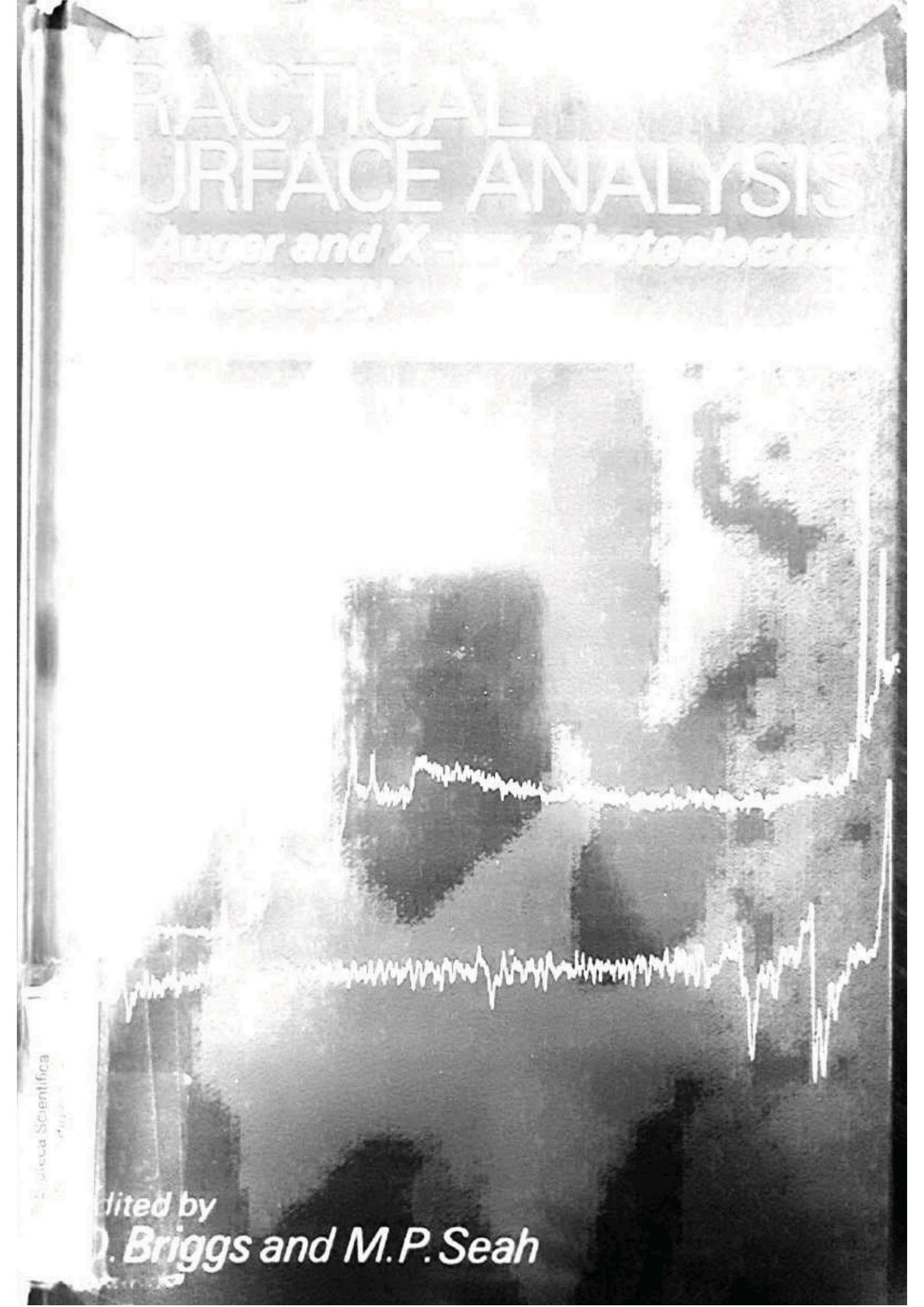
RACTICAL URFACE ANALYSIS

Auger and X-ray Photoelectron



dited by
. Briggs and M. P. Seah

Contents

Preface

The amazing growth and diversification of surface analytical techniques began in the 1960s with the development of electron spectroscopy—first Auger electron spectroscopy (AES), closely followed by X-ray photoelectron spectroscopy (XPS or ESCA). Today these two complementary techniques still dominate the surface analysis scene with the total number of operational instruments well over one thousand. Since the 'heart' of either technique is an electron energy analyser, AES and XPS are frequently found in combination, with essentially no loss of performance in current fourth generation instruments. AES in the form of high resolution scanning Auger microscopy adds the surface compositional dimension to scanning electron microscopy. The addition of an ion source for sputter removal of surface layers allows either XPS or AES to perform composition depth profiling.

In recent years these techniques have rapidly matured with the generation of a large body of literature, the growth of 'user' groups in Europe, the United States and Japan, the systematic improvement of standards of operation and procedures for quantification, and so on. AES and XPS are now routinely used in a large number of industries; indeed, in some cases their development has been intimately connected with the growth of new industries (e.g. AES and the microelectronics industry). Many centres offering contract research facilities have appeared. Each year more and more people come into contact with the techniques through the need to solve problems relating to surface or interface composition and many people became practically involved in using the techniques without the advantage of relevant formal academic training. Although literature reviews of aspects of AES and XPS are legion, no text exists which reflects the new maturity of these techniques in a way which is useful *practically*, especially for newcomers to applied surface analysis.

