronic trajectories in the arc plasma (schematic)

The electrons in the arc plasma continuously take up kinetic energy due to the effect of the external electric field, such that a movement in a definite direction is superimposed upon their disorderly thermal movements. This results in electron concentrations. The velocity of the latter is small in comparison with the velocity of the thermal motion. The electrons frequently collide with neutral particles during their streaming movement, which causes a change in the direction of the movement. The electrons therefore tend to move in a zig-zag manner in the direction of the stream (Fig. 2.4), although these departures from the main direction are relatively small. The average mean distance between two collisions is equal to the mean free path of the electrons. In spite of this the electron distribution is approximately even in the plasma. This is because only $10^{-13}-10^{-14}$ of a second is necessary for the reestablishment of the Maxwell-type velocity distribution of electrons. This is the time necessary for establishing the following relationship between the relative number $\Delta n/n$ of electrons of speed c, and the maximum value of velocity $x = c/c_{max}$, as compared with the most frequent value, c_{max}

$$100\frac{\Delta n}{n} = \text{const. } x^2 e^{-x^2} \tag{2.2}$$

The temperature corresponding to the above velocity distribution of the electrons is the *electron temperature*. Its value determines, for example, the value of the degree of ionisation, or the relative number of electrons or ions. On the other hand, the *gas temperature* is the one at which the velocity distribution of the neutral particles corresponds to the Maxwell law. The velocity distribution of the neutral particles is, however, established from the collisions among the neutral particles. This is so because the relative mass of the electrons, and their relative kinetic energy, are not

* The probability for the collision of more than three particles is very low.

sufficient to alter significantly the kinetic energy of the neutral particles during collisions. This explains why there is a relatively slow equalisation between the electron temperature and the gas temperature in the plasma. This equalisation necessitates relatively frequent collisions between electrons and neutral particles.[7-9]

References

1. LINHART, J. G., *Plasma Physics*, 2nd edn, Amsterdam (1961)
2. THOMSON, W. B., *An Introduction to Plasma Physics*, Oxford–London–Paris (1962)
3. PFENDER, E., *Umschau*, **64**, 290 (1964)
4. ROMPE, R., THOURET, W. and WEIZEL, W., *Z. Phys.*, **1**, 122 (1924)
5. ORNSTEIN, L. S. and BRINKMAN, K., *Physica, 's Grav.*, **1**, 797 (1934)
6. MANNKOPFF, R., *Z. Phys.*, **140**, 440 (1955)
7. KOCH, O., *Z. Phys.*, **126**, 507 (1949)
8. BURHORN, F., *Z. Phys.*, **140**, 440 (1955)
9. WIENECKE, R., *Z. Phys.*, **143**, 118 (1955)
10. FINKELNBURG, W. and MAECKER, H., *Elektrische Bögen und thermisches Plasma*, in *Handbuch der Experimentalphysik*. vol. **XXII**. Gasentladungen **2**, Berlin–Göttingen–Heidelberg (1956)
11. GRIEM, H. R., *Plasma Spectroscopy*, New York–San Francisco–Toronto–London (1964)
12. BAUMANS, P. W. I. M., *Theory of Spectrochemical Excitation*, Hilger and Watts Ltd., London (1966)

2.2.3 Migration of particles in the arc plasma[15-21]

It has been shown in the previous discussion that in addition to their unorientated thermal motion, the particles in the arc plasma perform an orderly migratory motion of definite orientation. The following forms of migrations are distinguished:

(1) Migration along the axis of the arc, or rather along the field lines, which is governed by the effect of the external electric field.
(2) Migration brought about by diffusion.
(3) Migration due to convection.

In the arc gap conductance of the current is made possible by the migration of the charged particles along the axis of the arc. This is different from the mechanism of current conductance in metals. While in metals current conductance is made possible solely by the migration of electrons in the direction of the current, the migration of the ions in the arc plasma also contributes to the conductance, in addition to electron migration. The migration of ions occurs in the opposite direction to that of the electrons. The electrons have much higher mobility than the ions. It can be shown theoretically that the relationship of the electron and ion mobilities, b_e and b_i, respectively, is given by

$$\frac{b_e}{b_i} = \frac{\Lambda_e}{\Lambda_i} \sqrt{\frac{m_e}{m_i}} \approx 100 \tag{2.3}$$

where Λ_e and Λ_i are the respective mean free paths of the electrons and ions, while m_e and m_i are the respective masses of the electron and the ion. In the above equation a negligible error arises due to neglect of the ion conductance in the calculation of the conductivity of the plasma. When, furthermore, it is assumed that the conductivity is determined only by the charge of the electron, e, the electron concentra-

tion, n_e, and the electron mobility, b_e, then the electrical conductivity of the plasma is given by the expression

$$\sigma = en_e b_e \tag{2.4}$$

The value of n_e, that is, the number of electrons in unit volume, is a function of the ionisation degree of the gas. The value of the latter is, however, a function of the temperature only, according to the Saha equation, Eq. (2.29), for a given type of gas and at a given pressure. Among the above mentioned conditions, therefore, the value of the electron concentration is also determined by the temperature.

It can also be shown theoretically that the electron mobility is given by

$$b_e = \frac{e\Lambda_e}{\sqrt{3m_e kT}} \tag{2.5}$$

where k is the Boltzmann constant. It can be seen from the above expression that the electron mobility is also a function of the temperature, although its value is influenced by the *free path value of the electrons* Λ_e as well. The theoretical value of the latter is

$$\Lambda_e = \frac{1}{n_0 . Q_0 + n_e . Q_e^e} \tag{2.6}$$

here n_0 is the concentration of the neutral particles, n_e is the concentration of the electrons, while Q_0 and Q_e^e are the *effective diameters* of the neutral particles and the electrons, respectively, that have to be taken into account in considering collisions between ions and electrons. The value of Q_e^e is relatively large. In the static plasma there is an excess of ions in the *immediate* vicinity of the electrons and a similar concentration of electrons is established around the ions due to the Coulomb attractive forces. Thus the diameter of the ions tends to be greatly increased.

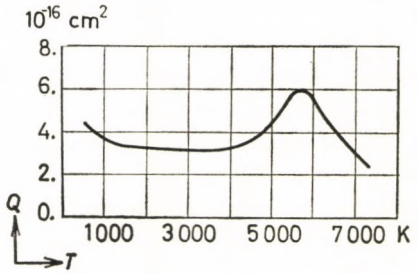


Fig. 2.5 Variation of the effective diameter of air particles with temperature[1]

The theoretically calculated diameters are rather different from the practically observed ones. In part this can be explained by the fact that the actual diameter measured depends on the temperature (Fig. 2.5). In practice the difficulty is further increased if it is taken into account that the arc plasma is composed of several gases instead of just one. Thus, not only the average concentration of the ions is uncertain but their mean diameter is also unknown. Consequently, the rule can be established that the electrical conductivity of the arc plasma is undoubtedly a function of the temperature. The value of the temperature is, however, influenced by so many factors that the effective relation between the intensity of the arc current and the temperature of the plasma cannot be established.

When the diameter of the ions in the gas mixture is about the same as the diameter of the neutral particles, the electrons tend to drag the gas atoms towards the anode. Consequently, the components of the gas mixture that are difficult to ionise are enriched in front of the anode, whereas the easily ionised components are concentrated at the cathode.

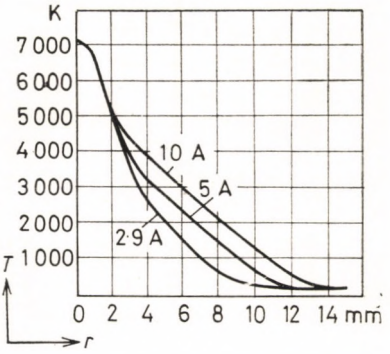


Fig. 2.6 Variation of temperature with distance (r) from the arc axis in the carbon arc plasma, for different arc-current values[2,3]

Migration of the charged particles in the plasma occurs not only under the influence of the electrical field, but also due to *diffusion*. This phenomenon is caused by the existence of concentration gradients that are brought about by the temperature differences in the plasma. Any temperature differences along the arc axis are not important, however, as they can only slightly influence the migration of the charged particles induced by the electric field.

More important is the migration of the charged particles perpendicular to the direction of the current. This process arises because there is a temperature gradient in the cross-section of the plasma, the temperature dropping with increasing distance from the arc axis (Fig. 2.6). This results in reducing the density of the charged particles in the outer layers of the plasma. Particles therefore migrate down the density gradient towards the outer parts of the plasma. The mobility of the electrons, or the diffusion coefficient D_e, is greater than the diffusion coefficient of the ions, D_i. According to the kinetic gas theory the following relation holds

$$\frac{D_e}{D_i} = \frac{b_e}{b_i} = \frac{\Lambda_e}{\Lambda_i}\sqrt{\frac{m_e}{m_i}} \tag{2.7}$$

Consequently an electric field is established perpendicular to the direction of the arc current. This will lead to a slowing down of the migration of the electrons, but a speeding up of the ions, so that at any time the migration velocity of the electrons and the ions will be equal. This phenomenon is called *ambipolar diffusion*. The corresponding ambipolar diffusion coefficient is given by

$$D_a = \frac{D_e b_i + D_i b_e}{b_e + b_i} \approx D_e \frac{b_i}{b_e} + D_i \tag{2.8}$$

However, as $b_i/b_e = D_i/D_e$, then the relationship $D_a \approx 2D_i$ is obtained. The ambipolar diffusion coefficient is therefore about twice as great as the diffusion coefficient of the ions.

The immediate consequence of the ambipolar diffusion is that charged pairs of particles (electrons and ions) migrate away from the arc plasma, while neutral

particles (atoms and molecules) migrate into the plasma from the surroundings. The indirect effect of this process is manifested not only by the exchange of the particles in the plasma, but also determines the time interval within which the particles can be found in the plasma.

The particles in the arc migrate not only under the influence of the electric field and diffusion, but also because of convection currents set up in the arc gap due to the liberation of thermal energy in the arc. Without these currents no stationary state could be achieved in the plasma, since the thermal energy set free by the arc current is taken away by these convection currents. That is steeper than that of the isotherms themselves.

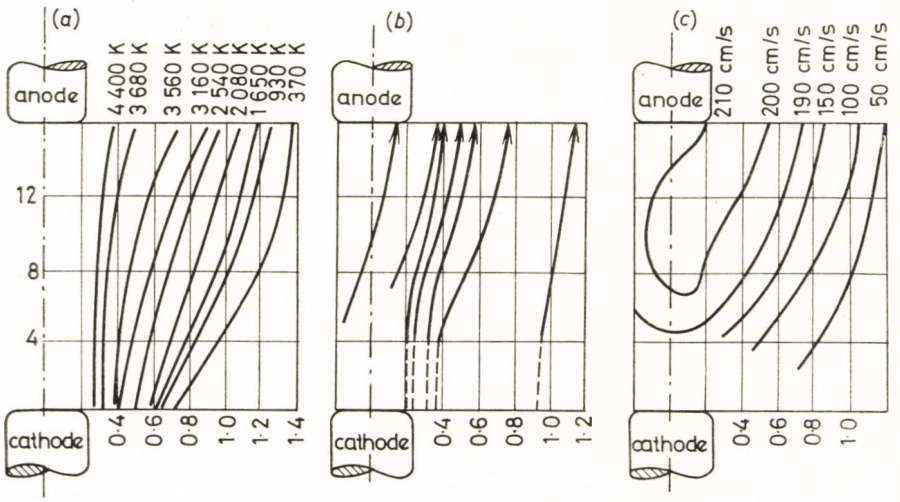


Fig. 2.7 (*a*) Isotherms in the 10 A direct current carbon arc, (*b*) direction of arc convection currents, and (*c*) velocity of the convection currents[4,5]

As illustrated in Fig. 2.7, the isotherms are very steep in the vicinity of the axis of the perpendicular arc, being almost parallel to the direction of the gas stream. Convection is therefore negligible in the vicinity of the arc axis. The further the isotherms are situated from the arc axis the greater curvature they will have, although the lower will be there the speed of the gas stream since the temperature is lower. The effect of convection becomes negligible at relatively small distances from the arc axis. These observations are in agreement with the following theoretically derived equation which gives the velocity of the convection currents

$$v_c = 2g \frac{\rho_0 - \rho}{\rho} z \tag{2.9}$$

where ρ is the air density at the temperature of the plasma at a distance z from the lower electrode, ρ_0 is the air density measured at ambient temperature and g is the gravitational acceleration.[6] This relationship not only provides a theoretical explanation for the facts mentioned above, but also corroborates the observation that the speed of the convection is independent of the form of the gas and the particles migrating within it.

In excitation it is important that convection promotes the outward migration of the particles from the plasma in a direct fashion, as well as increasing the effect of ambipolar diffusion. The convection currents remove the charged pairs of particles that are migrating out by ambipolar diffusion from the plasma. This leads to an increase in the concentration differences that cause diffusion.

It becomes clear that convection is mainly effective in the outer, cooler parts of the plasma. In these cooler zones, atoms of elements having low ionisation potentials will be ionised. On the other hand, it can be established experimentally that it is just those elements of low ionisation potential that are enriched close to the arc axis. On these grounds it can therefore be assumed that the particles are mainly removed from the arc plasma by diffusion and only a small portion is removed by convection.

In the arc plasma migration leads to particle *separation*,[7,8] since the plasma does not obey the law of mass effects. This is because within the plasma only local equilibria can be established under the effect of the external electric field and the temperature gradients. The concentration of the different particles depends on the geometrical position of those particles. Under such conditions only a stationary state will cause all the forces influencing this condition (external forces, pressure gradients, frictional forces, thermo-diffusion forces, etc.) to equalise with respect to each component in the plasma.

The ionisation energy of the atoms present in the plasma is generally rather variable. Thus, for example, there is a great difference between the ionisation energy of the metal atoms and that of the gases in the air. In the study of such composite systems it is advisable, in principle, to divide the temperature region under examination into two parts. It is assumed that one of these, the *filling-gas* region, contains high ionisation energy components exclusively in the atomic state, while the other region contains the low ionisation energy elements in completely ionised form. The latter ions diffuse ambipolarly together with a number of electrons corresponding to their charge number. On a theoretical basis it can be established that the diffusion velocity of the filling gas can be neglected. The following facts can be justified both practically and theoretically:

(1) Owing to the effect of the electric field the components that are difficult to ionise are enriched in the direction of the current flow in the neighbourhood of the anode, while the easily ionised components are enriched near the cathode.
(2) Owing to the effect of the temperature distribution, the easily ionised particles tend to accumulate in the cooler parts of the plasma (when viewed perpendicular to the current direction).

This process is influenced by frictional forces in such a manner that they promote the accumulation of the easily dissociable particles in the cooler, outer parts of the plasma. The frictional forces, however, do not effect the neutral filling gas. This can be explained by considering that in the case of collisions among ions the diameter is greater than in the case of collisions between ions and neutral particles.

The theoretical considerations given above are also justified by observation; that is, the easily ionised particles are, in fact, enriched in the cathode space. When the radiation spectrum coming from the space in front of the cathode is compared with the spectrum from other parts of the radiation source, it can be shown that the intensity of different spectral lines in the former spectrum is enhanced to varying degrees. For example, silver and copper show an enhancement factor of 7 and 10,

respectively, while for the alkali metals this factor is about 100. On the other hand, no such phenomenon is observed for such elements as carbon, mercury, boron, beryllium, and zinc.

The above observation can be explained by remembering that the greatest migration speed in the arc plasma towards the cathode is shown by those elements with the highest degree of ionisation (x) and the greatest density. From this it follows that if m is the atomic weight of the element in question, the mass migrating to the cathode within unit time is proportional to the ratio x/m. Therefore the accumulation of the elements in the cathode space depends only on their atomic weight and degree of ionisation. The value of the latter, however, changes rapidly with the ionisation potential. Furthermore it can be stated that in the stationary state the number of particles migrating out from the cathode space is equal to the number of particles diffusing into it. As the velocity of diffusion depends on the value of the diffusion coefficient and the concentration gradient, it can be calculated that the ratio of the atomic density D_0 in the cathode space and the atomic density of the plasma D is given by

$$\frac{D_0}{D} = 1 + xr_D \frac{\text{const}}{D} \tag{2.10}$$

where r_D is the distance of the plasma having atomic density D from the burning spot on the cathode.

The above relationship leads to the following conclusions:[11,12]

(1) Elements do not accumulate in the cathode space, unless their atoms are ionised. (In this case $x = 0$, thus $D_0/D = 1$.)
(2) The atomic density D_0 of the cathode space varies linearly with the distance x and the atomic density D of the plasma.
(3) The atomic density of the cathode space is, in principle, independent of the atomic weight of the atoms.
(4) It can be theoretically established that the greater the accumulation of the elements in the cathode space, the lower is their ionisation potential, but the higher is their concentration in the plasma. Notably, the value of x changes rapidly with the ionisation potential, being inversely proportional to the latter.

The theoretical conclusion of point (4) apparently contradicts experimental observation. When Table 2.1 is studied (it contains the elements and their ionisation potential in the order of increasing accumulation in the cathode space) certain exceptions are found. For example, the elements beryllium, zinc, cadmium, and magnesium are enriched to a greater degree in the cathode space then would be expected on the basis of their ionisation potential. This can be explained by considering that they have a metastable high energy level from which ionisation can be more easily effected.

The time during which a particular element can be found in the arc plasma can be determined experimentally. The presently known values show that this residence time is generally about 10^{-3} s. From Table 2.2 it is seen that the residence time increases with increasing ionisation potential of the elements. The more difficult it is to ionise an element, the more slowly it will migrate towards the mantle of the arc plasma.

Table 2.1

Elements in the order of increasing cathode space enrichment, and their corresponding ionisation potential[9]

Atom	Ionisation voltage (V)	Atom	Ionisation voltage (V)	Atom	Ionisation voltage (V)
C	11·26	Cd	8·99	Mg	7·64
Hg	10·44	Si	8·15	Ca	6·11
Au	9·23	Ti	6·83	Li	5·39
B	8·30	Bi	8·00	Sr	5·70
Be	9·32	Fe, Ni, Co	7·90 7·63 7·86	Ba	5·21
Zn	9·39	Al	5·98	Na	5·14
Cu,Ag	7·72 7·58	Ga	6·00	K	4·34

Table 2.2

Residence time of elements of different atomic weight and ionisation potential in the 10 A carbon arc[15]

Element	Ionisation potential (V)	Atomic weight	Residence time (s)
Lithium	5·39	6·940	$1{\cdot}2.10^{-3}$
Sodium	5·14	22·997	$1{\cdot}7.10^{-3}$
Thallium	6·10	204·38	$2{\cdot}5.10^{-3}$
Mercury	10·44	200·61	$5{\cdot}0.10^{-3}$

References

1. MAECKER, H., *Z. Phys.*, **128,** 289 (1950)
2. HÖRMANN, H., *Z. Phys.*, **97,** 539 (1935)
3. SCHMITZ, G., *Z. Phys.*, **126,** 1 (1946)
4. STERLING, J., *Z. Phys.*, **128,** 209 (1950)
5. HAGENAH, W., *Z. Phys.*, **128,** 279 (1950)
6. ENGELST, V. S., and SPEKTOV, V. A., *Bull. Acad. Sci. USSR Phys. Ser.*, **26,** 887 (1962)
7. FRIE, W., and MAECKER, H., *Z. Phys.*, **162,** 69 (1961)
8. FRIE, W., *Z. Phys.*, **172,** 99 (1963)
9. MANNKOPFF, R., *Z. Phys.*, **76,** 396 (1932)
10. MANNKOPFF, R., PETERS, CL., *Z. Phys.*, **70,** 444 (1931)
11. MAECKER, H., *Ann. Phys.*, **18,** 475 (1956)
12. MASTRIL, F., and WIESE, W., *Z. Astrophys.*, **44,** 259 (1958)
13. RAIKHBAUM, YA. D., and MALIK, V. D., *Optika Spektrosk.*, **9,** 421 (1960); **10,** 524 (1961)
14. MALIK, V. D., and SORD, M. A., *Optika Spektrosk.*, **16,** 368 (1964)
15. GRIEM, H. R., *Plasma Spectroscopy*, New York–San Francisco–Toronto–London (1964)
16. VUKANOVIC, V., PAVLOVIC, N. and IKONOMOV, N., *Proc. XIV. Coll. Spectr. Intern.*, Debrecen, 1967, 515
17. VUKANOVIC, D., VUKANOVIC, V. and GEORGIJEVIC, V., *Proc. XIV. Coll. Spectr. Intern.*, Debrecen, 1967, 509

18. VUKANOVIC, D., VUKANOVIC, V. and GEORGIJEVIC, V., *Proc. VIII. Intern. Conf. on Phenomena in Ionized Gases*, Vienna, 1967, 248
19. VUKANOVIC, D., *Spectrochim. Acta*, **22**, 815 (1966)
20. VUKANOVIC, D., *Proc. VII. Intern. Conf. on Phenomena in Ionized Gases*, Beograd, 1965, vol. I, 762
21. BOUMANS, P. W. J. M., *Theory of Spectrochemical Excitation*, Hilger and Watts Ltd., London (1966)

2.2.4 Phenomena occurring at the electrodes of the arc

The temperature of the arc plasma varies generally between 6000 and 10 000 K. On the other hand, when the electrode material has poor heat conductance, as in the case of carbon, the temperature of the arc is generally not higher than 4000 K. For tungsten electrodes which are difficult to vaporise, this temperature is about 3000 K, although for easily vaporised metal electrodes, having a good thermal conductivity, the temperature does not even reach this figure. There is, therefore, a significant temperature difference between the plasma and the electrodes in the interface area. The relatively cool electrodes effectively reduce the temperature of the plasma that comes into contact with them. The gas layer in contact with the electrodes is highly ionised so that conductance of the current is made possible by the charged particles in that area.

There is a different explanation for the above in the case of the cathode and the anode. In both cases, however, the relatively great potential drop between the electrodes (the cathode drop and the anode drop), is such that the arc plasma is contracted in front of the electrodes. Also to be taken into account is the work necessary for the electrons to leave the cathode and to reach the anode, since the electrical conductance of the plasma is mainly due to the electrons.

Among the phenomena occurring on the electrodes and in their vicinity, the simplest are found in the case of the anode. The anode potential drop is caused by the fact that the electrons migrating towards the anode are amassed in front of the electrode and thereby produce a negative space charge. The layer thickness of the anode drop, just as in the case of the cathode drop, is smaller than the free path of the electrons. The kinetic energy of the electrons that pass through the anode potential drop without collision is so much enhanced that the thermal energy liberated during their collision with the electrode surface strongly heats up the interfacial area between the anode and the plasma. Thus the temperature of the gas in contact with the relatively cold electrodes remains at a level that is enough to ionise the atoms sufficiently to secure electrical conductivity. Heating up is further increased by the thermal energy corresponding to the work function liberated when the electrons enter the anode. On the other hand, the energy used for the ionisation of the atoms in the plasma layer, which is cooled by the electrode, is negligible. This ionisation is a relatively infrequent process since the mobility of the ions is generally smaller, by a factor of about 100, than that of the electrons. Thus the plasma layer in front of the anode is quasi-neutral only in the case when there is a single atom ionised for every hundredth electron arriving from the plasma.

In proof of the above statements, it can be calculated that for a carbon arc of intensity 10 A, for which the anode potential drop is about 10 V, the heat liberated as a consequence of the anode drop corresponds to about 100 W/s. The electrons entering the anode yield about 40 W thermal energy per second, provided the average

value is taken. Altogether about 140 W are used per second for the warming up of the anode and to make up for the heat losses of the plasma in contact with the anode, provided that the 1 W necessary for ionisation is neglected.[1]

The thermal energy liberated at the surface of the anode in the way described above leads to a significant warming up of that part of the anode which is in contact with the plasma, provided there is a contraction of the plasma diameter immediately in front of the anode. This usually occurs, since for a given arc current of intensity I the product of the average current density and the diameter of the plasma F is a constant throughout the whole plasma height, thus

$$I = en_e b_e EF \tag{2.11}$$

where b_e is the mobility of the electron. As the field strength E is relatively high in the anode potential drop area, for a small electronic density n_e it is necessary that the diameter F of the plasma in this location should be much smaller than the mean value.

The intensive heat of the *burning spot* evaporates the anode material. The vapour arising in this way contains those atoms that are ionised in the field of the anode potential drop. (It is not necessary, therefore, that the atoms ionised in the anode drop should originate from the plasma.) The vapours emanating from the anode cause the burning spot to wander around on the frontal part of the anode in such a manner that the diameter of this circular path is roughly the same as the average diameter of the plasma.[2] Upon strongly increasing the arc current it could happen that the diameter of the plasma becomes greater than the frontal part of the anode. (This is indicated by the intensive sizzling noise of the arc.) When the current intensity becomes very great – greater than 80 A – a vapour stream is ejected from the anode whose velocity has the same order of magnitude as that of the streaming speed of the electrons. Simultaneously, there is a significant increase in the anode potential drop and the temperature of the vapor emanating from the anode reaches 6000 K.[3]

It is well known that no arc is ignited when the potential between the electrodes is lower than the breakdown voltage. The arc can still be ignited provided the electrodes are brought into contact. The explanation for this is that due to the contact resistance the electrodes are heated up so much that the ejection of electrons becomes possible. Another, more appropriate method for igniting the arc is to superimpose a high-frequency alternating current upon the arc current such that the voltage of the former is higher than the breakdown voltage of the electrode gap. The series of sparks produced in this way renders the arc gap conducting for the lower voltage arc current. Both procedures are based on the fact that no arc can be established till the thermionic work function of the electrons emanating from the cathode is surpassed.

It follows from the previous discussion that the processes occurring in the cathode have a very important role in the origin of the arc. Among them, those listed below are the most important:

(1) If the surface of the cathode has a high enough temperature it will emit electrons (the so-called *thermionic electrons*). This is the *thermionic emission.*
(2) The positively charged ions migrating from the plasma towards the cathode establish such a strong electric field in the cathode space that the ejection of electrons is rendered possible.

(3) Immediately in front of the anode the atoms are continuously ionised to such an extent that the electrons produced in this way replenish those electrons entering the anode from the side of the plasma.

By calculations it can be demonstrated that the last process is the most important out of the three listed above. This process also provides the explanation for the *cathode potential drop*. This is caused by the ionisation of the atoms immediately in front of the cathode in the course of the ignition of the arc. The mobility of the electrons arising is much higher than that of the ions. The electrons quickly migrate away from the cathode space into the plasma, while the positive ions left behind establish a positively charged region immediately in front of the cathode. The recombination of the ions migrating to the surface of the cathode happens in such a way that they attract the electrons from the cathode.

On the basis of the previously given assumption, the spatial dimensions of the cathode potential drop can be calculated. In order to simplify the calculations we take into account only those changes that occur along the arc axis, then by using the *Poisson* equation $\mathrm{d}E/\mathrm{d}x = 4\pi n_i e$, the relation for the current density $j = eb_i n_i E$, and furthermore the formula for the field strength $E = -\mathrm{d}U/\mathrm{d}x$, the following result is obtained (after double integration) for the cathode drop

$$\Delta U = \frac{2}{3}\sqrt{\frac{8\pi j}{b_i}} \cdot (\Delta x)^{3/2} \tag{2.12}$$

where Δx is the layer thickness of the cathode drop and b_i is the ion mobility.[4,5] When the average measurement data, $U \approx 10$ V and $j \approx 500$ A/cm^2, are substituted into the above expression it can be calculated that the layer thickness of the cathode drop is less than 5.10^{-4} cm, while the free path of the electrons is about twice as big. There is a high probability, therefore, that electrons passing through the cathode drop region will not collide with other particles.

The above conclusions lead to the result that the energy necessary for the ionisation of the atoms in the cathode drop region is provided exclusively by the thermal energy produced there. The surplus thermal energy not only makes up for the heat losses to the external environment of the gas, but also causes the warming-up and evaporation of the cathode. It can be calculated that in the case of a carbon arc of intensity 10 A, about 40 W are used for the work function of the electrons and about 60 W are necessary for the compensation of heat conduction[1] out of the total energy (100 W) of the cathode drop.

The recombined atoms on the cathode surface return to the gas of the cathode drop area heated to a temperature at which their ionisation is sufficient to produce a supply of electrons migrating towards the anode and thus to secure the conduction of the arc current. This state is reached only when the temperature of the plasma increases in the direction of the cathode. It is necessary that the diameter of the plasma should be reduced simultaneously to such an extent that no disadvantageous effects emerge owing to the cooling effect at the frontal part of the cathode and that this should not result in a too intensive cooling of the plasma.[6] During the plasma contraction a small burning spot area is therefore absolutely necessary. This statement is justified by the observation that by artificially increasing the temperature of the cathode the area of the burning spot is also increased.

It may happen that the temperature of the burning spot of the cathode becomes so high that *thermionic emission* is brought about. [7,8] This could provide an explanation for the observation that immediately after igniting the arc the burning spot undergoes rapid movement on the frontal part of the cathode. However, the burning spot becomes fixed in a certain position (being 'burnt into' the electrode), when it becomes so hot that it approaches the initial temperature of the thermionic emission.

References

1. MAECKER, H., *Ergebn. exakt. Naturw.*, **25**, 350 (1951)
2. SCHMICK, H. and SEELINGER, R., *Z. Phys.*, **29**, 168 (1928)
3. ROHLOFF, E., *Z. Phys.*, **124**, 264 (1948); **126**, 175, 224 (1949)
4. ROMPE, R. and WEIZEL, W., *Z. Phys.*, **119**, 366 (1944)
5. WEIZEL, W., ROMPE, R. and SCHÖN, M., *Z. Phys.*, **3**, 45 (1924)
6. NAGADA, H. and SEGIURA, Y., *Jap. J. Phys.*, **3**, 45 (1924)
7. BECKER, O. and SEELINGER, R., *Annln. Phys.*, **24**, 605 (1935)
8. BECKER, O. and SEMMERMEYER, K., *Z. Phys.*, **102**, 551 (1936)

2.2.5 Light emission of the arc plasma [1]

At the temperature of the arc plasma a very great number of gas atoms (or molecules) are continuously brought to their excited state. These emit their excitation energy mainly in the form of light quanta and in doing so return to their original ground state. The energy emitted this way is, however, only a few percent of the total energy transformed in the arc.

In the arc plasma the atoms are mainly excited, so the spectrum of the emitted light will consist dominantly of atomic lines. This is why the atomic lines are frequently called *arc lines* and why the spectrum composed mostly of atomic lines is called the *arc spectrum.*

In the plasma the following processes occur in connection with the excitation:

(1) During *electronic excitation* the atoms (or molecules) are excited thus

$$A + e_f \rightleftharpoons A^* + e_s \tag{2.13}$$

During the course of an *inelastic* collision of an atom, A, with an electron of sufficiently large kinetic energy, e_f, the atom is brought into the excited state, A^*, while the electron, with reduced kinetic energy, e_s, continues on its way. The probability of such a process is greatest when the kinetic energy of the electron just surpasses the excitation energy. The so-called *superelastic* or *secondary collision* has the same probability.

(2) In *triple collisions* the ion, A^+, is excited only when it collides with two electrons of smaller kinetic energy, such that the following reaction occurs

$$A^+ + e'_s + e''_s = A^* + e_f \tag{2.14}$$

The opposite process can be triggered by an electron whose kinetic energy is greater than the ionisation energy.

(3) When an *electron is in collision with a molecule*, the diatomic molecule dissociates according to the reaction scheme

$$A_2 + e_f = A^* + A + e_s \tag{2.15}$$

One of the atoms in the molecule thus becomes excited. Dissociation may also occur because of collision; that is

$$A_2 + e_f = A + A + e_s \tag{2.16}$$

This reaction is, however, far more complicated for polyatomic molecules.

(4) The '*spontaneous emission*' of radiation occurs when an excited atom, A^*, collides with another atom, B_f, that is travelling at a high speed. Owing to this collision the atom A^* becomes excited by the reaction:

$$A^* + B_f = A^{**} + B_s \tag{2.17}$$

The excited atom is then brought into the metastable state A^{**}. This atom then gets rid of the energy absorbed in the form of light quanta, according to the reaction

$$A^{**} = A^* + hv \tag{2.18}$$

It is much more probable, however, that the excited metastable atom is brought back to the excited state in the course of a process which is the opposite of reaction (2.17), and not by (2.18). This happens because the probability of (2.17) is greater by several orders of magnitude than that of (2.18). (The metastable state can, however, exist for comparatively longer periods in an undisturbed condition.)

The *radiation intensity* of a given element (see Section 1.3.1) depends on the number of atoms in the appropriate excited state. It can be shown theoretically that in a plasma of temperature T the proportion of all atoms of a given element in the excited state is given by

$$x_m = Cg_m \cdot e^{-E_m/kT} \tag{2.19}$$

where E_m is the appropriate excitation energy of the atoms, k is the *Boltzmann* constant, and g_m is the 'statistical weight' of the excited state. The value of the constant, C, is equal to $1/\sigma_0$; provided the partition function of the neutral atoms is given by

$$\sigma_0 = \Sigma g_m \cdot e^{-E_m/kT} \tag{2.20}$$

When it is considered that the number of atoms having the excitation energy, E_m, in unit volume is given by $n_m = n_0 x_m$, and that the number of neutral particles in unit volume is $n_0 = p_0/kT$, we obtain the following expression

$$n_m = \frac{1 - x}{1 + x} \cdot \frac{pg_m}{\sigma_0 kT} \cdot e^{-E_m/kT} \tag{2.21}$$

where x is the degree of ionisation and p the pressure and $1 - x$ and $1 + x$ are quantities proportional to the unionised and the total amount of the particles, respectively. Furthermore, when it is assumed that the probability of the transition from quantum state m to state n is w_{mn}, then the number of transitions in unit volume and unit time is $n_m w_{nm}$. Finally, as in each transition the energy transmitted is hv_{mn}, it can be calculated from the above that the intensity of light of frequency v_{mn} is given by

$$I_{mn} = \frac{1 - x}{1 + x} \cdot \frac{hv_{mn}\, pgw_{mn}}{\sigma_0 kT} \cdot e^{-E_m/kT} \tag{2.22}$$

Since the degree of ionisation of the atoms in question depends on the temperature of the plasma, it becomes evident from the above equation that the intensity of a given spectral line depends not only on the temperature of the plasma but also on the pressure of the gas. The value of the latter is influenced by the concentration of the element under determination in the electrode. It can, therefore, be concluded that the intensity of a given spectral line depends partly on the temperature of the arc and partly on the concentration of the element.

It follows from Eq. (2.22) that the relative intensity of two spectral lines of a given element is

$$\frac{I_1}{I_2} = \frac{g_1}{g_2} \cdot \frac{w_1}{w_2} \cdot \frac{\nu_1}{\nu_2} \cdot e^{-(E_1 - E_2)/kT} \tag{2.23}$$

That is, the relative intensity depends on the relative statistical proportions of the two lines, the relative probability of the transitions, the ratio of the wave numbers of the lines, the difference between their excitation energies, and, finally, the temperature of the arc plasma.

The arc plasma emits not only line spectra but also a *continuous spectrum.* The origin of the latter can be traced back to three processes. Radiation of the continuous spectrum is emitted when the electrons are recombined with ions or when the kinetic energy of two colliding electrons is transformed into electromagnetic energy. The third possibility is when the energy of the excited atoms is approximately equal to the ionisation energy. In all three cases the intensity of the emitted light is proportional to the square of the number of charged particles. In the plasma the concentration of the electrons and that of the positively charged particles are equal. Thus for the intensity of the emitted continuous spectrum the following formula can be given

$$I_{\text{cont}} \sim n_e^2 \sim p e^{-eU_i/kT} \tag{2.24}$$

The value of U_i should be somewhat smaller than the ionisation potential. Analogously to the relationship given by Eq. (2.21) the following formula can be written for the intensity of the continuous radiation

$$I_{\text{cont}} = A' p e^{-eU_i/kT} \tag{2.25}$$

where A' summarises the constants in Eq. (2.25). The intensity of the radiation of the continuous spectrum therefore increases with the temperature of the plasma and indirectly with an increasing arc current.

Reference

1. BOUMANS, P. W. J. M., *Theory of Spectrochemical Excitation*, Hilger and Watts Ltd., London (1966)

2.2.6 Relationship between the temperature of the arc plasma and the degree of ionisation

The character of spectra is determined by the degree of ionisation, x, of the atoms in the plasma, or in other words by the number of ionised atoms, n_i, in unit volume when compared with the number of atoms originally present, n. In a unit volume

of the plasma there are, therefore, $n_i = n_e = x.n$ ions or electrons. The number of the non-ionised atoms is $n_0 = (1 - x)n$, if it is assumed for the sake of simplicity that the ions are singly charged. The partial pressure of these particles, that is, the value of the *electron pressure*, p_e, or the *atom pressure*, p_0, is

$$p_e = p_i = xnkT \qquad \text{and} \qquad p_0 = (1 - x)nkT \tag{2.26}$$

where k is the Boltzmann constant and T is the absolute temperature of the plasma. The total pressure of the gas is therefore

$$p = p_e + p_i + p_0 = (1 + x)nT \tag{2.27}$$

The relative pressures are

$$p_e = p_i = \frac{x}{1 + x} p, \qquad \text{or} \qquad p_0 = \frac{1 - x}{1 + x} p \tag{2.28}$$

It follows from the above that the degree of ionisation is a function of temperature and pressure. This relationship is given by the Saha equation.[1] By using an appropriate thermodynamic calculation the following result is obtained

$$\frac{x}{1 - x} p = \frac{(kT)^{5/2}(2\pi m)^{3/2}}{h^3} \, e^{-eU_1/2kT} \tag{2.29}$$

where k is the Boltzmann constant, h is the Planck constant, and m is the mass of the electron. As the equation shows, the degree of ionisation is greatly enhanced when the temperature increases. Its value however is generally smaller than 1 % at the temperature of the arc plasma, although above 10 000 K it increases steeply.

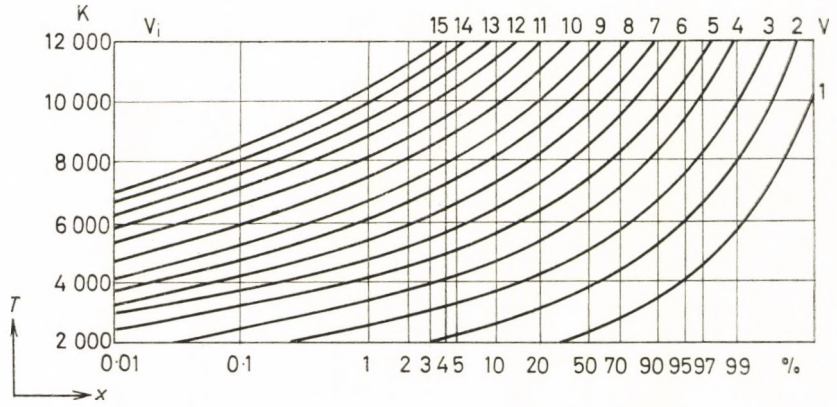


Fig. 2.8 Variation of the percentage ionisation (X) with temperature (T) at different ionisation voltages. Pressure is atmospheric[2]

On the other hand, at constant temperature the value of the degree of ionisation decreases with increasing pressure. In vacuum, therefore, there is greater ionisation than at atmospheric pressure, other factors being equal. At constant pressure and temperature the higher the ionisation potential, the lower the degree of ionisation (Fig. 2.8).

For practical purposes the above relationship is too complicated.[2] In most cases the value of the degree of ionisation, or when it is known, the value of the corresponding temperature, can be obtained to a sufficient degree of accuracy by the following simplified relationship which is applicable at atmospheric pressure

$$\log \frac{x}{1-x} = 2{\cdot}5 \log T - \frac{5040\, U_i}{T} - 6{\cdot}5 \tag{2.30}$$

References

1. SAHA, M. N., *Phil. Mag.*, 6th ser., **4**, 472 (1902)
2. AHRENS, L. H., and TAYLOR, S. R., *Spectrochemical Analysis*, 2nd edn, London–Paris (1961) p. 25

2.2.7 Measurement of the plasma temperature by spectra

In order to measure the plasma temperature, various procedures based on different principles can be used. Among these procedures the most simple are those that do not require additional equipment, but where the temperature can be directly established by using the spectra produced during analysis.[1–6]

The so-called *two-line method* meets the above requirement and is used generally for temperature measurements. The method is based on the observation that the relative intensities of two lines of an element selected for the measurement depend on the temperature of the plasma. The relationship, obtained by a transformation of Eq. (2.23) based on the Boltzmann law, is

$$\frac{I_1}{I_2} = \frac{w_1}{w_2} \cdot \frac{g_1}{g_2} \cdot \frac{\lambda_2}{\lambda_1} \cdot e^{-(E_1 - E_2)/kT} \tag{2.31}$$

When the Boltzmann constant is substituted into the above equation the following formula is obtained for the temperature of the plasma

$$T = \frac{5040(E_1 - E_2)}{\log w_1/w_2 + \log g_1/g_2 - \log \lambda_1/\lambda_2 - \log I_1/I_2} \tag{2.32}$$

where E_1 and E_2 are the excitation potentials in eV of the two spectral lines, respectively, w_1 and w_2 are the transition probabilities of the two lines, respectively, g_1 and g_2 are the statistical weights of the excited states, respectively, λ_1 and λ_2 are the respective wavelengths of the lines, and I_1 and I_2 are the corresponding intensity values. Provided the above constants of the selected lines are known, then from the intensity relation of the two lines the plasma temperature can be calculated.

In the application of the two-line method the following requirements emerge for the selection of the line pairs:

(1) The intensity relationships of the chosen lines should be very sensitive to the temperature changes of the plasma.
(2) The line intensity should not be influenced by any effects other than those of temperature. Consequently, the following types of lines cannot be used:
   (a) those for which the transition probability is influenced by the electric field; that is, for which the Stark effect leads to profile broadening of the spectral line, and

(b) those whose intensity is influenced by self-absorption, even if this effect is small.

(3) The excitation energies of the chosen lines should be as different as possible. The attainable precision of the measurement is greatly increased the larger this difference is.

(4) If the sample to be examined does not contain an element that meets the above requirements, then it is not enough that an element added to the sample for the measurement meets these requirements. Furthermore, it is necessary that the ionisation energy of the added element is high, so that its presence would not influence the plasma temperature to any degree.

The above requirements are met by many line pairs, although the method using the intensity relations of the line pair for magnesium at 2782·97 and 3336·68 Å has proved the most dependable.[6] In this case it follows from Eq. (2.32) that if the logarithmic difference of the intensities is

$$\log I_1 - \log I_2 = \Delta Y_{1,2}$$

then

$$\Delta Y_{1,2} = \log \frac{177}{13{\cdot}7} + \log \frac{3336{\cdot}7}{2783{\cdot}0} - \frac{5040.0{\cdot}75}{T}$$

that is

$$T = \frac{3770}{1{\cdot}19 - \Delta Y_{1,2}} \text{ K} \tag{2.33}$$

The solution of the problem is aided by Fig. 2.9.

In principle, the two-line method only yields precise data when the light source is homogeneous. This requirement is not fulfilled in the case of the arc, or for the other sources used in spectrochemical analysis. Better precision is obtained if the excitation energies of the lines are as different as possible. Most of the spectral lines having different excitation potentials are emitted from different locations in the

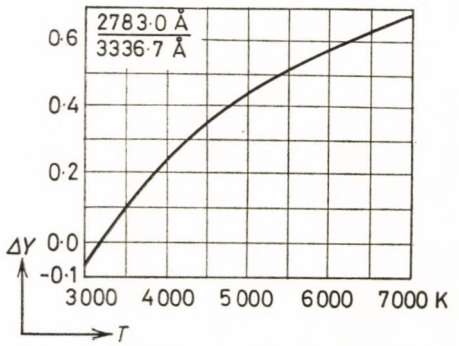


Fig. 2.9. Variation of the logarithmic intensity ratio (ΔY) for the Mg 2783·0/Mg 3336·7 Å line pair with the absolute temperature (T) of the plasma

plasma which have different temperatures. Thus measurements at different heights in the arc gap give different results. Unfortunately, the precision of the measurements is influenced by the uncertainty of the values of the transition probabilities as well.[7] In spite of all these drawbacks the precision attainable by the method is generally sufficient for practical purposes. In practice it is normally enough to know the relative plasma temperature. At present there are 70 elements in the spectra for which the relative transition probability of several lines is known with adequate precision.[8]

References

1. ORNSTEIN, L. S. and BINKMANN, H., *Physica, 's Grav.*, **1**, 799 (1934)
2. ORNSTEIN, L. S. and ROLLWAGEN, W., *Spectrochim. Acta*, **1**, 66 (1941)
3. KATZ, E. and ORNSTEIN, L. S., *Physica, 's Grav.*, **8**, 10 (1941)
4. DIKHOF, J. A. M., *Spectrochim. Acta*, **11**, 162 (1957)
5. SMITH, I. A. M, *Physica, 's Grav.*, **12**, 683 (1964)
6. HEGEMANN, F. and v. SYBIL, C., *Z. Wiss. Photogr.*, **50**, 463 (1955)
7. KESSLER, W. and DOBENECK, V. D., *Z. analyt. Chem.*, **198**, 35 (1963)
8. CORLISS, C. H. and BOZMANN, W. R., 'Lines of Seventy Elements', *Nat. Bur. Stand. Monogr.* (Phys. Chem.) **53** (1962)

2.2.8 Arc temperature and the first ionisation potential of the elements

It can be shown experimentally that there is a practically linear relation between the arc temperature as measured in absolute degrees and the first ionisation potential of the elements, that is

$$U_1 \approx KT \tag{2.34}$$

where K is a constant whose value can be established experimentally (Fig. 2.10). The values of the arc temperature belonging to the ionisation potential of several elements are summarised in Table 2.3.[1] The percentage degree of ionisation of some elements at different arc temperatures is shown in Table 2.4.[2]

It can be seen from Tables 2.3 and 2.4 that at relatively low arc temperatures there is only a slight ionisation of the non-metallic elements, semi-metals, and even some of the metallic elements. It is also evident that the ionisation degree of the technically more important metals is almost the same. The first ionisation potential of these elements is, with a few exceptions, around 6–9 eV. Therefore, the conclusion is that the spectra of all metals present in the material under study can be, in principle, *simultaneously* excited, provided their concentration is above a certain limit.

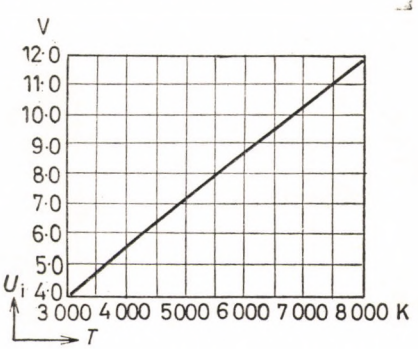


Fig. 2.10 Relationship between the ionisation potential (U_i) and the arc temperature (T), (The spectroscopically determined values refer to a 7 A current carbon arc)[1]

The above-mentioned regularities are even more evident if the first ionisation potentials of the elements are grouped according to the periodic system of the elements (Table 1.7). From such a presentation it becomes clear that the lowest excitation energies belong to the alkali metals, while the highest ones characterise the noble gases. For elements belonging to the same period, the first ionisation potential, as well as the lowest excitation energy, generally increases with the atomic number of the element in question. This observation can be interpreted by pointing out that in the alkali metals belonging to the first group there is only one relatively

loosely bound electron in the valence shell, while the outermost shell of the noble gases is saturated.

From Table 1.7 it can be seen that for elements belonging to the same group, the smaller the first ionisation potential and the lower the excitation energy is *generally*, the higher is the period number which characterises the element. This regularity can be illustrated by the Bohr atomic theory, which states that the number of periods is the same as the number of the electron shells belonging to those periods. The greater is the value of the latter, the further are the valence electrons to be found from the nucleus and thus the lower is the energy required to ionise

Table 2.3

The first ionisation potential of some elements (U_i) and their arc temperature (T) determined from the diagram in Fig. 2.10

Element	U_i (V)	T (K)	Element	U_i (V)	T (K)	Element	U_i (V)	T (K)
Cs	3·90	2 900	Zr	6·96	4 900	Si	8·15	5 700
K	4·34	3 200	Mo	7·18	5 200	B	8·30	5 900
Na	5·14	3 700	Sn	7·33	5 200	Sb	8·64	6 000
Li	5·39	3 800	Pb	7·42	5 200	Pt	8·97	6 200
La	5·6	4 000	Mn	7·43	5 200	Au	9·23	6 400
Al	5·98	4 300	Ag	7·58	5 400	Be	9·32	6 400
Ca	6·11	4 300	Mg	7·64	5 400	Zn	9·39	6 500
Y	6·57	4 700	Cu	7·72	5 400	As	9·81	6 700
Ce	6·6	4 700	Co	7·86	5 500	Hg	10·44	6 900
Cr	6·76	4 800	Fe	7·90	5 500	P	10·95	7 500
Ti	6·83	4 800	W	7·98	5 600	C	11·26	7 700

Table 2.4

Approximate percentage ionisation of the more important elements at different arc temperatures

Elements	Arc temperature (K)				
	4 000	6 000	8 000	10 000	12 000
K	3	40	85	97	>99
Na	1	20	70	95	98
Al, Ca	0·5	8	40	80	96
Ti, Cr, Zr	0·1	4	25	70	91
Sn, Pb, Mn, Mo	0·1	2	18	60	86
Ag, Mg, Cu	<0·1	1·5	15	50	83
Fe, Co, W	<0·1	1	12	38	80
Si, B		0·8	10	30	75
Pt		0·5	6	25	65
Au, Be, Zn		0·3	4	20	60
Hg, As		0·1	2	10	40
P		<0·1	1·2	8	30
C		<0·1	1	6	22

or excite the atoms. The conspicuously high ionisation energy of hydrogen is explained by the fact that the valence electron of hydrogen is orbiting very close to the nucleus on a relatively small orbit.

The line in Table 1.7 that separates the non-metallic elements from the others demonstrates that the first ionisation potential of the definitely metallic elements is essentially lower than that of the metalloids. Thus, in the excitation of elements like carbon, nitrogen, oxygen, phosphorus, sulphur, which have a non-metallic character, certain difficulties will be encountered such that normal methods of excitation cannot be used.

References

1. SEMENOVA, O. P., *Dokl. Akad. Nauk SSSR*, **51**, 683 (1946)
2. MASON, R. C., *A.S.T.M. Spec. Tech. Publ.*, No. **76** (1948)

2.2.9 Simultaneous element excitation

The relationships given in the previous chapters are, strictly speaking, valid only when the material to be excited contains one element. In the case of composite systems composed of several elements, the degree of ionisation is determined simultaneously by all the elements present in the plasma. Thus all elements play a role in establishing the relative intensities of the spectral lines appearing. These relationships are not only extremely complicated, but contain parameters whose values are generally not known. The arc temperature is determined not only by the processes taking place in the plasma, but also by the chemical and physical properties of the material to be studied. Moreover, the quality and pressure of the gas surrounding the plasma are also important. It is understandable therefore that only very general rules can be theoretically established for the excitation of composite systems and experimental observations can only explain their gross features.

From the point of view of spectrochemical analysis it is very important to remember that both the average intensity of the spectra of each element present in the sample and the intensity of the lines are more or less influenced by the other elements present. Therefore, it is essential to know what changes occur in the concentration of the latter. This phenomenon is referred to as the *interelement effect*.[1]

Interelement effects can be grouped into two categories. The first group is composed of those where the concentration of the vaporised elements in the plasma is not quite the same as the chemical constitution of the sample. These are known as *vaporisation effects*. The second category arises because thermodynamic laws interfere with the relationships existing between the intensity of spectral lines of certain elements and the concentration of the elements in the plasma. These are called *thermodynamic effects*.[2]

In connection with vaporisation effects it can be experimentally established that during arc excitation the greater the weight loss of the homogeneous sample, the greater is the chemical affinity between the sample and the surrounding gas.[3] Furthermore, it is evident that the greater the loss of weight of different samples, the smaller is the enthalpy of vaporisation.[4,5] From these observations it follows that in the examination of composite samples, the elements enter the plasma with variable velocities which may depend on their boiling point and the form of the surrounding

gas, etc. A difference can therefore be brought about between the chemical composition of the sample and that of the plasma gas. This discrepancy is less evident, however, the higher the energy available for vaporisation. At a sufficiently high excitation energy level elements will enter the plasma independently of their location in the periodic system, such that the surrounding gas only dilutes their vapour.[6]

In the examination of the thermodynamic effects it should be remembered that all relationships cease to exist between the electrons in the plasma and the dissociated atoms that produced the electrons in question. This fact implies that the value of the temperature T in the Saha equation (Eq. 2.29), valid for separate individual elements, is identical for all elements.[7] This is only achieved when the total electron pressure (p_e) that is composed of the contributions of the elements present is proportional to their ionisation.

The above statement implies that at comparatively low plasma temperatures, for example in the presence of three elements of variable ionisation potential, the total electron pressure is governed mainly by the element having the lowest ionisation potential, even when it is present at a very low concentration. When the temperature is increased, however, the electrons necessary for producing the electron pressure at the given temperature are provided mainly by the element that has medium ionisation potential and is of higher concentration. When the temperature is further increased this role will be assumed by the element with the highest ionisation potential, provided it is present at a high enough level of concentration.

The relationship between the relative ion pressure, p_e/p, of different elements and the plasma temperature, T, is depicted by Fig. 2.11. As can be seen, when the plasma temperature is higher than 10 000 K the relative ion pressures of elements of different ionisation potential are approximately the same. This is the theoretical basis of the observation that at temperatures above 10 000 K (spark plasma temperatures) the interelement effect is generally small. The latter cannot be neglected when it is caused by vaporisation effects and not thermodynamic ones. At temperatures in the range 3000–10 000 K (arc excitation) the interelement effect, therefore, cannot generally be neglected.[2] This effect becomes important the lower the ionisation potential of the disturbing element (Fig. 2.12).

It follows from the Saha equation that the total radiation intensity in the line spectra of a given element is given by the formula

$$I_\Sigma = C\frac{1-x}{1+x} \cdot \frac{h\nu_a p}{kT} \cdot w_a e^{-eU_a/kT} \tag{2.35}$$

where the average frequency of the radiation of the element is ν_a, the mean probability of its spectral lines is w_a, the mean excitation potential is the ratio of the excitation energy to the electronic charge $U_a = E_a/e$, while C is a constant characteristic of the element in question.[8]

From the above equation the following inferences can be drawn, which are also proved by practice.

The total intensity of the light radiated by the easily excited elements at lower plasma temperatures is greater than that of elements difficult to ionise, even when the pressures, p, of the elements' vapours are identical. This happens because the mean probability of spectral lines of elements that can be easily excited is relatively great, while their mean excitation potential is relatively low.

The previously given equation also demonstrates the theoretical explanation for the practical observation that it is not necessary for the spectral lines of an element that is present only in trace amounts to appear in the spectrum, even when the excitation potential of this element is advantageously low. The vapour pressure of such an element may be so low that for small values of atomic density the excitation

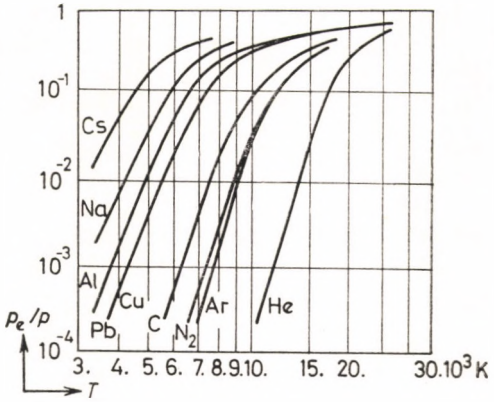


Fig. 2.11 Variation of the relative ionic pressure (p_e/p) of some pure elements with the plasma temperature[2]

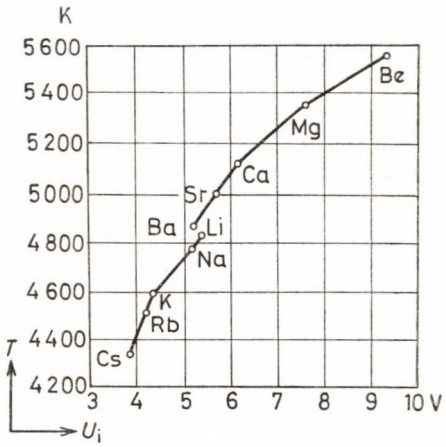


Fig. 2.12 Variation of the average arc temperature of the Corliss copper arc[9] with the ionisation potential of different metals[10]

probability of its spectral lines is also low. As a consequence of this fact the average intensity of the spectrum of such an element is remarkably low as well. From the practical point of view, if it is not possible to find the lines of an element in the spectrum, this fact in itself cannot be regarded as proof of the absence of the element in the sample.

The above discussion can be used for interpreting the so-called *buffer effect*.[9] This manifests itself in the phenomenon that when the sample contains an easily excited and highly volatile element at a sufficiently high concentration, the inter-element effect is practically eliminated. In this case the partial vapour pressure of the 'buffer element' is so high in the gas of the sample that its value is practically unchanged, even if the partial pressures of the other element present are changing. Thus the temperature and ionisation degree of the plasma are almost exclusively determined by the partial vapour pressure of the buffer element. The changes in the concentration of the disturbing element do not significantly influence either the

temperature of the plasma or the intensity relationships of the spectral lines of the elements studied.

Equation (2.35) provides the explanation for the following observation as well. When a sample to be excited has the same composition as that of the original material, except for a higher concentration of the easily excited element, then the spectrum is more atomic in character and the intensity of the lines of elements having higher excitation potential is smaller. This occurs because in the presence of the more easily excited element, the elemental composition of the sample vapour is changed in such a manner that the partial pressure of the element originally present in the sample is reduced. This, in turn, leads to a reduction of the total radiation intensity of these elements according to Eq. (2.35). The change in the character of the spectrum, on the other hand, can be explained by the fact that the presence of the easily excited element leads to a cooling of the plasma due to the buffer effect and thereby (according to the Saha equation) to a reduction of the ionisation degree of the plasma. In this way the ionisation of atoms of the elements originally present in the plasma is restricted, so the probability of the excitation of atomic lines becomes higher than that of ion lines. This is the reason why the spectrum becomes more like an atomic one.

References

1. WOHNBANK, F., *Metallk. Giesswes.*, **32,** 430 (1940)
2. KREMPL, H., *Z. analyt. Chem.*, **198,** 21 (1963)
3. SCHÖNTAG, A., *Microchim. Acta*, **1955,** 376
4. BUNGE, K., *Spectrochim. Acta*, **10,** 133 (1957)
5. RICHTER, G., *Proc. N VIII. Coll. Spectr. Inter.*, Luzern–Aarau, 1959
6. VAN CALKER, J., *Proc. VIII. Coll. Spectr. Inter.*, Luzern–Aarau, 1959
7. LOCHTE-HOLTGREVEN, W., *Z. analyt. Chem.*, **198,** 1 (1963)
8. WEIZEL, W. and ROMPE, R., *Theorie elektrischer Lichtbögen und Funken*, Leipzig, 18 (1949)
9. CORLICH, C. H., *J. Res. Nat. Bur. Stand.*, **66A,** No. 1 (1962)
10. KESSLER, W. and DOBENECK, D., *Z. analyt. Chem.*, **198,** 35 (1963)

2.2.10 Contours of spectral lines[1–20]

The contour or profile of a spectral line is the intensity distribution along the wavelength (or wave number) scale. The profile of a spectral line is therefore given by the intensity–wavelength curve. In this sense the spectral line of pure monochromatic light would be a perpendicular line. In fact spectral lines have a definite *width* and a characteristic *form*.

The physical profile of a spectral line is determined by the intensity distribution of the light emitted by the monochromatic light source. Under ideal conditions, that is, when no disturbing effects operate, the form of the spectral line is given by the Gaussian distribution curve. The half-width of the spectral line is equal to the wavelength difference at the half-height of the distribution curve and is denoted by $\Delta\lambda_{1/2}$ (Fig. 2.13).

The *natural line width*,[1–4,20] according to classical physics, originates from the radiation energy dissipated by the electron, as a harmonic oscillator with charge e, mass m, and frequency ν_0. On this basis the intensity distribution and the half-width

of the radiation can be established

$$\Delta\lambda_{1/2\,(\mathrm{nat})} = \frac{4\pi e^2}{3mc^2} \tag{2.36}$$

By substituting the values of the constants it can be calculated that the natural line width, that is, the smallest possible half-width of the spectral lines, is $1{\cdot}19.10^{-4}$ Å. This exact value applies only to resonant lines (see Section 2.2.11). This value for the natural line width is extremely small compared to the wavelength of the light although measurable in the X-ray domain. The natural line width and line profile are only of theoretical interest. In reality the form of the spectral line is always influenced by a variety of different effects. As a consequence the profile of the spectral line is symmetrically or asymmetrically broadened and the intensity maximum is shifted to longer or shorter wavelengths.

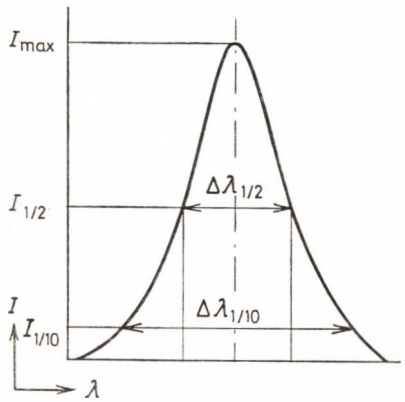


Fig. 2.13 Natural shape of spectral lines

Doppler broadening is caused by the thermal motion of the light-emitting particles in the plasma.[5,20] The *Doppler effect* arises from the fact that the wavelength of the moving light source, that originally has the value λ_0, changes proportionally with the velocity of its motion. In the stationary state the radiating atoms move in all directions in space. Particles moving with different velocities in different directions will emit wavelengths of slightly different values. It can be theoretically established that spectral lines retain, under such conditions, their Gaussian distribution and the extent of line broadening can be given by the expression

$$\Delta\lambda_{1/2\,(\mathrm{Doppler})} = 0{\cdot}716.10^{-6}\lambda_0\sqrt{T/A} \tag{2.37}$$

As a consequence of the Doppler effect the intensity maximum of the spectral lines is not shifted although their profile is broadened. The greater the broadening, the longer is the wavelength, the higher is the temperature of the plasma, and the lower is the atomic weight, A, of the radiating atom. For example, the half-width of spectral lines at 3000 Å and 6000 Å for an element of atomic weight of about 50 is increased by 0·02 Å and 0·04 Å, respectively, when the excitation temperature is 5000 K. For comparison, the broadening of the spectral lines at the same wavelengths in the case of the lightest element, hydrogen, at an excitation temperature of 300 K, is 0·04 Å and 0·08 Å, respectively.

In line broadening under the Stark effect a distinction should be made between broadenings caused by constant strength and varying strength electric fields.[6,20] *Stark broadening* is caused by the splitting of the terms and the spectral lines in the electric field and by the shift of the mid-point intensity of the lines.[5] This shift is proportional to the field strength in the case of *the first-order or linear Stark effect*, while it is proportional to the square of the field strength in the case of the *secondorder or quadratic Stark effect* caused by the force field of the ions.[7] In the practical application of spectral analysis Stark-type line broadening should be taken into account only when d.c. excitation is used. In this case Stark broadening is responsible for a negligible change of the line profile only. (In such excitation methods the field strength does not usually exceed 100 V/cm.) On the other hand, for discharge-tube excitation, the first-order Stark effect leads to line shifts which are not negligible, although the second-order Stark shifts are still hardly noticeable. For example, the effect of a field strength of 1000 V/cm [8] can shift the 4340 Å line of hydrogen by 0·47 Å for the first-order Stark effect, while the second-order Stark effect leads to a shift of only 0·00035 Å for sodium *D*-lines.

Contrary to the line broadening caused by Stark effects, there is a much greater profile change brought about by the *internal Stark effect*.[7,9,10] This arises from the fluctuating microfields brought about in the plasma, where the radiating atoms are surrounded by charged ions and electrons. The statistical distribution of the fields arising this way leads to the splitting of the spectral lines and consequently results in the broadening of the latter. It can be theoretically proved that a different mechanism causes the broadening due to the ions relative to the electrons. The slow-moving ions establish a 'quasi-constant' microfield giving rise to 'statistical' line broadening, while the fast-moving electrons lead to the so-called collision broadening. These two effects simultaneously determine the profile of the spectral lines.

The above theory is supported by the fact that the theoretically calculated profile of the Balmer series hydrogen lines, taking into account the linear Stark effect, is in satisfactory agreement with the experimental findings. It can furthermore be established that spectral lines broadened according to the first-order Stark effect easily undergo broadening even when the ion concentration is relatively low. The internal Stark broadening arising in this way is generally greater than all other types of broadening originating from other effects.[11]

It can be shown theoretically that the line broadening caused by the second-order (ionic) Stark effect depends on the relative velocity, v, of the atoms and ions and on the ion concentration, n, provided the half-width of the line is given by

$$\Delta\lambda_{1/2\,(\mathrm{Stark})} = a_1 \frac{\lambda_2}{2c} C^{2/3} v^{1/3} n \tag{2.38}$$

and the line shift is given by

$$\Delta\lambda_{(\mathrm{Stark})} = a_2 C^{2/3} v^{1/3} n \tag{2.39}$$

where C is the Stark constant and a_1 and a_2 are the constants corresponding to each equation.[6,12,13] Experimentally it is found that the ratio a_1/a_2 has a value in the range 2–10 as opposed to the theoretical value. The ion concentration values for spark and other excitation methods are not known[9] and the above relationships cannot be generally used for the calculation of line broadening and line shifts. They

can still be utilised, however, for drawing inferences from the experimentally obtained line profiles with respect to the ionic density and plasma temperature, etc. The above two equations also illustrate the experimentally observed fact that due to the internal Stark effect the line broadening essentially increases with increasing wavelength, while the line shift is independent of the wavelength.

The internal field strength can be as high as 10 000 V/cm, although in the case of spark excitation or relatively high current arc excitation it can reach several times this value. The line broadening caused by this factor has the order of magnitude of an ångström.[14] It can therefore be established that the greatest line broadening is caused by the Doppler effect and by the internal Stark effect. In the case of the Doppler effect the line profile is symmetrical, while the asymmetrical line profile and line shift are caused by the Stark effect.[11]

Among the other effects leading to the distortion of the spectral lines is the so-called *pressure broadening* effect. In this case a distinction should be made as to whether it is caused by the pressure of the gas itself or by the pressure of a foreign gas. Both effects are manifested by an asymmetrical distortion of the line profile and a shift of the intensity centre to the longer wavelengths.[5]

Pressure broadening can be explained by pointing out that in the plasma neighbouring atoms mutually disturb one another's electrical fields, a phenomenon especially important when the atoms collide. One of its results is that the average lifetime of the *excited* atoms or molecules is reduced. The collision causes the atoms to emit light. This has the consequence that according to the Heisenberg uncertainty principle $-\Delta E = h/\Delta T-$ the line width of the emitted light is increased. (Here E stands for energy and T for temperature.)

In the Lorentz theory the increase in the half-width (in Å) is[15]

$$\Delta\lambda_{1/2\,(\text{pressure})} = \frac{nvd^2\lambda_0^2}{c}\,10^8 \tag{2.40}$$

The greater the pressure broadening, the larger the number of atoms in a unit volume of space, n, with an average velocity, v, and an effective collision diameter, d, and the longer is the wavelength of the light. A further consequence is that the pressure broadening is proportional to the gas density, and pressure, and also to the square of the absolute temperature. The effective collision diameter is several times greater than the gas kinetic diameter value (about 3–100 times as big) and according to the Bohr theory it increases sharply in the higher excitation levels of the atom.[16] It follows then that lines of the higher series are broadened more extensively than those of the lower series.

Furthermore, pressure broadening is of an order of magnitude of 10^{-4}–10^{-3} Å in the case of the usual arc and spark excitation and only *leads to a blurring of the spectral lines.*[8]

With *discharge tube excitation,* which takes place in a closed space, pressure broadening varies proportionally with the pressure of the gas itself. It also depends on the elementary composition of the gas and on the nature of the spectral line in question.[5] The broadening becomes asymmetric, however, only above 400 torr. The pressure shift amounts to about 0·03 Å/atm up to about 10 atm pressure. Foreign gases lead to such a change of spectral line profiles that the half-width of the lines changes proportionally with the gas pressure.[17] On the other hand, the line shift is proportional to the square of the gas pressure.[17] The density, N, of the charged

particles present in the gas leads to an alteration of the line broadening that is proportional to $N^{2/3}$.This proportion can be further influenced by the Stark effect of the line considered.

In the case of gas excitation, *resonance broadening* should also be taken into account. This emerges because of the great number of interactions between the excited and ground state atoms of the same species. These exchange forces are inversely proportional to the third power of the distance between the atoms and rapidly drop as the concentration is lowered. In accord with this fact it can be established both theoretically and in practice that the resonance broadening changes proportionally with the gas density.[18,19] The measure of broadening due to resonance effects is greater by several orders of magnitude than the pressure broadening caused by foreign gases. This is illustrated by the fact that the resonance broadening observed on lines in the visible region, for gases having a pressure lower than 10 torr, is about 0·1–0·2 Å.

In summary, it can be established that the lower the pressure and temperature of the gas and the lower the current density, the less marked the fine profile distortion becomes.

References

1. Bohr, N., Kramers, H. A., and Slater, J., *Z. Phys.*, **24,** 69 (1924)
2. Weisskopf, V. F. and Wigner, E. P., *Z. Phys.*, **68,** 54 (1930)
3. Margenau, H. and Watson, W. W., *Rev. Mod. Phys.*, **8,** 22 (1936)
4. Low, F., *Phys. Rev.*, **88,** 53 (1952)
5. Kohlrausch, F., *Praktische Physik*, 20th edn, Leipzig, vol. I (1955)
6. Stark, I., *Jb. Radioakt. Elektronik*, **12,** 349 (1916)
7. Lindholm, E., Dissertation, Uppsala (1942)
8. Polansky, S., *High Resolution Spectroscopy*, London (1947)
9. Vainstein, E. E., Kolosnikov, V. G., Mazing, M. A., Mandelstam, S. L., and Sobelman, I. I., *Bull. Acad. Sci. USSR, Phys. Ser.*, **22,** 718 (1958)
10. Sobolyev, I. I., and Vainstein, E. E. *Optika Spektrosk.*, **6,** 440 (1959)
11. Lochte-Holtgreven, W., *Z. anal. Chem.*, **198,** 1 (1963)
12. Weisskopf, V. F., *Z. Phys.*, **75,** 287 (1932)
13. Breene, R. G., *The Shift and Shape of Spectral Lines*, Oxford (1961)
14. Holtsmark, J., and Trumpy, B., *Z. Phys.*, **31,** 803 (1925)
15. Lorentz, H. A., *Proc. Sect. Sci. K. Ned. Akad. Wet.*, **8,** 591 (1906)
16. Bohr, N., *Phil. Mag.*, **26,** 1 (1914)
17. Schulz, P., *Z. Phys.*, **39,** 412 (1938)
18. Minkovski, R., *Z. Phys.*, **36,** 839 (1926)
19. Korff, V., *Astrophys. J.*, **76,** 124 (1932)
20. Walker, S., and Straw, M., *Spectroscopy*, London, vol. 1 (1961)

2.2.11 Self-absorption

In the above sections it has been tacitly assumed that the light emitted reaches the surroundings without loss. This assumption is not justified in general, since the light emitted during excitation could be reabsorbed by neighbouring atoms. All ground-state atoms of a given element can absorb light of the same wave number as that emitted by the atoms of the same element in the excited state. The size of the effective diameter of the absorbing atoms is approximately equal to the square of the wave number of the absorbed light. It can be calculated from this relationship that there can be no absorption when the number of atoms present in 1 cm^3 is smaller

than about 10^{10}. Practical observations show that no self-absorption can be detected even at much higher atom concentration values. In general, it can be stated that the plasma does not absorb light if the radiation density emitted by its surface is essentially lower than that of the blackbody.

Usually, the radiation strength of the arcs used in spectrochemical analysis is much smaller than the above-quoted value and the arc plasma can therefore be considered to be practically transparent. On the other hand, light absorption occurs when the plasma is surrounded by a cooler vapour layer that contains the ground-state atoms of the emitting element in higher concentrations.

It is known that the atoms can be excited not only by thermal, kinetic, or electrical energy, but also by light. The prerequisite is that the light to be absorbed should have the same wavelength as the light radiated upon excitation. This process, if it occurs, is called *self-absorption* and the radiation arising this way is called *resonant radiation.* The spectral lines that show this effect are called *resonant lines* (these lines are generally denoted by the letter *R* in spectral line atlases and tables).[1,2]

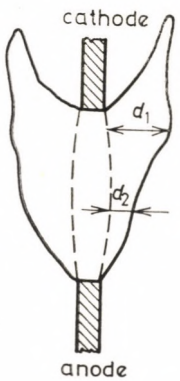


Fig. 2.14 Cross-section of a direct-current arc

In arc or spark excitation self-absorption is caused by the radially-disposed temperature gradient in the plasma. The light source proper is then surrounded by a lower temperature cloud containing ground-state atoms (Fig. 2.14) that absorb the light emitted by the plasma proportional to their concentration. The energy taken up during absorption is transformed mostly to resonant light radiation. As a consequence of this secondary light sources are produced in the cloud of atoms surrounding the plasma and radiate light in all directions. The intensity of light in the *original direction* is reduced by self-absorption in such a way that the following relation holds according to the Lambert–Beer law

$$I = I_0 e^{-ad} \tag{2.41}$$

where I_0 is the original light intensity, a is the absorption coefficient, and d is the path length of light in the absorbing medium.[3] The value of the absorption coefficient depends on the average concentration of the vaporised ground state atoms. Thus the greater the light loss caused by self-absorption, the higher is the mean concentration of the absorbing element in the vapour and the greater is its layer-thickness.

Self-absorption can also alter the profile of the spectral lines. Lines that have significant broadening due to Doppler and Stark effects are flattened by self-absorption, which increases their half-width still further (Fig. 2.15). The measure of this

self-reversal broadening is proportional to the square root of the number of absorbing atoms, $\sqrt{N}$, in the light beam. The value of N is also proportional to the partial pressure of the absorbing vapour and to the absorbing layer thickness.

Self-absorption not only has the consequence that the intensity of the emission lines is reduced and their half-width is increased, but it also increases the strength of the absorption line. Thus, a relatively narrow absorption line occurring in an atomic cloud of relatively low concentration cuts the broadened emission line into two halves (Figs 2.16 and 2.17). The half-width of the absorption line is also proportional to the value of $\sqrt{N}$. The wavelength of the absorption line, that is, the wavelength belonging to the minimum intensity of the inverted *Gaussian* distribution curve, completely agrees with that of the emission line. This has the practical advantage of making the determination of the wavelength of broad resonant lines relatively easy but it is always harmful from an analytical point of view.

The intensity of spectral lines continuously increases with the concentration of the element to be determined in the sample. This, however, is not the case for resonant lines, which due to the effects of self-absorption become less intense when the concentration of the element is increased. At high concentration levels it is possible that the intensity of the lines capable of self-absorption is not increased any further but is reduced. This phenomenon is called self-reversal (Fig. 2.16).

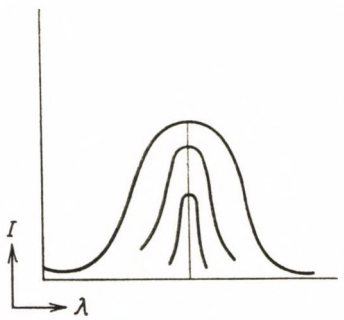


Fig. 2.15 Changes in the profile of spectral lines with increasing self-absorption

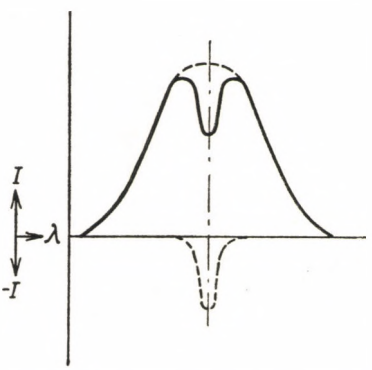


Fig. 2.16 Spectral line self-reversal (continuous line depicts the emission line profile and the dotted line the absorption profile)

The previous treatment suggests that self-absorption can only occur when the light absorbed by the atoms has the required energy. A further condition is that there should be a high enough concentration of the absorbing ground state atoms

surrounding the light source. The former requirement is generally met by the easily excited atomic lines, that is the fundamental lines, although some degree of self-absorption should be considered even in the case of the less easily excited atomic lines. On the other hand, the second requirement is met by those elements whose boiling point or sublimation temperature is low. Lines corresponding to transitions

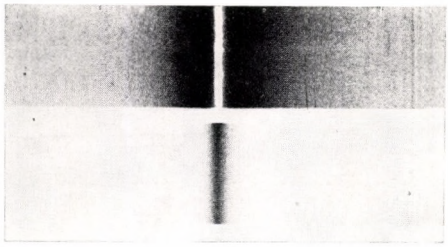


Fig. 2.17 The line Cd 2288·018 with (*a*) and without (*b*) self reversal broadening

between higher levels cannot be self-reversing, only those whose lower level is that of the ground state. It can therefore be stated that *the resonant lines of elements of low boiling point will usually show the effects of self-reversal and splitting.*

References

1. KAYSER, H. and RITSCHL, R., *Tabellen der Hauptlinien der Linienspektra aller Elemente*, Berlin (1939)
2. HARRISON, G. R., *Wavelength Tables of 100 000 Spectrum Lines*, New York (1939)
3. DIEKE, G. H. and CROSSWHITE, H. M., *J. Opt. Soc. Am.*, **23**, 425 (1943)

2.2.12 Background radiation

Line spectra are always accompanied by the continuous spectrum. The intensity measured at the location of a spectral line is composed of the true line intensity plus that of the background formed by the continuous spectrum.

In the complete spectra the intensity of the background alters with the wavelength such that no general relationship can be established between them. The background originates from a number of effects that are variously influenced by the experimental conditions. Among these the most important factors are:

(1) the continuum at the series limit of the spectral lines,
(2) the emergence of band spectra,
(3) the so-called free-free electron radiation, and
(4) the continuous spectra of incandescent solid, and liquid bodies. (This last effect is independent of the excitation and depends only on other experimental conditions!)

Continua at the series limit of spectral lines begin at definite wavelength values in the spectrum and cover only a narrow spectral region. Their intensity is relatively low, too. Their origin depends partly on the composition of the material studied and partly on the magnitude of the excitation energy. It can be generally stated that the series limit continua do not necessarily affect the state of the background.

Contrary to the above case, *band spectra* play an important role in the origin of the background. These bands are produced by excited gas molecules in the plasma and their presence depends on the chemical composition of the surrounding gas. Different bands emerge when the excitation is performed in the atmosphere or in an inert gas and when the surrounding gas forms certain compounds with the electrode material. It is generally observed that the background develops maximum intensity at the location of the band heads. This property is difficult to observe, or cannot be observed at all, for some compounds. Accordingly, it is only possible to have a schematic knowledge of the effect of molecular spectra upon the background, or at least to know the influence of the more important compounds. The most frequent bands are the following.[1]

The background in the *spectra of arcs burning in air* is due mainly to nitrogen. The most intense nitrogen bands are found at wavelengths 3159 and 3371 Å, whereas the NO bands are located at 2370, 2680, and 2722 Å. (The latter bands disappear when the excitation is performed in a nitrogen, or carbon dioxide atmosphere.) The band at 3062 Å, arising in the aureola of the arc, is very likely an OH band caused by the moisture content of the air.

In the *spectra of carbon arcs* the CN bands cause very short intervals of radiation between 3500 and 4600 Å, which lead to minor inequalities in the background intensity. These bands are in fact the extensions of the CN band system found in the longer wavelength regions of the spectrum.

The wavelength ranges of the most important CN bands are the following: 3590·0–3500·0; 3883·4–3730·0; 4216·0–4050·0; 4606·1–4500·0 and 4736·8–4600·0. It is to be noted that the upper limits indicate the band heads, while the lower limits are only approximate since they depend largely on the intensity of the source. Finally the intensity of the CN band corresponding to the last range is much lower than for the former ones.

The third source of the background is the so-called *free-free electron radiation.* This radiation is emitted during the free-free transitions of the electrons and in the recombination of the electrons with ions. The intensity of this type of background depends on the electron density of the plasma. It can be experimentally established that the intensity of the background is influenced not only by the electron density, but also by the presence of certain elements. There is a rule according to which the intensity of the free-free electron radiation is independent of the wavelength up to a certain *limiting wave number* value, ν_h, while at greater wave number values it drops according to the relationship, $\exp[-h(\nu_h - \nu)/kT]$[2,3] (Fig. 2.18). That part of the background arising from the free-free electron radiation increases exponentially with the decrease of wavelength up to a certain value and then remains steady at shorter wavelengths. Neglecting effects of secondary importance it can be calculated that this intensity, which is independent of the wavelength, is given by

$$I_{0,h} = K \frac{n_e \cdot n_i}{kT^{1/2}} \tag{2.42}$$

where n_e and n_i are the electron and ion density values, respectively. The value of the limiting wave number depends on the atomic number of the most concentrated element in the material under study.[4] On this basis the elements can be grouped into three classes:

(1) To the first group belong certain elements of the alkali metals, alkaline earths and Group III. These have the characteristic that the radiation intensity is found in the infra-red region of the spectrum.
(2) For elements in the second group the radiation intensity is found in the region of the spectrum that extends from the infra-red to 4000 Å. To this group belong the elements of Groups IV–VII and the rest of Group III.
(3) For the transitional elements, which belong to the third class, the radiation has constant intensity extending far into the ultra-violet region.

(It follows from the above, for example, that at some not-too-short wavelength, in the spectra of the alkaline earth elements the intensity of the background is lower than in the spectrum of iron.)

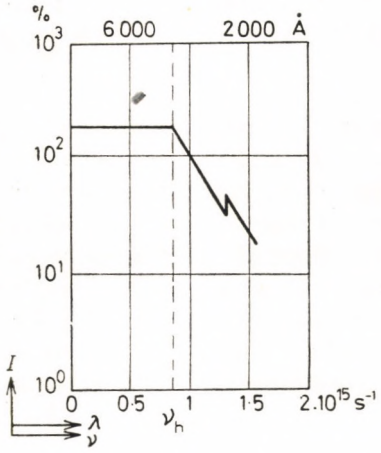


Fig. 2.18
Schematic variation of the relative intensity (I_{rel}) of the background with the wave number and the wavelength of light

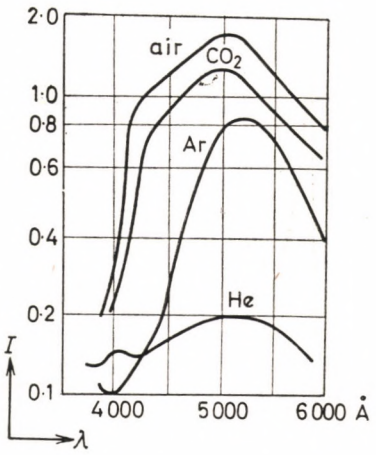


Fig. 2.19 Variation of the background intensity (I) for the carbon arc in different gaseous environments[5]

The intensity of the background can also be influenced, indirectly, by the experimental conditions as well. This is clearly shown by Fig. 2.19, which illustrates the effect of different gases surrounding the arc upon the background intensity. As can be seen, the background intensity is substantially reduced when the carbon arc is burnt in a noble gas atmosphere, instead of in air. This phenomenon originates in the fact that in the presence of noble gases molecular band spectra hardly appear and therefore cannot contribute to the background. The effect of the noble gases on the arc temperature may also play some role in this particular phenomenon.

References

1. Roes, R. L., and Smit, J. A., *Spectrochim. Acta*, **11**, 119 (1957)
2. Golling, E., *Spectrochim. Acta*, **11**, 125 (1957)
3. Lochte-Holtgreven, W., *Spectrochim. Acta*, **11**, 111 (1957)
4. Maecker, H., and Peters, T., *Z. Phys.*, **139**, 448 (1955)
5. Thiers, R. E., and Valle, B. L., *Spectrochim. Acta*, **11**, 179 (1957)

2.3 Electric parameters of arcs

2.3.1 Arc voltage and arc current[1–3]

For the conductance of an arc the voltage drop, U_a, between the electrodes in the arc gap depends on the current, I, flowing in the circuit, unlike that in the case of ohmic resistance. Therefore, it can be written that

$$U_a = U(I) \tag{2.43}$$

It can be experimentally shown that upon increasing the current, there is initially a steep increase that is followed later by a more gradual change of the arc conductance, until there is an approximately linear change of the conductance with the current intensity (see curve *a* in Fig. 2.20). At the same time the initial arc voltage drops sharply although later this becomes more gradual (curve *b*). It can be established experimentally that for carbon arcs the arc voltage is about 40 ± 3 V, while for metal electrodes it is usually 45–50 V, when the electrode distance is 2 mm and an arc current in the range 2–25 A. It can also be proved that the arc voltage depends on the chemical composition of the vapour. Thus in the presence of alkalis, the arc voltage drops to between 20 and 24 V.

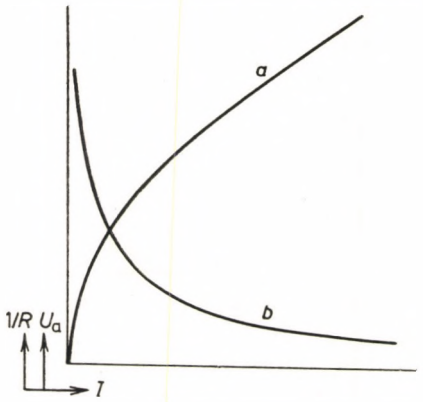


Fig. 2.20 Relationship between the arc current, I, and the conductivity, $1/R$, of the arc (curve *a*) and also the arc voltage, U_a (curve *b*)

The previously mentioned phenomena can be explained by the initial steep increase of the plasma temperature upon increasing the arc current. The reason for this is that initially the density of the charged particles in the plasma is very low and the resistance to the arc is accordingly large. Upon increasing the arc temperature, however, there is an increase in the degree of ionisation and the density of the conducting particles that increases the conductance of the plasma. The rising temperature of the plasma entails an increasing energy loss, brought about by the loss of heat via radiation, thermal conductance, etc. Consequently, a further increase in the current will

induce a less steep increase in the conductance of the plasma. Finally, a state is reached when a proportionality is established between the current and the temperature increase limited by heat losses.

Table 2.5

Values of the constants in the Ayrton relationship at different atmospheric pressures

Atmospheric pressure (torr)	*a*	*b*	*c*	*d*
760	35·7	3·0	114·8	1·8
740	38·5	2·15	54	6·1
200	35·5	1·84	38	8·3
50	33·7	1·51	26	10·8
10	27·0	1·35	19	13·4
5	23·7	1·20	0	15·7

Table 2.6

The Ayrton constants in air and different gaseous environments at 760 torr

Gas	*a*	*b*	*c*	*d*
air, stationary	35·7	3·0	114·8	1·8
air, streaming	44·1	2·6	17·8	1·8
argon	24·8	0·9	10·2	0·0
carbon dioxide	44·5	1·7	18·2	8·7
nitrogen	48·2	2·6	23·3	5·3

The relationship between the arc current and arc voltage can be calculated from the Ayrton equation, when the electrode distance, l, the atmospheric pressure, and the chemical composition of the gas surrounding the arc are all changing. The following relationship holds for the arc voltage

$$U_a = a + bl + \frac{c + dl}{I} \tag{2.44}$$

where I is the current, while a, b, c, and d are constants whose values for carbon arcs at different air pressure values are summarised in Table 2.5. In Table 2.6 values of these constants are listed for different gases at a pressure of 760 torr. A similar relationship can be obtained when the arc is created between different metal electrodes.

The relationship between the arc current and arc voltage can be established not only by the above formula, but also by using the Nottinghan equation as well.

According to the latter it can be shown that

$$U_a = a + \frac{b}{I^x} \tag{2.45}$$

where a and b are constants whose values depend on the form and relative position of the electrodes, and the value of the exponent, x, is determined by the material of the electrode. For example, x, is 0·67 for copper, 1·06 for carbon, and 1·38 for tungsten, etc. Using the value of the exponent for carbon electrodes it can be calculated that the relationship describing the change of the conductance of carbon arcs (that is, their reciprocal resistance) with the current is the following

$$\frac{1}{R} = \frac{I^2}{aI - b} \tag{2.46}$$

The Nottinghan equation (2.45) also shows that when the arc current increases the resistance of the arc drops. This has the consequence that the *arc is unstable when not controlled* and steadily increases.

According to the Kaufmann stabilisation condition, an arc will become stable when a resistance of R ohms is placed in the circuit of an arc operated by a source of terminal voltage, U_0, and where the following condition is met

$$R + \frac{dU_0}{dI} > 0 \tag{2.47}$$

In practical terms this means that the arc cannot be operated without an appropriately large series resistance (Fig. 2.21). Without such a resistance the current passing through the arc can, in principle, reach infinitely high values. With the proper choice of a resistance in series combination the arc current can be controlled.

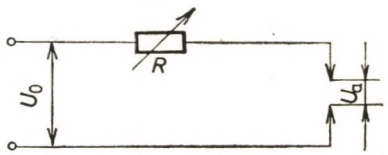


Fig. 2.21 Switching circuit of the electric arc

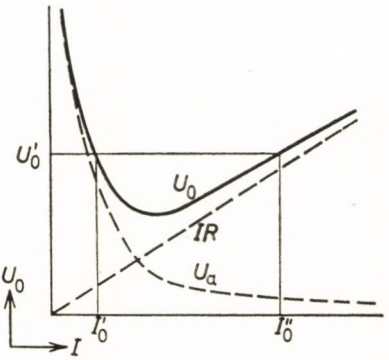


Fig. 2.22 Variation of the terminal voltage, U_0, of the arc circuit with the current, I

The resistance in the arc circuit is composed of two parts: the resistance of the arc and the ohmic resistance of the series resistance. The voltage between the input terminals of the arc circuit is therefore given by

$$U_0 = U_a + IR \tag{2.48}$$

where I is the arc current. It follows from the above relationship that the curve representing terminal voltage–current variation passes through a minimum (Fig. 2.22). Thus two current values (I'_0 and I''_0) correspond to the voltage U'_0 applied

to the circuit. The stationary condition can only be achieved at one of these current values.

Let us take I for the current value before the establishment of the stationary state, and I_0 for the stationary state.

Also, in both the stable and the unstable states, the voltage, U_0, applied to the circuit is equal to the sum of the voltage drops brought about by the elements of the circuit. Therefore, it can be written that for the unstable state immediately preceding the stable one, the terminal voltage of the circuit is

$$U_0 = U(I) + RI - L\frac{\mathrm{d}I}{\mathrm{d}t}$$

where L is the inductance and t is the time necessary for establishing the current value I. On the other hand, when the state is realised the following relationship holds

$$U = U(I_0) + RI_0$$

From the two equations above it follows that

$$U(I) - U(I_0) = R(I_0 - I) + L\frac{\mathrm{d}I}{\mathrm{d}t}$$

When it is considered that the difference between the voltages for the stationary state and the one immediately preceding is $(\mathrm{d}U)_0$, and the corresponding difference of the current values is $(\mathrm{d}I)_0$, then by modifying the previous formula we obtain

$$U(I) - U(I_0) = (I - I_0)\left(\frac{\mathrm{d}U}{\mathrm{d}I}\right)_0 = (I_0 - I)R + L\frac{\mathrm{d}I}{\mathrm{d}t}$$

As $(\mathrm{d}U/\mathrm{d}I)_0$ is the resistance, R_a, of the stationary arc, it can be established that

$$L\frac{\mathrm{d}I}{\mathrm{d}t} = (I - I_0)(R_a + R) \tag{2.49}$$

In the above formula, L, R_a and R are positive quantities. This accounts for the experimentally observed fact that when the arc is switched on, the current value corresponding to the stationary state is always reached from the direction of higher current values. (In Fig. 2.22, therefore, the larger of the current values, I_0' and I_0'', corresponds to the stable condition of the arc.)

When the distance between the electrodes is increased, the conductance of the arc is reduced while the arc voltage U_a is increased, although the latter change is not completely linear. The arc voltage changes very steeply when the gap between the electrodes is relatively small. It then slowly becomes less steep and upon reaching a certain limiting distance (4–5 mm) the change becomes linear (Fig. 2.23).

The value of the arc current for a given source voltage U_0 is determined from the value of the series resistance, R, and the arc resistance, R_a. The following relationship is valid

$$I_0 = \frac{U_0}{R + R_a} \tag{2.50}$$

This formula is of practical importance as it shows that fluctuations in the arc resistance are of less importance when the series resistance value is large or, indirectly,

when the terminal voltage of the source is greater. From the practical point of view *it is advantageous to have a high value for the series resistance and for the terminal voltage of the source.* On the other hand, it should be remembered that to obtain a certain value of the arc current, one needs a greater terminal voltage when the series resistance is large.

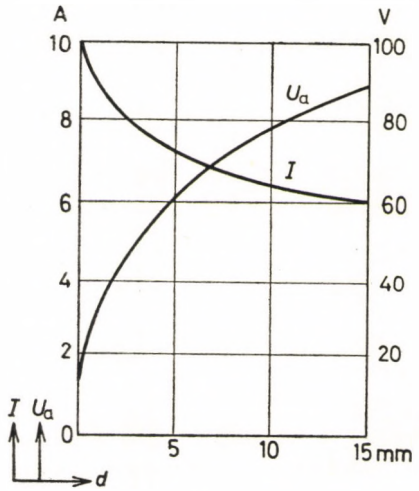


Fig. 2.23 Variation of the arc voltage, U_a and, arc current, I, with the electrode distance, d. Supply voltage is 220 V [3]

References

1. KOHLRAUSCH, F., *Praktische Physik*, 19th edn, Leipzig (1955)
2. SCHALLREUTER, W., in *Grimsehls Lehrbuch der Physik*, 13th edn, vol. 2 (1955)
3. WEIZEL, W. and ROMPE, R., *Theorie der elektrischen Lichtbögen und Funken*, Leipzig (1949)

2.3.2 Breakdown voltage

There is a certain minimum voltage, U_b, called the *breakdown voltage*, that is necessary to bring about an arc between two electrodes. The value of the latter changes with the distance, d, between the electrodes and also depends on the material, form of the electrodes, the pressure, temperature, and even on the ionisation conditions of the gas in the arc gap.[1]

It can be shown experimentally that the breakdown voltage changes rapidly when the electrode distance is increased. When the gap reaches a value of about 0·01 mm, the rate of increase becomes approximately linear. If the electrodes have a spherical surface and the pressure of the air at room temperature is 760 torr, then, as can be seen from Fig. 2.24, the breakdown voltage is 8–15 kV, for an electrode gap of 2–4 mm. Furthermore, it can be established that the *field strength* shows an opposite effect to the breakdown voltage, first dropping quickly when the electrode gap is increased and later showing only a slight decrease.

The value of the breakdown voltage is also influenced by the geometrical parameters of the front portions of the electrodes. The breakdown voltage tends to be much lower when the electrodes have an edge or are sharpened than when composed in spherical form. The difference with respect to the spherical electrode form is smaller if only one of the electrodes has an edge or point, while the other one has

a spherical or semi-spherical form. In the latter case, however, the polarity of the electrodes influences the value of the breakdown voltage and is lower when the sharpened or pointed electrode is the cathode. This happens because the inhomogeneous field created around the strongly curved surface of the electrode facilitates the removal of electrons.

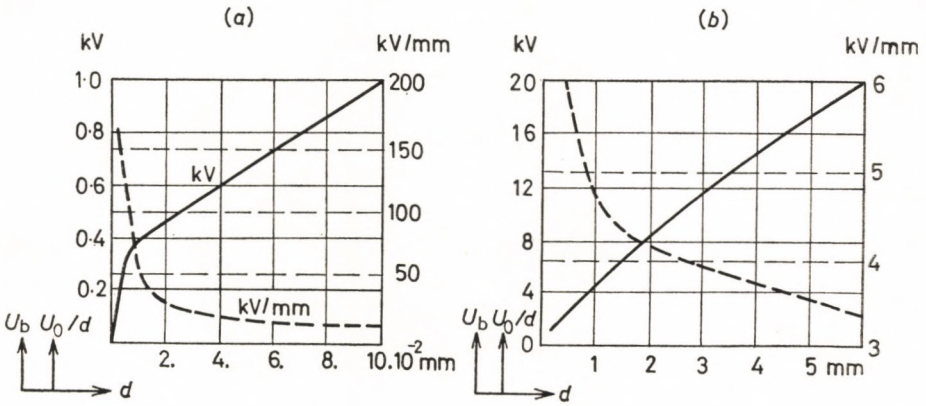


Fig. 2.24 Variation of the breakdown voltage, U_b, and field strength, U_0/d, for a spherical electrode at atmospheric pressure with variable electrode distance from 0·0 to 0·1 mm (*a*) and between 0·1 and 6 mm (*b*)

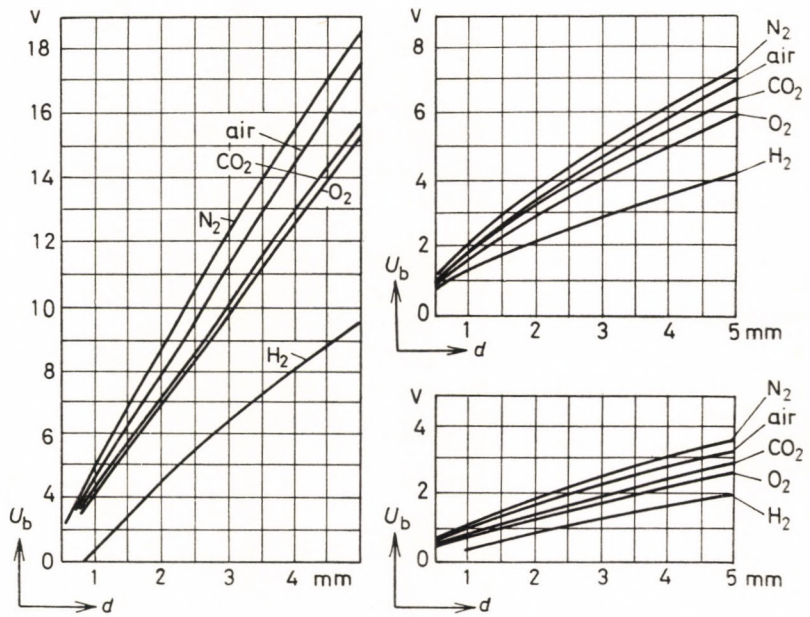


Fig. 2.25 Influence of different gases on the breakdown voltage, U_b, at variable electrode distances[1]

The influence of the gas in the electrode gap is illustrated by Fig. 2.25. As seen from the diagrams the breakdown voltage is highest in an atmosphere of nitrogen when compared to the values for air, nitrogen, oxygen, carbon dioxide and hydrogen.

There is an approximately linear decrease of the breakdown voltage on reducing the pressure (Fig. 2.26) for gases of different chemical composition. The moisture content of the air influences the value of the breakdown voltage only when the distance of the electrodes is greater than 10 cm, that is when the gap is much wider than is normally used in spectrochemical analysis. Thus humidity changes can be generally disregarded.

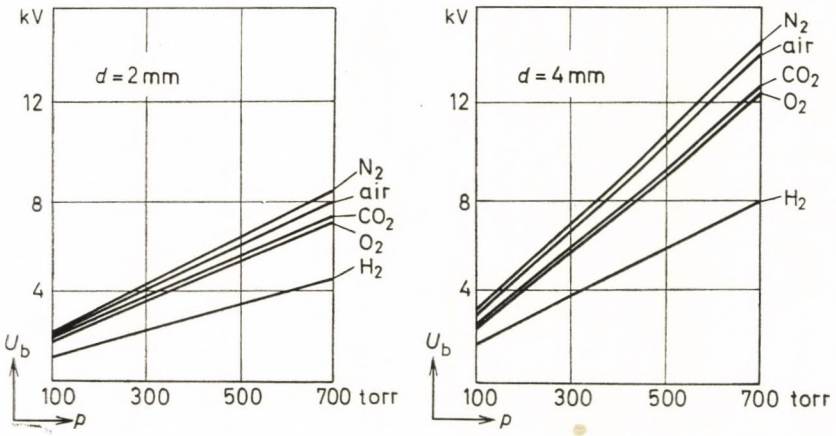


Fig. 2.26 Variation of the breakdown voltage, U_b, with the pressure of air or hydrogen at electrode distances of 2 and 4 mm, respectively[1]

Strictly speaking the above statements apply only to d.c. arcs. It can be proved that slight changes occur in the value of the breakdown voltage if the arc is established by an a.c. supply of frequency lower than 10^7 Hz.

The arc is '*self-ignited*' when the voltage applied to the electrodes is higher than the breakdown voltage for a given arc-gap value. If this voltage is not attained, the arc can be established only by igniting it mechanically or by electrical means using the so-called 'ignition current'.

For the mechanical ignition of the arc the front parts of the electrodes are set at a distance where the actual breakdown voltage is lower than the voltage applied to the electrodes.

The arc established in this way is not disrupted when the electrodes are then separated to a distance at which the arc current follows the relationship shown by Eq. (2.50)

$$I_0 = \frac{U_0}{R + R_a}$$

In order to keep the arc burning, however, it is necessary that the condition set out in Eq. (2.48) is met

$$U_0 = U_a + IR$$

From the two relationships above it follows that in order to keep the arc operating a supply is necessary whose terminal voltage obeys the relationship

$$U_0 \geqq \left[1 + \frac{R}{R_a}\right] U_a \qquad (2.51)$$

In practice, the arc is ignited simply by touching the front parts of the electrodes together and then separating them. During this operation the supply is not short-circuited as this is prevented by the resistance placed in the circuit. This procedure is based on the fact that when the front parts of the electrodes are brought into contact they become so hot, due to the large contact resistance, that the work function of the electrodes is reduced to such an extent that the ignition of the arc is assured.

According to current practice the arc is not usually ignited mechanically but instead a special ignition current is used. In the latter case the so-called *ignition current* is required in addition to the *maintaining current* that maintains the operation of the arc. The voltage of the ignition current is higher than the breakdown voltage of the arc gap, but its intensity is too low to keep the arc in operation. When the ignition current is applied, a sufficient degree of ionisation is achieved which overcomes the resistance of the arc gap. The voltage difference between the electrodes is then reduced to a value low enough for the operating current to maintain the arc.

Reference

1. WIEN, W. and HARMS, F., *Handbuch der Experimentalphysik*, vol. 12, Part 3, Leipzig (1927) p. 226

2.3.3 Effect of alternating current

In the above section it has been assumed that the arc current is direct current, although in practice this is the exception and a.c. mains voltage is generally used for the operation of the arc. In the application of a.c. mains supply the most important feature to be considered is that the voltage changes periodically. In so doing the voltage curve reaches the zero value twice and the polarities of the electrodes are inverted during each period. In the case of the a.c. arc, the excitation processes are repeated so that the orientation of the electrical field is changed in every half-period.

An a.c. arc can thus only be self-supporting or maintain its operation without the application of an ignition current, when the voltage of the operating current is higher than the breakdown voltage of the electrode gap. Exceptions are *carbon arcs* where the intensity of the arc current has a sufficiently high value. This is explained by the fact that the front parts of the carbon electrodes are maintained at a temperature that allows the ejection of electrons by the mains voltage even when no current is flowing. This exceptional case is limited to carbon electrodes that have relatively poor thermal conductance. When the electrodes are made of metal with a good thermal conductance, the use of an ignition current cannot be avoided when the voltage of the operating current is lower than the breakdown voltage value. There are also other advantages in the application of the ignition current. With an appropriate setting of the ignition time, the arc can be initiated at any instant within the relatively broad phase interval, or, if necessary, ignition can be achieved only once in a period or even less frequently.

It can be experimentally demonstrated that in the course of a.c. arc generation the *voltage difference between the electrodes* follows the power supply voltage increase until the voltage reaches the value of the breakdown point; that is until the arc is initiated by the ignition current. At the instance of ignition the voltage value drops suddenly to the value of the glimm voltage (i.e., the smallest voltage difference that induces gas discharge in the gas between the electrodes), U_g, and then after a short

time it drops further to the arc voltage level, U_a (Fig. 2.27*a*). In the first phase of arc initiation, the current conductance is provided by charged particles produced partly by collision ionisation in the arc gap and partly by collision of the ions with the cathode. As the electrodes heat up, electrons ejected from the cathode assume an increasingly important role in the current conductance mechanism. As a consequence, the resistance of the arc gap is reduced until a static thermal state is reached and the voltage between the electrodes is reduced to the arc voltage level. There is no voltage change until the arc current is no longer able to maintain the operation of the arc. This happens because the voltage drops in every half-period of the a.c. supply, not only reducing the arc current, but also lowering the temperature of the arc plasma and thus reducing the degree of ionisation. The arc becomes gradually transformed into a glimm discharge and is soon extinguished. Within this period the voltage between the electrodes increases from the level of the arc voltage to the value of the glimm voltage and then follows the periodic voltage changes of the power supply. The above process is repeated as the polarity changes in the following half-periods.

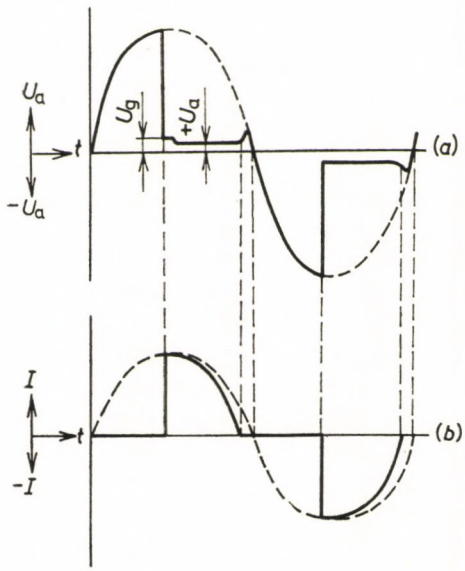


Fig. 2.27 Schematic voltage–time curve (*a*) and current–time curve (*b*) of an a.c. arc

It follows from the above discussion and from Eq. (2.48) that the a.c. arc current is not constant even in the course of the burning period of the arc. The value of the arc current is not only zero until the arc is ignited, but has a negligible value until the glimm discharge is transformed into a proper arc. When the latter process takes place the value of the arc current suddenly increases to the short-circuit level (Fig. 2.27*b*). Later, the arc current changes according to the value of the short-circuit current during the period when the resistance of the arc can be neglected in comparison to the other resistances in the circuit. On the voltage drop of the power supply in half-periods, however, the current is reduced. The arc current thus increasingly deviates from the value of the short-circuit current. The arc current–time curve will, therefore, reach zero earlier than the voltage–time curve.

The burning time of the arc in a half-period is generally about a quarter of the period; for example, when the frequency of the a.c. supply is of 50 Hz, the burning time is about 5.10^{-3} s. The burning periods alternate with non-burning periods that have

roughly the same duration. Thus the temperature of the a.c. arc strongly fluctuates with the a.c. frequency, such that resultant temperature changes can reach the value of a few thousand degrees centigrade for a frequency of 50 Hz. This temperature fluctuation is significantly reduced when the a.c. frequency is increased. It can be proved experimentally that when the frequency is 500 Hz, the temperature fluctuations are so small that the arc will have practically the same temperature as the d.c. arc.[1] In the plasma of the a.c. arc and also of the d.c. arc the same physical and chemical processes are taking place, except for the differences brought about by the temperature fluctuations of the former plasma.

Reference

1. ORNSTEIN, L. S. and BRINKAM, H., *Physica, 's Grav.*, **1**, 777 (1933)

2.4 Arc generators

2.4.1 Classification of arc generators

The equipments used for producing arcs are called arc generators. They can be classified in various ways depending on different parameters as seen in the scheme below:

(1) According to the *frequency* of generation
  (*a*) d.c. arc generation
  (*b*) a.c. arc generation
(2) According to the *voltage* of the power supply
  (*a*) mains voltage generation
  (*b*) high-voltage generation
(3) According to the *burning time*
  (*a*) continuous arc generation
  (*b*) interrupted arc generation
(4) According to the *type of ignition*
  (*a*) mechanical ignition
  (*b*) high-voltage, high-frequency current ignition
  (*c*) electronically-controlled ignition

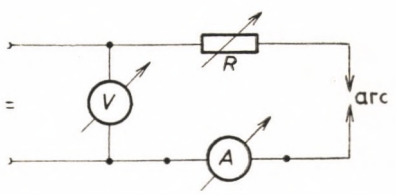


Fig. 2.28 Circuitry of arc generators

It becomes clear from this scheme that by combining various generation features, different types of arc generator can be constructed. All arc generators are built from essentially the same components (Fig. 2.28) and are simple in their basic construction. The basic components are: (1) a continuous or step-wise variable, limiting resistance, R, (2) an ammeter, A, for measuring the arc current, and (3) a voltmeter, V, for checking the terminal voltage of the power supply.

2.4.2 Power supplies for arc generators

Arc generators currently in use generally utilise the mains power supply. In the case of d.c. arc generators, the direct current is normally obtained by using semiconductor rectifying units, mercury-vapour rectifiers, or rotary rectifiers. The required current is about 10–15 A for an a.c. supply of 220 V. The advantage of copper oxide or silicon dry rectifiers is that they are convenient and simple to use. By using one unit it is only possible to utilise one polarity of the a.c. circuit, thus those systems which are based on double rectification or Graetz circuitry have a distinct advantage as both half-periods of opposite polarity can be utilised (Fig. 2.29). The voltage produced by the latter method is not uniform but fluctuates strongly as is shown in Fig. 2.30*b*.

The voltage–time curve has less of a zig-zag character if the rectification is performed by a three-phase mercury-vapour rectifier (Fig. 2.30*c*). The use of this relatively heavy-duty rectification method is only economical if the d.c. power supply is used for other equipment as well as the arc generator.

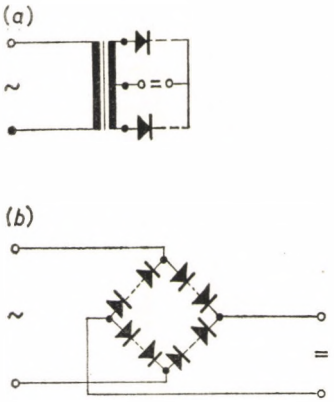


Fig. 2.29 Circuit diagrams for (*a*) full-wave rectification and (*b*) Graetz-circuit rectification

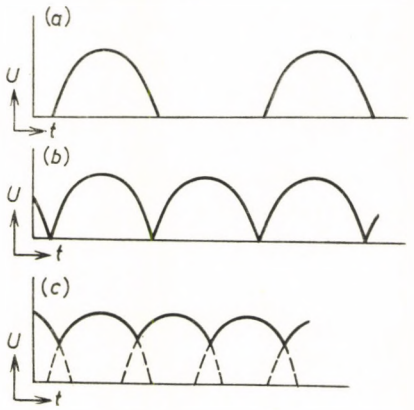


Fig. 2.30 Voltage–time curves of rectified a.c.: (*a*) half-wave rectification, (*b*) full-wave rectification, and (*c*) by the use of the three-phase mercury vapour rectifier tube

The periodic change in the voltage of a current rectified in the above way can be smoothed if a condenser of appropriate capacitance or a series-connected iron-core induction coil is used (Fig. 2.31*a*, *b*). When the voltage is higher than the average level in a high capacitance condenser, the electrical energy accumulates in such a way that the voltage changes of the rectified current influence the mean terminal voltage of the condenser only to a small degree. The required capacitance of the condenser depends, therefore, on the maximum intensity of the arc current, and also on the periodic voltage changes of the rectified current. In the induction coil the

electrical energy accumulates in the magnetic field of the coil during the period when the current increases and is released when the current drops. The induction coil is capable of efficiently smoothing the current fluctuations only when the iron core of the coil is not saturated. In the design of the latter, therefore, one should always consider the intensity of the maximum current that could occur in practice.

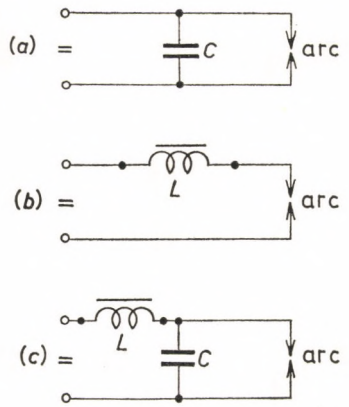


Fig. 2.31 Equalisation of the voltage variations of the arc current by (*a*) condenser, (*b*) induction coil, and (*c*) Steinmetz circuitry

In the case of *rotary rectifiers*, a direct current can be obtained whose voltage is particularly constant if the dynamo is driven by a directly-coupled, synchronous motor. When this method is applied, current fluctuations stemming from unavoidable changes of the arc resistance can only be smoothed if the arc is connected in series with an iron-core induction coil.

The power supply of a.c. arc generators is usually the mains current, while a transformer is needed only in the case of high-voltage arc generators. When a.c. circuitry is used, it is advisable to use the Steinmetz circuit for the stabilisation of the arc current (Fig. 2.31*c*). A successful current stabilisation can only be achieved by the latter circuit if the inductance of the induction coil, L, and the impedance of the capacitor, C, are equal. In this case the intensity of the current is independent of the impedance of the arc.

2.4.3 Arc ignition

In the case of high-voltage arc generation the arc is, as a rule, self-ignited. On the other hand, in mains voltage arc generation the ignition of the arc should be achieved in some other way.

Among the usual ignition methods, mechanical ignition of the arc can only be used for d.c. arc generation. When mains voltage a.c. arc generation is used the arc should be reignited every half-period. For the initiation of the arc Tesla current is generally used. The three main characteristics of the Tesla current are high voltage, high frequency, and low consumption. After the high voltage has secured the breakdown of the analytical gap, the arc gap becomes ionised and the working current becomes operative. The high frequency renders it possible to separate the mains from the Tesla current by a simply constructed filter. Finally, the low power of the ignition current secures a very small ignition spark, which has practically no effect on the emission of the working arc.

The design of the apparatus for producing the current (Fig. 2.32) is in principle the same as that of spark generators (see Sub-section 2.5.3). The only difference is that the consumption of the former is lower and that it also contains a Tesla transformer. Through a buffer resistance, R, the transformer, Tr (of about 5–10 kV secondary voltage) is connected to the mains and charges the condenser, C (of about 1000 pF). The charged condenser is discharged through a primary coil of the Tesla transformer, Te, connected in series across the secondary spark gap S. The discharge occurs in that phase instant when the voltage difference in the secondary spark gap, which is identical to the terminal voltage U_s of the condenser C, reaches the value

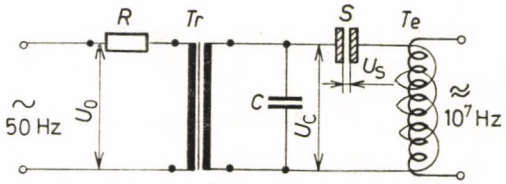


Fig. 2.32 Principal circuit of the device used for the production of Tesla current

of the breakdown voltage U_b, of the spark gap. The voltage, U_c, is lagging somewhat behind the mains voltage, U_0, due to the phase-lagging effect of the capitance, C (Fig. 2.33). The condenser, C, is recharged following the discharge and the whole discharge–charge process repeated until the voltage, U_c, drops below the level of the discharge voltage (Fig. 2.33*b*). The duration of the charging process and the number of recharges and discharges are determined by the circuit data; mainly the values of R and C. The discharge of the condenser is a high-frequency damped vibration phenomenon. The frequency of the latter is determined by the value of the capacitance, C, and the inductance of the primary coil of the Tesla transformer. [See Eq. (2.71).] (The d.c. resistance of the discharge circuitry can be neglected.)

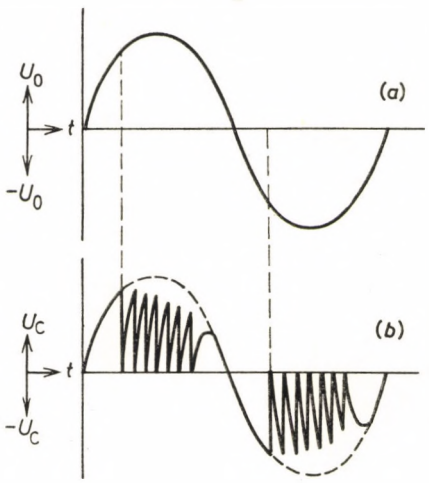


Fig. 2.33 Condenser voltage–time curve of the apparatus used for Tesla current generation

The corresponding frequency is about 10^7 Hz, while the duration of total damping is about 10^{-5} s. Owing to the effect of the high-frequency current, a current will be induced in the secondary winding of the Tesla transformer. Most of the magnetic lines of force induced by the current passing through the primary winding of the

Tesla transformer are enclosed in such a way as to avoid the turns of the secondary winding. Thus the efficiency of the transformer is very poor, especially at low frequencies. The discharging current of the condenser, C, contains many overtones, so that the current induced in the secondary winding of the transformer, *Te*, will be composed of just the high-frequency components. It can also be seen that demands for high voltage and low consumption of the ignition current are satisfied.

We have seen that the terminal voltage, U_c, of the condenser, C, can reach the breakdown voltage level, U_s, of the spark gap, S, several times during every half-period (see Fig. 2.33*b*). For a given set of electrical parameters, the number of breakdowns depends on the distance between the electrodes in the spark gap. If the electrodes are separated, such that the breakdown voltage is identical to the peak voltage, only one breakdown will occur in each half-period. In general, the smaller the distance between the electrodes, the greater will be the number of breakdowns. The arc is initially started by the ignition spark from the first breakdown, so subsequent breakdowns are unimportant in this respect. It is important to remember that when the spark gap is reduced the point of ignition is earlier. As the condenser has a delaying effect on the phase, the instance of ignition will, in practice, not occur earlier than the phase instant belonging to the extreme value of the mains voltage.

The *Tesla* current, when used as the ignition current in a.c. arc generators, can be taken over by the working current circuit in two ways. One of these methods is based on the Pfeilsticker method[1] in which the secondary winding of the Tesla transformer is connected in series with the analytical gap (Fig. 2.34a). The arc current, therefore, also passes through the secondary winding of the Tesla transformer. The ignition

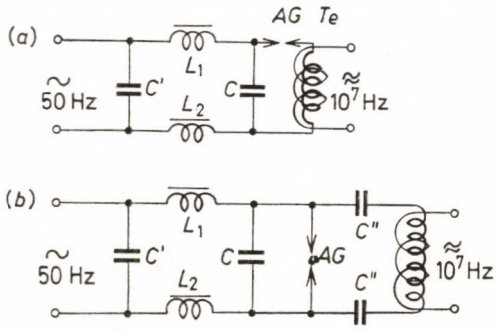


Fig. 2.34 Superposition of the Tesla current on the working current by (*a*) the Pfeilsticker method, and (*b*) the Calker method

current is circuited through the analytical gap and the condenser, C. The latter has practically no resistance towards the high-frequency current, although it blocks the 50 Hz frequency mains current. The resistance of the condenser is lowered when the frequency, ν, is increased according to the relationship $R_c = 1/2\pi\nu C$.

In the application of the Pfeilsticker process it is important that the high-frequency ignition current does not get into the mains. This is avoided by using a filter circuit consisting of coils, L_1 and L_2, having a self-inductance with the magnitude, H, and a condenser, C', with capacity of some μFs. The inductance of the induction coils is increased as the frequency, ν, of the current is increased according to the relationship $R_L = 2\pi\nu L$. The L_1 and L_2 coils, therefore, allow the free passage of the 50 Hz arc current, while allowing only a very small proportion of the high-frequency current to pass. The latter is short-circuited through the condenser, C', and thus is not admitted into the mains circuitry.

Another method for the transfer of the high-frequency ignition current is the Calker principle.[2] In this system the ignition current is transferred to the main circuit by two condensers whose capacitance is of the order of a few picofarads (Fig. 2.34*b*). In the method based on the Calker principle the arc current does not pass via the secondary winding of the Tesla transformer. The high-frequency ignition current, however, is connected through the arc gap only. Here it can only start the arc, when the inductance of the L_1 and L_2 induction coils in the circuit is higher than the resistance of the arc gap. (The high-frequency current that occasionally passes these chokes cannot get into the mains circuit, since the latter is short-circuited through the condenser.)

A drawback in the above process is that the arc is ignited at the instance when the voltage of the Tesla exciting condenser reaches the breakdown level of the arc gap. The value of this breakdown voltage is generally uncertain. It depends upon among other factors, the particular condition of the front part of the electrodes at that time. One further disadvantage is that the point of arc ignition can be controlled only with difficulty. Furthermore, it is only possible to control it within very narrow intervals. Such control can only be effected by changing the electrode distance of the arc gap, *AG*, or changing the charging speed of the condenser, *C*, that is the value of the resistance, *R*, in Fig. 2.34.

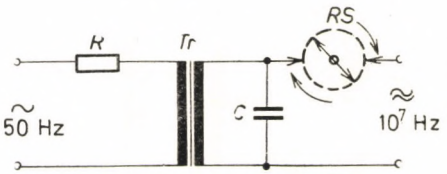


Fig. 2.35 Rotating switch control of the ignition current

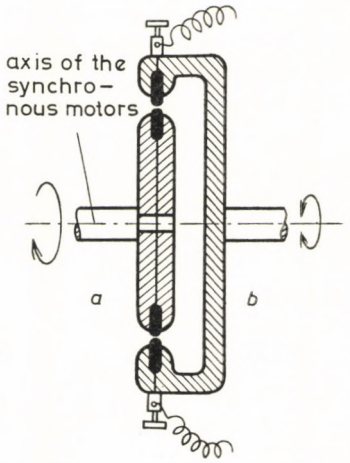


Fig. 2.36 Design of the rotating switch: (*a*) rotating part, (*b*) stationary part

The above disadvantage is partly eliminated *if the ignition current is mechanically controlled.*[3] This can be achieved when an arc gap, forming part of the discharge circuit of the condenser, is replaced by a rotating arc gap driven by a synchronous motor (Fig. 2.35). This system is built basically of two parts—one stationary and the other moving. The stationary part is a disc (or double arm) made of some insulating material, with electrodes connected by a conductor placed around the circumference. These electrodes are rotated in front of two 'stationary' electrodes situated in the plane of the disc (Fig. 2.36). When the condenser is discharged and the high-frequency current produced starts the arc, the total breakdown voltage of the two spark gaps, formed by the rotating and stationary electrodes, becomes equal to the terminal voltage of the condenser. (The control principle is the same as for the Feussner system spark generators that will be described in Sub-section 2.8.4.)

The advantage of having a rotating control for the ignition current is that the instance of arc ignition can be reproduced to a high degree of precision and that

the ignition can be controlled simply by turning the stationary electrode pair. In principle, this scheme also has the advantage that the rotary switch avoids the multiple discharge of the condenser in every half-period.

There is little doubt that arc ignition by electronic control is the most precise and dependable method available. There are many instruments that can be used for this purpose and although of different construction they are basically composed of two parts—a system to produce an electrical control signal and an ignition circuit.

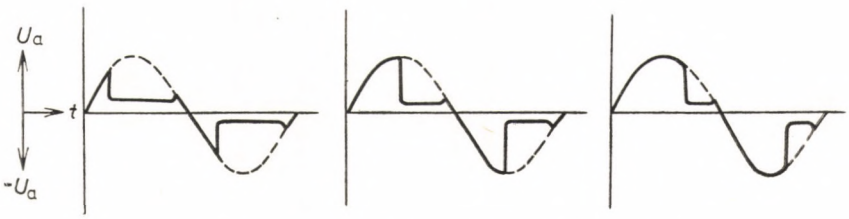


Fig. 2.37 Voltage–time curves of the a.c. arc for ignitions with different phase positions

The electrical control signal is generally produced by a 'pulse generator', an electric device producing current pulses of very short duration. The instant ignition of a pulse generator can be easily moved with respect to the phase of the mains current (Fig. 2.37). The voltage forms arising after each stage, indicated on the flow diagram of a pulse generator, can be seen in Fig. 2.38.[4] The purpose of the first unit called 'phase shifting' is to ignite the arc in the desired phase position. The second unit, the 'multivibrator', produces voltage impulses in each half-period having a polarity that is the same as that of the half-periods. The third unit, the 'pulse rectifying unit', rectifies the pulses. This is necessary because only positive polarity voltage pulses can be used for switching the current with a thyratron tube. During this rectification the pulse with positive polarity occurs in only every second half-period. This unwanted effect is eliminated by the next unit, called a 'flip-flop circuit', such that the rectified signal is transformed into two branched signals.

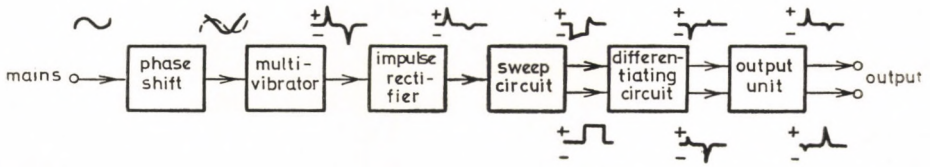


Fig. 2.38 Block flow diagram of the Bardócz-system pulse generator[4]

These, approximately, square-wave signals are already shifted with respect to each other by a half-period, but have negative polarity. These are then transformed by the *differentiating circuit* such that the pulses in the two branches will have negative sign. The latter are changed to the necessary positive polarity by the 'output unit'.

The signals of positive polarity coming from the pulse generator are used to ignite the arc either via various resistors or via the control grid of the *thyratron* tube. Their negative grid bias voltage of these heavy-duty, gas-filled triodes generally stops the current flow. At a certain grid bias voltage, however, their anode current suddenly

starts, and the thyratron tube becomes conducting if the anode voltage is sufficiently high at the same time. The intensity of the anode current of the thyratron tube is increased by collision ionisation occurring in the space between the grid and the anode, such that glimm discharges are produced. The current arising in this way is sharply terminated when the anode voltage drops below a certain level—the so-called extinguishing potential value.

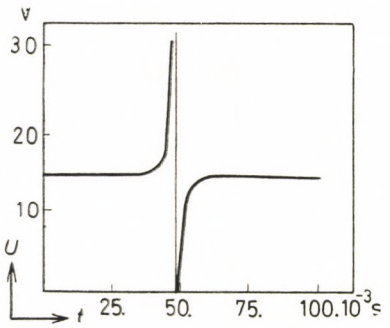


Fig. 2.39 Electrical signal produced by magnetic effects[5]

The electrical signal used for starting the arc can be produced by other techniques' as well as by pulse generators. One of these more recent techniques is the production of a signal by *magnetic means*.[5] In principle, this method involves the induction of an electrical voltage when an iron disc is rotated in front of a magnetic coil. The signal produced arises in exactly equal time intervals if the iron disc is rotated by a synchronous motor immediately in front of the magnetic coil (Fig. 2.39). The voltage induced in the winding is taken through a condenser and ohmic resistors to the grid of the control tube the grid having an appropriate bias voltage. The phase position for the ignition is the central line of the rotating disc of the magnetic coil driven by the synchronous motor.

The above principles are combined in the design of equipment used in the *electronic control of arc* (or spark) generation. The basic circuitry of one of the best designs is shown in Fig. 2.40. In this case the controlling spark gap is formed by two series-connected, identical spark gaps, S_1 and S_2, both spanned by two series-connected resistors R_1 and R_2. These distribute the terminal voltage of the ignition condenser, C, evenly over the S_1 and S_2 spark gaps. The breakdown voltage, U_s, of the latter should be chosen such that the terminal voltage, U_c, of the condenser is smaller than the total breakdown voltage of the two spark gaps, although one of the spark gaps could be safely broken down by the total terminal voltage of the condenser. (It is therefore necessary that in the first case the condition $U_c < 2U_s$ is met and in the second, that the condition $U_c \gg U_s$ is obtained.) In this way the discharge of the condenser through spark gap S_1 only occurs when the positive pulse of the pulse generator, PG, ignites the thyratron tube, Ty, and thereby short-circuits spark gap S_2. When the resistance, R_3, in the anode circuit of the thyratron tube is sufficiently great, the current produced when spark gap, S_1, is broken down develops a voltage difference at the terminals of resistor R_2, which is greater than the breakdown voltage of spark gap S_2. Therefore, the current breaks down this spark gap as well. In this way the condenser is discharged through both spark gaps and a high-frequency ignition current is generated in the secondary circuit of the Tesla transformer, Te. Consequently, the voltage difference at the terminals of resistor, R_2, is reduced so

much that the anode current of the thyratron tube is terminated. Thus the condition that prevailed before the production of the positive polarity signals from the pulse generator is re-established.

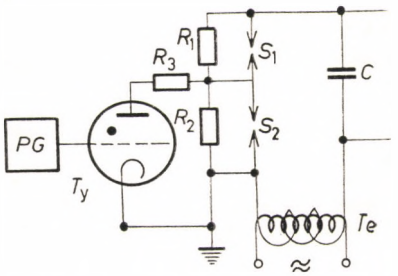


Fig. 2.40 Circuit diagram of the Bardócz electronically controlled generator[6]

Figure 2.41 illustrates an electronic control system that is much simpler than the one described above, although it is not so popular. In this system a low capacity condenser, C, is charged by a high-voltage d.c. power supply. The condenser cannot, however, be discharged until the positive signal of the pulse generator, *PG*, renders the thyratron tube, T_y, conducting. When this happens the condenser, C, is discharged through the primary winding of the air-core transformer, designed as an

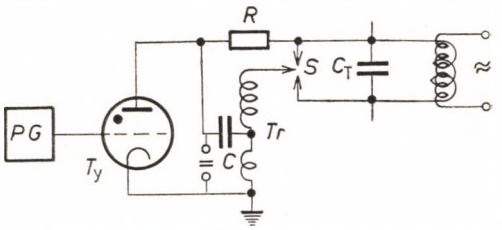


Fig. 2.41 Electronic control of arc generation according to the Kovács system[7]

autoconnected transformer, *Tr*. A great voltage difference is thereby induced in the secondary winding of the Tesla autoconnected transformer such that the current generated breaks down the arc gap, *S*. The spark gap is then ionised and discharge of the condenser, C_T, is made possible. A high-frequency current pulse then generates the ignition current via the Tesla transformer, *Te*.

References

1. Pfeilsticker, K., *Z. Elektrochem.*, **43**, 719 (1937)
2. Van Calker, J., *Z. analyt. Chem.*, **105**, 396 (1936)
3. Török, T., *Acta Chim. Hung.*, **1**, 289 (1950)
4. Bardócz, Á., *Spectrochim. Acta*, **7**, 307 (1956)
5. Bardócz, Á., and Varsányi, F., *Z. Naturf.*, **10**, 1031 (1955)
6. Bardócz, Á., *Acta Phys. Hung.*, **11**, 65 (1955); **13**, 3 (1955)
7. Kovács, F., unpublished results

2.4.4 Direct current continuous-arc generators

The basic circuitry of d.c. continuous-arc generators is in principle the same as that shown in Fig. 2.28 assuming the arc is mechanically ignited and the power supply is d.c. of appropriate voltage. When the power supply is a.c., rectification is carried out using the Graetz system, while the rectified voltage is smoothed by a condenser

of suitable capacitance (Fig. 2.42). Furthermore, in d.c. continuous-arc generation it is advisable to ignite the arc by Tesla current, instead of a mechanical procedure.

The supply voltage of normal d.c. continuous-arc generators is usually the same as that of the mains supply (220 V or occasionally 110 V), while the arc current is somewhere between 3 and 15 A. There is a definite advantage in using a high voltage, as the reproducibility of the spectral line intensity is improved.

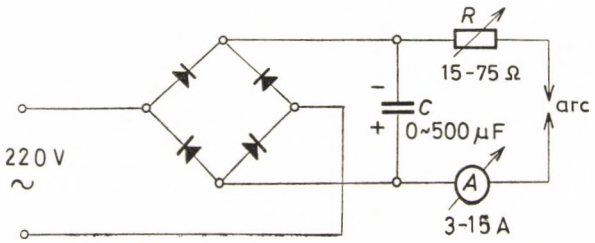


Fig. 2.42 Circuit diagram of rectified-current, continuous-arc generators

The great advantage of d.c. continuous-arc generation is that it is the most sensitive method of all the normally used excitation techniques. By its application, elements present in low concentration can be analysed that would not be detected by any other excitation means. This is mainly due to the fact that the spectra are composed of relatively few, easily excited atomic lines (fundamental lines). On the other hand, this method has the disadvantage that the electrode tips become strongly heated during arcing.

It can be theoretically shown that the two experimental observations mentioned above are not independent of each other. According to Eq. (2.22) the change in intensity of a spectral line is proportional to the partial pressure of the corresponding elements' vapour in the arc plasma. On the other hand, if the experimental conditions are identical, then the greater the partial pressure, the greater is the amount of material evaporated in unit time, or, indirectly, the higher is the temperature of the electrode tips.

The strong heating suffered by the front parts of the electrodes can lead to fractional distillation.[1] As the internal parts of the electrode are at a much lower temperature than the exposed tip, high-boiling-point components are more extensively evaporating from the external portion of the electrode than from the deeper internal layers. However, low-boiling-point elements distil to the arc plasma from the deeper internal parts as well. In the course of arc excitation, therefore, *varying* concentration differences emerge between the original constitution of the electrodes or sample and the composition of the vapour phase. *The effect of fractional distillation is such that a reproducible change in the intensity of the spectral lines cannot be obtained, although there can be an increase in the detectability of the low-boiling-point elements.* The former factor is a great disadvantage in quantitative analysis, while in certain cases the latter may prove an advantage.

Non-reproducible evaporation of the electrodes or samples is further increased by a change in the chemical composition of the foremost portion of the electrode during arcing. For example, certain oxides and nitrides can be formed, while other compounds present can undergo decomposition. Such chemical changes can lead to variations in the vaporising speed of certain elements and change the work func-

tion of the electrons, etc. Sudden variations in the arc current during excitation can be partly explained by the above phenomena, which accounts for the non-reproducible nature of the spectral line intensity.

Due to large fluctuations of the arc current the relative analytical reproducibility attainable by continuous-arc generation is at the most $\pm 10\%$. The validity of this statement is upheld by the fact that the relative analytical error is reduced to about $\pm 3\%$ when the arc current is stabilised.[2] Sudden changes in the arc current can be almost completely eliminated by using appropriate electronic current control equipment.[3] Experiments proved, however, that this control, in itself, does not result in substantial improvements. There is no important increase in the reproducibility of the intensity ratio of the spectral lines at different excitation potentials. Reproducibility can only be realised when the voltage of the stabilised arc current is also increased at the same time.

In d.c. continuous-arc generation it can be observed that the temperature of the burning spots arising on the front part of the anode is much higher than that of the corresponding spots on the cathode. It can be established that after arcing for a short time, the burning spots on the anode are deeper and more localised than those on the cathode (Fig. 2.43). This difference is less evident when the electrodes are examined after the arc has been operating for a longer period. The spectral properties of the radiation coming from the plasma mass located in the anode and cathode spaces are not quite the same, even though the two electrodes are chemically and

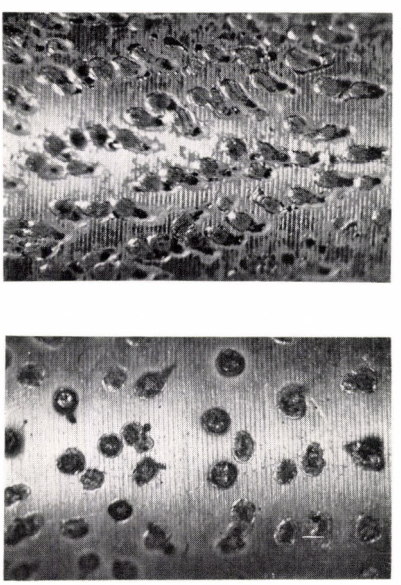


Fig. 2.43 Discharge spots by polarised d.c. arc with 0·005 s burning times ((*a*) cathode, (*b*) anode)[4]

physically identical. The radiation that originates immediately from the cathode space is completely different from the average radiation of the total plasma. This so-called *cathode layer radiation* comes from the very strong electrical field produced by the great voltage drop in front of the cathode.[5]

References

1. EBERHARDT, G., *Sber. Preuss. Akad. Wiss.*, **1908,** 851
2. FASSEL, V. A., *J. Opt. Soc. Am.*, **39,** 187 (1949)
3. BELYAKOV-BODIK, I. B., and MANDELSTAM, S. L., *Zh. techn. Fiz.*, **14,** 400 (1944)
4. TÖRÖK, T., SZAKÁCS, O., and SZABÓ, Z. L., *Magyar Kém. Foly.*, **66,** 487 (1960)
5. MANNKOPFF, R., and PETERS, CL., *Z. Phys.*, **70,** 444 (1931)

2.4.5 Mains voltage a.c. arc generators[1]

Mains voltage a.c. arc generators are composed of the *arc current circuit* (or work current circuit) and the *ignition circuit* (or the Tesla current generator) (Fig. 2.44). The mains voltage is generally 220 V or may occasionally be 380 V, while the arc current is usually about 3–15 A. With respect to changes in the voltage in the analytical circuit and in the arc current with time, the reader is referred to Subsection 2.3.3 and Fig. 2.27.

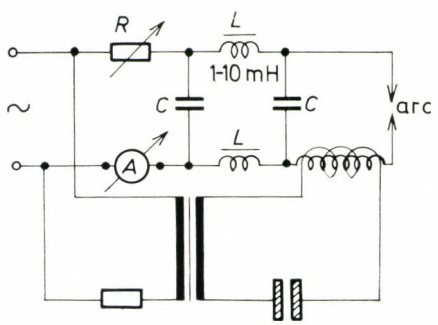


Fig. 2.44 Schematic circuit diagram of mains voltage a.c. arc generators with Tesla transformer transmission of the ignition current

The advantage of mains voltage a.c. arc generation over d.c. continuous-arc generation is that the electrodes are allowed to cool down somewhat during the intermission in the operation of the arc. Thus the difficulties encountered due to the strong heating-up of the electrodes are partially removed. Thus the relative analytical reproducibility attained is slightly better than that of d.c. continuous-arc generation.

Reference

1. PFEILSTICKER, K., *Elektrochem.*, **43,** 718 (1937)

2.4.6 High-voltage a.c. arc generators and high-voltage rectified arc generators [1–3]

In arc generators of these types the mains a.c. is first transformed into an alternating voltage of several kilovolts, which can then be used directly for arcing or first rectified and subsequently used for arc generation. In such methods there is no need for the separate ignition of the arc, provided the operating voltage of the arc is higher than the breakdown voltage of the electrode gap.

A.c. high-voltage arc generators are of relatively simple construction, with a circuit composed of a transformer, a variable resistor in series (or an iron core variable choke), and an ammeter in the secondary circuit of the transformer (Fig. 2.45). The arc current can be regulated by means of the ohmic or inductive resistance. The

arc voltage, on the other hand, can be varied via different branches from the secondary winding of the transformer. In normal high-voltage arc generators the arc current (as measured on the secondary side of the transformer) is generally about 2–5 A, while the voltage can be altered from 1 to 5 kV.

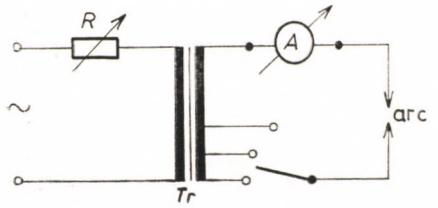


Fig. 2.45 Circuitry of high-voltage a.c. arc generators

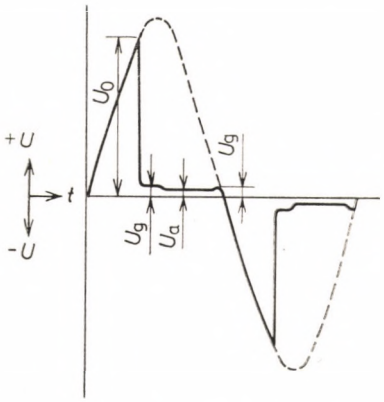


Fig. 2.46 Voltage–time curve of high-voltage a.c. arcs

When high-voltage a.c. arc generators are used, a voltage rise in the secondary current takes place at the beginning of every half-period, until its value reaches the level of the breakdown voltage, U_b, of the arc gap. Thereafter, the voltage suddenly drops to the glimm discharge level U_g, until finally it reaches the value of the arc voltage, U_a (Fig. 2.46). The arc radiation produced in this way is extinguished when the current is reduced to such an extent by periodic voltage changes that it is not enough to keep the arc operating. The burning time of the arc is therefore about a quarter of the main period. On the extinction of the arc radiation, a short-duration glimm discharge is developed. The voltage increases somewhat before dropping to zero owing to the periodic change of the external voltage. The whole process is then repeated such that the polarity of the electrodes is changed. The value of the current is zero during the arc intermissions, but suddenly increases when the arc is ignited. In the course of the operation of the arc, the arc current changes in a sinuous manner.

For certain practical applications it is advantageous to have definite polarity for the electrodes during the arc generation. This requirement is met by rectified high-voltage arc generators. The appropriate arc current can be simply obtained from these instruments by transforming the mains a.c. into a high-voltage current and rectifying it by either the single-channel or the double-channel method (Fig. 2.47).

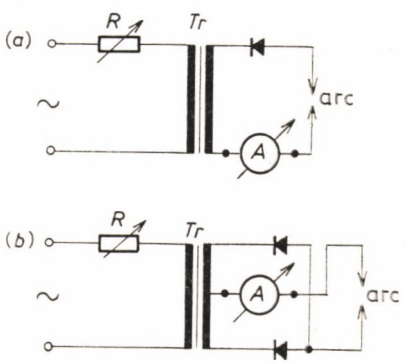


Fig. 2.47 Circuit diagram of high-voltage, rectified arc generators: (*a*) for half-wave operation, and (*b*) full-wave operation

For high-voltage, rectified arc generation, the voltage of the rectified current is not stable, but changes by a half sine wave in each half-period. There are breaks between these half-periods when single-channel rectification is used, although no intermissions are developed by the double-channel rectification method. The rectified current produced is, therefore, not smooth, but has a 'pulsating' form. The current is not constant, but also pulsates in half sine wave forms that follow the voltage changes. Moreover, in the case of single-channel rectification there are current-free breaks lasting for a half-period.

In the procedure described above the arc is ignited again when its voltage reaches the breakdown level of the arc gap. With respect to the analogous voltage–time curve for high-voltage a.c. arc generation, the only difference is that in the present case there is no change in the electrode polarity (Fig. 2.48). The burning time of the

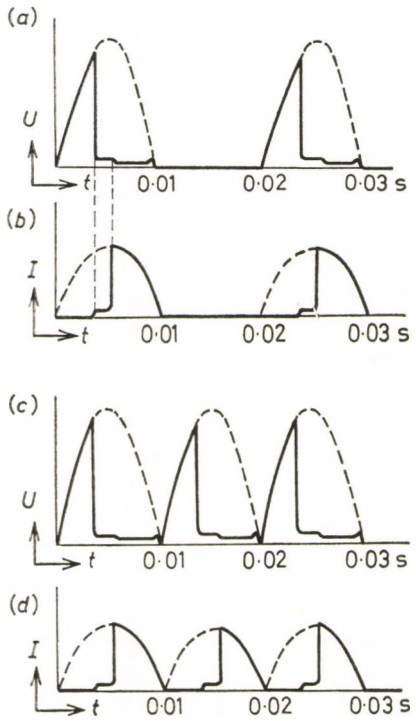


Fig. 2.48 Voltage–time and arc current–time curves of high-voltage, rectified arc generators in the case of (*a*) and (*b*) half-wave rectification and (*c*) and (*d*) full-wave rectification

arc is always about 5.10^{-3} s as well. Each burning period is separated from the next by about the same time break when the double-channel rectification method is used, although the break is three times as long for the single-channel method. Thus, there is about the same length of time available for the electrodes to cool down with the double-channel method as there is with other a.c. arc generation techniques. However, in the single-channel rectification technique much longer cooling periods will occur.

The sensitivity of high-voltage arc generation techniques is generally lower than that of low-voltage, continuous-arcing methods, although the sensitivity is much higher than that of spark generation techniques. Compared with low-voltage arcing techniques, a high-voltage technique has the advantage that the arc current is more stable and, thus, the analytical reproducibility is better. This fact can be accounted

for by pointing out that in high-voltage arc generation unavoidable changes of the arc resistance are virtually negligible compared with the ohmic or inductive resistance used for controlling the arc current.

References

1. DUFFENDOCK, O. S. and THOMSON, K. B., *Proc. ASTM*, **36**, (II), 301 (1936)
2. SAWYER, R. A. and VINCENT, H. B., *Proc. 6th Summer Conf. Spectroscopy*, New York-London, 54 (1939)
3. DUFFENDOCK, O. S. and LARUE, J. M., *J. Opt. Soc. Am.*, **31**, 146 (1938)

2.4.7 Mechanically controlled, interrupted-arc generators

During the application of the arc generation techniques described above, the electrodes become considerably heated. This can be avoided if the mains current is periodically interrupted, such that the electrodes are allowed to cool down. By this means they are not excessively heated up even at relatively high current values. The spectra obtained are then much less populated by bands and the background tends to be reduced. Furthermore, there is less self-absorption as during the breaks only the vaporised atoms and not the excited atoms completely leave the arc gap. Provided the breaks are long enough partial evaporation does not cause any problems. The excitation is generally very sensitive and of good reproducibility.

There are different types of interrupted-arc generators. The only structural difference between these generators and d.c. generators is that they contain some form of interruption device. The interruption or breaking of the arc current can be achieved either by direct on-and-off switching or by periodic switching of the ignition current. The latter technique is recommended, as the switching of a low-current ignition circuit is more dependable than that of a relatively high-current arc circuit. Interrupted-arc generators are classified according to whether they are mechanically or electronically controlled. Depending on the working current, a.c. and d.c. interrupted-arc generators have been developed. All types of interrupted-arc generators have one feature in common – they are operated on mains voltage work current.

In the first mechanically interrupted d.c. arc generators, one of the electrodes was fixed, while the other was made periodically to touch it via an eccentric device driven by a motor.[1] When the electrodes are separated an arc is established that is kept working until the electrode distance becomes so large that, owing to the increased resistance of the arc gap, the current becomes insufficient to maintain the arc. As the burning time of the arc is not only uncertain but cannot be limited in any way, this method is rarely used now.

Another obsolete system for interrupted-arc generation is when the arc established between fixed electrodes is periodically short-circuited by a rotary switch via a secondary circuit connected in parallel with the arc gap.[2,3] In this system the duration of the burning and the intermissions can only be varied with difficulty. It is a distinct disadvantage that during the interruption of the relatively high-current circuitry the switching device is not precise in its operation and leads to an ambiguous value for the arc current intensity.

In general, modern equipment *almost exclusively uses the a.c. interrupted-arc generation technique*. There is nothing to gain in using d.c. for the arc interruption

devices. When it is necessary to have a definite polarity for the electrodes, the arc current can be rectified relatively simply.

In modern a.c. interrupted-arc generators the burning time of the arc is regulated by the periodic interruption of the ignition current. The ignition can be controlled either by a mechanical switching device or by electronic circuitry.

The mechanically controlled interrupted-arc generator[4] differs from the mains voltage a.c. arc generator only in that the ignition current is periodically switched on and off by a rotating switch operated from the mains supply (Fig. 2.49). By means of this rotating switch, the ratio of the burning time to the breaks between the burning periods can be normally changed in the range 1 : 1 – 1 : 50. The total duration of the burning and intermission times of the arc can be regulated by varying the rotation speed of the motor and can normally be varied in the range 0·1–1 s, that is, the time for 10–100 half-periods.

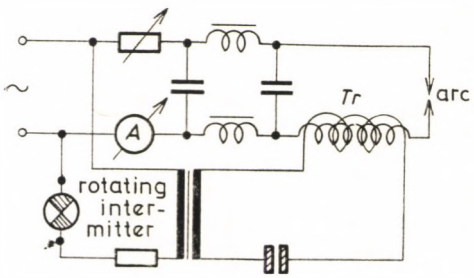


Fig. 2.49 Circuit diagram of the mechanically-controlled, interrupted arc generator

There is a phase shift between the working current and the transformed high-voltage current used for ignition. In general, the break device does not switch on the ignition current at the optimal point for the arc ignition. The uncertainty of operation derived from this fact is, however, only of theoretical importance as in practice the various deviations statistically compensate one another.

In the case of the interrupted-arc generators mentioned above, it is impossible to operate the arc for a very short period, for example for only a half-period of the main current. If this is attempted the arc ignition period becomes uncertain. However, when the arc is operated for a relatively long time excessive heating of the electrodes can only be avoided by allowing long ignition breaks. This procedure leads to a pronounced variation in the electrode temperature that can be harmful to the reproducibility and precision of the analysis.[5–9]

References

1. GERLACH, W., and RUTHARDT, K., *Z. anorg. allg. Chem.*, **209**, 337 (1932)
2. GERLACH, W., and ROLLWAGEN, W., *Techn. Physik*, **15**, 451 (1934)
3. VAN CALKER, J., *Z. analyt. Chem.*, **105**, 396 (1938)
4. PFEILSTICKER, K., *Z. Elektrochem.*, **43**, 719 (1937)
5. PFEILSTICKER, K., *Metallkunde*, **30**, 211 (1932)
6. KOSBAHN, T., *Ann. Phys.*, **25**, 625 (1936)
7. KAISER, H., and KECK, P. H., *Spectrochim. Acta*, **2**, 92 (1941)
8. BALZ, G., *Metallk. GiessWes.*, **33**, 260 (1941)
9. KAISER, H. and SOHM, M., *Spectrochim. Acta*, **2**, 81 (1941)

2.4.8 Electronically controlled a.c. arc generators

The uncertainties encountered in mechanically ignited arc generators can be eliminated by using an electronic control of the interrupting equipment. In this way the arc is always ignited exactly at the desired phase instant of the alternating current. In most modern equipment the burning and cooling periods can be controlled such that their duration, given in half-periods, can be varied within wide limits. The usual limits are from 2 + 0 to 2 + 100, or 1 + 1 to 1 + 100 (Fig. 2.50). (The paired

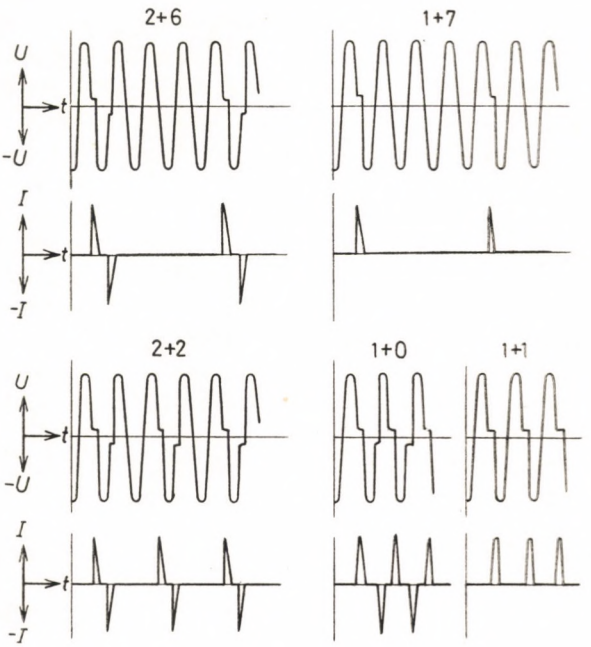


Fig. 2.50 Voltage–time and current–time curves of modern electronically controlled, a.c interrupted-arc generators[1]

numbers give the number of the half-periods corresponding to the burning and the cooling periods, respectively.) With electronic control the ratio of the burning and cooling periods can be varied over a large interval range, when compared with interrupted-arc generators ignited by Tesla current. The fact that the duration of the burning period can only be slightly changed is a drawback, although the arc burning time within a period can be varied relatively widely in electronically controlled devices. By an appropriate phase-shift in the ignition current it is possible to allow the arc to burn, either for a long time during a half-period, or only for an instant. The latter possibility has the advantage that during the very short burning time there is no extensive warming up of the deeper layers of the electrodes. Also, the use of electronic controls enables polarised arc discharges to be produced.

In comparing the electronically controlled interrupted-arc generators with those operated by a mechanically controlled ignition current, the former generators are certainly more up-to-date in their design and allow a more precise operation procedure. These advantages can only be obtained, however, via relatively complicated and expensive equipment. For all practical purposes it can be proved experimentally that the use of *mechanically controlled interrupted-arc generators produces satisfactory analytical results.*

Reference

1. BARDÓCZ, Á., *J. Opt. Soc. Am.*, **42**, 357 (1952)

2.4.9 Arc generation by high-frequency current

In exceptional cases it is an advantage if the arc current is high-frequency, high-voltage a.c. When current of this type is applied, the high-frequency a.c. flows mainly in the surface layer of the conductor (the 'skin-effect') instead of flowing across the whole cross-sectional area. According to the Thomson formula, $d = 10^3/2\sqrt{f\kappa\mu}$: the thinner the surface layer, d, the higher is the frequency, f, of the current and the greater are the conductivity, κ, and permeability, μ, of the layer, or, considering that $\varepsilon = 1/\mu$, the lower is the dielectric constant of the material under examination.

It follows from the above that in high-frequency arc generation only a very thin layer of the front part of the electrode is heated up to a temperature at which the material evaporates. The amount of material lost through evaporation is very small owing to the relatively large loss of energy via thermal conductance and to the comparatively small efficiency of the current used for this purpose. Thus, the intensity of the light sources produced by high-frequency arc generation is relatively low. Taking all these factors into account it can be stated that high-frequency arc generation can only be of use in exceptional cases, such as when the surface of a sample requires examination.

The current supplies used for high-frequency arc generation are either Tesla current generators or high-frequency generators using electron tubes. The optimal voltage of the high-frequency current is determined mainly by the breakdown voltage of the electrode gap, while its frequency is set by industrial limits.

2.4.10 Excitation by d.c. plasma jet

The plasma jets used as light sources in spectrochemical analysis are composed of two electrodes made of a high-temperature resistant material (e.g. graphite), placed one above the other and surrounded by a water-cooled 'burning chamber' (Fig. 2.51). When the jet is in use an arc of relatively high d.c. current value (usually about 20 A, or occasionally 100 A) is established between the two electrodes. An inert gas is leaked into the burning chamber under pressure, such that the arc plasma is blown out through the bore in the upper electrode to produce a flame-like glowing cone a few centimetres high. As the sample is sprayed in liquid form into the light source the d.c. plasma jet is really a combination of d.c. arc generation and flame excitation.[1,2,11–13] The plasma jet, however, has only the advantages of these two methods and not their usual drawbacks.

The above statement, which is also justified by experiments, can be explained by the fact that the d.c. plasma jet is actually a 'stabilised' d.c. arc.[3] For plasma jets of this type any fluctuations with time in the light intensity of the arc are generally eliminated if the volume of the plasma is limited by thermal means. This can be achieved by reducing the conductivity of the outer layer of the plasma (the 'mantle'), by cooling, and also by removing the plasma from the electrode gap by a gas stream.

An important source of error in d.c. arc generation is that the relative analytical precision of arc generation is strongly influenced by the chemical and physical

properties of the electrode surfaces. With plasma jets this error is almost entirely eliminated. The reason for this is that the sample is introduced into the light source in dissolved (or finely powdered) form and is therefore homogeneous, as well as being continuously renewed during excitation. It is partly due to this fact that in the application of the plasma jet very good analytical reproducibility can be achieved.

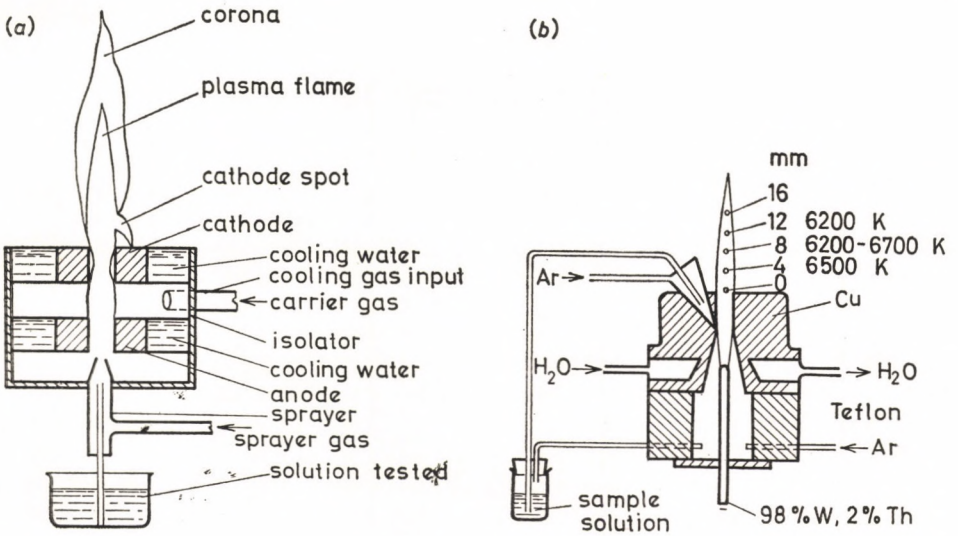


Fig. 2.51 Constructions of plasma burners: (*a*) usua construction, (*b*) less common construction[6]

It follows from the previous discussion that the gas flow-rate and the size of the bore in the upper electrode are important factors in the analytical application of plasma jets. By the selection of appropriate conditions most of the charged particles in the arc are forced through the bore of the upper electrode (the cathode) and into the 'flame'.

