J. F. JAMES
R. S. STERNBERG

The Design of
Optical Spectrometers

CHAPMAN AND HALL LTD
11 NEW FETTER LANE LONDON EC4
Contents

Authors' Preface ................................................................. page ix

1. INTRODUCTION: THE IDEAL SPECTROMETER
   1:1 The Classification of Spectrometers .......................... 1
   1:2 The Ideal Spectrometer ........................................... 5
   1:3 Instrumental Profile .............................................. 7
   1:4 Luminosity ........................................................... 10
   1:5 Choice of a Spectrometer ........................................ 12

2. SOME PRINCIPLES OF OPTICAL DESIGN
   2:1 Introduction ......................................................... 14
   2:2 Gaussian Optics .................................................... 14
   2:3 Aberrations .......................................................... 20
   2:4 Aplanatic and Isoplanatic Systems ............................ 27
   References ........................................................................ 30

3. ELEMENTARY FOURIER THEORY AND ITS APPLICATION TO OPTICAL SPECTROMETRY
   3:1 Fourier Series ........................................................... 31
   3:2 Fourier Transforms .................................................. 32
   3:3 Convolutions .......................................................... 34
   3:4 The Algebra of Convolutions ...................................... 36
   3:5 Physical Applications of Convolutions ........................ 36
   3:6 Fourier Pairs of Interest in Spectrometry .................... 37
   References, Bibliography ................................................. 40
### 4. PRISM SPECTROMETERS

4.1 Introduction 41
4.2 The Focal Curve Theorem 43
4.3 The Prism Spectrometer 45

References 50

### 5. SPECTROMETERS EMPLOYING PLANE DIFFRACTION GRATINGS

5.1 The Grating Equation 51
5.2 Resolving Power 52
5.3 Blazing 55
5.4 Angles of Incidence Different From the Blaze Angle 55
5.5 Overlapping of Orders and Free Spectral Range 57
5.6 Order Sorters 58
5.7 Grating Ghosts 59
5.8 Dispersion 61
5.9 The General Grating Equation – Oblique Incidence 62
5.10 Differential Dispersion 68
5.11 Luminosity and Resolving Power 69
5.12 Mountings for Plane Gratings 70
5.13 Grating Spectrometers in the Infrared 79

References 80

### 6. SPECTROMETERS EMPLOYING CONCAVE DIFFRACTION GRATINGS

6.1 The Rowland Grating 81
6.2 Geometrical Theory of the Concave Grating 83
6.3 The ‘Japanese’ Mountings 85
6.4 Polychromators 89
6.5 Practical Design Considerations 93

References, Bibliography 94

### 7. THE FABRY–PÉROT SPECTROMETER

7.1 Theory of the Fabry–Pérot Étalon 95
7.2 The Fabry–Pérot Étalon as a Spectrometer 98
7.3 Effective Finesse 100
7.4 The Luminosity-Resolving Power Product 102
7.5 Alternative Methods of Order Sorting 105

### 8. MULTIPLEX SPECTROMETERS

8.1 Multiplex Systems and the Multiplex Advantage 116
8.2 The Resolving Power 119
8.3 Truncation and the Sampling Theorem 120
8.4 Aliasing 123
8.5 Practical Methods for Fourier Transformation 125
8.6 Apodizing Functions 127
8.7 Practical Details of Multiplex Spectrometers 128
8.8 Positions of Aperture and Field-Stops 136
8.9 Field Division 138
8.10 The Problem of Noise in Multiplex Spectrometry 140
8.11 Other Multiplex Spectrometers 142

References, Bibliography 145

### 9. DETECTORS

9.1 Introduction 146
9.2 Types of Detectors 147
9.3 A Basic Distinction 147
9.4 Noise in Detectors 148
9.5 Criteria of Performance 150
9.6 Descriptions of Individual Detectors 152

References 166

### 10. AUXILIARY APPARATUS

10.1 Source Optics 167
10.2 Detector Optics 170
10.3 Source at Infinity 171
10.4 Source and Detector Systems with Reflecting Optics 173
10.5 Double-beam Spectrophotometry 175
Contents

10:6 Monitoring the Source 179
10:7 Multiple Pass Absorption Cells 181
References 182

11. GENERAL PRINCIPLES OF CONSTRUCTION

11:1 Constructional Materials 183
11:2 Modes of Construction 186
11:3 Mounting of Optical Parts 187
11:4 Linear Wavelength Drives 191
11:5 Slit Jaws 194
11:6 Slit Mechanisms 195
11:7 Michelson Interferometers 200
References, Bibliography 206

12. OPTICAL COMPONENTS

12:1 Optical Components 207
12:2 Transparent Optical Materials 207
12:3 Materials for Mirrors and Grating Blanks 211
12:4 Filters 212
12:5 Reflecting Coatings 215
References 217

Appendix 1 Fundamental Design Formulae 218
2 Alignment of Spectrometers 222
3 Readily Available Wavelengths for Laboratory Use and Wavelength Standards 225
4 (i) The Sinc Function (ii) The (Sinc)^2 Function (iii) The Airy Function 227

Index 231

Authors' Preface

Although a number of excellent books on the general subject of 'spectroscopy' have appeared during the last half-century, none of them appears to deal in any detail with the design or construction of spectroscopic instruments. Indeed, only a few go so far as to suggest to the reader the sort of instrument that might be suitable for his particular investigation. As the quantity of physical, astronomic and chemical research continues to increase, it is becoming more apparent that the researcher can afford to devote less time to the understanding of his tools, and he has to consider the spectrometer, like so many other pieces of scientific equipment, as just another 'black box' which he can use in his work. He expects to obtain his results without having to inquire too closely into its inner workings. The manufacturers of optical equipment have, to a considerable extent, fulfilled his needs. A wide variety of spectrosopes, spectrographs and spectrometers can be seen in the catalogues of several firms. From among these instruments it is possible to choose something suitable for most needs of a routine or semi-routine nature, but clearly it is impossible to provide something that is tailored to one particular experiment, particularly when that experiment is on the verge of what is physically possible. It is precisely this sort of experiment that people need to do when they are working on what has been called the 'frontiers' of physics and chemistry, and it commonly happens that the commercially made spectrometer is not adequate for the task.

Lord Rutherford is said to have made a dictum, that in experimental science one should never attempt anything difficult: the modern research student might reply, with just a touch of bitterness, that all the easy things have been done and that the difficult things have been saved up for him. What is implied really is that one should not lightly attempt
research in a field where the experimental technique is poorly developed so that all one's time is spent in persuading recalcitrant apparatus to perform. Fortunately there are some people who take a deliberate interest in doing the difficult things: who make it their business to persuade and coax inanimate matter and to advance the experimental art. They too have made their contribution to the advance of knowledge, despite a certain amount of condensation by the purists. Quite an array of respectable precedents can be quoted for making this a method of research: to choose a technique, to advance it as much as possible and only then to search for a physical problem that it can be applied to.

Spectroscopy (or spectrology, to introduce here a word coined by Fellgett) belongs to this branch of physics. The quantum theory of the atom and the molecule followed the discovery of line and band spectra, and followed at some distance for most of the time. Band spectra were only discovered because someone invented a bigger and better spectroscope so that the 'fluted' spectra, as they had been known, resolved themselves into lines. This is one of the classical examples of the physical discovery following advances in technique. Others might be quoted, but we do not intend this as an apology. We address ourselves to those who are interested in the spectroscopic art or the science of spectrology, whether for its own sake or as a means to an end: also to the worker who wishes to build a bigger and better spectrometer to see what he can do with it in his own line, and, in particular, to the chemist or astronomer who wishes to look at some source of light that he thinks may be interesting but inaccessible to commercially available instruments. We hope too that we may be able to help those who are only faintly interested in instrumentation to choose the right sort of spectrometer for their needs, even if they intend to buy one rather than make one. The modern limits of the art are far wider than most people imagine, and resolving powers and sensitivities are available today that were undreamed of even twenty years ago.

We have deliberately ignored the subject of spectrography. The techniques of design are quite different from those used for spectrometers, and for most non-routine problems they almost always give a performance which is inferior to that of a properly designed spectrometer. Occasionally it is found that there is an application for a spectrograph in which the photographic plate has been replaced by an image tube, and such an instrument is probably superior to a spectrometer for the exploration of a completely unknown spectrum. But for the study of the detail of a source whose broad outline is already known, the spectrometer wins because it can go to regions of resolution and of photometric accuracy where the spectrograph would be lost, even with an image tube. Even more potent, though less well understood (except to the Initiated) is the multiplex spectrometer, which can go as far as the conventional spectrometer in resolving power, but can tackle regions of very feeble intensity where any other instrument would be lost in the signal/noise swamp. Thus it is that we have devoted some considerable space to a discussion of the high performance grating spectrometer and to the multiplex spectrometer. We have not attempted to say the last or the latest word about them. Improvement is continuous as it is in most other branches of science. But the basic details of these instruments are now settled and it has been our design to provide the reader with the background to the more recent advances. Since many of these advances are marginal improvements in resolving power, sometimes purchased at the expense of much analysis and computer time, we have not attempted to include all of them or even to list them. We do not wish to convey the impression that it is impossible to design a respectable spectrometer without an intimate knowledge of all these latest developments. There should be enough information here to permit a research student to design, without any outside help, an Ebert spectrometer, for example, which is likely to do anything that a research student might be reasonably expected to do; and there should be enough information on the other types of spectrometer to provide an introduction to the material to be found in the journals. We cannot, as yet, pontificate about multiplex spectrometers. Only in the last few years has it become apparent that some techniques are better than others, and we have written only about those parts of the subject that are reasonably well settled. But again there should be enough information to allow a multiplex spectrometer to be built without help from other optics books.

There are some types of spectrometer, from the familiar to the bizarre, that we have ignored. This is either because, in our opinion, the technique is not well enough understood or because it has become apparent that the method is not likely to become of general interest. For example, it may be possible, with some particular apparatus, to achieve a resolving power of $10^{12}$ as a sort of tour-de-force. As a tool for asking questions of nature, however, such an instrument may not amount to much. It may well be too difficult to use or too delicate to adjust to be a standard piece of laboratory equipment. The field of study that uses devices like this is probably so specialized as to be of interest to half-a-dozen people or so, and they will be people who need no instruction from us.
THE DESIGN OF
OPTICAL SPECTROMETERS

J. F. James and R. S. Sternberg

Lecturers in Physics in the University of Manchester

CHAPMAN AND HALL LTD
11 NEW FETTER LANE, LONDON EC4
Introduction: The Ideal Spectrometer

1.1 The Classification of Spectrometers

There are many ways of classifying spectroscopic instruments. They can be considered according to the physical principles that they embody, whether they depend for their working on differential refraction, diffraction or interference; or alternatively according to the function that they are to carry out. That is, whether they are intended for the routine analysis of steel samples or the accurate measurement of wavelengths in an unknown source, or for searching for the radiation from a very feeble astronomical object.

From the point of view of the experimentalist the latter classification is the more useful, since it is according to function that he will select his instrument. Thus he will need to compare the merits only of those instruments that are suited to his purpose. The instrument designer, on the other hand, may well prefer the former type of classification, since each of his types follow a certain set of laws, and many of the design problems are common to the whole class. For this reason we shall here use the designer's classification, and the chapters of this book are divided accordingly. We shall not neglect the needs of the experimentalist, and at the appropriate points are critical tables outlining the main function of each type.

Many of the principles of design are common to all types of spectrometer, and indeed to all optical instruments. These will be considered first.

Fundamental Concepts

(i) Resolving Power. This is most easily understood as it applies to monochromators. These are instruments which select and transmit a certain range of wavelengths in the spectrum, while absorbing or reflecting all the rest. If the range selected is from $\lambda - (\Delta \lambda / 2)$ to...
\( \lambda + (\Delta \lambda /2) \), centred on wavelength \( \lambda \), the resolving power is defined as
\[ R = (\lambda / \Delta \lambda) \]
and is a dimensionless number. There are two kinds of criterion for resolving power, and these will be considered in the appropriate places. The concept can be applied to other instruments. For example, in a photographic spectrometer, if two wavelengths \( \lambda \) and \( \lambda - \Delta \lambda \) just appear as two separate lines on the photographic plate, the same formula applies.

(a) Luminosity. Despite the use of this term in photometry it has come into common use when speaking of the performance of an optical system. It has no connection here with the brightness of a source, but is used instead to describe the 'light-gathering-power' of an instrument. It is somewhat akin to 'speed' in a photographic system but is more general than this. Other terms, synonymous with luminosity in this sense, are 'étendue', 'throughput', 'light-grasp' and 'collecting-power'. The term denotes the amount of light that passes from the source, through the instrument and on to the detector. ('Detector' in this instance could mean a photosensitive cell or a grain of a photographic emulsion.) The amount will depend on:

(a) The area of the entrance aperture – usually the entry slit,
(b) The solid angle subtended at the entrance aperture by the collimator or dispersing element – whichever is the smaller – or alternatively on any apertures in the optical system which are conjugate to these. The concept of luminosity is not confined to spectrometers and in the general case of an optical system can be defined as:
\[ L = \text{(area of field stop)} \times (\text{solid angle subtended by the aperture stop at the field stop}) \]

In a spectrometer the luminosity can be defined as above, or equally well by the product of the area of the exit slit and the solid angle subtended there by the final focusing lens. The point here is that the luminosity is conserved as the light passes through an imaging system, \( L \) being a constant when measured at any pair of neighbouring stops. Normally one of the aperture stops – possibly the only one – will be the prism or diffraction grating that produces the dispersion. The focusing lenses or mirrors associated with it will be larger than this in order that they themselves shall not limit the stop. In any case, as will be seen in more detail in the section on geometrical optics, the aperture stop must be in one particular plane, and focusing elements which are too small will not only reduce the luminosity (and waste valuable resolving power) but will cause vignetting.

(3) Resolving Power-Luminosity Product. This is commonly known as the 'resolution-luminosity product', although the term is defined by \( E = R.L \). It is also known as the 'efficiency' of a spectrometer.

For many types of spectrometer the efficiency, defined thus, is a constant or a function of the wavelength that is being measured. In the Fabry–Pérot étalon, for example, we can write simply:
\[ E = R.L = 2\pi A \]
where \( A \) (a constant) is the area of the aperture stop of the system. For a grating spectrometer, on the other hand, the formula for efficiency is:
\[ E = Am\lambda/50a \]
where \( \lambda \) is the wavelength transmitted, \( m \) the order of diffraction, the slit length is \( 1/50 \)th of the focal length, and \( a \) is the grating constant.

The chief usefulness of the concept is as a means of comparing the relative merits of two competing systems. For a given instrument at a given wavelength the efficiency is fixed, but it is possible to trade resolution for luminosity or vice versa as required. The limit set to this trading is the theoretical diffraction limit to the resolving power of the system. In a simple prism spectrometer the light throughput can be increased simply by opening the slit, or the resolution can be increased by closing it. There comes a point, though, when closing the slit further does not increase the resolution, although the luminosity continues to fall. (In fact it falls faster than before.) Before commencing a design or choosing an instrument one should decide what resolving power and what luminosity are required. The formulae given later then provide an easy means of computing what a given type of instrument will do, and whether it will be suitable for the required purpose.

These remarks apply only to photoelectric spectrometers. In the case of the photographic spectrometer or spectrograph the situation is somewhat different. The luminosity is fixed by the area of one grain of the emulsion and the solid angle subtended by the camera lens. An exposure is made when approximately 1,000 photons have passed through this system. Thus nothing is to be gained, during the examination of an emission source, by opening the slit. The amount of light per unit area falling on to the plate is not changed.

This reveals a fundamental difference between photographic instruments and photoelectric instruments. For high sensitivity in a photographic instrument a small focal ratio is needed in the camera, whereas with a photoelectric instrument the rate at which power falls on to the detector depends only on the area of the aperture stop. A large focal ratio can be employed together with a larger exit slit, to avoid the
complications of geometrical aberrations which appear at low focal ratios. After passing the exit slit the light can be condensed by crude optics on to a small detector area.

The gain in sensitivity in a photographic instrument at low focal ratios is more illusory than real. For a given size of aperture, a small focal ratio implies a small image of the entry slit, and so a finer grain film to realize the same resolving power. The finer grain film in turn implies a slower film, and it turns out that there is almost exact compensation for the gain in speed due to the small focal ratio. For this reason spectographs as well as spectrometers tend to have large focal ratios, usually between $F.12$ and $F.25$.

All the above arguments apply to the case of a line emission spectrum. If there is a continuum emission, then the product of resolving power and flux is no longer constant. Instead, a doubling of the slit width will not only double the luminosity but also the spectral range that is transmitted. This allows four times as much flux through and it is $RIL$ rather than $RL$ that is conserved. Similarly, in a spectrograph, doubling the slit width will double the flux per unit area on the plate, so that there is an advantage in using a wide slit when a crude study of continuum emission features is to be made. Naturally any fine absorption detail, such as the Fraunhofer lines in the solar spectrum, will be 'filled in' and lost.

To sum up, there is little to be gained except the saving of space and weight by designing spectrometers with small focal ratios. Only in the most unusual circumstances, such as spectrometers designed for satellite or space-probe experiments, is it worth while accepting the complication that results from larger geometrical aberrations and the need for finer optical tolerances.

(4) INFORMATION BANDWIDTH. This is a piece of jargon taken from information theory, and is used here to describe the rate at which the spectrometer collects information about the spectrum. If each resolved element of the spectrum is regarded as one information channel then, clearly, a monochromator has a small information capacity compared with a polychromator, since only one channel is open at one time. The rate at which information is collected will depend on several things, of which the number of open channels is only one. The noise of the detector, the power in the channel and the noise (photon shot noise) in the channel are the others. The data-processing system should not contribute any significant amount of noise to the total. The chapter on detectors indicates how detector noise can be limited, and a high efficiency spectrometer is the obvious way to provide the maximum power for the detector. The point to be made here is that the whole information system from the source of light to the detector read-out must be analysed as a unit, in order that the optimum sizes and shapes may be chosen for each of the components. Such an analysis is not really necessary for those who seek more or less routine instruments, but if one is on the limit of sensitivity, an analysis, even at a relatively elementary level, is well worth while.

If the subject is considered in detail, the design of the system will depend on a priori knowledge of the spectrum. A scanning spectrometer, for example, may be programmed to avoid those parts of the spectrum that are known to be empty or without interest. The study of information theory as it is applied to spectrometric systems is far from complete, and those who would design systems for the extreme limits of spectrometric information would do well to consult the research journals.

When considered as an information-gathering device, the photographic spectrometer is far less at a disadvantage than appears at first sight. In spite of the very low sensitivity (= high noise level) of the photographic emulsion, there is the fact that every one of perhaps two or three thousand information channels is open simultaneously. For the examination of a spectrum where there is no a priori knowledge, a spectrograph compares favourably with many monochromators. Where it fails to compete is in the study of a small part of a spectrum at high resolution, a case which frequently arises. There are few information channels, possibly twenty or thirty, and the low sensitivity of the emulsion is not compensated by the simultaneity advantage. This would be the case for instance if the profile of an emission or absorption line were being studied in detail.

As we shall not be considering photographic instruments to any great extent in later chapters, it is worth parting from the subject with one valedictory word. As a store of information, the photographic plate is unrivalled in its capacity per unit cost, or per unit weight, or per unit storage time. None of these things are lost by those whose business it is to design photometric systems which go to the limit.

1-2 The Ideal Spectrometer

We will consider here the essential elements that go to make up a spectrometric system. Despite the apparent complexity of many spectrometers, much of the optical system is for auxiliary purposes, and that
6 · Design of Optical Spectrometers

part of the optical system that is concerned with the dispersion of the light is in principle, quite simple.

The Ideal Spectrometer consists of:

1. **An Entrance Aperture.** In a conventional spectrometer this consists of a long narrow slit. It may be of any shape in fact, and the shape and size depend on the particular means that are used to disperse the light.

2. **A Collimating Element.** This is to make parallel all the rays passing through one point of the entry slit or field-stop. This element may be a lens or a mirror, or it may be an integral part of the dispersing element, as in a concave grating spectrometer or a Féroprism spectrograph.

3. **A Dispersing Element.** This is any device which alters the intensity that passes through the system, in a way that depends on the wavelength.

4. **A Focusing Element.** This is to form an image of the entry field-stop at some convenient focal plane.

5. **An Exit Aperture.** This is a stop at the focal plane which transmits the light from the image that the focusing system has formed. There need not be a real aperture there. The exit field-stop could be, and sometimes is, defined by the detector. The system is designed in fact so that the sensitive area of the detector forms the exit field-stop.

It is doubtless possible to define a spectrometer in terms more general than this, but this definition will cover most of the types that will be discussed in this book. There are a few types which are relatively common and which fall outside this definition; in particular the 'slitless' spectrograph used by astronomers for obtaining the spectra of stars. In this case the entry field-stop is the star, and is already at infinity, so that the optical section in front of the disperser is superfluous.

Let us now consider the properties of an ideal spectrometer such as we have described. Apertures, fields and luminosity have already been discussed and defined. Resolving power has been mentioned but no criterion has yet been given to enable it to be described. In order to do this we must first define the 'instrumental profile' of a spectrometer.

---

**Introduction: The Ideal Spectrometer**

1-3 Instrumental Profile

If a spectrum consists of a single emission line at a wavelength \( \lambda_t \) with unit intensity, then, when it is examined with a spectrometer, the output from a perfect instrument would be \( S(\lambda) = \delta(\lambda - \lambda_t) \). No real instrument is perfect, of course, and the spectrum that is actually measured will be of the form:

\[
S(\lambda) = K \cdot F(\lambda - \lambda_t)
\]

where \( K \) is a constant which depends on the transmission and luminosity of the spectrometer, and the function \( F(\lambda) \) is called the instrumental profile.

It is normal to regard the instrumental profile function as having the same form at all wavelengths. This is not generally true, but the change of shape of a spectrum line due to instrumental defects is usually small over any practical range of wavelengths.

If the ideal spectrometer has no optical defects, then, when a finite slit width – or its equivalent – is allowed, the profile due to a monochromatic line will be triangular. This assumes that the dispersion is linear, and that the exit and entry field-stops match perfectly. Beyond the feet of the triangle the intensity should be zero everywhere. That is to say, no stray light passes through the instrument. This is something else that never happens in practice unfortunately, and the ratio of the areas of the spectral profile that lie inside and outside the feet of the ideal triangle provide a measure of the quality of the design. Stray light may be unimportant when emission spectra are to be studied, but in absorption spectrophotometry it is of the greatest importance that it should be kept to a minimum. Elaborate precautions must be taken to exclude it. A stray light ratio of 10,000 : 1 or more must be sought (and can be achieved) when high-quality absorption spectra are to be made.

A quick test of a spectrometer for the visible and nearby regions is to use it at its best resolving power to measure the solar spectrum. If Fraunhofer lines can be seen which are less than 0.1 Å wide, this indicates a high-quality instrument, so far as scattered light is concerned. This may well be too stringent a test for many purposes, and for many spectrometers.

The triangular shape of \( F(\lambda) \) will be obtained usually at comparatively modest resolutions. When approaching the theoretical resolving power of the system, diffraction effects will begin to modify the line shape, and the observed profile is not subject to a theoretical treatment, (There are exceptions to this, as for example, the Fabry-Pérot spectrometer.) A separate experiment is normally needed to determine the line profile.
for a given system. The emission from a laser or from a Mercury Hg lamp can be used, with the assumption that it is, to all intents and purposes, monochromatic.

In practice a spectrum may be described as having a power distribution with wavelength $S(\lambda)$. This function can be defined by saying that the surface brightness of the source is $S(\lambda)\, d\lambda$ watts/cm$^2$/steradian in the spectral interval $\lambda - (d\lambda/2)$ to $\lambda + (d\lambda/2)$. The power that is transmitted through the spectrometer at a given slit setting and hence a given instrumental profile can then be derived as follows:

Assuming that the instrumental profile $F(\lambda)$ is the same at all parts of the spectrum, we can write down an expression for the power that is apparently present at wavelength $\lambda_2$ due to an element of spectrum at $\lambda_1$. The intensity at $\lambda_1$ is $S(\lambda_1)$ per unit spectral interval, and we consider the elementary spectral range $\lambda_1$ to $\lambda_1 + d\lambda_1$ as a $\delta$-function, of area $S(\lambda_1)\, d\lambda_1$ at $\lambda_1$. The contribution that this element will make to the apparent intensity at $\lambda_2$ will be

$$dI(\lambda_2) = S(\lambda_1)\cdot F(\lambda_2 - \lambda_1)\, d\lambda_1$$

and the sum of all the contributions to the total at $\lambda_2$ will be

$$I(\lambda_2) = \int_{\lambda=0}^{\lambda=\infty} S(\lambda)\cdot F(\lambda_2 - \lambda)\, d\lambda$$

(1.1)

Thus the observed spectral power distribution is the convolution of the true spectral power distribution and the instrumental profile.

By using a computer to find the Fourier transforms of $I(\lambda)$ and $F(\lambda)$ and using the convolution theorem, it should be possible to recover the power distribution by division. In practice it frequently – indeed usually – happens that the instrumental profile, when transformed, contains zeros at regular intervals. The division is thus not valid at these points.

To take a specific example, suppose that the instrumental profile were of the sinc$^4$ form.* The Fourier transform of this is a triangle and this triangle multiplies the Fourier transform of the spectrum (what in multiplex Fourier spectrometry we would call the interferogram), and hence eliminates the higher frequency components. These high frequency components correspond to the fine structure in the spectrum, and this is irremediably lost: no amount of deconvoluting is going to recover it. If on the other hand the instrumental profile is of the Gaussian form, the Fourier transform is too. On performing the division, however, one finds it necessary to divide the experimental data points by

* The function sinc $(x)$ is defined as $(\sin x)/x$.
Design and Imaging

The apparent Luminosity of an object is a measure of the brightness of the object as it appears to the eye. It is defined as the ratio of the light emitted by the object to the light emitted by a standard object, such as a blackbody at a known temperature. The Luminosity of an object is a function of its temperature and its surface area. The higher the temperature of the object, the higher its Luminosity will be.

Figure 1.4(b). The effect of the instrumental profile on the recorded spectrum. (a) The Gaussian profile. In this case the Fourier transform has no zeros and the high frequency part of the spectrum transform is recorded, albeit with considerable attenuation. The fine structure can then be deduced by deconvoluting provided the signal/noise ratio in the recorded spectrum is high enough, and that the actual shape of the instrumental profile is well known.

1.4 Luminosity

The fact that there is an invariant property of an optical system, governing the amount of light that passes through it, can be derived from the Helmholtz–Lagrange theorem. Suppose that (figure 1.4) \( h_0 \) is the height of an object and that \( h_1, h_2, \ldots, h_i \ldots \) are the heights of successive images through the system. Let the marginal rays from the axis at the image point \( p_i \) to the next aperture form an angle \( \theta_i \), and let the refractive index of the medium where \( h_i \) was formed be \( \mu_i \). Then the quantity \( \theta_i h_i \mu_i \) is constant for all values of \( i \) through the system. This quantity is called the Helmholtz–Lagrange invariant. Occasionally it is found under other names, such as the 'Smith–Helmholtz invariant' or just the 'Helmholtz invariant'. It is important to remember that the invariant is two-dimensional. That is to say, if two perpendicular planes are drawn, both containing the axis of the system, there is a separate invariant for each. The two invariants are not necessarily the same. In a spectrometer with long narrow slits they are considerably different. The result of this is that if one tries to change the shape of the field for some reason (for example to couple a Fabry–Pérot étalon to a grating spectrometer) the shape of the aperture adjacent to the field will change also. An anamorphic lens will change a square object into an oblong image, but a circular cone of light from a point in the object is simultaneously changed to an elliptical cone. Since a spectrometer slit is commonly one thousand times as long as it is wide, an attempt to change its shape radially will run into trouble with the divergence of one of the ray fans. The limit comes when the angle \( \theta_i \) approaches \( 180^\circ \).

This invariant, considered in two planes, results in the theorem of conservation of étendue. It can be derived on purely geometrical arguments, but a more convincing and fundamental physical argument derives from the second law of thermodynamics. If étendue were not conserved in an otherwise perfect optical system, then with a black body as object, an image might be formed with a different surface brightness. If it were less than the object's surface brightness some light would have been lost in the system. If it were greater, the equilibrium temperature would be higher than that of the object. Neither possibility is
allowed: the first because of the assumption of an otherwise perfect system, and the other because the second law does not allow it.

1.5 Choice of a Spectrometer

The type of spectrometer that is chosen for a task depends on several factors, some of which are uncontrollable. One may be limited by what is available in the store room, or again by the price that one is prepared to pay. These factors are not within our scope. Technically the instrument should be adequate for its particular purpose, whether it is to measure a bright source at high resolution, or a very feeble source at a lower, but still fixed resolution. Sometimes the source is transient, and this in turn governs the total energy that will enter the instrument. 'Transient' may mean anything from a microsecond spark, to an hour or two of Aurora Borealis. Either way the choice of instrument is restricted to those which can collect enough information in the given time.

While we might use many pages of words to compare and contrast the performances of various types of spectrometer, we prefer to draw a simple graph which will present the two basic factors that govern the choice, namely resolving power and flux gathering power, or étendue. The other factor to be borne in mind while making a choice is simplicity and its concomitant, reliability. Nothing is simpler than a prism spectrometer. The net result is of course that most of the useful spectroscopy that can be done with such an instrument has already been done! The extra complication of a grating spectrometer with a motor to drive it and a photomultiplier to measure the signal is very small these days, and few pieces of electrical equipment are more reliable than a photomultiplier. We therefore put grating spectrometers near the top of the reliability list. Proceeding downwards we come to the Fabry–Pérot étalon, which has no moving parts (unless mechanical scanning is used) and then to the Michelson spectrometer. This is still not a 'routine' instrument that can be taken down off the shelf and plugged in. It may well become so as the usefulness of other instruments becomes more and more limited, so that people have to turn to it for solutions to the more interesting new problems. But at the moment (1968) there is no acknowledged 'best' way of making one and as much work is being done on it as with it.

For quick reference, therefore, we have constructed a diagram (figure 1.3). The boundaries are necessarily vague and sometimes represent the limits of convenience rather than the limits set by what is possible 'in principle'. It is assumed that the light source is 'extended', so that it can be made to fill both the field and the aperture with light. Stars generally do not count as extended sources, unless large telescopes are used where their 'seeing' disc is comparable in size with the slit width.

![Graph showing performance limits of various spectrometers.](image)
2.1 Introduction

Unless some quite extraordinary features are required in the optical performance, no particularly complicated optical design is required for the imaging system of a spectrometer. As a rule quite simple optical elements are sufficient to ensure adequate performance, and more complicated arrangements than simple single-element lenses or spherical or paraboloidal mirrors confer little advantage other than savings of space and weight.

For this reason there is no extensive inquiry here into the darker corners of geometrical optics, but an outline is given of the main arrangements that should be made, of the more familiar tricks of the trade and of the pitfalls that are to be avoided. Wherever possible, simple formulae and recipes are given for the design of optical components, based on the assumption that the reader is not an expert in geometrical optics, and furthermore, has no particular desire to be one.

A further simplification that can be made when considering optical spectrometers is that one of the conjugates is nearly always infinite. This results in simplified formulae for focal length and for the various aberrations that are encountered. We offer no excuse, then, for quoting these formulae at the appropriate point without, in most cases, any attempt at derivation. Both the derivation and the more general version of the formulae are to be found in any textbook on optical design.

2.2 Gaussian Optics

This is the name by which the simple theory of lenses is known. Aberrations are ignored, and the theory deals with refraction and reflection at surfaces, where angles of deviation are small and the approximation \( \sin \theta = \theta \) can be made. The result is the usual 'schoolboy' optics.

For refraction of a ray at a single surface we use the equation:

\[
\mu \left( \frac{1}{R} - \frac{1}{L} \right) = \mu' \left( \frac{1}{R} - \frac{1}{L'} \right) \quad (2.1a)
\]

where \( \mu, \mu' \) are the refractive indexes, \( L, L' \) are the object and image distances, and \( R \) is the radius of curvature of the surface. It is usually assumed that \( \mu = 1 \) for lenses in air. This is not always accurate enough and on a few occasions it is necessary to take into account the refraction of air.

Using this equation twice one may find a formula for the focal length of a thin lens in terms of its refractive index and the radii of curvature of its surfaces:

\[
\frac{1}{f} = (\mu - 1) \left( \frac{1}{R_1} - \frac{1}{R_2} \right) \quad (2.1b)
\]

The extension of this formula to a thick lens is:

\[
\frac{1}{f} = (\mu - 1) \left( \frac{1}{R_1} - \frac{1}{R_2} \right) + \frac{(\mu - 1)^2}{\mu} \frac{d}{R_1 R_2} \quad (2.2)
\]

Equation (2.1b) can be used to derive the resultant focal length of two thin lenses separated by a distance \( d \). The derivation is simple if one remembers that the angle of deviation of a ray on passing through a thin lens depends only on the height at which the ray meets the lens. For all angles of incidence the angle of deviation \( \theta \) is given by

\[
\theta = h/f.
\]

The resultant focal length is then

\[
\frac{1}{f} = \frac{1}{f_1} + \frac{1}{f_2} - \frac{d}{f_1 f_2} \quad (2.3)
\]

This focal length is to be measured from a point on the optic axis
known as a 'principal point'. For a ray coming from the left the principal point concerned is at a distance

\[ x_1 = \frac{d(f_1 - d)}{f_1 + f_2 - d} \]

to the right of the first lens.

There are two principal points for any lens system or for any lens of finite thickness. They are two of the six 'cardinal points' of the system. The other points are the focal points and the nodal points. Planes through these points, perpendicular to the optic axis, are the principal planes, focal planes and the nodal planes. They are well-defined only in the Gaussian approximation.

The cardinal points and planes are most easily defined with the aid of a figure (figure 2.3).

A ray, coming from the left, parallel to the axis and at a distance \( h \) from it, will appear to emerge from the system at the second principal plane at the same height \( h \) above the axis. It will then pass through the focal point \( \phi_2 \) of the system. Any ray coming from the left and meeting the first principal plane at a height \( h' \) will appear to emerge from the system at the same height \( h' \) at the second principal plane.

The nodal points \( n_1 \) and \( n_2 \) are specified by a ray, coming from the left and meeting the axis at the point \( n_1 \). It will emerge from the system at \( n_2 \) and parallel to its original direction. For lens systems in air the nodal points are coincident with the principal points. This is not so if the final refractive index is different from the initial one.

The positions of the principal points and focal points of a single thick lens are not readily available and are given here. Measurements are made from the actual lens surfaces. The distance of the first focal point \( \phi_1 \) from the first surface is

\[ z_1 = -f \left[ 1 + \left( \frac{\mu - 1}{\mu} \right) \frac{d}{R_2} \right] \]

and if \( z \) is positive it is to the right of the first surface. The distance of the second focal point \( \phi_2 \) from the second surface is

\[ z_2 = f \left[ 1 - \left( \frac{\mu - 1}{\mu} \right) \frac{d}{R_2} \right]. \]

The distance of the principal points from their respective surfaces is, similarly:

\[ z_1 = -f \left( \frac{\mu - 1}{\mu} \right) \frac{d}{R_2}; \quad z_2 = -f \left( \frac{\mu - 1}{\mu} \right) \frac{d}{R_1}. \]

The sign convention is the Cartesian convention, the lens pole at the origin, and radius positive if the surface is concave to the right.

The Gaussian theory may be said to apply only to a system of mirrors and lenses of infinitesimal aperture. In practice, of course, apertures are
finite and some further definitions are needed to specify an optical system.

The 'aperture' of a lens is self-explanatory and simply means the diameter. Occasionally it is taken to mean the transparent area.

The 'field' defines that part of the focal plane where an image is to be formed. Again it may be measured as an area but more usually it means the diameter of the patch on which the image is formed. This diameter may be measured linearly or, more usefully as a rule, as an angle.

The optical system will possess two 'stops'. These may be real or virtual and may be thought of as two holes in screens which have been placed perpendicular to the optic axis. The entrance pupil serves to define the diameter of the largest bundle of parallel rays that can pass through the system. It can itself be defined as the smallest actual aperture in the system, as seen from the entry side by an observer at infinity; that is to say, the image of that aperture produced by all the components that come before it. The exit-pupil is similarly defined by looking through the system from the other direction. The pupil may be different when looked at from different parts of the field. A pupil which is circular when seen on the axis will usually become elliptical when seen from the edge of the field. Again different real apertures, or portions of them, may play a part in defining the pupil in the off-axis case. When this happens there is said to be 'vignetting' in the system. This is something to be avoided whenever possible; it produces undesirable variations in the intensity of the image at the edges of the field. Occasionally it is impossible to avoid it.

The pupils may be inside or outside the actual physical system. In the case of a telescope, for example, the entrance pupil is the objective itself, and the exit pupil is the image of the objective produced by the compound eyepiece. It is usually situated a few millimetres outside the last component. It is the place at which the eye pupil should be placed in order to capture all the light that has passed through the objective and the rest of the telescope (figure 2.4).

In a spectrometer the exit pupil is the image of the grating or prism produced by the combination of the focusing lens and the field lens. It will be somewhere after the exit slit. It is the point at which the detector sensitive area should be placed. Normally in spectrometer design the dispersing element, being the most expensive item, will determine the aperture, and the images of it in the fore and aft optics form the entrance and exit pupils. Field lenses are normally placed at these pupils and are used to image the previous aperture on to the next aperture. They make use of the fact, derived from eqn. (2.3) that a lens placed at the focus of a previous lens has no effect on the focal length of the system. The field lens need not usually be of good optical quality. The field lens at the entrance pupil serves to image the source on to the dispersing element, and that at the exit pupil images the disperser on to the detector. Regarding the system as a whole, there is as a result of using these field lenses an image of the source on the detector. Because the quality need not be high, small focal ratios can be chosen, especially for the exit field lens, and this in turn means that a small detector can be used. Mirrors are not uncommon as exit field elements and can have focal ratios substantially less than unity. The focusing elements and the field lenses must be made large enough so that there is no vignetting. This normally means that the field lens must have a diameter equal to or very slightly greater than the slit length.

The chief-ray of a bundle is a somewhat ill-defined concept which is sometimes useful in tracing the bundle through a system. It may be thought of in the context of the definitions above, as that ray of a parallel bundle which passes through the centre of the entrance pupil. In the case of an oblique bundle it passes through the entrance pupil on the axis of the system.

Stop are often placed in a system in order to help control the aberrations. For example the iris diaphragm in a photographic lens is placed most carefully when the lens is designed. The most notable example of aberration control by a stop is to be found in the Schmidt
camera (figure 2.5). Coma can sometimes, as in this instance, be eliminated by the proper placing of a stop, or alternatively the distortion can be removed. There are of course other ways of handling these defects.

2.3 Aberrations

Only in the Gaussian approximation do ray bundles converge conveniently to points. In practice they do not and the images produced of point objects suffer from a variety of defects. These defects depend on

1. The aperture of the system. Such defects disappear for points on or off the optic axis when the aperture is infinitesimal.
2. The field. These disappear for a finite diameter ray bundle producing an image on the optic axis.
3. Combinations of the two which disappear only for an infinitesimal diameter bundle producing an image on the optic axis.

Only the gross aberrations will be considered here. These are the defects which appear when the Gaussian approximation is replaced by:

\[ \sin \theta = \theta - \theta^3/6. \]

The theory which results is called the 'Seidel theory' or alternatively the 'First-order' theory or occasionally the "Third-order" theory.

The aberrations can be classified in decreasing order of importance. They are:

1. **Spherical Aberration**
   This arises when a finite bundle of rays, all parallel to the optic axis, are brought to a focus by a lens or mirror system. The marginal rays meet on the axis at a different point from the Gaussian or paraxial focus. If the focal length for these marginal rays is less than the paraxial focal length the spherical aberration is said to be positive. Otherwise it is negative. At no point on the axis is a perfect image formed from a point source, but there is a position, close to the Gaussian focus, where a screen can be placed so that there is a 'circle of least confusion'. This position is for most purposes half-way between the Gaussian focus and the focus for marginal rays. The diameter of the circle is given below for several elementary cases:
   
   \[ \Delta = \begin{cases} 
   0 & \text{for a paraboloidal mirror} \\
   f/64F^3 & \text{for a spherical mirror} \\
   f/2R_1^2 & \text{for a single element lens}
   \end{cases} \]

   where \( D \) is the lens diameter and \( R_1 \) is the radius of curvature of the first surface. It can be shown that for an object at infinity the spherical aberration is least in a lens when it is approximately plano-convex (or plano-concave) with the curved side towards the infinite conjugate. This
Design the deviation of the rays equally between the two surfaces. Taking \( \mu = 3/2 \) the diameter of the circle of least confusion is:

\[
\Delta = \frac{7}{48} \frac{f}{F^2}
\]

A further useful criterion for assessing spherical aberration is the condition:

\[
\Delta < \left| \frac{8\lambda}{F^2} \right|
\]

where \( \Delta \) is the value of the expression in eqn. (2.4). If this criterion is satisfied, the image of a point object is indistinguishable from that produced by a system without spherical aberration. That is, the diameter is the result of diffraction and not of system defects [1].

Spherical aberration can be removed if non-spherical surfaces are allowed in the system. It can be removed on an ad hoc basis by 'figuring' one or more surfaces, using a Twyman-Green interferometer to assess the quality. Alternatively corrector plates can be made and introduced into the system. A corrector plate will possess enough spherical aberration to cancel that due to the rest of the system but at the same time it introduces no extra power. The Schmidt system uses such a figured plate, and the Maksutov system uses a meniscus which has virtually zero power, but negative spherical aberration equal and opposite to the positive aberration of the main spherical mirror.

(2) COMA

This is a defect which is most easily described by a diagram (figure 2.7). Consider an annular element of the lens, and pairs of rays from opposite ends of various diameters. Each pair will in general converge to a different point on the focal plane, those from AA' going to 'a', those from BB' going to 'b' and so on. The Gaussian focus is at 'g'. The whole annulus will contribute a small circle to the image, and different annuli will contribute circles of different diameters and in different positions. The result will appear as in figure 2.7.

Although the distribution of intensity is not uniform in the coma patch it is worth giving an expression for the height of the patch. This is:

\[
\Delta_c = \frac{3}{8} \frac{Dh}{f} \left\{ Z \left[ \left( \frac{\mu}{\mu - 1} \right)^3 \frac{1}{f^2} - \left( \frac{2\mu + 1}{\mu - 1} \right) \frac{1}{R_1 f} + \left( \frac{\mu + 2}{\mu} \right) \frac{1}{R_1} \right] \right.
\]

\[
+ \left( \frac{\mu}{\mu - 1} \right)^2 \frac{1}{f^2} - \left( \frac{\mu + 1}{\mu} \right) \frac{1}{R_1} \right\}
\]

As before, \( R_1 \) and \( f \) refer to the first radius and focal length of the lens (or mirror); \( h \) is the distance from the optic axis of the gaussian image point; \( D \) is the diameter of the entrance pupil and \( Z \) is the distance from the pole of the first surface to the point where the entrance pupil crosses the axis. Following the sign convention, \( Z \) is negative if the stop is to the left of the first surface (rays assumed to be coming from the left). If the stop is to the right of the first surface, then the distance is the virtual distance to the image of the actual physical stop, as required by the definition of the entrance pupil.

For a single spherical mirror eqn. (2.5) reduces to a simple form:

\[
\Delta_c = \frac{3}{16} \frac{h}{F^2} \left( 1 - \frac{Z}{2f} \right)
\]

where \( F \) is the focal ratio, which implies that the height of the coma patch is \( 3h/16F^2 \) when the stop is the rim of the mirror; and, more interestingly, that the coma is zero when the stop is set at a distance \( 2f \) from the pole. This latter case may be recognized as the Schmidt system. Since a spherical mirror has no particular axis of symmetry, a stop placed at the centre of curvature ensures that every parallel ray bundle reaching the mirror is symmetrical about its chief-ray, and this symmetry implies that there is no coma.

In the more general expression, a similar condition can be found for the removal of coma. It is generally known as 'Fraunhofer's condition'.

**Figure 2.7**: Positive coma from the annulus of a lens. Rays from the points marked with a capital arrive at the foci labelled with the corresponding lowercase letter.
Design of Optical Spectrometers

If we write:

\[ \Delta_e = \frac{3}{8} D f^2 (Z.A + B) \]  

(2.5b)

where \( A \) and \( B \) are the two expressions in eqn. (2.5), then evidently coma vanishes when \( Z = -B/A \). Thus with a lens, as with a Schmidt system, it is possible to find a position for the entrance pupil that will remove coma. In the particular case of a plano-convex lens arranged with the convex face towards the infinite conjugate (the arrangement for minimum spherical aberration) with a refractive index of 1.5 it is readily calculated that

\[ Z = f/7 \]

which is the position of the virtual stop, and, being positive, is to the right of the lens: and that the actual position of the physical stop is a distance \( f/8 \) to the right of the lens.