The aim of this book is to correct this omission and to present, in one volume, all of the important concepts and tabulated data. A brief introduction gives the historical background to AES and XPS and sets them in the perspective of surface analytical techniques as a whole. The essentials of technique are covered in chapters on instrumentation, spectral interpretation, depth profiling and quantification. The remaining chapters are intended to give an insight into the major fields of application, both in terms of the special attributes of AES and XPS and the contribution they have made. These fields are

microelectronics, metallurgy, catalysis, polymer technology and corrosion science. Throughout, the underlying electron spectroscopy link between AES and XPS is stressed. Aspects of technique which have fundamental importance in day-to-day operation such as instrument calibration, XPS binding energy referencing and XPS data processing (especially complex curve resolution) are discussed in the Appendices. Finally, there are full tabulations of major peak positions in AES and XPS, relative sensitivity factors for XPS and binding energy/Auger parameter data for elements and compounds.

Wilton
Teddington
February 1983

D. BRIGGS
M. P. SEAH

Practical Surface Analysis
by Auger and X-ray Photoelectron Spectroscopy
Edited by D. Briggs and M. P. Seah

Chapter 1

A Perspective on the Analysis of Surfaces and Interfaces

M. P. Seah
Division of Materials Applications, National Physical Laboratory, Teddington, Middlesex, UK

D. Briggs
ICI PLC, Petrochemicals and Plastics Division, Wilton, Middlesbrough, Cleveland, UK

1.1 Introduction

Surface analysis and surface science are evolutionary disciplines. They are where they are at present as a result of a number of small steps by a very large number of researchers. Progress along the whole front has not been even and advance has been rapid in certain areas but non-existent in others. Some of the pressures for advance come from mission-orientated studies for industry and some from the curiosity-motivated studies of academe. These pressures have led to an amazing plethora of techniques for surface analysis, each with its bureaucratic acronym, and to a very wide range of studies and data. Not all of these techniques and studies are relevant to researchers who wish to understand and solve surface-related problems in applied science. Over the years, Auger electron spectroscopy (AES) and X-ray photo-electron spectroscopy (XPS) have been found to show the greatest applicability. In this chapter, therefore, we chronicle the evolution of these techniques and show how they fit into the jigsaw of modern surface spectroscopies.

We have already mentioned surface analysis without defining what we mean. In its simplest sense we require the elemental composition of the outermost atom layer of a solid. Having found that, there will be immediate requests for detailed knowledge of the chemical binding state, precise sites of atoms in relation to crystal structure, surface homogeneity and the state of adsorbates. For many people surface science still concerns the complete