In d.c. plasma jets the electrical connection between the two electrodes is mainly established by material streaming from the tip of the arc plasma, or 'flame', onto the upper surface of the cathode. Here, it forms a burning spot or so-called *cathode spot*, on the top of the electrode, that wanders in a haphazard fashion over the surface. This has the drawback that the arc current and the light emission, in the given direction, tend to fluctuate in a non-reproducible manner.

This source of error can be eliminated if a secondary electrode (tungsten or carbon rod) is inserted into the 'flame cone' above the cathode and is connected with the cathode via a resistance.[4,5] In this way the arc current will be carried to the cathode mainly by the secondary electrode. Thus the emergence of the cathode spot is eliminated and the arc does not attack the electrode surface so much. The advantage of using a secondary electrode is that the contamination of the plasma by electrode material is reduced.

The plasma can be made completely free of contamination, however, if the Kranz-type, horizontal-axis plasma jet[5] is used instead of the common perpendicular-axis plasma jet previously mentioned (see Fig. 2.52). In the Kranz-type system, the arc plasma is led through two water-cooled copper rings (the stabilising rings), placed

in front of the cathode and the anode, before reaching the 'burning head'. The cooling gas, which is leaked tangentially into the stabilising rings, pushes the arc plasma through a perpendicular outlet in the burning head. The atomised solution or disintegrated sample is introduced into the burning chamber through a perpendicular bore that ends in the lower part of the burning head. In this system the

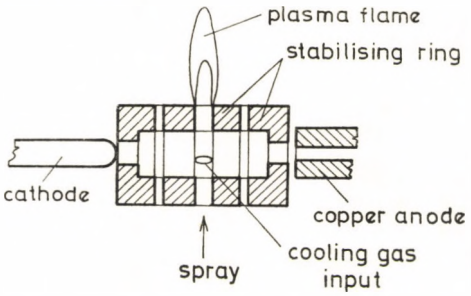


Fig. 2.52 Scheme of the Kranz-system plasma burner with horizontal arc axis (water cooling and insulation not included)

'flame' of the plasma jet is not brought into contact with the electrodes at all, because the 'contaminated' plasma in front of the electrodes is removed via special vents. By appropriate control of the experimental conditions and in particular of the gas flow-rate, the streaming of the plasma jet 'flame' becomes almost laminar.

This fact clearly illustrates that the construction of the plasma jet influences the operation of the jet and the temperature of the 'flame', etc.[9,10] With respect to the latter the following general points can be established. For perpendicular-axis plasma jets the temperature of the 'flame' is increased by narrowing the bore in the cathode which also increases the effluence rate of the gas. A limit is placed on the minimum size of the bore due to the fact that when the plasma is excessively narrow, salts precipitate out from the aerosol and block the bore. (The optimum size of the cathode bore is usually about 3–5 mm.) The distance apart of the electrodes only slightly influences the temperature of the 'flame' and the stability of the arc. It does, however, determine the smallest possible terminal voltage of the current supply.

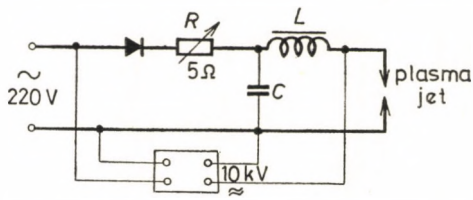


Fig. 2.53 Circuit diagram of the power supply of the d.c. plasma jet

No special demands are made concerning the current supply. It is generally sufficient to rectify the mains a.c. by a single-channel method and to smooth the pulsations of the rectified current by a condenser. In modern plasma jets, however, the arc is initiated by a high-frequency, high-voltage current instead of short-circuiting the electrodes. In this method (Fig. 2.53) the choke, L, in the arc current circuit is only effective in igniting the arc if its a.c. resistance is much greater than the resistance of the arc gap.[5] The current value of d.c. plasma jets used for analytical purposes is usually 20 A.

It is an advantage to use a carrier gas for the plasma jet whose excitation potential is fairly high. When this is done no problems arise from the presence of bands in the

spectrum. Accordingly, the best carrier gases are argon and helium and, less commonly, nitrogen. The disadvantage of the latter is that OH bands of the solvent (water) as well as nitrogen bands complicate the spectrum.[7] On the other hand, if argon is the carrier gas only argon spectral lines are present in addition to the OH bands. In principle it is advantageous if the flow-rate of the cooling gas is high, although the optimum flow-rate depends on the construction of the plasma jet. Thus the optimum rate can only be found by experiment.

For the atomisation of the solutions it is advisable to use atomisers that have been successfully used in flame photometry. In perpendicular-axis plasma jets the spray is carried into the plasma through the bore in the lower electrode. In some other designs the spray is carried into the burning chamber by the in-flowing carrier gas. It cannot be established with certainty, however, which is the more appropriate design, although it is simpler to use the same gas for both the atomisation of the solution and cooling the plasma. This is not possible, however, when the cooling gas is helium, as its low specific gravity does not bring about a sufficient pressure drop within the atomiser.[1]

The optimal volume of the analytical solution to be sprayed per minute can only be found by experiment, as the volume is influenced by the construction of both the atomiser and the plasma jet. From this point of view the physical properties of the sample solution (surface tension, viscosity, etc.) are important factors. It should also be taken into account that at very low spraying levels the intensity of the radiated light may be far too low. On the other hand, the plasma jet may be extinguished if the volume used is too great. In general, the optimal consumption volume for perpendicular-axis plasma jets is about 1–2 ml/min. This value can be reduced to 0·1–0·3 ml/min for special, direct-operation atomisers constructed explicitly for horizontal-axis, Krantz-system plasma jets.

The form, height, and even the temperature of the plasma flame are influenced by all the factors mentioned in the above discussion. It is, therefore, understandable that reproducible analytical results can only be obtained when all the factors mentioned are kept constant during a series of experiments.

The temperature distribution in the plasma flame is not uniform. There is a radial temperature variation in addition to changes of temperature along the axis of the flame. The highest temperature is found for both the horizontal and perpendicular axis plasma jets at the location of the plasma outlet, or for secondary electrode stabilised flames at the top of the electrode. At increasing distances away from these points along the flame axis, the temperature drops slowly, while at right angles to the axis it drops rapidly (Fig. 2.54). Rays of different wavelengths, therefore, reach their maximum intensity in the lower region of the plasma flame, where the excitation potential of the radiation for a given wavelength is the highest (Fig. 2.55).

In practice, the spatial inhomogeneity of the radiation in the plasma flame does not lead to errors, provided the emission from the entire plasma flame or from a certain portion of it is taken into account. It can, however, be proved experimentally that for the Krantz system, horizontal-axis plasma jet, only a 10 cm long section of the 50 cm flame column can be used for analytical purposes.[6]

In summary, the d.c. plasma jet combines the advantages of arc and flame excitation in such a way that the drawbacks of neither method appear. For example, it combines the advantages of arc excitation sensitivity with the reproducibility of flame excitation. Furthermore, the plasma jet is not only of high luminosity, but the spectra

obtained from it are also relatively free from background. Thus with the plasma jet more accurate results can be achieved and the technique lends itself to a wider range of applications than the arc excitation method. One further advantage is that the temperature of the plasma flame is higher than that of the arc plasma, but lower than that of the spark plasma. The plasma jet is, in this respect, an intermediate

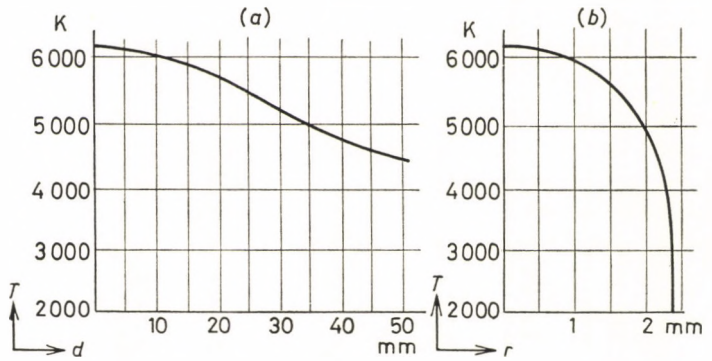


Fig. 2.54 Temperature distribution in the flame of the Kranz-system plasma jet; along the plasma axis (curve *a*) and perpendicular to the plasma axis (curve *b*)[6]

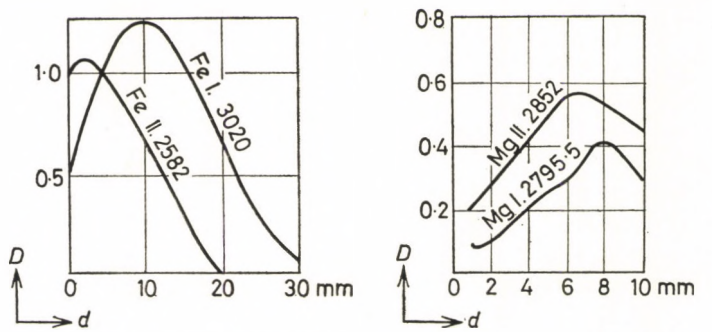


Fig. 2.55 Density variations of atomic and ion lines with the height of the plasma jet[2,5]

solution between the above two excitation methods. This has the consequence that by using the plasma jet not only the alkaline earths and the alkali metals can be excited, but most of the elements that are more difficult to excite as well. According to the few data available, the application of the d.c. plasma jet secures a relative analytical precision of 3–4%.[1,5,8–10] It can, therefore, be expected that the plasma jet will assume an important future role as a light source in emission analysis.

References

1. Margoshes, M. and Scribner, B. F., *Spectrochim. Acta*, **15**, 138 (1959)
2. Korolyev, V. V. and Vainstein, E. E., *Zh. analit. Khim.*, **14**, 658 (1959)
3. Maeker, H., *Z. Naturf.*, **11a**, 157 (1956)
4. Owen, L. E., *Appl. Spectrosc.*, **15**, 150 (1961)

5. NEEB, K. H. and GEBAUHR, W., *Z. analyt. Chem.*, **190,** 92 (1962)
6. KRANZ, E., *Mber. dt. Akad. Wiss. Berl.*, **5,** 308 (1963); *Emissionspektroskopie*, Berlin, p. 60 (1964)
7. LENSING, J. and VAN CALKER, J., *Z. analyt. Chem.*, **198,** 9 (1963)
8. KOROLYEV, V. V. and VAINSTEIN, E. E., *Zh. analit. Khim.*, **15,** 413, 686 (1960)
9. SERIN, P. A. and ASHTON, K. H., *Appl. Spectrosc.*, **18,** 166 (1964)
10. LERNER, R., *Spectrochim. Acta*, **20,** 1619 (1964)
11. GREENFIELD, S., JONES, I. L. and BERRY, C. T., *Analyst, Lond.*, **89,** 713 (1964)
12. SIROIS, E. H., *Analyt. Chem.*, **36,** 2389, 2394 (1964)
13. GOTO, H. and ATSUYA, I., *Z. analyt. Chem.*, **225,** 121 (1967)

2.4.11 Excitation by high-frequency plasma jet

When plasma jets are used the excitation can be effected not only by direct current but also by high-frequency current.[1–5] These latter light sources are different from the d.c. plasma jets, not only in their construction, but in their fundamental operating principles as well.

In high-frequency plasma jets the gas stream carrying the atomised solution passes through a strong high-frequency electric field. The electrons present in the streaming gas adsorb energy owing to the presence of the high-frequency electric field. Their kinetic energy is thereby increased to such an extent that when they collide with the particles in the gas stream, some of the atoms (or molecules) present are partially excited and ionised (or dissociated). The plasma arising in this manner is pushed out from the plasma jet by the gas streaming into the electric field, such that the excited atoms and molecules form a luminous 'flame' above the orifice of the jet. The high-frequency plasma jet is, therefore, a stabilised high-frequency arc, in which the volume of the plasma is controlled mainly by electromagnetic forces. In the latter process, however, thermal effects also play some role (see below).

It can be shown experimentally that a high-frequency plasma flame is produced only if the gas flowing into the high-frequency electrical field contains a sufficiently large number of free electrons. Thus, the high-frequency plasma jet should be initiated either by a high-frequency arc or by a Tesla spark. The electrons and ions produced by the ignition current are accelerated in the electric field such that numerous collisions occur and the number of charged particles in the carrier gas increases rapidly. The charged particles then ionise and excite a sufficient number of the particles present in the in-coming spray to start the plasma flame.

It can be shown experimentally that the plasma flame will only produce a stable burn if the distribution of the electrical condition between the internal and external fields is satisfactory and the gas flow-rate and volume of spray per minute are adequate. It is, therefore, absolutely necessary to provide a means for changing this spatial distribution. It is also important to see that the internal electrode is properly cooled or the plasma flame will be extinguished.

Experiments with high-frequency plasma jets show that the electrons necessary for maintaining the plasma flame do not originate from the electrodes, but are produced from within the plasma.[7] The electrons vibrate in the high-frequency electric field with an amplitude in the order of 10^{-2} mm and generally take up energy at each collision independent of their direction of motion. Under such conditions the replenishment of the electrons is less important in the high-frequency plasma jets than in the case of d.c. arc generation. In the latter case, the electrons can only take

up or release energy during their migration between the cathode and anode and from the point of view of excitation they are lost at the cathode. In arc excitation, therefore, the replenishment of the electrons is absolutely necessary, although high-frequency plasma jets can be sustained without a supply of additional electrons.

The construction of the devices necessary for the production of high-frequency plasma flames varies with the frequency used. (In practice the value of the frequency is limited to ranges normally allowed for industrial purposes – about 27 460 or 2400 Hz.) The usual equipment is always built from two parts: the high-frequency generator and the plasma jet proper (Fig. 2.56).

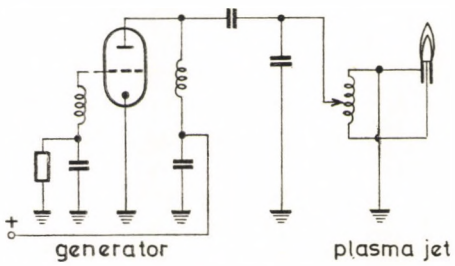


Fig. 2.56 Circuit diagram of the 27·12 MHz generator and the plasma jet[7]

In the case of high-frequency excitation techniques the plasma burner is composed of two electrodes which form the 'internal and external connections'. These electrodes are connected to the high-frequency generator in such a way that a very strong (about 1 kV/cm) electric field is established between them. The gas carrying the atomised particles of the sample solution is then made to pass through this field (Fig. 2.57). The plasma flame is ejected from the tip of the water-cooled internal electrode. The electrode tip is usually made of some refractory metal like molybdenum.

The high-frequency plasma flame consists of well-separated zones (Fig. 2.57). The 'core' of the plasma is the actual light source and is surrounded by a less luminous 'mantle'. Right above the tip of the internal electrode a glimm layer is situated. As the different parts of the plasma flame have different temperatures, the excitation conditions will also be different.

The height of the high-frequency plasma flame is influenced by several factors. First of all it can be established that the height of the plasma flame is increased proportionally with the input electrical power to the plasma. When the latter is increased the temperature of the plasma is elevated as well. Due to the effects of flow phenomena there is a smaller increase in the flame height when the flow-rate of the carrier gas is greater. When the flow-rate is too high the plasma flame drifts away from the burner and if it is too low electrical discharges will develop directly between the internal and external electrodes. In this latter case, a necessary condition for creating the plasma flame, that is, the quick removal of the charged particles from the electrode gap, is not satisfied.

In addition to the above factors the height and temperature of the high-frequency plasma flame are particularly influenced by the chemical composition and the physical properties of the carrier gas. Among the physical properties involved, the thermal conductivity plays an especially important role. For example, when a good thermal conductor such as helium is used, the 'cold' and relatively short flame that develops

is due to the property mentioned above. The temperature and the geometrical form of the plasma flame are also influenced by the dissociation energy of the gas. It is, therefore, an advantage to use a gas, such as oxygen, that has a comparatively low dissociation energy (5·08 eV) to produce a plasma flame of higher temperature and greater length.

From another point of view it is important to secure for as long as possible the lifetime of the metastable state of the carrier gas. As the energy released by the high-frequency generator drops to zero in a half-period, each plasma flame should be reignited many times a second. This condition, which is rather difficult to meet, becomes unnecessary if the lifetime of the metastable atoms in the gas is greater than the length of the intermission between the burning periods. This happens because the excitation energy accumulated by the metastable atoms in the gas is enough to last over the intermission periods and any secondary collisions that occur are sufficient to bring about reignition. From this point of view it is an advantage to use nitrogen as the carrier gas as the metastable lifetime of nitrogen is about one second.[8]

In order to keep the background intensity of the spectra at a minimum a carrier gas should be selected that has a high excitation potential, for example argon.

Up till now, nitrogen or air was used in high-frequency plasma jets, as both gases retain atoms with relatively long metastable lifetimes. In many instances it is preferable to use oxygen if a high-temperature plasma flame is required. The temperature of such high-frequency plasma flames varies along the axis of the flame from between 4000 K and 3000 K such that the highest temperature is observed at the bottom of the flame.[6]

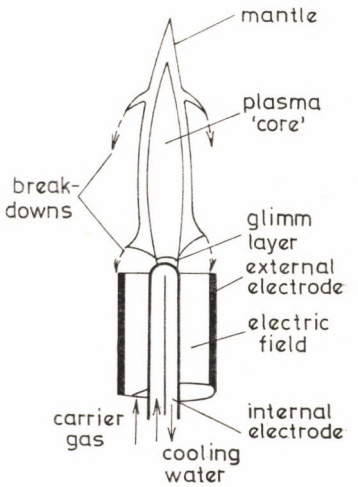


Fig. 2.57 Construction of the high-frequency plasma jet, and the structure of the plasma flame

At present there are few data available for deciding the analytical applicability of high-frequency plasma jets.[9] The following statements, however, are important in this context:

(1) The *flame temperature* of high-frequency plasma jets is lower than the d.c. plasma flame, although higher than the temperature of flames used in flame excitation techniques.

(2) As the intensity fluctuation of the spectral lines with time is generally about 3–4%, excitation stability meets the normal analytical requirements.
(3) The spectral background is low. The origin of this background can be traced to spectral bands produced by the carrier gas and the solvent.
(4) *Interelement effects* only appear for elements of low ionisation potential; no effects are seen for elements of high ionisation potential.
(5) The detection limit is generally much better when high-frequency plasma jet excitation is used, than that obtained by the flame-photometric procedure.
(6) The *alkali metals are exceptions* with respect to point 5. This anomaly can probably be traced to the fact that the alkali atoms are in a highly ionised state in the plasma flame. The relative number of non-ionised atoms necessary for the excitation of their fundamental lines is therefore small.

References

1. COBINE, J. D., *J. Appl. Phys.*, **22**, 835 (1951)
2. MOLLWO, L., *Ann. Phys.*, VII. **2**, 97 (1958)
3. CRISTESCU, G. D., *Revue Phys., Buc.*, **1**, 103 (1956); **4**, 153 (1959)
4. SCHMIDT, W., *Elektron. Rdsch.*, **13**, 11, 404 (1959)
5. SCHOLZ, O., *Umschau*, **59**, 716 (1959)
6. TAPPE, W., and VAN CALKER, J., *Z. analyt. Chem.*, **198**, 13 (1963)
7. JECHT, U., and KESSLER, W., *Z. analyt. Chem.*, **198**, 27 (1963)
8. FONER, S. N., and HUDSON, R. L., *J. Chem. Phys.*, **37**, 1662 (1962)
9. GOTO, H., KICHINOSUKE, H., and MITSUYASU, S., *Z. analyt. Chem.*, **225**, 130 (1967)

2.5 Spark excitation

2.5.1 Spark discharges[1–15]

In all the analytical techniques that come under this general heading of spark excitation the excitation is brought about by a series of sparks. The phenomenon usually called a spark is in fact a quickly terminated electrical discharge between two current-conducting bodies (in this case the electrodes), provided that the potential difference between them exceeds the value of the breakdown voltage. The spark discharge is produced by charged particles that are always present in the electrode space. These particles are accelerated by the electrical field such that they are able to ionise those atoms with which they collide. This process takes place mainly in the area immediately in front of the electrodes, where the field effect is the greatest. On the electrodes a weakly emitting discharge layer of very short duration, the so-called *corona discharge*, is established. As the ionisation process quickly spreads over the whole of the electrode gap, the number of charged particles and the frequency of collisions are increased rapidly. Thus the electrode gap suddenly becomes conducting, and the 'spark' proper is produced.

In spectrochemical analysis the only sort of electric discharge that can properly be called a spark is that produced when a condenser is discharged. The discharge of a condenser which has a sufficient amount of energy lasts until the work of the electrons passing through the spark gap and the ohmic resistance of the discharge circuit (that is, the energy loss of the condenser) is smaller than the amount of energy instantaneously stored within the condenser. The condenser plates change

their charge after the first discharge. The condenser is now again discharged and the polarity of the plates again changes until the terminal voltage of the condenser is reduced so much that the potential becomes insufficient to bring about the breakdown of the spark gap. The whole process normally lasts for about 10^{-5} s.

In a condenser discharge the following processes can be distinguished: ignition, formation of the spark plasma, full discharge of the condenser, afterglow, and finally after-effect.[1]

The ignition of the condenser discharge occurs when the voltage between the electrodes that form the spark gap reaches the value of the breakdown voltage of the electrode gap. The value of the breakdown voltage depends not only on the distance between the electrodes, but also on their geometrical form, on the type of electrode material used, and on the pressure and temperature of the gas between the electrodes. The breakdown voltage also depends on the ionisation state of the electrode gap. The ionisation condition of the electrode gap is generally uncertain, especially as it changes during the condenser discharge. The ignition voltage is, therefore, somewhat unpredictable, even for certain fixed experimental conditions. The spark gap process at the beginning of the condenser discharge can be summarised as follows.[2,3] Initially a narrow, thread-like 'discharge channel' is established between the electrodes and a very strong detonation wave – whose strength exceeds 1000 torr/cm – is produced like in an explosion. The wave-front of the latter first ionises the atoms of the gas present and then causes the ionised particles to move with a velocity of 2–3 km/s. The diameter of the breakdown channel remains relatively small during this process. Thus, the current density in the breakdown channel of the spark attains very high values, of about 10^5–10^6 A/cm^2, as compared to the usual current density in arc generation methods of 10^2–10^3 A/cm^2. Owing to the effect of this high current density the breakdown channel is heated up to a temperature greater than 50 000 K in a very short time; generally about 10^{-8} s. The diameter of the channel is not increased since the processes leading to broadening (radiation diffusion, ambipolar diffusion) are relatively slow. In this initial stage of the condenser discharge, therefore, the resistance of the plasma is comparatively great ($>10\ \Omega$), despite the high rate of ionisation caused by the very high plasma temperature. Thus, the current intensity passing through the discharge channel is relatively low and the amount of electrical energy used for the ignition is comparatively small. Accordingly, there is only a very slight drop in the condenser voltage in the course of the ignition (Fig. 2.58).

It follows from the above that a condenser charged up to the discharge voltage is able to discharge itself. On the other hand, the condenser charged to a not too high potential is discharged by an ignition current. This is called the *ignition* of the spark.

Although the plasma temperature remains high, the breakdown channel starts expanding after about 10^{-7} s or less, and assumes the form of the spark plasma. In the meantime, the resistance of the spark plasma is continually reduced until it reaches its minimum value somewhere in the region of 0·1 Ω. Simultaneously, the voltage drops to the level of the 'spark voltage', which is nearly the same as the 'arc voltage', that is, about 50 V. The whole process, that is, the formation of the spark plasma proper, occurs within a half-period of the cycle of the condenser discharge. The processes taking place during this time, that is, the variation of the discharge current, I, the voltage, U, the resistance, R_S, of the spark plasma and the radius, r, of the latter, are depicted in Fig. 2.58.

The dilatation of the spark plasma is caused mainly by radiation diffusion and ambipolar diffusion.[3] The former effect is caused by rays having a wavelength shorter than 1500 Å, being absorbed by the external part of the plasma in such a manner that the energy released is used for excitation and ionisation. This process, which lasts for a few times 10^{-8} s, is repeated until no more light quanta are absorbed.

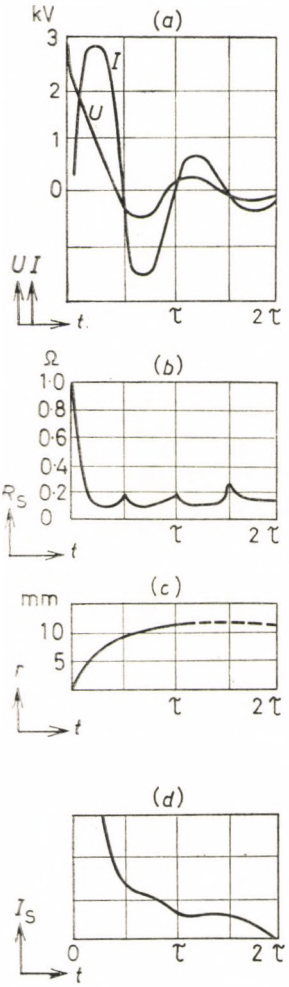


Fig. 2.58 Formation of the spark plasma: (*a*) current–time, and voltage–time curves, (*b*) spark resistance–time curve, (*c*) plasma radius–time curve and (*d*) relative optical performance–time curve. Excitation: ^{3}Ar (400 torr), $U = 3$ kV, $C = 32$ μF, $\tau = 34$ μs, $2\sqrt{L/C} = 0{\cdot}35$ Ω

The ambipolar diffusion of the charged particles is caused by the temperature drop within the outer layers of the plasma, such that an electrical field is established on the outskirts of the plasma that delays electron diffusion and speeds up ion diffusion until the two diffusion rates become equal. It can be shown experimentally that dilatation occurring in this way obeys the relationship between the radius, r, of the luminous part of the discharge channel and the duration, t, of the discharge as follows

$$r \approx \sqrt{t} \tag{2.52}$$

There is no further dilatation of the discharge channel when the plasma does not take up any more energy and thus the plasma temperature drops due to energy loss.

The rate of both types of diffusion is strongly influenced by the temperature. It is the lowering of the temperature that causes the plasma to contract upon the reduction of ionisation.

At the beginning of the formation of the spark plasma the concentration of the charged particles in the discharge channel is not very high. The resistance of the spark gap is, therefore, significant compared with the inductive resistance of the discharge circuitry. The condenser discharge produced in this way is an aperiodic process that is apparently not influenced by the self-induction of the circuitry. This restriction characterises the initial stage of the condenser discharge, during which the formation of the spark plasma occurs.

It is also characteristic of the previously mentioned discharge phase of the condenser that the dimensions of the discharge channel are not altered by the transformation of the thermal plasma into spark plasma. This can be explained by the charged particles, especially the electrons, migrating very fast along the axis of the discharge channel. On the other hand, the radiation diffusion and the ambipolar diffusion of the charged particles are comparatively slow processes. Thus these processes are practically ineffective during the transformation period of the plasma.

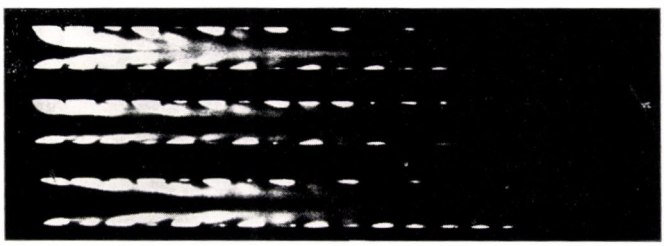


Fig. 2.59 Time sequence pictures of spark discharges in a time interval of 2.15^{-4} s

During the condenser recharges that occur in the course of the formation of the spark plasma, the number of charged particles and thus the current value are significantly increased. As a consequence, the electrodes are heated up to such a degree that gas emanations occur on the surfaces. The vapour clouds arising in this way, which emit light for a longer period than the recharging period of the condenser, provide the radiation of the electrode material. The vapour emanations leave the electrode gap at a very high speed (100–1000 m/s) and their light emission remains after the termination of the discharge for about 10^{-5} s (Fig. 2.59).

In the third stage of the condenser discharge the condenser is fully discharged. That is, the electric energy accumulated within the condenser is directly and almost entirely transformed into thermal energy. This process is, however, influenced by circuit parameters as well. During this discharge process the condenser releases only that part of the accumulated energy that is not dissipated in the external circuit. The plasma reaches its maximum temperature, or its maximum degree of ionisation only after the lapse of 10^{-7} s.

The light energy radiated from the plasma immediately after the ignition of the spark is only a very small portion of the energy stored within the plasma. The portion of the energy that appears as light increases as the plasma temperature is elevated, although it is still relatively small even at the end of the full discharge of the con-

denser. This happens because the plasma starts radiating light in greater proportions only after the condenser has practically been discharged. This phenomenon is the so-called *afterglow* and constitutes the fourth stage of the condenser discharge.

The major portion of the energy of the arc plasma is transformed into heat. Only a small portion of the total radiation is responsible for the infra-red, visible, and ultra-violet spectrum. The far ultra-violet rays, which form a significant proportion of this light, are intensely absorbed by the gases surrounding the discharge channel of the spark and only a small fraction escapes the 'gas mantle'. Due to the effect of the absorbed light energy, however, excited atoms, ions, and electrons are created in the gas mantle that diffuse outwards in all directions, such that the plasma is extended in space. The same phenomenon develops as that also caused by thermal conduction or by the charged particles that diffuse perpendicularly to the axis of the plasma.

The energy loss of the plasma during the dilatation process is only caused by the light energy actually emitted. Therefore the energy content of the plasma does not change owing to the dilatation, although it does extend over a wider area, such that the temperature of the plasma is reduced as a consequence of the dilatation. In spite of this, as can be shown theoretically, the maximum light emission of the spark discharge happens only after the dilatation of the plasma, or when the temperature of the plasma is not at its maximum.

During the afterglow the temperature of the plasma is essentially reduced. As a consequence of this the intensity of the longer-wavelength portion of the emitted light is enhanced and the ratio of the energies emitted to the energy retained in the plasma is increased. During the afterglow, therefore, the so-called *internal cooling* of the plasma also takes place. The charged particles present in the plasma release their kinetic, excitation, and ionisation energies such that the kinetic energy is increased of the atoms that are originally present or that arise in the meantime. The result of this energy transformation is that the electron temperature of the plasma drops and the gas temperature increases until the two temperatures become equal.

The fifth effect of condenser discharges – the so-called *after-effect* – is caused by the gas in the electrode space not returning to its original condition even after the total cooling-down of the discharge channel, when the latter disappears. The condition before the ignition of the spark is, therefore, not reached again. The equilibrium between the temperature and the ionisation degree is reached comparatively slowly and only when the charged particles are sufficiently recombined or migrate into the outer gas region. The *after-effect* is seen for a short time when the ionisation in the electrode gap is greater than it should be according to the temperature prevailing there. From the practical point of view the after-effect has the important feature that it facilitates the ignition of the subsequent spark after a short time interval.

It can be experimentally established, provided the initial discharges of the condenser are neglected, that during the entire discharge period of the condenser the voltage between the electrodes is practically the same as the arc voltage during the 'burning' time of the spark. The voltage–time relationship of the condenser accordingly forms a meandering curve during the charging of the condenser plates to opposite charges, while after the termination of the condenser discharge it transforms into a curve of variable shape (Fig. 2.60).

From the above it could be assumed that spark excitation is a form of a.c. arc generation in which the periodically recurring burning times are extremely short (of

the order of 10^{-6}–10^{-5} s. This assumption would, however, be correct only when the stationary states characterising the arc plasma occur in spite of the very quick and extensive changes in the voltage and current brought about by spark discharges.

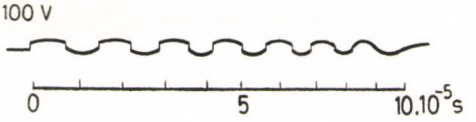


Fig. 2.60 Voltage–time oscillogram of the spark discharges between Mg electrodes

It can be shown by calculation that the following intervals are necessary for the completion of the different processes taking place in the plasma or for the establishment of the equilibrium state:[1]

for the Maxwell-type velocity distribution of the electrons	about 10^{-13} s
for the occurrence of the ionisation	about 10^{-12} s
for the development of the excitation energy	about 10^{-12} s
for the energy transfer among the electrons and the atoms	about 10^{-5} s
for the Maxwell-type velocity distribution of the atoms	10^{-10}–10^{-3} s

In the case of spark excitation, the current passing through the electrode gap increases in about 10^{-6} s from zero to its maximum value and then drops again to zero. During this time the temperature of the spark plasma fluctuates very strongly. Therefore, those processes that have reached their equilibrium state in the arc plasma have not usually done so in the spark plasma. In spark discharges a distinction should be made between the electron temperature, corresponding to the Maxwell-type velocity distribution of the electrons, and the *gas temperature* corresponding to the Maxwell-type velocity distribution of the neutral particles. The electron temperature determines the translational kinetic energy of the electrons, as well as the ionisation and excitation energies. On the other hand, the average energy of the atoms and their thermal conduction are functions of the gas temperature. It follows from the above that *the spark generally refers to the 'non-stationary' arc state.* In spark discharges any approximate stationary condition can be achieved only if the discharge is damped by a suitable ohmic or inductive resistance. This situation does not exclude the possibility of the occurrence of local and instantaneous thermodynamic equilibria within the spark plasma.[11]

From time–resolution curves of high-voltage spark discharges[4-6] it can be established that at the beginning of the condenser discharge a strong background is produced in the spectrum, followed later by the spectral lines of the gas in the electrode gap (Fig. 2.61). These *atmospheric lines* appear only for a very short time, but before their disappearance the spectral lines of the vapour stemming from the electrode metals appear in the spectrum. These metal-vapour lines appear initially in the neighbourhood of the electrodes and can be observed only in the radiation originating from the central parts of the spark gap when the background and the air lines disappear from the spectrum.

It can also be established experimentally that the intensity of spectral lines having a high excitation level more or less follows the fluctuations of the intensity of the current during the spark discharge. These lines, therefore, disappear every time when the current reaches zero due to the polarity interchange of the electrodes between half-periods. On the other hand, the spectral lines of those elements that can be easily excited exist throughout the discharge of the spark. Furthermore, it can be

shown that the momentary fluctuations of the light intensity are always greater in front of the negative electrode than the positive electrode. This observation is interesting, since the glowing metal vapour mainly comes from the positive electrode. The vapour is cooled down so much during its travel that only the easily excited spectral lines appear in front of the negative electrode. The vapour emanations alternately appear and disappear from the electrode area every half-period. This is the cause of the intensity fluctuation of the radiation that originates from the electrode spaces.

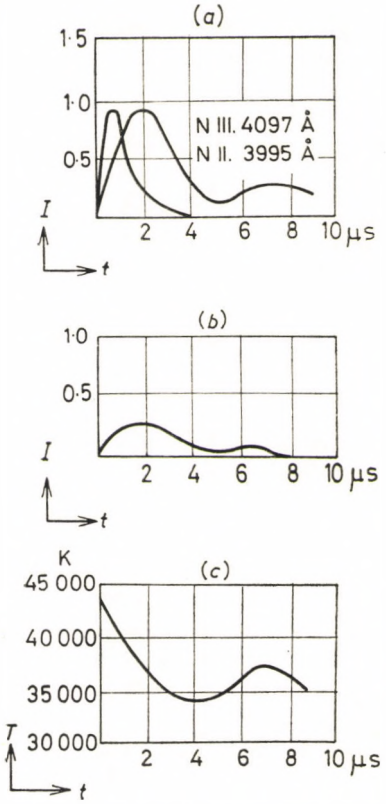


Fig. 2.61 Schematic variations with time of (*a*) the intensity shown by nitrogen lines of different excitation potential, (*b*) the intensity of the background and (*c*) the electron temperature of the plasma at the beginning of the spark discharges. $C = 0{\cdot}25\ \mu$F, $L = 10\ \mu$H, and $U = 15$ kV [2]

The intensity of the spectral lines originating from the central part of the spark gap does not change in a periodic manner, provided the material of the two electrodes is the same. This observation can be accounted for by the fact that the metal vapours arrive at the centre of the spark gap from the two electrodes with a phase difference of 180°. Thus, the supply of the vapour to the middle part of the electrode gap is twice as much as that to the space immediately in front of the electrodes. Another factor that plays a role here is that the cooling effect of the electrodes does not manifest itself in the middle of the spark gap. The time resolution of the spark discharges also proves that following the termination of the electrical discharge there is an afterglow period.[4-6]

From data obtained by different measuring methods it can be calculated[3,9] that the following values hold for the spark plasma:

average gas density	$\rho \approx 4.10^{-3}$ g cm^{-3},
mean gas temperature	$T_g \approx 40\,000$ K,

average concentration of free electrons $c_e \approx 10^{17}$ electrons/cm^3.

Also, the resistance of the plasma is proportional to[2]

$$R_s \approx \sqrt{\frac{pd^2}{1/2CU_0^2}}$$

where, provided the external inductance can be neglected, p is the gas pressure, d is the distance apart of the electrodes, C is the capacity of the condenser and U_0 is the charging voltage.

It also follows from the above that the temperature of the spark plasma changes both in time and space during the discharge of the condenser. This temperature change influences the degree of local ionisation in the plasma and, thus, the light

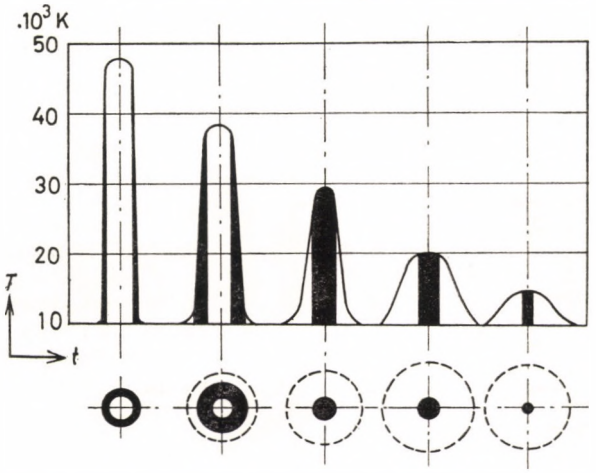


Fig. 2.62 Changes in the place of origin for the spectral line Al III 3601 (the black part of the diagram) during the time of the condenser discharge[11]

emitted by atoms of different elements in different ionisation states (Fig. 2.62). The emission coefficient, (T), of a spectral line having an excitation potential, U_g, is proportional to the number, n_i, of particles emitting the particular radiation, that is

$$\varepsilon(T) \approx n_i e^{-U_g/kT} \tag{2.53}$$

Accordingly, the spectral structure of the radiation from the spark plasma changes both in time and space during the discharge of the condenser.

It can be established by interferometric measurements[3] that after the lapse of 2·3 μs following the initiation of the condenser discharge, the gas temperature of the spark plasma and its relative gas density are distributed in the spark discharge channel in the manner shown in Fig. 2.63. As shown, the radial change of both the temperature and the gas density is more or less like a Gauss distribution curve. (The step features appearing at the fringes of the discharge channel are caused partly by the shock wave and partly by the plasma mantle.) The highest temperature, T_0, that arises in the centre of the plasma is called the *axis temperature*.

It can be proved both theoretically and experimentally[12,13] that during the condenser discharge the plasma takes up the greater part of the energy during the very first microsecond. The cooling period of the spark plasma commences immediately after this event. During this cooling period the temperature of the spark plasma

fluctuates continuously between the initial very high value and the final temperature (which is the same as that of the surroundings). As can also be shown,[11] the cooling curve of the spark plasma core is identical to the independent cooling curve of the incandescent metal filament; that is, in the first period of the cooling process the core temperature follows the formula

$$T_0 = T_{00}\, e^{-t/\tau} \tag{2.54}$$

where T_{00} is the initial temperature and τ is the time constant. The above formula can only be applied up to the point where the temperature of the plasma is comparatively high, or in other words, while the dilatation of the spark discharge channel

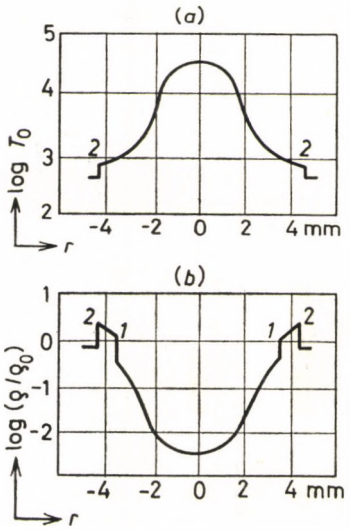


Fig. 2.63 The radial distribution of (*a*) the gas temperature and (*b*) the gas density in the discharge channel of the spark, when $t=2{\cdot}3$ μs, $L=2$ μH, $C=0{\cdot}25$ μF, and $U=10$ kV. The number 1 on the curves is the place of origin of the shock wave, while 2 is that of the plasma mantle

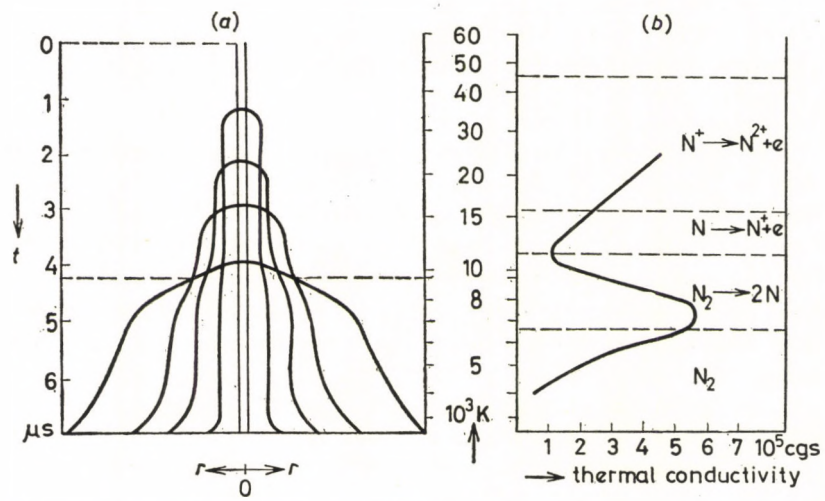


Fig. 2.64 Schematic cooling of the spark plasma in air. (*a*) Variation of the radial temperature distribution with time, and (*b*) variation of thermal conductivity with temperature for nitrogen[14]

is not yet significant. This is because at lower temperatures, during the cooling of the spark plasma core, an increasingly wider cold mantle is established around it (Fig. 2.64*a*). The dilatation of this cold mantle is especially great around 10 000 K in the presence of nitrogen. Consequently, the mantle broadening is transformed into a dome-shape, at that temperature which favours an increasing association of the nitrogen atoms into N_2 molecules. The latter process is characteristically strongly exothermal and, consequently, the classical thermal conductivity of the gas is greatly increased (Fig. 2.64*b*). Thus the core of the plasma, which has a relatively high temperature, transfers its energy to the cooler dome-shaped mantle, not only via ambipolar diffusion but largely classical thermal conductance. In the last period of the cooling of the plasma the core is demolished at a greater speed than the dome, since the cooling of the latter is delayed by the thermal energy set free in the formation of the N_2 molecules.[14]

References

1. WEITZEL, W. and ROMPE, R., *Theorie elektrischer Lichtbögen und Funken*, Leipzig (1949)
2. GLASER, G., *Z. Phys.*, **143**, 44 (1955)
3. MANDELSTAM, S. L., *Spectrochim. Acta*, **11**, 245 (1956)
4. BARDÓCZ, Á., *Magyar Fiz. Foly.*, **7**, 195 (1959)
5. BARDÓCZ, Á., *Acta Tech. Hung.*, **40**, 179 (1962)
6. BARDÓCZ, Á., *Spectrochim. Acta*, **16**, 1173 (1960)
7. FRIND, G., *Z. angew. Phys.*, **12**, 251, 515 (1960)
8. BARDÓCZ, Á., *Proc. IX. Coll. Spectr. Intern.*, Lyon, 144 (1962)
9. MEEK, J. M. and GRAGGS, I. D., *Electrical Breakdown of Gases*, Oxford (1953)
10. HULDT, L., *Spectrochim. Acta*, **7**, 264 (1955)
11. KREMPL, H., *Z. Phys.*, **167**, 302 (1962)
12. WEITZEL, W., *Z. Phys.*, **135**, 639 (1953)
13. ABRAHAMSON, I. S. and GEGETSCHKORI, N. M., *Abh. Sowj. Phys.*, vol. III, Berlin (1953)
14. MAECKER, H., *Z. Phys.*, **157**, 1 (1959)
15. GRIEM, H. R., *Plasma Spectroscopy*, New York–London (1964)

2.5.2 Determination of the temperature of the spark plasma

The 'two-line' method, generally successful in the measurement of the arc temperature, is not applicable for the determination of the temperature of the spark plasma. It can be shown that this method gives temperature values that are too low.[1,2] This can be interpreted as being due to the emission coefficient of the atomic lines applicable for the two-line method as having already passed through a maximum between 5000 and 10 000 K (Fig. 2.65). Thus the two-line method can only be used for the measurement of plasma temperatures essentially higher than 10 000 K, provided the determination is performed using the spectral lines of multiple-ionised atoms.[11] The application of these lines is hindered, however, as either the corresponding transition probability values are unknown or the precision of the published data is not sufficiently high.

The above problems are eliminated when in the determination of the temperature of the spark plasma the so-called *characteristic temperature* values of the spectral lines ('norm temperature') are taken into account. This is a temperature characteristic of the spectral line in question, at which the emission coefficient of the spectral line passes a maximum.[3,6] It is generally denoted by $\tilde{T}$, or, for atomic lines, $\tilde{T}_0$, and

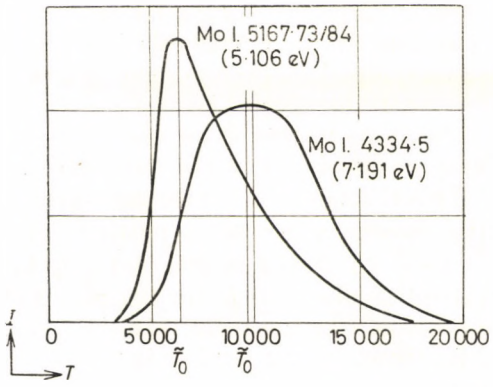


Fig. 2.65 Variation of atomic line intensity with plasma temperature

for ion lines $\tilde{T}_i$. This value is more or less the same temperature at which the intensity of the spectral line is considered to be at its maximum value. These values can be calculated, for example, by using the formula from Eq. (2.53).[3] Table 2.7 shows the 'characteristic temperature' of certain spectral lines, as determined by other methods.[4]

Table 2.7

The 'characteristic temperature' ($\tilde{T}$) of different spectral lines, and the duration of the intensity maximum ($\tilde{t}$)

Spectral line	$\tilde{T}$ (10^3 K)	$\tilde{t}$ μs	Spectral line	$\tilde{T}$ (10^3K)	$\tilde{t}$ μs	Spectral line	$\tilde{T}$ (10^3 K)	$\tilde{t}$ (μs)
Al I 3 961	8·1	4·4	Pb II 2 203	17·0	2·3	Pb III 2 562	29·5	1·7
Pb I 2 663	8·8	4·5	Al II 2 816	18·4	2·8	C III 2 297	35·5	1·2
Pb I 2 393	9·4	7·5	Si IV {4 089 / 4 116	25·9		Pb IV 2 864	40 0	0·75
Pb I 2 170	9·5	4·2	C II 2 512	28·5	1·55	C III 4 651	41·5	
C I 2 478	12·8	3·4	Al III 2 601	29·0	1·7	N III 2 097	48·0	0·75

The data listed in the above table relate to the case when the spark plasma is composed only of the element in question. According to Eq. (2.53), however, the intensity of the spectral line and the pertaining characteristic temperature as well are functions of the partial pressure of the element under consideration. Thus, it can be expected that upon diluting the vapour of the element in question by other elements in the plasma, the 'characteristic temperature' is shifted. It can be shown that this *shift* of the characteristic temperature is considerable at the lower temperatures that characterise the arc plasma. On the other hand, at the higher temperatures of the spark plasma this shift is negligible. This difference is due to the fact that while at relatively low temperatures the electron pressures of various elements are different by many orders of magnitude, at temperatures exceeding 15 000 K it only amounts to a few per cent.[4] Considering that the characteristic temperature is only a function of the logarithm of the electron pressure, it follows that at high plasma temperatures the presence of the other elements will be practically negligible.[5] Accordingly, it can be stated that at high temperatures the relative intensity of the spectral lines is only

slightly influenced by the elementary composition of the plasma. Quite probably the other observation that interelement-effect in the plasma is less of a problem in spark excitation than in arc generation, can also be traced to the previous reason.[4]

When it is considered that the temperature of the plasma drops continuously from the initial highest value to the final lowest one during the cooling period, then it is evident that in a given time the temperature, $\tilde{t}$, will reach the characteristic value $\tilde{T}$ for a given spectral line. Therefore if the time, $t = \tilde{t}$, is measured from the initial breakdown of the spark gap, until the maximum of a suitably chosen spectral line is reached, and if the appropriate value of $\tilde{T}$ is known, a cooling curve for the plasma can be constructed. If it is assumed that the light emitted, corresponding to the spectral line, originates from the core of the plasma, then according to Eq. (2.54) there is a linear relationship between $\log \tilde{T}$ and $\tilde{t}$; see also Fig. 2.66. It follows from the above that the temperature of a spark plasma can only be established by resolving the time of emission by the plasma. This has already been achieved.[6–11] It is, however, absolutely necessary to determine a distribution measurement of the spark discharge as well. This necessity emerges since the intensity of the spectral lines is

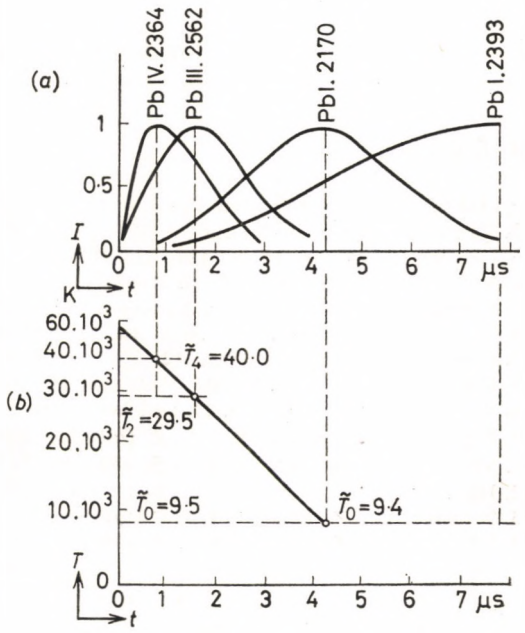


Fig. 2.66 Cooling curve of the plasma[11]

determined by both the axis temperature, T_0, of the spark plasma and the radial distribution of the temperature. The radial resolution of the radiation is rendered impossible, however, by the unavoidably erratic movement of the spark.

The problem described above can be eliminated by using the so-called 'trial-and-error' method. In this procedure a spark channel of radius r is divided into several parts and only those values of $T = T(T_0 r)$ are taken into account that show the best agreement with the experimental results. It can be theoretically established that the best results are obtained when the discharge channel is divided into hollow elementary parts whose volume is proportional to the product $r.dr$.[10] When the latter is multi-

plied by the appropriate emission coefficient $\varepsilon(T)$ and the product is integrated, the following relationship is obtained

$$I(T_0) = \text{const} \int_0^\infty \varepsilon[T(r)]r.\mathrm{d}r \tag{2.55}$$

Provided Eq. (2.53) is regarded as valid, the left-hand side of the above equation is known. The distribution $T(r)$ and the channel dilatation $r(t(T_0))$ are, however, unknown quantities. Among the several different possibilities the most useful one is to regard the distribution as being identical to the temperature distribution of the d.c. arc plasma,[12–14] that is, the relation

$$T = \frac{T_0}{1 + r^m} \tag{2.56}$$

is taken into account. In the case of the channel dilatation, on the other hand, the experimentally established relation, $r \sim \sqrt{t}$ [Eq. (2.52)], is taken to be valid. Consideration of all these relationships allows the previously given Eq. (2.55) to be solved by graphical integration.

The relation $I(T_0)$ obtained by the previous method adequately fits the observations, provided $T_0 > \tilde{T}$. At lower temperatures, however, it only proves to be an approximate description of the real situation.

References

1. HULDT, L., *Electrical Breakdown of Gases*, Oxford (1953)
2. VAN CALKER, J. and BRAUNISCH, H., *Spectrochim. Acta*, **7**, 264 (1955)
3. LARENZ, R. W., *Z. Phys.*, **129**, 327 (1951)
4. KREMPL, H., *Z. Phys.*, **167**, 30 (1962)
5. SPATH, H. and KREMPL, H., *Z. angew. Phys.*, **12**, 7 (1960)
6. BARDÓCZ, Á., *Spectrochim. Acta*, **16**, 1173 (1960)
7. FRIND, G., *Z. angew. Phys.*, **12**, 251, 515 (1961)
8. LOCHTE-HOLTGREVEN, W. and MAECKER, H., *Z. Phys.*, **105**, 1 (1937)
9. BRANDO, C. J., CRAGGS, J. D. and WILLIAMS, G. O., *Spectrochim. Acta*, **3**, 540 (1949)
10. BLITZER, L. and CADY, W. M., *J. Opt. Soc. Am.*, **41**, 440 (1951)
11. STETTER, A., *Spectrochim. Acta*, **5**, 350, 460 (1953)
12. HÖRMANN, H., *Z. Phys.*, **97**, 539 (1935)
13. SCHMITZ, G., *Z. Phys.*, **126**, 1 (1949)
14. SPERLING, J., *Z. Phys.*, **128**, 269 (1950)

2.5.3 Basic principles of spark generators

Spark generators are electrical devices used for the production of spark discharges at atmospheric pressure during excitation. The most important parts of the equipment are the damping resistance, R_d, the transformer, Tr, the self-induction coil, L, the condenser, C, and the resistance, R (Fig. 2.67).

For the above purpose another device is required such that a *single* discharge of a condenser charged to a sufficiently high voltage (say, 10 000 V) and of sufficiently high capacity (say, 1000 μF) would effect the excitation. This simple solution would not, however, meet the practical demands of the method. In such a case the non-reproducibility of the physical and chemical processes influencing the sample evap-

oration and the excitation of its vapour cannot be neglected. This source of error becomes statistically irrelevant when the excitation is brought about by a series of capacitor discharges.

Using spark generators, applicable for spectrochemical analysis, the excitation can be effected by a series of condenser discharges of relatively low energy. The condenser should be automatically recharged to the same voltage following each discharge.

Spark generators meeting the above demands are composed of three main parts –

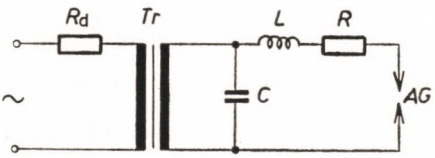


Fig. 2.67 Circuit diagram of spark generators

the charging circuitry of the condenser, the discharge circuitry of the condenser, and the control unit.

Spark generators only meet the practical requirements when, in addition to the above factors, the electrical processes taking place in the two main circuits are in proper balance during the excitation. In some cases irregularities that appear during the condenser discharge (partial discharges, etc.) bring about large analytical errors. This source of error is eliminated by the use of the control unit that has the task of precisely *timing* the condenser discharges. Although the charging and discharging time of the condenser can be controlled by selecting the correct parameter values of the charging circuitry, this sort of control is not entirely free of accidental quirks and is therefore unreliable.

In certain types of spark generators, namely the so-called medium-voltage and low-voltage spark generators, the terminal voltage of the condenser is much lower than the breakdown voltage of the spark gap. The control unit of such equipment not only determines the time of the condenser discharge but also causes the ignition of the spark.

2.6 Processes taking place in the discharge circuit of spark generators

2.6.1 Preliminary remarks

In spark generation a condenser of capacity C charged to voltage U_0 is generally discharged through the induction coil, L, the ohmic resistance, R, and the analytical gap, S. Among the parameters influencing the condenser discharge the values of C, L, and R are constants independent of the current, I, of the discharge. On the other hand, the value of the resistance, R_S, of the spark gap is a function of the discharge current, I, both of which change with time.

It follows from the above that the treatment of the processes taking place in the discharge circuitry of the condenser would be rather complicated unless the time-variation of the value of R_s is neglected. Difficulties arise also if it is taken into account that the spark plasma has a capacitance and inductance that depend on the discharge time and current. It is therefore expedient in the discussion of the electrical processes of spark excitation to neglect the variable parameters of the analytical gap for the sake of simplicity. Thus, it can be assumed that in the discharge circuit of the

condenser the spark gap is short-circuited and its variable resistance, capacitance, and inductance have been included in the values of R, C, and L in the form of some suitably chosen time average. Generally, the parameters of the spark gap can normally be neglected when compared with the R, C, and L values of the circuit. For example, the resistance of a 2 mm spark gap is about 2 Ω when the current is 20 A, while at current values of 60–70 A, it is about 1 Ω and at 200 A the resistance is smaller than 0·5 Ω. Since in spark discharges the current is normally very high the above omission is usually allowed.

2.6.2 Discharge of a condenser*

Let us assume for the sake of simplicity that the discharge circuit of the condenser is induction-free, that is, $L = 0$ and its ohmic resistance is R. Furthermore, that the charge of a condenser of capacity C and terminal voltage U is Q. Then it can be written that

$$Q = CU$$

During the discharge of the condenser the value of the current, I, has the following relationship with the change of the charge of the condenser, Q, with respect to time:

$$I = \frac{\mathrm{d}Q}{\mathrm{d}t}$$

When it is taken into account that $U = Q/C$ then

$$\frac{Q}{C} + RI = 0 \tag{2.57}$$

That is, the arithmetic sum of the potential differences across the two elements of the circuitry is equal to zero. It follows from the differentiation of the last two equations that

$$\frac{1}{C} I + R \frac{\mathrm{d}I}{\mathrm{d}t} = 0 \tag{2.57a}$$

Thus at time t the current value I is given by

$$I = I_0 \mathrm{e}^{-t/RC} \tag{2.58}$$

provided the maximum intensity at the start of the condenser discharge is denoted by I_0. The formula given in Eq. (2.58), therefore, shows the variation of the discharge current of the condenser with time.

The variation with time of the condenser voltage is obtained by taking into account the relationship $I_0 = U_0/R$, and thus we derive from Eq. (2.58) the following

$$U = U_0 \mathrm{e}^{-t/RC} \tag{2.59}$$

where U_0 is the initial terminal voltage of the condenser. It can be seen from the previous two equations that the intensity of the discharge current and the terminal

* In the calculations that follow only the initial data and the final results are given. This will be our practice unless the derivations cannot be found in the relevant textbooks.[1–4]

voltage of the condenser drop exponentially with the time lapse of the discharge. The rate of this drop depends in both cases on the value of the product RC.

The time constant of the discharge circuitry is the lapse of time, τ, during which the initial current value and the initial terminal voltage are reduced to their e-th fraction, that is, when they become reduced by a factor of 2·718. This reduction occurs according to Eqs (2.58) and (2.59) when $t = RC$.

In spectrochemical analysis the time interval, t_v, during which the initial voltage, U_0, of the condenser drops to a certain minimum level, U_{min}, is important. This duration of the condenser discharge is given by the following expression, according to Eq. (2.59)

$$t_v = RC \ln \frac{U_0}{U_{min}} \tag{2.60}$$

For spark excitation it is also important to know the amount of energy accumulated within the condenser. This is identical to the Joule heat released from the resistance, R, during the discharge of the condenser. Its value is therefore

$$E = \int_0^\infty RI^2 \mathrm{d}t \tag{2.61}$$

Taking Ohm's law into account, the above equation can be rewritten as

$$E = \int_0^\infty \frac{U^2}{R} \mathrm{d}t \tag{2.62}$$

Finally, it can be calculated from the former relationship and Eq. (2.59) that

$$E = \frac{1}{2} CU^2 \tag{2.63}$$

The energy accumulated within the condenser is therefore proportional to its capacitance and to the square of its terminal voltage.

In the above it has been assumed for the sake of simplicity that the discharge circuitry of the condenser is induction-free. In reality this is not so, for when the discharge circuit of the condenser does not contain a switched-on induction coil, it is not free of its own so-called 'remanent inductance'.

When the discharge circuit of the condenser has *inductance* too, then there is a voltage drop across the induction element during the discharge. Taking this voltage drop into account, Eq. (2.57) is modified as follows

$$\frac{Q}{C} + RI + L \frac{\mathrm{d}I}{\mathrm{d}t} = 0 \tag{2.64}$$

Following the substitution $Q = C.U$ and differentiation, we obtain

$$\frac{1}{C} I + R \frac{\mathrm{d}I}{\mathrm{d}t} + L \frac{\mathrm{d}^2 I}{\mathrm{d}t^2} = 0 \tag{2.65}$$

This is a second-order homogeneous differential equation which can be solved via the substitution

$$I = k\mathrm{e}^{\gamma t} \tag{2.66}$$

Having performed the necessary mathematical operations the following is obtained

$$L\gamma^2 + R\gamma + \frac{1}{C} = 0 \tag{2.67}$$

From this second-degree equation it can be calculated that

$$\gamma_{1,2} = -\frac{R}{2L} \pm \sqrt{\left(\frac{R}{2L}\right)^2 - \frac{1}{LC}} \tag{2.68}$$

The discharge circuit of the condenser can be looked upon as an oscillator between whose damping resistance, R, and critical resistance, R_c, there is the relation, $R \leqq R_c$, and also

$$R_c = 2\sqrt{\frac{L}{C}} \tag{2.69}$$

Taking this relationship into consideration together with Eq. (2.68), three cases can be distinguished for the solution of the differential equation (2.65) and thus for the discharge of the condenser of spark generators:

(1) *The discharge of the condenser is periodically damped* if the resistance of the discharge circuit is smaller than the critical resistance: in this case $R < R_c$ and γ_1 and γ_2 are (conjugate) complex numbers.
(2) *The discharge is aperiodically damped* or overdamped when the ohmic resistance of the discharge circuit is larger than the critical resistance. Then $R > R_c$ and Eq. (2.68) has two real roots.
(3) *The discharge is critically damped* if the resistance of the discharge circuit is equal to the critical resistance: $R = R_c$. In this case $\gamma_1 = \gamma_2 = \gamma$, which is a negative real number.

References

1. FRAENKEL, A., *Theorie der Wechselströme*, Berlin (1930)
2. JOOS, G., *Lehrbuch der theoretischen Physik*, Leipzig (1945)
3. BARKHAUSEN, H., *Lehrbuch der Elektronröhren*, vol. I, Leipzig (1945)
4. SCHÖNFELD, H., *Die wissenschaftlichen Grundlagen der Elektrotechnik*, Leipzig (1951)

2.6.3 Periodically damped discharge of a condenser[1,2]

In spark excitation the resistance of the discharge circuit of the condenser is normally smaller than the critical value, that is $R < R_c$. Generally, therefore, the condenser discharge is periodically damped. This is illustrated by the fact that the condenser plates are recharged several times during the time lapse of the full discharge.

If the discharge circuit did not contain ohmic resistance then it would be an undamped oscillator whose angular frequency is

$$\omega_0 = \frac{2\pi}{T_0} = \sqrt{\frac{1}{LC}} \tag{2.70}$$

where the value of T_0 is the same as that obtained from the Thomson formula

$$T_0 = 2\pi\sqrt{LC} \tag{2.71}$$

In practice the condenser discharge circuit always has some ohmic resistance. In this general case the discharge of the condenser has a damped discharge whose angular frequency is

$$\omega = \frac{2\pi}{T} = \sqrt{\frac{1}{LC} - \left(\frac{R}{2L}\right)^2} \tag{2.72}$$

It follows from the above two equations that the larger the ohmic resistance of the discharge circuit, the lower will be the angular frequency of the damped vibrations compared with that of the undamped vibrations.

In the case of the periodically damped discharge of the condenser the *discharge current* varies according to the formula

$$I = -\frac{U_0}{\omega L} e^{-tR/2L} \sin \omega t \tag{2.73}$$

while the *voltage of the condenser* varies as a function of the discharge time, t, as follows

$$U = U_0 \frac{\omega_0}{\omega} e^{-tR/2L} \sin (\omega t + \varepsilon) \tag{2.74}$$

As can be seen from the previous equations, the amplitude of both the discharge current and the terminal voltage of the condenser drops exponentially (Fig. 2.68) such that the rate of diminution is determined by the value of the ratio $R/2L$. It becomes clear that the current lags behind the voltage by the phase angle ε. The value of the phase angle is given by

$$\tan \varepsilon = \frac{2L\omega}{R} = \sqrt{\frac{4L}{R^2C} - 1} = \sqrt{\left(\frac{R_c}{R}\right)^2 - 1} \tag{2.75}$$

The time necessary to reach the discharge current maximum is

$$t_{max} = \frac{\varepsilon}{\omega} \tag{2.76}$$

The above relationships show that in the case of the periodically damped discharge of the condenser the values of I, U, and ε depend on the resistance, R, of the discharge circuit. When this resistance is very small, that is, when $R \ll R_c$, then, according to Eq. (2.75) $\tan \varepsilon \approx \infty$ and thus $\varepsilon \approx \pi/2$. Under the same conditions, according to Eq. (2.72) $\omega \approx \sqrt{(1/LC)}$ and thus, according to Eq. (2.76) $t_{max} \approx \pi\sqrt{(LC/2)} = = T/4$. In summary, it can be stated that when the condenser discharge is only slightly damped the phase-lag of the current with respect to the voltage is $\varepsilon = \pi/2$, while the current reaches its maximum value after the lapse of the first quarter-period.

It is possible that the resistance of the discharge circuit of the condenser is approximately the same as the critical value, that is, *the discharge is strongly damped.* In this case, however, according to Eq. (2.75) $\varepsilon \approx 0$, that is, $\varepsilon \approx \tan \varepsilon$. Thus by taking into

account Eq. (2.76) it can be established that for reaching the current maximum $t_{max} \approx 2L/R$ time is necessary. In such conditions the form of the current–time function is determined first of all by the exponential term in Eq. (2.73) that overrides the sine variation.

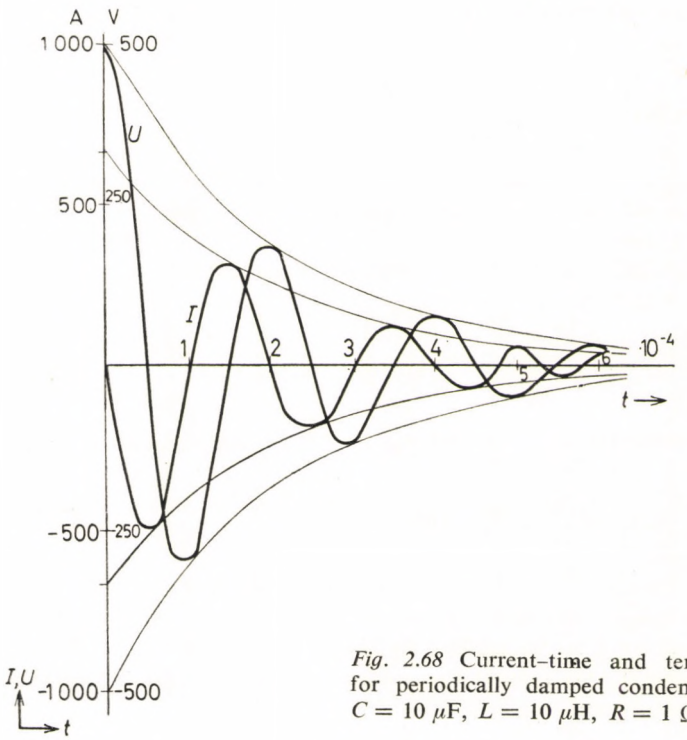


Fig. 2.68 Current–time and terminal voltage–time curves for periodically damped condenser discharge ($U_0 = 1$ kV, $C = 10\ \mu$F, $L = 10\ \mu$H, $R = 1\ \Omega$)

From Eqs (2.72), (2.74), and (2.75) it can be calculated that when the condenser discharge is damped to an extent that cannot be neglected, the maximum current value is given by

$$I_{max} = -U_0 e^{-R\varepsilon/2L\omega} \sqrt{\frac{C}{L}} \tag{2.77}$$

In spark excitation it frequently happens that the resistance of the discharge circuit is negligible as compared to the value of the critical level, that is, $R \approx 0$. In this case it follows from the relationship in Eq. (2.77) that for a slight damping of the discharge the current maximum is given by the formula

$$I_{max} \approx -U_0 \sqrt{\frac{C}{L}} \tag{2.78}$$

From the previous two equations it becomes clear that the maximum value of the discharge current increases with increasing terminal voltage or capacitance of the condenser, while it decreases with increasing self-inductance in the discharge circuitry. With the change of the current across the spark gap, however, there is a change in the temperature of the spark plasma and its degree of ionisation and thus in the

character of the spectrum. On increasing the terminal voltage and capacitance of the condenser, therefore, the spectrum becomes more ionic in character, while an increase in the self-induction of the discharge circuitry shifts it in character towards the atomic spectra.

It may happen that the resistance of the discharge circuit becomes practically equal to the critical value, that is, $R \approx 2\sqrt{L/C}$. In this case, as we have shown previously, $\varepsilon \approx \tan \varepsilon$. Then, according to Eqs (2.75) and (2.77)

$$I_{\max} \approx -\frac{U_0}{e}\sqrt{\frac{C}{L}} \tag{2.79}$$

Thus, as shown previously, when the resistance of the discharge circuit is increased from zero to the level of the critical damping, the maximum discharge current drops to 1/2·71 of its original value.

In the previous simplified calculations it has been assumed that the time-dependent resistance of the spark gap can be neglected. This simplification allowed the calculation of $t_{\max}$ and $I_{\max}$ in the initial period of the discharge. The resistance of the spark gap is very small when the current is high during the discharge. However, no practically useful results can be obtained using this method to find the relationship between the parameters of the circuit and the discharge duration t_s. This difficulty can be eliminated by the following calculations.[1]

It should be first taken into consideration that the following relationship exists between the energy, E_L, accumulated in the magnetic field of the induction of the oscillator and the current

$$E_L = \frac{1}{2}LI^2 \tag{2.80}$$

when

$$I = I_{\max} f(t) \sin \omega t \tag{2.81}$$

In the former relation, $f(t)$ is the desired damping function. Obviously the discharge of the condenser continues until $I = 0$ is realised, that is, when $f(t) = 0$. This means that the discharge keeps on functioning until the magnetic field becomes totally dissipated by the analytical gap and the ohmic resistance. The energy accumulated in the magnetic field of the induction of the oscillator reaches its first extreme value at the instant $t_{\max} \approx \pi/2\omega$, or approximately at the end of the first quarter-period. The smaller second extreme value is attained at the instant $t_{\max} + T/2$, or approximately at the end of three-quarters of a period. The energy difference, ΔE_L, between these two energy values is dissipated by the ohmic resistance of the spark gap and the circuitry. Therefore

$$-\Delta E_L = \Delta E_s + \Delta E_R \tag{2.82}$$

where ΔE_s and ΔE_R denote the energy dissipation during the interval $T/2$ by the spark gap and the ohmic resistance, respectively. In order to calculate the values of the terms in the above equation, the following expression can be developed from Eqs (2.80) and (2.81)

$$E_L = \frac{1}{2}LI_{\max}^2 f^2(t)$$

provided the value of $\sin \omega t$ is taken to be unity, and $\omega t_{max} \approx \pi/2$. By differentiating the previous equation

$$\frac{dE_L}{dt} = LI^2_{max} f(t) f'(t)$$

Substituting $dE_L = \Delta E_L$, and $dt = T/2$ we obtain

$$\Delta E_L = LI^2_{max} f(t) f'(t) \frac{T}{2} \tag{2.83}$$

The value of ΔE_s is equal to the time integral of the work of the current, I, in the spark gap, that is

$$\Delta E_s = \int_{\pi/2\omega}^{3\pi/2\omega} U_{min} I_{max} f(t) . \sin \omega t . dt$$

where U_{min} is the burning voltage of the spark. In the integration it should be taken into account that, within the proper limits, the signs of $\sin \omega t$ and U_{min} change.

The result of the integration is

$$\Delta E_s = U_{min} I_{max} f(t) \frac{T}{\pi} \tag{2.84}$$

The value of ΔE_R can be obtained by using the effective current

$$\Delta E_R = R \left[\frac{I_{max}}{\sqrt{2}} f(t) \right]^2 \frac{T}{2}$$

therefore

$$\Delta E_R = RI^2_{max} f^2(t) \frac{T}{4} \tag{2.85}$$

By substituting the values of ΔE_L and ΔE_R according to Eqs (2.83) and (2.85) the following expression is obtained from Eq. (2.82)

$$-LI_{max} f(t) f'(t) \frac{T}{2} = U_{min} I_{max} f(t) \frac{T}{\pi} + RI^2_{max} f^2(t) \frac{T}{4}$$

After some simplification and transformation the above can be written in the form

$$-\frac{2L}{R} f'(t) = \frac{4U_{min}}{\pi R I_{max}} + f(t) \tag{2.86}$$

The above differential equation is solved as follows

$$f(t) = -\frac{4U_{min}}{\pi R I_{max}} \pm k e^{-tR/2L} \tag{2.87}$$

With regard to the initial circumstances where $t = \pi/2\omega$, that is, when $\sin \omega t = 1$ and $I = I_{max}$, it is necessary for the function $f(t)$ to have the value

$$f(0) = 1 \tag{2.88}$$

From Eqs (2.87) and (2.88) we obtain

$$\pm k = \frac{4U_{\min}}{\pi R I_{\max}} + 1$$

By substituting the above expression into Eq. (2.87) the following relationship is obtained

$$f(t) = -\frac{4U_{\min}}{\pi R I_{\max}} + \left[\frac{4U_{\min}}{\pi R I_{\max}} + 1\right] e^{-tR/2L} \tag{2.89}$$

Considering the requirement expressed by Eq. (2.87) we can write that

$$\left[\frac{4U_{\min}}{\pi R I_{\max}} + 1\right] e^{-t_s R/2L} = \frac{4U_{\min}}{\pi R I_{\max}}$$

Thus the time interval, t_s, within which the discharge is damped can be calculated from the following formula

$$t_s = \frac{2L}{R} \ln\left[\frac{\pi R I_{\max}}{4U_{\min}} + 1\right] \tag{2.90}$$

From the above relationship it becomes evident that the larger the self-inductance of the discharge circuit and the higher the maximum current of the discharge current, the longer will be the time required for the total damping of the sparking. The damping period number can be calculated by the following formula

$$p = \frac{t_s}{T} \tag{2.91}$$

using the parameter t_s and the oscillation time T according to Eq. (2.72).

The time variation of the condenser current can be calculated from the following equation, by using Eqs (2.81) and (2.89)

$$I = \left\{-\frac{4U_{\min}}{\pi R} + \left[\frac{4U_{\min}}{\pi R} + I_{\max}\right] e^{-tR/2L}\right\} \sin \omega t \tag{2.92}$$

If the ohmic resistance of the oscillator circuit is neglected [Eq. (2.86) is multiplied by R, and its last term is dropped], then the following differential equation is obtained[2]

$$-2If'(t) = \frac{4U_{\min}}{\pi I_{\max}} \tag{2.93}$$

By using the above formula and Eq. (2.81) it can be calculated that when the ohmic resistance of the discharge circuit is neglected, or we have pure spark damping, the discharge current of the condenser is given by the formula

$$I = I_{\max}\left[1 - \frac{2U_{\min}}{\pi I_{\max}} t\right] \sin \omega t \tag{2.94}$$

It can be established on the basis of the above formula that when pure spark damping occurs the amplitude of the current drops linearly with time; that is, when the ohmic resistance of the discharge circuit of the condenser is negligibly small (Fig. 2.69).

On the other hand, according to Eq. (2.73), current variation becomes exponential when the damping is purely ohmic or when it is both ohmic and spark damping; that is, when the ohmic resistance of the circuit cannot be neglected (see Fig. 2.69). Taking Eq. (2.88) into consideration, it follows from Eq. (2.93) that

$$f(t) = -\frac{2U_{\min}}{\pi L I_{\max}} t + 1 \tag{2.95}$$

The condenser discharge lasts until $I = 0$, or according to Eq. (2.94) the duration of the spark discharge in the case of pure spark damping is

$$t_s = \frac{\pi L I_{\max}}{2U_{\min}} \tag{2.96}$$

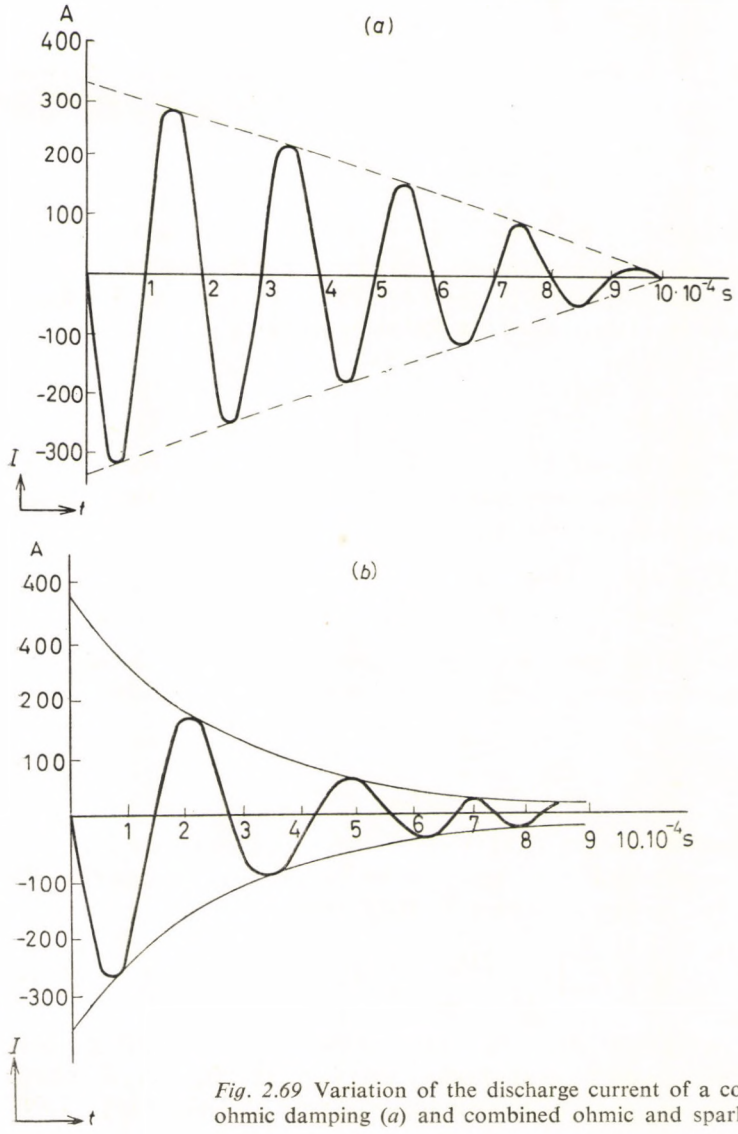


Fig. 2.69 Variation of the discharge current of a condenser with time for ohmic damping (*a*) and combined ohmic and spark damping (*b*)

Using the previous formula and Eqs (2.78) and (2.91) plus the Thomson formula, we obtain the following expression for the period number in the case of pure spark damping

$$p = \frac{U_0}{4U_{\min}} \tag{2.97}$$

References

1. TÖRÖK, T., *Proc. V. Congress on Material Testing III. Section, Spectrochem. Analysis*, Budapest, 3 (1970)
2. KAISER, H., and WALLRAFF, A., *Ann. Phys.*, **34**, 297 (1939)

2.6.4 Over-damped and critically damped discharge of a condenser

The discharge of a condenser is *over-damped* when the resistance, R, of the discharge circuit is greater than the critical damping value, R_c. In this case the voltage drops in a continuous fashion and not periodically, until the condenser is totally discharged. In such circumstances the variation of the discharge current of the condenser with the discharge time, t, can be described by the following equation

$$I = -\frac{U_0}{\sqrt{R^2 - 4L/C}} (e^{\gamma_1 t} - e^{\gamma_2 t}) \tag{2.98}$$

In the exponent of the above equation the following inequality holds for the γ values according to Eq. (2.68): $\gamma_2 < \gamma_1 < 0$. The discharge current of the condenser is, therefore, proportional to the difference of two exponential expressions, one of which drops faster than the other. This has the consequence that the current value is zero at the beginning of the condenser discharge when $t = 0$, and then rises very rapidly to the maximum value, $I_{\max}$. The time necessary for attaining this state is given by

$$t_{\max} = \frac{1}{\gamma_1 - \gamma_2} \ln \frac{\gamma_2}{\gamma_1} \tag{2.99}$$

Having reached the maximum value, the discharge current at first drops rapidly, then later slows down, such that the discharge of the condenser occurs continuously without oscillations.

As in the case of periodical damping, it becomes clear from Eq. (2.98) that the current of the discharge current increases with increasing initial terminal voltage and capacitance of the condenser for an over-damped discharge and that it also drops with increasing ohmic resistance and self-induction of the discharge circuit.

In the over-damped discharge of the condenser the voltage of the condenser varies with the discharge time, t, according to the formula

$$U = \frac{U_0}{2} [(k + 1)e^{\gamma_1 t} - (k - 1)e^{\gamma_2 t}] \tag{2.100}$$

provided that

$$k = \frac{R}{\sqrt{R^2 - 4L/C}} = \frac{R}{\sqrt{R^2 - R_c^2}} \tag{2.100a}$$

The terminal voltage of the condenser drops continuously to a voltage minimum, U_{min} (as is shown by the curve in Fig. 2.70), at which the voltage is just enough to let the discharge current pass across the spark gap.

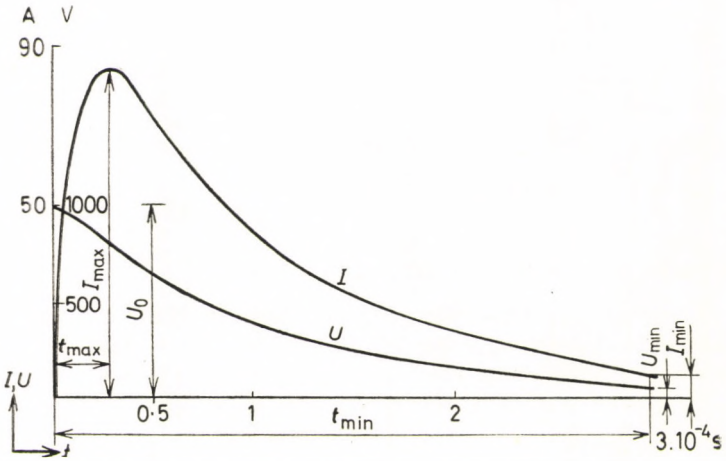


Fig. 2.70 Current-time and voltage-time curves of an over-damped condenser discharge ($U_0 = 1$ kV, $C = 10$ μF, $L = 10$ μH, $R = 10$ Ω)

The *discharge time* of the condenser could be calculated by a series expansion technique from Eq. (2.100). If it is taken into account that $\gamma_2 < \gamma_1$, it becomes clear that the second exponential term in Eq. (2.100) drops much faster with time than the first term does. The longer the discharge time, the smaller will be the error committed by neglecting the value of the second exponential term as compared to the first. By this simplification it can be easily calculated from the former two equations that the discharge time of the condenser or the duration of the spark is

$$t_s = \frac{1}{\gamma_1} \ln 2 \frac{U_{min}}{U_0} \frac{\sqrt{R^2 + R_c^2}}{R + \sqrt{R^2 - R_c^2}} \tag{2.101}$$

The above formula shows that in the case of over-damped condenser discharge, the longer the discharge time, the smaller is the initial terminal voltage of the condenser, and the smaller is the difference between the ohmic resistance of the discharge circuit and the critical resistance.

In the exceptional case when the ohmic resistance of the discharge circuit of the condenser is equal to the critical value, the condenser is aperiodically critically damped. Under such conditions, however, according to Eqs (2.68) and (2.69)

$$\gamma_1 = \gamma_2 = \gamma = -\frac{R_c}{2L} \tag{2.102}$$

In this case the variation of the discharge current with the discharge time of the condenser is given by the following formula

$$I = -\frac{2U_0}{R_c\sqrt{LC}} t e^{-t/\sqrt{LC}} \tag{2.103}$$

The current value at the beginning of the discharge quickly increases from zero to the level of I_{max} within the time

$$t_{max} = \sqrt{LC} \tag{2.104}$$

and then it starts dropping, quickly at first and later more slowly, as shown by Fig. 2.71. The value of the maximum current is given by

$$I_{max} = -\frac{2U_0}{R_c} e^{-1} = -0{\cdot}736 \frac{U_0}{R_c} \tag{2.105}$$

According to the above equation the maximum current value depends on the initial terminal voltage of the condenser and the critical resistance of the circuit. The length of time during which the maximum current is attained is influenced by the capacitance of the condenser and by the self-induction of the discharge circuit. When it is considered that the critical resistance is given by $R_c = 2\sqrt{L/C}$, then it becomes clear that in this case also the greater the maximum current, the larger is the capacitance of the condenser, or the smaller is the self-induction of the discharge circuit.

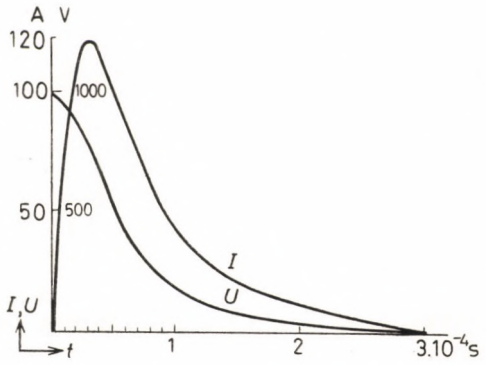


Fig. 2.71 Aperiodic discharge current (critical-damped)–time and voltage–time curves of a condenser ($U_0 = 1$ kV, $C = 10$ μF, $L = 10$ μH, $R = Rf = 2$ Ω)

It can also be calculated that in the case of aperiodic critical damping of the condenser discharge the terminal voltage of the condenser is given by the equation

$$U = U_0 \left(1 + \frac{t}{\sqrt{LC}}\right) e^{-t/\sqrt{LC}} \tag{2.106}$$

and

$$U = U_0 \left(1 + \frac{2t}{R_c C}\right) e^{-2t/R_c C} \tag{2.107}$$

The time for the full discharge of the condenser[1] is given via the substitution of

$$\frac{t_s}{\sqrt{CL}} = x \tag{2.108}$$

in Eq. (2.106). Thus

$$\frac{U_{min}}{U_0} e^x = 1 + x \tag{2.109}$$

This equation, however, cannot be solved either in closed form or by a series expansion technique, although it is possible to solve it graphically, given the following transformation

$$\log \frac{U_{\min}}{U_0} + 0{\cdot}434x = \log (1 + x) \tag{2.110}$$

The graphical solution of this equation is illustrated in Fig. 2.72, such that the quantities on either side of the equation are plotted as functions of the variable x. The x-value corresponds to the intersection of the curve and the straight line. When the value of x is known the quantity t_s can be simply calculated by using Eq. (2.108).

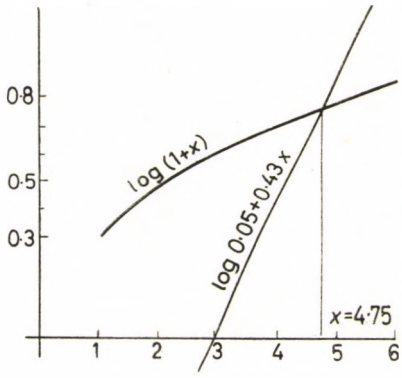


Fig. 2.72 Graphical determination of the value of t_s ($U_{min}/U_0 = 50/1000$, $C = 10^{-5}$ F, $L = 10^{-5}$ H, $t_c = x\sqrt{1L} = 4{\cdot}75.10^{-5}$ s)

Reference

1. TÖRÖK, T., *Proc. V. Congress on Material Testing III. Section Spectrochem. Analysis*, Budapest, 3 (1970)

2.6.5 The influence of the parameters of the discharge circuit

From the numerical data of the discharge circuit of the condenser the discharge parameters can be calculated[1] using the equations derived in the previous sub-sections. The values obtained in this way are summarised in Tables 2.8–2.10, utilising the values of U_0, C, and L that are used in practice. In the selection of these values 1 μH stands for the approximate remanent induction of the circuit, while 0·1 Ω is the value of the remanent resistance. In the calculation of the critical resistance the relationship: $R_c = 2\sqrt{L/C}$ was used and for the value of U_{min} the normally occurring average voltage (50 V) was taken into account. The values of the quantity U_0 correspond to the condenser terminal voltage of the high-, medium-, and mains-voltage generators. In the case of the three different values of the terminal voltage, the energy accumulated within the condenser is 0·05 and 0·5 Ws for high-voltage spark generators, while for medium- and mains-voltage excitation this energy value is 0·5 and 5 Ws, respectively.

From the data given in Tables 2.8–2.10 the following conclusions can be drawn, all of which strongly influence the character of the spectrum:

Table 2.8

The electrical parameters of the discharge circuit at charging voltage: $U_0 = 10$ kV ($U_{min} = 5.10^{-2}$ kV)[1]

C (nF)	L (μH)	R (Ω)	R_c (Ω)	E (Ws)	T (s)	t_{max} (s)	I_{max} (A)	t_s (s)	p
1	1	0·1	63	0·05	2.10^{-7}	5.10^{-8}	317	8.10^{-6}	40·5
		1			2.10^{-7}	5.10^{-8}	314	$3{\cdot}5.10^{-6}$	17·5
		10			2.10^{-7}	$4{\cdot}5.10^{-8}$	256	$7{\cdot}5.10^{-7}$	4
		100				$2{\cdot}5.10^{-8}$	86	5.10^{-7}	
1	100	0·1	630	0·05	2.10^{-6}	5.10^{-7}	32		49
		1			2.10^{-6}	5.10^{-7}	32	8.10^{-5}	41
		10			2.10^{-6}	5.10^{-7}	31	4.10^{-5}	18
		100			2.10^{-6}	$4{\cdot}5.10^{-7}$	26	$7{\cdot}5.10^{-6}$	4
1	1000	0·1	2000	0·05	6.10^{-6}	$1{\cdot}5.10^{-6}$	10	3.10^{-4}	49·5
		1			6.10^{-6}	$1{\cdot}5.10^{-6}$	10	3.10^{-4}	46·5
		10			6.10^{-6}	$1{\cdot}5.10^{-6}$	10	2.10^{-4}	30
		100			2.10^{-6}	$1{\cdot}5.10^{-6}$	9	$5{\cdot}5.10^{-5}$	27·5
10	1	0·1	20	0·5	6.10^{-7}	$1{\cdot}5.10^{-7}$	990	2.10^{-5}	30
		1			6.10^{-7}	$1{\cdot}5.10^{-7}$	930	$5{\cdot}5.10^{-6}$	9
		10			2.10^{-7}	1.10^{-7}	550	2.10^{-7}	1
		100				$4{\cdot}5.10^{-8}$	94	$5{\cdot}5.10^{-6}$	
10	100	0·1	200	0·5	6.10^{-6}	$1{\cdot}5.10^{-6}$	100	3.10^{-4}	44·5
		1			6.10^{-6}	$1{\cdot}5.10^{-6}$	99	2.10^{-4}	28·5
		10			2.10^{-6}	$1{\cdot}5.10^{-6}$	93	$1{\cdot}5.10^{-6}$	8
		100			2.10^{-6}	1.10^{-6}	55	3.10^{-6}	1·5
10	1000	0·1	630	0·5	2.10^{-5}	5.10^{-6}	32	9.10^{-4}	50
		1			2.10^{-5}	5.10^{-6}	32	8.10^{-4}	40·5
		10			2.10^{-5}	5.10^{-6}	31	4.10^{-4}	18
		100			2.10^{-5}	$4{\cdot}5.10^{-6}$	26	$7{\cdot}5.10^{-5}$	4

(1) The maximum current of the discharge of condensers charged to the same energy level is largest if the circuits have the same inductive resistance and the same ohmic resistance when the initial terminal voltage of the condenser is about 1 kV, that is, when it is about the same as the condenser voltage of medium voltage spark generators.

(2) Upon increasing the induction of the circuit the maximum current value of a condenser discharge is generally lowered. This effect is reduced, however, when the difference between the ohmic resistance of the circuit and the critical resistance is small and when the initial terminal voltage of the condenser is low. For example, in the case of a terminal voltage of 0·1 kV for low-voltage spark excitation, the inductive resistance of the discharge circuit does not change the current value of the discharge current, provided the resistance of the circuit is larger than 10 Ω.

(3) The ohmic resistance of the discharge circuit influences the maximum current value of a condenser discharge only when the value of the resistance approaches or exceeds the critical resistance of the circuit. Any larger ohmic resistance leads to rapid decrease of the current. However, the influence of the ohmic resistance is small when the induction of the circuitry is large. (That is due to the fact that the value of the critical resistance increases with increasing induction of the circuit.) The above

Table 2.9

The electrical parameters of the discharge circuit at charging voltage: $U_0 = 1$ kV ($U_{min} = 5.10^{-2}$ kV)[1]

C (μF)	L (μH)	R (Ω)	R_c (Ω)	E (Ws)	T (s)	t_{max} (s)	I_{max} (A)	t_s (s)	p
1	1	0·1	2	0·5	6.10^{-6}	$2{\cdot}0.10^{-5}$	736	2.10^{-3}	3
		1			7.10^{-6}	1.10^{-6}	550	$4{\cdot}5.10^{-6}$	0·5
		10				5.10^{-7}	95	3.10^{-5}	—
		100				7.10^{-8}	10	3.10^{-5}	
1	100	0·1	20	0·5	6.10^{-5}	2.10^{-5}	100	3.10^{-4}	4·5
		1			6.10^{-5}	$1{\cdot}5.10^{-5}$	95	2.10^{-4}	3
		10			2.10^{-3}	1.10^{-5}	55	$4{\cdot}5.10^{-5}$	—
		100			—	5.10^{-6}	10	3.10^{-4}	
1	1000	0·1	63	0·5	2.10^{-4}	5.10^{-5}	30	10^{-3}	5
		1			2.10^{-4}	5.10^{-5}	30	8.10^{-4}	4
		10			6.10^{-4}	$4{\cdot}5.10^{-5}$	25	3.10^{-4}	0·5
		100			—	3.10^{-5}	10	3.10^{-4}	—
10	1	0·1	0·6	5	2.10^{-6}	$4{\cdot}5.10^{-6}$	2560	6.10^{-6}	3
		1			—	3.10^{-6}	865	3.10^{-5}	—
		10			—	7.10^{-7}	100	3.10^{-4}	—
		100			—	9.10^{-8}	10	3.10^{-4}	—
10	100	0·1	6·3	5	2.10^{-4}	5.10^{-5}	325	8.10^{-4}	4
		1			2.10^{-4}	$4{\cdot}5.10^{-5}$	255	3.10^{-4}	1·5
		10			—	3.10^{-5}	85	3.10^{-4}	—
		100			—	7.10^{-6}	10	3.10^{-3}	—
10	1000	0·1	20	5	6.10^{-4}	$1{\cdot}5.10^{-4}$	100	3.10^{-3}	4·5
		1			6.10^{-4}	$1{\cdot}5.10^{-4}$	95	2.10^{-3}	3
		10			7.10^{-4}	1.10^{-4}	55	$4{\cdot}5.10^{-3}$	0·5
		100			—	5.10^{-5}	10	3.10^{-3}	—

phenomenon is emphasised in high-voltage spark generation when the terminal voltage of the condenser is around 10 kV. On the other hand, the lower the terminal voltage of the condenser, the smaller is the influence of the ohmic resistance of the discharge circuit.

(4) It is also evident from the above tables that the maximum current of the discharge is less influenced by the amount of the energy accumulated within the condenser (that is, the capacitance of the condenser), the greater the ohmic resistance of the discharge circuit relative to the value of the critical resistance. This is generally the case in low-voltage spark excitation when the initial terminal voltage of the condenser is in the order of 0·1 kV.

(5) The time interval, t_{max}, during which the discharge current reaches its maximum value, increases with the induction of the discharge circuit, the capacitance of the condenser and its initial voltage. Any increase in the ohmic resistance of the circuit will lead to a shortening of this time interval, although the variation is generally small.

(6) It is not only true that the oscillation time, T, of the condenser discharge can vary within wide limits (10^{-8}–10^{-2} s), but damped oscillations are also possible during which the recharging of the condenser plates to the opposite charge does not occur.

Table 2.10

The electrical parameters of the discharge circuit at charging voltage $U_0 = 0{\cdot}1$ kV ($U_{min} = 5.10^{-2}$ kV)[1]

C (μF)	L (μH)	R (Ω)	R_c (Ω)	E (Ws)	T (s)	t_{max} (s)	I_{max} (A)	t_s (s)	p
100	1	0·1	0·2	0·5	7.10^{-5}	1.10^{-5}	550	1.10^{-5}	0·2
		1			—	5.10^{-6}	94	7.10^{-5}	—
		10			—	5.10^{-7}	9	7.10^{-6}	—
		100			—	7.10^{-8}	1	7.10^{-6}	—
100	100	0·1	2	0·5	6.10^{-4}	$1{\cdot}5.10^{-4}$	93	3.10^{-4}	0·4
		1			7.10^{-4}	1.10^{-4}	55	1.10^{-4}	0·2
		10			—	5.10^{-5}	9	7.10^{-4}	—
		100			—	7.10^{-6}	1	7.10^{-4}	—
100	1000	0·1	6·3	0·5	2.10^{-3}	5.10^{-4}	31	1.10^{-3}	0·5
		1			2.10^{-3}	$4{\cdot}5.10^{-4}$	26	$6{\cdot}5.10^{-4}$	0·3
		10			—	3.10^{-4}	9	$7{\cdot}5.10^{-4}$	—
		100			—	7.10^{-5}	1	7.10^{-3}	—
1000	1	0·1	0·06	5	—	3.10^{-5}	2580	2.10^{-5}	—
		1			—	7.10^{-6}	99	7.10^{-4}	—
		10			—	9.10^{-7}	10	7.10^{-4}	—
		100			—	1.10^{-7}	1	7.10^{-4}	—
1000	100	0·1	0·06	5	2.10^{-3}	$4{\cdot}5.10^{-4}$	316	8.10^{-4}	0·4
		1			—	3.10^{-4}	85	$1{\cdot}5.10^{-3}$	—
		10			—	7.10^{-5}	10	7.10^{-3}	—
		100			—	9.10^{-6}	1	7.10^{-3}	—
1000	1000	0·1	2	5	6.10^{-2}	$1{\cdot}5.10^{-3}$	93	2.10^{-2}	0·3
		1			7.10^{-3}	1.10^{-3}	63	4.10^{-3}	0·2
		10			—	5.10^{-4}	9	7.10^{-4}	—
		100			—	7.10^{-5}	1	7.10^{-3}	—

The oscillation time increases with the increasing induction of the discharge circuit and the condenser capacitance. On the other hand, no significant change occurs when *only* the ohmic resistance of the discharge circuit is changed. Reduction of the initial voltage of the condenser by several orders of magnitude, however, leads to a sizeable increase in the oscillation time.

(7) The damping time of the condenser discharge, that is, the time for the complete discharge of the condenser, is generally increased with an increase in the induction of the discharge circuit. It also increases when the energy accumulated within the condenser or the condenser capacitance increases. The effect of the ohmic resistance of the discharge circuit is not so simple. It is generally established that the damping time only increases as the ohmic resistance decreases, when the value of the latter is smaller than the critical resistance.

(8) The period number for the complete discharge of a condenser is increased not only by increasing the initial voltage of the condenser, but also by increasing the induction of the discharge circuit. The period number is reduced, however, upon decreasing the capacitance of the condenser or increasing the ohmic resistance of the discharge circuit. The condenser discharge becomes aperiodically damped when the ohmic resistance approaches or surpasses the value of the critical resistance. The

lower the initial terminal voltage of the condenser the earlier this phenomenon occurs. Aperiodically damped discharges, therefore, generally arise in the case of spark generators operating on mains voltage. In extreme cases it could also happen when the initial terminal voltage of the condenser is about 10 kV.

In summary, when spark excitation is used, the discharge current of the condenser can be varied within very wide limits (between 1 A and 2500 A) by an appropriate choice of the parameters of the discharge circuit. The time interval for the complete discharge of the condenser can also be varied between very wide limits (between 10^{-7} and 10^{-2} s). Consequently the ratio of the spark burning time and the intermission between burning periods can also be changed within wide limits, namely between $1 : 10^5$ and $1 : 1$.

Thus, it is possible in high-voltage spark excitation to produce an excitation in which the current passing through the spark gap can reach several hundred amperes and the discharge time of the condenser is in the order of only 10^{-8} s. The ratio of the burning and cooling times is such that the electrodes do not warm up excessively in spite of the high current value. Accordingly, only small portions of the electrodes evaporate extremely rapidly, such that fractional distillation does not occur and the vapour composition in the spark gap is very close to that of the sample. The concentration of the components in the vapour space, however, is still not very high. Due to the high current the temperature of the plasma becomes so elevated that the atoms of the elements present are mostly in the ionised state. Accordingly, the intensity of the ion lines is comparatively high when compared with the intensity of the atomic lines. The spectrum arising in this way is, therefore, definitely ionic in character. From the data given in the former tables, it is evident that in medium-voltage excitation condenser discharges can be produced which are totally damped and whose lifetime is about 10^{-3} s. Under such excitation conditions, processes taking place in the electrode gap are more or less the same as those occurring in the case of a.c. arc generation. The spectrum obtained will, therefore, have an atomic character. The conclusion reached is that spark excitation produces not only spectra of definite ionic character, but also spectra of atomic character, provided the parameters influencing the discharge of the condenser are appropriately chosen.

Reference

1. Török, T., *Proc. V. Congress on Material Testing, III. Section Spectrochem. Analysis*, Budapest, 3 (1970)

2.7 Electric phenomena occurring in the charging circuit of a condenser

2.7.1 Preliminary remarks

From the points discussed in Section 2.5.3 it is clearly advantageous from the analytical point of view if spark excitation can be produced by a series of condenser discharges within the energy range 0·05 – 5 Ws. The condenser should be recharged by the same amount of energy before each subsequent discharge during the whole excitation procedure. This is the task of the *charging circuit* of spark generators.

Among the processes taking place in the charging circuit of spark generators only the time variation of the condenser terminal voltage is important in spectrochemical analysis. For a given capacitance of the condenser the excitation energy is determined only by the value of the terminal voltage at the instance of the discharge. In this respect distinction should be made as to whether the condenser is charged by the mains current via a transformer or in a direct way. The former method, frequently applied in practice, is used with high-voltage or medium-voltage spark generators, while the latter is used in low-voltage spark generators.

2.7.2 Condenser charging via a transformer[1,2]

In this system the condenser, C_1, of the spark generator is charged via the critical resistance, R_d, and the transformer, *Tr*, by the a.c. mains current (Fig. 2.73*a*). The processes taking place in the charging circuit can be studied by considering a typical circuit (Fig. 2.73*b*). The value of R in the latter is equal to that of the critical resistance,

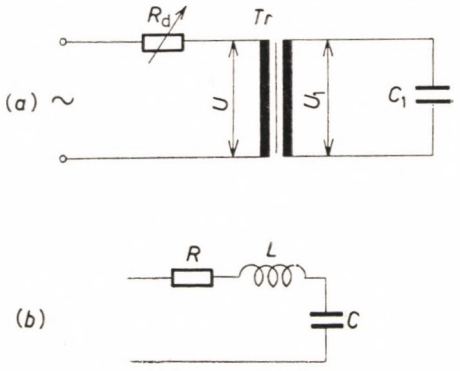


Fig. 2.73 Circuit diagram of the high-voltage and medium-voltage spark generator charging circuit (*a*) and the substitution circuit of the latter (*b*)

R_d, in the spark generator, while the value of C in the simple circuit is the same as the value of the capacitance, C_1, as recalculated for the primary winding of the transformer. The recalculation is based upon the equivalence of the energy accumulated within these two capacitors, that is, $1/2\,CU^2 = 1/2\,C_1U_1^2$. Thus, it can be written that

$$C = C_1\left(\frac{U_1}{U}\right)^2 \tag{2.111}$$

where U and U_1 stand for the primary and the secondary voltages of the transformer, respectively, while the coil, L, in the simple circuit substitutes for the transformer inductance. This induction arises from the lines of force of the magnetic field, established by the current passing through the primary winding of the transformer, being partially closed within air instead of the secondary winding of the transformer. These lines of force, therefore, do not induce any current in the secondary winding. The role of the stray magnetic field is the same as an induction coil connected in series that has an appropriate induction factor.

The electrical processes taking place in the above circuit can be studied if it is taken into account that the voltage drop brought about by a current of intensity I [as expressed in Eq. (2.64)] is equal to the mains voltage. When P_{max} is the peak voltage of the mains current, ω_n is its angular frequency and ψ is an appropriately

selected phase angle, then the value of the voltage, P, at an instant t, is given by the relation, $P = P_{max} \cdot \sin(\omega_n \cdot t + \psi)$. For the circuit processes the following differential equation, therefore, applies*

$$RI + \frac{Q}{C} + L\frac{dI}{dt} = -P_{max} \sin(\omega_n t + \psi) \tag{2.112}$$

By differentiating relative to time, the following formula is obtained

$$R\frac{dI}{dt} + \frac{1}{C}I + L\frac{d^2I}{dt^2} = -\omega_n P_{max} \cos(\omega_n t + \psi) \tag{2.113}$$

Note that $dQ/dt = I$, that is, the change of time relative to the charge is equal to the instantaneous value of the current.

The solution of the above differential equation is the sum of the two partial solutions that is

$$I = I_s + I_e \tag{2.114}$$

where I is the actual current value, I_s is the so-called stationary current and I_e is the so-called equalising current. Thus, it follows that

$$\frac{1}{C}\int I dt = \frac{1}{C}\int I_s dt + \frac{1}{C}\int I_e dt \tag{2.115}$$

Furthermore, as the time derivative of the charge, Q, of the condenser is equal to the current, I, and also the charge accumulated within the condenser is the product of the capacitance and the terminal voltage of the condenser, then

$$\frac{dQ}{dt} = I \quad \text{and} \quad Q = CU \tag{2.116}$$

Thus according to Eq. (2.115)

$$U = U_s + U_e \tag{2.117}$$

The actual voltage, U, of the condenser is, therefore, composed of the so-called *stationary voltage*, U_s, and the *equalising voltage*, U_e. Due to the current change brought about in the course of the condenser discharge the intensity of the magnetic field is also changed. Thus, a voltage is induced that opposes the resistance within the discharge circuit.

The *stationary* value of the charging current and the condenser voltage is determined by the following equations

$$I_s = -\frac{P_{max}}{R}\cos\phi \sin(\omega_n t + \psi + \phi) \tag{2.118}$$

and

$$U_s = -\frac{P_{max}}{RC\omega_n}\cos\phi \cos(\omega_n t + \psi + \phi) \tag{2.118a}$$

* The negative sign on the right-hand side of Eq. (2.112) is only significant, in that the end result will have a positive sign. The value of the circular frequency of the mains is: $\omega_n = 2\pi\nu$. Corresponding to European and American usage, ν is 50 or 60 and ω_n is 314·1 or 376·9, respectively.

provided the phase shift is given by

$$\tan \phi = \frac{1/C\omega_n - L\omega_n}{R} \tag{2.119}$$

The value of the *equalising current* is obtained by reducing Eq. (2.113) to zero

$$R \frac{dI}{dt} + \frac{1}{C} I + L \frac{d^2I}{dt^2} = 0 \tag{2.120}$$

The above equation, (2.120), is in complete agreement with Eq. (2.65), which was used as a starting point in the study of the short-circuited discharge circuit of a condenser. Concerning this equation, therefore, all the points expressed on page 137 are valid, with the exception that the initial conditions determining the values of the integration constants are different. Accordingly, the solution of the above equation is different, although in a similar fashion the value of the equalising current, I_e, varies aperiodically when $R > R_c$ and periodically when $R < R_c$.

According to the above the solution of Eq. (2.120) for an aperiodic equalising current is the following

$$I_{e(R>R_c)} = \gamma_1 k_1 e^{\gamma_1 t} + \gamma_2 k_2 e^{\gamma_2 t} \tag{2.121}$$

where the values of γ_1 and γ_2 are determined by Eq. (2.68), while k_1 and k_2 are constants whose values depend on the initial conditions. On the basis of the above equation, (2.121), and using Eqs (2.115) and (2.116), the equalising terminal voltage of the condenser can be calculated from

$$U_{e(R>R_c)} = \frac{1}{C}(k_1 e^{\gamma_1 t} + k_2 e^{\gamma_2 t}) \tag{2.122}$$

In the case of critical damping, that is, when the damping resistance of the charging circuit is equal to the critical resistance (that is, $R = R_c$), the value of the equalising current is given by the equation

$$I_{e(R=R_c)} = [k_2 - \omega_0(k_1 + tk_2)]e^{-t\omega_0} \tag{2.123}$$

where k_1 and k_2 are integration constants, while ω_0 is the angular frequency of the undamped oscillations, whose value is defined by: $\omega_0 = \sqrt{1/LC}$. Using the former equation and Eq. (2.115) it can be calculated that

$$U_{e(R=R_c)} = \frac{1}{C}(k_1 + tk_2)e^{-t\omega_0} \tag{2.124}$$

In the case of periodic damping ($R < R_c$), it can be calculated from Eq. (2.120) that the value of the equalising current is given by

$$I_{e(R<R_c)} = -A\omega_0 e^{-tR/2L} \sin(\omega t + \delta - \varepsilon) \tag{2.125}$$

while the voltage is given by

$$U_{e(R<R_c)} = \frac{A}{C} e^{-tR/2L} \sin(\omega t + \delta) \tag{2.126}$$

where A and δ are integration constants and the value of the phase angle, ε, is determined by the relationship $\tan \varepsilon = 2L\omega/R$.

It is advisable in the selection of the values of the integration constants in Eqs (2.121) – (2.126) to choose as the initial time, ($t = 0$), the instant right after the complete discharge of the condenser. This is the instant when the charging circuit is 'switched on'. When it is assumed, for the sake of simplicity, that the extremely short condenser discharge interval can be neglected, then the instant belonging to the maximum terminal voltage of the condenser can be taken as the initial time. (This assumption is allowable, since during the very short interval of the condenser discharge there is practically no change in the value of the stationary voltage, U_s.) With this assumption, however, the value of the cosine term in Eq. (2.118) becomes unity at the instant $t = 0$, thus

$$\psi = -\phi \tag{2.127}$$

In fixing the initial conditions it should be taken into account that at the initial instant, ($t = 0$), the current flowing in the charging circuit and the terminal voltage of the condenser are both zero. Thus, Eqs (2.114) and (2.117) must be modified in such a manner that the following relationships hold

$$I_{(0)} = I_{s(0)} + I_{e(0)} = 0 \tag{2.128}$$

$$U_{(0)} = U_{s(0)} + U_{e(0)} = 0 \tag{2.129}$$

where the subscript (0) denotes the value corresponding to the instant $t = 0$.

The stationary values in the above two equations are obtained by first finding the value of ψ from Eq. (2.127) and substituting it into Eqs (2.118) and (2. 118*a*) together with $t = 0$. The equalising values that refer to the three different types of damping are obtained by substituting $t = 0$ into Eqs (2.121) and (2.122). When these values are obtained they are then substituted into Eqs (2.128) and (2.129), such that two equations are obtained for each of the three damping types. From these equations k_1 and k_2 and the integration constants A and δ can be calculated. When the latter are then substituted into Eqs (2.122), (2.124), and (2.126), the U_e values belonging to the three different damping cases can be obtained. When the values of U_e and the U_s values determined from Eqs (2.118a) and (2.127) are substituted into Eq. (2.117), the corresponding U values can be calculated for the three damping types.

From the previously described calculations, the variation of the terminal voltage of the transformer during the charging is given by the following equations.

(1) In *aperiodic damping*, that is, when $R > R_c$

$$U_{(R>R_c)} = \frac{P_{max}}{RC\omega_n} \cos\phi \left[\cos\omega_n t + \frac{1}{\gamma_1 - \gamma_2}(\gamma_2 e^{-\gamma_1 t} - \gamma_1 e^{-\gamma_2 t})\right] \tag{2.130}$$

(2) In *critical damping*, that is, when $R = R_c$

$$U_{(R=R_c)} = \frac{P_{max}}{RC\omega_n} \cos\phi \left[\cos\omega_n t - (1 + \omega_0 t)e^{-\omega_0 t}\right] \tag{2.131}$$

(3) In *periodic damping*, that is, when $R < R_c$

$$U_{(R<R_c)} = \frac{P_{max}}{RC\omega_n} \cos\phi \left[\cos\omega_n t - e^{-tR/2L}\frac{1}{\cos\varepsilon}\cos(\omega t + \varepsilon)\right] \tag{2.132}$$

It is clearly demonstrated by the above three equations that the terminal voltage

of the condenser is determined by the resultant of the stationary voltage and the equalising voltage varying according to a cosine curve. At the instant $t = 0$, the signs of the equalising and stationary components are opposite, while their absolute value is the same.

A special consideration should be given to the exceptional case when the frequency of the charging circuit equals the mains frequency, that is, when the charging circuit is in resonance with the mains.[2] The resonance condition, according to Eq. (2.72), is the following:

$$\sqrt{\frac{1}{LC} - \left(\frac{R_r}{2L}\right)^2} = \omega_c$$

where R_r is the value of the resistance belonging to the given C and L values. It then follows from the above formula that

$$R_r = 2\sqrt{\frac{L}{C} - L^2\omega_n^2} \tag{2.133}$$

According to this relationship and Eq. (2.70) a further condition of resonance is that

$$\frac{R_r}{2L} = \sqrt{\omega_0^2 - \omega_n^2} \tag{2.134}$$

where ω_0 is the angular frequency of the charging circuit having no damping, that is, $R = 0$.

When the charging circuit is in resonance with the mains current the variation with time of the terminal voltage of the condenser can be calculated using Eq. (2.132) such that the following substitutions are made: $R = R_r$ and $\omega = \omega_n$. The terminal voltage is then given by

$$U = \frac{P_{max}}{R_r C\omega_n}\cos\phi\left[\cos\omega_n t - \frac{1}{\cos\varepsilon}e^{-tR_r/2L}\cos(\omega_n t + \varepsilon)\right] \tag{2.135}$$

A little damping can be achieved when $R_r/2L$ is not too great, that is, when according to Eq. (2.134) $\omega_0 \approx \omega_n$, and also $R_r \approx R_c$, such that from Eq. (2.75) $\varepsilon \approx 0$. The stationary and equalising components of the voltage, U, therefore, have a phase difference of 180° according to Eq. (2.132). The former equation must be modified so that the result is as follows

$$U_{(R_r = R_c)} \approx \frac{P_{max}}{R_r C\omega_n}(1 - e^{-tR_r/2L})\cos\phi\cos\omega_n t \tag{2.136}$$

The variation of time with the stationary voltage, U_s, the equalising voltage, U_e, and the resultant value, U, is depicted in Fig. 2.74*a*. As can be seen, the value of the resultant voltage, U, is so small that no new discharge is developed in the half-period. Thus, when the charging circuit is in resonance with the mains current the excitation becomes controlled. By the appropriate selection of the parameters it can be achieved that the resultant voltage, U, does not reach the value of the terminal voltage at which the condenser is discharged (Fig. 2.74*b*). By the proper choice of parameters an $R_r/2L$ value is obtained for which the value of the multiplier: $(1 - e^{-tR_r/2L})$ in Eq. (2.136) significantly differs from unity even in a time interval of

Table 2.11

Variation of the value $(1 - e^{-tR_r/2L})$ with the parameters C, L, R_r and t.
(C and L are the capacitance and induction, respectively,
as recalculated for the primary side of the transformer)

C (F)	L (H)	R_r (Ω)	R_n (Ω)	$R_r/2L$	t (s)	$1 - e^{-tR_r/2L}$
10^{-3}	3.10^{-2}	9·48	10·93	$1{\cdot}58.10^{2}$	1.10^{-2}	0·79
					2.10^{-2}	0·96
					3.10^{-2}	0·99
10^{-3}	5.10^{-2}	10·5	14·2	$1{\cdot}05.10^{2}$	1.10^{-2}	0·65
					2.10^{-2}	0·88
					3.10^{-2}	0·95
10^{-3}	10^{-1}	14·0	20·0	$0{\cdot}70.10^{2}$	1.10^{-2}	0·52
					2.10^{-2}	0·75
					3.10^{-2}	0·88

a half-period, that is, 10^{-2} s (Table 2.11). Thus, it is possible to obtain and control a type 1 + 0, in addition to a type 1 + 1 condenser discharge within a period. From the data given in Table 2.11 it can be seen that when the values of C, L, and R_r are appropriately chosen, the condenser discharge can be eliminated not only in a single period but also for several periods. This type of control can only be realised, however, if the value of the capacitance, C, as calculated for the primary winding of the transformer, is in the order of 10^{-3} F. When the value of the capacitance is smaller,

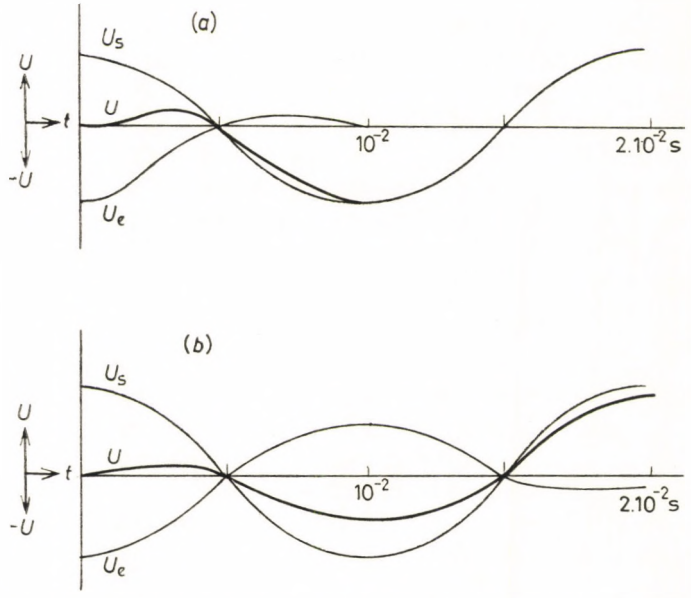


Fig. 2.74 Resonance variation of the condenser voltage for the charging circuit and the mains in discharges of the type (*a*) 1 + 1 in a period and (*b*) 1 + 0 in a period

for example, 10^{-4} or, 10^{-5} F, the value of ω_0 and also the value of $R_r/2L$ is so high that according to Eq. (2.70) U reaches the level of the stationary voltage within 10^{-2} s. This drawback is eliminated when a variable transformer with a relatively high ratio (for example, 40/15 000 V) is used. The primary and secondary portion of the iron core of the latter is bridged through the air gap by an iron core. Thus, the induction, L, of the transformer is increased or decreased when the air gap is increased or decreased, respectively.[3]

References

1. TÖRÖK, T., *Explle Tech. Phys.*, **7**, 49 (1959)
2. TÖRÖK, T., unpublished results
3. SCHEIBE, G., *Chemische Spektralanalyse* (BÖTTGER, W., *Physikalische Methoden der analytischen Chemie*, Leipzig, vol. I), 27 (1933)

2.7.3 Direct condenser charging via the mains current[1]

In mains-voltage spark generators the condenser is charged directly through the damping resistance by the mains current, such that a transformer is not required (Fig. 2.75). The charging circuit of the generator, therefore, has no inductive resist-

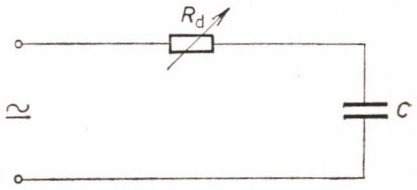


Fig. 2.75 Schematic circuit diagram of the charging of spark generators for the direct charging of the condenser

ance. By substituting $L = 0$ in differential equation (2.112) the following relationship holds for the charging circuit

$$R\frac{\mathrm{d}I}{\mathrm{d}t} + \frac{1}{C}I = -\omega_n P_{max}\cos(\omega_n t + \psi) \tag{2.137}$$

The actual intensity of the current, I, in the circuit is still composed of the stationary current I_s and the equalising current I_e such that the solution of the equation for the stationary current is as follows

$$I_s = -\frac{P_{max}}{R}\cos\phi\sin(\omega_n t + \psi + \phi) \tag{2.138}$$

The corresponding condenser terminal voltage is then obtained from Eqs (2.138) and (2.116)

$$U_s = \frac{P_{max}}{RC\omega_n}\cos\phi\sin(\omega_n t + \psi + \phi) \tag{2.139}$$

provided the value of the phase-shift is

$$\tan\phi = \frac{1}{RC\omega_n} \tag{2.140}$$

The value of the equalising current can be obtained by making Eq. (2.137) equal to zero as follows

$$R\frac{\mathrm{d}I_e}{\mathrm{d}t} + \frac{1}{C}I_e = 0 \tag{2.141}$$

As this equation is the same as Eq. (2.57*a*) the solution will be identical. According to Eq. (2.59) the voltage corresponding to the equalising current is given by the formula

$$U_e = U_{e(0)}e^{-t/RC} \tag{2.142}$$

where $U_{e(0)}$ is the integration constant referring to $t = 0$. Its value can be calculated by using Eqs (2.139) and (2.140), taking into account that $t = 0$ and the condition $\psi = -\phi$ (see Eq. 2.127). On substituting the obtained value of $U_{e(0)}$ into Eq. (2.142) and using the relationship $U = U_s + U_e$, we can obtain via Eq. (2.139) the following

$$U = \frac{P_{max}}{RC\omega_n}\cos\Phi(\cos\omega_n t - e^{-t/RC}) \tag{2.143}$$

From the above equation it is seen that the stationary component of the condenser terminal voltage varies according to a cosine curve during the charging process, while the equalising component drops exponentially. In Figure 2.76 it can also be

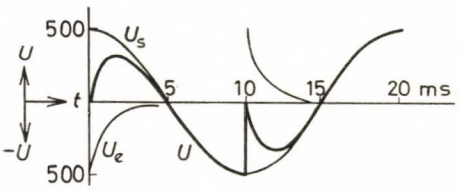


Fig. 2.76 Condenser voltage variation during mains voltage spark generation; U_s: variation of the stationary voltage, U_e: variation of the equalizing voltage

seen that *after the discharge of the condenser, the absolute value of the condenser voltage initially increases owing to the effect of the equalising current and then passes through a maximum until changing according to the stationary voltage.*

Reference

1. TÖRÖK, T., *Spectrochim. Acta*, **6**, 228 (1954)

2.7.4 Influence of the damping resistance

The damping resistance contained in the charging circuit of spark generators has the following purposes: (1) it has a damping effect on the equalising process arising in the charging circuit, (2) it counteracts any harmful overvoltage arising at the terminals of the condenser, and (3) it determines the number of condenser discharges occurring during the half-period.

In spark excitation the condenser is recharged during the half-period immediately following a complete discharge. It is therefore necessary that the equalising process that is taking place in the charging circuit is fully damped during this period. In some cases it is possible that the condenser may not attain the opposite polarity value of

its terminal voltage, or, may in fact exceed that level. This drawback can be eliminated if the equalising voltage drops during the half-period, that is, during the interval where $t = \pi/\omega_n$, to at least a hundredth of its initial amplitude value.

It follows from the previous chapters that the above requirement is met under different conditions, such as when the mains current charges the condenser through a transformer or charges the condenser directly. Different conditions will also exist if the damping resistance in the charging circuit of the condenser is smaller than or greater than, or again equal to the critical resistance.[1]

According to Eqs (2.130) to (2.132) for high-voltage and medium-voltage spark excitation, that is, when the condenser is charged through a transformer, the equalising current drops to one per cent of its initial value during a mains half-period provided:

(1) for aperiodic damping ($R > R_c$)

$$\frac{1}{\gamma_1 - \gamma_2} [\gamma_2 e^{\gamma_1 \pi/\omega_n} - \gamma_1 e^{\gamma_2 \pi/\omega_n}] \leqq \frac{1}{100} \tag{2.144}$$

(2) for critical damping ($R = R_c$)

$$\left(1 + \frac{\pi\omega_0}{\omega_n}\right) e^{-\pi\omega_0/\pi\omega_n} \leqq \frac{1}{100} \tag{2.145}$$

(3) for periodic damping ($R < R_c$)

$$\frac{1}{\cos \varepsilon} e^{-\pi R/2L\omega_n} \leqq \frac{1}{100} \tag{2.146}$$

No simple form of equation can be devised for the inequality expressed by Eq. (2.144), even when the term containing the exponent γ_2 is neglected. On the other hand, the expansion of the inequality expressed by Eq. (2.145) leads to the simple result that in the case of critical damping

$$\omega_0 \geqq 2{\cdot}11\omega_n \tag{2.147}$$

When it is considered that for the European frequency value of 50 Hz, $\omega_n = 314$, the angular frequency of the non-damped oscillations $\omega_0 = \sqrt{1/LC}$, and $R_c = 2\sqrt{L/C}$, then it follows from the above equation that

$$R = R_c \geqq 1325L \tag{2.148}$$

That is, in critical damping the minimum value of the damping resistance is determined by the induction of the transformer. This requirement is easy to meet in the case of an efficient transformer with a low stray value. As the value of L is generally in the order of 10^{-3} H, it is sufficient if the damping resistance has a magnitude of only a few ohms.

From Eq. (2.146) it follows that for periodic damping the value of the damping resistance is determined by the relation

$$R_{(R<R_c)} = \frac{2L\omega_n}{\pi} \ln \frac{100}{\cos \varepsilon} \tag{2.149}$$

When it is taken into account that $R_c = 2\sqrt{L/C}$, and that according to Eq. (2.75) $\varepsilon \approx 0$ in this case, then it can be established for $\nu = 50$ Hz that

$$R_{(R<R_c)} \geqq 920L \tag{2.150}$$

The minimum value of the damping resistance is, therefore, determined by the induction of the transformer for periodic damping as well. According to the above inequality and Eq. (2.148) it can be established that the above damping requirement can be easily achieved when an efficient transformer is used.

The one condition for the proper functioning of the charging circuit is that the equalising process should be adequately damped within the prescribed time. It is also essential that no overvoltage should arise in the condenser. It becomes clear from the diagrams in Fig. 2.77 that no overvoltage can arise in the case of critical damping. On the other hand, for periodic damping, high overvoltages can occur which can lead to the breakdown of the condenser. This problem can be avoided when the condenser is charged through a transformer in such a way that critical damping is applied during this operation.[2]

It is evident that the more the damping resistance exceeds the critical damping level, the less will be the likelihood of an overvoltage arising. An excessively large damping resistance is a disadvantage, however, from a practical point of view, since according to Eqs (2.118a) and (2.119) it reduces the maximum value of the stationary condenser voltage too much (Fig. 2.78).

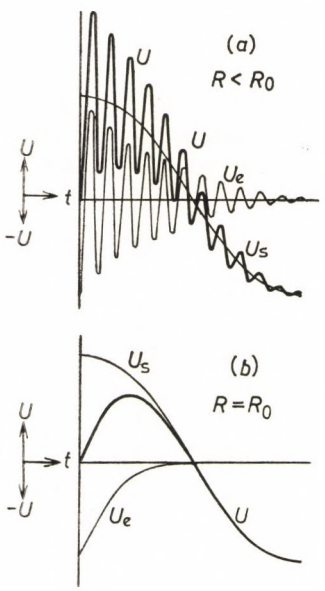


Fig. 2.77 Variation of the stationary voltage (U_s), equalizing voltage (U_e) and condenser voltage (U) with time for (*a*) periodic and (*b*) critical damping

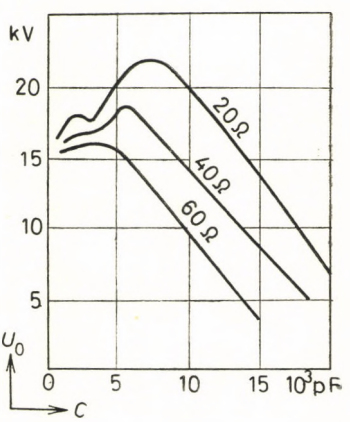


Fig. 2.78 Effect of the damping resistance on the initial voltage, U_0, of the condenser for different values of the condenser capacitance (C)[3]

As there should be a damping of at least 1 % during the half-period, the following relationships should hold

$$e^{-\pi/RC\omega_n} \leqq \frac{1}{100} \tag{2.151}$$

that is

$$RC \leqq \frac{\pi}{4{\cdot}6\omega_n} \tag{2.152}$$

This condition is achieved for an a.c. of frequency 50 Hz when

$$RC \leqq \frac{1}{460} \tag{2.153}$$

This condition can be met if the charging current is kept at a minimum (see Eq. 2.118), and the damping resistance is at the maximum allowed by Eq. (2.152). Thus according to Eqs (2.138), (2.140), and (2.152) we obtain

$$U_{\max} = 0{\cdot}8P_{\max} \tag{2.154}$$

Therefore, for mains-voltage spark generation the maximum absolute value of the condenser terminal voltage is 0·8 times the peak voltage of the mains current. From

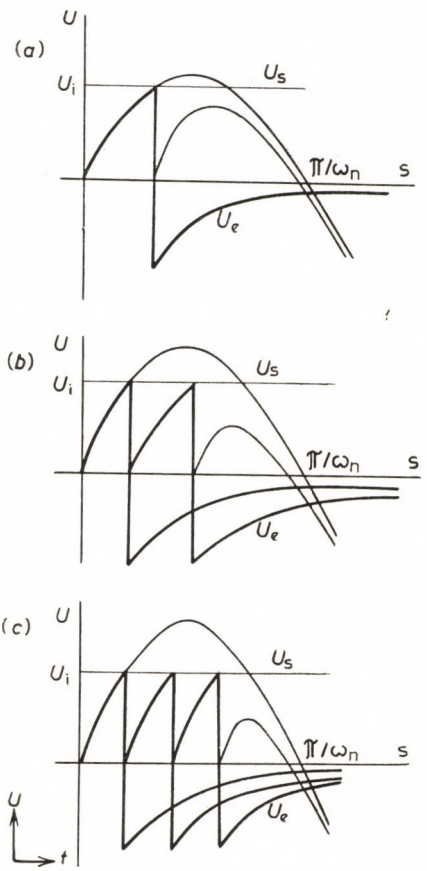


Fig. 2.79 Increase in the number of condenser discharges during a half-period

Figure 2.77 and according to Eq. (2.139) it can be stated that in this case no overvoltage can arise.

In the above discussion it has been assumed that the condenser only discharges once every half-period. This situation is very frequently not satisfied. If the terminal voltage of the condenser reaches the breakdown voltage, U_i, of the spark gap again within the same half-period, then the condenser will discharge repeatedly in that half-period.

In the discussion of the processes that take place in the charging circuit of the condenser, it was shown that the variation with time of the terminal voltage, U, of the condenser, is determined by the resultant of the stationary, U_s, and equalising voltages, U_e, (see Eq. (2.117)]. When, however, the value of the damping resistance, R, is much greater than that of the critical resistance, the equalising voltage, U_e, has only a relatively slight variation with time. Thus, the resulting voltage does not reach the value of the breakdown voltage, U_i, of the spark gap again in the same half-period. The condenser will, thus, discharge only once in a half-period (Fig. 2.79*a*). When the damping resistance, R, is reduced the maximum value of the stationary voltage is increased, although owing to the smaller damping the variation of the equalising voltage is speeded up. This results in an increase in the number of discharges in a half-period to two, three, or more (Fig. 2.79*b*, *c*). Thus, by the appropriate alteration of the value of the damping resistance the number of discharges in a half-period can be changed.

References

1. TÖRÖK, T., unpublished results
2. BARDÓCZ, Á., *Nature, Lond.*, **171**, 1156 (1953)
3. HÖFERT, H. J., and KECK, P. H., *Zeiss Nachr.*, **4**, 37 (1941)

2.8 Principal design of high-voltage spark generators

2.8.1 Classification of spark generators

Spark generators can be classified either according to the condenser voltage determined by their charging circuit or according to the method of control.

According to the charging voltage of the condenser the following categories exist:

(1) High-voltage (10–30 kV)
(2) Medium-voltage (about 1 kV)
(3) Mains-voltage (110–380 V) spark generators

According to the method of control the following classification is possible:

(*a*) Non-controlled discharge circuit
(*b*) Controlled discharge circuit
(*c*) Controlled charging and discharge circuit spark generators

Technical data for spark generators classified according to the charging voltage of their condenser are summarised in Table 2.12. (The induction of 1 μH and the resistance

of 0·1 Ω, given in the table, are the remanent induction and remanent ohmic resistances of the discharge circuit, respectively.)

Table 2.12

Technical data of commonly used spark generators

	High voltage	Medium voltage	Mains voltage
Condenser			
charging voltage, U_0 (eff)	10–30 kV	~1 kV	110–380 V
capacitance, C	1–20 μF	1–60 nF	10–250 μF
Discharge circuit			
inductance, L	1 μH–1 mH	1–400 mH	0·1–100 mH
damping resistance, R	0·1–10 Ω	0·1–300 Ω	~0·2 Ω

2.8.2 High-voltage, non-controlled spark generators

The principal circuitry of the non-controlled spark generators is given in Fig. 2.80. The mains a.c. charges the condenser, C, through the damping resistance, R_d, via the high-voltage transformer, Tr. The condenser is discharged through the induction coil, L, the ohmic resistance, R, and the analytical gap, AG.

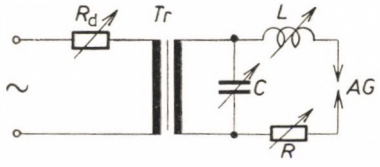


Fig. 2.80 Schematic circuit diagram of a high-voltage, non-controlled spark generator

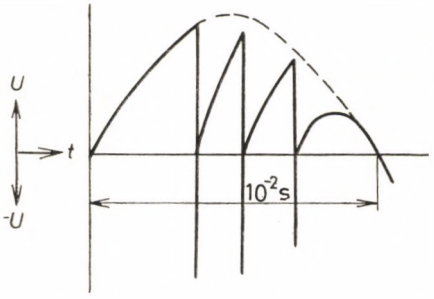


Fig. 2.81 Variation of the terminal voltage of the condenser as a function of time for non-controlled spark generators (schematic)

The electrical processes taking place in the charge and discharge circuit of the generator have already been given in Section 2.6 and Sub-sections 2.7.2 and 2.7.4. Thus, we shall be primarily concerned with the operation of the unit and more specifically with the problems of its control.[1]

When the charging circuit is properly designed, the terminal voltage of the condenser varies sinusoidally, or almost so, in every half-period (Fig. 2.81). The voltage difference between the electrodes in the spark gap also increases in the same way.

Prior to the discharge of the condenser there is no current flowing in the discharge circuit of the condenser. Accordingly, $I = 0$ and the voltage drop, $RI + L\mathrm{d}I/\mathrm{d}t$, is also zero. The condenser is then discharged at the instant when the terminal voltage value reaches the level of the breakdown voltage of the analytical gap. The duration of the discharge arising in this manner is generally in the range 10^{-3}–10^{-7} s, that is, it is smaller by 1–5 orders of magnitude than the duration of the mains half-period.

(The duration of the condenser discharge is so short that in the time scale shown on Fig. 2.81, the damped variations of the condenser terminal voltage during this interval can be illustrated as perpendicular lines.)

The discharged condenser is usually recharged in the same half-period and in so doing its terminal voltage again reaches the breakdown voltage level. The condenser may, therefore, be discharged and recharged a number of times in the same half-period and produce what are referred to as *partial discharges*. The second partial discharge usually occurs at a lower terminal voltage of the condenser than the first. The 'after effect' of the spark generated by the first discharge facilitates the ignition of the next spark.

It follows from the previous paragraph that the number of partial discharges is strongly influenced by the breakdown voltage of the electrode gap. The value of the breakdown voltage is, however, determined by many factors (see Sub-section 2.3.2). Among these factors the so-called *peak effect* can have an important role. It is possible that during the sparking process a small raised peak is formed at a particular point on the electrode tip. Due to the high electron density on the peak, a comparatively larger electron emission is produced here relative to the peak surroundings.

As the volume of gas above the peak is more strongly ionised than the surrounding gas, the discharge channel of the spark tends to start from the peak area. Accordingly this part of the electrode will be excessively heated such that the local value of the breakdown voltage is further reduced. Therefore, there is no change in the initial position of the spark discharge channel as it will always start again from the peak. In the course of the sparking process, the material forming the peak can be either totally evaporated or mechanically removed from the electrode surface. The peak effect can, therefore, be terminated and the breakdown voltage of the electrode gap allowed to increase and reach its normal value.

The apparent burning of the spark at the electrode can be caused by other effects, as well as that due to the formation of a peak. The burning at the surface can be traced to the fact that on the heterogeneous electrode tip the less volatile particles are left behind when the surface material evaporates. On the originally homogeneous surface, conglomerations of such particles (for example, oxides) are able to develop that alter the geometrical form of the electrode surface. The burning can, furthermore, be caused by various local chemical and thermal effects acting on the surface exposed to the sparks, which lead to a *local* lowering of the breakdown voltage.

According to Eq. (2.77) the variation of the breakdown voltage during sparking alters the maximum current of the condenser discharge. Thereby the temperature of the spark plasma is changed, which causes a corresponding change in the relative and absolute intensity of the spectral lines as well. Thus, it is an experimental fact that the intensity fluctuations of the spectral lines during sparking are twice as large for non-controlled spark excitations as for controlled generators.[1] From the diagram shown in Fig. 2.82 it is also evident that certain spectral line intensity values

are much more frequent in the case of non-controlled spark generation than for controlled sparks. It can be shown experimentally that these particular intensity values are established by the 'surface burning' phenomenon occurring on certain parts of the electrode tip. As we have seen, this leads in turn to a change in the breakdown

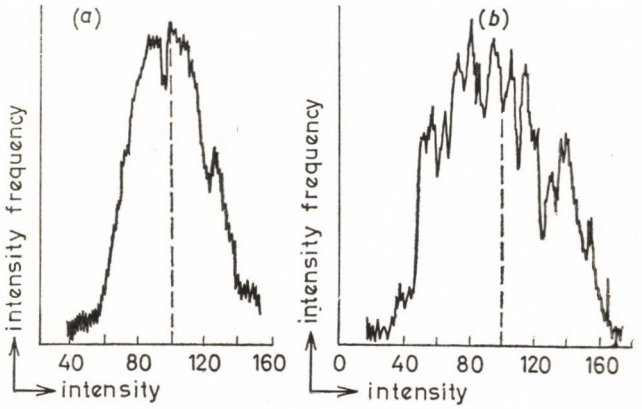


Fig. 2.82 Intensity fluctuations of the spectral line C II 3920: (*a*) controlled and (*b*) non-controlled spark generation[2]

voltage of the spark gap, together with the initial voltage of the condenser discharge and the temperature of the spark plasma. A further increase in the intensity fluctuation is due to the fact that when there is a major change in the breakdown voltage the number of condenser discharges in a half-period can also change.

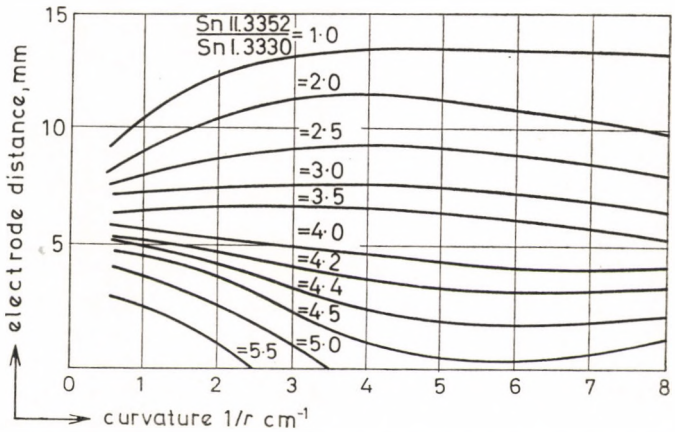


Fig. 2.83 Influence of the electrode distance and the radius of curvature of the electrode surface on the relative intensity of the spectral lines[3]

In the case of non-controlled spark generation it should also be taken into account that the breakdown voltage, and also the character of the spectrum obtained, depend on the distance apart and the form of the electrodes (Fig. 2.83). The quality of the finish of the electrode tip and the nature of the material used for the electrode are

also important in this respect. The errors that arise from these factors can be eliminated in practice, however. Those errors produced by changes in the physical or chemical properties of the electrode tip cannot be influenced by changing the practical conditions. Still another error could be caused by the incorrect design of the charging circuit and the damping resistance of the generator. Considering all these drawbacks it is understandable that non-controlled spark generators are little used nowadays.

The so-called *resonant spark generators* are the intermediate types between non-controlled and controlled spark generators.[4,5] In the practical application of resonant generators the control is obtained by setting the proper electrode distance, switching in the necessary capacitance and self-inductance, and adjusting the damping resistance of the spark generator and the transformer inductance until the breakdown of the spark gap becomes uniform. The precise setting is facilitated by using a variable resistor free of self-inductance in the discharge circuit of the condenser. The governing principle of the method is to set the frequency of the charging circuit equal to the frequency of the mains by the appropriate setting of the parameter values mentioned above. From the design point of view, resonant spark generators are different from non-controlled generators in that the mains current is connected to the transformer through a potentiometer. This transformer has a comparatively large variable inductance ratio (1 : 400). (The electrical processes in the charging circuit of the resonant generator have been discussed in Section 2.7.2.)

The handling of resonant spark generators is relatively complicated and as their application requires considerable experience, they tend to be little used today.

References

1. KAISER, H., *Spectrochim. Acta*, **1**, 477 (1941)
2. KREMPL, H., and SPATH, H., *Arch. EisenhüttWes.*, **31**, 473 (1960)
3. SCHEIBE, G., and SCHÖNTAG, A., *Arch. EisenhüttWes.*, **8**, 533 (1935)
4. SCHEIBE, G., *Chemische Spektralanalyse* (BÖTTGER, W., *Physikalische Methoden der analytischen Chemie*, Leipzig, vol. I, 27 (1933)
5. CORLISS, C. H., *Spectrochim. Acta*, **5**, 19 (1952)

2.8.3 High-voltage spark generators having a controlling spark gap[1–4]

The reproducibility of spark generation can be greatly improved if the time of the condenser discharge is determined by the breakdown voltage of a so-called controlling spark gap, connected in series to the analytical spark gap instead of being determined by the latter. The circuit designs of spark generators having a double spark gap and the Raiski spark generator which is constructed on their principle are shown in Fig. 2.84. Both constructions differ from non-controlled spark generators in that the condenser is discharged through two spark gaps instead of one. The difference between the two designs is that the analytical spark gap is bridged by a large resistance (measured in megohms) or by a large inductive resistance (of about 0·1 mH).

The *controlling spark gap* S_c is usually made of smooth-surfaced tungsten or platinum electrodes that are water-cooled and whose distance can be reproducibly set using a micrometer screw. A constant-rate air stream blown into the controlling electrode gap removes the charged particles left after the discharge of the condenser. Thus, the breakdown voltage of the controlling spark gap becomes entirely independent

of various uncertain factors, such that for a fixed electrode distance its value remains constant. Accordingly, the condenser discharge will always occur at a predetermined terminal voltage. The ohmic or inductive resistance bridging the analytical spark gap does not influence the potential difference between the electrodes of the controlling spark gap. The condenser is, therefore, discharged when the voltage difference

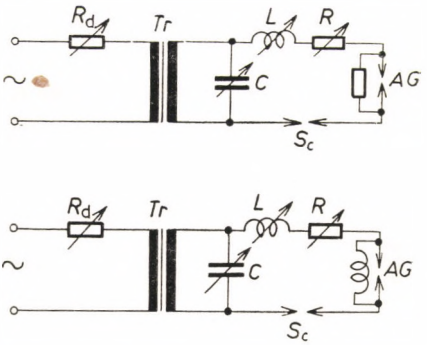


Fig. 2.84 Circuit diagrams of high-frequency spark generators with controlling spark gap: (*a*) tandem spark gap and (*b*) Raiski spark generator

between the controlling gap electrodes reaches the value of the breakdown voltage. When the voltage difference between the electrodes of the analytical gap is about 50 V, the bridging resistance presents such a high resistance towards the discharge current that practically no current is passed during the discharge of the condenser.

High-voltage spark generators with controlling spark gap have other advantages in addition to the fact that the condenser is always discharged at a predetermined voltage defined by the electrode distance of the controlling spark gap. For example, it is an advantage that the number of condenser discharges in a half-period can be varied between wide limits (usually 1–15) via the ohmic and inductive resistors contained in the primary circuit of the transformer. On the other hand, it is a drawback from the theoretical point of view that the energy accumulated within the condenser is dissipated in two spark gaps instead of being used exclusively in the analytical gap.

References

1. ZEHDEN, W., *J. Soc. Chem. Ind.*, London, **59**, 230 (1940)
2. KAISER, H., *Spectrochim. Acta*, **1**, 477 (1941)
3. WOLFE, R. A., and JEMAL, E. J., *Met. Ind.*, London, **65**, 274 (1945)
4. LEVY, S., *J. Opt. Soc. Am.*, **35**, 221 (1945)

2.8.4 Mechanically controlled, high-voltage spark generators

In mechanically controlled, Feussner-system, high-voltage spark generators, the discharge of the condenser is effected by a motor-driven rotating switch, S_r, at the instant of a free charge (Fig. 2.85).[1]

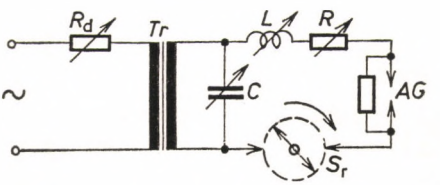


Fig. 2.85 Circuit diagram of the Feussner mechanically controlled, high-voltage spark generator

In setting the rotating switch it should be taken into account that there is generally a phase-shift between the mains voltage and the secondary voltage of the transformer and thus the terminal voltage of the condenser. The value of the phase angle is influenced by all the properties of the charging circuit. The phase-shift is given by the following formula

$$\tan \phi = \frac{1}{R_d}\left[L\omega_n - \frac{1}{C\omega_n}\right] \tag{2.155}$$

where R_d is the damping resistance, C is the capacitance of the condenser as recalculated for the primary side (see Sub-section 2.7.2), and L is the inductance of the transformer. Thus, the greater the capacitance of the condenser the further the rotating switch should be set in the direction opposite to its rotation. When the rotating switch is incorrectly set the condenser is discharged at the moment when its terminal voltage has not yet reached the maximum value (Fig. 2.86). The smaller variations in the parameters of the charging circuit can be neglected in this context. This is so since the phase angle has the least effect on the sine-form of the voltage when the latter is in the neighbourhood of extreme values.

The anomalies stemming from the variation of the phase angle can be eliminated by either mechanical or electrical means. In the mechanical procedure the position of either the arm of the rotating switch or the stationary electrode in relation to the synchronous motor should be changed. The latter procedure achieves the best results as the stationary electrodes can be easily rotated during the sparking procedure.

Modern Feussner-system spark generators are constructed such that the variation of the phase angle, as a function of the capacitance of the condenser, is automatically compensated by electrical means. In this equipment the capacitance of the condenser is varied such that those parameters that govern the phase angle of the synchronous motor are changed simultaneously. There are several ways to solve this problem.

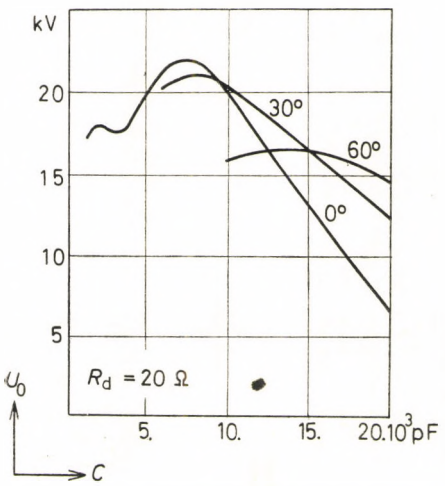


Fig. 2.86 Variation of the discharge voltage of the Feussner-system spark generator as a function of the condenser capacitance, at different rotary switch settings. Damping resistance 20 Ω [2]

It is relatively simple when the three-phase synchronous motor is powered not directly from the mains, but when appropriately selected condensers of identical capacitance are inserted between the three phase points.[4] The condensers and the

motor windings connected in series with the former represent an L–C circuit within which the phase angle of the current increases on increasing the value of the capacitance. Thus, the phase lag of the spark generator charging current can be equalised by phase-delaying the current operating the synchronous motor, provided the capacitance of the three condensers is altered according to the condenser capacitance of the spark generator. In this relatively simple solution it should also be considered that spark generators are usually powered by a single-phase mains, whereas the previously described method can only be applied when the synchronous motor driving the rotating switch is three-phase. This difference can easily be eliminated by applying artificial phases for operating the synchronous motor. These phases can be simply produced by the condenser phase-shift method. The relative phase angle of the artificial phase is 120°.

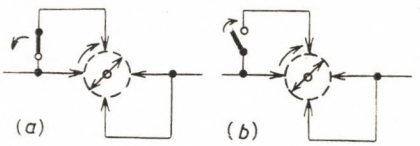


Fig. 2.87 Circuit diagram of the electrode arrangement for the rotary switch using a synchronous motor speed of 1500 rms: (*a*) type 1 + 1 discharge and (*b*) type 1 + 0 discharge

The rotating switch of Feussner-system spark generators is made of a short-circuited, tungsten electrode pair fastened on an arm perpendicular to the shaft of the synchronous motor, with another, usually 'stationary', tungsten electrode pair connected into the discharge circuit of the condenser. When, for example, the revolutions per minute of the synchronous motor, driven by a 50 Hz mains current, number 3000, the condenser is discharged only once in a half-period. That is, a discharge of type 1 + 1 is produced. On the other hand, the discharge type will be 0 + 1 if the speed of revolution of the synchronous motor is 1500 revolutions per minute. In the application of the previous type of synchronous motor a discharge of type 1 + 1 can be established provided a 'stationary' switching device, consisting of *two* electrode pairs at right angles, is used (Fig. 2.87).

The condenser of Feussner-type spark generators is discharged through three spark gaps – the analytical gap and the two spark gaps of the rotating switch. Thus, the collective resistance of the three spark gaps influences the value of the condenser current. The energy dissipated within the rotating spark gap governs that portion of the energy accumulated within the condenser that will go to the analytical spark gap. Therefore it indirectly influences the temperature of the spark plasma and the spectral character of the emitted radiation. The collective breakdown voltage of the three spark gaps determines the instant that the condenser discharge commences at. An analytical error can therefore occur if the distance between the electrodes of the rotating switch or the geometry of the electrodes is changed during the spark generation process. The latter factor cannot be eliminated during the long-term operation of the rotating switch, since an oxide layer is built up on the front portion of the tungsten electrodes. Furthermore, the peripheral velocity of the rotating switch or the speed of revolution of the synchronous motor is also an important factor after the discharge has been initiated.

In principle, the exact time of the condenser discharge is also influenced by the breakdown voltage of the analytical gap, even in the case of Feussner-type spark generators. The uncertainty of this instance can be eliminated when the analytical

gap is bridged by a large resistance.[3] In this way, the terminal voltage of the rotating control spark gap becomes equal to that of the condenser before the discharge starts, such that the influence of the analytical spark gap vanishes. During the burning period of the spark the electrical conductivity of the plasma is so great that the bridging resistance does not at all influence the excitation. By taking all these factors into account it can be stated that in Feussner-type spark generators the condenser discharge is controlled exclusively by the rotatory spark gap system.

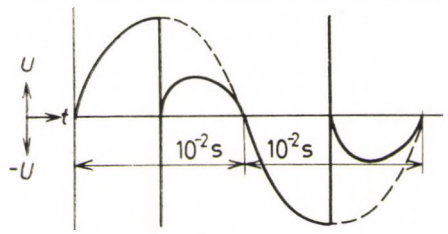


Fig. 2.88 Condenser voltage–time curve of a Feussner spark generator ($\nu = 50$ Hz)

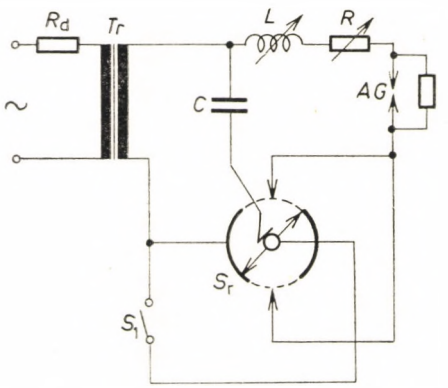


Fig. 2.89 Circuit diagram of a high-voltage spark generator with a mechanically controlled charging circuit[4]

In Feussner-type spark generators the discharge of the condenser is terminated when the continually dropping terminal voltage of the condenser is not sufficient to sustain the sparking process at all three electrode gaps. This process lasts for about 10^{-6}–10^{-3} s and should, of course, be taken into account when computing the diameter of the rotating switch. It is also important to consider that, following the instant of the discharge the charging of the condenser restarts (Fig. 2.88) and if this process is too rapid the condenser may be discharged again before the rotating switch disconnects. This disturbance can be eliminated by the appropriate selection of the damping resistance.

It can be established that Feussner-system spark generators usually meet all the practical demands normally encountered in the case of high-voltage spark generators. The spectra obtained by the spark generator are sufficiently reproducible and the structure and handling of the equipment are relatively simple. Thus, it is understandable that Feussner-system spark generators are still the most popular sparking device.

In the case of mechanically controlled charging circuit, high-voltage spark generators the partially delayed charging of the condenser previously mentioned does not occur. This apparatus is, therefore, different from the Feussner-type generator in that the rotating switch controls not only the discharge current of the condenser, but

also its charging current. The two switches are driven by the same synchronous motor.[4]

In mechanically controlled, high-voltage spark generators (Fig. 2.89) the rotating switch, S_r, connects the transformer with the condenser at that instant when the stationary voltage is zero, although the stationary current itself is not zero. The equalising process that develops does not cause any trouble, since it is adequately damped by the damping resistance in the charging circuit during the quarter-period

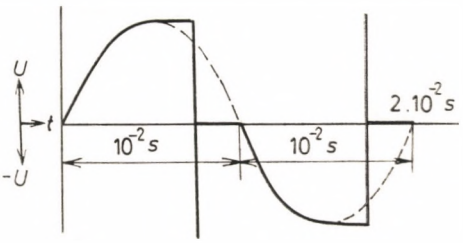


Fig. 2.90 Condenser voltage–time curve of a high-voltage spark generator with a controlled charging circuit, using a type 1 + 1 discharge ($\nu = 50$ Hz)

available for the charging of the condenser. The rotating switch disconnects the charging circuit after the quarter-period, when the value of the stationary voltage is at maximum. At this moment the charging current is zero and no new equalising process can occur, that is, the condenser is completely charged. The S_r rotating switch connects the *discharge circuit* of the condenser to the analytical spark gap with the fully charged condenser after a quarter-period, at which time the condenser will be discharged.

The condenser voltage–time curve of mechanically controlled, charging circuit spark generators has the form shown in Fig. 2.90, when the discharge type is 1 + 1. This occurs when the stationary electrodes of the rotating switch of the charging circuit are short-circuited by the switch denoted by S_1 in Fig. 2.89. When this switch is in the open position the discharge type will be 1 + 0.

The rotating switch of the apparatus functions partly as a mechanical rotating rectifier. This operation, as we shall see in the next Sub-section, can be performed by means of two-channel electronic rectification as well. The mechanical version has two advantages, however, over the electronic method. Firstly, the corresponding mechanical elements can be easily combined with the rotating switch such that their operation is perfectly dependable. On the other hand, the mechanical solution makes it possible to produce discharges of type 1 + 1 without doubling the number of electrical elements connected to the secondary winding of the transformer and having another secondary winding as well (see Fig. 2.93*a*).

The advantage of controlling the charging current is that the discharge of the condenser is entirely independent of the charging circuit. The danger of striking a high voltage arc after a relatively long discharge is therefore eliminated.