(3) ASTIGMATISM AND FIELD CURVATURE

It was assumed in the preceding section on coma that the rays from different parts of the annulus in figure 2.7 met on the Gaussian focal plane, albeit at different points. In the absence of coma all such pairs would have met at the same point. The defect of astigmatism describes the fact that these pairs may come to a focus before or after the Gaussian plane, and that the focal length may be different for different pairs.

Provided that the lens or mirror possesses an axis of symmetry the astigmatism must be zero on the axis, and so it is a defect of field. Two focal surfaces can be defined, as in figure 2.8, by the points where two perpendicular fans of marginal rays meet. They are known as the sagittal and tangential focal surfaces respectively. In spectrometer design it is often possible to accept the defect of astigmatism, since an image is to be formed of a long narrow slit, and we need then concern ourselves only with the tangential focus. It is tolerable if the points of the entry slit are imaged as short-line segments at the exit slit, provided that the line segments are parallel to each other and to the exit slit. This would not do in the case of spectrograph design, where image intensity is an important factor, and one cannot allow the image to spread longitudinally.

The defect of field curvature is treated simultaneously with astigmatism since the two are intimately connected. If astigmatism is absent, so that the tangential and sagittal foci coincide, an expression can be obtained for the radius of curvature of this field. In Seidel theory the field is spherical of radius \( R \), where:

\[ \frac{1}{R} = \sum_i \frac{1}{\mu_i f_i} = \sum P_i \]

the sum being over all the elements of the system. The separations of the components and the stop positions are not involved, only the refractive indexes and the focal lengths. This expression is well known to optical designers as the 'Petzval sum' for the system. Its remarkable simplicity makes the correction of the field quite easy. A 'field flattening' lens may be placed at the image point where it makes no contribution to the power of the system. Its focal length is chosen to make the Petzval sum equal to zero.

When astigmatism is present, the sagittal and tangential surfaces, which are both spherical, must be treated separately. By convention the astigmatism is said to be positive if the tangential focal length is less than
the sagittal focal length. The difference \( \Delta_s \) in the focal lengths is given, for a thin lens with the object at infinity, by:

\[
\Delta_s = \frac{h^3}{Z^2}(Z^2.A + 2.Z.B + 1)
\]

(2.6)

where, as before, \( Z \) is the distance to the pole of the lens from the entrance pupil, and \( A \) and \( B \) are the same expressions as in eqn. (2.5b).

If the expression in the brackets of eqn. (2.6) has real roots in \( Z \) there are two positions for the entrance pupil for which the astigmatism vanishes. As a rule neither of these coincides with the position of the coma removing pupil.

For a spherical mirror the astigmatism is given by:

\[
\Delta_s = \frac{R}{2}h^2(Z^2 + 1)
\]

(2.7)

Some care is needed in this case: if the concave mirror faces the incoming beam of light, the radius is negative (although by convention the focal length is positive). Hence \( \Delta_s \) is positive and the tangential focal length is less than the sagittal focal length.

If the pupil is placed at the centre of curvature, as in the Schmidt system, so that \( Z = R \), the astigmatism vanishes just like the coma.

To return to the general case of the astigmatic lens: three radii of curvature can be defined. They belong to:

(a) The tangential focal surface.
(b) The Petzval surface.
(c) The sagittal focal surface.

If the curvatures (the reciprocals of the radii) are denoted by \( C_t \), \( C_p \) and \( C_s \), it follows from eqn. (2.6) that

\[
C_t - C_s = \frac{1}{f}(Z^2.A + 2.Z.B + 1)
\]

The following relation exists between the Petzval sum and the astigmatism:

\[
\delta C = 2(C_s - \Sigma P_s)
\]

\( \Sigma P_s \) is the Petzval sum from which it follows that

\[
C_t = \Sigma P_s + 3/2.\delta C
\]

(2.8a)

\[
C_s = \Sigma P_s + 1/2.\delta C
\]

(2.8b)

Thus the distances between corresponding points of the (fictitious) Petzval surface and the two focal surfaces are in the ratio 3 : 1. Also it can be seen that the sequence of the surfaces is either T-S-P or P-S-T according to whether the astigmatism is positive or negative. For a thin lens with the entrance pupil at the lens aperture the equations become:

\[
C_t = \frac{1}{2f}(3\mu + 2)
\]

(2.9a)

\[
C_s = \frac{1}{2f}(-\mu + 2)
\]

(2.9b)

When a flat tangential field is needed and the system is astigmatic the field-flattening lens must be chosen, not to make the Petzval sum vanish, but to make it equal to \(-\frac{3}{2} . \delta C\).

Thus for a spherical mirror with the stop at the aperture, the value of \( \delta C \) is:

\[
-\frac{R}{f^3} = \frac{R}{f^3} + \frac{2}{f}
\]

And so a field-flattening lens to be placed at the focus of the mirror will have a focal length \( f' \), where \( f' = -f/2\mu \).

2.4 Aplanatic and Isoplanatic Systems

These are means of avoiding spherical aberration and coma in lens systems. They are of use chiefly when a small focal ratio is desired and are not of great deal of interest in the design of the main spectrometer optical system. There are occasions, however, when the very smallest detector area must be used, and it is then useful to be able to reduce the final image of the grating to the smallest possible size. By the law of conservation of étendue this entails a small focal ratio, and the aberrations are not confined to the Seidel region and can become gross unless care is taken in the design.

In the elementary case of a spherical surface separating two media of different refractive indexes the spherical aberration becomes zero in three cases:

(i) If the object is at the centre of curvature and no refraction takes place at the surface.
(ii) If the object is at the vertex of the surface.
(iii) If the object and image distances \( L \) and \( L' \) are connected with the radius of curvature by the relations:

\[
L = R(1 + \mu); \quad L' = R(1 + 1/\mu)
\]

It can be shown that in case (iii) the system is free from spherical aberration and coma of all orders. This is because the two conjugate points obey the Abbé sine condition, a subject which it is not necessary to pursue in this book. The two points are called the aplanatic points of
Design of Optical Spectrometers

28

The aplanatic points of a refracting surface.

Figure 2.9. The aplanatic points of a refracting surface.

the system. As case (i) is also a trivial case of an aplanatic pair of points, it is possible to make an aplanatic lens, using two aplanatic surfaces each either of case (i) or (iii). The four possibilities are then:

1. A thick lens with both surfaces concentric to the object. In this case the image coincides with the object and the magnification is +1.
2. A thick lens whose first surface is concentric with the object and whose second surface is aplanatic with the object point. Such a lens acts as a converging lens. Its radii are

\[ R_1 = L; \quad R_2 = \frac{\mu}{\mu + 1} (L - t) \]

where \( t \) is the thickness of the lens. Its focal length is

\[ f = - \frac{\mu^2}{(\mu - 1)(\mu L + t)} L \] (Note that \( L \) is negative.)

3. A thick lens whose first surface is aplanatic with respect to the object point, and whose second surface is concentric with the image produced by the first surface. This lens acts as a diverging lens. Its radii are

\[ R_1 = \frac{L}{\mu + 1}; \quad R_2 = \frac{L}{\mu} - t \]

and its focal length is

\[ f = \frac{1}{(\mu - 1)} \left[ \frac{L - \mu t}{L - t(\mu + 1)} \right] L \]

4. A thick lens, both surfaces of which are aplanatic. This lens produces a parallel shift of the rays emerging from the object point.

The four types are illustrated in figure 2.10.

In practice, cases (1) and (4) are trivial, and of the other two case (2) is the more important. The usual application is in high-power microscope objectives where the initial focal ratio is very small, and in

Figure 2.10. The four types of aplanatic lens. Only types (2) and (3) are significant as optical components, and of these type (2) is the more useful.
condensers for radiation detectors. This is the case of interest in spectrometer design.

In no case can an aplanatic lens produce a real image of a real object and so a sequence of such lenses must terminate in one non-aplanatic member. This last member will introduce coma and spherical aberration, unless the precaution is taken of placing its entrance pupil in the proper position (see eqn. (2.5b)). This removes the primary coma, and the resulting system (which still has spherical aberration) is called isoplanatic.

REFERENCES


3-1 Fourier Series

It may well be that the reader is already acquainted with the theory of Fourier transforms. If so, he will probably find that the formalism adopted in this chapter is slightly different from that found in most textbooks on Fourier theory. This is due to the fact that it is convenient to employ actual physical quantities such as wavelength, \( \lambda \); wave-number \( k = 1/\lambda \); and path changes or differences, usually denoted by \( \Delta \). The general expression for the amplitude of a plane monochromatic wavefront moving in the positive \( x \)-direction can be written:

\[
A(x,t) = A_0 \exp \left( 2\pi i(x - ct + \delta) \right)
\]

where \( \delta \) is a constant of integration called the ‘phase’ of the wave. Coordinates can usually be chosen to make \( \delta = 0 \), or to give it some other convenient value.

Thus it is found that the arguments of the sines, cosines and complex exponentials used in optical theory contain a factor \( 2\pi \) that is often omitted from the more mathematical treatments of Fourier theory. In consequence of this, the limits of integration are different, and the normalizing factors of \( 2\pi \) or \( \sqrt{2\pi} \) are missing.

The theory to be used here is set out briefly below.

A function \( F(\Delta) \) which is periodic, so that \( F(\Delta + \Delta_0) = F(\Delta) \) can be represented by the sum of an infinite series of sines and cosines with suitably chosen amplitudes and with periods \( \Delta_0, \Delta_0/2, \Delta_0/3 \), etc. Written formally the function is

\[
F(\Delta) = \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \cos \frac{2\pi n \Delta}{\Delta_0} + \sum_{n=1}^{\infty} b_n \sin \frac{2\pi n \Delta}{\Delta_0}.
\]
32 - Design of Optical Spectrometers

Fourier's theorem then permits us to recover the values of the amplitudes by:

\[ a_n = \frac{1}{\Delta_0} \int_{-\Delta/2}^{+\Delta/2} F(\Delta) \cos \frac{2\pi n \Delta}{\Delta_0} \, d\Delta \]
\[ b_n = \frac{1}{\Delta_0} \int_{-\Delta/2}^{+\Delta/2} F(\Delta) \sin \frac{2\pi n \Delta}{\Delta_0} \, d\Delta \]

3.2 Fourier Transforms

The transition from a periodic function to a non-periodic one can be imagined by allowing the period to tend to infinity, so that sines and cosines are added with only infinitesimal differences between their periods and with only infinitesimal amplitudes. Thus if \( F(\Delta) \) is symmetrical about the origin, so that \( F(\Delta) = F(-\Delta) \), the function can be constructed entirely out of cosines:

\[ F(\Delta) = \int_{-\infty}^{+\infty} a(\nu) \cos 2\pi \nu \Delta \, d\nu \]

while if it is antisymmetrical, so that \( F(\Delta) = -F(-\Delta) \), it can be constructed out of sines:

\[ F(\Delta) = \int_{-\infty}^{+\infty} b(\nu) \sin 2\pi \nu \Delta \, d\nu \]

where \( a(\nu) \, d\nu \) and \( b(\nu) \, d\nu \) are the infinitesimal amplitudes of the cosine and sine components whose period is \( 1/\nu \).

By a suitable combination of sines and cosines a function without symmetry can be synthesized:

\[ F(\Delta) = \int_{-\infty}^{+\infty} [a(\nu) \cos 2\pi \nu \Delta + b(\nu) \sin 2\pi \nu \Delta] \, d\nu \]

and as before, the Fourier inversion theorem permits us to recover the functions \( a(\nu) \) and \( b(\nu) \):

\[ a(\nu) = \int_{-\infty}^{+\infty} F(\Delta) \cos 2\pi \nu \Delta \, d\Delta, \quad b(\nu) = \int_{-\infty}^{+\infty} F(\Delta) \sin 2\pi \nu \Delta \, d\Delta \]

and, if desired, the functions \( a(\nu) \) and \( b(\nu) \) can be combined into a complex function \( c(\nu) \):

\[ c(\nu) = a(\nu) + i \cdot b(\nu) = \int_{-\infty}^{+\infty} F(\Delta) \cdot \exp(2\pi i \nu \Delta) \, d\Delta. \]

There is then a fundamental reciprocity between \( c(\nu) \) and \( F(\Delta) \):

\[ \text{if} \quad c(\nu) = \int_{-\infty}^{+\infty} F(\Delta) \cdot \exp(2\pi i \nu \Delta) \, d\Delta, \]

then

\[ F(\Delta) = \int_{-\infty}^{+\infty} c(\nu) \cdot \exp(-2\pi i \nu \Delta) \, d\nu \]

and \( F(\Delta) \) and \( c(\nu) \) are said to be the Fourier transforms of each other. Alternatively they are called a 'Fourier pair'.

Further theorems follow immediately:

If the Fourier transform of \( F(\Delta) \) is written \( \tilde{F}(\nu) \), the reciprocity is expressed as:

\[ F(\Delta) = \tilde{F}(\nu) \Rightarrow \tilde{F}(\Delta) = F(\Delta) \]

If \( F_1(\Delta) \) and \( F_2(\Delta) \) are two functions, then:

(i) \( \tilde{F}_1(\nu) + \tilde{F}_2(\nu) = [F_1(\nu) + F_2(\nu)] \)

(ii) \( \int_{-\infty}^{+\infty} F_1(\Delta) \cdot F_2(\Delta) \, d\Delta = \int_{-\infty}^{+\infty} [\tilde{F}_1(\nu) \cdot \tilde{F}_2(\nu)] \, d\nu \)

and a special case of this is Parseval's theorem:

(iii) \( \int_{-\infty}^{+\infty} |F(\Delta)|^2 \, d\Delta = \int_{-\infty}^{+\infty} |\tilde{F}(\nu)|^2 \, d\nu \)

(iv) if \( \int_{-\infty}^{+\infty} F(\Delta) \cdot \exp(2\pi i \nu \Delta) \, d\Delta = \tilde{F}(\nu) \)

and

\[ \int_{-\infty}^{+\infty} F(\Delta + \Delta_0) \cdot \exp(2\pi i \nu \Delta) \, d\Delta = \tilde{F}(\nu) \]

then

\[ F_0(\nu) = \tilde{F}(\nu) \cdot \exp(2\pi i \nu \Delta_0) \]

which is called the 'shift theorem'.

Elementary Fourier Theory · 33
Combining (ii) and (iv) we obtain the Convolution theorem:

\[ F_3(\Delta) = \int_{-\infty}^{\infty} F_1(\Delta) \cdot F_2(\Delta') \cdot d\Delta \]

then

\[ F_3(v) = F_1(v) \cdot F_2(v). \]

### 3:3 Convolutions

The process of convolution of two functions is one which occurs not only in optical theory but throughout physical science. It is of particular importance in the theory of optical instruments as well as in the theory of spectra and the theory of communications, and it is best explained by means of a simple illustration.

A spectrometer scanning a monochromatic wave train of wave-number \(v_0\) records it so that it appears to contain a finite range of wave-numbers. That is to say, a spectrum which is really represented by \(I(v) = \delta(v - v_0)\) is recorded incorrectly by the instrument as

\[ I(v) = A(v - v_0). \]

**Figure 3.1(a)**. Convolutions in spectrometry. The instrument changes a purely monochromatic line into one with a profile \(A(v)\) centred at \(v_0\).

The function \(A(v)\) is called the instrumental profile of the spectrometer. (In electrical theory it would be called the 'impulse response' of the filter.)

If now the spectrometer is used to scan a complete spectrum represented by \(S(v)\), it will treat each elementary wave-number range from \((v_1 + \frac{1}{2} \delta v_1)\) to \((v_1 - \frac{1}{2} \delta v_1)\) as a \(\delta\)-function of area \(S(v_1) \cdot d\nu_1\), and will record it as \(S(v_1) \cdot A(v_2 - v_1) \cdot dv_1\), so that the intensity recorded at wave-number \(v_2\) due to light in the spectral element at \(v_1\) is:

\[ \frac{d}{dv_2} = S(v_1) \cdot A(v_2 - v_1) \cdot dv_1 \]

and the total amount of light recorded at \(v_2\) due to all such elements in the spectrum will be

\[ J(v_2) = \int_{-\infty}^{\infty} S(v_1) \cdot A(v_2 - v_1) \cdot dv_1 \]

and \(J(v)\) is said to be the convolution of \(S(v)\) and \(A(v)\). The normal notation for the convolution process is:

\[ J(v) = S(v) \ast A(v). \]

**Figure 3.1(b)**. The element of spectrum at \(v_1\) contributes an element at \(v_2\). The total contribution at \(v_2\) from all elements is the convolution of the true spectrum and the instrumental profile.
The process is symmetrical in the two components since the integral has infinite limits.

There may be several causes contributing to the change in profile of a spectrum line. Some of them may be physical, such as Doppler broadening in the emission from a hot gas, and others may be instrumental, not only in the spectrometer, but in other parts of the optical system. Each such modification may be represented by a function which convolution the original spectral intensity function \( S(v) \), and the net result can be represented by a multiple convolution:

\[
G(v_n) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} S(v_1) \cdot A(v_2 - v_1) \cdot B(v_3 - v_2) \cdot C(v_4 - v_3) \cdots K(v_n - v_{n-1}) \, dv_1 \, dv_2 \cdots dv_{n-1}
\]

which can be written conventionally as:

\[
G(v) = S(v) \ast A(v) \ast B(v) \cdots \ast K(v).
\]

And the convolution theorem holds in the extended form:

\[
\bar{G}(\Delta) = S(\Delta) \cdot \bar{A}(\Delta) \cdot \bar{B}(\Delta) \cdots \bar{K}(\Delta).
\]

### 3.4 The Algebra of Convolutions

We have already the equivalences:

\[
\phi(v) = F(v) \ast I(v) \quad \text{then} \quad \bar{\phi}(\Delta) = \bar{F}(\Delta) \ast \bar{I}(\Delta)
\]

and it is easy to show other similar equivalences such as:

\[
\phi(v) = [F(v) \ast I(v)] \ast [J(v) \ast K(v)]
\]

then

\[
\bar{\phi}(\Delta) = \bar{[F(\Delta) \ast I(\Delta)]} \ast \bar{[J(\Delta) \ast K(\Delta)]}.
\]

It is, however, important to realize that the associative law does not hold in a mixture of convolutions and multiplications. Thus, for example

\[
[F(v) \ast I(v)] \ast J(v) \neq F(v) \ast [I(v) \ast J(v)].
\]

### 3.5 Physical Applications of Convolutions

Convolutions are of widespread interest in spectrometry, and indeed everywhere where Fourier theory plays an important part in the experimental theory. The obvious application is as aid to the recovery of the true spectrum from the observed spectrum. If a monochromatic or nearly monochromatic line is passed through the spectrometer, the instrumental profile can be determined, and if the Fourier transform of the observed spectrum is divided, point by point by the Fourier transform of the instrumental profile, the quotient, when re-transformed should give the true spectral profile. This does not work properly in practice for several reasons. The most important of these (and the most obvious) is that the Fourier transform of the instrumental profile will contain regions of low value or, quite likely, zeros. At such places the quotient is either indeterminate, or very poorly known, because of the inevitable noise that enters any experimental measurement. As a rule only an approximate correction can be made. A more important application is in ensuring that the instrumental profile is of a convenient shape, and that it does not, for instance, contain large side-lobes such as are found in the profile of a Michelson spectrometer. Optical masks can be used in appropriate places to modify the instrumental profile, and these may prevent the obscuration of weak satellite lines by such side-lobes. Such a process is called *apodizing*. In Fourier spectrometry a similar effect is obtained by the use of \('truncating functions\' which multiply the interferogram, and whose Fourier transforms convolute the spectrum that is obtained. The optical mask acts as a spatial frequency filter while the truncating function acts as an actual frequency filter.

### 3.6 Fourier Pairs of Interest in Spectrometry

Some functions turn up frequently when considering various aspects of spectrometer design and performance. The more interesting ones are described below. In every case the transform is performed by elementary methods, either by direct integration or by using one of the theorems mentioned earlier. The conjugate variables are \( v \) and \( \Delta \).

1. **The 'Top-hat' Function** (in some American literature the 'Box-car' function)

   This function will be denoted here by the symbol \( \Theta(a, D) \), where \( a \) is the coordinate in \( \Delta \)-space of the centre, and \( 2D \) is the width. The function

   ![Figure 3.2. The 'top-hat' function \( \Theta(a,D) \)]
Design of Optical Spectrometers

is defined by the conditions:
\[
\Theta(a,D) = \begin{cases} 0, & -\infty < \Delta < (a - D) \\ \frac{1}{2}D, & (a - D) < \Delta < (a + D) \\ 0, & (a + D) < \Delta < \infty \end{cases}
\]

The Fourier pair of \( \Theta(0,D) \) is \( \frac{\sin (2\pi \Delta)}{2\pi \Delta} \).

This function is known as a ‘sinc’ function and written sinc \((2\pi \Delta)\).

The Fourier pair of \( \Theta(a,D) \) then follows from the shift theorem (see eqn. (iv), section 3.2). It is sinc \((2\pi \Delta)\), cos \((2\pi \alpha)\).

(2) THE TRIANGLE FUNCTION

This will be denoted by \( \text{Tr}(a,D) \) with \( a \) and \( D \) defined as above.

It is defined by the conditions:
\[
\text{Tr}(a,D) = \begin{cases} 0, & -\infty < \Delta < a - 2D \\ \left(1 - \left|\frac{a - \Delta}{2D}\right|\right), & a - 2D < \Delta < a + 2D \\ 0, & a + 2D < \Delta < \infty \end{cases}
\]

The Fourier pair of \( \text{Tr}(a,D) \) is \( D \cdot \text{sinc}^2(2\pi \Delta) \cdot \cos(2\pi \alpha) \).

![Figure 3.3 The triangle function \( \text{Tr}(a,D) \)](image)

A moment’s reflection will show that the triangle function is the convolution of a top-hat function with itself, and the convolution theorem then gives the Fourier pair immediately. This function has played some part in the theory of apodization. Light from a point source at infinity, brought to a focus by a lens with a rectangular aperture, will show a spatial distribution which is the convolution of a sinc\(^2\) function with a delta function. This distribution has side-lobes, as is well known from the theory of the plane diffraction grating. All this follows from the basic theorem of Fraunhofer diffraction, that the amplitude in the image plane is the Fourier transform of the amplitude distribution in the diffracting plane. If now the lens aperture is covered with a mask which changes the transmission from a top-hat curve to a triangle curve, the image of the point source will change from a sinc\(^2\) distribution to a sinc\(^4\) distribution. The width of this sinc\(^4\) curve will be greater than that of the original sinc\(^2\) curve, but the side-lobes will be suppressed. In practice the mask over the lens may well be a diamond-shaped aperture. This particular technique is not applied, as it might be, to plane-grating spectrometers, since it is found in practice that the side-lobes due to grating ghosts are much larger and more troublesome than these secondary diffraction maxima. The apodizing process is not worth while.

This is not the case, however, in Fourier spectrometry. The interferogram is multiplied by a top-hat function, so that the spectrum that is recovered is convoluted with a sinc-function. The side-lobes in this case are very large and negative, and it is essential to use some form of apodizing. One possibility that has been used is to multiply the interferogram by a triangle function, so that the spectrum is convoluted with a sinc\(^4\) curve. This is not the only possibility however, and other curves that have been used to multiply the interferogram include such things as \( \frac{1}{2}(1 + \cos x) \) and \( 0.54 + 0.46 \cos x \). All these result in different curves that convolute the spectrum. They are discussed further in the chapter on multiplex spectrometry.

(3) THE ‘GAUSSIAN’ CURVE

This has the curious property, which it shares with all the Hermite polynomials, of having a Fourier pair of the same form, i.e., another Gaussian curve. In \( \Delta \)-space and \( \nu \)-space as before, the pair can be defined as
\[
\text{erf}(\Delta, \alpha) = \frac{1}{\sqrt{\pi}} \exp(-\Delta^2/\alpha^2) \quad \text{and} \quad \text{erf}(\nu, \beta) = \frac{1}{\pi} \exp(-\nu^2/\beta^2)
\]

where \( \beta = 1/\pi \alpha \).

From this it can readily be deduced that the convolution of two Gaussian curves yields another Gaussian curve with a half-intensity width which is the diaphragm sum of the component widths.

(4) THE ‘LORENTZ’ CURVE

This is the profile of a spectrum line which results from the ‘natural’ decay of an excited state. In \( \nu \)-space it can be defined by
\[
L(\nu, \nu_0) = \frac{1}{\left[\nu^2 + (\nu_0^2/2)\right]}
\]

where \( \nu_0 \) is the half-intensity width of the line. Its Fourier pair is
\[
\overline{L}(\Delta) = \frac{2\pi}{\nu_0} \exp(-\pi \nu_0 |\Delta|).
\]
40 - **Design of Optical Spectrometers**

From this it follows that when two Lorentz curves are convoluted the resulting half-width is the sum of the two component half-widths.

In practice the observed profile of a spectrum line, excluding effects caused by the instrument, is often the convolution of a Gaussian curve and a Lorentz curve. This yields a curve known as a 'Voigt profile' [1]. Tables of Voigt profiles with various proportions of Gaussian and Lorentz curves in the mixture have been published by Posener [2] and are useful for fitting to experimental graphs of spectral lines to distinguish, for example, between natural line broadening and Doppler broadening due to thermal motions of the emitting particles.

**REFERENCES**


**BIBLIOGRAPHY**


---

**Prism Spectrometers**

4.1 **Introduction**

The earliest scientific record of the use of a prism as a dispersing element appears to be that of Newton. In his experiments the prism was used in convergent light, with the sun acting both as source and as entry slit. A small hole in a blind acted as collimator and aperture stop. The final images of the sun in different wavelengths appeared on the wall of Newton's chamber. For obvious reasons the resolving power was not very high.

The earliest record that we have been able to trace of the use of a slit and collimating lens to achieve a respectable resolving power is that of the Scotsman Wollaston, who discovered with his device the 'Fraunhofer lines' in the spectrum of the sun. It was Fraunhofer who first studied these lines and gave a correct conjecture of their cause. Hence, quite rightly, they bear his name. It is to Fraunhofer that we owe the design of the classical prism spectroscope that we find in school teaching laboratories. For serious work at high resolutions this design is largely superseded by the plane-grating spectrometer, but for certain applications it possesses some advantages over grating instruments. The absence of overlapping spectral orders and the freedom from 'ghosts' make it suitable for Raman spectroscopy, where a search must be made for very faint emission lines very close to a strong parent line. With careful selection of the prism material (to be free from striae) and careful polishing of the refracting surfaces to be free from scratches and other blemishes, the stray light level can be brought down to a lower level than is possible with the best grating spectrometer, and hence there is still an application in absorption spectrometry. However, the luminosity-resolution (LR) product of even a large prism spectrometer is far below that of a grating spectrometer of modest size, and the highest resolution that is possible is an order of magnitude less than that which is possible from a grating spectrometer.
The dispersing element usually consists of one or more prisms each with a 60° apex angle, or a 30° prism with a plane reflector behind it. Variants such as the Pellin–Broca prism are really of this form, although apparently different. There is no particular magic about the choice of 60° for the prism apex, but it is not possible to exceed this angle greatly since the shortest wavelength that can be refracted corresponds to the condition:

$$\sin \left( \frac{\theta}{2} \right) = \frac{1}{\mu}$$

where \(\theta\) is the apex angle of the prism and \(\mu\) is the refractive index associated with this shortest wavelength. In the case of the 60° prism this refractive index is 2, which includes the region of most available prism materials. On the other hand, the angle of incidence of the radiation on to the prism surface should be kept as small as possible to avoid reflection losses, which become prohibitive at large angles of incidence. These reflection losses depend on the direction of polarization of the incident beam of radiation, with the result that the transmitted beam is partly polarized even if the incident radiation is not. Thus, for some specific applications it may well be possible to use a prism with a refracting angle of less than 60°, for example as an order sorter to a grating spectrometer.

If reflection losses are omitted for the moment the criterion of performance of a prism spectrometer is usually the LR product. The resolving power is obtained from the equation:

$$R = -b \cdot \frac{\delta \mu}{\delta \lambda}$$  \hspace{1cm} (4.1)

where \(b\) is the length of the base of the prism, and \(\frac{\delta \mu}{\delta \lambda}\) is the dispersion. The proof of this equation is to be found in any elementary text. The equation is remarkable for the quantities that it does not contain. \(R\) does not depend on the prism apex angle, and so not on the length of the refracting side either. Neither does the refractive index appear but only its rate of change with wavelength.

The luminosity of a prism spectrometer is obtained by multiplying the exit slit area by the solid angle subtended by the refracting face of the prism at the exit slit. If the length of the exit slit is taken as \(1/50\)th of the focal length (the usual criterion) of the focusing lens, the luminosity comes to:

$$L = \frac{b \cdot T \cdot \delta \mu}{50 \cdot \frac{\delta \lambda}{\Delta \lambda}}$$

where \(T\) is the thickness of the prism parallel to the refracting edge (or, if you like, the 'height' of the prism).

Both the resolving power and the luminosity quoted here are at the maximum possible resolving power that can be achieved, limited by diffraction in the system.

The LR product then comes to:

$$E = R \cdot L = -\frac{b \cdot T \cdot \lambda \cdot \delta \mu}{50 \cdot \frac{\delta \lambda}{\Delta \lambda}}$$  \hspace{1cm} (4.2)

We may take it that the refractive index varies with wavelength according to:

$$\mu = \left( a + \frac{b' \cdot \lambda^2}{\lambda^2} \right)$$

where \(a\) and \(b'\) are the Cauchy constants of the prism material, and then the LR product comes to:

$$E = \frac{T \cdot b \cdot b'}{25 \cdot \lambda^2}$$

Thus the efficiency falls rapidly at longer wavelengths implying that only low resolutions and low luminosities are available in the infrared.

### 4.2 The Focal Curve Theorem

This theorem states that if the deviation is small so that in the elementary formula:

$$\mu = \frac{\sin \left( \frac{a + \lambda}{2} \right)}{\sin \left( \frac{a}{2} \right)}$$

where \(a\) is the angle of incidence and \(a'\) is the angle of deviation. From the law of refraction:

$$\frac{\sin \theta}{\sin \theta'} = \frac{\sin \left( \frac{a + \lambda}{2} \right)}{\sin \left( \frac{a}{2} \right)}$$

we see that the ratio of the sine of the deviation to the sine of the angle of incidence is a constant for a given material. This ratio is important in the design of spectrographs.

The following diagrams illustrate the general nature of the focal curve for different numbers of prisms. The curves are plotted with the deviation in microradians on the horizontal axis and with the height of the prism on the vertical axis.
The classical prism spectroscopy arrangement, drawn to illustrate the focal curve theorem. This theorem applies only in the small deviation approximation. In practice the focal length will be slightly less than the theorem indicates for shorter wavelengths and the focal surface will curve gently towards the prism.

The problem to be studied and the resolving power that is needed. The size of the prism can be calculated immediately from the formula given for resolving power, and from published tables of partial dispersions. The focal length of the collimating lens and the focusing lens are determined partly by the size of the prism (each must be large enough in diameter so that the prism and not either of the lenses is the aperture stop), and in addition by the photographic material that is to be used. The dispersion should be such, at the focal surface, that one resolved element covers a width greater than one grain of the emulsion. Usually it should cover about five emulsion grains. There is a balance to be struck here between large linear dispersions (needed if the emulsion is fast, and hence coarse grained) and the large focal ratio, and the consequent low photographic speed that a large focal ratio implies. It turns out in practice that there is no apparent advantage in using fast or slow emulsions, since with a slow emulsion a smaller dispersion and hence a smaller focal ratio is possible. This is based on the observed fact that a grain in an emulsion, irrespective of size, requires about 1,000 photons to make it developable. Hence if the focal length is doubled, one can afford to use grains which are twice the linear dimensions, and four times the area, exactly balancing the fourfold reduction in the rate of photon arrival at the plate which the larger focal ratio implies. In general, then, prism spectrographs will be made as short as possible, with low focal ratios, the limit being set by geometrical aberrations which will interfere with the resolving power and will need expensive elements to correct them.

The point might be made here that it is important to distinguish between high resolution and high dispersion. A high dispersion means only that the instrument may be more crudely made, whereas it presents no advantage at all so far as resolving power or luminosity are concerned. Ultimately the performance will depend on (a) the base length and height of the prism, and (b) the ratio of speed to grain size of the photographic emulsion.

4.3 The Prism Spectrometer

The remarks above have been concerned with prism spectrographs, where the detector is a photographic plate, or, possibly, an image tube. It is possible however, to make a prism spectrometer, or, more accurately, a prism monochromator. It is feasible to convert a spectrograph to a monochromator by mounting a photoelectric device on a carriage which is then carried along the focal surface. Since the criteria for the design
of a spectrometer are quite different from those for a spectrograph, this
is not usually a satisfactory way of making a monochromator.

One normally thinks of a monochromator as a 'black box' with a hole
for the radiation to enter, a hole for the selected wavelength to emerge
and a knob to turn to select the chosen wavelength.

Such an arrangement can be achieved in several different ways, but
one of the most satisfactory, and for many years one of the most popular
ways, was by means of the Pellin-Broca prism, as used in the Hilger
wavelength spectrometer. The prism is the normal 60° prism, but it is

![Diagram of the Hilger wavelength spectrometer](image)

**Figure 4.3.** The Hilger wavelength spectrometer. (Schematic, not to scale.)

in effect split into two parts with a reflecting element in between. At any
one position of the prism there is just one wavelength that is refracted
and reflected through 90°, and consequently the collimator and camera
can be permanently mounted. Different wavelengths appear at the exit
slit as the prism is rotated (by means of a tangent screw, for example)
and this screw can be calibrated directly in terms of wavelength. The
calibration will need to be non-linear, to match the change of refractive
index of the prism material.

Alternatively a constant-deviation spectrometer can be made using
a normal prism. The prism is allowed to rotate, and carries with it a
plane reflecting mirror. Collimation and focusing can be by lenses as in

![Diagram of the Wadsworth constant-deviation mounting](image)

**Figure 4.4(a).** The Wadsworth constant-deviation mounting (I).

the Wadsworth mounting (not to be confused with the Wadsworth
stigmatic mounting for concave gratings), or an Ebert-like system can be
used, as described by Schick and Winter [1]. By proper choice of the
positions of the prism and the plane mirror, a very compact arrangement
can be made. If the angle between the base of the prism and the mirror
surface is \( \phi \) the deviation from the incoming to the outgoing beam is
\( (\pi - 2\phi) \). As usual with Ebert-type mountings, the mirrors should be
spherical rather than paraboloidal.

**The Littrow Mounting** has been popular for many years as a
spectrograph mounting (figure 4.5). For use as a spectrometer, either
the detector must be moved along the focal plane, or the usual lens focusing
element must be replaced by a mirror. If high performance is needed,
it may be necessary to use an off-axis paraboloidal mirror.
THE EBERT MOUNTING is probably superior (figure 4.6). The mounting is described in detail in the chapter on the plane-grating spectrometer. The grating is replaced by a prism and, provided short straight slits are used, the same focusing properties hold. Scanning is done by rotating the prism. If a single-mirror Ebert system is used, the centre of the mirror must be obscured, as with grating spectrometers of this type, to prevent transmission of the incoming beam via the centre of the mirror to the exit slit.

MULTIPLE PRISM SPECTROMETERS
Greater resolving power can be obtained by using more than one prism. If a train of $n$ prisms is used, the net resolving power comes to

$$R = \sum_{i=1}^{n} \frac{b_i \delta \mu}{\delta \lambda}.$$
the 'Féry circle'. Like the Rowland grating the system is astigmatic and no use can be made of a Hartmann diaphragm. It is most unlikely that anyone will make use of the Féry prism as a spectrometer mounting, but if so, then arrangements similar to those for the concave grating are indicated.

REFERENCES


5-1 The Grating Equation

The characteristics of a plane diffraction grating can be most easily illustrated by examining the simple case of a plane transmission grating, where there are alternate opaque and transparent strips, the width of each opaque strip being b and the width of each transparent strip, a.

Consider monochromatic light incident normally to the plane of the grating, and a diffracted beam at an angle of diffraction \( \theta \). The intensity in the diffracted beam is given by the equation

\[
I(\theta) = \frac{I(0)}{n^2} \left( \frac{\sin nku}{\sin ku} \right)^2 \left( \frac{\sin \theta}{\theta} \right)^2
\]

(5.1)

where \( u = \left(2\pi/\lambda\right) \frac{1}{\lambda} \sin \theta, k = (a + b)/a \) and \( n \) is the total number of transmitting lines in the grating.

This equation is to be found in all textbooks on physical optics, and is easily derived by remembering that if a converging lens follows the grating, the spatial amplitude distribution in the focal plane of the lens is the Fourier cosine transform of the amplitude distribution in the plane of the grating. The formula generalizes immediately to the case of the reflecting grating and to the case of other than normal incidence.

The first term in the intensity distribution will be recognized as that due to interference by \( n \) equally spaced point sources, and the second term is that of the diffraction of a plane wavefront by any one of the slits of the grating. The curve of \( I(u) \) against \( u \) is illustrated in figure 5.1. The maxima and minima of the function \( I(u) \) are of interest. Minima with \( I(u) = 0 \) occur when \( nku = 0 \), provided that \( \sin ku \neq 0 \) simultaneously.

This is when

\[
u = \frac{\pi}{nk} \frac{2\pi}{nk} \frac{3\pi}{nk} \cdots \frac{(n-2)\pi}{nk} \frac{(n-1)\pi}{nk} \frac{(n+1)\pi}{nk} \cdots
\]
but if $u = \pi/k, 2\pi/k \ldots mn/k \ldots$ then $(\sin nk \sin ku)/k \cos ku$ is indeterminate and de l'Hôpital's rule is applied:

$$\lim_{u \to mn} \frac{nk \cos nk}{k \cos ku} = n; \quad I(u) = I(0), \frac{(\sin \frac{u}{k})^2}{u^2}$$

so that there is a principal maximum of the function $I(u)$ at each of these points. These values of $u$ determine the directions of emergence of the diffracted beams, and $m$ is called the order of diffraction.

$I(u)$ can still be zero if the second term should happen to be zero as well, that is, if $u = s\pi$, i.e., if $m$ has a factor $k$.

Thus if $k$, the 'mark-space' ratio of the grating, is an integer, certain of the principal maxima will be suppressed. If $(a + b)/\pi = 3$, say, then the principal maxima for which $m = 3, 6, 9 \ldots$ will be missing. In the extreme case, for which $k = 1$, all the principal maxima are missing except for $m = 0$. This is when the grating has become a plane piece of glass.

5.2 Resolving Power

Consider the function $I(u)$ close to a principal maximum, and ignore for the moment the slowly varying $(\sin u/k)^2$ term. The criterion for resolution is that the principal maximum for one wavelength lies just on top of the first minimum of the closest wavelength that can just be resolved. At the point $A$ on the diagram (figure 5.2) we have, for wavelength $\lambda$

$$u = mn/k = \frac{\pi}{\lambda} a \sin \theta$$

and for the adjacent wavelength

$$u = (mn - 1)\frac{\pi}{\lambda} = \frac{\pi}{\lambda} a \sin \theta.$$

Dividing these two to eliminate $\sin \theta$

$$\frac{\lambda + \Delta \lambda}{\lambda} = \frac{mn - n}{n} \frac{nk}{k} (mn - 1) = \frac{mn}{mn - 1} = 1 + \frac{1}{R}$$

where

$$R = \frac{\lambda}{\Delta \lambda}$$

whence

$$R = m \cdot n.$$

so that the resolving power is the product of $n$, the total number of lines on the grating, and $m$ the order of diffraction.
Then, using \( I = (\sin n\kappa /\sin \kappa)^2 \); \( \sin n\kappa = \sin (mn + 3/2)\pi \);
\( \sin \kappa = \sin (m\pi + 3\pi/2n) \simeq 3\pi/2n \), and so \( I(u) = I(0).4/9\pi^2 \).
Thus the height of each first subsidiary maximum is about 4 per cent of the height of its adjacent principal maximum.

It is possible in principle to suppress these weak satellite lines by the process of apodizing (see section 3.6). In practice this would mean that the grating was covered by a suitably shaped mask, so that the full length of the middle ruling was exposed and only a fraction of each of the end rulings. Figure 5.3 (a–c) illustrates some typical apodizing masks. The

![Figure 5.3](image_url)

**Figure 5.3.** Typical apodizing masks. The letter below each refers to the function with which each spectrum line is convoluted. In the vicinity of a single line the function

\[
\frac{\sin(n\kappa)}{\sin(\kappa)}
\]

can be approximated by \( \text{sinc}(n\kappa) \).

(a) the 'zero-order' mask. Each line is convoluted with \( \text{sinc}^4(n\kappa) \), i.e. the usual grating equation gives the line profile. The \( \text{sinc}^4 \) term is omitted as it has no relevance when discussing the vicinity of a line.

(b) The 'first-order' mask. The convoluting function is \( \text{sinc}^4(n\kappa) \).

(c) The 'second-order' mask. The profile is of the form \( A + B \cdot \cos \theta \) and the function which convolutes each spectrum line is then:

\[
[A + \text{sinc}(n\kappa) + B \cdot \text{sinc}(n\kappa + \pi) + B \cdot \text{sinc}(n\kappa - \pi)]^2
\]

The effect of such masks is to redistribute the light among the secondary maxima. For a more complete discussion on this method of apodizing see Chapter 8 and reference 8.6.

---

**Spectrometers Employing Plane Diffraction Gratings**

full theory of such masks is discussed further in connection with multiplex spectrometers in section 8.6. In the case of the diffraction grating the process of apodizing, while easier to visualize, is less useful practically, since the other defects of the grating, such as the ghosts, are more important sources of 'noise' on the spectrum.

### 5.3 Blazing

It is the custom when ruling a reflecting grating (and sometimes when ruling a transmission grating) to cut the grooves so that the reflecting surface of each lies at a fixed angle to the grating surface. The angle between the groove normal and the grating normal is called the *blaze angle* of the grating. The effect of this process is to change the way in which light is reflected from each groove. There is now an extra phase change across the width of the groove and this affects the theory and changes the second term of eqn. (5.1). This second term defines the envelope of the function (shown dotted in figure 5.1) and for a blazed grating the principal peak of the envelope is moved sideways. It will be obvious that with a blazed grating it is possible to make the opaque strip width \( b \) almost equal to zero so that \( k = 1 \). Then for the wavelength whose first principal maximum lies under the centre of the envelope, all orders of diffraction except the first are found to lie under minima of the enveloping function and so are suppressed. All the incident light of this wavelength must be diffracted in the first order. This is naturally a very desirable state of affairs. Similarly, if light of one half of this wavelength is diffracted the order \( m = 2 \) lies under the centre of the envelope, and the envelope obligingly contracts to suppress all the orders \( m = 0, 1, 3, 4 \ldots \).

Thus reflecting gratings are now made with a variety of blaze angles for different purposes. The device known as an *echelon*, no longer used to any great extent, can be thought of as a grating with a small number of lines and a very large blaze angle. The high resolution that it achieves is the result of having \( m \) very large instead of \( n \) in the equation \( R = mn \). However \( k \) is not equal to unity and so there is no suppression of unwanted orders.

### 5.4 Angles of Incidence Different from the Blaze Angle

Most gratings are arranged in their mountings so that the angle of incidence is very nearly equal to the angle of diffraction, and if this angle
happens to be the same as the blaze angle there will be complete diffraction of the appropriate wavelength into that order. Under these conditions the phase angle \( \psi \), the phase difference between waves arriving at the centre of a ruling from those arriving at an edge, is zero, instead of \( \left( \pi / \lambda . a . \sin \alpha \right) \) in the normal-incidence case. If the grating is turned so that the angles of incidence and reflection are the same for some other wavelength the condition \( ku = m\lambda \) is maintained for adjacent rulings but the phase variation over the surface of a single ruling is not zero as was

\[
\left( \frac{2}{2m + 1} \right) \lambda_0 < \lambda < \left( \frac{2}{2m - 1} \right) \lambda_0
\]

where \( \lambda_0 \) is the blaze wavelength in first order, and \( m \) is the order in which the grating is to be used. It is found in practice that the polarization effects can be marked. Unpolarized light incident may be reflected with as much as 30 per cent polarization in a very unfavourable case, and no polarimetry should be attempted without first determining the polarizing property of the grating.

### 5.5 Overlapping of Orders and Free Spectral Range

The complete extinction of all orders of diffraction except the chosen one is clearly a theoretical ideal which is not achieved in practice simply by governing the shape of the rulings. When a spectrometer is adjusted to transmit a wavelength \( \lambda \) in first order there will be a certain amount of radiation transmitted simultaneously from wavelength \( \lambda / 2 \) in the second order, from \( \lambda / 3 \) in the third order and so on. In order that an unambiguous measurement shall be made these unwanted wavelengths must be excluded by some auxiliary means. Often simple colour filtering will be sufficient. In the visible region a filter which excludes all radiation below 3,500 Å will permit the whole of the visible range to be examined without overlap provided that the first order of diffraction is used. This is not always convenient. In the design of a spectrometer it is often desirable to use a higher order for the sake of the extra resolving power that is available, or because the most suitable grating has a large blaze.
angle. (The grating may have been chosen for its large blaze angle, in fact, to obtain a large efficiency.) Then either a set of filters must be chosen to isolate the desired wavelength region, or the spectrometer must be preceded by another spectrometer of lower dispersion but of the same étendue, which will remove the unwanted wavelengths before the radiation is passed to the main instrument. Such an attachment is called an order sorter.