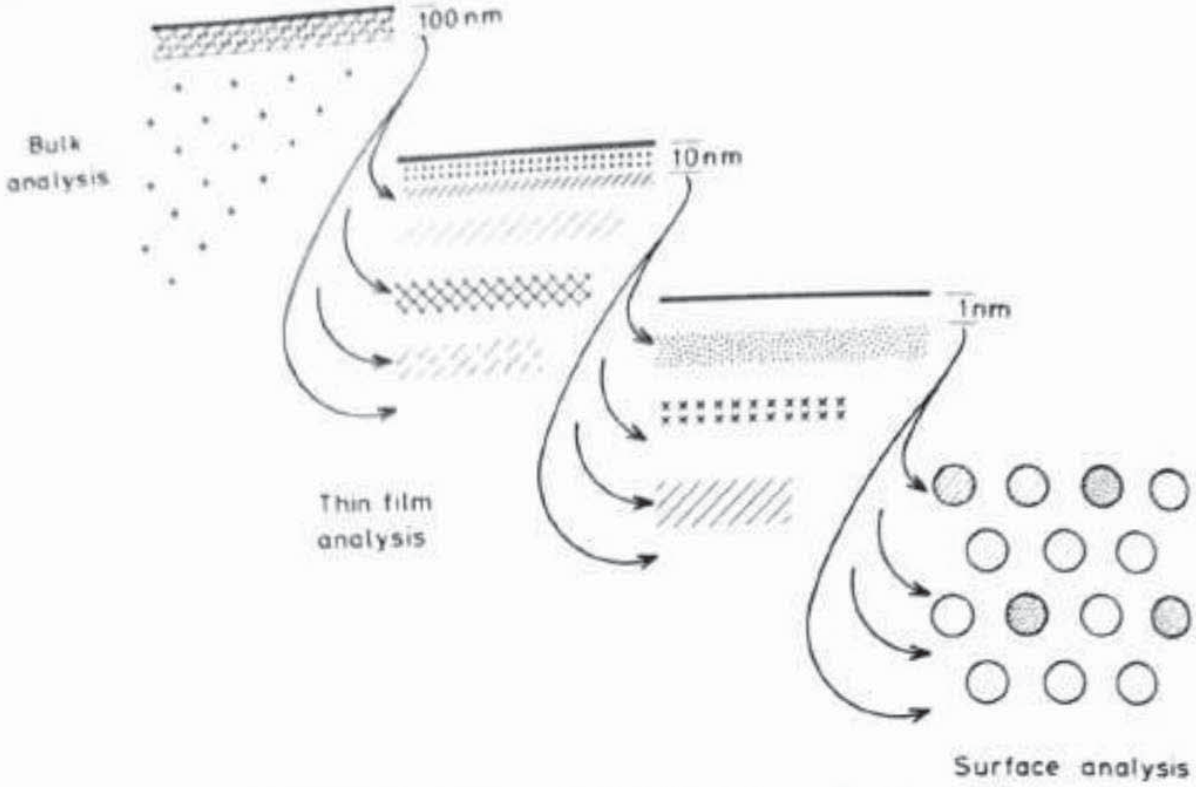


Figure 1.1 The regimes of surface analysis, thin film analysis and bulk analysis

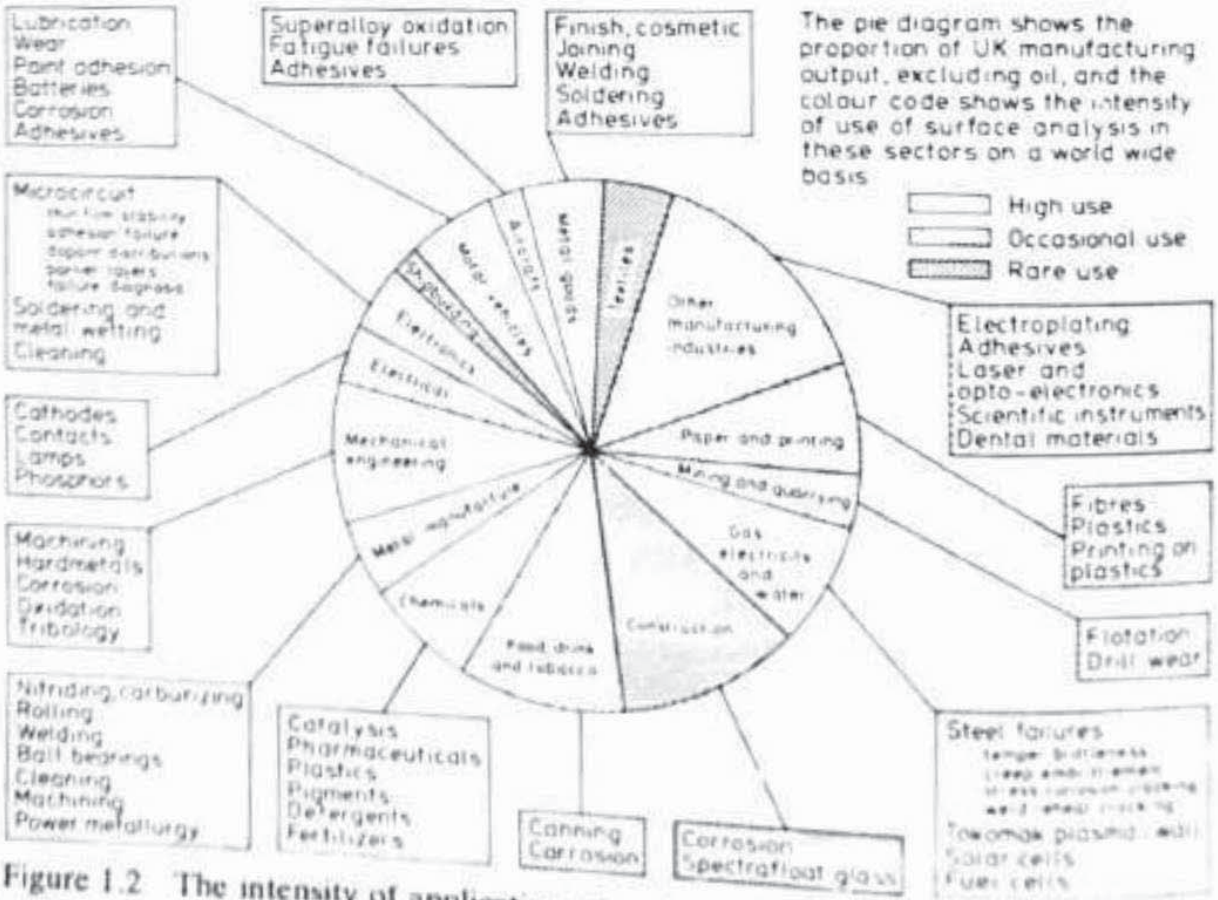


Figure 1.2 The intensity of application of surface analysis, illustrated by the manufacturing sectors of the United Kingdom. After Seah[1]

characterization of clean, low index, metal surfaces in vacuum; for others, the occasional adsorbed molecule of carbon monoxide may be added. However, in applied studies, the surfaces will be far more complex and the characterization will not generally be complete—simply adequate for the purpose. In applied studies all of the above will be required but, in addition, there will be requests for similar information for the atom layers below the surface as a function of depth, to depths of 1 μm or so, as shown in Figure 1.1. Each of the many surface analysis techniques approaches one or more of these aspects better than the others so that, in principle, each has a particular advantage. However, the value of a technique to the user depends not only on the theoretical advantage but also on the available experience in that technique, the back-up data and examples of similar approaches by previous workers. This books seeks to provide such information for the two principal surface analysis techniques, AES and XPS.