References

1. Feussner, O., *Z. tech. Phys.*, **13**, 573 (1939); *Arch. Eisenhütt Wes.*, **6**, 551 (1953)
2. Höfert, H. J., and Keck, P. H., *Zeiss Nachr.*, **4**, 37 (1941)
3. Kaiser, H., *Spectrochim. Acta*, **2**, 98 (1941)
4. Török, T., *Explle Tech. Phys.*, **7**, 49 (1959)

2.8.5 High-voltage spark generators with a rectified charging circuit

In these devices the condenser is charged through the rectifying tube. The rectification can be of one- or two-channel type. In the former case, the condenser is charged and discharged only once during a period, while in the latter charging and discharging occur twice.

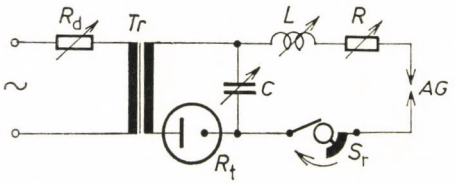


Fig. 2.91 Schematic circuit diagram of a high-voltage spark generator with a charging circuit controlled by a rectifying tube (half-period rectification)

In the case of *unidirectional rectification* the condenser is charged through the rectifying tube, R_t, placed in one branch of the charging circuit (Fig. 2.91). The rectifier transmits the secondary current of the transformer only in one half-period, while during the other half-period the tube acts as a closed valve. The effect of the rectification tube is, therefore, exactly the same as a rotating switch in the charging circuit. The rotating switch in the discharge circuit of the condenser is required to allow the condenser to discharge during the half-period following the initial discharge of the condenser. In this manner discharges of type 0 + 1 are produced and the electrodes retain their polarity during the excitation. Accordingly, the condenser voltage–time curve has the form as shown in Fig. 2.92.

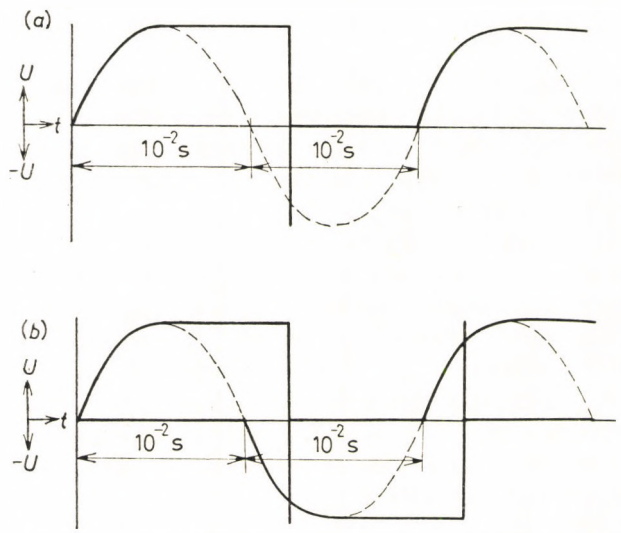


Fig. 2.92 Condenser voltage–time curve of a high-voltage spark generator with a rectified charging current: (*a*) simple half-wave rectification and (*b*) full-wave rectification

A condenser discharge of type 1 + 1 can only be produced by a spark generator that has *duodirectional rectification.* For such a rectification, however, a transformer

is necessary that has a double secondary voltage and is centre-tapped. In addition, two more, oppositely connected rectification tubes, plus two condenser batteries, are necessary, that consist of condensers of exactly the same capacitance (Fig. 2.93*a*). In this way a discharge is obtained whose condenser voltage–time curve is shown in Fig. 2.92*b*. On the other hand, discharges of type 2 + 2 can be obtained in such a way that the number of rectification tubes and condenser batteries is doubled together with the number of stationary electrodes in the synchronous switch (Fig. 2.93*b*).

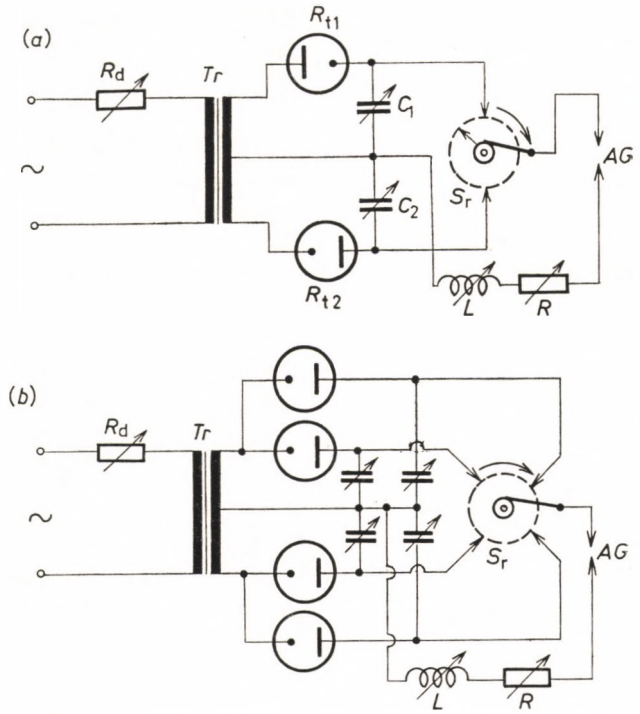


Fig. 2.93 Circuit diagram of a high-voltage spark generator with a full-wave rectified charging circuit for the production of: (*a*) type 1 + 1 discharges and (*b*) type 2 + 2 discharges

The rectification-tube spark generators described above are basically intermediate forms of electronically controlled spark generators. The latter are more expensive than mechanically controlled spark generators as the equipment is more delicate and of more modern construction. Their application has the advantage, just as is the case with mechanically controlled charging circuit spark generators, that the discharge of the condenser is not disturbed by processes arising in the charging circuit.

A special version of rectified charging circuit, high-voltage spark generators is the so-called *filament spark generator*.[1–4] In this equipment a rectified secondary current of 6–10 V from the transformer charges a small-capacity (25–100 pF) condenser (Fig. 2.94). Under such conditions the condenser is charged only in every even half-period, while it is discharged in every other half-period. The discharge current during the even half-periods passes through a rectification tube connected

in series with the spark gap that is called the *counter-voltage equalising tube*. By this method a *rectified condenser discharge* is produced (Fig. 2.95).

During the experimental procedure described above special excitation conditions are created. Owing to the relatively low capacitance of the condenser and the small induction of the discharge circuit, the frequency of the condenser discharge current is extremely high and could exceed 20 MHz, such that its vibrational time is of the

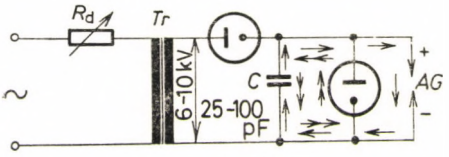


Fig. 2.94 Circuit diagram of a filament spark generator → direction of condenser in an odd half-period, → → direction of condenser in an even half-period

order of only 10^{-8} s. When a condenser is charged at a low energy level, the maximum current developed on discharge is relatively low, about 10–15 A. This has the consequence that only a small amount of electrode material is evaporated within unit time. Thus the resistance of the spark plasma is very large and an extremely large potential drop (about 2000 V) occurs. Due to the latter fact and to the short vibration time of the condenser discharge, very high energy levels are excited. The character of the emitted light is definitely that of ion spectra, although its intensity is low. Even in the most favourable circumstances the intensity does not reach one-fifth of the average intensity obtained by common high-voltage spark generators.[5]

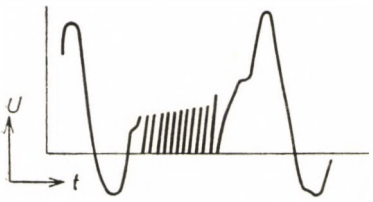


Fig. 2.95 Condenser voltage-time oscillogram generated by a filament spark generator[5]

A further characteristic of filament spark generation is that the spark plasma does not dilate but keeps its filamentary form all through the excitation. Thus, it is possible to achieve a burning area smaller than 0·01 mm^2. Due to the low evaporation speed of the electrode material, at electrode separation distances of about 10 mm, the radiation of the vapour from the two electrodes is completely separated in space. In this way the emission of the materials from the two electrodes is not mixed.[6] Furthermore, it can be shown experimentally that the reproducibility of the spectra obtained by filament spark excitation is better than mechanically-controlled high-voltage spark generators.[5]

References

1. Scheibe, G. and Martin, J., *Spectrochim. Acta*, **1**, 47 (1939)
2. Murray, W. M., Gettys, B. and Ashley, S. E., *J. Opt. Soc. Am.*, **31**, 433 (1941)
3. Brijan, F. R., and Nahstoll, G. A., *J. Opt. Soc. Am.*, **37**, 311 (1947)
4. Hurwitz, J. K., *J. Opt. Soc. Am.*, **42**, 484 (1952)
5. Mika, J. and Péter, L., unpublished results
6. Paksy, L., *Acta Chim. Hung.*, **28**, 17 (1961)

2.8.6 Electronically controlled, high-voltage spark generators

The discharge of the condenser of mechanically controlled and rectified charging circuit spark generators is regulated by a rotating switch. It is, however, a disadvantage that the variation of the time of the condenser discharge is rather complicated and cannot be performed exactly enough. Also the moment of spark initiation depends on the condition and distance apart, etc., of the electrodes of the rotating switch. All these drawbacks can be removed by the application of electronic control. In electronic control the phase instant of the condenser discharge is remarkably stable and can be easily varied. However, the equipment is comparatively complex as well as expensive. Where the reproducibility of the phase instant of the condenser discharge is an important aspect this method is indispensable.

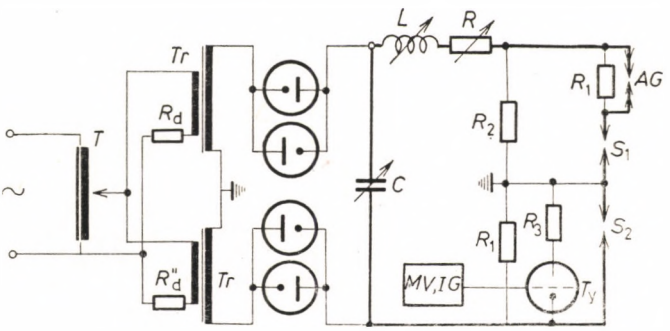


Fig. 2.96 Simplified principal circuit diagram of the Bardócz electronically controlled, high-voltage spark generator[6]

There is a number of ways of electronically controlling high-voltage spark generators.[1-5] Among these the Bardócz system has proved to be the best. The principle of control is the same as that discussed in Sub-section 2.4.3 and will be only briefly outlined here. The electronically controlled spark generator differs from the non-controlled types in that the discharge circuit of the condenser contains two spark gaps (S_1 and S_2) connected in series to the analytical gap, *AG* (Fig. 2.96). In this system the condenser is discharged when the controlling multivibrator, *MV*, ignites the thyratron tube, T_Y, via the pulse generator, *IG*, and thereby short-circuits spark gap, S_2. In this way either 1 + 1 or 1 + 0 type discharges can be produced depending on the appropriate circuit design for the rectifying tubes. The advantage of the generator is that it controls both the discharge and the charging circuit. On the other hand,it is a drawback that condenser discharges of type 1 + 1 can only be produced by doubling the high-voltage elements of the generator (that is, the secondary winding of the transformer, and the number of rectifying tubes).

This problem is eliminated when a specially designed induction transformer is used whose iron core quickly becomes saturated on increasing the primary current, *I* (Fig. 2.97*a*). This means that the magnetic flux, Φ, prevailing in the iron core assumes constant values during its rapid variation with time (Fig. 2.97*b*). Since the voltage difference, *U*, at the terminals of the secondary winding of the transformer is proportional to the value of $d\Phi/dt$, there emerges in every half-period a voltage peak

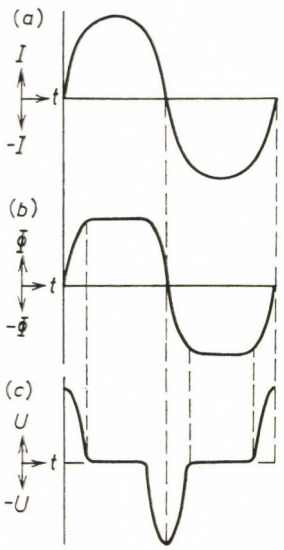


Fig. 2.97 Current–time curve (*a*), flux–time curve (*b*), and voltage–time curve (*c*) of the Bardócz-system special stray transformer[7]

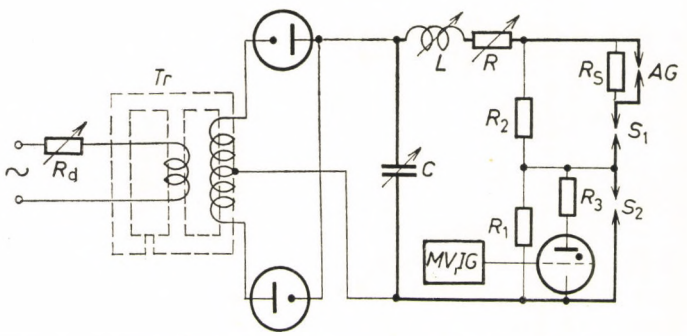


Fig. 2.98 Schematic circuit diagram of the simplified Bardócz-system electronically controlled high-voltage spark generator[7]

lasting for a comparatively short time (Fig. 2.97*c*). By fixing the starting signals of the discharge circuit correctly, the condenser can be discharged every half-period, such that a discharge of type 1 + 1 is produced. (The circuit design of this type of electronically controlled high-voltage spark generator is shown in Fig. 2.98)[7].

References

1. Malpica, J. T. M., and Berry, T. M., *Gen. elect. Rev.*, **39**, 298 (1940); **44**, 563 (1941)
2. Shirley, H. T., Oldfield, A., and Kitchen, H., *J. Iron and Steel Institute*, **166**, 329 (1950)
3. Bardócz, Á., *Nature, Lond.*, **171**, 1156 (1953); *Acta Phys. Hung.*, **7**, 307 (1953); *Acta Tech. Hung.*, **8**, 99 (1954); *Appl. Spectrosc.*, **10**, 183 (1956)
4. Bardócz, Á., *Acta Tech. Hung.*, **13**, 3 (1955)
5. Bardócz, Á., *Acta Tech. Hung.*, **40**, 179 (1962)
6. Bardócz, Á., *Acta Phys. Hung.*, **4**, 91 (1954)
7. Bardócz, Á., *Z. angew. Phys.*, **9**, 82 (1957)

2.8.7 Pulse-spark generators

By using mechanically or electronically controlled spark generators, the phase instant of the discharge and the number of discharges occurring during a unit period, or the distribution with time of the discharges, can be regulated. On the other hand, the amount of energy dissipated during the individual discharges within a unit time, that is, the value of $\Delta E/\Delta t$ cannot be controlled. This naturally reduces the reproducibility of the spectra. So-called pulse-spark generators do, however, meet this latter requirement as well.

Pulse-spark generators are, in fact, modified (mainly in their efficiency) pulse generators that originally served some other purpose, for example in radar equipment, etc., and therefore are complicated electronic devices. Among the equipment that has proved ideal for spectrochemical analysis are those whose basic frequency is 400 Hz and where the duration of individual pulses can be set between 0·2 and 5 μs, with an output of 2 MW.[1] In these generators the number of 'sparks' occurring within unit time is about four times as large as in the common spark generators; that is, the number of the pulses is the same as for type 4 + 4 condenser discharges. The ratio of the time duration of the pulses and the intervals between them is about 1 : 1000, while the energy of a single pulse is about 1 J.

From the above data it follows that the excitation energy of pulse-spark generators is very high. Accordingly, the sensitivity of the excitation is very high, too. Very minute amounts of elements that are difficult to excite can be detected (for example, about 10 ppm. As or Te). Furthermore, it can be shown experimentally that the reproducibility of the excitation is excellent. The scatter obtained from replicate measurements is about 50% lower than the corresponding scatter of results obtained by Feussner-type spark generators. The major drawback of the equipment is its costly nature.

Reference

1. LEUTWEIN, F., *Chem. Tech., Berl.*, **4**, 512 (1952); *Explle Tech. Phys.*, **1**, 173 (1953)

2.9 The principal design of low- and medium-voltage spark generators

2.9.1 General remarks

Medium-voltage spark generators are devices in which the condenser is charged by a current of about 1 kV. On the other hand, the condenser of the *low-voltage*, more appropriately called *mains-voltage*, spark generator is charged directly by d.c. or a.c. mains. In the latter case the effective value of the condenser terminal voltage – apart from the voltage drop occurring due to the damping resistance – is generally 220 V, or occasionally 380 or 110 V. The important structural difference between the generator types is that a transformer forms part of the charging circuit of medium-voltage generators, while low-voltage equipment functions without the application of a transformer. This difference is also important when operational principles are concerned, as will become clear from Sub-sections 2.7.1 to 2.7.4.

As seen from Tables 2.8–2.10, (pp. 149–150) by correctly choosing the capacitance of the condenser it is possible to make the energy accumulated within the condenser of high-, medium-, and mains-voltage spark generators the same, irrespective of the terminal voltage of the condenser.

As opposed to high-voltage spark generators, the terminal voltage of medium- and low-voltage generators is usually insufficient to produce a breakdown of the spark gap. In these generators, therefore, the condenser discharge can only be initiated when a high-frequency, high-voltage ignition current is used for this purpose. This necessitates the insertion of a valve-circuit between the charging and discharging circuits of both medium and mains voltage generators that isolates the high-frequency current from the charging circuit.

2.9.2 Medium-voltage spark generators

Both arc excitation and spark excitation have essential properties that render them indispensable for practical analysis. Therefore, it is understandable that there is a tendency to devise generators that are equally applicable for producing atomic, ion, or any intermediate type of spectra. This goal can be achieved by using medium-voltage spark generators.

The above statement can be justified either theoretically or experimentally. The initial, maximum current of the discharge of a condenser is theoretically higher for medium-voltage spark generators than for high-voltage spark excitation, provided the ohmic resistance of the discharge circuit is small (see Table 2.8, p. 149). Under identical experimental conditions, the damping period of the discharge is longer in the case of medium-voltage spark generation than for high-voltage excitation. Taking this point into consideration, it can be assumed that the temperature and the ionisation degree of the spark plasma and thus the character of the emitted spectra, are practically identical under both excitation methods. It can also be shown theoretically that the initial current value of the discharge is about 10 A for medium-voltage spark excitation. This is about the same value as in a.c. arc generation, when the ohmic and inductive resistance of the discharge circuit of the condenser is large. The duration of the aperiodically damped discharges brought about under the above conditions is roughly equal to the burning time of the a.c. arc in a half-period. It follows that by means of medium-voltage spark generation, spectra can be obtained whose character is identical to that of the ion spectra produced by high-voltage spark generation. On the other hand, atomic spectra can also be excited whose character is identical to that of spectra produced by a.c. arc generation. This statement is proved by the data plotted in Fig. 2.99. As shown, the value of the logarithmic intensity ratio of the variant line pair considered does not differ for spectra obtained by medium- and high-voltage spark generation. On the other hand, when the discharge is strongly damped the value of the logarithmic intensity ratio that belongs to the same line pair is essentially lower for a.c. arc generation. Medium-voltage, strongly damped discharge spark generators render unnecessary the use of an arc generator in addition to a spark generator for the accomplishment of different analytical tasks.

Among the known medium-voltage spark generators, the multi-source unit[1,2] is mainly used in America (Fig. 2.100). In this equipment a transformer converts the mains current into a current of 1000 V. The mains current can be controlled by

a damping resistance. The rectified secondary current of the transformer charges a condenser of relatively large capacitance (variable between 1 and 60 μF) every other half-period. The condenser is discharged in alternate half-periods, when the rotating switch driven by a synchronous motor switches on the high-voltage, low-cur-

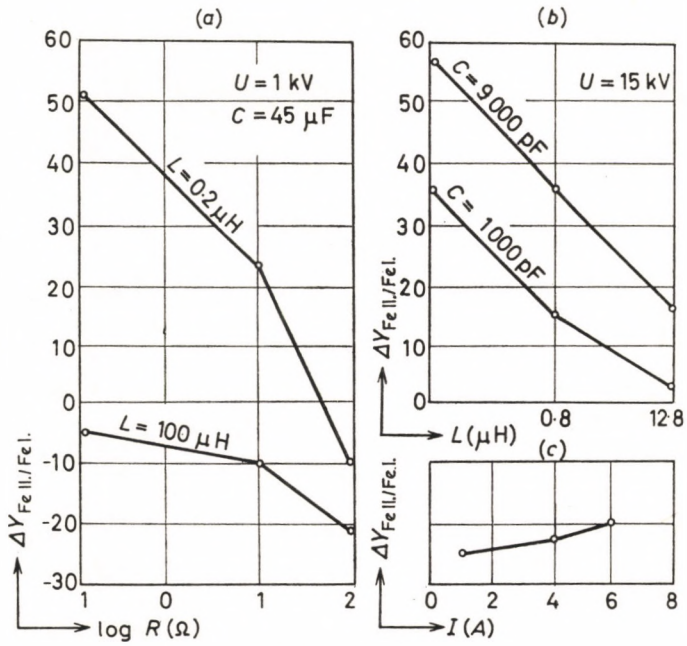


Fig. 2.99 Variation of the logarithmic intensity relations of the variant line pair Fe II 2941·40/ Fe I 2941·34 Å for (*a*) medium-voltage spark generation, (*b*) high-voltage spark generation, and (*c*) a.c. arc excitation[3]

rent ignition current. The phase point of the ignition can be controlled by the setting of the rotating switch. Although in this case only one discharge is possible within a period, it has the advantage that the condenser is fully discharged even when over-damping of the discharge occurs (Fig. 2.101). The character of the spectrum can be changed by means of a variable inductive resistance within the discharge circuit

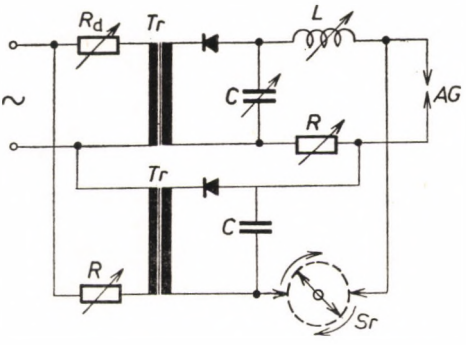


Fig. 2.100 Circuit diagram of the 'multi-source unit', medium-voltage spark generator

of the condenser or by a variable ohmic resistance. The range of the former is usually 0–400 μH, that of the latter 0–300 Ω.

The advantage of the equipment is that the charging and discharging processes are entirely independent of each other and can be separately controlled. A further advantage is that the electrodes of the spark gap have a definite polarity. The dominant characteristic of the multi-source unit is that it is very versatile and spectra of any desired character can be produced, ranging from pure ionic spectra to atomic spectra. The reproducibility of the spectra is far better than that obtained by using mechanically controlled, high-voltage spark generators.

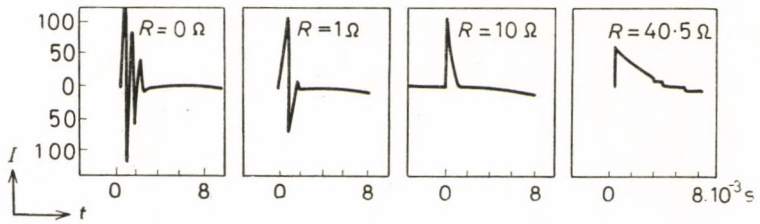


Fig. 2.101 Current–time oscillograms of the 'multi-source unit' spark generator for different ohmic damping values

The difference between mechanically controlled, charging and discharging circuit, medium-voltage spark generators and similar high-voltage spark generators is not only confined to differences in the terminal voltage and capacitance of the condenser, and the ohmic or inductive resistance within the discharge circuit. It is also an important difference that the rotating switch regulates not only the charging and discharging of the condenser, but also the instance of spark ignition as well.[3]

There is no doubt that an electronic control instead of a mechanical control is much more dependable. For electronically controlled, medium-voltage spark generators the electronics only cause ignition of the spark.[4] On the other hand, Fig. 2.102 shows a circuit diagram in which both the charging and discharge circuits of the condenser are electronically controlled. The pulse generator, *IG*, renders the triode, T_1, conducting at the appropriate instant and thereby charges the condenser, *C*.

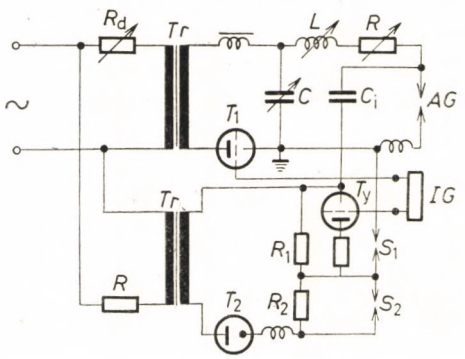


Fig. 2.102 Circuit diagram of the Bardócz-system electronically controlled medium-voltage spark generator[4]

The discharge of the condenser occurs when the thyratron tube, T_Y, short-circuits the secondary spark gap, S_1, such that the condenser, C_i, charged to some kV, is discharged through the analytical spark gap, *AG*.

In summary, it can be stated that neither the mechanically controlled, nor the electronically controlled medium-voltage spark generators described in the above Sub-section have found widespread use. On the other hand, the multi-source unit, medium-voltage spark generator is very common, especially in the United States of America.

References

1. HASLER, M. F. and DIETERT, H. W., *J. Opt. Soc. Am.*, **33**, 218 (1943)
2. HASLER, M. F. and KEMP, J. W., *J. Opt. Soc. Am.*, **34**, 21 (1944)
3. MIKA, J. and PÉTER, L., *Öntöde*, 53 (1958) (in Hungarian)
4. BARDÓCZ, Á., *Acta Tech. Hung.*, **13**, 3 (1955)

2.9.3 Mains-voltage spark generators

In the case of mains-voltage, or low-voltage,[1] spark generation the terminal voltage of the condenser is the same as the voltage of the mains current. The discharge of the condenser is, therefore, made possible by using a high-frequency, high-voltage, ignition current (Tesla current) as in mains-voltage a.c. arc generators. The circuit design of the usual equipment is different from that of arc generators in that their working current circuit contains a condenser of larger capacitance (for example 200–2000 μF). The transfer of the Tesla current is effected by using the Pfeilsticker[1] or Calker[2] principle. The circuitry of a mains-voltage spark generator of the former type is shown in Fig. 2.103. The circuit design for the discharge circuit and for the transfer of the ignition current is depicted in Fig. 2.104.

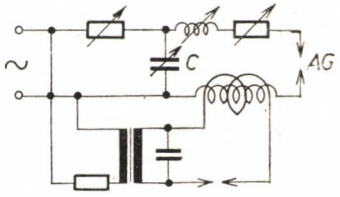


Fig. 2.103 Circuit diagram of the mains-voltage Pfeilsticker-ignition-system spark generator

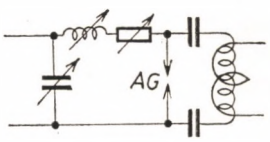


Fig. 2.104 Discharge circuit diagram of the mains-voltage spark generator and the Calker-type transfer of the ignition current

For both Pfeilsticker- and Calker-type ignition the inductance of the discharge circuit of the working condenser cannot be smaller than 25 μH. This level of inductance is represented in the former case by the secondary winding of the Tesla transformer. In the latter case, it is the smallest value of the switched-on inductance that avoids the short-circuiting of the Tesla current through the working condenser and thus secures the breakdown of the analytical gap. This constrained inductance, however, reduces the value of the discharge current. Provided it is possible to reduce the inductance of the discharging circuit right down to its remanent level of about 1 μH, then, when $R \approx 0$, the value of the maximum current would be increased fivefold according to Eq. (2.79).

The discharge of the condenser at remanent induction level can be achieved by the ignition method shown in Fig. 2.105.[3] This method differs from the Calker-system (Fig. 2.104) in that the ignition current passes through two secondary spark gaps to the analytical gap, instead of reaching the latter directly. This represents a significant difference from the operational point of view. The Tesla voltage difference, U_{Te}, brought about across the analytical spark gap is given by

$$U_{Te} = L_0 \frac{dI_{Te}}{dt} \tag{2.156}$$

where L_0 is the remanent induction of the discharge circuit without ohmic resistance and switched-on inductance, and I_{Te} is the value of the Tesla current in the circuit. When the electrodes of the secondary spark gap are separated such that the Tesla spark is just able to break it down, then the sudden current change results in high values for dI_{Te}/dt. Thus, the product value in Eq. (2.156) can reach the analytical gap breakdown voltage level even in the case of very small L_0 values.

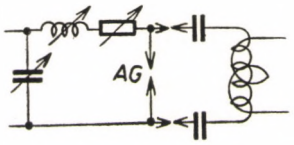


Fig. 2.105 Schematic diagram for the transfer of the ignition current by a secondary spark gap

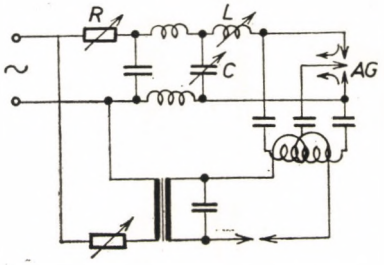


Fig. 2.106 Circuit diagram of the modified mains-voltage spark generator[4]

The above conditions can also be realised in the so-called *triple secondary spark gap*, mains-voltage spark generator.[4] The characteristic of this system is that the secondary winding of the Tesla transformer consists of two series-connected, oppositely wound coils (Fig. 2.106). It can be achieved by this method that the two ends of the secondary winding always have the same polarity, which is the opposite of the polarity of the tapped centre of the winding. Thus, the Tesla current passes through the spark gap at a certain phase instant in the direction shown by the arrows in Fig. 2.106. The ignition current, therefore, effectively avoids the generator condenser.

It can be seen that in the application of the Tesla ignition current the condenser of the mains-voltage spark generator is discharged – independently of the method of transferring the ignition current – at the phase instant when the condenser terminal voltage of the Tesla current generator reaches the breakdown voltage of the controlling spark gap. The *phase instant of the condenser discharge* of mains-voltage spark generators is determined by the phase relations of the terminal voltage of the condenser belonging to the Tesla-generator and by the breakdown voltage of the control-

ling spark gap. The phase instant of the discharge of the working condenser can, therefore, be controlled by the resistor in the charging circuit of the Tesla-generator and by the electrode distance of the controlling spark gap inserted into the discharge circuit. The frequency of the condenser discharges in a period depends on the capacitance of the condenser of the spark generator and the equivalent value of the damping resistance in the charging circuit. When these values are chosen such that the controlling spark gap is broken down only at extreme voltage values of the spark generator condenser, then discharges of type 1 + 1 are produced. Discharges of type 2 + 2 can be obtained, however, if the electrode distance of the controlling spark gap is reduced.

Only one condenser discharge is produced in a period if the relative charging current of the condenser of the spark generator (that is, the relation between the current value of the charging circuit and the capacitance of the condenser) has a value previously established experimentally.[5] By using a higher relative charging current value, or by the appropriate reduction of the damping resistance of the spark generator, discharges of type 2 + 2 can also be produced.

In the case of mains-voltage spark generators controlled in the way described above, condenser discharges of type 0 + 1 can theoretically be produced only when an asymmetrical controlling spark gap is used. One electrode is made of a poor heat conducting material (for example, carbon) and is sharpened while the other electrode is made of some metal that has good thermal conductivity and has a large and smooth tip.[6] At an appropriately large electrode distance the discharge only occurs in the half-period when the sharpened electrode of poor thermal conductivity is the cathode. The latter remains incandescent during the interval between the discharges. (The emission of electrons is easier from the hot sharpened electrode than from the cold one with a large flat surface.)

The previous method used for controlling the operation of mains-voltage spark generators by controlling the spark gap is not entirely predictable. Difficulties can

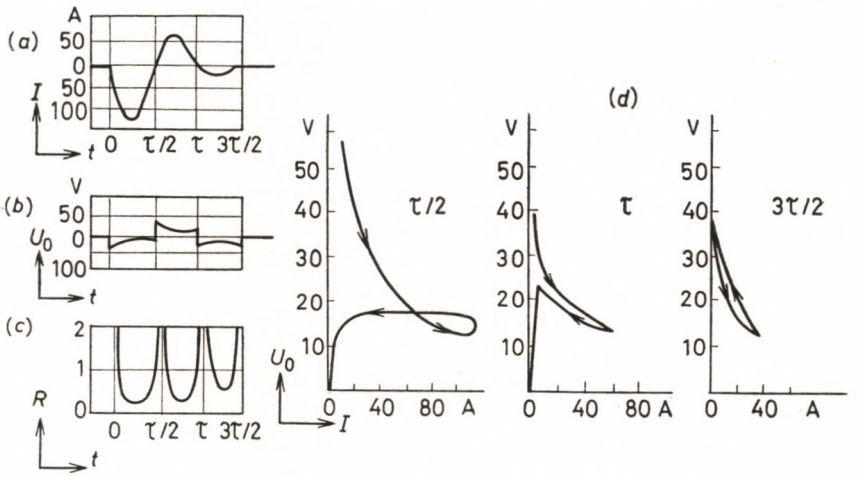


Fig. 2.107 Electrical characteristics of mains-voltage spark generation for supplying a large capacitance and a large inductance: (*a*) current–time curve, (*b*) voltage–time curve, (*c*) resistance–time curve, and (*d*) voltage–current curve (dynamic characteristics)[8]

be largely eliminated if the controlling spark gap is replaced by a rotating switch driven by a synchronous motor, that is, when the ignition circuit of the generator is being mechanically controlled.[7] This type of control of mains-voltage spark generator is the same as the method used for controlling the multi-source unit, medium-voltage spark generators (Fig. 2.100). A further advantage of this system is that discharges of type 0 + 1 can be produced with essentially greater certainty than in the case of the asymmetrical controlling spark gap.

It is seen from the diagrams given in Fig. 2.107 that the duration of the condenser discharges becomes much too long in mains-voltage generators when high damping is applied. In such cases it is absolutely necessary to control the charging circuit as well.[3] This can be achieved most effectively by a rectifying tube or a silicon diode. The rectification of the mains current has the consequence that discharges of type 0 + 1 are produced. Thus, the condenser discharges will be entirely free of disturbances since they last over the next half-period.

For mains-operated spark generators it is advisable to have some form of electronic control of the discharge. The control principle of such generators is identical to that of electronically controlled a.c. arc generators.

From a more detailed study of mains-voltage spark generation the following statements are relevant:[8]

(1) When the capacitance of the condenser and the inductance of the discharge circuit are relatively large, then during the first two periods of the condenser discharge radiation is emitted whose spectral character is the same as that obtained by very high current arc generation.
(2) In spectra obtained under the previously described experimental conditions, strong continua are observed with more or less broadened emission spectral lines superimposed. On the other hand, the resonant lines of low-boiling-point elements appear as *absorption lines*. This has the consequence that many spectral lines of elements present at high concentration (mainly their fundamental lines) cannot be used for spectrochemical analysis.
(3) The reproducibility of the spectra is good. In spectra obtained by a single condenser discharge the average scatter of the *absolute* intensity of the lines is about $\pm 10\%$. Thus, the scatter of the *relative* intensity of the spectral lines does not exceed the allowable level, even when it has been derived from the mean of relatively few condenser discharges.
(4) The above feature can be partly accounted for by the formation of a very thin molten layer on the tips of the metal electrodes that is pushed away by the shock wave created during the spark breakdown. In this manner the surface is continually renewed. Thus, the disturbances caused by the chemical and other changes occurring on the front electrode surface are eliminated.
(5) The sensitivity of mains-voltage spark generation is about the same as that for arc generation, provided the experimental factors influencing the spectral character are properly chosen.
(6) In the case of excitations at reduced atmospheric pressure the broadening of the spectral lines and the disturbing effect of the background are greatly reduced.
(7) This can be partly explained by the previously mentioned phenomenon that *at about 10* mm Hg *pressure the non-metallic elements that are difficult to excite* (Cl, Br, I, S, Se, H, N) *can be detected and analysed.*[1,9]

When the performance of high-, medium-, and mains-voltage spark generator is comparatively evaluated, it can be established that the initial, maximum current produced at the discharge of the condenser is the smallest in the case of mains-voltage spark generation. At the same time the number of recharges occurring during individual condenser discharges is also at a minimum (see Table 2.9). The damping time of the discharges is relatively long and can reach or even exceed the burning times encountered in a.c. arc generation. The discharge current, however, is greater by several orders of magnitude than the usual current level of the arcs. From this fact we can conclude that the mains-voltage spark generation technique is an intermediate method between the other two excitation methods.

References

1. PFEILSTICKER, K., *Spectrochim. Acta*, **1**, 424 (1940); *Metallk. GiessWes.*, **33**, 267 (1941)
2. VAN CALKER, J., and MAAS, W., *Z. angew. Phys.*, **4**, 161 (1952)
3. VAN CALKER, J., *Spectrochim. Acta*, **5**, 19 (1952)
4. TÖRÖK, T., *Spectrochim. Acta*, **6**, 228 (1954)
5. MIKA, J., MACHER, F., and VORSATZ, B., *Analytica Chim. Acta*, **3**, 228 (1949)
6. TÖRÖK, T., *Acta Tech. Hung.*, **4**, 429 (1953)
7. TÖRÖK, T., *Explle Tech. Phys.*, **7**, 49 (1954)
8. VAN CALKER, J., *Spectrochim. Acta*, **11**, 260 (1957)
9. PFEILSTICKER, K., *Mikrochim. Acta*, **1955**, 358

2.10 Excitation in discharge tubes

2.10.1 Fundamentals of the method[1–7]

Discharge tube excitation methods are characterised by the feature that the charged particles of the test material gain their kinetic energy for atom excitation almost exclusively from the direct effect of the electrical energy instead of from the thermal mechanism. A further feature of this type of excitation is that it can only be performed in an enclosed environment – the discharge tube.

Discharge tubes are evacuated, cylindrical or variously shaped glass or sometimes quartz vessels that contain two electrodes. The use of these devices is based on the phenomenon that free electrons and ions, always present in gases due to radioactive radiation, etc., are accelerated by an electric field and migrate towards the electrodes when there is a sufficiently large voltage difference between them in a gas-filled tube. The electrons, travelling at a high speed owing to their small mass, collide with the gas atoms and ionise them. The free electrons created in this way, in turn ionise further gas atoms, provided their mean free path and the electrical field strength are of an appropriately high value. Thus, the number of charged particles is increased rapidly in the gas volume between the electrodes.

Among the charged particles the mobility of the electrons is much greater than that of the ions, which have a comparatively large mass. Thus the electrons reach the anode quicker than the positively charged ions travelling to the cathode. Between the electrodes, therefore, a positive space charge is created that produces a voltage drop that increases towards the cathode. Accordingly, the space in front of the cathode not only promotes ionisation, but the ions within it are also highly accelerated. Thus, they attain kinetic energy and upon collision with the cathode numerous

electrons are released. Under such circumstances current conductivity is mainly obtained by the newly created particles and not by those originally present. When this happens the discharge tube is *ignited*, that is, the original quiet discharge in the tube is transformed into a self-sustaining glimm discharge. The necessary voltage difference is called the *ignition voltage*.

The value of the ignition voltage depends on the discharge tube form and dimensions, as well as the type and pressure of the gas. The electrode material, in particular, that of the cathode is also important from this aspect. For the case of simple cylindrical discharge tubes having parallel electrodes, Paschen's law applies. According to this law the ignition voltages of two discharge tubes filled with the same gas are identical when the product of the gas pressure and the electrode distance is the same for both tubes.

The charged particles necessary for current conductance in glimm discharges are created partly by collision ionisation and partly by the collision of the ions with the cathode. These processes occur quite frequently, provided the pressure of the gas is not too high. However, the mean free path of the charged particles is so short that any increase in their kinetic energy that the collisions produce, cannot reach the level necessary for the initiation of further ionisation even at high field strength values. Under these conditions a glimm discharge is not developed even at great voltage differences, as a spark discharge will occur only when the potential difference reaches the value of the breakdown voltage. Therefore, the prerequisite of glimm discharges is that the pressure of the filling gas should be appropriately low – in fact, it should be lower than 100 mm Hg but greater than 0·001 mmHg. Furthermore, it is necessary that the voltage difference between the two electrodes should be sufficiently high. (The value of the latter should be between a hundred and a few thousand volts, depending on the dimensions of the tube.)

It is characteristic of glimm discharges that:

(1) There is always light emission.
(2) The cathode of the discharge tube is not usually warmed up.
(3) The spectrum of the emitted light depends solely on the chemical composition of the filling gas and not on the material of the electrodes. (This is because electrodes are not heated and so they cannot evaporate.)

It can be shown experimentally that the current–voltage curve of the discharge tube is not linear (Fig. 2.108). The value of the current, I, varies in a linear manner with the terminal voltage, U, in the initial stages, although when it reaches the ignition voltage, U_i, the current remains at a constant level as the voltage is increased further. The fact that the current becomes independent of the voltage at this stage can be explained as follows. *Before* the attainment of the ignition voltage the charged particles are increasingly replenished owing to the effect of the increasing voltage, as well as the number of the ionised atoms. Thus as the number of charged particles increases this causes a corresponding increase in the current. When the ignition voltage is attained, a state of saturation is realised since, in spite of the voltage increase, the number of atoms ionised per unit time is the same as that formed by recombinations with electrons.

When the voltage in the discharge tube is further increased a level is reached at which the current again starts rising steeply. At this terminal voltage the ions of

charge Q assume a kinetic energy of QU owing to the effect of the voltage difference U which is equal to the ionisation energy of the neutral particles. As a result of ionisation the number of collisions and thereby the number of charged particles rapidly increases, so that the current passing through the discharge tube also increases steeply. Simultaneously, the ions possessing the appropriate amount of kinetic energy collide with increasing frequency with neutral molecules and ions, thereby dissociating or exciting them. Upon increasing the voltage further the ion-lines of the gas present appear in the spectrum of the emitted light. When, however, the current increases the cathode is progressively heated to higher temperatures. This causes an increase in the number of spectral lines of the elements constituting the electrode material to appear in the spectrum of the emitted light. Thermal emission will, therefore, assume an increasing role in the electrical conductance, such that the discharge gradually transforms into an arc.

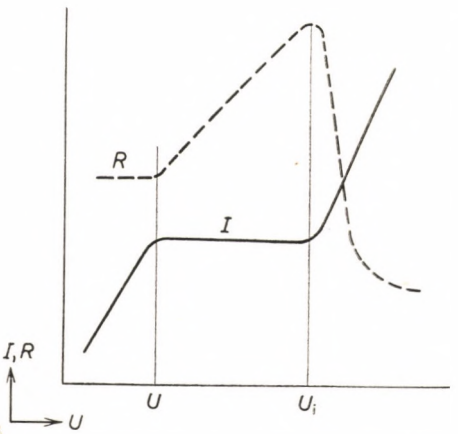


Fig. 2.108 Schematic current–voltage (solid line) and resistance–voltage (dotted line) curves of discharge tubes

As seen in Fig. 2.108 the resistance of the discharge tube does not change as the voltage across the tube increases until the ignition voltage level is reached. In the saturation region the resistance of the tube is greatly increased. At this stage of the discharge the current passing through the tube is independent of the voltage. The resistance of the discharge tube, however, drops very quickly when the increasing terminal voltage of the tube reaches the ionisation potential, U_i, of the gas present. The current, therefore, cannot be controlled without the use of an appropriately chosen ohmic resistance, connected in series. Without such a resistor the discharge is not stable and the gas discharge will be transformed into an arc that can damage the tube.

It is clear from the previous discussion that the processes taking place in the discharge tube are influenced by the value of the free path of the particles and indirectly by the pressure of the gas as well. When an appropriately high voltage (a few kV) is applied to the electrodes of the discharge tube and the pressure is reduced to about 50 torr, a gas discharge is established between the electrodes of the tube which is somewhat similar to a spark discharge. Initially a waving light beam appears between the electrodes. When the pressure is decreased further the light becomes extended until, at about 10 torr pressure, it occupies the whole diameter of the discharge tube. At the same time dark zones appear along the axis of the tube (Fig. 2.109). The positive light column appearing in front of the anode is broken into successive dark

and light layers as the gas pressure is further reduced until, at about 1 torr pressure, it becomes finely laminated. When the pressure is still further reduced the number of layers at first decreases and then the light column slowly contracts. Simultaneously, the glowing cathode space occurring in front of the cathode and the Faraday dark field are extended. Finally at a pressure of about 10^{-3}–10^{-4} torr, light emission from the gas is practically terminated.

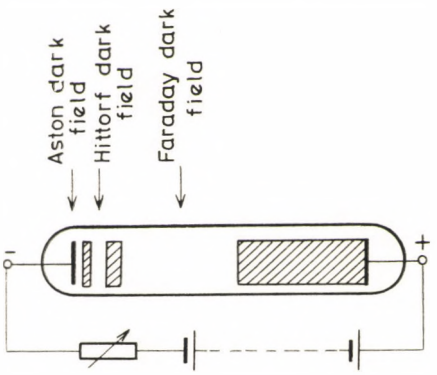


Fig. 2.109 Stratification arising in discharge tubes at a gas pressure of about 10 torr

Among the light emission phenomena occurring in the discharge tube are the following features:

(1) The so-called *Aston dark field*, which is completely free of light emission, does not appear for large cathode drops, nor in the case of molecular gases.
(2) The intensity of the '*luminescent cathode layer*' is very low.
(3) The *Hittorf* (or *Crookes*) *dark field* is developed next between two neighbouring luminescent layers.
(4) The *negative glow* is one of the most intensely emitting parts of the cathode space of the discharge tube and is terminated by a relatively sharp *glow-blanket* towards the cathode end.
(5) The light intensity of the *Faraday dark space* is very small. Sometimes the part facing the anode merges with the positive column.
(6) The *positive column* is the most intensively emitting part of the whole discharge tube. This column does not extend right up to the anode when the gas pressure is about 0·5 mmHg, but is separated from the electrode by a dark layer and a thin *positive glow layer* that is attached to the anode (Fig. 2.110*a*).

It follows from the various features listed above that among the different layers of the discharge tube the positive column and the negative glow can be utilised for spectrochemical analysis.

The *voltage drop* between the two electrodes of the discharge tube is not uniform (Fig. 2.110*b*). The greatest voltage drop, U_c, or the so-called *cathode fall*, is observed between the cathode and the negative glow. A significant, but smaller voltage drop is the anode fall, U_a. On the other hand, the smallest variation in the voltage occurs in the Faraday dark space. The cathode fall can be explained by the positive charge in the cathode space. This is due to the mobility of the electrons being greater than that of the ions, such that the positively charged ions accumulate within the cathode

space. On the other hand, the anode fall is explained by the negative space charge emerging in front of the anode (Fig. 2.110*c*).

The so-called *normal cathode fall* arises when the value of the current passing through the discharge tube is comparatively small. It is characteristic of the latter that the negative glow does not completely cover the cathode. The cathode fall is independent of the value of the current and the pressure of the gas, although it does depend on the electrode material and the chemical composition of the gas. The degree of the normal cathode fall is smallest when the work function of the electrode material is lowest and the gas is most electropositive.

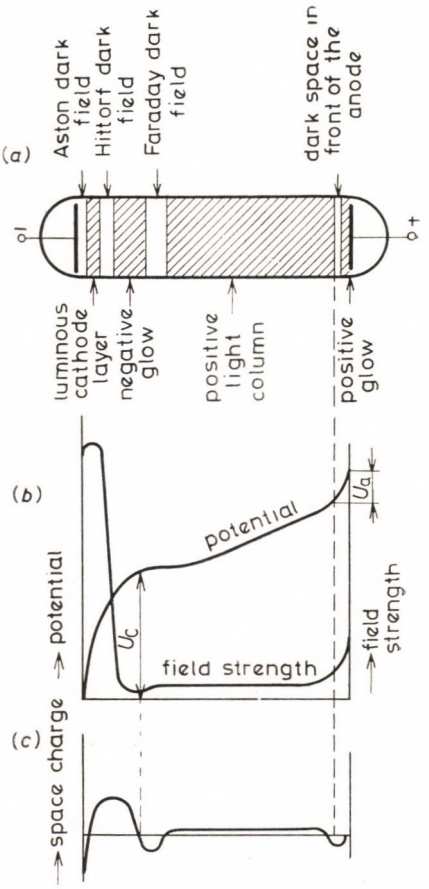


Fig. 2.110 The discharge tube: (*a*) stratification of the light phenomena at about 0·5 torr pressure, (*b*) distribution of the voltage drop and field strength, and (*c*) charge distribution in space[2]

By an approximative increase of the current or by reducing the gas pressure it is possible to make the negative glow cover a progressively larger area of the cathode. In the case of an anomalous cathode fall the glow covers the whole cathode, although the *Hittorf* dark field is extended at the same time. The anomalous cathode fall arising in this way is greater than the normal cathode fall, but in contrast to the latter it is dependent on the current and the gas pressure. When an anomalous cathode fall arises due to the reduction of the gas pressure, a field strength is created which requires several kilovolts to sustain the discharge.

The space charge is not evenly distributed in the discharge tube. The largest negative space charge is developed right in front of the cathode, while the lowest space charge is found in the positive glow. At the border of the Aston dark field the negative space charge is transformed into a relatively large *positive* space charge whose maximum lies at the border between the Hittorf dark field and the negative glow. The latter, however, is totally extinguished in the Faraday dark field. In the positive column a weak positive space charge is established that does not influence the processes that occur there.

From the above discussion it can be seen that the electrons ejected from the cathode are accelerated to such an extent, due to the effect of the high field strength prevailing in the negative glow area, that the excitation within the discharge tube is mainly caused by electronic collisions. It can also be observed that in the negative glow area the electrons coming from the cathode direction, which have a relatively large kinetic energy, ionise atoms and excite ions in this region by collision. In the region of the negative glow a radiation emerges whose spectrum mainly consists of lines of positive ions. The spectrum is, therefore, of ionic character.

The kinetic energy of the electrons in the positive column zone is generally smaller than the energy of those electrons that produce light emission from the negative glow. In the positive column, therefore, neutral atoms and molecules are mainly brought into the excited state. Accordingly, the spectrum of the radiation from the positive column is predominantly composed of atomic lines and bands. The spectrum has an atomic character, or in other words it is similar to spectra obtained by arc excitation, except that the number of lines is greater. At a relatively low current density the spectral lines are very sharp. Since excitation is also caused by electronic collisions in the positive column, an increase in the strength of the electric field promotes the excitation of spectral lines of higher excitation energy. The same result can be achieved by the reduction of the gas pressure, since in this way the free path value is increased.

References

1. FINKELNBURG, W. and MAECKER, H., *Elektrische Bögen und thermisches Plasma* in *Handbuch der Experimentalphysik*, vol. XXII Gasentladungen, **2**, Berlin–Göttingen–Heidelberg (1956)
2. KOHLRAUSCH, F., *Praktische Physik*, Leipzig, vol. 2, 350 (1951)
3. JOOS, G., *Lehrbuch der theoretischen Physik*, Leipzig (1950)
4. BOCHKOVA, O. P. and STREIDER, YE. YA., *Spektralbnyi Analys Razovyk Smesei*, Moscow (1955)
5. MEEK, J. M. and CRAGGS, J. D., *Electrical Breakdown of Gases*, Oxford, Clarendon Press, (1953)
6. LOEB, L. B., *Fundamental Processes of Electrical Discharge in Gases*, New York, (1947)
7. THOMSON, J. J. and THOMSON, G. P., *Conduction of Electricity through Gases*, Cambridge, (1928)

2.10.2 Excitation in Geissler tubes[1,2]

Geissler tubes are special types of discharge tubes that allow us to analyse gaseous mixtures spectrochemically. The excitation of gaseous elements cannot be performed by the previously mentioned arc and spark excitation methods. This is largely because the excitation potential of gases is higher than most conducting materials that are

difficult to excite. For example, the difference between the excitation potential of the most easily excitable elementary gas, bromine, and the metallic element, most difficult to excite, tellurium, is: 7·82−5·45 = 2·37 V! Therefore, only the vapour of the sample or electrode material is usually excited by arc and spark generation methods. This difficulty does not, however, occur if the gases are excited at low pressures in discharge tubes.

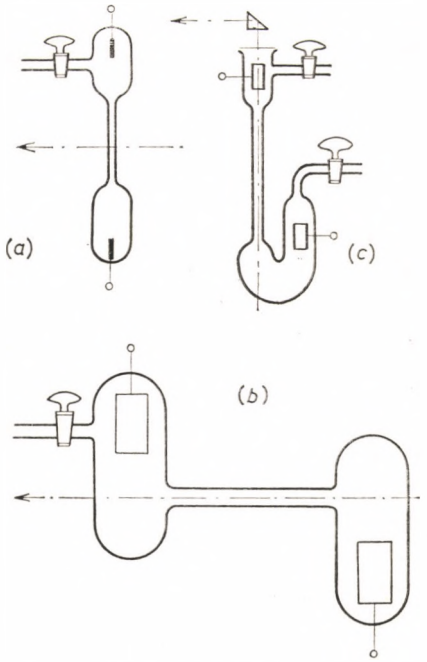


Fig. 2.111 Geissler tubes used for gas analysis

Geissler tubes used for gas excitation are essentially made of two cylindrical vessels connected by a capillary having a length of 20–60 mm and an internal diameter of 2–5 mm (Fig. 2.111). The electrodes are rods made of platinum or molybdenum welded into the glass. Geissler tubes used for the examination of the ultra-violet spectral region are made completely of quartz, although in general they are composed of hard glass. The shape of common Geissler tubes depends partly on whether the radiation is examined along the axis of the capillary or perpendicular to it, or whether the gas under study flows through the apparatus or is only used as a filler.

In the capillary of the Geissler tube a positive column radiation emerges when the pressure of the gas varies between 1 and 10 torr. The intensity of the emitted light is increased if the plasma is strongly concentrated and restricted by the narrow bore of the capillary. Thus, the current density and the number of excited atoms in unit volume are increased.

It is an advantage, from the point of view of intensity enhancement, if the radiation passing along the axis of the capillary is utilised. However, the main drawback of this method is that self-absorption effects must be taken into account. On the other hand, self-absorption does not disturb the resonant lines when the capillary of the Geissler tube is in a position perpendicular to the optical axis of the spectrograph. This arrangement has the further advantage that Doppler broadening of the spectral

lines is also slight, since the thickness of the gas layer is relatively small. However, one disadvantage of this method is that the intensity of the radiation in this direction is essentially smaller than that of the radiation in the direction of the capillary axis.

It can be shown experimentally that the spectral lines are sharper when a small current is passed through the Geissler tube, or to be more exact, the smaller the current density in the capillary. In the opposite case, one has to consider the Stark and iintrnal Stark broadening of the spectral lines. In choosing the most appropriate value of the current density it should be taken into account that upon reducing the current the intensity of the spectral lines is drastically lowered. Thus the value of the optimum current depends on the shape and dimensions of the Geissler tube. (The optimum current is usually around 1 mA, while its maximum value is generally about 50 mA; the required voltage is 1–2 kV.)

Generally, it is advisable to use a Geissler tube in which the sample gas (mixed with an appropriate carrier gas) is forced through at about 1 mmHg pressure. Thus, memory effects caused by foreign gases adsorbed on to the internal surface of the Geissler tube are eliminated. By the appropriate choice of the carrier gas, errors caused by self-absorption can be largely reduced. The carrier gas dilutes the sample and thus reduces the probability of excited and non-excited species of the same kind of atom colliding.

In the choice of the carrier gas it is essential that its presence should not lead to interferences in the analysis. It is, therefore, an advantage if the excitation potential of the carrier gas is higher than that of the gas to be determined. When this difference is sufficiently great it is possible for the carrier gas in the positive column of the discharge tube not to be excited at all. In this context, helium or occasionally argon is generally used as a carrier gas (Table 2.13). When helium is used the spectrum of the sample gas has an ionic character, while if the carrier gas is argon the character of the spectrum is much more atomic. This can be explained by remembering that in gas-

Table 2.13

Lowest excitation potential of gaseous elements, the wavelength of the corresponding spectral line, and the ionisation potential

Element	Fundamental line		Ionisation potential (V)
	Excitation potential (V)	Wavelength (Å)	
Br	7·82	1 574·9	11·84
Xe	8·27	1 469·6	12·13
Cl	8·90	1 389·2	13·01
O	9·11	1 302·0	13·61
Kr	9·86	1 235·8	13·99
H	10·15	1 215·7	13·60
N	10·29	1 135·0	14·54
Ar	11·56	1 066·7	15·75
F	12·66	954·8	17·42
Ne	16·50	735·9	21·56
He	20·86	584·3	24·58

discharge tubes the excitation is caused by secondary collisions. Under such conditions the character of the spectrum is influenced by the ionisation potential of the carrier gas.

Geissler tubes can be used, besides their applications for the analysis of gases, for the examination of low-boiling-point compounds, powdered solid organic compounds, etc., when appropriate modifications are carried out. Moreover, they can also be used for the continuous analysis of gases, although only gases whose partial pressures are comparatively high can be excited in the gas mixture under study.

The power supply of Geissler tubes is usually a.c. or d.c. of mains frequency, transformed to 1·5 kV. It is important to note that they cannot be used without an appropriate ohmic resistance connected in series.

References

1. WIEN, W., and HARMS, F., *Handbuch der Experimentalphysik*, vol. 21 Leipzig (1927)
2. JOOS, G., and ANGERER, E., *Spektroskopie*, Leipzig (1927)

2.10.3 Excitation by hollow-cathode tubes

The hollow-cathode discharge tube is a specially constructed discharge tube that has a particular type of cathode.[1–3] The most important feature of the apparatus is a closed space filled by a noble gas, at a pressure of 0·1–10 mmHg, with a cathode having a cylindrical cavity adjacent to a ring or rod-shaped anode (Fig. 2.112). The d.c. terminal voltage between the two electrodes is generally 1·5–2 kV. The optimum current value, controlled by a series resistance, is usually 0·2–0·5 A or very rarely 2–3 A. The breakdown voltage of the electrode gap is between 250 and 400 V, depending on the construction of the apparatus. (The relatively high terminal voltage of the current supply is due to the fact that the stability of the discharge is improved by a relatively large, series-connected ohmic resistance.) Non-gaseous samples are placed within the hollow-cathode cavity for excitation purposes.

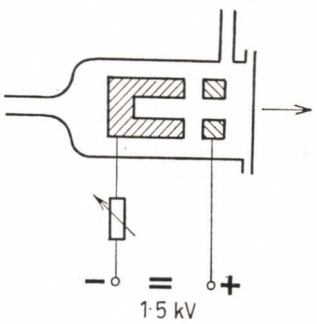


Fig. 2.112 Diagram of a hollow-cathode discharge tube

Various light phenomena due to the d.c. voltage can be observed in the cavity of the cathode. Along the wall of the cavity a barely visible '*luminous cathode layer*' is created, while almost the whole diameter of the cavity is filled with a strongly emitting *negative glow*. Between the previous two features, the *Hittorf dark space* is situated (Fig. 2.113). The other light phenomena seen in the main forms of discharge tubes cannot be found in the case of the hollow cathode. This fact is explained by

the small distance between the two electrodes. In the hollow cathode the light source proper is the negative glow that has an almost homogeneous intensity distribution.

The excitation mechanism in the hollow cathode is as follows. The electrons ejected at right angles to the surface of the cathode are strongly accelerated due to the great potential gradient in the Hittorf dark field. They release their kinetic energy in the almost-zero field-strength region of the negative glow in such a way that collisions

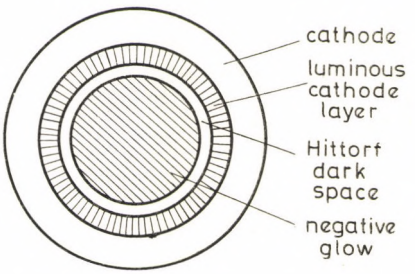


Fig. 2.113 Light phenomena arising in a hollow-cathode[3]

with water vapour atoms lead to their ionisation. The secondary electrons created move towards the anode and during their travel produce further ionisation by collision. The number of charged particles increases until an equilibrium state is reached and the number of simultaneously created and recombined ions becomes equal. The electrons moving towards the anode are ordered by the field distribution into a narrow beam and leave the cavity of the cathode in this form.[3]

The positively charged carrier gas ions do not leave the electrode cavity, in contrast to the electrons, but are accelerated during their travel to the cathode under the effect of the high field strength. Gas ions of comparatively large mass thus attain a significant amount of kinetic energy. This energy is dissipated in the collision with the cathode such that the heat produced by the collision heats up the cathode. Another part of the kinetic energy of the gas ions brings about cathode sputtering effects.

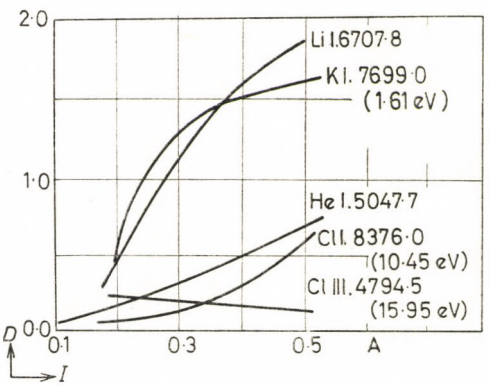


Fig. 2.114 Variation of the density of the spectral lines (D) as a function of the current value (I) for a 1 torr pressure of helium[8]

Thus, owing to the effect of these two processes the solid sample placed within the cavity of the cathode is continuously evaporated or sputtered off. Carrier gas ions of high energy excite the atoms of the sample brought into the gaseous phase by the collisions. Thus, the whole excitation process happens within the cavity of the hollow cathode. The radiation emerging in this process leaves the apparatus

along the axis of the ring-shaped anode and passes through a quartz or glass window with parallel surfaces placed opposite the cathode.[3–5]

When the current is increased the intensity of the emission is also generally increased (Fig. 2.114). There is a limit set to this increase by the fact that, when the current density is too high in the work area of the cathode an anomalous cathode fall is created. The highest allowable current density depends on the quality and pressure of the carrier gas as well as on the dimensions of the cathode cavity, etc. This minor disadvantage can be easily eliminated by using hollow cathodes in which the length of the cathode cavity is variable.[6,7]

The intensity of the light emitted by the hollow cathode is also influenced by the *gas pressure*. The ideal pressure value is about 1–5 mmHg (Fig. 2.115), although this value is slightly different for various shapes and dimensions of the hollow cathode.

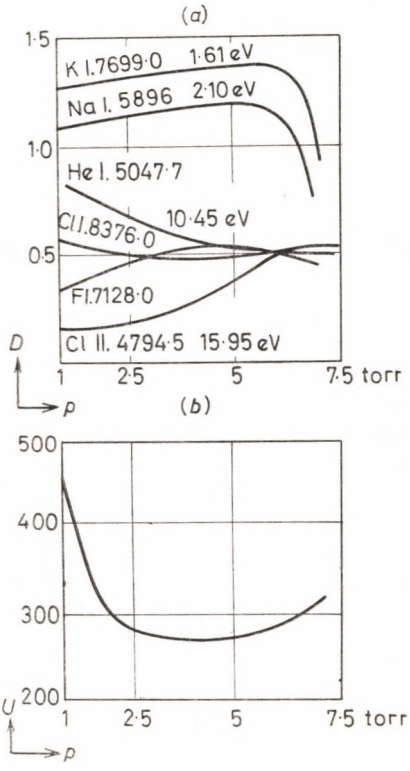


Fig. 2.115 Variation of the density (D) of different spectral lines (*a*) and of the voltage (U) between the electrodes (*b*) as a function of the pressure of the helium carrier gas[8]

The lowest gas pressure at which the negative glow vanishes and only a very poorly emitting, positive glow can be seen also depends on the experimental conditions. This phenomenon usually occurs at about 0·2 torr.[3]

In choosing the carrier gas one should first take into account that the excitation is produced by secondary collisions. This has the consequence that very strong bands appear in the spectra when the gas is diatomic (for example, hydrogen, nitrogen, oxygen, etc.). It is, therefore, a definite advantage to use noble gases, as only the spectral lines of the particular noble gas appear in the spectra. In the selection of the carrier gas it should also be taken into account that the character of the spectrum is affected by the ionisation potential of the carrier gas. In general, the higher the

ionisation potential of the carrier gas the more ionic in character will be the spectrum. This is explained by the fact that the higher the ionisation potential of the element in question, the greater is the excitation energy of the carrier gas atoms that produce the excitation. Accordingly, if helium is the carrier gas, a spectrum of more ionic type is produced, while argon gives a spectrum of atomic character.[8] This rule is not generally valid. An exception is xenon, which in spite of its low ionisation potential has a spectrum of more ionic character than argon.[9]

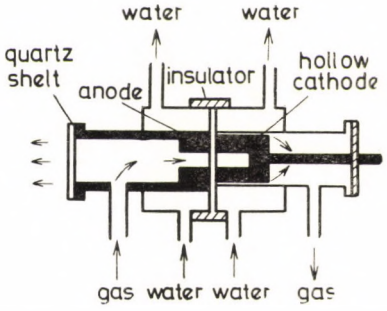


Fig. 2.116 Diagram of the cooled hollow cathode[5]

It is evident from the above discussion that a considerable amount of heat is liberated during the excitation in the cavity of the hollow cathode. This will heat up the cathode unless appropriate cooling is provided. A distinction can therefore be made between the so-called cold (cooled) hollow cathode and the so-called hot (non-cooled) hollow cathode.

In the case of *cooled-cathode excitation* both the cathode and the anode are cooled by circulating water (Fig. 2.116). This technique has two advantages. The first is that the electrodes can be manufactured from materials with good electrical and thermal conductivity, for example copper, aluminium, nickel, stainless steel, or occasionally graphite. The second advantage is that there is no evaporation of the cathode material during the excitation, even when the melting point of the cathode is relatively low.

In the cooled hollow-cathode excitation method the ions having the greatest kinetic energy produce by collision a sputtering of the material in the cathode cavity. The 'sputtered' material, therefore, reaches the excitation space in a very finely divided form. By means of cathode sputtering it is possible to study solid samples by using cooled hollow-cathode excitation. The excitation process proper then follows the formerly mentioned course.

It can be shown experimentally that in the case of cooled hollow-cathode excitation, the intensity of the emitted radiation is comparatively low. Only the spectral lines of relatively low excitation potential appear in the spectra, although they are generally very sharp. No self-absorption occurs and practically no background is observed in the spectra.[5]

In the application of the non-cooled hollow-cathode method the cathode becomes intensely heated during the excitation, such that its evaporation must be taken into consideration.[10,13,14] Due to cathode sputtering the amount of cathode material reaching the vapour space is increased by the evaporation. Elements in the sample and cathode material, therefore, appear together in the spectrum obtained. The evaporation of the cathode material is reduced by lowering the current or by increasing the gas pressure. In the case of the former method, the plasma temperature and

thus the intensity of the emission are reduced, while the character of the spectrum is shifted towards the atomic form. It is an advantage, therefore, if the cathode is made of a material of high thermal conductivity and possibly high melting temperature; for example, chromium steel.

The non-cooled hollow-cathode excitation method effectively produces a vacuum arc. The emission is mainly produced by the thermal electrons. Thus, the temperature of the plasma could drop to as low as 2000 K.[11] The concentration of the atoms in the plasma is relatively large since they are brought into the gaseous phase by thermal means and also by cathode sputtering. Accordingly, the radiation emitted by the non-cooled hollow cathode is generally very intense. Lines appear in the spectra that would not be observed if a cooled hollow cathode were used. In principle, therefore, the sensitivity of detection is significantly increased. Practically, however, this is not generally seen to be the case, as when the cathode temperature is increased the intensity and extension of the continuous spectrum also increase. The intensity of the background is, therefore, relatively high in non-cooled hollow-cathode excitation. On the spectral lines known to exhibit self-absorption or self-reversal, both of these phenomena are often shown.

The drawbacks of the above t wo methods are largely eliminated when a so-called transitional-type hollow cathode is used.[11] Here the sample is placed into a *sample holder* that is closed at one end and is made of a material with good thermal and electrical conductivity. A large area of the outer surface of the sample holder is in contact with the cooled cathode (Fig. 2.117). Thus, by varying the degree of surface contact (that is, the form of the sample holder), the temperature can be varied between wide limits. For example, it could be reduced to as low a temperature as 1500 °C. The temperature of the material to be studied can, therefore, be chosen so as to produce the most advantageous evaporation conditions. A further advantage of this method is that the spectral lines obtained are sharp, self-absorption is practically negligible, and the spectral background has a low intensity. By the correct choice of the experimen tal conditions it is possible to obtain the spectral lines of high-excitation-potential, non-metallic elements in the spectrum.[11]

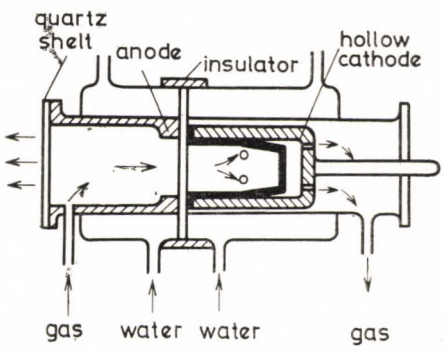


Fig. 2.117 Diagram of the transitional-type hollow cathode[11]

Owing to cathode sputtering in the hollow-cathode excitation method, solid material continuously contaminates the gas phase. This can lead eventually to the deposition of solid material by the carrier gas flowing through the cavity. It is therefore advisable to apply some form of *circulation system*, as well as to purify the gas

leaving the system by a freezing-out technique and if necessary to remove any hydrogen produced by preliminary burning (Fig. 2.118).

It is also essential that the gas, which enters along the axis of the cathode, is removed from the cathode cavity through the radial bores (Fig. 2.119).

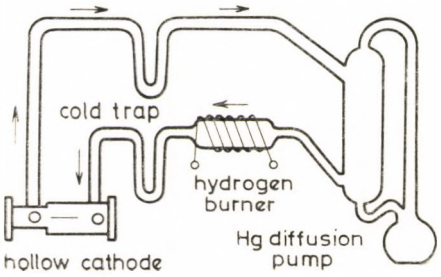


Fig. 2.118 Gas circulation system of the hollow cathode[2]

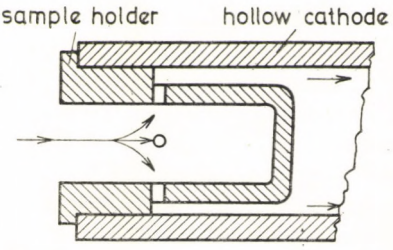


Fig. 2.119 Inlet for gas stream into hollow cathode

This way, a 'gas umbrella' is formed at the orifice of the cavity that reduces the outward migration of the sample vapour. This has the advantage that in excitations of long duration, even spectral lines of low-boiling-point elements, especially those that tend to undergo fractional distillation, retain their intensity. (For example, 10 mg of strontium chloride drops by only 10% during a half-hour excitation period.)

In hollow-cathode excitation it is a distinct advantage that *samples of a comparatively small size can be analysed with a reasonable degree of precision.* In most cases a few milligrams will be enough for analysis, although with the aid of specially made hollow-cathode tubes samples as small as a few micrograms can be examined.[5] In selecting the best weight of sample to use, it should be borne in mind that during long excitation periods the intensity of the spectral lines of the element to be determined can drop considerably. This intensity reduction can be traced back to the fact that during excitation the surface of the sample becomes progressively covered by more and more of the sputtered cathode material. This effectively reduces the evaporation speed of the sample. (This fact provides an explanation for a 30% drop in the emission observed when 10 mg of $SrCl_2$ is subjected to a half-hour excitation period.[11]) In general, the smallest amount of sample that can be examined in the hollow cathode is largely influenced by the effect of cathode sputtering, or indirectly by the experimental conditions and the duration of the excitation. The precision of hollow-cathode excitation is much less influenced by the chemical inhomogeneity of the sample surface than in the case of arc and spark excitation.

A further advantage of hollow-cathode excitation is that the excitation is very stable. The absolute fluctuation of the emission is as low as $\pm 0{\cdot}5\%$, even when a non-stabilised power supply is used. The reproducibility of the intensity of the spectral

lines obtained in this way is better than 1%, provided the experimental conditions influencing the excitation are held at a sufficiently constant level.[15] This exceptional stability is partly due to the excitation occurring in a closed space, such that variation in the surroundings (convections, etc.) are practically ineffective. Also, one other important factor is that the geometrical location of the radiation origin is very extended and stable compared with arc and spark excitation.

In many cases it is an advantage of the hollow-cathode excitation method that the profile of the spectral lines is influenced only by Doppler broadening. This type of broadening is reduced upon cooling the hollow cathode.[3,4] For example, the half-width of the hydrogen spectral line at 6562·79 Å with a cathode temperature of 300 K is 0·1 Å, whereas if the hollow cathode is cooled to 81 K by liquid air the half-width is only 0·03 Å. These values are further reduced as the atomic weight increases. Resonance broadening and pressure broadening of the lines can be ignored as compared with the small Doppler broadening. External Stark effect is not observed since there is no anomalous cathode fall in hollow-cathode excitation.[2] Also the internal Stark effect does not arise since the concentration of the ions and electrons is so low in the low-pressure vapour phase that their electric field does not have any effect on the excitation. At the gas pressure usually applied in hollow-cathode excitation, the resonance broadening of the spectral lines is only of the order of 0·01 Å. The line broadening is still essentially smaller than the previous value since the carrier gas greatly reduces the partial pressure of the sample vapour. When it is considered that the pressure broadening of spectral lines is always smaller than the resonance broadening (even at atmospheric pressure), it is reasonable to neglect pressure broadening at the relatively low pressures found on hollow-cathode excitations.

In summary it can be stated that the little-used hollow-cathode excitation method is very important from the analytical point of view. It is excellent for the examination of non-conducting materials. Among the commonly used excitation methods the hollow-cathode method is the most useful for the determination of elements that are difficult to excite, but are comparatively easy to evaporate (halogens, sulphur, etc.). It is also useful in the examination of semiconducting materials (even certain metals) in trace quantities. At present it cannot be reliably established in which cases the sensitivity of hollow-cathode excitation is higher than that of d.c. arc excitation.

References

1. Paschen, F., *Ann. Phys.*, **50**, 901 (1916); **71**, 142 (1923)
2. Schüler, F., *Z. Phys.*, **22**, 264 (1921)
3. Schüler, F., *Z. Phys.*, **35**, 323 (1926); **59**, 149 (1930); **72**, 423 (1931)
4. Wittke, H., *Z. Phys.*, **116**, 547 (1940)
5. Schüler, H., and Gollnow, H., *Z. Phys.*, **93**, 611 (1935)
6. Tolansky, S., *Proc. R. Soc.*, *A*, **152**, 663 (1935)
7. Tolansky, S., *Proc. R. Soc.*, *A*, **158**, 110 (1937)
8. Birks, F. T., *Spectrochim. Acta*, **6**, 169 (1954)
9. Mitchell, K. B., *J. Opt. Soc. Am.*, **51**, 846 (1961)
10. Rosen, B., *Revue Univ. Mines*, **9**, 445 (1953)
11. Eichhoff, H. J., Rozpravy narod. techn. muz. Praze, *Fys. Chem. rada*, **13**, 7 (1964)
12. Chorovin, Yu. I. and Lipis, L. V., *Optika Spektrosk.*, **5**, 334 (1958)
13. Chorovin, Yu. I. and Lipis, L. V., *Proc. 2. U.N. Intern. Conf. Peaceful Uses Atomic Energy*, **28**, 604 (1958)

14. MASSMANN, H., *Proc. IX. Coll. Spectr. Intern., Lyon, 1961*, vol. III, 170
15. ADAMS, K. B., and BURNS, K., *Research Memo*, 6-94602 — 9.M 37

2.10.4 Excitation in flash tubes

Essentially the flash tube is a glass or quartz tube filled with a gas at a pressure of about 100 mmHg. Between the electrodes that are welded into the two ends of the tube, the potential difference is much lower than the breakdown voltage; usually about 1–3 kV. The electrodes are connected to a condenser having a capacitance of 10–100 μF, charged by a d.c. supply. The discharge is brought about via an external or internal electrode, the former of which is called the *ignition electrode*. The flash discharge emerging in this way is of extremely short duration and generally lasts for about 10 μs, although its energy is relatively large (>10 W s).[1–3, 7]

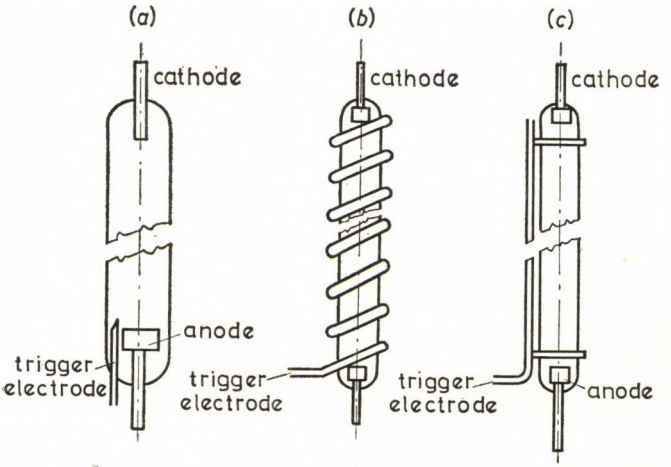


Fig. 2.120 Flash tubes with secondary electrodes: (*a*) internal type, (*b*) spiral type, (*c*) straight external ignition type

If the ignition of the discharge tube is secured by an internal pointed, *secondary electrode* (Fig. 2.120*a*) it is usually connected to the anode. Thus, the electric field is concentrated in between the cathode and the point of the ignition electrode, although no breakdown is possible yet. When, however, a large enough positive voltage pulse is passed into the ignition electrode, the field strength and thus the kinetic energy of the electrons in the neighbourhood are rapidly increased, such that the gas is ionised. Along with the ionisation process secondary electrons are liberated and photons produced. The ionisation front and the radiation become extended towards the anode, until the discharge fills the whole length of the tube.

When the *external ignition electrode* is used, either the electrode surrounds the discharge tube in a spiral form or the electrodes are fastened to the whole length of the tube by means of rings (Fig. 2.120*b*, *c*). It can be shown in both cases that the ionisation is initiated *along* the trigger electrode at the beginning of the discharge when the positive voltage pulse is passed into the trigger electrode.

The *ignition pulse* is generally a strongly damped oscillation. The best ignition is produced when it occurs during the first maximum and when the duration of the latter is the shortest. The duration of the gas discharge is, however, influenced by subsequent oscillations as well. When the anode of the discharge tube is earthed it is

necessary to use a positive ignition pulse. Ideal ignition is only obtained when the ignition voltage is significantly higher than the breakdown voltage. (For example, in the case of discharge tubes whose working voltage is 2·5 kV, the ignition voltage is usually 3 kV. Thus, at the instant of the ignition a potential difference of 5·5 kV is produced between the anode and the cathode.)

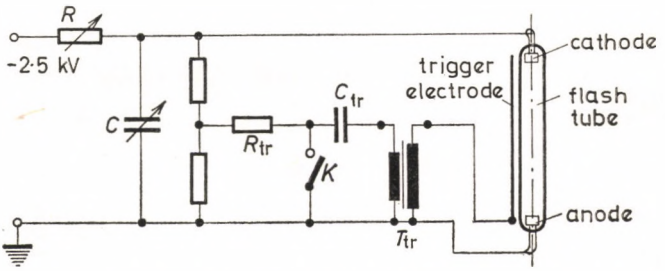


Fig. 2.121 Circuit diagram of theele ctrical power supply for a flash tube

The principal circuitry of flash ignition is illustrated in Fig. 2.121. The rectified, high-voltage current charges the variable capacitance condenser, C, through the limiting resistor, R, while the ignition condenser, C_{tr}, is charged through the voltage divider resistor to about half-voltage. Upon closing the switch the condenser, C_{tr}, is discharged through the primary winding of the autotransformer, T_{tr}. Thus a current pulse is produced that establishes a positive voltage pulse of short duration, that is, the ignition voltage in the secondary winding of the transformer. The effect of the latter is that the gas in the discharge tube is ionised and becomes conductive such that the discharge of the condenser, C, is made possible.

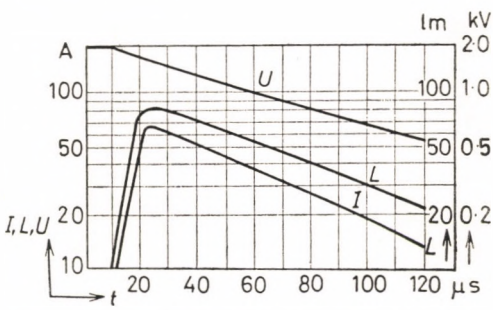


Fig. 2.122 Schematic current–time, I, voltage–time, U, and radiance–time, L, curves of the flash tube

From the current–time curve of the flash discharge (curve I in Fig. 2.122) it is evident that the current passing the discharge tube increases rapidly during the ignition, then, after reaching a maximum intensity, it drops exponentially. After the formation of the plasma the terminal voltage of the discharge tube also drops exponentially, such that the resistance of the tube does not change in the meantime. The impedance of the flash discharge tube is therefore a purely ohmic resistance.[4] The trend of the light flux–time curve is more or less the same as that of the current–time curve. The only difference is that the current–time curve passes through a maximum, while the maximum light flux can be observed for a somewhat longer period, a few microseconds.

The spectrum of the emitted light consists of ion lines, and atomic lines, as well as a continuum stretching from the infra-red to the ultra-violet. It follows from the current–time curve that the constitution of the spectrum varies during the different stages of the formation of the plasma.[5] Initially, the field strength and thus the kinetic energy of the emerging charged particles is very high. The positively charged, excited ions of the ionised gas initially produce a spectrum of definitely ionic character. Because the electrons have a relatively high kinetic energy, thermal dissociation continues and the ionic spectrum persists for several microseconds, in spite of the fact that the current passing through the sharp maximum drops very rapidly. Presumably, the electronic temperature of the plasma exceeds 7000 K at this stage of the discharge.[6] In the next stage, the terminal voltage of the condenser, and thus the ionic density of the plasma and the current along the discharge tube, are reduced to the level where the voltage becomes so low that the discharge terminates. In the meantime a further loss of ionic density is caused by the diffusion of the ions to the wall of the discharge tube where recombination occurs. *Atomic lines* only appear later during the afterglow period. Their appearance can be related to the fact that the energy liberated during the recombination of the electrons and ions is used for the excitation of the atoms. The intensity of the continuous spectrum varies in the same fashion as the ionic density of the plasma during the discharge, being highest at the formation of the plasma and later dropping quickly. This is explained by the electrons being decelerated in the ion field, such that 'deceleration radiation' is produced.[5] The intensity of the latter drops with the decreasing ionic density and the electronic temperature.

The flash tube is mainly applicable for the examination of noble gases. It cannot be used for gases that attack the electrodes.

References

1. Robinson, N. W., *Philips tech. Rdsch.*, **16**, 164 (1945)
2. Edgerton, H. E., Germeshausen, J. K., and Grier, H. E., *Br. J. Appl. Phys.*, **8**, 219 (1937)
3. Hoyt, G. D., and McCormig, W. W., *J. Opt. Soc. Am.*, **40**, 658 (1950)
4. Lamparte, M. J., *Br. J. Appl. Phys.*, **8**, 332 (1937)
5. Olsen, H. N., and Huxford, W. S., *Phys. Rev.*, **87**, 992 (1952)
6. Craggs, J. D., and Meek, J. M., *Proc. R. Soc. A*, **154**, 241 (1934)
7. Fayolle, F., and Naolin, P., *Photographie instantanée et cinématographie ultra rapide* Paris (1950)

2.10.5 High-frequency excitation at reduced pressures

Excitation can be produced in discharge tubes by a high-frequency current as well.[1–3] This has the constructional advantage that to produce excitation it is enough to surround the discharge tube by a coil of a few turns that is connected to a tube generator of frequency 10^7 Hz and 100–500 W (Fig. 2.123).

In high-frequency excitation the electrodes that form part of the discharge tube are not necessary, so that the shape of the discharge tube can be greatly simplified. In most cases a quartz or hard-glass tube is used that can be evacuated and one end of which is closed by a parallel-sided quartz window that transmits ultra-violet light (Fig. 2.124). The simple construction of the discharge tube is not only welcome from

the glass-blowing point of view but it also allows the apparatus to be evacuated easily. (For the same reason it is preferable to have a large internal diameter, not only for the discharge tube, but also for its associated equipment.)

A further special feature of high-frequency excitation, as compared to the excitation methods described in the previous Sub-sections, is that the charged particles do not migrate along the axis of the tube. Therefore, the light phenomena that can be

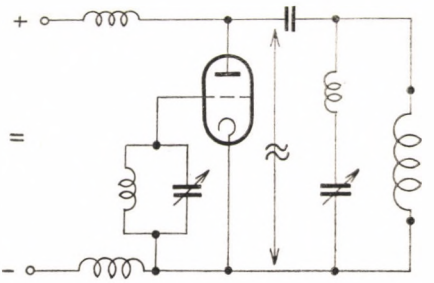


Fig. 2.123 Schematic circuit diagram of a high-frequency generator containing a tube generator unit[8]

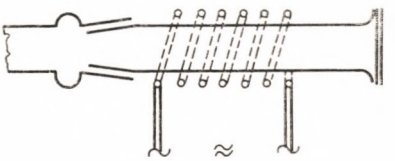


Fig. 2.124 Discharge tube used for high-frequency excitation

observed in the different parts of the discharge tube during d.c. excitation do not occur under high-frequency excitation. In the latter method, the direction of movement of the charged particles is determined collectively by disoriented thermal movements plus the orderly force-line variations of the high-frequency electric field. As a result a whirling convective motion occurs. The collisions occurring during this motion ionise the atoms and excite the ions produced.

The intensity of the radiation emerging not only depends on the electric field strength, but also varies with the gas pressure and the shape and dimensions of the discharge tube. The kinetic energy of the charged particles is also influenced by their free path length, that is, by the gas pressure. At comparatively high pressures the average free path length of the particles is relatively small. Thus, relatively few particles possess kinetic energy great enough to produce extensive collision ionisation, as well as to excite the atoms and ions. At gas pressures that are too low, on the other hand, the probability of collisions that lead to excitation is very small. Both factors lead to the reduction of the intensity of the emission from the light source. It can be experimentally shown that in high-frequency excitation the optimum gas pressure is between 10^{-2} and 10^{-3} mmHg. Disadvantageous effects occur if the pressure is much higher than 1 mmHg or when the gas pressure is lowered to 10^{-5} mmHg.[5]

Apart from the low emission intensity, a further consequence of low gas pressures and the relatively large free-path lengths developed is that the exciting particles cause excitation of the higher energy levels. The spectrum produced is therefore of ionic character. Consequently, low gas pressures in high-frequency discharge tubes cause

elements that are normally difficult, to excite, become excited. Thus, the elements As, Cl, Br, I, Se, S, and P are also brought to the excited state.[1–4, 6, 7] It can furthermore be proved that the excitation of the above elements is greatly increased if the carrier gas has a high excitation potential. This is so since in high-frequency excitations the secondary collisions play an important role.

A further advantage of a low gas pressure is that the natural profile of the spectral lines is only altered by Doppler broadening. The spectral lines are therefore sharper than when Geissler discharge tubes are used. At low gas pressures not only the internal Stark effect can be neglected (it is caused by the electric field of the ions), but also the external Stark effect. The latter is produced by the electric field arising due to the high-frequency current.

With a slight modification the discharge tube can be used in high-frequency excitation for the examination of solid samples as well as gases. The solids are strongly heated in the high-frequency electric field. The partial pressure of their vapour is enough to excite the elements to a sufficient degree.

In the examination of electrically non-conducting materials, the heating of the sample is caused partly by dielectrical hysteresis and partly by so-called displacement currents. In an electric field the molecules of dielectrics are polarised in such a way that the separate charges do not traverse a long distance in the medium and thus do not reach the surface of the material. The exciting current, which varies in a sinusoidal fashion, not only produces a similarly varying electrical field, but the electrical polarisation of the molecules also changes in a sinusoidal fashion. In the latter process, *dielectric hysteresis* causes the evolution of heat that warms the dielectric materials. The heating produced is increased by *displacement currents*. Due to the polarising effect of the electric field the molecules of the dielectrics are elastically displaced. The small movement of the charges brings about an electric current of very short duration. The energy necessary for the displacement is eventually transformed into heat and this in turn increases the temperature of the dielectrics. It can be calculated that the amount of heat, P_I, that is turned into thermal energy per second by the electric field, E, in dielectrics of volume, v, is given by

$$P_I = \frac{E^2 \varepsilon_a \, v\omega^2 \tan \delta}{8\pi^2(1 + \tan^2 \delta)} \text{ watts} \tag{2.157}$$

where ε_a is the absolute dielectric constant of the dielectrics, δ is the loss angle that depends on the quality and temperature of the material, and ω represents the angular frequency. According to the above equation it can be established that heat production and thus the efficiency of the excitation are largest when the electrical field strength and frequency are highest.

If the sample is an electrically conducting material then the heating and evaporation of the sample during high-frequency excitation are caused by various effects, as well as those described above. In this case, the relatively minor *skin effect* also plays a role. Due to this effect the high-frequency secondary current that is producing heat is found only in the layers close to the surface of the metal. Therefore, the resistance of the surface layer of the metal is important in this respect. The interior of the metal is thus heated by thermal conduction, as opposed to the case of electrically non-conducting materials. It is therefore an advantage if the specific surface area of metallic samples is as large as possible. The value of the power produced, which is completely

converted into heat, is given via

$$P_{II} = UI \cos \phi$$

where U is the induced voltage, I is the current due to the former voltage, and ϕ is the phase angle. From the above fundamental relationship and using the elementary relationships: $U = \Phi\omega/\sqrt{2}$, $I = U/\sqrt{R^2 + L^2\omega^2}$, and $\cos \phi = R/\sqrt{R^2 + L^2\omega^2}$ we obtain

$$P_{II} = \frac{\Phi^2\omega^2}{2} \cdot \frac{R}{R^2 + L^2\omega^2} \tag{2.158}$$

where Φ is the value of the flux established when the magnetic field strength, B, penetrates the surface f according to the formula: $\Phi = \int B \, df$ and where R and $L\omega$ are the ohmic and inductive resistances, respectively, and ω is the angular frequency. It can be stated that for efficient excitation by a high-frequency current, the higher the frequency the better.

The advantages of the above excitation method are as follows.[1-4, 6, 7]

(1) The amount of the sample necessary for analysis is relatively small; generally about 10–20 mg.
(2) The sample can be examined in powder form.
(3) The lower limit of detectability of metallic elements is about the same as the sensitivity of normal arc excitation. Detection is reduced, however, when there is intensive gas evolution during the excitation. This is due to interference by the easily excited bands of the gases produced (H_2, NH_3, N_2, CO, CO_2, etc.).
(4) Non-metallic elements that are difficult to excite (halogens, S, Se) can be detected when present in the concentration range 10–1000 p.p.m. The *analytical accuracy obtained is good;* for example, the error involved in determining 100 p.p.m. Cl or I is $\pm 10\%$.

In contrast to the previously mentioned advantages, it is a pity that the procedure cannot be applied in general and not only for the examination of easily evaporated samples.

References

1. FENNER, F., *Spectrochim. Acta*, **1**, 164 (1941)
2. GATTERER, A., and FRODL, V., *Ric. Spectrosc. Lab. Astrofis. Specola Vatic.*, **1**, 201 (1946)
3. GATTERER, A., *Spectrochim. Acta*, **3**, 214 (1947)
4. SALPETER, E., *Ric. Spectroscop. Lab. Astrofis. Specola Vatic.*, **2**, 1 (1952)
5. TOLANSKY, S., *High Resolution Spectroscopy*, London (1947)
6. SCHROLL, E., and ROCKENBAUER, W., *Spectrochim. Acta*, **11**, 338 (1956)
7. KELLER, R. E., and SMITH, L., *Analyt. Chem.*, **24**, 796 (1952)
8. KOHLRAUSCH, F., *Praktische Physik*, 2nd edn, Leipzig, vol. 2 (1951)

2.11 Excitation by laser

2.11.1 Fundamental principle of the procedure[1–5]

The word 'laser' is an abbreviation standing for 'light amplification by stimulated emission of radiation'. It is, in fact, a light amplification brought about by *induced emission*, whereas the light emitted by the light sources used in spectroscopy is produced by *spontaneous radiation*. Thus there is an important difference between the origin and properties of the two types of radiation.

The light emitted by spontaneous radiation light sources is radiated by a great number of atoms in a haphazard fashion. Accordingly, the light emerging in this way travels in every direction and is composed of rays that usually have different frequencies and are always of different phase. According to Einstein a further characteristic of spontaneous radiation is that it is typically a random and chaotic process. Thus, only the *probability* of the individual states and the *average* lifetime of the excited states can be given. The latter is independent of the radiation energy density, but depends on the frequency of the radiation. (For example, in the case of an optical frequency the average lifetime is in the order of 10^{-8} s.)

In contrast to spontaneous emission, the *induced emission* of an atom is radiated at a particular frequency. Accordingly, the radiation produced is coherent, and is composed of rays having the same phase in addition to the same frequency. An important difference is that the process of spontaneous emission cannot be controlled, whereas induced emission, being initiated by *external* factors, can be controlled. According to Einstein, however, the probability of induced emission is the same as the probability of exciting an atom from a lower energy level to a higher one. The two processes are therefore reversible. In contrast to spontaneous radiation, the probability of induced emission changes proportionally with the energy density of the radiation field in which the emitting atom is located.

It follows from the above discussion that a prerequisite of induced emission concerns *absorption*. Since absorption is the reverse process of emission it follows that the same energy relations are valid for both processes. The atom can only absorb radiation of a particular frequency corresponding to the condition: $E = h\nu$. The prerequisite of induced emission, therefore, is the *resonant absorption*.

Due to resonant absorption the number of atoms in the lower energy level, N_1, is reduced while the number of atoms in the upper level, N_2, is increased. In a closed system, however, there is a statistical equilibrium between the numbers of excited and non-excited atoms. According to Boltzmann's law this happens when the following condition is met

$$N_2/N_1 = c \cdot e^{-\Delta E/kT} \tag{2.159}$$

where $\Delta E = E_2 - E_1$, the Boltzmann constant $k = 1{\cdot}38 \,.\, 10^{-16}$ erg/deg, T is the absolute temperature, and c is a constant.

It follows from the above formula that when $E_2 > E_1$, then $N_1 > N_2$. This means that the number of non-excited atoms is generally greater than the number of excited ones. Resonant absorption is terminated when the number of atoms in the two energy states (or their relative populations) is the same, that is, when $N_1 = N_2$. This is the so-called *saturation* state. On the other hand, *population inversion* is the state when the number of atoms in the higher energy state is greater than in the lower state, that is, when $N_2 > N_1$. Without the inversion of the populations no induced emission can occur.

In lasers used for spectrochemical analysis (for example, the ruby laser) the population inversion is made possible by a simple process based on the *three-level system.* In order to explain this process we take an example of an element whose atom has two very closely associated levels, E_2 and E_3, in addition to a level E_1. In this case, $E_1 \ll E_3 > E_2$ and the relatively small energy produced in the transition $E_3 \rightarrow E_2$ is converted to heat in a radiationless way. The population of these energy states according to Boltzmann's law is *generally* $N_1 \gg N_3 < N_2$.

When the atoms with the above described energy level structure are irradiated by a very intense *pumping light* of frequency $\nu_{3,1} = (E_3 - E_1)/h$, it is possible to make the energy level, E_3, saturated; that is, its population, N_3, becomes equal to N_1. When the relaxation time for the transition $E_3 \rightarrow E_2$ is short, then the population ratio of the two energy levels cannot deviate very much from Boltzmann's law, that is, $N_3 < N_2$. In the saturation state, where $N_1 = N_2$, it follows that if the relaxation time of the transition $E_2 \rightarrow E_1$ is relatively long, then $N_2 \gg N_1$ and the number of electrons in this metastable state can be very large when compared to the ground state. In this 'activated state' the energy level E_2 can contain an excess of energy, $E_2 - E_1 = h\nu$, sufficient for the production and maintenance of induced emission. This state can only occur, however, when the intensity of the external 'pumping light' exciting the laser is sufficiently high.

A further prerequisite for laser action is the existence of an appropriate 'optical resonator'. This is an optical arrangement that provides the maximum number of atoms emitting coherent radiation. The basic principle of the resonators used for this purpose is the same as that of the Fabry–Perot interferometer.

The original non-modified resonators were made from a single cylindrical crystal of ruby with exactly parallel ends. The crystal ends are effectively mirrors, one of which reflects all the light passing along the axis of the resonator, while the other is only partially reflecting. The distance between the two mirrors is a whole-number multiple (which is very large) of the half-wavelength of the light produced by induced emission.

Only solid-phase lasers are used for spectrochemical purposes. These are usually single crystals (or glasses) that contain in the lattice an appropriate metal cation or so-called *optical activator.* Its concentration is generally between a few and several per cent. The material of solid lasers is, therefore, composed of an 'active' elementary system and a 'neutral' optical medium. In the former the absorption and emission processes take place, while the latter is essentially a matrix within which the active atomic system is homogeneously distributed.

A proportion of the atoms of the active element is brought into the excited state when a *pumping light* of appropriate wavelength and intensity irradiates the resonator. A large proportion of the atoms excited in this way release the absorbed excitation energy as spontaneous radiation, such that the light emerging leaves the resonator in different directions. (This radiation, therefore, does not take part in the later processes.)

On the other hand, the rays passing parallel to the resonator axis and reflected from the end mirror force the atoms that are still in the excited state to produce induced emission, so that coherent light of a steadily increasing intensity is emitted. Thus the radiation field strength can be so high in the system that it becomes 'independent' for a period; that is, no more external energy is necessary for the maintenance of the induced emission. The greater the dissipation of energy, the shorter is

the period for which this ideal state prevails. Energy losses are caused by light being absorbed by the *'neutral' material* of the laser, lattice faults, etc.

The amplified light emerges only when the following relationship exists between the product of the wavelength, λ, and the integer k, with the distance, l, between the reflecting surfaces and the refractive index, n

$$k\lambda_1 = 2ln \tag{2.160}$$

When the above condition is met, the 'resonator' is in a resonant state with respect to the wavelength of the radiation emerging due to laser action. This state is characterised by the fact that the waves passing between the reflecting surfaces are always in phase at any particular location in the radiation field, such that standing waves are produced. This state can only be reached, however, when the two reflecting surfaces are *exactly* parallel. Thus, the light beam leaving the resonator at a very small solid angle is coherent both in time and space and has a very high energy density.

References

1. LENGYEL, B. A., *Lasers*, New York–London (1963)
2. MOENKE, H. and MOENKE-BLANKENBURG, L., *Einführung in die Laser-Mikro-Emissionsspektralanalyse*, Leipzig (1966)
3. HAPP, H., *Umschau*, **61**, 599 (1961); **62**, 133 (1962)
4. ROTHE, H., *Umschau*, **64**, 486 (1964)
5. GÖRLICH, P., LEHMAN, R., and LÜDKE, W., *Feingeräte-Tech.*, **14**, 227, 241 (1965)

2.11.2 Solid-phase lasers used for excitation[1]

The commonest type of laser used for excitation purposes is the ruby laser or, less frequently, the Nd^{3+}-glass laser. The advantage of using the former type is that the laser radiation is visible, while the light from the neodymium-glass laser lies in the infra-red region. The latter property, therefore, renders the alignment of the laser or that of the induced radiation difficult. Apart from this aspect, the basic application principles of the different types of laser described below are the same.

The ruby laser is a single crystal of aluminium oxide within which a small proportion of the aluminium ions are replaced by Cr^{3+} ions. The active substance of the ruby laser is the chromium, which amounts to about 500 ppm Cr.

When the ruby laser is irradiated with green light of wavelength around 5500 Å the light energy absorbed excites the Cr^{3+} ions via a transition from the $^4A_2$ level to the $^4F_2$ level. The average lifetime of this excited level is very short and the ions are reduced to the ground state via the metastable ^2E level, instead of directly (Fig. 2.125). This intermediate transition, however, produces no radiation, as the energy released is heat taken up by the crystal lattice. In the transition between the metastable level and the ground state a laser radiation consisting of a red doublet at wavelengths 6943/6929 Å occurs at a temperature of 300 K (that is, about room temperature).

The original ruby lasers were usually single cylindrical crystals of ruby having a diameter of 3–20 mm and a length of 20–250 mm, with ends ground parallel to within at least 1 minute.[2] The silvered ends act as mirrors and *one of them* has a transmittance of about 2–4%.[3] In the original design the ruby laser is illuminated by the green light of a xenon flash tube. The ruby crystal is positioned such that it lies along

the axis of a spiral form of the flash tube. The ruby laser and the xenon flash tube, plus the reflector, constitute the so-called *laser head* (Fig. 2.126). The other part of the apparatus is the supply unit for the flash tube. The latter is normally operated by the discharge current of a high-capacity condenser with a terminal voltage of

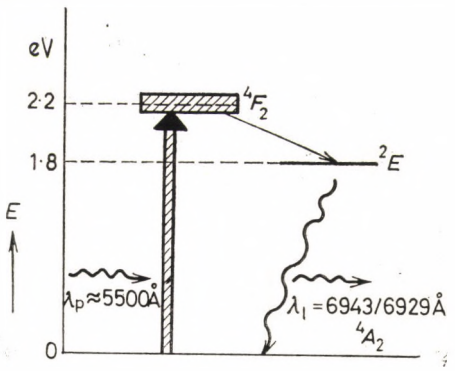


Fig. 2.125 Simplified excitation scheme of the ruby laser[2]

2–4 kV. Light pulses lasting for about 10^{-3} s emerge. The intensity of the laser radiations initiated by the flashes is, however, uneven. The light emitted during a single pulse consists of pulses of non-uniform intensity having a duration of about 10^{-6} s interspersed with longer intervals when no radiation is produced. These pulses are called 'spikes'. As the duration of the spikes cannot be controlled, the original ruby laser system was of the *non-controlled pulse type.*

The drawbacks of the original ruby laser are eliminated by a number of modifications. For example, the reflecting surface formed by a layer of silver coated onto the parallel resonator ends has not been found practically useful. The unavoidable

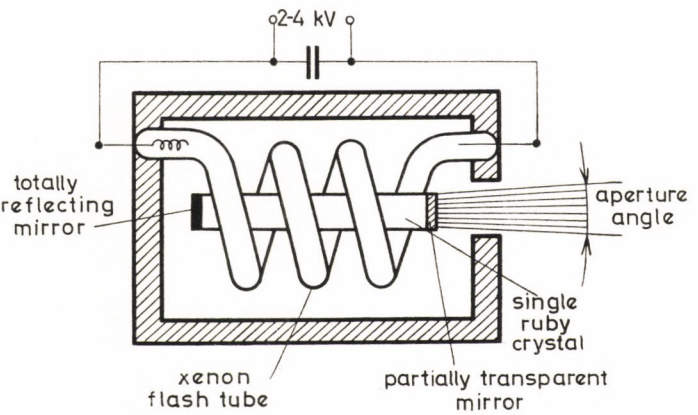


Fig. 2.126 Schematic drawing of the original ruby-laser head

heating of the resonator progressively deteriorates the reflectivity of the silver mirror and thus the efficiency of the laser. This fault can be eliminated when *dielectrically coated surfaces* are used. Here the reflecting surface is formed by thin alternating layers having different refractive indices. With this modification the reflectivity can reach as high a value as 99·9%, which is much better than the reflectivity of silver.

In another common modification, the previously mentioned disadvantage is eliminated by using a separate interferometer plate, or a totally reflecting frontal surface instead of the reflecting surface of the resonator (Fig. 2.127).

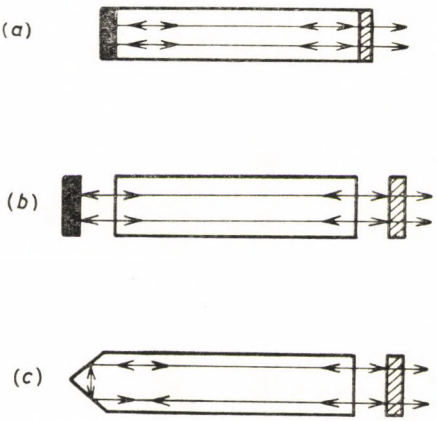


Fig. 2.127 Different reflection methods of solid-phase lasers: (*a*) original arrangement, (*b*) with separate interferometer, and (*c*) resonator equipped with a deviation prism

The efficiency of the original ruby laser – and in fact of all solid-phase lasers – is very small and does not usually reach the value of 10^{-3}. For example, when the flash pulses have an *input energy* of 1000 W s the emerging radiation or *output energy* of the ruby laser has an energy of less than 1 W s. The efficiency can be significantly increased if the transmittance of the output side of the resonator is increased to 4–15%.

The spiral-form flash tube also reduces efficiency, because only a small proportion of the flash radiation falls upon the resonator. Thus, a pumping light of sufficient energy can be obtained only if the input energy is significantly greater than 1000 W s.

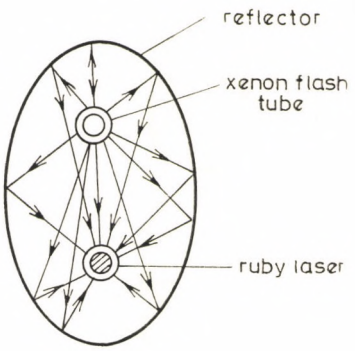


Fig. 2.128 Illumination of the solid-phase laser by a cylindrical mirror of elliptical cross-section

This requirement can, however, be reduced to a few 100 W s if the rod-form flash tube is used for the illumination of the resonator. In this system, one of the focal axes of an elliptical mirror-mantle coincides with the axis of the flash tube, while the other coincides with the axis of the ruby laser (Fig. 2.128). In the non-controlled, solid-phase laser the energy accumulated during a single pulse of the flash tube produces a series of light pulses consisting of a great number of rapidly developed spikes (Fig. 2.129). The spike oscillation can be inhibited by an optical switch situated between the resonator and the mirror. This has the advantage that the single pulse that emerges contains all the output energy accumulated within the laser.

The output power in this case can be about a hundred times bigger than that of the non-controlled laser. This modification of the ruby laser is accordingly called the *giant pulse laser.*[4]

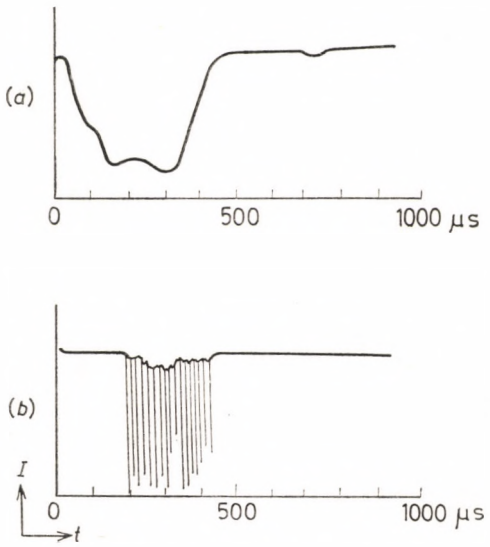


Fig. 2.129 Light pulse–time curve of a xenon flash tube (*a*) and a non-controlled pulse laser (*b*)

The emergence of the giant pulse is explained by the fact that most of the atoms are not excited during the optical pumping until the so-called optical switch (or *Q-switch*) allows the rays to be reflected from the mirror surface. Although it is possible that photons may reach the resonator reflecting surface and be reflected, the optical switch excludes the possibility of multiple reflections between the two reflecting surfaces. Thereby, any significant induced emission is precluded. Since the energy lost in this way is relatively small and can be easily compensated for by the pumping process, the energy accumulates in the resonator until the optical switch allows the induced emission that is travelling along the optical axis of the resonator to escape as a giant energy pulse. The duration of the light pulse emerging in this way is in the order of 0·1 μs (Fig. 2.130).

A Kerr-device or rotating deviation prism or mirror can be used as the optical switch.

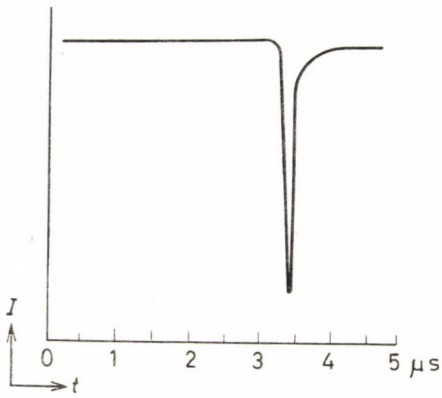


Fig. 2.130 Light pulse–time curve of the 'giant-pulse' laser

In the optical switch based on the Kerr effect, the light leaving the resonator passes through a polariser and a Kerr cell to the semireflecting mirror (Fig. 2.131*a*).[4] It is well known that when an appropriate voltage is applied to the condenser plates of the Kerr cell, the nitrobenzene becomes birefringent. By the appropriate variation of the voltage the optical switch either blocks or transmits the polarised beam. For

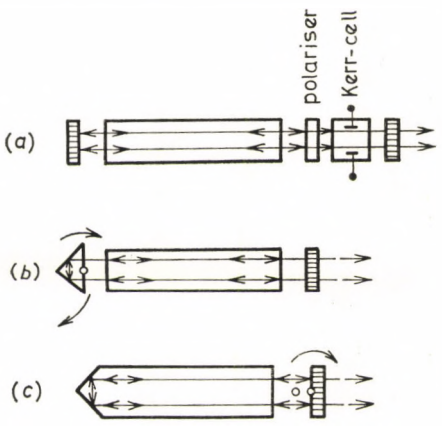


Fig. 2.131 Commonly used optical switches: (*a*) based on the Kerr effect, (*b*) rotating-prism device, and (*c*) rotating interferometer plate device

the operation of the device an electronic unit is necessary to control the flash tube and the terminal voltage of the Kerr cell.

In the rotating prism optical switch the totally reflecting mirror of the resonator is replaced by a rotating deviating prism (Fig. 2.131*b*). The prism is rotated by a synchronous motor at a very high speed. For example, by using a deviating prism rotating with a speed of 30 000 rev/min (or 10·8 minutes of an arc per microsecond) it is possible that the radiation along the resonator axis is reflected only for a duration of 0·1 μs. When the rotating prism optical switch is used, the flash pulses are electronically controlled by the light pulses of the secondary light source being projected by a *trigger-mirror* onto a photocell.

There are two versions of the *rotating mirror optical switch* in use. One of them is different from the previously described optical switch in that the rotating prism is replaced by a rotating dielectric mirror. A more important difference is that the revolution speed of the mirror is only 6000 rev/min (that is, its angular velocity is only 2·16 minutes of an arc per microsecond). Accordingly, during a period of about 10 μs 8–10 giant pulse spikes occur.[6] In another version of the rotating mirror system, the interferometer plate having a 50 % transmittance rotates with a speed of 1500 rev/min (that is 0·54 minutes of an arc per microsecond).[7, 8]

In the application of the rotating mirror optical switch the ignition of the flash tube is also controlled electronically with the aid of a trigger mirror. It is evident that both versions of the rotating mirror system represent transitions between the giant pulse and the non-controlled pulse methods.

In the use of solid-phase lasers it should also be taken into account that the actual diameter of the resonator, d, and the *effective diameter*, d_e, are not identical, but that the following relationship exists between them

$$d = d_e n \tag{2.161}$$

where n is the refractive index of the 'neutral optical material'. For example, the effective diameter of the ruby laser is equal to $d/1{\cdot}75$.

References

1. MOENKE, H. and MOENKE-BLANKENBURG, L., *Einführung in die Laser-Mikro-Emissionsspektralanalyse*, Leipzig, 22 (1966)
2. BRECH, F., *X. Coll. Spectr. Intern., College Park, 1962*, Washington (1963)
3. MAIMAN, F. H., *Nature, Lond.*, **187**, 493 (1960)
4. MCCLUNG, F. J. and HELLWARTH, R. W., *J. Opt. Soc. Jap.*, **33**, 828 (1962)
5. *Jarrel Ash Co. Bulletin* 45–600
6. FELSKE, A., HAGENAH, W. D. and LAQUA, K., *Z. analyt. Chem.*, **216**, 50 (1966)
7. LAQUA, K., *Arbeitstagung Optik und Spektroskopie aller Wellenlengen*, Jena (1965)
8. HAGENAH, W. D., *Z. angew. Math. Phys.*, **16**, 130 (1965)

2.11.3 Application of lasers

The laser can be used for spectrochemical analysis when its radiation is directed onto the surface of the sample. At the surface, the intense high-energy light beam is absorbed and transformed into thermal energy. This results in the evaporation of the sample and establishes a *microplasma* in the immediate neighbourhood of the irradiated surface. The radiation from the microplasma acts as the light source for spectral analysis.

The greater the energy density of the laser radiation impinging on the sample, the higher the temperature and, thus, the degree of ionisation of the microplasma. The value of the energy density depends partly on the input energy, E, of the laser and partly on the diameter, r, of the illuminated area. When f is the focal distance of the illuminating lens system and α is the solid angle of the light beam incident on the surface, then the following relationship holds for an illumination diameter of r

$$r = f \cdot \tan \alpha$$

Accordingly, it can be written that

$$\text{energy density} = \frac{E}{\pi r^2} = \frac{E}{\pi f^2 \tan^2 \alpha} \tag{2.162}$$

Therefore, at a given input energy the energy density and thus the temperature of the microplasma can be increased by reducing the focal distance and the aperture of the lens system. The limit to the shortening of the focal distance is determined by the '*working distance*' of the lens system; that is, by the distance between the front surface of the lens system and the sample. This distance cannot be reduced below a certain level for practical reasons. According to the diffraction theory the solid angle of the laser radiation is given by

$$\alpha = 1{\cdot}22 \frac{\lambda}{a} \tag{2.163}$$

where a is the diameter of the light beam. The solid angle of common solid-phase lasers is about 10 seconds of an arc. In practice this value is significantly greater and is usually several minutes of an arc.

Furthermore, it should be noted that owing to the effect of the laser radiation the size of the *crater* created on the surface of the sample is always bigger than the focus-spot of the lens system. The size of the crater depends on the surface properties of the sample, especially on the absorption, the type of the material, and the evaporation temperature,[1] as well as on the output energy of the laser and the parameters of the lens system.

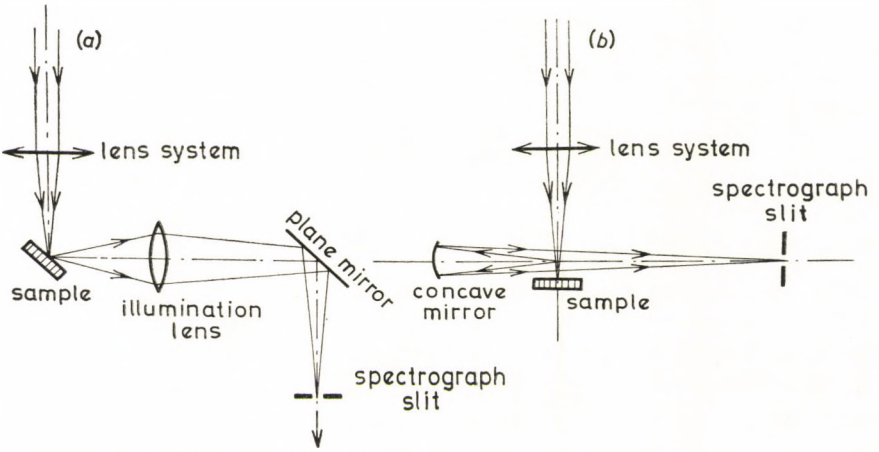


Fig. 2.132 Illumination of a spectrograph by a microplasma, using: (*a*) a collecting lens and (*b*) a concave mirror

The above process can be utilised for the purposes of emission analysis in two ways. In the method less frequently used the microplasma created directly by the output energy of the laser is the light source.[2,3] In this case, the image of the microplasma impinges on the slit of the spectrograph (Fig. 2.132). Thus, laser pulses with a duration of 0·1–0·5 μs produce a microplasma of a sufficiently high radiation density (about 10^{10} W/cm^2) which lasts for 20–100 μs.[4] The appropriate radiation density, however, can be achieved only if the laser radiation is concentrated by a microscope objective of a working distance sufficiently large to produce an evaporation crater with a diameter of about 10 μm.

Examination of the emission shows that the microplasma, produced by the laser pulses, is strongly ionised at the onset, although it rapidly cools with time.[5] Accordingly, while spectral lines with a high degree of ionisation, for example, Al III, Sn IV, Ge IV, have a very short lifetime, singly ionised spectral lines can be observed in-between individual laser pulses. The atomic lines can also be still observed for a few microseconds after the termination of the laser radiation. On the other hand, oxygen and nitrogen lines or bands do not arise even when the excitation is performed in the atmosphere. Owing to the self-absorption of the atomic lines the inference can be drawn that the neutral atoms appear about 10^{-7} s after the initiation of the laser radiation; that is, when the temperature of the area irradiated has still not reached a state of white incandescence. From the presence of the highly ionised spectral lines we can conclude that the electronic temperature of the microplasma can be as high as 20 000 K.

In a modification of the previously described process, the laser radiation is projected onto the surface of the sample material by a simple focusing lens with a short

focal length, instead of a microscope objective lens. Under such conditions the diameter of the evaporation crater can be as large as 200 μ. This leads to the formation of what is called a 'macro-vapour space'. In this process – in the case of excitation both in air and in vacuum – the spectrum of the radiation emerging shows not only lines, but bands as well if the sample vapour formed contains molecules in addition to atoms.[6] For example, in the presence of a great amount of carbon, cyanogen bands appear.

The most frequent application of laser radiation in practical spectroscopy uses a secondary spark gap for the excitation process.[7] In this method the laser radiation mainly causes the evaporation of the sample, while the discharge current of the condenser performs the excitation proper. The light source is, therefore, the microplasma. This method has an advantage over arc and spark excitation in that the laser radiation, and thus the location of the sample evaporation can be precisely controlled.

In the application of the above method the spark gap is formed by two pointed carbon electrodes 1–3 mm apart. The centre of the spark gap coincides with the axis of the laser beam (Fig. 2.133). The sample to be examined is situated at a distance of 1–2 mm below the centre of the spark gap. During excitation a d.c. voltage (1–5 kV) is applied to the electrodes, such that the condenser, C, is discharged only when the strongly ionised vapour, produced by the laser radiation, renders the spark gap conductive. (The value of the d.c. voltage depends on the form and distance apart of the electrodes and also on the analytical procedure used.) Usually the capacitance of condensers used for excitation is relatively small and varies between 1 and 2·5 μF. Thus, the maximum excitation energy is only 37·5 Ws.[8] This can be explained by the fact that the energy accumulated within the condenser is used exclusively for the excitation and not for the evaporation of the material under study. (The laser energy used for the latter purpose is generally about 0·2 Ws.[9]) The character of the spectrum obtained can be modified by the inductor, L, within the discharge circuit of the condenser. In the design of the damping resistance, R, it is necessary to bear in mind that no arc should be established between the electrodes following the discharge of the condenser.

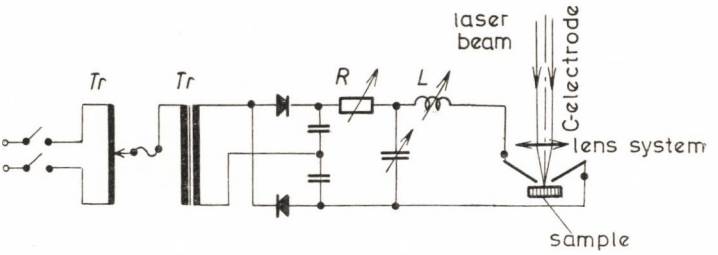


Fig. 2.133 Excitation by a secondary spark gap

The combination of laser radiation with a secondary spark gap has the advantage that the utilisation of the radiation can be increased a hundredfold. As the amount of sample necessary for analysis is only 0·1 μg this technique is essentially a micro-analytical method.[10] In many cases it is also an advantage that the spectrograph exposure time is only of the order of microseconds, as compared with a minimum of 10 s in arc and spark excitation. A further advantage is that the spectra obtained are extremely clean. Faint cyanogen bands only appear between 3800 and 3890 Å.[11]

On the other hand, the vapour of the secondary electrodes dilutes the vapour of the sample and thus influences the sensitivity of the method. It should also be borne in mind that the 'burning spot' of the spark is constant in its position, unless the laser radiation is moved to another location. Thus, the spectrum produced is not a composite one, but represents the local composition of the sample only. Furthermore, it can be established experimentally that in order to obtain good reproducibility it is better to use a single, giant pulse for the excitation. Also, the sensitivity of the detection is greater than in the case of non-controlled laser pulses.[12]

If local analysis is performed and the element composition of the sample at a certain point is required, the laser beam should be exactly focused at that location. The focus area of the radiation is in the order of microns. Precise focusing, therefore, cannot be done without the appropriate optical and mechanical alignment. Among the many methods used the most appropriate utilise the rotating stage of the microscope.[13] The laser radiation enters the eyepiece of a microscope fastened to the optical bench of the spectrograph (Fig. 2.134). The diameter of the incident beam is adjusted by

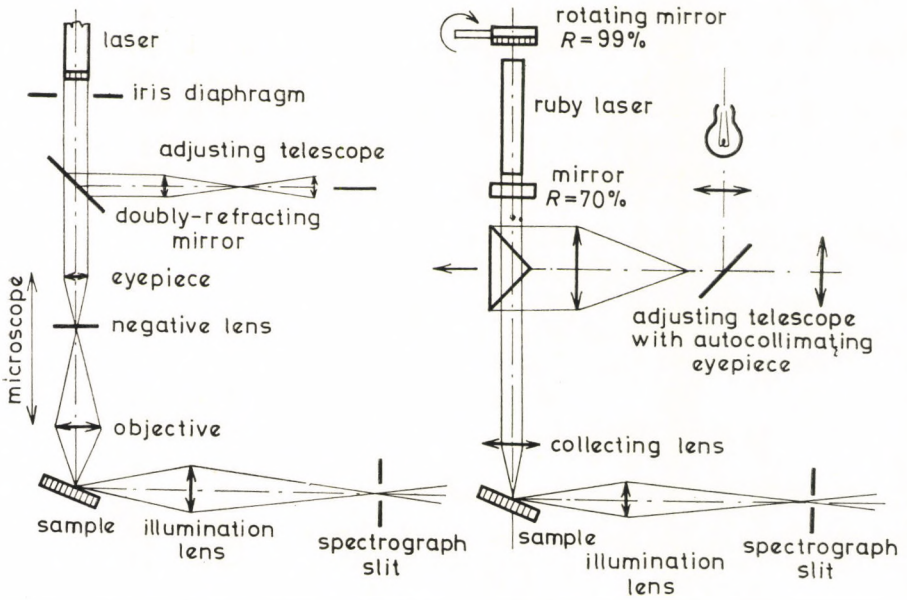


Fig. 2.134 Alignment of laser radiation by a measuring microscope

a diaphragm such that objective lenses (eyepieces) of different magnification or focal length can be used. A semitransparent mirror, forming an angle of 45° with the optical axis, allows the operator to find that part of the sample onto which the laser radiation is to be directed. In another method, a prism which can be removed from the laser beam substitutes for the semitransparent mirror (Fig. 2.134*b*). The advantage of the latter method is that it also allows adjustment of the laser to take place.[14]

In the examination of inhomogeneous samples the methods described do not secure the degree of analytical accuracy obtained by spark excitation.[15] Also, as lasers require much more complicated and expensive equipment than the spark excitation method, it is advisable to use them only in the following cases:

(1) Local analyses – that is, where the problem is not the determination of the average constitution of the sample, but the analysis of a certain part of the sample surface. For work of this type it is an advantage that the burning spot is extremely restricted and generally produces a hemispherical depression of about 10–40 μ diameter.
(2) The direct examination of samples having poor or no electrical conductance. In this case it is very convenient that there is no need for the time-consuming preparation of the sample. A great advantage is that a source of error, the contamination of the sample during the preparatory operations, is eliminated.
(3) The examination of materials where an analytical error is caused by non-reproducible changes in the chemical composition of the sample surface during the excitation.
(4) In analyses that are performed in an inert gas atmosphere and where non-volatile compounds are produced in the course of excitation.

In the examination of heterogeneous substances an average analytical result for the sample can be obtained if the laser pulses scan the surface of the sample. That is, during the excitation the laser beam is not held at a single location on the sample, but moves from place to place in a systematic way.[14] This can be achieved by using two servomotors for moving the sample in two mutually perpendicular directions.

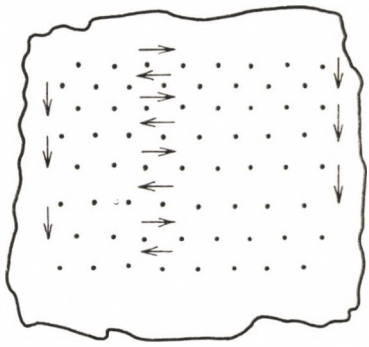


Fig. 2.135 Distribution of the sequential burning locations on the surface of the sample

These motors are electronically controlled and move the sample up or down and from side to side in such a way that about four or five burning craters are established per sq. mm (Fig. 2.135). Thus, the composite spectrum produced represents the average of more than a hundred radiation pulses from the laser.

From the analytical results obtained by this procedure, the lower limit of detectability for different elements is about the same as that achieved in ionic spark excitation. The relative deviation of the analytical results is about 8–10%. The application of the procedure, however, is strongly influenced by self-absorption.[14]

References

1. HONIG, R. E. and WEELSTON, J. R., *Appl. Phys. Lett.*, **2,** 138 (1963)
2. RUNGE, E. F., MINCK, R. W. and BRYAN, F. R., *Spectrochim. Acta*, **1964,** 733
3. DEBRAS-GUÉDON, J., *X. Coll. Spectr. Intern., College Park, 1962*, Washington, 60 (1963)
4. HAGENAH, W. D., LAQUA, K. and MOSETTI, V., *XI. Coll. Spectr. Intern., Beograd, 1963*, Abstract B-24 (unpublished)
5. ARCHBOLD, E., HARPER, D. W. and HUGES, T. P., *Brit. J. Appl. Phys.*, **15,** 1321 (1964)

6. STOICHEFF, B. F., *X. Coll. Spectr. Intern., College Park, 1962*, Washington, 399 (1963)
7. BRECH, F. and SCHUCK, K., *Lecture held at the X. Coll. Spectr. Intern., College Park, 1962* (unpublished)
8. BERNDT, M., KRAUSE, H., MOENKE-BLANKENBURG, L. and MOENKE, H., *Jenaer Jb.*, **1**, 45 (1965)
9. MOENKE, H. and MOENKE-BLANKENBURG, L., *Prakt. Chem.*, **16**, 225 (1965)
10. HAGENAH, W. D., LAQUA, K. and v. EMDE, H., *XI. Coll. Spectr. Intern., Beograd, 1963*, Abstract B-23
11. MOENKE, H., and MOENKE-BLANKENBURG, L., *Einführung in die Laser-Mikro-Emissionsspektralanalyse*, Leipzig, (1966), p. 54
12. RASBERRY, S. D., SCRIBNER, B. F. and MARGOSHES, M., *XI. Coll. Spectr. Intern., Beograd 1963* (unpublished)
13. PEPPERS, N. A., *Appl. Optics*, **4**, 555 (1965)
14. FELSKE, A., HAGENAH, W. D. and LAQUA, K., *Z. analyt. Chem.*, **216**, 50 (1966)
15. FELSKE, A., HAGENAH, W. D. and LAQUA, K., *XI. Coll. Spectr. Intern., Beograd, 1963*, Abstract B-22 (unpublished)

3 Dispersion of light

3.1 Classification of procedures used for the dispersion of light

3.1.1 Different dispersion methods

The light produced by excitation is composite, in the sense that it is composed of rays of different wavelengths. For analytical purposes the composite light must be resolved into its component parts. The instruments used for this purpose are called *dispersion devices.*

The basic operating principle of dispersion devices is as follows. With the aid of a narrow slit a divergent beam is selected from the rays coming from the light source and then made parallel by means of a lens or mirror. The parallel beam of composite light is resolved by an appropriate dispersion element into monochromatic components that are projected in different directions. The direction of projection of the monochromatic beams depends, however, on the wavelength. Finally, the different beams are brought together to form a spectral line by means of a lens or mirror. The lines are in fact the monochromatic images of the slit.

The relative performance of dispersion devices depends on the degree the originally parallel rays of different wavelength are caused to be diverted from one another. That is, the better the performance, the greater the angular dispersion (S_a) produced. The latter is the angle at which the originally parallel or coincident rays of unit wavelength difference leave the dispersion device

$$S_a = \lim \frac{\delta_1 - \delta_2}{(\lambda_1 - \lambda_2)_{\lambda_1 \to \lambda_2}} = \frac{d\delta}{d\lambda} \tag{3.1}$$

where $(\delta_1 - \delta_2)$ is the angle at which the two rays of wavelength difference $(\lambda_1 - \lambda_2)$ leave the dispersion unit.

The dispersion devices used for spectral analysis are based on two different principles, (1) the refraction of light and (2) the interference due to light diffraction. *Prism instruments* belong to the former group, while *grating instruments* belong to the latter.

The common dispersion instruments previously mentioned cannot be used for the dispersion of light having very small wavelength differences. For such purposes other devices are used that are based on interference phenomena. In spectrochemical analysis stepped gratings or Fabry–Perot interferometers are used for dispersion of small wavelength differences.

A less frequent dispersion method used in spectrochemical analysis is with the aid of filters. This method is based on removing the unwanted wavelengths by absorption of light by the filters. This relatively simple method can only be applied successfully in the case of spectra having only a few lines.

The spectral lines obtained by dispersion are in fact the optical images of the slit. Accordingly, dispersion devices can be classified depending on whether the spectral lines are true images of the slit or have been produced astigmatically. Thus, there are *stigmatic and astigmatic image dispersion instruments.* In prism dispersion the spectral lines are mainly brought about stigmatically. In grating dispersion, however, both stigmatic and astigmatic line imaging are common.

A further classification of dispersion devices is based upon whether the spectrum is examined directly by eye, indirectly by photography or by a photoelectric effect. Accordingly, there are spectroscopes, spectrographs and spectrometers.

3.1.2 Performance of dispersion devices

In the classification of dispersion devices the performance of the instruments should also be taken into consideration. The performance is generally characterised by the dispersion, spectral resolution, and illumination of the devices and by the breadth of the spectra expressed in wavelength units. In this connection it should be noted that the frequently used expression 'dispersion' is not a precise enough term. A distinction should be made between the angular dispersion and linear dispersion of the dispersion device. Generally, the term 'dispersion' only implies linear dispersion. In practice when we are talking about the dispersion (practical dispersion) of a device we in fact mean the reciprocal quantity or the physical dispersion. The two terms or concepts are frequently mixed up.

The angular dispersion of dispersion devices cannot be easily determined experimentally. For this reason spectrochemical instruments are usually characterised by the *linear dispersion. The value of the latter is the separation distance (in millimetres) in the spectrum corresponding to a wavelength difference of 1 Å.* When, for example, there is a separation of l mm in the spectrum between lines of wavelengths λ_1 and λ_2, then the linear dispersion is given by

$$S_l = \lim \frac{l}{(\lambda_1 - \lambda_2)_{\lambda_1 \to \lambda_2}} = \frac{\mathrm{d}l}{\mathrm{d}\lambda} \quad \text{mm/Å} \tag{3.2}$$

The greater the linear dispersion, the farther will the lines of unit wavelength difference be separated. The value of the linear dispersion, however, depends not only on the angular dispersion of the dispersion device but also on the distance between the spectral line and the lens (or mirror) forming the image of the latter. Thus, it follows that the linear dispersion value does not by itself characterise the performance of the instrument.

The reciprocal value of the linear dispersion gives, in the above sense, the wavelength difference corresponding to a distance of l mm as measured on the spectrum.

Its value is therefore

$$\frac{1}{S_l} = \lim \frac{\lambda_1 - \lambda_2}{l} = \frac{d\lambda}{dl} \quad \text{Å/mm} \tag{3.3}$$

The dispersion usually varies with the wavelength and is represented by the *dispersion curve*. Using the latter the dispersion level of a given dispersion device can be simply and quickly established for any wavelength. The dispersion curve also shows the spectral range in which the given dispersion device can be successfully used. In a similar manner to linear dispersion, *the dispersion curve of a given dispersion device cannot unequivocally specify the performance of the instrument.*

As opposed to linear dispersion, the value of the *resolving power is a particular characteristic of the dispersion performance. The theoretical resolving power* (R_0) of the dispersion device is given by the ratio: $\lambda/d\lambda$, where the wavelength difference, $d\lambda$, at wavelength λ, corresponds to the *theoretically* just resolved lines. Its value is

$$R_o = \frac{\lambda}{\Delta\lambda} \tag{3.4}$$

In the determination of the theoretical resolving power it should be taken into account that the strictly monochromatic image of the infinitely narrow slit cannot be a geometrical straight line. This is so because the image of the spectral line is disturbed by interference phenomena caused by the diffraction of light. At the wavelength value of the line an intensity maximum appears, while on either side minima and successively decreasing maxima are found. In the case of two adjacent spectral lines two series of interference maxima appear that are shifted with respect to one another. *According to the Rayleigh principle two adjacent spectral lines can be regarded as being resolved when the main intensity maximum of the first line coincides with the first intensity minimum of the other* (Fig. 3.1).

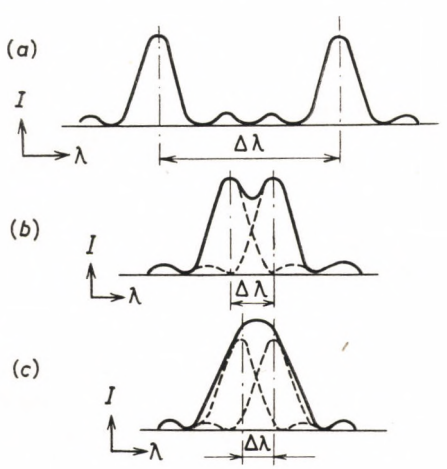


Fig. 3.1 Variation of light intensity along the spectrum: (*a*) for two neighbouring spectral lines, (*b*) two partly resolved lines, and (*c*) two unresolved lines

In practice, the theoretical resolving power of a dispersion device can never be realised. One reason is that the slit is not infinitely narrow, and the light beam producing the spectral line does not reach the plane of the spectrum at right angles. Therefore the *true resolving power* is always lower than the theoretical value.

A further characteristic of dispersion devices is their *illumination*. The illumination is a ratio that unfortunately does not give the amount of light impinging per unit area of the spectrum, but at least allows comparisons to be made between different devices. A high illumination value is advantageous since this secures a better detectability for low-intensity spectral lines.

3.2 Fundamental principles of prism dispersion

3.2.1 Refraction of light[1-3]

Dispersion by prisms is caused by the refraction of light. As is known, light refraction is produced in different optical media by rays of monochromatic or polychromatic light travelling at different velocities. As a result of this fact the wavefront of parallel rays of the same wavelength is broken, such that the direction of propagation of light is changed upon passing from one medium into another (Fig. 3.2). The extent of this

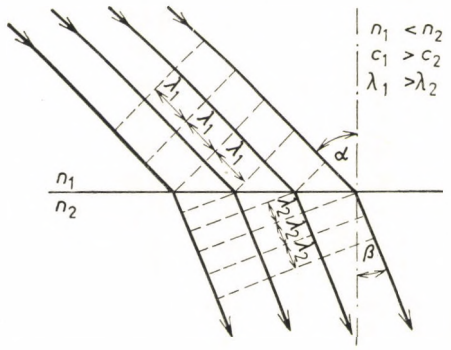


Fig. 3.2 Principle of light refraction according to the velocity of propagation

deviation from the original direction is determined not only by the velocity of the light but indirectly by the wavelength value. The following relations are valid: $v = c_1/\lambda_1 = c_2/\lambda_2$. Except in a vacuum, therefore, the velocity of light depends on its wavelength.

The *ratio* of the propagation (phase) velocity values measured *in vacuo* (c_0) and in a transparent medium (c_λ) of light having a wavelength of λ, is called the *absolute refractive index* of the medium. This value characteristic of the medium is normally designated by the capital letter N, that is

$$N = \frac{c_0}{c_\lambda} \tag{3.5}$$

In practice it rarely occurs that light enters the refractive medium from a vacuum. Therefore, the refractive index is usually referred to air for the sake of simplicity and is generally designated by the small letter n. Strictly speaking, the value of the latter depends on the atmospheric pressure and humidity, in addition to the temperature. The effect of the former two factors is so small that any deviation caused can be practically neglected in spectrochemical analysis. Thus, apart from far ultra-violet radiation the difference is negligible.

It can be stated that when a ray of wavelength λ is incident at an angle α on the plane interface between media of refractive indices n_1 and n_2, then the angle formed

between the ray and the normal to the interface is given by the relationship

$$n_{1,2} = \frac{\sin \alpha}{\sin \beta} = \text{const.} \tag{3.6}$$

From the above law (the Snellius–Descartes law) the following three important facts can be deduced:

(1) *The greater the angle of refraction, the greater is the angle of incidence.*
(2) *Incident light* normal to the surface *passes through the media without refraction.*
(3) The passage of the light is independent of the direction of propagation; that is, *the direction of the light can be reversed.*

It follows from the previous fact that when the light passes in the opposite direction – from the medium towards air – no light passes from the medium *when the angle of refraction is greater than the so-called limiting angle, such that the light is reflected back into the medium.* This phenomenon occurs when the sine of the refraction angle, β, is greater than the value of the reciprocal of the refractive index, that is, $\sin \beta > 1/n_{1,2}$. It follows from Eq. (3.6) that when $\sin \beta = 1/n$, then $\sin \alpha = 1$, thus $\alpha = 90°$. The limiting angle is about 42° for a glass–air interface. *Deviation prisms* (Fig. 3.3) are constructed on this principle.

The dispersion curve of the transparent media illustrates the relation between the refractive index of the material in question and the wavelength of the light; that is, it shows the variation of n as a function of the wavelength λ (Fig. 3.4).

For convenience 'optical materials' are usually characterised by their respective refractive indices at the wavelength of the *Fraunhofer* lines, instead of their dispersion curves. Thus, the *specific dispersion* is the refractive index difference determined at the wavelengths of H 3968·5 Å and C 6562·8 Å. This value is usually denoted by ϑ,

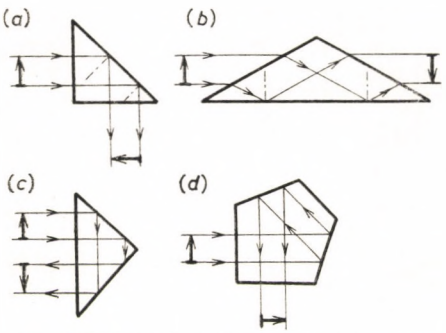


Fig. 3.3 Deviation prisms: (*a*) 90° prism, (*b*) 180° prism, (*c*) image-reversing prism, and (*d*) pentaprism

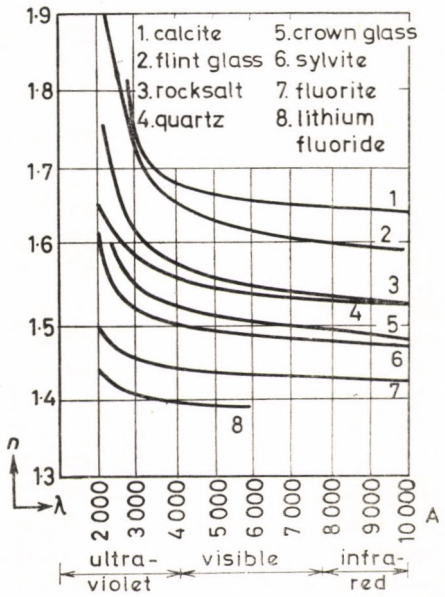


Fig. 3.4 Dispersion curves of different refractive media used in spectral analysis, between 2000 and 10 000 Å

thus $\vartheta = n_H - n_C$. On the other hand, the *mean dispersion* is the difference $n_F - n_C$, which is measured at wavelengths 4861·3 Å and 6562·8 Å. More characteristic than these dispersion quantities is the *Abbé number*, which is the value of the ratio

$$v = \frac{n_D - 1}{n_F - n_C}$$

It is evident from the above discussion that data used for the characterisation of refractive media are obtained from refractive indices of the visible part of the spectrum. These results are therefore satisfactory only for the characterisation of optical glasses.

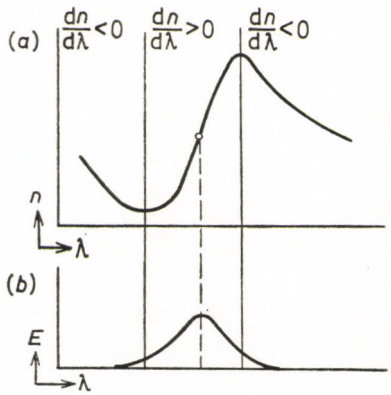


Fig. 3.5 Variation of the refractive index (*a*) and the extinction (*b*) in the normal and anomalous dispersion regions

The refractive index of materials used for optical purposes at varying wavelengths can be calculated by means of formulae that are both empirically derived and theoretically proven. Thus, for example, the refractive index of quartz between 2000 Å and 40 000 Å can be calculated by means of the following expression, provided the value of λ is taken in millimicrons

$$n_\lambda^2 = 3{\cdot}4629^2 + \frac{0{\cdot}010654}{\lambda^2 - 3{\cdot}010627} - \frac{111{\cdot}47}{100{\cdot}77 - \lambda^2} \tag{3.7}$$

Instead of the above the following also empirically derived formula can be used

$$n_\lambda = n_0 \frac{C}{(\lambda - \lambda_0)^\alpha} \tag{3.8}$$

This is the *Hartmann formula*. By using this formula the refractive index of any optical material can be calculated with sufficient accuracy, provided the refractive index n_0 determined at wavelength λ_0 and the values of the constants C and α are known. These constants can, however, be easily obtained via experimentally determined refractive indices at certain wavelengths. The value of the exponent generally varies between $0{\cdot}5 < \alpha < 2$, although it can usually be taken to be 1, especially when the wavelength interval to which it refers is not very broad. This is also true for the constant C if the wavelength interval is not too broad.

For normal dispersion the value of the refractive index increases with decreasing wavelength, that is, $dn/d\lambda < 0$. *In the case of prism dispersion, therefore, the ultraviolet rays are refracted to a greater extent than the visible rays.* The variation of the

dispersion is not even, but increases as the wavelength approaches that part of the spectrum where the prism material absorbs strongly (Fig. 3.5). At the beginning of the spectral interval the dispersion curve passes through a maximum, then the value of the refractive index starts dropping, or in other words *anomalous dispersion* is encountered.

Anomalous dispersion is characterised not only by the fact that the refractive index decreases with decreasing wavelength, that is, $dn/d\lambda < 0$, but that the refractive medium increasingly absorbs the light as the wavelength drops in this interval. The point of maximum absorption coincides with the inflection point of the dispersion curve (Fig. 3.5*b*). *In the dispersion devices used in spectrochemical analysis anomalous dispersion should be taken into account.* This is explained by the fact that in practical analysis one always works at wavelength intervals that lie in between the absorption bands of the ultra-violet and infra-red regions.

References

1. SCHALLREUTER, W. in *Grimsehls Lehrbuch der Physik*, 13th edn, Leipzig, vol. 3 (1955)
2. KOHLRAUSCH, F., *Praktische Physik*, 19th edn, Leipzig, vol. 2 (1955)
3. JOOS, G., *Theoretical Physics*, New York (1934)

3.2.2 Prism dispersion[1,2]

In instruments based on the refraction of light, optical prisms are used for the resolution of the polychromatic light into its components. The prism is a wedge-shaped body of some transparent material that is limited by at least two intersecting plane surfaces. The angle ε formed between these two surfaces is the *refractive angle of the prism* (Fig. 3.6). The line of intersection of the two planes is the edge of the prism, while the *main section of the prism* lies in a plane perpendicular to the edge. The side of the prism opposite the refractive angle is called the *base of the prism.*

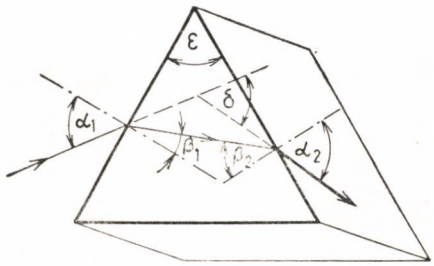


Fig. 3.6 Monochromatic light ray passing through the principal section of a prism

Incident monochromatic light enters the prism from the air in such a way that its path is directed towards the normal of incidence, while the direction of the light leaving the prism bends away from the normal of the interface. This double change in the direction of propagation of the monochromatic light passing through the prism determines the *deviation angle* δ.

In the plane of the *principal section of the prism* the rays traverse the prism according to comparatively simple rules. In the first case, those rays that fall upon the prism in the plane of the principal section remain in this plane after refraction. The mono-

chromatic light follows a path through the prism such that the sum of the angles β_1 and β_2 formed with the normal of incidence and the normal of departure, respectively, is equal to the angle of refraction of the prism, that is

$$\varepsilon = \beta_1 + \beta_2 \tag{3.9}$$

In the given triangle the following relation holds:

$$\varepsilon + (90° - \beta_1) + (90° - \beta_2) = 180°.$$

The deviation angle of monochromatic rays passing in the principal section of the prism can be calculated easily when the triangle has a complementary angle to the angle of the deviation δ. The sum of the internal angles of this triangle is $(180° - \delta) + (\alpha_1 - \beta_1) + (\alpha_2 - \beta_2) = 180°$, thus

$$\delta = \alpha_1 + \alpha_2 - \varepsilon \tag{3.10}$$

The deviation angle of monochromatic light passing through the principal section of the prism depends, therefore, on the angle of incidence, the refractive angle of the prism, and the refractive index of the prism material.

The passage of rays that fall obliquely onto the principal section of the prism are not easy to trace within the latter. Therefore, from the practical point of view, *the refractive edge of the prism should be perpendicular to the direction of the incident rays.*

The laws of prism dispersion are further simplified if the passage of the monochromatic light in the prism is *symmetrical* ($\alpha_1 = \beta_2$). In the case cited the light path is parallel to the base of the prism (Fig. 3.7). It then follows that *in the case of a symmetrical light path the angle of deviation of the prism is the smallest possible.* The prism is, therefore, in a condition of minimum deviation with respect to the light of a particular wavelength and is said to be in the so-called *minimum position. In the minimum position the dispersion of the prism is at its highest.*

The above-mentioned rules can be validated in the following way. According to the Snellius–Descartes law [Eq. (3.6)] $\sin \alpha_1 = n \,.\, \sin \beta_1$ and $\sin \alpha_2 = n \,.\, \sin \beta_2$. By the addition of the two equations we obtain

$$\sin \alpha_1 + \sin \alpha_2 = n \,(\sin \beta_1 + \sin \beta_2)$$

or by the application of trigonometric transformation

$$2 \sin \frac{\alpha_1 + \alpha_2}{2} \cos \frac{\alpha_1 - \alpha_2}{2} = 2n \sin \frac{\beta_1 + \beta_2}{2} \cos \frac{\beta_1 - \beta_2}{2}$$

As Eq. (3.9) applies in this context, we obtain the expression

$$\sin \frac{\alpha_1 + \alpha_2}{2} = n \sin \frac{\varepsilon}{2} \frac{\cos \frac{1}{2}(\beta_1 - \beta_2)}{\cos \frac{1}{2}(\alpha_1 - \alpha_2)} \tag{3.11}$$

For a symmetrical path $\alpha_1 = \alpha_2 = \alpha$, thus $\beta_1 = \beta_2$. In this case, however, the value of the ratio which is the multiplier of $\sin \varepsilon/2$ is unity. If this is true, then

$$\sin \alpha = n \,.\, \sin \frac{\varepsilon}{2}$$

Thus, according to Eq. (3.10) $\delta = 2\alpha - \varepsilon$ which is the minimum value. This can be justified as follows. In the opposite case, that is, when the prism is not in minimum position, $\alpha_1 > \alpha_2$, and $\beta_1 > \beta_2$. Since otherwise $\alpha_1 > \beta_1$ and $\alpha_2 > \beta_2$, thus it follows that $(\alpha_1 - \beta_1) \geqq (\alpha_2 - \beta_2)$, thus $(\alpha_1 - \alpha_2) > (\beta_1 - \beta_2)$, so that

$$\cos\frac{1}{2}(\beta_1 - \beta_2) > \cos\frac{1}{2}(\alpha_1 - \alpha_2)$$

that is

$$\frac{\cos\frac{1}{2}(\beta_1 - \beta_2)}{\cos\frac{1}{2}(\alpha_1 - \alpha_2)} > 1$$

The value of the cosine function increases as the deviation angle decreases. It follows from the previous relation on the basis of Eq. (3.11) that $\alpha_1 + \alpha_2 > 2\alpha$, which then means, according to Eq. (3.10) that the value of the deviation angle, δ, is greater than the one calculated in the previous case. It is thereby proved that the *smallest deviation angle is obtained in the case of a symmetrical light path.*

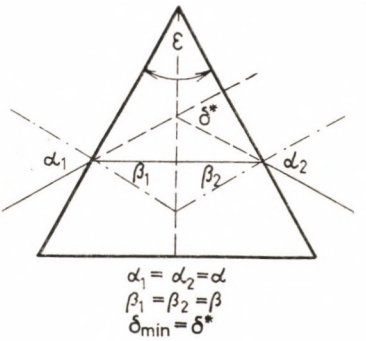


Fig. 3.7 Symmetrical arrangement of rays

In the application of the 'prism equation', valid for the case of the minimum position of the prism, it should be taken into consideration that for a symmetrical path the angle of incidence is $\alpha = 1/2(\delta + \varepsilon)$ and the angle of refraction is $\beta = \varepsilon/2$. On the other hand, the value of the refractive index (from the Snellius–Descartes law) is, $n = \sin\alpha/\sin\beta$, thus

$$n = \frac{\sin\frac{\delta_{\min} + \varepsilon}{2}}{\sin\frac{\varepsilon}{2}}$$

By the transformation of the prism equation we obtain the formula

$$\sin\frac{\delta_{\min} + \varepsilon}{2} = n\sin\frac{\varepsilon}{2} \tag{3.12}$$

From the above equation it is seen that the deviation angle of the prism in its minimum position depends exclusively on the refractive angle of the prism and its re-

fractive index or, indirectly, on the wavelength of the light. The shorter the wavelength, the greater is the angle of deviation for any given prism.

It follows from Eq. (3.10) that the relationship given below exists between the angle of deviation, δ_{min}, and the angle of refraction ε, of a prism in its minimum position

$$\delta_{min} = 2\alpha - \varepsilon \tag{3.13}$$

where α is the angle of incidence and departure. For any given prism the angle of incidence and departure is known to depend on the wavelength of the light. Rays of given wavelength pass through the prism in a symmetrical fashion only when the latter has a certain refractive index value and a certain refractive angle.

The collective angle of deviation of a prism system consisting of a number of prisms in their minimum positions is the sum of the individual angles of deviation (Fig. 3.8).

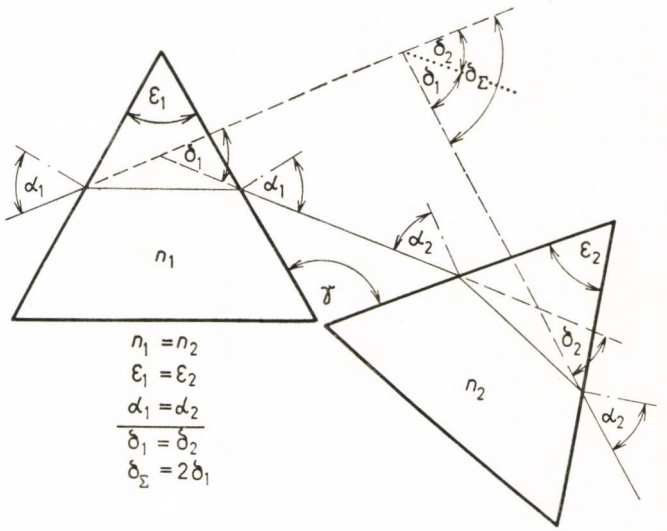


Fig. 3.8 Total deviation angle of two prisms

In the summation, however, all the deviation angles are positive only when the individual directions of deviation are all the same. Taking this into account we obtain

$$\delta_\Sigma = \delta_1 + \delta_2 + \ldots \delta_n \tag{3.14}$$

Prisms of the same geometry and having the same index of refraction are in the minimum position only when the angle between the surfaces of neighbouring prisms, γ, satisfies the following relationship

$$\sin\frac{\gamma}{2} = n \sin\frac{\varepsilon}{2} \tag{3.15}$$

In the *calculation of the angular dispersion of a prism* our starting point is Eq. (3.1), such that the angular dispersion can be written as

$$S_{a(p)} = \frac{d\delta}{d\lambda} = \frac{d\delta}{dn}\frac{dn}{d\lambda} \tag{3.16}$$

The left-hand member of the product depends on the geometrical form of the prism, while the right-hand member depends on the dispersion of the prism material.

In the calculation of the left-hand differential we obtain by the differentiation of the 'prism equation' (3.12) that

$$\frac{\mathrm{d}\delta}{\mathrm{d}n} = \frac{2 \sin \frac{\varepsilon}{2}}{\cos \frac{\delta + \varepsilon}{2}} \tag{3.17}$$

From Fig. 3.9 it can be seen that

$$2 \sin \frac{\varepsilon}{2} = 2 \frac{b/2}{c} = \frac{b}{c}$$

or according to Eq. (3.13)

$$\cos \frac{\delta + \varepsilon}{2} = \cos \alpha = \frac{a}{c}$$

By substitution of the above two values and by using the relationships obtained in Eqs (3.16) and (3.17) we see that

$$S_{a(p)} = \frac{\mathrm{d}\delta}{\mathrm{d}\lambda} = \frac{b}{a} \frac{\mathrm{d}n}{\mathrm{d}\lambda} \tag{3.18}$$

The greater the angular dispersion of the prism in minimum position, the greater is the value of the ratio b/a and the dispersion of the prism material. Strictly speaking, Eq. (3.17) is only valid at the wavelength for which the passage of the light is symmetrical.

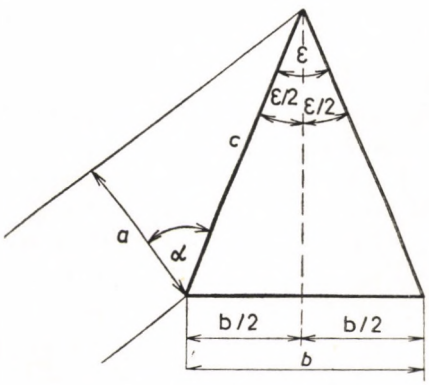


Fig. 3.9 Angular dispersion of a prism

A more general relationship is found if Eq. (3.17) is transformed by the following trigonometrical relationship

$$\cos \frac{\delta + \varepsilon}{2} = \sqrt{1 - \sin^2 \frac{\delta + \varepsilon}{2}} \tag{3.19}$$

Furthermore, according to Eq. (3.12)

$$\sin\frac{\delta+\varepsilon}{2} = n\sin\frac{\varepsilon}{2}$$

thus, Eq. (3.17) can be modified to have the form

$$\frac{\mathrm{d}\delta}{\mathrm{d}n} = \frac{2\sin\frac{\varepsilon}{2}}{\sqrt{1-n^2\sin^2\frac{\varepsilon}{2}}} \tag{3.20}$$

The dispersion $\mathrm{d}n/\mathrm{d}\lambda$\ n Eq. (3.18) can be easily calculated by using the simplified Hartmann formula

$$n = n_0 + \frac{C}{\lambda-\lambda_0}$$

in which n_0 is the refractive index at the appropriately chosen wavelength λ_0. C is a constant whose value depends on the prism material and can be experimentally determined. When the value

$$\frac{\mathrm{d}n}{\mathrm{d}\lambda} = \frac{-C}{(\lambda-\lambda_0)^2}$$

calculated from the above equation and the value $\mathrm{d}\delta/\mathrm{d}n$ given in Eq. (3.10) are substituted into Eq. (3.16), then for the *angular dispersion of a prism* the following formula is obtained

$$S_{a(p)} = \frac{\mathrm{d}\delta}{\mathrm{d}\lambda} = \frac{-2C\sin\frac{\varepsilon}{2}}{(\lambda-\lambda_0)^2\sqrt{1-n^2\sin^2\frac{\varepsilon}{2}}} \tag{3.21}$$

From this formula it follows that *the angular dispersion of a prism in its minimum position varies with the wavelength of the light solely as a function of the refractive angle of the prism.*

It is easy to see that when light passes in a symmetrical fashion through any number of geometrically similar prisms of the same material, then *the angular dispersion of the prism system is proportional to the number of prisms.*

The above relationships are much simpler if the prism is a 60° prism having an equilateral section. In this special case Eq. (3.21) will have the form

$$S_{a(p)\,60^\circ} = \frac{-C}{(\lambda-\lambda_0)^2\sqrt{1-\frac{n^2}{4}}} \tag{3.22}$$

From an optical point of view, the 30° or so-called *Littrow prism* should also be regarded as one of refractive angle 60°. This type of prism is based on the principle of *autocollimating dispersion.* This is different from simple prism dispersion in that the light entering the prism does not immediately leave, but is first reflected from

its silvered back-surface (Fig. 3.10). The prism, therefore, first disperses the incident composite rays and then reflects the different rays produced in such a manner that they are dispersed further. Consequently *the angular dispersion of the Littrow prism is equal to that of a 60° prism that has a base length twice that of the Littrow prism.* Another consequence of its special construction is that the Littrow prism produces a deviation of the light of almost 180°.