Before describing order sorters, we will make some remarks about the wavelength range that is available, since the range is limited by grating theory rather than by the particular type of order sorter chosen.

If the shortest wavelength to be examined is \( \lambda_1 \) and the longest that can be examined simultaneously (i.e., without adjustment of the spectrometer other than the wavelength control) is \( \lambda_2 \), then the wavelength range \( \lambda_2 - \lambda_1 \) is called the free spectral range of the instrument. In first order, for example, the range will extend from \( \lambda_1 \) to \( 2\lambda_1 \), since at wavelengths above \( 2\lambda_1 \), \( \lambda_1 \) will appear again in second order. The free spectral range in this case is \( \lambda_1 \). Again, in second-order working, the range may extend from a shortest wavelength \( \lambda_1 \) to \( 3\lambda_1/2 \), since at wavelengths greater than \( 3\lambda_1/2 \), \( \lambda_1 \) will appear in third order. The free spectral range is then \( \lambda_1/2 \).

In the general case we may suppose that all wavelengths below \( \lambda_1 \) have been excluded by some external means. If the spectrometer is working in order \( m \), then at the wavelength setting for this shortest wavelength, \( \lambda_1 \) will appear, as required in order \( m \), and superimposed on it will be \( m\lambda_1 \), in first order, \( (m/2)\lambda_1 \), in second order, and \( (m/r)\lambda_1 \), in the \( r \)th order. Finally \( [m/(m - 1)]\lambda_1 \) will appear in the \( (m - 1) \)th order. All these longer wavelengths must be excluded and the range that can be examined without ambiguity is from \( \lambda_1 \) to \( [m + 1]/m \lambda_1 \) at which point \( \lambda_1 \) appears again in the \( (m + 1) \)th order. The free spectral range is then \( \lambda_1/m \).

5.6 Order Sorters

The simplest form of order sorter, the colour filter, has already been mentioned. In the visible region of the spectrum these will nearly always suffice. In the ultraviolet one almost always works in the first order, so that, for example, the range 1,800 Å to 3,600 Å can be examined provided that radiation below 1,800 Å is excluded. This is easy to do since air begins to absorb at about this wavelength. For the examination of wavelengths below 1,800 Å a vacuum spectrometer is needed, and then various crystalline materials can be used as filters. These are discussed in the chapter on concave grating spectrometers.

The real need is felt most keenly in the infrared region. Here the spectrum may cover several octaves, for example from \( \lambda = 10 \mu \) to \( 50 \mu \). In this case a more elaborate device is needed, and may take the form of a complete auxiliary spectrometer, of comparatively low dispersion, either feeding, or fed by, the main instrument. This auxiliary may be grating or a prism instrument.

A prism instrument has some advantage, since there is only one order of dispersion, but since this dispersion is non-linear, there may be some complication in coupling the scanning of the two instruments. If a second grating instrument is used, it should work in a different order from the first instrument. The effect of the two grating instruments is to combine the high resolution of one with the large free spectral range of the other. Whichever is used, care must be taken over the optical matching of the two, including a field-lens (q.v.) to image one disperser on to the other.

5.7 Grating Ghosts

These are the faint satellite lines that are seen on either side of the principal maximum of a diffracted line. They are not to be confused with the secondary maxima of the diffraction pattern, since their distance from the parent line is very much greater. They may be confused with genuine fine structure in a line from an atomic source. There may be a whole series of ghosts equally spaced on either side of the parent, decreasing in intensity as their separation from the parent increases. In a good quality grating the brightest ghost is less than 0.05 per cent of the parent intensity, but even so, with a sensitive detector, up to twenty or thirty ghosts can be detected on either side. The cause of these ghosts is usually a periodic error in the screw of the ruling engine on which the grating was made, or a progressive change in the shape of the ruling point.

Theory of Grating Ghosts

The grating eqn. (5.1) is derived, using Fourier theory, by considering the Fourier transform of a \( \theta \)-function of width \( a \) centred on a point \( x \). Using the summation property of Fourier transforms, this is summed for \( x = 0 \), \( (a + b) \), \( 2(a + b) \ldots \) up to \( m(a + b) \). The amplitude distribution in the focal plane follows immediately.

In practice, a periodic error in the screw of the ruling engine will
result in uneven spacings between the rulings and the above sum, instead of being taken at equal intervals \((a + b)\), will be at intervals:

\[
\begin{align*}
  x &= 0, (a + b) + \alpha \sin [2\pi \beta(a + b)], \\
  2(a + b) + \alpha \sin [2\pi \beta.2(a + b)], \\
  3(a + b) + \alpha \sin [2\pi \beta.3(a + b)], \\
  \ldots
\end{align*}
\]

The position error in each groove will then vary cyclically, the maximum displacement from the proper position being \(\pm \alpha\) and the period of the error being \(1/\beta\). That is, the pitch of the lead screw is \(1/\beta\) cm.

The amplitude distribution in the image plane will then be

\[
F(v) = \sum_{n=0}^{\infty} \exp \left\{ 2\pi i n [a + b] + \alpha \sin (2\pi n(a + b)) \right\}
\]

where \(v = m/\lambda\), and we have considered each slit to be a \(\delta\)-function of unit area. Both \(\alpha\) and \(\beta\) are considered to be small compared to \((a + b)\).

A result from Bessel function theory can be used here:

\[
\exp (ix \sin y) = \sum_{p=-\infty}^{\infty} J_p(x) \exp (ipy).
\]

Applying this:

\[
F(v) = \sum_{n=0}^{\infty} \exp \{2\pi i m(a + b)\} \sum_{p=-\infty}^{\infty} J_p(2\pi nx) \exp \{2\pi ipn(a + b)\}
\]

\[
= \sum_{n=0}^{\infty} \sum_{p=-\infty}^{\infty} J_p(2\pi nx) \exp \{2\pi i n(a + b) \cdot (p\beta + v)\}
\]

and for small values of the argument of \(J_p(x)\) that is, for \(J_0(x) = 1, J_1(x) = x/2, J_2(x) = x^2/4\) etc.:

\[
F(v) = \sum_{n=0}^{\infty} \exp \{2\pi i nx(a + b)\}
\]

\[
\quad + \sum_{n=1}^{\infty} 2\pi nx \exp \{2\pi i n(a + b) (v + \beta)\} + \ldots
\]

\[
- \sum_{n=1}^{\infty} 2\pi nx \exp \{2\pi i n(a + b) (v - \beta)\} + \ldots
\]

Summing for \(n\):

\[
F(v) = \frac{\sin \pi n(a+b)}{\sin \pi a} \frac{\sin \pi n(a+b)}{\sin \pi a} + \frac{\sin \pi v - \beta(a+b)}{\sin \pi v + \beta(a+b)} + \ldots (5.4)
\]

\[
\sin \phi = \frac{m \lambda}{ka} = \frac{m \lambda}{a + b}
\]

The condition for a principal maximum is \(u = m\pi\) and so, substituting, we obtain the formula:

\[
\sin \theta + \sin \phi = \frac{m \lambda}{a + b}
\]

relating the angle of incidence \(\phi\) and the angle of diffraction \(\theta\). A sign convention is necessary here since the equation is sometimes written

\[
\sin i - \sin r = \frac{m \lambda}{a + b}
\]

The equation:

\[
\sin i + \sin r = \frac{m \lambda}{a + b} \quad (5.5)
\]

will be employed here, on the understanding that the sign of the angle changes as it crosses the grating normal. If there is any doubt in a particular case, specular reflection \((m = 0)\) requires that \((\sin i + \sin r) = 0)\)
and hence either \( i \) or \( r \) is negative. That is, they lie on opposite sides of the grating normal.

5.9 The General Grating Equation – Oblique Incidence

The consequences of oblique incidence are most easily shown with the aid of Cartesian geometry. Consider a plane grating, and let there be a rectangular coordinate system with the grating lying in the plane \( x = 0 \). Consider a ray incident on the grating at the origin, with direction cosines \( L_1, L_2, L_3 \). Let there be another ray, parallel to the first, meeting the grating at a point \( P(o, y, z) \). Let them both be reflected with direction cosines \( L_1', L_2', L_3' \). Draw lines \( PA, PB \) from \( P \) to the first incident and diffracted rays, meeting them perpendicularly. Then \( PA \) lies on an incoming wavefront, and if the outgoing rays are in the direction of a principal maximum, \( PB \) lies on an outgoing wavefront. Therefore \( AOB \) must be an integral number of wavelengths long. By geometry:

\[
AO = L_2y + L_3z; \quad BO = L_2'y + L_3'z
\]

so that the condition becomes:

\[
y(L_2 + L_2') + z(L_3 + L_3') = m\lambda
\]

(5.6)

Now suppose that the grating rulings lie parallel to the \( y \)-axis. Then \( z \) can be taken as the distance between two adjacent rulings, and \( y \) is an arbitrary point on one ruling. The equation must hold for any value of \( y \), and so we obtain the first condition for \( OB \) to be an outgoing ray. We require that \( L_3 = -L_3' \) and also that \( L_2 + L_2' = m\lambda/(a + b) \). If now we transform the system to polar coordinates \( (r, \theta, \phi) \) as in the diagram, we find that the two conditions become:

\[
\theta = \theta' \quad \text{and} \quad \cos \theta \sin \phi + \sin \phi' = \frac{m\lambda}{a + b}
\]

(5.7)

The second condition is clearly the usual grating equation, with a modification to allow for oblique incidence. Several things of interest follow from this condition. We can ignore, for the purposes of this discussion, the optical components of the spectrometer, and consider only the chief-ray incident at \( O \).

We consider the following cases:

(i) Radiation incident, normal to the grating, from a straight slit. Then \( L_3 = 0 \) for all incoming chief-rays, from all points on the slit. Therefore \( L_3' \) must take the same value for all the outgoing chief-rays,

![Figure 5.5](image)

![Figure 5.6](image)

and all these chief-rays must therefore lie on a circular cone whose axis is the \( z \)-axis. If the focal plane of the spectrometer is perpendicular to the grating surface and rulings the images of the entry slit will be circular. Otherwise they will be elliptical or hyperbolic.
(ii) Radiation incident from a straight slit not on the grating normal. This case, like the general case, is best dealt with by spherical trigonometry. Consider the grating to be at the centre of a sphere of unit radius. (The radius is not important since we are only concerned with angles of diffraction.) Let the grating normal cut the sphere at latitude \( \theta \) and longitude \( \phi \), and let the rulings be parallel to the polar axis. All the incoming rays from a straight slit will have the same longitude at the points where they cut the sphere. One of these rays will cut the sphere at latitude \( \theta \) and longitude \( \phi \). The point of emergence of the corresponding outgoing ray will be governed by the equation:

\[
\cos \theta (\sin \beta + \sin \beta') = \frac{m\lambda}{a}
\]

where \( a \) is the grating spacing (instead of \( a + b \) as previously).

![Diagram](image)

Figure 5.7(a). Radiation at oblique incidence from a straight slit not on the grating normal.

![Diagram](image)

Figure 5.7(b). Radiation at oblique incidence from a straight slit. Mercator projection of the unit sphere, showing the formation of a parabolic image of the slit.

Spectrometers Employing Plane Diffraction Gratings

The coordinates of the point of emergence are thus: latitude \(-\alpha\), longitude \(\phi\), where \(\sin \beta' = (m\lambda/\alpha) \sec \alpha \sin \beta\). If \(\beta'\) is the emergent longitude for the equatorial ray \((\phi = 0)\) and if \(\alpha\) is small so that the approximation \(\sec \alpha \approx 1 + (\alpha^2/2)\) is valid then

\[
\sin \beta' - \sin \beta' = (m\lambda/2\alpha)\alpha^2
\]

and

\[
\Delta\beta' = \sec \beta' - (m\lambda/2\alpha)\alpha^2.
\]

This is the equation of a parabola, provided that the slits are not long compared with the focal length of the instrument. In the particular case where the entry and exit beams are at the same angle to the grating normal, then \(\beta' = \beta'\) and the expression further simplifies to:

\[
\Delta\beta' = \alpha^2 \tan \beta' - \alpha^2.
\]

This case is quoted at some length since it is a common way of using a plane grating, that is, with a straight entry slit and, hopefully, a straight exit slit. Unless the exit slit is curved to fit the diffracted image there may be a serious loss in resolving power. To make things more difficult still, the shape of the parabola changes with wavelength, so that there is no particular curved exit slit that can be fitted to match a straight entry slit over an extended range.

As we shall see later, the maximum length \(2\alpha\) of the straight slit that can be matched to another straight slit is connected with the resolving power required by the equation: \(2\alpha^2 = 1/R\).

This is a result more easily derived from the theory of circular curved slits.

(iii) The general case. Consider again the unit sphere with the grating at its centre, oriented as before, with its normal at latitude and longitude \(\theta\) and \(\phi\), and rulings parallel to the polar axis.

Take a point \(A(\theta, \phi)\) on the equator, and let \(OA\) be the optic axis of the spectrometer. Draw a small circle on the surface of the sphere centred at \(A\). The coordinates of a point \(B(\theta, \phi)\) on the circumference of this circle are connected by the equation:

\[
\cos \theta \cos (\phi - \beta) = \cos \omega
\]

where \(\omega\) is the angular radius of the small circle.

Two points diametrically opposite on this circumference are \(B\) and \(B'\) with coordinates \((\theta, \phi)\) and \((-\theta, \pi + \phi)\). If \(B\) and \(B'\) are also points where ingoing and outgoing chief-rays cut the surface of the sphere we must also have the relation:

\[
\cos \theta [\sin \phi + \sin (\pi + \phi - \phi)] = \frac{m\lambda}{a}.
\]
This equation can easily be changed into:
\[ 2 \cos \theta \sin \beta \cos (\phi - \beta) = \frac{m\lambda}{a} \]

or,
\[ 2 \cos \omega \sin \beta = \frac{m\lambda}{a} \]

and since the left-hand side of the equation is a constant, so will be the right-hand side. That is, every pair of diametrically opposite points on a circle will be chief-ray intersections for the same wavelength. The actual wavelength will depend on \( \sin \beta \), i.e. on the angle the grating normal makes with the optic axis of the system.

**Figure 5.8.** The coordinate system for the general case of curved entry and exit slits.

Thus any part of such a circle can be used as an entry slit, and the corresponding exit slit will be the diametrically opposite part of the same circle. In principle it should be possible to use an entire semicircle as the entry slit and the other semicircle as the exit slit. In practice this is not feasible since the slits need a finite width and the dispersion is in one direction. There is, however, no restriction on the size of the circle. It can be as large as convenient. In practice, since the dispersion is along the equatorial plane, there is no advantage in using more than about 120° of arc (of the small circle) for the length of the slit. It is difficult to form an image of an external source on a long curved slit, and equally difficult to remove the light from a long curved exit slit and focus it on to a detector. In photographic photometry there is no point in using a long curved slit at all. It is strictly a device to allow monochromators to yield a high luminosity.

As mentioned above, it is quite usual to approximate to the ideal condition by using short straight slits which are chords of the ideal circle. The wavelength focused at \( A \) and \( A' \) is given by the equation:
\[ 2 \cos \omega \sin \beta = \frac{m\lambda}{a} \]

and the wavelength focused at \( B \) and \( B' \) is similarly given by:
\[ 2 \cos \omega' \sin \beta = \frac{m(\lambda + \Delta \lambda)}{a} \]

Now by geometry, \( \cos \omega = \cos \theta \cos \omega' \) and these three equations yield:
\[ \frac{\Delta \lambda}{\lambda} = \frac{\theta^2}{2} = \frac{1}{R} \]

so that the resolving power is limited in this way by the lengths of the slits.

More frequently seen is the equation connecting the total angular lengths of the slits with the resolving power:
\[ \frac{\theta^2}{8} = \frac{1}{R} \]

\( \theta \) now being the angular length of the slits, as subtended at the collimator.

It is no coincidence that this formula is identical with that giving the
resolving power of a Fabry-Pérot spectrometer in terms of the field-stop diameter. The conditions of dispersion are analogous. It would be possible to use circular slits of a diameter so small that the inner diameter is zero, and the entry and exit slits together make a circle. This would be rather far from the optimum condition for a grating spectrometer.

5.10 Differential Dispersion

Consider the rays entering and leaving the instrument in the equatorial plane. If one of the slits has a finite width and the other is infinitesimal, the instrumental profile, that is the line profile that is obtained when strictly monochromatic light passes through, will be top-hat shaped. If the width of the exit slit is made equal to the width of the image of the entry slit, the profile is triangular, but the distance on the wavelength scale is the same, between the half-intensity points, as in the top-hat profile. The wavelength range passed by the instrument is then given by differentiating the grating equation.

Take the case of a finite entry slit:

\[ \sin i + \sin r = \frac{m\lambda}{a} \]

from

\[ \cos i \cdot di = \frac{m}{a} \cdot d\lambda \]

giving the angular slit width corresponding to a spectral bandwidth \(d\lambda\). The same bandwidth would be passed by an exit slit of width \(dr\) where

\[ \cos r \cdot dr = \frac{m}{a} \cdot d\lambda \]

so that the corresponding entry and exit slit widths are related by the equation:

\[ \cos r \cdot dr = \cos i \cdot di. \]

Thus there is some magnification effect when \(i\), as is usual, is not the same as \(r\). This becomes important when we consider spectrometers with multiple exit slits. In the Ebert mounting too, with single entry and exit slits, the angles of incidence and diffraction may differ by several degrees, thus giving some magnification, and as the grating is rotated to scan through different wavelengths this magnification changes. When very high precision spectroscopy is contemplated this small change must be compensated by an adjustment to the width of one or both of the slits.

Spectrometers Employing Plane Diffraction Gratings

An interesting, though rather superfluous point is that the relative slit widths are determined by the law of conservation of étendue: the narrower slit sees a larger projected area of the grating than does the wider slit.

5.11 Luminosity and Resolving Power

Consider again the simple case of a system with short straight slits. In practice a figure of \((\text{focal length})/50\) is acceptable for the length of a straight slit. With a focal length \(f\), the slit width for a wavelength range \(\Delta\lambda\) is:

\[ w = \frac{fm}{a} \frac{\Delta\lambda}{\cos i} \]

and the length is \(f/50\) so that the slit area is:

\[ \frac{f^2 m}{50} \frac{\Delta\lambda}{a \cos i} = A_s. \]

The projected area of the grating as seen from the slit is \(A_s \cos i\) and the luminosity of the system is thus:

\[ L = \frac{A_s A_r \cos i}{50} \frac{m}{a} \frac{\Delta\lambda}{co}. \]

The resolution-luminosity product can then be written as

\[ E = RL = \frac{A_s}{50} (\sin i + \sin r). \]

(In the particular case of an Ebert spectrometer where the optic axis makes an angle \(\beta\) with the grating normal, and \(\omega\) is the Ebert angle, \(\sin i + \sin r\) becomes \(2 \sin \beta \cos \omega\) and then the \(RL\) product is:

\[ RL = \frac{A_s}{25} \sin \beta \cos \omega. \]  \(5.9\)

Thus, for maximum efficiency, the grating should be used with the largest possible blaze angle and the smallest possible Ebert angle (e).

Luminosity and resolving power are interchangeable up to the diffraction limit set by the criterion \(R = mn\). As usual, by doubling the slit width the luminosity is doubled and the resolving power halved. In practice it is possible to gain luminosity faster than resolving power is lost since the slit length can be increased. Instead of taking the slit length as a constant \(1/50\)th of the focal length we can use the criterion developed earlier, that the angular length is \(2\sqrt{2/R}\). With the same procedure as before we find a new equation connecting the resolving power and the luminosity:

\[ R^{1/2} L = 2.83 A_s (\sin i + \sin r) \]  \(5.10\)
The further caution can be added at this time. These formulae only apply when emission lines are being measured. When a continuum is studied, a doubling of the slit width will not only increase the luminosity by a factor 2, but will also double the spectral range admitted. The total power passed through the instrument is quadrupled. For a fixed slit length we would have $R^2L = \text{constant}$, and with the more advanced criterion, $R^{9/2}L = \text{constant}$.

5.12 Mountings for Plane Gratings

The remarks about chief-rays and the laws governing their behaviour apply equally well to any method of mounting the plane grating as a spectrometer. The other factors that govern the image formation are geometrical, and a variety of mountings are available which may be selected because they are compact or because the aberrations are small.

The plane transmission grating has such a poor performance by comparison with the reflecting grating that it is used only in the cruder type of instrument. For this reason no separate space will be devoted here to transmission gratings, but we will consider instead the mountings that can be used with plane reflecting gratings to make high-performance spectrometers.

The basic mountings all use one or more concave mirrors as the collimating and focusing elements. In this way the chromatic aberrations of lenses are avoided. Thereafter, mountings can be classified into one of two groups, each with a number of variants to suit particular needs.

1. THE LITTROW MOUNTINGS

A single concave mirror is used, and it acts both as collimator and focusing element. Normally this mirror should be an off-axis paraboloid, but a spherical mirror can be used, although it incurs some extra aberration. The chief advantage of the mounting is its cheapness and compactness. The whole spectrometer can be packed into a tube not much wider than the ruled width of the grating. On the other hand its optical performance is comparatively poor if used over a range of grating angles greater than a few degrees. An analysis of the geometrical optics of the Littrow mounting has been made by Kudo [3] who gives formulae for the radius of curvature of the line image under various conditions, with the assumption of a straight entry slit. As we have noted earlier in this chapter, the radius of curvature of the image will vary with the angles of incidence and diffraction at the grating, so that no single exit slit will match the entry slit over a wide wavelength range.

Kudo finds a condition for minimum astigmatism in the system:

(a) The entry and exit slits must lie on a line perpendicular to the axis of the paraboloid.

(b) The distance of the grating from the mirror must be equal to the focal length of the mirror.

It goes without saying that, both from the point of view of compactness, as well as optical aberrations, the distance of the grating from the optic axis should be as small as possible.

To avoid undue complication in the vicinity of the focal plane a small plane mirror is sometimes used to image the spectrum lines onto an exit slit suitably removed from the entry slit. This mirror may also be made to serve as a fine wavelength control, by attaching a bar and a micrometer to its stalk.

LITTROW SPECTROMETERS USING A LENS AS A FOCUSING ELEMENT. When a diffraction grating is used to obtain dispersion, it is only under the most extraordinary circumstances that the use of a lens should be contemplated. While a lens can be quite satisfactory for a prism-Littrow spectrometer, it should be remembered that even an apochromatic lens will not make the focal surface flat, so that there will need to be a complicated link system to move the focal plane as the grating is tilted.
The form of the link itself will depend on the optical properties of the glass in the lens, and would presumably involve the manufacture of cams and similar devices.

(2) THE EBERT MOUNTING
This type of mounting is the most important for the construction of high-performance spectrometers and spectrographs. It was first described by H. Ebert [4] in 1889 and subsequently rediscovered, with advantages, by W. G. Fastie [5] who described it in 1952. For this reason it is often known as the Ebert–Fastie mounting. The basic difference between it and the Littrow mounting is that two different optical elements (or two different parts of the same element) are used to collimate and to focus the radiation. There are several variants, some of which use two separate mirrors while others use only one.

The basis of the system is shown in figure 5.11. The entry and the exit slits, S₁ and S₂, can be either in the plane of dispersion (the 'side-by-side' Ebert) or in a plane which is perpendicular to the diffraction directions, and which contains the axis of symmetry of the system.

The outstanding feature of the system is that, although the mirrors are used off-axis, the aberrations generated by the first are to some extent compensated by those of the second. In particular the primary coma introduced by the first mirror, although it might be quite large, is mostly removed by the opposite coma of the second. The distorted wavefront (which ideally should be plane) coming from the first mirror is reflected by the grating with approximately the same shape, and this shape is the right one to be reconverted into a spherical wave by the second mirror. Spherical aberration is not corrected, and the coma correction will hold less and less well as the Ebert angle becomes larger.

\[ \text{Figure 5.11. Basic layout of the Ebert-Fastie spectrometer.} \]

However, if the Ebert angle is kept small, and the focal ratios used are large, the chief aberration, the one which limits the performance of the Littrow mounting, coma, is largely eliminated. This makes it possible (a) to realize most of the theoretical resolving power of the grating, and (b) to use long curved slits to improve the luminosity.

The chief variants of the Ebert system are:

(i) THE CZERNY–TURNER MOUNTING. Two separate spherical mirrors are used, side-by-side, and the slits are placed on either side of

\[ \text{Figure 5.12.} \]

(a) The Czerny–Turner variant of the Ebert mounting.
(b) The crossed Czerny–Turner variant. Both of these variants offer some advantages over the original Ebert system. For details, see text.

the grating. This is the most common variant, and is the one that will be described in greatest detail later on.

(ii) THE CROSSED CZERNY–TURNER MOUNTING. This is identical in most of its essential details to the previous variant, but it offers some advantages in practice. In particular it is easier to suppress scattered
light by means of internal baffles, and to prevent the entry slit from ‘seeing’ the second mirror, or the exit slit from seeing the first.

(iii) **The Newtonian Ebert Mounting.** This reduces the off-axis aberrations even further, and paraboloidal mirrors can be used to remove spherical aberration. This is an important point when the highest possible resolution is required. There may be a five or ten per cent loss of luminosity due to the central obstruction, but this is tolerable. A small but important design point may be mentioned here. The Newtonian mirror for the emerging ray, and its mounting, must both be clear of the parallel section of the incoming ray, and both must lie in the shadow cast by the Newtonian mirror for the incoming ray. The initial alignment of this type of mounting is not particularly easy, and some little thought should be given to the design of the adjustments on the Newtonian mirrors.

(3) **The Pfundt Mounting**

This may be classified as a hybrid of the two previous mountings and the original Ebert design. The two extracting mirrors can be side-by-side, as in figure 5.14, or they can be above and below the plane of the diagram. In the latter case they can, if desired, face the same way, in which case the mounting is a variant of the Littrow mounting. In practice the mirrors cause a large obscuration of the grating area when reasonably long slits are used, and this makes the mounting unsuitable for many purposes.

![Diagram of the Pfundt system spectrometer](image)

**Figure 5.14.** The Pfundt system spectrometer. As in the previous figure the dimensions have been exaggerated. The central obstruction would be much smaller in practice. An alternative system would have two small prisms, one above the other, to extract the beam from the system.

(4) **The Ebert Spectrograph Conversion**

The Ebert spectrograph is designed to allow the focal surface to be a plane, at the expense of having images of the entry slit which are curved. The entry slit is then below the grating and the photographic plate above. If the instrument is converted to a spectrometer, it implies either that the detector is moved bodily along the focal plane, or that the grating is tilted while the detector remains fixed. In either case the curvature of the lines changes during the scan, and very high performance is consequently not to be expected. This mounting is not considered further here.

**Details of the Czerny-Turner Mounting**

Many of the geometrical optical properties of the Czerny-Turner mounting are shared by the other variants. This is the most flexible variant, in that there are a larger number of design parameters. Of these parameters, the most important are: (a) the position of the grating in relation to the foci of the mirrors; (b) the separation of the mirror centres; (c) the positions of the centres of curvature of the mirrors; and (d) the focal ratios of the mirrors.

**The Choice of the Focal Length.**

Once the required resolving power has been fixed, the angular widths of the entry and exit slits are determined by the equations:

\[
\delta r = \frac{2}{R(\cot \beta + \tan \omega)}; \quad \delta r = \frac{2}{R(\cot \beta - \tan \omega)}
\]
where $\beta$ is the angle between the grating normal and the main optic axis of the system, and $\alpha_0$ is the Ebert angle.

The specification of the grating decides the diameters of the beams that are to be focused by the two mirrors. It then remains to choose a focal ratio such that the geometrical aberrations do not harmfully degrade the image of a monochromatic line. In a spherical mirror the diameter of the smallest circle that a point source will produce is obtained from the equation:

$$\Delta_{\text{geom}} = \frac{f}{64F^2}$$

where $f$ is the focal length, $F$ is the focal ratio.

The diffraction limited width of the image of a point source is given by:

$$\Delta_{\text{diff}} = \frac{2f \tan \beta}{R}$$

and if we allow $2\Delta_{\text{geom}} = \Delta_{\text{diff}}$ we have a criterion for the focal ratios:

$$F > \frac{4}{5}(R \cot \beta)^{1/3} \quad (5.11)$$

Other forms of this equation can be obtained after some algebra. Another possibility is:

$$F > \frac{4}{5} \left( \frac{2\pi R}{\lambda R_{\text{max}}} \right)^{1/3}$$

where $\omega$ is the projected width of the grating, $R$ the required resolving power and $R_{\text{max}}$ is the theoretical limiting resolving power of the grating.

In all these expressions the focal ratio $F$ is the ratio of the focal length of the mirror to the projected width of the grating as seen from the centre of the mirror.

**Choice of the Mirror Positions**

One obvious choice is to follow the geometry of the original single-mirror Ebert mounting, and to place the mirrors with their centres of curvature at the same point. The main result is that the incoming and outgoing chief-rays are parallel. This has an important consequence, since, if long curved slit pairs are used, the two slits lie in the same plane, the plane normal to the chief-rays.

The relative positions of the mirrors and the grating are also important. Sassa [6] has made an analysis of the behaviour of the parallel marginal rays leaving a grating and focused by part of a single Ebert mirror. There is some astigmatism, and it is the tangential focus that is important to spectrum line formation. Coupled with the astigmatism is some distortion and field curvature. Sassa was able to show that there

are two particularly interesting positions for the grating in relation to the mirror. One of these results in a short straight slit being imaged as a short straight slit, although the focal surface is not a plane. The other results in a flat focal surface, but in a curved image of the slit. This latter condition is likely to be of use in spectrograph design.
Spectrometers Employing Plane Diffraction Gratings

The two positions are as follows. Consider a set of Cartesian coordinates with the main optic axis as the x-axis. The Ebert mirror has its pole at the origin and the centre of the grating is at (G, 0). Any two rays incident on to the grating at angles $\alpha$ and $\gamma$ and leaving at angles $\beta$ and $\phi$ will come to a focus at $(X, Y, Z)$ where $X, Y$ and $Z$ are given by the expressions:

$$X = \frac{R}{2} + \frac{R}{4}[1 - 3\left(\frac{G}{R} - 1\right)^2 - 3\beta + \sin^2 \phi]$$

$$Y = \frac{R}{2} \sin \beta \cos \phi + \left(\frac{G}{R} - 1\right)\frac{G - R}{R} \sin^2 \beta + \sin^2 \phi \sin \beta$$

$$Z = \frac{-R}{2} \sin \phi - \left(\frac{G}{R} - 1\right)\frac{G - R}{R} \sin^2 \beta + \sin^2 \phi \sin \beta$$

If now we put $G = R(1 - 1/\sqrt{3})$ we find that $X$ is a constant. Then the focal surface is a plane.

Alternatively if we put $G = R/4$ then $Y$ is a constant. Each image is curved in the $XZ$ plane, but lies entirely in a plane parallel to the $X-Z$ plane.

This second condition would, in general, be more suitable for a spectrometer. However, there is the difficulty that the ray bundle from the mirrors must be able to pass the grating, and a simple geometrical argument will show that this type of mounting requires an Ebert angle twice as great as that in a mount in which the grating is close to the focal plane.

The expressions given above are the early terms of an expansion in powers of the sines, and so will hold only for small angles, that is for short slits. The effects are small anyway, and the most usual position for the grating in practice seems to be in the vicinity of the focal plane of the instrument. That is, close to the position required by condition (a). Whether the grating is closer to the mirrors than the focal plane or more distant depends on practical circumstances. If there is no shortage of space it may well be found convenient to have the focal plane behind the grating, thus leaving a clear space for auxiliary apparatus that needs to be in the focal plane. On the other hand, a considerably more compact spectrometer can be made if the focal plane is in front of the grating and the beams are injected and extracted by a small pair of $45^\circ$ mirrors close to the edge of the grating.

Very large, permanently mounted spectrometers are perhaps best arranged with the focal plane behind the grating. It is then possible to arrange for several different input and output positions with removable kinematically mounted mirrors that select the particular input and output channels for a particular experiment. Experimental sources and apparatus can then be constructed separately and attached to the instrument without disturbing another experiment.

5.13 Grating Spectrometers in the Infrared

The visible region of the spectrum covers less than one octave, and so it is possible to talk about the choice of a resolving power, on the assumption that, once the slit width is set, the resolving power is determined. This cannot be done either in the infrared or the ultraviolet since in both of these regions the experimental technique covers several octaves. At a constant slit width the resolving power varies considerably over the working range. As a rule the information that is needed about a spectrum is the spectral energy density $I(\lambda)$ or $I(\nu)$, where $I(\lambda).d\lambda$ is the total energy transmitted in the range $\lambda - \lambda + d\lambda$.

If the slit width is $d\lambda$ the luminosity of the instrument will be

$$L = \frac{A \cos \theta . d\lambda}{50}$$
80 - Design of Optical Spectrometers
when short straight slits are used, and the total flux transmitted through
the instrument will be:
\[ F(\lambda) = \frac{A \cdot \cos \, i \cdot di}{50} \cdot I(\lambda) \cdot d\lambda. \]
Since \( i \) and \( di \) determine \( d\lambda \), we have:
\[ F(\lambda) = \frac{2A \cdot \cos^2 \, i \cdot di^2}{50 \cdot m} \cdot I(\lambda) \]
and so \( F(\lambda) \) will be proportional to \( I(\lambda) \) if the slit can be made to vary
so that \( di = (\text{constant}). \sec \, i \).
To make \( F(\nu) \) proportional to \( I(\nu) \) it is easy to show by a similar
argument that the slit width must be proportional to \( \tan \, i \). Thus \( I(\nu) \) is
measured if the resolving power of the instrument is kept constant.
For most purposes in the infrared it is \( I(\nu) \) that will be most interesting,
and the slit can be coupled to the grating turntable by a simple
sliding link that will give a displacement proportional to \( \tan \, i \).

REFERENCES

6.1 The Rowland Grating
The invention of the concave diffraction grating is due to H. A. Rowland,
who first gave an account of its unique focusing property in a short note
in the Philosophical Magazine of June 1882. A more detailed account
followed in the September issue of that year, and it excited some contro-
versy, chiefly because the method of ruling the grating was not under-
stood. The Rowland grating consists of a series of parallel rulings made
on the surface of a concave spherical mirror. The spaces between the
lines are not constant along an arc of the surface, but are equal when seen
from infinity, that is they are equally spaced along a chord of the arc.
Alternatively the rulings may be described as the intersections with the
sphere of a series of equally spaced parallel planes. One of these planes –
the one through the centre of the grating surface – also contains the
optical axis of the grating.

If a plane is taken through the centre of the grating, perpendicular
to the rulings, a circle can be drawn, touching the grating at its centre
and with a diameter equal to the radius of curvature of the grating.
This circle is called the Rowland circle. The focusing property of the
grating is that a source of light placed at a point on the Rowland circle
is diffracted and refocused to form an image of the source at some other
point on the Rowland circle. Different wavelengths form images at different points. If a ray is traced from the source, via the
centre of the grating surface to the image, then the usual grating equation \( \sin \, i + \sin \, r = m\lambda/a \) governs the angles of incidence and
diffraction.

Because of the focusing property, no other optical elements are
needed. This makes the grating eminently suitable for the study of those
regions of the spectrum where optical transmission is low. In particular,
concave gratings are used for the study of the spectrum below \( 1,200 \, \text{Å} \),
where reflection coefficients are likely to be as low as 10 or 20 per cent and any solid thicker than a few hundred Å is quite opaque.

Commercially made concave gratings are likely to be blazed for the first order of some short wavelength, and the blaze angle is sometimes changed as the ruling progresses across the surface. The 'tripartite' rulings of some of the Bausch and Lomb gratings are of this type.

As has been pointed out in other chapters, it is always desirable to have as few adjustments as possible on a spectrometer, and in particular it should be possible to change the wavelength appearing at the exit slit by means of only one adjustment. This is particularly the case when considering concave grating spectrometers, since they are always likely to be enclosed in a vacuum-container for the study of the far ultraviolet.

The original Rowland gratings were intended for spectrographs where no adjustments were needed except for a gross alteration of the region of the spectrum to be studied. Such photographic mountings then consisted usually of simple applications of the elementary geometrical theory of the Rowland circle. The main exception was the Wadsworth mounting which will be described below. In recent years some vacuum spectrometers have been made by converting spectographs, and these conversions have sometimes involved linkages or multiple adjustments. However, a series of theoretical studies by several Japanese workers [1, 2, 3, 4, 5] has resulted in the evolution of several mountings for the concave grating which require only one simple rotation as the wavelength adjustment, and which frequently give rise to simple and compact instruments that can be placed in small vacuum-containers. Most modern concave grating spectrometers are based on one or other of these mountings.

These mountings have their limitations. They usually give a good focus only over a restricted range of wavelengths. The most restricted ranges are those of the Johnson-Onaka mountings [6], although these are the ones that are most compact. The Seya mounting has a much more extended range, but needs a container of a more awkward shape. All of the mountings are strictly monochromators; there is no extended range of wavelengths that can be focused on a surface simultaneously. Polychromators, should they be desired, must be based on the Rowland focusing property.

To see how these mountings have evolved, and at the same time to describe them and the more conventional mountings, we will examine a simplified version of the Beutler theory of the concave grating. The classic paper on the subject is that of Beutler [7], and he has been followed by Namioka, Seya, Miyake, Greiner and Schaffer [8, 9] and several others. It is mostly from Miyake's papers that we shall be quoting here.

### 6.2 Geometrical Theory of the Concave Grating

The theory is based on the 'Eikonal' method of geometrical optics. Beutler considered the imaging properties of the grating for photographic work, and was concerned with obtaining an intense image. Aberrations such as astigmatism, which spread the light from a point source out into a straight-line image, will seriously affect the overall speed of a spectograph but are less troublesome in a spectrometer, where the designer is only concerned with resolving power per se, and does not mind if he has to collect the light from a long thin exit slit.

In the general case considered by Beutler there is a system of Cartesian coordinates, with the grating centre at the origin, the \(x\)-axis as the optic axis of the grating (so that the \(y-z\) plane is a tangent plane) and the rulings parallel to the \(z\)-axis.

There is an object point at \(O(x, y, z)\), and an image point at \(I(x', y', z')\). A ray is then considered, which travels from \(O\) to an arbitrary point \(H(u, v, l)\) on the grating surface and thence to \(I\). The total path \(OH = F\) and \(HI = F'\) is then written down as a power series in the variables \(u, v, l\). The resulting expression is somewhat involved, and is given explicitly by Namioka [2]. The condition that I shall be an image of \(O\) is that the optical path \(F + F'\) shall change by one wavelength as \(w\) increases by one ruling space, and that \(F + F'\) is

![Figure 6.1](image)

**Figure 6.1.** The simplified Eikonal. The path OPI must differ from the path OQI by an integer number of wavelengths, one for each groove between P and Q.
shall be independent of \( l \). These two requirements imply that all the coefficients of \( \omega, \omega^2, \ldots \), and of \( l, l^2 \ldots \) in the expansion should vanish. In practice they do not, of course, but instead give rise to the various aberrations. These aberrations are discussed by Beutler.

In the simplified case that we shall consider, the third dimension, \( z \), is ignored, and everything is supposed to take place in the \( x-y \) plane.

For convenience we choose coordinates \( \rho, \rho', i \) and \( r \) and write the point \( O \) as \( O(\rho \cos i, \rho \sin i) \). Similarly \( I = I(\rho' \cos r, \rho' \sin r) \). The equation of the grating surface is \( x^2 - 2Rx + y^2 = 0 \). With this as ammunition, and after some tiresome elementary algebra we get an expression for the Eikonal \( E \) as follows:

\[
E = \rho + \rho' - \omega(\sin i + \sin r) + \frac{\omega}{\rho}
\]

\[
+ \frac{\omega}{2} \sum_{n=0}^{\infty} \left[ \left( \frac{\sin i}{\rho} \left( \frac{\cos^2 i}{\rho} - \frac{\cos i}{R} \right) + \left( \frac{\sin r}{\rho'} \left( \frac{\cos^2 r}{\rho'} - \frac{\cos r}{R} \right) \right) \right] \tag{6.1}
\]

The vanishing of the various coefficients then gives the conditions for an image to be formed at \( I \).

(i) If the coefficient of \( \omega \) vanishes, then \( \sin i + \sin r = \frac{\omega}{\rho} \), the standard grating equation.

(ii) If the coefficient of \( \omega^2 \) vanishes, then:

\[
\frac{\cos^2 i}{\rho} - \frac{\cos i}{R} + \frac{\cos^2 r}{\rho'} - \frac{\cos r}{R} = 0 \tag{6.2}
\]

which is satisfied by \( \rho = R \cos i, \rho' = R \cos r \). This is the original Rowland condition, which requires that the entry and exit slits lie on the Rowland circle.

(iii) If the coefficient of \( \omega^3 \) vanishes, then:

\[
\frac{\sin i}{\rho} \left( \frac{\cos^2 i}{\rho} - \frac{\cos i}{R} \right) + \frac{\sin r}{\rho'} \left( \frac{\cos^2 r}{\rho'} - \frac{\cos r}{R} \right) = 0 \tag{6.3}
\]

Conditions (ii) and (iii) can be satisfied simultaneously by the auxiliary condition:

\[
\left( \frac{\sin i}{\rho} - \frac{\sin r}{\rho'} \right) \left( \frac{\cos^2 i}{\rho} - \frac{\cos i}{R} \right) = 0 \tag{6.4}
\]

This yields either the Rowland condition or, as Miyake has pointed out, the alternative requirement:

\[
\frac{\sin i}{\rho} = \frac{\sin r}{\rho'} \tag{6.5}
\]

Several mountings, including the Wadsworth, derive from this condition of Miyake.

Combining eqns. (6.2) and (6.5) we get:

\[
\rho = \frac{R(\sin r \cos^2 i + \sin i \cos^2 r)}{\sin r (\cos i + \cos r)} \tag{6.6}
\]

and a similar expression for \( \rho' \).

Two mountings derive from this equation immediately:

(i) If \( r = 0, \rho = \infty \) and \( \rho' = \frac{R}{1 + \cos i} \). This is the Wadsworth stigmatic mounting.

(ii) If \( r = 0, \rho = \infty \) and \( \rho' = \frac{R}{1 + \cos i} \). This is the Wadsworth stigmatic mounting.

### 6.3 The Japanese Mountings

Other mountings are derived by Miyake by taking a point \( P(X, Y) \) in the plane and rotating the grating about it, the slits meanwhile being kept fixed. Neither \( O \) nor \( I \) remains on the Rowland circle (unless \( P \) is at the centre of the Rowland circle) but under certain conditions an image of \( O \) is formed at \( I \). The condition is that \( P \) should lie on a particular straight line. The derivation of the Cartesian equation of this line is fearsome rather than tiresome, but Miyake gives the equation as:

\[
\frac{2X}{R} + \frac{Y}{R (\sin i + \sin r)} \left( 1 - 2 \cos i \cos r \right) = 1 \tag{6.7}
\]

In order that, during rotation, the grating should not move too far in relation to the incoming and outgoing chief-rays, the point of rotation of the grating should lie on a line perpendicular to the bisector of the chief-rays. The intersection of such a line with the one above fixes the optimum point for the centre of rotation. This point is given by Onaka as being at a distance \( L \) from the pole of the grating, on a line perpendicular to the bisector of the lines joining the pole to the entry and exit slits. The distance \( L \) is:

\[
L = \frac{R \sin \theta}{[1 - \tan \alpha (\tan i - \tan r)/2]} \tag{6.8}
\]

where \( \theta \) is the angle between the grating normal and the bisector of the chief-rays, \( 2 \alpha = \text{angle OGI} \), and \( i \) and \( r \) are the angles of incidence and reflection when the grating is set for some standard wavelength (usually the wavelength in the middle of the required range).

This is the Johnson-Onaka mounting.

The equation of the line (6.7) can be derived by constructing an eikonal for the condition where the grating has been rotated about the point \( P \) through an angle \( \omega \). In cartesian coordinates as before, the
expansion is now in powers of \( w \) and \( \omega \) and in products of these powers. The focusing condition is obtained by requiring coefficients of \( w^m\omega^n \) to vanish as well as those of \( w^m \).

The Johnson-Onaka mounting is derived by requiring the coefficient of \( w^4\omega \) to vanish, the condition which results in the straight line (6.7) above. One may proceed further in this direction. If, in addition the coefficient of \( w^4\omega \) is to vanish, a different straight line results. The two lines become identical provided that \( i - r = 0 \) or \( \pi/2 \). The first of these gives the Eagle mounting again, and the other requires that the grating pole, the entry slit and the exit slit lie on the circumference of a circle. This circle will be coincident with the Rowland circle for one position of the grating. The two (coincident) lines now become the line joining the entry and exit slits, and the pivot may be anywhere on this.
Spectrometers Employing Concave Diffraction Gratings · 89

line. Miyake suggests that a point be chosen to give two positions satisfying the Rowland condition, as illustrated in figure 6.5.