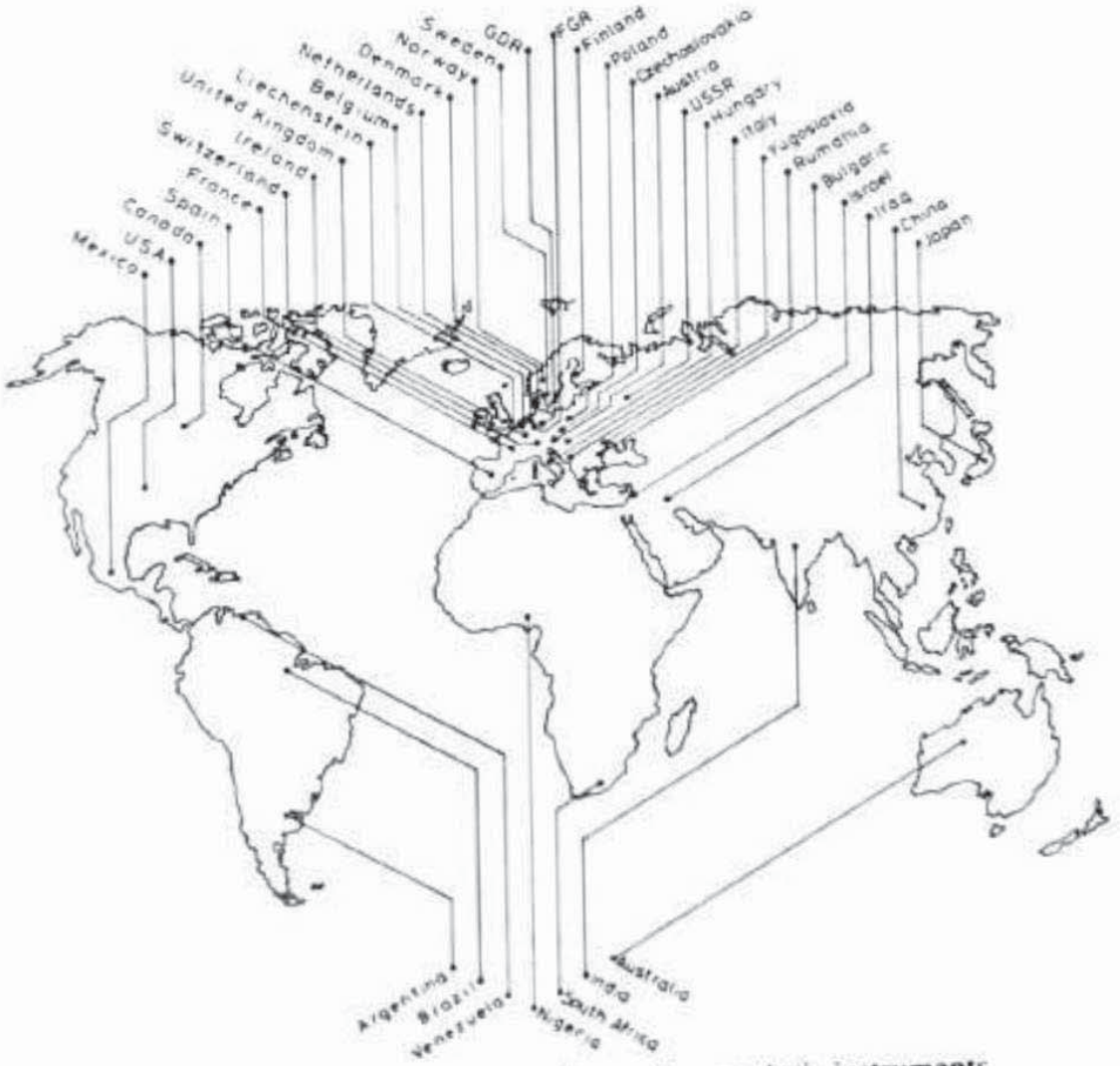


Figure 1.3 Countries operating surface analysis instruments

In the later chapters, examples are given of the use of AES and XPS in the major areas of microelectronics, metallury, catalysis, polymers and corrosion science. Figure 1.2 shows these and other areas in which surface analysis has made considerable contributions, ranging from the wear of cutting edges in strip metal production and the design of lubricant additives to opto-electronics and the architecture of integrated circuits. The ubiquity of surface analysis applications is matched by the range of countries now operating surface analysis instruments, as shown in Figure 1.3. This overall development has occurred in the years since 1970. Some developments took place before 1970 but most are concentrated into the last few years.

1.2 The Background to Electron Beam Techniques

1.2.1 Auger electron spectroscopy

The start of much of the work can be placed in the early 1920s when C. Davisson and H. E. Farnsworth were studying the secondary emission properties of metal surfaces in high vacuum under bombardment by electrons with energies of up to a few hundred electronvolts. Davisson's interests at that time partly concerned the properties of the secondary electron emission of anodes in gas-filled and vacuum triode valves. During his experiments with L. H. Germer, an accident occurred to the glass vacuum system when the nickel target was being outgassed at a high temperature. This led to oxidation of the target which was then cleaned by prolonged heating in hydrogen, causing the originally fine grained target to recrystallize to a few large grains. Low energy electron diffraction (LEED) was observed from these crystals by Davisson and Germer in 1927[2] and was correctly interpreted in terms of de Broglie's publication on wave mechanics. Davisson, Germer, Farnsworth[3] and others then worked on LEED through the 1930s using glass vacuum systems and their own designs of Faraday cup detectors to monitor the LEED beams. The experiments were very difficult but these workers still managed to study many low index single crystal metal surfaces, under what were probably ultra-high vacuum (UHV) conditions, as a function of temperature and gas adsorption. The work had a very limited appeal until Germer[4] revived the display LEED concept of Ehrenberg[5] and the concept was marketed commercially in a metal UHV system by Varian in 1964. Within a few years other manufacturers were producing commercial systems and many laboratories were set up to study single crystal metal surfaces, by LEED, in relation to a whole range of notionally applied problems that today, with the benefit of hindsight, may appear a little far-fetched. Many of the laboratories were, unfortunately, not in a position to handle the theoretical requirements of LEED.