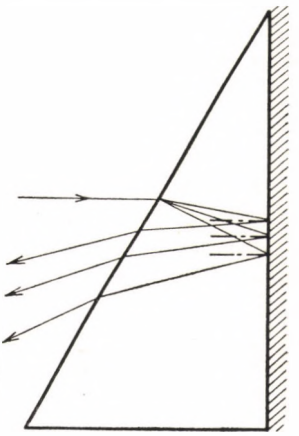


Fig. 3.10 Schematic representation of light dispersion in a Littrow prism

The relationships given above are valid when the light passes nearly symmetrically through the prism. This can occur provided the angle of incidence is appropriate in relation to the prism material, the refractive angle of the prism, and the wavelength value. The value of the angular dispersion depends also on the angle of incidence. Thus the angle of deviation of rays having the same wavelength is different if these rays reach the refractive surface of the prism at different angles of incidence. *The prerequisite of prism dispersion is that the rays impinging on the prism should be parallel.*

References

1. KAYSER, H., *Handbuch der Spektroskopie*, Leipzig, vol. I (1903)
2. NACHTRIEB, N. H., *Principles and Practice of Spectrochemical Analysis*, New York–Toronto–London (1950)

3.3 Fundamental principles of prism dispersion instruments

3.3.1 The construction of prism dispersion instruments

The necessary geometrical condition for the formation of the spectral lines is that the light beam should enter the dispersion device through a narrow slit. This is necessary since the spectral lines are the images of the *illuminated* slit. In the case of prism dispersion the spectral lines are in general the true images of the slit, so they usually arise stigmatically. Apart from the dispersion produced by the Fery prism, astigmatic imaging only occurs in unfavourable conditions in prism dispersion equipment.

From the discussion in Sub-section 3.2.2 it is evident that the angle of deviation produced by the prism depends, among other things, on the angle of incidence.

The deviation angles of rays of the same wavelength that have different angles of incidence are therefore different. This undesirable effect can be eliminated by a *collimating lens* placed before the prism such that the rays coming from the slit are made parallel. This can only be achieved if the distance between the slit and the collimating lens is equal to the focal length of the lens.

The deviation of the parallel incident rays is such that the monochromatic transmitted rays leaving the prism in different directions tend to be parallel; that is, they keep their parallel orientation after prism dispersion. This advantageous effect is only realised when the rays are parallel to the principal section of the prism. Therefore, it is important to place the prism in an appropriate position relative to the incident light.

In principle, the parallel rays leaving the prism in various directions according to their different wavelengths are brought together to form spectral lines, provided the distance between the slit and the collimating lens are properly determined. The sharpness of the spectral lines and thus the quality of the spectrum are considerably improved if the light is condensed to form spectral lines by means of a *camera lens* situated behind the prism. The camera lens projects the images of the slit onto its focal plane in such a way that their relative positions correspond to the wavelength of the dispersed rays coming from the prism. These *images* are, therefore, the *spectral lines* and their collection is the *spectrum.*

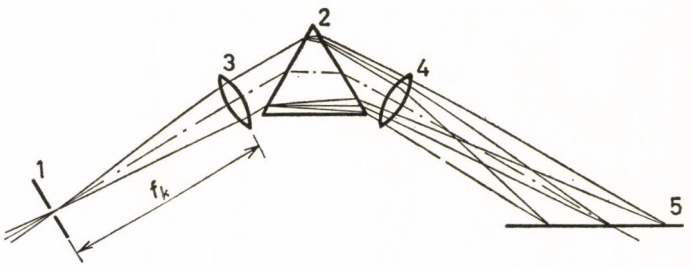


Fig. 3.11 Schematic optical diagram of a single-prism light dispersive apparatus

The spectrum produced in the way described above can be utilised in several ways. When a photographic plate is placed at the focal plane of the camera lens, a photograph of the spectrum can be obtained. These instruments are called *prism spectrographs.* The spectrum can, however, be directly observed when the image formed by the condensing lens is viewed by an eyepiece. The two lenses thereby form a telescope set to infinity. These instruments are called *prism spectroscopes.* The spectrum projected by the camera lens can also be examined photoelectrically. The intensity of the light reaching slits placed at appropriate locations on the focal plane can be determined by sensing elements – usually photomultipliers. Accordingly, these devices are called *polychromators,* while the complete equipment is called a *spectrometer.*

The *main components of prism dispersion devices* are the following: (1) slit, (2) prism, (3) collimating lens, (4) camera lens, and (5) photographic plate holder (for spectrographs), or an *eyepiece* (for spectroscopes), or *exit slits* (for spectrometers) (Fig. 3.11).

3.3.2 The slit

The spectral lines are generally the true images of the slit in prism dispersion. The form, width, and height of the spectral lines, therefore, depend on the form, width, and height of the slit.

The edges of the ideal slit are straight lines, perpendicular to the optical axis of the dispersion device and parallel to the refractive edge of the prism. The slit is formed by two metal sheets whose parallel edges are ground to have a wedge shape and which are mirror images of each other.

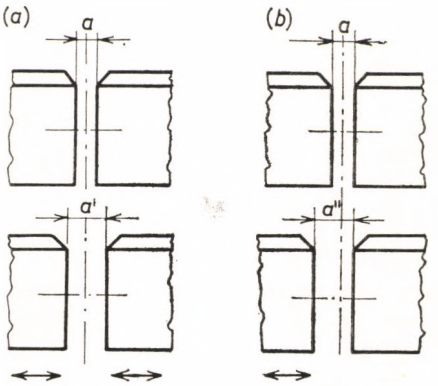


Fig. 3.12 Schematic drawing of a symmetrical slit (*a*) and an asymmetrical slit (*b*) at different slit widths

The distance between the edges of the slit is the *slit width.* This can be altered by a micrometer screw in modern equipment such that the central line of the slit always remains on the optical axis of the instrument. The width of the slit can be reproducibly set to an accuracy of 0·001 mm and can also be measured to the same degree of accuracy. As opposed to the case of the *symmetrical slit* just described, the width of an *asymmetrical slit* can be altered by only moving one of the slit jaws, while the other remains fixed. Thus the central line of the slit does not remain on the optical axis of the instrument during changes of the slit width (Fig. 3.12).

In the *selection of the slit width* it should be taken into account that in prism dispersion the spectral lines are generally the true images of the slit. Upon reducing the slit width, therefore, the photocurrent on unit area of the spectral line does not change until the light losses, due to diffraction on the slit, become significant. It can be shown experimentally that when the slit is only a few microns wide the effect of diffraction is so great that only a small proportion of the photoenergy falling on the slit reaches the collimating lens.

In the calculation of the smallest optimum slit width, $a_{\min}$, one starts from the consideration that, due to diffraction occurring at the slit, the maximum photocurrent density is found in the direction of the optical axis of the slit. This is the

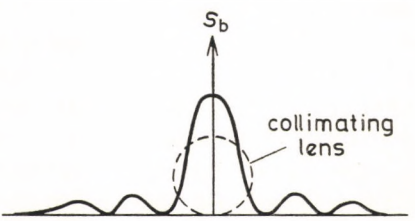


Fig. 3.13 Distribution of diffraction radiance on a slit in the plane of a fully illuminated collimating lens

most intense maximum and is surrounded by minima and gradually decreasing maxima on both sides (Fig. 3.13). It is easy to see that upon a significant reduction of the slit width, light is not lost until the width of the most intense maximum in the plane of the collimating lens is equal to, or smaller than, the diameter of the lens. According to Fig. 3.14 this occurs when the proportion $AC : CD = BE : EF$ is realised and when the distance CD corresponds to the half-wavelength, $\lambda/2$, of the light. In this case the allowable smallest slit width, $d_{\min}$, can be determined from Fig. 3.14

$$d_{\min} = \lambda \frac{f}{d}$$

as where $AC = d_{\min}$, $2EF = d$, the diameter of the collimator lens, and $BF = f$, the focal length of the latter. Considering that f/d is the aperture, k, of the collimator lens, it can be written that

$$d_{\min} = k\lambda \tag{3.23}$$

Therefore, in principle, the larger the aperture of the collimating lens and the longer the wavelength of the light, the greater is the smallest optimum slit width. However, from practical experience it is generally a drawback if the slit width is smaller than 0·01 mm in the ultra-violet region or smaller than 0·02 mm in the visible region of the spectrum.

Too large a *slit height* is a disadvantage both theoretically and practically. This is because the collimating lens only makes those rays that are focused by the lens parallel. In reality no trouble is caused if rays also reach the collimating lens where the distance from the focus is negligible compared to the focal length. For this reason it is *advisable to have the focal length of the collimating lens as large as possible*

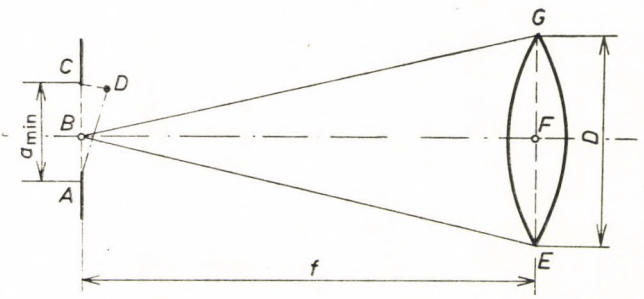


Fig. 3.14 Principle of determining the optimum smallest slit width

When the height of the slit is too large the collimating lens projects the rays onto the prism such that those rays that fall on the slit furthest from the optical axis have a different angle of incidence from those coming from the middle part of the slit. The former rays do not pass through the principal section of the prism and leave it at a different angle compared to those emerging from the middle portion of the slit. As a consequence, the camera lens forms curved images of identical wavelength rays instead of straight ones. As these curved parts of the spectral lines arise astigmatically, they are, therefore, *not* true images of the slit. In prism dis-

persion the curvature of the spectral lines is such that the convex side always points towards the longer-wavelength end of the spectrum (Fig. 3.15). A further consequence of astigmatism is that the intensity distribution of the spectral line is not uniform along its length. The highest intensity is found in the central portion and drops



Fig. 3.15 Detail of the spectrum obtained at high slit value during the prismatic dispersion of light

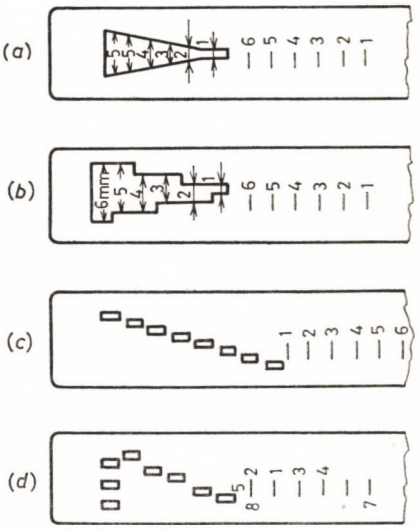


Fig. 3.16 Commonly used slit diaphragms: (*a*) wedge type, (*b*) gradual type, (*c*) stepped type, and (*d*) comparative type

towards either end of the line. These irregularities only occur for very great slit heights. Generally, the best conditions are obtained with non-magnifying spectrographs when the length of the spectral lines in the focal plane of the camera lens is about 1 mm. The optimal slit height increases with the magnification of the spectrograph.

The slit height is controlled by a *diaphragm* in prism spectrographs. Diaphragms are made of thin metal sheets that slide perpendicular to the vertical axis of the slit. The diaphragm is situated immediately in front of the slit and the openings cut into it leave certain portions of the slit uncovered. Slit diaphragms may have openings of different form and arrangement for use in measurement of different types (Fig. 3.16).

The required slit height, h_s, can be calculated when the *magnification of the dispersion device*, that is, the ratio of the focal lengths of the collimating lens, f_{col}, and the camera lens, f_{cam}, is known. Thus the slit height is given by the formula

$$h_s = h_l \frac{f_{col}}{f_{cam}} \tag{3.24}$$

where h_l is the desired length of the spectral line.

3.3.3 The prism

The most important part of the dispersion device, which is based on the refraction of light, is the prism. The resolving power, the illumination, and the performance of the equipment depend on the dimensions of the prism. The material of the prism

defines the spectral region the instrument can be used in. The purity of the prism material and its method of construction greatly influence the operation of the equipment. For example, if the prism is incorrectly located there will be an increase in astigmatism.[1] In prism dispersion the astigmatism is smallest when: (1) the edge of the prism is completely parallel to the slit, (2) the rays passing through the slit are parallel with the optical axis of the prism, (3) the rays pass through the principal section of the prism, and (4) the prism is in its minimum position.

Generally, the refractive angle of quartz prisms is 60°. The Littrow prism, used in autocollimated equipment, has a refractive angle of 30°, although its function is principally the same as that of the 60° prism. Thus, in principle, the Littrow prism can also be regarded as having a refractive angle of 60°. The refractive angle of the usual, simple glass prisms is normally 63°.

Table 3.1

Angular dispersion of prisms having different angles of refraction and refractive indices[2,5]

Refractive index	Refractive angle				
	45°	60°	60°56′	64°	67°22′
1·5	0·935	1·512			2·00
1·6	0·968	1·667		2·00	
1·7	1·008	1·899	2·00		

Table 3.1 lists the angular dispersion, in the minimum position of prisms that have different refractive angles and refractive indices, while Fig. 3.17 shows the light losses occurring when prisms of different refractive angles and having a refractive index of 1·5 are used. From these data it can be seen that the value of the

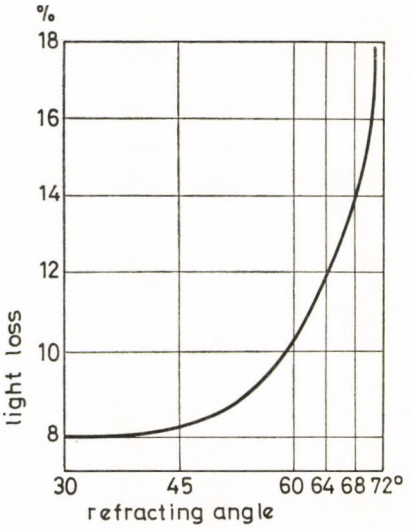


Fig. 3.17 Relationship between the refractive angle of a prism having a refractive index of 1·5 and the light loss caused by its application[2]

angular dispersion increases as the refractive angle increases, while at the same time there is a rapid increase in the loss of light when the refractive angle is greater than 60°. The optimum angular dispersion value, $d\delta/d\lambda$, is realised at a particular refractive angle when the light reflected from the surface of the prism is totally polarised. Thus, it is advisable to obtain a value of 2·00 for the refractive angle of quartz or glass prisms.

Requirements concerning the *prism material* are quite strict. It is absolutely necessary to have an optically perfect, homogeneous, flawless, and possibly isotropic material for this purpose. It is also necessary for its transmittance to be high over a broad spectral interval and for its dispersion to be large and the temperature coefficient small. A further requirement is that the surface of the prism material must be resistant to atmospheric effects (humidity, etc.) and easily workable to produce an optically flat surface.

The *applicability of dispersion devices* for specific tasks is determined by the prism material and especially by the refractive index, the wavelength dependence of the refractive index, and the transmission and temperature coefficients of the prism material. The prism material should therefore be chosen according to the spectral interval to be studied.

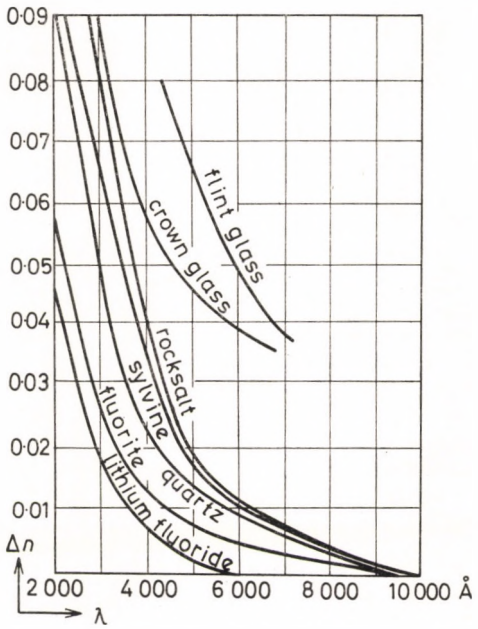


Fig. 3.18 Correction curves for the calculation of the refractive index of some important prism materials

Refractive index data of the most important prism materials are summarised in Table 3.2. Depending on the wavelength, the value of the correction Δn can be read from Fig. 3.18. On the other hand, the data in Table 3.3 refer mainly to the D-line of Na (5889·95 Å) or the visible part of the spectrum.

As the spectral lines generally used in analysis are found in the ultra-violet part of the spectrum, the most important prism material is quartz. The simple quartz prism is, however, not applicable for dispersion owing to the birefringence of the quartz crystals. The refractive index of the ordinary ray is $n_{D_o} = 1{\cdot}5442$ at a wavelength of 5889·95 Å, while the refractive index of the extraordinary ray is $n_{D_e} = 1{\cdot}5533$.

This has the consequence that monochromatic light passes along two separate paths within the simple quartz prism. For example, light of wavelength 5889·95 Å leaves with a deviation angle of 0·8° for a 60° quartz prism whose crystal optical axis is parallel to the refractive edge. Thus, when such a prism is used, all spectral lines are doubled.

Table 3.2

Refractive indices of the more important prism materials between 2000 and 10 000 Å[4]

Prism material	Refractive index	Prism material	Refractive index
lithium fluoride	$n = 1{\cdot}39 + \Delta n$	rock salt	$n = 1{\cdot}53 + \Delta n$
fluorite	$n = 1{\cdot}43 + \Delta n$	crown glass	$n = 1{\cdot}48 + \Delta n$
quartz	$n = 1{\cdot}53 + \Delta n$	flint glass	$n = 1{\cdot}70 + \Delta n$
sylvine	$n = 1{\cdot}48 + \Delta n$		

Table 3.3

Optical data of the more frequently used prism materials
[n_D = refractive index measured at 5889·95 Å, $n_F - n_C$ = average dispersion, $(d\delta/d\lambda)_D$ = dispersion of 60° prism measured at 5895·6 Å, $(dn/dt)_D$ = temperature coefficient of refractive index n_D[5–8]]

Prism material	n_D	$n_F - n_C$	$(d\delta/d\lambda)_D$	Transmittance (Å)	$(dn/dt)_D$
lithium fluoride	1·39177	0·0040	$0{\cdot}286.10^{-5}$	1 000–150 000	$-2{\cdot}3.10^{-5}$
fluorite	1·43385	0·0045	$0{\cdot}333.10^{-5}$	1 250– 95 000	$-1{\cdot}0.10^{-5}$
quartz, fused	1·45848	0 0067	$0{\cdot}517.10^{-5}$	1 850– 35 000	$-0{\cdot}6.10^{-5}$
quartz, natural	1·54426	0 0078	$0{\cdot}628.10^{-5}$	1 800– 40 000	$-0{\cdot}5.10^{-5}$
'Uviol' crown glass	1·5035	0 0078	$0\,616.10^{-5}$	2 100– 8 000	$\sim -0{\cdot}5.10^{-5}$
crown glass	1·5271	0 0087	$0{\cdot}700.10^{-5}$	3 500– 20 000	$\sim -0{\cdot}5.10^{-5}$
rock-salt	1·5443	0·0127	$0{\cdot}938.10^{-5}$	1 750–145 000	$\sim -3{\cdot}7.10^{-5}$
light flint glass	1·5804	0·0139	$1{\cdot}144.10^{-5}$	3 800– 25 000	$\sim -0{\cdot}5.10^{-5}$
heavy flint glass	1·6555	0·0190	$1{\cdot}703.10^{-5}$	3 800– 25 000	$\sim -0{\cdot}5.10^{-5}$

The above fault is eliminated if a composite quartz prism, the so-called *Cornu prism*, is used (Fig. 3.19). This is a composite quartz prism of refractive angle 60° made from quartz crystals of right-hand and left-hand rotation. The two 30° prisms are joined such that their optical axes are perpendicular to the plane bisecting the refractive angle of the prism. The contact surface of the two halves is optically flat to within a fraction of the wavelength of the light applied. In this way it is possible to ensure that practically no light reflection occurs at the contact surface.

For the construction of quartz prisms *natural quartz crystals* are best. Suitable crystals are very rare as the majority are not of the correct size, nor are they optically perfect. Only recently has it proved possible to manufacture *fused quartz* crystals that can be used for the construction of good prisms. These artificial crystals are

relatively economical to make and also have the useful property of not being birefringent. In place of the composite Cornu prism, therefore, a simple 60° prism can be used. In comparison to natural quartz, fused silica has the drawback that the lower limit of acceptable transmission lies at a wavelength 100 Å, greater than in the case of natural quartz.

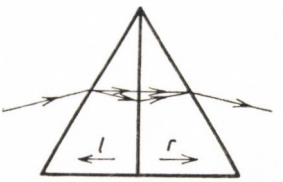


Fig. 3.19 Path of monochromatic light in the Cornu prism

For the dispersion of visible light, glass prisms are used. As can be seen from the above data, the dispersion of glass is much greater in this region than that of quartz. On the other hand, the optical glasses generally employed usually absorb light with a wavelength shorter than 3600–3800 Å. Glass prisms are, therefore, normally used in the visible and near-ultra-violet part of the spectrum. Just recently, special optical glasses have been produced that can also be used in the region 3100–6000 Å, that is, in the ultra-violet region.

In glass-optics spectrographs, and especially in spectroscopes, *composite prisms* are frequently used instead of 60° prisms. Composite prisms are usually made with the components stuck together by an adhesive (Canada balsam) whose refractive index is approximately the same as that of the optical glasses. Composite prisms, therefore, have the advantage that there is a much smaller light loss due to reflection, than in a prism system of several individual components having the same angular dispersion.

In many instances, the correct choice of the refractive angles, refractive indices, and relative base lengths of the prisms will enable *constant-deviation prisms* to be produced. These are composite prisms that disperse the incident light at a definite angle, such that a ray passing through the prism in a symmetrical fashion leaves at a constant angle. By rotating the prism or by appropriately altering the angle of incidence it is possible that light of any wavelength always leaves at the same angle.

Among the composite glass prisms most widely used is the *Abbé prism*, or as it is sometimes called, the *Pellin–Broca* prism.[9–15]

This is composed of two 30° refracting prisms (*ABC* and *ADE*) and a 90° deviating prism (*BCD*) (Fig. 3.20*a*). The light entering the prism is dispersed by the first refracting prism and the dispersed rays are totally reflected at side *BD* of the prism. Thus, those rays which travel approximately parallel to the base of the refracting prism *ABC* are deviated by an angle of 90°. The deviated rays are further dispersed by the other refracting prism such that they are caused to deviate in the opposite direction compared to the first deviation effect. The deviations produced by the two prisms therefore compensate each other. Thus, the path of the emerging light is determined *solely* by the reflection occurring at the 90° prism, which is independent of the wavelength. When the prism is rotated about the axis O then the light of any wavelength leaves the prism always with a deviation of 90°. The light always traverses the 60° prism, made from two 30° refracting prisms, in a symmetrical fashion. The angular dispersion of the Abbé prism, therefore, corresponds to that of a 60° prism.

Its resolving power is determined by the total base length of the two refracting prisms. As the Abbé prism is a glass prism it can only be used in the visible part of the spectrum.

For the examination of the ultra-violet part of the spectrum the *Straubel 90° constant-deviation quartz prism* can be used. The construction of this prism is identical to that of the Abbé prism (Fig. 3.20*b*). The refracting prism, *ABC*, is made of quartz

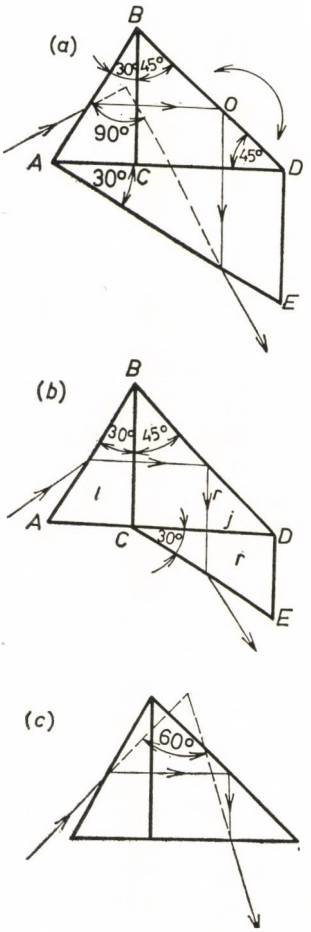


Fig. 3.20 Constant-deviation prisms: (*a*) 90° Abbé prism, (*b*) 90° Straubel quartz prism, and (*c*) Calbier 60° glass prism

such that the optical axis is parallel to the surface *BC*. On the other hand, the optical axes of the deviating prism, *BCD*, and the refracting prism, *CDE*, are perpendicular to the former. Thus, the error caused by the birefringence of natural quartz is eliminated.

Less frequently used is the 60° constant-deviation prism, known as the *Calbier* prism, that is made from two prisms – a 30° refracting prism and a 90° deviating prism (Fig. 3.20*c*). Compared with the Abbé prism, both the deviation angle and the angular dispersion of the Calbier prism are smaller.

The Abbé prism is mainly used in high-performance spectroscopes. This prism is also being used in conjunction with the *Försterling compound prism* as the dispersing

element of medium- and high-performance glass spectrographs. The latter prism system is composed of two 63° glass prisms and an Abbé prism placed between them (Fig. 3.21). The use of the 63° prisms is justified because glass has a greater refractive index than quartz. Thus, a large-refraction-angle prism, that gives good angular dispersion, can be used and will avoid the unfavourable large light losses arising from reflection.

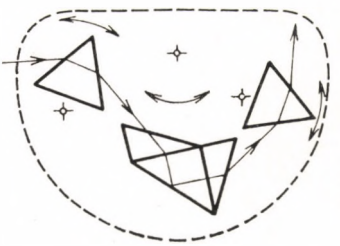


Fig. 3.21 Försterling prism system

In the lower performance spectroscopes the so-called *direct-vision prism* is used for dispersion purposes. Direct-vision prisms are made of crown and flint glass that have very different refractive indices and dispersion. By the proper choice of the refractive angle it is possible that the dispersion of one prism is partly or totally compensated (depending on the wavelength) by the dispersion of the second prism

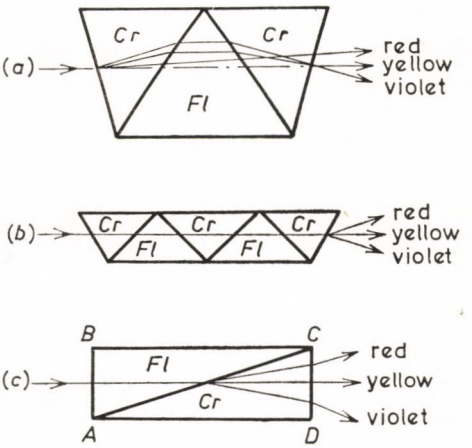


Fig. 3.22 Direct-vision glass prisms: (*a*) and (*b*) Amici prisms; (*c*) principal section of a Zenger prism (*Cr*: crown glass; *Fl*: flint glass)

of opposite refractive angle. Thus rays of medium wavelengths emerge from the prism in a direction that is the continuation of the path of the incident rays. Rays with wavelengths shorter or longer than the average are deviated to the right or left, respectively, relative to the direction of medium-wavelength rays.

Among the direct-vision prisms the commonest is the *Amici prism*. In this arrangement a flint-glass prism with a high refractive angle is surrounded by two crown-glass prisms with smaller refractive angles (Figs. 3.22*a* and *b*). The refractive angles of the latter prisms are such that yellow light passes through the system practically undeviated, while red and violet rays are deviated upward and downward, respectively. The resolving power of the compound prism is determined by the *difference*

between the total resolving power of the crown-glass prisms and the total resolving power of the flint-glass prisms. In this case it is better to have the same resolving power obtained by a greater number of short-base-length prisms than by a smaller number of large prisms. Commercially available Amici prisms are, therefore, not composed of three prisms, but of five, cemented together with Canada balsam.

Another, but less frequently used, direct-vision prism is the *Zenger prism* (Fig. 3.22*c*). It has the advantage that its principal section is a square. The light can thus impinge on the prism perpendicularly; that is an angle at which light loss by reflection is minimal. It is, however, still more important that the light does not enter directly, but through the flint-glass prism, ABC, before first passing into the crown-glass refracting prism, ACD. Thus, the light enters the refracting prism at a large angle, an angle at which there would be large reflection losses if entry occurred at the air/prism interface. This has the advantage that the *longer* side of the refractive prism will be the base. This is the prism side that influences the resolving power of the prism.

Quartz prisms are not applicable for the examination of the far ultra-violet region as quartz will absorb the rays in this radiation range. *Fluorite prisms* made of calcium fluoride are better for this purpose as there is total absorption only at a wavelength of 1250 Å. There is no problem in connection with this material as fluorite prism material can be made synthetically. In the same spectral region *lithium fluoride prisms* can also be used since total absorption only occurs at a wavelength of 1050 Å.

In the manufacture of large prisms difficulties may arise due to the problem of finding optically faultless pieces of a sufficiently large size. Under such circumstances *liquid prisms* can occasionally be used.[12,13] These are prisms made of parallel-sided glass or quartz sheets whose interior is filled with a liquid of an appropriate refractive index. Liquid prisms are mainly used for the investigation of the visible and ultra-violet regions. The advantages are that their construction is relatively simple, they are optically homogeneous, and they meet all the optical demands completely. On the other hand, it is a drawback that the temperature coefficient of liquid prisms is greater by an order of magnitude compared with that of solid prism materials (Table 3.4). No general use is found for liquid prisms in practice.

Table 3.4

Optical data of the more important 60° liquid prisms at a temperature of 20 °C and a pressure of 760 mmHg[14,15]

Prism material	n_D	$n_F - n_C$	$-(dn/d\lambda)_D$	Wavelength region of transmittance (Å)	$(dn/dt)_D$
water	1·3330	0·0060	$0{\cdot}416.10^{-5}$	1 850<	$-0{\cdot}8.10^{-4}$
carbon disulphide	1·6276	0·0341	$2{\cdot}885.10^{-5}$	1 850 2 400–2 800 3 500<	$-8{\cdot}0.10^{-4}$
monobromo naphthalene	1·6576	0·0325	$2{\cdot}501.10^{-5}$	3 500–8 000	$-4{\cdot}5.10^{-4}$

References

1. KAYSER, H., *Handbuch der Spektroskopie*, Leipzig, vol. I (1903)
2. WADSWORTH, F. L. O., *Astrophys. J.*, **2**, 264 (1895)
3. PICKERING, E. C., *Am. J. Phys.*, (2), **45**, 301 (1968); *Phil. Mag.*, 4th ser., **36**, 39 (1968)
4. D'ANS, J., and LAX, E., *Taschenbuch für Chemiker und Physiker*, 2nd edn, Berlin–Göttingen–Heidelberg, vol. 2, 1187 (1949)
5. HOHLY, H. W., *Ann. Phys.*, **29**, 433 (1937)
6. PERRIN, F. H., *J. Opt. Soc. Am.*, **28**, 86 (1938)
7. FORREST, J. V., *J. Opt. Soc. Am.*, **32**, 382 (1942)
8. STEPHEN, R. C., *Bur. Stand. J. Res.*, **43**, 81 (1949)
9. ABBÉ, E., *Gesammelte Abhandlungen*, Jena, vol. 1, 1 (1904)
10. PELLIN, P., and BROCA, A., *Ann. Phys.*, **8**, 314 (1899)
11. FORSYTHE, W. E., *Astrophys. J.*, **45**, 278 (1917)
12. PULFRICH, C., *Ann. Phys.*, **45**, 238 (1892)
13. GIFFORD, J. W., *Proc. R. Soc.*, **A 91**, 319 (1915)
14. DUCLAUX, J., *Rev. opt. théor. instrum.*, **2**, 384 (1923)
15. BREWSTER, D., *Eding. Trans.*, **8**, 285 (1914)
16. FLATOW, E., *Ann. Phys.*, **12**, 11 (1903)

3.3.4 The collimating lens[1,2]

The collimating lens renders parallel the rays of different wavelengths coming from the slit. The camera lens, on the other hand, combines the parallel beams emerging from the prism with varying degrees of deviation into spectral lines representing different wavelengths. Accordingly, the two lenses have different requirements, although they form a unit from the optical point of view and can be regarded as a compound lens.

First among the requirements for the collimating lens is that any chromatic aberrations should be taken into account. A further requirement is that the collimating lens should be free of spherical aberrations and coma.

Chromatic aberrations of lenses are caused by the fact that the refractive index is dependent on the wavelength. In the case of simple lenses, therefore, the focal length is shorter for the rays of shorter wavelength, while it is longer for those of longer wavelength (Fig. 3.23*a*). Chromatic aberrations can partly be eliminated by specially made compound lenses or achromats as they are called. The achromats are made from a crown-glass collecting lens and a flint-glass dispersing lens, that are adjusted so that the focal lengths of rays of *two* different wavelengths (λ_1 and λ_2) are the same when the light incident upon the lens is parallel to the optical axis of the lens. On the other hand, the focal lengths of rays having wavelengths different from λ_1 and λ_2 are still different (Fig. 3.23*b*). These deviations cause *secondary chromatic aberrations* (Fig. 3.23 *c*).

Taking into account the fact that the light path is reversible, it follows from the above considerations that rays of different wavelengths, coming from the slit, are not made strictly parallel by the collimating lens, unless the latter is a specially adjusted compound lens. In principle, however, not even the achromat lens can produce parallel rays that do not come from the exact location of the intersection point of the optical axes of the slit and the collimating lens. This latter requirement is less important the higher the slit.

Spherical aberrations are encountered with poorly made lenses of large diameter when the latter are illuminated by rays parallel to, but some distance away from,

the optical axis of the lens. In such cases the rays do not intersect at the focal point of the lens, but at a point closer to the latter. The further the vertical distance of the rays from the optical axis the greater is this difference (Fig. 3.24*a*). The different

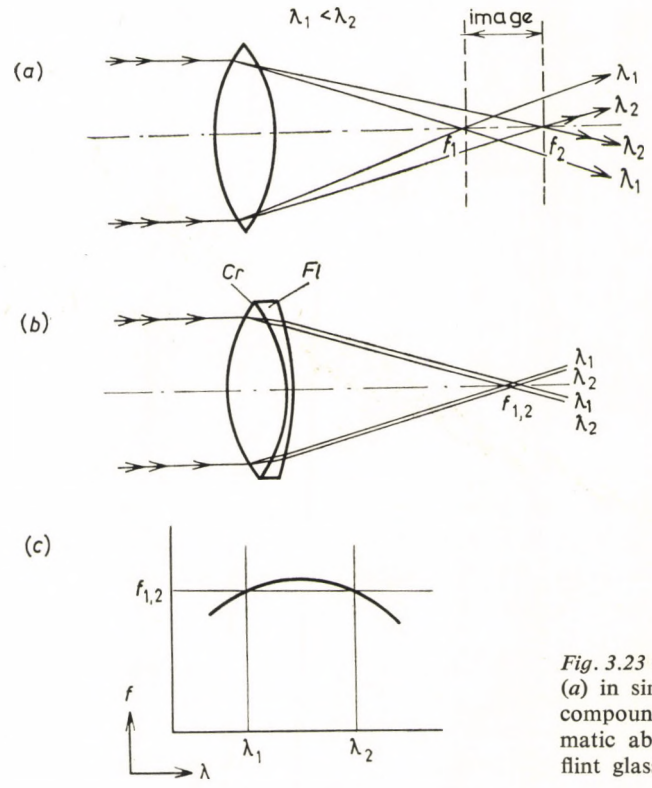


Fig. 3.23 Chromatic aberration of lenses: (*a*) in simple lenses, (*b*) in achromatic compound lenses, (*c*) secondary chromatic aberration (*Cr*: crown glass; *Fl*: flint glass)

concentric zones of the lens, therefore, have apparently different focal lengths and is generally referred to as the *zone error*. Consequently, the image of a point on the optical axis is not a discrete dot, but a disc-shaped patch with diffuse edges instead (Fig. 3.24*b*).

Spherical aberrations can be eliminated by means of a diaphragm of appropriate diameter, such that the free aperture of the lens is restricted. With proper optical

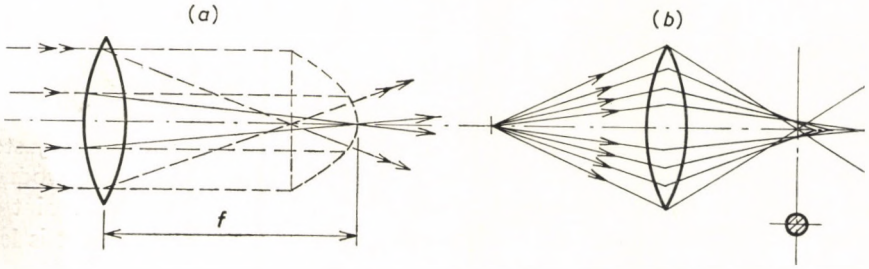


Fig. 3.24 Spherical aberration of lenses: (*a*) uniting rays parallel to the optical axis, and (*b*) a point image lying in the optical axis

adjustments, however, the same result can be achieved with appropriately designed compound lenses. In addition, spherical aberrations can also be avoided by simple lenses ground to have a parabolic shape. These adjustments are based on the consideration that the focal length of the lens zones progressively farther from the optical axis tends to increase together with the radius of curvature. Coma in the collimating lens is manifested by the presence of smeared, elliptical, partly focused spots instead of a clearly defined point-image (Fig. 3.25*a*). Coma can be regarded as the spherical aberration of rays falling *obliquely* on to the lens. This aberration can be reduced when a diaphragm is placed in front of the lens at a certain distance towards the light source.

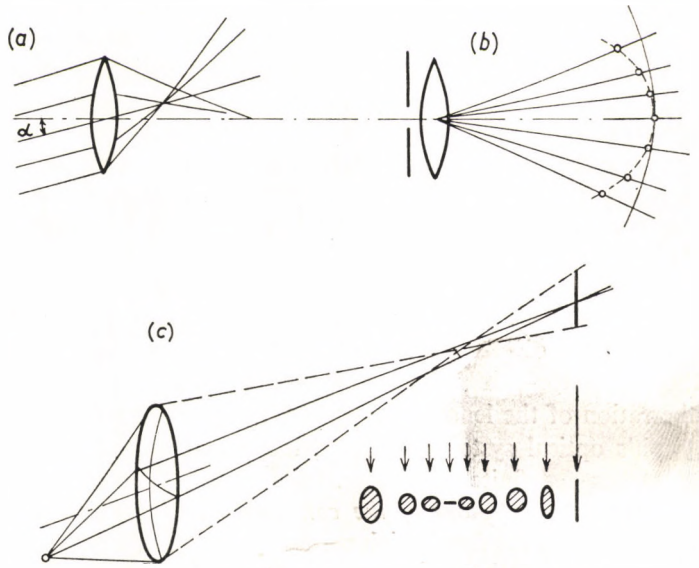


Fig. 3.25 Astigmatic error: (*a*) origin of the coma error, (*b*) the two astigmatic focal planes formed by rays of different angles of incidence, and (*c*) origin of the images of a point-source

Lenses corrected by diaphragms can still suffer *astigmatism*. The intersection points of parallel rays incident at an angle α to the lens axis lie on two differently shaped concave surfaces. The rays travelling in vertical planes give rise to images on the surface lying closest to the lens, while the rays in the horizontal planes intersect on the surface further away (Fig. 3.25*b*). The line image of the point object is produced at the end of a series of images at increasing distances from the lens. First a perpendicular ellipsoid, then a circular and later horizontal ellipsoid image is brought about. Upon increasing the distance the latter is transformed into a relatively sharp horizontal line. Still farther from the lens, a diffuse, horizontal ellipsoidal image appears, followed by a circular one and a perpendicular elliptical image. It is only farther away still that a relatively sharp, line image of the point object is found.

Astigmatism has the consequence that the relatively sharp image of the object is distorted. This fault cannot be eliminated by restricting the free aperture of the lens as the diaphragm only improves the sharpness of the image. Astigmatism can only

be removed by using a lens system built from lenses of the appropriate refractive index and curvature. The *astigmats* are compound lenses consisting of two lenses whose concave component has a lower refractive index than the convex one. Astigmats, however, do not eliminate the faults caused by spherical aberrations. These can also be eliminated when a compound lens of four components is used.

Optical systems corrected for the above-mentioned faults are still not completely free of image distortion. Distortion can often be observed in the focusing of non-parallel beams. Thus it can be neglected in spectrochemical analysis.

These faults can be eliminated by applying composite lenses having a diaphragm between the components.

The requirements for a good collimating lens can be relatively easily produced for glass spectrographs (spectroscopes). There are no major difficulties in manufacturing appropriately adjusted compound lenses when properly selected optical glasses are used. On the other hand, adjusted compound lenses necessary for quartz spectrographs can only be produced with difficulty. These can only be made of fused quartz and fluorite or of materials having refractive indices within narrow limits. A further factor rendering the manufacture of corrected lens systems difficult is that the adjustment of the collimating lens of quartz spectrographs should be done for a generally wide interval of about 3000 Å in width. It is therefore evident that the tendency is to make the collimating lens of quartz spectrographs exclusively of quartz; that is, to use a collimating lens made from a single component.

Chromatic aberrations caused by the collimating lens do not lead to faults when the collimator–camera lens system is properly aligned. In this case, the chromatic aberration of the lens only causes the rays to strike the focal plane of the spectrograph's optical system obliquely, instead of at right angles. As we shall see, this effect has a definite advantage.

Spherical aberrations of the collimating lens are practically eliminated, provided the *aperture of the lens* is high enough in a simple lens system. That is, when the ratio of the focal length, f, and the free diameter, D, of the lens

$$F = f/D \tag{3.25}$$

has a high enough value. Increasing the aperture of the collimating lens decreases the solid angle of the light beam coming from the point light source and thereby reduces the illumination of the lens. The value of the latter varies proportionally with the reciprocal of the aperture, that is, proportionally to the square of the *aperture ratio* of the lens. Owing to these various effects the aperture of the collimating lens is normally between 10 and 15 mm (or occasionally 20 mm) in good quality, single-prism quartz spectrographs.

It also follows from the above discussion that the *coma of non-compound collimating lenses* can also be neglected, provided the slit height is not too great. For commonly used collimating lenses with a focal length of 500–600 mm a deviation of 1 mm does not produce any coma. For slit heights of 6–8 mm one has to consider the effects of astigmatism.

From what has been said it is understandable that the comparatively expensive compound quartz lenses are not normally used as collimating lenses. For the study of the far ultra-violet region quartz–fluorite, calcium-lithium-fluoride achromats can be used to advantage as collimating lenses in vacuum spectrographs.

In place of the relatively expensive quartz achromats, *parabolic metal mirrors*, which are comparatively easy to manufacture, can be used instead. These are produced by vacuum evaporation of aluminium or selenium onto an appropriately ground glass surface. Their great advantage is the complete absence of chromatic aberrations. When the solid angle of the parabolic mirror is small, the rays parallel to the optic axis are concentrated at the focal point of the mirror independently of the wavelength (Fig. 3.26*a*). On the other hand, *parabolic metal mirrors with large solid angles* do not even make parallel the rays coming from the focal point. This is due to the parallel rays that lie some distance from the optical axis being reflected from the parabolic surface such that the intersection points of the neighbouring rays lie on a curved focal surface (Fig. 3.26*b*). Thus spherical aberration does occur when parabolic metal mirrors with very large solid angles are used.

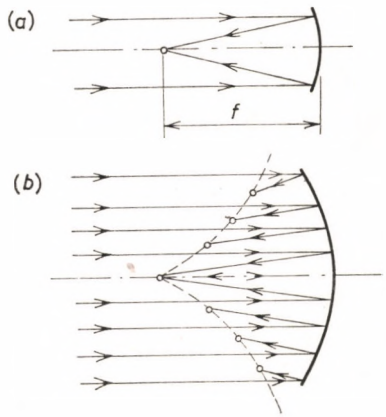


Fig. 3.26 Illumination of paraboloid metal mirror by rays parallel to the optical axis: (*a*) at a small aperture angle, and (*b*) at a large aperture angle

Since the illumination of concave mirrors varies proportionally with the square of their aperture ratio, *the aperture of the collimating mirror is the same as that of common collimating lenses.*

When a concave metal mirror is used in place of a collimating lens, one of the more important differences between them is that the mirror will reflect the rays coming from the focal point on the optic axis in a manner parallel to the optic axis of the mirror. The slit, therefore, cannot be situated *exactly* at the focus of the concave mirror since the incident and reflected rays would be mixed up. In oblique illumination of the concave mirror, however, one should first take into account the astigmatism and then any *distortion effect* that may occur.

The above-mentioned faults can be eliminated in two ways. The simplest solution is to make the angle between the incident beam from the slit and the reflected beam from the mirror as small as possible (about 2°). In this way it is possible that any problems arising from the oblique incidence can be practically neglected or they can be sufficiently corrected for by an appropriate modification of the parabolic surface of the mirror (Fig. 3.27*a*)

In another solution to the problem, coma and astigmatism can be practically eliminated by replacing the collimating lens and the camera lens by concave mirrors, such that the collimating mirror reflects the light to the left, while the camera mirror reflects it to the right (Fig. 3.27*b*). In this arrangement the second concave mirror

compensates for the astigmatism of the first one, so that both astigmatism and coma can be neglected, even when the light reaches the collimating mirror at a relatively large oblique angle instead of a small one.

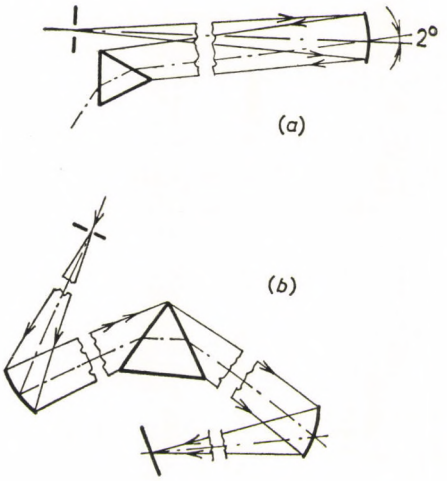


Fig. 3.27 Elimination of the astigmatic error of a concave collimating mirror

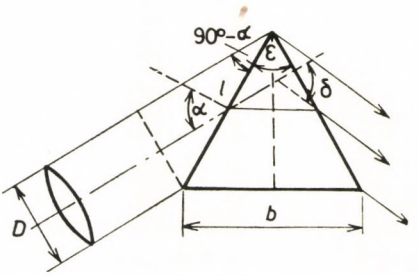


Fig. 3.28 Determination of the permissible diameter of the collimating lens

The free diameter of the collimating lens (or collimating mirror) should be chosen according to the dimensions of the prism. The side of the prism, l, in its minimum position will be fully illuminated by a collimating lens of free diameter, D, when, according to Fig. 3.28

$$D = l.\sin(90 - \alpha) = l.\cos\alpha$$

Furthermore, when it is taken into account that for a prism in its minimum position $\alpha = (\varepsilon + \delta)/2$, then

$$D = l.\cos\frac{\varepsilon + \delta}{2} \tag{3.26}$$

Performing the transformation $\cos x = \sqrt{(1 - \sin^2 x)}$ and taking into account Eq. (3.12) it can be calculated that

$$\sin\frac{\varepsilon + \delta}{2} = n\sin\frac{\varepsilon}{2}$$

where n stands for the refractive index. By substituting these values in Eq. (3.26), the smallest free diameter of the collimating lens is given by

$$D = l\sqrt{1 - n^2 \sin^2 \frac{\varepsilon}{2}} \tag{3.27}$$

where ε is the refractive angle of the prism.

It follows from the above relationship that for common 60° prisms the smallest free diameter of the collimating lens is given by the relationship

$$D_{60°} = b\sqrt{1 - \frac{n^2}{4}} \tag{3.28}$$

where b is the base length of the 60° prism. When it is considered that the average refractive index of prism materials is about 1·5, then the final result for the permissible minimum free diameter of the collimating lens is in general

$$D_{\min} = 0{\cdot}7b \tag{3.29}$$

References

1. POHL, R. W., *Optik und Atomphysik*, Berlin–Göttingen–Heidelberg (1954)
2. STRONG, J., *Concepts of Classical Optics*, San Francisco (1957)

3.3.5 The camera lens[1,2]

It follows from the above discussion that the camera lens combines the parallel beams emerging from the prism into spectral lines corresponding to different wavelengths, on the focal surface. Accordingly, the camera lens should be free of zone and coma effects and also properly corrected for spherical aberrations. Finally, the camera lens should be corrected for chromatic aberrations in such a way that spectral lines of different wavelengths are formed on an essentially planar surface. In other words, the focal surface of the camera lens, especially along the strip corresponding to the location of the spectrum, should be practically flat. This is necessary since the photographic plates can only be very slightly bent.

In the case of *glass spectrographs* it is not particularly difficult to construct compound lenses that meet the above requirements. A simple lens of sufficiently long focal length is generally free of zone and coma effects. By properly matching the collimating and camera lenses any deviation of the focal surface of the camera lens from a planar surface becomes so small that the bending of the photographic plate should be permissible even in the case of high-performance instruments. It is not possible, however, to eliminate the consequence of chromatic aberration in that the focal plane of the camera lens is not perpendicular to the optic axis of the lens. The resulting focal length of the coupled collimating camera lens is shorter for rays of shorter wavelength than for those of longer wavelength. The longer wavelength portion of the spectrum, therefore, lies farther from the collimating lens than the shorter wavelength portion. This is not always a drawback. Along with the slanting of the focal plane, the line dispersion of the instrument increases, although simultaneously the practical resolution is reduced. This is the explanation of the fact that the camera lens of commercial quartz spectrographs is usually a non-compound

or simple lens. Only recently have quartz spectrographs with a compound camera lens gained general acceptance, in order to utilise the better resolving power of the prism.

The focal length of the camera lens not only influences the zone and coma defects of the lens. For example, where the slit height is large the radius of curvature of the spectral lines, as measured on a plane perpendicular to the axis of the camera lens, is given by the formula

$$r = \frac{n_\lambda^2 f_\lambda^2}{2(n_\lambda^2 - 1)} \cot \alpha \tag{3.30}$$

where n_λ is the refractive index of the lens material measured at wavelength λ, f_λ is the focal length of the lens at the same wavelength, and α is the angle of incidence of a ray of this wavelength.[1,2] The radius of curvature of the spectral line increases with increasing focal length of the camera lens. When the value of the refractive index decreases, the wavelength of the light increases. The spectral lines are, therefore, more bent in the shorter-wavelength region of the spectrum than in the longer wavelength portion. From the above relationship it is clear that the spectral lines tend to be more curved when the angle between the focal plane and the optical axis of the camera lens is small, or when the angle of incidence is large. Finally, it can be mentioned that in the calculation of the effective curvature, the slanting angle of the focal plane should also be taken into account, since Eq. (3.30) refers to the case of normal incidence only.

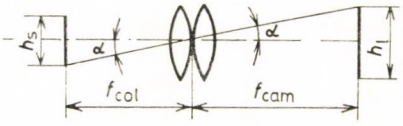


Fig. 3.29 Fundamental principle of the magnification of prism spectrographs

In certain spectrographs the camera lens is also replaced by a concave metal mirror. If this mirror is used as shown in Fig. (3.27*b*), then not only chromatic aberration but also the astigmatism of the camera mirror is largely eliminated. Spherical aberration can be sufficiently corrected for when the camera mirror is not a spherical one, but is an appropriately made parabolic mirror with a small aperture.

The optimum free *diamete* of the camera lens or camera mirror is mainly determined by the dimensions of the prism. It is important that the camera lens is illuminated by the *full* beam emerging from the prism, otherwise the resolving power of the prism cannot be fully utilised. This condition also determines the optimum focal length of the camera lens, as well as the free aperture of the camera lens. In the case of good-quality equipment the value of the free aperture varies between 1 : 10 and 1 : 12.

The free diameter of the camera lens can be reduced by means of a diaphragm in many spectrographs, being frequently reduced to 1 : 30. When an aperture is used, however, it should be taken into account that by reducing the free aperture of the camera lens the illumination of the instrument and also its resolving power are reduced.

In prism dispersion the spectral line can be defined as the image of the slit. The *magnification of the spectrograph* is, therefore, given by the ratio of the length of

the spectral lines, h_l, and the height of the slit, h_s; that is, the magnification of the spectrograph is $\beta'_{sp} = h_l/h_s$. As the rays leaving the collimating lens and incident on the camera lens are parallel, it follows that the magnification is not influenced by the dispersive part of the spectrograph. The collimating and camera lenses can, therefore, be regarded as a pair of lenses that cause no deviation of the rays passing through them. Considering Fig. 3.29 it can be shown that

$$\tan \alpha = \frac{h_s/2}{f_{col}} = \frac{h_l/2}{f_{cam}}$$

that is

$$\beta'_{sp} = \frac{f_{cam}}{f_{col}} \tag{3.31}$$

The magnification of prism spectrographs is therefore determined by the ratio between the focal lengths of the camera lens and the collimating lens.

References

1. KAYSER, H., *Handbuch der Spektroskopie*, Leipzig, vol. I (1903)
2. WIEN, W. and HARMS, F., *Handbuch der Physik*, Leipzig, vol. 21 (1927)

3.4 Performance of prism dispersion instruments

3.4.1 Linear dispersion and dispersion curve of prism spectrographs

The linear dispersion of prism spectrographs can be calculated from Eq. (3.2) which gives the linear dispersion as $S_l = dl/d\lambda$, and Eq. (3.1) where the angular dispersion is given by $S_a = d\delta/d\lambda$. In this case then

$$S_l = \frac{dl}{d\lambda} = \frac{dl}{d\delta}\frac{d\delta}{d\lambda} = S_a \frac{dl}{d\delta}$$

If the light is perpendicular to the focal plane of the camera lens, then, as is shown in Fig. 3.30, $dl = f_{cam}.d\delta$, provided f_{cam} is the average focal length of the camera lens. In reality, however, the normal to the focal plane of the camera lens forms

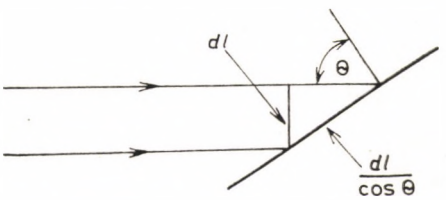


Fig. 3.30 Determination of the linear dispersion of prism spectrographs

an angle, Θ, with a ray of medium wavelength passing through the lens. In this case $dl = dl'.\cos \Theta$ and the linear dispersion of prism spectrographs is, therefore, given by

$$S'_{l(p)} = \frac{dl'}{d\lambda} = S_a \frac{f_{cam}}{\cos \Theta} \tag{3.32}$$

The greater the linear dispersion of prism spectrographs, the greater is the average focal length of the camera lens, the larger is the angle between the normal to the focal plane of the camera lens and the optic axis of the spectrograph (the tilt angle of the focal plane), *and the greater is the angular dispersion of the prism.* The value of the latter rapidly increases with decreasing wavelength. *Therefore, the larger the linear dispersion, the shorter is the wavelength.*

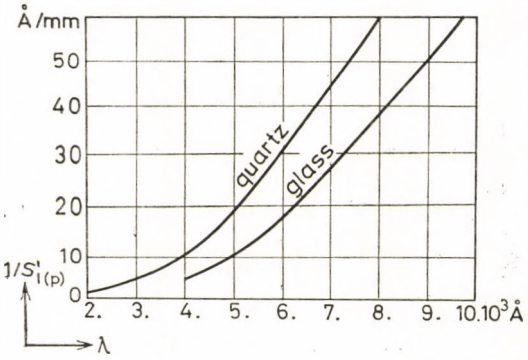


Fig. 3.31 Reciprocal linear dispersion of high-performance quartz and glass spectrographs

In the case of multiprism spectrographs, the linear dispersion depends on the number of prisms. *The greater the sum of the refractive angles of the individual prisms and the dispersion of the prism material, the greater is the linear dispersion.*

It follows from the above discussion that the value of reciprocal linear dispersion also varies with the wavelength of light. This variation is shown in Fig. 3.31. As is evident from the graph, *the reciprocal linear dispersion of prism spectrographs first increases slowly, then more rapidly, as the wavelength increases.* The longer the wavelength of the light, the shorter the distance becomes between lines representing the same wavelength difference. When the wavelength is increased, lines representing the same wavelength difference become so densely packed that the longer wavelength portion of the spectrum becomes practically useless. In the case of quartz spectrographs this effect occurs at shorter wavelengths than for glass spectrographs. Therefore, for the examination of the visible part of the spectrum, glass spectrographs are to be preferred, while, for the study of the ultra-violet region, quartz spectrographs are better.

3.4.2 Theoretical resolving power of prism spectrographs

The concept of the theoretical resolving power has already been defined in Subsection 3.1.2; that is

$$R_0 = \lambda/\Delta\lambda$$

We have seen that $\Delta\lambda$ is the wavelength difference for which the first secondary maxima of rays of wavelength λ coincide with the main maximum of rays of wavelength $\lambda + \Delta\lambda$ in the interference pattern.[1]

The basic relationship is given by

$$R_0 = \frac{\lambda}{\Delta\delta}\frac{\Delta\delta}{\Delta\lambda} \tag{3.33}$$

where $\Delta\delta$ is the angular difference corresponding to the wavelength difference $\Delta\lambda$ (Fig. 3.32). The quantity, $\Delta\delta$, simultaneously meets two separate conditions. First, it is identical to the direction difference between the main maximum and the first secondary maximum of rays having the same wavelength λ. Secondly, it is identical to the direction difference between the rays of wavelength λ and $\lambda + \Delta\lambda$ produced by the angular dispersion of the prism. Let us now write the two conditions in mathematical form.

As it can be seen from Fig. 3.32 and according to the first difference

$$W \,.\, \Delta\delta = \lambda \tag{3.34}$$

where W stands for the width of the beam and $\Delta\alpha$ is given in radians. To the direction of the first secondary maximum there belongs a full wavelength difference.

The second condition can be formulated on the basis of Eq. (3.18), which relates to the angular dispersion of the prism

$$\frac{\Delta\delta}{\Delta\lambda} = \frac{b}{a}\,\frac{\mathrm{d}n}{\mathrm{d}\lambda} \tag{3.35}$$

where b is the base-length of the prism and $\mathrm{d}n/\mathrm{d}\lambda$ is the dispersion of the prism material. It should be added that the value of $\Delta\delta/\Delta\lambda$, in the above equation, can be taken to be equal to $\mathrm{d}\delta/\mathrm{d}\lambda$ in Eq. (3.17), as these quantities are very small.

According to Eqs (3.33)–(3.35) we obtain

$$R_{0,p} = b\,\frac{\mathrm{d}n}{\mathrm{d}\lambda} \tag{3.36}$$

where, in contrast to the notation R_0 given in the general relationship (3.33), the theoretical resolving power of prism instruments is denoted by $R_{0,p}$. The longer the prism base and the dispersion of the prism material (or, indirectly, the shorter

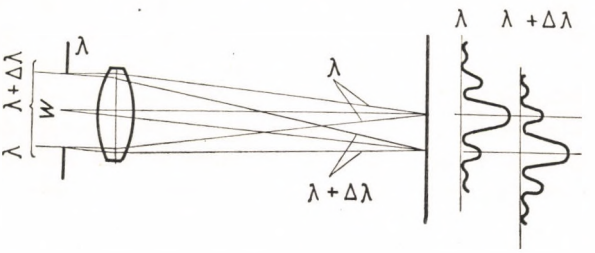


Fig. 3.32 Idealised ray path for the calculation of the theoretical resolving power of prism spectrographs

the wavelength), the greater is the theoretical resolving power of prism spectrographs. The theoretical resolving power is *not* influenced by the refractive angle of the prism, nor by the focal length of the camera lens, which is an important factor in the linear dispersion of the spectrograph.

The value of the refractive index changes approximately with the wavelength of light, λ, according to the Cauchy relation

$$n = A + \frac{B}{\lambda^2} + \frac{C}{\lambda^4} + \ldots \tag{3.37}$$

where the constants A, B, and C are characteristic of the refractive material.[2] When

for the sake of simplicity the higher-order terms are neglected, we obtain

$$\frac{dn}{d\lambda} \approx -\frac{2B}{\lambda^3}$$

From the above equation and Eq. (3.36) it follows that

$$R_{0,p} \approx \frac{K}{\lambda^3} \tag{3.38}$$

where K is a constant. Accordingly, *the theoretical resolving power of prism spectrographs varies approximately inversely with the third power of the wavelength.*

The previously given relationships for the theoretical resolving power of prism spectrographs only relate to single-prism instruments. *To establish the theoretical resolving power of multiprism spectrographs* it should be taken into account that the collective angular dispersion of the prism system is equal to the angular dispersion sum of individual prisms. According to the previous discussion, however, the value of the theoretical resolving power is proportional to the angular dispersion. From this it follows that the collective resolving power of prisms of the same material, in their minimum position, is given by

$$R_{0,\Sigma,p} = \Sigma b \frac{dn}{d\lambda} \tag{3.39}$$

Under the above conditions *the theoretical resolving power of prism systems varies linearly with the sum of the prism base lengths,* provided the material of the prisms is the same.

In the above it has been assumed that the rays coming from the collimating lens illuminate all the front part of the prism. The resolving power, however, is reduced when the rays do not completely illuminate the prism or prisms. That is, the illumination is such that the effective base length of the prism is reduced. The resolving power is therefore reduced if a diaphragm is used to cut down the size of the beam incident on or emerging from the prism. To reduce the free aperture of the camera lens is therefore a disadvantage from the point of view of resolving power. On the other hand, there is no change in the resolving power when the illumination of the prism is reduced by a diaphragm placed behind the collimating lens such that only the height of the illuminated prism area is reduced.

References

1. RAYLEIGH, R., *Phil. Mag.*, 5th ser., **9**, 40 (1880)
2. BALY, E. C. C., *Spektroskopie*, Berlin, 56 (1908)

3.4.3 Practical resolution of prism spectrographs

In the previous calculations it has been assumed for the sake of simplicity that the focal plane of the camera lens is perpendicular to the optic axis and that the slit of the spectrograph is infinitely narrow. In practice neither of these conditions is satisfied. The true (or practical) resolving power of a spectrograph is always lower than the theoretical. The greater the tilt angle of the focal plane of the camera lens

and the larger the width of the slit, the greater is the difference between the two values. Depending on the width of the slit the practical resolution of a spectrograph is variable. The relative performance of different equipment can only be compared when the practical resolution figures refer to the same *defined* slit width.

It is well known that on a progressive reduction of the slit width, the spectral lines can only be narrowed down to a certain minimum width ($d_{\min}$). The value of $d_{\min}$ depends on the wavelength as well [as shown by Eq. (3.23)]. This factor can lead to uncertainties in the comparative performances of different spectrographs.

This uncertainty is removed when the slit width, b', is taken into account, if the following condition is met[1]

$$d' = \lambda F_{\text{col}} \tag{3.40}$$

The corresponding resolving power is normally called the 'spectral purity of the spectrum', and is denoted by P. The spectral purity is the practical resolution when the slit width is equal to the product of the wavelength and the aperture of the collimating lens. (According to another definition[2] the spectral purity is given by $P = R_{0,\,\text{p}}\,[\lambda/(\lambda + bW/f_{\text{col}})]$, where $R_{0,\,\text{p}}$ is the theoretical resolution, W is the width of the beam and f_{col} is the focal length of the collimating lens.) Note that the spectral purity is also a function of the wavelength of light.

Furthermore, it can be established that the ratio of the resolving power, $R_{0,\,\text{p}}$, and the practical resolving power, R_{p}, of prism spectrographs is given by

$$\frac{R_{0,\text{p}}}{R_{\text{p}}} = \frac{1}{\lambda}\frac{b'}{F_{\text{col}}} + \frac{\lambda}{\lambda + 2\dfrac{b'}{F_{\text{col}}}} \tag{3.41}$$

Taking into account that $R_{\text{p}} = P$ and the relationship in Eq. (3.33), we obtain

$$P = \frac{3}{4} R_{0,\,\text{p}} \tag{3.42}$$

Therefore, the spectral purity is three-quarters of the theoretical resolving power. Its numerical value is generally in the order of 10^{-2}. This is only true when the focal plane of the camera lens is perpendicular to the optic axis of the lens. In practice $s_{\min}$ also depends on the tilt angle of the focal plane. As a consequence the practical resolution corresponding to the slit width of $d_{\min}$ is lower than P.

A limit is set to the resolving power of spectrographs by the resolving power of the photographic emulsion. As emulsions are grainy, only those rays that fall upon separate grains can form discrete spectral lines. The grain size of the commonly used emulsions is about 1 μm. To achieve the best resolution it is necessary that the distance between two spectral lines is ten times the above value. Eq. (3.32) is valid for the linear dispersion of prism spectrographs. Thus, any spectrograph whose resolving power is given by the following expression will satisfactorily resolve the photographic emulsion

$$R_{\text{p}} = \frac{\lambda}{\Delta\lambda} = \lambda \frac{S_{\text{a}}}{\Delta l \cos\phi} f_{\text{cam}} \tag{3.43}$$

where Δl is the distance corresponding to the practical resolving power of the photographic emulsion.

Equipments in which the above requirements are taken into account are called *normal spectrographs*. The resolving power of the latter is *just* fully realised. On the other hand, spectrographs with a theoretical line width smaller than 10 μm are called luminous spectrographs. The latter can be used to best advantage when the luminosity of the light source is low. Accordingly, these instruments are not used very frequently in practice.

The opposite extreme is represented by the so-called *over-magnifying spectrographs*. In these spectrographs the theoretical line width is greater than 10 μm. This equipment has the advantage that there is a better statistical elimination of the error coming from unavoidable local irregularities in the photographic emulsion. However, over-magnifying spectrographs have the disadvantage that their illumination is lower. Their application does not secure an overall advantage, but they can be usefully employed as multichromators.

Spectrographs whose linear dispersion is large compared with their resolving power are said to have no magnification. In this case the resolution of the photographic emulsion does not play a part. The application of instruments of these types is not generally advised. In principle the linear dispersion can be increased at will by increasing the magnification of the camera lens, that is, by lengthening the focal length of the latter. This, in itself, however, does not enable closely associated lines in the spectrum, which are below the resolving power of the instrument, to be separated. On the other hand, it is a drawback that the luminosity drops in proportion to the square of the focal length. Therefore there is an upper limit set to the magnification of the camera lens by the resolving power of the dispersion system. From this, the fact emerges that the performance of spectrographs is unequivocally determined by the resolving power and not by the linear dispersion.

Furthermore, it can be established that it is generally no advantage if the practical resolution of the spectrograph is fully utilised. In the optimum case the practical resolving power is somewhat higher than absolutely necessary.[3] In the opposite case the spectral lines are not sharp enough. This is explained by pointing out that diffraction occurring at the slit could lead to the appearance of first-order maxima, or, less frequently, second-order and third-order maxima, in addition to the main maximum when the resolving power is just fully utilised. These additional maxima render the edges of the spectral lines diffuse.

This undesirable phenomenon also occurs when the ratio of the diameter of the beam, emerging from the prism, to the focal length of the camera lens (W/f_{cam}) reaches a certain Ω value (Table 3.5).

Table 3.5

Ω-values belonging to different wavelengths

Å	$\Omega = W/f_{\text{cam}}$	Å	$\Omega = W/f_{\text{cam}}$
2 000	1 : 100	4 000	1 : 50
2 500	1 : 80	5 000	1 : 40
3 000	1 : 67	6 000	1 : 33
3 500	1 : 57	7 000	1 : 29

The diagram shown in Fig. 3.33 illustrates the relationship existing between the practical resolving power of the more commonly used spectrographs and the wavelength of light. On the other hand, Fig. 3.34 illustrates for different wavelengths the minimum wavelength difference which is necessary for the resolution of two nearby spectral lines. (Also on this figure the corresponding curve for a grating spectrograph is shown.) As seen from the figures the practical resolving power of prism spectrographs decreases very steeply with increasing wavelength, before becoming constant for quartz and glass spectrographs at about 4000 Å and 7000 Å, respectively. The longer the wavelength, the larger is the wavelength difference between two neighbouring lines that are just resolved. (On the other hand, for grating spectrographs this wavelength difference is almost independent of the wavelength.)

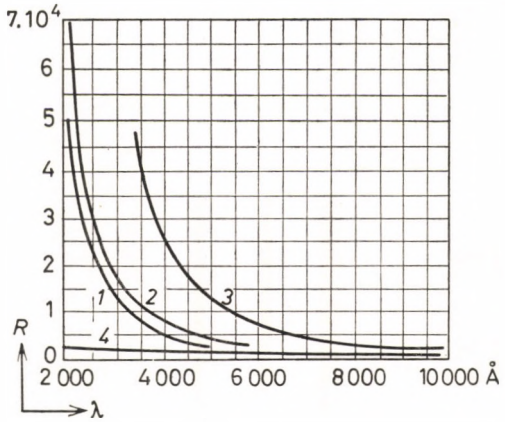


Fig. 3.33 Variation of the true resolving power of common prism spectrographs as a function of the wavelength of light: (*1*) medium performance quartz spectrograph, (*2*) Littrow-system autocollimation spectrograph with quartz optics, (*3*) the same equipment with glass optics, and (*4*) grating spectrograph

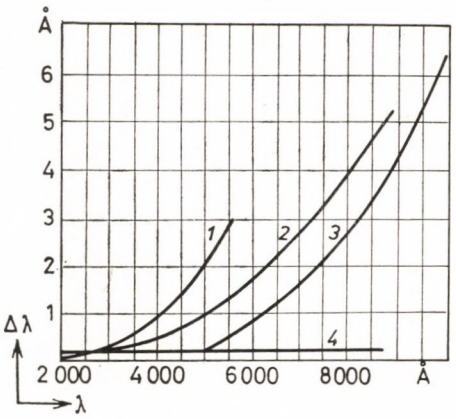


Fig. 3.34 Variation of the smallest wavelength difference of the resolved lines as a function of the wavelength of light: (Curves 1–4 refer to the same apparata as in Fig. 3.33.)

The construction of the curves previously mentioned is very complicated. The practical resolving power can, however, be established relatively quickly, albeit approximately, if the distance between the *appropriate* lines is measured on a spectrum specifically produced for this purpose. For example, the resolving power of quartz spectrographs can be established with the aid of the iron triplet at 3100 Å.

The estimation of the resolving power is made easier by the fact that at 3099·97 Å a single line appears in spectra produced by low-resolution spectrographs, while at

high-resolution two lines of wavelengths 3099·897 and 3099·971 Å can be observed (Fig. 3.35). In the visible spectral range, the D-doublet of sodium at 5889·96 and 5895·93 Å, or the line pairs in the iron spectrum at wavelengths 4066·98–4067·28 Å, 4871·33–4872·15 Å, and 6408·04–6411·67 Å, can also be used for the same purpose.

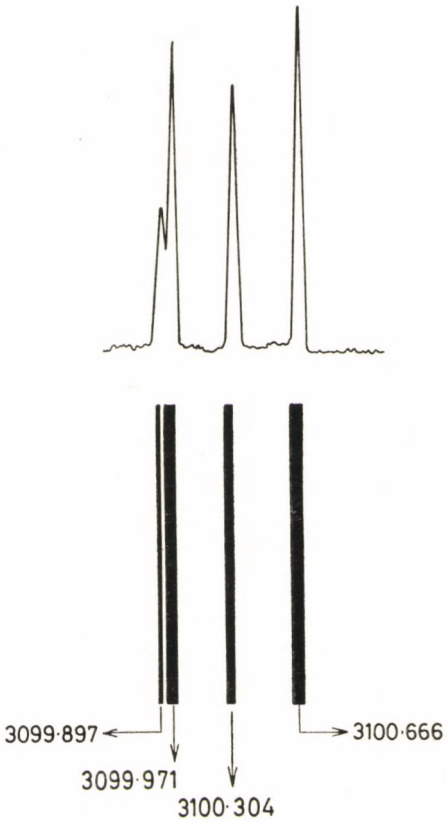


Fig. 3.35 Spectra and registrogram of the iron 'triplet' by a high-resolution Ebert spectrograph [4]

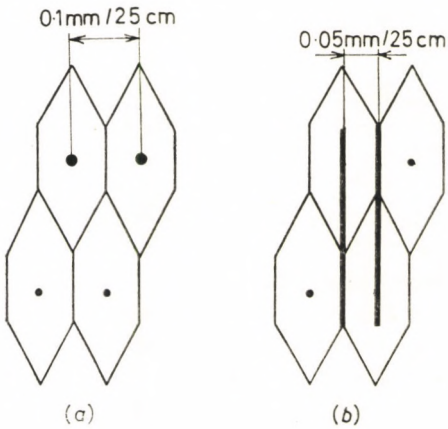


Fig. 3.36 The image on the retina of two resolved points (*a*) and two parallel lines (*b*)

The practical resolving power of spectroscopes will also depend on the resolving power of the human eye as well. A measure of the latter is the smallest distance of approach registered by the human eye between two closely associated but resolved points or dots from a distance of 25 cm. This corresponds to about one moment of an arc (about 0·1 mm distance). On the other hand, two parallel lines (for example, spectral lines) can be distinguished when the angular difference is about 30 seconds of an arc (0·05 mm). This can be explained by considering the arrangement of the visual (optical) nerve ends on the retina. For the resolution of two very closely lying points, two neighbouring nerves are necessary, while for the differentiation of two lines, two nerves lying one below the another are enough (Fig. 3.36).

The above data should be taken into account in the design of spectroscopes, as the resolving power of the photographic emulsion is important in the construction of spectrographs.

References

1. SEITH, W. and RUTHARDT, K., *Chemische Spektralanalyse*, 4th edn, Berlin, 13 (1949)
2. NACHTRIEB, N. H., *Principles and Practice of Spectrochemical Analysis*, New York–Toronto–London (1950)
3. HAMMER, A., *Spectrochim. Acta*, **2**, 365 (1944)
4. The Compact High Performance Plane Grating Spectrograph B.S. Catalogue Optica Milano (1962)

3.4.4 Photographic speed of prism spectrographs

The speed of prism spectrographs is influenced by many factors. Factors such as light reflection and absorption at the surface of optical elements, etc., are discussed in detail in the next section. A numerical value for the spectrographic speed, $\frac{1}{F_p}$, is rather complicated to establish. Theoretically, it can only be shown that the speed is influenced by the F-number and the tilt angle ϕ of the focal plane of the camera lens, such that the following relation holds

$$\frac{1}{F_p} \approx \frac{\cos \phi}{F^2_{cam}} \tag{3.44}$$

However, the above formula can be used in the comparison of similar types of equipment.

On consideration of the above, prism spectrographs can be classified according to their F-number. Spectrographs with a camera lens F-number of 10–15 are normally referred to as of *normal speed*. On the other hand, the F-number of *low-speed*, low-luminosity spectrographs is greater than 20. (To this category belong the Littrow-system spectrographs.) In practice, *high-speed* spectrographs have a camera lens F-number smaller than 10.

In certain spectrographs the free opening of the camera lens can be reduced by means of a diaphragm. In this case the speed of the equipment is determined by the effective F-number of the lens, that is, the ratio of the focal length of the camera lens to the diameter of the beam passing through the lens (f_{cam}/W). Here, the effective speed varies with the square of the effective F-number. By altering the effective speed it should, however, be remembered that by reducing the free diameter of the

camera lens the resolving power of the spectrograph is reduced. This occurs because the diaphragm excludes a part of the rays coming from the prism and thereby shortens the effective base length of the latter.

3.5 Error sources of prism spectrographs

3.5.1 Main sources of error

It has already been mentioned (see previous section) that only a part of the light flux (see Sub-section 5.1.1) entering the spectrograph reaches the final spectrum. Most of it is absorbed within the lenses and the prism of the instrument, as well as being reflected at the surfaces of the lenses and the prism. If the light flux entering the spectrograph is taken as unity, then the following relationship exists between the absorbed flux, Φ_A, the reflected flux, Φ_R and the flux passing through the apparatus, Φ

$$\Phi = 1 - (\Phi_A + \Phi_R) \tag{3.45}$$

The above light losses are caused by absorption and reflection. This loss depends, however, on many factors, such as the quality, geometrical form, and physical state of the surface of the refracting material. Light loss is also influenced by the wavelength of the light and the angle of incidence of the rays.

Light losses as such cannot be regarded as errors as they only influence the luminosity (speed) of the instrument. One factor is definitely harmful, in that light scattering is produced by reflection. In prism dispersion errors derived from alteration of the atmospheric pressure, temperature, etc., cannot be neglected.[1]

Reference

1. KAYSER, H., *Handbuch der Spektroskopie*, Leipzig, vol. I (1903)

3.5.2 Consequences of light absorption

To a certain extent every transparent material absorbs light passing through it and transforms the absorbed light energy into thermal energy. There is no clear-cut boundary, however, between transparent and non-transparent materials. Transparent bodies only absorb a small portion of the light when they are very thin. On the other hand, bodies that are definitely opaque (for example, metals) allow a partial transmission when they are thin enough.

The absorption loss can be simply calculated according to Bouguer–Lambert's law which states that

$$I = I_0 e^{-al} \tag{3.46}$$

where I_0 is the intensity of the incident light, I is the transmitted light intensity, a is the absorptivity of the absorbing medium, and l is the thickness of the latter.

As a consequence of absorption, the intensity of light drops exponentially with the thickness of the medium. The value of the absorptivity, a, and thus the measure of the intensity change, not only as a function of material quality, but also with the wavelength of the light. Media for which $1/a < \lambda$ are weak absorbers or 'transparent', while those with $1/a > \lambda$ are strong absorbers or 'opaque'. As seen, the

above media property is also a function of the wavelength. This has the practical consequence that the spectral interval, within which a certain dispersion medium can be used, does not solely depend on the refractive index of the material but depends also on its absorptivity.

In prism dispersion errors can occur if the absorption coefficient varies with the wavelength. In principle, this variation could lead to a distortion of the intensity relationships of the spectral lines. As seen from Fig. 3.37, however, it is evident that this error can be practically neglected.

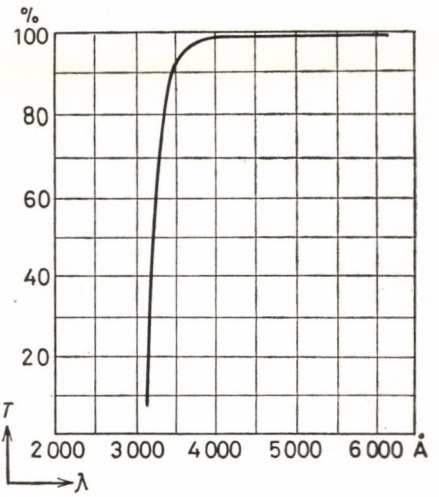


Fig. 3.37 Variation of the percentage transmittance by the prism material (crown glass) as a function of the wavelength of light[1]

In prism dispersion absorption losses are caused mainly by the prism.[2] In general, the optical paths of the parallel rays incident on the prism are not the same. The intensity of the rays emerging from the prism in the vicinity of the refractive edge is higher than that of rays passing along the base of the prism, even when the prism is uniformly illuminated. This factor renders the exact determination of the light loss caused by the prism rather difficult, since the absorption varies exponentially with the pathlength. Generally, it is advisable to determine any light loss at some average pathlength; for example, the half-base length of the prism in its minimum position ($b/2$) is usually taken as the standard.

According to the above discussion the average intensity of a monochromatic beam emerging from a composite prism consisting of m pieces of identical components, placed behind one another, in their minimum positions, is given by the formula

$$I = I_0 e^{-am\frac{b}{2}} \tag{3.47}$$

Reflection losses are neglected. According to Eq. (3.36), however, the resolving power of the same prism system, apart from its dependence on wavelength, depends *only* on the collective base length of the prisms. It follows with regard to light absorption that the number and shape of the prisms used is immaterial for the realisation of a certain level of resolving power.

The average width of spectrograph lenses is not large compared with the half-base length of the prism. Thus, light absorption by lenses involves an even smaller loss than in the case of the prism.

References

1. KOHLRAUSCH, F., *Praktische Physik*, 19th edn, Leipzig, vol. 2, 528 (1955)
2. KAYSER, H., *Handbuch der Spektroskopie*, Leipzig, vol. I (1903)

3.5.3 Consequences of light reflection[1]

In prism spectrographs, light losses caused by reflection at the surfaces of the lenses are much greater than losses produced by absorption. The larger the reflecting surface, the greater is the reflection loss. With regard to reflection losses it is therefore important to consider how many prisms are required to achieve a certain resolution level. It is well known that natural, non-polarised light, incident normally upon the plane surface of a transmitting medium, only suffers a small reflection loss. For example, in the case of an air–glass interface the loss is about 4%. The percentage of the light reflected, however, depends on the refractive indices of the two media in contact. Upon increasing the angle of incidence the intensity of the reflected light does not change until a certain incidence angle (the *polarisation angle*) is reached. (This angle is about 55° in the case of an air–glass interface.) On increasing the angle of incidence past the polarisation angle, the reflection increases steeply until total (or almost total) reflection occurs (Fig. 3.38*a*).

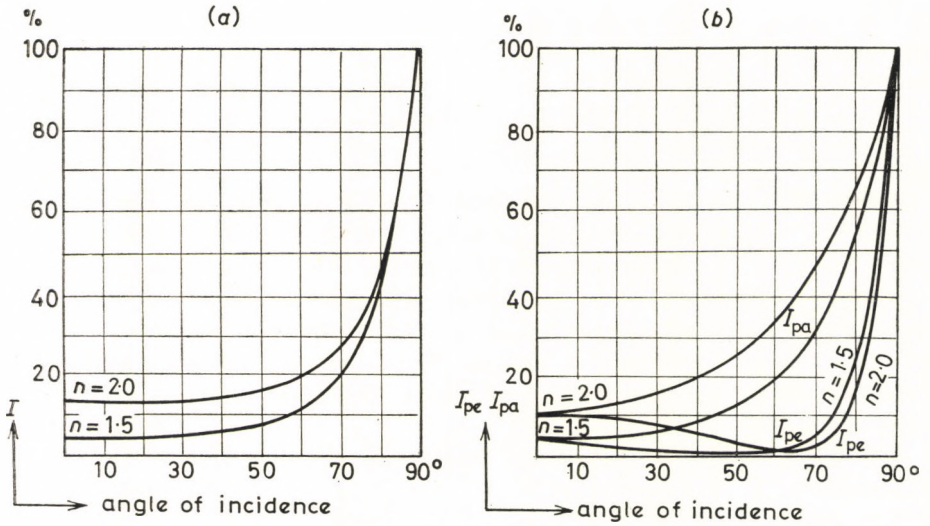


Fig. 3.38 Relationship between the relative intensity of reflected light and the angle of incidence: (*a*) natural light, (*b*) light polarised parallel to the plane of incidence (I_{pa}), and polarised perpendicular to the plane of incidence (I_{pe})

According to the Brewster law, degree of polarisation at an angle of incidence α_r when the path of the reflected ray is perpendicular to the ray that is refracted on entering the transparent medium (Fig. 3.39). The value of this *polarisation angle* depends on the refractive indices of the two media in contact, n and n', respectively, such that the following relation holds

$$\tan \alpha_r = \frac{n'}{n} \tag{3.48}$$

At this limiting angle the rays reflected from the surface are polarised in the plane of incidence, while the rays entering the reflecting medium are polarised in a plane at right angles to the reflected rays (Fig. 3.39).

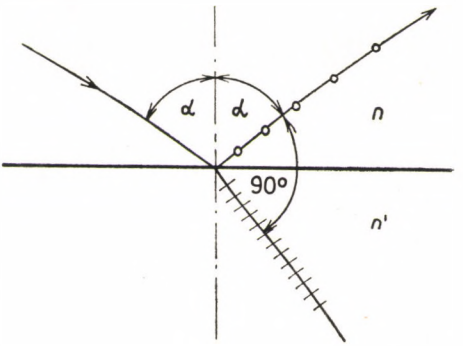


Fig. 3.39 The polarisation angle

Polarised light incident normally upon the surface of a transparent medium is reflected so that the intensity of the reflected ray depends exclusively on the refractive indices of the media in contact. This case is similar to that of natural light. On the other hand, with oblique incidence of polarised light, the intensity of the reflected light not only depends on the angle of incidence, but also on the angle between the plane of polarisation and the plane of incidence. In Fig. 3.38 it can be seen that the relative intensity, of the reflected beam rapidly increases when the angle of incidence is increased, provided the plane of polarisation of the reflected light is parallel to the plane of incidence.

The relative intensity, I_{pe}, of the beam polarised perpendicularly to the plane of incidence is comparatively low at low angles of incidence. Upon increasing the angle of incidence the intensity initially drops to zero. In this particular state the angle of incidence is equal to the angle of polarisation. When the angle of incidence is further increased, relative to the angle of polarisation, the relative intensity of the light polarised perpendicularly to the plane of incidence increases steeply.

The above-mentioned theoretical facts have the following practical consequences.[2]

The *collimating lens* is normally illuminated by natural light and the rays passing through the lens are practically non-polarised. This occurs since the light falls approximately at right angles onto the collimating lens. Light losses only occur because of reflection on the lens surface. The value of these relatively low light losses is influenced by the refractive index of the lens material.

The surface of the prism is illuminated by rays, rendered parallel by the collimating lens, at an angle corresponding to the minimum position of the prism. The natural light beam, therefore, falls obliquely on the refractive surface of the prism. This has the consequence that both the light reflected from the prism surface and the light passing through the prism are polarised. The plane of polarisation of the light within the prism is parallel to the plane of incidence, while that of the reflected light is perpendicular to the other beam.

The relative energy of the polarised beam reflected from the surface of the prism can be calculated from the Fresnel equations. According to these equations the proportion of natural light falling on the prism that is reflected with a plane of

polarisation perpendicular to the plane of incidence is given by

$$X_{pe} = \frac{\sin^2(\alpha - \beta)}{\sin^2(\alpha + \beta)} \tag{3.49}$$

where α is the angle of incidence and β is the angle of refraction. On the other hand, the proportion of natural light that is reflected with a plane of polarisation *parallel* to the plane of incidence is given by the formula

$$X_{pa} = \frac{\tan^2(\alpha - \beta)}{\tan^2(\alpha + \beta)} \tag{3.50}$$

From the above discussion the fraction of natural light that passes through the prism is given by the formula

$$(1 - X_{pe}) + (1 - X_{pa}) = 2 - (X_{pe} + X_{pa})$$

This is so provided absorption losses are not taken into account. As half the natural light falling on the prism is polarised parallel to the plane of incidence and the other half polarised perpendicular to it, then it follows from the above relationship that the reflected fraction of the light is given by $1/2.(X_{pe} + X_{pa})$. Accordingly, the percentage loss of light occurring at the prism is given by the formula

$$L_p = 50\,(X_{pe} + X_{pa})\% \tag{3.51}$$

The former statements relate to light losses occurring at a *single* prism when only one reflection is taken into account. In the case of a prism system consisting of m prisms of the same size and refracting material in their minimum position, the light losses arising from reflection are given by the equation

$$L_{mp} = 1/2\,(X_{pe}^m + X_{pa}^m) \tag{3.52}$$

It follows from the above relationship that on increasing the number of prisms there is a significant increase in the loss of light arising from reflection. This reflection loss depends not only on the number of prisms in minimum position, but on their refractive angle and the refractive index of the prism material as well. This correlation is illustrated in Table 3.6. On the other hand, according to Eq. (3.33) the required resolution can only be achieved by prisms of small refractive angle and when the base length is comparatively large. Enlargement of the prism, however, involves a corresponding increase in the diameter of the lenses. If the diameter of the lens is too large there is an increase in the optical faults, and one disadvantageous consequence is that the optimum F-number of the collimating lens can only be obtained by using a lens that has a very long focal length. In this case the speed of the spectrograph will be reduced. Thus, the advantage gained by using a prism with a smaller refractive angle is lost.

From the data given in Table 3.6 it is evident that when a prism with a refractive angle greater than 60° is used there is an increase in light losses due to refraction. Prisms with much larger refractive angles cannot be used satisfactorily because the angle of incidence of the light on the prism is larger than the angle of polarisation. As a consequence, reflection losses become intolerably high. Therefore, in the final analysis, the optimum conditions are achieved when the refractive angle of the prism is about 60°.

Table 3.6

Percentage light loss arising from reflection for prism systems with variable numbers of prisms and different refractive angles and indices (in their minimum position)[3]

Refractive index, n	Light losses (%)			Refractive angle of prism
	1-prism system	2-prism system	3-prism system	
1·5	8·4	15·9	22·6	
1·6	10·8	20·1	28·1	45°
1·7	14·1	25·5	34·7	
1·5	10·5	18·9	25·8	
1·6	14·7	25·2	32·8	60°
1·7	19·9	31·9	39·2	
1·5	13·7	23·7	30·9	67°22′*
1·6	18·2	29·8	37·1	64°0′
1·7	22·0	34·3	41·2	60°56′

* For the three refraction angles given here the angle of incidence of light upon the prism is equal to the angle of polarisation.

On the basis of the data in Table 3.6 prism systems made from a number of individual non-composite prisms have reflection losses that can be said to increase with the number of prisms. There is about 70–80% greater loss of light with a prism system made from two 60° prisms than with a single prism of double the base length. A third prism would cause a further 20–40% loss. The losses produced in this way increase as the refractive index of the prism material increases. For *composite prisms* the reflection losses are naturally much lower.

References

1. SCHALLREUTER, W. in *Grimsehls Lehrbuch der Physik*, 13th edn, Leipzig, vol. 3, 115 (1955)
2. KAYSER, H., *Handbuch der Spektroskopie*, Leipzig, vol. 1 (1903)
3. PICKERING, E. C., *Phil. Mag.*, **36**, 39 (1968)

3.5.4 Effect of stray light

In prism dispersion the error caused by *stray light* is usually more significant than that produced by reflection. Stray light is caused by reflection occurring at the prism and lens surface which alters the original direction of the light path so that the reflected rays travel in all directions.

Any stray light that reaches the focal plane of the camera lens increases the background of the spectra. A significant fraction of the light getting into the spectrograph is reflected at the surfaces of the dispersion elements and turns into stray light. However, well-designed spectrographs allow only a very small portion of the

stray light to reach the focal plane of the camera lens, so that little serious trouble is caused.[1]

Reflection and thus stray light are not only produced when light passes from air into the refractive medium, but also in the reverse sense.

The lenses of prism spectrographs are usually illuminated by rays that are parallel to the optical axis of the lens. Of these rays only those that are normal to the convex surface of a *convex lens* travel exactly in the direction of the optical axis of the lens. The greater the angle of incidence with respect to the convex surface of the lens, the greater is the distance of the rays from the optical axis of the lens (Fig. 3.40*a*).

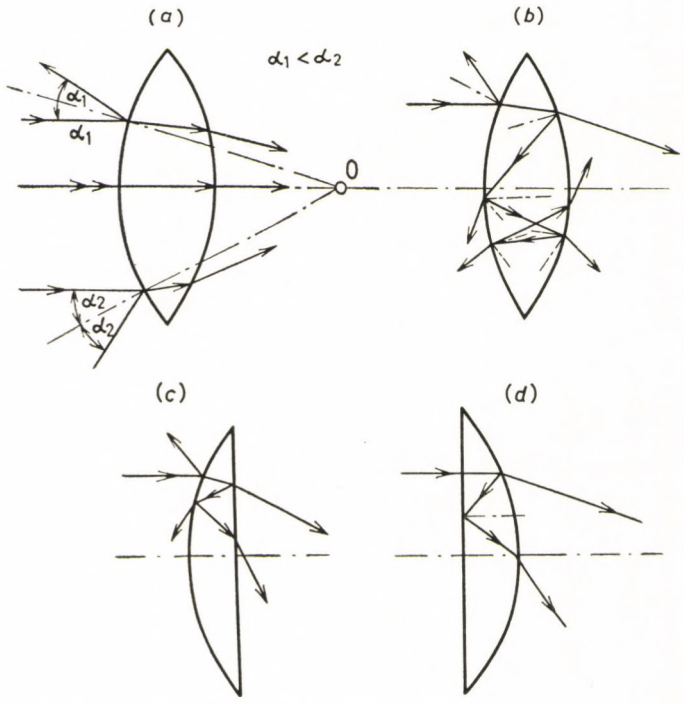


Fig. 3.40 Stray light arising on the surfaces of lenses: (*a*) reflection occurring at the interface between a convex lens and air, (*b*) multiple reflection occurring within a biconvex lens, (*c*) reflection arising when the convex side of a plano-convex lens is illuminated, and (*d*) reflection arising when the plane surface of the same is illuminated

Accordingly, the rays are reflected in different directions at the convex surface of the lens and stray light will be produced. However, it is only a small portion of the light reaching the convex surface of the lens that is reflected as most of it will enter the lens. A large proportion of the rays that enter the lens will leave at the other side, although a small fraction is again reflected at the lens–air interface. These rays are reflected within the lens at the interfaces, such that at each reflection a part of the light leaves the lens on either side (Fig. 3.40*b*). Thus, on both sides of a biconvex lens polarised stray light emerges.

A smaller amount of stray light is produced, however, when the lens is plano-convex (Fig. 3.40*c*). Such lenses are especially useful when light illuminates the plane surface of the lens. (In reality this surface is always somewhat curved.) In the case of a lens made from a material of appropriate refractive index and ground as necessary, the rays reflected at the lens–air interface may be totally reflected at the plane surface of the lens. Therefore, it is possible that at the plane surface of a plano-convex lens practically no stray light is produced (Fig. 3.40*d*).[2]

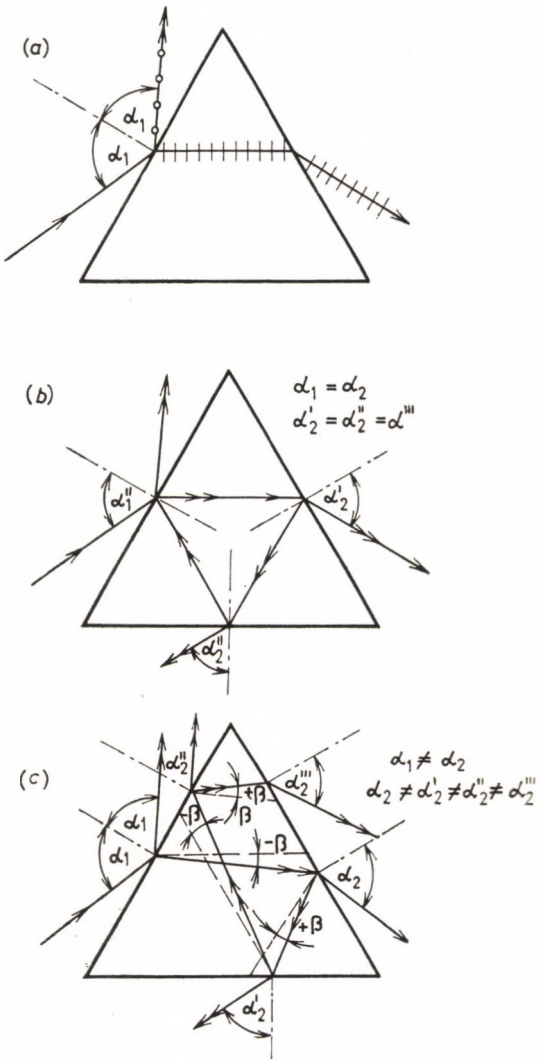


Fig. 3.41 Reflection of light on the surfaces of the 60° prism: (*a*) at the air–prism interface, (*b*) reflection from the prism–air interface of a ray having a wavelength corresponding to the normal position of the prism, and (*c*) reflection of a ray having a different wavelenght

Parallel natural light emerging from the collimating lens falls obliquely onto the prism. The rays reflected at the air–prism surface will be polarised in such a way that their plane of polarisation is perpendicular to the plane of incidence. When 60° prisms are used the direction of these rays tends to deviate considerably from the direction of the rays emerging from the prism (Fig. 3.41*a*). These reflected rays are only troublesome when some other surface reflects them back into the prism.

Some of the rays travelling within the prism, that are already dispersed and are polarised parallel to the plane of incidence, do not leave the prism, but are reflected at the prism–air interface. When the prism is in its minimum position with regard to the wavelength in question, then the direction of the rays repeatedly reflected at the medium interface is always parallel to the opposite side of the prism. That part of the reflected ray which leaves the prism only after the third reflection emerges at an angle identical to that of the original non-reflected ray (Fig. 3.41*b*). The direction of these emerging rays is always definite and no stray light is produced directly. They can, however, be a problem if the rays emerging through the base or the refractive surface of the prism are reflected back into the prism or the camera lens.

A more serious problem is when the rays have a wavelength for which the prism is not in the minimum deviation position. These rays do not travel parallel to the base of the prism, but form an angle $\pm\beta$ with the latter (Fig. 3.41*c*). Thus the reflected rays fall at different angles on the next prism–air interface encountered and also leave the prism at different angles.[3] It can be easily demonstrated that a ray that has been reflected three times at the prism–air interface in a 60° prism leaves on the same side as the primary ray at an angle of α_2'. For this ray the following relationship holds

$$\sin \alpha_2' = n \sin (30° \pm \beta) \tag{3.53}$$

The difference between the angles α_2 and α_2' depends on the wavelength of the light and the refractive index of the prism. Thus, rays with the same wavelength that directly leave the prism and those that emerge from it after three reflections fall at different angles on the camera lens. The lens will not unite these rays into a sharp spectral line at its focal plane. (The intensity of the rays leaving after such a triple reflection is so small that the error arising from this source is practically negligible.)

The stray light arising in the above way produces varying levels of trouble depending on the construction and design of the spectrometer. Generally, most of the light reflected by the lenses and the prism falls on the internal surface of the equipment casing or onto the clamps used for fixing the optical components. The surface of the latter should be machined so that the stray light is absorbed or at least is not reflected onto the focal plane of the collimating lens. This latter requirement is difficult to meet where there is an acute bend in the optical axis, such that the focal plane of the collimating lens and the slit are placed in proximity and not at opposite ends of the instrument.

The effects of stray light can largely be eliminated when the spectrograph is made in separate parts; that is, where there is maximum separation of the collimating lens, the prism housing, and the focal plane of the collimating lens. This design is, however, not feasible for certain instruments. In this case the passage of the stray light should be blocked by shields situated at appropriate positions. It is also an advantage if the internal surface of the equipment is not glossy, but mat and painted black.

It should finally be mentioned that certain types of quartz and fluorite crystals produce fluorescence under the effect of ultra-violet light. The fluorescent light produces the same sort of trouble as stray light.

References

1. KAYSER, H., *Handbuch der Spektroskopie*, vol. I, Leipzig (1903)
2. WIEN, W. and HARMS, F., *Handbuch der Experimentalphysik*, vol. 21, Leipzig, 1927; JOOS, G., and ANGERER, E., *Spektroskopie*, I, p. 234
3. EINHORN, H. D. and COHEN, A. E. I., *J. Opt. Soc. Am.*, **44**, 232 (1954)

3.5.5 Effects of temperature and barometric pressure changes[1–6]

Changes in temperature influence prismatic dispersion in two ways. First, there is a change in the refractive index of the optical materials and, secondly, changes can occur in the dimensions of the instruments that influence the dispersion of the light.

The absolute refractive index of solid bodies changes with the temperature, such that an increase of 1 °C reduces the value of the refractive index by 4–5 digits in the sixth decimal place. Accordingly, the refractive index of, say, quartz in air, measured at 5890 Å changes by $0{\cdot}539.10^{-5}$ for every 1 °C.[1] The absolute refractive index of gases, including air, is also influenced by the humidity level, in addition to temperature and pressure. The well-known relationship between the refractive index, n_0, of a gas, measured at n.t.p., and the refractive index, n, at a certain pressure and temperature, is given by

$$\frac{n-1}{n_0-1} = \frac{\rho}{\rho_0} \tag{3.54}$$

where ρ and ρ_0 are the densities of the gas referred to the examined state, and to the normal state, respectively.

It follows from the previous treatment that the value of the refractive index drops with increasing temperature. Thus, the focal length of the lenses and the angular dispersion of the prisms are reduced. In order to discuss these changes let us consider the following formula

$$\frac{\mathrm{d}\lambda}{\mathrm{d}t} = \frac{\mathrm{d}n/\mathrm{d}t}{\mathrm{d}n/\mathrm{d}\lambda} \tag{3.55}$$

It follows from the above relation that theoretically the change of wavelength with temperature is such that it increases when the temperature coefficient of the refractive index ($\mathrm{d}n/\mathrm{d}t$) increases, but decreases with an increasing value of the dispersion ($\mathrm{d}n/\mathrm{d}\lambda$).

Thus any effect caused by a temperature change is a function of the wavelength as well.[2–4]

In practice, not only changes in the refractive index, but also changes in the prism dimensions with temperature should be considered, together with the related optical properties.

In the case of anisotropic crystals different degrees of dilatation will occur along the different crystal axes. For a quartz prism it can be stated that the refractive angle changes with the temperature and also, according to Eq. (3.21), the angular dispersion is altered.

In prism spectrographs the distances between the slit and the collimating lens and also between the camera lens and the spectral plate change with temperature.

In principle, these changes influence not only the sharpness of the spectral lines, but also the linear dispersion of the instrument, as seen from Eq. (3.32). Therefore, because of temperature changes the position of the spectral lines on the focal plane of the camera lens changes. The longer the wavelength of light, the greater is the shift produced (Fig. 3.42). It can be seen from the diagram that for medium-performance spectrographs a significant shift of the spectral lines is caused by a temperature change of 2 °C.

The shift of spectral lines also depends on the design of the equipment in addition to the physical constants. Values obtained for certain types of spectrograph can only be established by experimental means and have no relevance for spectrographs of a different design.

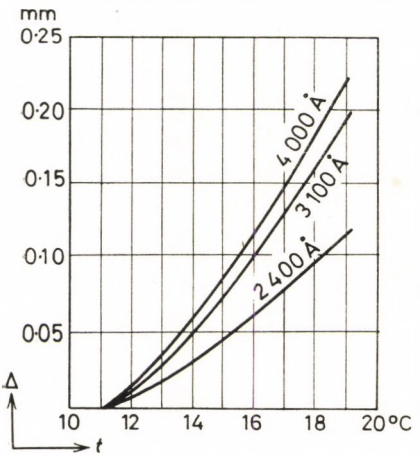


Fig. 3.42 Variation of the Δ shift of spectral lines of different wavelengths with temperature for the Zeiss Q 24 spectrograph[5]

The above statement is also supported by the given diagram. The curves shown in Fig. 3.42 refer to a medium-performance spectrograph of classical design (Zeiss Q 24) whose optical axis (and thus the central line of the supporting casting) is bent at a wide angle. This arrangement has the consequence that the line shift produced by the thermal dilatation of the casting contributes to the shift caused by the optical effects. However, this only occurs up to that value of the wavelength for which the deviation angle of the rays is smaller than the angle enclosed by the optic axis of the equipment. At long wavelengths the two effects add together, while at shorter wavelengths they partially compensate each other.

The absolute refractive index of solid bodies is hardly influenced by *atmospheric pressure changes*. On the other hand, atmospheric pressure influences the absolute refractive index of air and thereby the angular and linear-dispersion of prism dispersion equipment in a way that cannot be neglected. This change is generally $3{\cdot}6.10^{-6}$ mm/torr.[4] The shift in the spectral lines that develops is generally 0·063 mm/10 torr in triple-prism spectrographs with a dispersion of 12·5 Å/mm at 6000 Å and 2·5 Å/mm at 4000 Å. An atmospheric pressure change of 10 torr therefore causes a shift corresponding to about 0·8 Å in the first case and about 0·15 Å in the second case quoted above.[5]

In photographic spectrochemical analysis such a shift of the spectral lines could give rise to errors, not only in the determination of the wavelength but also in the determination of the intensity of the spectral lines. The first type of error can be

demonstrated when the standard iron spectrum and the spectrum of the material to be examined are photographed at different temperatures or at significantly different atmospheric pressures. The second error, which occurs less frequently, is encountered during very long exposures when there is a change in the temperature or atmospheric pressure. In such a case, any shift that occurs leads to a broadening

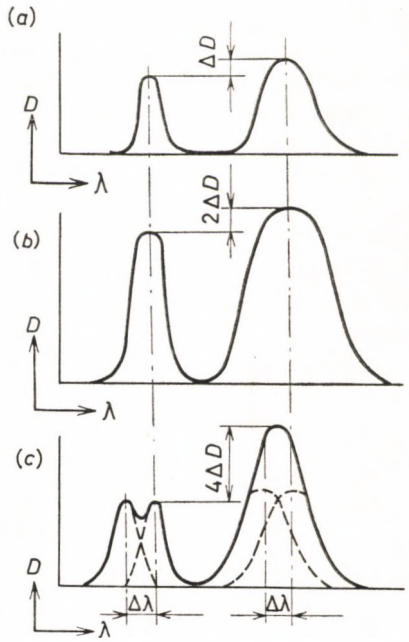


Fig. 3.43 Error arising from the shift of the spectral lines when comparing lines of different blackening (density) value (*a*) blackening-wavelength curve for two spectral lines of different intensity, (*b*) the same, when the spectral lines have shifted for a double exposure of the photographic emulsion, (*c*) the same, but when the spectral lines have shifted during the longer exposure[5]

of the spectral lines. This has the consequence that the light energy per spectral line is distributed over a larger surface of the photographic emulsion. This, in turn, results in a reduction of the line blackening. This phenomenon is especially evident when spectral lines having very different intensities are being compared[5] (Fig. 3.43). A difference in the blackening value emerges because of a line-shift between two lines of the same initial blackening value, but with different profiles.

References

1. D'ANS, J., and LAX, E., *Taschenbuch für Chemiker und Physiker*, 2nd edn, Berlin, vol. 2, 1187 (1949)
2. BLASERNA, J., *Poggendorffs Annalen*, Göttingen–Heidelberg, **143**, 655 (1871)
3. KRÜSS, J., *Ber. chem. Ges. Frankfurt*, **17**, 2732 (1884)
4. MARTON, T. R., *Proc. R. Soc.*, **A, 113**, 704 (1927)
5. SCHÖNTAG, A., *Z. Wiss. Photogr.* **46**, 99 (1951)
6. ELLIOT, A., and CAMERON, W. H. B., *J. Scient. Instrum.*, **14**, 25 (1938)

3.6 Prism dispersion instruments

3.6.1 Single-prism quartz spectrographs

Most quartz spectrographs in normal use are medium-performance, single-prism instruments, that can just resolve the iron lines at wavelengths 3099·97 Å, 3100·31 Å, and 3100·67 Å (the theoretical resolution is 9000–10 000 at 3000 Å). Quartz spectrographs with a lower level of performance (for example, Zeiss Q 12, Hilger Small Quartz) are nowadays little used, although they can still be applied to the examination of materials with simple spectra. The high-performance, multiprism quartz spectrographs have become more widespread now that large prisms and lenses can be manufactured from good-quality fused quartz. The dispersion curves of typical quartz spectrographs are shown in Fig. 3.44, while the practical resolution of the equipment and the minimum wavelength difference of the resolved lines are shown in Figs 3.32 and 3.33 (pp. 255 and 259) respectively.

In addition to the resolvability of the iron triplet, other features are also characteristic of medium-performance quartz spectrographs. For example, the resolved spectrum can be photographed *at once* on a photographic plate that has usually a length of 24 cm. The prism base length is normally about 60 mm (in exceptional cases, as in the ISP 28 equipment, this length is 47 mm) while the base length of fused quartz prisms (for example, Fuess Q 110 M) can be as much as 80 mm.

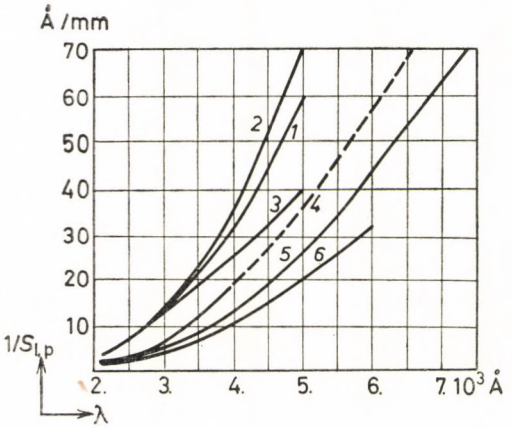


Fig. 3.44 Reciprocal linear dispersion curves of some characteristic quartz spectrographs: (*1*) Zeiss Q 24, (*2*) ISP 28, (*3*) Fuess Q 110 M, (*4*) Huet UV 120 (double-prism), (*5*) Fuess Qu 110 H (four-prism), and (*6*) Littrow-system autocollimating quartz spectrograph

The lens diameter is usually about 50 mm and the focal length of the collimating lens is normally about 550 mm, while the *F*-number of the camera lens normally varies between 1 : 10 and 1 : 12.

In certain medium-performance quartz spectrographs (for example, the Zeiss Q 24, Hilger Medium Quartz, etc.) the optical axis of the instrument is bent at a relatively small angle, an angle determined by the direction of the medium-wavelength rays (Fig. 3.45*a*). A construction of this type has the advantage that the housing of the spectrograph can be separated such that the part of the spectrograph between the slit and the collimating lens, and the prism housing and the photographic camera are fully enclosed. No stray light can, therefore, get from one part to the other. On the other hand, such an arrangement is comparatively voluminous and requires a length of more than 2 m together with the optical bench.

The designers of more recent medium performance spectrographs normally consider such an excessive instrument length a disadvantage. The length of the total equipment can be significantly reduced if a mirror is built into an appropriate place in the spectrograph such that a sharp bend in the optical path is produced. In constructions of this type it should, however, be borne in mind that unwanted reflection may cause problems. When the angle of reflection is too acute, errors of this type can only be eliminated by placing light barriers at appropriate places to block off the unwanted radiation.

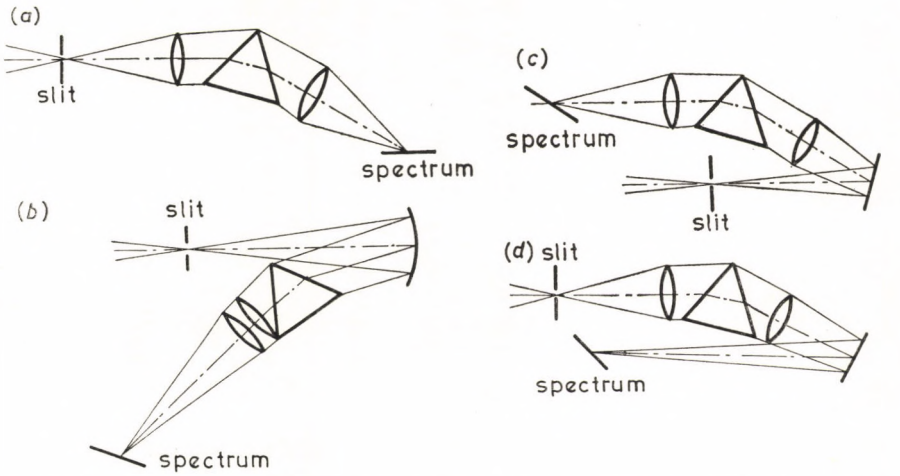


Fig. 3.45 Optical scheme of different types of single-prism quartz spectrographs: (*a*) Zeiss Q 24, (*b*) ISP 28, (*c*) Fuess Q 110 M, and (*d*) Gaertner spectrograph

Among the spectrographs whose construction is not classical are those where the collimating lens is replaced by a concave metal mirror. Such a design is found in the ISP 28 and the ISP 30 quartz spectrographs (Fig. 3.45*b*). In this equipment the collimating mirror reflects the light onto the prism at a very acute angle; usually about 2°. An acute bend is necessary since at higher angles such a high degree of astigmatism is produced that would be very difficult to correct. The application of a small deviation angle has, however, the drawback that the prism cannot be placed immediately behind the collimating mirror. A relatively large distance is necessary to achieve a sufficient separation of the rays falling on and reflected by the collimating mirror, even if a relatively small prism is used. The latter construction has the drawback that the path of the rays within the spectrograph is thus greatly increased. This causes a significant drop in the intensity of the light getting into the instrument and no reduction in the volume of the spectrograph is achieved.

The problems mentioned above are not encountered in medium-performance quartz spectrographs (for example, the Fuess Q 110) where the rays coming from the slit are reflected by a planar mirror onto the collimating lens (Fig. 3.45*c*). A further advantage of this construction is that the units of the spectrograph are separate from one another and the harmful effects of stray light can therefore be easily eliminated. In the previously quoted equipment it is also an advantage that the prism and lenses are made of fused quartz. Thus it becomes feasible to increase the size of the

prism and the lenses compared with prisms made from natural quartz. This is certainly advantageous, from the aspects both of the resolving power and of the photographic speed of the spectrograph. By the insertion of a mirror, the volume of the spectrograph can be reduced to about half that of classically constructed spectrographs having the same performance. Also, the fact that the slit and the plate holder of the spectrograph are closely situated renders the use of the instrument much easier.

The volume of spectrographs can also be reduced by inserting a plane mirror into the light path so that the camera lens projects the image onto the spectral plate via the mirror (Fig. 3.45*d*). This design has the additional advantage that by appropriately setting the 60° Cornu prism and the metal mirror, the ray corresponding to the spectral line in the middle of the photographed spectrum passes symmetrically through the prism. Gaertner-system quartz spectrographs are built according to the former principle. As the lenses of this equipment have relatively long focal lengths (1700 mm), their linear dispersion is so large that the length of the spectrum between 2000 and 8000 Å is about 35·5 cm. In this case the spectrum cannot be photographed all at once. The required spectral interval can be 'automatically' adjusted by means of a setting screw placed below the slit of the spectrograph, such that the wavelength limits of the interval can be read off on a projection screen fitted above the spectrograph. The handling of the equipment is therefore extremely simple.

3.6.2 Autocollimating quartz spectrographs

As single-prism quartz spectrographs are only of medium performance, their resolving power is not always sufficient for the study of dense spectra. A limit is set to the resolution obtained since it is difficult to produce sufficiently large, optically perfect quartz prisms. The performance of the spectrograph can, however, be increased either by the application of the autocollimation principle or by using more prisms. The former method is commonly used.

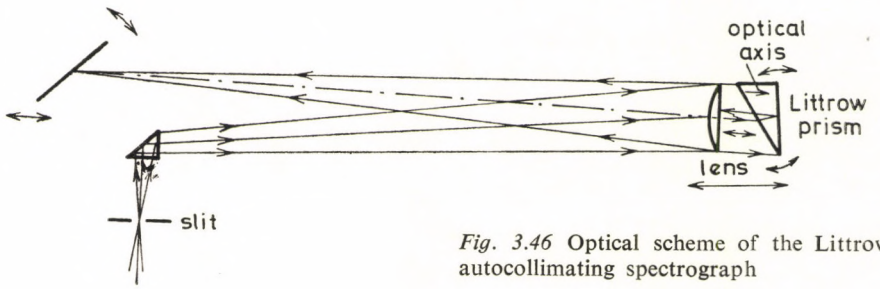


Fig. 3.46 Optical scheme of the Littrow-system autocollimating spectrograph

The base length of the 30° refraction angle, Littrow prism in Littrow-system, autocollimation quartz spectrographs is about 70 mm. Optically, the prism corresponds to a 60° prism with a base length of 140 mm. A further advantage inherent in its application is that any error produced by the birefringence of the quartz is compensated, as in the case of the Cornu prism. The focal length of the plano-convex lens that is used as the collimating lens or camera lens is relatively large (about 2000 mm) and has an *F*-number of about 1 : 35.

When a Littrow prism is used the refracted rays leave the prism on the same side as the incident rays. This property determines the structure of Littrow-system autocollimating spectrographs; that is, these spectrographs do not have a separate collimating or camera lens (Fig. 3.46).

These lenses are replaced by a single combined collimating-camera lens. This lens collimates the rays coming from the slit, and then combines the transmitted parallel beams, which diverge according to their wavelength, into spectral line images. To meet the necessary requirements in making such a lens involves a certain amount of difficulty. This provides a partial explanation for the fact that the *F*-number of the lens of Littrow-system spectrographs is always small compared with that of other spectrographs. This, in turn, has the disadvantage that the speed of Littrow-system autocollimating spectrographs is relatively low.

The chromatic aberrations of the long-focal-length quartz lenses used in Littrow-system spectrographs cannot be corrected to such a degree that little variation in the focal length takes place. A sharp spectrum, therefore, can only be achieved when the distances between the slit and the lens, and between the lens and the spectral plate, as well as the tilt angle of the spectral plate, are adjusted according to the wavelength of the spectral range under study. When the tilt angle of the spectral plate is changed the prism should also be rotated, so that the medium-wavelength rays of the spectral range travel symmetrically through the prism.

A major difficulty with Littrow-system autocollimating spectrographs is that the spectrograph housing cannot be separated into different parts. Accordingly, it is difficult to eliminate *stray light* and its harmful effects. Thus, problems arise because the rays coming from the slit are partially reflected at the convex, air–lens interface (Fig. 3.47). These divergent rays will arrive undispersed at the spectral plate. Another source of error is that the light reflected from the front surface of the prism is also undispersed. Due to the oblique position of the front part of the prism, however, the light reflected in this way will only fog the spectral plate when it is reflected again from other surfaces.

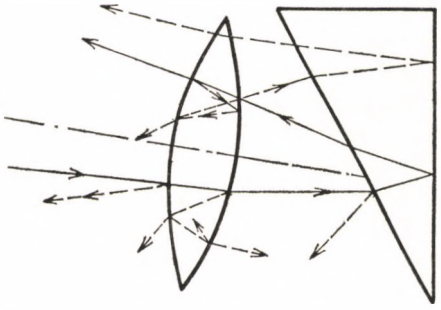


Fig. 3.47 Reflection of light arising on the prism and lens of Littrow-system autocollimating spectrographs

Another source of trouble may arise from the presence of rays that are already dispersed but are reflected from the prism-facing side of the *biconvex lens* onto the prism. The path of these rays is not parallel to that of the original, non-dispersed rays made parallel by the lens. Consequently, the optical path of the reflected and dispersed rays within the prism and thus the direction at which they emerge are not the same as the original rays having the same wavelength. Thus, the sharpness of the spectra is reduced. This error can be largely eliminated when the lens is plano-

convex and its plane surface directed towards the prism. Under such conditions the rays coming from the prism strike the plane surface of the lens almost at right angles and are therefore not reflected.

A further source of error is inherent in the construction when the light enters the instrument through a slit placed at the side of Littrow-system spectrographs. The light is then projected by a 45° deviation prism (or plane mirror) onto the lens. The photographic plate is situated in the neighbourhood of the slit. Thus, the paths of the initial non-dispersed rays falling on the lens and the rays dispersed by the lens are not only parallel, but almost coincide. Reflected light can, therefore, very easily cause spectral plate fogging. This error can be eliminated by placing the lens at a slight angle within the instrument. In this construction, however, one should bear in mind that, because of astigmatism the spectrum is not entirely sharp. Another and better construction is where the unwanted radiation is shielded from the spectral plate by appropriately placed light-traps in the blackened interior of the equipment. In certain cases (for example, Hilger F478) a horizontally directed shield is placed in front of the lens on the side facing the slit. The shield is narrow and is oriented parallel to the direction of dispersion. This construction naturally entails a certain reduction in resolving power and photographic speed.

The focal length of the lens of Littrow-system spectrographs and thus the linear dispersion of the equipment are relatively large. The 2000–8000 Å portion of the spectra produced by these quartz spectrographs is generally about 60 cm long. For optical and technical reasons, however, only a third or a quarter of this length can be photographed at once. In order to photograph the different spectral ranges the following changes should be carried out:

(1) Change the distance between the lens and the slit.
(2) Change the angle of incidence by turning the prism.
(3) Change the distance between the spectral plate and the lens.
(4) Change the tilt angle of the spectral plate.

The necessary adjustments can be made with automatic, autocollimating spectrographs using a single screw situated in the vicinity of the spectral plate holder or cassette. The spectral range available for photographing can be read off from a scale.

It is an advantage of Littrow-system spectrographs that they usually have easily replaceable quartz and glass optics, so that they can be used for both the ultraviolet and the visible range. The interchange of the two optical systems can be carried out externally in some modern equipment (for example, Hilger E 478) without the necessity of rearranging the whole instrument.

The Krampl–Glock quartz spectrograph (a non-commercial model) has a construction that is optically interesting and practically very useful.[1] In this arrangement, based on an autocollimating single-prism spectrograph, the light is dispersed by a 60° Cornu prism, instead of a Littrow prism. Here, the rays emerging from the prism are reflected back again into the prism by means of a plane mirror placed immediately behind it (Fig. 3.48). The prism, therefore, effects a double deviation of the rays according to their wavelength. The theoretical resolving power of the prism is twice as large as that of single-prism spectrographs having the same base length.

In a spectrograph of the above type the light enters the instrument through a slit at the side. The light is reflected by the deviation mirror onto a concave metal mirror

instead of a lens. The slightly obliquely-set concave mirror, acting as the collimating mirror, projects the parallel rays onto the Cornu prism in such a way that these rays pass below those coming from the slit. The beam that has passed through the prism is then reflected back to the latter by means of a plane metal mirror placed behind it, so that the rays emerging from the prism fall on the concave metal mirror at a different height. The concave metal mirror, which acts as the camera lens, reflects these rays back again at a different height and combines them into line images on the spectral plate.

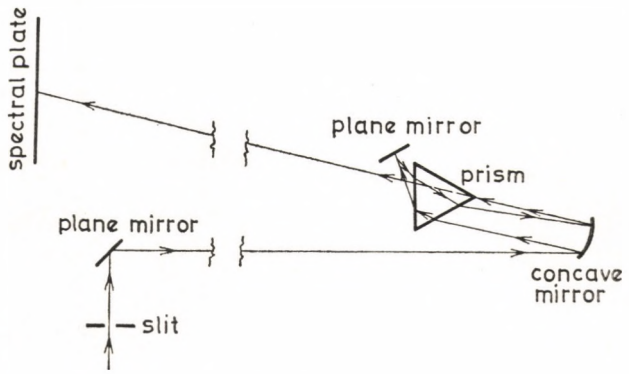


Fig. 3.48 Optical scheme of the Krampl–Glock-system quartz spectrograph

The great advantage of this equipment is that the focal length of the concave mirror is independent of the wavelength. Thus, it is possible to photograph a 40 cm long portion of the spectrum from 2000 Å to 5000 Å in one exposure, if necessary. Compared with the Littrow-system autocollimating spectrograph, it is a distinct advantage that any portion of the spectrum can be photographed without the need to rearrange the optical elements. In this case the photographic plate is placed at the appropriate location of the focal surface of the concave mirror. The complicated, precisely made equipment necessary to change the spectral range in 'automatic' Littrow spectrographs can therefore be dispensed with.

Reference

1. DEBOL, C. H. and HAULE, W., *Arch. EisenhüttWes.*, **28**, 127 (1958)

3.6.3 Multiprism quartz spectrographs

As seen from Figs 3.32 and 3.33 (pp. 255 and 259) the resolving power of Littrow-system quartz spectrographs is about double that of the usual single-prism quartz spectrographs. The performance of the autocollimating spectrographs cannot, however, be significantly increased above the present level. The limits are set by the dimensions of the prisms. A significantly higher resolution level can be attained only by using a larger number of prisms. In multiprism dispersion a given level of resolution can be obtained by using much less prism material than in the case of single-prism dispersion. The quantity of material needed for the construction is inversely proportional to the square of the number of prisms. This fact is especially important in the case of quartz spectrographs. The difficulties encountered in making an optically perfect prism of large enough size are naturally reduced when smaller prisms are to be used.

A further advantage of multiprism dispersion is that the adjustment of the two separate lenses – the collimating lens and the camera lens – can be made more easily than the optical corrections necessary for the single, common lens in the Littrow-system spectrograph. As a consequence of this the spectral lines are sharper. It also becomes unnecessary to have a very large *F*-number for the lenses and a large tilt angle for the spectral plate. Thus, multiprism spectrographs are much faster photographically than Littrow-system equipment having the same level of performance. Furthermore it is important that with multiprism dispersion there is practically no increase in the background level due to stray light. This is essential for quantitative trace analysis.

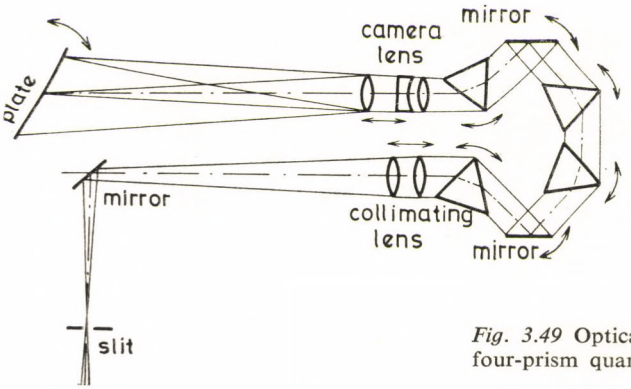


Fig. 3.49 Optical scheme of the Fuess Q 110 H four-prism quartz spectrograph

In the design of multiprism quartz spectrographs it should be taken into account that composite quartz lenses cannot be corrected to such an extent that they meet all the requirements necessary for the collimating and camera lenses. Therefore, in high-performance quartz spectrographs with more than two prisms, the photographed spectral range can only be changed if the positions of the collimating and camera lenses, and the tilt of the spectral plate, are simultaneously altered.

In the construction of multiprism quartz spectrographs it has to be borne in mind that the equipment should not be too large or too difficult to operate. These requirements can only be satisfied by placing mirrors at different points in the spectrograph to bend the optical path.

The modern, very-high-performance, four-prism quartz spectrograph whose optical scheme is shown in Fig. 3.49 has been manufactured on consideration of the above points. The resolving power of this equipment is so high that the apparent iron triplet at 3100 Å is resolved into four component lines. Thus, at this wavelength the resolving power is the same as a fourth-order spectrum of a grating that has 45 000 lines. According to the diagram given in Fig. 3.44 (p. 274) the dispersion of the four-prism spectrograph, in the medium-wavelength portion of the available spectral range, is lower than that of Littrow-system spectrographs. This contradiction can be explained by considering that the focal length of the lens in the latter spectrograph is about twice as long as the focal length of the lenses of the four-prism quartz spectrograph. Littrow-system spectrographs, therefore, suffer from over-magnification and accordingly their speed is essentially lower than that of four-prism quartz spectrographs.

Another advantage of the equipment is that it is easy to operate, as well as relatively small. This is achieved by inserting two plane deviating mirrors between the four fused quartz prisms. Also, as the spectral range is changed the prisms and the mirrors between them rotate automatically, so that the medium-wavelength rays of the spectral range studied passes through the prisms at the optimum minimum deviation angle. The 'automatic' adjusting device also sets the appropriate focal length for the collimating and camera lenses and the tilt of the spectral plate as well. The 2100–8000 Å portion of the spectrum is about 80 cm long.

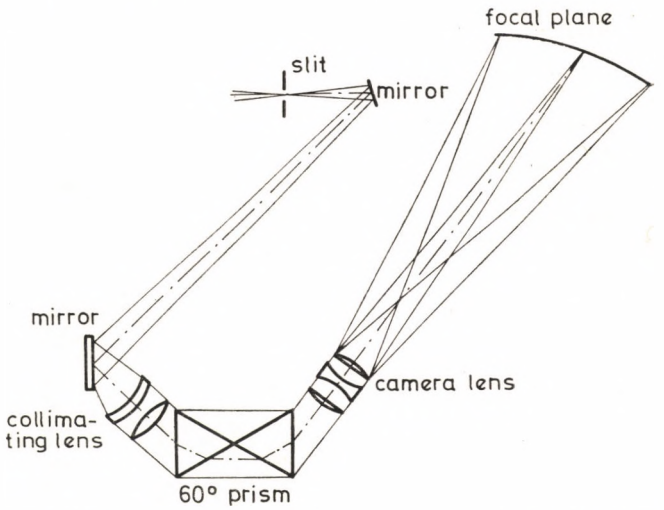


Fig. 3.50 Schematic optical diagram of the Huet (Paris) UV 120, double-prism quartz spectrograph

The modern, double-prism quartz spectrograph shown in Fig. 3.50 has composite lenses whose focal length is relatively long (about 1200 mm). To save space the rays travel mainly in the perpendicular position instead of passing horizontally. The 2100–4000 Å portion of the spectrum has a length of 40 cm and can be photographed on two spectral plates having a length of 24 cm, while the spectral region between 4000 and 10 000 Å is photographed on spectral plate of 24 cm long.

3.6.4 Low- and medium-performance glass spectrographs

Low-performance glass spectrographs are single-prism instruments. The only difference between their optical construction and the design of single-prism quartz spectrographs is that the lenses of the former are always composite. Moreover, these lenses are carefully adjusted and corrected for both chromatic and spherical aberrations. Accordingly, the plane of the spectral plate is perpendicular to the optical axis of the camera lens. A further characteristic of these instruments is that they are able to resolve the sodium D-doublet 5890·0 and 5895·9 Å; that is, their resolving power is about 1000 at this wavelength.

Medium performance glass spectrographs are multiprism devices with dispersive systems either made on the Försterling prism principle or composed of three 63°

prisms. In the former case, the optical path is bent at 90°, while in the case of three-prism dispersion the rays are deviated by almost 180° in the instrument. A further characteristic of medium-performance glass spectrographs is that their camera and camera lens can be easily interchanged with another camera and lens of different focal length. Thus, by using the same glass spectrograph spectra can be produced that are of different magnification and have a different linear dispersion. The resolving power of the instrument remains unaltered (curves *1* and *2* in Fig. 3.51).

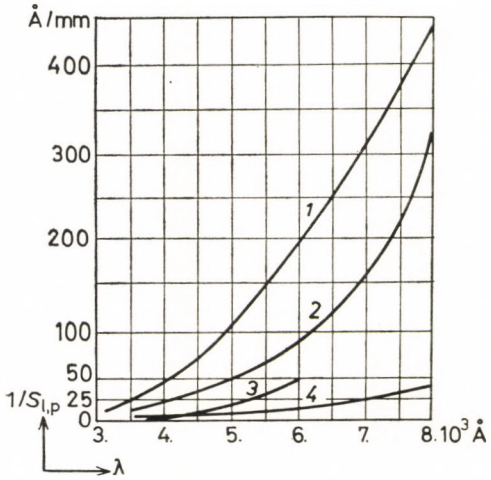


Fig. 3.51 Reciprocal linear dispersion of glass spectrographs: (*1*) medium-performance equipment with variable spectral ranges (f_{col}= 120 mm), (*2*) the same, but f_{col} = 270 mm, (*3*) Littrow-system spectrograph (f_{col} = 1700 mm) and (*4*) high-performance autocollimation spectrograph with *Försterling* prism system (f_{col} = 2075 mm)

The medium performance glass spectrographs can be grouped into two classes. The first group constitutes equipment that provides for the overall photographing of the total spectral range. In the second group only a certain portion of the spectrum can be studied at any one time. Thus, there are medium-performance glass spectrographs of fixed and variable spectral range.

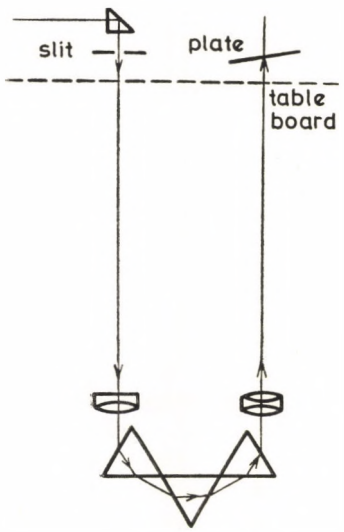


Fig. 3.52 Schematic optical diagram of the Steinheil, G.B. Universal glass spectrograph, with non-variable spectral range (*perpendicular* section)

The advantage of fixed-spectral-range glass spectrographs is that the lenses and prisms of the instrument are fixed. The handling of the spectrograph is therefore very simple. On the other hand, a major drawback is that this equipment can only be used within a relatively narrow spectral interval – usually between 4000 and 6000 Å. This limitation is made necessary partly by optical and partly by photographic factors. The first reason is that chromatically corrected lenses cannot be manufactured such that the focal plane of the camera lens is entirely planar in a spectral range larger than that given above. Furthermore, no photographic emulsions are produced that would be appropriately sensitised for a wider spectral interval.

Among commercial, fixed-spectral-range glass spectrographs, the model shown in Fig. 3.52 has an interesting construction. The object was to design a very small instrument. This is secured by the fact that a 90° deviation prism projects the horizontal rays coming from the light source vertically onto a slit with a horizontal axis. The prism system, made from three 63° glass prisms, deviates the dispersed rays almost perpendicularly upwards. Thus, the focal plane of the camera lens, situated in the neighbourhood of the slit, forms a relatively small angle with the horizontal. (An advantage of this construction is that the dispersion device is situated directly below the top of the working bench.)

By using medium-performance, variable-range glass spectrographs the 3600–8000 Å (or perhaps 3600–10 000 Å) spectral range can be photographed in detail (Fig. 3.53). Optically, it is an advantage that the prisms can always be set to the minimum deviation position corresponding to the intermediate wavelength. As the wavelength range is varied, the distance between the slit and the collimating lens, and between the camera lens and the spectral plate must also be changed. This is necessary since the lenses cannot be chromatically corrected for the full wavelength range of the spectrograph.

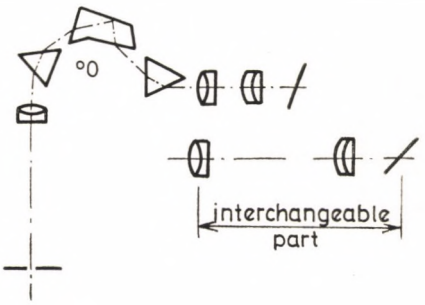


Fig. 3.53 Optical scheme of the variable spectral range, glass spectrograph

In commercially available medium-performance, variable-range glass spectrographs, the dispersive element is generally a Försterling-prism system that has a total base length of about 210 mm. Modern equipment is generally 'automatic'. The prisms can be collectively rotated by means of a calibrated adjusting screw. The calibration scale usually shows the medium wavelength of the spectral range available for photographing. The distance between the camera lens and the plate holder, and between the slit and the collimating lens, as well as the tilt angle of the plate holder, should be adjusted separately.

In medium-performance, variable-range glass spectrographs the length of the spectral interval that can be photographed at any one time is comparatively short. Depending on the magnification of the camera lens it is usually only about 5 or 10 cm

The usual length of the photographic plates employed is normally 6·5, 9 of 12 cm. The theoretical resolving power of these spectrographs is usually 67 000 at 4000 Å, or 16 000 at 6000 Å.

Littrow-system spectrographs equipped with glass optics are essentially medium-performance glass spectrographs, in spite of the fact that their linear dispersion is about the same as high-performance glass spectrographs (see curve 3 of Fig. 3.50). This fact can be explained by considering that the theoretical resolving power of glass Littrow prisms with a refractive angle of 26° and a base length of 55 mm is only abo ut 40 000 at 4000 Å and 7000 at 6000 Å; that is, it is significantly lower than the reso lving power of the previously described medium performance glass spectrographs.

In summary it can be stated that the most useful medium-performance glass spectrographs for spectrochemical analysis are those with a resolving power between 60 000 and 70 000 at 4000 Å and a reciprocal dispersion of at least 20 Å/mm at the same wavelength. Thus, glass spectrographs with a higher performance level are only recommended for the study of spectra with very many lines.

3.6.5 High-performance glass spectrographs

Without exception all of these instruments contain some form of autocollimating device. The only principal difference between them and Littrow-system, autocollimating spectrographs is that the Littrow prism is replaced by a higher performance Försterling system, or by a system composed of three 63° glass prisms. In either case the rays emerging from the last prism are projected by a plane mirror back into the prism system, as shown in Fig. 3.54. The final prism–mirror combination,

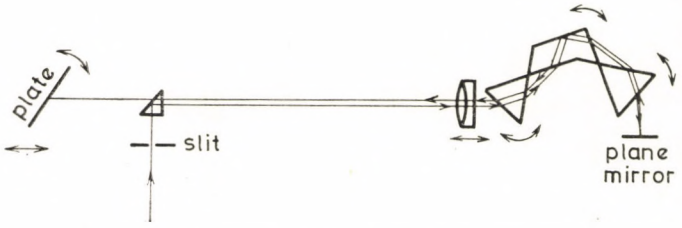


Fig. 3.54 Optical scheme of the Försterling prism-system, autocollimating glass spectrograph

however, can be replaced by a Littrow prism as well. The theoretical resolving power of the prism system is thereby increased twofold. The high resolving power can only be produced, however, when the linear dispersion of the combined camera–collimating lens is sufficiently great. In this type of spectrograph the same errors are found as in Littrow-system autocollimating spectrographs. Light losses arising from reflection tend to increase, because the number of air–prism interfaces encountered are much greater. Thus, in the design of high-performance autocollimating spectrographs stray light must be properly eliminated.

In high-performance spectrographs the required spectral interval can only be changed if the interchanging device alters the position of the lens and the prisms, the distance between the lens and the spectral plate, the tilt angle of the latter, and

finally rotates the plane mirror or Littrow prism behind the prisms, by an appropriate amount. The interchange device is, therefore, rather complicated.

The dimensions of the prisms in the Försterling system cannot be increased above a certain practical limit. Thus, the resolving power of autocollimating glass spectrographs equipped with a Försterling prism system is limited. On the other hand, the size of the 63° glass prisms can be increased twofold over that of usual glass prisms. The resolving power of a spectrograph equipped, for example, with three prisms having a base length of 150 mm can reach the performance level of Försterling-system autocollimating spectrographs even without any autocollimation. The theoretical resolving power of glass spectrographs having a Försterling prism system with a total base length of 420 mm is 134 000 at 4000 Å. However, the resolution figure for three-prism equipment with a total prism base length of 450 mm, is 140 000 without autocollimation and 280 000 with autocollimation, at the same wavelength as before. This high performance is only realised when the reciprocal dispersion of the equipment is sufficiently great; that is, smaller than 1 Å/mm at 4000 Å. Such a high spectrographic dispersion can only be achieved if the focal length of the lens

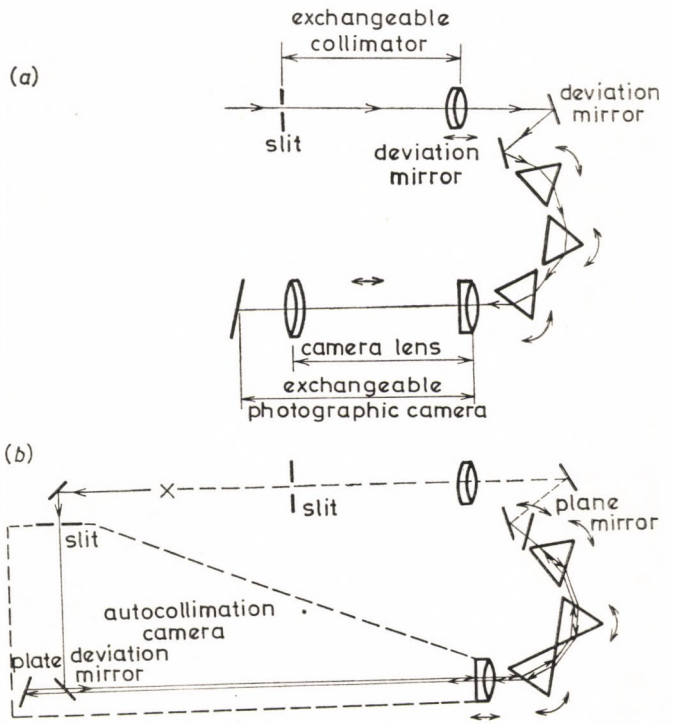


Fig. 3.55 Schematic optical arrangement of the high-performance, three-prism glass spectrograph (ISP 67): (*a*) without the autocollimation camera, and (*b*) with the autocollimation camera

is very large as well (about 3000 mm). The high-performance instruments are not only voluminous but are characterised by their low speed. Accordingly, high-performance autocollimating glass spectrographs are only used for spectrochemical

analysis in exceptional circumstances. Thus commercially designed, high-performance glass spectrographs can operate both with or without the autocollimating facility. In this case a Littrow prism cannot be used instead of the prism–mirror combination. The equipment in Fig. 3.55 has this form of construction. Most of the medium-performance glass spectrographs can be turned into high-performance instruments by placing a plane mirror instead of the camera lenses. These camera lenses are usually interchangeable. In addition the equipment should be fitted with an autocollimating camera.

3.6.6 Liquid-prism spectrographs

It should first of all be mentioned that liquid-prism spectrographs are not generally used. One of the main reasons for this is the relatively large temperature coefficient of the refractive index of liquids compared with solid substances (see Tables 3.3 and 3.4 pp. 240 and 244).

Quartz spectrographs cannot generally be used below 2000 Å, neither can spectrographs with fused quartz optics below about 2100 Å. This is due to the absorption of the shorter-wavelength radiation by the optics of quartz spectrographs. This drawback can be practically eliminated if the lenses are thin and the prisms are liquid, enclosed by thin quartz sheets. Taking into account that the lower limit of optical transmittance of quartz crystals is about 1850 Å, liquid prisms, filled with an appropriate liquid, can be used down to about this limit. The optimum upper limit is set by the solid quartz spectrographs.

According to the data given in Table 3.4 (p. 244) water is a satisfactory prism liquid as it has the properties required to attain the spectral range mentioned above. Between 1860 and 1900 Å the dispersion of water is about 20% higher than that of quartz.

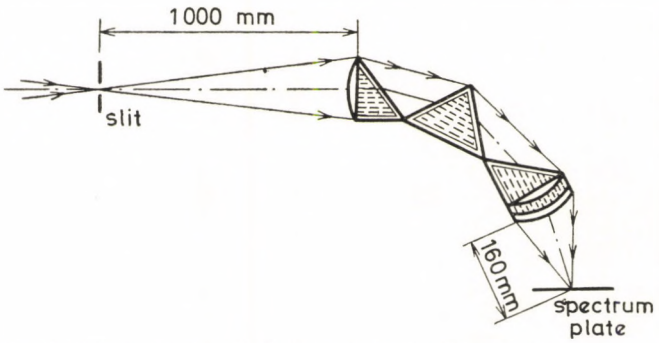


Fig. 3.56 Schematic optical diagram of the Kipp and Zonen (Delft) quartz–water optics spectrograph

At the same time the temperature coefficient of the refractive index of water is the smallest of the liquids generally used as refractive media. As no significant light losses are produced even at an angle of incidence of 70°, the refractive angle of quartz–water prisms is generally greater than 60°, or more specifically 72°.[1] On consideration of these facts it follows that spectrographs with quartz–water optics can be recommended for use between 1800 and 2200 Å.

Among the commercially available quartz–water-prism spectrographs the design shown in Fig. 3.56 is interesting. It has a prism system that is composed of a prism

with a refractive angle of 72° and two others with angles of 36° each. From the optical point of view the three prisms are equivalent to two 72° prisms. The use of two half-prisms allows the front part of one half-prism to be used as the collimating lens, while leaving the other half-prism to act as the camera lens. By this design the number of reflecting surfaces and absorbing quartz layers can be reduced. It is also worth mentioning that the focal length of the collimating lens is relatively large, while that of the composite camera lens is relatively very short. This is due to the spherical and chromatic aberrations of the collimating lens being eliminated by the long focal length of the lens. The achromat camera lens can be easily corrected in the relatively narrow spectral interval of 1800–2200 Å, even if the focal length of the lens is short. A lens of short focal length will of course increase the speed of the spectrograph.

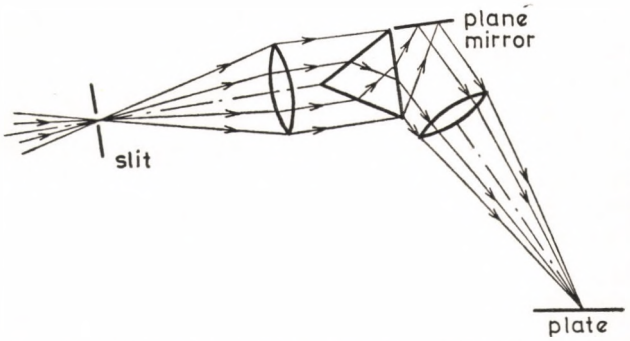


Fig. 3.57 Schematic optical diagram of the liquid-prism spectrograph showing so-called "economic prism arrangement"

Another interesting liquid-prism spectrograph is one in which the light beam made parallel by the collimating lens illuminates two sides of the liquid prism (Fig. 3.57). Thus, the light beam is effectively halved. Accordingly, those rays that pass symmetrically through the prism are parallel either to one side of the prism or to the other. The prism, therefore, virtually has two bases and its theoretical resolving power is much higher than that of normally used prisms. A certain resolution can, therefore, be realised by using a smaller prism.[2]

References

1. DECLAVA, J. and JEANTOT, P., *Revue Opt. Théor. Instrum.*, **2**, 384 (1923)
2. RUMSEY, N. J. O., *Nature, Lond.*, **195**, 168 (1962)

3.6.7 Prism vacuum spectrographs

The previously described prism spectrographs, apart from the very rarely used quartz–water types, are not applicable for the study of spectral regions below 2000 Å. This part of the spectrum is absorbed not only by quartz but also by the air. The lower transmittance limit of quartz is about 1850 Å. The air absorbs a large fraction of the light below 2000 Å, while there is almost complete absorption within a few millimetres below 1800 Å.[1] To study the spectral region below 2000 Å, therefore,

specially constructed *vacuum spectrographs* are available. These are spectrographs that can be evacuated and whose optics freely transmit the radiation in the 'vacuum-ultra-violet region' that is, below 2000 Å.

The applicability of vacuum spectrographs is mainly determined by the optical properties of the prism material. As seen from the data in Table 3.3 fluorite, artificial calcium fluoride, or even lithium fluoride can be applied for this purpose.

The transmittance of good-quality natural fluorite (CaF_2) quickly decreases at wavelengths shorter than 1250 Å. Crystals not pure enough will absorb light below 1350 Å. In principle, the use of synthetic calcium fluoride is better than natural fluorite as it can be used down to 1100 Å. On the other hand, natural fluorite is harder and produces a good finish, whereas artificial calcium fluoride is difficult to grind flat.

Likewise, synthetic lithium fluoride can also be used to about 1100 Å. It has, however, the disadvantage that after prolonged and intense ultra-violet irradiation the lower transmittance limit shifts towards the longer wavelengths by a few hundred angstroms. This does not occur with natural fluorite. Therefore, for the study of the vacuum-ultra-violet spectral range the ideal prism material is natural fluorite. The only drawback is that natural fluorite crystals of appropriate size and quality are very rare.

In principle, the lower wavelength limit of prism vacuum spectrographs is about 1300 Å. However, commercial prism vacuum spectrographs can be used generally only between 1600 and 1850 Å. This limitation is necessary because the lenses of spectrographs with fluorite optics can only be chromatically corrected in a narrow spectral range. For technical reasons, the operations required to change the spectral range are very difficult in the case of the vacuum spectrographs.

The normal prism vacuum spectrographs are usually multiprism or, less frequently, Littrow-system, autocollimating instruments.[2] This is because of difficulties encountered in the manufacture of prisms large enough to give the required resolution level. The optical system of prism vacuum spectrographs is principally the same as that of other spectrographs, although there is a significant difference in the overall construction.

With vacuum spectrographs, not only the spectrograph housing has to be evacuated, but it is also necessary to eliminate absorption of the short-wavelength radiation in the section between the light source and the slit. Thus, it is necessary for the light source to be placed in an area that can be evacuated.

For the construction of the equipment it should be remembered that although high-performance (heavy-duty) vacuum pumps can very quickly evacuate large volumes, the gases adsorbed on the surfaces of the evacuated spectrometer can only be removed with difficulty. It is therefore advisable to have only a small space that requires evacuating with little surface area.

Concerning the above requirements, the casing of modern vacuum spectrographs and the housing of the light source should be evacuated separately. For this reason these two housings are separated by a sheet of calcium fluoride. Thus, when the samples are changed only the light source housing need be evacuated again. It has been found in practice that much time can be saved if instead of evacuating the light source housing, the air is expelled by flushing with a gas that does not absorb in the vacuum-ultra-violet radiation. A noble gas, such as argon, can be used for this purpose. This gas does not emit continuous radiation in this spectral range (Fig. 3.58).

The slit of vacuum spectrographs is generally fixed at about 0·01 mm or as appropriate for the given case. It is advisable to have a curved slit as the lenses of vacuum spectrographs cannot be sufficiently corrected for spherical aberrations. Provided the slit has the proper curvature the spherical aberration of the lenses produces straight spectral lines.

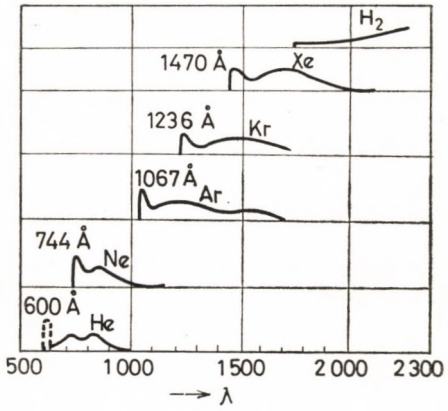


Fig. 3.58 Continuous radiation spectrum of the noble gases[3]

It is advisable to place the plate cassette of vacuum spectrographs into a small, evacuable enclosure. Thus the rest of the equipment remains evacuated during plate replacement.

The optimum refractive angle of the prisms of vacuum spectrographs is not 60°, but 66°. This angle gives a better angular dispersion and no harmful reflection occurs at the vacuum–fluorite interface in this case. The dimensions of the fluorite prisms are small, their base length being usually about 25 mm.

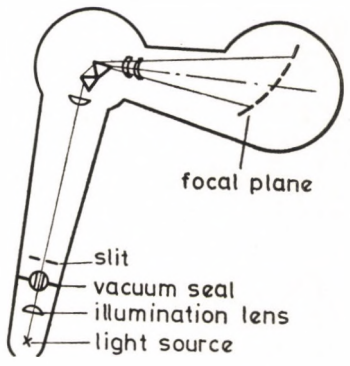


Fig. 3.59 Schematic optical diagram of the vacuum polychromator (Hilger Polyvac) with fluorite prism[4]

Corresponding to the small dimensions of the prisms, the free diameter of the lenses is also relatively small. Sufficient illumination can therefore be secured only by a relatively short focal length of about 100–150 mm. The distance between the plano-convex collimating lens and the slit can be adjusted in certain equipment such that it does not exactly correspond to the focal length of the medium wavelength in the spectral interval, but differs from the latter by about 1%.[2] Thus, the convergent rays passing through the prism give rise to astigmatism. In such a way it is possible that there is only a very small difference between the plane and the focal surface of the camera lens. A further prerequisite is that the camera lens be

a composite one corrected for chromatic aberrations. Even with the most careful corrections to the camera lens it is still not possible to produce a spectral width in excess of 5–6 cm without bending the spectral plate (Fig. 3.59).

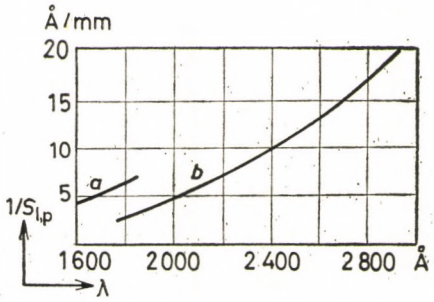


Fig. 3.60 Dispersion curve of vacuum spectrographs: (*a*) three-prism, vacuum spectrograph (collective base length 75 mm, $f_{cam} = 1{\cdot}52$ mm) and (*b*) the polychromator shown in Fig. 3.59 (collective base length 89 mm)

The difficulties mentioned previously are not encountered if the light intensity of the analytical lines is determined by photoelectric means. In this case no problems occur due to the curvature of the focal surface of the camera lens. The exit slits of the instrument can be set without any difficulty even if the focal surface is strongly bent. The equipment based on this principle and illustrated in Fig. 3.59 can be used for a relatively broad spectral range; for example, between 1700 and 3100 Å. The dispersion curve of two-prism instruments based on this principle is shown by curve *b* in Fig. 3.60.

References

1. SCHNEIDER, E. C., *Phys. Rev.*, **45,** 152 (1934); **49,** 314 (1936)
2. ANGERER, E., *Apparate und Methoden der Spektroskopie*, Leipzig (1927)
3. PRICE, W. C., Spectroscopy in the Vacuum Ultraviolet, in THOMPSON, H. W. (ed.), *Advances in Spectroscopy*, New York, vol. I, 58 (1959)
4. Hilger Catalogue C. H. 414

3.6.8 The Fery spectrograph

The Fery spectrograph is an autocollimating quartz spectrograph that is based on an optical principle very different from that employed in other prism spectrographs. This difference is not only manifested by the optical construction of the instrument but the formation of the spectral lines is also different.

The dispersion part of the Fery spectrograph is very simple and consists merely of a slit and a special autocollimating prism that replaces both the collimator and the camera lens (Fig. 3.61). This special prism has curved refractive surfaces rather than the usual planar surface. Furthermore, it is necessary that the convex rear

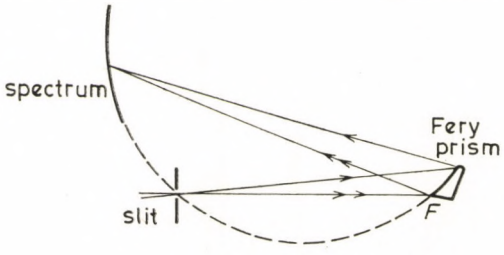


Fig. 3.61 Schematic optical diagram of the Fery spectrograph

side of the prism acts like a mirror. The Fery prism is therefore an advanced version of the Littrow prism.

In the Fery spectrograph, the spectral lines are not true images of the slit, but are astigmatic focal lines. Therefore, at a point in the slit only a certain part or portion of the spectral line is seen. These line portions overlap such that the spectral line is formed by mutually overlapping portions.

The spectral lines only arise as astigmatic focal lines when the axes of the two curved surfaces of the Fery prism intersect the principal section of the prism at points O' and O'' on the circle that passes through points A' and B' of the prism (Fig. 3.62). The curved surface passing through these two points and the internal concave side of the Fery prism coincide. Furthermore, the spectral lines can only arise along the previously mentioned circle, when the slit of the spectrograph is situated at any point on this circle such that the central line of the slit is perpendicular to the principal section of the prism. (For example, a ray of wavelength λ_x is brought to the point X on the circle when the slit is placed at point R.) Finally, it is necessary that the axis of the slit is exactly parallel to the refractive edge of the prism.

To prove the previous statements, let us assume that rays of wavelength λ, starting from point R on the circle, fall on the $A'B'$ refractive surface of the Fery prism. Rays RA' and RB' form identical angles of incidence with the incidence normals $O'A'$ and $O'B'$. Thus, according to Eq. (2.44), $n_{1,2} = \sin\alpha/\sin\beta = \text{const.}$, they will also form identical angles of refraction.

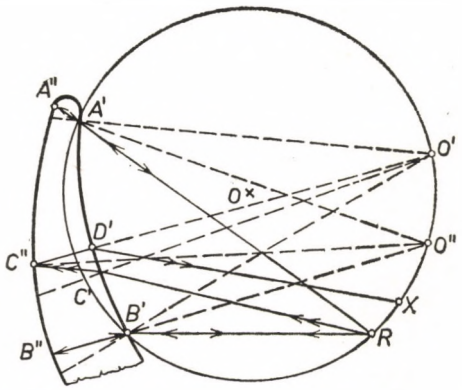


Fig. 3.62 Ray paths in the principal section of the Fery prism

In principle, there could exist a ray with a wavelength whose deviated directions $A'A''$ and $B'B''$ within the prism coincide with the $O''A''$ and $O''B''$ incidence normals of the reflecting prism surface $A''B''$. These rays are reflected back along the same route to points A' and B' on the front part of the prism and thus leave the latter along $A'R$ and $B'R$. The spectral line corresponding to a ray of wavelength λ_1 therefore appears at point R.