All these mountings depend on the vanishing of coefficients of \( w \), \( w^2 \), \( w^3 \) and of \( w^4 \omega \) and \( w^5 \omega \). No mention has been made of higher powers of \( \omega \). By deliberately choosing the centre of rotation to be at the pole of the grating and requiring \( w^4 \omega \) coefficient to vanish, Seya [1] derived another mounting, which was useful for larger values of \( \omega \) and hence covered a greater spectral range. Again there is much algebra, and it results in a transcendental equation for \( i_0 \), the angle of incidence at the position when the slits are on the Rowland circle. The requirement is that \( \sin^2 i_0 = 1/3 \), so that \( i_0 = 35°15' \). \( r_0 \) has the same value. The optimum value of \( i_0 \) is found by Namioka [2] to be close to this value, his requirement being that there shall be the smallest possible variation of \( \rho \) and \( \rho' \) during rotation. The eventual figures given by Namioka are:

\[ i_0 - r_0 = 70^°15'; \rho = \rho' = 0.8156. R \]

This is the Seya–Namioka mounting.

6.4 Polychromators

If it is necessary to examine several wavelengths simultaneously then a mounting must be chosen that fulfils the Rowland condition at all times. This is not easy. One may mount the grating in a fixed position and move the detectors around the Rowland circle, or the detectors may be fixed and the grating moved. This will require the beam entering the instrument to be large compared to the size of the grating. On the other hand, moving the detectors is frequently incompatible with the vacuum requirements. (Fixing the detectors and moving only the slits is most undesirable, since for this process one normally requires field optics near the exit slit to image the grating on to the detector.)

There seem to have been comparatively few attempts to construct polychromators for laboratory use, and those have made use of the Rowland focusing property. As good an example as any is described by Tousey et al. [10] Polychromators have proved more popular as space-borne instruments, where the design restrictions imposed by the vacuum requirements are removed. The Paschen–Runge mounting appears to be the most popular, with fixed entry slit and grating and a bank of detectors moved around the Rowland circle.

For instruments of this type a study must be made of the full Beutler Eikonal. The region of usefulness of these mountings will depend on the image quality that can be obtained, and this is found to be limited
by astigmatism and coma. The astigmatism limits the useful spectral range, and the coma limits the resolution. The astigmatism is treated by Beutler with a result in a closed form: A point source on the Rowland circle, with angle of incidence $\alpha$ and angle of diffraction $\beta$ will yield an image which is a line of length $z'$, where:

$$z' = l \left[ \sin^2 \beta + \sin^2 \alpha \frac{\cos \beta}{\cos \alpha} \right] = l \Gamma$$

(6.9)

where $l$ is the length of the rulings of the grating.

The astigmatism arises because the Rowland circle (or cylinder rather, if three dimensions are to be considered) is the tangential focal surface of the system. The sagittal surface is a cylinder, similar to the Rowland cylinder, but perpendicular to it.

The question of coma is somewhat more complicated than in simple optical imaging theory. Coma would normally be represented by the term in the Eikonal which was the coefficient of $w. l^2$ where $w$ is the

$y$-coordinate of the point where the ray meets the grating, and $l$ is the $x$-coordinate. There is another factor to consider. It is the disturbance to the chief-ray due to oblique incidence from points at the ends of the slit, a topic discussed at length in the previous chapter. It results in a curved image being produced of a straight slit, or at least it would do were it not for the astigmatism. The combination of this astigmatism and the line curvature results in a widening of the image and gives another coma-like term. The effect is illustrated in figure 6.8.

In practice two separate cases can be distinguished:

(i) The image of a point source on the Rowland circle produced by a grating with a finite length of ruling.

(ii) The image of a long straight slit on the Rowland cylinder produced by a grating with a short length of ruling.

To ease the amount of algebra to be written we use $\Gamma$ as defined in eqn. (6.9) and we define a quantity $\nu$ by:

$$\frac{1}{\nu^2} = (-\sin^2 \alpha + \sin \beta \tan \beta + \sin \alpha \sin^2 \beta \tan \beta) - \sin 2\alpha \sin \beta \tan \beta.$$

Then with $\Delta p$ as the width of the image along the direction of dispersion and $R$ as the radius of curvature of the grating surface, we have

(i) For a long slit on the Rowland cylinder

$$\Delta p = \frac{l^2}{2R \nu^2} \quad \text{and} \quad l = (2R \Delta p)^{\frac{1}{2}} \nu.$$

$l$ is the length of the longest rulings that the grating can have.

**Figure 6.7.** The coma of the concave grating produced by rays from the extremities of rulings of finite length. The departure $\Delta p$ of the marginal focus from the Gaussian focus is described in the text.

**Figure 6.8.** The astigmatic curvature. A fan of rays diverges from an object point on the Rowland circle and is reflected in one of the rulings. The converging conical fan crosses the Rowland cylinder in a short arc $AB$ and fans from other rulings cross through the same short arc to different points on the line $OC$. The result is a short curved image of a point source. Its curvature is called the astigmatic curvature of the grating.
Design of Optical Spectrometers

(ii) For a short slit (or point source) on the Rowland circle:

\[ \Delta \rho = \frac{z^2}{2R} \frac{\Delta \rho}{\rho^2} \quad \text{and} \quad z < (2R \Delta \rho)^{\frac{1}{3}}. \]

\( z \) is the longest slit length that can be used.

These formulae take account of all the significant coma-like terms that follow from the above causes and can be used to compute the useful spectral range of a particular grating in a particular mounting. In practice the useful range is likely to be restricted by other considerations such as loss of efficiency away from the blaze angle. These are discussed further in section 5.4. In each case \( \Delta \rho \) is in centimetres, and the equivalent resolution available is to be obtained by differentiating the grating equation:

\[ \sin \alpha + \sin \beta = \frac{m \lambda}{a} \]

\[ \cos \beta \Delta \beta = \frac{m \Delta \lambda}{a} \]

and with \( \Delta \rho = R \Delta \beta \) we find:

\[ \Delta \lambda = \frac{a}{m} \cos \beta \frac{\Delta \rho}{R} \]

(6.10)

From these formulae one may estimate the performance limitations of Rowland-circle based designs.

6.5 Practical Design Considerations

Most concave grating spectrometers will be required for the vacuum ultraviolet, since it is only there that they show a significant advantage over the Ebert and Czerny–Turner mountings. Constructional methods are conventional, but the vacuum requirements affect the designs slightly in the following ways:

(i) It should be possible to complete the optical alignment under vacuum conditions, or steps must be taken to ensure that no possible misalignment can arise from the application of the vacuum.

This may result in the slits being on the ends of flexible and compressible tubes, or in the whole instrument being built on a baseplate which is then suspended kinematically in a vacuum chamber.

(ii) It must be possible to control and monitor the grating position without disturbing the vacuum. A single control, working through a rotating seal, will meet this requirement, but the control should not be firmly attached to the grating mount. There should be a kinematic linkage somewhere, since the control axle, being attached to the wall of the vacuum chamber, may change its position relative to the grating as the vacuum is applied. Sine-bar drives with spring returns are suitable. The grating position monitor should be attached to the turntable, and a variety of methods will permit the angle to be read accurately. A simple protractor and vernier read through a window will serve many purposes, but it may be desirable in more elaborate instruments to employ coded discs and photocells. All this can be placed in vacuum chamber.

The problem of initial alignment is made somewhat easier if the grating drive will reach far enough that visible spectra can be seen. It is worth extending the drive for just this purpose. All the adjustments of the grating can then be made visually. The alternative – adjustment by ultraviolet light – can be a very protracted operation.

The two obvious alternative methods of construction which present themselves are:

(i) To build the instrument with the vacuum container as the struc-
Design of Optical Spectrometers

tural element which maintains the optical alignment or at least supports the optical components.

(ii) To build the whole instrument complete on a baseplate and insert it in toto into a vacuum chamber.

Of these two, the latter method is to be recommended.

Usually the instrument will require its own vacuum pumping system. This is not the place to stint on pumping capacity, and diffusion pumps with diameters of 6 or 9 in. should be considered in preference to any smaller ones. In the experience of one of the authors, vacuum spectroscopy will provide sufficient experimental difficulties without giving any hostages to fortune.

REFERENCES


BIBLIOGRAPHY


7.1 Theory of the Fabry–Pérot Étalon

The Fabry–Pérot étalon [1] is described in nearly every textbook on physical optics. In practice it usually consists of two parallel discs of glass or quartz, with the inner faces worked optically so that the gap between them is constant over the whole area to better than 1/50th or sometimes 1/100th of a wavelength of green light. These flat faces are coated with some material so that most of the light incident upon them is reflected and only a small fraction is transmitted. Coating materials are chosen to have as small an absorption as possible.

\[
I(\delta) = \frac{I(0)}{1 + \frac{4\pi}{(1 - \pi^2) \sin^2 \left( \frac{\delta}{2} \right)}}
\]
The independent variable in this equation is \( \delta = (2\pi/\lambda)2d \cos \theta \).

where \( \lambda \) is the incident wavelength

- \( d \) is the optical thickness of the gap
- \( \theta \) is the angle of incidence of the light beam in the gap
- \( r \) is the reflection coefficient for intensity of the coatings on the flat surfaces.

By allowing \( \lambda, d \) or \( \theta \) to vary while the other quantities are kept fixed, we can describe with this formula:

- (i) With \( \lambda \) as variable, how the transmitted intensity varies with wavelength when white light is incident from a fixed direction.
- (ii) With \( d \) as variable, how the transmitted intensity of monochromatic light varies as the separation of the plates is changed.

(iii) With \( \theta \) as variable, how the transmitted intensity of monochromatic light varies with angle of incidence.

The Airy function plotted graphically with \( \delta \) as variable appear in figure 7.2. The half-intensity width of each peak on this graph depends on the value of \( r \), the reflection coefficient, and the peaks become narrower as \( r \) increases.

The familiar treatment of the theory of the Fabry–Pérot étalon leads to the description of the ring system formed when light is transmitted after coming from an extended monochromatic source. The rings can be seen at infinity by looking through the étalon or they can be projected on to a screen by a lens.

With \( \theta \) as variable in the Airy function, maxima of intensity are found for \( \sin^2 \delta/2 = 0 \), that is \( \delta/2 = n\pi \) or:

\[
\cos \theta = n\lambda/2d
\]  

(7.2)

**Figure 7.2.** The Airy function.

\[
I(\delta) = \frac{I(0)}{1 + \frac{4r}{(1 - r)^2} \sin^2 (\delta/2)}
\]

with \( \delta = \frac{2\pi}{\lambda}2d \cos \theta \)

and if the rings are projected on to a screen with a lens of focal length \( f \) the diameter of the ring formed in the \( m \)th order of interference is \( D_m \), where:

\[
D_m = 8 mf^2[1 - (n\lambda/2d)]
\]

With \( \lambda \) as variable, maxima are found when \( \lambda = 2d/\pi \) and the graph of intensity against wavelength is as in figure 7.3. Notice that \( \lambda \) decreases from left to right.

**Figure 7.3.** Resolving power and free spectral range. The general variable increases from left to right, but if the function is plotted with \( \lambda \) as variable, then \( \lambda \) decreases from left to right.

This graph is the instrumental profile of the instrument when it is used as a spectrometer. It suffers, like the grating spectrometer, from the overlapping of orders. We can define, in the usual way:

**Resolving power** \( R = \lambda/\delta \lambda \) where \( \delta \lambda \) is the half-intensity width of one of the peaks.

**Free spectral range** \( \Delta \lambda \) being the distance between neighbouring peaks on the wavelength scale.

There is also a further quantity, the Finesse, \( F \) the ratio \( \Delta \lambda/\delta \lambda \), being the number of resolvable spectral elements in one spectral range. As we shall see later, this is a most important parameter and is a function of the reflection coefficient \( r \) and the surface flatness of the plates. \( \delta \lambda \) may be computed by setting \( I = 1/2 \) in the Airy formula. We find that

\[
\frac{2\sqrt{r}}{(1 - r)} \sin \delta/2 = \pm 1
\]

so that there are two values of \( \delta, \delta_1 \) and \( \delta_2 \) satisfying this condition:

\[
\frac{2\sqrt{r}}{(1 - r)} \sin \delta_1/2 = 1 \quad \text{and} \quad \frac{2\sqrt{r}}{(1 - r)} \sin \delta_2/2 = -1.
\]

Subtracting, and remembering that \( \delta_1 \) and \( \delta_2 \) are nearly equal to \( 2\pi\tau \) since they are close to a maximum of intensity, we find:

\[
\frac{2\sqrt{r}}{(1 - r)} \frac{(\delta_1 - \delta_2)}{2} = 2
\]
and if, at the maximum between these two positions \( \delta = (2\pi/\lambda) 2d \cos \theta \) then:

\[
\delta_1 = \frac{2\pi}{[\Lambda + (\delta \Lambda / 2)]} 2d \cos \theta \quad \text{and} \quad \delta_2 = \frac{2\pi}{[\Lambda + (\delta \Lambda / 2)]} 2d \cos \theta
\]

Substituting for \( \delta_1 - \delta_2 \) and putting \( 2d \cos \theta = n\Lambda \) we obtain finally:

\[
R = \frac{n\pi \sqrt{r}}{(1 - r)}
\]  

(7.4)

The free spectral range is found more easily. If at one peak we have \( \delta = (2\pi/\lambda - \Delta \lambda) 2d \cos \theta = 2n\pi \), then at an adjacent peak

\[
\delta = (2\pi/\lambda) 2d \cos \theta = 2(n + 1)\pi.
\]

Eliminating \( \theta \) we get immediately:

\[
\Delta \lambda = \lambda / n
\]  

(7.5)

The finesse then follows from eqns. (7.4) and (7.5) as

\[
F = \frac{\Delta \lambda}{\delta \lambda} = \frac{\pi \sqrt{r}}{(1 - r)}
\]  

(7.6)

Sometimes the Airy formula is written as

\[
I(\delta) = \frac{I(0)}{1 + (2F/\pi)^2 \sin^2 \delta / 2}
\]

and the quantity \( 2F/\pi \) is known as the coefficient of finesse. This is a form suited to the computation of practical systems.

There is now a set of formulae to describe most of the important quantities involved in the Fabry–Pérot spectrometer. They can be amplified with some elementary algebra which is omitted here.

(i) Free spectral range: \( \Delta \lambda = \frac{\lambda}{n} \frac{2d}{2d} = \frac{2d}{n^2} \)

(ii) Resolving power: \( R = \frac{\lambda}{\Delta \lambda} = \frac{n\pi \sqrt{r}}{(1 - r)} \)

(iii) Finesse: \( F = \frac{\Delta \lambda}{\delta \lambda} = \frac{\pi \sqrt{r}}{(1 - r)} = \frac{R}{n} = \frac{R \lambda}{2d} \)

7.2 The Fabry–Pérot Étalon as a Spectrometer

Unlike the grating and the prism, the Fabry–Pérot étalon disperses the light passing through it in two directions, both perpendicular to the optic axis. If after transmission the rings are brought to a focus on a screen, different wavelengths produce rings of different radius. Thus the dispersion is radial, and by cutting a circular hole in the screen some wavelengths can be selected for transmission while all the others are rejected. This hole plays the same part as the exit slit in a grating spectrometer. In general the hole could be annular, but in practice for obvious reasons of simplicity it is an annulus with the inner radius zero. The diameter of the hole determines the luminosity and the resolving power of the instrument.

Let us consider the general case where the hole would be annular. The inner and outer edges should have diameters which correspond to the half-intensity points of the wavelength desired for transmission. If a plot is made of \( I \) against \( \theta \), we obtain figure 7.4. The proper positions for the edges of the stop can be seen, and the angular radii are obtained from the equation for the angular radius of the ring produced by a wavelength \( \lambda \).

\[
\frac{\theta_2}{2} = 1 - \frac{n \lambda}{2d}
\]

then:

\[
\frac{\theta_1^2}{2} = 1 - \frac{n(\lambda - \delta \lambda / z)}{2d} \quad ; \quad \frac{\theta_2^2}{2} = 1 - \frac{n(\lambda + \delta \lambda / z)}{2d}
\]

so that

\[
\frac{\theta_1^2 - \theta_2^2}{2} = \frac{n \delta \lambda}{2d} = \cos \theta / R
\]

and if the annulus is a central hole so that \( \theta_2 = 0 \) and \( \cos \theta = 1 \) then

\[
\frac{\theta_1^2}{2} = \frac{1}{R}
\]  

(7.7a)

If the usual practice is followed of writing \( \theta \) for the angular diameter of the hole, then

\[
\frac{\theta^2}{8} = \frac{1}{R}
\]  

(7.7b)
100 · Design of Optical Spectrometers

As a spectrometer the étalon suffers from the defect mentioned before—the small spectral range and the consequent overlapping of orders. Some sort of order-sorting device is needed just as in the case of the grating spectrometer. In this instance a grating spectrometer itself can be used as the order-sorting device. The effective resolution of the spectrometer can be as low as $\Delta \lambda$, the spectral range of the étalon, and so there is a gain in luminosity. The effective resolution of the combined spectrometer and étalon is $\delta \lambda$, the resolution of the étalon. Thus if the two devices are matched optically the RL product is increased by a factor $F$, the finesse of the étalon plates.

The gain in luminosity or resolving power is thus quite considerable, although not as great as the RL product of the étalon itself.

7.3 Effective Finesse

The finesse given by eqn (7.6) may be called the ‘theoretical’ finesse. It presumes that the field stop is infinitesimal in diameter and that the plates are perfectly flat. Neither of these conditions is satisfied in practice so that the instrumental profile is somewhat different from the theoretical one.

Effect of the Finite Field-stop

If we consider the field-stop to be made up of successive annuli, it is easy to show that annuli of equal area cover the same fraction of an order between their inner and outer radii. Each annulus will contribute an infinitesimal intensity with the wavelength distribution as in figure 7.3, but each contribution will be slightly displaced from its neighbour. The resulting effect will be to convolute the Airy profile with a $\Theta$-function of unit height, and width equal to the fraction of an order covered by the whole stop. This convoluted curve will be the true instrumental profile of the Fabry–Pérot spectrometer. It will represent the apparent profile of a strictly monochromatic line. It corresponds to the instrumental profile of a grating spectrometer when allowance is made for the diffraction limited profile of an infinitesimal slit and the fact that a finite slit width is used. The effect of the finite stop diameter is to reduce the effective finesse, which is now determined as the reciprocal of the fraction of an order corresponding to the radius of the stop. In the limiting case when the stop size is determined by the criterion $\theta_0/8 = 1/R$, the effect on the finesse will be small and can be ignored.

A more serious effect is that due to the unevenness of the gap between the plates.

Effect of Variations in the Gap

If the gap is specified as $d \pm \lambda/50$, we can assume that the gap will be within these limits for most of its area, say 95 per cent. At any point the most probable value of the gap thickness is $d$ and the number of places where it departs from this by a large amount will be very small. The total area where the gap departs from $d$ by an amount between $x$ and $x + dx$ can be expressed in a number of ways. As convenient as any is the Gaussian formula:

$$dA(x) = \frac{bA}{\sqrt{\pi}} \exp \left( -\frac{b^2x^2}{} \right) dx$$

where $b$ is a ‘figure of merit’ which increases as the gap becomes more and more uniform. Thus different parts of the gap will contribute to the total intensity at slightly different orders of interference and the effect reveals itself as a convolution of the theoretical Airy profile with a Gaussian curve. Actually to perform this convolution requires the services of a computer, but this becomes unnecessary if we assume that near a maximum, the Airy profile is approximately the same shape as a Gaussian curve, and if we remember that the convolution of two Gaussians is another Gaussian with a half-intensity width which is the Pythagorean sum of the component half-intensity widths. Now we can determine the reflection coefficient that the plates should have, provided that we know the flatness of them. If they are flat to $\lambda/50$, then the best possible finesse is $F = 50$. This would be with a reflection coefficient of 1-oo and there would be no luminosity. If the reflecting layers have a reflection such that the reflection finesse would be 50, the effective finesse will be $50/\sqrt{2}$, or about 35. The absorption that is encountered with reflecting layers of higher reflectivity is such that the RL product begins to decrease if higher finesse is used. In practice, then, one tends to employ reflecting coatings that would themselves give a finesse equal to the best obtainable due to the gap variations,
and the experimentally determined finesse is 0.7 of this quantity. Again in practice, the problem usually resolves itself into the question of whether to use 5-layer or 7-layer coatings of cryolite and zinc sulphide.

7.4 Luminosity-Resolving Power Product

The luminosity is given by the usual equation:

\[ L = A_1 A_2 / f^2 \]

where \( A_1 \) is the useful area of the étalon plates, \( A_2 \) is the area of the stop and \( f \) is the focal length of the lens.

\( A_2 \) is thus equal to \( \pi \cdot f^2 / 4 \)

and so the luminosity comes to

\[ L = A_1 \cdot \pi \cdot f^2 / 4 \]

Since \( R = 8/\theta^2 \), we have for the RL product:

\[ RL = 2 \pi A_1 \quad (7.8) \]

This is several hundred times greater than the RL product of a grating spectrometer with a grating of the same area.

In practice it is not possible to use the full luminosity of the étalon in a simple way. The light selected by the étalon emerges from a circular stop, while the order-sorting spectrometer needs a long narrow slit. Conversion of the circular stop to a long thin ellipse by an anamorphic lens is no help since, because of the Helmholtz-Lagrange theorem, the beam would emerge from the slit at a very large angular spread in the plane perpendicular to the major axis of the ellipse. The only simple way of matching the étalon to a grating gains only a factor \( F \) in the RL product of the system, over the RL product of the grating spectrometer alone. The focal length of the étalon lens is chosen so as to make the diameter of the central stop equal to the available slit length of the spectrometer.

Because of the low resolution of the spectrometer, a field lens can be used to image the étalon on to the grating, thus making sure that no light is lost. Using the resolution criteria for the étalon and the grating spectrometer, we can find the size of the smallest grating that can be used as an order-sorter. Let:

\[ f_1 \] the focal length of the étalon lens
\[ R_1 \] the resolving power of the étalon
\[ f_2 \] the focal length of the spectrometer collimator
\[ l \] the length of the spectrometer slit
\[ R_2 \] the resolving power of the grating spectrometer alone

\[ \theta = \pi / f_2 R_2 \]

Then

\[ l^2 = f_2^2 \theta^2 = 8 (f_2^4 / R_2) \]

so that

\[ (f_2 / f_1)^2 = (R_1 / R_2) = F \]

The finesse of the étalon.

\[ (f_2 / f_1)^2 \]

A typical value of \( F \) is 25, so that \( f_1 / f_2 = 5 \) and the étalon can be matched to a grating only 1/25th its area.

This is not the optimum size of grating. If a larger grating is used, the focal lengths are chosen to image the étalon on to the grating through the slit in such a way as to fill the grating with light. The width of the slit in relation to the length used is then greater and a larger fraction of the light from the étalon is passed into the grating spectrometer. The
limiting case is when the diameter of the étalon stop is equal to the width of the spectrometer slit. In this case all the light from the étalon passes into the spectrometer and the full luminosity of the étalon is used. This requires in practice an impossibly large grating, and so, in order to obtain the maximum possible luminosity, the largest available grating is used. The crux of the problem is that there is always a greater length of slit available than can be used, and so one aims at as great a slit width as possible. In practice several factors conspire to decrease the luminosity. The figures quoted above ignore the losses of light that occur due to

(i) Inefficiency of the grating.
(ii) The need for some sort of crude filtering with dye filters.
(iii) Losses in transmission through the étalon coatings.
(iv) Losses sometimes incurred due to a large number of transmissions through air-glass surfaces, such as occur in pressure-scanned étalons.

The authors have found in practice that it is as well to divide the theoretical transmission by three when calculating the necessary parameters of such an instrument. The peak transmission through the étalon coatings is likely to be between 0.5 and 0.7 and the efficiency of the grating will be between 0.6 and 0.85. Transmission losses due to air-glass surface reflection can be radically reduced by means of antireflection coatings and the number of surfaces is usually such that this is worth while. It is impossible to lay down any exact rule, since so much depends on the wavelength regions used and the actual values of luminosity and resolution that are required.

The theory given so far assumes that the spectrum to be investigated is unknown and that it may contain wavelengths very close together that must be resolved. If it is known beforehand that the spectrum consists of a number of widely spaced lines (or multiplets within one spectral range) then a large increase in luminosity can be achieved without loss of resolution. The principle can be demonstrated by referring to the extreme case where the spectrum consists of only one line. Then the spectrometer slit can be opened as widely as desired without fear of order overlap. In fact there is no need for a spectrometer at all in this particular case. When several lines are present in the spectrum all that is needed is that the spectrometer shall prevent unwanted lines from passing through the étalon. Hence the slits can be opened until the images of the unwanted lines begin to overlap the image of the line under examination.

7.5 Alternative Methods of Order-Sorting

Interference filters. Filters are available, at the time of writing, with pass bands in the visible region of 10 Å or less, with peak transmissions in the region of 0.8. Filters can be ordered for a particular problem with the centre of the pass band sited at any wavelength to ±3 Å. (It may take several attempts to hit the target, but this is part of the price one pays.) This peak wavelength can be altered to shorter wavelengths by tilting the filter. The basic function of the filter is to convert the spectrum under examination to a line spectrum which consists of a single line about 10 Å wide. The fine structure of this line is what is observed with the étalon. The shape of this line, which is the transmission curve of the filter, can be determined by a conventional spectrometer with white light, but is best done by measuring the spectrum produced by the complete Fabry–Pérot spectrometer when white light is incident on it. Modifications to this profile are then produced by the unknown spectrum, which can be recovered by simple proportionality.

7.6 Multiple Étalon Spectrometers

The interference filter is a type of Fabry–Pérot étalon with a very high finesse and with a very small separation between the plates. It is usually working in first order. Isolation of the experimental range can similarly be achieved by using several Fabry–Pérot étalons in series [2], each one having a different space between the plates. Two methods have been used: (i) The spacers of successive étalons are multiples of the first, so that the instrumental profile is as in figure 7.8. (ii) The spacers are in the ratio of 1 : (1 + a) : (1 + 2a) etc., where a is small compared with the thickness. The value of a will depend on the resolving power required and on the finesse that is available. The optimum value is chosen so that there is good suppression of wavelengths outside the pass-band of the instrument. The major fault of the Fabry–Pérot spectrometer is that its instrumental profile does not fall to zero outside the pass band as does that of a grating spectrometer (apart from scattered light). Consequently, when white light is incident, not only is the desired wavelength passed in full, but all other wavelengths are present in a minute degree. The integrated effect of all these can be large, even larger than the signal from the desired wavelength [13, 14].

The system with the spacers ratioed as above operates on the same principle as the vernier scale. It ensures that if all étalons are chosen to transmit a given wavelength, their maxima will not coincide again until
Design

**Figure 7.8.** The multiple Fabry-Pérot spectrometer. When the spacers are fractions of the largest one, the spectral ranges are multiples of the smallest one. Below is a curve showing the instrumental profile resulting from using two étalons in series, the larger spacer being four times the smaller.

A large number of orders have passed. This will be at a wavelength far removed from the experimental region. The instrumental profile is then as in figure 7.9.

7.7 Scanning in Wavelength

So far we have described the way in which one narrow band of wavelengths can be selected for transmission through the system, while all unwanted wavelengths are removed. In order that the apparatus shall function as a spectrometer it is necessary to be able to change the wavelength at will and in a known way. The three variables in the Airy formula are at our disposal, but since $\lambda$ is the dependent variable, either $d$ or $\theta$ must be altered to produce the necessary change. Changing $\theta$ implies altering the shape of the field-stop or tilting the étalon bodily. The latter method can be used but results in a drastic loss of luminosity. It is easy to show that, given a uniform extended source, the same total flux will be transmitted in each order. If the étalon is tilted the stop will ‘see’ only a small fraction of a ring, and if the off-axis angle is at all considerable it will see a small fraction of a number of rings, so that both resolution and luminosity may be lost. Tilting as a means of changing the wavelength is usually only feasible for an interference filter.

7.8 Pressure-Scanned Fabry-Pérot Spectrometers

Pressure scanning relies on the fact that over a considerable range of pressures, the refractive index of a gas is proportional to the pressure. This is expressed by the equation:

$$\mu = 1 + kp$$
and the constant \( k \) depends on the particular gas. In order to make the available scanning range as large as possible for a given pressure change it is necessary to choose a gas with as large a value of \( k \) as possible. This is a major criterion in the choice of a gas, but is not the only one. The gas must be chemically harmless to the étalon coating and the other materials in the pressure chamber. Air or nitrogen is sometimes used for this reason, but over moderate ranges of pressure propane has been found suitable. It may be unwise to employ propane at low temperatures since its vapour pressure falls rapidly, and not only will the available range of pressures be reduced, but condensation may take place on the étalon plates, with most unfortunate effects on the multilayer coatings. At normal laboratory temperatures it is most satisfactory.

The value of \( k \) for propane is 0.0011/atm. From the equation above we obtain

\[
k p = \Delta \lambda / \lambda_0
\]

where \( \lambda_0 \) is the wavelength transmitted at zero pressure. In a typical case \( \lambda_0 = 5,000 \text{ Å} \) and \( \Delta \lambda = 5.5 \text{ Å/atm} \). Air, having a \( k \)-value of 0.0003, requires correspondingly greater pressure changes to scan through the same spectral range.

The wavelength change per atmosphere pressure change is independent of the order of interference, and thus of the resolving power of the instrument. With propane, pressures of up to 10 atm. can be achieved without elaborate pressure systems, and so a scanning range of 50 or 60 Å can be obtained. Air, having a lower \( k \)-value, is less efficient. High pressure systems are needed for long scanning ranges. As with other materials, the refractive index falls with longer wavelengths, and to operate in the infrared, chambers capable of withstanding several tens of atmospheres pressure difference may be needed. [3]

**FACTORS IN THE DESIGN OF PRESSURE CHAMBERS**

The design is, in general not complicated. There is, as a rule, no need to make the strength marginal, and an adequately robust structure should be chosen. The following design points are worth noting:

(i) The étalon must be attached inside in such a way that distortions of the chamber walls are not transmitted to the étalon.

(ii) There should be provision for the alignment of the étalon from outside the chamber while the pressure inside is still high.

(iii) There should be a leak valve so that the chamber can be flushed out quickly with the refracting gas. This leak should be as far as possible from the entry port.

(iv) The chamber should be easily opened, so that the étalon can be quickly removed for servicing.

(v) Focusing lenses should be outside the pressure chamber, since the focal length of a lens will alter as the gas pressure changes. This is not generally a serious effect, and this criterion is not as stringent as the others. There will be a change in focal length of about 2 per cent at 10 atmospheres of propane.

7.9 Piezoelectric Scanning

This is a way of changing the gap width physically. It relies on the well-known piezoelectric effect which some materials possess, of altering their length when an electric field is applied across two opposite faces. Transducers can be made from rods or annular rings of suitable materials such as sintered aggregates of barium titanate, lead zirconate or a mixture of both. The annular ring, with one plate of the étalon attached to each end, has the merit of simplicity but one must rely on the homogeneity of the material to maintain parallelism during the gap-changing process. Three separate rods, or sandwiches of discs stuck together to make rods, make it possible not only to vary the gap with moderate voltages but also to consider servo-controlling the parallelism and the gap thickness. This has been achieved by Ramsay [4]. The expansion of such a transducer (i.e. the change in length) is proportional only to the voltage applied, because the strain, or change of length per unit length, is proportional to the voltage gradient. The material can be polarized so that the expansion or contraction is in the same direction as the applied field. Since, for a piece of material of thickness \( l \) between electrodes \( \delta l = d(V/l) \), it follows that \( \delta l = d \times V \). The constant \( d \) is quoted by manufacturers in various units. It may be in e.s.u. of charge per dyne. This, when multiplied by 300 gives centimetres per volt. Alternatively \( d \) may be quoted as coulombs/newton and this is the same as metres/volt. Typical values are between 50 and 350 \( \times 10^{-14} \) dyne/e.s.u. or between 1.5 and 10 \( \times 10^{-10} \) metres/volt.

To avoid inconveniently high voltages Ramsay [5] employs sandwiches of ceramic discs. Typical dimensions are shown in figure 7.10. Each face of each disc is coated with a conductive epoxy resin and there are alternate discs of barium titanate and brass, stuck together in a jig under moderate pressure. When set, the whole transducer is encapsulated in resin, and electrical contacts are made by drilling through to the brass discs and inserting wires. Phosphor-bronze is a suitable material for these wires. In order that both ends may be at
earth potential there should be an even number of barium titanate discs and an odd number of brass discs in each complete transducer. With this sandwich construction each transducer can be used, with suitable voltages at suitable points in the stack, to provide a servo-signal, to vary the gap for scanning and to receive a control signal. The power taken by these transducers is very small and they can be regarded as loads of infinite resistance and about 100 $\mu$F impedance. If necessary they can be driven at high frequencies provided that the étalon plates are kinematically mounted. The possibility then arises of rapid scanning (e.g. 10 milliseconds per spectral range) by applying a sawtooth waveform simultaneously to each transducer. Thus digital methods can be applied to this sort of spectrophotometry. The detector output could, for example, be digitized and fed to a core-store, with an address signal coming from the sawtooth generator. The spectral profile can be allowed to accumulate until the observer is satisfied with the signal/noise ratio. Alternatively the traditional method can be used quite simply, with a slow time-base type of circuit to provide both the piezoelectric signal and the signal to one axis of an X–Y recorder, while the detector provides the other signal.

The servo-control of the parallelism of the gap developed by Ramsay solves the major problem which has hitherto faced the designer of Fabry–Pérot spectrometers, namely, the stability of the system. Previous to this, much ingenuity has been used by very many people to make Fabry–Pérot étalon assemblies which were relatively insensitive to temperature and mild shock. These could often be relied on to stay in adjustment for hours or days, but only permanently mounted étalons required no attention at all. These are of course not adjustable in gap for different experiments. For many purposes in the laboratory it will not be worth while going to the elaboration of a complete servo system, and the cylindrical piezoelectric transducer will be an adequate method for scanning in wavelength. The gap can be maintained separately by mechanical means.

7.10 Magnetostrictive Scanning

This offers similar advantages to piezoelectric scanning. Current has to be provided instead of voltage and this raises problems of Joule heating and heat dissipation. The principle is that if a coil is wrapped around a bar of material such as nickel or a nickel–iron alloy, then a current passing through the coil will cause the length of the bar to change. The change may be positive or negative, depending on the material of the bar, and the effect is non-linear and even reverses at high field strengths. Saturation effects also occur. Dupeyrat [6] has reported on experiments with magnetostrictive scanning of a Fabry–Pérot étalon, and quotes a total scan of eight orders at 5461 Å with 600 amper-turns around a nickel bar 10.7 cm in length. This was the total scan before saturation occurred. More recently, Slater, Betz and Henderson [7] have described a complete scanning spectrometer using magnetostriction, which has been successfully used in the field as well as in laboratory conditions.

The chief advantage that magnetostrictive scanning has over piezoelectric scanning is in the case where it is desired to scan through many orders. It might well happen then that the voltage or the number of layers needed in a piezoelectric sandwich becomes excessive.

**MECHANICAL SCANNING METHODS**

The parallel strip spring method of producing parallel motion has been thoroughly investigated by R. V. Jones [8] and has been applied to the motion of Fabry–Pérot étalon plates by Bradley and Tolansky [9]. As with many other mechanical methods, it can be made to work provided it is made carefully enough, and if the materials for manufacture are properly chosen. The mechanical tolerances are generally those of angle. During the motion the moving plate may be allowed to tilt with respect to the fixed one by an angle which does not allow the gap to depart from a constant value by more than $\lambda/2F$, where $F$ is the
observed finesse. Thus the tolerance is much stricter for a 10-cm diameter étalon than for a 2-cm diameter étalon. Clearly if high resolution is required and there is no shortage of light a small diameter will be used. With a 10-cm diameter étalon the tolerable angle change is likely to be 0.02° of arc, and this is difficult to maintain reliably. The question of reliability is an important one. If the étalon is readily accessible so that it can be inspected and adjusted easily the reliability is less important than in the case, say, of an instrument to be sent up in a rocket. Thus simple mechanical scanning is not likely to be chosen for large étalons which are to be used in arduous conditions where accessibility is poor and thermal control may be difficult. They will be even less likely to be considered now that a reliable servo mechanism exists for automatic control of the gap. Nevertheless the parallel strip spring motion, within these limits, is cheap and easy to make.

More elaborate forms of mechanical scanning have been devised, for example that of Chabbal and Soulet [10] and also a hydraulic system [11] in which the moving étalon plate is mounted on a deformable disc, and the disc is made to be the lid of an oil-filled metal cylindrical box. Parallelism is achieved by ‘figuring’ the lid with emery cloth until uniform motion has been achieved.

All these methods of mechanical scanning, with the exception of the parallel strip springs, are more or less complicated and delicate and in the design, construction and alignment are, in the end, as complicated and costly as a servo-control mechanism.

7.11 Mounting the Étalon Plates

The actual attachment of the étalon plates to the mechanical part of the apparatus needs some care. The classic method was to manufacture a sandwich consisting of the étalon plates as the bread, and a spacer, or a set of three spacers, as the jam. For large separations (>2 mm) a quartz ring was common, with raised sections at intervals on its edges as in figure 7.11. The gap was then determined by the thickness of this spacer. The raised flats were polished so that, if a speck of dust was inadvertently trapped between the spacer and the plate, a set of Newton's rings could be seen. It is more common to use three separate spacers, especially for large-diameter étalons, and for spacings from 5 mm down to 0.1 mm ball-bearings or watchmakers' end-stones are suitable. From a gross of ball-bearings it is possible, with some patience, to select three which have the same diameter to ±λ/2 of green light. If end-stones are used, these will need grinding and polishing with carborundum and rouge, using the étalon plates as the test-pieces for equal thickness. The sandwich is held together in an annular metal container, the lower plate of the étalon resting on three balls and the upper plate kept in contact with the spacer (and hence the lower plate) by pressure applied through three further balls. It is most important, if this method of assembly is used, that the lower ball, the spacer and the upper ball should be collinear, otherwise the étalon plates are distorted, and this distortion appears (very easily) as non-uniformity of the gap.

Final adjustment of the parallelism of the gap is then achieved by changing the pressure on the three top balls, using reduction levers or springs controlled by screws. There must be some care to ensure that turning the screw or otherwise making an adjustment does not cause the ball to slide over the upper surface. It must be constrained to move in a direction normal to the surface. A typical adjustment device is shown in figure 7.12. Almost everyone who has built his own étalon has devised a different adjustment. All of them work, but some work better than others.

This method of assembling an étalon would be considered by some modern workers as a rather barbarous technique. The alternative is to mount each plate separately in its own metal ring, and then to attach the rings to a rigid mechanical frame of some sort, in such a way that small adjustments to the gap thickness and to the parallelism can be made through rigid and stable levers. Since there is no question of pressing or distorting the plates themselves there is no need for springs to transmit the motion. The procedure for mounting is similar to that for mounting mirrors and lenses. It is useful to have a shoulder on the plate so that constraints can be applied in all directions without interfering with the flat surface. Alternatively the mounting can be purely
kinematic, with three ‘cat’s ears’ of quartz or glass as needed, fixed to the rim of the plate with an epoxy resin. This is a process that must be carried out during the manufacture of the plates, with particular care to see that the radius of the ‘cat’s ear’ matches the radius of the plate. The three ‘cat’s ears’ should have a hole, slot and a plane which match into three small ball-bearings. The technique is described by Ramsay [12]. It is a matter of mild controversy as to whether the plates of the étalon should have their optic axis vertical or horizontal, and there is authoritative support for both factions. When a spacer is used it probably does not matter very much, but if the plates are held separately then the question arises of how to hold them firmly so that they do not fall out, without at the same time gripping them so tightly that they distort. This is when a shoulder is useful, so that a vertical restraining force can be applied at points other than the flat surfaces.

7:12 Reflecting Coatings

In the early days of the Fabry-Pérot étalon the usual coatings were silver, deposited chemically. Later these gave way to the lower reflection but better durability of aluminium. The reflection coefficient from either of these materials is not particularly high by modern standards, being in the region of 0.8 to 0.88, depending on who makes the coatings. The finesse that is obtainable is not, in practice, likely to be greater than about 18. The situation is much improved by the use of multi-layer coatings. These consist of alternate layers, each of $\lambda/4$ optical thickness usually of zinc sulphide and cryolite. Five such layers (three of zinc sulphide, two of cryolite) give a reflection coefficient of approximately 0.90, while seven such layers give 0.94. The corresponding finesses are 30 and 50.

These high reflectances are obtained only in the region for which the optical thicknesses are $\sim \lambda/4$. Dispersion in the materials of the layers will restrict this range to about 250 Å on either side of the design wavelength, so that it is not possible to make a Fabry-Pérot spectrometer that will adapt to any part of the spectrum. The peak transmission through an étalon with these layers is likely to be about 0.7. If there is no shortage of light in the source it is possible to achieve a much higher finesse at the cost of low transmission.

REFERENCES

8.1 Multiplex Systems and the Multiplex Advantage

The term 'multiplex' has been borrowed from communications theory, where it is used to describe a certain way of transmitting information. In spectrometry it denotes a type of spectrometer in which one detector simultaneously receives signals from different elements of the spectrum, and encodes these signals in such a way that they can be transmitted and recorded by a single information channel. The advantage of doing spectrometry by this rather complicated method is that noise is introduced into the system only by a single detector, rather than by all the detectors that would be needed for separate simultaneous recording of all the elements of the same spectrum. In particular the infrared part of the spectrum becomes more accessible, since the signal/noise ratio available from a conventional single-channel spectrometer can become very low indeed at long wavelengths.

Two methods are recognized for the multiplex transmission of signals. They are called 'time multiplexing' and 'frequency multiplexing'. One may envisage an idealized communications system in which eight microphones have their leads attached to points on the rim of a disc, and a pointer rotates at high speed, connecting each of them in turn to a single telephone wire. A similar pointer at the receiving end routes each message into one of another eight channels and so to one of eight earphones. Thus eight messages are transmitted simultaneously. This is time multiplexing. We shall see later that for all the information in each message to be transmitted, the pointer must rotate at a rate equal to twice the highest frequency that is present in the messages.

Frequency multiplexing requires each message to modulate a different carrier frequency, the messages being decoded at the other end by passing the whole combined signal through an appropriate electrical filter. Different communications channels will then be kept separate by different filters. The type of multiplexing found in spectrometry is an extreme case of frequency multiplexing, where each element of the spectrum requires a different carrier frequency, but each message consists of only one number - the intensity of that element.

The classic example of a multiplex spectrometer is the Michelson interferometer in which one of the reflectors is able to move at a steady speed or in a series of small jumps, so as to change the interferometer path difference at a uniform rate. The elementary theory of what happens is as follows:

In figure 8.1 let the incident radiation have amplitude $A$ and wave-number $\nu$ at the moment when it reaches the beam-splitter. After division there will be two beams, one transmitted and one reflected, with amplitudes $(A/\sqrt{2}) \exp(i\phi_1)$ and $(A/\sqrt{2}) \exp(i\phi_2)$ where $\phi_1$ and $\phi_2$ are the phase changes on reflection and transmission.

After passing along the two paths the beams return to the beam-splitter, where the amplitudes before recombination will be:

$$(A/\sqrt{2}) \exp(i\phi_1 + 2\pi\nu d_1)$$

and

$$(A/\sqrt{2}) \exp(i\phi_2 + 2\pi\nu d_2)$$

where $d_1$ and $d_2$ are the paths in the interferometer arms.

Part of the beam which was transmitted will now be reflected and vice versa, so that the resultant transmitted amplitude will be:

$$(A/2) \exp i(\phi_1 + \phi_2) \exp(2\pi\nu d_1 + 2\pi\nu d_2)$$

and this, multiplied by its complex conjugate, will give the resultant transmitted intensity, which is:

$$I(\Delta) = (A^2/2)[1 + \cos(2\pi\nu\Delta)]$$

(8.1)

where $\Delta$ is the difference in the lengths of the two paths traversed by the two beams through the interferometer.
Thus the transmitted intensity varies sinusoidally with $\Delta$. The amplitude and frequency of this wave depend on the intensity and wave-number of the incident radiation.

If the input radiation consists of a number of different wave-numbers, each with its own intensity, the output will be a number of cosine waves, each of different amplitude and period. The detector will measure the sum of all these. If, in the interval $\nu \to \nu + d\nu$ the intensity is $S(\nu) \cdot d\nu$, then the resultant output will be:

$$I(\Delta) = \frac{1}{2} \int S(\nu) \cdot (1 + \cos 2\pi \nu \Delta) d\nu$$

$$= \frac{1}{2} \int S(\nu) d\nu + \frac{1}{2} \int S(\nu) \cdot \cos 2\pi \nu \Delta \cdot d\nu.$$  \hspace{1cm} (8.2)

The quantity $I(\Delta) = 2I(\Delta) - I(0)$ is now recognizable as the Fourier pair of $S(\nu)$. If $I(\Delta)$ can be measured, then $S(\nu)$ can be recovered by one means or another, usually by getting a digital computer to perform the transform. $I(\Delta)$ is called the interferogram.