In 1967 Harris published two reports from GE, subsequently appearing in the regular literature,[6,7] which followed the notion of Lander[8] that Auger electrons from solids could be used for surface analysis. Briefly, the principle

of the technique is that the sample is bombarded by an electron beam of 1–10 keV energy and this ejects core electrons from a level E_1 in atoms in a region of the sample up to 1 μm or so deep. The core hole is then filled by an internal process in the atom whereby an electron from a level E_2 falls into the core hole with the energy balance taken by a third electron from a level E_3. This last electron, called an Auger electron after Pierre Auger who first observed such events in a cloud chamber,[9] is then ejected from the atom with an energy E_A, given very approximately by†

$$E_A = E_1 + E_2 - E_3 \tag{1.1}$$

The electrons thus have energies unique to each atom and, if the energy spectrum from 0 to 2 KeV is measured, the energies of the Auger electron peaks allow all the elements present, except hydrogen and helium, to be identified. The reason that AES is a surface-sensitive technique lies in the intense inelastic scattering that occurs for electrons in this energy range, so that Auger electrons from only the outermost atom layers of a solid survive to be ejected and measured in the spectrum. Measurements of Auger electrons had been made much earlier, but not at these energies and not in relation to surface analysis.[10] Harris realized that the direct energy spectrum, with its small peaks on a large background, could be much easier to present and analyse if differentiated. He thus developed the potential modulation of the analyser to present the differential mode spectra used ever since. Harris also demonstrated the great sensitivity of the technique in measuring surface contaminants with a signal-to-noise ratio of over 200.

In 1967 Harris discussed his work with Peria who immediately realized that the standard LEED apparatus could be modified for AES work with the addition of a small amount of standard electronics. Weber and Peria's publication[11] appeared before those of Harris due to unfortunate refereeing problems.[12] Some of the early euphoria for LEED was fading and so standard LEED apparatuses were rapidly converted to AES work. Many early AES instruments were thus ideally suited for UHV work on single crystal surfaces but, because of the inherent differences of the LEED system as an electron spectrometer compared with the 127° deflection spectromer of Harris, the LEED system was generally used at poor energy resolution. The poor resolution led many workers to believe that Auger electron peaks were broad and that they contained little chemical information. Thus AES developed from the beginning as an elemental analysis technique but was well established for carefully prepared single crystal surfaces in UHV. This enabled Palmberg and Rhodin[13] to show, by depositing single atom layers on low index crystal faces, that AES is characteristic of the surface to a depth of only 5–10 Å. The next major step came with the introduction of the cylindrical mirror analyser by Palmberg, Bohm and Tracy[14] which significantly improved the signal-to-noise

†In Eq (1.1) the energies are given algebraically, ie bound states have negative energy. Often, for convenience, the energies are tabled with positive values. Thus in Eq (3.1), where that convention is used, the opposite signs occur.

ratio available with AES and which opened the door to the wide spectrum of applications we see today.

1.2.2 Other techniques

Concomitant with these developments have been the developments of a number of parallel electron beam analytical techniques which have their own particular following but which do not have such general appeal as AES. Techniques such as appearance potential spectroscopy (APS),[15] in which the onset of excitation of the core levels is measured as the beam energy is increased, were popular as cheaper alternatives to AES but have poor sensitivity and so are only used now for solid-state studies. Ionization loss spectroscopy (ILS),[16] in which the electron losses associated with the onset of ionization of a core level are measured, similarly has poor sensitivity but now is being revived as a technique appropriate for very localized chemical analysis in the scanning transmission electron microscope (STEM).[17] Other variants on these techniques exist but the only other analytical electron beam technique to have established itself is high resolution low energy electron loss spectroscopy (HRLEELS).[18,19] In this technique the electron loss events from a highly monochromatic electron beam are measured with an energy resolution in the 5–20 MeV range. These losses are associated with the excitation of the vibrational modes of atoms and molecules in specific sites and orientations on metal surfaces. Reasonable quantities of data are now established so that the technique is becoming more general in application; however, polar compounds cannot be studied as the spectra become swamped by losses due to the easy excitation of optical phonons. A final technique which analyses structure rather than chemistry, but which is finding strong application today in the growth of epitaxial layered materials, is reflection high energy electron diffraction (RHEED). Like LEED, RHEED can be used to analyse the crystal structure of the outermost layers of a solid. To achieve a similar surface sensitivity to the 10–200 eV electrons used in LEED, the 10–30 keV RHEED electron beam is arranged to strike the surface at grazing angles of incidence. The analysis of diffracted intensities here is more straightforward than is the case for LEED and partly explains its current popularity.

1.3 The Background to Photon Beam Techniques

1.3.1 X-ray photoelectron spectroscopy[20]

XPS has a very illustrious pedigree in which the Auger effect plays a part. Its long history is intricately bound up with the developments of wave particle duality and the early days of atomic physics. For the more discerning reader, with a little more time to invest, this history is excellently and vividly

chronicled by the workers at LaTrobe University[20-23] whose articles contain interesting historical perspectives and anecdotes.

XPS has its origins in the investigations of the photo-electric effect (discovered by Hertz in 1887) in which X-rays were used as the exciting photon source. For example, Innes described, in 1907,[24] experiments involving a Rontgen tube with a platinum anti-cathode and subsequent magnetic analysis of the emitted electrons by velocity selection using two Helmholz coils and photographic detection.