In contrast to the ray of wavelength λ_1, the directions of incidence and emergence of a ray of wavelength λ_2 coming from point R are not the same. This occurs because the prism refracts the ray having a direction RC' in such a way that the refracted ray reaches the mirror surface of the prism $A'B''$ at point C'' instead of continuing along the incidence normal. It will then be reflected from this point such that the angles of incidence and reflection, relative to the incidence normal $O''C''$, are identical. This reflected ray, therefore, emerges at point D' on the front part of the

prism. The emerging ray makes an angle with respect to the normal, $O'D'$, which is equal to the angle between the ray RC' and the incidence normal $O'C'$. Thus the ray of wavelength λ_2 intersects the circle that passes through the points R, B', and A', at X. Since C' is an arbitrary point on the front part of the prism, it follows that rays of wavelength λ_2 falling on the whole refracting surface of the prism meet at X and form a spectral line.

It is evident that the facts concerning Fery prism dispersion are the same as those given for concave grating dispersion (Section 3.8). With a concave grating the rays coming from the slit situated on the Rowland circle are united by astigmatism into spectral lines on the same circle provided the concave grating is also located on the Rowland circle. From the practical point of view this comparison is important in that the astigmatic origin of the spectral lines produces the same sort of disadvantageous effects.

Astigmatic line formation is caused by the fact that the height of the spectral lines cannot be limited by a slit diaphragm appropriate for Fery spectrographs. The line height can only be controlled by a diaphragm situated immediately in front of the spectral plate, with edges at right angles to the spectral lines. Different spectra can be photographed below one another simply by lowering or raising the variable width of the diaphragm. A further consequence of the astigmatic line formation is that the intensity of the spectral lines cannot be attenuated by a stepped filter or rotating step sector placed immediately before the slit of the spectrograph.

The advantages of Fery spectrographs can be summarised as follows:

1) There is only a single refracting or reflecting optical element. As this is relatively thin, light losses are very small. The Fery spectrograph is, therefore, comparatively fast.
(2) In this spectrograph type the harmful effect of stray light is reduced to the absolute minimum. The reflection background is therefore practically negligible.
(3) The spectral lines are very sharp.
(4) The linear dispersion of the equipment is comparatively high. (For example, in the case of an instrument where the radius of curvature of the concave surface of the Fery prism is 3600 mm, the value of the reciprocal dispersion at 2200 Å is 1·0 Å/mm, while at 3200 Å it is 5·1 Å/mm. It therefore reaches the performance level of four-prism quartz spectrographs.)

In summary it can be stated that Fery spectrographs are most useful for the special study of the region between 2000 and 2600 Å, for poorly illuminated spectra. In this spectral region, light absorption by quartz is already significant.

References

1. Fery, O. K., France Licence 10 330 (1910)
2. Mathien, F. C., *Spectrochim. Acta*, **5**, 174 (1952–53)

3.6.9 Spectroscopes

The term *spectroscope* refers to those dispersion devices where the spectra are studied directly by eye. As prism dispersion is almost exclusively used instead of grating dispersion for analytical purposes, only prism spectroscopes will be dealt with. The

optical construction of these instruments is different from that of prism spectrographs in so far as the photographic camera is replaced by an eyepiece and a telescope set to infinity by means of the camera lens. The main components of a spectroscope are the slit, collimating lens, prism, telescope lens, and eyepiece.

The principal use of spectroscopes is that from the visible spectra obtained the presence of certain elements can be established directly and their intensity estimated. The identification of the spectral lines can be made by comparing the examined spectrum with a standard one, or by the determination of the wavelength of the spectral lines. Spectroscopes used for this purpose are, therefore, usually fitted with auxiliary instruments that facilitate this work.

For the comparison of spectra a deviating prism (or sometimes mirror) is normally placed in front of the spectroscope slit which projects light from the standard spectrum onto half of the slit. In this scheme the rays that enter the spectroscope through one half of the slit are parallel to the optical axis, while through the other half the rays are perpendicular to the optical axis. Thus, in the field of vision the spectra of the two sources appear below one another. Lines of the same wavelength will coincide on the two spectra and appear as continuous lines in the total slit width.

The wavelength of the spectral lines can be simply established with an instrument having an appropriate optical system, that projects the wavelength calibration onto the eye of the observer so that the image of the scale appears at the same location as the spectrum image. In order to avoid unnecessarily complicated systems, the arrangement previously described is used only when the whole of the visible part of the spectrum can be observed at once. This arrangement can therefore be used only with low-performance spectroscopes.

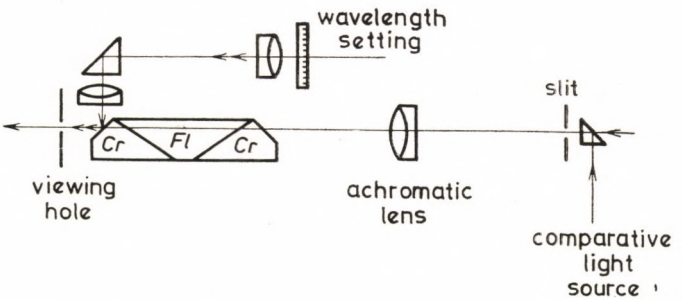


Fig. 3.63 Schematic optical diagram of a hand spectroscope (*Cr*: crown glass; *Fl*: flint glass)

The low-performance hand spectroscopes (Fig. 3.63) are constructed according to the previous principles. In such equipment the rays passing through the adjustable, symmetrical slit fall on a chromatically corrected collimating lens. The parallel rays are then dispersed by a direct-vision (Amici) prism. The rays emerging in different directions from the prism form parallel beams of variable wavelength. These beams are seen as spectral lines by the eye, which is set at an infinite distance away. The wavelength of the lines can be established by auxiliary optical instruments placed beside the direct-vision prism. The latter projects the rays, which pass through the transparent wavelength scale, onto the front part of the Amici prism, such that the eye sees the sharp image of the scale together with the spectrum.

By using hand spectroscopes the whole of the visible spectrum can be observed *at the same time*. The wavelength of the spectral lines can generally be determined with an accuracy of about ± 10 Å, provided the projected wavelength scale has previously been calibrated by means of a setting screw and a spectral line of known wavelength, for example, the sodium D-doublet.

With medium- and high-performance spectroscopes only sections of the visible part of the spectrum can be observed at any one time. Usually, the dispersion element of these instruments is the 90°, constant-deviation Abbé prism. By slowly rotating this prism the visible part of the spectrum can be continuously scanned. If the setting screw is combined with a reading drum having a helical wavelength scale then the determination of the wavelength of the lines also becomes possible. The optical scheme of this type of spectroscope is shown in Fig. 3.64. In order to reduce the dimensions of the equipment the light path is bent twice by means of a mirror and a pentagonal prism. The pentagonal prism has two reflecting surfaces and ensures that the wavelength of the spectral lines increases towards the right-hand side of the eyepiece. If a simple, right-angled reflecting prism were used instead of the pentagonal prism, then the direction of dispersion would be in the opposite sense and would not conform to internationally accepted usage.

The resolving power of the spectroscopes described above does not reach the level necessary for the analysis of the complicated spectra of steel; in this case spectroscopes with a high performance are needed.

High-performance spectroscopes usually have two forms of construction. They both have two prisms, although one type is autocollimating and the other is not. Since, from the optical point of view, there is no significant difference between a spectrograph and a spectroscope with the same dispersion elements, this group of spectroscopes will be dealt with here only briefly.

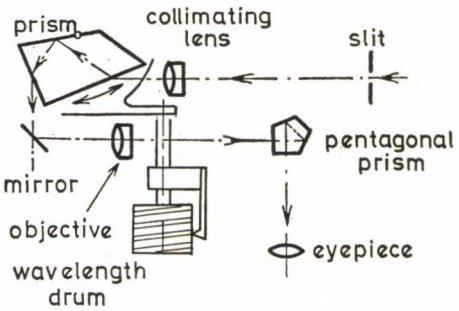


Fig. 3.64 Scheme of the Amici-prism spectroscope (TB 2, FOK-GYEM, Budapest)

The most widespread of the high-performance spectroscopes used has a combined 60° prism and a Littrow prism. Owing to the autocollimation principle this system is equivalent to the performance of a three-prism system and has a very high resolving power. The scanning of the spectrum is always made by rotating the Littrow prism (see, for example, the Polish-made SLP2 and the USSR model SL 11).

The non-autocollimated, two-prism system tends to be used less frequently (for example, the Hilger Steeloscope and the Fuess Metallspektroskop). With two 60° autocollimated prisms, very high resolution, corresponding to a four-prism system, can be achieved (for example, the Zeiss Industriespektroskop).

The resolving power achieved by the various systems discussed meets all the practical demands encountered in spectral analysis by spectroscopes. An increase in the number of prisms would only lead to unnecessary complications, reduce the speed of the spectroscope, and increase the effects of stray light.

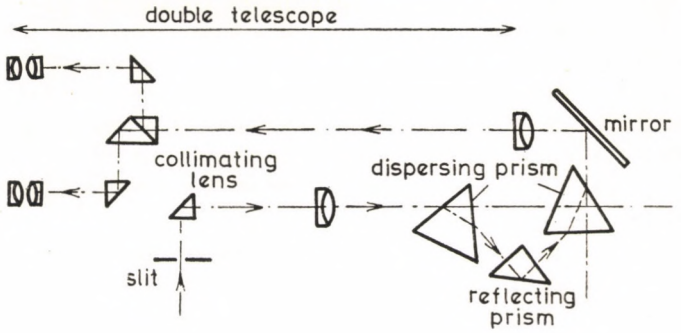


Fig. 3.65 Optical construction of a binocular (double-eyepiece) spectroscope (TB 3, FOK 3, GYEM)

All the equipment mentioned above are without exception monocular spectroscopes. Modern requirements, however, have demanded the development of a binocular spectroscope. This is especially advantageous if, along the intensity measurements of the lines, a combined mass analysis can also be determined. Binocular viewing greatly facilitates the observation of the spectrum, reduces eye fatigue, and increases the accuracy and dependability of the analysis.

Figure 3.65 illustrates the optical construction of the two-prism binocular spectroscope. For dispersing the light two fixed prisms are provided. Between the two prisms a reflection prism is situated and after the second dispersion prism is a plane mirror. The numerous bends in the optical path reduce the dimensions of the instrument. The scanning of the spectrum is effected by rotating a mirror connected to the helical wavelength scale.

The most important part of the instrument is the double telescope. The converging beam leaving the objective of the double telescope is symmetrically split by a composite prism having a semi-transparent, partially reflecting surface. With the aid of the two fully reflecting prisms, the true image of the two identical spectral parts is formed on the front of the two eyepieces. These images are seen by the eyes through the eyepieces as magnified, complete images. The distance between the eyepieces can of course be set to the distance between the eyes of the individual user. The angle at which the eyepieces lie with respect to the horizontal can also be simply adjusted. It is odd that while in the field of microscopy binocular vision is widespread, monocular spectroscopes, which cause a lot of eye trouble, are still in favour.

Those spectroscopes that measure subjectively the intensity of the analysis lines or the intensity ratio of the analysis line pairs are called *intensity measuring spectroscopes.* In addition to the optical elements mentioned so far, these spectroscopes also contain intensity-measuring devices. They will be discussed in detail in Section 5.13 which deal with the intensity measurement of spectral lines.

Usually, the two-prism, autocollimated spectroscopes are used for visual analytical spectroscopy, although in simple cases the Abbé prism spectroscope is adequate. The use of these instruments can be extended by intensity-measuring elements.

3.7 Illumination of prism spectrographs

3.7.1 General remarks

Illumination is the term given for the process whereby the light coming from the source is led into the spectrograph. In the illumination process the light available at the source and the resolving power and speed of the spectrograph can only be fully utilised provided the extreme rays of the beam, passing through the slit, form an angle that corresponds to the *F*-number of the collimating lens. When the angle is greater, a part of the light is lost, while if it is smaller the prism is not fully illuminated (Fig. 3.66). In the illumination of spectrographs it is necessary that the collimating lens of the prism and the stigmatic line-formation devices are just fully illuminated.

In addition to the above requirement in quantitative analysis, it is essential that the intensity of the spectral lines is uniform along their full length. Furthermore, it is necessary in qualitative analysis, as well as in certain quantitative problems,

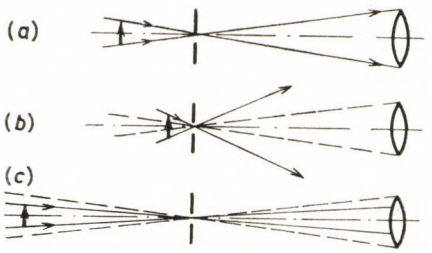


Fig. 3.66 Illumination of prism spectrographs: (*a*) correct and (*b*, *c*) faulty

that only rays from a certain part of the light source are allowed to enter the spectrograph. The requirements concerning spectrographic illumination are therefore not always the same. Thus, in many cases different illumination methods should be used during spectrographic analysis.

It should also be added that the factors governing the illumination of spectrographs are only valid when the axis of the light source is parallel to the slit of tht spectrograph. If the two are mutually perpendicular then the diameter of the lig h source (electrode gap) and its height are obviously interchanged.

3.7.2 Direct illumination of spectrographs

In this simple illumination method the source is placed in front of the spectrograph slit so that its centre is on the continuation of the optic axis of the spectrograph. In this case the collimating lens, etc., is fully illuminated, as shown in Fig. 3.67. When considering the section *perpendicular to the slit*, the distance between the light source and the slit is given by formula

$$X_m = d\frac{f_{col}}{D_{col}} \tag{3.56}$$

where d is the diameter of the source and it is assumed that the slit width is negligible compared with the diameter of the lens, D_{col}, of focal length, f_{col}. (This assumption

is always valid in practice.) On the other hand, when the section passing through the slit axis is taken into account, it is clear that the collimating lens cannot be illuminated by rays coming from all those points situated in the cross-hatched area on Fig. 3.67*b*. In this case, however, the distance between the intersection point, *C*, and the slit is given by

$$X_p = h \frac{f_{col}}{D_{col}} \tag{3.57}$$

where h is the height of the slit. The height of the source is neglected in comparison with the diameter of the collimating lens. The requirements determined by the above two equations are only obtained in the exceptional case when the diameter, d, of the light source is equal to the slit height. Therefore, in general the spectrograph can only be illuminated directly in exceptional cases.

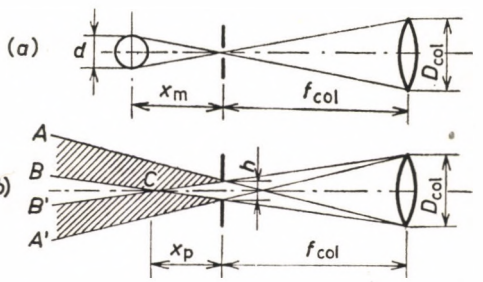


Fig. 3.67 Direct illumination of the slit: (*a*) perpendicular to the slit, (*b*) in a section parallel to the slit axis

From Fig. 3.67 it can be seen that the amount of light entering the spectrograph is not influenced by the fact that, within the area limited by BC and $B'C$, the perfectly homogeneous radiating source is lying close to or some distance from the slit. This can be explained by considering that, although the intensity of the light drops with the square of the distance, simultaneously the utilised area of the light source increases quadratically.

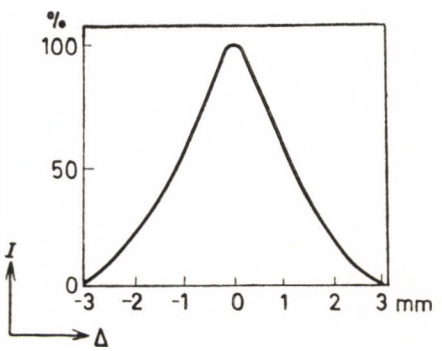


Fig. 3.68 Light intensity variation in a section perpendicular to the optical axis of the light source on direct illumination of the slit ($d = 3$ mm)

A further drawback of the direct illumination of spectrographs is that the illumination of the slit and thus the blackening of the spectral lines are not uniform along the length of the lines if the source radiation is not homogeneous. Furthermore, it can be established experimentally that the intensity of the light entering the spectrograph changes substantially if a light source with a homogeneous intensity distribution is not positioned exactly on the optical axis of the spectrograph (Fig. 3.68). However,

it is well known that neither in arc nor in spark excitation can the spatial and time variations of the plasma, or its random movement in general, be eliminated. Finally, therefore, it can be stated that in quantitative spectrochemical analysis the direct or complete illumination of the spectrograph is not desired.

3.7.3 Illumination of spectrographs by a collecting lens

Spectrographs are generally illuminated by a collecting or so-called *illumination lens*, instead of being directly illuminated. This has the advantage that without any sizable loss of light the source can be set further from the slit. Also, it is an advantage that the image of the source can be projected either onto the slit of the spectrograph or onto the collimating lens, as required.

In the first case, of focusing on the slit (sometimes called external focusing), the illumination lens projects the true image of the source onto the plane of the slit (Fig. 3.69). Thus, the rays passing through the different points of the slit at various heights are united into spectral lines. By using an appropriate slit diaphragm it is possible that the spectrum is formed not by rays from the whole area of the source, but only from a certain position or height. Thus, not only can any chosen part of the light source be examined, but it is possible to stop rays entering the spectrograph that come from unwanted parts of the light source; for example, the continuous spectrum of the incandescent electrodes.

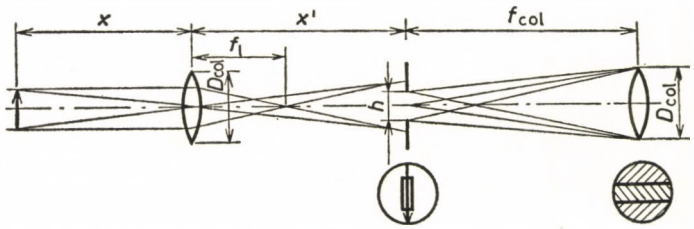


Fig. 3.69 Illumination of a spectrograph by external focusing

A further advantage of focusing on the slit is that the light completely illuminates the collimating lens. In its application, therefore, the resolving power and speed of the spectrograph are fully utilised. On the other hand, it is a drawback that this process does not homogeneously illuminate the slit, such that an even blackening of the total length of the spectral lines is not obtained. The conclusion is reached, therefore, that this method is not applicable to the purposes of quantitative analysis when a stigmatic spectrograph is used. On the other hand, as we shall see later, the longitudinal blackening of the spectral lines can be very homogeneous in astigmatic line formation, so that even the focusing on the slit method can be used in quantitative analysis.

An illumination lens of focal length f_1 projects the m-fold magnified image of the source on to the plane of the slit, if the distance between the slit and the illumination lens is given by

$$x' = f_1(m + 1) \tag{3.58}$$

The distance between the lens and the light source is x'/m.

An illumination lens of free diameter D_l, in the position previously specified illuminates a collimating lens of focal length f_{col} and free diameter D_{col} completely, if in neglecting the slit width the condition $D_l : x' = D_{col} : f_{col}$ is met. That is, if the free aperture of the illumination lens is given by

$$D_l = f_l(m+1)\frac{D_{col}}{f_{col}} \tag{3.59}$$

Another result is obtained if the height h of the slit is also taken into account. In this case, the prerequisite for the full illumination of the collimating lens is that the free diameter of the illumination lens be given by

$$D'_l = h + f_l \frac{(m+1)(D_{col}+h)}{f_{col}} \tag{3.60}$$

From Eqs (3.59) and (3.60,) it follows that the F-number of the illumination lens chosen should depend on that of the collimating lens. It is also evident that the magnification of the projected image and the height of the slit influence the optimum value of the free diameter of the illumination lens. The height of the slit is so small in practice that it is not necessary to take it into account. If the image of the source is formed on the slit, then at large slit-height values a part of the light does not reach the collimating lens.

In the opposite case of 'internal focusing', the image of the source is formed by the illumination lens in the plane of the collimating lens. This can be achieved by placing an illumination lens of appropriate focal length in the neighbourhood of the slit, while the source is positioned at a suitable distance from the latter (Fig. 3.70). Thus the illumination of the slit is homogeneous, which in principle should produce homogeneous blackening of the spectral lines along their full length. The method of internal focusing is therefore mainly used in quantitative spectral analysis.

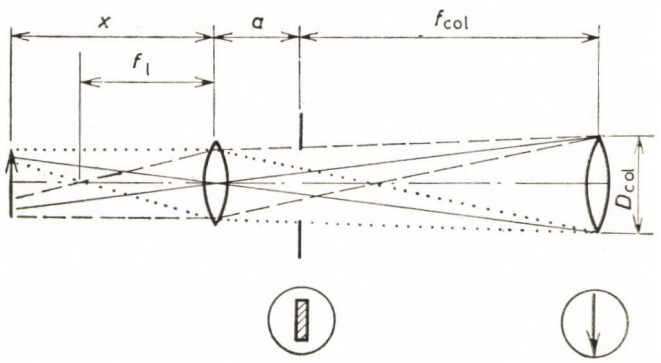


Fig. 3.70 Illumination of spectrograph by internal focusing

It follows from the above discussion that in the case of internal focusing the focal length, diameter, and position of the illumination lens should be chosen in such a way that the true image of the source appears in the plane of the collimating lens and also that the collimating lens should be just fully illuminated. The dimensions of the illumination lens should be chosen with regard to the F-number of the collimat-

ing lens in this case as well; that is, only an illumination lens designed for the spectrograph in question can be used.

The distance x between the light source and the illumination lens and the distance a between the lens and the slit both depend in internal focusing on the focal lengths f_1 and f_{col} of the illumination and collimating lenses, respectively. The approximate position for the light source can be calculated by means of the following formula

$$\frac{1}{f_{col}+a}+\frac{1}{x}=\frac{1}{f_1} \tag{3.61}$$

In modern spectrographs the illumination lens is frequently placed immediately in front of the slit, or within the slit head (Fig. 3.71). In this case, $a = 0$ and the distance between the source and the slit head is given by

$$x_{a=0}=\frac{f_{col}f_1}{f_{col}-f_1} \tag{3.62}$$

The maximum speed is obtained if the height of the source, that is, the distance between the electrodes, is

$$l_{a=0}=D_{col}\frac{f_1}{f_{col}-f_1} \tag{3.63}$$

Finally, with internal focusing the image of the source should, strictly speaking, be formed on the principal plane of the collimating–camera lens system, instead of on the plane of the collimating lens alone. However, since the distance between these two lenses can be neglected when compared with the focal length of the collimating lens, the position of the image is usually given simply as the plane of the collimating lens. The above-mentioned principal plane is situated between the two lenses.

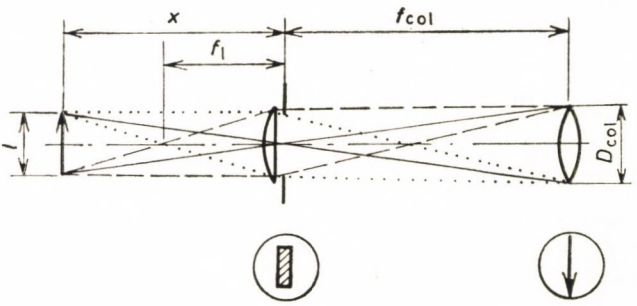


Fig. 3.71 Internal illumination of spectrograph by a lens positioned immediately in front of the slit

The illumination lenses generally used are in fact simple spherical lenses. They are not of composite type in the case of quartz spectrographs, although recently chromatically corrected quartz–fluorite achromats have been used. On the other hand, for the illumination of glass spectrographs achromats are always used.

In some cases the illumination lens is cylindrical in shape and projects the rays from the light source in the form of perpendicular bands onto the slit of the spectro-

graph[1,2] (Fig. 3.72). It should, however, be noted that the projected image has a horizontal width, since the light source (arc or spark) also has a horizontal width. This type of focusing has the advantage that no illumination error occurs, even if the source is not situated exactly on the optical axis of the spectrograph. The illumination error will only be eliminated, however, when the geometrical axis of the cylindrical surface of the lens is exactly parallel to the slit of the spectrograph and if the

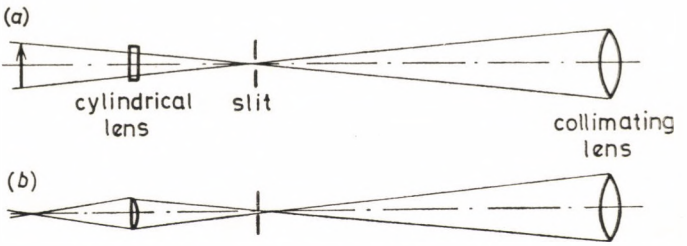


Fig. 3.72 Application of the cylindrical illumination lens: (*a*) in a section parallel to the slit, and (*b*) in a section perpendicular to the slit

lens was specifically designed for the spectrograph in question. Cylindrical lenses are not very commonly used in prism spectrographs, although they are frequently used for the illumination of grating spectrographs.

References

1. HANSEN, G., *Z. Phys.*, **29**, 356 (1924)
2. KINDER, W., *Z. InstrumKde*, **56**, 356 (1936)

3.7.4 Illumination of spectrographs with intermediate focusing

In Zeiss-system intermediate focusing illumination an illumination lens (I) of short focal length projects the magnified image of the source onto the intermediate diaphragm (Fig. 3.73).[1] The latter is a disc that is perpendicular to the optical axis and that can be rotated about a horizontal axis. There are slots of different size on the

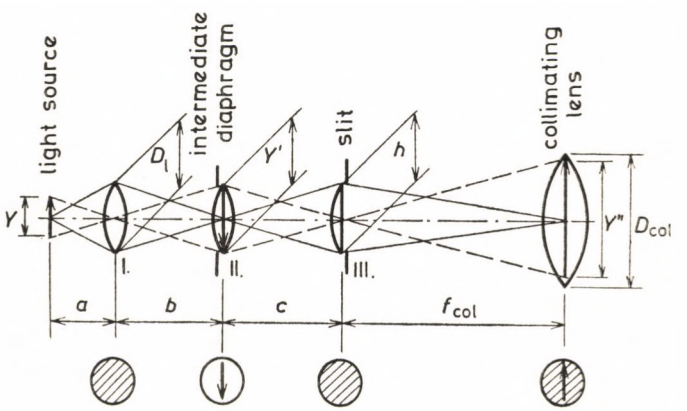


Fig. 3.73 Illumination of a spectrograph by the Zeiss-system intermediate focusing method[1]

disc. By the appropriate setting of the diaphragm it is possible to have either all the radiation of the source enter the spectrograph, or only the rays coming from a certain position. The intermediate image of the source produced is projected onto the collimating lens by a slit lens of long focal length situated immediately before the slit (III), such that in the plane of the collimating lens an image of the source framed by the intermediate diaphragm is developed. The long-focal-length lens (II), that is placed immediately behind the intermediate diaphragm, is used to project the image of the homogeneously illuminated lens (I) onto the slit of the spectrograph—as is shown by the dotted line on the figure. Lens (II) therefore secures the homogeneous illumination of the slit, while also greatly increasing the luminosity. The photographic speed developed in this way is the same as that of the internal focusing method, where the distance between the source and the slit is the same as the distance, *a*, between the source and the first illumination lens (I) used in intermediate focusing.

The advantages of the intermediate focusing method are summarised as follows:

(1) The source homogeneously illuminates the slit of the spectrograph.
(2) By means of the intermediate diaphragm, radiation coming from certain areas of the source can be stopped from entering the spectrograph.
(3) The radiation of the source can be evenly attenuated by a 'grey filter' fastened to the lens of the intermediate diaphragm.
(4) Intermediate focusing facilitates the alignment of the electrode gap onto the optical axis of the spectrograph.

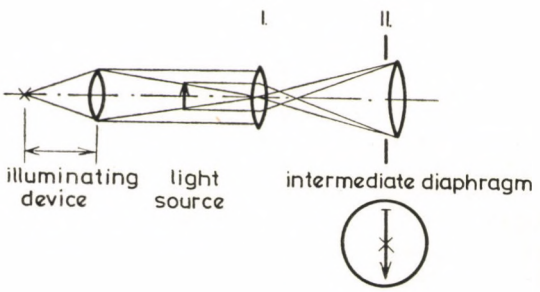


Fig. 3.74 Alignment of the light source by intermediate focusing along the optical axis

The intermediate focusing method combines the advantages of the external and internal focusing techniques but does not have their disadvantages. It also satisfactorily meets the requirements of quantitative spectrochemical analysis.

The adjustment of the light source can be achieved as shown in Fig. 3.74. A light illuminates the area of the electrode gap from behind. Under such conditions, lens (I) of the intermediate focusing illumination device projects the shadow of the electrode gap onto the intermediate diaphragm. The position of the source is satisfactory when the centre of the shadow coincides with the sign that indicates the position of the optical axis on the intermediate diaphragm.

The distances between the light source, the lenses, and the slit are only constant when the lenses are sufficiently corrected for chromatic aberration. This requirement can be relatively easily achieved when glass optics are used, although difficulties are encountered with quartz spectrographs. Accordingly, simple quartz lenses are fre-

quently used for the purposes of intermediate focusing. When non-achromatic quartz optics are used, the distances between the lenses, etc., should be made such that the focusing is sharp for the medium wavelengths of the spectral range examined.

When illumination lenses that are not corrected for chromatic aberration are used, one should take into consideration the error when the electrically excited light source does not fully illuminate the collimating lens, even though the previously given requirements are met.[2] When the F-number of the illumination lens (I) and the magnification, Y''/Y, of the illumination system are considered, then it can be seen from Fig. 3.73 that the necessary condition for the total illumination of the collimating lens is

$$\frac{D_1}{a} = \frac{Y''}{Y}\,\frac{h}{f_{\text{col}}} \tag{3.64}$$

As is well known, the random motion of the burning point of the spark or arc cannot be avoided in electrical excitation. Thus variation of the distance a and the F-number of the first illumination lens (I) is also impossible to avoid. This means that one should take into account the variation of the illumination flux of the collimating lens. This, in turn, has the consequence that the intensity of the spectral lines is not even along their full length. This error is, however, only significant for those spectral lines whose wavelength is considerably different from that used to adjust the distances between the lenses. According to the above relationship this anomalous behaviour is lessened by a shorter focal length of illumination lens (I). The use of a lens of longer focal length is not an advantage, since the luminosity is reduced.

In the illumination method described the positions of the lenses should be changed when the spectral range is altered. In the method of 'achromatic intermediate focusing' this drawback is eliminated.[2] In this illumination process the first illumination lens [lens (I) in Fig. 3.73] is corrected for ultra-violet chromatic aberration, while the other lenses are of simple form. Thus, the intensity of the spectral lines does not change along their length even for high slit values. The illumination of the collimating lens is only complete when the requirements given by Eq. (3.64) are met. In principle, therefore, the height of the light source (the electrode gap) can only be increased when the achromatic lens is simultaneously replaced by another lens of correspondingly longer focal length and the distance between the lens and the light source is appropriately increased. The latter requirement is easily obtained for usual electrode distances by means of two achromatic lenses.

In intermediate focusing the illumination can be increased by about 75% if a concave metal mirror is placed behind the source[3] (Fig. 3.75). The combination of

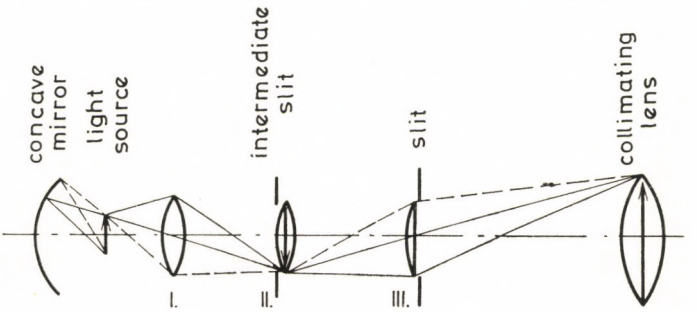


Fig. 3.75 Application of a concave mirror in intermediate focusing illumination

the intermediate focusing method and a concave mirror is especially useful when the spectrograph is illuminated by radiation coming from a small portion of a weakly emitting source.

In principle, the collimating lens can be illuminated to any degree, independent of the electrode distance, if the lens placed behind the intermediate diaphragm is removed; that is, if only two lenses are used for intermediate focusing.[4] In this case, however, it is necessary to alter the distance a between the source and the first illuminating lens, and the distance $(b + c)$ between the lens and the slit (Fig. 3.76). For this system an illumination lens of comparatively large F-number is necessary.[2]

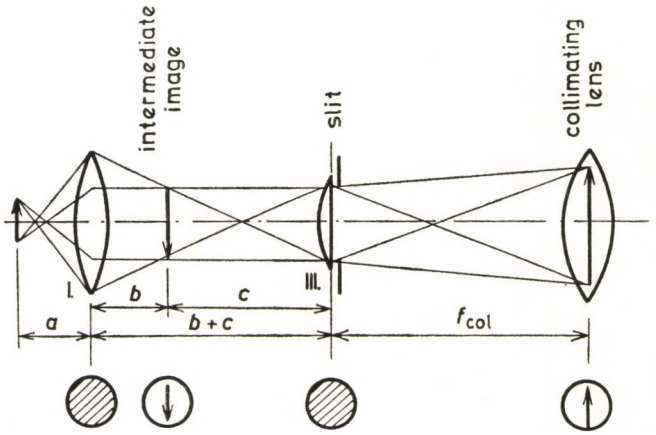


Fig. 3.76 Illumination using intermediate focusing via the Twyman-system double-lens arrangement[4]

The error arising from geometrical changes of the light source can be completely eliminated, if in the previously described two-lens intermediate focusing method a multiple-lens condenser is added to the first illumination lens[3,5] (Fig. 3.77). The multiple-lens condenser is an optical element with a relatively low precision level.

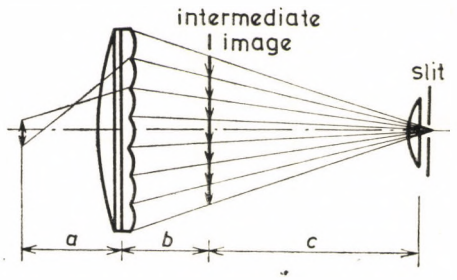


Fig. 3.77 Horizontal section of the Preuss-system lens-raster illumination arrangement[5]

One of its surfaces is planar and the other is composed of a number of smaller lens surfaces with a definite radius of curvature. These smaller lenses are packed in a regular square lattice pattern. The lens-raster and the adjoining biconcave lens can, therefore, be regarded as an array of small lenses placed adjacent to one another, such that the focal length is the resultant of the focal lengths of the large and the small lenses.

In the lens-raster focusing method the intermediate image is formed not by a single projection of the source, but as a result of many small images formed in the same plane. Since the slit lens forms the intermediate image on the plane of the collimating lens, a great number of such images will appear on the collimating lens in this case. Two advantages of this system are immediately obvious. First of all, the unavoidable random motion of the source during electrical excitation methods does not cause any trouble. Secondly, as the whole surface of the collimating lens is used the practical resolving power of the spectrograph is also fully utilised.

References

1. KAISER, H., *Spectrochim. Acta*, **3**, 529 (1949)
2. MIELENZ, K. D., *Spectrochim. Acta*, **10**, 99 (1959)
3. NORDMEYER, M., *Spectrochim. Acta*, **7**, 128 (1955)
4. TWYMAN, F., *Metal Spectroscopy*, London, 80 (1951)
5. PREUSS, E., *Heidelb. Beitr. Miner. Petrogr.*, **4**, 163 (1954)

3.7.5 Illumination of spectrographs by a concave metal mirror

A concave metal mirror can also be used in the illumination of the spectrograph, as well as a collecting lens. The advantage of a mirror is that the focus and thus the paths of the reflected rays are independent of the wavelength. When the wavelength range is changed the position of the illumination mirror is therefore not altered. On the other hand, it is definitely a disadvantage that the application of the metal mirror is frequently made difficult for geometrical reasons; for example, the structures of the various auxiliary devices get in the way, etc. This is the main reason why this method has not gained general acceptance.

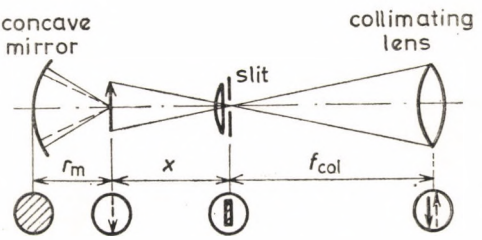


Fig. 3.78 Internal focusing illumination of a spectrograph by means of a lens and concave mirror combination

When illuminating mirrors are used, it is not generally advisable to substitute them for the illumination lens placed immediately in front of the slit of the spectrograph. The chromatic aberration of this lens can generally be entirely neglected.[1]

In one method that uses a combined concave metal mirror and illumination lens placed in front of the slit, the mirror concentrates the image of the light source within the source itself. The optical axis of the mirror is the continuation of the optical axis of the spectrograph (Fig. 3.78). In this arrangement the concave mirror operates *only* as a reflector. By its application the luminosity of the source is practically doubled. The error caused by the erratic movement of the plasma in arc and spark excitation, with respect to the optical axis of the spectrograph, is mainly eliminated.[1] The fluctuations of the rays reflected from the mirror are of the opposite sense to those coming from the source directly.

The method previously discussed is in effect an internal focusing and illumination of the spectrograph, or, rather, it could be used for this purpose. Its advantages are the same as those of the method described previously. The source homogeneously illuminates the slit of the spectrograph in both methods. The drawback of the latter method is that self-absorption is increased in spectra obtained in this way.

In another method, using the combination of the concave mirror and the lens, the concave mirror combines the light of the source into an intermediate image that is projected onto the illumination lens placed in front of the spectrograph slit (Fig. 3.79). This method is, therefore, a modification of the 'two-lens' illumination method described in the previous section.[2] There is, however, a significant difference in that when a concave illumination mirror is used the source cannot be positioned on the optic axis of the spectrograph. The source must be placed in such a way that it does not obstruct the reflected radiation. Accordingly, the optical axis of the concave mirror is not a continuation of the optical axis of the spectrograph, but forms a sharp angle with the latter. The intermediate diaphragm should be placed such that a 2–3 times magnified image of the source (or a particular part of it) appears in the opening. This occurs when the source is situated between the value of the focal length f_t and twice the focal length value, that is when $f_t < a < 2f_t$. The illumination system of the spectrograph is properly aligned only when the ray coming from the upper (or lower) edge of the mirror and touching the upper (or lower) edge of the intermediate diaphragm, just touches the lower (or upper) point of the collimating lens of the spectrograph. The value of the opening of the concave mirror should be chosen not only with respect to the *F*-number of the collimating lens but also with regard to the maximum electrode distance and the maximum magnification of the intermediate image. It is an advantage, from a practical point of view, that when the magnification is varied only the position of the light source and the mirror should be changed; that is, only the distances b and a.

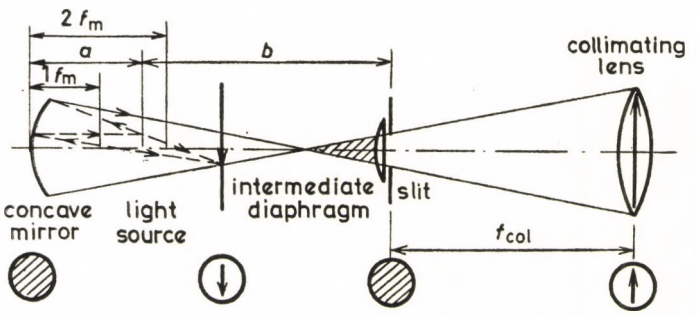


Fig. 3.79 Illumination of a spectrograph by the combination of concave mirror and lens for intermediate focusing of the light source (perpendicular section of the optical arrangement)[1]

It can be experimentally shown that under the above conditions the beam of light entering the spectrograph is already homogeneous before the slit is reached – at the positions indicated by the cross-hatching on the above figure. Thus, the intensity of the spectral line is completely even along its full length. A small but generally negligible error is caused by the fact that the source image that appears in the plane

of the intermediate diaphragm is not quite symmetrical because of the position of the source.[1]

Compared with the previously described illumination method, a more complicated and also less advantageous technique uses so-called Z-position mirrors for illumination purposes[3] (Fig. 3.80). This process is principally a modification of the illumination system based on the Zeiss-system intermediate image formation. The only

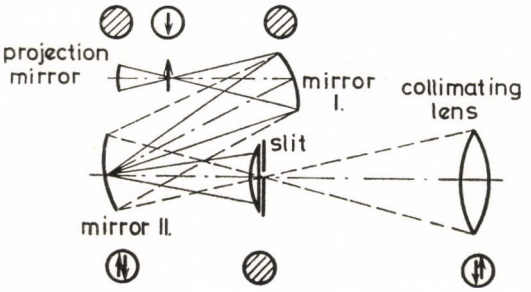


Fig. 3.80 Illumination of a spectrograph by concave mirrors in the Z-position

advantage of the method is that the Z-position of the mirrors reduces the asymmetry error. On the other hand, this insignificant error can only be eliminated if at least three or more mirrors are used. Furthermore, the illumination of the slit is not homogeneous when the reflectivity of the mirror surface is not quite perfect, because of some contamination or local fault, etc.[1] Thus, this method is not recommended for spectroscopic purposes.

References

1. NORDMEYER, M., *Spectrochim. Acta*, **7**, 128 (1955)
2. TWYMAN, F., *Metal Spectroscopy*, London, 80 (1951)
3. CZERNY, H. and TURNER, A. F., *Z. Phys.*, **61**, 192 (1930)

3.7.6 Time-resolved illumination of spectrographs

In spark or a.c. arc excitation the spectrograph is illuminated by a series of periodically recurring light phenomena. In certain cases it is an advantage if only light emerging in a certain phase of excitation is allowed into the instrument. This can be achieved by illuminating the spectrograph indirectly by means of a device that evenly disrupts the periodic light pulses, such that a form of time resolution is developed.

During excitation light pulses are repeated every 0·01 s at a mains frequency of 50 Hz. In low- and high-voltage spark excitation, however, the burning time of a partial discharge (T), apart from strongly damped discharges, is generally of the order of 10^{-5} s, while in a.c. arc excitation it is about 10^{-3} s. With spark excitation, therefore, time-resolving equipment should be used that allows the spectrograph to be illuminated in each half-period for about 10^{-5} or 10^{-6} s, while in the case of a.c. arc excitation a time resolution of about 10^{-4} s is enough. The first requirement can be met most simply by using either a rotating disc with a slot in it or a rotating mirror. In a.c. arc excitation it is sufficient if the oscillating mirror method is used.

In the application of the rotating-disc technique, a synchronous motor drives a disc that is placed immediately before the spectrograph slit, and that has a slot of

appropriate size and form along its diameter.[1-4] The duration of the illumination of the spectrograph, Δt_{disc}, under these conditions depends on the value of arc, α, corresponding to the slit width and the speed of revolution, U, that is

$$\Delta t_{\text{disc}} = \frac{\text{arc}\,\alpha}{U} \tag{3.65}$$

while the mean slit width, b, according to Fig. 3.81, is

$$b = r\,\text{arc}\,\alpha \tag{3.66}$$

where r is the radius of the slot.

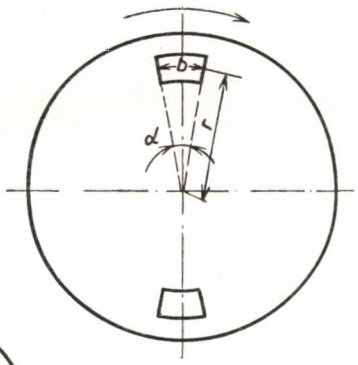
It becomes evident from the above data that by means of the rotating disc a time resolution greater than 10 μs cannot be achieved. There is a limit set to a further decrease of the time duration by the fact that a slot width of less than about 0·1 mm cannot be used, not only for technical reasons, but mainly because of diffraction phenomena. The height of the disc slot is determined by the length of spectral line required. By using a number of slots of different mean radius, r, the lines produced at different heights of the spectrum will characterise the spectral composition of the radiation emerging at different phase instants of the condenser discharge. When, however, the mean radii of the slots are very different, one should consider the possible error that can occur because the resolution does not vary linearly along the lenght of the spectral lines owing to variation of the peripheral velocity.[4] (Deviation due to this source for large slit heights, for example if $r = 40$ mm mean radius, is about 1%, while if $r = 80$ mm it is about 0·5%.[4]) From this it is seen that by using rotating discs with different slots very different degrees of time resolution can be realised very simply (Fig. 3.82).

In general, the prerequisite for the production of time-resolved spectra is that the spectrograph is a stigmatic instrument, with a spark generator controlled by an electronic unit. The third requirement is that the examined phase of the spark discharge and the instant of illumination of the spectrograph should be exactly synchronised. In the opposite case, the emission arising from repeated condenser discharges at the same phase instant would not be brought to the same point of the slit.

In the application of the rotating disc the timing of the spark discharge can be precisely synchronised with the illumination of the spectrograph if the rotating disc controlling the illumination simultaneously interrupts a light beam from an auxiliary light source (Fig. 3.83). The latter source illuminates a photocell whose photocurrent is input to a trigger generator. The voltage drop obtained controls the spark generator.[4] Under such conditions the signal, N, from the rotating disc illuminated by the spark discharges is *apparently* at standstill. When the time interval between the starting trigger signal of the photocell and the initiation of the condenser discharge is electronically altered, the apparent location of signal N also changes. It can be definitely established by observing which B-setting is the phase instant of the condenser discharge at which the illumination of the spectrographs commences.

The great advantage of the described rotating disc method is that it only requires a comparatively simple device. On the other hand, the time resolution effected by rotating mirrors is rather complicated.[4-6] In this latter system, the spark discharge illuminates the spectrograph by means of a concave mirror placed on the optical axis of the spectrograph. This is driven by a synchronous motor whose axis of

Fig. 3.81 Fundamental principle of time resolution by using a rotating disc



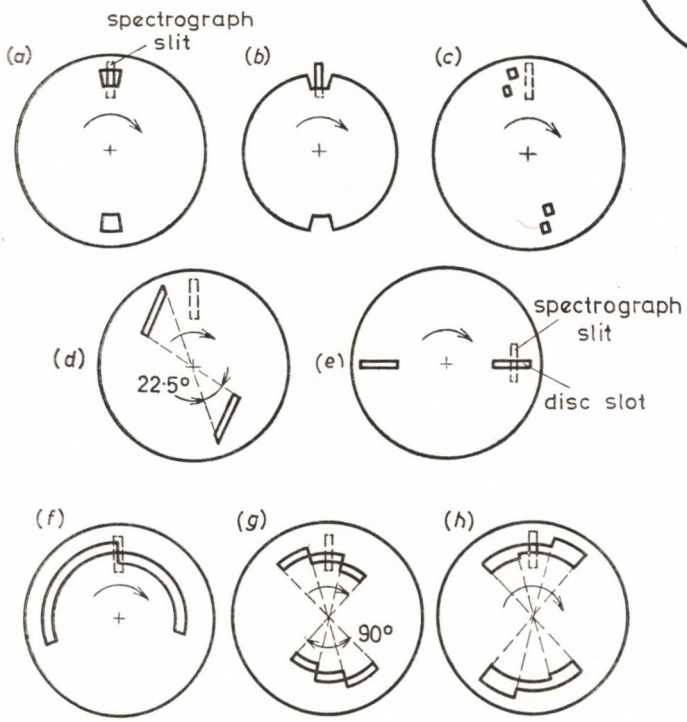


Fig. 3.82 Rotating discs used for different types of time resolution: (*a*) starting at a certain phase instant, (*b*) time resolved and non-resolved, (*c*) spectrum arising from two time ranges, (*d*) and (*e*) for continuous time resolution, (*f*) strongly damped discharge of type 0 + 1, (*g*) and (*h*) disc used for the study of a.c. arc excitation[3]

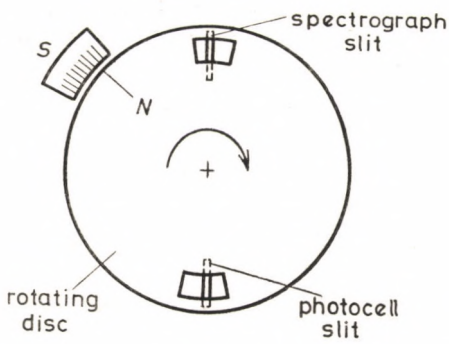


Fig. 3.83 Setting of the phase position of a rotating disc

rotation is perpendicular to the optical axis of the instrument (Fig. 3.84). The extent of time resolution is determined by ω, the angular frequency of the rotating mirror. the distance r between the mirror and the spectrograph, and by the slit width (or height), h. The time duration for a single illumination of the spectrograph is given by the formula

$$\Delta t_{\text{mirror}} = \frac{h}{\omega r} \tag{3.67}$$

The angular frequency of the rotating mirror is, however, determined by the frequency of the mains. For example, in the case of a 50 Hz mains current the former equation is modified as follows

$$\Delta t_{50} = 0{\cdot}02 \frac{h}{2\pi r} \tag{3.67a}$$

The best results are therefore obtained if the distance r is large; that is, when the focal length of the rotating mirror is large. This factor is limited by the fact that the illumination of the spectrograph is complete only when the aperture angles of the concave mirror and of the collimating lens of the spectrograph are the same, that is, when: $D_{\text{col}}/f_{\text{col}} = D_t/r$. Excessive lengthening of the distance r is disadvantageous, because the intensity of light entering the spectrograph drops quadratically with increasing distance and the use of a large rotating mirror produces technical difficulties. At about 1 μs time resolution can be easily attained. (In this case, from the above equation, $r = 2.10^4\, h$, and thus for a slit width of $h = 0{\cdot}025$ mm, $r =$ 500 mm. In the case of a collimating lens with a relative aperture of 1 : 10, it is sufficient if the free diameter of the concave mirror is 50 mm.)

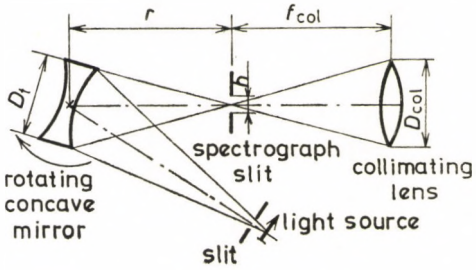


Fig. 3.84 Diagram of time resolution by a rotating disc in the perpendicular section

The direct illumination of the spectrograph by means of a rotating mirror can be achieved in two ways. In one method, a concave mirror with an axis of rotation perpendicular to the slit forms the image of the source in the plane of the slit. If in such an arrangement a slit is placed in front of the source with a slot perpendicular to the optical axis of the source and also to the slit of the spectrograph, then the source image framed by the slit is scanned along the slit of the spectrograph during the discharge of the condenser (Fig. 3.84). Therefore, when a slit of appropriate height is used, a spectrum is obtained whose lines are formed by continuous time resolution. Thus, bands at different heights in the spectrum are the spectra of the condenser discharge at different instants. On the other hand, if a horizontal slit is placed in front of the spectrograph slit then a spectrum is obtained that corresponds to a certain phase instant of the condenser discharge.

In the other application of the rotating concave mirror, the rotation axis of the mirror is parallel to the slit of the spectrograph (Fig. 3.85). During the condenser discharge the vertical image of the source produced has the same orientation as the intermediate slit parallel to the slit of the spectrograph. By adjusting the width of the intermediate slit it is possible that only a part of the time-resolved radiation of the condenser discharge illuminates the spectrograph for a predetermined duration. The horizontal slit placed in front of the source allows only rays coming from a certain height in the light source to enter into the spectrograph.

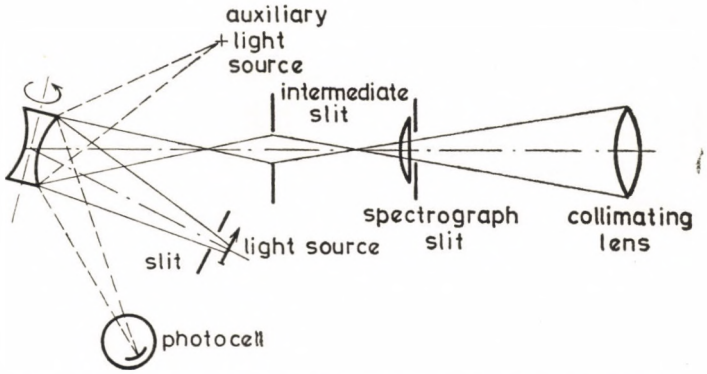


Fig. 3.85 Illumination of a spectrograph by rotating concave mirror with intermittent light dispersion (solid line) and production of the optical trigger signal for an electronically controlled spark generator (dotted line)

In the case of the rotating concave mirror method, the coupling between the electronically controlled spark generator and the time resolving equipment is also established by means of a photocell. Here, the light of the auxiliary source is projected at the right moment by the rotating concave mirror onto the photocell (Fig. 3.85). The resultant pulse-like photocurrent produces the necessary starting trigger signal for the condenser discharge.[5]

In a.c. arc generation, time resolution for a burning period of approximately 0·005 s can be achieved by using an oscillating mirror, provided its oscillation frequency is the same as the mains frequency of the arc generator.[7] In this arrangement the image of the narrow, horizontal slit, placed in front of the source, is formed by the oscillating mirror in the plane of the perpendicular spectrograph slit. Thus, the radiation produced during the half-period of the alternating current illuminates the slit in such a manner that during the subsequent half-periods the rays emerging at the same phase instant always fall on an identical part of the slit. In this manner a time resolution is brought about whose time axis has the same direction as that of the slit of the spectrograph.

In order to apply the above procedure, an electrical device is necessary to control the amplitude of the concave illuminating mirror (oscillating with the mains frequency) and also the variations of the phase-shift relative to the mains current (Fig. 3.86). In this device the concave mirror is attached to an iron plate that is caused to oscillate by a low-inertia coil placed in the magnetic field of a permanent

magnet. The value of the mains current and thus the elongations of the mirror can be controlled by the potentiometer P_1, while the relative phase of the oscillations is adjusted by means of the potentiometers P_2 and P_3.

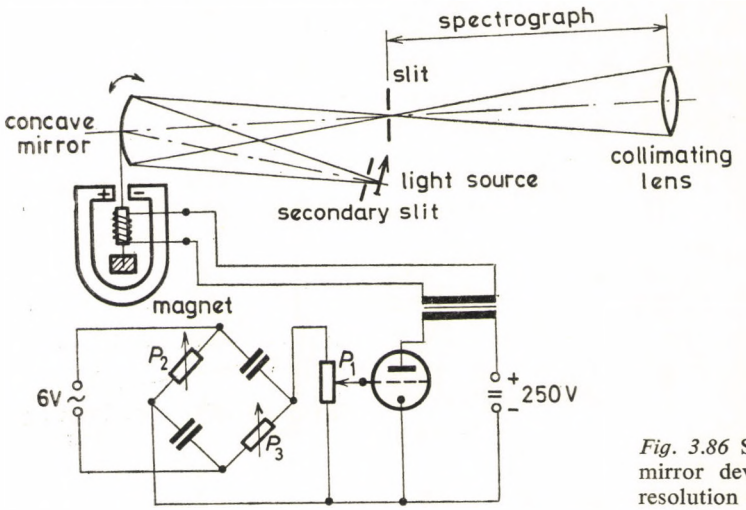


Fig. 3.86 Scheme of the vibrating mirror device used for the time resolution of arc spectra

References

1. DEINUM, H. E., and DE BOER, F., *Spectrochim. Acta*, **2**, 318 (1943)
2. VAN CALKER, J., *Naturwissenschaften*, **40**, 434 (1953)
3. GHERSEN, P., *J. Opt. Soc. Am.*, **48**, 712 (1958)
4. BARDÓCZ, Á., *Acta Tech. Hung.*, **40**, 179 (1962); *Spectrochim. Acta*, **16**, 1173 (1960)
5. BARDÓCZ, Á., and VARSÁNYI, F., *Z. Naturf.*, **10**, 1031 (1955); *Nature, Lond.*, **177**, 222 (1956)
6. BARDÓCZ, Á., *Z. angew. Phys.*, **9**, 82 (1957); *Appl. Spectrosc.*, **11**, 167 (1957)
7. VORSATZ, B., *Emissionsspektroskopie*, Berlin, 215 (1964)

3.7.7 Attenuation of the illumination of spectrographs

The purpose of simple radiation attenuators is to reduce the density of the light flux independently, if possible, of the wavelength. In practice, it frequently occurs that the light source has such a high intensity that without attenuation the photographic plate can only be exposed for a very short interval. The object of the more sophisticated stepped attenuators, however, is to divide the homogeneous illumination at the slit into a number of sections such that the ratio of the light flux passing through the different sections is a known quantity. When this form of attenuator is used with stigmatic-type spectrographs, spectra are obtained in which each line is composed of several portions having definite intensity ratios. (These attenuators are used in, e.g. quantitative analysis when the intensities of the compared spectral lines are very different and also in the determination of the intensity ratios of the lines.)

Light-attenuating devices are generally constructed according to different principles. For spectrographs either grey filters, rotating sectors, or diaphragms are used.

Grey filters are parallel-faced plates whose transmittance is, in principle, independent of the wavelength. In practice, however, perfectly 'grey' or 'neutral'

filters cannot be produced. Generally, therefore, the filters are neutral only in limited wavelength intervals.

For filters the extent of attenuation is determined by the value of the filter constant. This is the ratio of the intensity of light transmitted by filter (I) to the intensity of the incident light (I_0). In practice, percentage attenuation is used, that is, $I/I_0 . 100$. For commercial grey filters an approximate value is given on the frame and usually refers to white light. The difference between this value and the true percentage transmission measured at different wavelengths is usually negligible within the spectral range used in quantitative analysis.

The grey filters used in the ultra-violet spectral range are parallel-faced quartz plates, their surfaces vacuum-coated or cathode-sprayed with some metal of the right properties and thickness, such as platinum, palladium, rhodium, or sometimes antimony, stainless steel, etc.[1] The thickness of the layer determines the transmission properties of the filter. The platinum filters are used predominantly for the 7500–10 000 Å range, rhodium filters between 4000 and 7000 Å, palladium filters in the interval 3000–5000 Å, and antimony filters between 2000 and 4000 Å.[2, 3] The above-mentioned grey filters are not perfectly neutral in a wider spectral range. Filters whose transmittance is practically independent of the wavelength over a longer interval can only be made by vacuum-coating the surface with several different metal layers.[5] In the visible spectral range, however, attenuators made of grey glass are better than the filters made in the above way (Fig. 3.87).

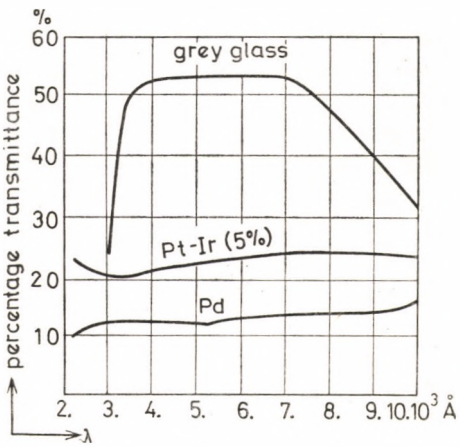


Fig. 3.87 Transmittance variation of commonly used grey filters as a function of wavelength: Schott, Jena N.G.5 grey glass 1 mm thick[1]; platinum grey filter containing 5% iridium and palladium grey filter[2,4]

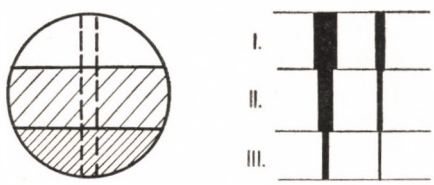


Fig. 3.88 Three-step filter

Stepped filters frequently used in quantitative analysis are quartz, or occasionally glass, plates fixed immediately before the slit of the spectrograph. They are vacuum-coated with layers of platinum or aluminium of variable thickness in the form of

stripes of equal width. If the stepped filter is fixed to the spectrograph so that the metallic stripes are perpendicular to the slit,[5] then in stigmatic line formation all lines of the spectrum are composed of several parts of different intensity (Fig. 3.88).

Table 3.7

Percentage step values of commonly used step filters

Filter type	I	II	III	IV	V	VI	VII
two-step	100	80					
	100	60					
three-step	100	50	10				
	100	20	4				
six-step	100	65	40	25	16	10	
	100	40	16	8·8	2.5	1	
'even'	100	45·5	21·5	9·8	4·65	2·2	1
'shortened'	100	21·5	4·65	1			

The percentage transmittance of common stepped filters is given in Table 3.7. Three-step filters are used when the analytical determinations are made using a quartz spectrograph. In the visible part of the spectrum it is not advisable to have a transmittance value lower than 20%. In the application of six-step filters, which are mainly used for the construction of the characteristic curve of the photographic emulsion, it is a disadvantage if the transmittance does not change evenly between the individual steps. This drawback is eliminated by the use of even-grade stepped filters, where the ratio of the transmittance value of the adjoining steps is practically constant; usually the value is about 2. In the application of these step filters, one disadvantage is that a relatively tall slit must be evenly illuminated. This problem can be removed by using a shortened step filter. The four steps of this filter cover the whole intensity working range of the photographic emulsion.[6,7]

In practice, it is advisable that the slit of the spectrograph be illuminated by a beam whose intensity flux is attenuated not stepwise but continuously. This can be achieved by means of a grey wedge. The better-quality grey wedges are parallel-faced quartz or glass plates upon which an appropriate metal layer is vacuum-coated in progressively increasing amounts so that the transmittance changes in a logarithmic manner.[16] According to the Lambert's law this can be secured if the thickness increases in a linear fashion.

If a grey wedge is used for light attenuation it is a possible source of error if it is contaminated. In this case, at different heights of the spectral lines the illumination does not correspond to the theoretical value of the filter constant.

A further source of error when grey filters are used is if the transmittance of the filter is not calibrated; that is, if the correlation between the wavelength and the transmittance is not established by a spectrophotometer or by some other method. This is especially important in the case of step filters, and is due to the fact that both the absolute value and the relative value of the transmittance of the individual steps vary with the wavelength (Fig. 3.89).

The above-mentioned error can be completely eliminated by a different method of light attenuation, where a rotating disc or rotating sector is used[8]. In this method a thin metal disc with a slot is rotated at an even speed immediately in front of the slit of the spectrograph. In this case there is no need to calibrate the disc, since the value of the filter constant can be calculated. The value is given by the ratio of the central angle α, belonging to the slot of the disc, relative to 360°; the percentage transmittance is therefore $\alpha/360.100$.

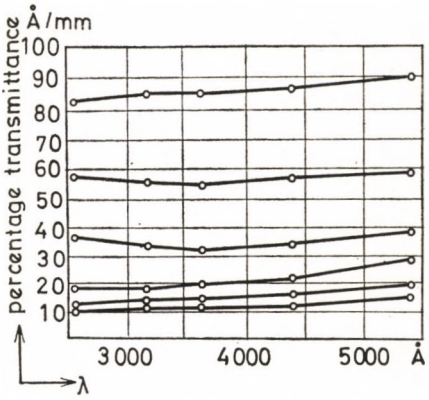


Fig. 3.89 Variation of relative transmittance of a six-step platinum filter as a function of wavelength

The common rotating discs used in light attenuation are schematically depicted in Fig. 3.90. By the use of the so-called two-disc rotating sector, radiation reaching the spectrograph can be reduced continuously. Both slots of the two discs have angles of 90°. The completely open disc, therefore, transmits half the light falling

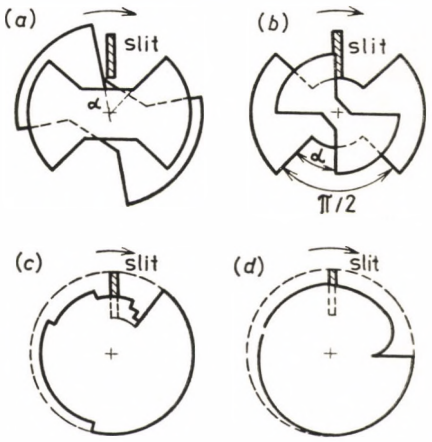


Fig. 3.90 Rotating discs: (*a*) two-disc type, (*b* two-step type with variable filter constant, (*c*) six-step type and (*d*) logarithmic rotating sector

upon it. By rotating the smaller disc, any opening smaller than 90° can be set. Thus, at an opening of α, the percentage transmittance of the two-disc rotating sector is $100\,\alpha/180\,\%$. In good-quality equipment the precision of setting is 0·02°. This is achieved by a nonius and in the case of the optimum opening (90°) produces an attenuation precision of about $\pm 0{\cdot}02\,\%$.[9]

A version of the two-disc rotating sector is the two-step, variable-constant, rotating sector. As seen from Fig. 3.90*b*, the lower part of the slit remains open for a period that is shorter by a factor of $2\alpha/\pi$ than the corresponding interval for the upper part.

The free opening, α, between the two sectors can be varied in such a manner that the transmittance of the lower field relative to the upper one can be read directly from a scale. This ratio can be generally set to a precision of 0·1 % if the transmittance is greater than 1 %. At lower transmittance levels diffraction phenomena and light reflection adversely affect the accuracy.[8,17]

The step filter can be replaced by the rotating step disc depicted in Fig. 3.90*c*. The central angles belonging to the individual steps usually vary in the ratio 1 : 2 : 4 . . . or accordingly 0·30–0·60–0·90 . . . as $n.\log 2$ increases, where $n = 1, 2, 3, \ldots$ [10] The advantage of a rotating disc of the latter type is that the same blackening differences are found between each pair of steps of the photographed spectrum, provided the density values fall on the straight portion of the characteristic curve of the photographic emulsion.

In the case of the logarithmic rotating disc (Fig. 3.90*d*), the disc attenuates the light falling on the slit of the spectrograph in a logarithmic fashion along the length of the slit. This is due to the fact that points on the spiral-form disc circumference are at a distance l from the rotational axis of the disc. For this quantity the following relation holds

$$-\log \alpha = C + kl \tag{3.68}$$

where C and k are constants.[11]

For stigmatic line formation, therefore, the intensity of the spectral lines varies logarithmically along their lengths.

In the application of the rotating sector it should be taken into account that it cannot be used for the reduction of the radiance. This is because the rotating disc reduces the light flux only by periodically blocking the beam. The amount of radiation getting into the spectrograph through the slit is reduced by the periodic shortening of the time interval. This fact should be taken especially into account in photographing the spectra. It is not immaterial whether a certain amount of light energy reaches the photographic emulsion in a continuous fashion or in separate quantities. Errors can also occur when the speed of rotation of the disc is not even. The degree of attenuation becomes uncertain for periodic discharges if the number of revolutions per unit time is an integral multiple of the frequency of the mains, especially in the case of a.c. excitation methods (stroboscopic effect).

In the third method of light attenuation the cross-section of the beam, composed of parallel rays, is reduced by a diaphragm. The advantage of this method is that the degree of attenuation is independent of the wavelength, but is proportional to the area of the uniformly illuminated diaphragm. If the shape of the diaphragm opening is simple, its area and thus the attenuation can be simply calculated. The iris diaphragm with a polygonal shape, frequently used for other purposes, is not appropriate for this purpose. It can only be used for the reproducible and measurable attenuation of light after calibration. The same remarks apply to the use of a black wire mesh covered with carbon black. In this case, owing to diffraction on the grid, the attenuation is not independent of the wavelength.

Radiation can only be measurably attenuated by a diaphragm when the light illuminates the diaphragm opening in a completely homogeneous fashion. This aim can be achieved if, for example, in the intermediate focusing method, the diaphragm is placed immediately behind and within the frame of the first illumination lens.[12] By this method the illumination of the spectrograph can be evenly reduced.

In the diaphragm attenuation method the step filter is replaced by the Hansen step diaphragm.[13,14] This is a composite diaphragm with steps as shown in Fig. 3.91. The diaphragm, set parallel to the lens and illuminated evenly, is placed behind the first illumination lens such that the increasingly wider parts of the diaphragm should be perpendicular to the slit of the spectrograph. The cylindrical lens between the diaphragm and the slit of the spectrograph is placed such that its axis is parallel to the slit (Fig. 3.91). The cylindrical lens and the spherical lens

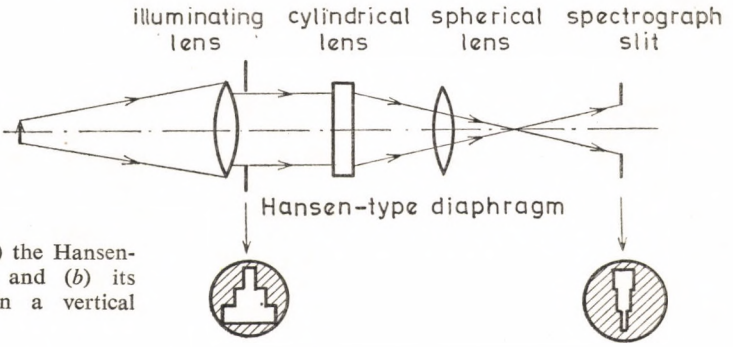


Fig. 3.91 Scheme of (*a*) the Hansen-type step diaphragm and (*b*) its optical arrangement in a vertical section

behind it form a composite lens in this arrangement that projects the narrowed image of the step attenuator onto the spectrograph slit. This results in a stepwise variation of the radiation entering the spectrograph, such that along the height of the slit the light flux varies proportionally with the width of the diaphragm steps. Thus, the relative exposure emerging between the steps can be easily calculated.

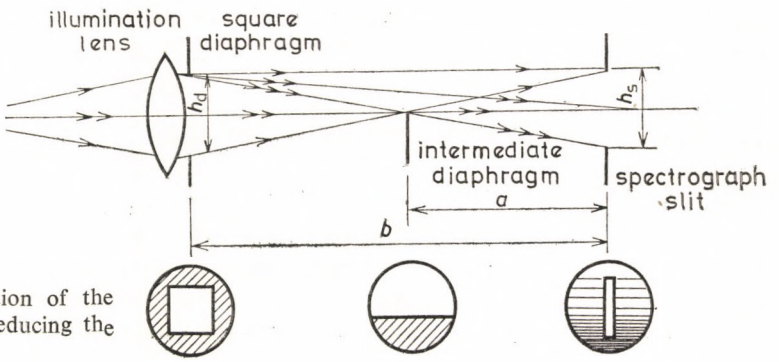


Fig. 3.92 Illumination of the slit by uniformly reducing the radiance

By using various optical devices it is possible that the intensity of the spectral lines, as attenuated by the diaphragm, drops linearly along their length.[15] In this method a square diaphragm is fastened to the rim of the illumination lens such that the two edges of the square are exactly perpendicular to the slit of the spectrograph (Fig. 3.92). Another diaphragm is then positioned such that the rays coming from the upper edge of the square diaphragm should illuminate the lower end of the slit, while those coming from the lower edge of the diaphragm should illuminate the upper end of the slit. However, the upper part of the slit is illuminated by radiation coming through the whole square diaphragm, the half-height of the slit receives

half the above radiance, while there is practically no light reaching the lower end of the slit. The distance a between the intermediate diaphragm and the slit can be easily calculated when the following quantities are known: h_d, the height of the square diaphragm, b, the distance between the square diaphragm and the slit, and h_s, the height of the latter. The distance a is given by

$$a = \frac{h_d b}{h_d + h_s} \tag{3.69}$$

References

1. KIEMLE, H., and SIEDENTOPP, H., *Z. Phys.*, **56**, 726 (1929)
2. ANGERER, E., *Apparate und Methoden der Spektroskopie*, Leipzig (1927)
3. LANGSTRITH, G. O. and MCRAE, D. R., *Proc. Fifth Summer Conference on Spectroscopy*, 105 (1935)
4. NAGY, Z and BENKŐ, K., *Spectrochim. Acta*, **19**, 2057 (1963)
5. TEISSING, H. and GOEBERT, M., *Z. tech. Phys.*, **21**, 305 (1940)
6. HANSEN, G., *Spectrochim. Acta*, **2**, 374 (1940)
7. KAISER, H., *Spectrochim. Acta*, **2**, 385 (1940)
8. KORTÜM, G., *Kolorimetrie, Photometrie und Spektrometrie*, vol. 2, 4th edn, Berlin–Göttingen–Heidelberg, 87 (1962)
9. KORTÜM, G., *Z. InstrumKde*, **54**, 373 (1934)
10. NACHTRIEBE, N. H., *Principles and Practice of Spectrochemical Analysis*, New York–Toronto–London, 86 (1950)
11. SCHEIBE, G. and NEUHÄUSSER, A., *Z. angew. Chem.*, **41**, 1218 (1928)
12. ORNSTEIN, L. S., MOLL, W. J. H. and BURGER, H. C., *Spektralphotometrie*, Braunschweig, 83 (1937)
13. HANSEN, G., *Z. Phys.*, **29**, 356 (1924)
14. KINDER, W., *Z. InstrumKde*, **56**, 393 (1936)
15. HIRSCH, R. and SCHÖN, M., *Z. Astrophys.*, **1**, 164 (1930)
16. NAGY, Z. and SÁMSONI, Z., *Spectrochim. Acta*, **24B**, 147 (1969)
17. RICHARDSON, H. M., FOWLER, R. C. and KOFFMANN, M. L., *J. Opt. Soc. Am.*, **43**, 873 (1953)

3.8 Basic principles of grating dispersion

3.8.1 Gratings

The dispersive element in grating light dispersion is the optical grating. This is a planar or concave, transparent or reflecting surface upon which there are rulings at regular distances. Thus, there are plane and concave, transparent or reflecting gratings. The distance between the rulings is the *ruling constant (grating constant)*.

Among plane gratings the low-performance types are transmission gratings, while the higher-performance types operate on the reflection principle. On the other hand, concave gratings are made only in the reflection mode.

3.8.2 Basic principles of plane-grating dispersion[1–4]

Generally, the same factors apply to both reflection and transmission diffraction gratings. Thus, for the sake of simplicity, plane transmission gratings will be discussed first. They can be regarded as a planar surface upon which there are a great number of parallel, regularly spaced, narrow slits. The spacing between the slits,

the grating constant, has the same magnitude as the wavelength of the radiation to be dispersed.

Let us first assume that a homogeneous, coherent beam of parallel rays falls normally on slits of width a. According to the Huygens principle every part of each slit becomes a spherical light source, that is, the different parts of the slits radiate light in all directions of space. Among these rays, however, only those that are in phase leave the slits in a direction perpendicular to the plane of the slit, that is, in the direction of incidence. These rays reinforce one another through mut interference, such that a radiance maximum is produced in the direction of incidence

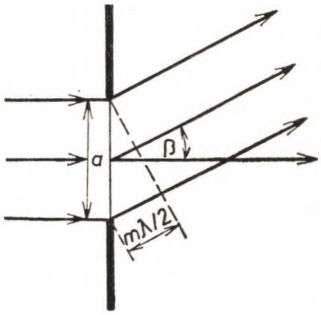


Fig. 3.93 Diffraction at a single slit

According to Fig. 3.93, however, it is evident that in addition to this maximum further radiance maxima are also produced in directions measured at an angle β relative to the normal. In these directions the condition

$$a \sin \beta = m \frac{\lambda}{2} \tag{3.70}$$

is satisfied, so that the value of m is an *even* integral number. Except for $m = 1$ on the other hand, for all odd values of m minima are produced. These rays differ by half a wavelength and extinguish one another as a result of interference (Fig. 3.94).

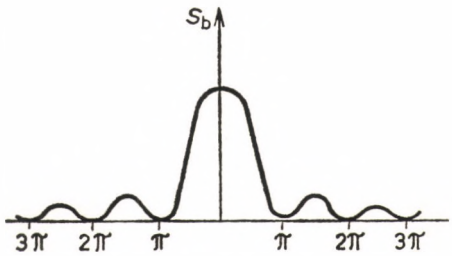


Fig. 3.94 Distribution of diffraction maxima and minima at a single slit

In the case of *two* adjacent slits of width a and spacing b, Eq. (3.67) is modified according to Fig. 3.95*a*, such that the radiance maxima and minima are determined by the following relationship for incidence rays normal to the grating

$$b \sin \beta = m \frac{\lambda}{2} \tag{3.71}$$

where b is the grating constant.

In the case of four slits, interference occurs not only among the rays coming from adjacent slits S_1 and S_2, or S_3, but also between rays from slits S_1 and S_3, or S_2 and S_4, as well as from S_1 and S_4 (Fig. 3.95*b*). The path difference between rays coming from two adjacent slits is $m \cdot \lambda/2$. Thus, radiance maxima are brought about when the value of $b \cdot \sin \beta$ is $0, 2\lambda/2, 4\lambda/2, \ldots$ and minima occur at values of $1\lambda/2, 3\lambda/2, 5\lambda/2, \ldots$. The path-difference between rays S_1 and S_2, and S_2 and S_4 is $2m\,\lambda/2$. The prerequisite for the emergence of radiance maxima is that the value of m is 0, 1, 2, while on the other hand minima are brought about when $m = 1/2, 3/2, 5/2, \ldots$ Thus, a maximum emerges if $b \cdot \sin \beta = 0, 2\lambda/4, 4\lambda/4, \ldots$

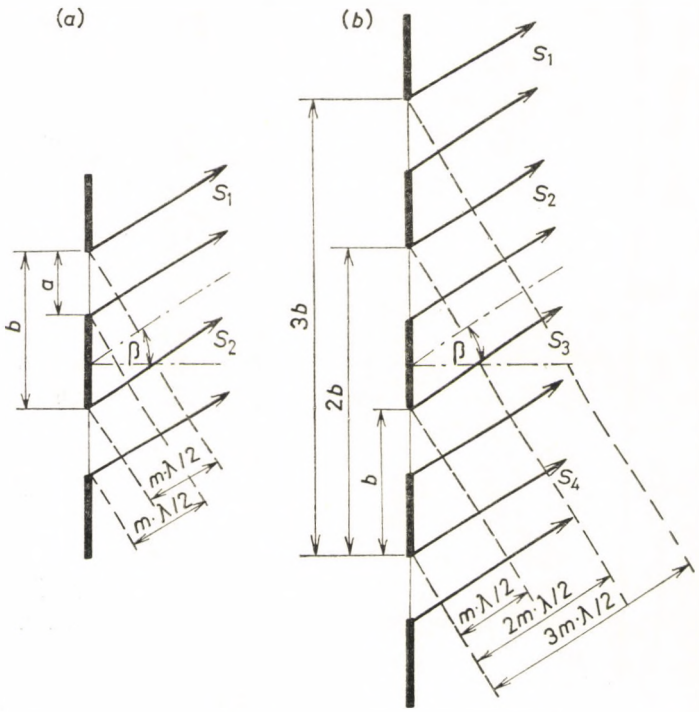


Fig. 3.95 Diffraction (*a*) at two slits, and (*b*) at four slits

Thus, in the general case the positions of the maxima produced by N slits are given by the formula $(2m + 1)\,\pi/N$, while the positions of the minima are given by $2m\,\pi/N$, where m is an integer.

If the positions of the maxima and minima produced by two, four, and eight slits, respectively (see Fig. 3.96), are compared, then it is seen that to the left and right of the main maximum (I), other main maxima arise (II). The greater the number of the slits, the greater the number of secondary maxima developed between these main peaks. By increasing the number of slits the radiance of the maxima is also increased. The position of the primary maximum radiance does not change. As the beam produced will be narrower the corresponding spectral line will be sharper. In addition, the peaks between the maxima are progressively decreased. When the distributions of the minima produced by two or four slits are compared, it becomes evident that in the case of diffraction from four slits a minimum arises at the same location where a maximum ought to occur. This apparent anomaly can

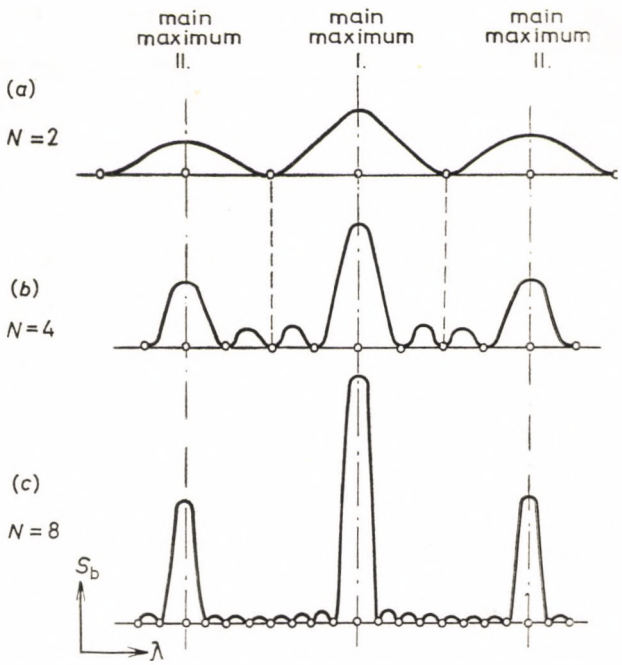


Fig. 3.96 Distribution of radiance maxima and minima: for (*a*) two slits, (*b*) four slits and (*c*) eight slits. (Only the diffraction arising on neighbouring slits is taken into account)

be explained by pointing out that at this position a minimum is produced by diffraction from *two* slits!

In the former treatment we have assumed that the light falls normally upon the diffraction grating, that is, $\alpha = 0$. This case, however, is the exception rather than the rule in practice. If the incident light is oblique, Eq. (3.70) is modified such that the following holds

$$b\,(\sin\alpha - \sin\beta) = m\lambda \tag{3.72}$$

According to Fig. 3.97 the path difference between the rays falling upon the grating at angle α and leaving the latter at angle β is $OC - AB = b\,.\,\sin\alpha - b\,.\,\sin\beta$. On the

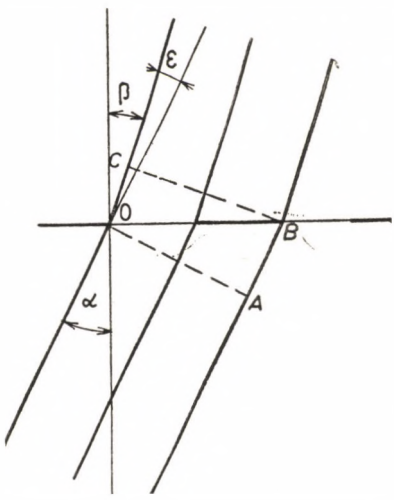


Fig. 3.97 Diffraction of light on a plane grating for oblique illumination of the grating

other hand, because of diffraction brought about by the grating of constant b, radiance maxima arise when the condition determined by the above equation is satisfied, such that m can be any positive or negative *integer.*

It is evident from the above equation that when coherent, parallel rays of wavelength λ fall on the grating, then spectral lines corresponding to the value of m are formed in the direction: $\pm\beta_1, \pm\beta_2, \ldots, \pm\beta_m$. Thus, in grating dispersion a spectral line of wavelength λ is formed, not only in one direction, but in $2m$ directions. Thus, depending on the value of m, first-order, second-order, ..., mth order spectra can be distinguished. Accordingly, the integer m is also the *spectral order number.*

According to general agreement the spectral order and thus the deviation angle ε are positive if the 'spectrum' arises on the opposite side of the direction of incidence, while if it lies on the same side it is designated by a negative sign. On the other hand, the angle of incidence of the ray always has a *positive* sign. Accordingly, Eq. (3.72) is modified in such a way that the following holds

$$m\lambda = b\,(\sin\alpha \pm \sin\beta) \tag{3.73}$$

According to the sign rule, when the rays falling upon the grating are found on the incidence side, which corresponds to positive β-angles, the positive sign is used, whereas in the opposite case the negative sign should be used. (With reflection gratings this rule is reversed.)

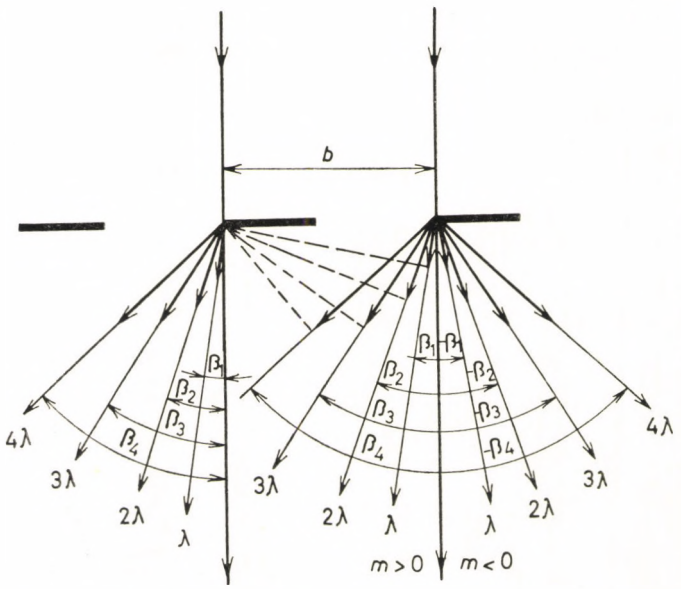


Fig. 3.98 Direction of emergence of different spectral orders using a transmission-plane grating

The above equation is the basic equation of grating dispersion. As will be seen later, this relationship is valid not only for plane gratings, but for concave grating dispersion as well.

It follows from the previous equation that for a transmission plane grating the zero-order spectrum arises from the continuation of the ray falling on the grating.

In this case $\beta = \pm\alpha$ since $b(\sin\alpha \mp \sin\beta) = 0$. A further characteristic of the zero order spectrum is that no dispersion occurs. The value of the product $m\lambda$ is independent of the wavelength λ and is always zero.

It follows from Eq. (3.73) that, for example, in the normal illumination of the grating (that is, when $\alpha = 0$), $\sin\beta = \pm m\lambda/b$. In grating dispersion the greater the angle of deviation for a ray of a certain wavelength (that is, the value of $\varepsilon = \alpha - \beta$) the smaller is the grating constant and the higher is the spectral order. The greater the order number of the spectrum in question, the further it will emerge either side of the zero-order spectrum (Fig. 3.98). According to the above relationship, the smaller the absolute values of $\sin\beta$ and angle β, the smaller is λ. Thus in grating dispersion, as opposed to prismatic dispersion, the shorter-wavelength lines are less deviated than the longer lines (Fig. 3.99).

The direction of emergence of a spectral line of wavelength λ can be determined by a graphical procedure based on the grating equation given above. In the application of this method a semicircle with a radius proportional to the grating constant b is drawn with a straight line through its centre in the direction of the angle of incidence, α. The intersection of the line and the semicircle is then projected onto the diameter (Fig. 3.100). Starting from point A on the diameter of the circle, points

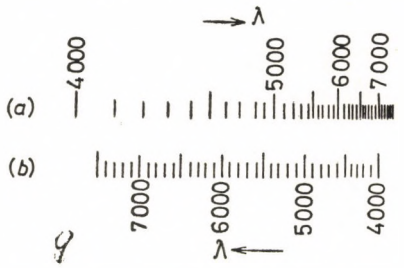


Fig. 3.99 Distribution of the wavelength (*a*) in prismatic dispersion and (*b*) in grating dispersion for a normald sectrum

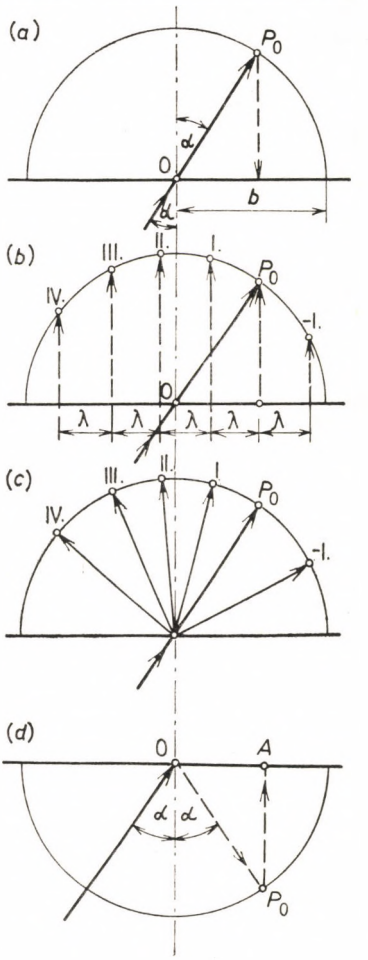


Fig. 3.100 Geometric procedure for the determination of the direction of emergence of different spectral orders: (*a*)–(*c*) transmission grating, and (*d*) reflection grating

are placed such that the distances between them are proportional to the wavelength of the spectral lines in question. Perpendicular lines are then drawn through these points to intersect the semicircle. These points determine the direction of emergence of spectral lines of wavelength λ and of the different orders, with respect to the centre of the semicircle (Fig. 3.100*b*). In the case of plane reflection gratings the above procedure is modified according to Fig. 3.100*d*.

Via the grating equation, the observation that in grating dispersion the spectra of different orders partly overlap one another can be justified. From Eq. (3.73) it follows that

$$\sin \beta = \frac{m'\lambda'}{b} - \sin \alpha = \frac{m''\lambda''}{b} - \sin \alpha \tag{3.74}$$

that is

$$\lambda'' = \frac{m'}{m''} \lambda' \tag{3.75}$$

In grating dispersion, therefore, the lines of the m'th-order spectrum coincide with the lines of the m''th-order spectrum for which the wavelength ratio λ''/λ' is the same as the reciprocal of the ratio of the appropriate orders m''/m'. For example, the 5000 Å line of the first order spectrum and the 2500 Å line of the second order will coincide. It follows from the above relationship that the disturbing effect of foreign spectral lines is lessened if the order number of the spectrum is low. When, for example, rays with a wavelength shorter than 2000 Å are not taken into account, the only range where spectral lines of foreign order do not cause any disturbance is between 2000 and 4000 Å of the first-order spectrum (Fig. 3.101).

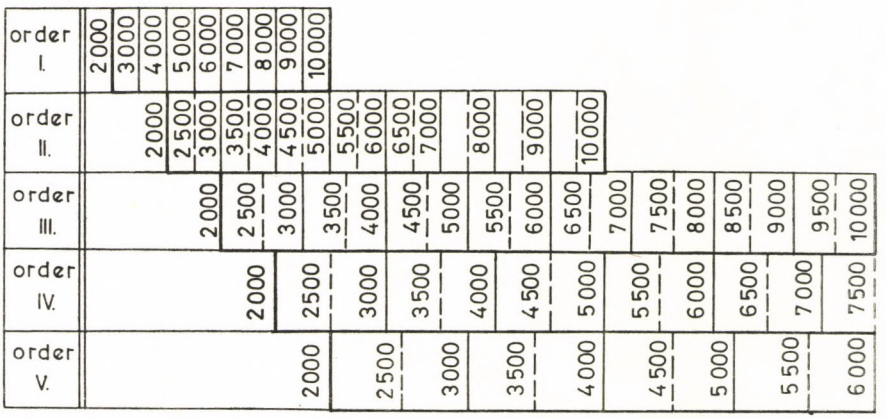


Fig. 3.101 Overlap between different spectral orders at 2000–10 000 Å

The angular dispersion of the plane grating can also be established by means of the differentiated grating equation (3.73). On differentiating the above expression we obtain

$$S_{a(g)} = \frac{d\beta}{d\lambda} = -\frac{|m|}{b \cos \beta} \tag{3.76}$$

In plane-grating dispersion, therefore, the greater the extent of angular dispersion, the higher the order of the spectrum and the smaller the grating constant, or the greater the angle of the transmitted ray. Accordingly, in contrast with prismatic dispersion, angular dispersion is greater in the larger-wavelength region of the spectrum than in the shorter-wavelength region.

The latter statement does not apply for the normal spectrum, that is for the spectrum that arises in the vicinity of the grating normal. In normal spectra, $\beta \approx 0$, thus $\cos \beta \approx 1$ and so is practically independent of the value of β, or indirectly of the wavelength of light. The characteristic feature of the normal spectrum is that its angular dispersion is practically independent of the wavelength of light. Since $\cos \beta$ has the greatest value in this case, the angular dispersion of the normal spectrum is the lowest in the same spectrum.

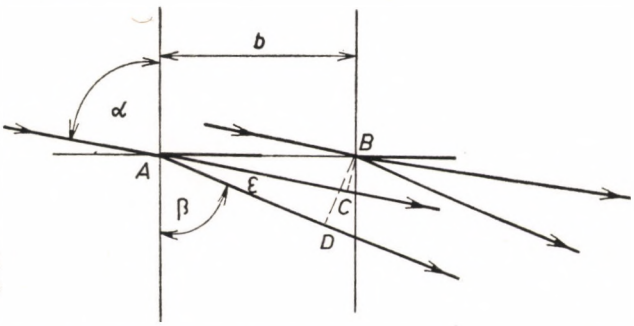


Fig. 3.102 Light dispersion caused by grazing incidence on a plane grating

In practice, light reaches the plane grating at large angles of incidence, or very obliquely. In this case, the grating equation (3.73) can be written in another form. From Fig. 3.102 we see that when the path difference between rays arising from point A on the grating and travelling in the direction of the mth-order and zero-order maxima is taken into account it follows that $AC - AD = b \,.\, \sin(\beta + \varepsilon) - b \,.\, \sin \beta$, or by an appropriate trigonometrical transformation: $AC - AD = b(\sin \beta \,.\, \cos \varepsilon + \cos \beta \,.\, \sin \varepsilon - \sin \beta)$. Considering that in this case $\varepsilon \ll \beta$, we can safely neglect the other terms and write

$$m\lambda = b \cos \beta \,.\, \sin \varepsilon \tag{3.77}$$

In this way, the angular dispersion of the grating for grazing incidence:

$$S_{a(gr)} = \frac{d\beta}{d\lambda} = -\frac{m}{b \sin \varepsilon} \tag{3.78}$$

If it is taken into account that at conventional illumination of the plane grating $S_{a(g)} = -m/b$, then it can be established that very oblique or *grazing* incidence has the same dispersion as if a grating constant of $b \,.\, \sin \varepsilon$ is used. As the value of ε is very small, that of $b \sin \varepsilon$ is also very small. It follows from this that $S_{a(gr)} \gg S_{a(g)}$. This has a particular importance in the region below 2000 Å where very small values of b would be otherwise needed. Therefore, by using oblique incidence the dispersion of the plane grating can be increased.

The larger the total number of grating rulings, the greater is the emerging radiance flux. It can furthermore be shown that for a certain mth-order spectrum, the value of the radiance maximum is proportional to the square of the number of rulings.

An increase in the number of rulings is an advantage from this point of view also. From the above relationship it is also clear that the dimensions of the slits and the absolute width of the unruled parts are immaterial.

The grazing incidence is customarily used only for reflection gratings and in the vacuum U.V. region. Namely, in this region the slight reflectance of the surface of the gratings is significantly enhanced by grazing incidence.[4]

Finally, the spectral lines of grating spectrographs are not curved, as opposed to the case of prism instruments.

References

1. HUNDT, F., *Einführung in die theoretische Physik*, Leipzig, vol. III (1950)
2. SCHALLREUTER, W. in *Grimsehls Lehrbuch der Physik*, 13th edn, Leipzig (1955)
3. NACHTRIEB, N. H., *Principles and Practice of Spectrochemical Analysis*, New York–Toronto–London (1950)
4. JAMES, A. and SAMSON, K., *Vacuum Ultraviolet Spectroscopy*, New York–London–Sydney (1967)

3.8.3 Basic principles of concave-grating dispersion[1–5]

The concave or so-called Rowland grating is a concave reflecting surface of large radius covered with exactly parallel rulings. Projection of these rulings onto the intersection plane shows that they are equidistant from each other.[5] According to Fig. 3.103 it can be seen that the relationship $b' = b/\cos\gamma$ holds, where the distance b is the grating constant. The characteristic feature of a concave grating is that sharp images of the spectral lines appear along the so-called Rowland circle.

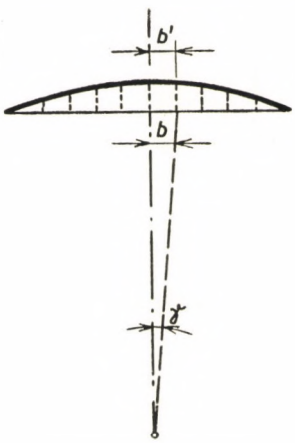


Fig. 3.103 Schematic drawing of a section perpendicular to the rulings of the Rowland grating

The diameter of this circle is equal to the radius of the grating, provided the grating is illuminated by rays coming from any point on the circle (Fig. 3.104).

A further prerequisite is that the grating rulings must be parallel to the slit, otherwise the spectral lines will not be sharp.

In order to derive a grating equation for concave-grating dispersion, let us assume that an element, AB, of a concave grating having a curvature radius O_rA is illuminated by the ray SA or SB coming from a point, S (the slit) on the Rowland circle, such that they form angles α or $\alpha + d\alpha$ with the incidence normals O_rA or O_rB

(Fig. 3.105). Rays AL or BL reflected at angles of β and $\beta + d\beta$, respectively, are united at point L into a spectral line when the difference between the paths SAL and SBL is equal to λ or to its integral multiple, $m\lambda$. It is therefore necessary that the following conditions are met:

$$(SA + AL) - (SB + BL) = (SC + DL) - (SB + BL) = BC + BD \approx$$

$$\approx AB \cdot \sin\alpha + AB \cdot \sin\beta = m\lambda$$

As an approximation, however, the arc AB, which is very small ($\varphi \approx 0$), can be taken as equal to the grating constant b. Thus it can be written

$$m\lambda = b \cdot (\sin\alpha \pm \sin\beta) \tag{3.79}$$

Thus, the basic equations of plane-grating and concave-grating dispersion are identical.

According to convention, the angle of incidence α is always positive in concave grating dispersion. On the other hand, the angle of reflection β is positive only if the direction of the reflected ray falls to the opposite side of the grating normal, while one falling to the left side is given a negative sign. The order of the spectrum is positive, if $|\beta| > \alpha$, that is, when the absolute value of the reflection angle is greater than the value of the angle of incidence.

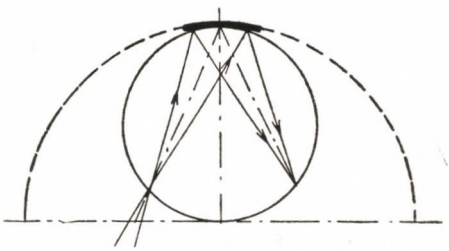


Fig. 3.104 Rowland circle

Eq. (3.79) is only valid when the grating, slit, and spectrum are situated on the Rowland circle, that is, when the grating is directly illuminated by light coming from the slit, relative to light coming from a parallel beam. In order to prove this, let us assume that the grating element, AB, of radius R is illuminated by rays

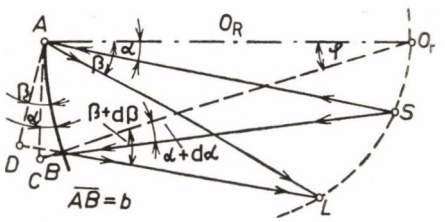


Fig. 3.105 Fundamental principle of concave-grating dispersion

coming from point S at a distance r_1 from the grating, in such a manner that the reflected rays form a spectral line in point L at a distance r_2 from the grating (Fig. 3.106). According to Eq (3.79) $d(\sin\alpha) \pm d(\sin\beta) = 0$; thus the following relationship is obtained

$$\cos\alpha\, d\alpha \pm \cos\beta\, d\beta = 0 \tag{3.80}$$

According to Fig. 3.106 it can furthermore be established that the relationship between the angles α, δ, and γ (the angles of the triangle ADS) is such that $\alpha + \delta + (180° - \gamma) = 180°$, therefore $\alpha + \delta = \gamma$. In a similar manner it can also

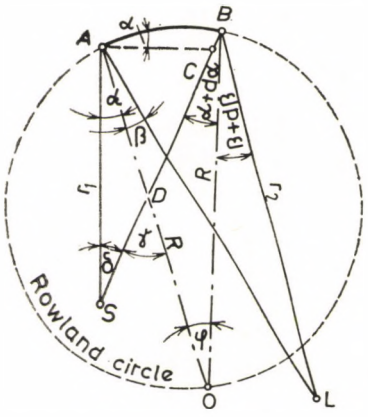


Fig. 3.106 Positions of the slit and the point of emergence of the spectral line relative to the Rowland circle

be shown that $\varphi + (\alpha + d\alpha) = \gamma$ and $d\alpha = \delta - \varphi$. Using a legitimate approximation, $\delta = AC/r_1$, while $\varphi = AB/R$, we obtain

$$d\alpha \approx \frac{AC}{r_1} - \frac{AB}{R} \tag{3.81}$$

similarly

$$d\beta \approx \frac{AC}{r_2} - \frac{AB}{R} \tag{3.82}$$

Substituting the two previous expressions in Eq. (3.82) we obtain

$$\cos\alpha\left[\frac{AC}{r_1} - \frac{AB}{R}\right] \pm \cos\beta\left[\frac{AC}{r_2} - \frac{AB}{R}\right] = 0$$

As the arc element AB can be regarded as a straight line and is, thus, perpendicular to the normal of the grating (the straight line O_rA), and as AC is perpendicular to the straight line AS, it follows that the angle formed by the lines AB and AC is equal to α. Thus $AC = AB \,.\, \cos\alpha$ and so

$$\cos\alpha\left[\frac{AB\cos\alpha}{r_1} - \frac{AB}{R}\right] \pm \cos\beta\left[\frac{AB\cos\beta}{r_2} - \frac{AB}{R}\right] = 0$$

that is

$$\cos\alpha\left[\frac{\cos\alpha}{r_1} - \frac{1}{R}\right] \pm \cos\beta\left[\frac{\cos\beta}{r_2} - \frac{1}{R}\right] = 0 \tag{3.83}$$

The above relationship is valid for all facets of concave-grating dispersion, even if the grating and the slit are not situated on the Rowland circle, or if the concave grating is illuminated by a parallel beam of light.

Eq. (3.83) is simplified if the spectrum emerges in the neighbourhood of the grating normal, that is, when $\beta \approx 0$. In this case

$$\cos \alpha \left[\frac{\cos \alpha}{r_1} - \frac{1}{R} \right] = 0 \tag{3.84}$$

This condition is satisfied only if $\cos \alpha / r_1 - 1/R = 0$, or $r_1 = R \cdot \cos \alpha$, that is, S is situated on the Rowland circle. Similarly, it can be shown that $r_2 = R \cdot \cos \beta$, that is, L is also situated on the Rowland circle. Accordingly, the grating equation (3.81) is only valid if the slit, the grating, and the spectrum are actually situated on the Rowland circle.

It follows from grating equation (3.81) that monochromatic rays reflected from the concave grating form spectral lines along the Rowland circle in the emergence direction β, if the angle of incidence of the ray coming from the grating satisfies the requirement

$$\sin \beta = \pm m \frac{\lambda}{b} - \sin \alpha \tag{3.85}$$

According to convention, spectra of different order will have a positive sign if they are formed to the same side of the 'direct image' or the position of the zero-order spectrum. The position of the direct image relative to the normal of the grating is determined by the angle of incidence of the ray according to the previous equation, such that $\sin \alpha = \sin \beta$, that is, $\alpha = \beta$. Accordingly, the position of the different spectral

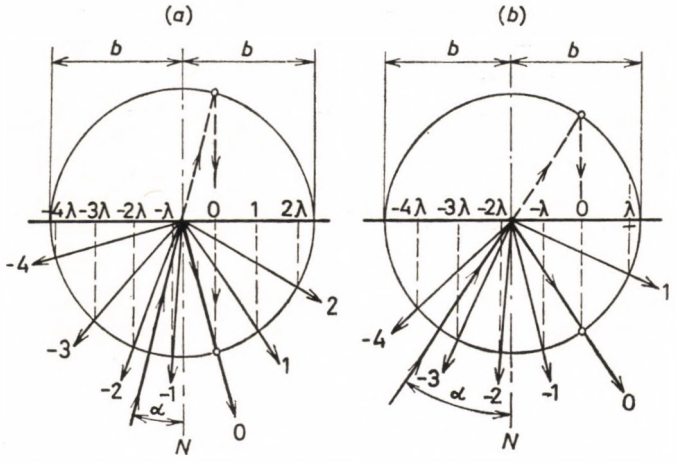


Fig. 3.107 Change of the emergence direction of the different spectral orders as a function of the angle of incidence

orders relative to the grating normal are moved around the Rowland circle as the angle of incidence varies. By appropriately changing the slit position, therefore, it is possible to produce at different locations on the Rowland circle different spectral orders or spectral lines of different wavelength (Fig. 3.107).

It becomes evident from the previous figure that the further the spectrum is situated from the grating normal, N, the greater is the arc length to which the same wavelength difference will correspond. This observation can also be demonstrated by using Eq. (3.79). When the angle of incidence, α, is constant, differentiation of the grating equation (3.79) yields

$$m\mathrm{d}\lambda = -b\cos\beta\mathrm{d}\beta \tag{3.86}$$

Thus, the angular dispersion for a concave grating is given by

$$S_{a(g)} = \frac{\mathrm{d}\beta}{\mathrm{d}\lambda} = -\frac{m}{b\cos\beta} \tag{3.87}$$

Furthermore, when it is taken into account that $\cos\beta = 1 - \beta^2/2! + \beta^4/4! + \ldots$, then the approximate relationship is obtained

$$S_{a(g)} = \frac{m}{b}\frac{1}{1-\beta^2/2} \tag{3.88}$$

Therefore, the greater the angular dispersion of the concave grating, the higher is the spectral order and the smaller is the grating constant. The value of the angular dispersion does, however, depend on the angle of reflection β of the ray forming the spectrum.

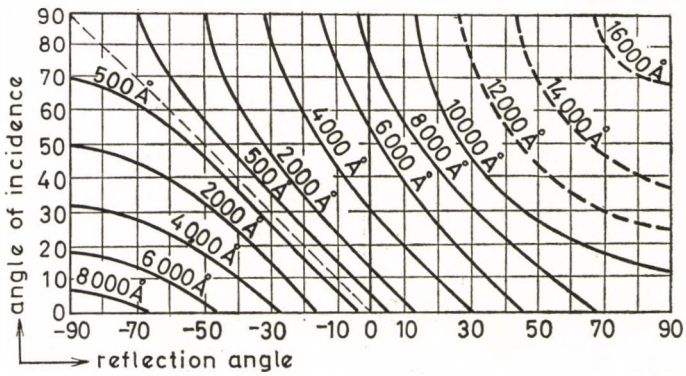


Fig. 3.108 Relationship of spectral lines of variable wavelength in the spectrum for different angles of incidence and reflection when a concave grating of ruling density 12 000 lines/cm is used

Considering that the value of the cosine function decreases with increasing angle values, it follows from the above equation that the value of the angular dispersion also increases with increasing angle β. Thus, in contrast to prismatic dispersion, a concave grating produces greater deviation at longer wavelengths than for shorter wavelengths. As a rule, therefore, the greater the increment of the angular dispersion, the longer the wavelength of light. The extent of the deviation does, however, depend on the value of the angles of incidence and reflection (Fig. 3.108).

It also follows from Eq. (3.88) that the smallest angular dispersion will be in the neighbourhood of the grating normal. In this case $\beta \approx 0$.

For the origin of normal spectra [$\beta \approx 0°$ in Eq. (3.79)] the grating equation is modified such that $m\lambda = b \,.\, \sin\alpha$. Considering the extreme values of the angle of incidence, α, it is possible to determine the limiting wavelength values given by

$$\lambda_l = \frac{b}{m} \sin\alpha \tag{3.89}$$

between which the grating performs the necessary dispersion. Since the possible lowest value of the angle of incidence is $\alpha = 0°$, while the maximum value is 90°, the wavelength limits according to Eq. (3.89) are

$$0 < \lambda_l < \frac{b}{m} \tag{3.90}$$

Table 3.8

Relationship between the wavelength limits of the mth order spectrum resolved by a grating of constant b and the extreme values of the angle of incidence and reflection

Appropriate		Spectral range	Sign of order number
angle of incidence	angle of reflection		
$\alpha \approx 0°$	$\sin\beta \approx \frac{m\lambda_l}{b}$	$0 < \lambda_l < \frac{b}{m}$	++
$\alpha \approx 90°$	$\sin\beta \approx \frac{m\lambda_l}{b} - 1$	$\frac{b}{m} < \lambda_l < \frac{b}{2m}$	+−
$\sin\alpha \approx \frac{m\lambda_h}{b}$	$\beta \approx 0°$	$0 < \lambda_l < \frac{b}{m}$	+−
$\sin\alpha \approx \frac{m\lambda_h}{b} - 1$	$\beta \approx 90°$	$\frac{b}{h} < \lambda_l < \frac{2b}{m}$	++
$\sin\alpha \approx 1 - \frac{m\lambda_h}{b}$	$\beta \approx -90°$	$0 < \lambda_l < \frac{b}{m}$	−−

The following inferences can be drawn from the data given in Table 3.8.

(1) The longer the wavelength and the lower the spectral order, the larger the grating constant should be.
(2) The optimum case is when the grating constant is somewhat larger than half the upper limit wavelength of the spectral range examined.
(3) The data given in the table are independent of the radius of curvature of the concave grating. The two previously given rules also apply to plane-grating dispersion. (This is because the plane grating can be regarded as a concave one whose radius of curvature is infinitely large.)

Astigmatic line formation causes the intensity of the spectral lines to vary along their height so that an even density is only found in the central portion. The focal

lines corresponding to the two ends of the slit tend to overlap less the closer they are to the ends of the slit.

The original grating equation (3.79) is not valid when the concave grating is illuminated by a parallel beam, instead of one coming from the slit of the spectrograph. In this case, instead of the distance between the slit and the grating [which is found from Eq. (3.83)] the value $r_1 = \infty$ should be taken into account. Thus, it can be calculated that the distance between the grating and the position of the spectral line is given by

$$r_2 = \frac{\cos^2\alpha}{\cos\alpha + \cos\beta} R \tag{3.91}$$

The spectral lines, therefore, do not arise along the Rowland circle when a parallel beam falls upon a concave grating.

The above relationship can be further simplified if it is assumed that the spectrum is always produced in the neighbourhood of the grating normal, as is seen in the Wadsworth-system concave-grating spectrographs. In this case, $\beta \approx 0°$, and the distance between the grating and the spectrum is given by the formula

$$r_2 = \frac{\cos\alpha}{1 + \cos\alpha} R \tag{3.92}$$

In the above equation the value of the ratio is always smaller than one. Thus, it can be established that the spectrum does not arise on the Rowland circle, but along a parabolic curve lying within it.

According to Eq. (3.87), the angular dispersion for a concave grating is influenced neither by the radius of curvature of the grating nor by the angle of incidence of light. Therefore as far as angular dispersion is concerned it is immaterial whether the rays falling on the concave grating are diverging or parallel. The relationship for angular dispersion by a concave grating is also valid in the above special case, when the concave grating is illuminated by a parallel beam.

References

1. NACHTRIEB, N. H., *Principles and Practice of Spectrochemical Analysis*, New York–Toronto–London (1950)
2. HUND, F., *Einführung in die theoretische Physik*, Leipzig, vol. III (1950)
3. STRONG, J., *Concepts of Classical Optics*, San Francisco (1957)
4. WIEN, W. and HARMS, F., *Handbuch der Experimentalphysik*, vol. 21
5. ROWLAND, H. A., *Phil. Mag.*, 5th ser., **13**, 469 (1882)

3.8.4 Dispersion of grating spectrographs[1–3]

The linear dispersion of grating spectrographs naturally depends on the angular dispersion of the grating used. It also depends on the optics that unite rays of different wavelengths into spectral lines.

In plane-grating dispersion the rays reflected from the grating are united into spectral lines by a concave mirror (or possibly a chromatically corrected lens). The value of the linear dispersion, therefore, depends on the focal length, f, of the concave mirror (or lens) as well as the angular dispersion of the grating. Thus, by

considering Eq. (3.78) the following formula can be written

$$S_{l(pg)} = fS_{a(pg)} = f\frac{|m|}{b\cos\beta} \quad \text{mm/Å} \tag{3.93}$$

In the case of concave-grating dispersion the value of the linear dispersion is also influenced by the radius, r, of the Rowland circle and the angular dispersion of the grating. Therefore, when spectral lines are formed on the Rowland circle, the linear dispersion is given by

$$S_{l(cg)} = rS_{a(cg)} = r\frac{|m|}{b\cos\beta} \quad \text{mm/Å} \tag{3.94}$$

On the other hand, in Wadsworth-system concave-grating spectrographs the dispersion of the spectral line arising in the neighbourhood of the grating normal is always influenced by the angle of incidence, α, instead of the angle of reflection, β. Thus, the linear dispersion is given by

$$S_{l(w)} = r\frac{m}{b(\cos\alpha + 1)} \quad \text{mm/Å} \tag{3.95}$$

It follows from the above that in grating dispersion the greater the linear dispersion, the higher the spectral order and the smaller the grating constant. In addition, for plane-grating spectrographs the linear dispersion increases with the focal length of the concave mirror, while for concave-grating equipment it increases with the radius of the Rowland circle.

In a given concave-grating spectrograph the radius r of the Rowland circle and the grating constant b or, in the case of a plane-grating instrument, the focal length f of the concave mirror and the grating constant b are constants. Thus, from the above equations we see that for the same mth-order spectrum, lines corresponding to the same wavelength difference are removed further from one another as the wavelength increases. As opposed to prismatic dispersion, therefore, the linear dispersion generally increases with increasing wavelength. Strictly speaking, this is also valid for the normal spectra. However, since the value of $\cos\beta$ changes only slightly with the angle β around 0°, the conclusion reached is that linear dispersion can be taken as constant for normal spectra.

From the above equations it follows that in grating instruments the greater the linear dispersion, the larger is β; that is, the larger is the angle of deviation or reflection of light. Thus, the greater the linear dispersion of grating spectrographs and, accordingly, the length of the different spectral orders, the higher is the order of the spectrum. In the mth-order spectrum, linear dispersion is m-times larger than for the first order. Accordingly, the first-order spectrum has the lowest linear dispersion and length.

It follows from Eq. (3.94) that in the Wadsworth arrangement of the concave grating the value of the expression $\cos\alpha + 1$ is approximately equal to 2, provided the angle of incidence upon the grating is small. At low angles of incidence, therefore, the linear dispersion of Wadsworth spectrographs is only about half the value of a Rowland-circle instrument when equipped with the same grating. With a greater angle of incidence the value of the linear dispersion increases, although it only reaches the level of the Rowland system when the angle of incidence is 90°. This

cannot be achieved for a number of practical reasons. Therefore, in the case of Wadsworth spectrographs as well, the longer the wavelength the greater is the linear dispersion.

References

1. NACHTRIEB, N. H., *Principles and Practice of Spectrochemical Analysis*, New York–Toronto–London (1950)
2. HUND, F., *Einführung in die theoretische Physik*, Leipzig, vol. III (1950)
3. STRONG, J., *Concepts of Classical Optics*, San Francisco (1957)

3.8.5 Resolving power and optimum slit width of grating spectrographs[1-4]

It can be easily shown that the theoretical resolving power of grating spectrographs depends solely on the mth-order of the spectrum and the total number, N, of the illuminated rulings of the grating. In order to prove this statement let us assume that (Fig. 3.109) parallel rays are falling upon a plane grating, AB. Among the rays reflected

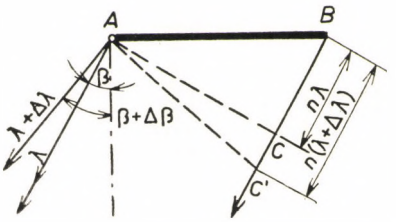


Fig. 3.109 Theoretical resolving power of grating spectrographs

from the grating at an angle β only those of wavelength λ will be united into spectral lines if for the first-order spectrum the distance $N\lambda$ and for the second order, $2N\lambda$ is equal to the path length, BC, starting from the Nth ruling (the last) of the illuminated grating (Fig. 3.109). In general, for an mth-order spectrum it is necessary that

$$mN\lambda = BC$$

When $\Delta\beta$ is the angular difference at which the rays of wavelengths λ and $\lambda + \Delta\lambda$ are just resolved, then it can be written that

$$mN(\lambda + \Delta\lambda) = BC'$$

The main maximum of a line of wavelength $\lambda + \Delta\lambda$, belonging to the mth-order spectrum, coincides with the first secondary maximum of a spectral line of wavelength λ, if

$$mN\lambda + \lambda = mN(\lambda + \Delta\lambda)$$

From this it follows that the theoretical resolving power of grating spectrographs is given by the expression

$$R_{0,\mathrm{gr}} = \frac{\lambda}{\Delta\lambda} = mN \tag{3.96}$$

The above equation is generally valid. Thus the higher the order of the spectrum and the larger the number of grating rulings, the higher the theoretical resolving power of any grating spectrograph (Fig. 3.110). The theoretical resolving power of

grating spectrographs is therefore influenced by neither the magnitude of the grating constant nor the wavelength. The theoretical resolving power is, furthermore, independent of the radius of curvature of the grating and the position of origin of the spectral lines. As opposed to the case of linear dispersion, the theoretical resolving power of Rowland and Wadsworth-type spectrographs with the same grating is identical.

On the above grounds it can be established that the theoretical resolving power of a modern grating 60 mm long with 1000 rulings/mm is $R_g = 60.1000.2 = 120\,000$ for a second-order spectrum. Thus, for example, at a wavelength of 3000 Å the wavelength difference of the theoretically resolved lines is 0·025 Å. The resolving power of good quality plane gratings can be as much as $R_g = 200\,000$.

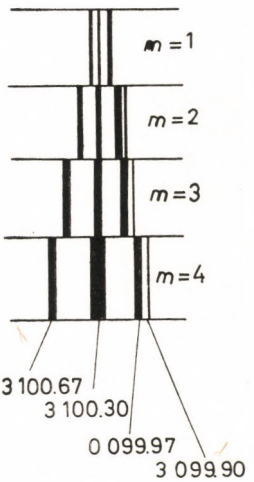


Fig. 3.110 Different-order spectra of the iron spectrum at 3100 Å produced by a grating having a ruling density 45 000 lines

Finally, it can be stated that according to practice the true resolving power of gratings approaches the value of the theoretical limit, especially in the ultra-violet spectral range.

The minimum optimal slit width is determined by the diffraction arising on the slit in concave-grating spectrographs. It can be theoretically proved that the minimum slit width in any system of concave-grating spectrographs at a wavelength λ is

$$d_{\min} = \frac{R_g}{B}\lambda \tag{3.97}$$

where R_g is the radius of curvature of the concave grating, while B is the width of the ruled portion.[4] The optimum slit width of concave-grating spectrographs is large when the ratio of the radius of curvature and the illuminated width of the grating is large and the wavelength is long.

The smallest optimum slit width of plane grating spectrographs is

$$d_{\min} = \frac{f}{B\cos\alpha}\lambda \tag{3.98}$$

where f is the focal length of the lens or concave mirror used, B is the working width of the grating, α is the angle of incidence and λ is the wavelength. The smallest

optimum slit width, therefore, increases with increasing focal length, wavelength and angle of incidence, while decreasing upon increasing working grating width.

References

1. NACHTRIEB, N. H., *Principles and Practice of Spectrochemical Analysis*, New York–Toronto–London (1950)
2. HUND, F., *Einführung in die theoretische Physik*, Leipzig, Vol. III (1950)
3. STRONG, J., *Concepts of Classical Optics*, San Francisco (1957)
4. PASCHEN, F., *Ann. Phys.*, **32**, 29 (1938)

3.8.6 Photographic speed of grating spectrographs

The speed of grating spectrographs depends on many factors. It can be generally stated that the speed increases fourfold as the total number of rulings on the grating increases. On the other hand, it generally decreases steeply with the spectral order. The speed also decreases quadratically with the distance between the grating and the place of emergence of the spectrum. In astigmatic line formation the speed increases with the height of the slit. In addition to the above generally valid statements, the speed of grating spectrographs also depends on the reflectivity of the grating surface, on the length and cross-section of the rulings, and on the performance level of the grating ruling process. The latter factors influence the radiance reflected from the grating to such an extent that there are no two optical gratings (that are virtually the same) with *completely* identical speeds.

By appropriately changing the profile of the grating rulings it is possible for the grating to reflect the dispersed rays, mainly in the direction of a spectrum of a certain order. Thus, by the use of step (or echelette) plane gratings, about 80% of the light energy falling upon the grating can be utilised for some higher-order spectrum.

The grating spectrographs in current manufacture have speeds that not only attain, but also surpass the speed of prism spectrographs having the same performance level.

3.8.7 Common diffraction gratings

It is evident from the previous chapters that the performance level of grating dispersion depends largely on the properties of the grating used. Accordingly, the requirements for diffraction gratings are exacting.

The reflectivity of the material of the grating should be very high and independent of the wavelength, while its thermal dilatation coefficient should be relatively small. Also, the material should be easy to work, but strong enough to resist external effects. The first requirement can be met by some metal like aluminium, or possibly silver (Fig. 3.111). Unfortunately, the thermal dilatation coefficients of these metals are relatively large; for example, the coefficients of aluminium and silver are $23{\cdot}1 \cdot 10^{-6}$ and $18{\cdot}7 \cdot 10^{-6}$ $°C^{-1}$, respectively.[1] It is also a disadvantage that these two metals can only be polished with difficulty. Thus, the rulings of modern diffraction gratings are made on an easily workable surface that also has a small dilatation coefficient. This surface is usually an optical glass that is first vacuum-coated by a very thin chromium layer and later by an aluminium layer. The thickness of the latter must

not be greater than the grating constant. The aluminium layer obtained in this way is resistant to atmospheric corrosion and is only attacked by acidic fumes and the vapours of chlorine and mercury. (The aluminium layer can, however, be easily removed from a damaged grating surface by a sodium hydroxide solution and the exposed chromium layer can then be recoated with aluminium.) The aluminium grating surface is not as efficient at reflecting the rays of wavelengths shorter than 1000 Å as a glass or quartz surface. Thus, for the study of the further ultra-violet range, glass or quartz gratings are more appropriate.

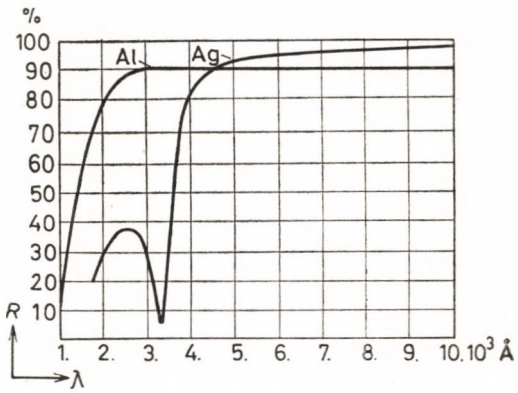


Fig. 3.111 Variation ofthe percentage reflectivity of the more important grating materials as a function of wavelength[1]

The practical ruling density or the optimum value of the grating constant is determined by the use to which the grating is put. Ideally, the grating constant should be somewhat larger than half the value of the longest wavelength utilised in the spectrum. The structure of diffraction gratings manufactured for analytical purposes and used in modern spectrographs is such that the ruling density varies between 300 and 1200 lines/mm. This density value cannot be significantly increased above 1500/mm.

The permissible error of ruling also depends on the type of application. The shorter the wavelength and the higher the spectral order, the more even should be the ruling distribution. For example, in the case of a grating that is mainly used at about 2500 Å, the permissible error is $\pm 250/2$ μm for the first-order spectrum, while for the second-order it is $\pm 250/4$ μm, and for the fifth-order $\pm 250/25$ μm. The depth and width of the rulings also depend on the application of the grating. A general rule is that the lower the order of the spectrum for which the grating is used, the finer and shallower the rulings should be. For higher-order spectra, rulings of greater depths and larger reflecting surfaces are needed.

The reflecting surfaces of common reflection gratings are formed by the unruled parts between the rulings. This has the drawback that the maximum radiance flux of the dispersed radiation will be found in the undispersed zero-order spectrum. The largest portion of the remaining energy is found in the first-order spectrum. The intensity of the spectra decreases rapidly as the spectral order increases. Thus, orders higher than the third are of so low a speed that they cannot be used for practical purposes.

This unfavourable situation can be completely eliminated by step or echelette-profile gratings.[2] These are special gratings where the rulings form a certain angle, α, with the plane of the grating reflecting surface. Therefore, echelette-profile gratings deviate a large proportion of the incoming radiation in a certain direction. When such a grating is illuminated by parallel light falling at right angles onto the reflect-

ing surfaces of the grating, then according to Fig. 3.112, the angle of incidence and reflection of light is equal to α. The value of the latter angle determines the so-called *blaze angle*, the direction of which is that of the mth-order spectral line whose speed is significantly higher than that of lines of similar wavelength arising in the other orders. The order of this spectrum can be calculated if we modify the generally valid relationship shown by Eq. (3.74) as follows

$$2b \,.\, \sin \alpha = m\lambda \tag{3.99}$$

The order of the spectrum formed in the direction of reflection has a high value when the angle of tilt of the reflecting surface is large and the wavelength is short. By the appropriate choice of the tilt angle, or the profile of the rulings, it is possible that the maximum intensity could even fall in the twentieth-order spectrum.

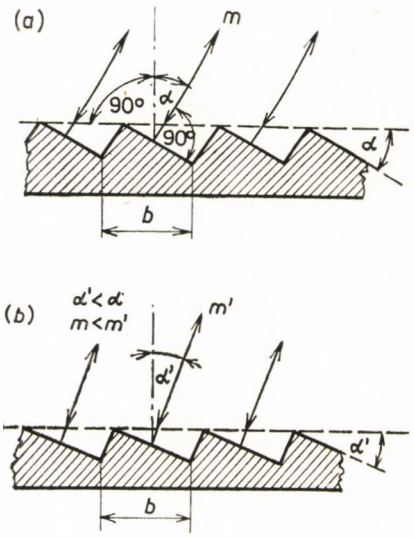


Fig. 3.112 Principle of grating dispersion by an echelette profile grating

Table 3.9 gives the relationship between the order of the spectrum and the blaze wavelength for the gratings of different ruling density and blaze angle. In order to compare the performance of the gratings the table also contains the theoretical resolving power, R_0, calculated for a grating width of 100 mm. It is to be noted that the blaze angle of an echellette grating usually refers to the wavelength belonging to the first-order spectrum. It is evident from Eq. (3.99) and also Table 3.9 that the wavelength values corresponding to the maximum intensity in any of the higher orders can be simply calculated from the blaze wavelength belonging to the first order. These intensities can be regarded as practically the same in a broad spectral range, and depend on the reflection characteristics of the material of the grating surface. Finally, it should be noted that a common reflection grating can be regarded as an echelette grating with a blaze angle of zero. In this case the maximum intensity falls in the zero order.

Master diffraction gratings are specially produced by very carefully designed ruling devices.[3] For the production of the rulings diamond heads are used that have a specially formed point. Not only are the requirements of these diamonds very strict, but the whole operation should be performed in a room free from vibration

the temperature of whose atmosphere is very precisely controlled. Thus, the production of master gratings is time-consuming and relatively costly.

The problems previously mentioned can be avoided with replica gratings. A number of different methods are used in the production of these gratings,[3] although the basic principles are generally the same. Here, a 'negative' ruling is made on glass and 'positive' replicas are produced by applying gelatine or some appropriate plastics to the ruled surface. The replica, which is then fastened onto a glass surface, is then vacuum-coated with a thick chromium or aluminium layer.

Even using the above process the quality of replica concave gratings is not the same as that of the master gratings. Their production, however, is cheaper, with costs as low as one-tenth of the production cost of master gratings. Accordingly, replica concave gratings are only used in low-performance spectrographs.

Table 3.9

Relationship between blaze angle, blaze wavelength, and order number, and also the theoretical resolving power, for two differently ruled gratings

$1/b$ (mm^{-1})	α (°)	m	λ_α (Å)	$R_0 . 10^{-4}$
600	20	1	11 400	6·0
		2	5 700	12·0
		3	3 800	18
		4	2 850	24
		5	2 280	30
200	40	1	64 260	2·0
		10	6 426	20·0
		20	3 213	40·0
		30	2 142	60·0
		40	1 606	80·0

Unlike replica concave gratings, replica plane gratings are not only significantly cheaper to produce than the original master grating, but are of such a quality that they completely meet all practical demands. This is because the ruling of the Merton plane gratings is practically error-free, provided the ruling density is not higher than 1000 lines/mm.[4] Thus, very few 'ghosts' are encountered when they are used.

In the production of Merton gratings a top-quality rotary bench is used for ruling an infinite spiral onto a perfectly polished metal cylinder by means of a sharp diamond. The advance rate is such that the desired grating constant is produced.[5,6] Even in this case, periodically recurring flaws cannot be avoided in the ruled spiral. In order to eliminate these errors the ruled part of the metal cylinder is fitted into a clamp lined with a comparatively soft plastics material and to which an appropriate cutting device can be fastened. The 'knife' of the latter is a pointed diamond whose shape corresponds to the desired ruling profile. The metal rod fixed in this way is evenly advanced such that at the other end a flawless spiral is produced. The cut produced completely fits the cross-section of the rulings of the grating. The negative replica of the grating is made by covering the spiral with a high-viscosity solution of a plastics material, for example polystyrene. The cylindrical plastics sheet is then cut open

parallel to the cylinder axis. The cutting takes place after the plastics film has hardened. The replica proper is then made by pressing the cylindrical surface against a thin layer of gelatine fixed to a glass sheet. After the gelatine has dried aluminium is vacuum-coated onto it. Sometimes sputtering methods are used for the metal coating process.

References

1. D'Ans, J., and Lax, E., *Taschenbuch für Chemiker und Physiker*, 2nd edn, Berlin–Göttingen–Heidelberg (1949)
2. Harrison, G. R. and Bausch, C. L., *Proc. of the Conf. on Optical Instruments, London*, 77 (1951)
3. Clark, G. L., *The Encyclopedia of Spectroscopy*, New York–London, 173 (1960)
4. Sayce, L. A., *Umschau*, **62**, 50 (1962)
5. Merton, T. R., *Proc. R. Soc.*, **A, 201**, 187 (1950)
6. Sayce, L. A., *Spectrochim. Acta*, **6**, 57 (1953)

3.9 Plane-grating spectrographs

3.9.1 Construction of plane-grating spectrographs

Basically, these instruments are identical to prism spectrographs. The difference in construction is that in the latter the prism is replaced by a plane grating and instead of the collimating and camera lenses there is a concave metal mirror. This system has the advantage that the focal length of the concave mirror is independent of the wavelength. Thus, the problems that arise from imperfect chromatic correction of the collimator, and camera lenses in prism spectrographs, are eliminated. Therefore, there is no need to change the distance between the slit and the collimating mirror or between the spectral plate and the mirror, even when the mean wavelengths of the photographed ranges are very different.

A further advantage of plane-grating spectrographs is that the spectrum is formed stigmatically, that is, the spectral lines are true images of the slit. The spectral lines are therefore straight and their height can be set by a slit-diaphragm. The intensity of the lines is independent of the slit height. The illumination of the spectrograph and the attenuation of the radiance are achieved in a manner similar to that for prism spectrographs. Furthermore, the spectrum is formed on a completely planar surface.

Another advantage is that the construction of plane gratings is much simpler and cheaper than the production of the concave types. Thus, it is understandable that, ever since the production of high-performance equipment, plane-grating spectrographs have gained wider acceptance than prism spectrographs and concave-grating spectrographs.

It follows from the above that the main components of plane-grating spectrographs are the slit, the collimating mirror, the plane grating, the camera mirror, and the photographic plate cassette. These are arranged so that the rays coming from the slit are projected onto the plane grating and the dispersed beams, reflected from the grating, are united into spectral lines by the camera mirror such that the rays are bent three times (Fig. 3.113). Thus, as opposed to those of single-prism spectrographs, the slit and the cassette are on the same side of the instrument. The construction of the spectrograph is therefore compact and its operation very simple.

The Czerny–Turner-system plane-grating spectrograph is based entirely on the principles previously given. The most popular types, however, are the Ebert-system spectrographs in which the collimator and the camera mirror are replaced by a common concave mirror. A further development of the latter is the high-performance Jaco–Ebert-system plane-grating spectrograph.

The requirements concerning the slit and the cassette of plane-grating spectrographs are completely identical to those of prism spectrographs. Thus, their construction is the same as well.

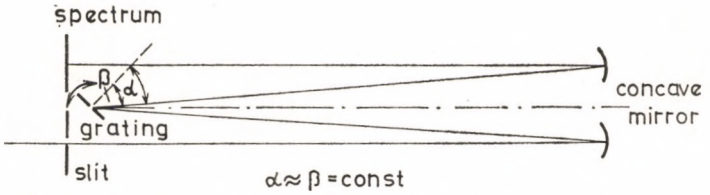


Fig. 3.113 Basic construction of plane-grating spectrographs

The collimating and camera mirrors (or the common mirror) always have a small aperture angle or a relatively large focal length. [The focal length or aperture number (*F*-number) of the camera mirror are frequently used for the designation of the instrument.] These mirrors are made of glass with a small temperature coefficient and vacuum-coated on the concave side with a layer of reflecting aluminium less than 0·001 mm thick. The focal length of the mirrors generally varies between 0·75 and 2 mm, depending on the performance of the equipment. At large focal lengths spherical aberration does not produce any important errors.

The ruling density of the gratings of plane-grating spectrographs also varies with the performance level of the equipment. Generally, the ruling density is between 300 and 1000 lines/mm. Gratings with higher density values, for example, 2000 lines/mm, are only exceptionally used.

3.9.2 Basic principle of plane-grating spectrographs

The handling of simple plane grating spectrographs is comparatively easy. When the spectral range is changed *only* the grating need be rotated; that is, only the angle of incidence, α, should be altered. According to the general grating equation (3.73), rays of wavelength λ are reflected back in the β-direction from a grating with a constant of b, towards the spectrum of order m, if the following condition is met

$$\lambda = \frac{b}{m}(\sin\alpha \pm \sin\beta) \tag{3.100}$$

However, it follows from the above equation that as the angle of incidence is increased a ray of a given wavelength leaves the grating at smaller and smaller angles. Upon increasing the angle of rotation of the grating, therefore, both the spectral lines and the spectral range are shifted towards the longer wavelengths. At each wavelength there is an angle of rotation at which the focusing is optimal. Instead of using very large angles of rotation (above 15–20°) it is optically more advantageous if the grating is replaced by another with a more suitable grating constant. This is

because the spectral range available for photographing also shifts to longer wavelengths if a plane grating with a larger constant is used. When an echelette-profile grating with an appropriate blaze angle is used no foreign-order lines will interfere with the spectrum.

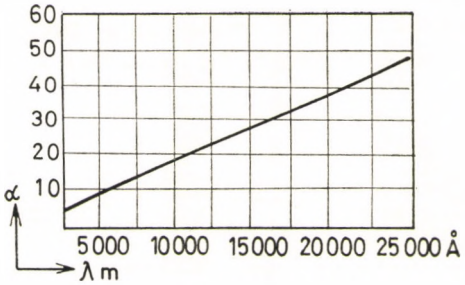


Fig. 3.114 Relationship between the grating rotation angle (α) and the wavelength of the ray falling on the centre of the grating for a spectrum of order λm

The gratings of commonly used plane-grating spectrographs where the Czerny–Turner system or the Ebert system are applied, have a 'symmetrical' position. The beam coming from the slit always impinges on the centre of the grating at the same angle at which the ray falling on the centre of the photographed spectrum leaves the centre of the grating. With regard to the fact that in this equipment the distance between the slit and the spectrum is always small (relative to the focal length of the concave mirror), the angle of incidence, α, and the angle of deviation, β, are regarded as being practically the same. Thus, the generally valid grating equation (3.73) can be modified such that

$$m\lambda = 2b \sin \alpha \tag{3.101}$$

It follows from the above relationship that the wavelength of a line in the first-order spectrum, that falls on the centre of the photographed spectrum, varies with the angle of rotation of the grating, as shown in Fig. 3.114. (The corresponding wavelength of a higher-order spectrum is obtained in such a way that the value read from the diagram is divided by the spectral order m.)

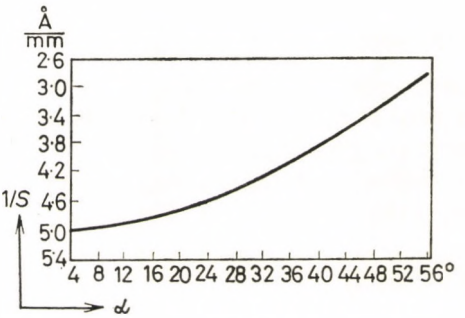


Fig. 3.115 Variation of the reciprocal linear dispersion of plane-grating spectrographs as a function of the grating rotation angle

According to Eq. (3.93) the reciprocal linear dispersion of plane-grating spectrographs is

$$\frac{1}{S_{l(pg)}} = \frac{b}{fm} \cos \alpha \quad \text{Å/mm} \tag{3.102}$$

where f is the focal length of the collimating mirror. By rotating the grating the reciprocal linear dispersion of a plane-grating spectrograph will decrease (Fig. 3.115).

From this point of view, therefore, a large angle of rotation is advantageous. However, when the angle is greater than 65–70° the speed of the spectrograph decreases rapidly.

Reference

1. GRAUE, G., ECKARD, S. and MAROTZ, R., *Arch. EisenhüttWes.*, **29**, 619 (1958)

3.9.3 Czerny–Turner-system plane-grating spectrographs[1–5]

These instruments are characterised by the fact that their collimating and camera mirrors are usually two similar concave metal mirrors, placed adjacent to each other. Another characteristic is that the beam entering the slit and the rays that fall on the focal surface of the camera mirror pass horizontally *beside* the grating. The advantage of using two similar concave mirrors is that the second mirror reduces the effect of any optical faults caused by the first mirror. A further advantage of this arrangement is that the rays only travel in planes perpendicular to the rulings of the grating. Therefore practically no polarisation occurs.

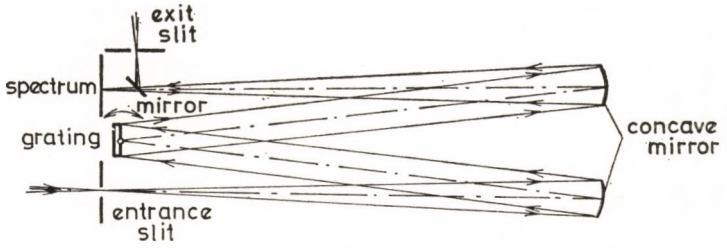


Fig. 3.116 Optical scheme of Czerny–Turner-system, plane-grating spectrographs

The grating of commercially available Czerny–Turner-system plane-grating spectrographs usually has a ruling density of 1200 lines/mm (or, rarely, 600 lines/mm). The focal length of the concave mirrors is generally 75 cm. The reciprocal linear dispersion observed in the first-order spectrum at 2000 Å is about 10·5 Å/mm, while at 14 000 Å it is about 6·5 Å/mm. In its linear dispersion the equipment is about the same as a medium-performance quartz spectrograph at about 2700 Å, while at about 3800 Å it is similar to the Littrow-system quartz spectrograph; and at about 5100 Å it is equivalent to a Littrow-system glass spectrograph. The resolving power is generally better than the spectrographs previously mentioned. Another very favourable aspect is that the photographed spectral lines are very sharp and straight, even at large slit heights. The speed of the spectrograph is comparatively high. To change the spectral range only the grating is turned. Thus, the spectrograph can be used for the study of a very broad spectral range between 1850 and 10 000 Å.

A further advantage of the Czerny–Turner-system spectrograph is that it can be easily transformed into a spectrometer. All that is required is to place between the camera mirror and the cassette a 45° deviation mirror that bends the rays at right angles and projects them onto the exit slit (Fig. 3.116). If a photomultiplier is placed behind the instrument exit slit and the grating is slowly rotated, then by using some appropriate chart recording system a plot is obtained that gives the relationship

between the light intensity and the wavelength. The resolving power of this photoelectric system is 0·1–0·2 Å. For example, the 3099·90–3099·97 Å wavelength 'line' of the so-called iron triplet is resolved into a doublet (Fig. 3.117).

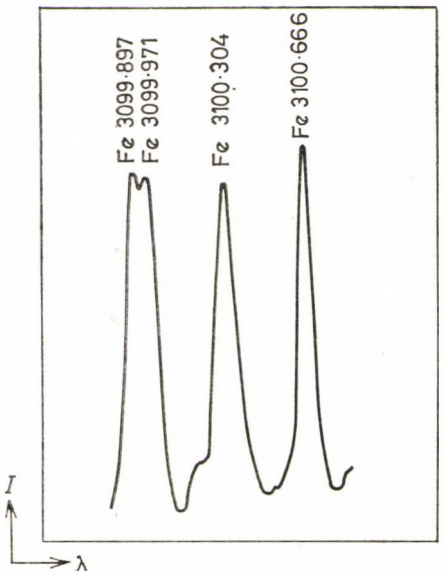


Fig. 3.117 Profile of the iron 'triplet' resolved by a Czerny–Turner-system spectrograph equipped with a grating of 1200 lines/mm ruling and a camera mirror of focal length 75 cm[6]

References

1. Czerny, M. and Turner, A. F., *Z. Phys.*, **61**, 792 (1930)
2. Leo, W., *Z. angew. Phys.*, **8**, 196 (1950)
3. Fastie, W. G., *J. Opt. Soc. Am.*, **42**, 648 (1952)
4. Shafer, A. B., Megill, L. R. and Dropplemann, S. L., *J. Opt. Soc. Am.*, **54**, 879 (1964)
5. Clark, G. L., *The Encyclopedia of Spectroscopy*, New York–London (1960)
6. What's New in Czerny–Turner Spectrometers, *The Spex. Speaker*, Metuchen, **9**, No. 3 (1964)

3.9.4 Ebert-system plane-grating spectrographs[1–5]

These instruments are different from the previously described Czerny–Turner-system plane-grating spectrographs in that the collimating and camera mirrors are replaced by the one concave metal mirror. This design can only be used if the rays coming from the slit pass almost parallel to the optic axis of the mirror. This requirement can be achieved in two ways. One solution involves the rays entering the instrument via a slit situated on the front part of the spectrograph and passing below the grating such that the lower part of the concave mirror is directly illuminated (Fig. 3.118). In the other construction the slit is found on the side of the spectrograph. The plane mirror deviates the rays coming through this slit by 90° in such a manner that they fall upon the lower part of the concave mirror. In construction of both types the concave mirror projects the parallel rays onto a plane grating whose perpendicular rotation axis intersects the optical axis of the concave mirror. The grating disperses the light in such a manner that the parallel beams, deviated

according to their wavelength, are reprojected onto the upper part of the concave mirror. The concave mirror reflects the beams of different wavelengths backwards above the grating so that they are united into spectral lines in the focal plane of the concave mirror.

It is evident from the above discussion that the rays travel in planes parallel to the grating rulings in Ebert-system spectrographs. The design of the common concave mirror that functions as a collimating and a camera mirror is not simple. The astigmatic error can be neglected if the focal length is sufficiently large.[6] The present Ebert-system plane-grating spectrographs are usually equipped with a grating having a ruling density of 600–620 line/mm, and a concave mirror of focal length varying between 2·00 and 3·50 m. Thus, the dispersion of the spectrograph varies between 5 and 7·5 Å/mm for a first-order spectrum. The dispersion is equivalent to a single-prism quartz spectrograph between 2200 and 2500 Å, while at about 3200 Å the dispersion is similar to that of a Littrow-system autocollimation quartz instrument. The theoretical resolving power of commercial Ebert-system spectrographs is about 45 000 and is significantly higher than that of Littrow-system autocollimated quartz or glass spectrographs.

From the previously given data it can be seen that by using Ebert-system plane-grating spectrographs most analytical work can be undertaken if the first-order spectrum is used. The advantage lies in the fact that in photography of the spectrum between 2000 and 8000 Å the grating need only be turned through an angle of between 5·5 and 9°. From the optical point of view it is an advantage if the angle of incidence or reflection of the light is kept small. Use of the first-order spectrum has the further advantage that the spectral range between 2000 and 4000 Å is practically undisturbed by spectral lines of other unwanted orders.

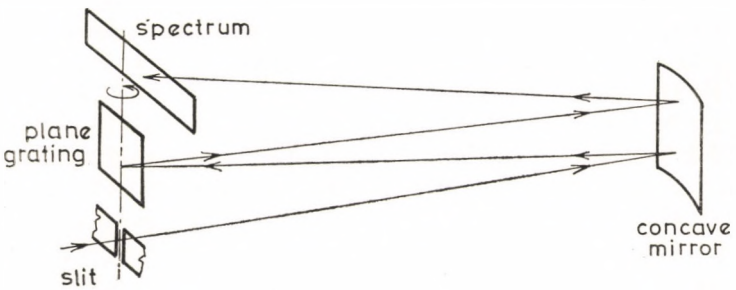


Fig. 3.118 Optical scheme of Ebert-system spectrograph

In Ebert-system plane-grating spectrographs, any interference caused by higher-order spectra can usually be removed by using a specially ruled grating that directs the rays falling upon it towards the blaze-line of the first-order spectrum. Thus, the higher-order spectral lines falling on the photographed spectrum are so weak that their disturbing effect can be practically eliminated. In addition, there is the advantage that the speed of Ebert-system spectrographs, equipped with such a special plane grating, can attain the speed of prism spectrographs with the same *F*-number. On the other hand, it is a drawback that in order to study a broader spectral range a number of gratings with different 'blaze' wavelengths are necessary.

For example, commercially available gratings can be used in the following wavelength intervals: a blaze wavelength of 3000 Å for a grating position of 5·5° in the

range 2000–5400 Å; a blaze wavelength of 3600 Å at a grating position of 6·5° for the range 2500–6400 Å, and a blaze wavelength of 5200 Å at a grating position of 9·7° for the range 3600–9300 Å. These special gratings can also be used at a reasonable speed level, in addition to the above optimum grating position, when the angle of reflection of the dispersed beams is very large (between 40 and 60°).

The performance of Ebert-system plane-grating spectrographs is doubled if they are equipped with a mirror to redirect the dispersed beam coming from the grating such that it is dispersed again before being projected onto the concave mirror. The dispersion of the secondary spectrum produced in the double path is twice the dispersion of the usual primary spectrum and corresponds to the dispersion or resolving power of second-order spectra, but the intensity lost is also noticeable.

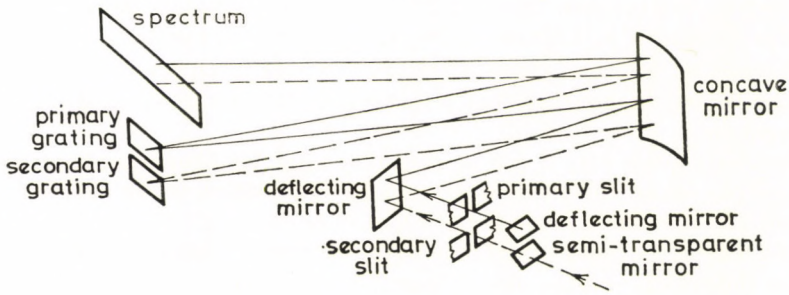


Fig. 3.119 Scheme of the Bausch–Lomb double-plane-grating spectrograph

In another variation of Ebert-system spectrographs, the so-called double-grating Ebert spectrograph, the rays coming from the light source fall on a semi-transparent deflecting mirror. This transmits and also bends the ray through 90° in an upward direction. The first ray enters the instrument through the lower slit of the spectrograph, while the other is projected by the deflecting mirror onto the upper slit. The two beams, passing one above the other, are made parallel by the common concave mirror and are then projected onto two rotatable plane gratings placed one above the other. One of the gratings has twice the ruling density of the other; that is, usually 600 or 1200 lines/mm. The rays are dispersed according to their wavelengths and are united into spectral lines by the concave mirror such that the two spectra, which can be photographed at the same time, are developed one above the other (Fig. 3.119).

There are a number of advantages in this method. Thus, e.g. two spectra can be made simultaneously of the same sample in two appropriately chosen spectral ranges and by different dispersions.

References

1. Ebert, H., *Wied. Ann.*, **38**, 489 (1889)
2. Jarrel, R. F., *J. Opt. Soc. Am.*, **45**, 259 (1955)
3. Fastie, W. G., *J. Opt. Soc. Am.*, **42**, 641 (1952)
4. Zeiss Catalogue No. 31-G. 3706-1, Jena
5. Clark, G. L., *The Encyclopedia of Spectroscopy*, New York–London (1960)
6. Fastie, W. G., *J. opt. Soc. Am.*, **42**, 648 (1952)

3.9.5 High-performance Jacob–Ebert-system plane-grating spectrographs, illuminated by a predispersion unit

The performance of grating spectrographs cannot be *significantly* increased such that the grating constant is reduced or the ruling number is increased. The only real solution to the problem is to use higher-order spectra instead of the first-order spectrum. According to Eqs (3.93) and (3.96) the linear dispersion and the theoretical resolving power increase m-times the first-order value in the mth-order spectrum.

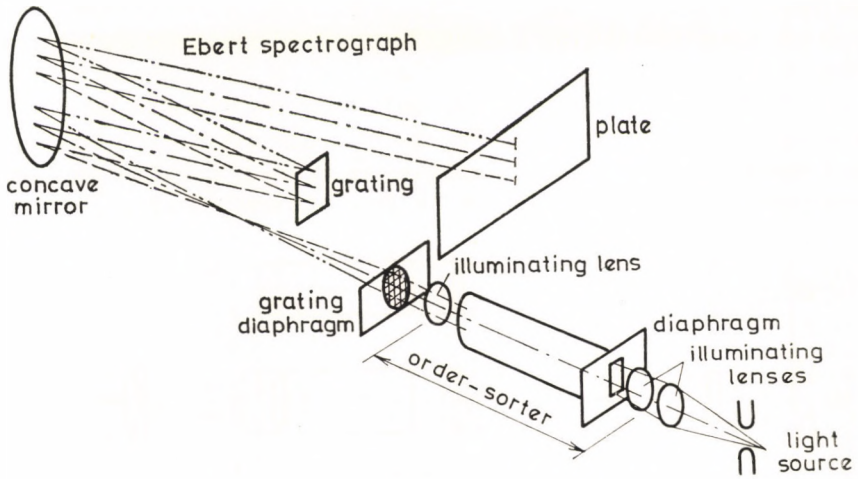


Fig. 3.120 Ebert-system plane-grating spectrograph illuminated by an order-sorter

The application of higher-order spectra has the further advantage that a grating with a large constant (which can be produced at low cost) can be used for light dispersion. In contrast to these advantages, one minor drawback is that the higher-order spectra are generally of slower speed than the lower-order spectra. However, the dominant disadvantage is caused by interference with the high-order spectra by lines of other spectral orders.

It is evident from the above that in Ebert-system spectrographs, the smaller the difference between the wavelengths of the spectral lines falling in the middle part of the spectrum the higher is the order of the spectrum.[3] The effect of unwanted lines interfering with the high-order spectra can therefore be eliminated by filters, or in certain parts of the spectrum by using selectively sensitive, photographic plates. This can only be achieved if the spectra of different order are separated in space.

The former requirement can be achieved if the slit of the spectrograph is not directly illuminated but via an order-sorter (or predispersion device) as shown in Fig. 3.120. This is a quartz–fluorite prism instrument with a horizontal slit. The axis of the spectrograph slit and the slit of the order-sorter are therefore at right angles. By means of this construction it is possible to focus the dispersed spectrum of the light source, produced by the order-sorter, onto the perpendicular slit of the Ebert-system spectrograph. By using an appropriately cut slit diaphragm it is possible to produce on the focal plane of the spectrograph various parts of the highly

dispersed spectrum one above the other. Therefore several, or more generally three, spectra can be photographed at the same time.

The fact that interference by spectral lines of unwanted orders is eliminated by the above method can be explained as follows. It follows from the relationship (3.100) that at a certain angle of grating rotation, α, not only the spectral line of wavelength λ_k of the first-order spectrum appears at the middle of the spectral plate, but also the spectral line of wavelength λ_k/m of the mth-order spectrum. For example, as shown in Fig. 3.119, for an angle of incidence of $\alpha = 42°$, the spectral line having a wavelength of 25 000/10 = 2500 Å in the tenth-order spectrum coincides with the line of wavelength 25 000/8 = 3125 Å in the eighth-order spectrum. This situation does not occur since the two lines will be produced one above the other, if the dispersion level of the order-sorter and the cut of the diaphragm in front of the slit are appropriate. Therefore it is absolutely necessary to obtain the best operating conditions for these parameters. For any particular spectrograph therefore only an order-sorter of a given dispersion curve can be used, such that only the spectral range required for analysis falls on the slit of the spectrograph.

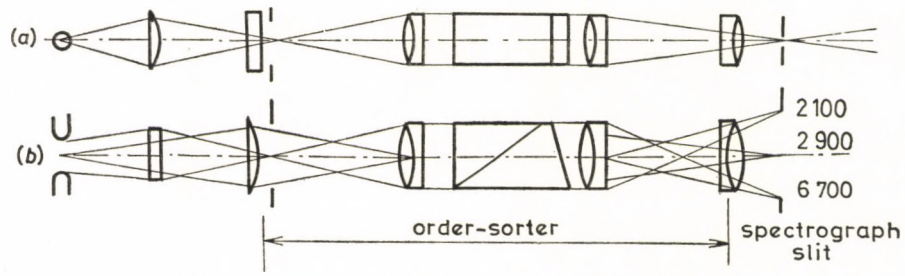


Fig. 3.121 Optical scheme of an order-sorting device: (*a*) horizontal section and (*b*) vertical section

The optical construction of the order-sorter is the same as that of a direct-vision spectroscope.[1,2] The important difference, however, is that its optics are made of fluorite and quartz instead of glass, so it can be used for the dispersion of ultra-violet rays as well. In the application of the equipment the radiation of the source is used to illuminate evenly the comparatively broad slit of the order-sorter. To achieve this, two cylindrical lenses are used, the first of which has a horizontal optical axis, while the second, placed immediately in front of the order-sorter slit, has a perpendicular axis (Fig. 3.121). By appropriate setting of the diaphragm placed behind the slit, the chromatically corrected collimating lens of the order-sorter is only illuminated by rays coming from the desired part of the light source. The parallel rays are dispersed by a composite prism made of fused quartz and fluorite so that the beams of different wavelengths leave the prism in a perpendicular plane. The deviated rays are then united into spectral lines at the slit of the spectrograph by the quartz–fluorite lens behind the prism. This lens is also corrected for chromatic aberration. The spectral lines are projected by the achromatic lens in front of the slit, through the slit diaphragm and onto the concave mirror of the Ebert-system spectrograph.

The dispersion plot of the order-sorter is curved (Fig. 3.122). It is mainly due to this fact that the photographed spectra of different orders are not parallel to each other on the spectral plate, but are slightly convergent. Those lines that are situated

below one another in the spectra of different orders are in reality somewhat displaced, although they are always formed at the same position for the same grating position. This anomaly occurs because the lines are formed astigmatically if the slit height of the spectrograph is too large.

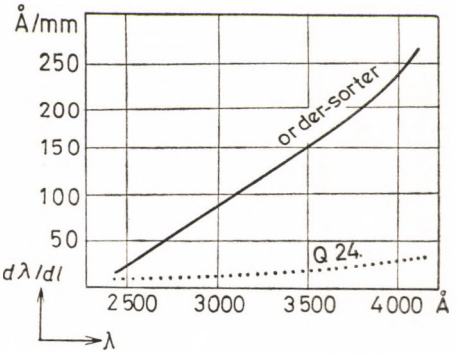


Fig. 3.122 Dispersion curve of an order-sorter

In the system described above the height of the slots of the slit diaphragm and the distance between the slots are the factors that determine the limits and the mean wavelength of the spectral range that can be photographed. By using an appropriate diaphragm design the disturbing effects of foreign-order lines in the photographed spectra can be prevented. The extent of the undisturbed spectral ranges varies with the order number of the spectra, the degree of rotation of the grating, and the curvature of the order-sorter dispersion curve. Therefore, slit diaphragms must be used where the slot heights and the distance between the slots can be varied as required.

It is for the above reasons that in commercial equipment the density of the grating rulings is such that either even-ordered or odd-ordered spectral ranges can be photographed. From Fig. 3.123 the factors mentioned in the previous text become apparent, as well as the fact that there are gaps between the spectral ranges. This is not generally a disadvantage since by rotating the grating the sections missing can also be photographed.

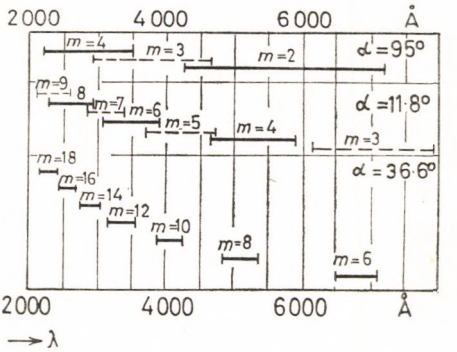


Fig. 3.123 Spectral ranges of different order that can be photographed simultaneously for different grating positions. The spectral range illustrated by a solid line is valid for a grating of 2880 lines/cm ruling, while the range given by the dotted line is valid for a grating of 5760 lines/cm. [3]

When Ebert-system plane-grating spectrographs illuminated by an order-sorter are used, it is a great advantage to the analyst that the performance of the equipment can be varied between wide limits without changing the grating. Their theoretical resolving power can be increased 10 to 20 times the resolution of the first-order spectrum, if the analysis is made in the spectral range between about 2100 and

4700 Å. The dispersion of the spectrograph can be varied, even exciuding the previous limitation, between 0·25 and 5·5 Å/mm. The same equipment can therefore be used as both a medium-performance and a high-performance spectrograph. It is a further advantage that the instrument has a sufficient speed even in the case of the higher-order spectra, provided gratings with an appropriate 'blaze' wavelength are used.