This is the core of the multiplex system. Half of all the radiation passing into the instrument is collected by one detector, and the detector signal can be recorded and processed at leisure. The instrument looks at all the elements of the spectrum all the time, instead of dividing its time equally among the elements.

If there are $N$ elements in the spectrum the instrument will show an improvement of the order of $\sqrt{N}$ in its signal/noise ratio, as compared to a single channel spectrometer of the same luminosity and resolving power. This figure of $\sqrt{N}$ is in the general case, where there is no a priori knowledge of the shape of the spectrum. The gain would be much less, for example, if the spectrum consisted of a small number of emission lines whose position was roughly known. The single channel spectrometer can skip the parts of the spectrum that are known to be empty, while the multiplex spectrometer must plod on through all the path differences from 0 to the maximum.

This signal/noise advantage is termed the multiplex advantage or the Fellgett advantage. It was first pointed out by Fellgett and reported in print by Strong [1].

In addition to the multiplex advantage (which only holds when detector noise is the principal source of noise in the system) the Michelson spectrometer has the same efficiency or $L-R$ product as a Fabry–Pérot étalon. Thus in the visible region it finds its application

in the examination of faint extended light sources, a much higher ratio of efficiency/cost being possible than for a Fabry–Pérot spectrometer.

The chief application has however been hitherto in the infrared, especially the far infrared where radiation fluxes from luminous sources are very low and detectors are very noisy. Here the multiplex advantage is most noticeable. The performance of detectors is such that in the visible region photon shot noise is the principal source of noise in the system, and the multiplex advantage disappears. A further development, the field-compensated Michelson spectrometer [2], promises greater luminosity still, comparable with that of a Fabry–Pérot spectrometer with plates of twenty times its diameter.

### 8.2 The Resolving Power

The elementary theory supposed that the incoming radiation was a beam parallel to the optic axis. In practice the instrument will have a collimator and therefore a finite field. The radius of the field-stop will limit the resolution in the same way, and for the same basic reason, as in the Fabry–Pérot spectrometer. Consider a ray from the edge of the field-stop, passing through the centre of the collimator lens. Let it make an angle $\theta$ with the optic axis. The path difference between the two rays through the instrument will be $\Delta \cdot \cos \theta$, where $\Delta$ is the path difference for paraxial rays. The modulation of a paraxial ray of wave-number $\nu$ is described by $\cos 2\pi \nu \Delta$, and the modulation of a marginal

![Figure 8.2](image_url)  
**Figure 8.2**. Paraxial and marginal rays in a Michelson interferometer. Not only is the path-difference dependent on the angle that the chief-ray makes with the axis, but as the path-difference increases there is successively more shearing of the wavefront against itself at the plane of recombination.
ray will be \( \cos(2\pi v\Delta \cos \theta) \). This rate is the same as that for a paraxial ray of wave-number \( v \cos \theta \). Hence if the angular radius of the field stop is \( \theta \), the instrument will be unable to distinguish between radiation of wave-number \( v \) and of wave-number \( v \cos \theta \). The resolution is then \( \Delta v \approx \Delta(\theta^2/2) \), and to translate this to the usual terminology, where \( \theta \) describes the field diameter:

\[
\frac{\theta^2}{8} = \frac{1}{R} \tag{8.3}
\]

8.3 Truncation and the Sampling Theorem*

We cannot know the exact value of \( J(\Delta) \) at all stations \( \Delta \). This would require an infinite amount of information. We inquire therefore how little information about the interferogram we need in order to be able to recover the spectrum. Neither can we extend the path difference to infinity, and so we inquire what will be the effect on the spectrum of terminating the interferogram.

The effect is, mathematically, to multiply the interferogram by a \( \Theta \)-function, and it follows that the spectrum that is recovered will be the convolution of the true spectrum with a sinc-function. If the maximum path difference is \( \Delta_1 \), the spectrum will be convoluted by \( \text{sinc}(2\pi v\Delta_1) \), thus limiting the resolution to \( \Delta v = 1/(2\Delta_1) \).

**THE SAMPLING THEOREM [3]**

Consider a spectrum \( S(v) \) extending from \( v = 0 \) to \( v = v_0 \). We suppose that there is no energy present at wave-numbers greater than \( v_0 \). Consider also a periodic function \( S_p(v) \) with the property that it is identical to \( S(v) \) in the range \( 0 \rightarrow v_0 \); is the mirror image of it in the range \( v_0 \rightarrow 2v_0 \), and thereafter is periodic with period \( 2v_0 \). Thus \( S_p(v) \) is symmetric and can be represented by a Fourier series of cosines:

\[
S_p(v) = \sum_{n=-\infty}^{\infty} a_n \cos \frac{n\pi v}{v_0} \tag{8.4}
\]

and by the inversion theorem:

\[
a_n = \frac{1}{v_0} \int_{-v_0}^{v_0} S_p(v) \cos \frac{n\pi v}{v_0} \, dv
\]

*This is more properly known as the cardinal theorem of interpolary function theory. The name ‘sampling theorem’ appears to have been coined by communications engineers.*

In the interferogram of \( S(v) \) we have, after subtracting the zero path difference term:

\[
J(\Delta) = \int_{-v_0}^{v_0} S(v) \cos 2\pi v\Delta \, dv
\]

and the integral extends only as far as \( v_0 \) since the integrand is zero from there on.

Clearly \( J(\Delta) \) is the same as \( a_n \), apart from a constant, at all points \( \Delta \) where \( \Delta = n/2v_0 \) (\( n = 1, 2, 3, \ldots \)).

If we write

\[
\Delta_0 = 1/2v_0 \quad \text{then} \quad J(n\Delta_0) = k_n a_n
\]

and so if the interferogram is measured (‘sampled’) at points \( 0, \Delta_0, 2\Delta_0, 3\Delta_0, \ldots \), \( S_p(v) \) can be reconstructed. As the maximum path difference is increased, that is, as more samples are taken, more harmonics are added to the sum for \( S_p(v) \), and so higher resolution is achieved.

Suppose that \( N \) samples are taken, apart from the one at \( \Delta = 0 \). Then the spectrum that is recovered is:

\[
S_p(v) = \frac{J(0)}{2} + \sum_{n=1}^{N-1} J(n\Delta_0) \cos \left( \frac{n\pi v}{v_0} \right) + \frac{J(N\Delta_0)}{2} \cos \left( \frac{N\pi v}{v_0} \right) \tag{8.5}
\]

The maximum path difference is now \( N\Delta_0 = \Delta_1 \), so that \( v_0/N = 1/(2\Delta_1) \) is the width of one resolved element in \( S_p(v) \). The \( m \)th point on the

---

*This is a detailed analysis of the principles underlying optical spectrum analysis, focusing on the sampling and reconstruction of interferograms. The mathematical derivations include the use of Fourier series and integrals to describe the relationship between the interferogram and the spectrum.*
Design of Optical Spectrometers

The spectrum will be \( S_\delta(mv_0/N) \), given by the sum:

\[
S_\delta(mv_0/N) = \frac{J(\omega)}{2} + \sum_{n=1}^{N-1} J(n\Delta_\delta) \cos \frac{2\pi mn}{N} + \frac{J(N\Delta_\delta)}{2} \cos m\pi \quad (8.6)
\]

This expression applies to any number of samples chosen.

Therefore:

1. Samples must be taken at intervals of path difference equal to the reciprocal of twice the highest wave-number in the spectrum. In telecommunications theory, where the sampling theorem applies to frequency instead of wave-number, this highest frequency is called the 'Nyquist frequency' of the signal. Here we might refer to the 'Nyquist wave-number' of the interferogram.

2. The number of resolved elements in the spectrum is the same as the number of samples taken.

8.4 Aliasing

Suppose that the spectrum does not in fact extend as far as \( v_0 \). Then part of the recovered spectrum will be empty, and there will be fewer resolved elements in the non-empty part. This is a wasteful use of time, and a larger sampling interval could have been chosen. The same is true if the spectrum extends from \( v_1 \) to \( v_2 \). A periodic function \( S_\delta(v) \) may now be defined such that the actual spectrum occupies a half-period which is not the first. The possibility is illustrated in figure 8.5.

![Figure 8.4](image1)

**Figure 8.4.** Examples of spectra and the corresponding interferograms.

(a) Monochromatic source \( S(v) \) and cosine wave interferogram.
(b) Two monochromatic lines. The interferogram exhibits 'beats'.
(c) Bounded 'White' spectrum. The interferogram is a sine-function multiplied by a cosine wave of period corresponding to a line at the mid-point of the spectrum.
(d) Typical natural spectrum. The interferogram is apparently different from the mean level only near the origin. Nevertheless most of the information content is out in the wings.

![Figure 8.5](image2)

**Figure 8.5.** A spectrum of the same shape as in figure 8.3, but confined between the limits \( 2v_2 \) and \( 3v_2 \). The corresponding interferogram is equal to the previous one multiplied by a cosine curve. This interferogram touches the original one (shown dotted) at the points where the samples are taken, so that the same set of samples results in both cases, and the same periodic spectrum \( S_\delta(v) \) is recovered.

Samples of the interferogram are now taken at intervals \( \pi(v_2 - v_3) \) and the spectrum, or rather \( S_\delta(v) \), is reconstructed as before.

This is possible because the interferogram of \( S(v) \) will be the same at the sample points as that of \( S_\delta(v) \), and the same as the interferogram...
of a fictitious spectrum of the same shape but occupying the first half-period. To demonstrate this:

Suppose that this fictitious spectrum is $S(v)$ and its interferogram is

$$J_f(\Delta) = \int_0^\infty S(v) \cos 2\pi v \Delta \, dv.$$ 

$S(v)$ is related to $S(v)$ by $S(v - v_1) = S(v)$, hence by the shift theorem:

$$J_f(\Delta) = J_f(\Delta) \exp (2\pi iv_1 \Delta)$$

and at the sample points, where $\Delta = n\Delta_0$

$$J_f(n\Delta_0) = J_f(n\Delta_0) \exp \left[ \frac{2\pi iv_1 n}{2(v_1 - v_2)} \right].$$

$J_f$ and $J$ are identical provided that

$$(v_2 - v_1) = v_1/2a \ (a = 1, 2, 3, \ldots) \quad (8.7)$$

Thus the spectrum that is being examined must be confined to an interval $v_1 - v_2$ which fulfills this condition. In practice $v_1$ and $v_2$ will be chosen to make $a$ as large an integer as possible.

![Figure 8.6](image.png)

**Figure 8.6.** The effect of various sampling intervals on the shape of the reconstructed spectrum.

(a) Original spectrum.

(b) Reconstruction after using sampling intervals which are too large.

(c) Reconstruction after using sampling intervals according to the requirements of the sampling theorem.

(d) Reconstruction after using a sampling interval smaller than required by the sampling theorem.

Referring back to the periodic function $S_c(v)$, it is impossible to tell, when reconstructing the spectrum from the samples, in which period the spectrum lies. Each possible spectrum is said to be an **alias** of the first period, and this first period is known as the **principal alias** of the series.

If the above confining condition is not fulfilled, the true spectrum may occupy parts of two periods. Because they are indistinguishable when reconstructed, they will both appear in every period when the interferograms are processed. Thus, if the sample interval $\Delta$ has been made too long, a situation will arise where the true spectrum will overflow its period, and the overflow must then appear in every period. This will result in the overflow appearing to be 'folded back' into the spectrum to give the effect shown in figure 8.6. This is normally quite undesirable, but may be tolerable in practice if the range cannot be reduced by optical filtering, and if a detailed study is to be made of a small portion of the spectrum at the end of the period remote from the folding.

### 8.5 Practical Methods for Fourier Transformation

The measurement of interferograms and the reconstruction of the spectrum has received a great deal of study. A primitive way would be by direct recording, on a recording voltmeter or ammeter, of the detector output, the measurement of the sample heights by hand, and the computation of the necessary cosines. This is not practical for high-resolution spectrometry, and automatic methods are normal. Although some work has been done on analogue computers for doing Fourier transforms, the normal method is to do a numerical computation using a digital computer.

If it is certain that the interferogram is symmetric, so that $J(\Delta) = J(-\Delta)$, the spectrum is simply recovered by forming, for each required spectral point, the sum:

$$S(\frac{mv_0}{N}) = \frac{J(0)}{2} + \sum_{n=1}^{N-1} J(n\Delta_0) \cdot \cos \left( \frac{2\pi vn}{N} \right) + J(N\Delta_0) \cdot \cos \frac{2\pi vn}{N}.$$ 

In practice it is rare that one can be certain that the interferogram has been sampled correctly, or that one particular sample is known to be at exactly zero path difference. Errors in sampling position are known as phase errors, and with a well designed spectrometer the phase error will be the same for each point. The experimental data, which should have consisted of samples $J(0), J(\Delta_0), J(2\Delta_0), \ldots$ will
in fact consist of $J(x_1), J(x_1 + \Delta y), J(x_1 + 2\Delta y), J(x_1 + 3\Delta y), \ldots$. The spectrum can be recovered by doing both sine and cosine transforms so that the spectrum is $S(v) = \sqrt{(S_x(v)^2 + S_y(v)^2)}$. This is a time-consuming process by normal computation methods, but the Cooley-Tukey algorithm [4] makes it possible to write computer programs with which the computation time is proportional to $N \log N$ instead of the normal $N^2$. The algorithm shows a way of computing the complete Fourier transform with the complex kernel, but is probably shorter than the more pedestrian programs even when only a cosine transform is needed. The program for computation requires that the number of interferogram points should be $2^m$ where $m$ is an integer. If this happens not to be the case in a given instance, the deficiency can be made up by adding zeros.

When more conventional programs are to be written, the need to compute both sine and cosine transforms can be eliminated by a subsidiary program which first of all finds the phase error and then computes, by interpolation, a new interferogram which is symmetric. The cosine transform can then be applied to this.

If $S_x(v)$ is the periodic equivalent of the spectrum, as before, then $J(\Delta)$, the Fourier transform of $S(v)$ is identical to the $n$th coefficient of the Fourier series for $S_x(v)$ at the point $\Delta = n\Delta_0$. The interpolation theorem [5a, 5b] then says that if we know $J(\Delta)$ at the points $n\Delta_0$ we know all there is to know about $S_x(v)$ and therefore about $S(v)$. Hence we can construct $J(\Delta)$ at all points $\Delta$ from a knowledge of the values at $n\Delta_0$. The theorem states:

$$J(\Delta) = \sum_{n=-\infty}^{\infty} J(n\Delta_0) \cdot \text{sinc}[2\pi v_0(\Delta - n\Delta_0)]$$

(8.8)

where $v_0$ is, as before, the highest wave-number present in the spectrum.

If the true centre of symmetry of the interferogram is at $e$, then the required new centre interferogram points are at $n\Delta_0 + e$, where the experimental points are at $e, \Delta_0, 2\Delta_0, \ldots$. By the interpolation theorem:

$$J(m\Delta_0 + e) = \sum_{n=-\infty}^{\infty} J(n\Delta_0) \cdot \text{sinc}[2\pi v_0(m\Delta_0 + e - n\Delta_0)]$$

(8.9)

Fortunately the sinc-function decreases rapidly away from zero, and it is necessary in practice to compute only a few terms of this series to obtain each new point. Before this can be done, the quantity $e$ must be found. Here again, points on either side of zero path difference must be used. The sine and cosine transforms of the observed interferogram are computed using only a few data points, to give:

$$S_x(v) = \frac{1}{2} \sum_{n=-\infty}^{\infty} J(n\Delta_0) \cdot \sin \left(\frac{\pi m n}{N}\right)$$

$$S_y(v) = \frac{1}{2} \sum_{n=-\infty}^{\infty} J(n\Delta_0) \cdot \cos \left(\frac{\pi m n}{N}\right)$$

and the phase error $\epsilon$ is obtained from:

$$\tan^{-1}\left(\frac{S_x(v)}{S_y(v)}\right) = \frac{2\pi v_0}{N} e = \frac{\pi m}{N\Delta_0} e$$

(8.10)

If the only error in the interferogram is the phase error, then $\epsilon$ will be a constant. If it is found to be dependent on $m$, that is on $v$, then there is a dispersion error, and the interferogram is genuinely not symmetric.

A further alternative method of recovering the spectrum is to compute the autocorrelation function, which is necessarily symmetric, and then to perform a cosine transform on it to give $[S(v)]^2$. The autocorrelation function is normally defined by:

$$A(\Delta') = \lim_{\Delta \to \infty} \frac{1}{\Delta} \int_0^{\Delta} J(\Delta) \cdot J(\Delta + \Delta'). \, d\Delta$$

(8.11)

and can be computed to sufficient accuracy from the interferogram data by:

$$A(n\Delta_0) = \sum_{m=0}^{\infty} J(m\Delta_0) \cdot J((m + n)\Delta_0)$$

(8.12)

### 8.6 Apodizing Functions

An advantage of this technique is that apodization can be done after the transform has been made. Blackman and Tukey [6] discuss various ways of doing this, known as *hannings*, *hamming*, etc., when applied by communications engineers. Truncating functions can be chosen to do various tricks, such as making the total area outside the main lobe a minimum, or making the first subsidiary maximum as small as possible.

Typical of these apodizing functions ('lag-windows' in the Blackman and Tukey terminology) are:

(i) The *top-hat* which is the window as given by the raw data.

(ii) The *triangle*, which results from the multiplication of the data by $(1 - |[x/N]|)$ at the $x$th sample, when there are $N$ sample points
between zero path difference and the maximum. The effect on the
spectrum is to convolute it by \( \text{sinc}^2 \) instead of \( \text{sinc} \).

(iii) Functions of the general type \([A + B \cos(\pi x/N)]\). These have
the effect of redistributing the spectrum in the secondary maxima,
and by suitable choices of the constants \( A \) and \( B \), the first subsidiary
maximum can be made very small or the signal can be equally dis-
tributed among the first six or seven subsidiary maxima. Blackman
and Tukey give two examples with \( A = B = 0.5 \); and with \( A = 0.54,
B = 0.46 \). The point now is that the interferogram need not be treated
with these functions. The transform is computed first, to give a set
of points in the spectrum \( S(0), S(1), S(2), \ldots, S(N) \). A new spectrum
is then computed from this by forming at each of the points \( 1, 2, 3, \ldots, N \)
the sum:

\[
S'(i) = A S(i) + B/2 [S(i - 1) + S(i + 1)]
\]  

(8.13)

Thus at the trivial cost of losing the first and last points of the
spectrum the apodizing is achieved at very small expenditure of com-
puter time. On the same set of data different values of \( A \) and \( B \) can be
tried until the best apodization for the particular job is found.

8.7 Practical Details of Multiplex Spectrometers

The theory given above is based on the Michelson interferometer as
the device for producing the interferogram. This instrument and its
derivatives are the most important devices for producing interferograms
but they are not the only ones. We must mention in addition and in
particular the 'Mock interferometer' of Mertz [7], and the 'coded disc'
spectrometers of Golay [8] and of Braddock and Wilcock [9]. These
instruments all provide ways of imposing modulating waveforms on
the spectrum elements, the different waveforms being members of an
orthogonal set. By far the greatest ingenuity and time has been spent
on the development of the basic Michelson interferometer, and it now
has numerous progeny, which have little in common except that they
are two-beam interferometers and are able to produce usable interfero-
grams.

The technique was first used in its entirety by Fellgett [10] and many
workers have since made multiplex interferometers for the far infrared,
where the multiplex advantage is greatest and the mechanical difficulties
are least. In the near infrared and the visible regions the basic Michelson
interferometer is unsuitable for mechanical reasons and it is here that
the greatest effort has been needed. Pre-eminent among the workers
in this field are P. and J. Connes, who have not only contributed most
of the basic theory [11] but have been the first to devise and use
practical machines for high-resolution spectrometry in these difficult
spectral regions [12].

We consider here that the 'far infrared' is the region of wave-numbers
below \( 500 \text{ cm}^{-1} \) and in this region the classical Michelson instrument
can be used. One of the mirrors can be displaced mechanically to vary
the path difference. Such instruments are mechanically straightforward
(although they provided enough headaches to their inventors at the
time!). The moving carriage may ride kinematically on plate-glass ways,
or on precisely ground steel cylinders. The movement can be controlled
by a mechanical link (e.g. a wobble-pin) to a micrometer screw. The
beam-splitter may be a thin film of a suitable plastic material such as
Melinex, stretched on an 'embroidery frame' mounting. The thickness
of the film must be chosen with care, since there is reflection both from
the front and rear surface. These reflections are coherent as usual in
thin films, and for some wavelengths the film will be an optical \( \lambda/4 \)
layer. Brier [13] has investigated the properties of beam-splitters with
this order of optical thickness, and finds that the performance for one
state of polarization is nearly perfect over part of the wavelength range,
but that it is much worse for the other polarization state. A reasonable
efficiency is maintained in the region from \( 50 \mu \) to \( 200 \mu \).

No suitable coating materials appear to have been found which will
improve the performance of these beam-splitters.

Refinements of this basic mechanical method of scanning have been
used to construct interferometers for wave-numbers as high as \( 2,000 \)
\( \text{cm}^{-1} \), but the mechanical tolerances become extremely severe. The
difficulty is in maintaining the optical alignment over large ranges of
path difference. Several inventions have been made to overcome this
trouble. Some replace the plane mirrors with different optical com-
ponents, while others involve a different way of changing the path
difference. In historical order, as applied to Fourier spectrometry they are:

(i) The 'Cube-corner' reflector [14].
(ii) The 'Cat's-eye' reflector [12].
(iii) The 'Tilting Michelson' interferometer [15].
(iv) The 'Möbius-band' interferometer [16].
(v) The 'Field-compensated' interferometer [17].

(i) The cube-corner, as its name implies, is an assembly of three
plane mirrors, each accurately perpendicular to the other two. One arm
Design of Optical Spectrometers

of a Michelson interferometer may have the usual plane mirror, and the other may contain the cube-corner and a second plane mirror as in figure 8.7. It is necessary that the cube-corner should be passed twice in order to avoid inverting one of the returning wavefronts with respect to the other. Since there are consequently six reflections at the cube-corner it is necessary that the mirror surfaces there should be flat to \(\lambda/30\). The plane mirrors should be flat to \(\lambda/10\).

(ii) The 'cat's-eye' reflector uses a concave spherical mirror with a secondary reflector at its focus. This secondary reflector is usually spherical convex, and occupies the focal surface. Any parallel beam entering is reflected back, and is inverted. Thus there should be a cat's eye in both arms of the interferometer. A typical focal ratio would be F.8, but larger focal ratios than this would be needed if a large aperture instrument is contemplated. The aberrations of the spherical mirror system are not themselves important since they are duplicated in the two arms, and the returning wavefronts, although not plane, have the same shape. The difficulty is that they should be exactly the same shape, to \(\lambda/4\), and the mechanical tolerances on the secondary mirror position and the question of thermal stability begin to make this type of reflector unattractive at apertures above 20 cm.

Although larger focal ratios can be used, the path length in each arm becomes long, and the field that can be accepted without vignetting is small unless a very large beam-splitter is used.

![Figure 8.7](image1.png)

**Figure 8.7.** The principle of the 'cube-corner' Michelson interferometer.

![Figure 8.8](image2.png)

**Figure 8.8.** Principle of the 'cat's-eye' Michelson interferometer. The two 'cat's-eyes' are Schmidt systems. Provided the two are identical in size there is no need for a corrector, as the distortions of the returning wavefronts will be identical, at least over moderate path differences.

![Figure 8.9](image3.png)

**Figure 8.9.** Principle of the 'Tilting' Michelson interferometer.
(iii) " Tilting Michelson", like the cube-corner, is based on the laws of reflection, but the sliding motion of the Michelson interferometer is replaced by a rotation about an axis. This axis does not need to be well defined and may change during the scanning. Any ray reflected from the beam-splitter and from mirror "A" (figure 8.9) emerges parallel to itself, irrespective of the angle \( \theta \). The path difference however, depends on \( \theta \) and so can be changed by tilting the carriage. This tilt does not disturb the alignment. A feasibility study by the authors has shown that an aperture of 20 cm is possible, although a beam splitter-compensator combination with dimensions of \( 31 \times 25 \) cm is needed.

(iv) The 'Mobius-band' interferometer is a device which uses the two parallel faces on a single piece of glass as the reflectors, and feeds the light to them in such a way that a tilt of the glass does not shear the two beams as they return to the beam-splitter. A tilt of the glass or a translation of it will alter the path difference, and, as in the case of the tilting Michelson interferometer, a separate optical method must be used to measure and control the path difference (as indeed it must in all these devices at short wavelengths) [16].

![Figure 8.10. Principle of the 'Mobius-band' interferometer.](image)

(v) The 'field-compensated' interferometer is a radical departure from some of the principles laid down for Michelson-type interferometers. Normally the field of the instrument is restricted by the condition \( \theta^2/8 = 1/R \). This is because the two rays derived from a marginal ray have a different path difference in the two arms from two rays derived from a paraxial ray. The intensity at the detector due to a monochromatic ray from the edge of the field-stop will thus be modulated at a different rate from that due to the same monochromatic ray coming from the centre of the field. The field-compensated device arranges that rays from all parts of the field suffer the same path difference in the two arms. The field is then not limited by the resolving power required. The action of the device is not obvious, and to describe it we first define three optical paths through a block of glass.

In figure 8.11 the physical path is ABCDEFG. This is the actual path in space followed by a ray. The equivalent geometrical path, sometimes called the reduced path is ABHFG. This is the path the ray would have to follow in order to give the same return rays, if the glass were removed. The optical path is \( AB + CD + DE + FG + \mu(BC + EF) \), and its length is \( \lambda \times \) (the number of wavelengths between A and G).

The interferometer action requires only that the optical paths should differ in the two arms, and that this difference be controllable. Field-independence can be achieved if the equivalent geometric path difference is maintained at zero, while the optical path is changed. The field-compensated interferometer achieves this.

The Connes-Bouchareine design [17] uses two identical prisms in
place of the plane mirrors of a normal Michelson interferometer. Each is silvered on the rear surface. The system is illustrated in figure 8.12. The dotted line indicates the plane of reflection for the equivalent geometrical path. If now one of the prisms is moved along this line, the optical path difference is changed, but the equivalent reflector is fixed, so that the geometrically equivalent paths do not change. At large optical path differences one of the prisms is, in effect, much thicker than the other. This makes the instrument equivalent to a normal Michelson interferometer at zero geometrical path difference, with a parallel-sided slab of glass in one arm. This slab is tilted with respect to the optic axis. Consider now a ray, starting from the edge of the field and reflected as in figure 8.13 to return to the conjugate point in the field. The path length OAB should be the same as the path for a paraxial ray CAC. It will not be the same, however, because of this tilted slab, which introduces a wavefront deformation similar to coma and astigmatism. The useful field of the instrument is limited by the condition that the marginal ray shall not suffer a path difference in the instrument greater or less than a paraxial ray by more than \( \lambda/2 \). This is the same limitation as in a normal Michelson interferometer. The field is restricted by astigmatism to about ten times the diameter of the field of a normal Michelson interferometer.

A further development of this invention by Mertz [2] removes this astigmatism. The slab of glass is, in effect, still there, but is now normal

\[
d\Delta = 2x \sin \alpha \frac{\Delta - 1}{\mu}
\]
to the beams. The field is now limited by spherical aberration, and, using the usual formula for the spherical aberration produced in a converging beam by a block of glass, it is possible to arrive at a formula for the permitted semi-diameter of the field. In radians, it is given by $\theta$, where

$$\theta^4 = 4\mu^3/R$$

(8.14)

$R$ being the resolving power required, and $\mu$ the refractive index of the slab.

8.8 Positions of Aperture and Field-stops

In order that there shall be no vignetting there must be a single aperture stop in the instrument. This is normally at the reflector face, but if a large range of path differences is contemplated the two stops will no longer be sufficiently close to be considered coincident. One of the reflectors must be made larger in aperture than the other to avoid vignetting; the smaller of the two should be the fixed mirror and should define the aperture of the system. The beam-splitter must be large enough to allow all the marginal rays through, and when a wide field is possible as in the field-compensated interferometer, the beam-splitter will be considerably larger than the reflectors. The thickness of the beam-splitter will need to be between $\frac{1}{4}$ and $\frac{1}{2}$ of the diameter, and due allowance must be made for the refraction inside. Since the beam-splitter is likely to be the most expensive item in the instrument and therefore the one that will determine its size, the angle at which it reflects must be chosen carefully. There will be an optimum, when considering anything but the plane mirror Michelson interferometer, since a larger angle of reflection means a wider beam but a longer path from the beam-splitter to the reflector. It is the luminosity of the system that is to be optimized. In the case of the tilting Michelson interferometer, for example, the optimum angle is 65°. This is a fairly flat optimum, but the tolerance is not more than 5°. Thus there is significantly more luminosity at 65° than there would be for the same size of beam-splitter at 55°.

The position and power of the focusing lenses follows normal spectrometric practice. There is no virtue, apart from space-saving, in using small focal-ratio lenses.

The use of Cassegrain systems for the focusing optics is also singularly unrewarding. One finds that either the obstruction due to the secondary mirror or the hole at the centre of the primary removes an unacceptably large fraction of the light from the beam. A simple ray tracing will convince the doubtful reader of this. Newtonian or Pfundt systems are tolerable, but the simplicity of the lens, which need not be of particularly high quality, has much to recommend it.

There should be a field lens at the exit field-stop which should image the aperture stop of the system on to the detector surface or, better, on to a mask which is cut to be the same shape as the image of the aperture stop. This prevents vignetting. In some cases it may not be possible to match the aperture stop to the detector. The conservation of étendue requires that the diameter of the stop multiplied by the divergence angle of the rays, should be a constant. The aperture stop may be of such a size that the convergence of the rays on to its smaller image at the detector may be greater than $\pi$. Rather than use a bigger detector,
with its extra noise, one should use an axicon (q.v.) as the final light-collecting element.

8.9 Field Division

The intensity transmitted through a two-beam interferometer is given by

\[ J(\Delta) = \frac{1}{2} \int S(v) (1 + \cos 2\pi v \Delta) \, dv \]

In astronomical spectrometry in particular, trouble may arise from the fluctuation with time of the general radiation level. Part of the transformation program is to subtract \( J(o) \) from each sample. Since the samples are taken at different times the appropriate \( J(o) \) may be different for each sample. Serious errors can arise if this is ignored.

The intensity reflected from the same interferometer is given by:

\[ J(\Delta) = \frac{1}{2} \int S(v) (1 - \cos 2\pi v \Delta) \, dv \]

And the difference between the reflected and transmitted intensities is then:

\[ J(\Delta) = \int S(v) \cos 2\pi v \Delta \, dv \quad (8.15) \]

so that the detector output can be transformed directly without any pre-subtraction. Variations in the total light level during the observation will act as truncating or apodizing functions. High-frequency fluctuations that can be detected by the system are a source of noise, and these are discussed later together with other sources of noise.

To measure the adjusted samples directly, the field of the instrument must be split. Sometimes the aperture is split as well, giving the impression of a pair of interferometers tied together. This is not necessary, but is sometimes convenient. The self-compensating interferometer using cube-corners and a single beam-splitter is of this type. The useful étendue is the same as if the beam-splitter had been divided into two and used as the normal beam-splitter compensator combination. Such a division of the field must be regarded as normal for astronomical Fourier spectrometry. Indeed, Connes divides the field into four parts in order to subtract the sky background emission [13].
Design of Optical Spectrometers

Because of defects in the beam-splitter and inequalities in the detectors the quantities measured will not be exactly those given by the equations above. If a fraction \( t \) is transmitted by the beam-splitter and a fraction \( r \) is reflected, the two signals from the detectors will be:

(i) Transmitted:

\[ J_1(\Delta) = k_1.2r.t. \int_0^\infty S(v).\cos 2\pi v \Delta \, dv \]  
\[ (8.16a) \]

(ii) Reflected:

\[ J_2(\Delta) = -k_2.2r.t. \int_0^\infty S(v).\cos 2\pi v \Delta \, dv + (r^2 + t^2) \cdot \int_0^\infty S(v) \, dv \]

\[ (8.16b) \]

The sensitivities \( k_1 \) and \( k_2 \) of the detectors can be matched, either on the equipment or later in the computer if the two channels are recorded separately. But \( r \) and \( t \) are fixed by the properties of the beam-splitter, so that the difference of the two outputs:

\[ J_1 - J_2 = 4r.t. \int_0^\infty S(v).\cos 2\pi v \Delta \, dv - (r - t)^2 \cdot \int_0^\infty S(v) \, dv \]

\[ (8.16c) \]

contains a small term which should be constant, but will change if the strength of the source changes. If \( r \) and \( t \) differ by less than 10 per cent the extraneous signal is less than 0.4 per cent of the total output. It is the variation of this extraneous signal that is troublesome. Leaving it in the data for computation may result in a large peak at \( v = 0 \) in the spectrum. Variations of this term will contribute their own spectrum superimposed on the genuine one.

8.10 The Problem of Noise in Multiplex Spectrometry

Although some theoretical studies have been made on the effects of various types of noise [18, 19, 20] we will consider it here from the practical point of view. Noise can be contributed to the system at several points. The noise produced by the detector is the chief contributor in infrared spectrometry, unless the source is extremely feeble, and this is why it is in this region that the Fellgett advantage is most marked. This advantage begins to disappear when significant amounts of noise are contributed by the source, and is removed almost completely when the source noise is dominant. This leaves only the luminosity advantage over a normal slit spectrometer. Even this advantage only applies when the source is (or can be made to be) of sufficient extent to fill both the field and the aperture.

A stellar image in astronomy, for example, would not be analysed in a noticeably shorter time by a Fourier spectrometer than by a large grating spectrometer, provided that the stellar image was small enough to go entirely into the entry slit. It is with large telescopes, where a 1" of arc seeing disc at a focal ratio of F:20 is between \( \frac{1}{4} \) and \( \frac{1}{2} \) mm in diameter, that one thinks seriously about Fourier spectrometry in the visible region. Most of the light from such an image is lost in conventional spectrometry.

The source noise can be analysed into several components.

(i) The genuine fluctuations in the brightness of the source itself.
(ii) ‘Seeing fluctuations’ (in astronomy, not in the laboratory).
(iii) Photon shot noise.

The effect that such noise has on the spectrum is quite different from its effect in single-channel conventional spectrometry. Since the signal fluctuates in amplitude because of this noise, the noise can be regarded as a function of time, and hence of path difference \( \Delta \). The noise itself will be transformed, and it is the frequency spectrum of the noise that will appear in the spectrum. The frequency scale at which it appears will depend on the rate of change of path difference. Thus it is important when such noise is present, to choose the spectral range and the scan rate so that there is an empty region of the spectral range where this noise can appear. The noise may not be in the same alias as the spectrum. If the scan rate is very fast for example, the noise might be expected to appear near zero wave-number, and the spectral range is then chosen so that there is an empty region at the low wave-number end. This point was made and discussed fully by Connes [12]. For this reason too, Mertz [21] has adopted rapid scanning, where the total scan time is of the order of one second. Adequate signal/noise ratio is obtained by repeated scanning up to one thousand times if necessary.

Seeing-noise and photon shot-noise will appear differently, because the noise amplitude now depends on the signal strength. The signal which would appear at the detector in the absence of noise is:

\[ J(\Delta) = \int_0^\infty S(v).\cos 2\pi v \Delta \, dv \]

But if there is noise present, \( S(v) \) becomes a function of \( \Delta \) because it is a function of time. In an elementary treatment one may assume that...
the whole spectral region is affected in the same way by the noise, so that we can write for the observed intensities:

$$J_{\text{obs}}(\Delta) = J(\Delta) \cdot [1 + n(\Delta)]$$

and the spectrum finally recovered will be the convolution of $S(\nu)$ and the Fourier transform of $(1 + n(\Delta))$. The effect of the noise is then to modify the instrumental profile. It would be easy to extend this theory to cases where the noise power varies with wave-number or is a function of the signal power other than a simple linear one as above. This has been done [18], but for practical purposes one may say that the ‘proof of the pudding is in the eating’, and we follow the advice of Connes [12]: to choose a spectral range with some empty region so that the noise can be assessed on its own.

### 8.11 Other Multiplex Spectrometers

We have dealt at length with the case of the Fourier spectrometer, since this is the device that makes most use of the multiplex method. The ‘mock interferometer’ of Mertz [7] is another device with similar multiplex properties, which is an adaptation of a normal grating spectrometer. In a Littrow or Ebert mounting, the usual entry slit is replaced by a large circular or rectangular aperture, and a similar aperture replaces the exit slit. The two apertures should be co-planar for the sake of simplicity, and should be close together so that both can be covered by a fine line screen, such as is used by process engravers for making printing blocks from photographs. The spectrometer will produce at the exit aperture an image of the screen over the entrance aperture, and the position of this image will depend on the wavelength of light that is used. If the screen has equal widths of transparent and opaque lines, the image will be covered by the other half of the screen, or partly covered or uncovered, depending on the wavelength. If the whole screen is now rotated, the bars of the image will cross the bars of the screen covering the exit aperture, and the rate of crossing will again depend on the wavelength of light. Thus the intensity of the transmitted signal will be modulated at a frequency that is wavelength dependent, and so a whole section of the spectrum can be observed with one detector. In order to maintain a constant frequency for a given wavelength, the grid must be rotated at a non-uniform rate, and Selby [22] has shown that the intensity transmitted at a wavelength $\lambda$ is given by the equation:

$$I(\theta) = \frac{I}{2\lambda} \Delta \sin \frac{\nu(\lambda - \lambda_0 + 2wd) \sin \theta}{2zd} + \phi$$

where $s$ is the width of a dark strip on the grid,

$w$ is the distance of the centre of rotation to the centre of the aperture,

$\theta$ is the angle through which the grid has been rotated,

$d$ is the dispersion of the grating,

$\lambda_0$ is the wavelength for which the modulation frequency is zero.

(This will depend on the grating setting and is one of the adjustable quantities in the instrument.)

$\Delta(\nu)$ is a function defined by Selby as a symmetrical periodic triangular function and illustrated in figure 8.18. This shape of modulation function implies that there will be odd harmonics of the fundamental frequency present in the data, so that the spectral range of the device is restricted to one and a half octaves of the spectrum.

Although simple and ingenious in principle, the device turns out to have several snags in practice. The chief of these is the magnification

![Figure 8.17](image)

**Figure 8.17** The aperture of the Mertz mock interferometer and its image in the masking grid.

![Figure 8.18](image)

**Figure 8.18**. The triangle function $\Delta^2(\theta)$. This corresponds to the $\frac{1}{2}(1 + \cos \theta)$ modulation produced by a single wavelength in a Michelson interferometer. Because the triangle contains odd harmonics of the fundamental period there is a restriction on the free spectral range.
effect that is present in all grating spectrometers, which in this case gives an image of the grid of unit magnification in one direction and \((\cos i/\cos r)\) in the other. To obtain high resolution it is necessary to make a correction and for this Selby (op. cit.) uses a cylindrical zoom lens.

It is doubtful whether this device can be considered a serious rival to the interferometer as a Fourier spectrometer, both in view of the complexity and size of instrument needed to obtain a given \(L.R.\) product. Even less is it likely to compete with the field-compensated interferometer.

The 'coded slit' spectrometer, like the mock interferometer, is an adaptation of a normal grating spectrometer or, better in this case, a spectrograph. It is a much simpler device than the mock interferometer, but has a much smaller degree of multiplexing, and hence a smaller Fellgett advantage. Such devices were first described by Golay [8] as a means of improving the signal/noise ratio in an infrared spectrometer, and a similar device has been used by Braddock and Wilcock [9] for astronomical spectrophotometry.

The entry slit of the grating spectrometer remains, but instead of an exit slit there is a transparent disc, on which are described a series of concentric circles. In each annulus is a series of alternate black and transparent sectors, the number being different for each annulus. This disc is carried normal to the outgoing beams, and with the focal plane lying along one radius of the disc. As the disc is rotated, each element of the spectrum corresponding to one annulus is modulated in intensity at a different rate. These frequencies are passed, via a detector, to a suitable store, and the spectrum can be reconstituted at will by means of a light source and a similar disc. Alternatively the spectrum can be displayed in 'real time' by placing a light source behind the opposite diameter of the disc, and allowing its intensity to be modulated by the detector output. Grainger [23] has pointed out the need for careful coding of the disc to avoid the possibility of one radius having transparent patches at each annulus. If this is avoided, the dynamic range of the detector can be smaller, and the detectivity higher. For the 'real time' display, the number of opaque/transparent pairs in each annulus needs to be even. To avoid cross-talk in the system the number of sector pairs in the outermost annulus must be less than twice the number in the innermost annulus.

The simplicity of this device and its ability to act as its own Fourier transformer make it very attractive for applications where there are only a few resolved spectral elements, so that only a low degree of multiplexing is possible.

REFERENCES


BIBLIOGRAPHY

There are several books which should be in the possession of anyone attempting multiplex spectrometry. Among these are:
And the numerous papers of J. and/or P. Connes on the subject should also be consulted.
9.1 Introduction

A spectrometer takes the light from a source, disperses it, and selects a small spectral region to pass on to a detecting and measuring device. The amount of energy in this spectral interval is usually very small, and hence the general requirements for the detector are:

(i) High sensitivity: The detector should produce a large output signal (generally electrical) for a small radiation input.

(ii) Low noise: Inevitably, in the absence of radiation a detector will produce some signal, i.e. 'noise'. This limits the smallest genuine signal that can be detected.

(iii) Linearity: The output of the detector should be proportional to the radiation input, so that accurate photometric measurements can be made. The constant of proportionality (gain) should be independent of the wavelength of the radiation.

The human eye is a radiation detector of surprisingly high sensitivity (in the visible region of the spectrum). The fully dark adapted eye can detect a light flash containing only a few quanta of radiation. As a quantitative instrument, on the other hand, the eye has a very poor performance. With the most careful precautions photometric comparisons between two light sources of the same wavelength can be made within a few per cent. If the sources are of different wavelengths, even worse results are obtained.

The photographic plate was until the advent of photoelectric detectors the principal means of recording radiation. Its chief characteristics are a low quantum efficiency (i.e. a large number of radiation quanta are needed to make one grain of the emulsion developable) and a very large information-storing capacity. Thus a photographic plate can simultaneously record the information in a large number of spectral channels. Depending on the application this may outweigh the disadvantage of the low quantum efficiency. Accurate photometry of photographic plates is difficult because the response is non-linear, and wavelength dependent. A number of effects, such as fog, solarization, Eberhard effect, etc., have to be taken into account.

9.2 Types of Detectors

All detectors used in spectrometers function by converting the incident radiation into an electrical signal which can be amplified and recorded by conventional apparatus.

Detectors can be classified into two main groups according to the mechanism of this conversion:

(i) Thermal detectors – detect the heating of an absorbing receiving element due to the absorbed radiation energy.

(ii) Photon detectors – Absorption of one quantum of radiation releases a quantity of electrical charge, which can be detected in one of two ways: pulse counting, which records directly the number of events taking place in the detector, or rate measuring, which integrates the charge released during a given time, and thus measures the rate of arrival of photons.

Photon detectors function in one of two ways:

(iia) External photoelectric effect – The energy of a photon is large enough to free a charge carrier (electron) completely from the sensitive receiver surface.

(iiib) Internal photoelectric effect – The energy of the photon is not sufficient for (iia) but is large enough to raise a charge carrier (electron or hole) in a semiconductor into the conduction band.

Three types of internal photoelectric effect are of practical importance:

(iib1) The charge carriers increase the conductivity of the semiconductor – photoconductive effect.

(iib2) The carriers are generated at a point in the semiconductor at which a potential barrier exists, and a voltage is produced by charge separation – photovoltaic effect.

(iib3) Charges are separated by diffusion in opposite directions in a magnetic field – photoelectromagnetic (PEM) effect.

9.3 A Basic Distinction

The response of an ideal thermal detector depends only on the amount of radiant power absorbed, and is independent of its spectral distribution. Hence a plot of response to unit power per unit wavelength interval
incident on the detector, versus wavelength, is a straight line parallel to
the axis.

A photon detector measures the rate of arrival of radiation quanta.
As the number of photons per second per unit power is proportional
to the wavelength \( E_{\text{photon}} = h c / \lambda \), its response to unit incident power per
unit wavelength interval increases with wavelength up to that wave-
length at which the photon energy is no longer sufficient to produce a
photoelectric event.

These statements, represented diagrammatically in figure 9.1, repre-
sent the ideal behaviour of detectors. In practice the number of signal
events produced by a photon detector is only a fraction of the number

![Figure 9.1. Response characteristics of ideal detectors.](image)

of incident quanta. This fraction\(^*\) is the quantum efficiency of the
detector. The quantum efficiency may be different for different wave-
lenghts, and thus the response may not be linear, or even a monotoni-
cally increasing function of wavelength. Further modification of the
response curve may be caused by the transmission characteristics of
possible detector envelopes or windows.