Rutherford's laboratory, in Manchester, was at this time ideally placed to develop the field, having much experience in the measurement of the β-ray spectra of radioactive materials by magnetic analysis and being at the forefront of the new field of X-ray spectroscopy. The early experiments were carried out by Moseley, Rawlinson and Robinson before the outbreak of World War I. Indeed, in 1914 Rutherford made a first stab[25] at stating the basic equation of XPS, which was subsequently modified to

$$E_K = h\nu - E_B$$

where E_K is the kinetic energy of the β-rays (photo-electrons), $h\nu$ the incident photon energy and E_B the electron binding energy. After the War, Robinson continued the work and Maurice de Broglie started research in France, both still using photographic detection. Understanding developed rapidly and in the early 1920s the photo-electron spectra of many elements, excited by a variety of high energy X-rays, had been obtained. Anomalous 'lines' (not obviously from core levels) were found to correspond with excitations by X-ray fluorescence photons characteristic of the atoms under study. The realization of the true nature of these electrons came only with Auger's demonstration of the radiationless transition,[9] which bears his name, in cloud chamber experiments. At this time XPS was seen as extending the study of atomic structure beyond the confines of X-ray spectroscopy (XES). However, it eventually became obvious that the limitations of the experimental XPS technique on the one hand and new advances in XES technique on the other had changed the situation heavily in favour of XES. Only Robinson continued research, though without any major advance in resolution or sensitivity, through the 1930s, and this effort too ended with the outbreak of World War II.

After the War, Steinhardt and Serfass at Lehigh University conceived the idea of reviving XPS as an analytical tool, particularly for studying surface chemical phenomena. Several instruments were constructed which, though failing to advance performance, introduced G-M electron counting and eventually a 127° electrostatic energy analyser. Despite the use of very high energy X-rays, surface effects on spectra were, however, discerned.[26] Steinhardt's PhD thesis had the prophetic title 'An X-ray photoelectron spectrometer for chemical analysis'.

Meanwhile decisive developments were underway in Uppsala, Sweden. Kai Siegbahn had developed, during the 1940s, β-ray spectroscopy to very high levels of precision and had realised that electron spectroscopy using X-ray excitation, rather than radioactive sources, might now be possible. An iron-free double-focusing spectrometer for accurate magnetic energy analysis was specially developed with a resolving power of 10^{-5}. In 1954 the first X-ray photo-electron spectrum from cleaved sodium chloride was obtained. Up to this point all previous spectra had consisted of a series of bands with a more-or-less well-defined edge followed by a tail. The Uppsala spectrum revealed for the first time a completely resolved line on the high kinetic energy side of each 'edge', representing electrons which had lost zero energy. The peak maximum (E_K) allowed the electron binding energy (E_B) to be measured accurately and the goal of using XPS for atomic structure investigation had been realized. Subsequently Siegbahn's group observed the chemical shift effect on core-level binding energies and went on to develop the whole field of electron spectroscopy during the period 1955–1970. Siegbahn coined the acronym ESCA (electron spectroscopy for chemical analysis) to underline the fact that both photo *and* Auger electron peaks appear in the 'XPS' spectrum.[27] A seminal book[28] on the subject was published with this title in 1967 which brought large-scale awareness of the potential of XPS. Commercial instruments started to appear around 1969–1970, although at this time, as a result of work with stearates by Siegbahn *et al.*,[28] it was thought that XPS characterized the outermost 100 Å of the surface. It was not until the work of Brundle and Roberts[29] in UHV that XPS truly became a surface technique whilst in 1972 Seah[30] pointed out that the thickness of the surface layer characterized must be very similar to that in AES.

1.3.2 Other techniques

Instead of using photons of kiloelectronvolt energy, much lower energy photons may be used to excite electrons in the solid. In ultra-violet photo-electron spectroscopy (UPS) the source of photons is a differentially pumped inert gas discharge 'lamp'. This produces discrete low energy resonance lines (e.g. He^{I}, 21.2 eV, and He^{II}, 40.8 eV) with an inherent width of a few millielectronvolts. Since there is only sufficient energy to emit electrons from the valence band, the technique has no analytical potential (in the sense of this book). UPS is, however, widely used in the study of the electron band structure of metals, alloys and semi-conductors and of adsorption phenomena.[31] A further source of photons for photo-electron spectroscopy is synchrotron radiation—the radiation emitted by accelerating electrons. Electrons circulated continuously in a storage ring at energies of ~1 GeV produce a continuous spectrum of photons with energies from a few electronvolts to several kiloelectronvolts. With a monochromator, therefore, a variable energy

photon source is provided and this has several attractions for photo-electron spectroscopy.[32,33] The synchrotron source is also ideal for carrying out extended X-ray absorption fine-structure spectroscopy (EXAFS). Here, information about local structure and short-range order is obtained from the pattern of oscillations on the high energy side of X-ray absorption edges. Essentially this is a bulk analytical technique to complement X-ray diffraction (XRD) but applicable to amorphous materials. However, 'surface' information is obtained, for instance, in the study of very small metal crystallites in supported metal catalysts.[34] True surface information can also be obtained by a variety of methods, e.g. by monitoring the Auger electron yield over the same photon energy range in the region of an absorption edge. Such variants are referred to as surface EXAFS or S-EXAFS.[35]

1.4 The Background to Ion Beam Techniques

Over a similar period to that of AES has been the development of two major ion beam techniques: secondary ion mass spectroscopy (SIMS),[36,37] and ion scattering spectroscopy (ISS).[38,39] In SIMS a mass filtered beam of argon, oxygen or cesium ions strikes the surface and removes surface atoms. Some of the emitted atoms are either positively or negatively ionized and these may be detected, very sensitively, by a mass spectrometer. For surface analysis a primary beam of argon, at current densities of the order of nanoamperes per square centimetre, is used with the beam spread out over some tens of square millimetres. In this mode, called static SIMS, the technique has a very high surface sensitivity, especially to electropositive elements, and can also give, in principle, detailed chemical state information and, with some effort, moderate quantification. Although now twelve years old, the technique has not yet found as much favour as AES and XPS. It is quite likely, however, that interest in the use of static SIMS will grow in the near future for samples already characterized by AES or XPS. The latter techniques rapidly provide an overall picture but SIMS is sometimes better at unravelling detailed aspects of chemistry, has a much higher sensitivity and furthermore can be used to observe hydrogen. In the alternative operating mode of dynamic SIMS the ion optics allow submicron images of the sample to be produced for different elements. The high primary ion flux densities of dynamic SIMS cause the surface to be eroded, as measurement occurs, in contrast to static SIMS where the top atom layer may remain for many hours before removal. In dynamic SIMS the alternative ion sources of oxygen and cesium are used as these enhance the yields for positive and negative ions, respectively, to increase the detectability. Here limits are given in parts per million (p.p.m.) or billion (US) instead of fractions of a percent as in AES and XPS. In dynamic SIMS the chemical information is lost and the instrument can be operated in one of two ways. In one mode the instrument behaves similarly to