References

1. FASTER, W. Q., *J. Opt. Soc. Am.*, **42**, 661 (1956)
2. JARELL, R. F., *J. Opt. Soc. Am.*, **45**, 279 (1958)
3. SCHUHKNECHT, W., *Z. analyt. Chem.*, **136**, 81 (1952)

3.10 Concave-grating spectrographs

3.10.1 Basic principle of the instrument

In concave-grating dispersion the spectral lines are brought about astigmatically along the Rowland circle, either when the slit and the grating are situated on the same circle, or when the rays coming from the slit fall *directly* on the grating. These spectrographs have a very simple optical design, being composed of the slit, the grating, and the photographic plate holder.

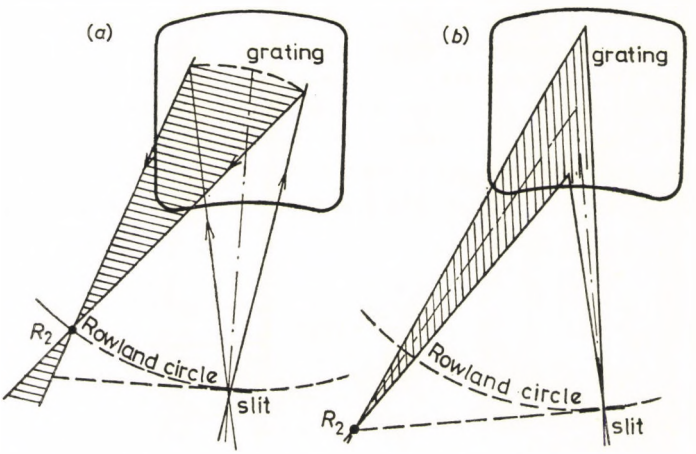


Fig. 3.124 Astigmatism of concave gratings: (*a*) path of rays perpendicular to the grating rulings and (*b*) path of rays parallel to the grating rulings

Astigmatic line formation is caused by the fact that only those reflected rays that are perpendicular to the grating rulings are united along the Rowland circle into first-order focal points (Fig. 3.124). Therefore, among the beams originating from a certain point of the grating only those of perpendicular incidence form points on the Rowland circle, while the beams parallel to the grating rulings appear as perpendicular lines along the Rowland circle. Thus, the spectral lines are not true images of the slit.

The extent of astigmatism is determined by the length, h, of the portion of the spectral line produced by the rays coming through a given point of the slit.[1,2] The value, h, depends on the angle of incidence, α, with which the ray impinges on this

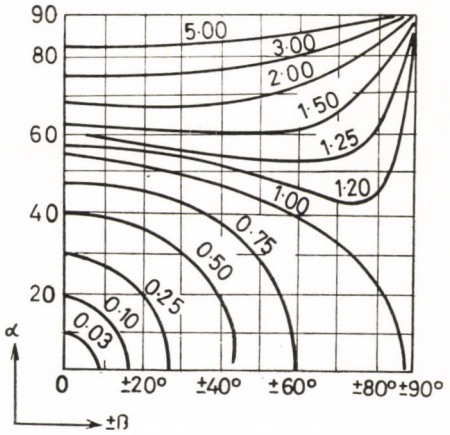


Fig. 3.125 Variation of the astigmatism of concave-grating spectrographs as a function of the angle of incidence, α, and the angle of reflection, β. (The variation of the value of the constant k is valid for the first-order spectrum of a grating of 1200 lines/mm ruling)

grating and also on the angle of deviation and the length, l, of the grating rulings in the following manner

$$h = (\sin^2 \beta + \sin \alpha \,.\, \tan \alpha \,.\, \cos \beta)\, l \tag{3.103}$$

or if the expression in brackets is denoted by k, then

$$h = kl \tag{3.104}$$

The value of the constant k varies with the angles of incidence and reflection according to Eq. (3.103) as shown by the diagram in Fig. 3.125. It can be seen that the astigmatism increases proportionally with the length of the grating rulings and the larger the angle of incidence or reflection, the greater the astigmatism. It also follows that if the degree of astigmatism is large the length of the spectral line can exceed the height of the slit.

Thus, in the case of astigmatic line formation the intensity of the spectral line depends on the slit height as well. This is due to the fact that the rays coming from all parts of the slit form not only points on the spectrum but also partly overlapping lines. The greater the slit height, the greater is the overlap and also the light energy, E, falling on the unit area of spectral lines. It can be shown that

$$E \approx A \left[\frac{\sin^2 \alpha \,.\, \cos \beta}{\cos \alpha} + \sin^2 \varepsilon \right] \tag{3.105}$$

where A is a proportionality constant. This relationship shows that the effects of astigmatism do not influence the relative speed of the spectral lines up to the point where the angles of reflection of rays of different wavelengths start to deviate strongly. The relative speed of spectral lines of similar wavelengths is therefore hardly influenced by astigmatism.

The intensity of the spectral lines is not entirely homogeneous along their length. An even intensity distribution is only found in the centre of the lines, where it represents the highest intensity level. This is explained by pointing out that the more the

point illumination of the lines is transformed into line illumination, the larger will be the area over which light is distributed and, therefore, the lower will be its intensity.

It is also possible to trace back to the phenomenon of astigmatism that spectral lines produced on the Rowland circle are only sharp when the axis of the slit is parallel to the grating rulings. From this it can be seen that no gap will be produced at the centre of the spectral line if the grating is masked such that only a narrow band remains uncovered at the upper and lower edges of the grating. On the other hand, spectral lines of the same wavelength merge into one another if the light illuminates parts of the slit that are not adjacent. The step slit diaphragm (see Subsection 2.3.2), therefore, cannot be used.

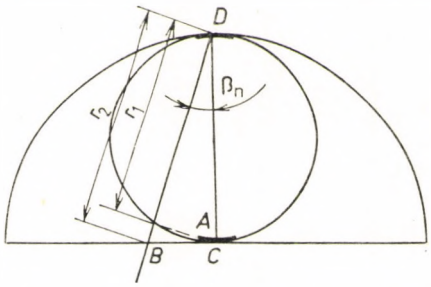


Fig. 3.126 Geometric interpretation of the Sirks astigmatic difference

It also follows from the above that in astigmatic line formation the heights of the spectral lines cannot be controlled by a diaphragm placed immediately before the slit. This can become possible, however, if the Sirks astigmatic difference[3] is taken into account. This is equal to the difference between the first-order, r_1, and second-order, r_2, focal lengths, or to the distance AB in Fig. 3.126. According to the figure, $r_1 = AD$ and $r_2 = BD$. Thus we obtain from the triangles ACD and BCD that $\cos \beta_n = r_1/R$ and $\cos \beta_n = R/r_2$, provided β_n is the angle of reflection of the ray travelling from the slit to the middle of the normal spectrum and R is the radius of the Rowland circle. According to the above relationships the following formulae can be derived

$$\left.\begin{aligned} r_2 - r_1 &= R(\sec \beta_n - \cos \beta_n) \\ \frac{1}{r_1} + \frac{1}{r_2} &= \frac{1 + \cos \beta_n}{R} \end{aligned}\right\} \qquad (3.106)$$

if the concave grating is taken to be a concave mirror. From these two equations the Sirks astigmatic difference can be calculated.

On the basis of the above it can be stated that in the case of astigmatic line-formation equipment, the use of diaphragms, step filters, rotating sectors, etc., can be successful only if they are situated at point B in Fig. 3.126 (the intersection of the lateral axis of the slit and a straight line that cuts the Rowland circle in the middle of the normal spectrum).

The greater the angle at which the ray coming from the grating falls on the centre of the spectral plate, the farther from the slit this optimum location lies. At large deviation angles this method is not recommended.[4]

In astigmatic line formation a number of factors should be taken into account that have no significance in stigmatic focusing. Astigmatic line formation cannot,

however, be regarded as definitely disadvantageous. It is an advantage, for example, that no error is produced by the uneven illumination of the slit, even in the case of quantitative analysis.

The previous difficulties are eliminated if the light coming from the slit illuminates the concave grating after having been made parallel, instead of directly. In this case, however, the spectrum arises within the Rowland circle, instead of upon it. This is true for the Wadsworth-system concave-grating spectrographs. On these grounds, concave-grating spectrographs can be grouped into two classes: astigmatic and stigmatic spectrographs.

References

1. RUNGE, C. and MANNKOPFF, R., *Z. Phys.*, **15**, 13 (1923)
2. DIEKE, J. J., *J. Opt. Soc. Am.*, **23**, 234 (1963)
3. SIRKS, J. L., *Astronomy and Astro-Physics*, **13**, 763 (1894)
4. THIERS, R. E., *J. Opt. Soc. Am.*, **40**, 849 (1950)

3.10.2 Different astigmatic concave-grating spectrographs[1–4]

The important optical components of the astigmatic concave-grating spectrographs are the slit, the grating, and the photographic plate, situated around the Rowland circle. These instruments can, therefore, be classified according to which part keeps its position or is not moved when the spectral range is changed.

From the above discussion various astigmatic concave-grating spectrographs, shown in Fig. 3.127, can be distinguished. In Rowland-system spectrographs only the slit position is fixed, while in Abney-system instruments only the slit position can be altered. On the other hand, in Paschen–Runge-system equipment both the slit and the grating have a fixed position. To photograph the different spectral orders the photographic plate or film is placed at the appropriate section of the Rowland circle. In Eagle-system spectrographs the slit has a fixed position. To change the spectral range the grating and the cassette must be rotated or the distance between the slit and the elements previously mentioned should be varied so that all three are situated on the Rowland circle.

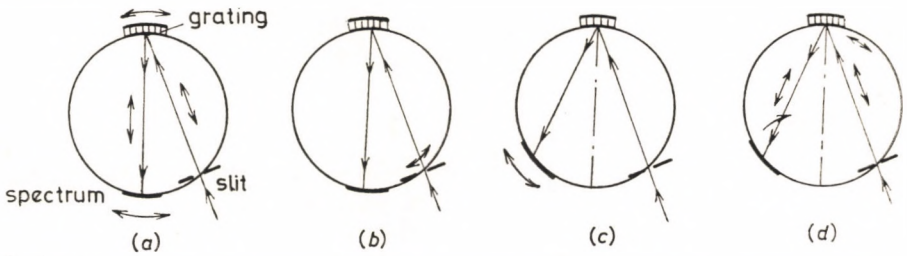


Fig. 3.127 Schemes of different astigmatic concave-grating spectrographs. (*a*) Rowland-system, (*b*) Abney-system, (*c*) Paschen–Runge-system, and (*d*) Eagle-system

Among the types listed, Paschen–Runge spectrographs and, to a lesser extent, Eagle-system instruments are commonly used for spectrochemical analysis. These spectrographs can be easily handled and are, therefore, applicable for the performance

of routine analysis. On the other hand, the features mentioned above are typical neither of Rowland-system nor of Abney spectrographs.

In the original Rowland-system spectrographs the spectral range could be changed by moving the grating and the photographic plate-holder in the vertical sense. The distance between the grating and the photographic plate is always equal to the radius

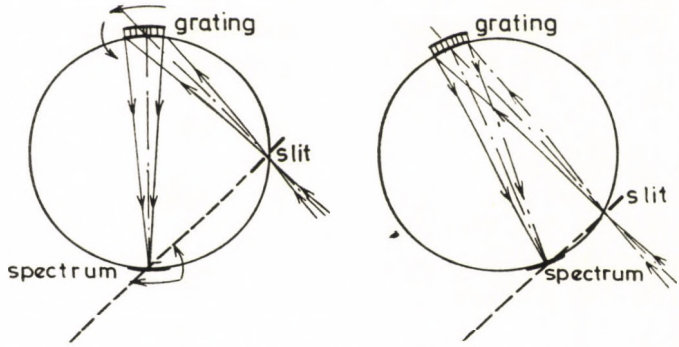


Fig. 3.128 Change of spectral range for Rowland-system spectrograph

of curvature of the grating. Under such conditions, slit, grating, and plate-holder always remain on the Rowland circle, while the middle portion of the photographed spectrum lies always on the normal of the grating (Fig. 3.128). A major drawback, however, is that handling this equipment is relatively complicated. A further disadvantage is that to obtain optimum conditions the spectrographs are made with

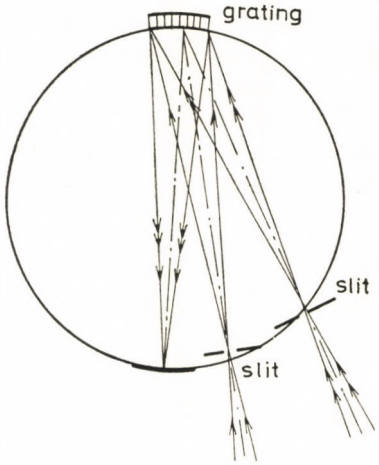


Fig. 3.129 Optical scheme of a two-slit, Abney-system spectrograph

a large Rowland circle that may have a radius greater than 6 m. Again, the fact that the astigmatism is comparatively large is another disadvantage. Thus, the light falls on the grating at angles that cannot be larger than 45°. Furthermore, the speed of the spectrograph is very low, especially for the higher-order spectra. Thus, Rowland-system spectrographs are used only for the determination of the wavelength of spectral lines. However, the wavelength of the spectral lines can nowadays be determined simply and more accurately by means of an interferometer.

The optical properties of the Abney-system concave-grating spectrographs are the same as those of the Rowland-system instruments. The spectrum appears on the normal of the grating. In the spectrum, therefore, the lines are evenly spaced if the wavelength difference is the same. Not only is the astigmatism of the spectrograph large, but it is complicated to use, in spite of the fact that to change the spectral range the positions of the grating and the cassette are not altered. It is difficult to rotate the slit around the Rowland circle such that the axis of the slit and the optical bench always pass *exactly* through the central point of the grating. Also, an error can develop when for all positions of the rotatable grating the slit axis and the rulings of the grating are not exactly parallel.

These unfavourable features can be eliminated when the spectrograph is fitted with a number of fixed slits and more than one optical bench, instead of a single rotatable slit (Fig. 3.129). This system, however, can be used for the photography of *certain* spectral ranges only. One difficulty is that the central lines of the slits should be made *exactly* parallel with the grating rulings. Any errors that may arise from non-parallelism, however, can be satisfactorily eliminated in low-performance equipment. Abney-system spectrographs have not, however, become very widespread.

References

1. SAWYER, R. A., *Experimental Spectroscopy*, 2nd edn, New York (1956)
2. HARRISON, G. R., LORD, C. and LOOFBOUROW, R., *Practical Spectroscopy*, New York (1949)
3. NACHTRIEB, N. H., *Principles and Practice of Spectrochemical Analysis*, New York–Toronto–London (1950)
4. WIEN, W. and HARMS, F., *Handbuch der Experimentalphysik*, vol. 21 Leipzig (1927)

3.10.3 Paschen–Runge-system spectrographs[1, 2]

Among the different astigmatic concave-grating spectrographs the Paschen–Runge-system instruments are the simplest and easiest to operate. In these spectrographs the position of neither the slit nor the grating can be altered. To photograph the different spectral ranges, therefore, only the photographic plate need be placed at the appropriate position on the Rowland circle (Fig. 3.130). In principle, the spectral range between 2100 and 9200 Å can be photographed with little instrumental alteration.

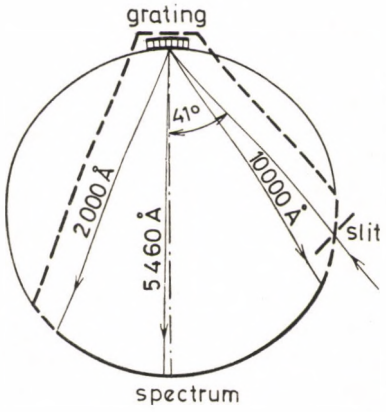


Fig. 3.130 Optical scheme of the Paschen–Runge-system spectrograph using a grating of radius of curvature 1·5 m and a ruling of 10 000 lines/cm

The optical properties of Paschen–Runge-system spectrographs have a number of distinct advantages. Their astigmatism is small relative to Rowland grating spectrographs and in the case of larger equipment the Sirks-type astigmatic difference can usually be neglected. It is advisable, however, to illuminate the spectrograph by the intermediate focusing technique to achieve the best results. Its application produces the primary emergence of the first-order spectrum; that is, the one that is the least disturbed by unwanted spectral order lines. Thus, this spectrograph can in many cases be used without filters. In practice, there is no disadvantage that the wavelength calibration of the spectrum is around 5500 Å; that is, in the spectral region that emerges in the vicinity of the grating normal. Although the equipment cannot be used for angles of incidence greater than 45°, this is no real problem. On the other hand, one disadvantage is that Paschen–Runge-system spectrographs are relatively voluminous. This disadvantage is reduced if the dispersion required is secured by means of a densely ruled Rowland grating, instead of one that has a large radius of curvature. The spectral plate, however, should be curved in this case and is generally more so than usual.

Commercial Paschen–Runge-system spectrographs are usually equipped with a Rowland grating 40 or 50 mm long, with a radius of curvature 1·5 or 2·0 m and of ruling density of 1200 lines/mm. The dispersion of the instrument is thus 7 Å/mm or 5·2 Å/mm, while the length of the first-order spectrum available for photography is 150 or 220 cm, respectively. The resolved minimum wavelength difference is given by the following values for a grating of radius of curvature of 1·5 m or 2·0 m:

at 2000 Å: 0·040 or 0·035 Å; at 6 000 Å: 0·125 or 0·100 Å;
at 3000 Å: 0·060 or 0·050 Å; at 8 000 Å: 0·165 or 0·115 Å;
at 4000 Å: 0·080 or 0·065 Å; at 10 000 Å: 0·205 or 0·150 Å.

For the purposes of spectrochemical analysis Paschen–Runge-system spectrographs are ideal and thus widely used by analysts.

References

1. RUNGE, C. R. and PASCHEN, F., *Abh. Akad. Wiss. Berl.*, Anhang **1** (1902)
2. NACHTRIEB, N. H., *Principles and Practice of Spectrochemical Analysis*. New York–Toronto–London, 67 (1950)

3.10.4 Eagle-system spectrographs[1–3]

In these instruments the different regions of the spectrum can be photographed without changing the slit position. In fact, the slit is not situated on the Rowland circle at all. Instead, it is attached on the side of the instrument inside the circle. The rays coming through the slit are projected onto the grating by a plane mirror such that the path length between the slit, the mirror, and the grating is always equal to the distance between the grating and the 'theoretical' slit lying on the Rowland circle. (It is therefore necessary that the distances denoted by Δ in Fig. 3.131 are exactly the same.) Accordingly, in order to change the spectral range it is first necessary to alter the distances between the deviating mirror and the grating, and between the grating and the central line of the cassette. Furthermore, the grating should be rotated about its perpendicular axis or about the central line of the cassette

(point S in the figure). All these operations should be done in such a way that the grating and the photographic plate remain on a Rowland circle. As the spectral range is changed the centre of the Rowland circle migrates, although it still passes through S – the position of the 'theoretical' slit.

The performance of the grating can be fully utilised only when the device used for the above process is very sensitive. The necessary settings are performed by precise servomotors. The automatic setting of the desired spectral range is achieved by a drum with a wavelength scale positioned in the neighbourhood of the slit.

One advantage of Eagle-system spectrographs is that a significant proportion of the spectrum can be photographed at one go. The photographic plate can occupy as much as 20–30° of the central angle of the Rowland circle. The spectrum, however, never emerges in the neighbourhood of the grating normal. Thus, the linear dispersion of the instrument increases somewhat with the wavelength of light.

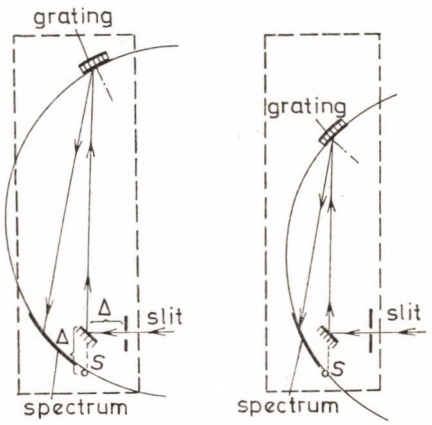


Fig. 3.131 Optical scheme of an Eagle-system concave-grating spectrograph set for the photography of two different spectral ranges

The astigmatism of Eagle-system spectrographs is greater than that of Paschen–Runge-system equipment, even under the most favourable conditions. From this point of view the best case is when both the angle of incidence and the angle of reflection are 40°. When the angle is larger the astigmatism increases, although the increase is not of such a magnitude that could significantly influence the usefulness of the instrument. The most frequently used spectral region has an incidence angle range of 10–45°. The longest wavelength line is produced with angles of incidence greater than 70°. In the latter case the extent of astigmatism is relatively large, but does not reach the level of Rowland- and Abney-system concave-grating spectrographs.

The advantage of Eagle-system spectrographs is that not only first-order but also higher-order spectra can be produced. By using a grating of the same size or ruling density, higher linear dispersion and theoretical resolving power can be obtained than with Paschen–Runge-system spectrographs. When the higher-order spectra are used, however, a concave grating with a large grating constant is necessary. A grating of this type can be manufactured relatively easily and modern instruments are made such that the grating can be changed if necessary. The optical design, similar to that of Littrow-system prism spectrographs, not only allows the use of gratings of radius of curvature larger than 6·5 m, but also has the advantage that the cassette lies in the vicinity of the slit.

A further advantage of Eagle-system spectrographs, compared with the previously described concave grating instruments, is that for the same performance level they occupy a significantly smaller area. On the other hand, it is a drawback that the performance of the grating is only fully utilised if the device operating the movement of the grating and cassette works extremely precisely. Thus Eagle-system spectrographs are generally used for research where high resolving power is absolutely necessary.

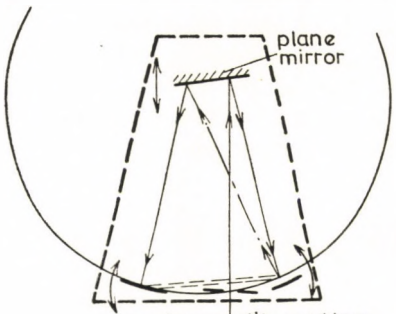


Fig. 3.132 Optical scheme of the Dowell–Menzies-system concave-grating spectrograph

The size of Eagle-system spectrographs can be reduced to about half, for the same performance level if the Dowell–Menzies-system is employed instead. In the case of this equipment the relatively large plane mirror, placed in the centre of the Rowland circle, deviates the rays coming from the slit in such a way that they are projected onto the concave grating situated on the Rowland circle (Fig. 3.132). The dispersed beams reflected from the grating are projected by the plane mirror onto the photographic plate positioned along the Rowland circle. To change the spectral range it is first of all necessary to alter the distance between the slit and the plane mirror such that the latter always remains parallel to its original position. Secondly, the grating and the cassette, which are rigidly fixed to one another, should be rotated around the central line of the spectrograph slit so that the former always remains on the Rowland circle. These movements are performed simultaneously by an automatic setting device. Dowell–Menzies-system spectrographs not only have the advantage of a reduction in space, but the higher-order spectra are also produced on both sides of the grating normal, instead of only on one side.

References

1. EAGLE, A., *Astrophys. J.*, **31**, 120 (1910)
2. CLARK, G. L., *The Encyclopedia of Spectroscopy*, New York–London, 185 (1960)
3. SAWYER R. A., *Experimental Spectroscopy*, 2nd edn, New York (1956)

3.10.5 Concave-grating spectrographs with stigmatic line formation

The spectral lines are true images of the slit when the concave grating is illuminated by parallel rays, instead of rays coming directly from the slit. A completely parallel beam is not entirely necessary, however.

Spectrographs whose Rowland grating is illuminated by parallel rays are, in fact, modifications of Paschen–Runge-system instruments. The only difference is that the

rays coming from the slit do not strike the concave grating directly, but pass through a cylindrical lens with a perpendicular optical axis. By using a lens of an appropriate focal length any astigmatic difference can be faded out. The spectral lines, therefore, emerge along the Rowland circle if the slit axis and the optic axis of the cylindrical lens are parallel to the rulings of the concave grating. In this equipment it is not possible to change the spectral range available for photography. This system can only be successfully applied in medium-performance spectrographs whose grating has a radius of curvature of 1·5 m, a ruling density of 415 lines/mm, a dispersion of 16 Å/mm, and a resolving power of 32 000.

In Wadsworth-system spectrographs the rays coming from the slit are made parallel by a concave metal mirror and are then projected onto the Rowland grating.[1] This has the consequence that the concave grating unites the reflected and dispersed beams into spectral lines within the Rowland circle according to Eq. (3.85). This fact does not influence the resolving power of the grating, but has the disadvantage that the linear dispersion is reduced to about half of the value of other concave-grating spectrographs equipped with the same grating. This does not occur if the radius of curvature of the grating is about twice as large as usual. This increases the space required by the instrument, although not very much. The spectrum arises within the Rowland circle along a parabolic curve that passes through the centre of the curvature of the concave grating. Any adjustment of the instrument is, therefore, rather complicated.

In the case of illumination of the grating by parallel rays, the spectrum is only stigmatic when the angle of reflection is not too large. Accordingly, Wadsworth-system spectrographs are constructed such that the spectrum always emerges on both sides of the grating normal and at angles of reflection smaller than 9°. Thus, $\beta \approx 0$ and the grating equation (3.73) is modified so that the following formula holds

$$\lambda = \frac{b}{m}(\sin \alpha \pm \beta) \tag{3.107}$$

The reciprocal angular dispersion is given by

$$\frac{1}{S_{a(w)}} = \frac{d\lambda}{d\beta} = \pm \frac{b}{m} \tag{3.108}$$

as the spectrum is produced on both sides of the grating normal in a symmetrical fashion (Fig. 3.133). On consideration of the above equations the so-called Wadsworth anomaly can be understood. This anomaly is due to the fact that in Wadsworth-system spectrographs the grating reflects the longer-wavelength rays at a smaller deviation angle than the rays of shorter wavelength. For spectra emerging in this way the longer the wavelength, the greater is the linear dispersion. Thus, from the diagram shown in Fig. 3.134 it becomes evident that at an angle of incidence of 10° and a wavelength of 4000 Å, the dispersion is 0·56 r/b, while at 6000 Å the dispersion is 0·66 r/b, where r is the radius of curvature of the grating and b is the grating constant.

In Wadsworth-system spectrographs the spherical aberration of the concave-grating increases rapidly with the angle of reflection. Thus, only a long spectrum can be used when its width is at most one-sixth the distance between the grating and the place of origin of the spectrum. In higher-performance equipment, therefore,

the whole spectrum cannot be photographed at once. To change the spectral range the following operations should be performed:

(1) The grating and the cassette should be bodily rotated around a perpendicular axis (denoted by *O* in Fig. 3.133) that passes through the central line of the grating such that the position of the latter does not change with respect to the grating normal.
(2) The curvature of the photographic plate should also be altered.

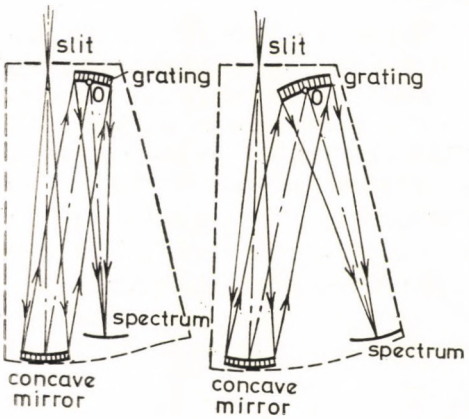


Fig. 3.133 Optical scheme of the Wadsworth-system concave-grating spectrograph set for the photography of two different spectral ranges

These operations are automatically performed in modern high-performance spectrographs by servo-motors, so that a sharp image of the spectrum is produced on the photographic plate. The perpendicular setting of the cassette, that is, the movement of the cassette necessary for the photography of the different spectra one above the other, is achieved by a separate motor. Thus, in spite of the fact that the slit and the cassette are to some distance from each other, the instrument can be entirely operated

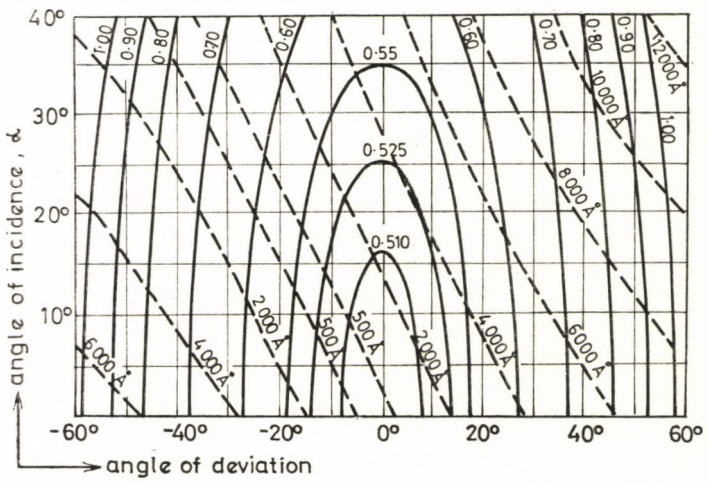


Fig. 3.134 Dispersion of Wadsworth-system spectrographs (in r/b units) for a grating of 30 000 lines/inch ruling[2]; on the ordinate: angle of incidence, on the abscissa: angle of deviation

from the neighbourhood of the slit. These high-performance spectrographs are usually equipped with a grating of radius of curvature of about 640 cm.[3] The dispersion of an instrument fitted with a grating having a length of about 15 cm and a ruling density of about 600 lines/mm is 5 Å/mm in the first-order spectrum and has a theoretical resolving power of about 63 000 $\lambda/\Delta\lambda$. Some Wadsworth-system spectrographs are also made with gratings that have a ruling density of 12 000 lines/cm, although optically they are the same. By using these instruments the 50 cm portion of the spectrum between 2100 and 7800 Å can be photographed in a number of sections.

For the study of spectra with only a few lines the lower-performance Wadsworth-system spectrographs are adequate. These are characterised by the fact that the whole spectrum can be photographed at once. They have a fixed optical construction that has the advantage that it is much simpler and, therefore, less expensive. On the other hand, the dispersion of an instrument equipped with a grating having a radius of curvature of 480 cm and a ruling density of 6000 lines/cm is only **7** Å/mm. This only surpasses the dispersion of Littrow-system quartz spectrographs when the wavelength is longer than 3600 Å.

In summary, Wadsworth-system spectrographs are often used and are quite popular for spectrochemical analysis. It is an advantage that the spectrum arises stigmatically, so that step filters can be used for their illumination.

References

1. WADSWORTH, F. L. O., *Astrophys. J.*, **3**, 54 (1896)
2. SAWYER, R. A., *Experimental Spectroscopy*, 2nd edn, New York (1956)
3. NACHTRIEB, H. N., *Principles and Practice of Spectrochemical Analysis*, New York–Toronto–London, 75 (1950)

3.10.6 High-performance spectrophotometers [1–5]

In the concave-grating instruments previously described the spectrum is photographed. High-performance spectrometers, on the other hand, are concave-grating instruments in which 'exit slits' placed along the Rowland circle separate the rays representing particular wavelengths from the rest of the spectrum. These automatic devices are built in such a way that concave mirrors project the rays coming from the slit onto photomultipliers.

For the above purpose only relatively high-dispersion devices built on the Paschen–Runge system and using a large Rowland circle can be used. To save space, in the newer equipment the light coming from the slit illuminates the grating via a plane mirror, instead of directly. In this construction the distance between the slit and the plane mirror should be chosen such that the slit is 'optically' on the Rowland circle (Fig. 3.135).

If the temperature changes, the dimensions of the spectrometer also change, in particular the setting of the exit slits. Any errors arising can be eliminated by using a monitor-system. Here, each group of adjacent exit slits is linked together by a common, dilatation-free frame that can be precisely moved around the Rowland circle. The temperature correction is effected by moving all the slits mounted on the frame until the photocurrent of the photomultiplier mounted behind the monitor slit reaches a maximum value. The frames carrying the exit slits can be set manually

or, in some spectrometers automatically as well. The great advantage of the latter construction is that the equipment need not be installed in an air-conditioned and temperature-controlled room.

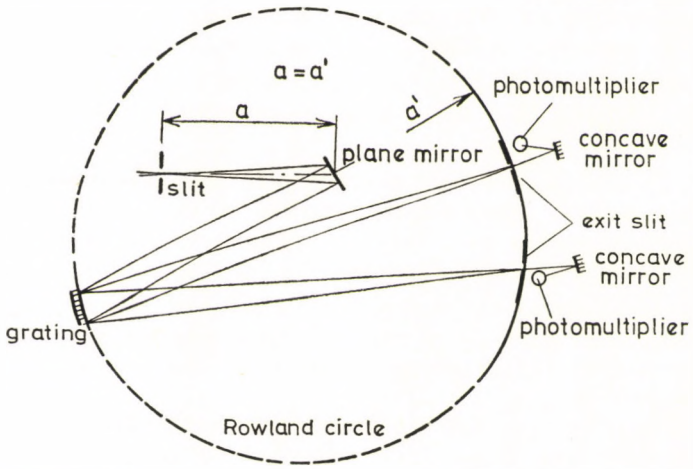


Fig. 3.135 Scheme of a high-performance spectrometer

In another construction the exit slits of the spectrometer are fixed, while the entrance slit can be moved. The disadvantage of this relatively simple solution is that to correct any temperature variation the whole spectrum must be moved as a unit. By collectively moving the exit slits, however, errors caused by variations in the dispersion cannot be entirely eliminated.

References

1. DUFFENDACK, O. S. and MORRIS, M. J., *J. Opt. Soc. Am.*, **32**, 8 (1942)
2. BOETTNER, E. A. and BREWINGTON, G. P., *J. Opt. Soc. Am.*, **34**, 6 (1944)
3. HASLER, M. F. and DIETERT, H. W., *J. Opt. Soc. Am.*, **34**, 751 (1942)
4. NAHSTOLL, G. A. and BRYAN, F. R., *J. Opt. Soc. Am.*, **35**, 646 (1945)
5. DIEKE, G. H. and CROSSWHITE, H. M., *J. Opt. Soc. Am.*, **36**, 192 (1946)

3.10.7 Concave-grating vacuum spectrographs with normal spectrum

The grating spectrographs described above are not applicable for the study of areas of the spectral region having wavelengths shorter than 2000 Å. Air absorbs far-ultra-violet radiation and also normal diffraction gratings are of little use in this region.

As mentioned before, optimum conditions are achieved for a diffraction grating if the grating constant is somewhat larger than half the longest wavelength of the spectral range under study. For far-ultra-violet studies, therefore, diffraction gratings with a very small grating constant must be used. A grating with a ruling density larger than 1800 lines/mm can only be manufactured with the greatest technical difficulty.[1] The gratings applied in the visible and medium-ultra-violet spectral ranges,

which have a ruling density of 600 and 1200 lines/mm, respectively, are in fact about a hundred times coarser than those required.

It should also be taken into account that the radiations coming from sources in the far-ultra-violet region are generally of low intensity. Therefore, it is advisable that the rays illuminate the grating at right angles and not obliquely. It should also be taken into consideration that at wavelengths longer than 1000 Å aluminium has the best reflectivity among the available grating materials, while at wavelengths below 1000 Å, for normal incidence, the reflectivity of speculum metal is 5·2%, glass 8·5%, and quartz 13·5%. In principle, therefore, the application of the latter should produce the best results.[2–5] From the manufacturing point of view, however, it is considerably easier if a glass grating is used to disperse the light of wavelengths shorter than 1000 Å.

The optimum conditions are achieved when the rays illuminate the grating of vacuum spectrographs at angles of incidence close to 90°. This has the advantage that the spectrum is always produced in the neighbourhood of the grating normal. Thus, linear dispersion is independent of the wavelength and astigmatism does not occur.

In principle the optimum case is when, for approximately perpendicular illumination of the grating, the cross-sections of the rulings are isosceles triangles. For perpendicular illumination, however, these gratings reflect only the shorter-wavelength rays to a small extent. Thus, these deeply ruled gratings can only be used for wavelengths between 600 and 500 Å.[6]

On the other hand, shallow-ruled gratings can be applied up to 140 Å.[7–9] The manufacture of shallow-ruled gratings with small grating constants involves various technical difficulties. For example, when the rulings are made it is not always possible to keep the non-ruled parts of the grating surface intact.

In practice the best shallow-ruled gratings are those with a ruling density of 500–1000 lines/mm. One of the disadvantages of shallow-ruled gratings is that they are very sensitive. Also their reflectivity tends to be reduced when the grating is exposed to air for a long period.[10]

The fundamental grating dispersion equation (3.73) can be modified for approximate perpendicular grating illumination to

$$\lambda = \pm \frac{b}{m} \sin \beta \tag{3.109}$$

Accordingly, when illumination is almost perpendicular to the grating, the spectral lines arise progressively farther from the grating normal as their wavelength increases.

In the case of concave-grating vacuum spectrographs with a normal spectrum, the angular difference between the rays falling on and reflected from the concave-grating is very small (Fig. 3.136). Unwanted stray light must be carefully eliminated

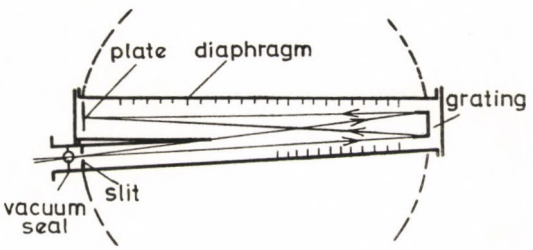


Fig. 3.136 Optical scheme of a concave-grating vacuum spectrograph for a normal spectrum

by means of diaphragms located at appropriate places in the spectrograph. After changing the spectral plate the spectrograph housing should be re-evacuated. Thus, it is advisable that the volume of the latter is as small as possible.

The dispersion of commercial equipment is, for example, 17·0 Å/mm for a grating of radius of curvature of 1 m and ruling density of 590 lines/mm and 4·3 Å/mm for a grating with constants twice those previously mentioned. (In the application of the latter, however, the grating should be rotated through a small angle when the spectral range is changed.)

References

1. SIEGBAHN, M. and MAGNUSSON, P., *Z. Phys.*, **62**, 435 (1930)
2. MILLIKAN, R. A. and SAWYER, R. A., *Phys. Rev.*, **12**, 118 (1918)
3. MILLIKAN, R. A. and BOWEN, I. S., *Phys. Rev.*, **23**, 1 (1924)
4. PHUND, A. H., *J. Opt. Soc. Am.*, **12**, 487 (1926)
5. PHUND, A. H., *Science*, **65**, 595 (1927)
6. ROSENTAL, J. E., *J. Opt. Soc. Am.*, **12**, 487 (1926)
7. MILLIKAN, R. A., *Science*, **50**, 138 (1914)
8. MILLIKAN, R. A., BOWEN, I. S. and SAWYER, R. A., *Astrophys. J.*, **53**, 150 (1921)
9. EKEFERS, E., *Z. Phys.*, **51**, 471 (1928)
10. BOWEN, I. S., *J. Opt. Soc. Am.*, **13**, 89 (1926)

3.10.8 Grating vacuum spectrographs illuminated by grazing incidence light

In the design of vacuum spectrographs belonging to this group it is necessary that the grating should not absorb the rays of the desired wavelength, but should reflect them instead.[1–3] This is achieved when the angle of incidence of the light is larger than a certain limiting value. This value can be calculated, for our purposes, by appropriately modifying the Drude–Lorentz relationship. Accordingly, the following condition should be satisfied

$$\sin \Theta = \lambda \frac{e}{c} \sqrt{\frac{N}{\pi m}} \tag{3.110}$$

where Θ is the angle of reflection (that is, the angle complementary to the angle of incidence α, such that $\Theta = 90° - \alpha$), λ is the wavelength of light, e is the charge on an electron, m is its mass, N is the number of electrons in unit volume, and c is the velocity of light. From the above relationship it can be seen that for total reflection the shorter the wavelength the smaller is the angle of reflection. In the dispersion of far-ultra-violet radiation, it is an advantage if the angle of incidence is close to its maximum value; that is, the grating is illuminated at a grazing incidence.

For grazing illumination of the grating the generally applicable grating equation $m\lambda = b(\sin \alpha \pm \sin \beta)$ must be modified. In Fig. 3.137 R and R' are the slit and its direct image, respectively, while δ is the angle of deviation of a ray of wavelength λ. From the diagram, $\alpha = \beta + \delta$ and $l = 2\delta \,.\, r/2 = \delta r$ for the case when r is the radius of curvature of the grating. Thus, the following relationship can be written

$$m\lambda = b \sin \alpha - b \sin \left(\alpha - \frac{l}{r}\right) \tag{3.111}$$

The above equation refers to a concave grating illuminated at grazing incidence.

By the differentiation of the latter we obtain for the reciprocal linear dispersion

$$\frac{1}{S_{l(cg)}} = \frac{d\lambda}{dl} = \frac{b}{mr} \cos\left(\alpha - \frac{l}{r}\right) \tag{3.112}$$

If the above expression is compared with the reciprocal relationship $d\lambda/dl = (b/mr) \cos \alpha$, given in Eq. (3.94), then the dispersion of the same concave grating is significantly higher with grazing incidence than with approximately perpendicular illumination. The dispersion of the grating is, however, influenced by the wavelength of the light as well. Thus, the linear dispersion of grating vacuum spectrographs illuminated at grazing incidence is not even.

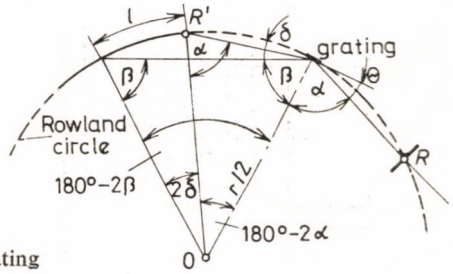


Fig. 3.137 Grazing illumination of a concave grating

From Eq. (3.112) it follows that the spectrum can also be produced between the direct image and the grating. For this negative-order spectrum the dispersion of the grating is at its highest. On the other hand, the resolving power of the grating is smaller than in the case of positive-order spectra. The negative-order spectrum is not used in practice.

The theoretical resolving power of grating spectrographs is: $R_{0(g)} = \lambda/\Delta\lambda = mN$, where m is the order of the spectrum and N is the number of illuminated grating rulings. If it is taken into account that the illuminated width of a grating of constant b is $L = Nb$, then $N = L/b$. According to the general grating equation, $m = b(\sin \alpha - \sin \beta)/\lambda$. Thus, the theoretical resolving power of a concave grating illuminated at a grazing angle is given by

$$R_{0(cg)} = \frac{\lambda}{\Delta\lambda} = \frac{L}{\lambda}(\sin \alpha - \sin \beta) \tag{3.113}$$

It follows from the above equation that for grazing angle illumination the theoretical resolving power of the same concave grating is greater than in the case of approximately perpendicular illumination.

The value of L in the former equation can be altered by partially covering the grating surface. If this is done, however, α and β are also changed, as well as $R_{0(cg)}$. The value of L at which $R_{0(cg)}$ reaches a maximum is called the optimum width. The optimum width depends on the wavelength of light, the angle of incidence, and the radius of curvature of the grating (Fig. 3.138). This theoretical change of the resolving power is, however, practically negligible.

During grazing illumination of the grating so-called 'grazing incidence ghosts' can emerge on both sides of the more intense spectral lines when the illuminated part of the grating is wider than the optimum value. This disturbing effect is more pronounced when the wavelength of the spectral line is short and the angle of incidence of the rays illuminating the grating is large. It follows from the previous

discussion that this effect does not emerge if both edges of the grating are sufficiently covered. (Ghosts only appear on one side of the spectral line when the grating is illuminated from one side or unevenly.) The partial illumination of the grating leads, however, to a reduction in the resolving power, such that excessive masking of the grating is a disadvantage. The optimum width of the grating depends on the wavelength of light, which increases with the latter (Fig. 3.139).

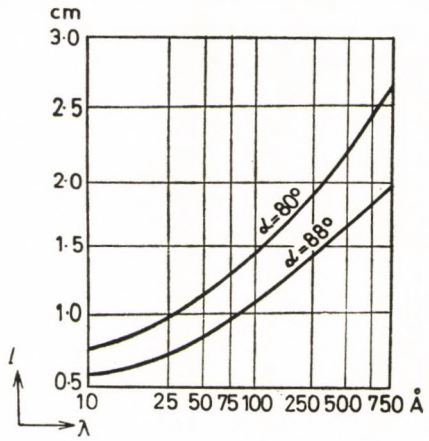


Fig. 3.138 Optimal grating width for a concave-grating of radius of curvature 1 m, and a ruling of 600 lines/mm, at angles of incidence of 80° and 88° [4]

The maximum, optimum resolving power of vacuum spectrographs based on the above principles depends on the slit width – in addition to the other factors mentioned – such that the following relationship holds

$$s_{\text{opt}} = r/L_{\text{opt}} \tag{3.114}$$

The optimum slit width, s_{opt}, therefore, increases with the radius of curvature of the grating, r, while L_{opt} decreases with the optimum grating width. In principle, there is a definite advantage in keeping the slit width at a minimum level. In practice, however, slit widths smaller than 0·001 mm cannot be used (owing to diffraction phenomena) even though the theoretical optimum slit width may suggest a smaller value than this.

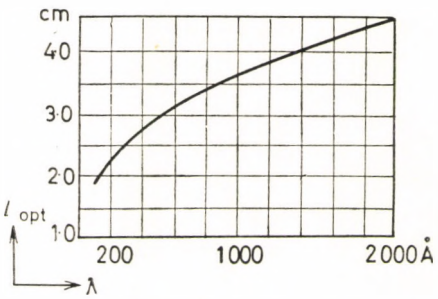


Fig. 3.139 Optimal grating width of a concave grating at grazing incidence for a grating of radius of curvature 1 m, and a ruling of 600 lines/mm

Astigmatism is very large for grazing illumination of the grating. If the length of the rulings illuminated is given by l, the angle of incidence by α, and the angle of reflection by β, then the slit will appear as a straight line whose length is given by Eq. (3.100). It can be calculated that for this degree of astigmatism the maximum

speed of the spectrograph is realised when the smallest slit height is given by

$$h_{s,\,min} = l\,(\tan\alpha \sin\alpha + \tan\beta \sin\beta)\cos\alpha \qquad (3.115)$$

It can be concluded from the above that the optimum dimensions of the slit depend not only on the length of the grating rulings but also on the angles of incidence and reflection, or, indirectly, on the wavelength as well. The calculation of the optimum dimensions is, however, rather complicated. The slit height of commercial instruments cannot, therefore, be altered.

Grazing-illumination grating vacuum spectrographs are usually applicable for the study of the spectral range 60–2000 Å. The spectral range can be simply altered by placing a spectral plate of the appropriate sensitivity on that part of the Rowland circle where the spectral range to be photographed emerges. The free opening of the diaphragm, placed in front of the grating, is also set to the value appropriate for the wavelength of the spectral range required (Fig. 3.140).

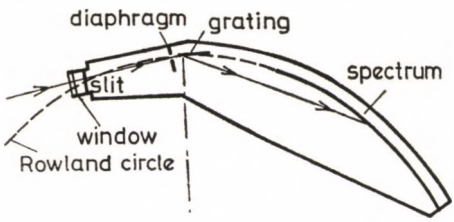


Fig. 3.140 Optical scheme of a grating vacuum spectrograph illuminated by grazing incidence

The dispersion of concave grating spectrographs with grazing illumination is about 3 Å/mm at 60 Å, and about 8 Å/mm at 2000 Å, provided the usable area of a concave grating with a radius of curvature of 1 m is 50×40 mm, or the ruling density on the glass surface is 600 lines/mm and the angle of incidence is 82°. The length of the spectrum between 60 and 2000 Å is about 37·5 cm.

References

1. MILLIKAN, R. A. and SAWYER, R. A., *Science*, **50**, 138 (1919)
2. HAAG, J. B., *Astrophys. J.*, **66**, 225 (1937)
3. COMPTON, A. H., *Phil. Mag.*, 6th ser., **45**, 1131 (1923)
4. EDLEN, B., MARK, J. E. and STEHN, J., *J. Opt. Soc. Am.*, **22**, 245 (1933)

3.11 Errors of grating spectrographs

3.11.1 The overlap between spectra of different orders

When grating spectrographs are used, one disadvantage is that spectra of different orders may partially or totally overlap. According to the general grating equation (3.74), the wavelength of a spectral line appearing at a certain place in the spectrum is not equivocal (Fig. 3.141). This situation, however, can be easily eliminated in practice by using an appropriate filter or photographic plate sensitised for the spectral range required.

In Fig. 3.141 it can be seen that the 2000–4000 Å portion of the first-order spectrum – the most important range for spectrochemical analysis – is not disturbed

by spectral lines of foreign orders. On the other hand, in the region between 4000 and 6000 Å the spectral lines of the second-order spectrum with wavelengths between 2000 and 3000 Å also appear. In the region between 6000 and 8000 Å, not only second-order but also third-order lines can be observed. Moreover, the 8000–10 000 Å region of the first-order spectrum is disturbed by fourth-order lines in addition to those of the second and third orders. No significant improvement is achieved when a photographic emulsion is used that is only sensitive to rays with a wavelength between 2000 and 4500 Å (Fig. 3.141). On the other hand, the portion of the first-order spectrum with wavelengths longer than 4500 Å is undisturbed if a filter absorbs rays with wavelengths shorter than 4500 Å. It is evident from these examples that by the combined use of an appropriate filter and photographic emulsion any portion of the first-order spectrum can be rendered free of unwanted lines.

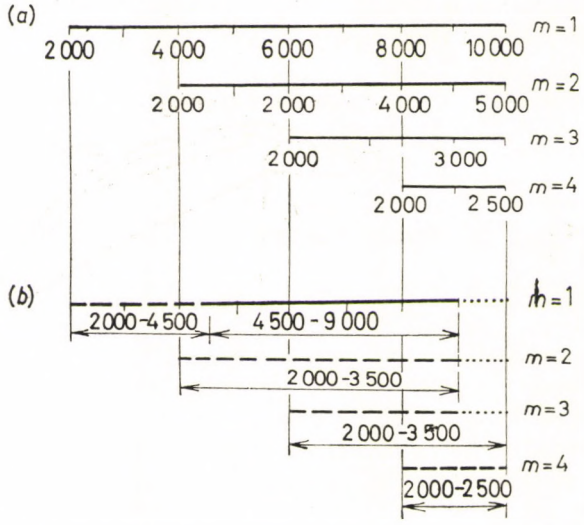


Fig. 3.141 Overlap among different spectral orders

It is due partly to the previous fact that concave grating spectrographs used for spectrochemical analysis have a design that mainly allows the study of the first-order spectrum. As the spectral order is increased there is a progressive increase in the degree of interference caused by spectra of different orders. When the order number is increased, the dispersion and resolving power of the diffraction grating is also increased, although the degree of increase slowly drops. The smaller the difference between the wavelength of a line in a certain mth-order spectrum, λ_m, and the coinciding line in the $(m + 1)$th-order spectrum, λ_{m+1}, the higher is the order number of the spectrum (Fig. 3.142). For example, a line of wavelength 3300 Å in the tenth-order spectrum coincides with a line having a wavelength of 3000 Å in the eleventh-order spectrum, while the 6000 Å wavelength line in the first-order spectrum is overlapped by a line of wavelength 3000 Å in the second-order spectrum.

In the case of high-order spectra produced by simple gratings, the error arising from spectral order overlap is reduced by the fact that the light energy in the higher-order spectra is significantly smaller than the first order. Thus, lines belonging to spectra whose order is higher by an order of magnitude of three or four have such a low intensity that they do not interfere. The gratings that project the incident beam in a direction that corresponds to the 'blaze' angle of the first-order spectrum

are therefore particularly useful in analysis. Interference effects caused by foreign-order spectral lines can be largely eliminated by the selection of an appropriate filter and photographic emulsion, or by an order-sorter for stigmatic line-formation.

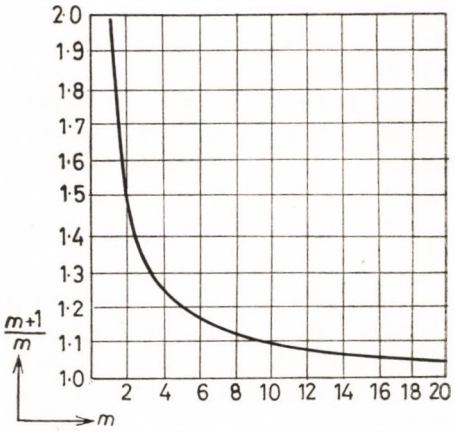


Fig. 3.142 Variation of the ratio $(m + 1)/m$ as a function of the mth-order number of the spectrum

3.11.2 Ruling errors of diffraction gratings

The ruling of diffraction gratings is only perfect when the rulings are perfectly straight and parallel, or, in the case of concave gratings, when the distance is constant between the rulings when they are projected onto a surface. For plane gratings the distance between the rulings must also be constant. Furthermore, it is necessary that the form and the depth of the rulings are of a uniform character. All kinds of deviations reduce the resolving power of the grating and allow the emergence of so-called ghosts.

It can be generally stated that anomalies appear during dispersion if the rulings are not parallel or when the distribution of the rulings is not even. (This error makes it difficult to determine the wavelength of the spectral lines.) On the other hand, astigmatism may arise in otherwise stigmatic spectrographs if the rulings are uneven.

The above ruling errors can be established with relative ease and any inaccuracies produced can be easily eliminated as well. However, in grating manufacture the requirement that the unruled parts between the rulings should be perfectly straight is more difficult to achieve. Such faults may even occur when the gratings are produced with the utmost care. Thus, gratings may differ even when produced by the same ruling machine.

Fig. 3.143 Rowland grating ghosts in the Balmer series of hydrogen spectrum



The statistically distributed, stochastic faults of the grating rulings do not generally produce problems. These faults usually enhance the intensity of the stray light. The scattered light produced in this way appears beside the spectral line even when the grating is illuminated by monochromatic light.

Periodic faults of diffraction gratings give rise to the Rowland ghosts. In this type of grating fault symmetrical ghost-lines appear on both sides of the spectral line (Fig. 3.143). The intensity of these lines is significantly smaller than that of the spectral

lines proper. The intensity difference is progressively reduced as the order number of the spectra increases.

The position of the Rowland ghosts can be calculated. Their wavelength is given by

$$\lambda' = \left(1 \pm \frac{m'}{mN}\right)\lambda \tag{3.116}$$

where λ is the wavelength of the original spectral line, m is the spectral order it belongs to, m' is the spectral order of the ghosts, and N is the total number of grating rulings.[1] From the previous equation we obtain

$$\lambda' - \lambda = \pm \frac{m'}{mN}\lambda \tag{3.117}$$

Thus the difference between the wavelengths of the true spectral line and the ghosts increases with the order of the ghost and also with the wavelength of the original spectral line, while it drops with the order number of the latter.

It can also be shown that the ratio of the intensity of the ghost line, I', and that of the original line, I, is given by

$$\frac{I'}{I} = \frac{m^2}{4}\left(\frac{2\pi e}{b}\right)^2 \tag{3.118}$$

where e is the distance between the periodically occurring ruling errors. The relative intensity of the ghosts increases with the square of the spectral order numbers.[5]

When modern diffraction gratings are used Rowland ghosts do not normally cause any trouble. Due to refinements in the construction of ruling machines the intensity of these ghost lines does not exceed 2% of the intensity of the original line. In good-quality gratings the relative intensity of the ghosts is at most only 1%.

Ruling errors of diffraction gratings produce ghosts not only in the immediate vicinity of the spectral line. Ghosts may also appear in those parts of the spectrum whose wavelength is a certain fraction of the wavelength of the original spectral line, for example: 2/5, 3/5, etc. These are the so-called Lyman ghosts,[2] whose intensity is very small and is generally less than 0·1–0·01% of the value of the original spectral line. They are thus of little importance. They can, however, be a source of trouble if the spectra have a low intensity, as in the case of the far-ultra-violet or infra-red regions.

The occurrence of Lyman ghosts can be traced to the additive effect of two periodic ruling errors. The existence of the Lyman ghosts can be easily studied via a spectrum of low intensity. Here, Lyman ghosts occur at regular intervals throughout the spectrum as groups of lines that otherwise would not be found in the spectrum concerned.[3,4] The good-quality gratings are always examined from this point of view and if free from such errors are labelled: 'Free of Lyman ghosts'.

References

1. ANDERSON, J. A., *J. Opt. Soc. Am.*, **6**, 434 (1929)
2. LYMAN, TH., *Phys. Rev.*, **12**, 1 (1901); **16**, 257 (1903)
3. RUNGE, C., *J. Opt. Soc. Am.*, **6**, 429 (1922)
4. RUNGE, C., *Ann. Phys.*, **71**, 178 (1923)
5. CLARK, G. L., *The Encyclopedia of Spectroscopy*, New York–London, 255 (1960)

3.11.3 Light losses and scattered light

In grating spectrographs light losses occur mainly at the grating. When the latter is a normal type and not an echelette type, light losses arise partly from the fact that some of the light falling upon the grating remains undispersed, as the zero-order spectrum, and is thus wasted. On the other hand, the dispersed light is distributed among the different spectral orders. With respect to the latter it can be established that the intensity drops with the spectral order according to the series: 1, 1/9, 1/25, ... Theoretically, therefore, when higher-order spectra are used, one should take into account that a significant amount of light will be lost, unless a type of grating is used that has a special ruling that reflects the light in the direction of the higher-order spectrum in question. This type of grating is the echelette grating.

In addition to the above reasons the extent of the light loss depends also on the surface of the grating and the way it is manufactured. Thus generally applicable numerical values cannot be given. In general, it can only be established that the loss of light is minimal if the covering of the grating is aluminium for the visible and near-ultra-violet regions, and glass for the examination of spectra of wavelengths shorter than 500 Å.

In grating spectrographs light losses can also arise owing to the plane mirror used for the deviation of the rays, or the concave mirror used to collimate the light in stigmatic line-forming spectrographs. The metal mirrors used at present reflect more than 90% of the light. By using these mirrors, therefore, the light lost is not generally significant. It can be generally stated that the photographic speed of good-quality, modern grating spectrographs attains, or in many cases exceeds, that of prism spectrographs with the same *F*-number.

In grating spectrographs stray light is mainly produced by rays reflected from the grating which belong either to the unresolved zero-order spectrum or to a foreign-order spectrum that is not utilised. The error arising from this source can, however, be largely eliminated by using properly located diaphragms or by the design of the spectrograph housing. On the surface of the plane and concave mirrors used in the spectrographs no stray light is produced, since the rays that fall on them are correctly orientated. In good-quality grating spectrographs, therefore, the harmful effect of stray light is significantly less than in prism instruments.

3.11.4 Errors caused by external experimental conditions

Errors can arise in grating spectrographs such that the dimensions of the grating and instrument change owing to various thermal effects. Change in the effective grating size is the more important of the two. Any such change in the grating dimensions causes a change in the grating constant and also in the angular and linear dispersions as well. The latter changes are inversely proportional to those of the grating constant. Owing to changes in temperature, therefore, the spectral lines are moved in such a way that the linear dispersion is reduced with increasing temperature, or, in other words, the dispersion of the spectrograph is increased. The error arising from this source is reduced if the carrier material of the grating has a low coefficient of expansion. From this point of view it is an advantage if glass is used instead of metal as the supporting material. The expansion coefficient of optical glasses is about one-and-a-half times smaller than that of metals. Moreover, the expansion coefficient

of Pyrex-type glasses, which have a high silica content, is only a third of the coefficient of metals.

When the temperature changes the dimensions of the spectrograph also change, in particular the distances between the slit and the grating and between the grating and the photographic plate. The latter change partly balances the error caused by any changes in the size of the grating. This factor is especially important in the case of large polychromators

It follows from the above discussion that a thermostatically controlled room with a constant temperature should be used during analysis. In particular this applies to automatic analysers set to certain spectral lines, as the error caused by any line-shift is only eliminated when the temperature is constant to $\pm 0{\cdot}1$ °C.

As the atmospheric pressure changes the refractive index of the air varies. This has the consequence that the wavelength of a ray with a definite wavenumber changes as well. In grating spectrographs this leads to a change in the position of a line of wavelength λ by $\lambda/10$ for each mmHg.[1]

The error caused by changes in the atmospheric pressure can be eliminated optically by placing an optically homogeneous quartz plate of refractive index n and width h behind the slit. This plate can be rotated on an axis parallel to the slit.[1] When this plate is turned the parallel shift of the rays is given by

$$\Delta = h \frac{n-1}{n} \tan \alpha \tag{3.119}$$

where α is the angle of rotation. The value of the latter is positive or negative, depending on whether the shift of the rays is to the right or the left. In this case it is a distinct advantage if the rotation of the plate can be performed automatically by some appropriate device.

Reference

1. BEUTLER, H. and FRED, H., *Phys. Rev.*, **57**, 69 (1940)

3.12 Light dispersion instruments used for the study of the fine structure of spectra

3.12.1 General remarks

The performance of the prism and grating spectrographs described cannot be significantly improved beyond the limits already attained. It is not advisable to increase the number and dimensions of the prisms any further, nor can the performance of the gratings be significantly increased either by reducing the grating constant or by increasing the ruling density. Gratings with a large radius of curvature are not really applicable for spectral analysis, as the equipment tends to become too large and thus has a low speed.

The difficulties previously mentioned can be eliminated if the performance of the dispersion equipment is increased by means of a stepped grating or a Fabry–Perot interferometer. Very-high-order spectra are produced by these devices, which can be used for the study of the fine structure of the spectrum. One difficulty, however, is that these dispersion devices can only be used over a very narrow spectral range

by themselves. They can only be applied for analytical purposes when a stigmatic spectrograph is used for the predispersion of the radiation. In this way, a spectrum is obtained in which the spectral lines of different wavelengths are arranged in two dimensions, such that the identification of the lines requires a great deal of practice.

It should be noted, however, that the significance of the two solutions mentioned above is greatly reduced when the commercially available Ebert-system plane-grating spectrographs equipped with blazed gratings are used. The speed of these instruments is so high that even twenty-fifth-order spectra have a sufficient intensity for practical purposes. The resolving power is also sufficient for the analysis of the rare earths.

3.12.2 The echelle spectrograph

The echelle grating or, as it is sometimes called, the reflection stepped grating, is a reflection version of the Michelson stepped grating. A reflection stepped grating consists of a series of quartz plates of the same thickness, joined together in a step-like manner. The front part of the quartz plates is the reflection surface (Fig. 3.144).

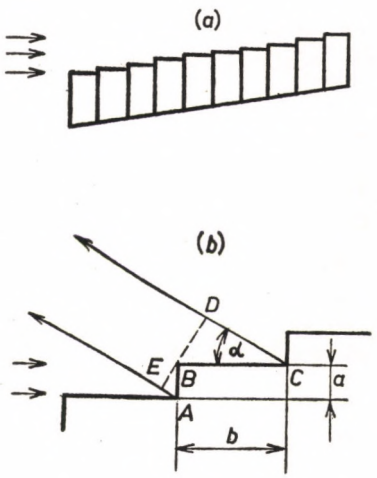


Fig. 3.144 Cross-section of a step grating (*a*) and dispersion developed by reflection from a step grating (*b*)

Under the above conditions the path length difference between the rays reflected in the direction of the angle α from points A and C is $AC = b$ before reflection and $b + CD - AE = b(1 + \cos\alpha) - a\sin\alpha$ after reflection. When it is considered that the two rays reinforce each other when the path length difference is m-times the wavelength, λ, then we obtain

$$b(1 + \cos\alpha) - a\sin\alpha = m\lambda \tag{3.120}$$

This is the fundamental equation of reflection-stepped-grating dispersion.[2]

By using stepped-grating dispersion, a practically useful speed can only be attained when the angle of deviation is very small; that is, when $\cos\alpha \approx 1$ or $\sin\alpha \approx \alpha$. Thus, it follows from the previous equation that the angle of deviation is

$$\alpha \approx \frac{2b - m\lambda}{a} \tag{3.121}$$

and the angular dispersion of the echelle grating

$$S_{a(eg)} = \frac{d\alpha}{d\lambda} = -\frac{2b}{a\lambda} \tag{3.122}$$

The greater the angular dispersion of the reflection stepped grating, the shorter is the wavelength and the greater is the ratio between the height and width of the steps. It is, therefore, an advantage if the series of steps is not particularly steep.

In stepped-grating dispersion the spectra of different orders are produced in such a way that they partially overlap. If the extent of the undisturbed spectral range is denoted by $\Delta\lambda$, and a line of wavelength λ of the mth-order spectrum coincides with the line of wavelength $\lambda + \Delta\lambda$ of the $(m - 1)$th-order spectrum, the following relationship holds: $m\lambda = (m - 1)(\lambda + \Delta\lambda)$. Therefore, the only part of the mth-order spectrum that is undisturbed is when $\Delta\lambda \ll \lambda/(m - 1)$. Furthermore, if we consider the simplified form, $2b \approx m\lambda$, of Eq. (3.120), then we obtain $\Delta\lambda < \lambda^2/2b - \lambda$. The value of λ can, however, be neglected relative to $2b$. The undisturbed part of the spectrum obtained by reflection stepped grating is, therefore, approximately

$$\Delta\lambda < \lambda^2/2b \tag{3.123}$$

The spectral range of extent $\Delta\lambda$ is not disturbed by foreign-order spectral lines if the width of the steps of the stepped grating is

$$b < \frac{\lambda^2}{2\Delta\lambda} \tag{3.124}$$

From the above two relationships we see that the undisturbed part of the spectrum is greatly increased as the wavelength increases. On the other hand, to secure an undisturbed spectral range of a certain extent, the coarser the ruling the longer is the wavelength of the light dispersed.

It can be calculated that the theoretical resolving power of a reflection stepped grating is given by

$$R_{0(eg)} = 2N\frac{b}{\lambda} \tag{3.125}$$

provided b is the height of the steps and N is their number. The theoretical resolving power, therefore, also increases with the number of the steps in stepped-grating dispersion, although its value decreases (unlike that of grating dispersion) on increasing the wavelength of light. When the optical data of an echelle grating used for spectrochemical analysis are taken into account (for example, $N = 40$, $a = 0{\cdot}1$ cm, $b = 1$ cm), we obtain that the theoretical resolving power of the stepped gratings is of the order 10^6. The true resolving power of good-quality gratings also reaches this value.

Furthermore, it is clear from the above that the unperturbed range of reflection stepped gratings is within the range 0·01–0·1 Å. This is the reason why echelle gratings cannot be used alone without some form of order-sorting.

According to the above the echelle spectrograph[1,3,4] is a Littrow-system instrument in which the back of the 30° prism is not a mirror, as the rays are reflected directly by a stepped grating placed immediately behind the prism (Fig. 3.145). Compared with the original Littrow-system spectrograph, the significant difference

is that the slit of the echelle spectrograph is not perpendicular but horizontal, that is, it is parallel to the steps of the echelle grating. The 30° prism deviates the rays according to their wavelength along the length of the spectrum, while the stepped grating effects a deviation in the perpendicular direction. Accordingly, a two-dimensional spectrum is produced (Fig. 3.146).

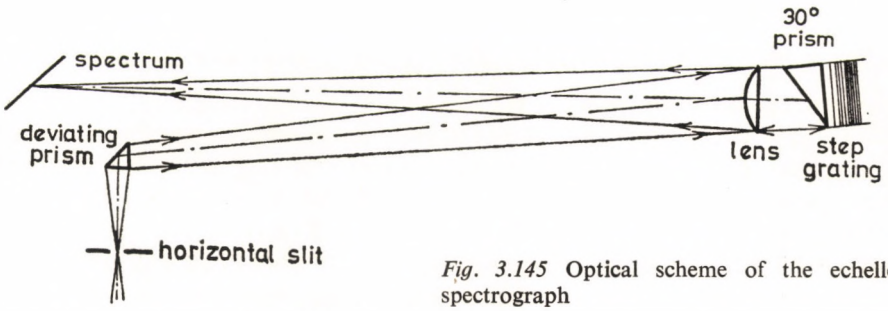


Fig. 3.145 Optical scheme of the echelle spectrograph

The horizontal linear dispersion of the two-dimensional spectrum is determined exclusively by the optical properties of the Littrow-prism and the lens, while the perpendicular linear dispersion depends on the angular dispersion of the stepped grating and the focal length of the lens of the spectrograph. By a proper combination of the latter two factors and the appropriate choice of the slit length it is possible to produce spectral lines in the form of columns lying side by side. Even with this arrangement the difficulty cannot be avoided that in addition to the *m*th-order spectrum, spectra of lower or higher orders are produced, such that certain lines in the spectrum are duplicated. Factors such as this render identification very difficult.

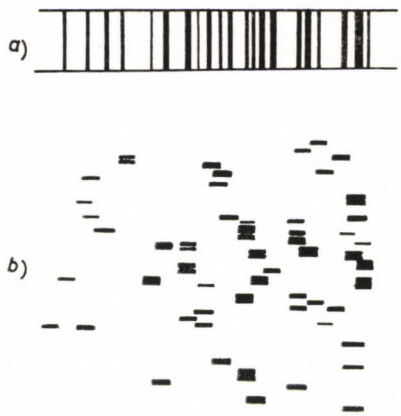


Fig. 3.146 Schematic diagram showing the spectrum obtained by a Littrow-system spectrograph (*a*) and the spectrum obtained by an echelle spectrograph (*b*)

One advantage, however, is that the resolving power of the echelle spectrograph is one order of magnitude higher than that of the high-performance spectrographs used for spectrochemical analysis. Furthermore, the echelle spectrograph has a high photographic speed. This is because the stepped grating reflects more than 50% of the light energy falling upon it. The energy falling on the foreign spectral orders is thus very small.

References

1. HARRISON, G. R. and BAUSCH, C. L., *Proc. of the Conf. on Optical Instruments*, London, 77 (1951)
2. KOHLRAUSCH, F., *Praktische Physik*, 19th edn, Leipzig (1955)
3. HARRISON, G. R., *J. Opt. Soc. Am.*, **39**, 522 (1949)
4. KIRCHGESSER, W. G. and FINKELSTEIN, N. H., *Analyt. Chem.*, **25**, 1034 (1953)

3.12.3 Combination of the Fabry–Perot interferometer and the spectrograph

Modern commercially produced Fabry–Perot interferometers (air-layer type) are made of two circular quartz or glass discs that act as semi-transparent metal mirrors.[1] The spacing between the two opposed, parallel reflecting surfaces is secured by an interchangeable invar metal ring between 2·5 and 50 mm thick. The supporting plates, which usually have a diameter between 20 and 60 mm, are not parallel-faced but form a truncated wedge with an angle of about half a degree between the two surfaces (Fig. 3.147). This slight oblique angle eliminates the error that would otherwise be caused by reflections occurring at the external surfaces of the plates.

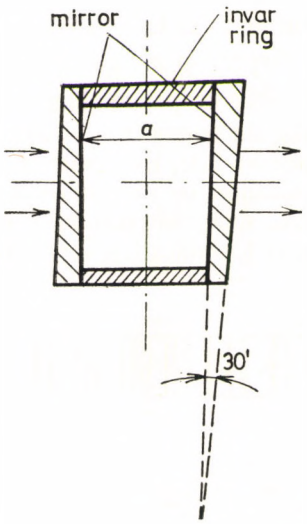


Fig. 3.147 Schematic drawing of the Fabry–Perot (air-layer) interferometer

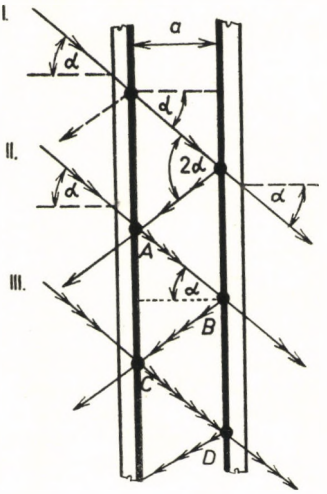


Fig. 3.148 Fundamental principle of the dispersion of light by the Fabry–Perot interferometer

The dispersion of the light is based on the interference of the coherent rays reflected from the two mirror surfaces (Fig. 3.148). When parallel rays are falling at an angle, α, on the interferometer plate then a portion of any ray, I, is reflected from the mirror surface and thus wasted. The part of the ray that enters the air layer travels through to the reflecting surface of the upper plate where some of the light is projected back again onto the reflecting surface of the lower plate. The paths

of rays II and III, etc., are roughly similar. The path length difference between the rays meeting at the nodal points A and B, etc., is therefore $\Delta = AB + CB = 2a/\cos\alpha$, where a is the thickness of the air layer. If the coherent rays leaving the instrument are assembled by a collection lens, then owing to the interference effect a light

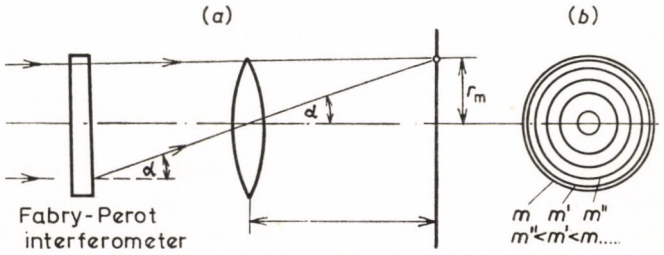


Fig. 3.149 Principle of emergence of the 'constant deviation curve' (*a*), and the ring system produced (*b*)

intensity maximum is produced when the path-difference is a multiple of the wavelength λ. Thus, the fundamental dispersion equation of the Fabry–Perot interferometer is

$$m\lambda = \frac{2a}{\cos\alpha} \tag{3.126}$$

where the integer m is the order number of the spectrum.

The coherent rays leaving the interferometer are parallel. Thus, these rays are collected by the lens along the circumference of a circle lying in the focal plane of the lens. The radius, r_m, of this circle depends on the exit angle of emergence, α_m, belonging to the maximum light intensity of the mth order. From Fig. 3.149 we obtain the following relationship: $\tan\alpha_m = r_m/f$. With respect to the fact that α_m is very small, the following can be written: $\tan\alpha_m = \sin\alpha_m = r_m/f$. Since $\sin^2\alpha + \cos^2\alpha = 1$, it follows from Eq. (3.126), by using the binomial theorem, that the following relationship is established between the mth-order spectrum of a line wavelength λ and radius r_m

$$m\lambda = 2a\left[1 + \frac{1}{2}\left(\frac{r_m}{f}\right)^2\right]$$

It follows from this equation that the radius of the interference fringes increases with the order number in such a way that the difference between the squares of the radii of the adjacent rings is given by

$$r_{m+1}^2 - r_m^2 = \frac{f^2}{a}\lambda \tag{3.127}$$

The squares of the radii, therefore, form an arithmetic series, or, in other words, the concentric interference rings converge the further they are from the centre; that is, in the direction of the lower order numbers.

In the above discussion it has been assumed that the Fabry–Perot interferometer is illuminated by monochromatic light. In practice, however, the light is generally of a composite nature. Owing to interference phenomena in this case, concentric rings are produced that belong to rays of different wavelengths.

The angular dispersion of the interferometer can be calculated via Eq. (3.126). By differentiating the latter we obtain

$$-\sin\alpha d\alpha = -\frac{2a}{m}\lambda^{-2}d\lambda$$

thus, the angular dispersion is given by

$$S_{a(F-P)} = \frac{d\alpha}{d\lambda} = \frac{2a}{m\lambda^2}\frac{1}{\sin\alpha} \tag{3.128}$$

In Fabry–Perot interferometric light dispersion, therefore, the higher the angular dispersion, the larger the layer thickness and the shorter the wavelength of the light. As opposed to the case of grating dispersion, the greater the angular dispersion, the smaller is the order number of the interference ring.

The relationship between the deviation angle α and the radius r of the interference ring can also be calculated by using Eq. (3.126). By the series expansion of $\cos\alpha$ in the formula, in such a way that only the first two members of the series are taken into account, since $\alpha \approx 0$, we obtain

$$\cos\alpha \approx 1 - \frac{\alpha^2}{2} = \frac{2a}{m\lambda} \tag{3.129}$$

When a collection lens with a focal length f is used, the radius of the interference rings is given by

$$r = \alpha f \tag{3.130}$$

If we take into account the change of r^2 relative to the wavelength change $\delta\lambda$, then

$$\delta r^2 = \delta\alpha^2 f^2 \tag{3.131}$$

From Eq. (3.129) we now obtain that

$$\delta\alpha^2 = \frac{4a}{m}\frac{\delta\lambda}{\lambda^2 + \lambda\delta\lambda}$$

In the above relationship, $\lambda\delta\lambda$ can be neglected when compared with λ^2. When the value of $\delta\alpha^2$, obtained in this way, is substituted into Eq. (3.131) we find that

$$\delta r^2 = \frac{4a}{m\lambda^2} f^2\delta\lambda \tag{3.132}$$

As the wavelength changes, therefore, the square of the radius of the interference rings changes proportionally; that is, by increasing the wavelength difference, the radius of the interference rings also increases rapidly.

It follows from the above treatment that if the composite light falling on the Fabry–Perot interferometer has a wavelength difference greater than $\Delta\lambda$ (the so-called working range), then interference rings of different orders will be mixed. The mth-order ring of wavelength λ coincides with the $(m + 1)$th-order ring of wavelength $(\lambda + \Delta\lambda)$, when the following formula applies

$$\frac{m+1}{\lambda} = \frac{m}{\lambda + \Delta\lambda}$$

that is, if $\Delta\lambda = \lambda/m$. Thus, from Eq. (3.126) we obtain

$$\Delta\lambda = \frac{\lambda^2}{2a}\cos\alpha$$

With regard to the fact that the deviation angle is generally very small, so that $\cos\alpha \approx 1$, we can write that

$$\Delta\lambda \approx \frac{\lambda^2}{2a} \tag{3.133}$$

The working range of the Fabry–Perot interferometer, therefore, increases with the square of the wavelength, whereas it drops if the layer thickness is increased.

Generally, the resolving power of Fabry–Perot interferometers is given as one-tenth of the working range. With good-quality plates and various appropriate adjustments, 10 intermediate rings can usually be distinguished lying between the inner and outer rings. As, in this case, $R_{0(F-P)} = (\lambda/\Delta\lambda)/10$, we obtain from Eq. (3.133).

$$R_{0(F-P)} = \frac{20a}{\lambda} \tag{3.134}$$

That is, the larger the layer thickness and the shorter the wavelength, the greater is the resolving power of the Fabry–Perot interferometer.

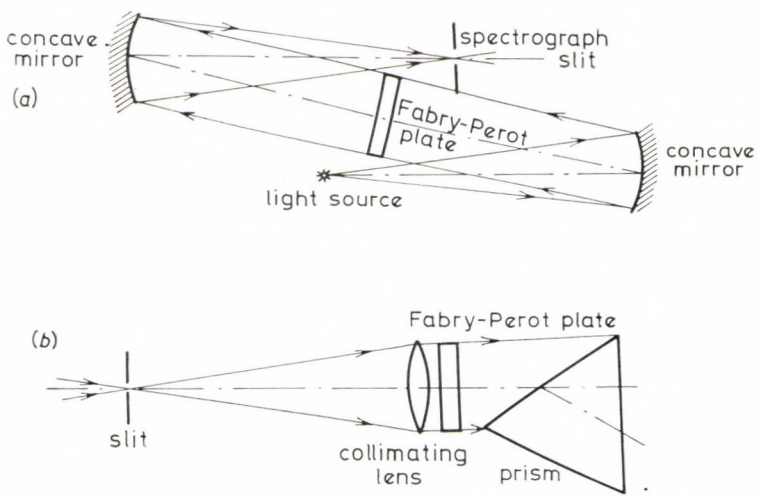


Fig. 3.150 Combinations of the Fabry–Perot interferometer and the stigmatic spectrograph

From the previous relationship, however, it is evident that the Fabry–Perot interferometer can only be used for spectrochemical analysis in combination with a stigmatic spectrograph. The most usual application is when the light of the source is made parallel by a concave mirror (or collecting lens), such that the interference rings are focused by another concave mirror (or illuminating lens) onto the slit of the spectrograph (Fig. 3.150*a*).[2,3] In another arrangement, which is better from the point of view of utilising all the light, the interferometer is placed between the

collimating lens and the prism of the instrument[4] (Fig. 3.150*b*). Thus, the rays falling on the interferometer are made parallel by the collimating lens of the spectrograph, instead of by a separate concave mirror (or collecting lens).

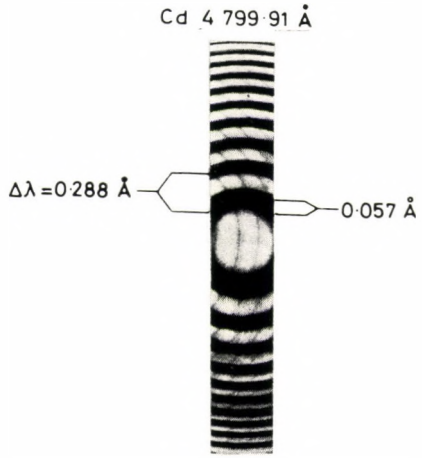


Fig. 3.151 Spectral lines obtained by combination of the Fabry–Perot interferometer and the stigmatic spectrograph

By the combined use of the interferometer and the spectrograph a spectrum is obtained with horizontally striped lines. From the intensity relationship of the stripes, deductions can be made concerning the intensity of the two types of radiation. The distance between the stripes increases with the wavelength of the lines (Fig. 3.151). At small wavelength differences that cannot be resolved by the spectrograph the spectral lines are doubly striped.

Unfortunately, the speed of the Fabry–Perot interferometer is slow, although the resolving power is in the order of 10^6.

References

1. KOHLRAUSCH, F., *Praktische Physik*, 19th edn, Leipzig (1955)
2. MEISSNER, K. W., *J. Opt. Soc. Am.*, **31**, 405 (1941); **32**, 185 (1942)
3. JACQUINOT, P., *Rep. of Progr. Phys.*, **23**, 267 (1960)
4. CLARK, G. L., *The Encyclopedia of Spectroscopy*, New York–London, 258 (1960)

3.13 Dispersion of light by filters

3.13.1 General considerations[1]

Filters transmit only a certain wavelength region of the composite light falling on them. In spectrochemical analysis there are two reasons why filters can be used. First, if the determination only requires rays of a certain wavelength to be isolated from the composite light emitted. Secondly, filters can be used to eliminate radiation of a given spectral range that interferes with the purpose of the analysis. The first type of use is encountered in flame photometry, the second type in the grating dispersion of light.

Filters are characterised by their spectral transmittance. Its value is given by the ratio of the intensities of the light incident on and transmitted by the filter, that is, I/I_0. This value is usually quoted as the percentage transmittance ($\% T = 100\, I/I_0$).

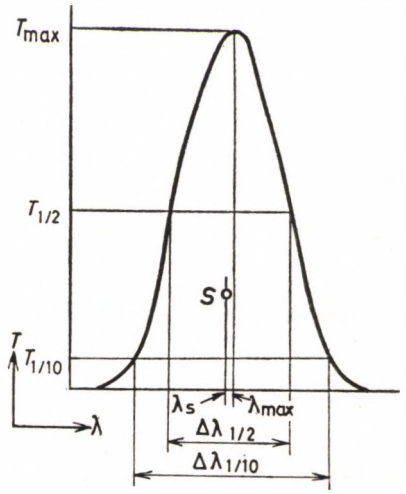


Fig. 3.152 Characteristics of narrow pass filters

The percentage transmittance of commercially available filters usually refers to white light. This is not, however, entirely characteristic of the filter, as the transmittance is dependent on the wavelength. Thus, what is required is the transmittance curve of the filter.

Filters can be classified into two groups, either 'wide' or 'narrow', depending on the form of the transmittance curve (Figs 3.152 and 3.153). The former transmit light through a broad spectral interval, the latter through a narrow interval.

For the characterisation of narrow transmittance filters it is frequently enough if the centre of the transmittance curve distribution is given instead of the full curve. This is the wavelength, λ_s, found at the centre of the area enclosed by the transmittance curve. In many instances, however, the filter is denoted by the wavelength, λ_{max}, at which the transmittance of the filter is at a maximum. In this case the highest percentage transmittance (T_{max}) of the filter is also given. These two figures do not fully characterise the filter, since the width of the filter is not known; this is generally taken at the half-peak height, or, occasionally, one-tenth the height of the spectral distribution (Fig. 3.152). These values are the wavelength differences $\Delta\lambda_{1/2}$ and $\Delta\lambda_{1/10}$, which are delimited by the transmittance curve of the filter at one-half and one-tenth the maximum transmittance level, respectively. These data are characteristic because the transmittance curve of narrow-pass filters is approximately a normal distribution curve. Thus, the distribution centre value of the filter and the maximum transmittance wavelength are usually the same.

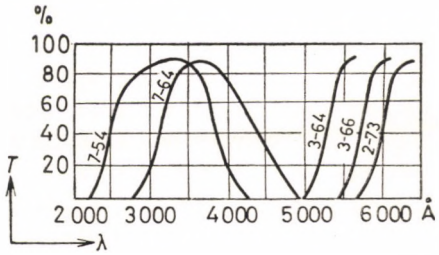


Fig. 3.153 Transmittance curves of some glass filters made by Corning (New York)

Wide band-pass filters have highly variable transmittance curves compared with the symmetrical curves of the narrow band-pass filters. Quite frequently the former absorb radiation at a wavelength longer or shorter than a certain wavelength value; see for example, curve 3-64 in Fig. 3.153. These filters can, therefore, be characterised by either their transmittance curve or their percentage transmittance at different wavelengths.

The different types of filters depend for their operation on various physical properties. Thus, filters used in spectrochemical analysis can be classified into two groups–absorption and interference filters.

Reference

1. KORTÜM, G., *Kolorimetrie-Photometrie und Spektrometrie*, 4th edn, vol. 2, Berlin–Göttingen–Heidelberg (1962)

3.13.2 Absorption filters[1]

The filters belonging to this group are composed of parallel-faced layers of some appropriate light-absorbing medium. For example, the absorbing medium in the visible spectral region is either coloured glass or gelatine, or in some cases a coloured solution.

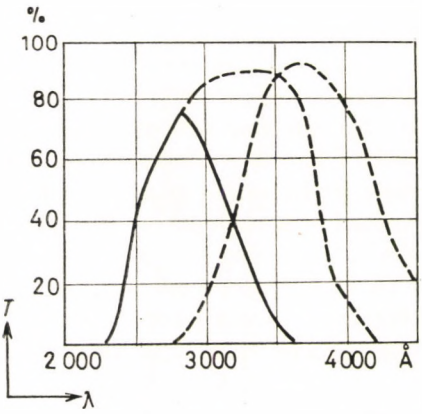


Fig. 3.154 Collective effect of two absorption filters

According to the Beer–Lambert law the logarithm of the absorption, that is, the extinction, E, changes proportionally with the thickness, d, of the homogeneous light absorbing medium, provided the rays fall perpendicularly on the filter. Therefore

$$E_\lambda = \log \frac{I_0}{I} = k_\lambda d \tag{3.135}$$

where k_λ is the *extinction coefficient*. This law only applies to monochromatic light and for a certain concentration range of the coloured material. In general, it can be stated that the lower the transmittance of the absorption filter, the greater is the layer thickness and the concentration of its colouring material (Beer law).

Simple absorption filters, which only contain one colouring material, absorb light so that their transmittance extends over a wide range of the spectrum. These absorp-

tion filters quite frequently only pass light having a wavelength shorter or longer than a certain limit and are, therefore, of the wide band-pass type.

Narrow band-pass filters of absorption type have in their absorbing medium a number of appropriately concentrated colouring materials, each with the correct type of absorption curve. On the other hand, these filters can be produced so that several filters with appropriate transmittance characteristics are placed one behind the other. (For example, by using the filter combination shown in Fig. 3.154, it is possible to make a composite filter with a half-width of about half the value of the individual filters used in the combination.) However, as seen from Fig. 3.154, the selectivity or width of absorption filters can only be reduced at the cost of reducing the percentage transmittance. This effect is emphasised by those absorption filters with a band-pass of only 200–300 Å at half-peak height, having a low transmittance level of 4–5%!

According to the above it is therefore advisable to use only absorption-type filters to select a large spectral interval; as, for example, in grating dispersion to eliminate spectral lines of foreign orders.

Reference

1. KORTÜM, G., *Kolorimetrie-Photometrie und Spektometrie*, 4th edn, vol. 2, Berlin–Göttingen–Heidelberg (1962)

3.13.3 Interference filters[1,2]

These filters are based on the same principle as the Fabry–Perot interferometers. The only difference is that between the two partially reflecting metal mirrors there is a suitable solid-phase optical material (for example, calcium fluoride) instead of air. The width, d, of the latter depends on the maximum transmittance of the filter such that the wavelength belonging to this transmittance value changes proportionally as the filter dimensions change. A further difference is that by means of a suitable optical adhesive, a coloured glass is often attached to the filter. The purpose of this glass is to absorb the radiation produced by high-order interference. Thus, the maximum transmittance of interference filters is given by

$$\lambda_{max} = 2(d + k) \tag{3.136}$$

where k is a constant characteristic of the filter. The above relationship only applies when the parallel rays fall *perpendicularly* upon the filter. In the opposite case, according to Eq. (3.126), we have

$$\lambda_{max} = 2(d + k) \cos \alpha \tag{3.137}$$

where α is the angle of incidence of the rays. The maximum transmittance of the filter, therefore, is moved towards the 'longer' wavelengths as the angle of incidence increases.

Interference filters can be used in the 3700–8000 Å spectral range. Within these limits a filter with any given centre of distribution value and half-width can be produced. The advantage of interference filters is that their maximum transmittance is very high compared with their width (Table 3.10). For example, at a half-width of 220 Å the maximum transmittance of band filters is 60%. The selectivity of the

filters, on the other hand, can only be increased by placing a number of filters of the same distribution centre value behind one another. In this case the relative total transmittance of the interference filters is given by the product of the relative transmittances of the individual filters.

Table 3.10

Average optical data for interference filters made by Schott (Mainz)

	T_{max} (%)	$\Delta\lambda_{1/2}$ (Å)	$\Delta\lambda_{1/10}$ (Å)
broad-pass filter	60	1 000	1 800
band-filter	60	220	400
double-band filter	38	160	256
line filter	40	120	385
double-line filter	20	80	185

References

1. GEFFCKEN, W., *Angew. Chem.*, **60**, 1 (1948)
2. KORTÜM, G., *Kolorimetrie-Photometrie und Spektrometrie*, 4th edn, vol. 2, Berlin–Göttingen–Heidelberg (1962)

4 Determination of the wavelength

4.1 Determination of the wavelength of optical radiation

4.1.1 Fundamentals

The internati n lly accep ed unit of optical wavelength is the international ångström (IÅ). This is equal to 1/6438·4696 of the wavelength of the red line of the cadmium lamp as measured in air of standard condition

$$1\ \text{IÅ} = \frac{\lambda_{\text{Cd red}}}{6438 \cdot 4696} \tag{4.1}$$

where the standardised air is dry, at a temperature of 15 °C, and pressure of 760 mmHg, as measured at a place whose gravitational acceleration is 980·665 cm s^{-2}. Instead of the notation IÅ, for the sake of simplicity the other notation Å is normally used. The relationship between the latter and metric units is

$$1\ \text{Å} = 10^{-8}\ \text{cm} \tag{4.2}$$

Recently the unit nanometer (nm) expressed in metric units has been internationally accepted as the wavelength unit instead of the Å. The relationship between the two quantities is

$$1\ \text{nm} = 10^{-9}\ \text{m} = 10\ \text{Å} \tag{4.3}$$

The former term for the nanometer was the millimicron (mμ).

In the wavelength tables of emission spectral analysis and working instructions the wavelength data are still generally expressed in ångströms.

The wavelength of the spectral lines examined in the spectra under study can be determined only if the relationship between the position of the lines in the spectrum and the wavelength is known. The lines used for this purpose are called *standard lines*. The position of the latter therefore forms the basis of wavelength determination, that is, of the scale.

First-order standard lines are those spectral lines whose wavelength could be determined with a precision greater than ±0·001 Å. These are the lines of different elements. Among these, of course, the most important line is the Cd line of wavelength 6438·4696 Å referred to above.

By means of the first-order standard lines the wavelength of the spectral lines of the spectra examined can only be determined by the *secondary spectra* of some appropriate elements, that is, by means of the spectra of a number of elements. This drawback can be eliminated by the use of *secondary standard lines*. These are spectral lines of a certain element whose wavelength can be determined either by direct measurement or by means of the primary standard lines to an accuracy greater than ±0·01 Å. For this purpose the lines of the noble gases were formerly used. More recently, however, carefully chosen lines of the iron spectrum have been used as secondary standard lines. At present there are a few hundred iron lines whose wavelength has been determined to an accuracy greater than ±0·001–0·002 Å by direct measurement. There is a further advantage in the spectral lines of iron being especially applicable for calibration work, since the secondary standard lines are quite evenly distributed in the easily obtainable spectrum of pure iron (Table 4.1). In the examination of the far vacuum-ultra-violet spectral range, however, it is advisable to use the spectral lines given in Table 4.2 for calibration wavelengths rather than the iron lines.

There are tables in the literature that contain the wavelengths and relative intensity of certain iron multiplet lines (see Section 5.3.2). These are especially carefully chosen lines with wavelength values given to two decimal places and so are quite applicable for wavelength calibration work as well.

The *spectral line tables* give the wavelength of the lines to a precision of at least ±0·01 Å,[2–4] accurate enough for the practical applications of spectrochemical

Table 4.1

The wavelength λ (IÅ)[1] and the relative intensity (I)[2] of the more important secondary standard lines of the Pfund iron arc

λ (IÅ)	I	λ (IÅ)	I	λ (IÅ)	I	λ (IÅ)	I
2 679·062	200	3 677·630	80	4 352·737	300	5 049·825	400
2 778·221	100	3 724·380	200	4 375·932	500	5 083·342	200
2 813·288	400	3 805·345	400	4 383·547	1 000	5 110·414	300
2 851·798	200	3 815·842	700	4 427·312	500	5 127·363	100
2 912·158	150	3 825·884	500	4 466·554	500	5 167·491	700
2 941·343	600	3 843·259	125	4 494·568	400	5 171·599	300
2 987·292	300	3 865·526	600	4 531·152	125	5 270·360	400
3 075·721	400	3 906·482	300	4 547·851	200	5 371·493	700
3 175·447	200	3 935·815	100	4 592·655	200	5 405·778	400
3 225·789	300	4 118·549	200	4 602·944	300	5 434·527	300
3 271·002	300	4 134·681	150	4 647·437	125	5 455·613	300
3 370·786	300	4 147·673	200	4 678·852	150	5 497·519	150
3 399·336	200	4 282·406	600	4 789·654	100	5 506·782	150
3 445·151	300	4 315·087	500	4 859·748	150		
3 513·820	400	4 325·765	1 000	5 012·071	300		

Table 4.2

Internationally recommended spectral lines in the vacuum-ultra-violet range, used for calibrating wavelengths,[1] and their respective relative intensities (I)[6]

Spectral line			Spectral line			Spectral line		
	λ (IÅ)	I		λ (IÅ)	I		λ (IÅ)	I
C I	1 930·897	10	N I	1 492·630	80	N I	1 176·504	15
N I	1 745·249	30	C I	1 481·760	30	N I	1 167·449	25
N I	1 742·730	60	C II	1 335·692	80	N I	1 134·981	25
N II	1 740·321	15	C I	1 329·587	40	N I	1 134·420	25
C I	1 658·123	20	C I	1 329·104	40	N I	1 134·172	20
C I	1 657·900	15	C I	1 277·279	40	N II	1 085·546	3
C I	1 657·378	10	C I	1 261·561	15	N II	1 084·580	30
C I	1 657·001	30	N I	1 200·708	30	N II	1 083·990	20
C I	1 656·260	15	N I	1 177·691	15	O I	990·797	2
N I	1 494·670	60				O I	990·207	8

analysis. By the usual procedures the wavelength of the unknown line can be established only at a lower accuracy level, which generally has an error of 0·1 Å. This is partly because the spectral lines are not theoretical ones but have a smaller or greater width. The accuracy of the determination also depends on the dispersion of the spectrograph; thus, in the evaluation of the spectrum made with a prism spectrograph, the accuracy also depends on the wavelength of the line (Table 4.3). In photographic spectral analysis a further error can be caused by the fact that after development, in the drying process, the photographic emulsion does not contract evenly.

In the practical application of spectrochemical analysis the wavelength of the unknown line is determined on the photographed spectrum either by interpolation or from the dispersion curve of the spectrograph used. In the former case the iron spectrum is used directly, while in the latter it is indirectly utilised in the construc-

Table 4.3

Estimation of the precision of measuring the distance between spectral lines at different wavelengths for medium-performance quartz spectrographs (Zeiss Q 24)[6]

Wavelength (Å)	Dispersion of the spectrograph (Å/mm)	Measurement precision	
		absolute (Å)	relative (%)
2 000	4	~0·1	~0·005
2 500	9	~0·2	~0·008
3 000	15	~0·4	~0·013
4 000	34	~1	~0·025
5 000	64	~1·5	~0·03
5 800	90	~2	~0·035

tion of the dispersion curve. Accordingly the wavelengths of the iron lines are compiled in separate tables, which usually also contain the so-called *iron spectrum atlas* to facilitate wavelength determination.[6–10] These are the magnified spectra of pure iron arc or spark discharges on which the wavelengths of certain characteristic iron lines, or of all visible lines, are given.

In the case of certain spectrographs the wavelength scale can be photographed onto the spectral plate. The photographed scale can, however, be used only for purposes of orientation. This scale indicates the wavelength of the spectral lines to a precision of a few ångströms.

In the visible part of the spectrum, that is in spectroscopes, the wavelength of the spectral lines can be established by means of *measuring drums*. Sufficient precision can be attained in this way only if the measuring drum is calibrated by standard lines.

References

1. *American Institute of Physics Handbook*, New York (1939)
2. HARRISON, G. R., *Wavelength Tables of 100 000 Spectrum Lines*, New York (1939)
3. KAYSER, H. and RITSCHL, R., *Tabellen der Hauptlinien der Linienspektren aller Elemente*, Berlin (1939)
4. ZAIDEL, A. N., PROKOFIEV, V. K., RAISKII, S. M., Spectrum Line Tables, Moscow–Leningrad, 1952 (in Russian)
5. SCHELLER, H., *Einführung in die angewandte spektrochemische Analyse*, Berlin, 28 (1953)
6. GATTERER, A. and JUNKES, J., *Funkenspektrum des Eisens von 4650–2242* Å, Castel-Gandolfo (1935)
7. GATTERER, A. and JUNKES, J., *Arc Spectrum of Iron from 8388–2242* Å, Castel-Gandolfo (1947)
8. SCHEIBE, G. and LINSTRÖM, C. F., *Tabellen des Funken- und Bogenspektrums des Eisens, Teil I, 3700–2300* Å, Berlin–Steglitz (1932)
9. SCHEIBE, G. and LIMMER, G., *Tabellen des Funken- und Bogenspektrums des Eisens, Teil II*, Berlin–Steglitz (1933)
10. SCHEIBE, G. and CORYELL, C. D., *Tabellen des Funken- und Bogenspektrums des Eisens, Teil III*, Berlin–Steglitz (1935)

4.1.2 Determination of the wavelength by linear interpolation

The process is based on the fact that the unknown wavelength λ_x of the spectral line X can be established by interpolation if the line in question is situated in between two such spectral lines A and B whose wavelengths λ_A and λ_B are known (Fig. 4.1*a*). In this case the following relation is satisfied: $(\lambda_B - \lambda_A) : d = (\lambda_X - \lambda_A) : x$, where d is the distance between two spectral lines of known wavelength, and x is the distance between the lines A and X. The unknown wavelength is therefore given by

$$\lambda_X = \lambda_A + (\lambda_B - \lambda_A)\frac{x}{d}$$

In the foregoing it has been assumed that the known and unknown lines belong to the same spectrum. In practice, however, it is rare for two spectral lines of known wavelength to be found in the spectrum studied. This difficulty is eliminated if an iron spectrum is coupled to the spectrum of the material under study, that is, if beside the spectrum of the sample an iron spectrum is photographed by means of

the slit diaphragm, without moving the photographic plate. Thus the lines of the same wavelength form exactly the continuation of one another (Fig. 4.1*b*). Then, in the determination of the wavelength no errors occur if the line of known wavelength is found in the iron spectrum instead of in the spectrum of the sample (Fig. 4.1*b*).

The wavelength can only be determined exactly by linear interpolation if the dispersion is constant between the lines *A* and *B*. This only occurs in the case of the normal spectrum of grating spectrographs. The situation is different with spectra made by prism spectrographs or with *non*-normal spectra of grating spectrographs.

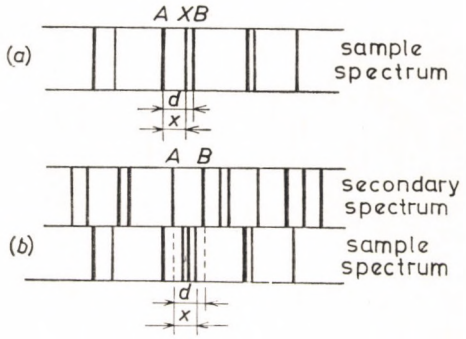


Fig. 4.1 Determination of wavelength by linear interpolation

The dispersion of these instruments changes with the wavelength of light. The linear interpolation technique therefore provides only an approximate result in theory. In the practical application of spectrochemical analysis, however, the process is of satisfactory accuracy if the wavelength difference between the two lines of known wavelength is not great and if the wavelength of the line to be determined does not fall on too steep a portion of the dispersion curve of the spectrograph. In the case of prism spectrographs a greater accuracy can be achieved at shorter wavelengths than at longer ones.

4.1.3 Determination of the wavelength by the dispersion curve

In this procedure the wavelength of the unknown line is determined directly from the spectrum examined. The coupling of a secondary spectrum is therefore not necessary. The dispersion curve of the spectrograph used should be known exactly, or at least the portion of the curve that corresponds to the wavelength of the line to be determined. In addition, the wavelength λ of a line that lies close to the unknown line should be known. On the spectrum the distance between the known line *A* and the unknown *X* should be measured (Fig. 4.2*a*).

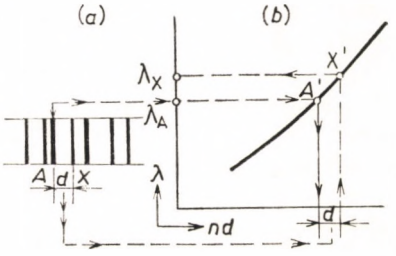


Fig. 4.2 Determination of wavelength by means of the dispersion curve

The wavelength to be found is determined so that the point A' of the dispersion curve is found that corresponds to the wavelength λ_A of the spectral line A, and then it is projected down to the horizontal axis of the diagram (Fig. 4.2*b*). To the value read on the latter the value corresponding to the distance d is added, and on the dispersion curve the point corresponding to the latter (X') is taken. The ordinate value of this point gives the wavelength λ_X sought.

For the construction of the dispersion curve an iron spectrum is needed that was produced by the same spectrograph as used for the spectrum itself. On this iron spectrum the distances d_1, d_2, etc., between the iron lines Fe_0, Fe_1, Fe_2, Fe_3, . . . , are measured. These lines should possibly be at equal distances from one another.

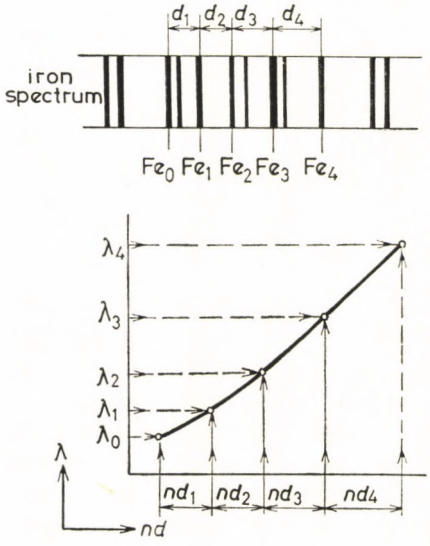


Fig. 4.3 Construction of the dispersion curve by means of the iron spectrum

The wavelengths λ_0, λ_1, λ_2, λ_3, . . . , corresponding to the iron lines chosen are established with the help of the iron spectrum atlas or table. From the data obtained the dispersion curve can be constructed (Fig. 4.3).

The accuracy of wavelength determinations by the method previously described depends largely on the accuracy of distance measurements and the construction of the diagram. The diagram should, of course, be constructed by using a scale that corresponds to the accuracy of the measurements. Then, however, the diagram is too large and not easy to handle even if it depicts only a part of the dispersion curve. For this reason it is better if a dispersion nomogram is used instead of the diagram. On such a nomogram the upper scale gives the wavelength, while the lower one is used for reading the value of the distance as measured in millimetres on the spectral plate.

In the use of the dispersion nomogram the distance d between a line of known wavelength and the unknown line is determined, as measured in millimetres. The latter value is added to the value read from the lower scale of the nomogram that corresponds to the wavelength of the known line. The wavelength corresponding to the sum of the two values, that can be read from the upper scale of the nomogram, gives the wavelength of the unknown line (Fig. 4.4).

In the construction of the dispersion nomogram the dispersion curve is used as starting point. On this curve the points corresponding to whole-wavelength values are found (Fig. 4.5). The perpendicular projection of these points forms the wavelength scale of the dispersion nomogram. The lower, evenly scaled axis gives the distance measured on the spectral plate in millimetres. It is helpful if the zero division of the latter scale coincides with the first whole-number value of the wavelength.

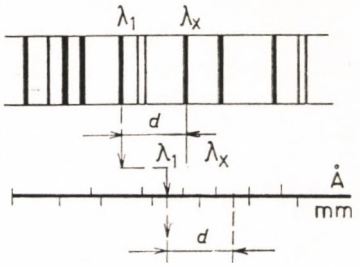


Fig. 4.4 Application of the dispersion nomogram

Both scales should be made in such a way that on the lower axis the distances read from the spectral plate can be established to an accuracy of 0·01 mm. Thus with the help of the dispersion nomogram a determination accuracy equal to that of the linear interpolation can be achieved.

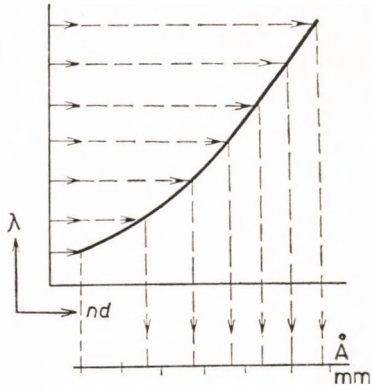


Fig. 4.5 Construction of the dispersion nomogram

In practice, *dispersion tables* are frequently satisfactory for wavelength determination. These give the value of $D_\lambda = \Delta\lambda/\Delta a$, that is, the wavelength difference corresponding to a distance of 1 mm on the spectral plate, valid for a certain spectrograph and for a certain wavelength. In the use of these tables the wavelength of the spectral line in question is given by the expression

$$\lambda_X = \lambda_A \pm a_{X-A} \cdot \frac{1}{S_1} \tag{4.4}$$

where λ_A is the wavelength of the known line, $1/S_1$ is the value of the dispersion found in the tables, and a_{X-A} is the distance measured in millimetres between the two lines in the spectrum.

In all the methods described previously it has been tacitly assumed that the wavelength changes linearly between the known and the unknown lines on the spectral plate. This assumption is, however, allowed in practical spectrochemical analysis

even for prism spectrographs when the distance between the two spectral lines is not so great that it would lead to errors due to the curvature in the dispersion curve.

4.1.4 Determination of the wavelength by the Hartmann formula[1]

As is known, the dispersion of prism spectrographs does not change linearly with the wavelength. The error arising from this source is largely eliminated if the Hartmann relationship is used for the determination of the wavelength. The following formula is a simplified one that is nonetheless satisfactory for the accuracy of chemical spectral analysis

$$\lambda_X = \lambda_0 + \frac{C}{d_0 - d_x} \tag{4.5}$$

where λ_0, d_0, and C are constants, and d_x is the distance of the line of wavelength λ_X from a given location in the spectrum. The only restriction is that the base-point must be on the short-wavelength side of the line, so that when λ_X increases d_x also increases.

The value of the constants found in formula (4.5) can be calculated with the help of the coupled iron spectrum, in which three lines of known wavelength are necessary for this purpose (Fig. 4.6). When the wavelengths λ_1, λ_2, and λ_3 of the iron lines Fe_1, Fe_2, and Fe_3 are known, and the distances of these lines d_1, d_2, and d_3 from a certain location in the spectrum are also known, then

$$\begin{aligned} \lambda_1 &= \lambda_0 + \frac{C}{d_0 - d_1} \\ \lambda_2 &= \lambda_0 + \frac{C}{d_0 - d_2} \\ \lambda_3 &= \lambda_0 + \frac{C}{d_0 - d_3} \end{aligned} \tag{4.6}$$

thus the three unknowns can be calculated from the above three relationships, provided we consider that

$$\lambda_0 = \lambda_1 - \left(\frac{C}{d_0}\right)$$

$$d_0 = a(\lambda_2 - \lambda_0)$$

$$C = \left(\frac{C}{d_0}\right) d_0$$

where

$$a = \frac{d_2 - d_1}{\lambda_2 - \lambda_1}, \qquad b = \frac{d_3 - d_1}{\lambda_3 - \lambda_1}$$

and

$$\frac{C}{d_0} = \frac{d_3 - d_2}{a - b}$$

When the three unknowns have been calculated the application of the Hartmann relationship is very simple. All operations consist of the subtraction $d_0 - d_x$, the division $C/(d_0 - d_x)$ and the addition $\lambda_0 + C/(d_0 - d_x)$. The work can therefore be done with a simple calculator.

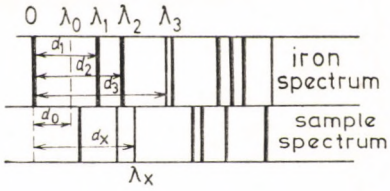


Fig. 4.6 Measurements necessary for the determination of the wavelength according to Hartmann

In the application of the Hartmann formula it should be noted that an error arises if the constants calculated are applied in a wider wavelength range. This is because the previously given relationship does not precisely describe the change of the dispersion of prism spectrographs. The error from this source is eliminated if

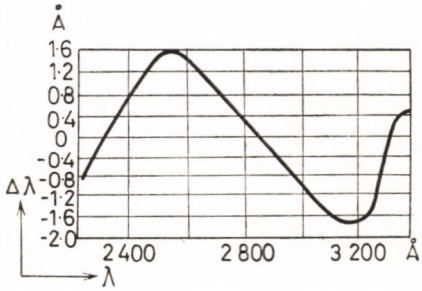


Fig. 4.7 Correction curve for the Hartmann formula

an experimentally established correction is used in calculating the wavelength. This is not a significant drawback if the correction curve relating to the spectrograph in question is prepared in such a way that the wavelength of the known lines is calculated in the above manner. The variation of the difference $\Delta\lambda$ between the value so obtained and the corresponding true wavelength forms the correction curve valid for the spectrograph in question (Fig. 4.7).

Reference

1. HARTMANN, J., *Astrophys. J.*, **8**, 218 (1898)

4.1.5 Determination of the wavelength using the normal spectrum of grating spectrographs

In the case of grating spectrographs the statement is only approximately valid that the angular dispersion and thus the linear dispersion of the normal spectra are independent of the wavelength. When, therefore, the wavelength of the lines needs to be *accurately* determined the simple interpolation process does not give satisfactory results even for normal spectra. This can only be achieved if the necessary correction is also taken into account. In this case the accurate value of the wavelength is

$$\lambda_X = x\frac{\lambda_1}{d} + C \tag{4.7}$$

where λ_1 is the wavelength of a certain line, d is the distance of the latter from the direct image of the slit as measured on the spectral plate, x is the distance between the line of unknown wavelength and the direct image of the slit, and C is the appropriate correction (Fig. 4.8).[1]

For the sake of simplicity let us assume that

$$x\frac{\lambda_1}{a} = \lambda' \qquad \text{and} \qquad \lambda'_n = c = \frac{\lambda_1}{a} \tag{4.8}$$

when c is the distance measured on the spectrum. This distance is determined by the line of wavelength λ_n defined by the grating normal and the direct image of the slit.

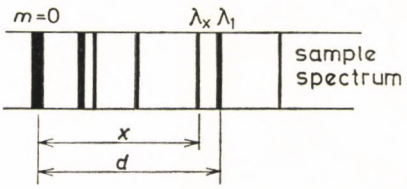


Fig. 4.8 Determination of the wavelength from the normal spectrum of a grating spectrograph

By taking into account these simplifications the value of the necessary correction is

$$C = \frac{\lambda'}{6b^2}\left[(\lambda_1^2 - \lambda'^2) - 3\lambda'_n(\lambda_1 - \lambda')\right] \tag{4.9}$$

where b is the grating constant. The correction therefore depends on the values of λ_1, a, and c in addition to the grating constant. These values can be regarded as constant in the use of a particular spectrograph. On the other hand, the value of λ' is different for each spectral line according to the relationship (4.8).

In spite of the above discussion the process is relatively simple. The C-values corresponding to the different λ'-values need to be calculated once only. The pairs of values obtained in this way can be collected in the form of a diagram from which the correction corresponding to λ' can be read directly.

Reference

1. HUTCHISON, R. O., *Astrophys. J.*, **58**, 280 (1923)

4.2 Optical devices used for the determination of wavelength

4.2.1 General considerations

To determine the wavelength of spectral lines equipment is used by means of which the distance between the spectral lines can be determined easily, quickly, and sufficiently accurately. In order to render the survey work easy the type of equipment required is such that not only the immediate vicinity of the line can be seen but a large area as well.

The sources of error of the above equipment can be grouped into two classes: one is the inaccuracy of measuring the distance between the lines set, and the other is the error of incorrectly setting the spectral lines in question. The first type is an

exclusively technical error, therefore, the inaccuracy can be reduced by technical improvement of the instrument. The accuracy of present equipment corresponds to the requirements of practical spectroscopy for the accurate determination of wavelength.

On the other hand, the error caused by the inaccurate setting of the spectral lines defining the distance to be measured cannot be reduced below a certain limit. The measurement error from this source is due to the fact that the spectral lines are not geometrically straight lines, but have a certain width depending on the fine structure of the lines, the intensity of the rays, etc. Thus, it can be easily shown that the accuracy of setting the spectral lines controls the accuracy of wavelength determination. With modern equipment, however, it is possible to ensure that this source of error does not significantly affect the accuracy of wavelength determinations, even in difficult cases. Also, this error can be significantly reduced – according to the probability laws of error distribution – by repeating the measurements and taking the average of the results.

For the determination of the distance between the spectral lines either a spectral projector or a measuring microscope is used. In practical spectral analysis the use of the former has become widespread.

4.2.2 Spectrum projectors

With equipment of this type the magnified image of the photographed spectrum can be projected onto a (usually horizontal) screen (Fig. 4.9). One advantage of spectrum projectors is that a relatively large portion of the spectrum can be seen

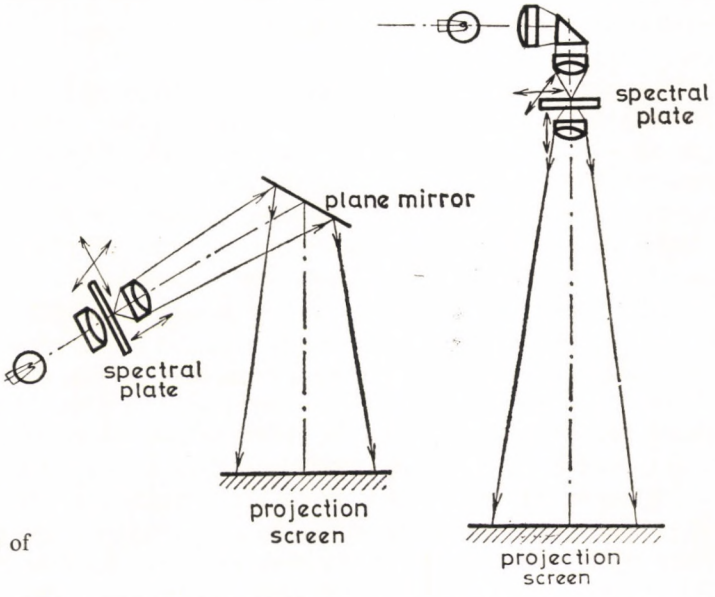


Fig. 4.9 Optical scheme of spectrum projectors

at a twenty-times magnification. This renders orientation much easier. The projected spectrum can be easily surveyed. It is a further advantage that the lines of the projected spectrum can be easily identified with and compared to the lines of the spectrum atlas placed on the screen. From this point of view the best projectors are those

that enable the magnification to be varied between certain limits. Then the two spectra to be compared can be easily examined even if the two spectra were made by spectrographs of different linear dispersions. These advantages account for the fact that in practical applications of spectrochemical analysis spectral projectors are used almost exclusively for the determination of the wavelength of lines.

In the case of different spectral projector systems the distance between the lines can be measured by different methods. In the case of the simple spectral projectors that were generally used in the past the distance on the projected spectrum is measured with a measuring stick. This method can achieve an accuracy of $\pm 0{\cdot}2$ mm in direct distance measurements, while at a magnification level of twenty times the accuracy is at most $\pm 0{\cdot}01$ mm, as measured on the spectral plate. The distance measured on the screen of the projector can only be exactly related to distances on the plate provided the magnification of the equipment is precisely known. A further source of error is that the lens of most simple spectrum projectors is not adequately corrected for spherical aberrations. This results in the 'measured' distance between the same lines being different on different parts of the screen.

The drawbacks of the simple spectrum projector are eliminated in newer equipment in which the spectral plate examined can be shifted in a precisely measurable way by a micrometer screw. With such equipment the distance between the lines can be established so that the spectral plate is shifted by means of the micrometer screw until the image of the two spectral lines in question exactly coincides with the narrow 'index line' drawn on the screen. The difference between the micrometer readings taken is the distance sought. With most spectrum projectors the setting of the measuring drum can generally be read to an accuracy of 0·001 mm. The error of the measurements is, however, always larger than this limit. This is because the image of the lines cannot be set with the naked eye to an accuracy better than 0·001 mm.

The error of setting and thus the error of measurement are reduced to the minimum in those modern spectrum projectors where the setting of the spectral line to the index line is effected by a photoelement instead of by the naked eye. In this case, however, the index line drawn on the projector table is replaced by a slit placed in front of the photoelement. The accurate setting of the line is indicated by the smallest deflection of the galvanometer connected to the photoelement. The setting is therefore entirely objective. The maximum density of the spectral lines generally indicates the centre of the lines, i.e. the place corresponding to the wavelength of the line in the spectrum. It can only be regarded as exceptional when the Eberhardt effect, etc., distorts the form of the spectral lines, so that the mode of setting mentioned previously gives rise to measurement error. A further advantage of the above-mentioned arrangement is that the apparatus can also be used as a line photometer. This type of dual-purpose equipment is, therefore, the most modern and appropriate.

The densitometers at present in use are generally constructed so that the spectral plate being examined can be shifted by a micrometer screw along the spectrum by measurable amounts. These densitometers may therefore be used to determine the distance between the spectral lines without any subjective error. However, a drawback of these instruments is that they show only a relatively small area of the spectrum at a time.

In many cases the wavelength of the spectral lines can also be established by comparing the spectrum to be examined with that of a known element photographed

onto another spectral plate. In the comparison the projected images of the spectral lines of the same wavelength form the continuation of each other if the two spectra are projected alongside one another in a proper way. The instruments used for this purpose are double-spectrum projectors (Fig. 4.10). These devices are well suited to the purpose only if both spectral plates can be shifted in both horizontal and vertical directions, separately and together. It is also necessary that the parts of the spectral plates to be compared are the only parts projected onto the screen.

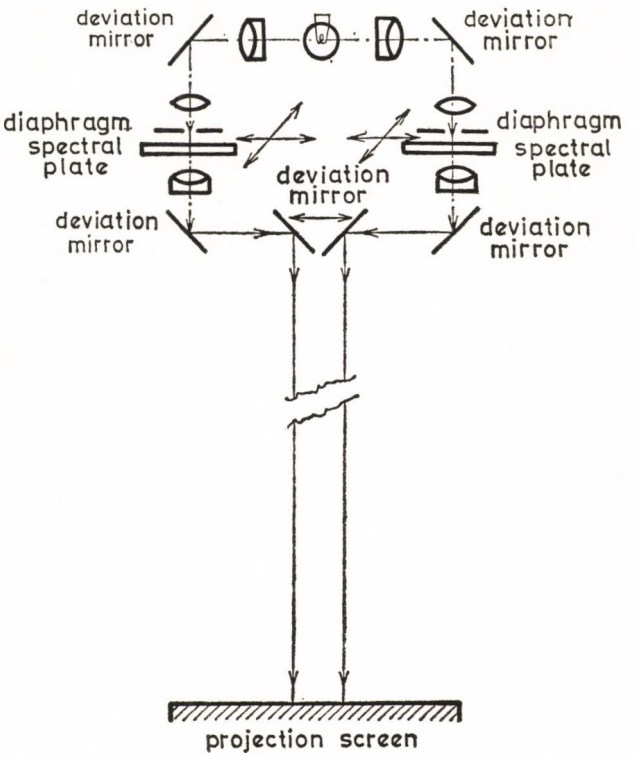


Fig. 4.10 Optical design of double-spectrum projectors

Otherwise, the excess parts of the spectrum affect comparison detrimentally. The excess parts of both spectra should therefore be covered by appropriate diaphragms. It can happen that the linear dispersion of the two spectra to be compared is not quite the same, when all the spectral lines of the same wavelength do not form the continuation of one another. This difficulty may be eliminated if the magnification of the projection system can be appropriately modified, that is, if the last deviation prism of the two optical arrangements (or the appropriate double mirror) can be properly set in the direction of the light. Thereby, one of the light paths is lengthened, and the other is shortened so that the change in the magnification of the projection system will bring about the necessary optical correction.

4.2.3 Measuring microscopes

By means of the measuring microscopes only a relatively small portion of the spectrum can be examined, usually at a magnification of 8–20 times. The viewing field of the measuring microscopes is small – as with all microscopes. This fact renders orientation of the spectrum difficult and therefore ineffective. A further drawback of the measuring microscopes is that they cause eye fatigue when working with them for any length of time. Thus it is understandable that in practical spectral analysis the measuring microscope is not often used.

When the distance between spectral lines is measured with a simple measuring microscope the spectral plate is fastened onto the cross-table mounted on the object-holding table. The setting of the spectral lines can thereafter be done by means of the narrow thread in the eyepiece of the measuring microscope, while the measurable shifting of the spectrum is done by the micrometer screw of the cross-table. The accuracy of the measurements is therefore affected by the backlash of the micrometer screw. It is a further disadvantage of these instruments that they are generally only usable for the determination of a length of 0–50 mm, to a precision of $\pm 0{\cdot}01$ mm.

The Abbé comparators (Fig. 4.11) provide a significantly better measurement accuracy and wider ranges than the simple measuring microscope. These comparators usually serve for the determination of distances between 0 and 200 mm, to an accuracy of about $\pm 0{\cdot}001$ mm. This high accuracy is due to the fact that the distance to be measured is *directly* compared to a very precise mark cut into glass. The advantage of this method is that the unavoidable backlash of the setting screw does not influence the result of the measurement. The accuracy of the measurements depends on temperature variations only to the extent that thermal expansion changes the distance to be measured. The measurement table is made of a metal the thermal coefficient of which is the same as that of glass.

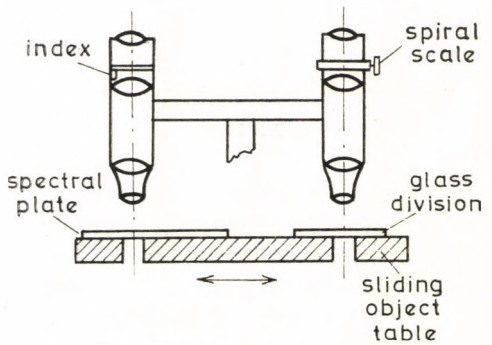


Fig. 4.11 Scheme of the Abbé comparator

In the use of the Abbé comparator the measurements are made by means of two microscopes. The measuring sign of the setting microscope is composed of parallel straight lines at different distances from one another but having the same central line. The alternative system is the 30° thread-cross. In both cases, either by the parallel lines or by the cross forming a comparatively small angle, spectral lines of different width can be delimited very accurately. The error of setting is therefore reduced to a minimum.

The measured distance can be established by the other, so-called *spiral microscope*. The structure of the eyepiece of the latter is constructed so that the glass-division set

can be directly read. Among the numbers visible the relevant one is found in the area limited by the double Archimedean spiral consisting of ten turns (Fig. 4.12). During the measurement the setting screw of the spiral microscope is turned until the next double spiral symmetrically brackets the previous sign. In this position the tenths of a millimetre can be read from the horizontal scale, while the hundredths and thousandths are found on the circular scale. The ten-thousandth parts can be estimated from the latter.

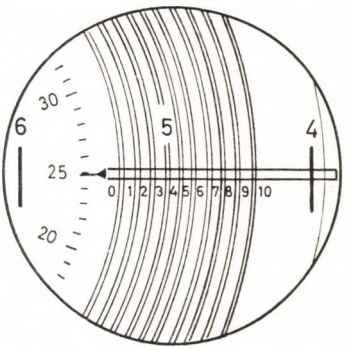


Fig. 4.12 Field of view of the spiral microscope. Value set is 5·3248 mm

In the Abbé comparator the distance measurement is based on the fact that the glass measuring scale is shifted together with the spectral plate fastened to the object table when the latter is shifted by means of the setting screw parallel to the spectral lines. Thus the measured distance is given by the difference of the two values read by means of the spiral microscope.

5 Measurement of the intensity of spectral lines

5.1 Fundamentals of light-intensity measurements

5.1.1 Basic principles[1–3]

The intensity of radiation and of radiation sources can be measured in the same units as those used for other electromagnetic radiations; that is, in watts or calories per second.

The radiation strength (intensity I) of the source is the flux irradiated in the unit solid angle, that is

$$I = \frac{\Phi}{\omega} \tag{5.1}$$

where Φ is the total energy emitted by the source in unit time and ω is the solid angle. The latter may be expressed as a ratio the value of which is the surface area cut from a sphere of unit radius by a cone of angle ω. Therefore, the solid angle corresponding to the total surface of the sphere is 4π. On the other hand, the unit solid angle cuts an area of 1 m^2 from the surface of a sphere of radius 1 m (Fig. 5.1).

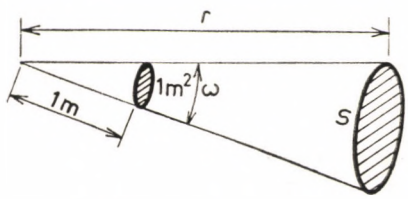


Fig. 5.1 Definition of the solid angle

Taking the above into account the total energy emitted by a point source situated in the centre of the sphere during unit time is

$$\Phi_{to} = 4\pi I \text{ erg/s} \tag{5.2}$$

provided the source radiates homogeneously in all directions. When the radiation is inhomogeneous the above relationship is valid with the modification that the value

of I is that of the average intensity. It should also be noted that in the case of composite light the radiation strength as referred to a certain wavelength – and thus the light flux as well – varies strongly with the wavelength of the light.

It furthermore follows from Eq. (5.1) that the light flux is directly proportional to the radiation strength and solid angle, that is

$$\Phi = I\omega \quad \text{erg/s} \tag{5.3}$$

It follows from this that the light flux falling on a surface of area S at a distance r from a point-like source is given by

$$\Phi = \frac{I.S.\cos\alpha}{r^2} \tag{5.4}$$

where the angle of incidence of light is α (Fig. 5.2), since the solid angle of the radiation-receiving area is: $\omega = S/r^2$; while for angle of incidence we have α, $\omega_\alpha = S.\cos\alpha/r^2$. The light current related to the unit area of the surface exposed to radiation is usually quoted as *irradiation (enlighting) strength.* Thus, its value is

$$E = \frac{\Phi}{S} \text{ erg/s.cm}^2$$

On the other hand, the flux emitted by the unit area of a light source of surface S is called *radiance* or *surface luminosity.* The value of the latter depends on the angle of emission ε according to the Lambert empirical law. The radiance of the source is given by

$$B = \frac{I}{S.\cos\varepsilon} \quad \text{erg/s.cm}^2 \tag{5.5}$$

According to Eq. (5.4) the proportionality

$$E_1 : E_2 = \frac{1}{r_1^2} : \frac{1}{r_2^2} \tag{5.6}$$

holds. The illumination strength E of surfaces at different distances from the source – illuminated normally by the point-like source – is therefore inversely proportional to the square of their distance. In other words: the intensity of the radiaton drops as an inverse function of the square of the distance (Lambert distance law).

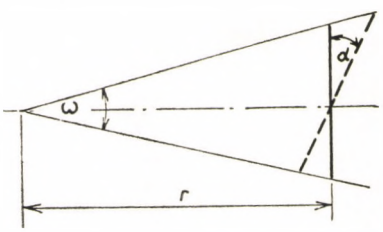


Fig. 5.2 Variation of the light flux as a function of the angle of incidence of light

In the case of oblique illumination the so-called *cosine law* applies, according to which the illumination strength of the surface drops proportionally with the cosine of the angle between the normal of the surface and the direction of the radiation

(Fig. 5.3). Accordingly

$$E_1 : E_2 = 1 : \cos \alpha \tag{5.7}$$

Eq. (5.5) proves furthermore the Lambert law, according to which the radiance of a surface radiating stray light – that is, radiating it evenly in all directions – is independent of the direction of light. This is because the *smaller* the radiance as referred to the apparent surface area, the *greater* is the apparent surface.

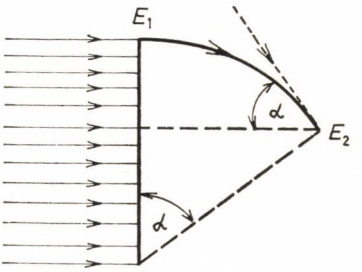


Fig. 5.3 Cosine rule

In practice, in the measurement of the radiation, etc., the so-called physiological (psychological) *units* are normally used instead of the physical ones. This is because the unit of radiance (and thus the other units as well) cannot be established by purely physical means. The physical and physiological notations and units for the related more important terms are collected in Table 5.1.

Table 5.1

Important physical units used in the measurement of radiation

Concept	Physical		Light phenomena	
	term	unit	term	unit
Light source				
energy per unit time: $\Phi = Q/t$	radiant flux	W	light flux	lm
total energy: Q	quantity of radiant energy	W s	quantity of light energy	lms
energy for a unit time and solid angle: $I = \Phi/\omega$	radiant intensity	W/s.r	luminosity	cd
energy for unit time, solid angle and surface area: $L = I/S . \cos \varepsilon$	radiance	W/s.r.m^2	surface luminosity	sb
Receptor				
energy for unit surface area: $H = Q/S$	radiant exposure	W s/m^2	illumination	lxs
energy for unit surface area and time: $E = H/t$	irradiance	W/m^2	illumination strength	lx

The interpretation of the units used in the practical treatments of radiation, and the related units, can be summarised as in the following.

The unit of light flux is the lumen (lm). One lm is the flux value equivalent to 0·001 57 W provided the wavelength is 5550 Å.

The light quantity (Q) of the practically utilised light energy is the product of light flux and time. It is usually measured in lumen hours (lmh). In photography the more usual unit is the lumen second (lms), that is, the quantity of light emitted by the source during one second at unit light flux level.

The unit of the luminosity of the emitting surface is at present the candela (cd). Its value has been established internationally so that the emission from a black body at the temperature of the solidification of platinum (2042·5 K) is 60 cd/cm^2. [The relationship between the candela and the formerly used Hefner candle (HK), and the physical unit is as follows: 1 HK = 0·92 cd = 941 erg/s cm^2.]

The surface luminosity of the emitting surface is measured by the stilb (sb). Its value is one-sixtieth of the radiance of a black body of temperature 2042·5 K, therefore 1 sb = 1 cd/cm^2, while the apostilb (asb) is a fraction of the former, that is, 1 sb = $\pi . 10^4$ asb.

The unit of *illumination strength* is the lux (lx). This is the illumination strength when 1 lumen falls on a surface of 1 m^2, therefore the value of the unit is: 1 lx = = 1 lm/1 m^2. In practice this is an inconveniently small unit. Accordingly, therefore, the flux value for 1 cm^2 is generally used. The unit value of the latter is the phot (ph), that is, 1 ph = 1 lm/cm^2 = 10^4 lx.

The illumination is of unit value when in one second, light of unit luminosity falls on a surface. Its unit value is 1 lux second (1 lx s).

The *mechanical light-equivalent* (M) is the light energy measured in watts that is needed to bring about a unit light flux of wavelength 5550 Å. Its value, as based on measurements, is

$$M = 0{\cdot}00147 \text{ W/lm} \tag{5.8}$$

The reciprocal of the former value is

$$K_m = 680 \text{ lm/W} \tag{5.9}$$

which is the photometrical radiation equivalent. (Both values given previously refer to yellow-green light of the wavelength to which the human eye is the most sensitive.)

Finally, it should be noted that the concept 'light intensity' entrenched in practical spectroscopy is not quite clear-cut. The 'light intensity' (I) used in spectroscopical photometry is, in fact, the flux (Φ) passing through a diaphragm of determined free opening, and its dimension is therefore the watt. On the other hand, the light flux density (S) – which is also called light intensity in practice – denotes the light flux passing perpendicularly through a unit cross-section and its dimension is therefore the W/cm^2. The *light quantity* is, on the other hand, $Q = \int I \, . \, dt$.

References

1. KOHLRAUSCH, K. W. F., *Optik*, Wien (1948)
2. MOHLER, H., *Chemische Optik*, Aarau (1950)
3. D'ANS, J. and LAX, E., *Taschenbuch für Chemiker und Physiker*, 2nd edn, Berlin–Göttingen–Heidelberg (1949)

5.1.2 Procedures used for the determination of the intensity of spectral lines

In the practical applications of spectral analysis it is generally not necessary to establish the light strength of rays of different wavelength in physical or physiological units. It is usually only necessary to know the intensity value of a line relative to that of another line, that is, the *relative* intensity or the *intensity ratio.*

The methods for the determination of the light intensity of spectral lines can be divided into two groups: the *objective and subjective measuring methods.* In the objective procedures the ability of the eye of the individual plays an insignificant role. On the other hand, the accuracy of the subjective methods is strongly influenced by the momentary and personal characteristics of the observer. The latter methods are generally called *visual procedures.*

There is no doubt that the objective methods are more dependable than the subjective ones that depend on personal characteristics. The use of objective procedures, however, necessitates more complicated and more expensive equipment. The objective intensity measuring methods can be grouped into *photographic* and *photoelectrical* methods. For the determination of spectral line intensity three types of objective methods are known:

(1) The *spectrographic* method based on photography.
(2) The *spectrometric* method based on photoelectrical measurements.
(3) The *spectroscopical* method based on subjective measurements.

In *photographic spectral analysis* the intensity of the spectral lines is determined indirectly, by making use of the blackening of the spectral lines. From the blackening value we can calculate the intensity ratio of the rays illuminating the photographic emulsion, that is, the relative strength of the spectral lines. The prerequisite of this procedure is to know the relationship between the blackening corresponding to a certain exposure and the light intensity giving rise to this blackening.

The blackening differences can only be determined by a comparative procedure that is very approximate when carried out by eye. The results obtained thus are only adequate for orientation purposes. Accurate results can only be obtained by photographic procedures when the blackening differences are established by photoelectrical methods with a spectral-line photometer. Accordingly, the photographic spectral analytical methods are further divided into two groups: the photographic semiquantitative methods, and the photographic quantitative methods proper.

For the photoelectrical measurement of light intensity, photoelectric cells – usually photomultipliers – are used, which are placed behind the exit slits of the polychromator. Thus the radiation energy of the rays leaving the slits of the polychromator is transformed into photoelectrical energy that can be *directly* measured. The photoelectrical energy is determined in such a way that the electrical energy given rise during a certain 'exposure' time is collected in a condenser, and the total amount is measured. This way the light-intensity fluctuations arising during the excitation are averaged. Therefore, these variations do not influence the accuracy of the analysis.

In the subjective spectroscopic methods the human eye reacts *directly* to the radiation. The intensity relationship of the spectral lines is established either by comparison or by some appropriate attenuating devices.

It follows from the previous discussion that the light-intensity measuring methods used in spectral analysis are classified as either indirect or direct measuring procedures. The photographic procedures are based on *indirect intensity determinations* while the photoelectrical and visual procedures are *direct measuring methods*. The former is an objective method while the latter is subjective.

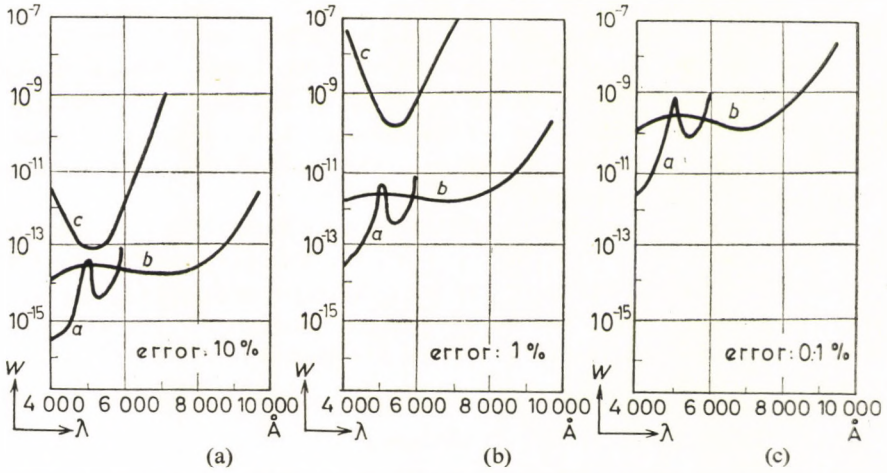


Fig. 5.4 Variation of the sensitivity of (*a*) photographic emulsion, (*b*) photoelectric cell, and (*c*) the human eye, for measurements of different accuracy level in the determination of the intensity of rays of different wavelengths

In the comparison of the performance of the above-mentioned three procedures there are a number of theoretical and practical questions to be considered. Among the theoretical questions the most important are the attainable measurement accuracy and the limits of applicability of the procedure. Numerical comparison is rendered difficult by the fact that the three measuring procedures are based on physically different phenomena. In the spectrographic process the photographic emulsion measures the product of the exposure time and the energy falling on a unit surface area, that is the quantity $t \,.\, Q/S$. On the other hand, in the photoelectrical procedure the electrical energy produced is directly proportional to the light quantity falling on the cathode of the photocell, that is, it is directly proportional to the value of Q. In the case of the human eye it is mainly the radiance, that is, the light energy falling on a unit area with a unit solid angle and within unit time: the value of the ratio

$$H = \frac{I}{S \,.\, \cos \varepsilon} = \frac{Q}{t.\omega \,.\, S.\cos \varepsilon}$$

is that matters. The value of the area S and those of ω and ε can be regarded as given.

It follows from all this that the three procedures mentioned above can be compared most appropriately by their Q-values. In fact, the measuring process is the more sensitive the smaller the value of Q, that is, the smaller the light quantity necessary for attaining the prescribed accuracy in the measurement of the intensity. This accuracy is also a function of the wavelength of the light. The diagrams in Fig. 5.4 are appropriate for the comparison of the performance of the three procedures

most frequently used for measuring the intensity of the spectral lines: the photographic, the photoelectrical, and the visual procedures.[1] From the comparison the following *principal rules* can be deduced:

(1) The *photographic emulsion* has the highest sensitivity, therefore it can detect the lowest light intensity. In order to increase the accuracy of the measurement by a factor of ten, however, it is necessary to increase the quantity of light by a factor of about a hundred. The sensitivity varies greatly with the wavelength of the light.
(2) The sensitivity of the *photomultiplier* is generally somewhat lower than that of the photographic emulsion. The spectral range, however, in that it can be used without large sensitivity variations is significantly wider. In order to increase the light quantity for increasing the analytical accuracy, however, the same factors are valid as for photography.
(3) From the points of view both of sensitivity and of accuracy the *visual* measurement procedure is the worst of all three methods. The performance of the visual method approaches that of the other two procedures only when the tolerated error of intensity measurement is at least 5%. (A measurement accuracy of 1% can only be achieved under favourable conditions.) A further disadvantage of this procedure is that it can only be used in a narrow part of the spectrum, the visible region. The sensitivity and accuracy of the method vary greatly with the wavelength.

In the comparison of the methods for the determination of light intensity the factors considered above, obtained mainly on a theoretical basis, are not the only important ones. The following practical aspects are also important:

In *photographic* spectral analysis it is a drawback that the determination of the light intensity – which forms the basis for quantitative spectral analysis – is not direct, but is indirectly based on the blackening of the photographic emulsion. Most of the sources of error are derived from the faults of the photographic plate. The quantity of silver produced on the plate during the development depends not only on the intensity of light but also on the length of the exposure, the method of illumination, the wavelength of the light, and the experimental conditions of the development procedure. Unavoidable plate defects also contribute to the sources of error previously mentioned. It is also a drawback that the relationship between the light intensity and the density of the emulsion is simple only within relatively narrow density limits. In many cases this drawback cannot be neglected. Also, the photographic procedure is comparatively lengthy and complicated.

In spite of these drawbacks the photographic method is still very widespread. This is because the procedure has very important advantages. The photographic emulsion accumulates and also averages the light energy falling onto it during the exposure. Thereby the analytical error arising from the momentary variations of the light intensity is eliminated. Also, the photographic emulsion is especially sensitive to ultra-violet light, which is most important from the analytical point of view. Furthermore, the photographic records can be preserved and kept for future evaluation.

In the *photoelectrical measurement* of the light intensity the light passing through the exit slit of the polychromator falls directly – or at most via a deviation prism or mirror – on the cathode of the photomultiplier. This solution eliminates many sources

of error. A further advantage is that the photoelectric energy is generally not measured directly but is first collected in a condenser. Thereby, the error due to the intensity variations is eliminated. An appropriate electronic device renders it possible to read the quantity proportional to the light intensity or intensity ratio directly from a scale. The measurement is therefore extremely quick. Another advantage is that in the application of modern photomultipliers the variation of the wavelength of the light does not influence the possibility of measurement in the range most important for spectral analysis, between 2000 and 6000 Å. It is also helpful that the characteristics of the photomultiplier are practically linear.

In the comparison of intensity measurement done by photography and directly with photomultipliers the following facts can be established:

(1) By the *photoelectrical procedure* it is possible to measure the light intensity or the intensity ratio without any later corrections and in a much wider intensity range than by photographic methods. The characteristic curve of photographic emulsion flattens out at low intensity values, which is a drawback. Because of this and because of the detrimental effect of the background, low intensity values can only be established when the measured density values are corrected by a relatively complicated calculation.

(2) By means of modern photomultipliers the light intensity can be measured to an accuracy of $\pm 0{\cdot}1\,\%$. On the other hand, the error sources of photographic emulsions lead to an error of 2–3%, even in favourable circumstances.

(3) In contrast to the previously mentioned advantages of photoelectrical intensity measurement, a marked disadvantage is that the equipment necessary for the measurements is much more complicated and more expensive than that used in the photographic method. A further drawback of photoelectrical instruments is that they can only be used for performing analytical tasks of only a certain number of types, known in advance. On the other hand, the photographic method is more flexible and can be generally applied. Moreover, the photographic emulsion records all details that may be irrelevant to the immediate purpose, but can be important later in the investigation of unforeseen troubles with the analysed material, etc.

By the visual measuring methods the light intensity is established with a significantly greater error than by the objective method. On the other hand, the equipment necessary for the measurement is simpler and less expensive than that necessary for the photographic method. It is also an advantage that the measurements can be done in a very short time. The visual intensity measurement method is therefore preferable for those jobs where speed and low cost are the prime factors and accuracy is of secondary importance.

Reference

1. HANSEN, G., *Mikrochim. Acta* **1955**, 709

5.2 Fundamentals of photographic radiation intensity determination

5.2.1 Fundamentals of photography[1-3]

The fundamentals of photography are formed by the physical and chemical processes taking place within the photographic emulsion. The photographic emulsion at present used for the measurement of light intensity is a gelatine layer containing colloid-size silver bromide particles and also some silver iodide and a few percent of sulphur compounds. Silver bromide is the most sensitive among the light-sensitive halides, while silver iodide hinders fogging.

In the modern theory of photography the lattice defects of silver bromide crystals play a very important role. (This is shown by the fact that if the silver bromide has a perfect crystal lattice without faults it is not discoloured even if it is exposed to daylight for a long period.) The lattice defects are the absence of silver or bromide ions at the appropriate lattice locations, or the occurrence of these ions at irregular places in the crystal lattice. An important lattice defect is when the place of the bromide ion is taken by a sulphide ion.

The photolytic effect of the above-mentioned crystal defects occurs because they bring about free charges at statistically distributed locations. These charges neutralise one another over the whole crystal. Thus in the crystal lattice of silver bromide places having uncompensated charges are found at the locations of the crystal defects. The signs of the charges are as follows:

in absence of silver ion:	negative	with surplus of bromide ion:	negative
in absence of bromide ion:	positive	bromide ion replaced by sulphide ion:	negative

The photochemical processes are initiated at these locations of free charges.

Under the effect of irradiation by light of appropriate intensity and wavelength, and of proper duration, the emulsion containing the silver halides becomes grey. This is due to the precipitation of silver. The prerequisite of this process is, however, that the silver halides should absorb an adequate amount of light energy. The yellow-coloured silver halides absorb only blue and ultra-violet light but do not absorb red light. The photochemical reaction, therefore, cannot occur under the effect of light of longer wavelength.

The quantity of light falling on the photographic emulsion during exposure is not enough to produce a visible amount of silver by 'photolytic action' within the emulsion. Only a latent image is created in the latter. A full image can only be achieved by a subsequent reduction; that is by the *development* process by which the reducible silver causes visible greyness at the location of the latent image.

The above phenomenon used to be interpreted as being due to the silver halide being slightly decomposed under the effect of light, the submicroscopic silver particles forming the latent image playing the role of a catalyst during development. This theory, however, fails to explain a number of points.

The modern theory of the latent image starts from the assumption that the bromide ion built into the crystal lattice of silver bromide yields an electron under the effect of light, thus the following processes give rise to silver atoms

$$Br^- + h\nu = Br + e^- \tag{5.10}$$

$$Ag^+ + e^- = Ag \tag{5.11}$$

Such a small quantity of silver is produced this way, however, that it does not result in a *visible* colour. In an appropriately large silver bromide crystal an adequate amount of light results in a blue colour that is due to an absorption band in the long-wavelength region. This band arises, however, only when the binding of electrons is looser in the crystals, that is, when an irregularity develops in the lattice. The reason for this is that during the process given by Eq. (5.10) the electron liberated by the bromine atom, owing to light irradiations, is replaced by the negative charge of the lattice defect. In the latter process especially the sulphide ions have an important role. The sulphide ion in the crystal lattice of the silver bromide has an excess of electrons.

The larger the amount of silver crystal seeds produced at a certain location of the photographic emulsion, the larger the amount of silver that will be reduced at this place during development, and the greater is the density.

The gelatin layer contains undeveloped silver halide particles following the development. If these were not removed the emulsion would develop a greyness; that is *fogging* would occur. After development, therefore, the silver halides left in the gelatin should be dissolved, that is, the developed image has to be *fixed*. For this purpose sodium thiosulphate solution is the best. Under the effect of the latter a soluble silver thiosulphate complex is formed from the insoluble silver halide:

$$AgBr + 2S_2O_3^{2-} = Ag(S_2O_3)_2^{3-} + Br^-$$

On the other hand, the silver produced during development is practically insoluble, provided the duration of the fixing procedure is not too long. Overlong fixing may give rise to troubles, because during fixing the silver is slightly dissolved so that the density of the emulsion is unevenly reduced. The correct duration for fixing depends on the concentration and temperature of the sodium thiosulphate solution. As Fig. 5.5*a* shows the necessary fixing time drops to only about 30% when the concentration of sodium thiosulphate is increased – which is unexpected; when, however, the concentration is significantly higher, the fixing time starts rising again. This anomaly can be explained by the fact that a too-concentrated solution hardens the

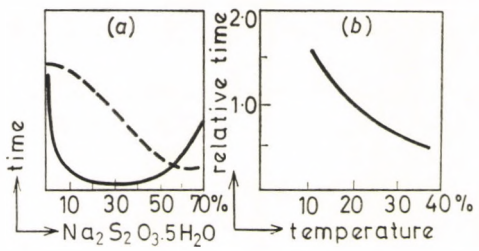


Fig. 5.5 Duration of fixing: (*a*) relation between the concentration of the solution and time (the dotted line shows the influence of concentration upon the swelling of the gelatin); (*b*) relation between the temperature of the solution and the duration of the fixing procedure[1]

gelatin, hinders its swelling and thus renders more difficult the diffusion necessary for bringing about the heterogeneous reaction. On the other hand, as would be expected, the fixing time falls significantly when the temperature of the solution is increased (Fig. 5.5*b*). This factor is reduced in importance by the fact that upon the increase of the temperature the solubility of the developed silver is also increased, and that at temperatures higher than 30° the gelatin layer is easily melted.

The processing of the photographic emulsion is not finished at the fixing stage. The gelatin layer absorbs the by-products of the fixing reaction and the chemicals

of the fixing solution. These compounds retained in the emulsion, especially the sulphur-containing ones, decompose after a certain time and the decomposition products give a brownish colour to the gelatin, thereby spoiling the record. The removal of the decomposition products is therefore absolutely necessary, so the fixed photographic emulsion must be thoroughly rinsed with clean water.

The last stage of the processing of the photographic emulsion, and a very important one, is the *drying*. The gelatin takes up water during the development, etc., and so swells to 10–20 times its original volume. Obviously the drying must be done carefully so that the gelatin layer regains the original condition it was in before processing. When this is not achieved not only the distance between the spectral lines, but also the density of the lines becomes uncertain. With careful drying, however, it is possible to obtain such good results that the differences in the line distances, as measured on the emulsion before processing and after drying, do not exceed $\pm 0{\cdot}001$ mm.

The best method of drying the emulsion is to irradiate it with infra red-light from the side of the carrier (glass or film). At the same time, filtered and preheated air should be passed over the emulsion. Before these operations the small water droplets on the emulsion surface should be gently removed with a plastics sponge of micropore structure. In the infra-red drying process the emulsion is dried 'from inside', as opposed to the case of normal drying, when the contraction of the gelatin surface renders the outward diffusion of water from the deeper layers difficult from the beginning of the drying process. Infra-red drying is not only more even but is also quicker.

The processing of photographic emulsions described above is based on heterogeneous reactions. Thus the reaction rate is strongly influenced by diffusion. The blackening of the photographic emulsion therefore strongly depends on the duration of development, on the way the reagent solution and the gelatin layer come into contact, and on the mode of solvent streaming.

References

1. ANGERER, E. and JOOS, G., *Wissenschaftliche Photographie*, 6th edn, Berlin (1956)
2. MEES, C. E. K., *The Theory of the Photographic Process*, 6th edn, New York (1952)
3. JAMES, T. H., *The Theory of the Photographic Process*, New York–London (1966)

5.2.2 The photographic characteristic curve

When the photographic emulsion is illuminated by light of intensity I for a length of time t and is developed (development, fixing, washing, and drying), a certain degree of blackening is obtained. In order to give a numerical characterisation to the blackening obtained the processed plate is illuminated by light of intensity J_0 and the value of light intensity J after passing through the plate is established. The *photographic transmittance* is the term expressing the ratio of the intensity of incident light to that of leaving light

$$T_P = \frac{J}{J_0} \tag{5.12}$$

The *transmittance percentage* (or percentage transmission) is a hundred times the

above value

$$100 \frac{J}{J_0} = T_P \, \% \tag{5.13}$$

As opposed to transmittance, the D blackening value *(density)** is the logarithm of the ratio of the intensity of incident light to that of transmitted light

$$D = \log \frac{J_0}{J} \tag{5.14}$$

By comparing Eqs (5.12) and (5.14) the relationship between the photographic transmittance and the density can be established

$$D = \log \frac{1}{T_P} = -\log T_P \tag{5.15}$$

In order to establish the intensity or intensity ratio of the spectral lines from the photographic transmittance or blackening (density) value one has to know the relationship between the exposure $I.t$ (light quantity) and T_P or D. As we shall see below (Sub-section 5.5.1) the density caused by the same doses of exposure is not independent of the values of I and t (deviations from the reciprocity law). In the spectrographic evaluation, however, the same time-factor always applies for the determination of the intensity of the spectral lines. Thus in the following treatment we shall deal exclusively with the relationship between light intensity I and blackening for a constant time duration.

The relationship between the exposure or light intensity on the one hand and the blackening on the other is very complicated. The physicochemical processes playing different parts are rather involved and so far not clearly understood. These factors can only be dealt with at best on a rather crude level of approximation.[2–5] Thus the exact relationship between the blackening and the intensity can only be established experimentally.

Figure 5.6 illustrates the graphical relationship between the intensity or its logarithm – the latter is denoted by Y according to international usage[6] – and T_P, $1/T_P$, and D. It can be seen from this that the graph that can be treated most simply is the one depicting the relationship between Y and D. This is because this graph is *linear* from the point Y_L, D_L towards the increasing values, at least up to the upper limit found on the plot. As we have seen in Sub-section 1.3.3 an exponential relationship exists between the analytical concentration and the intensity of the spectral lines, hence this becomes linear in the logarithmic scale ($\log c$, $\log I$). Thus the relationship between the intensity of the lines and the analytical concentration becomes very simple when the linear portion of the plot Y, D is used.

The foregoing may be summarised thus: the most suitable graphical relationship for the purposes of emission spectrochemical analysis is the one between Y and D. This graph is called the *density curve*.

From the density curve shown on Fig. 5.7 it is seen that a very small density arises *without* the exposure of the photographic emulsion. This is the *fogging density* (D_f). On the other hand, a very small exposure does not alter the density of the emulsion.

*Instead of the expression *density* or *photographic density*, IUPAC has recently suggested the term *blackening* (symbol: S) which is increasingly being used.

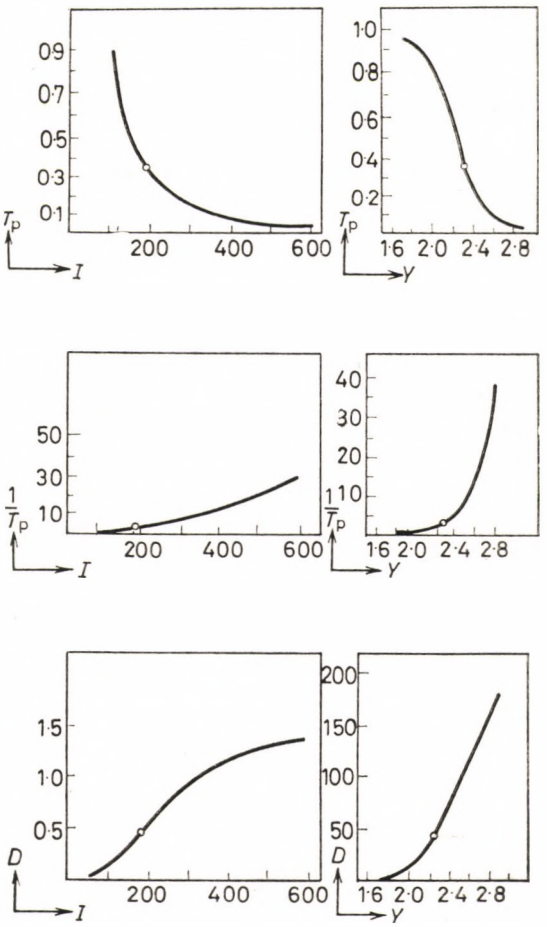


Fig. 5.6 Graphical relationship between I resp. Y, and T_p, $1/T_p$, and D, for the Gevaert Scientia 23 D 56 emulsion at $\lambda = 350$ nm [6]

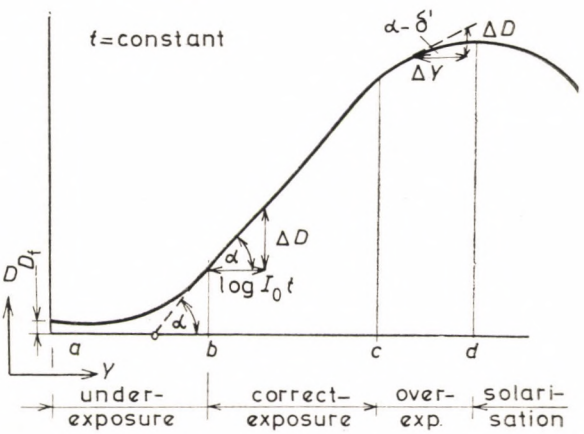


Fig. 5.7 Characteristic curve of the photographic emulsion

This is because an exposure corresponding to a threshold value a is necessary for first a slight, later a stronger increase of the density, until having reached the light intensity b the density change is proportional to the logarithm of the exposure (light quantity). However, this proportionality vanishes after attaining the intensity limit c. Upon further increase of the light intensity the increase in density steadily drops, and above the value d the density is reduced. Thus an S-curve is formed. The previously mentioned parts of the characteristic curve are called by different names: the a–b part is called the *under-exposure*, the b–c part is called the *correct exposure*, the c–d part is named *over-exposure*, while the portion coming after d is called *solarisation*.