Similarly, the response of a thermal detector need not be independent
of wavelength. The absorbivity of the receiving element may vary so
that at different wavelengths different fractions of the incident power are
utilized.

9.4 Noise in Detectors

Because of the difficulty of amplifying a steady (d.c.) signal, and also
because some detectors respond only to signal changes, some device is

\(^*\) This fraction can occasionally be vulgar, if photoelectric events in the far
ultraviolet are included.

nearly always used which interrupts the incident radiation periodically
('chopper'), and so produces an a.c. signal at the detector output. This
is amplified, rectified and finally recorded in some convenient form.
An advantage is often gained by synchronous (homodyne) rectification
at the chopping frequency. A synchronous rectifier responds only to
the chopping frequency, and its odd harmonics. Hence, if the noise
spectrum of the detector extends over a considerable frequency range,
a significant increase in signal/noise ratio can be obtained by homodyne
rectification.

The smallest radiation signal which can be seen is determined by
random fluctuation at the output of the complete detection system
(detector + amplifier + ... + recording device). This 'noise' is made
up of a number of contributions, some or all of which may be present
in any specific system.

**Photon Noise.** The number of quanta received by a detector in a
given time interval is subject to statistical fluctuations. If this number
is \( N \) then the standard deviation of \( N \) is \( \sqrt{N} \). If the quantum efficiency
is \( \eta \), the standard deviation is smaller by \( \sqrt{\eta} \) and the signal is smaller
by \( \eta \). The \( S/N \) ratio is thus worse by a factor \( \sqrt{\eta} \). This fluctuation in
the input constitutes the ultimate limit in any detection system. Its
effect can be reduced by increasing \( N \), either by employing a brighter
source (which usually is not possible), or by observing over a longer
time interval.

**Johnson Noise.** Any electrical detector is in effect an electrical
resistance. Thermal fluctuations of electrons in a resistance \( R \) at
absolute temperature \( T \) produce, in an electrical bandwidth \( \Delta f \), an
r.m.s. voltage:

\[
V_f = (4 k T R \Delta f)^{1/2} = 1.3 \times 10^{-19} (R \Delta f)^{1/2} \text{ volts.}
\]

**Current-induced (CI) Noise.** Certain detectors, in fact nearly
all semiconductor detectors, require a polarizing current, and noise
due to this current is present. Its spectrum is such that noise power per unit
bandwidth is proportional to the reciprocal of the frequency (1/\( f \) noise).
Hence in the presence of CI noise one should use as high a chopping
frequency as possible.

**Dark Current Noise.** Some detectors produce a signal even in
the absence of illumination (e.g. photomultipliers). This signal, the dark
current, is of the same nature, and subject to similar statistical fluctua-
tions as a signal due to radiation. Hence its fluctuations contribute to
the overall noise of the system and can become a serious limitation at very low levels of illumination, where the number of ‘dark counts’ is of the same order of magnitude as that of ‘light counts’. The dark current can often be reduced by cooling the detector.

OTHER SOURCES OF NOISE. There are sundry other contributions to the noise of a detection system, particularly in the case of semiconductor detectors. A very detailed treatment of all sources of noise is given by Kruse et al. [1]. As far as noise produced by the electronic amplifying system is concerned, it is usually possible by careful design to make this small compared with photon and detector noise.

9-5 Criteria of Performance

The noise equivalent power, \( P_n \), is defined as the r.m.s. value of sinusoidally chopped radiant power falling on to a detector, which will give rise to an r.m.s. signal voltage equal to the r.m.s. noise voltage. Unit electrical bandwidth is usually assumed, and the characteristics of the radiation source (temperature, in the case of a black body) and chopping frequency must be specified. Note that the merit of the detector increases as \( P_n \) decreases.

A second criterion is the detectivity, \( D \). This is defined simply as the reciprocal of \( P_n \), at a given source temperature and chopping frequency.

Another frequently used figure of merit is the noise equivalent input (NEI), defined as the radiant power per unit area of detector giving rise to a signal/noise ratio of unity. Thus:

\[
\text{NEI} = \frac{P_n}{A_d} = \frac{i}{DA_d}
\]

where \( A_d \) is the detector area.

Many detectors exhibit a \( P_n \) which is proportional to the square root of the detector area. Therefore an area independent figure of merit can be obtained by dividing \( P_n \) by the square root of the detector area. The reciprocal of this figure has become known as \( D^* \) (pronounced dee-star); thus:

\[
D^* = \frac{1}{(\text{NEI})A_d^{1/2}} = \frac{A_d^{1/2}/P_n}{D} = DA_d^{1/2}
\]  

(9.1)

The units of \( D^* \) are cm-\((c/s)^{1/2}\)/watt. Measurement conditions must again be specified, like \( D^*(T, f, \Delta f) \), \( T \) being the source temperature (black body), \( f \) the chopping frequency, and \( \Delta f \) the electrical bandwidth. This latter, in specifying \( D^* \), is almost invariably taken as \( 1 \) c/s.

Some confusion has arisen over the term ‘detectivity’. The detectivity \( D \), the reciprocal of \( P_n \), is very rarely used nowadays, having been superseded by \( D^* \). It has become fairly common usage now to refer to the latter as detectivity, or as ‘dee-star’ [3].

The figures of merit discussed above refer to the response of a detector to radiation from a black body at some specified temperature. For many detectors the response is wavelength dependent, and it is possible to determine \( P_n, \text{NEI}, D \), and \( D^* \), in terms of the detector’s response to monochromatic radiation. In this case the wavelength must be given, and the figures of merit are referred to as \( P_d, \text{NEI}_d, D_d \) and \( D^*_d \). In fact, the fullest information about a detector’s performance is given by a plot of one of these quantities against wavelength.

The responsivity, \( R \), of a detector is the ratio of the signal obtained to the incident power. As noise does not affect the responsivity, it is not necessary to specify bandwidth. In addition, so long as the chopping frequency is small compared with the response time (see below) of the detector, it does not affect responsivity either. Knowledge of source temperature or wavelength and detector area is still needed. Denoting the incident power by \( P \) and the signal voltage by \( V_s \):

\[
R = \frac{V_s}{P} \text{ volts/watt.}
\]

\( R \) is related to \( P_n \) and \( D^* \) by:

\[
R = \frac{V_n}{P_n(\Delta f)^{1/2}} = \frac{D^*V_n}{(A_d\Delta f)^{1/2}}
\]

where \( V_n \) is the r.m.s. noise voltage.

One can also specify the monochromatic responsivity, \( R_\lambda \), to radiation of a single wavelength. Unless otherwise given, this usually refers to that wavelength for which \( R_\lambda \) is largest.

The response of a detector to a suddenly applied signal is not instantaneous, but can be represented approximately by an exponential rise. Hence, at sufficiently high chopping frequencies the output will no longer follow the input signal. In fact, the detector can be likened to a low-pass filter, with a response:

\[
R(f) = R_0 \left( \frac{1}{1 + \frac{4\pi^2f^2\tau^2}{\lambda}} \right)^{1/2}
\]

(9.2)

where \( R(f) \) is the responsivity at chopping frequency \( f \), and \( R_0 \) that at frequency zero. The quantity \( \tau \) is the response time of the detector.

At low frequencies \( (f \ll 1/(2\pi\tau)) \) the responsivity is nearly frequency independent. At \( f = 1/(2\pi\tau) \), \( R \) drops to \( 0.71R_0 \), and if \( f \gg 1/(2\pi\tau) \),
the responsivity is inversely proportional to frequency. At the frequency \( f = 1/4\pi \) the responsivity becomes \( 0.53 R_\theta \), and this is usually taken as the upper frequency limit of a detector.

\( D^* \) can be obtained by dividing the product of responsivity and the square root of the detector area by the noise voltage per unit bandwidth. For detectors whose noise spectrum is frequency independent ('white noise'), the dependence of \( D^* \) on the chopping frequency is the same as that of \( R \). However, for current \((1/f)\) noise limited detectors this is no longer true, and \( D^* \) is given by:

\[
D^*(f) = \frac{k f^{1/2}}{1 + 4\pi^2 f^2 T_h^{1/2}} \tag{9.3}
\]

where \( k \) is a constant of proportionality.

There is now a value of frequency which maximizes \( D^* \). This can be found by the usual methods of calculus, and is:

\[
f_{\text{max}} = \frac{1}{2\pi T} \tag{9.4}
\]

Current-noise limited detectors should be operated at this frequency to give the highest value of \( D^* \).

### 9.6 Descriptions of Individual Detectors

Characteristics of radiation detectors will be described in this section. For this purpose detectors will be divided into thermal and photon detectors, and the latter classified according to the spectral region in which they are used.

**Thermal detectors** are sensitive to all wavelengths which are absorbed by their receivers. In practice thermal detectors are used principally in the medium and far infrared regions of the spectrum; in the near infrared, the visible and ultraviolet photon detectors in general have a superior performance.

1. **The bolometer**

The operating principle of the bolometer is the change of electrical resistance of a material due to the temperature change caused by the absorbed radiation.

Bolometers may be of three types: metal, semiconductor, and superconducting bolometers. Metal bolometers, in the form of thin wires or a thin metal film on insulating substrates, are hardly used nowadays on account of their low responsivity.

Certain semiconductor materials exhibit a relatively large change of resistance with change of temperature, and bolometers made from one of these materials have become known as thermistors. Contrary to the behavior of metals, the resistance of a thermistor falls as the temperature rises.

Thermistors are in the form of flakes, approximately \( 10 \mu \) thick, and are of varying sizes, mounted on heat-dissipating substrates. Radiation falling on the flake (which is blackened) warms it. On removal of the radiation the flake returns to its original temperature. Its time-constant depends on the thermal coupling between the flake and the substrate, and on the radiation interchange with the surroundings. A fast response is obtained by mounting the flake directly on to a heat sink. The flake may also be mounted in a gas or in a vacuum, with a slower response but a higher \( D^* \).

The change of resistance is detected by making the flake one arm of a bridge circuit. A second flake, shielded from radiation, in the other arm of the bridge, guards against drifts due to changes of ambient temperature.

A typical current/voltage characteristic of a thermistor is shown in Figure 9.2.

![Figure 9.2. Current voltage characteristic of a thermistor.](image)

At low applied voltages the power dissipation in the thermistor is small, and hence heating effects are negligible. In this region the thermistor behaves approximately as an ohmic device. As the voltage is increased, joule heating becomes larger and the resistance of the thermistor falls. Eventually, the dynamic resistance (rate of change of current with voltage) becomes negative, and unless a ballast resistor is used the thermistor will burn out when used in this region.

When employed as radiation detectors, thermistors are always used on the left of the maximum of the curve of fig. 9.2, usually at a bias voltage of about 60 per cent of the peak of the curve. For a thorough discussion of thermistors the reader is referred to Shive [2].
(2) THE SUPERCONDUCTING BOLOMETER

The mode of operation of this detector is based on the phenomenon of superconductivity, i.e., the complete disappearance of the electrical resistance of certain materials at temperatures near absolute zero. The typical change of resistance of a superconductor with temperature is shown in figure 9.3.

The superconducting transition from finite to zero resistance takes place over a very narrow temperature interval, typically of some tenths or hundredths of a degree. Thus the slope of the resistance/temperature curve in this interval is exceedingly steep, and a bolometer operated in this region has a high responsivity.

Superconducting bolometers have been described by Andrews [3], Milton [4], and Fuson [5].

The control of the operating temperature of a superconducting bolometer presents a severe technical problem. Because of the small interval over which the transition takes place the temperature of the bolometer must be controlled within very close limits (of the order of 10⁻⁴ degrees). Any temperature fluctuations appear as noise in the detector output.

The superconducting bolometer has the advantage over metal and thermistor bolometers of reduced thermal noise, due to its low operating temperature. However, the problems caused by the need for accurate temperature control have prevented it from being widely used.

Typical performance figures of bolometers are given in the table below:

![Resistance Temperature Characteristic of a Superconducting Bolometer](image)

**Figure 9.3.** Resistance temperature characteristic of a superconducting bolometer.

<table>
<thead>
<tr>
<th>THERMISTOR</th>
<th>SUPERCONDUCTING BOLOMETER</th>
</tr>
</thead>
<tbody>
<tr>
<td>Operating temperature</td>
<td>295°C</td>
</tr>
<tr>
<td>$D^*(500°C, f, 1)$</td>
<td>$2 \times 10^8$ (10 c/s)</td>
</tr>
<tr>
<td>$\tau$</td>
<td>1.5 msec</td>
</tr>
<tr>
<td>Optimum chopping frequency</td>
<td>10 c/s</td>
</tr>
</tbody>
</table>

(3) THE RADIATION THERMOCOUPLE

The thermocouple is a detector in widespread use, particularly in the infrared part of the spectrum. It is always operated at room temperature. A thermocouple consists of a low-mass junction of two materials of different thermoelectric powers, mounted on a blackened radiation receiver. To improve detectivity (by reducing convection losses) the whole device is mounted in an evacuated enclosure with a suitable window to admit the radiation. This construction entails a fairly long response time, necessitating low chopping frequencies.

The resistance of thermocouple detectors is always small (of the order of 10 ohms), and a step-up transformer is needed to match the detector to its amplifier. The design of radiation thermocouples is well described by Hornig and O'Keefe [6], and their ultimate performance limit has been discussed by Fellgett [7].

Typical operating and performance data of a thermocouple are:

<table>
<thead>
<tr>
<th>Operating temperature</th>
<th>295°C</th>
</tr>
</thead>
<tbody>
<tr>
<td>$D^*(500°C, 5, 1)$</td>
<td>$1.7 \times 10^6$</td>
</tr>
<tr>
<td>$\tau$</td>
<td>20 msec</td>
</tr>
<tr>
<td>Chopping frequency</td>
<td>$\sim 5$ c/s</td>
</tr>
<tr>
<td>Noise mechanism</td>
<td>thermal</td>
</tr>
</tbody>
</table>

(4) THE GOLAY DETECTOR

The Golay pneumatic detector (named after its inventor) is a thermal detector of high detectivity, slow response, and a rather fragile construction. It detects through an optical lever device the expansion of a gas in a chamber, due to the heat absorbed from the radiation. Full details of its construction and operation are given by Golay [8]. Its performance characteristics are:

<table>
<thead>
<tr>
<th>Operating temperature</th>
<th>295°C</th>
</tr>
</thead>
<tbody>
<tr>
<td>$D^*(500°C, 10, 1)$</td>
<td>$1.7 \times 10^6$</td>
</tr>
<tr>
<td>$\tau$</td>
<td>20 msec</td>
</tr>
<tr>
<td>Chopping frequency</td>
<td>$\sim 5$ c/s</td>
</tr>
<tr>
<td>Noise mechanism</td>
<td>thermal</td>
</tr>
</tbody>
</table>

(5) THE PYROELECTRIC DETECTOR

This is a novel device, developed by Putley [9] at the Royal Radar Establishment, and is based on a well-known principle.

Certain acentric crystals have a permanent electric dipole moment. Normally, the field produced by this is neutralized by surface charges attracted to the faces of the crystal. A temperature change will cause a change of lattice dimensions in the crystal, and hence a change in the
permanent dipole moment. If the material of the crystal is an insulator, a relatively long time is required for the surface charges to rearrange themselves, and an electric field appears. Hence, if the crystal is illuminated with chopped radiation, the resulting signal can be detected with a pair of electrodes normal to the polar axis of the crystal, and can subsequently be amplified.

The rate of change of dipole moment with temperature is the pyroelectric coefficient, which should be as large as possible. Other requirements are a large absorption coefficient for the radiation to be detected, large resistivity, small specific heat and density, and resistance to atmospheric attack.

The most suitable material found so far is lithium niobate (LiNbO₃) and Putley has reported detectors with values of D⁺ of the order of 10⁸, comparable that is, with thermistors.

A virtue of the pyroelectric detector is that the size of the active element is limited only by the size of crystal that can be grown. It also has a rapid response, making it suitable for use in systems with a bandwidth of about 100 KC/sec.

Photon detectors, in contrast to thermal detectors, are useful over a limited spectral range only. The long wavelength limit is set by the minimum photon energy (hν) required to produce a photoelectric event, i.e. the ejection of an electron from a photocathode, or the excitation of a charge carrier into the conduction band of a semiconductor.

1) PHOTOEMISIVE DETECTORS

If sufficiently energetic photons fall on to a surface of a conductor it is found that electrons are emitted from the surface. The energy of these electrons is given by:

\[ E = h\nu - p \]

where \( h \) is Planck's constant, \( \nu \) the frequency of the light, and \( p \) a constant characteristic of the conductor surface. In fact, \( p \), the work-function, is a measure of the energy lost by the electron as it leaves the surface. Hence this photoemissive effect can only take place if \( h\nu - p > 0 \).

There is a maximum wavelength \( \lambda_0 \), given by:

\[ \lambda_0 = \frac{12394}{p} \]

where \( \lambda_0 \) is measured in microns, and \( p \) in electron volts. Thus, in order to build a photoemissive detector with a long wavelength response it is necessary to find a material with a low work-function.

The number of the emitted electrons is a measure of the intensity of the incident light. To measure it, the emissive surface is made the cathode of a vacuum tube. The emitted electrons are accelerated towards the anode, and give rise to current pulses in an external circuit. These pulses may be counted by means of digital circuitry, or the total charge released over a given time interval may be integrated to give a measure of the rate of emission of photoelectrons.

A figure of merit for photoemissive surfaces is the quantum efficiency, \( \eta_p \), the number of photoelectrons released per incident radiation quantum. Quantum efficiencies vary between approximately 0.2 and 10⁻². The quantum efficiency of a photocathode is also a function of wavelength, as can be seen in figure 9.4, which shows the spectral characteristics of a number of commercially available photocathodes.

The simple photodiode, as described, has been superseded by the photomultiplier. In this device the electrons released from the photocathode are accelerated towards a second electrode by a positive
potential of about 100 volts, and the surface of this second electrode is treated to give a high secondary emission ratio for electron impact. Electrons emitted by the impact of the primary photoelectron are then accelerated by a similar voltage to a third electrode where further multiplication takes place. There may be as many as eighteen of these secondary electrodes or ‘dynodes’ but the usual numbers are between nine and thirteen. The multiplication factor is about 2.5 at each one (depending on the impact voltage and the particular type of photomultiplier) so that a single photoelectron may result in a cascade at the final electrode of up to $5 \times 10^7$ electrons. These provide a substantial current pulse for any subsequent electronics. Due to the fluctuation in secondary emission in individual cases there is a considerable spread in the actual sizes of pulses, and a frequency distribution graph of pulse height $h$ versus $N(h)$ is shown in figure 9.5. Pulses also appear due to random events in the tube itself. Possible causes are (a) thermal emission from the photocathode or one of the dynodes, (b) Čerenkov radiation due to cosmic radiation, (c) Čerenkov events due to decay of naturally radioactive constituents of the tube (usually $K^{40}$ in the glass), (d) fluorescence radiation caused by electron impact with the tube walls, and (e) fluorescence due to collision of positive ions with the walls or with the photocathode.

Any of these causes will produce a small rate of pulses even when the tube is in complete darkness, and the resulting current is known as the dark-current.

Photometry with a photomultiplier can be carried out relatively easily by measuring the current that appears at the anode as a result of the multiplication process. At very low light levels there is a great deal of noise when this method is used. This is not only photon shot-noise (which is not the fault of the multiplier) but multiplication noise due to the wide variations of pulse height at the anode from identical single emission events at the cathode.

It is possible to do photometry without the multiplication noise by counting the pulses that are emitted by the anode, irrespective of their sizes. The chief advantages of this method are:

(i) Loss of multiplication noise.

(ii) No need for ultra-stable E.H.T. supplies for the photomultiplier.

(iii) Ease of recording the photometric data. A digital counter can be used, with a tape or paper punch; or if a graph is required, a rate meter can be used, or a digital-analogue converter.

The arrangement for pulse counting is as in figure 9.6.

A preamplifier with unit gain is followed by a pulse amplifier with a bandwidth of about 20 megacycles. This amplifier should give out

![Figure 9.5](image-url)  

*Figure 9.5.* The pulse height distribution from an end-window photomultiplier. (Type: E.M.I. 9502. at 1800V E.H.T. Cooled to $-80^\circ$C.)

![Figure 9.6](image-url)  

*Figure 9.6.* Complete pulse-counting apparatus. For simplicity a ratemeter and chart recorder can be used, but for more elaborate data gathering system the spectrometer drive can be in small steps, with the total count at each position after a given time printed or punched ready for analysis.
rectangular pulses of standard width (say 1 μsec), and height proportional to the original pulse height from the photomultiplier anode. Good linearity is not essential. The pulse amplifier is followed by a pulse height discriminator which is set to reject the many small pulses that are emitted by the pulse amplifier, and which originate inside it. The setting is found by shorting the input terminals of this pulse amplifier and adjusting the discriminator level until there is less than one count every 10 seconds. It is the custom, with some users, to set the bias level much higher than this to discriminate against pulses originating in the photomultiplier at a later stage than in the photocathode. This is done to improve the ratio of light counts/dark counts, but can be shown to be incorrect [10] since the information rate is greater at lower bias levels. Although the light/dark count rate is lower, the actual rates themselves are much higher and the necessary extra counts for a given photometric accuracy can be acquired in a shorter time. If in a time \( t \) there are \( U.t \) counts with the light on and \( D.t \) counts with the light off, there will, in any similar period be \( U.t \pm \sqrt{(k_u.Ut)} \) and 
\[
D.t \pm \sqrt{(k_D.Dt)} \]
counts respectively. \( k_u \) and \( k_D \) should be equal to unity if the electron emission process were quite random. In practice it is not, and the \( k \)'s may be significantly greater than unity. The actual values should be found by experiment.

The signal/noise ratio is then:
\[
(U - D)t/\sqrt{([k_u.U + k_D.D]t)} = \rho
\]
and the time to produce a signal/noise ratio \( \rho \) is
\[
T = \rho^2 \frac{(k_u.U + k_D.D)}{(U - D)^2}.
\]

**Photomultiplier Preparation.** The inter-dynode voltages are usually applied via a resistor chain as in figure 9.7. It is found occasionally that the last few resistors are shunted by capacitors because large pulses may drain away a sufficient number of electrons to alter the voltage at that stage significantly. This, however, is intended for an entirely different type of pulse counting. Pulses are also counted when the photomultiplier is part of a scintillation detector in nuclear physics. In this case a pulse is due to an event outside the tube which may release simultaneously a large number of photons. These processes result in very much larger pulses than the ones that we deal with in photon counting. Care should be taken to make this distinction. Such capacitors are not necessary in this case. The range of pulse heights is altogether smaller.

The resistors should be soldered carefully to the appropriate pins and after testing there should be an application of a suitable insulating varnish. The following points are relevant when a tube is prepared for loading into the apparatus:

(i) There must be clinical cleanliness. The tube is degreased by washing in benzene or trichlorethylene followed by acetone.

(ii) A coating of colloidal graphite is applied to the outside of the envelope to within \( \frac{1}{4} \) in. of the bottom, and this is connected by a streak to the cathode pin. (This applies even if the tube is run with the cathode at earth potential and the anode at positive E.H.T.)

(iii) There should be a polythene or similar insulating sheath around the tube to prevent arcing or corona discharge to the walls of the container.

(iv) The dynode resistor chain should be cleaned and degreased.

(v) The tube should be loaded into its container with silica-gel or a similar dehydrating agent.

(vi) The container should be constructed to support the envelope at either end. Plastics such as those from which the base is probably made
have very large thermal expansion coefficients. At $-80^\circ$C the tube will be quite loose in the metal part of its base.

(vii) No E.H.T. should be applied at normal room lighting level. Such action would probably destroy the tube.*

(viii) An operating temperature of $-20^\circ$C should be sufficiently low to remove the significant noise. Often a dry ice ($-78^\circ$C) bath is convenient or a Peltier-effect cooler can be used.

(ix) The tube must be given time to settle to its best operating condition. It may take several hours on the first occasion of use before the dark current has diminished to a steady state. In normal routine conditions it should take between a half and one hour.

(x) The E.H.T. should not be switched off in short intervals between use. It may take up to an hour to recover from even a momentary switching off.

(xi) One should avoid an accidental application of a positive E.H.T. to the photocathode.

A tube selected for low dark-count rate by the manufacturer should show in the steady state and at the optimum operating condition a count rate in complete darkness at sea level of between 0.5 and 2.5 counts/sec. If the rate is less than this one should check the discriminator bias level to see if it is too high.

The photocathodes of end-window photomultipliers come in various sizes from 1 cm diameter upwards. The larger ones have greater dark-count rates and hence more shot-noise. The effective area can be reduced to the minimum necessary for the rest of the spectrometer optics with the aid of a magnetic field of the proper shape applied near the photocathode. This field serves to deflect photoelectrons and thermal electrons from the unwanted region of the photocathode to the walls of the tube so that they do not contribute to the signal. The magnetic field may be applied by permanent magnets with shaped pole-pieces, or by a properly wound coil through which a current passes. Such systems have been described by Frommhold and Feibelman [11], and by Farkas and Varga [12].

The quantum efficiency can be improved by introducing the light from such a direction that there is total internal reflection in the glass face of the photocathode. Most photocathodes are fairly transparent since if they were opaque enough to absorb all the light they see, they would be thick enough to prevent electrons from escaping from them. The photoelectrons are liberated on the glass side of the photocathode.

* The authors have no direct experience of this, and go by the legends that they have heard, which are lurid and have lost nothing in the telling!

and have to travel through it without being absorbed in order to reach the vacuum side.

The method illustrated (which is obviously only one of many) allows the photocathode to frustrate the total reflection on one side of the glass, and admit more of the incident light to the photocathode material.

For large pulse rates it is possible to use the photomultiplier with d.c. amplification. Due to the high internal gain of the detector the stability requirements on the external amplifier are such that commercial d.c. amplifiers (e.g. AVO 1338B) can be used successfully to measure photocurrents down to below $10^{-10}$ amp, with time constants of a few seconds to smooth out noise.

Typical performance figures for a photomultiplier are:

<table>
<thead>
<tr>
<th>Operating temperature</th>
<th>$200^\circ$K</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wavelength of peak response</td>
<td>4,000 Å</td>
</tr>
<tr>
<td>$D_s*(500^\circ$K, 1000, 1)</td>
<td>$&gt; 5 \times 10^{14}$</td>
</tr>
<tr>
<td>$\tau$</td>
<td>$\sim 0.01$ μsec</td>
</tr>
<tr>
<td>Chopping frequency</td>
<td>frequency independent to $\sim 100$ Mc/sec</td>
</tr>
</tbody>
</table>

(2) **SEMICONDUCTOR PHOTON DETECTORS**

As can be seen from figure 9.4, detectors using the photoemissive effect can be used only in the ultraviolet, visible, and very near infrared part of the spectrum. The reason for this is that no materials are known with a work function lower than that of the Ag-O-Cs (S.1) cathode.

Making use of the internal photoelectric effect it is possible to devise photon detectors for longer wavelengths, up to about 50μ. The internal
photoelectric effect consists of excitation of a charge carrier in a semiconductor from the valence band or from an impurity level into the conduction band. The energy required to do this is less (in some cases very much less) than that needed to produce photoemission. It depends on the semiconductor material and any doping element(s) present. There are a large number of combinations of these two factors, and a detailed study is beyond the scope of this chapter. The reader is referred to the excellent treatment by Kruse [1].

A consequence of the small excitation energy is that thermal excitation takes place much more readily than in the case of photoemissive detectors. Hence cooling of semiconductor detectors is always advantageous, and often essential. Some of the extreme long-wave detectors must be operated at liquid helium temperatures.

Semiconductor detectors operate most frequently in the photovoltaic and PEM detectors are much less common. The characteristics of some of the more frequently used detectors are shown in Table 9.1. It can be seen that the $D^*$ values of the best of them approach that of a photomultiplier, and are generally superior to those of thermal detectors.

The effective $D^*$ of a detector can be increased by means of immersion optics. In this technique a material of high refractive index (e.g. germanium, $\mu = 4$) is ground to a hemisphere or hyper-hemisphere so that the detector, which is placed in optical contact with the flat surface, is at one of the aplanatic points. Rays refracted by the hemisphere or hyper-hemisphere form an image reduced in area by $\mu^2$ in the first case, and by $\mu^4$ in the second, compared with the unrefracted image. Hence, in order to intercept the same amount of radiation, the detector area can be reduced by this factor. This leads to an improvement of...
the effective $D'$ by a factor of $\mu$ or $\mu^2$ respectively. Figure 9.9 shows diagrams of these two immersion systems.

REFERENCES

[13] JONES, R. C. (1953), Advances in Electronics, 5, 1; (1959) 11, 87

Auxiliary Apparatus

10-1 Source Optics

The purpose of source optics is to ensure that the maximum possible amount of radiation is collected from the source and transmitted by the spectrometer.

From considerations of conservation of energy it has been shown (Chapter 2) that the product of object (or image) size with the solid angle subtended by the entrance (or exit) pupil at the object (or image)

![Figure 10.1. Source optical system. Lens $L_1$ forms an image of the source $S$ on the spectrometer slit. The field lens $L_2$ images the lens $L_1$ on to the collimator objective.](image)

is a constant throughout any optical imaging system. This constant is the ‘etendue’, and limits the amount of light which the system can transmit from a source of given brightness.

In the design of source optical systems it is important to ensure that the etendue of the system is at least as large as that of the spectrometer, so that the amount of light transmitted is limited only by the latter. (An exception to this rule occurs in astronomical spectrometry, and will be discussed later.) In general it is not difficult to achieve this matching of source optics and spectrometer, and the principles of design are illustrated in figure 10.1.

Here the condenser lens $L_1$ forms an image of the source $S$ on the entrance aperture of the spectrometer (slit, or circular aperture in the
Design a are and size a chosen such that the size of the source image is equal to, or larger, than this entrance aperture.

A field lens $L_2$ placed close to the entrance aperture, and of a size sufficient to cover the aperture, forms an image of the condenser lens on the collimator lens of the spectrometer. The focal length of the field lens has to be chosen so as to give the correct conjugates, and the diameter of the condenser lens must be chosen so that its image is equal to, or larger than the diameter of the collimator lens.

If the source image is exactly equal to the entrance aperture, and the condenser lens image is exactly equal to the collimator lens (strictly that part of the collimator lens from which light can reach the dispersing element), all the light collected by the source system is transmitted by the collimator of the spectrometer. This is the case of exact matching of the étendues of the two systems. If the source image or the condenser lens image is larger than stated above, more light is collected from the source, but the part which is transmitted is exactly the same as in the case of exact matching. In the opposite case of smaller images, less light is collected and transmitted. Hence it is important always to 'fill' or 'overfill' both aperture- and field-stops. It must be stressed that once this condition is fulfilled, it is not possible by any optical means to increase the light flux through the spectrometer.

It is often possible to omit either the condenser- or the field-lens. The resulting configurations are shown in figures 10.2(a) and 10.2(b). In figure 10.2(a) the field-lens images the source directly on to the collimator objective. Hence the angular subtense of the source at the field-lens must be at least as large as that of the collimator objective. In the case of large aperture collimators and/or small sources this may require a field-lens of very short focal length, or of impractically low $F$ number.

Special problems are posed in the design of source optics for Raman spectroscopy. Here the source is of three-dimensional extent — the Raman cell — and because of the feebleness of the Raman scattered light it is important to collect as much light from the interior of the cell as possible, and to exclude any light scattered from the walls of the cell. This is achieved by imaging the interior of the cell into the collimator in such a way that the front face of the cell (the face nearest the slit) is imaged on to the collimator objective, and the rear face of the cell on to the slit. Correct choice of magnification ensures that no light scattered from the walls of the cell is transmitted by the collimator. The detailed design of such a system depends largely on the geometry of the Raman cell and of the spectrometer. No hard and fast rules can be laid down (apart from the above), and each case must be considered separately. In general, at least four lenses are needed for the source system, as indicated in figure 10.3.

If the aspect ratio of the cross-section of the cell differs from that of the dispersing element of the spectrometer, systems of crossed cylindrical lenses (anamorphic systems) can be of considerable benefit.
10.2 Detector Optics

The purpose of detector optics is to direct all the light passed by the exit collimator of the spectrometer on to the sensitive area of the detector. Hence the function of detector optics is the same as that of source optics in reverse, and very similar design considerations apply. Any of the systems of figures 10.1 or 10.2 can be used, with the detector replacing the source, and the direction of the light being reversed. The system of figure 10.2(a) is often preferred, as it is relatively immune from the effect of variations of sensitivity over the area of the detector.

Instead of using image-forming systems as described above, it is possible to employ 'light pipes' to guide the light from the spectrometer to the detector. A light pipe (often also called 'cone channel condenser' or 'axicon') is a conical tube whose internal surface is polished, and, if necessary, coated with a reflecting material. The diameters at the two ends of the tube are chosen so as to match the sizes of the exit aperture of the spectrometer, and of the sensitive area of the detector. Any light entering one end of the tube will undergo successive reflections on the wall, and will eventually emerge from the other end, provided that the inclination of the incident rays to the axis of the tube does not exceed a certain angle. This limiting inclination can be found by means of a simple geometrical construction, as shown in figure 10.4.

ABCD is a cross-section of the light pipe, and it is replicated so as to obtain the pattern shown, bounded by two concentric polygons (it is immaterial whether these polygons are closed or not). The condition that a ray entering the face CD should, after a number of reflections, emerge from the face AB is that the continuation of the incident ray should intersect the inner polygon. Thus, the ray 'a' in figure 10.4 meets this condition, while ray 'b' does not, and will eventually emerge from the face AB again. Light pipes, although not image-forming, are still subject to the law of conservation of etendue. In principle, a light pipe can always be replaced by a condensing lens. However, in many cases this lens might require an extremely low focal ratio on the detector side, leading to large aberrations of the image, and consequent loss of light. In these cases light pipes have a distinct advantage.

10.3 Source at Infinity

The case of a source at infinity arises principally in astronomical spectrometry. The source may be either a star, which can be regarded as a point source, or an object of finite angular extent such as a nebula or a planet. In any case, one deals with a source of fixed (and usually low) surface brightness, and the aim must be to collect and pass through the spectrometer as much radiation as possible.

The source optical system in astronomical spectrometry is usually the telescope itself. The amount of light collected by the telescope from a given source is proportional to the etendue, that is, to the product of the telescope area and the solid angle subtended by the source. In the case of a point (stellar) source this is therefore directly proportional to the area of the telescope objective (or primary mirror). The spectrometer, and any other optics used, must be designed in such a way that all the light accepted by the telescope is eventually passed on to the detector. The condition for this is that the focal ratio of the entrance collimator of the spectrometer should be equal to that of the telescope. If the collimator has a larger F-ratio, only part of the telescope aperture is used, with consequent loss of light, and if it has a smaller focal ratio, only part of the aperture of the collimator and of the dispersing element of the spectrometer is illuminated. Hence, provided that telescope and spectrometer are matched correctly, there is no optical advantage to be gained by using the telescope in its long focus (Coudé or Cassegrain) configuration in preference to the Newtonian or prime focus.
The telescope forms a diffraction image of a star, the central peak of this image having a diameter of \( \frac{2.2\lambda}{A} \), where \( A \) is the diameter of the telescope objective or mirror. If the focal ratios of telescope and spectrometer are equal, and the width of the spectrometer entrance slit is such that it will just accept this central peak, it can be shown that the full spectral resolving power of the spectrometer is attained. However, because of atmospheric turbulence (seeing), the angular size of a stellar image is usually at least 1" arc, which is much larger than the size of the diffraction image (for a telescope of reasonable size). Hence, either the width of the entrance slit of the spectrometer has to be increased to accept all the light from this larger image, with consequent loss of spectral resolution, or, if the slit is kept the same, there is an appreciable loss of light.

Similar considerations apply to the use of slit spectrometers in the study of extended astronomical objects. The spectral resolution is again determined by the width of the entrance slit, and this in turn determines the area on the celestial sphere from which the telescope/spectrometer combination collects light. In effect, the telescope projects an image of the entrance slit on to the sky, and only light from parts of the object within this slit image passes through the spectrometer. Thus, a certain amount of spatial resolution is possible, principally in a direction at right angles to the slit, and, depending on the length of the slit, in a direction along the slit. If the width and/or length of the slit is increased to accept more light, both spectral and spatial resolution suffer.

An entirely different situation prevails in the case of interference spectrometers of the Fabry–Pérot or Michelson types. Here the entrance aperture of the spectrometer is circular, and its angular diameter is of the order of \( \sqrt{R} \) times larger than the angular width of the slit of a grating or prism spectrometer of equal spectral resolution \( R \). Hence, in the case of stellar spectrometry, the loss of light or of spectral resolution caused by 'seeing' no longer occurs, provided that the star image (degraded by scintillation) is smaller than the entrance aperture. This condition can always be satisfied. The geometrical relations are demonstrated in figure 10.5.

D is the entrance aperture of the interferometer, and its angular diameter satisfies the relation
\[
\theta^2 = \frac{\lambda}{R}
\]

\( R \) being the spectral resolving power. The angular subtense of this aperture at the telescope mirror is \( \phi \), and this is also the angular diameter of the area on the sky from which the telescope/interferometer combination accepts light. \( L_2 \) is a field lens which images the telescope mirror \( L_1 \) on to the collimator \( L_3 \), and ensures that all the light from \( L_1 \) passing through \( D \) is transmitted by \( L_2 \). Then, if \( A \) is the diameter of the telescope mirror, and \( a \) is the diameter of the illuminated area of the interferometer, we have:

\[
\phi = \frac{f_a}{f_A} \sqrt{\frac{\lambda}{R}} = \frac{a}{A} \sqrt{\frac{\lambda}{R}} \tag{10.1}
\]

Hence, for a given telescope (that is given values of \( A \) and \( f_0 \)), the spatial resolution of an extended object can be increased (smaller \( \phi \)) by reducing the focal length of the interferometer collimator, which in turn entails a reduction of the étendue of the interferometer. In fact, taking the spatial resolution to be proportional to \( 1/\phi^2 \), the product of étendue and spatial resolution is a constant which depends only on the diameter of the telescope mirror. It is therefore always possible to increase spatial resolution at the expense of étendue, the limit being set by the faintness of the source.

10.4 Source and Detector Systems with Reflecting Optics

For reasons of compactness it is often advantageous to use reflecting optics instead of lenses for source and detector systems. The use of reflecting optics is essential in wavelength regions where no transparent refracting materials exist.

The exact layout of a reflecting system depends on the overall size of the detector (or source), and on the size of the sensitive element of the detector. The simplest case is that shown in figure 10.6, which uses a single concave mirror to image source or detector on to the spectrometer slit. This system is used whenever the overall size of detector or
source is small enough, so that the loss of light caused by the obstruction in the beam is negligible (say less than 10 per cent). If good imaging properties are required, the figure of the mirror should be an ellipsoid of revolution, with foci at the slit and detector (or source). Good imaging is more frequently needed in detector systems, because it is desirable to keep the sensitive area of the detector as small as possible. Hence an aberration free image of the exit slit must be formed to prevent light spilling over the edges of the sensitive area and being lost. If some aberration can be tolerated, one may use a (very much cheaper) spherical mirror.

If the detector is too large to permit the use of this simple on-axis system, one of the arrangements of figures 10.7 or 10.8 can be employed. Figure 10.7 uses an off-axis ellipsoidal mirror, with its foci again at the slit and detector. This mirror is an expensive component (although replica methods exist for production, once an exact convex master has been made), and must be mounted with some precision so that its foci lie at the required positions.

The system shown in figure 10.8, the so-called Pfundt system, is optically equivalent to figure 10.6, but uses a flat mirror to fold the light beam to prevent obstruction by the detector. Some loss of light still occurs, because the flat mirror must be pierced by a hole to allow light to pass to the detector. Figure 10.10 shows the arrangement of a spectro-meter, using reflecting optics throughout.

10.5 Double Beam Spectrophotometry

The purpose of double beam spectrophotometry is to eliminate errors due to fluctuations of the source, when measuring absorption, or to enable two sources to be compared with each other.

In either case it is usually the ratio of intensities at a particular wavelength which is desired, and so it is important that extraneous effects should be eliminated. The principles of the method are most easily demonstrated in the case of absorption spectrophotometry.

The basis of the method is to use a single source, and to direct beams from this source into one or two spectrometers, one of the beams being passed at some stage through the absorbing medium. The two beams are received either simultaneously on two detectors, or sequentially on one, and the intensities are compared electrically.

Systematic errors can arise from a number of causes, of which the most significant are:

(i) The use of two separate detection systems which may have different spectral responses.
(ii) The use of separate sources, or of two separate beams from one source, which may have different spectral intensity distributions.

(iii) The passage of the two beams along paths which are optically not equivalent, i.e. which entail different numbers of reflections, or transmissions, or intermediate imaging stages.

By suitable calibrations allowance can be made for these factors; but the apparatus can be enormously simplified, and its reliability improved, by eliminating as many as possible of these sources of error.

Consider then a system of double beam spectrophotometry which comes fairly close to the ideal, and which illustrates how the above difficulties can be avoided. It should be emphasized that this system is intended to demonstrate the important points of design, and is not necessarily applicable in all cases that may arise in practice.

The system, illustrated in figure 10.9, is an attachment to a normal spectrometer. It operates on the sequential system. One source is used to illuminate one detector, and the double-beam effect is obtained by switching between two paths.

A condenser lens C forms an image of the source on a reflecting chopper K. This chopper is a disc with an odd number of reflecting sectors and transparent gaps of width equal to that of the sectors. It rotates about an axis A. The lenses L₁ and L₂ collimate the light, which is alternately reflected and transmitted so that the beam passes through one of the sections P₁ or P₂, and is then refocused on to the chopper by the lenses L₃ and L₄. The final image is then re-imaged on to the entrance slit of the spectrometer in the usual way. The absorption cell is placed in one of the sections P₁ or P₂, and an exactly similar cell not containing any absorbing material is placed in the other. The system may be arranged so that the two cells can be interchanged easily. The detector then sees an alternating signal of approximately square wave shape, which can be amplified and recorded or displayed as required. Phase-sensitive rectification enables negative absorption coefficients to be dealt with (negative absorption is encountered in lasers and in materials where optical pumping takes place, and with fluorescent materials), but for ordinary purposes this is not essential. The rectified output of the detector is a measure of the ratio of the intensities of the two beams. The accuracy of this measurement depends on the linearity of the detector response, and on the linearity of the electronic circuits used for signal processing. Independence of linearity can be achieved by converting the system to make a null-measurement. For this purpose an attenuating element is introduced into one of the beams, and adjusted until the alternating output of the detection system is zero. The position of the attenuator is then a measure of absorption in the sample beam. Usually the attenuator takes the form of a comb with tapering teeth, so that if the comb is moved across the light beam the transmitted intensity varies in a linear manner. The comb can be moved across the beam by a servo-motor, the error signal for the servo system being derived from the detector output. This arrangement calls for a fairly advanced knowledge of servo-mechanism design, and may be eschewed by spectroscopists lacking expert knowledge or advice. The comb must be placed near an aperture stop in the system.

The important points illustrated by this system are:

(i) A single source, spectrometer, and detector are used.

(ii) The same source beam is used to provide both reference and sample beams, making the system independent of directional variation of emission from the source, and of intensity fluctuations, provided these are slow compared with the chopping frequency.

(iii) Both beams undergo identical optical processing.

(iv) Both beams enter the spectrometer along the same path, making the system insensitive to angular variation of the light transmission of the spectrometer.

(v) The source is imaged on to the chopper blades. This means that the blades need not be finished with the same accuracy as would be required to reflect a collimated beam. Furthermore, the chopped signal more nearly approaches a square wave shape. It is worth remembering that if the detection system is tuned sharply to the chopping frequency (as it should be to obtain the best signal/noise ratio), it will select the
fundamental of the chopping wave. This is sinusoidal, of course, and has an amplitude $4/\pi$ times that of the square wave. This is a slight gain, which is lost as the chopped signal becomes trapezoidal, or even triangular.

A disadvantage of this system is the large number of optical components needed. Considerable loss of light can be caused by this, and in particular in the vacuum-ultraviolet this system is unacceptable. A further drawback is that the optimum division of intensity cannot be achieved in cases where heavy absorption is expected. The signal/noise ratio can be shown to be a maximum when equal amounts of light are transmitted along each path.

The attachment may be placed before or after the spectrometer. In general it will be placed before it, so that secondary effects, such as fluorescence of the sample, are recorded in their proper positions in the spectrum.

The optical design of a commercial double beam spectrophotometer is shown in figure 10.10 (by courtesy of Beckman Instruments Inc.). In this instrument two synchronously rotating chopper discs are used to direct light from the source along one of the two paths shown in full or dotted lines. The radiation is then analysed in a double monochromator. The effect of this is to enhance the spectral purity of the transmitted light. At the same time the prism acts as an order sorter for the grating section of the monochromator. This arrangement naturally requires accurate synchronization of the scanning motions of the two sections of the monochromator by means of precisely machined cams (because of the non-linear dispersion of the prism).