the electron probe X-ray microanalyser (EPMA) for bulk microanalysis and imaging or, in the other mode, by rastering the primary beam over an area of approximately 50 × 50 μm, high depth resolution concentration–depth profiles of dopants in semi-conductors can be attained.

The second ion beam technique, ISS, relies on a simple billiard ball collision concept in which the primary ion is scattered into the spectrometer with an energy defined by the mass of the struck target atom. The technique does not have the broad applicability of AES, XPS and SIMS but has the unique advantages that it gives information of the outermost atom layer only and also, by varying the polar and azimuthal angles of the interrogating ion beam with respect to a single crystal target, of the adatom site position.

Two other ion beam techniques should be mentioned, sputtered neutral mass spectroscopy (SNMS)[40] and Rutherford back-scattering spectroscopy (RBS).[41] SNMS uses the same principles as SIMS but, recognizing that most of the signal is lost because most of the emitted particles are neutralized on leaving the target surface, seeks to re-ionize them in a plasma just above the surface. This should greatly increase the sensitivity over that of SIMS but, due to the difficulty of operating the plasma and of obtaining spatial resolution, the technique has not yet gained popularity. On the other hand, RBS is more akin to ISS, using the billiard ball collisions and energy analysis of the reflected beam ions. By using ions of a few MeV energy the depth distributions of different elements can be measured non-destructively over depths up to 1 μm with a depth resolution of 20 nm. Heavy elements of adjacent atomic number are difficult to resolve but sensitivities below 1 per cent. of a monolayer are claimed. The equipment is simple and rugged and suitable Van der Graaff accelerators are becoming available as surplus to requirements from the nuclear measurements field, where higher energy machines are being purchased. Ion neutralization spectroscopy (INS),[42] in which low energy ions neutralized at the surface produce electrons which are characteristic of the valance band, is not discussed here since, like UPS, it has not general analytical capability.

1.5 Conclusions

The broad scope of some of the surface analysis techniques* is shown in Table 1.1, where some techniques are summarized in terms of their input and measured radiations. Fuller versions of such a table and a more detailed discussion of many of the techniques may be found in the review of Coburn and Kay,[43] articles in *Methods of Surface Analysis*[44] and the review of Honig.[45] Table 1.1reminds us that for each input radiation there are many simultaneous pro-

*Little mention has been made of vibrational spectroscopy techniques which can provide surface analytical information, since they do not meet a basic requirement of identifying *elements*. However, in the study of adsorption of molecules or of organic and polymeric surfaces they have an important role to play. Some of these techniques are mentioned in Chapter 9.

Table 1.1 Easy reminder of the principles of surface techniques

		Incident Radiation			
		Electrons	Ions	Neutrals	X-rays
Radiation detected	Electrons	AES LEED HREELS	INS		UPS XPS
	Ions	ESD	SIMS ISS RBS	FABMS	
	Neutrals		SNMS		
	X-rays	(EPMA)	(PIXE)		S-EXAFS

cesses occurring in the sample. Thus AES instruments with high spatial resolution can usefully have added X-ray detectors, as in the electron probe X-ray microanalyser (EPMA), but electron stimulated desorption (ESD)[46] must also be recognized. Since a measurement in AES requires a certain dose of electrons and the fraction of surface atoms removed in ESD concerns the dose per unit area, it is clear that ESD sets limits on spatial resolution in AES.[47] Some extra acronyms are given in Table 1.1, where those in brackets are not true surface analysis techniques but are added for completion. Particle-induced X-ray emission (PIXE)[48] uses MeV α-particles, or protons, to irradiate the sample and, by analysing the emitted X-rays gives p.p.m. detectability for elements with $Z > 12$ over depths comparable with the EPMA. Fast atom bombardment mass spectrometry (FABMS)[49] replaces the ion beam of SIMS with energetic neutrals and so overcomes any charging effects of insulators.

The more important of these techniques are presented in some detail in Table 1.2 where some of the less easily quantified aspects are presented. Cost is not included since the bulk of the cost is often not with the optics of the particular technique but with the specimen handling arrangements, computer analysis, laboratory overheads and social security payments. Ease of specimen preparation and an understanding of the measurement are generally more important. Ease of specimen preparation is the main limitation in the application of the atom probe field ion microscope[50] although, for microanalysis at interfaces, it is doubtful if alternative techniques can approach within an order of magnitude of its spatial resolution. Sharp tips of material must, of course, be prepared; however, the range of possible materials is continually increasing.