The most important part of the characteristic curve is the linear part corresponding to the correct exposure. In this region there is a simple, that is linear, relationship between the *logarithm of the light intensity* and the *blackening* it causes. It is advantageous that on this part of the characteristic curve the largest density difference ($\Delta D_{ab} = D_m$) corresponds to a certain intensity ratio I_a/I_b. To the light intensity relationship $\log I_a/I_b = \log I_a - \log I_b = Y_a - Y_b = \Delta Y_{a,b}$, the corresponding density difference ΔD_m is largest just on the linear part of the characteristic curve.

The photographic emulsion is characterised by that γ-value that is equal to the tangent of the angle between the linear part of the characteristic curve and the axis of exposure, that is

$$\gamma = \tan \alpha = \frac{\Delta D_m}{\Delta Y_{a,b}} \tag{5.16}$$

The intersection point P of the elongated linear part of the characteristic curve with the axis $\log I.t$ is called the *inertia of the emulsion*. This is used for the characterisation of the sensitivity of the emulsion. The value decreases as the sensitivity increases.

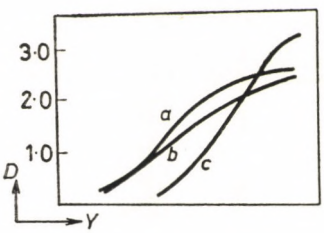


Fig. 5.8 Characteristic curve of emulsions of different contrast: (*a*) normal, (*b*) soft, and (*c*) hard[7]

The slope of the linear part of the characteristic curve is extremely important both for measurement of light intensity and for spectral analysis. The higher the value of γ the greater will be the density difference corresponding to the same change in the light intensity. For the same exposure, therefore, the light-intensity differences will be reproduced by the greater density differences in the photographic emulsion (that is, the contrasts will be the higher) with the higher value of γ (Fig. 5.8). The greater the slope value the higher the contrast of the emulsion: a high-contrast emulsion may also be referred to as 'hard' and a low contrast emulsion as 'soft'. Accordingly, when

$\gamma = 0{\cdot}4$–$0{\cdot}9$ the emulsion has 'soft' or *low* contrast
$\gamma = 0{\cdot}9$–$1{\cdot}5$ the emulsion has 'normal' or medium contrast
$\gamma = 1{\cdot}5$–$3{\cdot}5$ the emulsion has 'hard' or *high* contrast

The emulsion is therefore generally of high contrast when the angle of slope of the linear part of the characteristic curve is greater than 45°.

In connection with the above 'commercial' notations it should, however, be noted that these generally refer to white light. Therefore, the notation (or designation) given corresponds to an average γ-value over the full *visible* spectral range.

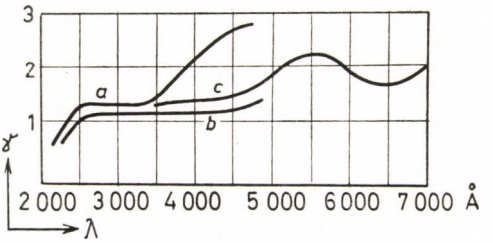


Fig. 5.9 Variation of the γ value of photographic emulsions as a function of the wavelength of light: (*a*) blue-sensitive, hard, (*b*) blue-sensitive, soft, and (*c*) commonly used photographic emulsion[8,9]

The γ-value of photographic emulsions generally varies with the wavelength of light. This variation, however, depends on a number of factors. It depends on the contrast of the emulsion; specifically, this variation of the γ-value is greater for high-contrast emulsions, the best for spectrochemical analysis, than for low-contrast emulsions (Fig. 5.9). All emulsions, however, have a spectral region where the slope of the characteristic curve is independent of the wavelength of light. For the purpose of light-intensity measurements this fact is very important. When the γ-values are identical, then from the density of the different spectral lines we can directly discover the relative intensity of the rays of different wavelength. This useful part of the spectrum is narrower in the case of the high-contrast emulsions than in the case of the lower-contrast ones. Thus, for example, the slope of the characteristic curve of the blue-sensitive emulsions always drops significantly at wavelengths shorter than 2500 Å. Towards the long-wavelength range, however, the slope increase generally occurs at those of shorter wavelengths, the higher the contrast of the emulsion. In the case of the high-contrast blue-sensitive emulsions in general use this increase of the γ-value usually occurs at 3100 Å, or exceptionally at 3300 Å. On the other hand, in the lower-contrast blue-sensitive emulsions this increase occurs at longer wavelengths. The relationship between the logarithm of the exposure and the corresponding blackening is a function of the wavelength of the light. A particular characteristic curve therefore applies only to the illumination of a certain photographic emulsion by light of a certain wavelength and for constant conditions of development.

The straight part of the characteristic curve starts generally at blackening (density) values of 0·3–0·9, while it extends to a density of 2·0 (or even greater values). The upper limit of the straight part of the characteristic curve and the following over-exposure part is not important practically. For the measurement of such high density values the error of measurement is generally unacceptably high. On the other hand, the under-exposure part of the photographic density curve can be utilised in just the same way as the linear part, provided that instead of the above-given 'true' density value a 'transformed density' is taken into account, established by empirical calculations (see Section 5.7).

According to one explanation for the form of the characteristic curve, on its lower part there is a linear variation of the density with the exposure. The likely reason

for this observation is that the size of the emulsion grains is not uniform. A small exposure first induces a photochemical effect in a small number of silver halide grains of the highest sensitivity. The reaction occurs for a larger number of less sensitive grains under the effect of further exposure. Thus the increase of density continues until there are no more grains having a sensitivity greater than average. The slope reduction of the over-exposed part of the characteristic curve can be accounted for by the fact that upon increasing the exposure the number of unexposed silver bromide grains of average sensitivity steadily drops. The same exposure, therefore, can only bring about smaller and smaller increases in density. Finally, the state is reached where all the silver halide particles become exposed. There can be no further increase in the photographic density.[1]

The above type of explanation cannot be applied to the phenomenon of solarisation, that is the phenomenon of a large exposure causing the characteristic curve to pass through a maximum. The reason for this phenomenon is not yet clear. According to one theory it is the result of a large exposure causing the silver (originally finely divided) that forms the latent picture to be brought together in bigger grains: consequently the crystal seeds formed during development will be reduced in number.[8]

In spectral analysis troubles can arise in principle owing to solarisation. On a photograph of too-intense lines a lighter band can appear at the centre of the lines. Thus solarisation can be mistaken for self-absorption. It should, however, be noted that in the case of modern, good-quality photographic emulsions solarisation occurs only as the result of very strong illumination, such as does not normally occur in spectrochemical analysis.

The fogging value of the photographic emulsion reduces the accuracy of the light-intensity determinations. Fogging always contributes to the density brought about by illumination. The error arising in this way has a relative value that is greater for high fogging values and low actual density levels. The fogging can therefore be neglected for high densities. On the other hand, when low density values are measured a correction should be applied.

The threshold value of the emulsion is significantly reduced and thereby its sensitivity is greatly increased if the emulsion is previously illuminated by low-intensity light. The same can be observed if the emulsion is stored for a longer time, but then the fogging value is significantly *increased.*

References

1. HURTER, F. and DRIFFIELD, V. C., *J. Soc. chem. Ind.*, London, **9**, 455 (1890)
2. ARRAK, A., *Appl. Spectrosc.*, **12**, 171 (1958)
3. ARRAK, A., *Appl. Spectrosc.*, **13**, 193 (1959)
4. FRIESER, H. and KLEIN, E., *Photogr. Sci. anal. Eng.*, **4**, 764 (1960)
5. TÖRÖK, T., *Emissionspektroskopie*, Berlin, 162 (1962)
6. KAISER, H., *Spectrochim. Acta*, **2**, 1 (1941)
7. TÖRÖK, T., *Proceedings of the XIV. Colloquium Spectroscopicum Internationale, Debrecen, 1967*, 6
8. ANGERER, E. and JOOS, G., *Wissenschaftliche Photographie*, 6th edn, Berlin (1956)
9. SAWYER, R. A., *Experimental Spectroscopy*, 2nd edn, New York, 125 (1956)

5.2.3 Designation of the sensitivity of photographic emulsions

In theory, the best way to describe the sensitivity of photographic emulsions would be by quoting that amount of light energy, measured in either ergs per square centimetre or lux-seconds, necessary to bring about a certain density on the linear part of the characteristic curve at a certain wavelength. The use of this unit has not become widespread in practice because knowledge of the sensitivity of the emulsion is only necessary for the right choice of exposure. For this purpose it is enough to know only the relative sensitivity instead of the absolute value. Internationally, however, no designation has yet been accepted that would uniquely characterise the relative sensitivity of photographic emulsion.

Formerly, the sensitivity of photographic emulsions was given on the so-called Scheiner scale whose approximate value is $10.\log 1/\text{lxs}$.[1] This designation has not, however, been accepted in practice, mainly because the values given for emulsion of different makes were not comparable.[2] This drawback is eliminated if the sensitivity of the emulsions is given in degrees DIN.[3,4] This is determined by establishing under strictly specified conditions the exposure, calculated in lux-seconds, necessary to bring about a density value of exactly 0·1 – the fogging value having been deducted. In practice, however, in the measurements the $D = 1{\cdot}40$ value is taken into account instead of the previously given density. The sensitivity so defined has the precise designation: 14°/10 DIN. The relationship between the DIN degree defined in the way described above and the lux-second is given in Table 5.2. From this table it is evident that as the sensitivity, in DIN units, of the photographic emulsion increases the exposure time necessary to attain a certain density value decreases.

Table 5.2

The exposure (E), measured in lux-seconds, corresponding to different DIN ratings

DIN/10	E (mlxs)	DIN/10	E (mlxs)	DIN/10	E (mlxs)
1	220	10	28·0	20	2·8
2	176	12	17·6	22	1·7
4	111	14	11·1	24	1·1
6	70	16	7·0	26	0·7
8	44·2	18	4·4		

For the characterisation of photographic emulsions the inertia value is also used (Fig. 5.10). The higher the value of the latter, the lower is the sensitivity of the emulsion, but the greater is the slope of the straight portion of the characteristic curve. This observation is the basis of the sensitivity designation by the HD number.[5] This is not a unique designation, however, since the value of the continental HD° is $34/i$, while that of the American HD° is $10/i$, where i is the inertia of the photographic emulsion in lux-seconds.[6] Furthermore, between the continental HD° and the DIN° the following relationship is valid:[7]

$$\log \text{HD}^\circ = 0{\cdot}1\ \text{DIN}^\circ + 1{\cdot}77$$

The designation ASA (American Standards Association, recently called United States Standards Association) is based on the Jones sensitivity criterion.[8] According to this system, there is a point on the characteristic curve – based on precisely processed spectral photograms made to exact specifications and by varying exposure depending on the experimental conditions specified – where the value of γ is 30% of the average gradient. This point is characteristic of the sensitivity of the emulsion.

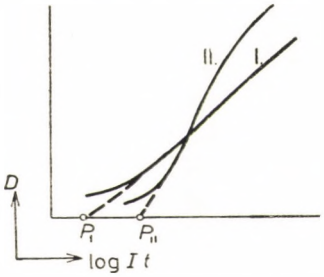


Fig. 5.10 Change of shape of the characteristic curve with the inertia of the emulsion

If the value of the exposure (in lux-seconds) belonging to the point obtained thus is E, then the *ASA Speed* sensitivity datum is obtained so that a zero is written in front of the reciprocal of E. (The relationship between the ASA Speed and the DIN° and the HD° values is summarised in Table 5.3.)[9]

Table 5.3

Numerical comparison of different sensitivity rating systems (ASA Speed, 10 DIN° and HD°)

ASA Speed	10 DIN°	HD°	ASA Speed	10 DIN°	HD°
0100	15	1 860	0400	21	7 400
0125	16	2 300	0500	22	9 300
0160	17	3 000	0650	23	11 750
0200	18	3 700	0800	24	14 800
0250	19	4 700	01000	25	18 600
0320	20	5 900	01250	26	

To summarise, according to a great number of comparative experiments, the most suitable unit of sensitivity is the degree DIN.[10] This statement, however, refers to general photographic uses and not specifically to spectrochemical analysis.

Photographic emulsions of different characteristics cannot be uniquely specified by a single number. Furthermore, from the point of view of spectrochemical analysis it is unfortunate that the designations for sensitivity described above are based on measurements where the experimental conditions are generally not the same as those for photographic spectrochemical analysis. These data refer to daylight, except for the Scheiner number which refers to the light of a benzine-burner flame. On the other hand, in spectrochemical analysis the spectrally resolved ultra-violet radiation generally is used.

References

1. SCHEINER, J., *Z. InstrumKde*, **14**, 201 (1894)
2. HANSEN, G., *Photogr. Ind.*, **36**, 53 (1938)
3. BILTZ, M., *Naturwissenschaften*, **21**, 734 (1933)
4. *Deutsche Industrie Normen* DIN 4512
5. HURTER, R., DRIFFIELD, V. C., *Eders Jahrbuch für Photographie*, 157 (1897)
6. LIEMPT, J. A. M. VAN, *Pays-Bas*, **53**, 163 (1939)
7. JONES, L. A., *Photogr. Ind.*, **39**, 163 (1939)
8. ASA Z-28 : 2.1 (1947) (American Standards Association)
9. EGGERT, J., *Schweiz. Photordsch.*, **12**, 193 (1948)
10. JAMES, T. H., *The Theory of the Photographic Process*, New York–London, 424 (1966)

5.3 Construction of the characteristic curve

5.3.1 Preliminary remarks

The relationship between the quantity of light falling on the emulsion and the resultant density, that is, the form of the characteristic curve, depends not only on the properties of the emulsion used, but also on a number of chemical and physical experimental conditions as well, as will be clear from the following. Thus the characteristic curves of emulsions made in the same batch can be different. This fact should be taken into account in the application of the characteristic curve.

In quantitative spectral analysis the intensity ratios of the analytical lines are used for the determination of concentrations. For the purposes of quantitative analysis therefore the only correctly applicable characteristic curve is one that has been constructed under conditions identical to those of the analysis. The further the divergence of the experimental conditions the greater is the inaccuracy.

According to the above a generalised characteristic curve is not applicable to spectral analysis, since it refers to white light instead of referring to the wavelength of the analytical line. Owing to the Schwartzschild effect the characteristic curve constructed with the help of spectra made at *different exposure times* is also unsuitable for spectral analysis. It should, however, be noted that if the exposure time is varied between narrow limits this effect is not significant.

The intermittence effect can be neglected. This effect causes the density value due to intermittent illumination to be different from that due to an equivalent amount of continuous illumination. In quantitative spectral analysis it is generally not the absolute intensity of the lines but their *ratio* that is required. Thus in the construction of the characteristic curve the unit of the $I.t$ exposure can be chosen at will. This fact does not affect the form of the curve but only leads to a parallel shift in the direction of the abscissa.

In principle, the characteristic curve is most suitable for spectrochemical analysis when the spectral photographs are made so that the stigmatic spectrograph used for the analyses is illuminated by light of continuous spectrum and with different slit widths.[1] Under such conditions the intensity of light falling on the photographic emulsion varies proportionally to the slit width even when the slit is so narrow that diffraction occurs. The advantage of the procedure is that in this way a characteristic curve valid for any wavelength of the spectrum can be obtained.

The specified accuracy can, however, be obtained only when the slit width is precisely set. It is furthermore important in the procedure that the intensity of the

light source (for example, high-pressure xenon burner, 'nitra' incandescent lamp used in photography) should not change while photographs are taken. This requirement is, however, very difficult to meet in practice. This difficulty is even important in the simple procedure in which the intensity of the light illuminating the photographic emulsion is changed by varying *the distance* between the light source and the slit (appropriately set) of the spectrograph. In this method the light intensity changes proportionally with the square of the distance. This method can only be used with success when the sufficiently powerful source has not only a constant intensity but is also point-like. In the opposite case the distances cannot be set with the required accuracy.

In both of the procedures mentioned above it is also a disadvantage that a continuous spectrum is used in the construction of the density curve instead of a line-spectrum source. The error caused during development should, therefore, be taken into account, especially the error caused by the Eberhardt effect (see Sub-section 5.6.5).

The difficulties mentioned previously do not occur when the characteristic curve is constructed by the methods that are most widespread at present; as described in the following. These have the further advantage that the necessary measurements can be performed with the equipment generally used in spectral analysis. Among the procedures given the ones based on multiplet lines and on two-step attenuation without a secondary curve are the most suitable.

It should furthermore be noted that the characteristic curves constructed according to the procedures to be described can only be used when the photographs necessary for the construction of the characteristic curve and for the analyses are processed under exactly the same conditions, and when the photographic plate has the same batch number and is, of course, of the same type. Even when this rule is observed it is recommended to check the characteristic curve periodically.

Reference

1. PREUSS, E., *Spectrochim. Acta*, **4**, 212 (1950)

5.3.2 Construction of the characteristic curve using multiplet lines

This method of constructing the characteristic curve is based on the fact that the relative brightness of the multiplet lines is greatly independent of the excitation. If the intensity relationships of the multiplet lines are known, then in favourable cases the characteristic curve can be directly constructed with the aid of the spectrum of the material investigated. The procedure is simply to measure the density of the multiplet lines. The values obtained this way are plotted on a diagram in which the distance between the parallel perpendicular lines is equal to $\log I$ belonging to the multiplet lines, that is, this distance corresponds to the difference of the Y-values (Fig. 5.11). The curve passing through these points is the characteristic curve of the emulsion in question.

In the selection of the spectral lines for the above procedure it should first be taken into account that their wavelength should be located in the range corresponding to the analytical line-pairs of the element to be determined. It is also necessary that the brightness of the lines should not be influenced by self-absorption, in other

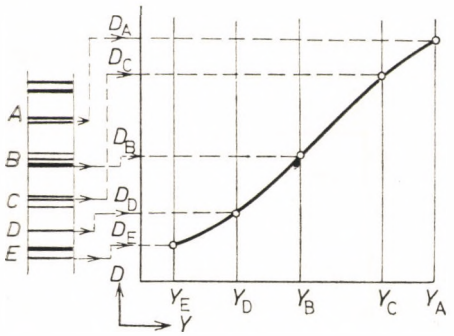


Fig. 5.11 Construction of the characteristic curve by means of multiplet spectral lines

words, the resonant lines are not suitable for this purpose. It is absolutely necessary that in the wavelength range of the multiplet lines used the slope value of the emulsion be constant. It is also advantageous if the distribution of the points determining the characteristic curve are, if possible, evenly scattered, that is, if the ΔY-values between the multiplet lines are almost the same. This requirement is only met by the presence of an element contained in the sample, for example, iron (Tables 5.4 and 5.5).

Table 5.4

Iron lines used for emulsion calibration and their corresponding logarithmic intensity values (Y) (spark excitation: 2/3 kVA, 0·32 MHz)[2,3]

Line group	Å	Y	Interfering lines	
2 500 Fe II	2 451·21	1·47	Ni	2 450·986
	2 461·28	2·49		
	2 461·86	2·58		
	2 463·28	2·14	Fe I	2 462·644
	2 464·01	2·32		
	2 464·90	2·21	Fe II	2 465·200
	2 465·91	2·26		
	2 470·41	2·28	Fe II	2 470·658
	2 497·30	1·31		
	2 503·32	2·35	V	2 503·020
			Fe II	2 503·556
	2 506·09	2·38		
	2 514·91	1·69		
2 800 Fe II	2 753·29	3·17	V	2 753·404
	2 779·30	2·82		
	2 783·70	2·92		
	2 793·89	2·13		
	2 799·29	1·99	V	2 799·451
	2 813·61	1·63		
	2 827·43	1·64		
	2 831·56	2·69		

Table 5.5

Iron lines used for emulsion calibration
and their corresponding logarithmic intensity values (Y)
(D.c. arc excitation: 2·2 A)[2, 4]

Line group	Å	Y	Line group	Å	Y
3 200 Fe I	3 151·35	1·909		3 623·19	2·051
	3 156·27	1·604		3 625·15	1·749
	3 157·04	1·997	3 800 Fe I	3 756·94	1·754
	3 157·89	1·863		3 774·83	1·663
	3 175·45	2·089		3 776·46	1·685
	3 178·02	1·968		3 805·34	2·449
	3 182·98	1·581		3 807·54	2·096
	3 199·52	2·283		3 836·33	1·929
	3 205·40	2·318	4 000 Fe I	3 971·33	1·882
	3 217·38	2·176		3 977·74	2·127
3 400 Fe I	3 369·55	1·978		3 981·77	1·706
	3 370·79	2·203		3 983·96	1·893
	3 394·59	1·722		4 009·72	2·004
	3 402·26	1·811		4 021·87	2·046
	3 415·53	1·543		4 024·74	1·573
	3 425·02	1·732	4 200 Fe I	4 175·64	2·018
3 600 Fe I	3 572·00	2·147		4 176·57	1·668
	3 594·64	2·093		4 184·90	1·920
	3 599·62	1·477		4 208·62	1·331
	3 612·07	1·822		4 210·35	2·134
				4 213·65	1·464

The procedure based on multiplet lines not only has the advantage that it is simple and fast. It also has the advantage that the spectra necessary for the construction of the characteristic curve can be made under exactly the same conditions as the spectra used for analysis. Thus the partial errors of the determination are largely eliminated. (For example, in steel analysis the characteristic curve can be *directly* constructed by means of the spectrum of the sample and by applying the above tables.) It is also useful that the procedure can be equally applied for both stigmatic and astigmatic spectrographs. On the other hand, it is a disadvantage that the method is limited in its applications because not all parts of the spectrum contain suitable multiplet lines.

References

1. DIEKE, G. H. and CROSSWHITE, H. M., *J. Opt. Soc. Am.*, **33**, 425 (1943)
2. ASTM Committee E-2 on Emission Spectroscopy, *Methods for Emission Spectrochemical Analysis*, 29 (1957)
3. DIEKE, G. H., 'A Study of Standard Methods for Spectographic Analysis', *Progress Report W-90, War Metallurgy Committee for the National Academy of Sciences, National Research Council (March 20, 1944)*
4. CROSSWHITE, H. M., *Spectrochim. Acta*, **4**, 122 (1951)

5.3.3 Construction of the characteristic curve using a multi-step attenuator

In this procedure the light falling on the slit of the spectrograph is attenuated in several steps along the length of the slit. In stigmatic line-formation the lines photographed are composed of a number of parts of different density value. The characteristic curve can therefore be determined directly if the densities of the line parts are determined and if the amount of attenuation of the different steps is exactly known (Fig. 5.12). The advantage of this procedure is that it is simple and speedy. The perpendicular lines corresponding to the ΔY-values between the different steps of the attenuator can be drawn on the diagram in advance. It is also advantageous that, in theory, by using the same spectral photograph any characteristic curve for any chosen wavelength – or maybe more than one – can be constructed. On the other hand, it is difficult to ensure that the large slit height is perfectly evenly illuminated. The large slit height causes not only optical errors, but also a development error due to the Eberhardt effect. The inaccuracy arising from this source can be eliminated only partially by the fact that during the analyses the spectrum of the sample examined is usually photographed by means of a two- or three-step filter.

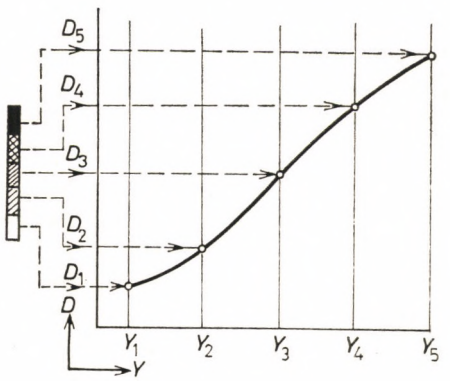


Fig. 5.12 Construction of the characteristic curve by means of a multi-step attenuator

The attenuation necessary for the construction of the characteristic curve can be carried out in many different ways. A simple and frequently applied procedure is that where the light is attenuated by a multi-step filter. In its application to the stigmatic spectrograph the multi-step grey filter is fastened immediately in front of the slit. The relative transmittance of the commercially available filter steps, the *nominal* value of the steps, is, however, normally given as referred to white light. The 'grey' filters used are also not quite grey. This does not, in theory, lead to errors if we know the variation of the relative transmittance of the individual steps as a function of the wavelength. For this purpose the relative transmittance of the steps is checked in the spectral range into which the wavelength of the analytical lines falls by means of a spectrophotometer (Fig. 5.13).

The calibration of the steps of the step filter can also be carried out spectrographically by Plsko's method. The essence of the procedure is that the intensity reduction between two adjacent steps can be compensated for by opening the slit, if the radiation is coming from a source of constant radiance and of continuous spectrum.

The other method of attenuation, which is also frequently used, requires a multi-step rotating disc (or step-sector). In this method the amount of attenuation is com-

pletely independent of the wavelength of light. The effect of the disc being rotated immediately in front of the slit is the same as that of the step filters, provided the speed of rotation is greater than the critical limit. Hence, the errors due to the large slit height are the same. When the speed of rotation is lower than necessary, the intermittency effect should also be taken into account.

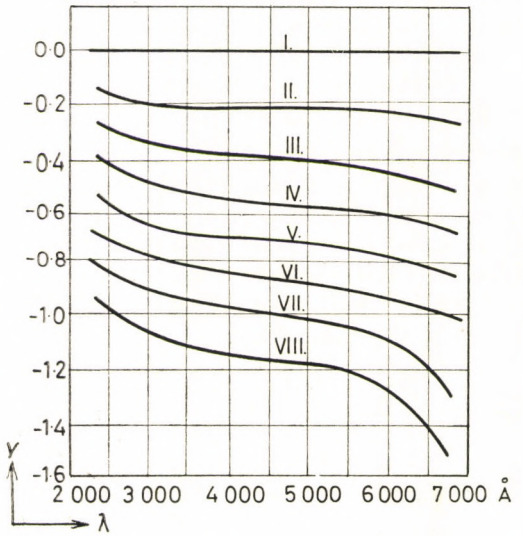


Fig. 5.13 Variation of the transmittance of an eight-step filter as a function of the wavelength

In both the procedures described previously it is advisable to attenuate logarithmically the light falling on the slit of the spectrograph because the measurement points are then located at equal distances along the *Y*-axis of the characteristic curve. This arrangement facilitates the construction of this curve.

In the third method of construction of the characteristic curve based on stepped light attenuation, the light from the spectral source is attenuated by a Hansen-type diaphragm (see Sub-section 3.7.7). This procedure has not, however, become widespread. The method is not only essentially more complicated than the two described earlier but it also involves the drawback that the light illuminates the slit of the spectrograph under circumstances entirely different from those of practical usage of the instrument.

Reference

1. Plsko, E. and Liska, M., *Acta geol. geograph. Univ. comenianae*, **6**, 583 (1962)

5.3.4 Construction of the characteristic curve using a single pair of multiplet lines

We have already seen (Sub-section 5.3.2) that the characteristic curve can be constructed by means of multiplet lines whose intensity relations are already known. The construction can, however, be done by using a single pair of lines, that is by using two lines of known intensity ratio.[1] The prerequisite of this procedure is, however, that the absolute intensity of the line pair should vary over a broad enough interval. This condition is easily met when the concentration of the element belonging to the line pair varies. For example, in the analysis of aluminium samples of variable iron content, in order to construct the characteristic curve necessary for the evaluation

of the spectra, a suitably chosen iron line pair may be used. Although the intensity of the latter varies with the concentration, the intensity ratio I_a/I_b, ($I_a > I_b$) remains the same. First of all, the transmittance values T_a and T_b, ($T_a < T_b$) belonging to these lines should be established. On the basis of these results and the knowledge of the intensity ratios the characteristic curve can be constructed according to the following.

First the so-called preliminary curve is constructed. This can be done by means of the pairs of measurement values so that the T_b values are plotted as a function of the T_a values (Fig. 5.14*a*). By means of the preliminary curve we find the series of value pairs T_a, T_b using the 45° projection lines and starting from a randomly chosen value pair T_{a1}, T_{b1}, for which the member T_a of the subsequent value pair is the same as the member T_b of the previous one ($T_{an} = T_{b(n-1)}$). With the help of the so-called value pairs T_{a1}, T_{b1}, T_{a2}, T_{b2}, . . . , T_{an}, T_{bn} the characteristic curve can be easily constructed, preferably with the help of the logarithmic intensity ratio $\log I_{a1}/I_{b1} = \Delta Y_m$ (Fig. 5.14*b*). Starting from the randomly chosen point on the Y-axis of the characteristic curve the value ΔY_m is measured as many times in the negative direction as there are value pairs T_a, T_b available. To the ordinates belonging to the points found thus the measurement results T_{b1}, T_{b2}, . . . , T_{bn} are measured. Connecting the points obtained in the way described above we obtain the characteristic curve Y, T.

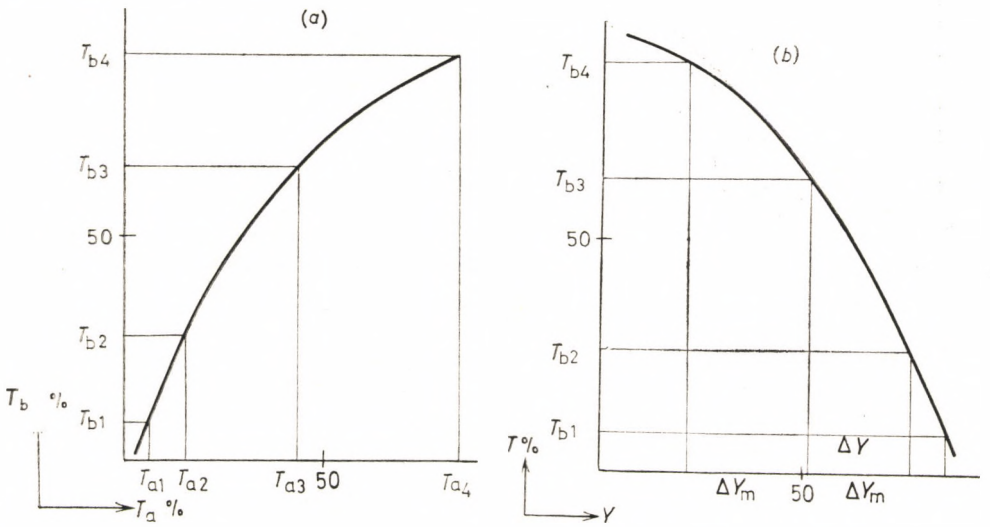


Fig. 5.14 Construction of the characteristic curve by means of a multiplet line pair

The advantage of the procedure is that it is sufficient to use a single line pair. A difficulty, however, is that only a few line pairs of known intensity ratio are given in the literature. A further difficulty is that the samples to be analysed contain either zero concentrations or else unsuitable concentrations of the element for which calibrated multiplet lines are available. Finally, the calibrated lines are perturbed by different sorts of foreign lines, or by the background, and thus their intensity ratio is altered. To eliminate the above difficulties the two-step filter method, which is dealt with in the following chapter, can be used.

Reference

1. CHURCHILL, J. R., *Ind. Engng Chem. analyt. edn*, **16**, 163 (1944)

5.3.5 Procedures based on the two-step attenuation of light

The characteristic curve can be constructed not only by the use of multi-step filters, but also by applying two-step filters. On spectral photographs made this way the intensity ratio between two different steps of the lines of different intensity to be compared is identical to the relative transmittance of the two filter steps. By the use of the density of the line portions belonging to the same intensity ratios the characteristic curve can be constructed.

In these methods the spectral photographs necessary for the construction of the characteristic curve are made by photographing the spectrum of the sample (containing the element to be determined in variable amounts) while using a two-step filter or rotating disc. In the spectral range of the analytical lines the density value of the chosen lines – preferably that of the analytical line – is measured in each step.

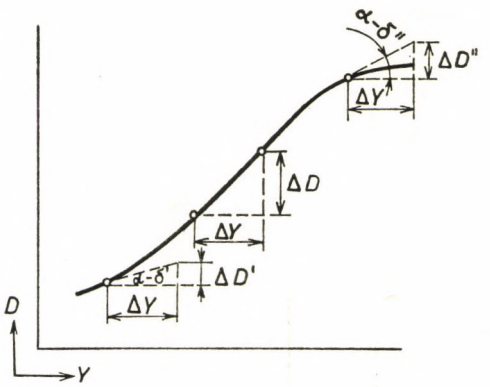


Fig. 5.15 Construction principle of the characteristic curve using the two-step attenuation method

The procedure is based on the fact that the D_a and D_b densities measured in the two steps of the spectrum correspond to an I_a or I_b intensity pair whose ratio I_a/I_b is equal to the filter constant. For all value pairs, therefore, the following relationship is valid

$$\log I_a/I_b = \log I_a - \log I_b = Y_a - Y_b = \Delta Y_{a,b}$$

In turn, to the above values the density difference $D_a - D_b = \Delta D_{a,b}$ belongs. (Note: in the following, instead of the designations $\Delta Y_{a,b}$ and $\Delta D_{a,b}$, the symbols ΔY and ΔD will be used for the sake of simplification.)

Furthermore, in the straight part of the characteristic curve the following relationship is valid – according to Fig. 5.15:

$$\Delta D = \tan \alpha . \Delta Y = \gamma \Delta Y = \text{const.} \qquad (5.17)$$

On the other hand, on the under-exposed and over-exposed parts of the characteristic curve the relationships are as follows

$$\Delta D' = \tan (\alpha - \delta') \Delta Y$$

and

$$\Delta D'' = \tan(\alpha - \delta'')\Delta Y \tag{5.18}$$

It is therefore true not only that on the under-, and over-exposed parts of the characteristic curve the density difference measured on the two filter steps is smaller than that measured on the straight portion, but also that the values of δ' and δ'' are not constants.

The *preliminary curve method* for the construction of the characteristic curve is based on the above statements.[1-3] This preliminary curve orders the density differences ΔD measured on the spectral lines in such a way that the value of ΔD is a function of D_a in the straight part and the over-exposed part of the spectrum, that is, a function of the density value measured on the lighter portions of the lines. On the other hand, on the under-exposed part the value of ΔD is a function of the density D_b measured on the denser part of the lines. This is due to the fact that ΔD achieves its maximum value on the under-exposed part when D_b attains the straight part.

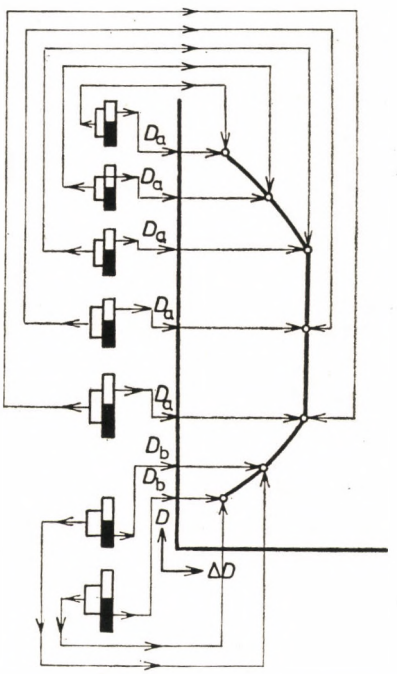


Fig. 5.16 Variation of the ΔD value as a function of D_a and D_b

On the other hand, ΔD starts dropping on the over-exposed part when D_a leaves the latter. The preliminary curve can therefore be constructed as shown on Fig. 5.16. With the aid of this preliminary curve the characteristic curve can be constructed as sketched on Fig. 5.17. The starting point A of the straight part of the curve is obtained so that a horizontal line is drawn across that point A of the preliminary curve where the straight part of the preliminary curve ceases. Any point on the latter straight line can be taken as the starting point A of the straight part of the characteristic curve. The slope of the straight portion is $\Delta D/\Delta Y$, according to Eq. (5.17). Accordingly, the direction of the straight part can be drawn through point C. The horizontal line drawn through point B' of the secondary line, on the other hand, indicates the termination point B of the straight part.

The *under-exposed part* of the characteristic curve can be constructed by the help of the preliminary curve so that the intersection point D_A formed by the horizontal line AA' and the perpendicular axis of the preliminary curve, denoted by $D_A D_B$, is projected at an angle 45° onto the secondary curve. Through the point D obtained thus a horizontal line is drawn. Afterwards the point A indicating the termination of the straight part of the characteristic curve is projected onto the Y-axis of the diagram. The first sought point D of the under-exposed part of the characteristic curve is formed by the intersection of the perpendicular line drawn through the point corresponding to the ΔY-value (calculated from the above-mentioned intersection point), and the horizontal line drawn through the previously constructed point D'.

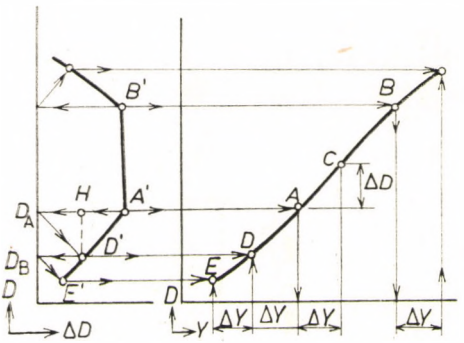


Fig. 5.17 Construction of the characteristic curve by means of the preliminary curve

By the 45° projection of point D_A the square $\Delta_A \Delta_B D'H$ is formed; thus $D'H$ equals $\Delta_B D'$. The distance $D'H$ is, therefore, equal to the density difference belonging to the two steps of the spectral line in question. The point E of the under-exposed part of the characteristic curve, and the rest of the points, can be constructed with the help of the preliminary curve points D', E', etc., in a manner identical to the procedure previously described. The point G belonging to the over-exposed part of the characteristic curve can be determined similarly.

From the construction described above it is evident that the number of points determining the characteristic curve becomes smaller as the filter constant becomes larger. It is therefore not advantageous to use an attenuator of too high a constant. When, in turn, the constant is too low, the construction of the characteristic curve becomes unnecessarily lengthy. The optimum value of the constant depends on the slope value of the straight part of the characteristic curve. According to experience, the best values are: when $\gamma \approx 1$ the two-step attenuator should be 100/60%, and when $\gamma \approx 2$ the two-step attenuator should be 100/80%, with respect to its transmittance. The first type is generally used in the ultra-violet range, while the latter is used in visible spectral range studies. In order to compensate for the variation of the slope with wavelength it is advisable to use the two-step rotating sector of variable sector constant, or the two-step filter of variable filter constant.[3] Owing to their simplicity the selective filters are especially useful to compensate for the slope-changes of certain types of emulsion.[4] This is because the wavelength-dependent filter constant varies in the opposite way to the slope value.

The characteristic curve can be constructed in a simpler and faster way without using the preliminary curve. To apply the procedure first select, in a spectrum made by two-step alternation, an analytical line (or another line in its *immediate* vicinity)

with a maximum density difference between the two steps. The value of this difference is used to determine the direction of the straight part of the characteristic curve, in the same way as described in the previous procedure, by using the value of $\Delta S/\Delta Y$ (Fig. 5.18). Furthermore, lines can be selected having density values in both steps that lie on the straight part of the characteristic curve, for example, the lines λ_A and λ_B given on the figure. Thereby the direction of the straight part can be determined. The under- and over-exposed parts of the characteristic curve can be constructed by using a line the density of one step of which falls on the straight part of the characteristic curve (for example, the line λ_C or λ_D in the figure). The point on the curved part of the characteristic curve (i.e. point C'' or D') can be determined by means

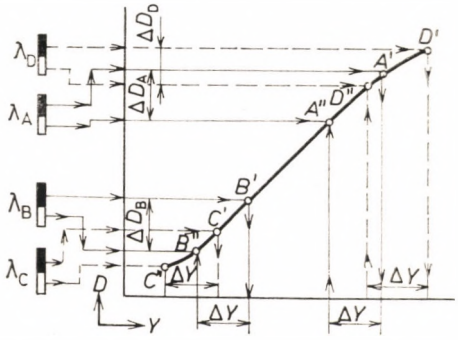


Fig. 5.18 Characteristic curve

of the density measured and the value of ΔY, that is, the value of the filter constants. A similar procedure can be used for the construction of those points corresponding to the value pairs D' and D'' that fall entirely on the curved part of the characteristic curve. In this operation, however, it should be borne in mind that one of the points of the value pairs should fall on a part of the characteristic curve that is already known.

Neither of the above-described methods of constructing the characteristic curve can be applied when the concentration of the element to be determined varies only between narrow limits, when the spectrum examined has few lines, or when γ varies strongly with the wavelength (generally above 320 nm). In these cases the density value of the analytical lines differs only slightly, or in other words, there are too few usable lines in the spectrum. These difficulties are eliminated when the spectra necessary for the determination of the characteristic curve are photographed by varying the exposure time. This way the number of analytical lines of different density values can be increased to an extent adequate for the exact construction of the characteristic curve. The fact that the spectra necessary for the characteristic curve and those necessary for the analyses are not prepared under quite the same experimental conditions does not give rise to any practically significant errors. This is because in order to construct the characteristic curve only the value pairs D_a, D_b and the $\Delta Y_{a,b}$ values are necessary. The exposure time does not play a role. Therefore neglecting the reciprocity law does not result in errors. However, it is still advisable for the exposure times to be not very different from those used in the analyses.

The construction based on the two-step filter method has the great *advantage* that the experimental conditions under which the spectral photographs, necessary for the construction of the curve, are made approximately closest to the experimental

conditions of the actual analyses. This is generally true, except for the method based on the multiplet lines. The two-step attenuation does not give rise to different development effects since the same attenuation is generally used in the analyses. A further advantage of the procedure is that the illuminated slit height is not too great. The even illumination of the slit can therefore be easily ensured. Also, it is useful that this method can be applied to the construction of characteristic curves valid for any wavelength of the spectrum.

A disadvantage of the method is that the light-attenuating device placed immediately in front of the slit can only be used with stigmatic spectrographs. It is to be noted, however, that most of the modern spectrographs are of stigmatic imaging construction.

References

1. TÖRÖK, T., *Annls. Univ. Scient. Rolando Eötvös Budapest Sect. Chem.*, **2**, 415 (1960)
2. TÖRÖK, T., *Optik und Spektroskopie aller Wellenlängen*, Berlin, 573 (1962)
3. VAN CALKER, J., *Z. analyt. Chem.*, **198**, 143 (1963)
4. NAGY, Z. and SÁMSONI, Z., *Spectrochim. Acta*, **24B**, 147 (1969)

5.4 Photographic emulsions used for intensity measurements

5.4.1 Spectral plates and films

The light-sensitive gelatin layer is carried either by a sheet of glass or by a film. The thickness of the spectral plates used for spectrochemical analysis is about 0·7 mm for thin sheets, and 1–1·5 mm for normal thickness. The thin sheets are used when it is necessary to bend the plates to a greater extent. With certain spectrographs this is the only way to ensure that the photographic emulsion is situated on the curved focal surface of the spectrograph. The length of the spectral plates is generally controlled by the length of the spectrum that can be photographed on one occasion. Their width is controlled by practical considerations. In series examinations it is more advantageous if the spectral plate is wide, while in quick informative analyses narrow plates are favoured. The more usual dimensions are 6×24, 9×12, 9×24 cm, or 4×10 in, etc.

The width of the films used for spectrochemical purposes varies between 0·15 and 0·3 mm. Generally the normal size, that is a width of 35 mm, is used and the edges of the film may be either perforated or left intact.

The layer thickness of the photographic emulsions used for spectral purposes is 15–20 mμ. This value increases ten-, or twentyfold during development, while in the course of drying the original dimensions are regained. The diameter of the silver halide grains suspended in the gelatin is usually about a few microns. The average dimensions of the grains of the highly sensitive emulsions are greater than those of the less sensitive emulsions. The surface density of the grains is usually 200 000–300 000/cm^2. The weight of silver bromide in an area of 1 cm^2 is 10–25 mg.

According to the Hartmann rule the best emulsion for the purpose of measuring intensity is one that produces the same density when the conditions of exposure and processing are identical. A further requirement of spectrochemical emulsions is that the relationship between the light intensity and the density of the emulsion should preferably be simple and well defined. It is also important that a large density differ

ence be produced by a small light-intensity difference. The only emulsions adequate for spectrochemical analysis are those manufactured specially for the purpose having a single layer, fine grains, freedom from halo, and possibly the most homogeneous emulsion available. Further needs are that the emulsion should be sufficiently sensitive and of a hard type.

As may be seen from the above discussion the emulsions made for amateur photographers are not suitable for light-intensity measurements and thus for quantitative spectrochemical analysis. These emulsions do not reproduce the intensity relationships but artificially level out the large intensity differences. These emulsions are not only 'soft' but are also of multilayer type. Their uppermost layer is the most sensitive, while the underlying layers are of successively lower sensitivity. Since the low-intensity rays enter only the upper layers, these produce a relatively larger density than the more intense rays which enter the deeper and less sensitive layers. This fact, of course, renders it impossible to use such emulsions for the determination of the intensity relations of spectral lines.

5.4.2 The halo effect[1-2]

This phenomenon is caused by either the reflection or the diffusion of light. Both effects are detrimental to intensity measurements.

The so-called *reflection halo* is caused by the reflection of light travelling through the spectral plate or film from the glass–air interface (or film–air interface in the case of films). These reflected rays re-enter the photographic emulsion from the back (Fig. 5.19). Depending on the position of the spectral plate the rays are either partially (*a*) or totally (*b*) reflected. Total reflection occurs when the angle of reflection is large, about 80°. In the case of partial reflection the spectral lines will have diffuse

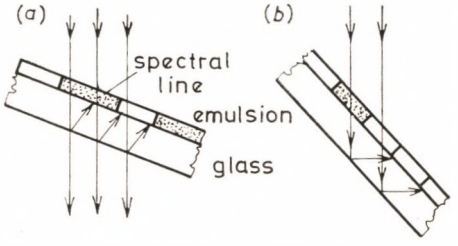


Fig. 5.19 Origin of the reflection halo

edges. When total reflection occurs a partial line is produced a certain distance from the genuine line. Such line perturbation occurs, however, only when the intensity of the reflected light is much higher than the threshold value of the emulsion.

The phenomenon mentioned above may be eliminated by using halo-free emulsions. In these emulsions an appropriately coloured, light-absorbing layer is placed between the light-sensitive emulsion and its carrier. Alternatively, the filter layer is placed behind the carrier. For this purpose the only appropriate colouring materials are those that do not interfere with the properties of the photographic emulsion and are decoloured or dissolved during development. In the case of plates made for photography of the ultra-violet spectral range, however, the glass itself greatly hinders the occurrence of the reflection halo. The glass base absorbs light of wavelength shorter than 3500 Å.

The *diffusion halo* is caused by part of the rays entering the photographic emulsion not being directly absorbed by the silver bromide but being reflected by the surface of the silver bromide crystals. Thus, scattered light appears within the emulsion.

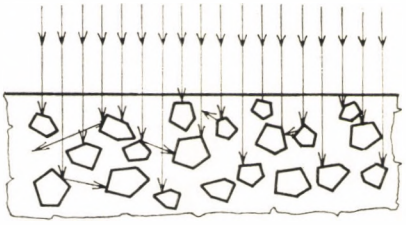


Fig. 5.20 Origin of the diffusion halo

This stray light is partly absorbed by the neighbouring silver bromide grains (Fig. 5.20). This has the consequence that the sharpness of the spectral lines is reduced and their width increased. This phenomenon occurs to a lesser extent in thinner photographic emulsion layers, hence the absorption of light is greater. Conversely, grainier emulsions produce more scattered light and therefore there is a greater tendency for diffusion halos to appear.

References

1. ANGERER, E. and JOOS, G., *Wissenschaftliche Photographie*, 6th edn, Berlin (1956)
2. MEES, C. E. K., *The Theory of the Photographic Process*, 6th edn, New York (1952)

5.4.3 Spectral sensitivity of photographic emulsion

The sensitivity of simple, non-sensitised emulsion commences at about 1800 Å and from 2500 Å onward increases rapidly. The maximum level is reached at about 3200 Å. It then decreases, slowly at first and then rapidly from 4800 Å; finally at about 5800 Å the sensitivity falls to zero (Fig. 5.21). The latter phenomenon occurs because the silver bromide does not absorb rays of wavelength longer than 5400 Å, so they have no effect. The intensity appearing towards shorter wavelengths is caused, in turn, by the fact that the gelatin increasingly absorbs rays of wavelength shorter than 2500 Å (Fig. 5.22).

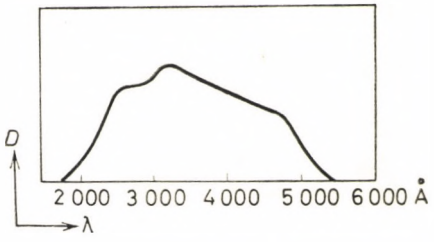


Fig. 5.21 Sensitivity curve of a non-sensitised silver-bromide photographic emulsion[1]

The upper limit of sensitivity is pushed to about 6500 Å if the photographic emulsion contains a sufficient quantity of *sulphur compounds* in addition to the silver bromide. This is because the silver sulphide easily releases electrons even as the result of thermal effects. The sulphide ion thus makes up for the electron loss of the bromide ion that occurs in the course of the latent image process (Eq. 5.10).[3]

Non-sensitised photographic emulsion can only be used between 2500 and 5000 Å. This limit can be shifted toward the shorter wavelengths by reducing the light-loss caused by absorption in the gelatin by using an emulsion containing only small amounts of gelatin. These are the so-called Schumann plates[4] that can be used between about 1000 and 3000 Å. Their drawback is that they are not durable. In a comparatively short time they become fogged, even if a specially prepared solution is used for development.

The previously mentioned drawbacks are eliminated in the so-called *u.v. plates*, sensitised for the ultra-violet spectral range. The sensitisation process of these plates is based on fluorescence, that is, on the fact that certain materials absorb shorter-wavelength light in such a way that they simultaneously emit longer-wavelength light.

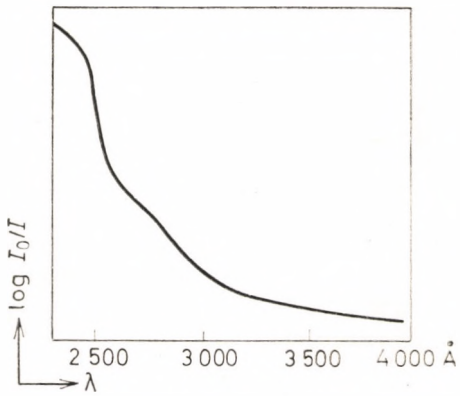


Fig. 5.22 Gelatin absorption curve[2]

In the case of emulsions covered by such a layer the blackening is brought about by the light emitted during fluorescence. Such commercially manufactured emulsions are preferably used for the 2000–2700 Å portion of the spectrum. Certain makes, for example the Agfa u.v. plates, can be developed directly. Their sensitisation is achieved by means of some water-soluble compound. Other makes, for example the Kodak u.v. emulsion plates, can only be developed so that the sensitising material is dissolved before development by benzine, benzene, etc.

Photographic emulsions can be sensitised for the far ultra-violet range in any laboratory. The compounds used for this purpose are normally aromatic or cyclic hydrocarbons. The following examples can be given: naphtha dissolved in benzene, alcoholic solution of esculine (a glycoside found in chestnut).[5] Recently, an alcoholic solution of sodium salicylate has been used very successfully.[5]

The sensitivity of the emulsions sensitised in the way described above does not extend to the visible or infra-red region of the spectrum. As has been mentioned, the insensitivity of the emulsion at the longer wavelengths is caused by the fact that these rays are not absorbed by the silver bromide. The emulsion can only be sensitised for this region by the use of compounds that are adsorbed on the silver bromide grains so that these will absorb the light of the required wavelengths. The dye probably forms with the silver bromide the type of surface system in which electrons are transferred from the dye into the crystal lattice of the silver bromide—as a result of the effect of the absorbed light. Also, the dye is continuously regenerated. On the electrons being transferred to the crystal lattice of the silver bromide the latent image appears in the usual way, as shown by the process (5.10).[3] There are variously de-

scribed commercially available emulsions, the names indicating their spectral sensitivity. The so-called *orthochromatic* emulsions are sensitive to about 5600 Å, starting from the ultra-violet. They are insensitive to red light (Fig. 5.23), so they can be developed in red light. The sensitivity of *panchromatic* photographic emulsions extends to the red region of the spectrum, that is, about 6400 Å. On the other hand, the *hyperpanchromatic* emulsions are even sensitive in the infra-red spectral region. It is striking that the sensitivity curves of the emulsions previously mentioned pass through a minimum at the wavelength of green light, that is at about 5000 Å. This 'green gap' is extremely useful as these emulsions can be developed in dark-green light. However, when measuring light intensity this factor is a drawback. The total elimination of the 'green gap' has so far not been successful, not even in the case of the so-called isopan emulsions. The commercially available *superpan* emulsions are mainly sensitised for red light. The latter two types of emulsion can only be developed in darkness.

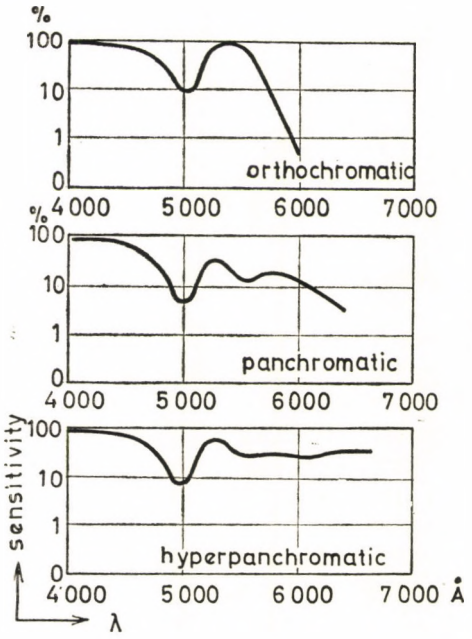


Fig. 5.23 Sensitivity curves of photographic emulsions used in practice, sensitised for the visible spectral range[3]

Emulsions sensitised for the invisible infra-red region are known as *'infra-red'* emulsions. Numbers are generally used to indicate the wavelength, in nanometers, at which the sensitivity of the emulsion is greatest. The general level of sensitivity and the durability of these emulsions become lower as the range of wavelength they are designed for extends farther into the infra-red region. A further characteristic is that their sensitivity to dark-green light is low, so they can generally be developed in dark-green illumination.

References

1. SCHEIBE, G., *Photographische Absorptionsspektrophotometrie*, Thesis, Munich (1929)
2. LUNDEGARDH, H., *Die chemische Spektralanalyse der Elemente* (1934)
3. ANGERER, E. and JOOS, G., *Wissenschaftliche Photographie*, 6th edn, Berlin (1956)

4. SCHOEN, A. L. and HODGE, E. S., *J. Opt. Soc. Am.*, **40**, 23 (1950)
5. EDER, J. M., *Rezepte, Tabellen und Arbeitsvorschriften*, Leipzig (1949)

5.4.4 Photographic emulsions used for spectral analysis

For spectrochemical purposes special *spectral plates* are commercially available. These are single-layer, halo-free, possibly homogeneous emulsions, sensitised for different spectral ranges. The sensitivity ranges of the 'Spektra' and 'Scientific' spectral plates made by Orwo and Kodak are shown in Fig. 5.24 so that the part

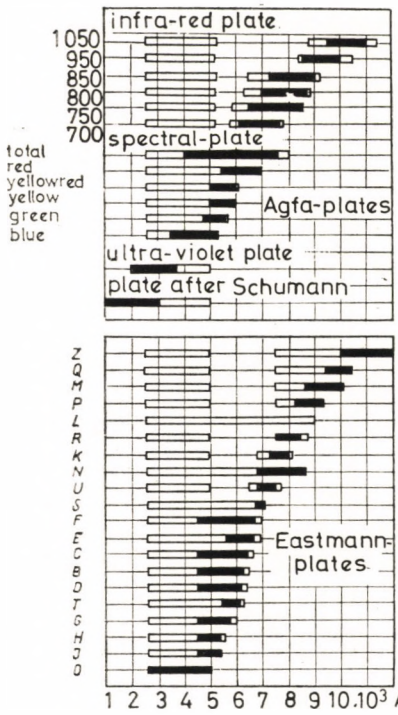


Fig. 5.24
Sensitivity table for Agfa and Kodak spectral plates

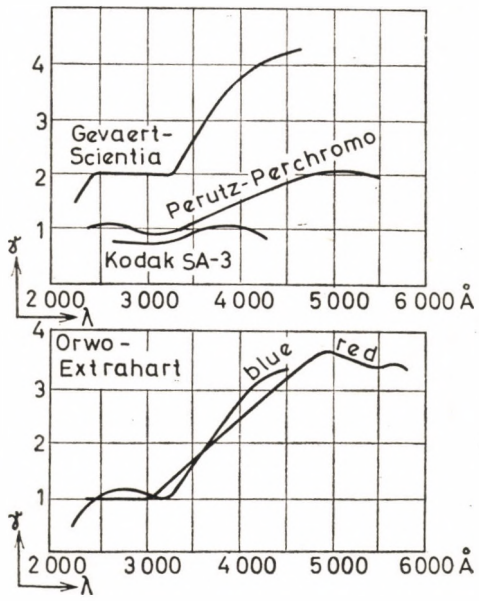


Fig. 5.25 Variation of the γ-value of some frequently used spectral plates as a function of the wavelength

within the frame contains the spectral range of moderate sensitivity, while the black part indicates the range of maximum sensitivity. These plates are available in different qualities and under different names: their relative sensitivity gradation γ, and graininess (see Sub-section 5.5.5) are contained in Table 5.6. The variation of the γ-value, as a function of the wavelength of some of the more frequently used spectral plates, is shown in Fig. 5.25.

Table 5.6

Important characteristics of different Agfa spectral plates

Notation	Ultrarapid	Rapid	Hart	Extrahart	Ultrahart
relative sensitivity	10	4	1	0·2	0·01
gradation (γ)	1·4	2·0	3·2	3·7	4·1
graininess	22	20	13	8	5

From Table 5.6 it appears that with increasing contrast (γ-value) of the emulsion the sensitivity drops and the graininess becomes finer.

Furthermore, it may be noted that among commercially available films the 'positive film' is very suitable for the purposes of spectral analysis. These low-sensitivity single-layer emulsions are used in film processing for the copying of the 'primary negative'.

5.5 Physical conditions influencing blackening or density

5.5.1 Deviation from the reciprocity law

According to the *Bunsen–Roscoe reciprocity law* the photographic density is a function of the product of the light intensity I and the exposure time t, that is

$$D = f(I.t) \tag{5.19}$$

According to this law the same density is produced independent of the intensity, provided the product of the intensity and exposure time is constant.[1] Thus a straight line is obtained which is parallel to the log I axis when the value of log It is plotted as a function of log I (Fig. 5.26*a*).

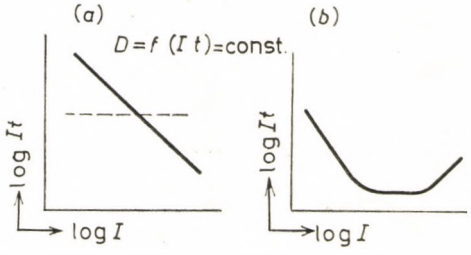


Fig. 5.26 The schematic diagrams for the reciprocity rule and the Schwartzschild equation (*a*), and variation of the value $I.t$ producing the same density level, as a function of intensity (*b*)

In reality, the reciprocity law applies only within narrow limits. It can be shown experimentally that usually a greater density arises if high-intensity light illuminates the photographic emulsion for a shorter time than when low-intensity light illuminates the emulsion according to the It law. There is therefore a real deviation from the reciprocity law.

The experimentally obtained results are generally explained if the Schwartzschild relationship

$$D = f(It^p) \tag{5.20}$$

is invoked and the value of the exponent p is restricted to between 0·7 and 0·95, that is, to less than one.[2–4] The curve corresponding to the equation is the downward sloping straight line in the diagram depicted on Fig. 5.26*a*.

The exponent p found in Eq. (5.20) – the Schwartzschild constant – can be calculated if we consider that

$$S = I_1 t_1^p = I_2 t_2^p$$

and thus

$$p = \frac{\log I_2 - \log I_1}{\log t_1 - \log t_2} \tag{5.21}$$

where I_1 is that intensity that results in the same density during exposure time t_1^p as the one brought about by light of intensity I_2 during the exposure time t_2^p.

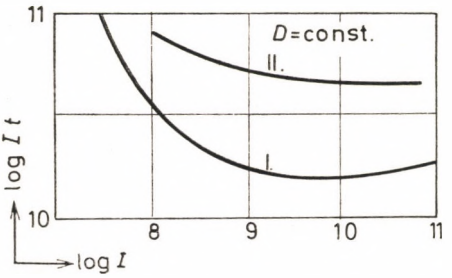


Fig. 5.27 Reciprocity failure: *I* — low-sensitivity, blue-sensitive emulsion and *II* — high-sensitivity panchromatic emulsion

The exponent p, that is, the Schwartzschild constant, is not really an invariable quantity. Its value changes not only with the quality of the photographic emulsion and the wavelength of light, but – under otherwise identical experimental conditions – it is also a function of the intensity of the light. This is because the Schwartzschild equation (5.21) is only valid for low intensity levels. Upon increasing the intensity level the curve depicting the variation of the product $I.t$ as a function of I goes through a minimum, as can be seen on Fig. 5.26*b*.[5] Therefore, the value of the exponent p at first drops linearly at low light intensity, but hardly changes upon further increasing the intensity, then on reaching the lowest value it curves upward again. The deviation from the reciprocity law is therefore smallest for a certain intensity level that corresponds to the lowest point of the curve. The $I.t$ exposure (light quantity) is also the smallest at this point. It follows, therefore, that the emulsion best utilises the exposure at this I_{min} intensity value, to produce a certain density value.

The actual magnitude of the deviation from the reciprocity law is depicted by Fig. 5.27. The curves of the two spectrochemical analytical emulsions show, however, that this deviation is only small. It follows that in the case of photographic emulsions used in spectrochemical analysis the deviation from reciprocity is usually negligible if the intensity of the light varies between limits in the proportion 1 : 100.

The deviation from reciprocity leads to the difference between characteristic curves made either by keeping the exposure time constant and varying the intensity, or by keeping the intensity constant and varying the exposure time. In the former case the equation of the straight part of the characteristic curve is

$$D_{t=\text{const}} = a + \gamma \log I \tag{5.22}$$

while that of the latter case is

$$D_{I=\text{const}} = a + \gamma \log t^p = a + \gamma p \,.\, \log t \tag{5.23}$$

The difference between the slopes of the two kinds of characteristic curve is therefore greater, the larger is the deviation of the Schwartzschild constant, p, from unity. However, p is generally smaller than unity in the intensity region used for spectrochemical analysis. It follows that the straight part of the characteristic curve referring to constant exposure time is generally of somewhat larger slope value than the straight part of the characteristic curve referring to constant intensity. From the above two equations it can also be established that the ratio of the slopes of the curves having either constant time base or constant intensity base is equal to the value of p.

References

1. ABNEY, W. W., *Proc. R. Soc.*, **54**, 143 (1894)
2. SCHWARTZSCHILD, K., *Publ. Kuffnerschen Sternw.*, **5** (1900)
3. STAR, J., *Ann. Phys.*, **35**, 461 (1911)
4. KRON, E., *Publ. Astr. Obs. Potsdam*, No. 67 (1913)
5. ANGERER, E. and JOOS, G., *Wissenschaftliche Photographie*, 6th edn, Berlin (1956)

5.5.2 The intermittency effect[1–8]

It is not unimportant in the blackening of the photographic emulsion whether the exposure $I.t$ reaches the emulsion in a continuous or intermittent manner, because

$$D = f(It) \neq f(\sum_t I\Delta t) \tag{5.24}$$

This phenomenon is caused by the so-called intermittency effect. The density brought about by the intermittent illumination is also influenced by the number of light pulses, and even by the length of the dark period between the pulses.

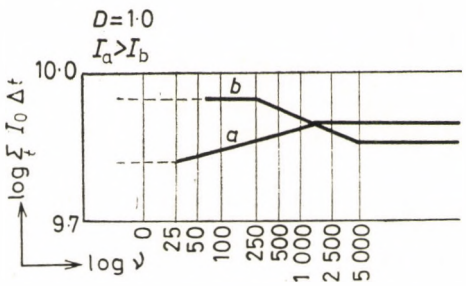


Fig. 5.28 Relationship between the intermittency effect and the frequency, in the case of low intensity (*a*) and high intensity (*b*), where the intensity unit is 4360 photon . $cm^{-2}s^{-1}$

Figure 5.28 illustrates the relationship between the intermittency effect and the frequency of the illumination. The two curves give the exposure $\sum_t I\Delta t$ that is necessary to bring about the same density $D = 1{\cdot}0$ by a larger (I_a) and a smaller (I_b) monochromatic intensity, with different frequencies of the rotating disc cutting off the light. It can be seen that when the intensity is lower an increased exposure has to be produced by increasing the cutting frequency, while when the intensity is higher a smaller exposure is needed. According to Fig. 5.29, when the frequency is small the intermittent illumination brings about a larger density, and when it is higher the continu-

ous illumination results in higher density. In each case, however, there is a 'critical frequency' above which the density is constant, that is, independent of the frequency. The value of this critical frequency depends on the intensity of the light and the wavelength. For the intensities and wavelengths usual in spectrochemical analysis the value of the critical frequency never exceeds 6000 per minute.

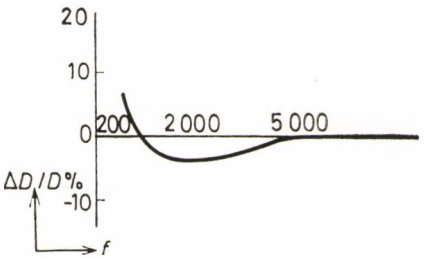


Fig. 5.29 Relationship between the *f*-number (chopping per minute) and the density difference ΔD for intermittent and continuous illumination[7]

The two curves in Fig. 5.30 illustrate the relationship between the densities produced as a result of equal intensities of light and various openings of the rotating disc. This is done by arranging that the rotation speed of one of the discs is twenty times that of the other. The frequency of illumination is therefore significantly higher in one case than in the other. As can be seen, the density becomes higher at the lower

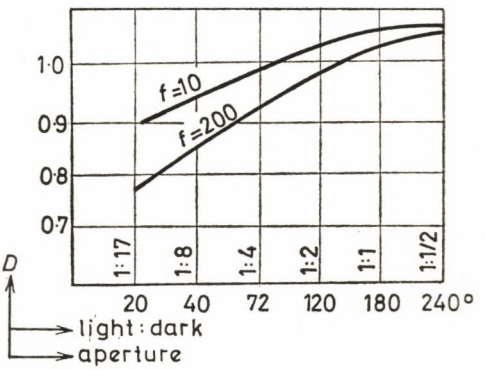


Fig. 5.30 Relationship between the density and the ratio of the light pulse and the black interval, for low and high frequencies[8]

frequencies of illumination. The difference is greater as the ratio of the illumination periods and the dark intervals becomes larger. Thus it is shown that the effect of intermittency is influenced not only by the ratio of the illumination time and the dark period, but also by the total time of these periods.

It has also been shown experimentally that the same density is brought about by intermittent illumination of frequency higher than the critical value, and continuous illumination equivalent to the same exposure.[4]

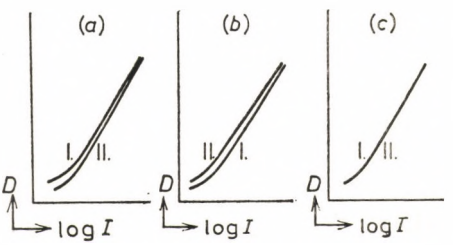


Fig. 5.31 Characteristic curves obtained by rotating disc (*I*) and by filter attenuation (*II*): (*a*) low, (*b*) medium, and (*c*) higher than critical chopping frequency[8]

From this observation the following three practical conclusions may be reached:

(1) In rotating-disc light attenuation the revolution speed of the disc is important because of the intermittency effect. It can therefore be a source of error if this speed is not constant, or if it is lower than the critical frequency.

(2) It is immaterial whether the attenuation is due to a disc rotating at a speed higher than the critical frequency or to an appropriate filter.

(3) The revolution speed of the rotating disc can influence the form of the characteristic curve. However, when the frequency is higher than the critical limit the same characteristic curve is obtained either with rotating discs or with filters (Fig. 5.31).

References

1. WEBER, A., *Ann. Phys.*, **45**, 801 (1914)
2. WEINLAND, C. E., *Phys. Rev.*, **30**, 203 (1927)
3. O'BRIEN, B. and PARKS, V. L., *Phys. Rev.*, **40**, 387 (1932); **43**, 105 (1933)
4. WEBB, J. M., *J. Opt. Soc. Am.*, **23**, 157, 316 (1933)
5. KOCH, A., *Z. wiss. Photogr.*, **36**, 97 (1937)
6. MALPICA, J. T. M., *Gen. elect. Rev.*, **43**, 288 (1940)
7. LANG, K., *Mikrochim. Acta*, **1955**, 278
8. PIERCE, W. C. and NACHTRIEB, N. H., *Ind. Engng Chem. Analyt. edn*, **13**, 774 (1941)

5.5.3 Effect of the background

The intensity measured at the location of each spectral line is composed of two kinds of contributions. The intensity measured is the sum of the true line intensity and the background intensity, which appears with smaller or greater intensity everywhere in the spectrum (Fig. 5.32). The true line intensity of a particular element line varies with the concentration of that element present in the sample, while the background intensity is independent of the concentration of that element; moreover, the intensity is constant under certain experimental conditions.[1] It is increasingly important not to neglect the background as the original brightness of the lines becomes lower and the intensity of the background becomes higher. Thus an appropriate correction is essential for low-intensity lines.

The essential correction, the background correction (see Sub-section 5.5.2), is quite simple when the intensity is determined by a procedure in which the data obtained from the measurement are directly proportional to the intensity of the light. In this case the true line intensity I_l is equal to the difference between the measured value I_m and the background intensity. In other words

$$I_l = I_m - I_b \tag{5.25}$$

The correction can be performed by a simple subtraction if the intensity measurement is performed at the location, and on both sides, of the line. Often, the background intensity measured at each side of the line is not the same (Fig. 5.32*b*). Then the most probable value, the mean of the two measurements, should be taken during the correction. However, this subtraction correction can only be used for photoelectric intensity measurements.

The background correction cannot be performed by a simple subtraction when the line intensity is determined by photography. The relationship between light intensity and density is not linear. The difference between the density directly measured on the spectrum and the true, corrected intensity of the lines is the greater, the lower is the density level. It follows that the background correction cannot be performed without taking into account the characteristic curve in the photographic intensity determination.

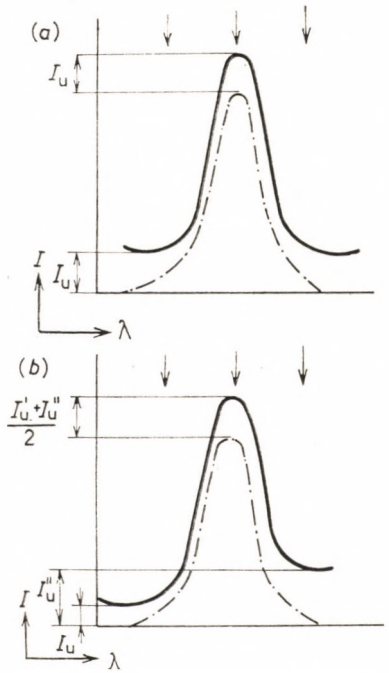


Fig. 5.32 Effect of the background for (*a*) symmetrical and (*b*) unsymmetrical backgrounds

Reference

1. PIERCE, W. C., and NACHTRIEB, N. H., *Ind. Engng Chem. Analyt. edn*, **12**, 41 (1940)

5.5.4 Callier effect[1]

The so-called Callier effect is due to the fact that the value of the measured density is greater if a parallel beam of light is used for the illumination of the photographic emulsion than if a diffuse beam is applied. The difference stems from the fact that in non-parallel illumination the silver grains in the emulsion scatter the light, thereby increasing the amount of light leaving the emulsion.

The ratio of the density D measured by a parallel beam (or by a beam whose angle of convergence is smaller than 1°) and the density measured in scattered light D' is called the Callier quotient, generally denoted by Q

$$Q = D/D' \tag{5.26}$$

The value of the quotient varies only slightly with the density but is affected by the

graininess of the emulsion. It can be generally stated that:

$$d = A.\log Q \tag{5.27}$$

This relationship applies when d is the mean grain size, measured in microns, and A is a factor depending on the experimental conditions.[2]

References

1. CALLIER, A., *Z. Wiss. Photogr.*, **7**, 257 (1909)
2. EGGERT, J. and KÜSTER, A., *Kinotechn.*, **16**, 127 (1934)

5.5.5 Graininess, resolving power, and emulsion micro-errors

The properties of photographic emulsion are less affected by the average size of the grains than by the statistical size distribution of the grains. The graininess of the emulsion causes the density of the virtually homogeneously black lines to be uneven if the cross-section of the beam passing through the emulsion is small, that is, of the order of a thousandth of a square millimetre.

The measure K of the graininess of the emulsion is usually characterised by the Callier quotient.[1] Its value is

$$K = 100 \log Q \tag{5.28}$$

The graininess of certain spectral plates is summarised in Table 5.6. It should be noted, however, that the graininess of the emulsion depends not only on the quality of the emulsion, but also on the developing process.

It can be established on the basis of the foregoing that the measured density is the average of the fluctuating 'micro-density' caused by the graininess of the emulsion. The measured average density therefore corresponds better to the intensity of the light falling on the emulsion the larger the illuminated area. Therefore, the graininess should be taken into account both in the photography of the spectra and in the measurement of the density values. In both cases the slit width should be greater for coarser-grained emulsions.

A further important characteristic of photographic emulsion is its *resolving power*. The higher this resolving power, the smaller is the distance apart at which the density of different areas of the emulsion can be distinguished, in other words, the greater is the number R of the resolved lines within one millimetre.

Table 5.7

Resolving power (R) of photographic emulsions in white light having variable average grain diameter (d) and distance (a) between neighbouring lines that are just distinguishable[2]

d (μm)	2·77	2·42	2·01	1·52	1·08	0·88
a (μm)	24	18	17	19	13	11
R (line/mm)	42	56	59	53	77	91

According to the data in Table 5.7 the resolving power of the emulsion is strongly affected by the graininess of the emulsion, but also by other factors. This is because both the resolving power and the gamma value become higher with increasing fineness of the grains. Accordingly, the resolving power is affected by all those factors that have an influence on the slope of the characteristic curve. Therefore the resolving power not only depends on the quality of the emulsion, but also on its processing – mainly the development process. The wavelength of the light is also important. The resolving power of the emulsion is greatest in the ultra-violet part of the spectrum and least for green light, that is at about 4850 Å. Nor can the role of the line density be neglected. Thus it can be established that the resolving power does not uniquely characterise the emulsion.

Local unevenness of the emulsion leads to the so-called 'emulsion micro-errors', since the density of the closely adjacent points is not the same even when the illumination of the emulsion is perfectly even. The cause is that a few per cent of scattered light can be observed between these points. The error from this source can, however, be significantly reduced when the density measurements are performed over a large area. The greater the illuminated area, the more likely it is that the local density differences will be equalised. From this fact it can be concluded that in practice it is advisable in quantitative spectrochemical analysis to photograph the spectra with the largest possible slit.

References

1. EGGERT, J. and KÜSTER, A., *Veröff. Agfa*, **4**, 49 (1935)
2. FRIESER, H. and LINKE, H., *Z. Wiss. Photogr.*, **37**, 19 (1938)

5.6 Chemical conditions influencing the blackening or density[1,2]

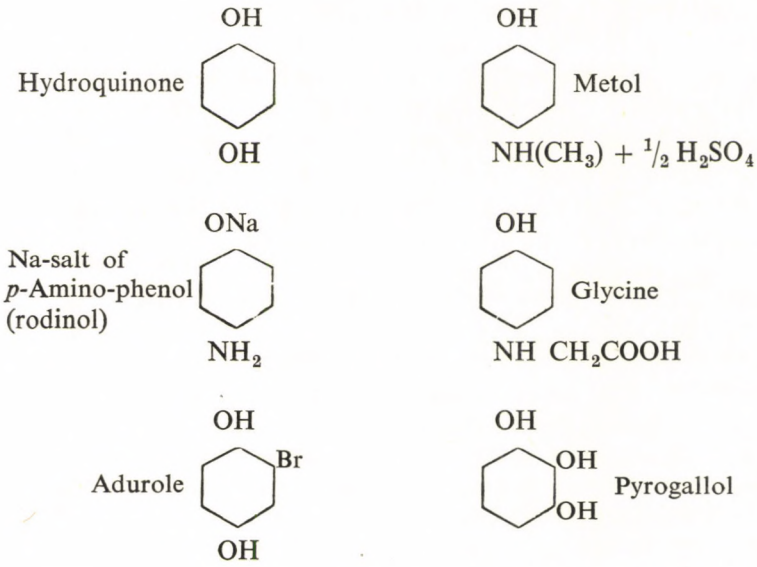
5.6.1 Developing

The density of the photographic emulsion depends not only on the exposure and the previously described physical factors but also on the chemical reactions that occur in the processing of the emulsion, and especially on the experimental conditions of development.

Chemically the essence of the development procedure is the reduction of the exposed silver halide to metallic silver: for example, the chemical reaction for silver bromide is

$$AgBr + e^- \rightleftarrows Ag + Br^- \qquad (5.29)$$

This reduction is preferably done by some weak reducing agents. The strong reducing agents cause the unexposed silver halides also to decompose. Not all weak reducing agents are suitable for this purpose. In practice – apart from the already obsolete iron(II) oxalate – the developers normally used are derivatives of benzene or naphthalene which contain either OH or NH_2 groups. Among these compounds only those are suitable which have at least two hydroxyl or two amino groups, or one hydroxyl and one amino group. The structural formulae of the usual compounds are as follows:



A solution of the compounds listed is not, in itself, a developer. It is necessary for the solution to contain, in addition, compounds that promote the developing process. This process, that is, the reduction of the silver halides to silver, occurs – as all reduction processes – together with the oxidation of the reducing compound. The end product of oxidation of the presently used developers, quinone or quinone–imine, is produced at the same time as hydrogen ions are liberated. Thus, for example, the frequently used hydroquinone is oxidised to quinone according to the following chemical reaction

$$2AgBr + C_6H_4(OH)_2 = 2Ag + C_6H_4O_2 + 2Br^- + 2H^+$$

This process is slowed down by the liberated hydrogen ions so much that the reaction occurs in practice only in alkaline solutions. Thus, for example, the effect of development is observed after 7 min at a pH of 9·2, while at a pH of 10 it can be observed after 2 min.[2] The development process can be seen to occur after only a couple of seconds when the pH of the solution is 13. On the other hand, the developing action is practically stopped in very dilute acetic acid solution. This observation is so general that other developers employed at present can also only be used in more or less alkaline solutions.

In alkaline solution the developer is easily oxidised by the air. Hence, the developer would quickly become useless, unless this process can be prevented. As a protecting agent, however, a compound can only be used that is oxidised more easily by air than the developing agent, i.e. is a stronger reducing agent. For this purpose at present sulphites are generally used, frequently sodium sulphite. This compound is oxidised to sulphate, which is neutral in the developing process.

The sulphite ions in the developer solution not only protect the developer but also promote the developing process by forming a complex compound (which also has a developing action) with the partly oxidised developer that arises from the reduction of the silver halide. The sulphite also quickens the development by slightly dissolving

the silver bromide. The silver ions produced thus are deposited upon the silver grains formed by the developer.

During the development bromide ions are also produced. The reduction of the silver halide in the heterogeneous system is increasingly slowed down locally by the presence of these ions with increasing local concentrations of the bromide ions in the emulsion. The developing error from this source is eliminated if the developer solution contains bromide ions in the appropriate concentrations. Then the amount of bromide ions produced during the reduction does not alter the original concentration of the bromide ions in the solution.

The foregoing may be summarised as follows: the developer solution normally contains, in addition to the developer proper, sodium sulphite as a protective agent, potassium bromide, and some other compound that ensures the appropriate alkaline pH.

References

1. ANGERER, E. and JOOS, G., *Wissenschaftliche Photographie*, 6th edn, Berlin (1956)
2. JAENICKE, H., *Photogr. Ind.*, **35**, 514 (1937)

5.6.2 Effect of the composition of the developer[1]

The effectiveness of the usual developers increases according to the so-called Watkins factor in the following order, and ratio, respectively

Hydroquinone	→	Glycine	→	Rodinal	→	Pyrogallol	→	Metol
1·0	:	1·6	:	6·0	:	16·0	:	20·0

This means that the developing process – under otherwise identical conditions – becomes quicker in the ratios given above.

The quick-action developers are the so-called *surface* developers. In their use the development process is already finished in the external layers of the emulsion when it has just been initiated in the deeper layers. This is because the fresh developer diffuses relatively slowly into the interior of the emulsion. On the other hand, the bromide ions produced by the reduction of the silver bromide that slow down the pace of the development leave the site of development only reluctantly. Thus the photochemical effect of the lower-intensity rays, which only enter the upper layers of the emulsion, is complete, whereas some of the energy of the higher-intensity rays that penetrate into the depth of the emulsion is lost. Developers of this type, therefore, partly level out the intensity differences, that is, they work *softly*. Accordingly, the straight part of the characteristic curve will be steeper if the concentration of hydroquinone relative to metol is increased in the developer solution.

Strong alkalis also speed up the developing process. The alkaline hydroxides (NaOH, KOH) are the most effective agents, while the alkaline carbonates (Na_2CO_3, K_2CO_3) do not produce such a large speeding-up effect. There is a much smaller speeding-up effect if sodium phosphate ($Na_3PO_4.12H_2O$) and especially borax ($Na_2B_4O_7.10H_2O$) is used. It follows that using the same developer a faster and higher-contrast type of developing solution is obtained if the solution is made alkaline by adding sodium or potassium hydroxide. However, this increases the coarseness of the grain.

By increasing the concentration of potassium bromide the speed of development is generally reduced. This effect is different for different developers. Thus, for example, to achieve the same rate of development about one-twentieth of the amount of potassium bromide is necessary if metol is used instead of hydroquinone. This effect of the potassium bromide is so characteristic that it used to be regarded as the measure of the reducing capability of developers.[2]

By increasing the potassium bromide concentration of the developer the straight part of the characteristic curve becomes steeper.[4] This is because the speed of development is more reduced by the bromide ion produced during the development (at places receiving the larger intensities) with a lower potassium bromide content of the developer. Since this effect is greatest at points with low densities, a lower gamma value is achieved. The retarding effect of the bromide ion produced during development can be neglected, even for places that have received high intensities, if the potassium bromide concentration of the developer is sufficiently high. However, larger density differences are produced, that is, the straight part of the characteristic curve will be steeper. This is because by increasing the concentration of the potassium bromide the development process is slowed down and thus the necessary development time is lengthened. This, in turn, necessitates the exchange of the used and fresh developer by diffusion.

An increase in the concentration of the potassium bromide has another effect; for any particular time of development (up to that time causing maximum density) the absolute density of the spectrum is lower, the higher the concentration of the potassium bromide. Therefore, to obtain the normally required levels of density higher light intensities are necessary when using high concentrations of potassium bromide in the developer. It follows that the fogging density of the emulsion becomes lower with increasing concentrations of potassium bromide. The foregoing may be summarised thus: by increasing the concentration of the potassium bromide the speed of development decreases but simultaneously the contrast of the emulsion is increased, while the fogging density is reduced.

It follows from this discussion that by appropriate variation of the concentration of the compounds mentioned above, developer solutions can be prepared whose properties can be altered within wide limits. In practice, however, it is not necessary to use all the compounds mentioned above for these developer solutions.[3] Developers appropriate for spectrochemical purposes, with different properties, can be made by using metol and hydroquinone, if the proportion and concentration of the two compounds, the potassium bromide content of the solution, and the alkalinity are varied as required.

It follows that the composition of the developer influences the relationship between light intensity and density. One of the fundamental controlling factors in the production of spectral photographs of reproducible density is the identical composition of the developer.

References

1. ANGERER, E. and JOOS, G., *Wissenschaftliche Photographie*, 6th edn, Berlin (1956)
2. NIETZ, A., *The Theory of Development* (1928)
3. CVONAT, J., *Z. Wiss. Photogr.*, **22**, 234 (1934)
4. TÖRÖK, T., *Annls Univ. Scient. Budapest Rolando Eötvös Sect. Chem*, **2**, 415 (1960)

5.6.3 Influence of the duration and temperature of development

Increased development times cause the density and also the slope of the straight part of the characteristic curve to increase and the threshold value to drop (Fig. 5.33).

It can also be generally shown that when relatively short development times are employed, the density of the low-intensity lines almost reaches a maximum while the density of the more intense lines hardly exceeds half the maximum value (Fig. 5.34). Thus, when an emulsion is developed to less than the greatest possible density there is a much smaller difference between the density of the spectral lines than that on an emulsion developed so as to attain the maximum density.[2]

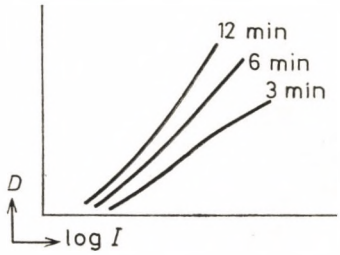


Fig. 5.33 Influence of development time on the characteristic curve[1]

The slope of the straight part of the characteristic curve, that is, the value of gamma, changes unevenly during the development process. The gamma value, γ_{max}, increases rapidly at the start of the development, then there is a continuing drop in this rate of increase until the maximum gamma value is reached. After a lapse of time t the value of γ_t will be

$$\gamma_t = \gamma_{max}[1 - \exp(-kt)] \tag{5.30}$$

where k is the rate constant of the development reaction.[3] Therefore, the value of gamma is influenced by all those factors that determine the rate of development (see Section 5.6.2). This variation in the gamma value is explained by the relationship

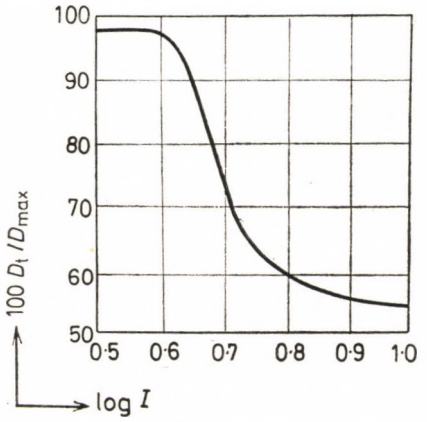


Fig. 5.34 Percentage ratio of the density D_t to the maximum density D_{max} for a two-minute development as a function of the intensity of light[2]

given earlier. This is a relationship between the intensity I of light, on the one hand, and the percentage ratio of the density D_t produced in a relatively short time t and the maximum density D_{max}, that is, 100 D_t/D_{max}. This is depicted on Fig. 5.34.

It can also be shown experimentally that the above rule, according to which the value of γ-approaches its maximum as the development time is increased, has a limited application. The value of γ-increases not only with development time, but also with an increase in the wavelength of light. It can happen that the γ-value valid for a longer wavelength drops after a 'critical' development time (Fig. 5.35). This anomaly is found, however, only in spectral ranges within which the value of γ-varies very strongly as a function of the wavelength.[4]

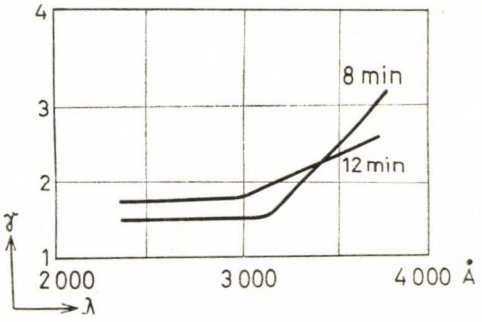


Fig. 5.35 Variation of the γ-value with wavelength for different development times of the 'Orwo blau extrahart' spectral plate[4]

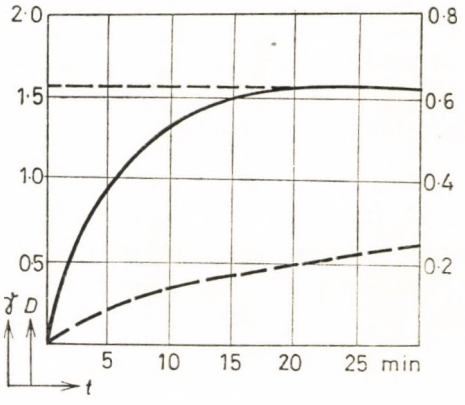


Fig. 5.36 Variation of the slope of the characteristic curve and the fog level as a function of the development time[1]

In Fig. 5.35 it can also be seen that by lengthening the development time the fogging density of the emulsion is continuously increased. An over-long development can cause the less intense spectral lines to disappear into the background. This gradual increase of the fogging density can be explained by the fact that the developer very slightly reduces the unexposed silver bromide grains. This effect is increased during the development period, the smaller is the number of exposed, but not yet reduced, silver bromide grains.[1]

From the foregoing it is apparent that lengthening the development time is partly advantageous and partly disadvantageous. However, a development time can always be found for which the slope of the straight part of the characteristic curve is sufficiently large but the fogging density is low. This is usually five minutes in normal spectral analyses.

From the foregoing it follows that if the specified development time is not correct, an error will develop in the photographic intensity determinations. Not only the absolute density, but also the relative density can deviate from the intensity or relative

intensity value. The development time should therefore be observed to an accuracy of $\pm 0{\cdot}1$ min.

The rate of the development reactions varies with the temperature of the developer in such a way that by increasing the temperature the development time necessary to attain a certain density is reduced. The effect of increasing the temperature is, therefore, the same as that of increasing the development time. Thus with increasing temperature not only the density of the spectral lines is increased – for a given development time – but the graininess and the fogging density of the emulsion are also increased. Therefore, a too high developer temperature – that is, one that deviates significantly from the usual temperature of approximately 20 °C – is harmful. On the other hand, when the temperature is less than 18 °C (say 16 °C) the development process is significantly retarded and may almost come to a halt. Furthermore, the variation of the development temperature affects to a different extent the rates of the reduction reactions of the different compounds used as developers. For example, hydroquinone is much more sensitive to the temperature being lowered than is metol. The development properties of the metol–hydroquinone solution generally used change when the temperature is reduced in such a way that the solution works as if it contained only metol. However, no error is caused when a certain development temperature, say 18, 19, or 20 °C, is chosen and maintained. One should, however, ensure that the temperature of the developer is exactly the one that is specified both at the beginning and at the end of the developing process. With increasing temperature there is a linear change in the γ-value between certain limits, for example, when using the emulsion 'ORWO blau extrahart' in the spectral range 2700–3000 Å, the value of γ-increases from about 1·0 to 2·2 between 17 and 23 °C.[5]

References

1. ANGERER, E. and JOOS, G., *Wissenschaftliche Photographie*, 6th edn, Berlin (1956)
2. STROCK, L. W., *Spectrochim. Acta*, **1**, 123 (1939)
3. NACHTRIEB, N. H., *Principles and Practice of Spectrochemical Analysis*, New York–London, 113 (1950)
4. ZIMMER, K., TÖRÖK, T. and BIEDL, J., *Acta chim. Acad. Sci. Hung.*, **51**, 17 (1966)
5. GERBATSCH, R. and BACHER, J., *Z. analyt. Chem.*, **225**, 81 (1967)

5.6.4 Agitation of the developer and emulsion macro-error

In all heterogeneous reactions it is important that there should be a constant renewal of the solution on the surface of the solid phase. Therefore, it is important in the developing process that the developer should be in a constant, and if possible even, motion during development. The compounds needed for the chemical processes are brought by diffusion into the swollen gelatin layer, while the end-products leave it also by diffusion. When the developer is agitated there is a steady renewal of the developer solution in contact with the gelatin. The agitation, therefore, makes the diffusion faster and more even. The diffusion process is only perfect, however, when the solution moves with the same speed over the whole surface of the emulsion. If this condition is not met concentration differences can develop in the solution in contact with the surface of the emulsion. This can lead to the diffusion rate and hence the development reaction rate being different from place to place. In this way locally different densities can develop even when the emulsion is evenly illuminated.

This difficulty may often be experienced in the development of photographic emulsions. The error due to this source is called the *macro-error.*

The presence of the macro-error is indicated by the density of the evenly illuminated emulsion not being even. This can occur even if the emulsion is developed in a tray that is large compared with the dimensions of the emulsion. The value of the macro-error from this source can reach 10–15% even in favourable cases. The distribution of this error is well illustrated by Fig. 5.37 which shows three-dimensionally the density values measured at different points of the emulsion. It can be seen that the density is greatest at the edges perpendicular to the direction of the movement of the developer, while it is smallest at the centre. The same phenomenon can be observed if the experiment is repeated with a sheet cut into halves at the centre. The unevenness of the density therefore is not caused by the unevenness of the emulsion but is due to a development error.[2]

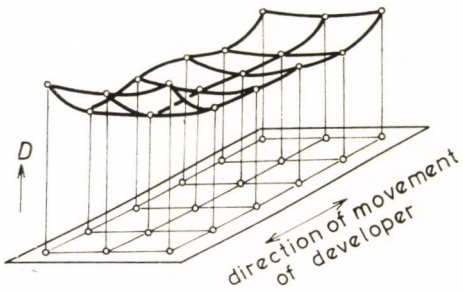


Fig. 5.37 Form of a macro-error in photographic emulsions[1]

The unevenness of the development stems from the fact that the agents necessary for development are present simultaneously only at a low concentration in the gelatin layer. The reaction products retarding development are most rapidly exchanged for fresh developer at the points where the agitation of the developer is most vigorous. This happens mostly at the edges perpendicular to the streaming direction of the developer solution. At these points whirling motions are induced in the developer when the developing tray is swayed to the other side and the streaming direction of the solution reverses. This is the reason why at this point of the evenly illuminated emulsion the density is higher than the average level.

This anomaly is explained by the fact that at the points where whirling motion occurs the bromide ions retarding the development process, and arising from the reduction of the silver bromide, are removed faster than at other points of the emulsion. This statement is proved by the reduction of the macro-error according to Fig. 5.38 when the concentration of the potassium bromide dissolved in the developer, as compared with the concentration of the other compounds present, and the time of development are increased in such a way that the same density is obtained.[1]

It follows from the above discussion that from the point of view of the density produced in the developing process it is important how and whether the developer is agitated. According to the figures given the disturbing effect of the macro-error is best eliminated when the solution next to the emulsion is streamed evenly perpendicular to the spectral lines, along the length of the spectral plate. From the diagrams given it can be seen that for photographic intensity determinations the photographing of spectra close to the edges of the emulsion should be avoided. In practice it is advisable to leave a 10–15 mm blank margin.

The developer can be agitated in a number of ways. The simplest way, but the least recommended, is to *rock* the developing tray; that is, the side of the tray parallel to the spectral lines is regularly lifted and lowered. The development process is much more even if the rocking of the tray is performed by an electric motor, for example one that drives an eccentric roller.

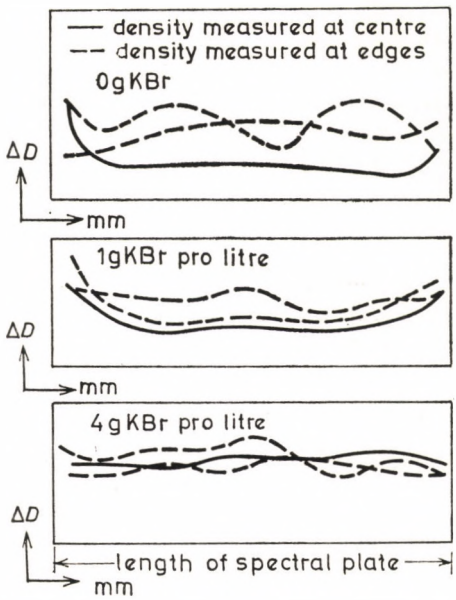


Fig. 5.38 Influence of KBr on a macro-error in the emulsion[1]

More homogeneous development can be achieved by *brush development.*[3] In this process the solution adjacent to the surface of the gelatin is removed by a wide soft brush that is regularly moved perpendicularly to the spectral lines. The drawback of the process is that it demands great skill and practice. This drawback is, however, avoided if the brush is moved by an electric motor. The difference in the efficiency of rocking development as opposed to brush development is such that the density values measured at different points of an emulsion (exposed to even illumination) can be different by as much as $\pm 15\%$ from the average after development by rocking, but only by $\pm 3\%$ after brush development.[4]

According to more recent experiments it seems that the smallest possible development error can be obtained by the turbulent streaming of the developing solution. This can be achieved most simply when the plate is held by an appropriate device, immersed in the developer solution, and vibrated by means of an electro-magnet.[5,6]

References

1. Török, T., *Annls Univ. Scient. Budapest Rolando Eötvös Sect. Chem.*, **2**, 415 (1960)
2. Schöntag, B., Dissertation, München (1936)
3. Dehio, H., *Der Phototechniker*, **10**, 3 (1941)
4. Moritz, H., *Spektrochemische Betriebsanalyse*, 2nd edn, Stuttgart, 78 (1956)
5. Török, T., Heltai, Gy. and Moharos, I., *Colloquium Spectroscopium Internationale, Madrid (1969)*. Proceedings of Plenary Conferences, London, p. 171
6. Török, T. and Heltai, Gy., *Spectrochim. Acta*, **27B**, 215 (1972)

5.6.5 The Eberhardt effect[1]

This sort of development error is encountered if closely adjacent points of the emulsion are illuminated by light of low and high intensity. The resulting Eberhardt effect is demonstrated by the phenomenon that the densities of adjacent points are not independent of each other. At the periphery of the area illuminated by the high-intensity light a darker strip is observed, while next to the area receiving the lower-intensity light a lighter strip is formed (Fig. 5.39). This uneven density distribution can spread over a width of about 1·5 mm, depending on the width of the illuminated area.

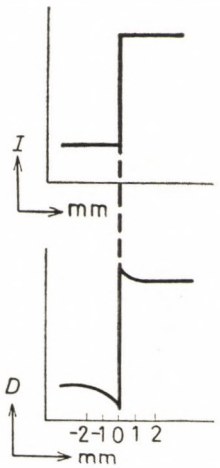


Fig. 5.39 Illustration of the Eberhardt effect

The Eberhardt effect is caused by diffusion. During development the fresh and used developer diffuses not only through the solution–emulsion border but also within the gelatin layer, provided there are larger concentration differences prevailing within the latter. Thus there is diffusion of the fresh developer towards the areas that had been strongly illuminated – which consume more developer solution – not only from the direction of the surface of the emulsion, but also from the neighbouring, less intensely illuminated emulsion. In this way the development process is made faster on the edge of the part illuminated by the more intense light. Therefore, in the case of development of a specific length of time a greater density is produced at these points. On the other hand, the used developer diffuses towards the less intensely illuminated parts of the emulsion and therefore contains a higher concentration of bromide ions, which in turn retards the development reaction. Thus a lighter strip is produced at the border of the less intensely illuminated part of the emulsion.

The most important consequence of the Eberhardt effect is that the density of spectral lines of equal intensity, but situated in different 'environments', is not identical. The density of lines in the neighbourhood of strong lines is smaller than that of equal-intensity lines situated in a free 'environment'. The Eberhardt effect can also cause the background density to drop.

The Eberhardt effect causes significant errors on spectral records made with step-filters (rotating sectors). The source of the error is such that the density measured at the edge of the neighbouring step does not exactly correspond to the intensity belonging to the step being measured. A further error may be introduced

by the density of the stronger background. (In quantitative analysis these anomalies can alter the slope of the evaluation curve.)

A further consequence of the Eberhardt effect is that the density profile of the broader spectral lines does not truly follow the actual intensity changes (Fig. 5.40).

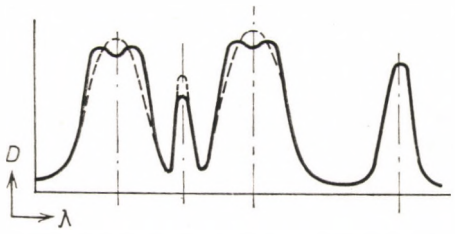


Fig. 5.40 Density profile of spectral lines distorted by the Eberhardt effect

The density reaches its maximum value somewhere near the edges instead of at the centre. This distortion of the line profile can cause an error in the determination of the wavelength if, instead of the centre of the lines, the maximum density is considered. The Eberhardt effect can distort the line profiles in a one-sided manner. This is called the Kostinsky effect. This effect can produce an error in the measurement of the distance between two lines, if the density profile of the two unsymmetrical lines is distorted in opposite directions (Fig. 5.41).

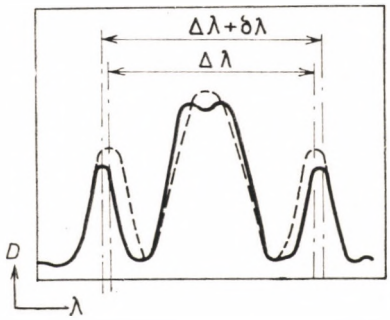


Fig. 5.41 Kostinsky effect

It may also be traced back to the Eberhardt effect that on the atomic and ion spectra of elements present at high concentrations in the sample the fundamental lines are not sharp but have *wings*. These wings are formed by diffusion at the expense of the density of the spectral lines. From this observation it follows that for the purpose of quantitative spectral analysis only sharp-profile analytical lines, with widths corresponding to the slit width of the spectrograph, may be used.

Reference

1. EBERHARDT, G., *Phys. Z.*, **12**, 288 (1912)

5.7 Straightening of the characteristic curve

5.7.1 General considerations

The straight part of the characteristic curve is characterised by the fact that to the same density differences ΔD the same differences of the logarithm of the intensity ΔY correspond. It may therefore be written

$$\Delta D = \gamma . \Delta Y \tag{5.31}$$

where γ is the slope of the straight part of the characteristic curve, and $Y = \log I$. Therefore, it would be advantageous if this simple relationship, which makes the measurement of intensity very easy, could be applied between broader limits.

In the spectral range that is most important from the point of light-intensity measurements, the middle ultra-violet range, the straight part of the characteristic curve starts at about 0·4–0·5, while its upper limit usually varies between 1·6 and 2·0. This statement applies if the normal density limits – that is, limits within which values can be measured to the required accuracy with the usual spectral slit widths and line widths – are taken into account. The usable part of the actual characteristic curve – the straight part – corresponds to only a unit density difference.

The under-exposed, bent part of the characteristic curve can be almost straightened if instead of the actual density the so-called *transformed density* is used. In this procedure a transformed density corresponding to about 0·1 density unit is utilised. Therefore, the transformation procedure increases by almost 100%, the Y-region belonging to the straight part of the characteristic curve.

In addition to the advantages formerly mentioned it is also somewhat favourable that the photographic errors are reduced by dropping the density of the spectral lines. And it is also advantageous that the lower density values can be measured to a much higher accuracy than the higher ones.

According to the foregoing, the essential thing in the density transformation is that to the straight and under-exposed part of the characteristic curve such P-values are assigned as are in a linear correlation with the Y-values. Accordingly, this procedure is called the *straightening* or *linearisation* of the characteristic curve.

The above assignment refers, in fact, to the functional relationship between the quantities D and P. Thus the following may be written

$$Y \pm \delta Y = l(P) \tag{5.32}$$

where the symbol l is the function maintaining the linear relationship between the quantities P and Y, and the additive term δY expresses the fact that the Y-values obtainable by the linear relationship and the true Y-values are mathematically not the same. However, it should be noted that no linear function l can be found that would make the physical quantity Y mathematically precise. One has to be content with a function that when applied does not cause the value of the error δY to exceed the allowed limit.

From the practical point of view it is an important further requirement that the function be as simple as possible, that is, its application should not require complicated operations. As a result, the transformation[1] has not become widespread in practice. The situation is radically changed when a computer is employed for the transformation. In this case more complicated transformation functions can be used. It is, further-

more, no problem if a number of parameters are required for the transformation.

Reference

1. CANDLER, C., *Spectrochim. Acta*, **8**, 262 (1956)

5.7.2 The *W*- or Seidel transformation[1,2]

The so-called *W*- or *Seidel transformation* is based on the observation that the lower, bent part of the characteristic curve valid for a given wavelength becomes straight if instead of the true density $D = \log I_0/I$ the following expression is used

$$W = \log\left[\frac{I_0}{I} - 1\right] \tag{5.33}$$

where I_0 and I stand for the intensities of the rays passing through the unexposed and exposed parts of the spectrum, respectively (Fig. 5.42).

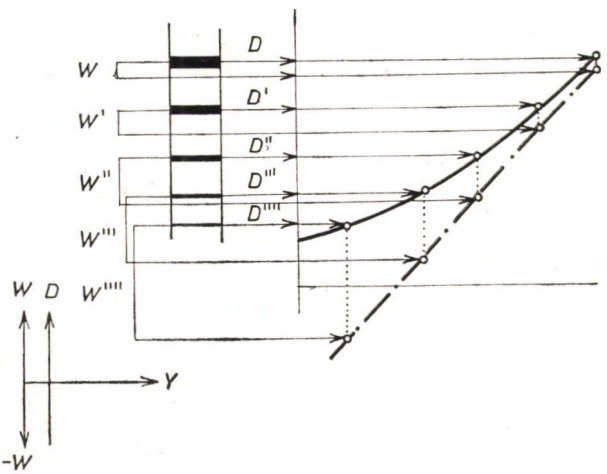


Fig. 5.42 The *W*-transformation: original characteristic curve (solid line) and transformed characteristic curve (dotted line)

The difference between the true and transformed densities is the Seidel correction, *G*. Its value is therefore

$$G = D - W = \log\left[\frac{I_0}{I_0 - I}\right] \tag{5.34}$$

and it also has a significant role in the background correction. From the above discussion it is apparent that the value of the Seidel correction is the same as the so-called Gauss differential logarithm. When it is assumed that $D = \log x$, then $W = \log(x - 1)$, thus $G = \log x - \log(x - 1) = D - \log(\text{num log } D - 1)$. The *W*-transformation, based entirely on observations, is limited in its application. Its applicability depends on a number of factors: the wavelength of the light, the properties of the photographic emulsion, and even on the method of processing the

emulsion. For example, in the case of the 'blue-sensitive' emulsion frequently used in spectral analysis, the W-transformation can only be applied in the range 3400–3900 Å.[2] In the range 2400–3300 Å an over-correction occurs, that is, the 'straightened' part of the characteristic curve is not linear, but convex from above. On the other hand, in the visible part of the spectrum the W-transformation only partly reduces the curvature in the lower part of the characteristic curve, so that it does not

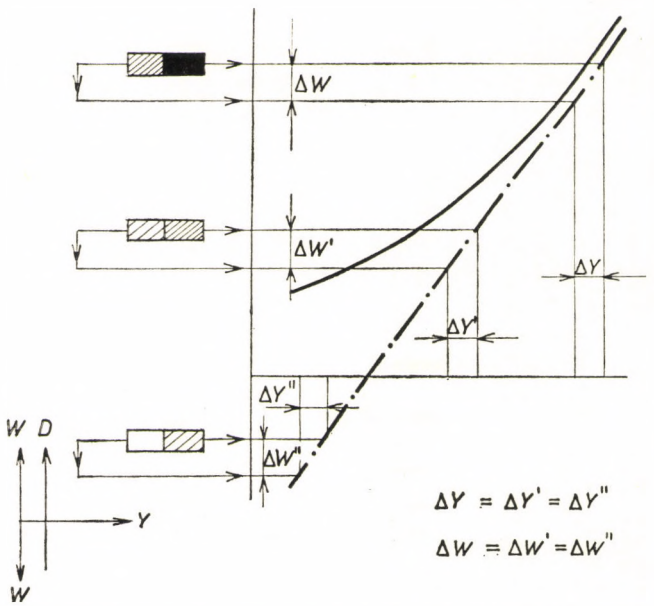


Fig. 5.43 Checking the W-transformation

provide the necessary correction. In the use of the W-transformation therefore one should always confirm experimentally that the corrected part of the transformed characteristic curve can be taken as a straight line. This can be controlled most simply by means of a two-step filter used in making the spectral record. The correction is appropriate at that part of the characteristic curve on which to the same Y-values the same $\Delta W' = \Delta W'' = \ldots$ values correspond (Fig. 5.43).

References

1. SEIDEL, W., *Sitzung des spektralanalytischen Ausschusses der Gesellschaft Metall und Erz* (1939) (Lecture)
2. KAISER, H., *Spectrochim. Acta*, **2**, 1 (1941)

5.7.3 The $P_{1/2}$-transformation[1]

The W-transformation leads to over-correction in just that spectral range that is most important for quantitative spectral analysis. In order to avoid this drawback the so-called $P_{1/2}$-correction was introduced. This is the arithmetical mean of the true density D and the corrected density W, according to Seidel (Fig. 5.44). Its value is therefore

$$P_{1/2} = (D + W)/2 = [D + (D - G)]/2$$

That is

$$P_{1/2} = D - \frac{G}{2} \tag{5.35}$$

The hundredfold W-, $P_{1/2}$-, and G-values corresponding to the hundredfold density values D are compiled in the form of tables.[1,3] It is seen from these tables that the W- and $P_{1/2}$-values can be negative as well as positive. This change of sign occurs in the W-transformation at density value $D = 0{\cdot}301$. When $W = 0$, then $D = \log 2$ and thus $W = \log(2 - 1) = 0$. On the other hand, the value $P_{1/2} = 0$ corresponds to $D = 0{\cdot}209$ [in this case $D + (1/2)G = 0$]. When the density value drops further the value of the transformed densities drops steeply, while finally at $D = 0$, $W = -\infty$, and thus $P_{1/2} = -\infty$ as well.

Furthermore, it can be established from the tables that, for example, to the density $D = 1$, W has a value of $0{\cdot}954$, that is $G = 0{\cdot}046$. To neglect this would lead to an error of about 5%, while to neglect the $P_{1/2}$-transformation would lead to an error of about 2.5 %. On the other hand, $D = 0{\cdot}500$ corresponds to $W = 0{\cdot}335$; therefore the difference is already $G = 0{\cdot}165$! It should be noted, however, that in quantitative spectral analysis not even deviations corresponding to values $D = 1$ can be neglected. Otherwise an analytical error of magnitude $\pm 10{\cdot}6$ and $\pm 5{\cdot}3$ relative per cent would be caused, provided the slope of the evaluating line (calibration curve) is 45°.[2]

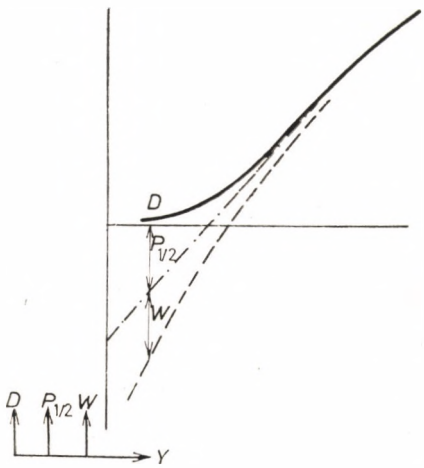


Fig. 5.44 The $P_{1/2}$-transformation

It can furthermore be shown experimentally that the $P_{1/2}$-transformation does not give an adequate correction in the upper ultra-violet spectral range. The corrected characteristic curve is not straight, but is concave from above.[2]

References

1. HONERJAGER-SOHM, M. and KAISER, H., *Spectrochim. Acta*, **2**, 396 (1944)
2. KAISER, K., *Spectrochim. Acta*, **3**, 159 (1947)
3. Zeiss-Sonderdruck, Mess. 41-360-1 Jena

5.7.4 The generally valid partial density transformation

The *P*-transformation can be used at any wavelength if instead of Eq. (5.35) the following equation is used:

$$P = D - \kappa G \tag{5.36}$$

that is, when G is multiplied by a κ-value, chosen according to the wavelength instead of one-half. This factor is called the *transformation constant*.[1]

The value of the transformation constant depends partly on the photographic emulsion used, the conditions of development, and on the wavelength. For the determinations two intensity signals I_a and I_b are necessary whose ratio is constant, that is, I_a/I_b = constant.

In the application of this procedure the optimum conditions are when the spectra are photographed by means of two-step filters. In this case the intensity signals are provided by the two filter steps. When the density values D'_a and D'_b, falling on the straight part of the characteristic curve are measured on appropriate spectral lines, while other lines are used for measuring the densities on the under-exposed part D''_a and D''_b, then – according to Eq. (5.36) – the following may be written

$$P'_a = D'_a - \kappa G_{D'a} \qquad P''_a = D''_a - \kappa G_{D'a}$$

$$P'_b = D'_b - \kappa G_{D'b} \qquad P''_b = D''_b - \kappa G_{D'b}$$

When the indices are left, on the basis of the above equations, the difference of the two transformed densities is given by: $\Delta P' = \Delta D' - \kappa \Delta G'$ and $\Delta P'' = \Delta D'' - \kappa \Delta G''$. Furthermore, when it is taken into account that $\Delta P' = \Delta P''$, then we obtain from the previous relationships for the value of the transformation constant

$$\kappa = \frac{\Delta D' - \Delta D''}{\Delta G' - \Delta G''} \tag{5.37}$$

Therefore, in order to determine the transformation constant it is necessary to establish the density difference, on two identical steps, of two spectral lines in the spectral range studied – on spectra made by means of a two-step filter. These are the values: $\Delta D' = D'_a - D'_b$ and $\Delta D'' = D''_a - D''_b$. In addition, it is necessary to establish the corresponding correction differences: $\Delta G'$ and $\Delta G''$. The procedure is relatively simple.

According to more precise studies it has been established that while in very different emulsions the value of the transformation constant lies in the region 0–2, as a func-

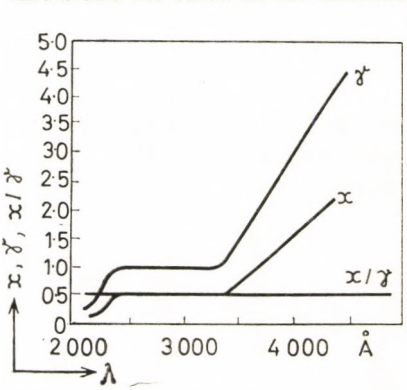


Fig. 5.45 Variation of the γ and κ values, and the ratio κ/γ of the Agfa 'Spektralplatte, blau extrahart' as a function of the wavelength

tion of the wavelength, the value of the slope γ varies between 0 and 5 on the straight part of the characteristic curve. Hence, the value of the ratio κ/γ is approximately independent of the wavelength (Fig. 5.45). The value of the latter varies between the limits 0·45–0·75 in widely differing emulsions.[2]

The *P*-transformation is, strictly speaking, not exact since the transformed characteristic curve produced by it is not straight.[3] This is partly due to the fact that the procedure also corrects the straight part of the characteristic curve. Since, however, the correction function is not linear, the resulting characteristic curve is also not linear. The deviation is also partly caused by the fact that the corrected line belonging to the under-exposed part of the characteristic curve is not only an imperfect straight line, but also does not form a straight continuation of the upper part of the transformed characteristic curve. Therefore, the slope of the two parts is different and on the transitional part between them there is usually an inflection. The density value belonging to the inflection point is usually the same as the limit of the straight and bent parts of the original characteristic curve. Furthermore, the characteristic curve obtained by the *P*-transformation deviates more from a straight line at longer wavelengths.

References

1. KAISER, H., *Spectrochim. Acta*, **3**, 159 (1947)
2. TÖRÖK, T. and ZIMMER, K., *Acta Chim. Hung.*, **29**, 273 (1961)
3. ZIMMER, K. and TÖRÖK, T., *Acta Chim. Hung.*, **28**, 59 (1961)

5.7.5 The *l*-transformation[1]

This procedure completely eliminates the faults of the *P*-transformation.[1] The corrected characteristic curve obtained by the *l*-transformation is a straight line formed by the straight part of the original characteristic curve and its continuation (Fig. 5.46).

The above requirement is met if Eq. (5.36) of the original *P*-transformation is modified so that the value of κ in the relationship $P_L = D - \kappa' G$ is *not* independent of the *D*-value of the density.[2] When the straight part of the original characteristic curve starts at D_L, then for $D > D_L$ the transformed characteristic curve P_L coincides with the straight part of the true characteristic curve only if the condition $\kappa G = 0$

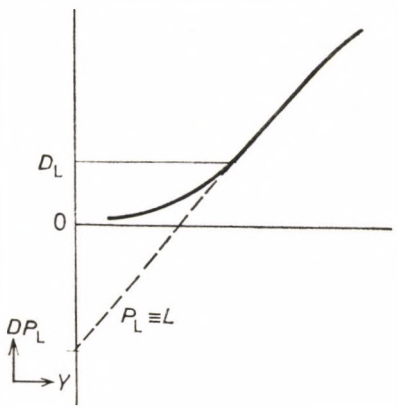


Fig. 5.46 Characteristic curve obtained by the *l*-transformation

is met, that is, when $\kappa' = 0$. On the other hand, in the case of densities falling on the lower part of the original characteristic curve $D < D_L$, thus it is necessary for the condition $\kappa > 0$ to be met. In the latter case the value of κ is not independent of the D-value.

The relationship between the κ- and D-values is influenced by the wavelength in so far as the slope of the straight part of the characteristic curve, or more precisely the γ-value, varies with the wavelength. It can be seen on the curves shown in Fig. 5.47*a* that the simple, linear relationship between κ and D is valid only when $\gamma = 1$. In the opposite case – that is, when $\gamma \neq 1$ – the relationship is complicated.

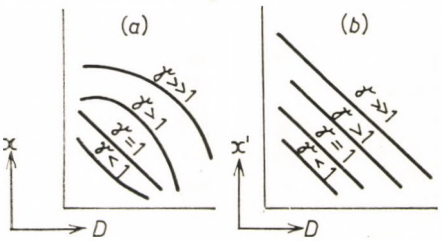


Fig. 5.47 Schematic variation of the κ- and κ'-values as a function of the density D, for different gamma values

The above relationship becomes linear when the Gauss differential logarithm is assigned to the values of D/γ instead of D, in other words, when instead of the D-correction function in the P_L-transformation equation the $G_{D/\gamma}$-value is used. If the value obtained thus is denoted by L then the equation of the full transformation is[3]

$$L = D - \kappa'' G_{D/\gamma} \tag{5.38}$$

As is seen from Fig. 5.47*b*, under such conditions the relationship between κ'' and D is linear for all wavelengths.

It is recommended that both sides of the above equation are divided by γ and that at the same time the following notation is introduced

$$\frac{D}{\gamma} = d, \qquad \frac{L}{\gamma} = l, \qquad \text{and} \qquad G_{D/\gamma} = g \tag{5.39}$$

when we obtain the following relationship

$$l = d - \frac{\kappa}{\gamma} g \tag{5.40}$$

It can be seen from Fig. 5.48*a* that – as opposed to the case of the L-transformation – the straight lines indicating the dependence of κ''/γ on d lie very close to one another. This means that the relationship previously given is almost independent of the gamma value and thus of the wavelength of the light.

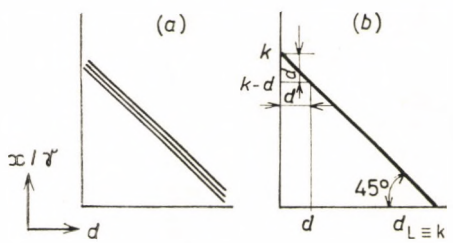


Fig. 5.48 Schematic variation of the quotient κ/γ as a function of D [1]

Closer studies have shown that the slope angle of the straight lines previously mentioned is 45°, not only in this case, but also in the case of the *L*-transformation.[1] This, in turn, has the result that the κ''/γ value obtained by extrapolation, and belonging to the $d = 0$ value, is numerically the same as $D_L/\gamma = d_L$, that is, if the notation k is introduced we obtain

$$D_L/\gamma = d_L = k \tag{5.41}$$

When the value of k is known, the value of κ''/γ belonging to the d-value can be written in the following way, according to Fig. 5.48*b*

$$\kappa''/\gamma = k - d \tag{5.42}$$

Thus the value of l belonging to d, that is, the equation of the l-transformation is

$$l = d - (k - d)g \tag{5.43}$$

provided the relationships (5.40) and (5.41) are taken into account.

The above transformation equation is valid only when $d \leqq k$. By careful studies it has been established that the value of k is entirely independent of the wavelength.[1,4]

The application of the l-transformation is made easier by the fact that the l-values corresponding to the different d-values are compiled in tables. There is a separate table for each reasonable k-value.[5] The value of k varies between 0·15 and 0·45 for emulsions of different make, sensitivity, and contrast. Its value reaches 0·55 only in the case of old, badly stored spectral plates of high fogging density.[6] Occasionally k may have an even higher value.

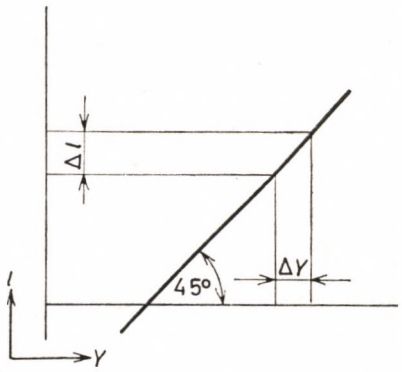


Fig. 5.49 Relationship between the values ΔY and Δl

With the help of the l-transformation the logarithm of the intensity relation of two spectral lines, that is the ΔY-value belonging to the line-pair, can be simply established. The slope angle of the transformed characteristic curve is 45°, thus it follows from Fig. 5.49 that

$$\Delta Y = \Delta l \tag{5.44}$$

In the application of the l-transformation one should first determine the γ-value belonging to the wavelength of the analytical line. This can be done most simply with the aid of the two-step filter method (described earlier), by taking into account the relationship $\Delta D_{max}/\Delta Y_m = \gamma$. Then the D_a and D_b-values are divided by γ, in other words the corresponding d_a and d_b-values are calculated. When the latter are known, an *iteration* procedure is applied to find those k-values at which the condition $\Delta l = \Delta Y_m$ is met, provided $d_b < k$. The value of k can be easily found. By

considering the ΔD-values the largest d_b-value, at which the value of $\Delta D_{a,b}$ starts dropping, can be established immediately. The d_b-value corresponding to this D_b-value is the same as the initial (approximate) k-value (Fig. 5.50).

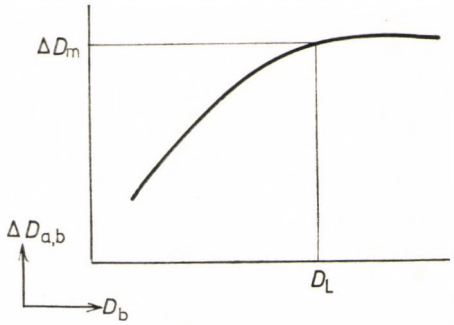


Fig. 5.50 Principle for the graphical determination of the D_b value necessary for obtaining the approximate value of k

It can be shown experimentally that the accuracy of the l-transformation is higher than that of the P-transformation. Thus, for example, in the determination of the $\Delta Y \approx 0{\cdot}30$ intensity ratio the standard deviation of the two transformations is: $\sigma_p = \pm 0{\cdot}016$ and $\sigma_l = \pm 0{\cdot}009$.[2]

References

1. TÖRÖK, T., ZIMMER, K. and RÉTI, S., *Microchim. Acta*, 611 (1962)
2. TÖRÖK, T. and ZENTAI, P., *Acta chim. Hung.*, **30**, 11 (1962)
3. TÖRÖK, T., *Emissionspektroskopie*, Berlin, 141 (1962)
4. TÖRÖK, T., *Chem. analit.*, **7**, 47 (1962)
5. TÖRÖK, T. and ZIMMER, K., *Quantitative Evaluation of Spectrograms by l-Transformation*, Budapest, 1971
6. ZIMMER, K., TÖRÖK, T. and ASZTALOS, I., *Chem. analit.*, **11**, 1065 (1966)

5.8 Background correction

5.8.1 General considerations

In photographic intensity determinations the background corrections can be made by various methods. In each method the apparent density of the spectral line – that is, the collective D_{l+u}-value of the line plus background density – is measured, in addition to the density D_u of the spectrum beside the line. The latter value cannot, however, be deducted directly from the apparent line density, since the density is a logarithmic quantity. The appropriate correction can therefore be effected only in such a way that the intensity values I_{l+u} and I_u belonging to the previously given densities are established.

In the procedure it should also be considered that the relationship between the density and the logarithm of the intensity is linear only in the linear part of the characteristic curve. On the other hand, the background density has a larger effect on the apparent density of the line, the lower is the true intensity of the latter. The lower is this intensity, however, the more likely it is that the density of the line will not fall on the straight part of the characteristic curve. On the bent part of the curve

the relationship between the quantities D and Y is complex. This difficulty is, however, avoided if during the background correction the transformed density is used instead of the true density.

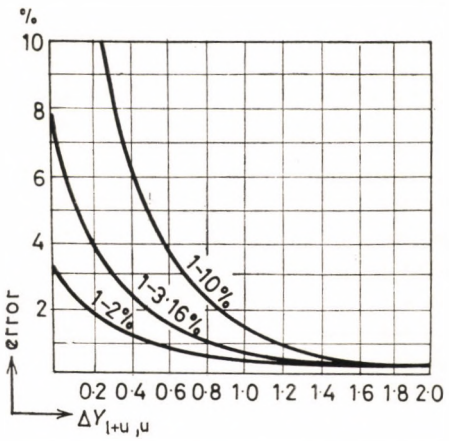


Fig. 5.51 The relative analytical error developed by neglecting the background correction, for different concentrations of the element to be determined

The importance of the background correction may be clearly appreciated from Fig. 5.51.[1]

Reference

1. KAISER, H., *Spectrochim. Acta*, **3**, 297 (1947)

5.8.2 Background correction by calculation[1]

The procedure is based on the application of the Gauss differential logarithm. Let us perform the following transformations

$$\log I_l = \log (I_l + I_u - I_u) + \log I_u - \log I_u + \log I_u(I_l + I_u) - \log (I_l + I_u) =$$

$$\log I_{l+u} - \left[\log \frac{I_{l+u}}{I_u} - \log \left(\frac{I_{l+u}}{I_u} - 1\right)\right] \tag{5.45}$$

where I_l is the true line intensity, I_u is the background intensity, $I_{l+u} = I_l + I_u$ is the apparent line intensity. When the ratio I_{l+u}/I_u in Eq. (5.45) is denoted by x then it becomes evident that the quantity in the square brackets, $\log x - \log (x - 1)$, is just the value of the Gauss differential logarithm $G_{\log I_{l+u}/I_u}$ belonging to the value of $\log I_{l+u}/I_u$. When finally the logarithm of the intensities is denoted by the usual symbol, while their difference is denoted by Y, then Eq. (5.45) can be written in the following final form

$$Y_l = Y_{l+u} - G_{\Delta Y_{l+u,u}} \tag{5.46}$$

Therefore, the true logarithmic intensity of the spectral line is obtained by subtracting from the apparent logarithmic intensity of the line that value of the differential logarithm that corresponds to the difference of the logarithmic intensity of the apparent line density and the background. The calculation of the true spectral line intensity should be performed to an accuracy of three decimal places. The G-values are compiled in appropriate tables.[1,2]

References

1. HOHNERJAGER-SOHM, H. and KAISER, H., *Spectrochim. Acta*, **2**, 396 (1944)
2. Zeiss-Sonderdruck, Mess., Jena, 41-360-1

5.8.3 Determination of the difference between the apparent density of the spectral lines and the background by the s.p.d. scale[1, 2]

The so-called *s.p.d.* (standard paper density) *scale* used for the determination of the density of the spectral lines is a strip of photographic paper on which lines of progressively higher density can be seen. This scale is used so that the spectral line in question as projected by the spectrum projector is compared to the lines of the scale. In this way the difference between the apparent density of the spectral line and the background can be established (Fig. 5.52).

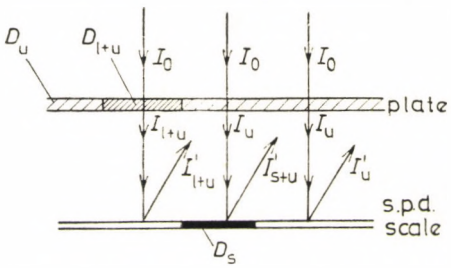


Fig. 5.52 The application of the s.p.d. scale

To justify the previous statement let us assume that the following relationship exists between the intensity of light falling on the table of the spectrum projector and the intensity of the light reflected from the latter into the observer's eye

$$I'_{l+u} = pI_{l+u} \qquad \text{or} \qquad I'_u = pI_u \tag{5.47}$$

where I_{l+u} is the intensity of light falling on the image of the spectral line, I'_{l+u} is that of the reflected light, and I_u and I'_u are the intensities falling on, and reflected from, the areas of background (Fig. 5.52). The symbol p is a multiplying factor whose value stays thus: $p < 1$. If it is also assumed that the intensity of light reflected from the line of the s.p.d. scale is $I'_{s+u} = qI'_u$ then

$$I'_{s+u} = p \,.\, q \,.\, I_u \tag{5.48}$$

The projected spectral line is seen to have the same density as that of the appropriate line of the s.p.d. scale, when $I'_{l+u} = I'_{s+u}$. However, in this case it is also true that

$$\log \frac{I_0}{I'_{l+u}} = \log \frac{I_0}{I'_{s+u}} \tag{5.49}$$

Furthermore, if it is taken into account that Eq. (5.47) applies and that the density of the spectral line is $D_{l+u} = \log (I_0/I_{l+u})$ then we obtain

$$\log \frac{I_0}{I'_{l+u}} = D_{l+u} - \log p \tag{5.50}$$

On the other hand, the value of the right-hand side of Eq. (5.49) is obtained in such

a way that Eq. (5.48) and the $D_u = \log I_0/I_u$ value of the background density is taken into account, that is

$$\log \frac{I_0}{I'_{s+u}} = D_u - \log p - \log q$$

If it is furthermore considered that the *apparent* density of the lines of the s.p.d. scale is given by $D'_{s+u} = \log (I'_u/I'_{s+u})$ then with the help of Eq. (5.47) and (5.48) it can be calculated that $D'_{s+u} = -\log q$. The previous relationship is therefore modified so that

$$\log \frac{I_0}{I'_{s+u}} = D_u + D'_{s+u} - \log p \tag{5.51}$$

By taking into account the relationships (5.49), (5.50) and (5.51) it can be shown that

$$D'_{s+u} = D_{l+u} - D_u \tag{5.52}$$

The lines of the s.p.d. scale are indicated by numbers so that the numbers increase with increasing density. These numbers are, however, only indices, therefore it is not possible to make any direct inferences about the measured density from their value. Therefore, the s.p.d. scale should be calibrated. This is all the more necessary since the difference between the *visible* density of the lines of the s.p.d. scale and that of the corresponding spectral lines becomes greater as the density value to be determined becomes higher. In the calibration procedure spectral lines are found on the projected, and preferably line-abundant, spectrum with density values that best match the densities of certain lines of the s.p.d. scale. The density of the selected lines is measured with a spectral-line photometer and the values obtained are compiled in a table or diagram.

It can be established experimentally that the background does not interfere with the use of the s.p.d. scale even if its value is $D_u > 0{\cdot}5$. The highest value of the measurable line density is $D_l \approx 1$. The standard deviation of the determinations is generally about $\pm 7\%$ when the value of the density to be determined is lower than 0·6. In measuring densities up to about 1·0, however, relative deviations larger than $\pm 10\%$ should be expected.

References

1. ADDINK, N. W. H., *Spectrochim. Acta*, **4**, 36 (1950)
2. DIKHOF, J. A. M. and ADDINK, N. W. H., *Mikrochim. Acta*, **1955**, 257

5.9 Determination of the density of spectral lines

5.9.1 Fundamental principles of the procedure

The density of spectral lines is the logarithm of the intensity ratio I_0/I, where I_0 is the intensity of light passing through the unexposed parts of the processed photographic emulsion and I is the intensity of light passing through the spectral line. The density is therefore $D = \log (I_0/I)$ when the original intensity of the two light-rays falling on the photographic emulsion is identical.

Usually, a photoelectrical instrument is used for the *exact* determination of the above-defined ratio, that is, a photoelectrical photometer. In some analytical tasks

it is enough, however, to know only the approximate value of the spectral-line densities. In these cases it is better to establish the density of the lines with the naked eye, by means of the *s.p.d.* scale in a simple and rapid way[1] (Sub-section 5.8.3). In practical spectrochemical analysis the procedure called 'density plastics' has not become widespread, although by this method densities much higher than $D = 3$ can be measured.[2-4]

For the measurement of spectral line densities the only type of photometer that can be used is one with which the value proportional to the intensity of the ray travelling through any line in the spectrum can be quickly determined to the required accuracy. For this purpose it is only practicable to use specially constructed photoelectrical instruments, *spectral-line photometers.*

The operation of spectral-line photometers is based on the following principle. An appropriate optical device projects the light from a constant-intensity source, through the required locations of the processed photographic emulsion, onto a photocell or photoelement. The galvanometer deflection corresponding to the resulting photocurrent is proportional to the intensity of the ray falling on the photocell or photoelement.

In principle, therefore, the required density value is obtained by reading first the galvanometer deflection G_0 corresponding to the intensity I_0 of the ray travelling through the unexposed part of the photographic emulsion. Thereafter, under identical experimental conditions, that galvanometer deflection G is recorded that corresponds to the intensity I of the light passing through the spectral line. The density value required is therefore $D = \log (G_0/G)$ or $D = \log G_0 - \log G$.

It follows that the determination of densities is very much facilitated and made faster if the galvanometer scale is not linear, but logarithmic. The scale of the modern spectral-line photometers is partly of negative logarithm character, that is $B = -\log G$. The scale is such that zero corresponds to the full deflection of the galvanometer, while the infinity sign corresponds to zero deflection. Then the measured density can be read directly from the scale. To achieve this it is necessary that the sensitivity of the galvanometer, or the intensity of the light illuminating the photographic emulsion, be adjusted so that the galvanometer deflection is exactly zero when the light travels through the unexposed part of the photographic emulsion. In this case $D = -(\log I - \log I_0)$ and thus $D = G$ since $-\log I \approx G$ and $-\log I_0 = 0$.

The necessary accuracy can only be achieved when using line photometers if the following conditions are met.

(1) For the density measurement a constant-intensity source is necessary suc that the maximum deflection of the galvanometer changes only by 1% during period of 10 min.

(2) The emulsion should be illuminated by direct light, otherwise the Callier effect gives rise to measurement errors. The error from this source can be as much as 30% in the case of fine-grain emulsions and 80–100% for coarse-grain emulsions.

(3) Scattered light should not be allowed to affect the measurement accuracy.

(4) Any part of the photographic emulsion should be reached quickly and accurately.

(5) It is necessary that the slit of the photometer should be adjusted by a micrometer screw so that the photocell is illuminated *only* by light transmitted through a single spectral line.

(6) The photoelement or photocell should be sufficiently sensitive and free of lag. The photocurrent should vary proportionally to the intensity of the light.

(7) It is necessary that the galvanometer while of the necessary sensitivity, should halt quickly, without deflections and oscillations.

(8) The relative measurement accuracy obtainable should be larger than 1%.

References

1. ADDINK, N. W. H., *Spectrochim. Acta*, **4**, 36 (1950)
2. LAU, E., KIND, E. G. and ROSE, G., *Optik*, **15**, 162 (1958)
3. SCHMIDT, H., *Optik*, **16**, 538 (1959)
4. SCHMIDT, H., *Umschau*, **62**, 749 (1962)

5.9.2 Construction of spectral-line photometers

The important parts of a spectral-line photometer are the following: source, lens system for the illumination, adjustable plate-carrier or table, projection system, photoelement, galvanometer, and reading equipment.

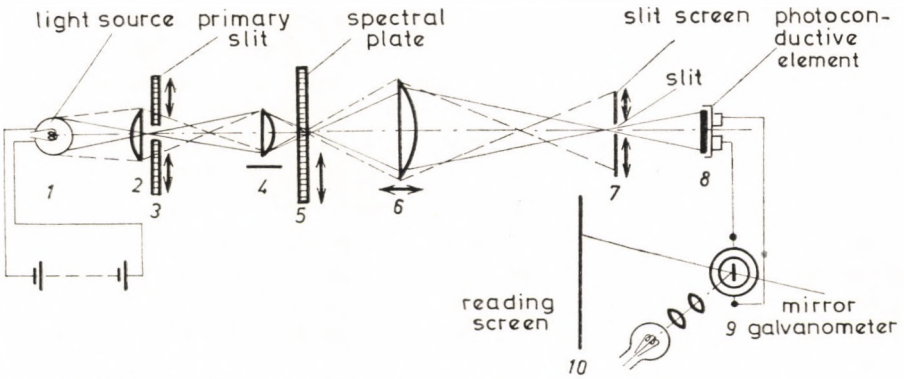


Fig. 5.53 Scheme of a spectral-line photometer

The schematic arrangement of the equipment is seen in Fig. 5.53. The image (*k*) of the spiral of the low-voltage incandescent lamp (*1*) is formed by the collector lens (*2*) on the illumination lens (*4*). Thus the light evenly illuminates the fore-slit (*3*) made of green glass in such a way that its sharp image is formed in the plane of the photographic emulsion (*5*). Thus it becomes possible that the projector lens (*6*) projects not only the magnified and sharp image of the spectral line, but also that of the foreslit onto the screen (*7*), of adjustable width and height, situated in front of the photoelement (*8*). Therefore, on this screen not only the image of the spectral line and the green fore-slit can be seen, but also the neighbourhood of the spectral line that has been brought in front of the slit by means of the plate-carrier – which can be moved in either direction. (This construction renders orientation in the spectrum much easier.) The deflection of the mirror galvanometer (*9*) can be read from the projection screen (*10*).

According to the requirements listed in the previous section it is necessary for the light source of the photometer to be point-like, that is, a low-voltage incandescent

lamp. The best solution to this problem is achieved when the power supply is a large-capacity lead accumulator. The radiance of the source is very sensitive to small changes in current. For this reason mains current can only be used for this purpose when either a high-quality current-stabiliser is inserted or when the mains-current fluctuations are levelled out by a lead battery of 'buffer circuitry'.

Stray light strongly affects the accuracy of the density measurements, therefore it can be used for qualifying the spectral-line photometers. The stray-light intensity adds to that of the rays coming through the photographic emulsion and thus it leads to an additive measurement error independent of the measured density. This error can be especially important in the case of narrow and high-density spectral lines. The extent of the disturbance can be most simply determined by regarding the shadow of a non-reflecting and non-transparent filament (for example a wire) as a *completely* black ($D = \infty$) spectral line. The density measured this way, however, reaches only about $D = 2$ in most photometers, and only exceptionally does it reach the limit $D \approx 3$.[1]

In most line photometers the harmful effect of stray light can be reduced by means of the *fore-slit*, that can be used to reduce the cross-section of the parallel beam coming from the light source. This solution is only successful, however, when the edges of the fore-slit form a sharp image on the photographic emulsion. Even more effective is the method when two *polarisation sheets* of the same transmittance form the fore-slit, while the third sheet placed in front of the slit of the photocell has its polarisation plane perpendicular to the other two sheets.[2] Stray light is completely eliminated in the method in which the source directly illuminates the spectrum through an *entrance slit* placed in front of the emulsion, so that the rays pass through only the spectral line under study.[3] The neighbourhood of the line cannot be seen, but this drawback is eliminated by using the *master plate*. The latter is a spectrum reproduction on which the location of the analytical lines has been indicated and whose magnified image is projected onto the slit-screen of the photometer. During the measurement procedure the master plate is moved simultaneously with the spectral plate under study. The apparatus is therefore a combination of the line photometer and the spectrum comparator.

In the density measurement of spectral lines the micro-error of the photographic emulsion becomes smaller the larger is the illuminated area of the spectrum. The latter, of course, cannot be larger than the surface of the spectral line. Accordingly, in practice the size of the photometer slit is generally chosen so that its height is somewhat smaller than, and its width about two-thirds of, the corresponding dimensions of the spectral line projected onto the slit-screen. (These dimensions can be simply calculated by multiplication provided the slit width and slit height of the spectrograph, and the magnifications of the spectrograph and of the photometer, are known.) However, it should be noted that when the illuminated area of the emulsion is varied the photocurrent also varies. It may be a source of error, therefore, if during the measurement of the relevant densities a change is effected in the dimensions of the slit in front of the photocell.

The density profile of the undistorted spectral lines is generally a bell-shaped curve the top of which coincides with the wavelength of the line in question. In the density measurement of the spectral lines, therefore, the photographic emulsion should be moved along the length of the spectrum so that the photocell is illuminated by the light of lowest intensity. It is therefore necessary to move the image of the spectral

line in front of the slit of the photocell until the smallest deflection of the galvanometer of the photometer is obtained. This operation can only be done by means of an appropriately fine adjusting screw. The 'fine' adjustment of the position of the photographic emulsion is not sufficient in itself. It is also necessary to be able to move the emulsion rapidly in the direction both of its length and of its width. If possible this should be done in such a way that the image of the spectrum projected on the slit-screen should remain sharp. For this purpose a 'plate-table' is needed that can be controlled in both directions by adjusting screws as required. It is also advisable to control the lengthways movement of the plate-table by a nonius scale or by a micrometer screw. In series examinations it is also advantageous to be able to limit the movement of the plate-table by adjustable stops. In this way the density measurement of spectra photographed directly below one another can be greatly speeded up. In the determination of individual spectral lines it is advantageous, on the other hand, to view a larger area of the spectrum – not only the immediate neighbourhood of the examined part of the emulsion.

Spectral-line photometers can be grouped into two classes according to their photoelectric equipment. One group consists of those instruments that use a photoelement, while the other group uses an alkali-metal photocell or photomultiplier for the measurement of the light intensity.

The photoelement system has the advantage that the measuring apparatus is relatively simple. The photocurrent of photoelements is so large that mirror galvanometers can be *directly* used for the density measurement of the spectral lines. On the other hand, the photoelements have properties that are bad for density determinations. First of all the lag of photoelements cannot be neglected. The lag prevents the galvanometer spot gaining its final position immediately when the photoelement is suddenly illuminated by high-intensity light. The instability of photoelements should also be considered. This characteristic causes the zero-point to wander during measurement. These drawbacks are eliminated when the measurements are performed so that lines of similar density are measured one after the other. (It is also due to the lag of photoelements that during the measurement the photoelement can only be illuminated when the image of the spectral line has already been set approximately in front of the slit of the photocell. Before and after the measurement the light path should be closed. It is advisable to do this operation at a regular rhythm, but this needs some practice.)

The alkali-metal photocell and the photomultiplier have the disadvantage that they cannot be used without an external power supply and fluctuations of the supply voltage can have an effect on the accuracy of the measurements. On the other hand, these devices have practically no lag effect. The error due to wandering of the zero-point is eliminated by the compensation of the dark-current – which is necessary anyway. A further advantage is that the photocurrent can be amplified and thus density readings can be recorded by using a chart recorder.

The maximum light sensitivity of photoelements, alkali photocells, and photomultipliers is generally different from that of the human eye. As a result, the spectral line adjusted to sharpness by eye is not usually 'sharp' to the photocell, when – as usual – white light is used for illumination and the optics of the photometer is not chromatically corrected. However, the Callier effect arising in this way does not lead to error if the 'sharpness' of the spectral line projected onto the slit of the photometer is adjusted by means of the minimal deflection of the galvanometer, and not by eye.

In practice, it is important that the galvanometer of the spectral-line photometer should quickly attain the final deflection value. This happens if the resistance of the external circuit of the galvanometer is equal to the so-called limiting resistance (or critical resistance). When the resistance of the external circuit of the galvanometer is larger than the critical value the galvanometer attains its final position only after having gone through a number of relatively large and weakly-damped oscillations. In this case the strength of the induced current is small and thus the damping is also ineffective. On the other hand, when the external resistance of the galvanometer is too low strong damping occurs and the galvanometer spot creeps slowly to its final position. Therefore, it is most important to choose an appropriate external circuit resistance.

On normal principles of measurement spectral-line photometers can be grouped into two classes. In one group of photometers the galvanometer deflection gives directly a quantity proportional to the measured intensity, and thus to the density of the photographic emulsion. These are *direct-measuring* instruments (Fig. 5.54*a*). In the other group of photometers, the so-called *compensation* instruments, the galvanometer is only used as a nulling device (Fig. 5.54*b* and *c*).

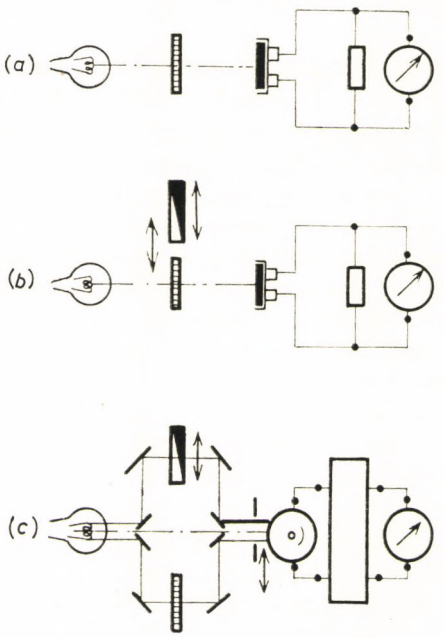


Fig. 5.54 Measurement principle of photoelectric photometers: (*a*) direct measuring method (deflection type), (*b*) compensation method (substitution type), and (*c*) alternating illumination method

In the application of the photoelectrical measuring methods it should be kept in mind that – as in the case of the extinction measurements – the relative measurement error (that is, the value of $\Delta D/D$) is not independent of the value of the measured density. The maximum measuring accuracy – except for substitution-type measurements – is obtained for the density level 0·43.[4] The measurement error increases, however, twofold at densities 0·1 and 1·1 and increases steeply when still higher density values are measured (Fig. 5.55).

The measuring arrangement of the direct-measuring spectral-line photometers is most simple when a photoelement is used for the measurement of the light intensity.

The entire set-up is composed of only three parts: photoelement, galvanometer, and damping resistance. This is the reason why most of the spectral-line photometers in current use are of this type. Instruments in which an alkali-photocell or photomultiplier is used instead of the photoelement are rare, despite the fact that it is possible with an instrument of this type to record the line profile by using a chart recorder.

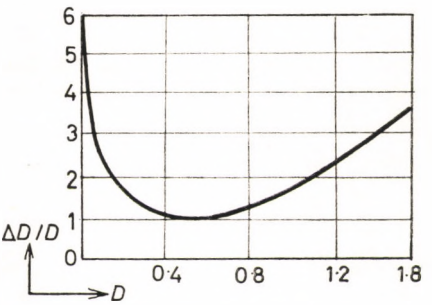


Fig. 5.55 Variation of the relative measurement error as a function of the measured density

In the reading part of modern direct measuring line photometers the mirror of the running scale galvanometer usually projects the image of the scale onto a stationary spot on a screen (Fig. 5.56). The advantage of this system is that the results of the measurements are always read at the same place. Thus measurement is less tiring than with older instruments where the scale spot is usually projected by the galvanometer mirror onto a relatively long scale fixed above the photometer.

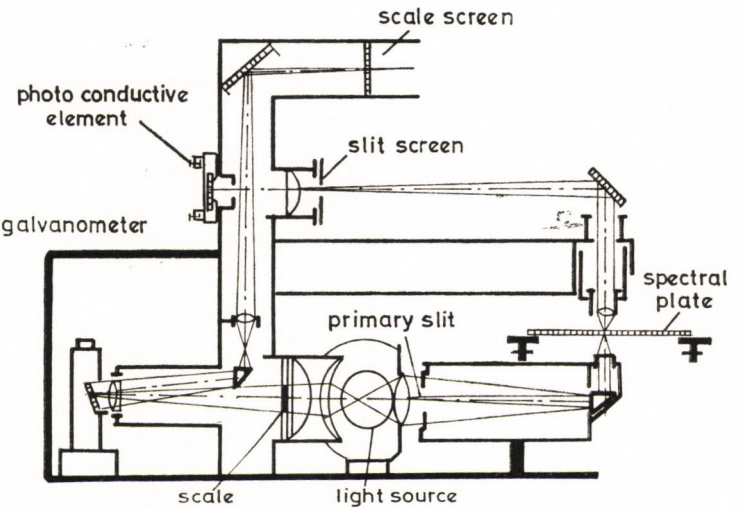


Fig. 5.56 Optical scheme of the 'rapid photometer' based on direct measuring principles[5]

There are two versions of the compensating-principle photometer. In one version the measurement is made in two separate steps: first the photographic emulsion, then the measuring attenuator is placed in the light path. By means of the latter, the light flux is reduced until the intensity of the photocurrent becomes equal to that observed in the first measurement (Fig. 5.54*b*). The drawback of this *substitution*

procedure is that errors occur when the light intensity of the source changes between the two measurements. Thus this method is no longer used.

The above drawback is eliminated in the version of the compensating-principle photometer in which the photocell is illuminated alternately by the ray from the light source that goes through the photographic emulsion and by the ray passing through the calibrated attenuator. The alternation frequency is high (Fig. 5.54*c*). In this alternating illumination method the equal intensity of the two beams is indicated by the fact that the alternating current arising when the two beams are of different intensity drops to zero-level and becomes direct current. The great advantage of the procedure is that voltage fluctuations do not affect the accuracy of the measurements.

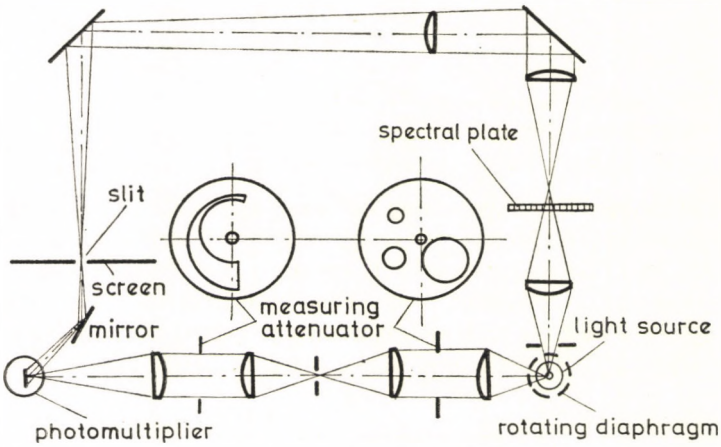


Fig. 5.57 Simplified optical diagram of the line-photometer based on alternating illumination measurement[6]

A further advantage is that the accuracy attainable is completely independent of the value of the density measured. Thus densities in the range 0–3 can be accurately measured. A much higher measuring accuracy can be achieved by this method than by any other. Its application, however, needs relatively complicated optical and electronic equipment (Fig. 5.57).

References

1. KECK, P. H., *Zeiss Nachr.*, **10**, 1 (1936)
2. ROLLWAGEN, W., *Congrès. Groupement pour l'Avancement des Méthodes d'Analyse Spectrographique des Produits Métallurgiques*, 207 (1953)
3. Description of the Applied Research Laboratories spectral-line photometer
4. KORTÜM, G., *Kolorimetrie, Photometrie und Spektroskopie*, vol 2, 4th edn, Berlin–Göttingen–Heidelberg, 247 (1962)
5. Zeiss-Druckschrift, CZ-32-G 661-1 Jena
6. Baird Atomic Associates, Inc., description of line photometer

5.10 Fundamentals of photoelectric intensity measurements

5.10.1 External photoelectric effect[1-6]

The external photoelectric effect is the phenomenon of electrons being emitted into a vacuum or gas from a solid-phase photosensitive cathode when it is irradiated by light.[7,8] In order to make up for the loss of electrons emitted from the cathode an appropriate electrical field is established between the cathode and the anode and a photocurrent is created between the two electrodes. The value of the current i is proportional to the performance Φ of the radiation, that is

$$i = C\Phi \tag{5.53}$$

provided the dark current and the isolation current are neglected.

The photocells or photomultipliers used for the measurement of spectral line intensities function on the basis of the external photoelectric effect. The principle of this effect is that all light quanta (photons) of sufficiently high energy content, hv release photoelectrons on the cathode, which is the illuminated surface. This happens, provided the work function $E_w = e_0\phi$ is smaller than the energy, hv, of the light quantum of frequency v, where e_0 is the electronic charge. The difference of the two quantities is responsible for the kinetic energy of the emitted photoelectron $mv^2/2$. Thus, according to the Einstein equation

$$hv = e_0\phi + \frac{mv^2}{2} \tag{5.54}$$

where ϕ is the voltage through which the photoelectron travels when it emerges from the interior of the metal to the surface.

Let us assume that the wavelength of the light absorbed on the cathode is exactly λ_0, just enough to supply the work function of the photoelectron. In this case the kinetic energy of the emitted photoelectron is zero, that is $mv^2/2 = 0$. It follows from the above equation that

$$hv_0 = e_0\phi = E_w = \frac{hc}{\lambda_0}$$

Thus

$$\lambda_0 \equiv \lambda_{max} = \frac{hc}{E_w} = \frac{1240}{\phi} \text{ nm} \tag{5.55}$$

Here λ_{max} is the limiting wavelength, or, in other words, the long-wavelength limit of the photoeffect. This is the wavelength at which in the case given no photoelectric effect is produced. It further follows from the above equation that the value of the limiting wavelength is higher, that is, the limiting wavelength shifts towards the longer values, as the work function decreases.

In the photoelectric effect the optimum case is when each photon absorbed on the cathode releases one photoelectron. Then the maximum utilisation of the photoelectron as calculated in the unit of radiation energy is given by the following formula

$$E_{max} = \frac{e_0}{hv} = \frac{\phi}{1240} \text{ A/W} \tag{5.56}$$

when $e_0 = 1{\cdot}60.10^{-19}$ C, and when the value of the work function ϕ is calculated in volts. If, instead of electrical units, light units are used, for example, at the wavelength at which the human eye is most sensitive (555 nm) when 1 W = 633 lm it follows that

$$E_{\max,\,555} = 707 \text{ A/lm} \tag{5.57}$$

In reality, not every absorbed light quantum releases a photoelectron. If the number of the electrons emitted from the cathode is compared to the number of the absorbed photons, then the quotient called *quantum yield* is obtained. Its value is therefore $\gamma = E/E_{\max}$, where E is the actual yield of the photoelectrons. This quotient characterises the light sensitivity of the photocell.

The quantum yield of each photocathode varies with the wavelength of the light. The curve depicting this variation is called the spectral distribution curve of the photocell. Unfortunately, the quantum yield decreases regularly with increasing wavelength, in the normal photoelectric effect. On the other hand, in the selective photoelectric effect the distribution curve passes through one or more maxima appearing in a relatively narrow wavelength range (Fig. 5.58). The value of the latter is much higher than that observed in the normal photoelectric effect.

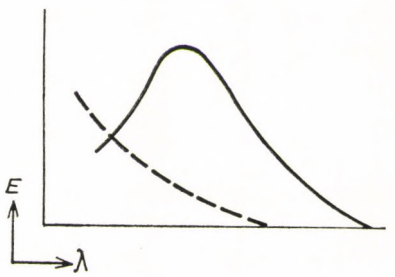


Fig. 5.58 Normal photoeffect distribution curve (dotted line) and the selective photoeffect distribution curve (solid line)

The external photoelectric effect has the result that when a sufficiently large voltage, U, is established between the light-sensitive cathode and the anode opposite it a current, I, emerges. The electrons emitted by the cathode are supplied by this photocurrent.

The product of the electronic charge and the potential difference between the cathode and the anode is equal to the work done against the electrostatic field. So

$$e_0 U_r = \frac{mv^2}{2}$$

and it follows from Eq. (5.54) that

$$U_r = \frac{hv}{e_0} - \frac{E_w}{e_0} \tag{5.58}$$

that is, there is a linear relationship between the retarding voltage U_r and the frequency v of the light. Therefore, an increase in the frequency produces a proportional increase in the decelerating voltage. When the relevant values of the two variables are plotted into a diagram, the straight line obtained intersects the v-axis at the point that has a corresponding frequency v_0, that is, the photoelectric threshold value (Fig. 5.59).