An altogether simpler device may be used, which records both beams simultaneously, but fails to meet some of the requirements of the ideal system. The device is an attachment which follows the spectrometer, and it consists of an absorption cell before which is placed a reflecting comb, or some other means of extracting part of the beam emerging from the spectrometer before absorption has occurred.

The device, shown in figure 10.11, is in fact a source monitor. It is essential that the two detectors have the same spectral response, and their outputs must be amplified separately, and compared. The comb can, in principle, be driven by a servomechanism, the servo signal being given by the difference between the two amplifier outputs. Because of its inherent disadvantages this system is one to be avoided unless no other is available. For example, in the vacuum ultraviolet, where all optical components are forbidden, the comb may carry a suitable phosphor (see for example ref [1]).

A slightly more acceptable variant is to have a chopper instead of a comb, followed by a synchronous rectifying system to detect the difference between the two output signals. If this scheme is adopted, exceptional care must be taken with the chopper design. Since the signal is trapezoidal in form, switching must occur at exactly the right moment in each cycle, so that the correct amount of signal appears in each half-cycle.

10.6 Monitoring the Source

Monitoring the source, in general, is bad practice, since the source spectral intensity distribution may well change (especially in a flame or gas discharge) in a manner quite unrelated to the fluctuations of the total intensity. Source monitoring may be acceptable when the source is monochromatic, as for example in the case of the filtered blue ($4358 \text{ Å}$)
radiation of a Raman source. But even in this case the intensity may fluctuate at different rates in different parts of the source, and it is essential, for the monitoring to be meaningful, that the monitor should view the same part of the source as the Raman sample. The ideal procedure would be to chop the output from the sample between two spectrometers, one of which sees the Rayleigh scattered exciting line, while the other searches for Raman lines. The same spectrometer might be used for both functions, and this is one of the rare instances where fixed optics and a moving exit-slit-and-detector assembly are called for. A fixed exit slit and detector monitor the exciting line continuously. The practical difficulties of this scheme are considerable, especially when low Raman frequencies are to be detected (when the Raman lines are very close to the exciting line). In view of the progress made in the development of Raman lamps of high stability, the problem is not a serious one and in most cases source monitoring can be dispensed with in Raman spectroscopy.

Another instance where source monitoring is beneficial occurs in astronomical spectrometry. Because of variations of sky transparency the intensity of an astronomical source may fluctuate considerably during a spectrometer scan. If some of the light from the source, before dispersion in the spectrometer, is directed on to a detector, the output of this detector provides a record of the fluctuations, and enables the record spectrum to be corrected. This is the easiest done by making the jaws of the entrance slit reflecting, and arranging that the image of the source is slightly larger than the width of the slit. Hence some of the light will be reflected from the jaws, and can be collected by a detector. This system does not allow for variations in the spectral intensity distribution caused by variations in atmospheric absorption, and hence it will still be present.

A preferred method of source monitoring in astronomical spectrometry is to obtain the comparison signal from light after dispersion by the prism or grating of the spectrometer. The light reflected from the inside surfaces of the jaws of the exit slit is collected by a detector. This light will cover a moderate range of wavelengths, centered on the wavelength passed by the slit. If the spectrum is continuous, comparison of the light reflected from and transmitted by the slit yields a photometric signal not affected by fluctuations in the spectral intensity distribution. The method is less satisfactory if the spectrum contains absorption regions of a width comparable with the range of wavelengths contained in the comparison signal. The effect is to suppress low spatial frequencies.

10.7 Multiple Pass Absorption Cells

Absorption coefficients of gases are usually very low. To obtain usable absorption spectra of gases, great path lengths in the absorbing medium are frequently needed. Clearly it is inconvenient to use absorption cells possibly many metres long. The need for these is avoided by bouncing the light back and forth many times through the same volume of gas.

A cell using this technique has been designed by White [2]. Its optical layout is shown in figure 10.12.

A, A', and B are three concave mirrors with the same radius of curvature. The centres of curvature of A and A' lie on the surface of B separated by a small distance. The centre of curvature of B lies in a position halfway between mirrors A and A'. Light leaving any point of A is after reflection on B brought to a focus at the corresponding point of A'. After a further reflection on B it is returned to its point of origin on A. In a similar way, light from a point on B is returned to B at a point slightly displaced from its point of origin. The amount of this displacement depends on the separation between the centres of curvature of the mirrors A and A'. Hence, light coming from a point S just beside the edge of the mirror B will travel back and forth between the three mirrors, as indicated in figure 10.12. Eventually its points of origin and A', and it may be as high as 80.

The mirrors are mounted at the ends of the absorption cell, and entrance and exit windows are provided close to the edges of the mirror B. Because of the large number of reflections the mirrors must be coated for maximum reflectance if large losses are to be avoided. If the cell is
to be used over a narrow spectral range only, multi-layer dielectric coatings give the best performance. Otherwise high quality metallic coatings must be used. A factor to be borne in mind is that the coatings should be immune from chemical attack by any gases likely to be used in the cell.

REFERENCES


11.1 Constructional Materials

Scientific instruments are traditionally made of brass. This metal is easy to machine, does not readily suffer corrosion and will take a pleasant finish. It is not particularly cheap. In recent years it has become much more common to make instruments of small and medium sizes (up to about one hundred pounds weight) from an aluminium alloy. Many such alloys exist and a selection can be made for good machining properties, creep resistance and strength/weight ratio. Large instruments tend to be made of steel, possibly with aluminium alloy accessories when differential thermal expansion is not a problem. Parts are commonly machined from solid stock, cast and, less commonly, welded. Forging is very rarely undertaken unless, for some special reason, an unusual strength/weight ratio is required. Casting is not a particularly advantageous way of forming parts for 'one-off' spectrometers unless the shape is so awkward that an exorbitant amount of machining from the solid would be needed. There is the added disadvantage that castings tend to 'creep' or slowly alter shape with time for up to several months after they have been made. Welding is a satisfactory process for steel but must be used with caution with light alloys since the metal near the welded joint becomes soft and loses strength. With brass and similar non-ferrous metals hard or soft soldering is convenient (as indeed it is with steel, although welding is usually easier) but again light alloys suffer from such processes and are better fixed by screws and dowel-pins, or by an epoxy resin such as 'Araldite'.

Magnesium and its alloys are less often used for instrument construction, and indeed there is little to recommend them over aluminium alloys. They would normally be considered where weight is a considerable factor in the design (as for instance in an astronomical spectrometer for use at the Newtonian focus of a telescope), and if structures
are being designed near the strength limit of the material it will be found that the strength/weight ratio of magnesium alloys is not significantly greater than that of aluminium alloys. Corrosion is a trouble with magnesium, especially at the points of contact with other metals. For this reason it tends to be used rather infrequently in the aerospace industry.

Wood and plastic materials have their uses, chiefly for thermal insulation and for making light-tight boxes to cover the optical table when such a design is used. They are not of course structural materials and mechanical stability is not to be expected. In particular, resin-bonded plywood ('marine ply') and expanded polystyrene foam should be considered. Materials such as nylon and terylene in sheet or rod form play the part these days that used to be played by cork: as padding material when mounting glassware, and generally as shock-absorbent cushions.

Most of the parts of a spectrometer will probably be made by machining. Since the cheapest and easiest form of machining is turning, it may well be worth while making arrangements in the design for as much as possible of the structure to be turned rather than milled or shaped.

Grinding is a process only to be employed when a precisely flat load-bearing surface is required. Often the need for grinding can be avoided by using plate glass for such surfaces, since plate glass departs from the plane by no more than a few tens of fringes of green light. Selected portions depart by no more than a few fringes. This is as flat as, if not flatter than any ground surface can be made. There is a much lower limit to the specific loading that it will bear, but interferometer slides can be satisfactorily constructed from plate glass. Sliding surfaces, with low specific loading, are to be preferred to rolling surfaces such as ball-bearings.

Glass has little merit other than this as a structural material. Glass fibre, resin bonded, is to be considered as a covering material. It needs moulding, which puts it at a disadvantage with respect to wood, but has good thermal insulation properties, and light weight.

Brass, mentioned earlier in its historical sense, is to be considered, in spite of its cost, when stability against creep is required. Most other structural materials change shape with time, some being worse than others. Brass is one of the best from this point of view and is to be considered for optical assemblies that are to stay in alignment for weeks or months without attention. Other materials should not, in general, be so trusted.

Large spectrometers have their main structural members made of steel. Other materials are not readily obtainable in large sizes as are steel girders, and the fabrication of a box-girder in light alloy is likely to be more expensive than is warranted. A spectrometer of two or three metres focal length is not likely to be a portable instrument, and there is no advantage to be expected from using a light alloy. Steel girders of 'T' section or 'U' section such as those used for the construction of steel frame buildings are particularly suitable. After the necessary machining has been done, such members will need stress relieving, as also will box girders that have been fabricated by welding steel plates together. This process will normally distort the shape. Due allowance must be made for this in the design of the instrument. The distortion is unpredictable and one must not, for example, expect to find two machined faces at opposite ends of the structure remaining parallel through the stress-relieving process. In any case, one would not normally design an instrument that needed machined parallel faces. The parts of the instrument that are to carry the optical components should be attached to the structure through raised bosses which can be welded on and machined flat locally. Alignment will be by shims and by adjusting screws, with locknuts or their equivalent for parts which are to be semi-permanent.

Corrosion proofing for steel members is by rust and scale removal, painting with an oil-based red lead or zinc chromate paint and finally with a protecting oil-based paint, usually white if the paint is the outer surface of the instrument.

A spectrometer of what may be called 'intermediate' size, or 'transportable' (as opposed to portable, which means that it has a carrying handle) may be of steel or light alloy. The choice will depend on how much the instrument is going to weigh. Steel is to be preferred on account of ease of fabrication, but it is at this point that an aluminium alloy casting should be considered. A fabricated aluminium structure is least likely to be suitable since the necessary strength will be achieved only at great cost in workshop time. If an aluminium structure is essential, then two \( \frac{1}{2} \) in. plates separated by 2 in. or 3 in. ‘U’-section alloy extrusions may be satisfactory. There must be diagonal members as well as edging if the necessary stiffness is to be obtained.

For permanently installed instruments in the one or two metre focal length range one may consider obtaining an old lathe bed from a scrap merchant. This will in fact be a large casting, possibly many years old, well aged and past its creeping days, and will have a tolerably flat surface machined on top of it, to which optical components can be attached. (The price is likely to be reasonable too, a strong factor in impecunious University departments.)
11.2 Modes of Construction

Four basic ways of constructing a spectrometer can be distinguished:

THE OPTICAL TABLE
The main structural member is a more or less flat table which may be a ribbed casting, or may be stiffened by fixing girders underneath. The optical components are screwed or bolted on to this surface, sometimes with parts protruding underneath through holes. For example, it may be convenient to place the grating rotating mechanism out of the way of the optical beams in this fashion. The optical table is then covered with a light-tight box, and there are slots in the side at the proper places so that light can get in and out.

THE VERTICAL TABLE
Similar to the table above, but turned on its side so that the various chief-rays through the instrument are in the same vertical plane. The advantage of this method is that such a box has superior stiffness for its weight, and that both sides of the table are accessible for servicing and adjustment.

MONOCOQUE
The main structural member is a drawn seamless tube which not only aligns the various optical components but acts as a light-tight cover. It has the merits of stiffness and compactness but it can be inconvenient for servicing and adjustment unless there are large holes in it at the appropriate places. An instrument built to this sort of design probably takes up less space than any other.

SPACE FRAME
In this type of design there are various optical sub-assemblies and these are held in alignment by struts. This method tends to involve more workshop time and requires generally better workmanship than the others, and is to be considered chiefly when strength/weight ratio is important. Castings are often indicated for the sub-assemblies.

There are several design points which are common to all these different modes.

(a) The complete structure should have lugs at appropriate places so that the instrument can be attached to other pieces of equipment. If the instrument is normally to stand on a table then these lugs can be terminated by detachable feet. Unless it is known beforehand that the instrument is always to be used for one specific purpose, it is worth

while making the feet detachable, so that the lugs are the points of attachment.

(b) The focal planes of the instrument, that is, the planes where the slits or other field-stops are to be placed, should be outside the main structure of the instrument. There are several reasons for this. The chief one is accessibility.

It is most useful, in practice, to be able to place the eye close to the slit so as to peer into the instrument. If a monochromatic source in the visible region is allowed to shine through the instrument, and the wavelength control is set to pass that wavelength, then a Foucault test can be made of the focus of the whole spectrometer.

With spectrometers intended for the ultraviolet, it is often an advantage to be able to take a photograph, using a Schumann-type emulsion if necessary, of the region near the exit slit while a suitable ultraviolet source is used as illumination. This serves to check the focus and alignment at the working region directly, rather than by inference from the adjustments made in the visible region.

(c) There should be attachment points for detectors, or absorption cells or other auxiliary equipment. It is unfortunate that no agreed standard exists for such attachments. For most purposes a set of studs equally spaced on a pitch circle diameter of 4 in. or 6 in. will be sufficient. Only for very large spectrometers will the useful slit lengths be great enough to make these dimensions embarrassingly small.

11.3 Mounting of Optical Parts

The basic condition to be satisfied is that there should be no pressure sufficient to cause distortion of the optical component. Three basic types of component can be distinguished, and they need slightly different treatment in mounting. They are (i) diffraction gratings and prisms; (ii) mirrors, plane and spherical and (iii) lenses. Of these, mirrors should perhaps be considered in two categories according to size. Large mirrors, with diameters greater than about 8 in. need more elaborate fixing arrangements than small mirrors.

In general all these optical components will be mounted on brackets or pedestals that will permit adjustments to be made to their positions. From time to time it will be necessary to realign the spectrometer and for this purpose it must be possible to rotate and translate each optical component about its various axes. These adjustments are usually not very delicate compared with those of an interferometer for example, and they can be carried out using screws with a fine thread. It would be
extraordinary if a thread of \( \frac{1}{4} \) in. diameter and 48 threads per inch were too coarse. Such a size is readily available. For many purposes 32 or 26 threads per inch will be adequate.

To allow these adjustments to be made, the optical component will be mounted on a metal part, and the constraints to hold it in position and the forces to move it will be applied to this metal part.

**Mirrors.** If smaller than 8 in. in diameter these can be carried satisfactorily on a plane aluminium face-plate, to which they are glued with a rubber-based glue such as 'Bostik-2'. An epoxy resin should not be used, not only because of its permanence, but because of differential expansion between the glass or quartz of the mirror and the metal of the face-plate.

The face-plate can be attached to the main body of the instrument in several different ways. If the mirror is small, less than about 4 in. in diameter, the 'hole, slot and plane' system of kinematic mounting can be used. The face-plate is retained in position against three ball-ended screws by a spring in tension.

This method tends to be unsatisfactory for large mirrors, since the strength of spring needed to retain the mirror against vibration gets excessive. Large mirrors as well as small ones can be fixed by three pairs of push-pull screws as in figure 11.1. Tilt adjustments can be carried out in a horizontal or a vertical plane as shown, and two mutually perpendicular sets of these push-pull screws permit all the necessary adjustments to be made.

Large mirrors are usually mounted in cells. These are cylindrical cups, deep enough to allow the back of the mirror to rest against a \( \frac{1}{4} \) in. pad of cork or Nylon while an annular cover can be screwed on to the front of the cell with three pressure pads under it bearing lightly against the edge of the mirror. Occasionally an O-ring carried in a groove in the annular cover is used instead of the three pads. The inside diameter of the cell is such that it will contain the mirror with a \( \frac{1}{4} \) in. thick band of Nylon wrapped around it.

If this arrangement is considered to be too elaborate, it is possible to mount the mirrors without cells, provided they are to rest permanently in one position, without any chance of them being tipped by a movement of the instrument. A vee-block is constructed with a 120° angle, in which the rim of the mirror sits. Three clips hold it at the edges against a back plate. The whole assembly is attached to a further back plate or to a table by three pairs of push-pull screws.

**Lenses.** These are invariably mounted in cells. The same basic design is suitable for all sizes. It is most unlikely that the diameter of any lens used in practice will be greater than 6 in., and most likely will be less than 4 in. The cell is annular and the lens is held against an internal shoulder by a counter-cell. This counter-cell can be threaded or bolted on. It is usually threaded. Thin rubber or plastic washers can be inserted between the glass and the metal, but they are not usually considered necessary. A typical lens cell design is shown in figure 11.2.

**Gratings and Prisms.** These pose altogether more elaborate problems in mounting since, as well as the position and attitude adjustments, there has to be a means of rotating them with considerable precision. The angle of rotation must be measurable, the axis of rotation must remain parallel to its original direction and the whole motion must be repeatable. The degree of precision depends upon the particular instrument. Most important
is the parallelism of the axis and in a three-metre Ebert spectrometer this needs to be controlled to about one second of arc. This can be achieved by a properly designed axle. Large ball-bearings or roller bearings should be used and the correct grade should be chosen on the assumption that there is to be no play in the bearing and that the inner and outer races are to be push-fits in their respective holders. If the axis of rotation is to be vertical the top bearing may be an angular contact ball-bearing and the lower one a roller journal bearing. If the axis is to be horizontal, both the bearings may be angular contact ball-bearings, with the inner (rotating) shaft held in tension by weak springs.

Such a system is illustrated diagrammatically in figure 11.3. This system of two opposed angular contact bearings is a most useful one in all parts of optical instrument design where a comparatively small part is to be rotated precisely at a low rate. It is not sufficiently accurate for use with interferometers where the angular defects can only be of the order of one tenth of a second of arc. The devices are about one order of magnitude too coarse for this application.

It goes without saying that all parts of the axle should be constructed of the same material. This is normally mild steel. These devices need better than average skill in manufacture if they are to have the highest precision.

11.4 Linear Wavelength Drives

It is normally useful to have a means of changing the wavelength setting of the instrument in a way which is linear in wavelength or, sometimes, in wave-number. That is to say, there is somewhere a shaft emerging from the instrument such that the wavelength that appears at the exit slit is proportional to some angle that the shaft has turned through. Methods for achieving this depend again on the size and precision of the instrument. For comparatively small grating instruments, up to one or perhaps two metres in focal length, a simple sine-bar drive is indicated. A rod protrudes from the back of the grating cell or from some point on the axle. At the end of this rod is a hardened steel ball. This may well be a ball-bearing that has been forced into a slightly undersized cup. This ball bears against a plane disc that has been fixed to the end of a screw, and has been hardened, ground and polished until it is flat to a few fringes of green light and normal to the axis of the screw to 20 or 30 seconds of arc. This can be done with the aid of fine emery, rouge and an autocollimator. The screw will usually be a standard workshop micrometer screw. The amount of travel will depend on the size of the spectrometer, but usually an inch is sufficient. In a practical case the authors have found that one of these will maintain a linear wavelength drive in which the linearity is such that the scale varies by less
than one part in 500 from 4,000 Å to 7,000 Å. Resetting can be achieved to ±1/20th Å at all points. The system is illustrated in figure 11.4.

For larger spectrometers a somewhat more elaborate method may be necessary. Instead of the micrometer screw, a larger screw with a precisely ground thread (not tapped or screw-cut) is used, and a captive nut moves along it without rotating. To this nut is attached a rod which is hinged to a part of the grating mounting. It is important that the distance from the axis of rotation of the grating to the axis of this hinge should be the same as the length of the rod to the captive nut, and also that the axis of the screw should intersect the axis of rotation of the grating. If this is so, the position of the captive nut on the screw is proportional to the wavelength at the exit slit. By a proper choice of dimensions the constant of proportionality can be something simple such as 100 Å per turn. The screw is mounted between two angular contact bearings. Obviously it can be turned by a motor, but this drive should be flexible to avoid overconstraint, and the position signal should be measured from the screw itself or from the grating axis.

**Linear wave-number drives.** These should be needed rather less frequently than wavelength drives and one may well pause to inquire whether such a thing is really necessary. Linear wavelength drives are achieved easily by geometrical linkages, but linear wave-number drives require more trouble. Their chief use would seem to be in Raman spectrometry. One suitable system is illustrated in figure 11.5.

![Figure 11.5](image)

**Figure 11.5.** Linear wave-number (cosec bar) drive. The movement of the micrometer screw is proportional to the cosec of the grating angle, that is, to the wave-number.

It has been used by Gee [1] in constructing a Raman spectrometer. The mechanism requires many constraints in order to be accurate, and this makes it somewhat complicated in practice.

**Partial Principles of Construction - 193**

A point in favour of the linear wavelength drive is that no drive system is *perfectly* linear, and a calibration curve or formula is needed if the very highest wavelength accuracy is required. The same is true of linear wave-number systems. Thus for large precise spectrometers there is no great saving in effort when a linear (or quasi-linear) wave-number drive is used.

**Grating cells.** Gratings that are bought from manufacturers are normally replicas of an original which have been mounted on glass blanks. These blanks may be circular or rectangular, and the ruled area may cover the whole surface or may be rectangular with a plane-reflecting rim around it. Gratings may be mounted in cells in the same way as mirrors, and this is the only feasible method if they are on circular blanks. If they are on rectangular blanks it is possible to stand them on edge on a table and to clamp them lightly in a pair of clamps similar to the 'G-clamps' used in woodwork. There should be ample quantities of cork or Nylon to act as cushioning material at all points where pressure is applied. This method again is suitable only where the grating is to be used with the chief-rays horizontal. The table on which the grating stands then has the usual adjustments. It may be attached to a table beneath it by push-pull screws, or it may have three feet, adjustable in height by screws, so that it can stand on a hole, slot and plane on the lower table. This latter method has the advantage that the grating is readily removable for dust-proof storage or for replacement by another grating with a different grating constant for another spectral range.

If the grating is to be permanently fixed in position it is as well to have a dust cover made for it. The same applies to the other optical components of course, unless the instrument has no easy access ports, but in this case dust will only accumulate slowly on the optical surfaces. Depending on the degree of contamination of the local atmosphere, it is not likely that a totally enclosed spectrometer will need stripping down for cleaning more often than once every two or three years.

**A note on the cleaning of diffraction gratings.** The surface is most likely a plastic replica of an original, which has been coated with aluminium and possibly overcoated with some other substance to improve its reflectivity in the far ultraviolet or to act as protection against weathering. In any event the surface is very delicate, and no solid material, and very few liquids for that matter, may be allowed to come into contact with it.
If the surface shows signs of stain, or if there is a film of dust that is so thick as to impair seriously the reflectivity, an attempt may be made to remove this contaminating material by pouring gently over the surface a slow stream of diethyl ether. This ether should be of anaesthetic or best analytical quality, so that it will leave no stain.

If this ether treatment is unsuccessful other liquids may be tried, preferably on a small corner at first, to see how the grating likes them.

Nevertheless, prevention is better than cure, and a grating cover and an access hole to make the grating accessible is to be preferred.

11.5 Slit Jaws

The function of the slit is to define the field of the instrument as a long narrow strip or a long arc of a circle. Since the slits will often be only a few tens of microns wide, the edges of the jaws must be made with some precision if the slit width is not to vary significantly from point to point along it. The thickness of the material at the edge of the jaws should be small so that the least possible amount of light is scattered inside the instrument from the ‘lands’ of the jaws. A typical cross-section of a slit jaw is shown in figure 11.6. It is often convenient to allow the flat surface to face the outside world, and to have this surface large, accessible, and of matt white texture. This makes it easy to focus the source on to the jaws.

Methods of manufacture of slit jaws differ according to whether the jaws are straight or curved (as for instance in the Ebert spectrometer). Straight jaws are most easily produced on a good surface grinder from flat bar material of suitable size. It is vital before commencing grinding of the jaw edges that one of the surfaces of the material (the surface used for locating in the grinding machine) should be accurately flat. If this is not the case, the material will distort on clamping, and after being released on completion of the grinding the edge will no longer be straight. This difficulty, which can be particularly serious in the case of long jaws, can be circumvented by setting the jaws in plaster-of-Paris in a suitable mould, after rough machining and before grinding to the final shape and size. If a precision higher than that obtainable by grinding is needed, it can be achieved by lapping mating pairs of jaws against each other, using as abrasive a suspension of flour emery (optical grade) in oil, or, preferably, one of the commercially available diamond-lapping pastes.

Curved slit jaws are most easily made on a lathe and/or cylindrical grinder. A disc with a radius equal to the desired slit radius, and with a suitable edge profile is turned, and from this disc a number of (convex) jaws can be cut. The concave jaws are made from a ring, the inside diameter of which is exactly equal to the outside diameter of the disc producing the convex jaws. Lapping together of mating pairs of jaws is almost invariably required, because it is next to impossible to match the radii of disc and ring to the precision usually required for accurate slits.

The choice of material for slit jaws is determined by their size, and by the machine shop equipment available. For small jaws (say up to 25 mm long) nickel silver is very suitable. It is stable, corrosion resistant, easily machined, and will take a high surface finish. Its drawback is that it is soft, and the edges are easily damaged by dust particles trapped between them when the slit is closed. Large slit jaws are usually made from steel. The obvious choice is stainless steel, on account of its corrosion resistance. Stainless steel, however, is not an easy material to machine, and obtaining a good surface finish by grinding is difficult. An alternative material is ordinary carbon tool steel (probably in the form of ‘gauge plate’), which should be rough machined, hardened, stress relieved, and then finish ground and lapped. Corrosion resistance can be obtained by a thin (flash) plating of chrome, which should not alter the dimensions or shape of the jaws in a significant way.

11.6 Slit Mechanisms

Commercially manufactured slits often feature dove-tail slides for the jaws (or jaw carriers), and some kind of screw mechanism to adjust the width of the slit. This type of design, while neat and compact, demands a very high degree of skill in machining and fitting, and is very likely beyond the capacity of the average workshop attached to a research laboratory.

The demands on skilled workmanship can be relaxed to a very large extent by applying the principles of kinematic design, and by using flexible leaf springs to produce accurate rotary or straight-line motions. Two basic mechanisms are illustrated schematically in figure 11.7 (a) and (b). Both these schemes are for unilateral slits, in which one of the jaws is stationary, and the other moves to adjust the slit width. It is not
difficult to duplicate and couple two such arrangements if bilateral slits - with both jaws moving - are required.

In the arrangement of figure 11.7 (a) it is obvious that the edge of the moving jaw moves along a circular arc rather than along the straight line desired. However, as the amount of movement is small (a few millimetres at most) this effect is not serious, provided the two jaws lie in a plane when the slit is almost closed.

In both these mechanisms the hinge pivots of the moving parts require careful attention. Conventional hinges - a hole in both stationary and moving parts, and a pin fitting these holes - invariably have some play, and are not sufficiently precise in all but the most crude applications. Two types of hinge, not suffering from this drawback, are illustrated in figure 11.8 (a) and (b). Fig. 11.8 (a) is a variant of the V-groove and knife edge hinge, designed in such a way as to avoid the difficulty of making a V-groove with a perfectly sharp angle at the bottom of the groove. The mode of operation of this hinge, first applied by Horace Darwin to the well-known Cambridge rocking microtome, is obvious from the figure. A variant of this hinge uses ball-bearings force fitted into a plate in place of the V-groove faces F and F'. This type of hinge is particularly suitable for the mechanism of figure 11.7 (a), with the provision of a suitable spring to keep the two parts of the hinge in contact. A simple slit designed on these lines is illustrated in figure 11.9.

The other type of hinge consists of a thin, flat spring, as shown in figure 11.8 (b). In order to provide adequate torsional stiffness, the spring should be fairly wide, or, preferably, a pair of springs, separated by an appropriate distance, should be used. The design of such spring mechanisms has been fully discussed by Jones [2]. A diagram of a slit using spring supports is shown in figure 11.10. It is seen that this design employs a single spring hinge for the slit width control, and two double hinges constrain the movement of the carriage bearing the moving jaw to follow very closely a straight line. The precision of this motion is largely dependent on the equality of the lengths, widths, and thicknesses of the various spring strips. These should therefore be cut...
General Principles of Construction

Springs

Moving jaw

Fixed jaw

Supports for fixed jaw

Micrometer

FIGURE 11.10. Cross-section of slit mechanism employing spring hinges.

from one piece of material, and the lengths should be controlled by a gauge block which can be inserted between the fixed base and the moving jaw carriage.

The width of the slit is usually adjusted by means of a screw. Commercially available micrometer heads are a convenient source of supply of precision screws at a moderate price, and are to be preferred unless important design considerations call for a specially made screw and nut. The micrometer screw can be coupled to the rest of the mechanism by means of a ‘wobble pin’ as shown in figure 11.10, eliminating the need for accurate alignment of the screw with the rest of the mechanism. The linkage to the slit jaw should be designed in such a way that screwing in the screw will increase the slit width; when the screw is retracted a spring causes the jaw to follow it and close the slit. This arrangement avoids the possibility of damage to the slit jaws which could be caused if the screw acted in the opposite sense, and forced the jaws together.
200 - Design of Optical Spectrometers

A point to be borne in mind in the design of the slit is that the micrometer scale should be in a position where it can be read easily, without the use of dentists’ mirrors, magnifiers, etc.

11.7 Michelson Interferometers

The use of Michelson interferometers as spectrometers poses a number of problems of mechanical design. It is necessary to provide precise and stable adjustments of the two mirrors $M_1$ and $M_2$ (figure 11.11), and to have an accurately rectilinear motion of translation for one of the mirrors. The tolerance on this motion is that in any position in its

\[ \text{figure 11.11. Optical diagram of Michelson interferometer.} \]

traverse the mirror plane should remain parallel to its original position within about $\lambda/10$ over its entire surface.

Mirror adjustments can be subdivided into coarse and fine adjustments. It is usual practice to provide coarse adjustments for both mirrors, and fine adjustments for one only. A typical system of coarse adjustment is shown in figure 11.12. The mirror is mounted in a cell in the usual way. Three ball-ended screws, threaded into the support

\[ \text{figure 11.12. Coarse adjustment for interferometer mirrors.} \]

member, engage with a hole, slot, plane location, or alternatively with three radial V-grooves, in the back of the mirror cell. A spring pulls the mirror cell against the screws, which, after adjustment, are locked by means of locknuts. The screws should have a reasonably fine thread, and the $\frac{1}{4}$ in. $\times$ 40 T.P.I. model engineer’s thread is suitable.

The type of adjusting mechanism described is suitable only for bringing the interferometer mirrors into rough alignment. For final adjustment one of the mirrors has to be set parallel to the image of the other one (as seen in the beam-splitter) within a small fraction of a wavelength. If a screw is used to effect this adjustment, some kind of mechanical demagnification of the screw movement is needed. Two basic ways of doing this are available. The first relies on the differential deflection of springs of different stiffness, and is shown schematically in figure 11.13. The mirror cell is gripped between two springs fixed

\[ \text{figure 11.13. Fine adjustment by means of differential spring deflections.} \]
to the support member. One of the springs, \( S_1 \), is fairly stiff. The second, \( S_2 \), is thin and flexible, and carries the adjusting screw. Movement of the screw will alter the pressure exerted by \( S_2 \) on the mirror cell, and so cause a small deflection of the stiff spring \( S_1 \). By altering the ratio of the stiffnesses of the two springs, any desired degree of demagnification can be achieved. Two such devices, together with a fixed third support, enable the mirror to be tilted in any direction. Practical experience shows that the long-term stability of this arrangement is not very good. Temperature changes, and relaxation effects in the springs tend to cause slow drifts of the alignment.

An alternative, and preferable solution is to use levers to demagnify the screw movement. This type of mechanism is shown in figure 11.14.

The mirror cell \( M \) carries a ball-ended pin \( B \), which rests in a conical recess in the lever \( L \). The latter is supported by two screws, \( S_1 \) and \( S_2 \). The distance between \( S_1 \) and \( B \) is small compared with the length of the lever, and hence a movement of the screw \( S_2 \) causes a much smaller displacement of the mirror. Screw \( S_1 \) acts without sensible demagnification, and serves as a coarse adjustment. A second, similar mechanism acting at right angles to the first one, and a third fixed support point complete the location of the mirror. A single spring keeps all members of the linkage in contact.

**Figure 11.14.** Fine adjustment of interferometer mirror, using lever demagnification.

In practice, fine alignment of the mirrors is carried out by observing the fringes produced by the interferometer. This, however, cannot be done when the beam-splitter is opaque to visible light, as for instance in an interferometer for use in the infrared. In this case it is necessary to have a small reference area of the beam-splitter coated with a film which is semi-transparent in the visible, and carry out alignment on this reference area. While this procedure will not necessarily guarantee adjustment over the whole mirror area, it must be remembered that in the infrared the wavelengths are much longer, and the alignment correspondingly less critical.

In order to achieve the necessary scanning motion one of the mirrors must be translated in a direction parallel to its normal, and remain accurately parallel to its original orientation. The magnitude of this scanning motion depends on the wavelength, and on the resolving power desired (see Chapter 8). Typically, at a wavelength of 1 \( \mu \), a movement of 5 mm will give a resolving power of 10,000. In general, a movement of 1 cm, preferably either side of zero path difference, is sufficient for most cases (an exception occurs in the very far infrared, where the wavelength may be several hundred microns).

Various ways of constructing a suitable mechanical motion are possible, and are illustrated in the following figures.

Figure 11.15 shows a double spring arrangement, similar to that
Design

General Principles of Construction

rotation of the bar about its axis is to be prevented, it can be fitted with an outrigger which contacts a plane surface parallel to the direction of movement. A bar of this type can be made parallel and cylindrical to \( \lambda/4 \), and with an optical surface finish. The drawback of the system is very high specific loading which occurs at the points of contact between ball-bearings and the bar. Some wear is inevitable, which will eventually destroy the precision of the system.

A similar arrangement is shown in figure 11.18. Here the ground bar is replaced by an optically worked prismatic component. The slide which travels along this component is constrained by five ball-bearings, and hence does not need a further constraint to prevent rotation.

Some means is always provided to keep the moving part in contact with the micrometer screw. In most cases a spring or system of springs will serve for this purpose. However, if the measurement of the displacement of the mirror is derived from the micrometer (and not from some optical reference system), the variation of spring tension with extension will introduce noticeable errors into the interferogram. In this case a constant force should be provided. This is most readily done by using a weight, acting on the moving carriage by means of a suitable system of strings and pulleys.

If, however, the position of the spectrometer is likely to vary, as may happen for example if it is mounted at the Newtonian or Cassegrain focus of an astronomical telescope, then a more elaborate technique may be justified. A fusee can be used, as is done by chronometer makers to provide a constant torque on the escape wheel as the mainspring unwinds. The spring is contained in a drum and its power is applied through a chain which is wound on to a spiral groove cut into a cone.
As the spring unwinds and its torque diminishes, the chain winds on to the cone towards the tip. The moment arm through which the chain works then diminishes in proportion to the reduced torque and the force applied through the chain is kept constant.

REFERENCES


BIBLIOGRAPHY

A bibliography would seem to be more generally useful than an inordinately long list of references. We have consulted the following books extensively when designing instruments:

Whitehead, T. N. The Design and Use of Instruments and Accurate Mechanisms, Constable.

We can also recommend the study of the other numerous papers of R. V. Jones which appear in the Journal of Scientific Instruments and which describe many ingenious mechanisms than can be applied to spectrometer design.

12.1 Optical Components

These are almost always bought rather than made. A description of the manufacturing methods is clearly outside the scope of this book, but a knowledge of the appropriate materials for a chosen purpose and of their particular properties is necessary for the proper design of a spectrometer. We consider separately (a) Transparent materials suitable for prisms and lenses and (b) Materials for making plane and spherical mirrors.

12.2 Transparent Optical Materials

The most common optically transparent materials are optical glass and fused silica. Optical glass differs from the more ordinary varieties of window and plate glass in being made under strictly controlled conditions of composition and production. The aim of this control is to produce the desired refractive index and dispersion, homogeneity of optical properties, absence of scattering defects such as bubbles and 'stones' (solid opaque inclusions) and maximum transparency. Transparency is largely determined by the purity of the raw materials going into the melt and by the avoidance of contamination during melting. Some coloration is often unavoidable and very dense flint glasses may be noticeably yellow.

The classical method of manufacture is to melt the ingredients in large clay crucibles. Bubbles and striae (regions where the optical properties differ slightly from those of the bulk material) are removed by prolonged stirring of the melt. It is then allowed to cool at a controlled rate to avoid setting up strains in the mass of the glass. When cold the pot is broken up and the glass cut into slabs of suitable size. These are heat treated (annealed) to remove any residual strain. In the
molten state all glasses attack their crucible to some extent, and this results in a contamination of the melt and limits the time for which it can be stirred. The modern method to limit this contamination is to use platinum crucibles to contain the melt under a controlled atmosphere. In this way glasses can be produced which are markedly superior to pot-melted glasses, and which can contain a wider variety of ingredients. Tinting can be almost completely eliminated.

For quantity production of lenses and prisms the glass is usually remelted and pressed into moulds to form blanks approximating to the size of the finished component. There is a suitable excess of material (about 1 mm) on all surfaces to allow for grinding and polishing. The moulded blanks are subjected to an annealing process.

When glass components are ordered for optical instruments the tolerable levels of defects should be stated. Complete freedom from defects can be demanded at a price, but in some types of glass an absence of bubbles is impossible to achieve except in small selected pieces. Occasionally such pieces might be as small as 1 c.c. in a dense Barium Crown.

The same sort of quality can be obtained in fused silica. Thermal Syndicate Ltd. give specifications for their "Spectrosil "A"", as follows: 'Free from bubbles greater than 0.01 mm mean diameter in components up to 200 c.c. volume. Components larger than 200 c.c. may contain not more than one bubble greater than 0.01 mm but none greater than 0.11 mm diameter. The total bubble cross-section is less than 0.03 mm² per 100 c.c.'

Apart from homogeneity, the important properties of transparent materials are (i) absorption coefficient, (ii) refractive index and (iii) dispersion. The absorption coefficient \( \alpha \) is defined by an equation for the intensity remaining in a beam of light that has traversed a thickness \( x \) in the material. The equation is:

\[
I(x) = I(0) \exp(-\alpha x)
\]

and is known as Beer's law.

Absorption can be ignored as a design factor if the coefficient \( \alpha \) is less than 0.003 cm⁻¹.

The refractive index is usually not a quantity of prime importance in the design of spectrometers. Its value has to be known, of course, in order that focal lengths may be calculated, but one does not choose an optical material primarily for any particular value of refractive index. One possible reason for preferring to make lenses out of high index materials is that the primary aberrations are smaller for a given focal ratio. It is true that lower index materials suffer lower reflection losses, but these can always be reduced to negligible amounts by coating the surfaces with multilayer anti-reflection coatings.

Dispersion is of primary importance. The resolving power of a prism spectrometer is proportional to it, and one naturally searches for the highest possible dispersion. In glass manufacturers’ catalogues of glasses for the visible region, a quantity known variously as the 'reciprocal dispersive power' or 'Abbe reciprocal dispersion' or 'constriction' is quoted. It is defined by the equation:

\[
v = \frac{\mu_D - 1}{\mu_R - \mu_C}
\]

where \( \mu_D, \mu_R \) and \( \mu_C \) are the refractive indexes for at 5,893 Å, 4,861 Å and 6,563 Å respectively.

For a prism, the reciprocal dispersion should be as low as possible. It would be of more use to the spectrometer designer if the first two Cauchy coefficients were quoted, but this is not done. The symbol \( v \) is not universal. Occasionally one finds \( V \) and \( v \) instead.

**Silicon Dioxide.** This occurs naturally as 'crystal quartz' in pieces large enough for optical components to be made. It is birefringent and expensive, and its use is normally confined to the manufacture of components which make use of this birefringence. It also occurs naturally as sand in various degrees of purity in different parts of the world. The 'fused silica' that is usually specified instead of crystal quartz is superior for most optical purposes.

**Glass.** Although the word covers a wide variety of materials and describes a particular state of matter rather than a particular set of chemical compounds or mixtures, its use here is confined to those materials formed by adding various metallic salts to molten silica, and stirring the result. There are two broad categories: crown glasses and flint glasses. The former are characterized by low refractive index and low dispersion (high \( v \)-value) and the latter by high index and low \( v \)-value. There are hundreds of different glasses available to the optical designer, each specified by its refractive index and \( v \)-value, but in general these are not the concern of the spectrometer designer. The latter is more concerned with the region of transparency, and for most of the usual glasses this extends from about 3,600 Å to 2.5 \( \mu \).

For ultraviolet spectrometry it is necessary to use silica for wavelengths down to 1,900 Å, calcium fluoride (or fluorite or fluor spar)
Design
double
useful
filled
211
centimetre
no
water
A
these
at
Working
surface
Collodion
clean
detailed
essential
the
water
(\textit{c.c.})
covered
1,400
dissolved
two
pressure
and
on
solution.
these
made
using
film
and

\section*{Optical Components}

These are normally made of glass or of fused silica. The working requirements are on the whole different from those of transparent materials. They hold in common the requirements of good thermal stability and ability to take a polish. Mirror materials on the other hand may contain any amount of striations, bubbles and stones. Thus the material chosen is usually a silica or near-silica glass. The boro-silicate glasses in particular are useful and they go under a variety of trade names. It is usually not necessary to specify anything more than a ‘low-expansion glass’ when ordering. Silica is only chosen when extreme stability is required, as for example when Fabry–Pérot étalon flats are ordered.

The optical system will normally require that the surface finish should be accurate to 1/8th of a wavelength. In the visible region it is usual to use the green line of mercury at 5,461 Å in an interferometer to test the optical quality. The finish then demanded is ‘1/8th wavelength of green light’. In the ultraviolet this may not be good enough, and it may be necessary to ask for ‘1/20th of a wavelength of green light’. Finer working than this is very expensive, and normally it is only plate-pairs for Fabry–Pérot étalons that are specified to less than 1/20th of a wavelength.

When very high optical tolerance is not required it is worth bearing in mind that ordinary plate glass has a surface flat to five or ten wavelengths and quite large pieces can be selected with the aid of a Fizeau interferometer which are flat to one or two wavelengths. These often

to references [1] and [2] for particulars of transmissions and refractive indexes. There are glass-like substances such as fused silica and arsenic trisulphide, a range of sintered materials made by the Eastman Kodak Co. under the name of ‘Irtran’, crystalline materials such as sodium chloride (rock salt), potassium bromide, calcium fluoride and so on, and a large range of semiconductors. These latter in particular are remarkable for their high refractive indexes (typically about 4.0) which result in high reflection losses. The losses from one component may well account for over half the total incident energy. This means that it is essential to ‘bloom’ these materials with anti-reflection coatings whenever possible.

In the very far infrared (\(\lambda > 50\) microns) polythene is a useful material for absorption cell windows and other components not requiring high precision.

to references [1] and [2] for particulars of transmissions and refractive indexes. There are glass-like substances such as fused silica and arsenic trisulphide, a range of sintered materials made by the Eastman Kodak Co. under the name of ‘Irtran’, crystalline materials such as sodium chloride (rock salt), potassium bromide, calcium fluoride and so on, and a large range of semiconductors. These latter in particular are remarkable for their high refractive indexes (typically about 4.0) which result in high reflection losses. The losses from one component may well account for over half the total incident energy. This means that it is essential to ‘bloom’ these materials with anti-reflection coatings whenever possible.

In the very far infrared (\(\lambda > 50\) microns) polythene is a useful material for absorption cell windows and other components not requiring high precision.

\section*{12-3 Materials for Mirrors and Grating Blanks}

These are normally made of glass or of fused silica. The working requirements are on the whole different from those of transparent materials. They hold in common the requirements of good thermal stability and ability to take a polish. Mirror materials on the other hand may contain any amount of striations, bubbles and stones. Thus the material chosen is usually a silica or near-silica glass. The boro-silicate glasses in particular are useful and they go under a variety of trade names. It is usually not necessary to specify anything more than a ‘low-expansion glass’ when ordering. Silica is only chosen when extreme stability is required, as for example when Fabry–Pérot étalon flats are ordered.

The optical system will normally require that the surface finish should be accurate to 1/8th of a wavelength. In the visible region it is usual to use the green line of mercury at 5,461 Å in an interferometer to test the optical quality. The finish then demanded is ‘1/8th wavelength of green light’. In the ultraviolet this may not be good enough, and it may be necessary to ask for ‘1/20th of a wavelength of green light’. Finer working than this is very expensive, and normally it is only plate-pairs for Fabry–Pérot étalons that are specified to less than 1/20th of a wavelength.

When very high optical tolerance is not required it is worth bearing in mind that ordinary plate glass has a surface flat to five or ten wavelengths and quite large pieces can be selected with the aid of a Fizeau interferometer which are flat to one or two wavelengths. These often
212 - Design of Optical Spectrometers

make adequate mirrors for fore-optics. The thermal stability is not good of course, but it will not need to be with this lesser optical tolerance.

**TABLE 12.1**

Physical properties of materials suitable for mirror and grating blanks.