According to Eq. (5.53) the photocurrent produced by the external photoelectric effect has an intensity that varies in proportion to the radiation energy Φ, that is, $i = C\Phi$, where C is a proportionality constant the value of which is determined by the sucking voltage U and the wavelength λ_0 corresponding to the photoelectric threshold value.

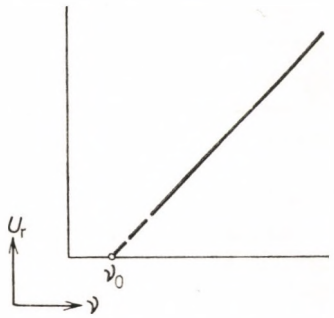


Fig. 5.59 Diagram illustrating the photoelectric threshold value

The sensitivity of the photocells is a quotient formed from the reading on the instrument used for measuring the photocurrent, i, or voltage u and the radiation energy, that is

$$s_i = \frac{i}{\Phi} \text{ A/W}, \quad \text{or} \quad s_u = \frac{u}{\Phi} \text{ V/W} \tag{5.59}$$

The sensitivity varies as a function of the wavelength of the light illuminating the cathode. The sensitivity determined at a certain wavelength is the so-called spectral sensitivity. Its denotation is

$$s_i^* = \frac{i}{\Phi}, \quad \text{or} \quad s_u^* = \frac{u}{\Phi} \tag{5.60}$$

When the above s_λ-value is compared with the spectral sensitivity s_0 of light of wavelength λ_0 chosen at will, then the so-called *relative spectral sensitivity* is obtained

$$s_{r,\lambda} = s_\lambda / s_{\lambda_0} \tag{5.61}$$

In practice the sensitivity of the radiation sensing elements, the function of which is based on the photoelectric effect, is given as a value in units (μA/lm) referred to the radiation of a black body at a specified temperature.

In addition to the photoelectric effect the thermal electron radiation always produces a low-intensity dark current (with current density of the order 10^{-9}–10^{-17} A/cm^2).[9,10] The current produced this way can be observed even when the cathode is not illuminated by light. The intensity per unit surface area is given by

$$i_T = AT^2 \, e^{-11\,600\phi/T} \text{ A} \tag{5.62}$$

where A is the emission constant, T is the temperature of the cathode measured in kelvins, and Φ is the work function of the thermal emission given in electronvolts.[11] (The value of the latter is similar to the Φ-value of the photoelectric work function for the alkali-metal cathodes.) The value of the emission constant depends on the type of metal forming the sensitive layer of the cathode. It may be added that in practice the dark current does not give rise to troubles even at low levels of light intensity since its effect can be simply eliminated by compensating the current.

On the other hand, the *shot noise*, or *shot effect*, which is harmful, cannot be eliminated. This effect is caused by the unavoidable statistical fluctuations of the electron radiation in time. The mean square average of the resulting current fluctuation is given by

$$i_s^2 = 3{\cdot}20 \,.\, 10^{-19}\, i_0^2 \, \Delta f \, \mathrm{A}^2 \tag{5.63}$$

where i_0 is the mean current value of the photocurrent, and Δf is half of the reciprocal time constant ($2\Delta f$) of the instrument. (The latter is the time necessary for the pointer of the instrument to attain its original position after having been deflected by some external event.) In practice the value of the shot effect determines the lower limit of the intensity measurement based on the photoelectric effect, but the shot effect does not give rise to other, more severe difficulties.

A further common failure of photoelectrical sensors is that their sensitivity decreases with time, that is *fatigue phenomena* occur. If the sensitivity decrease occurs very slowly, but irreversibly, it is an *ageing process*. This effect does not generally cause difficulties or errors in the intensity measurements. When, however, the sensitivity decrease occurs quickly, during the measurements, it is a *fatigue phenomenon* that could give rise to measurement errors. The latter effect is not usually permanent. After a certain time of rest the photocell regains its sensitivity.

References

1. GÖRLICH, P., *Die lichtelektrischen Zellen*, Leipzig (1951)
2. ZWORYKIN, V. K. and RAMBERG, E. G., *Photoelectricity and Its Application*, New York (1950)
3. SOMMER, A., *Photoelectric Tubes*, 2nd edn, London (1951)
4. SIMON, H. and SUHRMANN, R., *Der lichtelektrische Effekt und seine Anwendungen*, 2nd edn, Berlin–Göttingen–Heidelberg (1958)
5. KOHLRAUSCH, F., *Praktische Physik*, 19th edn, Leipzig (1955)
6. KORTÜM, G., *Kolorimetrie, Photometrie und Spektrometrie*, 4th edn, Berlin–Göttingen–Heidelberg (1962)
7. SMITH, W., *Am. Scient.*, **5**, 304 (1873)
8. ELSTER, J. and GEITEL, H., *Ann. Phys.*, **38**, 497 (1889); **41**, 161 (1890)
9. SCHAETTI, N. and BAUMGARTNER, W., *Helv. phys. Acta*, **25**, 605 (1952); **26**, 380 (1953)
10. SCHAETTI, N., *Z. angew. Math. Phys.*, **4**, 450 (1953)
11. RICHARDSON, O. A., *Phil. Mag.*, (6) **24**, 570 (1912); **27**, 476 (1914)

5.10.2 Photocells[1–4]

Photocells consist of glass or quartz tubes that are evacuated or filled with low-pressure gas, and have the internal surface coated by a very thin, light-sensitive, metal layer. Inside the tube there is the usual loop-form anode opposite a light-esnsitive cathode. When a high enough voltage is established between the two electrodes, then light falling on the cathode causes a current in the external circuitry of the photocell the value of which can be used to deduce the value of the light flux falling on the cathode (Fig. 5.60).

Photocells can be grouped into two types; vacuum cells and gas-filled cells. The properties of these two types are very different, and so are the curves correlating the illumination strength with the photocurrent.

In vacuum photocells the photocurrent varies linearly with the illumination strength of the cathode (Fig. 5.61). On the other hand, the value of the photocurrent

hardly increases with increasing voltage, U_s, when the latter is in excess of a certain level or the so-called saturation voltage. A difference can, however, be observed in the relationship between the photocurrent and the illumination strength when the voltage is significantly increased.

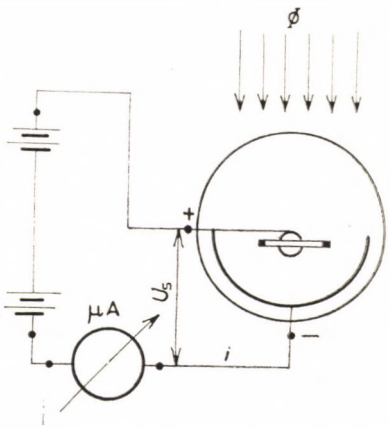


Fig. 5.60 Photocell circuit diagram

These observations can be explained by considering that the initial bend in the photocurrent–sucking voltage curve is caused by the fact that when there is a small sucking voltage some of the electrons emitted from the cathode do not reach the anode. When the saturation voltage, usually about 30 V, is reached this effect completely disappears. Certain small increases of the photocurrent that remain are mainly caused by the Schottky effect—if we neglect the possible effect of the remaining gas. This is shown by the fact that the work function of the electrons is reduced as the voltage is increased. This effect, however, does not appear when the voltage is high, although it can damage the photocell.

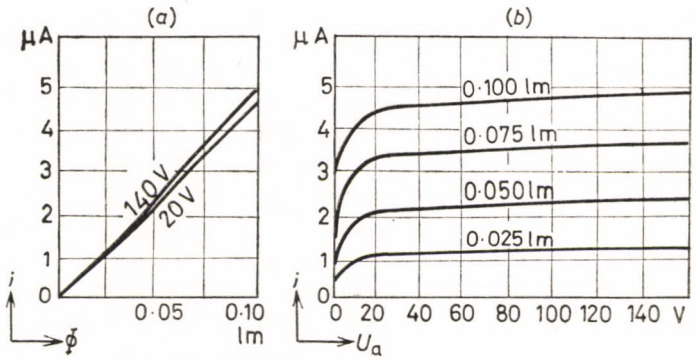


Fig. 5.61 Variation of the photocurrent of vacuum photocells (*a*) as a function of the cathode, illumination strength and (*b*) as a function of the acceleration voltage

The photocurrent–light flux and photocurrent–voltage curves of the gas-filled photocells are significantly different from the corresponding curves of the vacuum photocells. The difference is that their photocurrent not only varies much more, but at low voltage levels the variation as a function of the illumination strength is

not quite linear (Fig. 5.62). Therefore, the sensitivity of the gas-filled photocells is much higher than that of the vacuum photocells. However, a drawback is that the photocurrent of the gas-filled photocells varies very strongly with the voltage even if the illumination strength is constant. When using gas-filled photocells this fluctuation of the voltage causes errors. This type of error does not occur in the vacuum photocells when the voltage is higher than the saturation voltage.

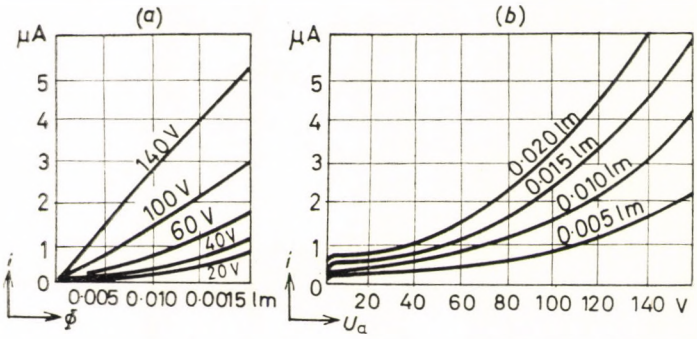


Fig. 5.62 Variation of the photocurrent of gas-filled photocells as a function (*a*) of the illumination strength, and (*b*) of the acceleration voltage

The higher sensitivity of the gas-filled photocells and the strong influence of the voltage can be traced back to the fact that the kinetic energy of the photoelectrons increases as the voltage is increased. Thereby, the collision ionisation of the gas present is increased. The current value of the external circuit is therefore also increased. For this reason, the sensitivity of gas-filled photocells is greatly increased when the voltage is increased. This phenomenon is, however, limited by the fact that at very high voltages (the glimm voltage) the gas is completely ionised. In conclusion it is the gas-discharge that mainly determines the current value through the photocell rather than the photoelectric effect. The great increase in the slope of the current–voltage curve has the disadvantage that voltage fluctuations strongly affect the current value and hence the accuracy of the light intensity measurements.

It also follows from the foregoing that the current passing through the photocell depends not only on the ionisation potential of the gas, but also on its pressure. When the pressure is too low the probability of collisions is very low. The mere presence of the gas does not greatly increase the photocurrent. When, on the other hand, the pressure is too high the number of collisions is so increased, and the free-path length of the electrons so reduced, that the electrons cannot take up the kinetic energy necessary for the ionisation of the gas. The pressure of the gas must therefore be carefully chosen if the highest possible amplification is required.

A further drawback of gas-filled photocells is that fatigue phenomena greatly affect the photocurrent value. There are only a few exceptional photocells in which this effect does not give rise to errors—even in the case of measurements of short duration. Taking all these factors into account the rule holds that only vacuum photocells may be used for intensity measurements.

The fatigue phenomenon of vacuum photocells is caused mainly by the fact that the gas molecules left behind in the cell following evacuation form positively charged

ions owing to electronic collisions. The positive ions collide with the cathode and bring about local alterations on its surface that change the work function of the electrons at these sites. Therefore, it is not advisable to have the kinetic energy of the electrons too high. Accordingly, it is not advisable to have the voltage (which gives rise to the kinetic energy of the electrons) much higher than about 20 V.

In general, the fatigue effect of vacuum photocells becomes greater with decreasing wavelength but increasing intensity of the light illuminating the cathode and at higher voltage levels. This fatigue effect is observed in the form of the work function of the electrons increasing with time, whereby the limiting wavelength and thus the spectral sensitivity of the photocell are shifted towards the shorter wavelengths. Local inhomogeneities of the light-sensitive layer of the cathode can also result in fatigue effect. In this case it can happen that a potential difference is produced between different parts of the cathode. As a result, the electrons emitted from one part of the cathode re-enter at another part of the cathode, giving rise to so-called secondary radiation.[5,6]

According to the foregoing the foremost requirement for vacuum photocells used for measuring the intensity of light is that they should be almost entirely evacuated. It is also advisable that the saturation current be produced at voltage levels below 20 V. The value of the voltage is important since it influences the minimum measurable intensity of the photocurrent. This value, however, also depends on the insulation resistance of the electrodes. In photocells made for technical purposes this resistance is in the range $10^{10}\ \Omega$, while for vacuum photocells made for measurement purposes it is in the range $10^{14}\ \Omega$. In the first case the measuring limit is generally about $3.10^{-6}\ \mu A$, and in the second case about $3.10^{-10}\ \mu A$.

It follows from Eq. (5.54) that the spectral sensitivity of photocells depends on the work function of the light-sensitive layer of the cathode and thus, indirectly, on the chemical composition of the latter. Accordingly, photocells in general use can be classified with regard to the spectral sensitivity of the cathodes.

Among pure metals the alkali metals have the lowest work function; this value is the lower the higher the atomic number of the alkali metal in question (Table 5.8).

Table 5.8

Atomic number and work function of the alkali metals[7]

Metal	Li	Na	K	Rb	Cs
atomic number	3	11	19	37	55
work function (eV)		2·46	2·24	2·17	1·9

Accordingly, the limiting wavelength and the spectral sensitivity of a photocell are shifted toward the longer wavelengths when, for example, instead of sodium of work function 2·46 eV, caesium of work function 1·9 eV is used for making up the light-sensitive layer of the cathode. On the other hand, the photocell is selectively sensitive to ultra-violet radiation when the cathode is made of a higher-work-function metal such as tungsten or platinum. In the use of the latter pure metals the limiting wavelength falls in the ultra-violet part of the spectrum (Table 5.9). No photocurrent

Table 5.9

Threshold values for some metals of practical importance[7]

Metal	λ_{max} (Å)	Metal	λ_{max} (Å)	Metal	λ_{max} (Å)	Metal	λ_{max} (Å)
Pt	2 330±480	Fe	2 870±280	Ti	3 130	Mg	3 900± 600
W	2 510±200	Ni	2 910±450	Cr	3 300	Na	5 900± 100
Au	2 560± 40	Bi	3 040±260	Pb	3 300±280	Ba	5 950± 550
Ag	2 630± 50	Sb	3 070	Tl	3 470±120	K	6 600± 500
Hg	2 660± 60	Sn	3 160± 35	Al	3 680±710	Cs	7 700±1 350

is caused by visible rays. Pure metal cathodes have the characteristic that the photo-effect is selective, that is, they can only be used in a relatively narrow spectral region. The sensitivity curve of this type of cathode passes through a steep maximum (Fig. 5.63).

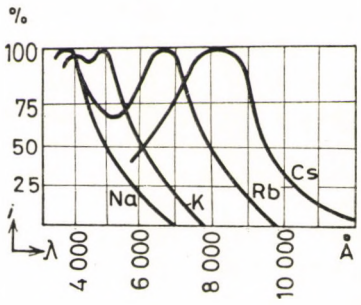


Fig. 5.63 Percentage sensitivity curves of alkali-metal photocells

The light-sensitive layer of the so-called composite cathodes is made either from alloys formed by the alkali metals and metals belonging to the fifth group of the periodic system, mainly antimony and bismuth,[8] or from semiconducting compounds formed by the alkali metals and elements in the sixth group (sulphur, selenium, tellurium and mainly oxygen.[8–10]) Alloy cathodes can generally be used over a wide spectral range. Thus, for example, the frequently used Sb–Cs cathode is not only one of the most sensitive photocathode systems used in the visible range, but can also be used very successfully in the short-wavelength region – the maximum quantum efficiency is 20%.[8–11] On the otherh and, the Te–Cs and Te–Rb cathodes are totally insensitive to the longer-wavelength rays, but can be used in the ultra-violet region as the quantum efficiency is about 30% for wavelengths shorter than 3000 Å.[12]

The third group of light-sensitive cathodes is formed by the so-called alkali-oxide cathodes. These are composite, multi-layer cathodes. The underlying metal layer of good conductivity (usually silver) is covered by a porous, semiconducting 'intermediary layer' the external surface of which contains adsorbed alkali-metal atoms. The electrical conductivity of the intermediary layer is usefully increased if some appropriate alkali-metal atoms are dispersed within the alkali-metal-oxide layer. These alkali-oxide and silver–caesium-oxide cathodes, called (Ag)–Cs_2O–Cs, or (Ag)–Cs_2O–AgCs, have the property that they can be used within a very broad range of the spectrum – for example, from the ultra-violet to the far infra-red. On the

other hand, their sensitivity varies with the wavelength in such a way that the sensitivity curve has several maxima.[13]

Finally, it may be mentioned that the limits of spectral sensitivity of alkali photocells can also be extended by sensitisers. These are compounds that transform the wavelength of the rays falling on the cathode into wavelengths to which the photocathode is sufficiently sensitive. Thus, for example, the alkali cathodes sensitised by sodium salicylate may also be used in the far ultra-violet range.[14]

In the use of photocells it should be kept in mind that their internal resistance is very high, of the order of several megohms. Direct measurement of the photocurrent is therefore very difficult. The best way is to connect a resistor of known resistance R_m in series with the photocell, then the voltage drop on the resistor is measured after appropriate amplification or by a compensation technique. In this case the current value is given by $i = U_m/R_m$. The results obtained this way are, however, free from error only when the circuitry is very carefully insulated.

References

1. GÖRLICH, P., *Die lichtelektrischen Zellen*, Leipzig (1951)
2. SOMMER, A., *Photoelectric Tubes*, 2nd edn, London (1951)
3. SIMON, H. and SUHRMANN, R., *Der lichtelektrische Effekt und seine Anwendung*, 2nd edn, Berlin–Göttingen–Heidelberg (1958)
4. KORTÜM, G., *Kolorimetrie, Photometrie and Spektrometrie*, 4th edn, Berlin–Göttingen–Heidelberg (1962)
5. PRESTON, J. S., *Revue Opt. Theor. instrum.*, **27**, 513 (1948)
6. BOER, I. H. and TEVES, M. C., *Z. Physik*, **73**, 192 (1932); **74**, 604 (1932); **83**, 521 (1933)
7. D'ANS, J. and LAX, E., *Taschenbuch für Chemiker und Physiker*, 2nd edn, Berlin (1949)
8. GOERLICH, P., *Z. Phys.*, **101**, 335 (1936); **109**, 374 (1938)
9. KOLLER, L. R., *Phys. Rev.*, **36**, 1639 (1930)
10. KLUGE, W., *Z. Phys.*, **67**, 497 (1931)
11. BURTON, J. A., *Phys. Rev.*, **72**, 531 (1947)
12. ALLEN, J. S., *Rev. Scient. Instrum.*, **18**, 739 (1947)
13. KLUGE, W., *Phys. Z.*, **34**, 115 (1933)
14. JOHNSON, F. S., WATANABE, K. and TOUSEY, R., *J. Opt. Soc. Am.*, **41**, 702 (1951)

5.11 Photomultipliers

5.11.1 Fundamental principles[1–8]

The sensitivity of the vacuum photocells described in the previous Section is generally about 3–60 μA/lm. This is inadequate for the direct measurement of the intensity of spectral lines to the required accuracy. External amplification of the photocurrent involves so many sources of error that the required accuracy generally cannot be easily obtained. These drawbacks are avoided by using photomultipliers.

Photomultipliers are vacuum photocells in which the photocurrent is amplified by built-in devices. Each photoelectron emitted from the light-sensitive cathode makes possible the emission of a number of secondary electrons (collisional electrons) from a suitable emission cathode – dynode – that is connected as anode. When the kinetic energy of these secondary electrons is appropriately increased by a further dynode then each electron colliding with the latter also renders the emission of a number of secondary electrons possible. By using a sufficient number of dynodes,

connected one after the other as anodes, it is possible to measure directly with a galvanometer the current produced by the photomultiplier.

The static photomultipliers used at present are basically composed of a light-sensitive cathode, a series of dynodes, and the anode – as described above. The electrons are accelerated between the dynodes by passing through equal electrostatic fields, that is, through amplifier stages that provide equal amplification and the most efficient multiplication of the secondary electrons. This *stepwise* amplification is necessary since photomultipliers work on the normal vacuum photocell current and the number of secondary electrons emitted from the dynode cannot be further increased if the sucking accelerating voltage has already attained the saturation level.

It follows from the foregoing that theoretically the current produced by photomultipliers depends partly on the emission coefficient – that is, on the ratio between the number a_s of secondary electrons emitted from the dynode, and the number a_p of primary electrons falling on the dynode – and partly on the number n of the amplification stages. Therefore, when the value of the emission coefficient is

$$\delta = a_s/a_p \tag{5.64}$$

and i_0 is the current leaving the cathode, the current produced by the photomultiplier is given by

$$i = i_0\delta^n \tag{5.65}$$

However, this relationship only applies in the ideal case when the value of the emission coefficient δ is the same in each multiplication stage.

The value of the emission coefficient depends on the kinetic energy of the primary electrons, and also greatly increases with increasing angles of incidence of these electrons. Its value is also dependent on the type of metal surface of the cathode; specifically, the emission coefficient is significantly greater (up to ten times more) for a composite cathode surface than for a simple one.

From the previous equation the rule follows that the current emerging from a photomultiplier increases exponentially with the number of multiplication stages. In theory, this means that when, for example, the value of the emission coefficient δ is *only* 4 and the number of amplifier stages n is 10, then an amplification of about 4^{10} (which equals 1 046 576) is obtainable with the photomultiplier.

In practice, the performance of photomultipliers may be characterised by the amplification factor S, that is, by the ratio of the anode current i_a arising from the photomultiplier, and the intensity i_0 of the entering current, hence $S = i_a/i_0$. The value of the amplification factor increases with the number of amplification stages and with the kinetic energy of the secondary electrons. In other words, it increases indirectly with increasing voltage between the photocathode, and the anode of the photomultiplier. Moreover, the amplification factor also depends on the system of the photomultiplier and even on the method of manufacture. Usually the variation of the amplification factor V as a function of the voltage U applied to the photomultiplier is given approximately by

$$V \approx \text{const} \cdot U^{\beta} \tag{5.66}$$

The value of exponent β varies with the system of the photomultiplier, etc. Thus, for example, in the case of ten-stage multiplication the value of β is usually between 3 and 5. In modern photomultipliers the amplification is usually about 10^8 or 10^9 and may exceptionally attain 10^{12}.[9]

References

1. GÖRLICH, P., *Die lichtelektrischen Zellen* (Leipzig) (1951)
2. SOMMER, A., *Photoelectric Tubes*, 2nd edn, London (1950)
3. SIMON, H. and SUHRMANN, R., *Der lichtelektrische Effekt und seine Anwendung*, 2nd edn, Berlin–Göttingen–Heidelberg (1958)
4. KORTÜM, G., *Kolorimetrie, Photometrie und Spektrometrie*, vol. 2, 4th edn, Berlin–Göttingen–Heidelberg (1962)
5. CAMMOY, C. G., *Electronics for Spectroscopists*, London (1960)
6. SCHAETTI, N., *Z. angew. Math. Phys.*, **2**, 241 (1951)
7. GLASER, G., *Gas- u. Hochvakuumtechn.*, **2**, 241 (1953)
8. PLOKE, M., *Arch. tech. Mess.*, 391 (1954)
9. SCHAETTI, N., *Helv. Phys. Acta*, **23**, 108 (1950)

5.11.2 Construction of photomultipliers[1–3, 6]

Photomultipliers are essentially vacuum photocells built from a light-sensitive cathode, a series of dynodes, and an anode. The following requirements should be met by photomultipliers used for intensity measurements:

(1) The sensitivity of the cathodes should be as high as possible, and the limiting wavelength should correspond to the function of the photomultiplier.
(2) The anode current should vary linearly as a function of the intensity of light shining on the cathode.
(3) Secondary electrons should only be produced by the effect of photoelectrons.

Most photomultipliers can be grouped into two classes according to the structure of the dynodes. There are photomultipliers with *net-type dynodes* and those with *plate-type dynodes*. The latter are further classified into photomultipliers of linear dynode arrangement and central dynode arrangement (Fig. 5.64).[4,5] The latter are the smallest. Accordingly, for the purposes of spectral analysis photomultipliers with a central dynode are the most suitable; hence only this type is used.

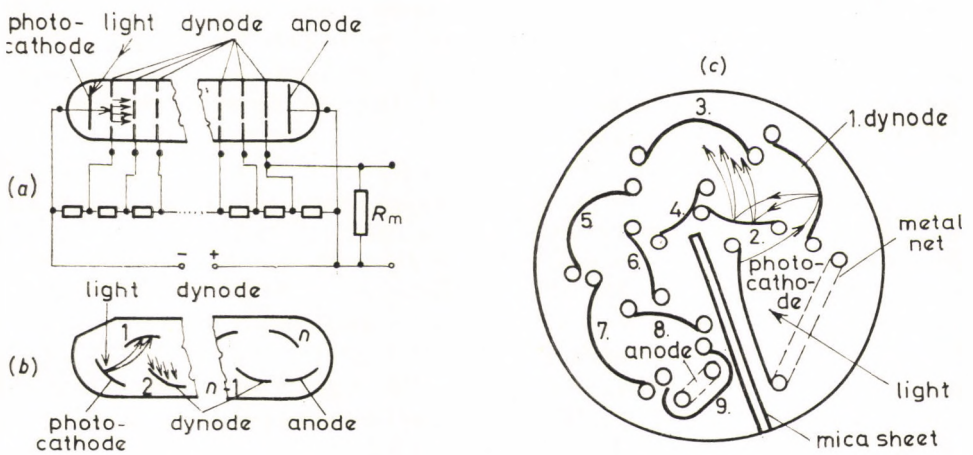


Fig. 5.64 Different types of photomultipliers: (*a*) net-type dynodes, (*b*) linear arrangement, and (*c*) central arrangement

The role of the photocathode of a photomultiplier is the same as that of the cathode of the other photocells. In theory, therefore, the same photocathode systems can be applied. In practice, however, only those are suitable that completely meet the requirements for photomultipliers given in the first two points above. As may be seen from the data in Table 5.10, photomultipliers in current use usually have an antimony–caesium photocathode, or Sb–O–Cs cathodes sensitised with oxygen. (See, for example, the EMI-made 9552 S and 9552 B photomultipliers.) The only disadvantage of these cathodes is that they are not sensitive to red light. This is due to the value of their limiting wavelength: $\lambda_{max} = 6000$ Å. For the examination of radiations of longer wavelength photomultipliers with a composite photocathode (for example type 7102 made by RCA) should be used.

The foremost requirement of the dynode of a photomultiplier is that its emission coefficient should be as large as possible. It is also an advantage if the value, at which changes in the accelerating voltage does not any longer essentially influence the emission coefficient, is at minimum. Again, the antimony–caesium surface dynode best meets these requirements, as may also be seen from Fig. 5.65. Accordingly, the dynode of modern photomultipliers is generally of the antimony–caesium surface type.

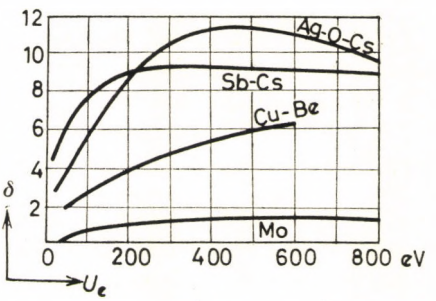


Fig. 5.65 Variation of the emission coefficient of normal dynode surfaces as a function of the energy of the primary electrons[1]

In practice the ratio of the secondary and primary electrons, that is, the value of the emission coefficient, is not determined solely by the chemical composition of the dynode surface. The value of the emission coefficient is also influenced by the angle of incidence of the electrons. It is also necessary that the form and arrangement of the electrodes should create an electrical field that reduces electron loss as far as possible. (Net-type dynode photomultipliers do not meet the latter requirement because their internal wall gains a charge that causes the secondary electrons to be partly diverted from the dynodes.)

In modern photomultipliers the undesirable scatter of the secondary electrons is almost entirely eliminated. In the photomultiplier shown in cross-section on Fig. 5.64*c* elimination of the stray secondaries is secured partly by the fact that the light illuminating the photocathode reaches it through a metal net at the same potential as the cathode. The mica sheet situated behind the photocathode eliminates unwanted 'feedback' since it stops secondary electrons returning to the photocathode from the dynode. In the sequential amplification stages the location and form of the dynodes create electrical fields that make loss of the primary and secondary electrons almost impossible. (From this point of view the construction of the dynode in front of the anode – indicated by (*9*) on the figure – is of particular interest.)

The number of multiplication stages is determined partly by the emission coefficient of the dynodes and partly by the required maximum intensity of the anode current,

Table 5.10

Important technical data* of the commonest types of photomultiplier: EMI (England), RCA (USA), VALVO (GFR), FEU (USSR), Carl Zeiss, Jena (GDR), Toshiba (Japan)

Product	Type	Cathode material	Cathode area (cm^2)	Wavelength range (nm)	Wavelength of maximum sensitivity (nm)	Main voltage (V)	Gain	Sensitivity ($\mu A/lm$)	Dark current (A)	Input noise equivalent** (W)
EMI	6256 S	CsSb	0·8	250–680	400	1 900	1.10^8	5 000	$1{\cdot}3.10^{-9}$	$4{\cdot}1.10^{-17}$
						1 700		2 000	$0{\cdot}5.10^{-9}$	
	9558 QB	NaKSbCs	15·0	200–850	280	1 500	2.10^6	2 000	$2{\cdot}0.10^{-8}$	$3{\cdot}7.10^{-16}$
						1 150	2.10^5	200	$2{\cdot}0.10^{-9}$	
RCA	IP-21	CsBs	1·6	300–650	420	1 250	2.10^6	80	$1{\cdot}0.10^{-7}$	$6{\cdot}0.10^{-15}$
	IP-22	BiCs	1·6	300–700	365	1 250	3.10^5	1	$2{\cdot}5.10^{-7}$	$2{\cdot}0.10^{-14}$
	IP-28	CsSb	1·6	250–650	340	1 250	1.10^6	20	$1{\cdot}0.10^{-7}$	$1{\cdot}0.10^{-14}$
	7265	NaKSbCs	13·0	200–850	420	2 400	1.10^7	1 400	—	$1{\cdot}0.10^{-15}$
VALVO	XP1110	CsSb	1·5	320–650	420	1 800	—	250	$2{\cdot}0.10^{-8}$	—
	XP1118	CsSb	1·5	350–680	400	1 800	—	250	$2{\cdot}0.10^{-8}$	—
	XP1116	AgOCs	1·5	300–1150	800	1 800	—	100	$1{\cdot}0.10^{-5}$	—
	XP1117	NaKSbCs	1·5	200–850	350	1 800	—	100	$1{\cdot}0.10^{-8}$	—
FEU	FEU-19	CsSbO	7·0	250–680	420	1 400	2.10^6	1 000	$4{\cdot}0.10^{-7}$	$3{\cdot}0.10^{-14}$
						1 150	2.10^5	100	$1{\cdot}0.10^{-7}$	
	FEU-62	AgOCs	0·8	300–1150	800	1 800	—	10	$6{\cdot}0.10^{-7}$	$3{\cdot}0.10^{-12}$
	FEU-64	CsSb	15·0	250–680	380	1 500	—	1 000	$5{\cdot}0.10^{-8}$	$8{\cdot}0.10^{-15}$
Carl Zeiss	K-14FQS-50	CsSb	14·0	250–680	380	2 800	2.10^7	5 000	—	$1{\cdot}5.10^{-14}$
	M-12FQC-52	NaKSbCs	15·0	200–850	360	1 500	3.10^5	120	—	$5{\cdot}0.10^{-14}$
Toshiba	PM-55	NaKSbCs	15·0	200–850	360	1 800	5.10^5	100	—	$2{\cdot}0.10^{-14}$

* All the data refer to 20 °C working temperature.

** At a bondwidth of 1 Hz. Thanks are due to the physicist Mr. G. Virághalmy for his help in compiling the table.

that is, by the sensitivity of the photomultiplier necessary for the intensity measurements. (Thus, for example, in order to achieve the same maximum anode current – under otherwise identical experimental conditions – 17 stages are needed when using Cu–Be dynodes, while only nine stages are required when Sb–Cs dynodes are used.)

The value of the voltage between the amplification stages and hence the value of the terminal voltage of the photomultiplier is determined by the voltage–anode current characteristic curve of the photomultiplier. The curvature of this curve increases as the amount of light flux illuminating the photocathode increases (Fig. 5.66).

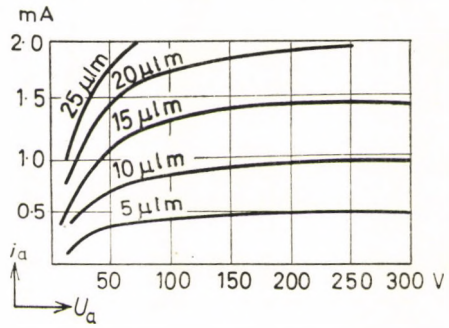


Fig. 5.66 Variation of the anode current of the RCA IP 21 photomultiplier, as a function of the voltage during amplification, at different illumination levels of the photocathode

However, the greater the curvature of this curve, the higher the voltage will be at which the optimum condition, of voltage fluctuations not significantly affecting the anode current, is realised. The voltage–anode current characteristic curve therefore gives the minimum value of the terminal voltage of the photomultiplier at which it can be satisfactorily used for intensity measurements, at a particular illumination level of the photocathode.

The anode current of a photomultiplier varies almost exponentially as a function of the voltage between the dynodes (Fig. 5.67). This means that the relative photo-

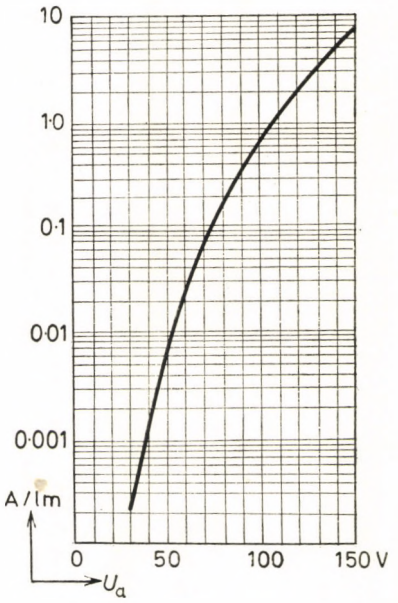


Fig. 5.67 Relationship between the sensitivity of RCA photomultipliers and the acceleration voltage between the dynodes

effect i/Φ – that is, the sensitivity of the photomultiplier – also varies exponentially as a function of the accelerating voltage. For example, when this voltage changes by 100 V, the value of the relative photoeffect changes by about three orders of magnitude. There is a practical limit to the increase of sensitivity as a result of the fact that when the voltage is increased excessively the effects that reduce the accuracy of the measurements also greatly increase. Also, the anode current cannot be increased above a certain limit – usually 1 mA – without having harmful effects on the photomultiplier.

In photomultipliers used at present for measurement purposes the greatest voltage difference between the subsequent amplification stages is usually 140 V, and 50 V between the last dynode and the anode. These values and the number of amplification stages determine the maximum terminal voltage of the photomultiplier.

The following rules are deduced from the above discussion:

(1) The sensitivity of the photomultiplier can be controlled by the terminal voltage. The sensitivity varies almost exponentially as a function of the latter.
(2) Errors occur when during the measurement series there is a small change in the terminal voltage of the photomultiplier, or in the voltage between the subsequent amplification stages.

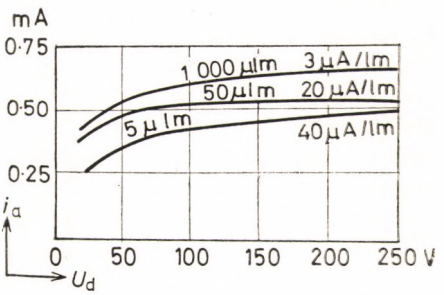


Fig. 5.68 Variation of the anode current of RCA IP 21, IP 22 and 931A photomultipliers, as a function of the voltage between the dynodes, for different illumination levels of the photocathode

(3) The voltage–anode current characteristic determines the appropriate use of any particular photomultiplier. It gives partly the optimum voltage applied to the photomultiplier, and it determines the recommended maximum illumination of the photocathode.
(4) By selecting a photomultiplier with a suitably sensitive photocathode optimum experimental conditions can be obtained for intensity measurements even when the light flux falling on the photocathode varies between wide limits. From Fig. 5.68 it can be seen that there is no significant difference between the voltage–anode current characteristics of the photomultipliers considered, despite the fact that the light flux falling upon them varies as follows

1000 µlm for a photomultiplier of sensitivity 3 µA/lm
50 µlm for a photomultiplier of sensitivity 20 µA/lm
5 µlm for a photomultiplier of sensitivity 40 µA/lm

References

1. GÖRLICH, P., *Die lichtelektrischen Zellen*, Leipzig (1951)
2. SOMMER, A., *Photoelectric Tubes*, 2nd edn, London (1950)
3. TURK, W. E., *Photoelect. Spectrom. Grp. Bull.*, **5**, 100 (1952)

4. ZWORYKIN, V. K., MORTON, G. A. and MALTER, L., *Proc. Inst. Radio Engrs*, **24**, 351 (1936)
5. ZWORYKIN, V. K. and RAYCHMAN, J. A., *Proc. Inst. Radio Engrs*, **27**, 558 (1936)
6. ENGSTROM, R. W., *J. Opt. Soc. Am.*, **27**, 420 (1947)

5.11.3 Dark current of photomultipliers

The external circuit of a photomultiplier is not current-free even when there is no light reaching the photocathode – as is generally the case with all types of photocells. This current is the so-called *dark current*. It is caused by the collective effect of a number of factors including:

(1) The thermal emission of the photocathode.
(2) The thermal emission of the dynodes.
(3) The so-called *ion current*, that is, the current produced by ionisation of the gas traces left behind in the tube and the resulting ions colliding with the dynodes.
(4) The isolation current arising from imperfections of the insulation.
(5) The emission arising from the effect of the electric field at the electrodes.

The last two factors are independent of temperature while the current resulting from the other effects varies as a function of the temperature. The influence of the last three factors can be reduced almost to nothing by evacuation of the tube and by its appropriate construction.

Conversely, the effect of the thermal emission arising on the photocathode and the dynodes cannot be altered by perfecting the photomultiplier. The value of the thermal emission depends on the quality of the light-sensitive layer of the photocathode and the dynodes. The thermal emission, and thus the intensity of the dark current, is the lower the shorter the threshold wavelength of the photoeffect produced by the electrodes. The currents produced by the effect of the thermal emission occurring at the photocathode and the dynodes appear collectively in a photomultiplier so that their separation is difficult. However, it can be established that the emission currents arising on the sequential dynodes appear with an amplification factor: $(\delta^n + \delta^{n-1} + \ldots + 1)$.

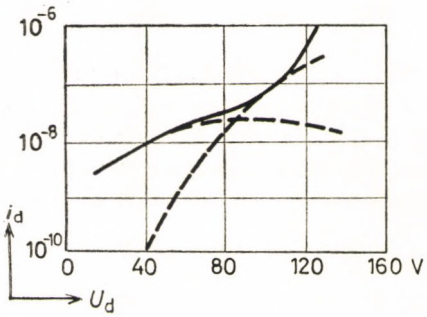


Fig. 5.69 Variation of the dark current (solid line), the thermal emission (steep dotted line), and the insulation current (almost horizontal dotted line) as a function of the voltage during the amplification stage of an RCA photomultiplier[1]

It follows from the foregoing, and it can also be shown experimentally, that the intensity of the thermal emission and of the isolation current – thus indirectly that of the dark current – increases with the voltage difference of the amplification stages (Fig. 5.69). While the variation of the isolation current approaches a limit upon increasing the voltage, the intensity of the thermal emission and thus indirectly that of the dark current increases steeply when a certain voltage level is reached. This

has a detrimental effect on the recommended voltage difference of the amplification stages.

When the photomultiplier is illuminated by intense light the intensity of the dark current is greatly enhanced. (This phenomenon occurs even when the amplification stages are under zero voltage.) The deviation arising this way becomes larger with longer exposures.[2,3] The normal, almost constant value of the dark current is resumed asymptotically and very slowly after the intense illumination of the photomultiplier;

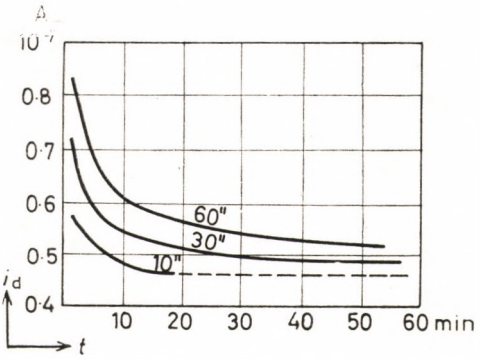


Fig. 5.70 Variation of the dark current of a photomultiplier with a Sb–Cs photocathode, as a function of the duration of illumination. (The dotted line gives the original intensity of the dark current)[4]

the time of recovery is longer, the longer the duration of the intense illumination (Fig. 5.70). Therefore, it is recommended that between measurements the photomultiplier is illuminated by light of approximately the same intensity as that measured. (A so-called 'fatigue-lamp' is used for this purpose.) A further source of error can occur if the photomultiplier is not stored in darkness.

References

1. ENGSTROM, R. W., *J. Opt. Soc. Am.*, **37**, 420 (1947)
2. SCHAETTI, N. and BAUMGARTNER, W., *Helv. Phys. Acta*, **25**, 605 (1952)
3. ECKART, F. *Annln Phys.*, **16**, 322 (1955)
4. SCHAETTI, N. and BAUMGARTNER, W., *Z. angew. Math. Phys.*, **4**, 159 (1953)

5.11.4 Shot effect of photomultipliers

When using photomultipliers the lower limit of the light intensity measurements is also determined by the shot effect of the photomultipliers. This is the mean square average of the time fluctuations of the anode current that are observed when the photocathode is illuminated by light of constant intensity. Its value is therefore

$$\overline{i_z^2} = \overline{(i - i_k)^2} \tag{5.67}$$

where i is the instantaneous current and i_k is the mean value of the current.[1,2]

The shot effect of photomultipliers is mainly caused by the following factors:

(1) The statistical fluctuation of the electron currents, that is, the so-called shot noise.
(2) The statistical fluctuations of the current distribution between the electrodes, the so-called *current-distribution noise*.
(3) The so-called *insulation noise*, that is, the statistical fluctuations of the isolation current produced by statistical fluctuations of the insulation.

It can be calculated that the shot effect of a photomultiplier is given by the following formula, when acceptable approximations have been made

$$\overline{i_z^2} = 2\Delta f e_0 S \sqrt{i_a} \tag{5.68}$$

where $2\Delta f$ is the reciprocal time constant of the measuring device, e_0 is the value of the elementary charge, S is the gain factor, and i_a is the intensity of the anode current. It follows from this – but it can also be justified experimentally – that the shot current of photomultipliers varies in proportion to the square root of the anode current (that is, proportionally to $\sqrt{i_a}$, usually between the limits 10^{-7} and 10^{-6} A. In some photomultipliers the shot current falls with falls in the gain, at higher anode current values.

The 'quality' of a multiplier is characterised by the signal to noise ratio. Its square is given by

$$\left(\frac{S}{N}\right)^2 = \frac{i_a^2}{i_z^2} = \frac{\Phi E}{2fe} \tag{5.69}$$

where i_a is the effective anode current of the illuminated photomultiplier, Φ is the light flux in lumens falling on the photocathode, and E is the sensitivity of the cathode in A/lm. It follows from the above relationship that the 'quality' of the photomultiplier is independent of the amplification factor. It also follows, however, that the 'quality' of a photomultiplier becomes higher with greater cathode sensitivity and higher light intensities. The first factor justifies the use of Sb–Cs photocathodes.

References

1. ECKART, F., *Annln Phys.*, **11**, 194 (1953)
2. BRUINING, H., *Die Sekundärelektronen-Emission fester Körper*, Berlin–Göttingen–Heidelberg (1942)

5.11.5 Principal faults of photomultipliers

Measurement errors can occur in the following circumstances:

(1) when the photocurrent of the photomultiplier does not vary linearly with the intensity of illumination of the photocathode, and
(2) if the photocurrent changes while the cathode is illuminated by a constant light flux.

Accordingly, error sources can be grouped into two categories.

The first category consists of those errors that are evidenced by non-linearities. It can be shown experimentally that the ideal relationship $\bar{i} = C\Phi$ between the photomultiplier photocurrent and the strength of illumination is not strictly applicable when measurements are made over a wide range. This deviation also depends on the wavelength of the light (see for example references 1–3).

The disproportionality can probably be attributed to those factors that cause the 'fatigue' of photomultipliers. The incorrect position of the electrodes can also give rise to errors. This can distort the electric field of the multiplication stages, depending on the value of the current passing through it, that is, depending on the illumination

strength of the photocathode. Disproportionalities can also be observed – as seen in practice – when the ratio of the current passing through the voltage divider belonging to the photomultiplier and the anode current is smaller than 0·1.

In practice, with some photomultipliers the disproportionalities can cause up to 5% error when the intensity range extends over three orders of magnitude.[4] On the other hand, among photomultipliers of the same make and of the same type some component can generally be found that makes it possible to achieve an accuracy of 0·1%. Obviously, only such photomultipliers can be used for spectrochemical purposes.

The time variation of the anode current is caused first of all by the 'fatigue effects' of the photomultiplier. This is either a reversible or an irreversible process. The irreversible process – that is, the ageing of the photomultiplier – occurs very slowly. It is not therefore a source of error when a series of measurements is performed over a fairly short period of time. In the opposite case this error source cannot be neglected. In practice, the ageing process causes the calibration curves used for quantitative analysis to be shifted parallel to the intensity axis of the graph, but their shape is not changed.

The reversible fatigue phenomenon of photomultipliers occurs either on the photocathode, or on the dynodes. It is indicated by the fact that over a few hours the anode current increases by about 10% of its original value.[5,6] Hence, this phenomenon could give rise to the same error that is caused by ageing. This effect can only be neglected therefore, when the measurements in a series are made within a short period of time.

The fatigue of photomultipliers can be attributed to a number of causes.[7] The role of the secondary emission produced by the positively charged ions arising from the gas traces left behind in the tube, by colliding with the cathode and the dynode of the previous multiplication stage, can be important. Secondary emission may also occur if local potential differences are established on the surface of the cathode or the dynodes. These are due to the fact that the surfaces of the cathode and the dynodes are not homogeneous as regards the work function, and the surface conductivity is not perfect over the whole area. In each case the electrons emitted from one point of the surface can re-enter at some other point on the surface, thus causing unwanted feedback. The same thing can also happen when the light does not evenly illuminate the entire surface of the photocathode. The current produced by the feedback causes a larger alteration of the anode current as the number of amplification stages increases, through which this current passes in the photomultiplier. Reversible fatigue of a photomultiplier can also occur when strong light occasionally falls on the photocathode of the photomultiplier.

As a contrast to the sources of error mentioned above an advantage is that the anode current of photomultipliers is practically independent of temperature variations.[4]

References

1. KORTÜM, G. and MAIER, H., *Z. Naturf.*, **8a**, 233 (1955)
2. HERMANN, W., *Z. Naturf.*, **12a**, 1006 (1952)
3. CANNON, C. G. and BUTTERWORTH, I. S., *Analyt. Chem.*, **25**, 168 (1953)

4. KORTÜM, G., *Kolorimetrie, Photometrie und Spektrometrie*. vol. 2, 4th edn, Berlin–Göttingen–Heidelberg, 169 (1962)
5. ENGSTROM, R. W., *J. Opt. Soc. Am.*, **37**, 420 (1947)
6. HILLER, M., *Br. J. Appl. Phys.*, **2**, 164 (1954)
7. PRESTON, J. S., *Rev. Opt. théor. Instrum.*, **27**, 513 (1948)

5.11.6 Power supplies for photomultipliers

It follows from the foregoing that the power supplies needed for the operation of photomultipliers must give a precisely stabilised voltage of the order of a thousand kV and at the same time the voltage between the individual amplification stages must be identical. The latter requirement can be most simply met by the use of a voltage divider composed of an adequate number of resistors connected in series. It is also necessary for the current passing through the voltage divider to be a few orders of magnitude greater than the anode current of the photomultiplier connected parallel to it. In this way the variations of the anode current do not affect the voltage drops occurring at the individual resistors.

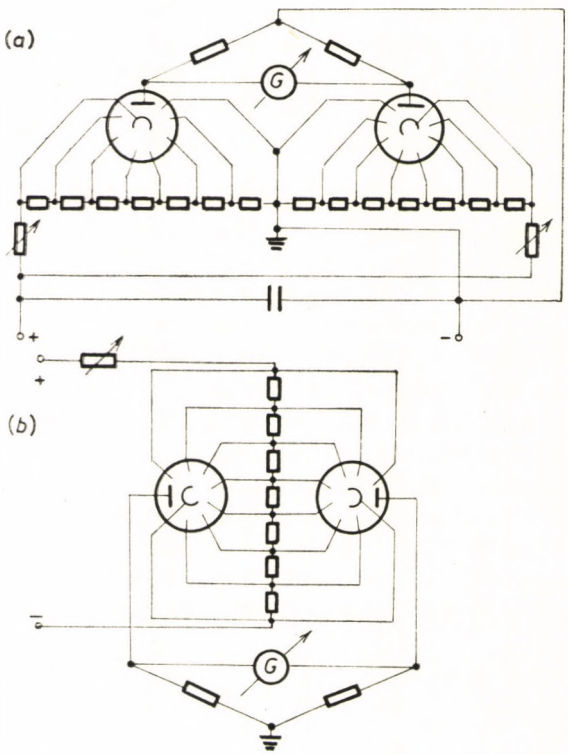


Fig. 5.71 Circuit diagram for a photomultiplier couple: (*a*) in series and (*b*) in parallel[1]

In quantitative spectrochemical analysis the intensity ratio of two lines is always what is to be determined. To perform such measurements one needs two photomultipliers. These are connected either in series or parallel to one another (Fig. 5.71). The advantage of the first arrangement is that the voltage applied to both photomultipliers can be controlled separately. In the second arrangement this advantage can only be achieved by means of more complicated equipment.

It is well known that the sensitivity of a photomultiplier varies almost exponentially as a function of the voltage applied to the individual multiplication stages. Hence, measurement errors are avoided only when the voltage level can be controlled to an accuracy of at least 0·001 % during a series of measurements. This requirement can only be met by means of a voltage stabiliser – that is, by a device that ensures a constant voltage level by a very rapid correction when there is a change in the mains voltage or output current. This is achieved by appropriate electronic apparatus. There are a number of systems for this purpose. The stabiliser systems used as the power supply for photomultipliers usually operate on the principle that the mains alternating current, approximately stabilised, is first transformed to a current of a few kilovolts voltage. The voltage of the direct current obtained by full-wave rectification of this alternating current is first smoothed by condensers of suitable capacitance, then stabilised by electron tubes whose anode voltage and grid potential have been appropriately stabilised (for example Fig. 5.72). Voltage stabilisation is not generally done entirely by means of electron tubes; additional stabilising devices, such as magnetic voltage-control units, glimm stabilisers, and occasionally iron–hydrogen resistors, are also necessary.[2,3]

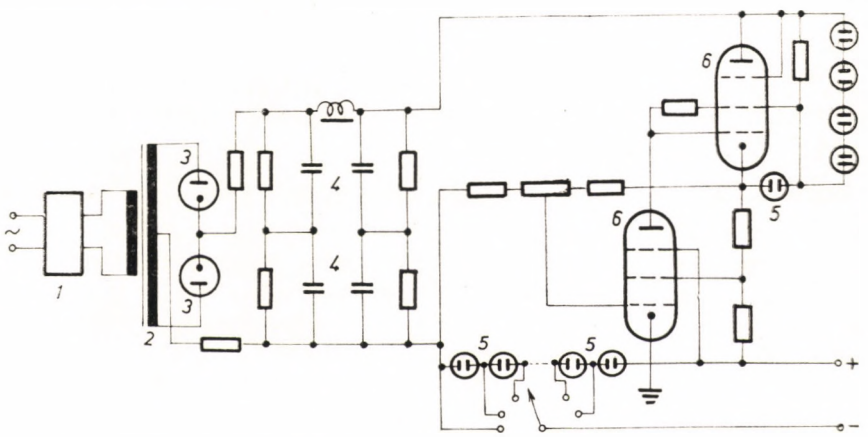


Fig. 5.72 Simplified circuit diagram of a voltage stabiliser that can be used as the power supply for a photomultiplier: (*1*) magnetic voltage stabiliser, (*2*) transformer, (*3*) rectifying tube, (*4*) smoothing condenser, (*5*) glimm stabiliser, (*6*) stabilising tube

Magnetic voltage stabilisers operate on the following principles. The secondary voltage of a transformer with an iron core at first linearly increases with the current; then, having reached a certain secondary voltage limit, it becomes independent of the current level. When such a transformer is coupled to another so that the secondary voltage is in the opposite direction, then – at a suitable ratio of transmissions of the two transformers – the difference of the two secondary voltages is, within certain limits, independent of variations in the mains voltage. Thus, for example, a $\pm 15\%$ fluctuation of the mains voltage may be reduced to ± 1–2%.

The current of the external power supply can also be stabilised to a certain extent by an iron–hydrogen resistor, that is, by an iron wire fastened into a tube filled with hydrogen at atmospheric pressure. This controls the current flowing through so that

its value becomes independent of the voltage variations, within certain limits. The temperature and thus the resistance of the iron wire varies as a function of the current – that is of the mains voltage – in such a way that the thermal energy liberated by the current is equilibrated with the heat losses resulting from the hydrogen being a good thermal conductor. Thus voltage variations of about $\pm 10\%$ can be reduced to about $\pm 1\%$ when the voltage belonging to the *average* current level falls at the midpoint of the voltage range controllable by iron–hydrogen resistors. However, iron–hydrogen resistors are rather sluggish and equilibrium is only reached after several seconds.

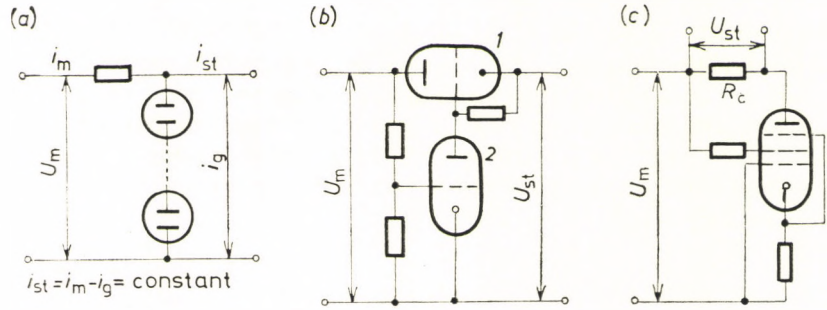


Fig. 5.73 Circuit principles of different voltage stabilising methods: (*a*) glimm stabiliser, (*b*) common electron-tube voltage stabiliser, and (*c*) current stabiliser used as a voltage stabiliser

Glimm stabilisers are gas-discharge tubes usually filled with helium at a pressure of 15 mmHg. They can therefore only be used for the stabilisation of suitable direct-current voltages. Their operation is based on the principle that their internal resistance and thereby the voltage drop across them decrease with increasing current – within certain limits. Accordingly, they can only be used combined with a series-connected load resistor. The voltage drop across the load resistor increases so much upon the increase of the current that if the resistor is properly chosen it compensates for the voltage drop through the glimm tube. Thus when the voltage to be stabilised varies by $\pm 10\%$ it is possible to reduce the voltage fluctuation to only $\pm 0{\cdot}1$–$0{\cdot}2\%$, if the glimm tube–resistor system is connected in parallel with the voltage to be stabilised (Fig. 5.73*a*). The glimm stabiliser is generally used for the stabilisation of the anode voltage and grid potential of the electron tubes (see, for example, the upper right part of Fig. 5.72).

Voltage stabilisation can be done most successfully and without any lag by means of electron tubes. Among the many possible solutions to the problem the one, for example, depicted on Fig. 5.73*b* can be used to advantage. This system is based on the fact that the stabiliser tube (*1*) of small internal resistance connected into the circuit plays the role of a variable resistor the value of which is controlled by the amplifier tube (*2*). When the mains voltage U_m increases upon the indirect effect of the latter the grid voltage of the tube is also increased and its anode voltage is therefore reduced. This has, in turn, the consequence that the grid voltage of the stabiliser tube (*1*) is shifted in the negative direction and thereby the internal resistance of this tube is increased. Thus the stabilised voltage U_{st} does not change. (The voltage stabiliser of the circuitry given in Fig. 5.72 is also based on this principle.)

Another frequently used method of voltage stabilisation – shown in Fig. 5.72 – is based on current stabilisation. It can be used for voltage stabilisation because the voltage drop across a certain resistance is constant if the current passing through it is also constant (Fig. 5.73*c*). The current value of the current flowing through the pentode depends within certain limits on the bias voltage of the control grid; that is, it depends indirectly on the value of the resistance R_k. This control is successful only when the voltage of the shielding grid is exactly constant, that is, when this is secured, for example, by a glimm stabiliser.

References

1. KORTÜM, W. and MAIER, H., *Z. Naturf.*, **8a**, 235 (1955)
2. CANNON, C. G., *Electronics for Spectroscopists*, London (1960)
3. KRETZMANN, R., *Industrial Electronics Handbook*, Eindhoven (1967)

5.12 Application of photomultipliers

5.12.1 Fundamental principles[1–7]

In spectrochemical analysis only a very small fraction of the light from the source falls on a single spectral line. This intensity is especially low when the element responsible for the line is only present in the sample in trace quantities. Thus the methods that use a photocell or photo-conductive cell for the light-intensity measurements can only be used in a limited way, mainly in flame photometry. On the other hand, the sensitivity of modern photomultipliers with amplification factors of about 10^{10} is so high that very low intensities can be measured which would only produce an observable density on the most sensitive emulsion after an exposure of several hours.

The inertia of photomultipliers is relatively small, so they can be used for recording very rapid variations of intensity. This is, however, no advantage for quantitative spectrochemical analysis – especially in the case of arc and spark excitation. In the examination of, for example, chromium steels by spark excitation the absolute intensity of the chromium lines varies by a few hundred per cent over a half-period, while the fluctuations in the intensity ratio of the Fe–Cr analytical line pairs can be as much as 100%. In addition, the intensity of the radiation produced during each half-period is not constant, but varies continuously with the temperature of the plasma. In order to attain maximum analytical reproducibility it is necessary to average thousands of discharges, that is, an integration period of a few tenths of a second is required. (Further prolongation of the averaging time does not increase the reproducibility.) Integration of the absolute intensity of the analytical line of the element to be determined can never obtain the accuracy that is obtainable by the averaging of the intensity ratio of the line pair.

In photographic spectrochemical analysis the photographic emulsion *automatically* averages the intensity fluctuations previously mentioned. This does not happen in the photoelectric measurement of the intensity, that is, when using photomultipliers. In this case, to achieve averaging a separate procedure is necessary.

The methods at present used for the determination of the mean value of the intensity ratio I_x/I_r belonging to the analytical line pair x and r can be divided into two categories: the *integrating-condenser* and the *graphical* procedures.

References

1. DUFFENDOCK, O. S. and MORRIS, S. H., *J. Opt. Soc. Am.*, **32,** 8 (1942)
2. HASLER, M. F. and DIETERT, H. W., *J. Opt. Soc. Am.*, **34,** 751 (1944)
3. BOETTNER, E. A. and BREWINGTON, G. P., *J. Opt. Soc. Am.*, **34,** 6 (1944)
4. HAHSTOLL, G. A. and BRYAN, F. R., *J. Opt. Soc. Am.*, **35,** 645 (1945)
5. SAUNDERSON, J. L., CALDECOURT, V. J. and PETERSEN, E. W., *J. Opt. Soc. Am.*, **35,** 680 (1945)
6. DIECKE, G. H. and CROSSWHITE, H. M., *J. Opt. Soc. Am.*, **35,** 471 (1945)
7. STEEN, I. M., WILLIAMS, A. J. and CLARK, W. R., *Trans. Am. Soc. mech. Engrs*, **66,** 271 (1944)

5.12.2 Principle of integrating-condenser measurements

The essential element of this method, at present most widely used, is that the anode current of the photomultiplier is used to charge the so-called *integrating condenser.* Thus the terminal voltage of the condenser is a measure of the average value of the light intensity during the period of charging of the condenser.

Considering that the intensity I of the photocurrent of the photomultiplier is *practically* proportional to the intensity I, of light falling on the cathode, the following can be written: $i = kI$, where k is an appropriate proportionality factor. If the average of the light intensity I from instant t_1 to instant t_2 is denoted by $\bar{I}$, then on the basis of the former equation

$$\bar{I} = \frac{1}{t_2 - t_1} \frac{1}{k} \int_{t_1}^{t_2} i \, \mathrm{d}t \tag{5.70}$$

The integral of the current is, however, identical to the charge Q of the integrating condenser. Thus the following equation is valid

$$\int_{t_1}^{t_2} i \mathrm{d}t = Q = CU \tag{5.71}$$

where C is the capacitance, and U is the terminal voltage of the condenser. From these equations it follows

$$I = \frac{1}{t_2 - t_1} \frac{C}{k} U \tag{5.72}$$

The mean current value of the photocurrent is therefore proportional to the terminal voltage of the condenser in question, and inversely proportional to the charging time of the condenser.

It follows from the previous equation that the ratio of the intensities $\bar{I}_x$ and $\bar{I}_r$ of the spectral lines x and r is

$$\frac{\bar{I}_x}{\bar{I}_r} = \frac{U_x}{U_r} \tag{5.73}$$

This relationship is only valid, however, when the capacitance of the integrating condensers belonging to the photomultipliers measuring the radiation of the spectral lines r and x, and also the proportionality constant k for the two photomultipliers used, are identical. In the opposite case – that is, when these two conditions are not

fulfilled – the suitable proportionality constant k' should be included, that is

$$\frac{\bar{I}_x}{I_r} = k' \frac{U_x}{U_r} \tag{5.74}$$

In both cases, however, the ratio of the average intensity of the two spectral lines is proportional to the ratio of the terminal voltage of the two integrating condensers.

For the measurement of the quotient U_x/U_r there are two methods that may be used, based on different principles, namely the *charging* and the *discharging* methods. The charging method is the simpler, thus it is the more frequently used method. In its practical applications the anode current of the two photomultipliers charges the two integrating condensers for a time that is determined in advance. An appropriate electronic device interrupts the charging of the condenser at that instant in which the current has already charged the integrating consenser of the line r of the reference element to a predetermined voltage level U_r. This method has the advantage that Eq. (5.73) is simplified to the following equation

$$\frac{\bar{I}_x}{\bar{I}_r} = U_x \tag{5.75}$$

if the voltage U_x is measured in U_r-units, that is, when $U_r = 1$. In this arrangement, therefore, the integration time is not constant, but varies as a function of the intensity of the reference line r. The advantage is, however, obtained that the intensity ratio of the two lines varies proportionally with the voltage U_x.

In the method previously described the terminal voltage of the integrating condenser is usually measured by an electronic device based on the *vibrating reed tube voltmeter*. This voltage measurement is based on the fact that the capacitance of the unit formed by the parallel-connected integrating condenser and vibrating condenser varies periodically if the plate of the latter condenser vibrates evenly. Considering, however, that the terminal voltage of the identically charged condenser is inversely proportional to the capacitance, as shown by Eq. (5.71), it can be understood that the terminal voltage of the condenser unit also varies periodically. The fluctuating voltage produced thus can be electronically amplified (without the slightest change of the terminal voltage of the integrating condenser) to such an extent that a suitable recorder – the self-balancing potentiometer – can be used to record the value of the voltage U_x.

The terminal voltage U_x of the integrating condenser can be determined not only by the method described above, but also by using decadic counting tubes. In this case the pulses necessary for the operation of the decatron tubes are supplied by a high-precision oscillator operated by a piezoelectric quartz crystal. The counting system is controlled directly by the difference of the terminal voltage U_x of the integrating condenser system and pulses whose height increases linearly with time. Thus counting stops at the instant when the steadily increasing reference voltage attains the level of the voltage to be measured. The value read from the decatron tube faces is directly proportional to the value of U_x.

The other version of the integrating-condenser method, the discharge method, is based on time measurements. In this procedure the integrating condenser corresponding to the line r of the reference element is discharged through the resistor R exactly when its terminal voltage becomes just equal to the terminal voltage U_x of

the integrating condenser corresponding to line x. Thus it can be written that

$$U_x = U_r e^{-t/RC} \tag{5.76}$$

where t is the time of discharge of the condenser of capacitance C. According to the above equation, therefore, $U_x/U_r = e^{-t/RC}$. Thus we find, according to Eq. (5.74), that the average intensity ratio of the two lines is

$$\bar{I}_x/\bar{I}_r = k' e^{-t/RC} \tag{5.77}$$

where k' is a constant whose value depends partly on the capacitance of the two integrating condensers and partly on the ratio of the proportionality constants for the two photomultipliers. It follows from the relationship that

$$\log I_x - \log I_r = \Delta Y_{xr} = k'' t \tag{5.78}$$

that is, in the discharge method the difference of the logarithmic intensities of the two spectral lines varies proportionally with the discharge time.

5.12.3 Principle of the graphical method

The procedure is based on a system in which the recorder draws the variation of the intensity ratio of the analytical lines as a function of time. The *record* so obtained is then graphically averaged.

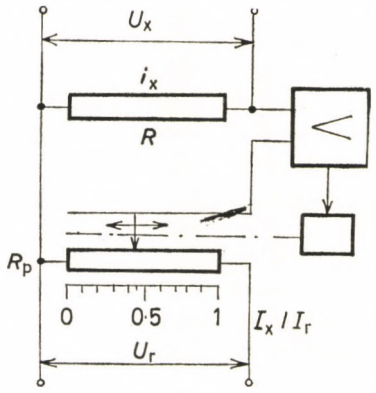


Fig. 5.74 Graphical procedure: fundamenta principle of the procedure

In the graphical method the two photomultipliers are set exactly at the spectral lines x and r. The photocurrents of the photomultipliers are amplified by an electronic device in such a way that the currents i_x and i_r obtained after the amplification are exactly proportional to the light intensities I_x and I_r falling at the cathodes of the two photomultipliers, provided these are connected to an identical resistor R and another resistor R_p formed as a potentiometer (Fig. 5.74). Thus voltage drops U_x and U_r are produced on the two resistors of the recording potentiometer that are proportional to the currents i_x and i_r, therefore proportional to the light intensities I_x and I_r falling on the two photomultipliers. When the servomotor M, controlled electronically, shifts the slider of the recording potentiometer until the voltage difference between the slider and the zero-point of the potentiometer becomes exactly U_x, then the intensity ratio I_x/I_r can be directly read from the potentiometer scale

or from the recording paper. This record is obtained by a second servomotor moving the recording paper evenly in the direction of the time-axis. It is recommended that this movement of the recording paper be relatively slow, that is the time constant of the recording potentiometer is so large that the record consists of only a few zigzags. This simplification makes the graphical evaluation easier.

In comparisons of the graphical and integrating-condenser methods it is apparent that – while the integrating-condenser method gives a 'blind average' – on the record one can clearly see the fluctuations of the intensity ratio during the integration time. Hence, in the graphical method it is an advantage that effects detrimental to the accuracy of the analysis can be observed (for example, inhomogeneity of the sample, contact troubles, etc.). On the other hand, a drawback of the graphical procedure is that it is relatively time-consuming as compared to the integrating-condenser method. The accuracy of the averaging is influenced by personal characteristics.

In the other version of the graphical procedure, the 'profilage' method, an appropriate device moves the spectrum of the material under study perpendicular to the slit in front of the photomultiplier, that is, along the length of the spectrum.[1] Thus, if the dispersive apparatus is stigmatic the recording system of the apparatus draws the profile of the line onto the evenly moving recording paper. From the maximum of the curve so obtained one can infer the intensity of the spectral line. The advantage of the method is that even when using spectrographs of relatively poor linear dispersion any line in the immediate vicinity of the line being measured does cause difficulties. On the other hand, a drawback of the method is that spectral lines can only be measured individually, therefore the intensity ratio of two lines cannot be directly established. This leads to analytical inaccuracy.

Reference

1. BRECKPOT, R. and CLIPPELEIR, K., *Congrès Groupement pour l'Avancement des Méthodes d'Analyse Spectrographique des Produits Métallurgiques*, **17**, 179 (1954)

5.13 Visual intensity measurements

5.13.1 Fundamental principles of the method

In visual intensity measurements the radiation receptor is the human eye. The properties of the eye therefore largely determine the achievable measurement accuracy, the optimum experimental conditions, etc.

It is known that the eye observes only the identity or difference of the illumination strength of two illuminated surfaces. It is not possible to determine the ratio of the illumination strengths by eye. This fact determines the basis of visual light-intensity measurements, the optimum experimental conditions, etc. The contrast sensitivity of the eye greatly influences the accuracy of the measurements.

According to the Weber–Fechner law the ratio of the difference ΔL of surface luminosity of two illuminated surfaces, which can be just established by eye, and the surface luminosity L – that is, *the contrast sensitivity of the eye* – is constant under certain circumstances, that is

$$\Delta s = \frac{\Delta L}{L} = \text{const.} \tag{5.79}$$

This law applies only when the surface luminosity of the illuminated area is greater than 20 asb and smaller than 1000 asb, and when the two surfaces of appropriate area are in contact. Over these limits the contrast sensitivity of the eye falls rapidly.

According to the König–Brodhun rule the maximum percentage contrast sensitivity of the eye is usually 1·7%, and is generally only half this value even in favourable conditions. It has furthermore been proved experimentally that when trying to determine the equal surface luminosity of two surfaces, a scatter of at least ±1% should be allowed for, even under optimum experimental conditions. This value also controls the level of accuracy attainable in visual intensity measurements

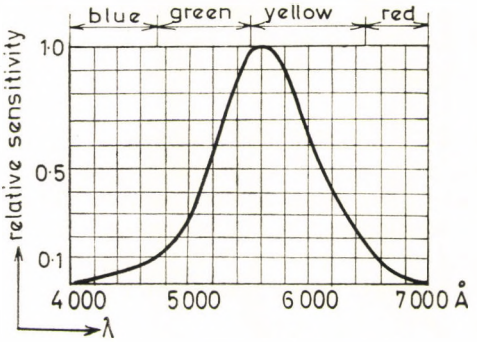


Fig. 5.75 Colour sensitivity curve of the human eye[1]

The above data refer to white light, or to wavelengths of light to which the eye is most sensitive. The sensitivity of the eye varies greatly with the wavelength of light (Fig. 5.75). The colour sensitivity of the eye (the so-called visibility curve) is at a maximum about 5500 Å, that is for yellowish-green light. At longer wavelengths the sensitivity falls rapidly; from 5100 Å to 6200 Å – that is from green to orange-yellow light – the sensitivity falls by 50%. In addition, the eye sensitivity curve moves towards shorter wavelengths when the surface luminosity is very low, that is, lower than about 0·001 asb. In these circumstances the maximum sensitivity occurs at about 5100 Å, after the eye has adapted to the low level of luminosity.

The fact that the sensitivity of the eye varies with wavelength is a drawback for intensity measurements as the contrast sensitivity of the eye and thus the accuracy of the measurements also depend on the wavelength of light. Hence, a lower accuracy is achieved with spectral lines of blue and red colour than with yellowish-green lines. A further disadvantage is that comparisons become uncertain when spectral-line wavelengths are slightly different, that is, when their colour differs visibly.

The contrast sensitivity of the eye also depends on the size and distance apart of the surfaces to be compared. The maximum accuracy is achieved when along a large length the two surfaces are in direct contact. In this case the equal surface luminosity of the two areas is indicated by the disappearance of the separating line. This phenomenon can be explained by the Crawford effect. The human eye sees rays of the same intensity to have a lower intensity as the distance increases from the centre of the pupil of the eye to the point where the narrow beam of rays enter the eye. This visual error can sometimes cause an error of 50%.

It follows that the absolute intensity of a line cannot be established by visual means. The unaided eye is only useful for establishing the equality or difference of intensity of two lines.

The intensity ratio of two spectral lines can only be determined by the flux of the more luminous line being quantitatively reduced until the intensity of the two lines appears to be identical. There are two methods that can be used for the measurable reduction of the light. The light flux entering the eye can be reduced either by a grey wedge or by a procedure based on polarisation effects. The other methods generally used for the quantitative reduction of light (such as those based on the rotating sector or on distance alteration) cannot be recommended for various technical reasons.

Reference

1. KORTÜM, G., *Kolorimetrie, Photometrie und Spektrometrie*, vol. 2, 4th edn, Berlin–Göttingen–Heidelberg (1962)

5.13.2 Measurement by the grey-wedge method[1,2]

Measurable reduction of light with a grey wedge is based on the absorption of light and hence on the Bouguer–Lambert law. According to this law, the absorbing homogeneous medium causes a loss of energy the value of which varies in proportion to the infinitesimal thickness of the medium, ds, and the light flux Φ of light falling on the medium

$$-\mathrm{d}\Phi = k_{\mathrm{n}\lambda}\Phi \mathrm{d}s$$

By integration of the above equation and by altering to a logarithmic base we obtain

$$E = \log \frac{\Phi_0}{\Phi} = k_\lambda s \tag{5.80}$$

The extinction E_λ of the homogeneous medium as referred to light of wavelength λ varies proportionally with the logarithm of the flux Φ_0 falling on the medium and the flux Φ leaving it. E_λ also varies with the layer thickness s of the medium and the so-called absorption coefficient k_λ, whose value depends on the absorbing medium and the wavelength of the light. The above relationship is a limiting case, that is, it applies only to layer thicknesses ranging from zero to a certain s value, and only to a certain wavelength λ. Hence the light flux can only be quantitatively reduced within certain limits.

The value of the absorption coefficient of the ideal grey medium is independent of the wavelength of light, that is, k_λ is a constant. When it is also taken into account that the entering flux Φ_0 and the leaving flux Φ are proportional to the intensity I of the spectral line in question it follows that

$$E \equiv \log I_0/I = ks \tag{5.81}$$

In spectral analysis the grey wedge is adjusted so that the intensity I_r of the more luminous line r is reduced, so that the intensity of the ray leaving the wedge is equal to that of the intensity I_x of the less luminous line x. In this case the previous relationship is modified so that

$$\log I_\mathrm{x} - \log I_\mathrm{r} = Y_\mathrm{x} - Y_\mathrm{r} = \Delta Y_{\mathrm{x,r}} = ks_\mathrm{r} \tag{5.82}$$

where s_r is the layer thickeness at which the previously mentioned condition is

satisfied. If the grey wedge is regarded as geometrical, and when the absorbing medium is homogeneous, then the ratio of the width s and the length l of the wedge is constant. Hence

$$k' = \frac{s}{l} k$$

or, by taking Eq. (5.82) into account

$$\Delta Y_{x,r} = k' l_r \tag{5.83}$$

where k' is a constant the value of which depends on the extinction coefficient of the absorbing grey material and also on the slope of the latter, and l_r is the distance, measured from the edge of the wedge, at which the flux passing through it is equal to the flux of the non-reduced ray (Fig. 5.76). The constant k' is, therefore, the optical slope of the grey wedge, that is the value of the increase in extinction for every millimetre of the wedge.

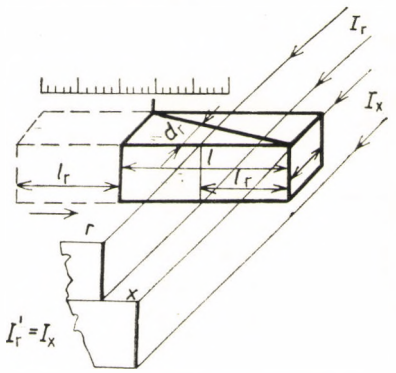


Fig. 5.76 Principles for the measurement of light attenuated by a grey wedge

It follows from the foregoing that when the grey wedge is used for the quantitative reduction of light it is an advantage if the logarithmic intensity ratio of the two spectral lines to be compared can be read directly from the linear scale. However, this requires that the optical slope of the grey wedge should be known, and that a pointer moving together with the grey wedge should indicate the movement of the latter.[2] A further advantage of the linear scale is that the measuring accuracy can be improved by using a vernier scale. The value shown is positive when $I_r < I_x$.

The grey-wedge light-reduction method can only be used in a limited fashion. This is due to the validity limits of the Bouguer–Lambert law, and also to the contrast sensitivity of the human eye. Adequate accuracy of measurement can only be obtained when the surface luminosity corresponding to the rays entering the eye is of medium value, that is when it ranges between 0·0007 and 0·14 sb. The best value also depends on the momentary visual capabilities of the eye, in addition to personal characteristics. For high intensity values, therefore, it is recommended first to reduce both light beams to be compared with either a multi-step-sector or a continuous attenuator, until the optimum level is achieved.

This initial reduction of the light rays is not necessary when the double grey wedge is used for visual measurement of the two spectral lines.[2] With this instrument each member of the line pair can be separately reduced in intensity (Fig. 5.77). In this procedure first the lower-intensity ray is reduced to an optimum level for

the measurement, then the other beam is reduced by means of the other wedge until the two beams appear to have the same intensity.

Let us assume that during the above operation the intensity of the spectral line R is reduced from the original value I_r, while the intensity of line X is also reduced from I_x to the same level I, provided the two beams travel the distances s_r and s_x, respectively, in two wedges of the same dimensions and absorption coefficient. By taking into account Eq. (5.83) we obtain the following

$$\log I_r/I = k'l_r \qquad \text{and} \qquad \log I_x/I = k'l_x$$

in other words

$$\Delta Y_{x,r} = k'(l_x - l_r) \tag{5.84}$$

The value $(l_x - l_r)$, that is, the extent of the shift of the two grey wedges, can be directly read from a linear scale, if the scale is fixed to one grey wedge, while the pointer is fastened to the other.

The use of the double grey wedge is advantageous not only because the measurement can always be done at the optimum intensity level, but also because the sign of the measurement can be directly read from the scale. In comparison with the simple grey wedge, however, it is a disadvantage that the double grey wedge requires more complex equipment.

The measurement accuracy is decreased when the light reduction is not entirely due to light absorption, but is also partly due to *reflection* from the surface of the grey wedge. This problem is eliminated if the parallel light beam falls perpendicularly on, and is reflected at right angles from, the surface of the grey wedge. This can be achieved by making the grey wedge of a parallel-faced sheet of glass composed of two parts: the grey wedge proper, and a 'completely' transparent glass wedge with the same refractive index as the grey wedge (Fig. 5.76).

The grey wedge can be made of grey glass and also from finely divided platinum or graphite embedded in gelatin. In the latter case the extinction of the grey wedge increases with the layer thickness of the gelatin. In another system a wedge-shaped layer of platinum is vacuum-coated onto parallel-faced glass sheets.[3] In the use of wedges made by this method care must be taken with regard to the Callier effect. The scatter of light on the platinum and graphite particles could be a major source of error when the geometric light path is altered, for example, when the scale of the grey wedge is calibrated under conditions different from the measuring conditions.

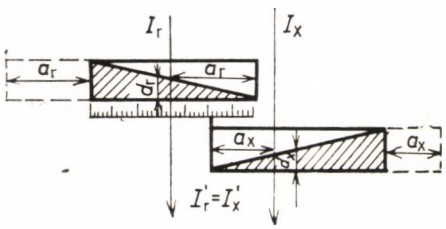


Fig. 5.77 Principles for the measurement of light attenuated by a double grey wedge

Generally speaking it is advisable to calibrate the scale of the grey wedge by using samples of known composition. The values read from the scale usually do not correspond to the actual values. This deviation could, however, be due to the personal characteristics of the observer, and therefore are not entirely due to the inhomogeneities of the grey wedge.

References

1. SCHELLER, H., *Jenaer Rundsch.*, **5**, 40 (1960)
2. TÖRÖK, T. and BARABÁS, J., *Acta chim. Hung.*, **19**, 52 (1959)
3. NAGY, Z. and SÁMSONI, Z., *Spectrochim. Acta*, **24B**, 147 (1969)

5.13.3 Measurement by polarisation techniques[1]

In the reduction of light flux by means of polarisation the linearly polarised light coming from the polariser passes through an analyser, that is, a rotatable polariser, whose angle of rotation ϕ can be measured. The light beam – except light reflected or absorbed – passes through the two polarisers without being reduced in intensity when the polarisation planes are parallel. Conversely, when the two polarisation planes are at right angles ('crossed') no light is transmitted; that is, there is complete extinction. In an intermediate position the extinction is given by the following formula

$$E = \log \Phi_0/\Phi = -\log \cos^2 \phi \tag{5.85}$$

In polarisation reduction of spectral-line intensities the beam from the slit of the spectroscope is divided so that the beams passing through the upper and lower parts of the slit each go through separate polarisers that are at right angles to one another. Thus the two spectra observable one above the other in the eyepiece of the spectroscope are produced by light rays that are perpendicularly polarised. When these rays pass through a common analyser whose plane of polarisation is parallel to the plane of one of the polarisers then one beam passes through the analyser without being reduced, while the other beam is completely extinguished. By rotating the analyser the lines in one spectrum can be reduced, while those belonging to the other spectrum will be increased in intensity. Hence, by appropriate rotation of the analyser the intensity of a spectral line x of one spectrum may be adjusted so that it is of equal intensity to that of a spectral line r of the other spectrum.

Let us assume that the rays of intensity I_x and I_r at the beginning become of equal intensity I when the angle of rotation of the analyser is ϕ. According to Eq. (5.85) this happens when $I_r = I_x.\cos^2 \phi$. Considering that the two light beams are perpendicularly polarised to each other, the relation $I = I_A \cos^2 (90° - \phi) = I_A \sin^2 \phi$ must also hold. It follows from the two equations that the two spectral lines to be compared will seem to have the same intensity when the intensity ratio of the two spectral lines is

$$\frac{I_x}{I_r} = \tan^2 \phi \tag{5.86}$$

The intensity ratio of the two spectral lines can therefore be read directly from the scale of the polarisation equipment, when it is graduated linearly in units of $\tan^2 \phi$.

By differentiation of Eq. (5.85) it can be calculated that in the simple reduction of light by polarisation effects the relative reading accuracy of the values read from the scale of the analyser is given by

$$\frac{dE}{E} = -0{\cdot}4343 \frac{\tan \phi}{\log \cos \phi} d\phi \tag{5.87}$$

that is, the relative accuracy varies as a function of the angle of rotation ϕ of the analyser as shown by curve (*a*) of Fig. 5.78. In the case of double polarisation reduction used in spectrochemical analysis this relationship refers to the reduction of one of the beams. Since the other beam is polarised perpendicularly the angular reading referring to it is given by $(90 - \phi)°$. The relative reading accuracy in this case varies as a function of the analyser rotation angle as shown by curve (*b*) in Fig. 5.78. Thus the reading accuracy achievable is given by the solid line in the figure. As may be seen the value of dE/E is at an optimum in double polarisation reduction when $\phi = 45°$, that is, when $\tan^2 \phi$ equals one. According to Eq. (5.86) this occurs when the original intensity of the two spectral lines to be compared is identical, that is, when $I_x = I_r$. As the difference in intensity of the two spectral lines to be compared increases the reading accuracy decreases. Therefore, it is *not* advisable to use the polarisation quantitative reduction method for the determination of the intensity ratio of two spectral lines of greatly different intensities.

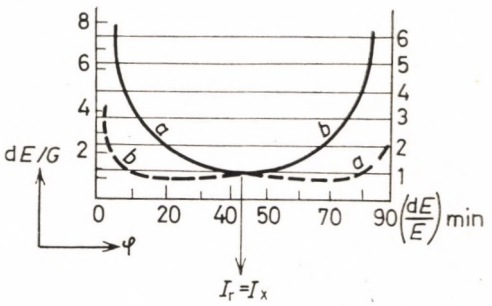


Fig. 5.78 Relative reading error of the attenuation method based on the polarisation of light

In the polarisation reduction method it should be remembered that during the measurements the intensity of both spectral lines is reduced, not only that of the stronger one. This loss of light is further increased by the fact that the polarisers absorb a large proportion of the light. This difficulty is, however, partly avoided in modern polarisation reducers by the fact that the polariser and analyser are polarisation filters instead of Nicol prisms. The filtering material consists of thin sheets containing herapatite crystals embedded in gelatin and subsequently magnetically orientated – chemically, these are quinine-sulphate-periodate crystals. This material has the great advantage that when the polariser and analyser are orientated parallel they absorb only 36% of white light, so that their transmittance is much greater than that of Nicol prisms.

Reference

1. *Fuess Katalog* 522-2 (1957), 522-3 (1958), Berlin/Steglitz

5.13.4 Auxiliary instruments used for visual intensity measurements

In the visual determination of the intensity ratio of spectral lines the maximum measurement accuracy can only be achieved when the two spectra to be compared have a sufficient length of contact. The accuracy of the comparison is reduced if stray light enters the observer's eye from the vicinity of the two spectral lines.

In practice the first requirement can be met in two ways. The two spectral lines can be brought by optical procedures either immediately one above the other, or

beside each other. The former method is relatively simple optically, but only a very narrow part (the width) of the spectral lines to be compared is brought into contact. The other method, which is doubtless better, has the disadvantage that a relatively complicated and expensive optical device is needed to bring the lines – which are some distance apart in the spectrum – close together in the eyepiece of the spectroscope.

The two spectral lines selected can be brought optically immediately one below the other only by the beam, passing through the slit of the spectroscope, being divided into two parts, a lower part and an upper part. Then two identical spectra appear in the eyepiece of the spectroscope. Their relative position is adjusted so that the two lines to be compared form a continuation of one another in the field of vision. To achieve this optical movement of the two spectral lines two procedures based on different principles are used. The two spectral lines can be brought optically one below the other by using a rhombic prism or by a sliding double slit.

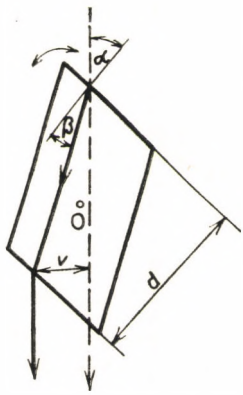


Fig. 5.79 Optical shift of spectral lines using a rhombic prism

The rhombic-prism spectrum displacement is based on the refraction of light. The rhombic prism is a glass body of rhombic form defined by parallel surfaces. The ray to be displaced parallel to its original direction enters the 'prism' through one of its parallel sides and leaves it in such a way that the planes of the incoming and outgoing beams remain parallel (Fig. 5.79). When α is the angle of incidence, β is the angle of refraction and d is the distance between the two parallel sides, then it can be calculated that the displacement, between the incoming and outgoing beams, caused by the a rhombic prism, is

$$v = d \frac{\sin(\alpha - \beta)}{\cos \beta} \tag{5.88}$$

When it is also considered that $\sin \alpha / \sin \beta = n$, the refractive index, then it follows from the above equation that the displacement caused by the rhombic prism increases as the distance between the parallel-faced sides and also the refractive index of the prism material increase. These two properties are constant in any individual prism. However, when using a rhombic prism, the parallel displacement of the light-beam can be increased by increasing the angle of incidence. The angle of incidence of the light can be varied by turning the rhombic prism, and hence the displacement of the spectrum can be measured.[1]

In using this method two rhombic prisms are usually mounted behind the slit of the spectroscope. One of these is fixed, while the other, placed either above or below the former, can be rotated about an axis parallel to the slit of the spectroscope (Fig. 5.80). By this arrangement the error due to light losses from reflections and absorption in the rhombic prism is eliminated. No gap occurs between the two spectra if the lower surface of the upper rhombic prism is situated slightly lower than the upper surface of the other prism. (This arrangement does not hinder rotation of the prism when the front surfaces of the two rhombic prisms are positioned in different planes.)

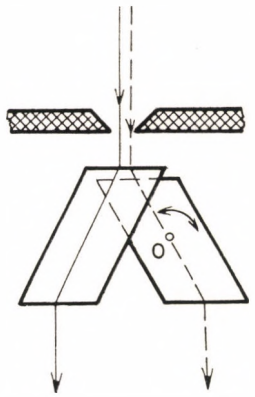


Fig. 5.80 Diagram showing the application of the rhombic prism

The sliding double-slit method is based on the principle that the two half-spectra produced in the eyepiece of the spectroscope are displaced with respect to one another when the light enters the instrument through two slits the central lines of which can be slid relative to one another. In practice, this is accomplished with two slits situated at different heights and of fixed widths. The central line of one of the slits is always located at the optical axis of the spectroscope, while the other can be moved. The two half-spectra are in close contact when the lower gap of one of the slits and the upper gap of the other are exactly in line (Fig. 5.81).[2]

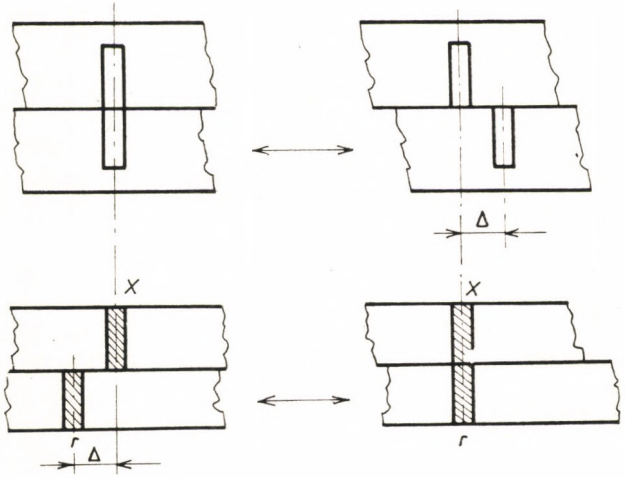


Fig. 5.81 Shift of spectral lines by means of the double slit method

The drawback of this procedure is that the intensities of identical lines of the two half-spectra are not identical even when the intensities are not reduced. The difference becomes greater as the central line of the sliding slit diverges further from the optical axis of the spectroscope. In practice this error is avoided by reading a 'correction' direct from the drum of the micrometer screw used for adjusting the slit. However, in theory this method is not entirely valid because measurement of the displacement and thus also the correction is affected by the wavelength of the light. This error can be neglected in comparison to other errors of visual intensity determinations. Also it is eliminated when the values read from the light reduction device are calibrated against standard samples.

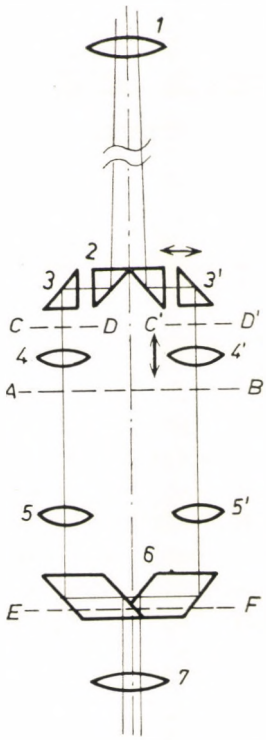


Fig. 5.82 Diagram for the side-by-side shift of spectral lines

In the sideways displacement of spectral lines it is the spectrum itself that is optically divided so that one analytical line is brought to one half-spectrum, while the other line is brought to the other. Thus the two spectral lines to be compared can be seen side by side when one of the spectra is optically displaced so as to be alongside the other. It follows that the spectra can only be displaced to appear side by side by means of a device located within the telescope of the spectroscope (Fig. 5.82).[3]

In this method the double deviating prism (*2*) divides the rays (that have been dispersed according to their wavelength by the spectroscope prism) before they are brought together by the telescope lens (*1*) into spectral lines in the focal plane *AB*. Thus the spectral lines are formed on the planes *CD* and *C'D'* when the divided rays are displaced parallel to one another by the deviating prisms (*3*) and (*3'*),

respectively. The rays coming from the true images formed on these focal planes are made parallel by the lenses (*4*) and (*4'*), respectively. Then lenses (*5*) and (*5'*) reunite the images so that the true image of the two spectral parts is formed in the focal plane *EF* side by side. The deviating prism (*6*) reflects the two light beams twice through ninety degrees. The two spectra can then be seen side by side through the eyepiece (*7*).

When using the equipment first, one of the spectral lines is brought into the field of vision of the eyepiece by appropriate rotation of the spectroscope prism. Then the distance between the deviating prism (*3'*) and the optical axis is altered until the required spectral line can be seen beside the first line. The images of the lines are adjusted to sharpness by means of lens (*4*).

Excess spectral lines can be easily removed from the field of view of the spectroscope by inserting an adjustable slit into the eyepiece. Both jaws, which must be parallel to the spectral lines, can be separately moved by adjusting screws.

To summarise, in order to visually determine the relative intensity of two spectral lines a variable-range spectroscope is needed by means of which the intensity of the two spectral lines can be quantitatively reduced. It is also necessary for the spectroscope to have a device by means of which the lines examined can be brought below one another or beside one another. Finally, it is necessary to avoid the effects of excess light.

References

1. *Hilger Catalogue*, F. I. London
2. TÖRÖK, T. and BARABÁS, J., *Acta Chim. Hung.*, **19**, 51 (1959)
3. *Spectrum Analysis Instruments*, Vsesojuznoe Exportno-Importnoje Objedinenije, Stankoimport CT-7, Moscow

Author Index

ABBÉ, E. *241*, 242, 245
ABNEY, W. W. *436*, 438
ABRAHAMSON, I. S. *122*, *129*, 131
ADAMS, K. B. *201*, 202
ADDINK, N. W. H. *464*, 465, *466*, 467
ÁG, A. 41, 42
AHRENS, L. H. *70*, 71
ALLEN, J. S. *480*, 481
ANDERSON, J. A. 370
ANGERER, E. *192*, 195, *269*, 271, *288*, 290, *313*, 318, *409*, 411, *415*, *416*, *431*, 432, *433*, 434, *437*, 438, *443*, 445, 446, *447*, *448*, 449
ARCHBOLD, E. *216*, 219
ARRAK, A. *412*, 416
ASHLEY, S. E. *175*, 176
ASHTON, V. H. *116*, *118*, 119
ASZTALOS, I. *461*, 462
ATSUYA, I. *114*, 119

BACHER, J. 449
BALY, E. C. *255*, 256
BALZ, G. 112
BARABÁS, J. *501*, *502*, 504, *507*, 509
BARDÓCZ, A. 103, *104*, 105, *113*, 114, *122*, *127*, *128*, 131, *133*, 134, *162*, 164, 177, 178, *181*, 182, 183, *308*, *310*, *311*, 312
BARKHAUSEN, H. *136*, 138
BARROW, G. M. *42*, 43
BAUMANS, P. W. I. M. *53*, 57
BAUMGARTNER, W. *475*, 476, 489
BAUSCH, C. L. *337*, 340, *374*, 376
BECKER, O. 67
BELYAKOV-BODIK, I. B. *107*, 108
BENKŐ, K. *313*, 318
BERNDT, M. *217*, 220
BERRY, C. T. *114*, 119
BERRY, T. M. *177*, 178
BEUTLER, H. 372
BIEDL, J. *448*, 449
BILTZ, M. *417*, 419
BINKMANN, H. *71*, 73
BIRKS, F. T. *196*, *197*, *198*, 201
BLASERNA, J. *271*, 273
BLITZER, L. *133*, 134
BOCHKOVA, O. P. *187*, 192
BOER, I. H. *479*, 481
BOETTNER, E. A. *361*, 362, *495*, 496
BOHR, N. *22*, 29, 44, *78*, *81*, 82
BOUMANS, P. W. J. M. *57*, 64, *67*, 69
BOWEN, I. S. *363*, 364
BOZMANN, W. R. *37*, 38, *72*, 73
BRAND, J. C. D. *42*, 43
BRANDO, C. J. *133*, 134
BRAUNISCH, H. *131*, 134
BRECH, F. X. *210*, *211*, 215, *217*, 220
BRECKPOT, R. 499
BREENE, R. G. *78*, *80*, 82
BREWINGTON, G. P. *361*, 362, *495*, 496
BREWSTER, D. *241*, *244*, 245
BRINKMAN, H. *55*, 57
BROCA, K. 241, 245
BRUINING, H. *489*, 490
BRYAN, F. R. *175*, 176, *216*, 219, *361*, 362, *495*, 496
BUNGE, K. *75*, 78
BURGER, H. C. *39*, 40, *316*, 318
BURHORN, F. 57
BURNS, K. *201*, 202

Italic page numbers indicate a citation by reference number only.

Burton, J. A. *480*, 481
Butterworth, I. S. *490*, 491

Cady, W. M. *133*, 134
Caldecourt, V. J. *495*, 496
Callier, A. 441, 442
Cameron, W. H. B. *271*, 273
Cammoy, C. G. *481*, 483
Candler, C. *454*, 455
Cannon, C. G. *490*, 491, *493*, 495
Chorovin, Yu. I. *198*, 201
Churchill, J. R. *424*, 426
Clark, G. L. *338*, *339*, 340, *343*, 344, 346, *356*, 358, 370, 380
Clark, W. R. *495*, 496
Clippeleir, K. 499
Cobine, J. D. *119*, 122
Cohen, A. E. *270*, 271
Compton, A. H. *364*, 367
Corlich, C. H. *37*, 38, *72*, 73, *77*, 78
Coryell, C. D. 388
Craggs, J. D. *133*, 134, *187*, 192, 204
Cristescu, G. D. *119*, 122
Crosswhite, H. M. *361*, 362, 422, *495*, 496
Cvonat, J. 446
Czerny, H. 307
Czerny, M. 343, 344

D'Ans, J. *240*, 245, *271*, 273, *336*, *337*, 340, *401*, 404, *479*, *480*, 481
De Boer, F. *308*, 312
Debol, C. H. *278*, 279
Debras-Guedon, J. *216*, 219
Declava, J. *286*, 287
Dehio, H. 451
Deinum, H. E. *308*, 312
Dieke, G. H. *83*, 85, *361*, 362, *421*, 422, *495*, 496
Dieke, J. J. *351*, 353
Dietert, H. W. *180*, 183, *361*, 362, *495*, 496
Dikhof, J. A. M. *71*, 73, *464*, 465
Dobeneck, V. D. *72*, 73, *77*, 78
Dorgelo, H. B. *39*, 40
Driffield, V. C. 416, *417*, 419
Dropplemann, S. L. *343*, 344
Duclaux, J. *241*, *244*, 245
Duffendock, O. S. *108*, 111, *361*, 362, *495*, 496

Eagle, A. *356*, 358
Eberhardt, G. *106*, 108, *452*, 453
Ebert, H. *344*, 346
Eckard, S. *342*, 343
Eckart, F. 489, 490
Eder, J. M. *433*, 435
Edgerton, H. E. *202*, 204
Edlen, B. *366*, 367
Eggert, J. *418*, 419, 442, 443
Eichhoff, H. J. *199*, *200*, 201
Einhorn, H. D. *270*, 271
Ekefers, E. *363*, 364
Elliot, A. *271*, 273
Elster, J. *473*, 476
Emde, H. *217*, 220
Engelst, V. S. *60*, 63
Engstrom, R. W. *483*, 488, 489, *491*, 492

Fassel, V. A. *107*, 108
Faster, W. Q. *348*, 350
Fastie, W. G. *343*, 344, *345*, 346
Fayolle, F. *202*, 204
Felske, A. *214*, 215, *218*, *219*, 220
Fenner, F. *204*, *206*, 207
Fery, O. K. 290, 291, 292
Feussner, O. *169*, 170, 171, 173
Finkelnburg, W. *22*, 29, *37*, 38, *51*, 53, *54*, 57 *187*, 192
Finkelstein, N. H. *374*, 376
Flatow, E. 245
Foner, S. N. *121*, 122
Forrest, J. V. *240*, 245
Forsythe, W. E. *241*, 245
Fowler, R. C. *316*, 318
Fraenkel, A. *136*, 138
Fred, H. 372
Frie, W. *61*, 63
Frieser, H. *412*, 416, *442*, 443
Frind, G. *122*, 131, *133*, 134
Frisque, J. *40*, 42
Frodl, V. *204*, *206*, 207

Gatterer, A. *204*, *206*, 207, *387*, 388
Gebauhr, W. *115*, *116*, *118*, 119
Geffcken, W. *383*, 384
Gegetschkori, N. M. *122*, *129*, 131
Geitel, H. *473*, 476
Georgijevic, V. *57*, 63, 64
Gerbatsch, R. 449
Gerlach, W. *111*, 112
Germeshausen, J. K. *202*, 204
Gettys, B. *175*, 176
Ghersen, P. *308*, *309*, 312
Gifford, J. W. *241*, *244*, 245
Glaser, G. *122*, *123*, *128*, *129*, 131, *481*, 483
Glennon, B. M. *37*, 38
Goebert, M. *313*, *314*, 318
Golling, E. *86*, 88
Gollnow, H. *197*, *198*, *200*, 201
Görlich, P. *208*, 210, *473*, 476, 481, 483, *484*, 487
Goto, H. *114*, 119, *121*, 122
Graggs, I. D. *122*, *128*, 131
Graue, G. *342*, 343
Greenfield, S. *114*, 119
Griem, H. E. *53*, 57, 63, *122*, 131
Grier, H. E. *202*, 204
Grimsehl, S. *46*, 48
Grosswhite, H. M. *83*, 85
Grotrian, W. 20, 22, *38*, 40

HAAG, J. B. *364*, 367
HAGENAH, W. D. *60*, 63, *214*, 215, *216*, *217*, *218*, 219, 220
HAHSTOLL, G. A. *495*, 496
HAMMER, A. *258*, 261
HANSEN, G. 301, *314*, 317, 318, *407*, 408, *417*, 419
HAPP, H. *208*, 210
HARMS, F. *37*, 38, *92*, *93*, *94*, 95, *192*, 195, *251*, *252*, 253, *269*, 271, *326*, 332, *353*, 355
HARPER, D. W. *216*, 219
HARRISON, G. R. *83*, 85, *337*, 340, *353*, 355, *374*, 376, *386*, 388
HARTMANN, J. *392*, 393
HASLER, M. F. *180*, 183, *361*, 362, *495*, 496
HAGEMANN, F. *71*, *72*, 73
HAULE, W. *278*, 279
HELLWARTH, R. W. *213*, *214*, 215
HELTAI, GY. 451
HERMANN, W. *490*, 491
HERZBERG, G. *42*, 43
HILLER, M. *491*, 492
HIRSCH, S. *317*, 318
HODGE, E. S. *433*, 435
HÖFERT, H. J. *162*, 164, *170*, 173
HOHLY, H. W. *238*, *240*, 245
HOHNERJAGER-SOHM, H. *456*, 457, *463*, 464
HOLTSMARK, J. *78*, *81*, 82
HONIG, R. E. *216*, 219
HÖRMANN, H. *59*, 63, 134
HOYT, G. D. *202*, 204
HUDSON, R. L. *121*, 122
HUGHES, T. P. *216*, 219
HULDT, L. *122*, 131, 134
HUNDT, F. *318*, 326, 332, 334, 336
HURTER, F. 416, *417*, 419
HURWITZ, J. K. *175*, 176
HUTCHISON, R. O. 394
HUXFORD, W. S. 204

IKONOMOV, K. *57*, 63

JACQUINOT, P. *379*, 380
JAENICKE, H. *443*, *444*, 445
JAMES, A. *318*, 326
JAMES, T. H. *409*, 411, *418*, 419
JARREL, R. F. *344*, 346, *348*, 350
JEANTOT, P. *286*, 287
JECHT, U. *119*, *120*, 122
JEMAL, E. J. *168*, 169
JOHNSON, F. S. 481
JOHNSON, R. C. *14*, 15, *42*, 43
JONES, I. L. *114*, 119
JONES, L. A. *417*, 419
JOOS, G. *136*, 138, *187*, 192, 195, *224*, 227, *269*, 271, *409*, 411, *415*, 416, *431*, 432, *433*, 434, *437*, 438, *443*, 445, 446, *447*, *448*, 449
JUNKES, J. *387*, 388

KAISER, H. 112, *138*, 143, 145, *165*, *166*, 168, 169, *172*, 173, *301*, 305, *314*, 318, *412*, *413*, 416, *455*, 456, 457, *458*, 459, 463, 464
KATZ, E. *71*, 73
KAYSER, H. *83*, 85, *227*, 233, *238*, 245, *251*, *252*, 253, 262, *263*, 264, *265*, 267, *268*, 271, *386*, 388
KECK, P. H. 112, *162*, 164, *170*, 173, *468*, 472
KELLER, R. E. *206*, 207
KEMP, J. W. *180*, 183
KEREKESNÉ, CSÉTI, S. 40, 41, 42
KESSLER, W. *72*, 73, *77*, 78, *119*, *120*, 122
KICHINOSUKE, H. *121*, 122
KICKERING, E C. 267
KIEMLE, H. *313*, 318
KIND, E. G. *466*, 467
KINDER, W. 301, *317*, 318
KIRCHGESSER, W. G. *374*, 376
KITCHEN, H. *177*, 178
KLEIN, E. *412*, 416
KLUGE, W. *480*, 481
KOCH, A. *438*, 440
KOCH, O. 57
KOFFMANN, M. L. *316*, 318
KOHLRAUSCH, F. *42*, 43, *46*, 48, *78*, *79*, *80*, *81*, 82, *88*, 92, *187*, *191*, 192, *205*, 207, *224*, 227, *263*, 264, *373*, 376, 380, *401*, 404, *473*, 476
KOLLER, L. R. *480*, 481
KOLOSNIKOV, V. G. *78*, *80*, 82
KORFF, V. *78*, 82
KOROLYEV, V. V. *114*, 118, 119
KORTÜM, G. *315*, *316*, 318, *380*, 382, 383, 384, *470*, 472, *473*, 476, 481, 483, *490*, 491, *492*, 495, *500*, 501
KOVÁCS, F. 105
KOVÁCS, J. *42*, 43
KRAMERS, H. A. *78*, 82
KRANZ, E. *115*, 117, *118*, 119
KRAUSE, H. *217*, 220
KREMPL, H. *75*, *76*, *77*, 78, *122*, *127*, *129*, *130*, 131, *132*, *133*, 134, *167*, 168
KRETZMANN, R. *493*, 495
KRON, E. *437*, 438
KRÜSS, J. *271*, 273
KÜSTER, A. 442, 443

LAMPARTE, M. J. *203*, 204
LANG, K. *438*, *439*, 440
LANGSTITH, G. O. *313*, 318
LAGUA, K. *214*, 215, *216*, *217*, *218*, 219, 220
LARENZ, R. W. *131*, *132*, 134
LARUE, J. M. *108*, 111
LAU, E. *466*, 467
LAX, E. *240*, 245, *271*, 273, *336*, *337*, 340, *401*, 404, *479*, *480*, 481
LEHMAN, R. *208*, 210
LENGYEL, B. A. *208*, 210
LENSING, J. *117*, 119
LEO, W. *343*, 344

LERNER, R. *116*, *117*, *118*, 119
LEUTWEIN, F. 179
LEVY, S. *168*, 169
LIEMPT, J. A. M., VAN *417*, 419
LIMMER, G. 388
LINDHOLM, E. *78*, *80*, 82
LINHART, J. G. *53*, 57
LINKE, H. *442*, 443
LINSTRÖM, C. F. *40*, 42, 388
LIPIS, L. V. *198*, 201
LISKA, M. *423*, 424
LOCHTE-HOLTGREVEN, W. *76*, 78, *80*, *81*, 82, *86*, 88, *133*, 134
LOEB, L. B. *187*, 192
LOMAKIN, B. A. *40*, 42
LOOFBOUROW, R. *353*, 355
LORD, C. *353*, 355
LORENTZ, H. A. *78*, *81*, 82
LOW, F. *78*, 82
LUNDEGARDH, H. *433*, 434
LÜDKE, W. *208*, 210
LYMAN, TH. 370

MABKER, H. *114*, 118
MCCLUNG, F. J. *213*, *214*, 215
MCCORMIG, W. W. *202*, 204
MACHER, F. *185*, *186*, 187
MCRAE, D. R. *313*, 318
MAECKER, H. *37*, 38, *51*, 53, *54*, 57, *58*, *61*, *62*, 63, *65*, *66*, 67, *86*, 88, *122*, *130*, 131, *133*, 134, *187*, 192
MAGNUSSON, P. *362*, 364
MAIER, H. *490*, 491, *492*, 495
MAIMAN, F. H. *210*, 215
MALIK, V. D. 63
MALPICA, J. T. M. *40*, 42, *177*, 178, *438*, 440
MALTER, L. *483*, 488
MANDELSTAM, S. L. *78*, *80*, 82, *107*, 108, *122*, *123*, *124*, *128*, *129*, 131
MANNKOPFF, R. *55*, 57, 63, *107*, 108, *351*, 353
MARGENAU, H. *78*, 82
MARGOSHES, M. *114*, *117*, 118, *218*, 220
MARK, J. E. *366*, 367
MAROTZ, A. *342*, 343
MARTIN, J. *175*, 176
MARTON, T. R. *271*, *272*, 273
MASON, R. C. *73*, 75
MASSMANN, H. *198*, 202
MASTRIL, F. *62*, 63
MATHIEN, F. C. 292
MAZING, M. A. *78*, *80*, 82
MEEK, J. M. *122*, *128*, 131, *187*, 192, 204
MEES, C. E. *409*, 411, *431*, 432
MEISSNER, K. W. *379*, 380
MERTON, T. R. 339, 340
MIELENZ, K. D. *303*, *304*, 305
MIKA, J. 176, *181*, *182*, 183, *185*, *186*, 187
MILLIKAN, R. A. *363*, 364, 367
MINCK, R. W. *216*, 219
MINKOVSKI, R. *78*, 82
MITCHELL, K. B. *198*, 201
MITSUYASU, S. *121*, 122
MOENKE, H. *208*, 210, 215, *217*, *218*, 220
MOENKE-BLANKENBURG, L. *208*, 210, 215, *217*, *218*, 220
MOHAROS, I. 451
MOHLER, H. *47*, 48, *401*, 404
MOLL, W. J. H. *316*, 318
MOLLWO, L. *119*, 122
MORITZ, H. 451
MORRIS, M. J. *361*, 362
MORRIS, S. H. *495*, 496
MORTON, G. A. *483*, 488
MOSETTI, V. *216*, *219*
MURRAY, W. M. *175*, 176

NACHTRIEB, N. H. *227*, 233, *257*, 261, *316*, 318, 326, 332, 334, 336, *353*, 355, 356, 361, *438*, *439*, 440, 441, *447*, 449
NAGADA, H. *66*, 67
NAGY, Z. *313*, *314*, 318, *428*, 430, *503*, 504
NAHSTOLL, G. A. *175*, 176, *361*, 362
NAOLIN, P. *202*, 204
NEEB, K. H. *115*, *116*, *118*, 119
NEUHÄUSSER, A. *316*, 318
NIETZ, A. 446
NORDMEYER, M. *303*, *304*, 305, *306*, 307

O'BRIEN, B. *438*, 440
OLDFIELD, A. *177*, 178
OLSEN, H. N. 204
ORNSTEIN, L. S. *39*, 40, *55*, 57, *71*, 73, 97, *316*, 318
OWEN, L. E. *115*, 118

PAKSY, L. 176
PARKS, V. L. *438*, 440
PASCHEN, F. 188, *195*, 201, *334*, *335*, 336, *355*, 356
PAULI, N. 33, 34
PAVLOVIC, N. *57*, 63
PELLIN, P. *241*, 245
PENNER, S. S. *42*, 43
PEPPERS, N. A. *218*, 220
PERRIN, F. H. *240*, 245
PÉTER, L. 176, *181*, *182*, 183
PETERS, C. L. 63, *107*, 108
PETERS, R. *52*, 53
PETERS, T. *86*, 88
PETERSEN, E. W. *495*, 496
PFEILSTICKER, K. 101, 105, 108, 112, 183, *186*, 187
PFENDER, E. *53*, *54*, 57
PHUND, A. H. *363*, 364
PICKERING, E. C. *240*, 245
PIERCE, W. C. *438*, *439*, 440, 441
PITTWELL, R. R. *41*, 42
PLOKE, M. *481*, 483

PLSKO, E. 423, 424
POHL, R. W. *245*, 251
POLANSKY, S. *78*, *80*, *81*, 82
PRESTON, J. S. *479*, 481, *491*, 492
PREUSS, E. *304*, 305, *419*, 420
PRICE, W. C. *289*, 290
PROKOFIEV, V. K. *386*, 388
PULFRICH, C. *241*, *244*, 245

RAIKHBAUM, YA. D. 63
RAISKII, S. M. *386*, 388
RAMBERG, E. G. *473*, 476
RASBERRY, S. D. *218*, 220
RAYCHMAN, J. A. *483*, 488
RAYLEIGH, R. 254, 256
RÉTI, S. *459*, *460*, *461*, 462
RICHARDSON, H. M. *316*, 318
RICHARDSON, O. A. *475*, 476
RICHTER, G. *75*, 78
RITSCHL, R. *83*, 85, *386*, 388
ROBINSON, N. W. *202*, 204
ROCKENBAUER, W. *206*, 207
ROES, R. L. *86*, 88
ROHLOFF, E. *65*, 67
ROLLWAGEN, W. *71*, 73, *111*, 112, *468*, 472
ROMPE, R. *51*, 53, *54*, 57, *66*, 67, *76*, 78, *88*, 92, *122*, *123*, *127*, 131
ROSE, G. *466*, 467
ROSEN, B. *198*, 201
ROSENTAL, J. E. *363*, 364
ROTHE, H. *208*, 210
ROWLAND, H. A. *326*, 332, 333
RUMSEY, N. J. O. 287
RUNGE, C. *351*, 353, *355*, 356, 370
RUNGE, E. F. *216*, 219
RUTHARDT, K. *111*, 112, *257*, 261

SAHA, M. N. *70*, 71
SALPETER, E. *206*, 207
SÁMSON, K. *318*, 326
SÁMSONI, Z. *314*, 318, *428*, 430, *503*, 504
SAUNDERSON, J. L. *495*, 496
SAWYER, R. A. *108*, 111, *353*, 355, *356*, 358, *360*, 361, *363*, 364, 367, *415*, 416
SAYCE, L. A. *339*, 340
SCHAETTI, N. *475*, 476, *481*, *482*, 483, 489
SCHALLREUTER, W. *88*, 92, *224*, 227, *264*, 267, *318*, 326
SCHEIBE, G. *40*, 42, *153*, *157*, 159, *167*, 168, *175*, 176, *316*, 318, 388, *432*, 434
SCHEINER, J. *417*, 419
SCHELLER, H. *501*, 504
SCHMICK, H. *65*, 67
SCHMIDT, H. *466*, 467
SCHMIDT, W. *119*, 122
SCHMITZ, G. *59*, 63, 134
SCHNEIDER, E. *287*, 290
SCHNETTLER, O. *40*, 42
SCHOEN, A. L. *433*, 435
SCHOLZ, O. *119*, 122
SCHÖN, M. *66*, 67, *317*, 318
SCHÖNFELD, H. *136*, 138
SCHÖNTAG, A. *75*, 78, *167*, 168, *271*, *272*, 273
SCHÖNTAG, B. *450*, 451
SCHROLL, E. *206*, 207
SCHUCK, K. *217*, 220
SCHUHKNECHT, W. *347*, *349*, 350
SCHULZ, P. *78*, *81*, 82
SCHÜLER, F. *195*, *196*, *197*, *198*, *200*, 201
SCHWARTZSCHILD, K. 436, *437*, 438
SCRIBNER, B. F. *114*, *117*, 118, *218*, 220
SEGIURA, Y. *66*, 67
SEIDEL, W. *455*, 456
SEITH, W. *257*, 261
SELINGER, R. 67
SEMENOVA, P. O. *73*, 75
SEMMERMEYER, K. 67
SERIN, P. A. *116*, *118*, 119
SHAFER, A. B. *343*, 344
SHIRLEY, H. T. *177*, 178
SIEDENTOPP, H. *313*, 318
SIEGBAHN, M. *362*, 364
SIMON, H. *473*, 476, 481, 483
SIRKS, J. L. *352*, 353
SIROIS, E. H. *114*, 119
SLATER, J. *78*, 82
SMIT, J. A. *86*, 88
SMITH, I. A. *71*, 73
SMITH, L. *206*, 207
SMITH, M. W. *37*, 38
SMITH, W. *473*, 476
SOBELMAN, I. I. *78*, *80*, 82
SOBELYEV, I. I. *78*, *80*, 82
SOHM, M. 112
SOMMER, A. *473*, 476, 481, 483, 487
SOMMERFELD, A. *37*, 38
SORD, M. A. 63
SPATH, H. *132*, 134, *167*, 168
SPEAKMAN, J. C. *42*, 43
SPEKTOV, V. A. *60*, 63
SPERLING, J. 134
STARK, I. 27, *78*, 80, 82, *437*, 438
STEEN, I. M. *495*, 496
STEENBECK, M. *51*, 53
STEHN, J. *366*, 367
STEPHEN, R. C. *240*, 245
STERLING, J. Z. *60*, 63
STETTER, A. *131*, *133*, 134
STOICHEFF, B. F. *217*, 220
STRAW, H. *17*, 20, *22*, 29, *37*, 38, *42*, 43, *78*, *79*, *80*, 82
STREIDER, YE. YA. *187*, 192
STROCK, L. W. *447*, 449
STRONG, J. *245*, 251, *326*, 332, 334, 336
SUHRMANN, R. *473*, 476, 481, 483
SYBIL, C. *71*, *72*, 73
SZABÓ, Z. L. *107*, 108
SZAKÁCS, O. *107*, 108

TAPPE, W. *121*, 122
TAYLOR, S. R. *70*, 71
TEISSING, H. *313*, *314*, 318
TEVES, M. C. *479*, 481
THIERS, R. E. *87*, 88, *352*, 353
THOMSON, G. P. *187*, 192
THOMSON, J. J. *187*, 192
THOMSON, K. B. *108*, 111
THOMSON, W. B. *53*, 57
THOURET, W. *54*, 57
TOLANSKY, S. *197*, 201, *205*, 207
TÖRÖK, T. *102*, 105, *107*, 108, *138*, *141*, 145, *147*, 148, *149*, *150*, *151*, 152, *153*, 159, 160, *161*, 164, *170*, *172*, *173*, 184, *185*, *186*, 187, *412*, *414*, 416, *427*, 430, 446, *448*, 449, *450*, 451, 459, *460*, *461*, 462, *501*, *502*, 504, *507*, 509
TOUSEY, R. 481
TRUMPY, B. *78*, *81*, 82
TURK, W. E. *483*, 487
TURNER, A. F. 307, 343, 344
TWYMAN, F. *304*, 305, *306*, 307

VAINSTEIN, E. E. *78*, *80*, 82, *114*, *118*, 119
VALLE, B. L. *87*, 88
VAN CALKER, J. *76*, 78, 102, 105, *111*, 112, *117*, 119, *121*, 122, *131*, 134, 183, 184, *185*, *186*, 187, *308*, 312, *427*, *428*, 430
VARSÁNYI, F. *104*, 105, *308*, *311*, 312
VINCENT, H. B. *108*, 111
VORSATZ, B. *185*, *186*, 187, *311*, 312
VUKANOVIC, D. *57*, 63, 64

WADSWORTH, F. L. O. *238*, 245, *359*, 361
WALKER, S. *17*, 20, *22*, 29, *37*, 38, *42*, 43, *78*, *79*, *80*, 82
WALLRAFF, A. *138*, *143*, 145
WATANABE, K. 481
WATSON, W. W. *78*, 82
WEBB, J. M. *438*, *439*, 440
WEBER, A. *438*, 440
WEELSTON, J. R. *216*, 219
WEINLAND, C. E. *438*, 440
WEISS, R. *51*, 53
WEISSKOPF, V. F. *78*, *80*, 82
WEIZEL, W. *51*, 53, 54, 57, *66*, 67, *76*, 78, *88*, 92, *122*, *123*, *127*, *129*, 131
WHITE, H. E. *17*, 20, *22*, 29
WIEN, W. *37*, 38, *50*, 51, *92*, *93*, *94*, 95, *192*, 195, *251*, *252*, 253, *269*, 271, *326*, 332, *353*, 355
WIENECKE, R. 57
WIESE, W. *37*, 38, *62*, 63
WIGNER, E. P. *78*, 82
WILLIAMS, A. J. *495*, 496
WILLIAMS, G. O. *133*, 134
WITTKE, H. *197*, 201
WOHNBANK, F. *75*, 78
WOLFE, R. A. *168*, 169

ZAIDEL, A. N. *386*, 388
ZEHDEN, W. *168*, 169
ZENTAI, P. *40*, 42
ZIMMER, K. *448*, 449, 459, *460*, *461*, 462
ZWORYKIN, V. K. *473*, 476, *483*, 488

Subject Index

Abbe comparator 398
Abbe number 226
Abbe prism 242
Abney spectrograph 355
Absolute line intensity 14, 37
Absolute refractive index 224
Absorption coefficient 262
Absorption filter 382
Achromatic intermediate focusing 303
Achromatic lens 245
'Activated state', laser 209
After-effect, spark discharge 126
Afterglow, spark discharge 126
Ageing process, photoelectric 476, 479
Alkali metal photocathode 479, 480
Alkali metals, work function 479
Alternating current, effect 95
Aluminium, term scheme 36
Ambipolar diffusion
  arc plasma 59
  spark plasma 124
Amici prism 243
Amplification factor, photomultiplier 482
Ångström unit 385
Angular dispersion 221
  concave grating 330
  Fabry–Perot interferometer 378
  plane grating 324
  prism 230
  Wadsworth-system spectrograph 359
Anode current, photomultiplier 426, 484
Anode process, arc 52, 64
Anomalous cathode fall, discharge tube 191
Anomalous dispersion 226
Antimony–caesium photocathode 480
Apostilb 404
Arc, 51–97
  anode drop 52
  aureola 52
  breakdown voltage 92
  burning spot 52
  cathode drop 52
  current and voltages 88
  effect of a.c . 95
  electrical parameters 88–97
  electrode-stabilised 52
  free-burning 52
  fundamental properties 51
  ignition 97, 99–105
    electronic 103
    mechanical 99, 102
    by Tesla current 99
  self-ignited 94
Arc generators 97–114
  classification 97
  high-voltage
    a.c. 108
    rectified 109
  interrupted
    electronically controlled 113
    mechanically controlled 111
  limiting resistance 90
  power supply 98
  principal circuitry 97
  series resistance, *see* Limiting resistance
Arc plasma 51, 53–57
  ambipolar diffusion 59
  burning spot 65
  cathode drop 54, 66, 68
  convection of particles 60

Arc plasma — *continued*
diffusion of particles 59
electrical conductance 58
electrons 55
free path value 58
temperature 56
electronic excitation 67
electronic trajectories 56
filling gas 61
gas temperature 56
heat conductivity 55
light emission 67
continuous spectrum 69
migration of particles 57, 59
particles 53
radiation intensity 68
residence of elements 63
secondary collision 67
separation of particles 61
'spontaneous emission' 68
superelastic collision 67
temperature 71, 73
thermal equilibrium, local 55
thermionic emission 67
Artificial phase, synchronous motor 171
ASA sensitivity rating system 418
Astigmatic concave-grating spectrographs 353
Astigmatic difference 352
Astigmatic error 247
Astigmatic image dispersion instruments 222
Astigmatic line imaging, Fery spectrograph 291
Astigmatism
of concave grating 350
of prism spectrographs 236
Aston dark field 190
Asymmetrical controlling, of spark gap 185
Atom line 15
density variation with plasma height 118
intensity variation with plasma temperature 132
notation 34
Atom pressure 70
Atomic spectrum 15
origin 15–22
Attenuation, of illumination of prism spectrographs 312
Aureola 52
Autocollimating prism
Fery type 291
Littrow type 232
Autocollimation, light dispersion 232
Autocollimation camera 285
'Average lifetime' 16
Ayrton equation 89
Azimuthal quantum number 25
group 29

Background 46, 85, 440
correction 440
by calculation 463
by s. p. d. scale 464
radiation 46
Balmer series 18
Band head 42
origin 45
Band-pass filters 382
Band series 42
Band spectrum 42, 85
origin 43
Binocular spectroscope 295
Black body radiation 46
Blaze angle 339
echelle-grating 338
plane-grating 345
Bouguer–Lambert law 501
Brackett series 18
Breakdown voltage, arc 92
Brewster rule 264
Brush development, photographic emulsion 451
Buffer effect 77
Bunsen–Roscoe reciprocity law 438
Burning
arc 52, 65, 66, 96
spark 166

Calbier prism 242
Calker transfer, of ignition current 102
Callier effect 441, 469
Callier quotient 442
Camera lens, prism spectrographs 251
focal length 252
free diameter 212
Camera mirror 252, 341
Candela 404
Carrier gas
Geissler tube 194
hollow cathode 196
plasma jet 120
Cathode fall, discharge tube 190
Cathode layer radiation, arc 107
Cathode potential drop, arc 52, 66
Cathode space enrichment, arc 61
Cauchy relationship 255
Characteristic curve, photographic 437, 438, 445, 447, 449
construction 419–430
by Hansen-type diaphragm 424
by multiplet lines 420, 424
by multi-step attenuator 423
correct-exposure part, 414
fogging part 412, 416, 448
γ (gamma) value 414, 430, 447, 448
over-exposed part 414
solarisation part 414, 416
under-exposed part 414
Chromatic aberration 245, 248
Coherent radiation 208

Collimator lens 245, 248
aperture 248
diameter 250
faults 245
focal distance 236
Collimator mirror 249
Collisions of particles, plasma 55, 67
Colour sensitivity curve, of human eye 500
Coma 247, 248
CO molecular spectrum, term scheme 45
Composite photocathodes 480
Concave-grating dispersion 326–332
Concave-grating spectrographs 350–367
Condenser
accumulated energy, 137, 149
capacitance 150
charging circuit, 152–164
damping resistance 160
equalising voltage 154
integrating 496
remanent inductance 137
stationary voltage 154
virtual resistance 101
Condenser discharge 136–138
after-effect 126
aperiodically damped 138
circuit 135–152
complete 125, 147, 151
critically damped 138, 146
damping time 151
oscillation time 150
over-damped 145
periodically damped 138
period number 143, 145, 151
phase shift 139, 155, 159
rectified 174
time interval 150
over-damped 146
periodically damped 143
pure spark damping 144
Constant-deviation prism 241
Continuous spectra 46
gases 47, 289
intensity, arc plasma 69
Contrast sensitivity, of human eye 499, 502
Controlling spark gap 168
Convection, arc plasma 60
Cornu prism 240, 241
Corona discharge 122
Cosine law 402
Critical damping 138, 146
Critical development time 448
Crookes dark field 190
Current–conductivity curve, arc generators 88
Current–voltage curve
arc generators 88
short-circuit 96
Czerny–Turner spectrographs 341, 342, 343
Cyanogen band 42, 86
Damping resistance, spark excitation 160
Dark current
photocell 475
photomultiplier 488
Dark field, discharge tube 190
D-blackening value 412
d.c. continuous arc generators 105–108
d.c. plasma jet 114–118
Deceleration radiation 204
Decadic counter method, photomultiplier 497
Density, photographic emulsion 412
Density transformation 454–462
generally valid, partial 458
l, 459
$P_{\frac{1}{2}}$ 456
Seidel-type 455
W 455
Deslandres branch formula 43
Developers 444, 445, 446, 451
agitation 449
composition 445
surface 445
Development, of photographic emulsion 409, 443, 448
brush 451
Deviation, angle
grating 322
prism 228, 230
Diaphragm, light attenuation 316
Dielectric hysteresis 206
Dielectrically coated surfaces, laser 211
Diffraction
of concave grating 326
of plane grating 318
on slit 235
Diffusion
in arc plasma 59
in spark plasma 124
Diffusion halo, photographic emulsion 432
DIN ratings 417
Direct illumination, of prism spectrographs 296
Direct image 329
Discharge channel 123
dilatation 124, 130
Discharge tubes 187
spectrum 192
voltage drop 190
Dispersion 222
angular 222
anomalous 226
light, *see* Light dispersion
linear, *see* Linear dispersion
normal 276
practical 222
in prism spectrographs 254, 282
specific 225
Dispersion curve 223
glass spectrography 254, 282

Dispersion curve — *continued*
optical materials 225
quartz spectrographs 254, 274
vacuum prism spectrographs 290
wavelength determination 389
Dispersion nomograms 391
Dispersion tables 391
Displacement currents 206
Doppler effect 79
Doppler line broadening 79
Double grey-wedge 503
Double spark-gap 168
Double spectrum projector 397
Doublet lines
relative intensity 39
wavelength 31
Dowell–Menzies-type spectrograph 358
Drude–Lorenz relationship 364
Dry rectifiers 98
Dynode, photomultiplier 483, 485
thermal emission 488

Eagle-system concave-grating spectrographs 353, 356
Eberhardt effect 452
Ebert-system spectrographs
double plane-grating 346
plane-grating 341, 344
Echelle grating 336
Economic prism arrangement 287
Einstein equation 473
Elastic collision, of particles 55
Electron
mobility 58
pressure 70
rotation 26
temperature 56, 57, 127
Electronic band 43
Electronic control
a.c. arc generators 172
high-voltage spark generators 177
by magnetic methods 104
mains voltage d. c. generators 108
medium-voltage spark generator 182
by pulse generator 103
Electronic excitation 67
Electronvolt 16
Emission coefficient 37
Empty magnification 258
Enlighting strength, *see* Irradiation strength
Enrichment
in anode space 61
in cathode space 59, 61
Equalisation, of voltage variation, of arc current 98
Equalising current, spark excitation 155
Error sources, of spectrographs 262–273
Evaporation effect 75
Excitation 49–220
arc 51–122
classification 49
in discharge tubes 187–207
in flash tube 202
in Geissler tubes 192
by high-frequency current 204
in hollow-cathode tubes 195
by plasma jet 114, 119
potential (lowest) of gaseous elements 194
simultaneous, of elements 75
spark 122–186
Excited state, statistical weight 37
External photoelectric effect 473
Extinction coefficient 382

Fabry–Perot interferometer 376, 379
Faraday dark field 190
Fatigue phenomena
photocell 476, 479
photomultiplier 491
Fery prism 291
Fery spectrograph 290
Feussner-type spark generator 169
Filament spark generator 175
Filter circuit 101
Filter constant 313
Fine structure, of spectra 21, 24
study 222, 372–380
First-order spectrum, *see* Normal spectrum
First-order standard lines 386
Flash discharge 202, 211
Flash tube 202, 211
circuit diagram 203
Fluorite prism 244, 270, 288
Fluorite, as prism material 239, 240
Fogging density 412
variation of, with development time 448
Four-prism quartz spectrograph 280
Försterling prism system 242
Försterling spectrographs 281, 284
Fractional distillation 106
Free–free electron radiation 58, 86
Free path value of electrons 58
Fresnel equation 265
Fuess quartz spectrographs 274, 280
Fundamental line 20
Fixing 410

Gaertner-system quartz spectrographs 275
γ (gamma) value 414
change 415, 435, 448
Gas, influence
on background 86
on breakdown voltage 93
Gas excitation potential 194
Gas-filled photocell 476, 477
Gas pressure
discharge tube 188, 189
flash tubes 202

Gas pressure — *continued*
Geissler tube 193
high-frequency excitation 205
hollow cathode 197
influence on, of breakdown voltage 94
Gas temperature
arc plasma 56
spark plasma 127, 218
Gaseous element, lowest excitation potential 194
Gauss differential logarithm 455, 463
Geissler tube 192
Ghost
'grazing incidence' 265
Lyman 371
Rowland 369
Giant pulse (GP) laser 213
Glass, optical 239, 240
Glass prism 241
composite 241
Försterling 242
Glass spectrographs 281–286
dispersion curve 254, 282
high-performance 284
low-performance 284
medium-performance 282
Glimm discharge 188
Glimm stabilisers 494
Glimm voltage 495
Graetz circuit rectification 98, 106
Grating
angle of deviation 323
angular dispersion 324, 330
aperture 351
blaze angle 338
common 326, 362
concave 326
constant
concave-grating 326
optimal 331, 337
plane-grating 318
direct image 329
dispersion 318, 322, 324, 325, 326, 327, 330, 338, 342, 359, 363
errors 369, 371, 372
illumination
by grazing incident light 325, 364
by parallel beam of light 358, 359
material 326
Merton 339
production 337
replica 339
Rowland 326
ruling 337, 363
Sirks astigmatic difference 352
slit width 335
transmission 318
zero-order spectrum 323, 329
Grating spectrographs 340–366
Abney-system 353, 354
astigmatic 352, 353
basic principle
concave-grating 350
plane-grating 340
Czerny–Turner-system 341, 343
dispersion 332
Dowell–Menzies-system 358
Eagle-system 353, 356
Ebert-system 341, 342, 344
errors 367–372
atmospheric pressure changes 372
light losses 371
scattered light 371
temperature changes 371
Jacob–Ebert-system 347
linear dispersion
plane grating 332
Wadsworth-system 359
'order-sorter' 349
Paschen–Runge-system 353, 355, 361
photographic speed 336
predispersion unit 347
resolving power
practical 335, 364
theoretical 334
Rowland-system 353, 354
stigmatic line formation, concave-grating 358–361
stray light 371
vacuum 362, 364
Wadsworth-system 359
Grating spectrophotometer, high-performance 361
'Green gap' 434
Grey filter 312
stepped 313
Grey-wedge method 501
Grotrian term scheme, *see* Term scheme
Ground state 17
Group quantum number 29

Halo-free emulsion 431
Hand-spectroscope 293
Hansen-type diaphragm 317
Hartmann formula 226
Hartmann relation 392
Hartmann rule 430
HD number 417, 418
Hefner candle 404
Heisenberg uncertainty principle 81
Helium, term scheme 333
High-frequency arc generator 114, 120, 205
High-frequency plasma jet 119
High-performance spectroscopes 294
High-speed spectrographs 261
High-voltage a.c. arc generator 108
High-voltage rectified arc generator 169
High-voltage spark generator 164, 165
Hilger quartz spectrographs 274

Hilger vacuum polychromator 289
Hollow cathode, transitional-type 199
Hollow-cathode excitation 195
Hydrochinone 444, 445
Hydrogen
  line series 17
  Rutherford atomic model 23
  Sommerfeld atomic model 24
  term scheme 21
Hittorf dark field 190, 195
HK number 404
Huet quartz spectrograph 281
Huyghens principle 319
Hyperfine structure 21
Hyperpanchromatic emulsion 434

Ignition current, arc generators 95, 99
  Calker principle 102
  electronic control 104, 105
  mechanical control 102
  Pfeilsticker method 101
  pulse generator 103
  rotating switch 102
  Tesla transformer 100
Ignition signal, produced by magnetic effect 104
Ignition voltage, discharge tube 168
Illuminating lens 298
Illumination
  of dispersion devices 224
  of prism spectrographs 296–317
    by collecting lens 298
    by concave mirrors 305, 307
    by cylindrical lens 300
    direct 296
    by external focusing 298
    by focusing on slit 298
    by intermediate focusing 301, 306
    by internal focusing 299
    by lens in front of slit 300
    by photocurrent attenuation 312–318
    by Preuss raster-lens 304
    by rotating concave mirror 310
    by time-resolution 307
    by Twyman system 304
    by vibrating mirror 311
    by Zeiss intermediate focusing method 301
  of spectral line photometers 467
Illumination strength unit 404
Induced emission 208
Inelastic collision 67
Infra-red emulsion 434
Infra-red spectral range 13
Input energy 212
Insulation noise 489
Intensity
  'light' 404
  of spectral lines 351
    absolute 37
    relative 38
Intensity rules
  qualitative 38
  quantitative 39
Intensity variation, of atomic lines, with plasma temperature 132
Interelement effect 75
Interference filters 383
Intermittency effect 438
Interrupted arc generators
  electronically controlled 113
  mechanically controlled 111
Ionic current 488
Ionisation 70
  and arc temperature 73
Ionisation energy 34
Ionisation potential 34
Ion line 34
Ion spectra 15, 34
Iron–hydrogen resistor 393
Iron lines 388
  for emulsion calibration 421, 422
  secondary standard 386, 387
  spectrum atlas 368
Irradiation strength 402
ISzP 28 quartz spectrograph 274, 275

Jacob–Ebert-system plane-grating spectrograph 347
Jones sensitivity criterion 418

Kaufmann stabilisation condition 90
Kerekes–Ág relationship 41
Kerr device, as optical switch 213
Kerr effect 214
Kipp and Zonen quartz–water optics spectrograph 286
Kohlrausch model, of diatomic molecule 44
König–Brodhun rule 500
Krampl–Glock quartz spectrograph 278
Kranz plasma jet 115

Lambert's absorption law 262, 282
Lambert's distance law 402
Laser 208
  alignment of radiation, by microscope 218
  application 215–220
    direct procedure 216
    local analysis 218
    scanning method 219
    secondary spark-gap method 217
  crater (burning spot) 216
  efficiency 212
  energy density 215
  energy losses 209
  fundamental principle 208, 210
  giant pulse 213
  head 211
  illumination, by flash tube 211, 212

Laser — *continued*
input energy 212
macro-vapour space 217
microplasma 215, 216
Nd^{3+}-glass 210
non-controlled pulse type 211
optical activator 209
optical switch 213, 214
output energy 212, 216
pumping light 209
Q-switch 213, 214
radiation pulses 211, 214
resonator 209, 214
ruby 210–212
excitation scheme 211
saturation state 208
secondary spark-gap excitation 217
solid-phase 210–215
'spikes' 211, 214
three-level system 209
working distance 215
Last lines 21
Latent image 409
Lenses
aperture 248
camera 251
collimating 245
illumination 296
light absorption 263
optical faults 245
astigmatic 247
chromatic aberrations 245
coma 247, 248
secondary chromatic 245
spherical aberration 245
zone error 246
raster 304
reflection 265
stray light 268
Light absorption
of gelatin 433
of lenses 263
of prism spectrographs 262
Light attenuation, *see* Photocurrent attenuation
Light dispersion 221–384
by autocollimation 232
by concave grating 326–332, 350–367
by Fabry–Perot interferometer 376–380
by Fery prism 290
by filter 380–384
by plane grating 318, 326, 340–350
by prism 227–295
by stepped grating 373–376
Light filters 380–384
absorption type 382
interference type 383
'narrow' type 381
stepped 313
'wide' type 381
Light 'intensity' determination
fundamentals 401–408
by photoelectric measurement 409, 473–499
by photographic procedure 405, 407, 409–472
by spectrographic method 405, 407, 409–472
by spectrometric method 405, 409, 473–499
by spectroscopical method 405, 408, 499–509
by visual measurement 405, 408, 499–509
Light losses
of grating spectrographs 371
of prism spectrographs 262
Light polarisation 264
Light quantity 404
Light radiation
arc plasma 67
discharge tube 190
dispersion 221–384
refraction 264
spark plasma 125–131
unit 403
Limiting angle, in light refraction 225
Limiting resistance, of arc generators 90, 97
Limiting wavelength
in external photoelectric effect 473
for metals 480
Limiting wave number, free–free electron radiation 86
Linear dispersion 222
of grating spectrographs 333, 359, 365
of prism spectrographs 253
Liquid prism 244
Liquid-prism spectrographs 286
Lithium fluoride prism 244, 288
Local analysis
by filament spark excitation 176
by laser excitation 218
Logarithmic rotating disc 317
Low-contrast developer 446
Low-contrast photographic emulsion 414
Low-voltage spark generators, *see* Mains voltage spark generators
l-transformation 459
Lumen 404
Lumen-second 404
Luminescent cathode layer 190
Lux 404
Lyman ghost 371
Lyman series 18
Lyman spectral range 13

Macro-error, of emulsion 450
Macro-vapour space 217
Magnetic quantum number 27
Magnetic voltage stabilisers 493

Magnesium, term scheme 35
Magnification, of dispersion device 237
Mains voltage a.c. generators 108
Mains voltage a.c. spark generators 165, 183
Maintaining current, arc 95
Malpica equation 40
Maxwell–Boltzmann distribution law 16
Measuring microscopes 398
Mechanical ignition, of arcs 99
Mechanical light-equivalent 404
Mechanically controlled ignition 102
Mechanically controlled spark generator
  high-voltage 169, 174
  mains voltage 185
  medium-voltage 182
Mechanically interrupted arc generator 111
Medium-performance spectroscope 294
Medium-voltage spark generator 165, 180
Mercury, term scheme 32
Mercury-vapour rectifier 98
Merton plane grating 339
Metal mirrors
  concave 250
  parabolic 249
Metastable state 15, 32
Metol–hydroquinone developer 446, 449
Micro-errors, of emulsion 443
Micro-plasma 216, 217
Minimum position, of prism 228
Molecular spectrum, *see* Band spectrum
Molecule–electron collision 67
Multiplet lines
  absolute intensity 37
  relative intensity 31, 39, 420
Multiplicity 21, 31
  maximal 30
Multiprism quartz spectrographs 279
Multi-source unit 180

Natural light 208
Natural line width 78
Natural quartz 240
Negative glow 190, 192, 195
Nd^{3+}-glass laser 210
Nitrogen, thermal conductivity variation with temperature 54, 130
Nitrogen bands 86
NO bands 86
Non-linearity errors, of photomultiplier 490
Normal cathode fall 191
Normal dispersion 226
Normal photoelectric effect 474
Normal sensitivity, of emulsion 414
Notation
  for atomic lines 34
  for azimuthal quantum numbers 25
  for group azimuthal quantum numbers 29
  for ion lines 34
Nottingham equation 90

Opaque media 262
Optical electron, *see* Valency electron
Optical resonator 209
Optical spectra 13, 14
Optical switch 213, 214
Output energy, laser 212, 216
Over-exposure 414

Panchromatic photographic emulsion 434
Partial discharges 166
Partial optical density transformation 458
Paschen law 188
Paschen–Runge-system spectrograph 353, 355
Paschen series 18
Pauli exclusion principle 33
Pellin–Broca prism 241
Pentaprism 225
Pfeilsticker method 101
Pfund standard lines 386
Phase angle shift 155, 159, 170
  elimination 170
Photocathode
  alkali-metal 469, 479, 481
  alkali-oxide 480
  composite 480
  pure metal 479
  sensitised 481
Photocells 476
  acceleration voltage 477, 479
  fatigue phenomena 478
  gas-filled 476, 477
  internal resistance 481
  photocurrent–light flux 477
  Schottky effect 477
  secondary radiation 479
  vacuum 476, 477
Photochemical reaction 409
Photocurrent attenuation
  by block wire 316
  by free opening of camera lens 261
  by grey filter 312
  by Hansen diaphragm 317
  by intermediate diaphragm 317
  by iris diaphragm 316
  by rotating sector 315
  by stepped grey filter 313
Photocurrent intensity 473
Photoelectric effect, external 473–476
Photographic density 411
  change
    due to background 440
    due to Callier effect 441
    due to chemical effects 443–453
    due to developing 443
  determination
    by line photometer 465
    by 'density plastics' 466

Photographic density — *continued*
by s. p. d. scale 464, 466
deviation from reciprocity 438
due to Eberhardt effect 452
due to intermittency effect 438
due to physical effects 438–443
transformation 454–462
Photographic emulsion 409
characteristic curve 411
construction 419–430
commercial notation 434
density 421
developing 443–451
diffusion halo 432
drying 411
films 430, 436
fixing 410
γ (gamma) value 414, 415, 435
graininess 442
'green gap' 434
halo effect 430, 431
hyperpanchromatic 434
latent image 409
low-contrast 414
macro-error 450
micro-error 433
non-sensitised 432
orthochromatic 434
panchromatic 434
reflection halo 431
resolving power 442
sensitised for
infra-red 434
ultra-violet 433
sensitivity designation 417
solarisation 414
for spectral analysis 435
spectral sensitivity 432
washing 410
Photographic speed
of grating spectrographs 336
of prism spectrographs 261
Photometer, spectral-line, *see* Spectral-line photometers
Photometric radiation equivalent 404
Photomultipliers 481–499
amplification factor 482
anode current 482, 486
application 495
charging method 497
decade counting system 497
discharging method 497
graphical method 498
integrating-condenser measurements 496
cathode material 484, 485
construction 483, 484
dark current 486
dynodes 483
emission coefficient 482
fatigue 490
faults 490
measurement errors 490
non-linearity 490
power supply 492
quality 490
relative photoeffect 486, 487
Sb–Cs dynodes 484, 485, 486, 490
secondary emission 491
shot effect 489
temperature variations 491
voltage–anode current characteristics 486, 487
wavelength 485
Physical spectral line profiles 78
Pfund secondary standard lines 386
Plane-grating dispersion 318–326
Plane-grating spectrographs 340–350
Planck radiation law 47
Plasma
arc 53
spark 123
Plasma jet, high-frequency 114
Polarisation angle 264
Polarisation filter 505
Polarisation technique, for visual intensity measurement 504
Population, of electronic levels 16
Population inversion, laser 208
Positive glow layer, discharge tube 190
Positive light column, discharge tube 190
Potassium bromide, effect 446
Power supplies
for a.c. arc generators 98
for d.c. arc generators 99
for d.c. plasma jet 110
for high-frequency plasma jet 120
for photomultiplier 442
Preuss-type lens-raster 304
Prisms 237–262
Abbe 241
Amici 243
angle of deviation 227, 230
angular dispersion 231
autocollimating (Littrow-type) 232
base 227
Calbier 242
composite 241–244
constant-deviation 242
Cornu quartz 240
deviation angle 227, 228
direct-vision glass 243
equation 229, 231
Fery 291
fluorite 244
Försterling-system 242
glass 241
image-reversing 225
light absorption 263
liquid 244, 286

Prisms — *continued*
lithium fluoride 244
Littrow-system 232, 277
materials 239
Pellin–Broca 241
reflection 264
refractive angle 227
optimum 238
Straubel 242
stray light 267, 277
Zenger 244
Prism spectrographs 224–292
autocollimating
glass 284, 285, 286
quartz 276
camera lens 251
collimator lens 245
construction 233
dispersion curve
glass instruments 254, 282
quartz instruments 254, 274
error sources 262–274
Fery-system 290
Fuess Q 110 H 274, 280
Gaertner-system 276
glass
high-performance 284
low-performance 281
medium-performance 281
Hilger Medium Quartz 274
Hilger Polyvac 289
Huet ISP28 274, 275
Huet ISP67 285
Huet UV 120 281
illumination 296–317
Krampl–Glock-system 278, 279
light absorption 262
light reflection 264
linear dispersion 253
liquid-prism 286
Littrow-system
glass 284
quartz 274, 276
magnification 252
multiprism 279
normal 258
over-magnifying 258
practical resolution 256–261
quartz 274, 281
single-prism 274
slit 235
spectral purity 257
Steinheil G. B. Universal 282
stray light 267
theoretical resolving power 254
Zeiss Q 24 274, 275
vacuum 290
Prism spectroscopes, *see* Spectroscopes
Profilage method 499
$P_{\frac{1}{2}}$-transformation 456
Pulse generator 103
Pulse-spark generators 179
Pumping light 209

Quantum number
azimuthal 25
group 29
inner 26
magnetic 27
principal 23
radial 25
spin 26
group 29
Quartz 239, 240
Quartz spectrographs 274–281, 286, 292

Radial quantum number 25
Radiation
induced 208
intensity 68, 76
spontaneous 208
units 403
Radiation diffusion 124
Raiski spark generator 169
Rayleigh principle 223
Rectified continuous arc generator 106
Rectified high-voltage spark generator 174, 177
Rectified medium-voltage spark generator 181
Rectifier 98
Reflection, of light 264
Refractive index 224
calculation 226
effect of atmospheric pressure changes 272
of optical materials 240
Relative intensity, of spectral lines 38
Remanent inductance 137
Replica gratings 339
Residence time, of elements 62
Resolving power
of human eye 261
of photographic emulsion 442
practical
of grating spectrographs 335
of prism spectrographs 256
of spectroscopes 261
theoretical 223
of echelle spectrographs 374
of grating spectrographs 334, 365
of prism spectrographs 253
Residence time of atoms, in arc plasma 62
Resonance broadening 82
Resonant absorption 208
Resonant lines 83
Resonant radiation 83
Resonant spark generators 168
Resonator, optical 209
effective diameter 214

Reversible fatigue phenomena 491
Rotary rectifiers 99
Rotating disc
  for light attenuation 315
  for time-resolved spectra 307
Rotating mirrors
  optical switch 214
  for time resolution 308
Rotating prism optical switch 214
Rotating sector, *see* Rotating disc
Rotational band 43
Rowland circle 327
Rowland grating 326
Rowland-system spectrographs 353, 354
Ruby laser 210
Russell–Saunders term symbol 27, 34
Rutherford atomic model 23
Rutherford–Bohr theory 22
Rydberg alternation rule 31
Rydberg–Ritz combination principle 16

Saha equation 70
Scheibe–Lomakin equation 40
Scheiner scale 417
Schottky effect 477
Schumann plates 433
Schumann spectral range 13
Schwartzschild relationship 436
Secondary collision 55
Secondary spark-gap excitation 217
Secondary standard lines 386
Secondary Stark effect 80
Seidel transformation 455
Selection rules 28, 32
Self-absorption, of spectral lines 82
Sensitised emulsions 433
Sensitivity curve, of human eye 500
Series formulae 17, 18, 20
Shot effect 489
Simultaneous element excitation 75
Sirks-astigmatic difference 352
Skin effect 206
Slit
  entrance 467, 468
  height 236, 336, 367
  of spectrographs 233, 235
    asymmetrical 235
    diaphragms 237
  optimum 236
  symmetrical 235
  width 235, 335
Solid angle, definition 401
Sommerfeld–Kossel displacement law 34
Sommerfeld theory 24, 25
Spark discharge 122, 128, 234
  after-effects 126
  afterglow 123, 126, 128
  atmospheric lines 127
  axis temperature 129
  background 128
  'burning' time 126
  channel 123
    dilatation 134
  cooling 129, 130, 133
  corona 122
  damping resistance, influence 160
  full, of condenser 123, 125
  electron temperature 126, 127
  gas density 128, 130
  gas mantle 126
  gas temperature 128, 129, 130
  ignition 123
  internal cooling 126
  light radiation 125, 129
  radiation diffusion 124
  radiation light energy 129
  resistance 123, 129
  spectral lines 129
  temperature 128, 131
  time-resolution 125, 127
  'trial-and-error' measurement 133
  voltage–time oscillogram 127
Spark generators 135–187
  basic principles 134
  circuit diagram 135
  classification 164
  discharge circuit 135–152
    critically damped 138, 147
    over-damped 138, 145
    periodically damped 138
  discharge current 135–152
  electronically controlled
    high-voltage 177
    mains-operated 186
    medium-voltage 182
  filament 175
  high-voltage 165, 179
    with controlling spark gap 168
    with double spark gap 168
    electronically controlled 177
    Feussner-system 169
    mechanically controlled 169
    non-controlled 165
    pulse-spark 179
    Raiski-type 168
    with rectified charging circuit 174
    'resonant' 168
  mains voltage 183–187
    with asymmetrical controlling spark gap 185
    Calker-system 183
    electronically controlled 186
    mechanically controlled 186
    Pfeilsticker-system 183
  medium-voltage 180–184
    electronically controlled 182
    mechanically controlled 181, 182
    'multi-source unit' 181

Spark plasma
afterglow 123, 126
ambipolar diffusion 124
axis temperature 129
'characteristic temperature' 132
cooling 126, 130, 133
dilatation 124
electron temperature 127
formation 124
free electrons 129
gas density 128, 130
gas mantle 126
gas temperature 126, 128, 129, 130
internal cooling 126
physical parameters 127, 128
radiated light energy 125–127, 129
radiation diffusion 124, 125
resistance 129
temperature 129
determination 131–134
s. p. d. scale 464, 466
Specific dispersion 225
Spectral-line photometers 467–473
based on alternating illumination measurement 472
based on direct measuring method 470
effect of stray light 368
illumination area 466
measurement principle 470
with photocell 469
with photoelement 469
'rapid photometer' type 471
Spectral lines 14, 324
atomic 15
'characteristic temperature' 131
curvature 237
density 412
displacement 80, 271, 272, 371, 372, 507
distortion
due to Doppler effect 79
due to Kostinsky effect 453
due to pressure broadening 81
due to resonance braodening 82
due to Stark effect 80
due to two unresolved lines 223
emission coefficient 37
intensity 14, 37, 38
ion 34
natural form 78
physical profile 78
resonant 83
self-absorption 82
self-reversal broadening 83
series 17
standard
first-order 386
secondary 386, 387
wings 453
Spectral plates 435
Spectral regions 13
Spectral sensitivity 432
relative 475
Spectroscopes 292–296
binocular 295
hand 293
high-performance 294
medium-performance 294
Spectrum 14
atomic 15
band 42–46
classification 13
continuous 46
fine structure 21
ion 15, 34–36
normal 325, 330
order number 322, 327
Spherical aberration
lens 248
parabolic mirror 249
Static photomultiplier 482
Statistical weight, of excited particles 37
Stepped light filters 313
Stray light
grating spectrograph 371
line photometer 468
prism spectrograph 267, 277
Stigmatic image dispersion, concave-grating spectrograph 358
Superelastic collision 67

Temperature
axis 129
characteristic 131
determination
arc plasma 73
spark plasma 131
electron
arc plasma 56
spark plasma 127
gas
arc plasma 56
spark plasma 127, 128
Term 16, 24
Term scheme 24
aluminium 36
CO molecule 45
Grotrian 20
helium 33
hydrogen 21
magnesium 35
mercury 32
sodium 28
Term symbol, Russell–Saunders 27
Tesla current
excitation 99
transfer
according to Calker principle 101, 183

Tesla current — *continued*
according to Pfeilsticker principle 101, 183
by secondary spark gap 184
Thermal excitation 50
Thyratron tube 103
Time-resolved spark discharge 125
Time-resolved illumination, of prism spectrographs 307
Transformation constant (*P*-transformation) 458
Transformer inductance 153
Transition probability 16, 37
Transitions, between possible energy levels 28
Transmittance
of light filters 381
of optical materials 259
Transparent media 262
'Trial-and-error' method 133
Triple collision 56, 67

Uhlenbeck–Goudsmith theory 26
Ultra-violet spectral region 13
Ultra-violet plates 433

Vacuum polychromator, with fluorite prism 289
Vacuum spectrographs
grating type 362, 364
prism type 287
Vacuum ultra-violet spectral region 13
Valency electron 17, 22, 31
Vaporation effect 75
Vibrating reed tube voltmeter 492
Vibrational–rotational band 43
Visual intensity measurement 499–510
by double grey-wedge 503
by grey-wedge 501
fundamental principles 499
instruments 505
by polarisation techniques 504
Voltage stabilising methods 493–495
by electron tubes 494
by glimm stabiliser 494
by iron–hydrogen resistors 493
by magnetic voltage stabilisers 493

Wadsworth anomaly 359
Wadsworth-system spectrographs 359
Watkins factor 445
Wavelength
determination 385–399
by dispersion curve 389
by dispersion nomogram 391
by double-spectrum projector 397
errors 394
by Hartmann formula 392
by interpolation 388
with line photometer 396
with measuring microscopes 398
by normal spectrum of grating spectrographs 393
unit 385
Weber–Fechner law 499
Wedge
double 503
grey 501
Wien displacement law 47
W-transformation 455

Zeiss quartz spectrographs 274
Zenger prism 244



Professor JÓZSEF MIKA started his career as a metallurgical engineer in the metallographic laboratory of the Iron Plant, Diósgyőr. He later moved to the Chemical Department of the High School for Mining, Metallurgy and Forest Engineering and when the latter became part of the Technical University of Budapest, he was appointed to the Chair of Analytical Chemistry, a position from which he retired in 1958. Prof. MIKA, a Kossuth Prize winner, has extensive research interest in microchemistry, electroanalysis, and optical emission and absorption spectroscopy. The Hungarian Academy of Sciences awarded him the title of Doctor of Chemical Sciences in 1952. He has published six books and some 50 original papers.



Professor TIBOR TÖRÖK graduated in chemistry, physics and mathematics in 1938 and obtained his doctor's degree in chemistry in 1939. Subsequently, he received a scholarship to study at the University of Munich under Professor Walter Gerlach. Prof. TÖRÖK, a Kossuth Prize winner, established several spectroscopic laboratories in both industry and universities, and in 1950 was appointed Assistant Professor at the University of Budapest. He obtained the title of Doctor of Chemical Sciences from the Hungarian Academy in 1952. In 1963 he was elected Professor of the Department of Inorganic and Analytical Chemistry. He is the member of the editorial advisory board of *Spectrochimica Acta*, *Hungarian Scientific Instruments* and of the 'Continuous Committee of the Colloquium Spectroscopicum Internationale' (C.S.I.); he is honorary member of the 'Groupement pour l'Avancement des Méthodes Physiques d'Analyse'. Prof. TÖRÖK has published some 150 original papers and is also the author of four books.