<table>
<thead>
<tr>
<th>Material</th>
<th>Density (Grams/cm³)</th>
<th>Strength/Weight Ratio (Dyne-cm/gram)</th>
<th>Coeff of Thermal Expansion (per°C)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aluminium</td>
<td>2.70</td>
<td>2.56 x 10¹¹</td>
<td>13.3 x 10⁻⁷</td>
</tr>
<tr>
<td>Beryllium</td>
<td>1.82</td>
<td>15.4 x 10¹¹</td>
<td>12.4 x 10⁻⁶</td>
</tr>
<tr>
<td>Cer-Vit</td>
<td>2.50</td>
<td>3.70 x 10¹¹</td>
<td>10 - 0.15 x 10⁻⁷</td>
</tr>
<tr>
<td>Fused Silica</td>
<td>2.20</td>
<td>3.18 x 10¹¹</td>
<td>5.5 x 10⁻⁷</td>
</tr>
<tr>
<td>Magnesium</td>
<td>1.74</td>
<td>2.59 x 10¹¹</td>
<td>14.0 x 10⁻⁶</td>
</tr>
<tr>
<td>Pyrex</td>
<td>2.35</td>
<td>2.89 x 10¹¹</td>
<td>3.2 x 10⁻⁶</td>
</tr>
<tr>
<td>Pyroceram</td>
<td>2.50</td>
<td>3.48 x 10¹¹</td>
<td>2.2 x 10⁻⁷</td>
</tr>
</tbody>
</table>

† The thermal coefficient of Cer-Vit varies from one blank to another, but is constant throughout a given blank. The value for each blank is measured by the manufacturers, Owens-Illinois, Toledo, Ohio, U.S.A.

12.4 Filters

These are used to isolate a more or less narrow region of the spectrum and can act as order sorters or as crude spectrometers in their own right. (The U-B-V system of star classification used for years by astronomers is a crude sort of spectrometry.) There are band-pass filters, band-stop filters and edge-filters. The filtering action may be due to absorption by chemical compounds (dye filters), naturally occurring absorption edges, as for example in some semiconductors like selenium, or due to Fabry–Pérot étalons made from two high-reflection layers with a single half-wave layer sandwiched in between. These latter are interference filters. Any combination of the above varieties may be selected to isolate a chosen region.

In the visible and closely adjacent regions dyed gelatine filters are both cheap and useful. They are available with a very large variety of transmission characteristics, and from a number of different manufacturers. For normal use the gelatine is mounted between two pieces of glass and the assembly is sealed at the edge. In particular, excellent edge-filters can be obtained for the visible region, with almost complete transparency at wavelengths greater than a certain value, and adequate filters can be made which will reject wavelengths above a certain value.

Because of the absorbing properties of dyestuffs, these are less efficient than the former. In other words there are excellent yellow filters which completely reject blue, but less efficient blue filters which will reject yellow and red, but have a lower and more uneven transmission in the blue and violet.

When really narrow bands are to be isolated (Δλ < 200 Å) it is necessary to use an interference filter. These filters are multilayer filters and are combined with suitable dye filters which serve to remove unwanted orders of interference. When a very low order of interference is used the number of layers in each reflecting stack can be made large, so that there is a high finesse, and the adjacent orders are several hundred Angstrom units away. It is usual for sideband cutters to be supplied by the manufacturer of the filter, unless he be specifically requested to omit them. It may, on occasion, be worth omitting the long-wave filter if the filter is to be used with a detector that is not sensitive to these longer wavelengths. There may well be a factor of two gained in overall transmission in the experimental region. Because these filters are Fabry–Pérot étalons the transmitted wavelength depends on the angle of incidence. The peak wavelength is shifted to the blue if the filter is tilted. Some care must consequently be used if the filter is inserted into a convergent beam. If, at normal incidence, the wavelength at the peak of transmission is λ, then at an angle of incidence θ from air on to the filter the peak transmitted wavelength will be λ − Δλ, where Δλ = (θ / μ)λ / 2. (μ is the refractive index of the material in the spacer. It is normally zinc sulphide in the visible region, with a refractive index of 2.37.)

When the filter is used in a convergent beam or in a collimated beam from a source of finite size, there is a broadening of the passband and some reduction of the peak intensity that is transmitted. The position of the filter and the design of the optical system must be arranged if necessary so that the range of angles at which light is transmitted through the filter is a minimum. A typical arrangement is to place a lens close behind the final field-stop with the filter close behind the lens. The focal length is chosen to be equal to that of the preceding lens so that the aperture is effectively at infinity. All parts of the filter then receive light from the same range of solid angles. A further lens will usually follow the filter in order to complete the field lens that is to image the aperture on to the detector. The arrangement is illustrated in figure 12.2. If the solid angle from the aperture is too large for the filter, the optical system must be changed. If the solid angle is reduced, the area of the filter will go up, in accordance with the law of conservation of étendue.
Design is semiconductors for sharp valence hole semiconductors. These absorptions are better at the order limit but is always approached if a filter is available. The principal properties required of a reflecting-coating material are (i) high reflectivity for the radiation concerned, (ii) good resistance to mechanical damage, (iii) good resistance to chemical attack, particularly from the atmosphere. These latter two requirements arise from the fact that most mirrors in precise optical instruments are 'first surface' mirrors, so that the reflecting material is not protected by being deposited on the rear surface of a transparent substrate.

Solid metal mirrors, the use of which can be traced back to antiquity, are rarely employed at the present time. The reasons for this are, firstly, that the reflectivities that are obtainable are lower than those that can be achieved with vacuum-deposited films, and, secondly, that the mechanical properties of most metals make it difficult to ensure long-term stability of the optical figure of the mirror. Mirrors of stainless steel or 'stellite' are sometimes employed in particularly corrosive environments, but their manufacture to high standards of precision is costly and difficult. Mirrors of light aluminium alloy and of beryllium have also been made for space-craft telescopes. A beryllium mirror might be expected to be very expensive and the long-term stability of mirrors of these types has not been explored at the time of writing. There is in any case little point in choosing such mirrors for any normal spectrometric application.

Mirrors usually consist of a glass or silica substrate that has been polished to its proper figure, on which is deposited a film of metal or a series of films of dielectrics, by sputtering or by vacuum evaporation. The thickness of such a reflector can be kept constant within a small fraction of a wavelength, and it is the substrate alone which determines the final figure of the mirror.

In the visible and adjacent regions a freshly deposited silver film has a reflection coefficient of 0.95. Silver films tarnish rapidly on exposure to the atmosphere, particularly an industrial atmosphere containing sulphur compounds, and so they are only used when they can be protected, as for instance in metal-dielectric interference filters.
Design of Optical Spectrometers

The most commonly used coating material is aluminium. This is easily evaporated from a tungsten wire and forms films of reflection coefficient in the vicinity of 0.87-0.90. On exposure to air, a film of aluminium oxide forms within a few hours and this gives both mechanical and chemical protection to the film. The protection can be improved by 'anodizing' the film or by depositing on it by further vacuum evaporation a film of silicon monoxide [3]. Instead of using pure aluminium, some people have tried different alloys with varying amounts of success. The best appears to be an alloy of 95 per cent aluminium $Al + MgF_2 (T = 250 \text{ Å})$ and 5 per cent chromium, which gives adequate reflectance together with improved abrasion resistance [4].

The reflection of aluminium mirrors is satisfactory down to wavelengths in the region of 1,600 Å. Below this the reflection coefficient falls rapidly, as in figure 12.3. At 900 Å it is down to 0.1 and at shorter wavelengths remains at this level until the soft x-ray region is reached. A significant improvement can be achieved between 1,500 Å and 950 Å by overcoating the aluminium with a layer of magnesium or lithium fluoride.

In the infrared, aluminium mirrors can be used, but a slightly higher reflection coefficient can be obtained from a film of evaporated gold. Although resistant to chemical attack, these gold films are delicate and have poor resistance to abrasive attack. This means that they cannot be cleaned by wiping with a cloth and so must be protected from dust by covering them when not in use.

Aluminium surfaces with a silicon monoxide overcoat will withstand considerable abrasive attacks, and can be cleaned with soap and water and similar materials.

REFERENCES

Appendix 1

Fundamental Design Formulae

(1) Plane Grating Spectrometers

Greatest resolving power:
\[ R = \frac{\lambda}{\Delta \lambda} = m \cdot n \] (order of interference, \( m \)) \times \frac{n}{(\text{total no. of lines in grating, } n)}

Free spectral range:
Working in \( m \)th order,
\[ \Delta \lambda = (\lambda_{\text{max}} - \lambda_{\text{min}}) = \frac{\lambda_{\text{min}}}{m} = \frac{\lambda_{\text{max}}}{m + 1} \]

Grating equation for oblique incidence:
(a) General case:
\[ (\sin i + \sin r) \cos \phi = m \lambda / a \]
where \( i \) and \( r \) are measured in the plane perpendicular to the rulings, and \( \phi \) in a plane perpendicular to that.

(b) Ebert spectrometers:
\[ 2 \sin i \cos \omega \cos \phi = m \lambda / a \]
where \( i \) is the angle between the grating normal and the axis of symmetry; \( \omega \) is the Ebert angle and \( \phi \) is as in the equation above.

Longest straight slit to match a given straight slit:
\[ a^2 / 8 = 1 / R \]
where \( a \) is the angular length of the slit, or the true length divided by the focal length of the appropriate collimator, \( R \) is the required resolving power.

Differential dispersion:
If \( \Delta i \) and \( \Delta r \) are the angular widths of the entrance and exit slits they are related by the equation:
\[ \frac{\Delta i}{\Delta r} = \frac{\cos r}{\cos i} \]

Relations involving the resolving power:
\[ \tan i = \frac{R \cdot \Delta i}{2} \]
\[ R_{\text{max}} = \frac{W}{\lambda} (\sin i + \sin r) \]
where \( W \) is the width of the grating.

Angular dispersion and wavelength are related by:
\[ \left( \frac{\Delta \lambda}{\Delta i} \right)^2 = \left( \frac{\lambda}{m} \right)^2 - \left( \frac{\lambda}{2} \right)^2 \]

(This equation is often useful as it does not involve the angles \( i \) and \( r \).)

Ghosts:
Ghost intensities vary as \((v \cdot m)^2\). \( v \) is the wave-number, \( m \) the order of interference.

Spacings from the parent line are:
\[ v' = v + p \beta \quad p = \pm 1, \pm 2, \pm 3, \ldots \]
or \( \Delta \lambda = \pm p \beta \lambda^2 / m \) where \( \beta = 1 / (\text{periodic error}) \) and is measured in cm\(^{-1}\)

Resolving-power Luminosity relationships:

(a) With emission spectra, and slit length = 1/50th of focal length:
\[ E = RL = \frac{A}{50} (\sin i + \sin r) \]

(b) With slit length limited by resolving power required:
\[ E' = R^{1/3} L = 2.83 A (\sin i + \sin r) \]

At the theoretical maximum resolving power, and appropriate slit length, the efficiency becomes:
\[ E = 2.83 \cdot H \cdot \lambda \]
where \( H \) is the length of one ruling.

(c) With slit length limited as in (b), and with continuous emission spectra, the efficiency becomes:
\[ E'' = R \cdot E' \quad \text{or} \quad R^{1/3} L \text{ is constant.} \]

Focal ratio of an Ebert spectrometer to ensure that the performance is limited by the grating quality:
\[ F = \frac{1}{4} \sqrt{\frac{R \cdot \cot \beta}{3}} \]

or
\[ F = \frac{1}{4} \sqrt{\frac{2 \pi \cdot R}{\lambda \cdot R_{\text{max}}}} \]
where $\beta$ is the angle between the grating normal and the axis of symmetry, $w$ is the projected width of the grating, seen from the mirror in question, $R$ is the required resolving power and $R_{\text{max}}$ is the theoretical limit of resolving power of the grating. (The two formulae are equivalent.)

Grating position for a flat focal surface, with curved line images:

at a distance $0.846 \times (\text{focal length})$ from the mirrors.

Position of the grating for straight line images but a curved focal surface:

half-way between the mirrors and the focal surface.

(2) Fabry-Pérot Spectrometers

Diameter of the ring at the $n$th order of interference:

$$D^2 = 8.f^2 \left(1 - \frac{n\lambda}{2d}\right)$$

Free spectral range: $\Delta \lambda = \frac{n\pi.r}{\lambda^2} \approx \frac{3n}{n^2}$

Resolving power:

$$R = \frac{\lambda}{\Delta \lambda} = \frac{n\pi.r}{\lambda^2} \approx \frac{3n}{n^2}$$

Finesse:

$$F = \frac{\Delta \lambda}{\lambda} = \frac{\pi.r}{\lambda^2} = \frac{R.\lambda}{2d} \approx \frac{3}{n^2}$$

Field-stop diameter:

$$\theta^2 / 8 = 1 / R \quad (R = \text{resolving power required};$$
$$\theta = \text{angular diameter of stop})$$

Resolving power-luminosity product:

$$E = R. L = 2\pi.A \quad (A = \text{available area of aperture})$$

(3) Michelson Multiplex Fourier Spectrometers

Field-stop diameter:

(a) Normal: $\theta^2 / 8 = 1 / R \quad (R = \text{resolving power required};$
$$\theta = \text{angular diameter of stop})$

(b) Field-compensated:

$$\theta^2 = 384\mu^3 / R$$

Resolving power-luminosity product:

(a) Normal: $E = R. L = 2\pi.A \quad (A = \text{available area of aperture})$

(b) Field-compensated:

$$E' = R. L^2 = 24(\mu A\pi)^2 \quad (\mu = \text{refractive index of glass})$$

Resolution and sampling:

Maximum sampling interval to avoid aliasing overlap:

$$\Delta = 1 / 2\nu_{\text{max}} = \frac{\lambda_{\text{min}}}{2} \quad (\nu_{\text{max}} = \text{highest wave-number in spectrum};$$
$$\lambda_{\text{min}} = \text{shortest wavelength})$$

If spectrum does not occupy the principal alias:

$$\Delta = a / 2\nu_{\text{max}} = a.\lambda_{\text{max}} / 2 \quad (\lambda_{\text{max}} = \text{longest wavelength in spectrum})$$

where $a$ is the greatest integer less than $\nu_{\text{min}} / \Delta\nu$ and the spectrum extends from $\nu_{\text{min}}$ to $\nu_{\text{max}} + \Delta\nu$.

Number of resolved elements in reconstructed spectrum:

$$= \text{to number of sample points taken in interferogram on one side of zero path difference.}$$

Resolution, $\delta\nu = 1 / (\text{greatest path difference}).$

(This figure is independent of any aliasing that is allowed).

Formulae for computation:

(1) Symmetric. No phase error:

$$S(N) = \frac{J(0)}{2} + \sum_{n=1}^{N-1} J(n\Delta\phi) \cos \left(\frac{\pi n}{N}\right) + \frac{J(N)}{2} (-1)^n$$

(2) Phase error correction needed:

$$S(N) = \frac{1}{2} \sum_{n=-N}^{N} J(n\Delta\phi) \sin \left(\frac{\pi n}{N}\right)$$

whence

$$e = \frac{1}{2\pi\nu} \tan^{-1} \left[ \frac{S(\nu)}{S(\nu)} \right] = \frac{N.\Delta\phi}{\pi^2} \tan^{-1} \left[ \frac{S(\nu)}{S(\nu)} \right]$$

$$J'(m\Delta\phi) = \sum_{n=-N}^{N} J(n\Delta\phi) \sin \left[ 2\pi\nu(e + (m - n)\Delta\phi) \right]$$

Then transform with the equation in the section preceding.
Appendix 2
Alignment of Spectrometers

The adjustment of a spectrometer is done in a number of steps, and the procedure can be speeded up considerably by proceeding in a systematic manner. Alignment methods will be described with particular reference to grating spectrometers. The following operations have to be carried out:

(1) Adjustment of the centre heights of all optical components and of entrance and exit slits.
(2) Adjustment of the angular orientations of the optical components.
(3) Focusing of the entrance and exit collimators, and then of the complete spectrometer.
(4) Adjustment of the grating rulings parallel to the axis of rotation of the grating.
(5) Adjustment of the exit slit to coincide with the monochromatic image of the entrance slit.

To make these adjustments the following equipment is needed:
(a) helium-neon laser, operating in single TEM mode.
(b) White-light source.
(c) Source giving a line spectrum (for instance a mercury lamp).
(d) Travelling microscope.
(e) Small plane mirror, some white cards, stands and clamps.

Method:

(1) With ruler, calipers, set-square and protractor, align all the optical components as closely as possible to their design positions.
(2) Place screens made from thin white card over the collimator mirrors and over the grating. Each of these screens is pierced by a small hole of about 5 mm diameter in the position where the chief-ray is intended to strike the mirror or grating (in the case of a single-mirror Ebert spectrometer the screen over the mirror has two holes, one for the incident and one for the diffracted chief-ray).
(a) With the slits fairly wide open, shine a laser beam through the centre of the entrance slit.
(b) Adjust the laser so that the beam passes through the hole in the screen over the first collimator mirror.
(c) Adjust the tilt of this mirror so that the beam passes through the hole in the screen over the grating.
(d) Adjust the grating to direct the beam to the hole in the exit mirror screen.
(e) Adjust the exit mirror so that the beam finally leaves through the centre of the exit slit.

The grating should be set to zero order, so that it acts as a plane mirror. The precision of these adjustments can be increased by placing a fine thread over the centre of the entrance slit and using the diffraction fringes produced by this as fiduciary marks.

(3) Place over the upper half of the entrance slit a mirror inclined at 45° to the focal plane and shine a beam of white light via this mirror into the spectrometer. The angular aperture of this beam should be sufficient to fill the aperture of the collimator. With the grating set normal to the beam reflected by the collimator mirror, an image of the upper half of the entrance slit is formed at the lower half and can be observed with a microscope. This image should lie in the same plane as the slit, and the distance between slit and mirror is adjusted until this is the case. Use the highest power microscope objective which can be accommodated to give a small depth of focus. The same procedure is repeated with the exit collimator, and finally the focus of the complete spectrometer is checked by observing the image of the entrance slit in the exit slit, using the grating in zero order. If large focusing adjustments have been made it is advisable to repeat step (2), and then check the focus again as above. A final, most sensitive check on focus can be carried out by means of the Foucault knife-edge test. With the eye as close as possible to the exit slit the light distribution over the collimator aperture is observed and the grating is slowly scanned through zero order. If the whole aperture appears to brighten and darken uniformly the focus is correct. If light or shadow appears to spread over the aperture from one side, further adjustment is needed. In a high-quality system at low focal ratio the characteristic ‘zone’ pattern of spherical aberration can be seen.

(4) Place a fine thread exactly across the centre of the entrance slit.
With white light, and in zero order, tilt the grating until the image of the thread appears at the centre of the exit slit. Then rotate the grating on either side of zero order as far as mechanical limitations permit, and observe the image at the exit slit with a microscope. If this image moves vertically (along the slit) rotate the grating in its plane until the image remains stationary and coincident with the thread over the exit slit.

(5) Blank off the centre part of the exit slit, leaving clear only two short sections at the top and bottom of the slit. Illuminate the spectrometer with a source giving a line spectrum, and scan slowly through one line using a detector and displaying the output on a chart recorder. If the exit slit is not parallel to the image of the entrance slit the recording will appear as a doublet. Adjust the slits until a single line of minimum width is recorded. This adjustment is particularly critical for spectrometers with long, curved slits. It should be checked at a number of wavelengths through the working range of the spectrometer. To obtain satisfactory adjustment over the required range it may be necessary to make small corrections to the adjustments described under heading (4).

Appendix 3

Readily Available Wavelengths for Laboratory Use and Wavelength Standards

(1) The Principal Solar Fraunhofer Lines

<table>
<thead>
<tr>
<th>Name</th>
<th>Wavelength</th>
<th>Element of Origin</th>
</tr>
</thead>
<tbody>
<tr>
<td>Y</td>
<td>8990.0</td>
<td></td>
</tr>
<tr>
<td>Z</td>
<td>8223.5</td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>7621.0</td>
<td>O</td>
</tr>
<tr>
<td>B</td>
<td>7594.0</td>
<td>O</td>
</tr>
<tr>
<td>C</td>
<td>6678.95</td>
<td>H</td>
</tr>
<tr>
<td>D</td>
<td>5095.94</td>
<td>D₁</td>
</tr>
<tr>
<td></td>
<td>5096.90</td>
<td>D₂</td>
</tr>
<tr>
<td>E</td>
<td>5289.56</td>
<td>Fe</td>
</tr>
<tr>
<td>b</td>
<td>5153.82</td>
<td>Mg</td>
</tr>
<tr>
<td></td>
<td>5172.70</td>
<td>Mg</td>
</tr>
<tr>
<td>d₂</td>
<td>5196.51</td>
<td>Fe</td>
</tr>
<tr>
<td>F</td>
<td>5184.33</td>
<td>Mg</td>
</tr>
<tr>
<td>G</td>
<td>4958.34</td>
<td>H</td>
</tr>
<tr>
<td></td>
<td>4960.48</td>
<td>H</td>
</tr>
<tr>
<td>G’</td>
<td>4070.91</td>
<td>Fe</td>
</tr>
<tr>
<td></td>
<td>4070.75</td>
<td>Ca</td>
</tr>
<tr>
<td>g</td>
<td>4226.74</td>
<td>Ca</td>
</tr>
<tr>
<td>h</td>
<td>4101.75</td>
<td>H</td>
</tr>
<tr>
<td>H</td>
<td>3966.49</td>
<td>Ca</td>
</tr>
<tr>
<td>K</td>
<td>3933.66</td>
<td>Ca</td>
</tr>
<tr>
<td>L</td>
<td>3933.66</td>
<td>Ca</td>
</tr>
<tr>
<td>M</td>
<td>3727.84</td>
<td>Fe</td>
</tr>
<tr>
<td>N</td>
<td>3541.21</td>
<td>Fe</td>
</tr>
<tr>
<td>O</td>
<td>3441.02</td>
<td>Fe</td>
</tr>
<tr>
<td>P</td>
<td>3361.19</td>
<td>Ti</td>
</tr>
<tr>
<td>Q</td>
<td>3286.77</td>
<td>Fe</td>
</tr>
<tr>
<td>R</td>
<td>3181.28</td>
<td>Ca</td>
</tr>
<tr>
<td>S₁</td>
<td>3100.58</td>
<td>Fe</td>
</tr>
<tr>
<td></td>
<td>3100.33</td>
<td>Fe</td>
</tr>
<tr>
<td>S₂</td>
<td>3099.94</td>
<td>Fe</td>
</tr>
<tr>
<td>s</td>
<td>3047.62</td>
<td>Fe</td>
</tr>
<tr>
<td>T</td>
<td>3021.07</td>
<td>Fe</td>
</tr>
<tr>
<td>t</td>
<td>2994.4</td>
<td>Fe</td>
</tr>
<tr>
<td>U</td>
<td>2947.9</td>
<td>Fe</td>
</tr>
</tbody>
</table>
(2) Lines of the Common Laboratory Sources

<table>
<thead>
<tr>
<th>Cadmium</th>
<th>Helium</th>
<th>Mercury</th>
<th>Sodium</th>
<th>Zinc</th>
</tr>
</thead>
<tbody>
<tr>
<td>6436.4686*</td>
<td>70, 580</td>
<td>39, 425</td>
<td>11, 404.2</td>
<td>6562.4</td>
</tr>
<tr>
<td>5086.42</td>
<td>70, 930</td>
<td>11, 220</td>
<td>11, 363.4</td>
<td>6314.6</td>
</tr>
<tr>
<td>4799.91</td>
<td>7055.2</td>
<td>10, 140</td>
<td>8184.33</td>
<td>8111.6</td>
</tr>
<tr>
<td>4678.15</td>
<td>6778.15</td>
<td>7729.2</td>
<td>6183.33</td>
<td>6102.5</td>
</tr>
<tr>
<td>5875.83</td>
<td>7092.0</td>
<td>5895.93</td>
<td>8021.3</td>
<td></td>
</tr>
<tr>
<td>4471.40</td>
<td>6807.5</td>
<td>5800.07</td>
<td>5894.6</td>
<td></td>
</tr>
<tr>
<td>4026.19</td>
<td>9790.98</td>
<td>4830.92</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3646.05</td>
<td>9790.98</td>
<td>4732.18</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5407.74</td>
<td>4660.14</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4368.55</td>
<td>4046.56</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3863.28</td>
<td>3654.83</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3650.15</td>
<td>3121.84</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3131.56</td>
<td>2556.53</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* The cadmium line at 6438 Å is an international standard wavelength. All wavelengths are vacuum wavelengths, and are in Angstrom Units. 
Ángstrom Unit = 10⁻⁸ cm.

Note: These are not intended as complete lists of wavelengths, but are intended as a quick reference to the more commonly used bright lines in laboratory sources. A more thorough and complete list of lines, which every spectrometric laboratory should possess, is to be found in the well-known Tabelle der Hauplinien der Linienspektren aller Elemente, H. H. Kayser, Springer, Berlin.

---

**TABLE I. A table of the function sinc x**

<table>
<thead>
<tr>
<th>x</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1</td>
<td>0.1</td>
<td>0.07</td>
<td>0.05</td>
<td>0.03</td>
<td>0.02</td>
<td>0.01</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>0.1</td>
<td>1</td>
<td>0.95</td>
<td>0.92</td>
<td>0.89</td>
<td>0.86</td>
<td>0.83</td>
<td>0.8</td>
<td>0.77</td>
<td>0.74</td>
</tr>
<tr>
<td>2</td>
<td>0.07</td>
<td>0.95</td>
<td>1</td>
<td>0.98</td>
<td>0.95</td>
<td>0.93</td>
<td>0.92</td>
<td>0.91</td>
<td>0.89</td>
<td>0.88</td>
</tr>
<tr>
<td>3</td>
<td>0.05</td>
<td>0.92</td>
<td>0.98</td>
<td>1</td>
<td>0.99</td>
<td>0.98</td>
<td>0.97</td>
<td>0.96</td>
<td>0.96</td>
<td>0.95</td>
</tr>
<tr>
<td>4</td>
<td>0.03</td>
<td>0.89</td>
<td>0.95</td>
<td>0.99</td>
<td>1</td>
<td>0.99</td>
<td>0.98</td>
<td>0.97</td>
<td>0.97</td>
<td>0.96</td>
</tr>
<tr>
<td>5</td>
<td>0.02</td>
<td>0.86</td>
<td>0.93</td>
<td>0.98</td>
<td>0.99</td>
<td>1</td>
<td>0.99</td>
<td>0.98</td>
<td>0.98</td>
<td>0.97</td>
</tr>
<tr>
<td>6</td>
<td>0.01</td>
<td>0.83</td>
<td>0.92</td>
<td>0.97</td>
<td>0.98</td>
<td>0.99</td>
<td>1</td>
<td>0.99</td>
<td>0.99</td>
<td>0.99</td>
</tr>
<tr>
<td>7</td>
<td>0</td>
<td>0.77</td>
<td>0.89</td>
<td>0.96</td>
<td>0.97</td>
<td>0.98</td>
<td>0.99</td>
<td>1</td>
<td>0.99</td>
<td>0.99</td>
</tr>
<tr>
<td>8</td>
<td>0</td>
<td>0.74</td>
<td>0.88</td>
<td>0.96</td>
<td>0.98</td>
<td>0.99</td>
<td>0.99</td>
<td>0.99</td>
<td>1</td>
<td>0.99</td>
</tr>
<tr>
<td>9</td>
<td>0</td>
<td>0</td>
<td>0.77</td>
<td>0.96</td>
<td>0.98</td>
<td>0.99</td>
<td>0.99</td>
<td>0.99</td>
<td>0.99</td>
<td>1</td>
</tr>
</tbody>
</table>
TABLE 3. The Airy function $I(\delta)$ for arguments $\delta = 0$ to $\delta = \pi$ for various values of the finesse $F$.

<table>
<thead>
<tr>
<th>Finesse $F$</th>
<th>10</th>
<th>20</th>
<th>30</th>
<th>40</th>
<th>50</th>
<th>60</th>
<th>70</th>
<th>80</th>
<th>90</th>
<th>100</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\delta = 0$</td>
<td>1.0000</td>
<td>1.0000</td>
<td>1.0000</td>
<td>1.0000</td>
<td>1.0000</td>
<td>1.0000</td>
<td>1.0000</td>
<td>1.0000</td>
<td>1.0000</td>
<td>1.0000</td>
</tr>
<tr>
<td>$\delta = \pi$</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
</tr>
</tbody>
</table>

Note: The table entries represent the Airy function $I(\delta)$ for different values of $\delta$ and $F$. The values are rounded for readability.
Subject Index

(Figures in bold type indicate major references.)

Aberrations, 20
Absorption cells, 176
   multiple pass, 181
      in spectrophotometry, 175, 179
Airy function
   in Fabry-Pérot Étalon theory,
      95 et seq.
      table of, 229
Alias, principal, 125, 221
Aliasing, 123
Alignment of spectrometer, 222
Anamorphic systems, 103, 169
Anti-reflection coatings, 211
Aperture, 18
Aperture division, 138, 139
Aperture, entrance, 6
Aperture, exit, 6
Aperture stop, 2, 18, 19
   of Michelson spectrometer, 136
Aplanatic
      lenses, 28, 29, 165
      points, 27, 28
      surfaces, 28
      systems, 27 et seq.
      for immersed detectors, 165
Apodizing, 37, 38, 54
Apodizing functions, 127, 138
   effect on line profiles, 37
Apodizing Masks, 54
Araldite, 183
Arsenic trisulphide, 211
Astigmatic curvature
   in concave grating images, 91
   Astigmatism, 24 et seq.
      of a concave grating, 90
      of an Ebert spectrometer, 76
      of a field compensated interferometer, 135
      of a spherical mirror, 26
      of a thin lens, 27
Astronomical spectroscopy
   interference spectrometers in, 172
   Michelson spectrometer in, 138
   source monitoring in, 180
   source optics for, 171
Attenuator, comb, 177, 179
Autocorrelation function, 127
Axicon, 138, 170
Axis, optic, 15
Axle, for grating table, 190
Beam-splitter, 117, 129, 136
   defects, 140
Beer's law, 208
Blaze angle, 55, 57
Blaze efficiency, 56
Blazing
   of a diffraction grating, 55 et seq.
Bolometer, 152
   Superconducting, 154
   Thermistor, 153
Bostik, 188
Box-car function, see Functions, 37
Calcium fluoride, 209, 211
Cambridge microtome hinge, 196
Cardinal planes and points, 16
Cardinal theorem, of interpolatory
function theory, 120
Cartesian sign convention, 17
Cassegrain system, in focusing
optics, 136
Casting, use in spectrometer con-
struction, 183
Cat's ears, as étalon supports, 114
Cat's eye reflectors, 20, 129, 130, 131
Cauchy constant, 43, 209
Cells, for mounting optical com-
ponents, 189, 193
absorption, 175
fine adjustment of, 201
multiple pass, 181
Cerenkov effect in photomulti-
pliers, 158
Cer-Vit, 212
Chief-ray, 19, 63, 76, 214
Circle of least confusion, 21, 22, 76
Coating, antireflection, 211 et seq.
Coatings, reflection, 215
Coded disc multiplex systems, 128, 144
Collecting power, 2
Collimator, 2, 6, 119, 163
Collodion films, for U-V windows, 210
Coma, 20, 22 et seq.
Fraunhofer's condition for the
removal of, 23
in concave gratings, 90
in Ebert spectrometers, 72
of a spherical mirror, 23
of a thin lens, 22
Coma-like terms, in concave
grating theory, 92
Comb attenuator, 177, 179
Condenser, cone-channel, 170
Condenser lens, 30, 167
Conservation of étendue, 2, 11, 213
Constringence, 209
Construction, materials of, 183 et seq.
Construction, methods of, 186 et seq.
Convolutions, 7, 8, 34 et seq.
Algebra of, 36
physical applications of, 36
Convolution theorem, 34
Cooley–Tukey algorithm, 126, 145
Corrector plate, in a Schmidt
system, 22
Corrosion proofing, for steel struc-
tures, 185
Criteria of choice of a spectrometer,
12
Crystal quartz, 209
Cube-corner reflectors, 129
Czerny–Turner Mounting for
Plane Grating, 73
crossed, 73
details of, 75
see also Ebert–Fastie mounting
D*, 150, 151, 153, 165, 166
table of, 164
Deconvolution, 7 et seq.
Delta (δ)-function, 8, 35, 122
Design formulae, 18
Design, optical, 14
Detectivity, D*, 150
for current limited detectors,
152
effective, for semi-conductor
photodetectors, 165
Detector optics, 170
detectors, 146 et seq.
characteristics of, table, 164
Gelat, 155
linearity of, 146, 177
photon, 147
photovoltaic, 147, 165
semi-conductor photon, 146 et seq.
thermal, 147, 152
upper frequency limit of, 152
Diffraction
limit to resolution of a system, 22
order of, 3, 52, 53, 55, 57, 61
Diffraction gratings
blazing of, 5
cleaning of, 193
concave, 81 et seq.
astigmatic curvature in, 91
Eikonal theory of, 83
mountings for
Eagle, 85, 86
Johnson–Onaka, 86
Miyake, 88
Paschen–Runge, 89
Seya–Namioka, 88, 89
Wadsworth, 82, 87
echelon, 55
mechanical mounting, 187 et seq.
mountings for, 70 et seq.
resolving power of, 52
Diffraction grating equation, 51 et seq.
Discriminator, pulse height, 160
Dispersion, 43, 45
of a grating, equation, 61
at oblique incidence, 62 et seq.
angular, 219
differential, 68, 218
partial, 44
reciprocal, 209
Doppler broadening, 36
Double-beam spectrophotometry,
175 et seq.
Dye filters, 212
Dyneo multiplications factor, 158
Dyneo resistor chain, 161
Eagle mounting, for concave grat-
ing, 85, 86
Ebert angle, 60, 72, 76, 79
Ebert–Fastie mounting
for Prism, 48
for Plane Grating, 69, 72, 218
choice of focal ratio for, 76, 219
choice of mirror positions for,
76, 220
Newtonian, for plane grating, 74
spectrograph conversion, 75
Echelon grating, 55
Edge filters, 212, 214
Efficiency, of a spectrometer, 3
quantum, 148
relative, of a grating, 56
Eikonal, 83
Emission spectrum, 4
Emulsion, photographic, 4, 5
Entrance aperture, 2, 6
angular diameter of, in étalon, 99
Entrance pupil, 23, 30
Enveloping curvature
in concave grating line images, 92
Epoxy resin, 183
Equivalent reflector, 135
Étalon, Fabry–Pérot, 3, 95 et seq.
finesse, defined, 97, 100, 220
free spectral range, 98, 220
matching, to grating spectrometer,
102
to optical telescope, 103
mounting methods, 112 et seq.
multifaceted, 105
vernier ratios of gaps, 106
order sorter for, 103, 105
resolving power of, 98, 220
resolving power–luminosity prod-
duct, 102, 220
ring diameters of, 99, 220
transmission losses in, 104
Étendue, defined, 2, 56, 173
conservation of, 2, 11, 27, 69,
137, 168, 171, 213
Exit Aperture, 6
Fabry–Pérot étalon, 95 et seq. (see
Étalon, Fabry–Pérot)
Fellgett advantage, 118, 140, 144
Féry Prism, 49
circle, 49
Field, 18
Field compensated interferometer,
119, 129, 133 et seq.
efficiency of, 221
semi-diameter of field, 136, 220
Field curvature, 24 et seq.
Field division, in Michelson, spec-
trometer, 138, 139
Field flattening lens, 25, 27
Field lens, 18, 137, 168
Field stop, 2, 9, 19, 100, 213
des of Fabry–Pérot spectrometer,
99, 220
of Michelson spectrometer, 136,
220
Filters
Dye, 212
edge, 212, 214
gelatine, 212
interference, 105, 106
Finesse, of a Fabry–Perot étalon, defined, 97, 220
coefficient of, 98
effective, 100
flatness, 101
First order aberrations, 20
Fluorspar, 209
Focal curve theorem, 43
Focal length
of a thick lens, 15
of a thin lens, 15
of two thin lenses, 15
Focal plane, 16, 187
Focal point, 16 et seq.
Focal ratio, 3, 4
of a cat’s eye reflector, 130
of an Ebert spectrometer, 76, 129
Focal surface, sagittal, 24
tangential, 24, 25
Focus, Gaussian, 21, 22
Fourier pairs, 33, 37
Fourier reciprocity, 33
Fourier series, 31
Fourier’s inversion theorem, 32, 120
Fourier transforms
complex, 33
cosine, 32, 51, 118, 126
formulæ for computation, 221
practical methods, 125
sine, 32, 126
Fraunhofer lines, 4, 41
as a test of performance, 7
list of principal, 225
Fraunhofer’s condition, for removal of coma, 23
Free spectral range
of a diffraction grating, 57, 58, 218
of a Fabry–Perot étalon, 98, 220
Functions
antisymmetrical, 32
apodizing, 54, 127
asymmetrical, 32
box-car, 37
complex, 33
Gaussian, 39
Lorentz, 39
periodic, 31
Θ, 37
top-hat, 37
triangle, 38, 143
truncating, 37
Fused silica, 208, 209, 212
Fusée, 205
Gap, étalon, variations in, 101
Gaussian focus, 21, 22
optics, 14 et seq.
profile, 8, 39, 101
Gelatine filters, 212
Geometrical path, equivalent, 133
Ghost, grating, 59 et seq., 219
Glass
crown, 209
flint, 209
as a structural material, 184
Goly detector, 155
Gratings, see Diffraction gratings
Grating table, design of axle for, 190
Hanning, 127
Hanning, 127
Helmholtz–Lagrange theorem, 19, 193
Homodyne rectifier, 149, 177
Images
curved, of a curved slit, 65
parabolic, of a straight slit, 64
Image tube, 45
Incidence, angle of, 42
Information
bandwidth, 4
channel, 4
rate, in photon counting photometry, 4, 160
system, 5
theory, 4, 5
Infrared, 59
luminosity of spectrometers in, 79
multiplex spectrometers in, 119
spectrometers for, 79
transparent materials for, 210
Instrumental profile, 7 et seq., 35, 37, 106
of a Fabry–Perot spectrometer, 97
of a grating spectrometer, 53
of a Michelson spectrometer, 127, 142
Interference filters, 212
pass bands of, 213, 214
Interference, order of, 97, 218
Interferogram, 128, 122, 123
Interferometer
field compensated, 119, 129, 133
et seq.
Michelson, 117
Twymon–Green, 22
Interpolation theorem, 126
Irtran, 211
Isoplanatic system, 30
Japanese mountings, for concave gratings, 85 et seq.
Johnson noise, 149
Johnson–Onaka mounting, 82, 85, 87
Kinematic mounting, 188
Kinematic hinges, 196 et seq.
Laboratory light sources, List of wavelengths, 226
Lens
anamorphic, 11
aplanatic, 28, 29, 165
condenser, 267
field flattening, 25
field, 18, 168
mechanical mountings, 189
thick, formula for focal length, 15
thin, formula for focal length, 15
Light grasp, 2
Light pipe, 170
Line profile, 7, 36, 39, 53
Linear wavelength drive, 191
Linear wave-number drive, 192
Linearity of detector response, 146, 177
Lithium fluoride, 210
Lithium niobate, pyroelectric detector, 156
Littrow
mounting for grating, 71
mounting for prism, 47, 71
Lorentz profile, 39
LR product, see Resolving power–luminosity product
Luminosity, 2, 10
in the infrared, 79
of a Fabry–Perot étalon, 102
of a grating spectrometer, 69
of a Michelson spectrometer, 136
of a prism spectrometer, 43
Magnetostriuctive scanning of an étalon, 111
Maksutov system, 22
Materials of construction, 183 et seq.
Mechanical scanning, of an étalon, 111
Methods of construction, 186 et seq.
Michelson interferometer, 117
Cassegrain system for, 136
construction of, 200 et seq.
in astronomical spectrometry, 172
practical details, 128 et seq.
cat’s eye, 131
cube-corner, 130
field compensated, 119, 129, 133
Möbius-band, 132
tilting, 131
Mirrors, 17 et seq., 215
ellipsoidal, 174
mechanical mountings, 188
paraboloidal, 21, 47, 70
plane, 52, 129
spherical
astigmatism of, 26
spherical aberration of, 21
coma of, 23
Reflectivity (contin.)
  of coatings in Fabry–Pérot étalon, 101, 115
Refraction at a surface, 15
Resolution
  defined, 9
  of concave grating, 93
  of a Michelson spectrometer, 120, 221
Rayleigh’s criterion for, 10
Resolving power
  defined, 2, 9
  diffraction limit of, 3
  of a Fabry–Pérot spectrometer, 98, 220
  of a grating spectrometer, 52, 218
  of a Michelson spectrometer, 119, 221
  of a prism spectrometer, 42
Resolving power-luminosity product
  defined, 3
  of a Fabry–Pérot spectrometer, 102, 220
  of a grating spectrometer, 69, 219
  of a Michelson spectrometer, 118, 220
  of a prism spectrometer, 43
Responsivity of a detector, 151
RL Product, see Resolving power-luminosity product
Rowland
  circle, 81 et seq.
  condition, 84
  cylinder, 90, 91
  gratings, 81 et seq.
Sagittal focal surface, 24, 25, 26
Sample points, 123
Sampling interval, 124
choice of, 142, 221
Sampling theorem, in multiplex spectrometry, 120, 123
Scanning
  of a Fabry–Pérot étalon magnetostriective, 111
  mechanical, 111
  piezo-electric, 109
  pressure, 107
  of a Michelson spectrometer, 129 et seq.
  of a spectrometer, 5
  Schmidt optical system, 20, 23, 24, 26, 138
  Schumann photographic emulsions, 187
Sciillation, 172
Seeing fluctuations
  in astronomical spectroscopy, 172
  in multiplex spectrometers, 141
Seidel theory of aberrations, 20, 25
Seya-Namioka mounting, for concave grating, 82, 88, 89
Shift theorem, 33
Shot-noise, 141
Side-band cutter, 213
Signal/noise ratio
  in multiplex spectrometers, 118, 141
  in photon counting, 160
Silver reflecting coatings, 215
Sinc function, 8, 36, 54, 120, 128
  defined, 38
  table of, 227
Sine^2 function
  table of, 228
Sine-bar, 93, 191
Slit
  straight, images of, 64
  theory of curved, 66
  jaws, 94
  bilateral, 196
  curved, 195
  materials for, 195
  mechanisms, 195 et seq.
  unilaterial, 195
Slit length, 3, 65, 67
Slit width, angular, 68
Slitless spectrograph, 6
Smith–Helmholtz invariant, 10
Source monitoring, 179
  in astronomical spectrometry, 186
Source optics, 167 et seq.
  for Raman spectrometry, 169
  in astronomical spectrometry, 171
  matching to spectrometer, 167
  reflecting, 173 et seq.
Sources, light, list of wavelengths, 226
Space-frame construction, 186
Spectrograph, 3, 4, 24, 45
  conversion, 75
Slitless, 6
Spectrometers
  Concave grating, 82 et seq.
  mountings of
    Eagle, 86
    Johnson–Onaka, 86
    Miyake, 86
    Wadsworth stigmatic, 85, 87
  Plane grating, 51 et seq.
  mountings of
    Czerny–Turner, 73
    Ebert–Fastie, 72
    Littrow, 71
    Newtonian Ebert, 74
    Pfundt, 74
  Prism, 41 et seq.
  mountings of constant deviation, 46
    Ebert, 48
    Péry, 49
    Littrow, 47
    Wadsworth, 47
  Spectrophotometry
    absorption, 175
    double-beam, 175
  Spectroscope
    classical, 41, 44
    Schuster, 49
  Spectrum, solar, Fraunhofer lines in, 225
  Spherical aberration, defined, 21
  of an aplanatic system, 27
  of a field compensated interferometer, 136
  of a concave mirror, 21
  of a thin lens, 21
  Stray light ratio, 7
  Strip spring hinges, 196, 199, 203, 204
  Superconducting bolometer, 154
  Surface brightness, 8

Synchronous rectifier, 149, 177
Table, optical, 184, 186
Tangential focal surface, 24, 25, 26
Telescope, 18
  astronomical, 103, 171 et seq.
  Thermocouple, radiation, 155
  Thermistor bolometer, 153
Thick lens, focal length of, 15
Thin lens, focal length of, 15
Throughput, defined, 2
  tilting Michelson interferometer, 129, 131, 132
  optimum angle of, 136
Top-hat function, 37, 38, 59, 100, 120, 127
Transparent materials for optical components, 207 et seq.
Triangle function, 7, 38, 68, 127
Tripartite grating rulings, 82
Truncation of interferograms, 120, 138
Ultra-violet, vacuum, 57, 82, 93, 179
  transparent materials for, 209
Vacuum ultra-violet, 57, 82
  colloidon films for, 210
  design considerations for, 93
  double-beam apparatus for, 179
  Vernier ratios, of étalon spacings, 106
Vignetting, 2, 18, 19, 131, 137
Voigt profile, 40
Wadsworth mounting, for prism spectrometer, 47
Wadsworth stigmatic mounting, for concave grating, 82, 84, 85, 87
Wavelength
  of a linear drive, 192
  list of laboratory and Fraunhofer line, 225, 226
  Wave-number, 31, 117 et seq.
  linear drive, 192
  White cell, 18
  Wobble-pin, 199