Before concluding this chapter it should be noted that the surface analysis instrumentation field is moving very rapidly and within a year Table 1.2 is likely to be out of date in relation to spatial resolution and sensitivity. AES and XPS will continue to have the greatest popularity since they are both

Table 1.2 Survey of the more popular techniques for surface and interface analysis

Technique	Information (E = elemental, C = chemical)	Spatial resolution	Sampling depth monolayers	Sensitivity (order of)	Quantification (✓ = easy)	Elements not covered	Popularity	Specimen preparation (✓ = easy)	Ease of use	Extent of support data	Effective take-off year
AES	E*	0.2 μm	3	0.3 %	✓	H,He	****	✓	****	*****	1968
Atom probe FIM	E†	1 nm	1	1 %	✓		*		*	**	1968
HREELS	C	1 mm	1	1 %			*		**	**	1970
ISS	E	1 mm	1	1 %	✓	H,He	*	✓	***	***	1967
RBS	E	1 mm	100	1 %	✓	H,He	*	✓	****	****	1967
SIMS (static)	C	50 μm	2	0.1 %			**	✓	**	**	1970
SIMS (dynamic, imaging)	E	0.5 μm	40	1 p.p.m.	✓‡		**	✓	***	****	1968
SIMS (dynamic, depth prof)	E	50 μm	40	1 p.p.m.	✓‡		**	✓	***	***	1975
UPS	C	3 mm§	3	1 %		N/A	**	✓	***	****	1969
XPS	C,E	0.2 mm§	3	0.3 %	✓	H,He	****	✓	****	*****	1967

*C is available, but not with high spatial resolution due to electron stimulated desorption effects.[45]
†C may generally be deduced.
‡When compared with a close standard.
§1 μm resolution has been reported in a new instrument concept using an He UPS source and resolutions below 100 μm are calculated for XPS energies.[51]

self-contained techniques which are available for any laboratory and for which there has consequently developed an enormous store of expertise and data. Both are limited in the sense that XPS has poor spatial resolution and AES tends to be rather destructive to weakly bound species. Count rates are likely to increase as spectrometer design improves and as multi-channel detectors become more popular. This will allow higher spatial resolution in XPS and reduce the damage associated with AES. Spatial resolutions of 100 μm should be possible in XPS using this approach although something below this figure may be possible using the axial magnetic field photo-electron spectromicroscope of Beamson, Porter and Turner.[51] This latter instrument is of a revolutionary rather than evolutionary nature and makes use of trapped electron orbits in the eight tesla field of a superconducting magnet.

The histories of many of our surface analysis techniques are measured in decades and yet we see from Table 1.2 that the effective take-off date for many of the techniques, when they mushroomed into many laboratories, is bracketed by the years 1968 to 1970. This is not a coincidence but most probably occurred through the maturing of UHV technology and the ready availability of complex UHV systems in the middle and late 1960s, so catalysing a general awareness of surfaces and their importance. From those days, with the careful handling of special vacuum systems by people who 'knew' about UHV, we have come today to systems of greater reliability, where samples may be inserted and removed through airlocks in minutes rather than days, and which may be operated by people no longer skilled in solving vacuum problems but skilled in solving surface problems.

In the chapters that follow the reader will find, first, the details of instrumentation, followed by the interpretation of spectra generated from solid surfaces by both AES and XPS. A major proportion of the applied work currently involves not just the outer surface but also the composition through reaction layers or coatings. The basic problems concerned with such depth profiling are discussed together with the status of the quantification in the spectra. This part of the book, together with the tabulations of data and technique in the voluminous appendices, provides the reader with all that should be needed to get the most from the techniques. The second part of the book, Chapters 6 to 10, illustrates the power of AES and XPS in the solution of problems in the major technological areas, as shown in Figure 1.2. These chapters show how the different aspects of the techniques, discussed in the earlier part of the book, may be turned to particular advantage in the solution of important problems.

References

1. M. P. Seah, *Surface and Interface Anal.*, **2**, 222 (1980).
2. C. Davisson and L. H. Germer, *Phys. Rev.*, **30**, 705 (1927).

3. H. E. Farnsworth, *Phys. Rev.*, **33**, 1068 (1929).
4. E. J. Scheibnet, L. H. Germer and C. D. Hartman, *Rev. Sci. Ins.*, **31**, 112 (1960).
5. W. Ehrenberg, *Phil. Mag.*, **18**, 878 (1934).
6. L. A. Harris, *J. Appl. Phys.*, **39**, 1419 (1968).
7. L. A. Harris, *J. Appl. Phys.*, **39**, 1428 (1968).
8. J. J. Lander, *Phys. Rev.*, **91**, 1382 (1953).
9. P. Auger, *J. Phys. Radium*, **6**, 205 (1925).
10. H. R. Robinson and A. M. Cassie, *Proc. Roy. Soc.*, **A113**, 282 (1926), and many later papers.
11. R. E. Weber and W. T. Peria, *J. Appl. Phys.*, **38**, 4355 (1967).
12. L. A. Harris, *J. Vac. Sci. Technol.*, **11**, 23 (1974).
13. P. W. Palmberg and T. N. Rhodin, *J. Appl. Phys.*, **39**, 2425 (1968).
14. P. W. Palmberg, G. K. Bohm and J. C. Tracy, *Appl. Phys. Lett.*, **15**, 254 (1969).
15. C. Webb and P. M. Williams, *Surface Sci.*, **53**, 110 (1975).
16. R. L. Gerlach, *J. Vac. Sci. Technol.*, **8**, 599 (1971).
17. R. D. Leapman, S. J. Sanderson and M. J. Whelan, *Metal Sci.*, **12**, 215 (1978).
18. H. Ibach and S. Lehwald, *J. Vac. Sci. Technol.*, **15**, 407 (1978).
19. H. Ibach and S. Lehwald, *Surface Sci.*, **76**, 1 (1978).
20. J. G. Jenkin, R. C. G. Leckey and J. Liesegang, *J. Electron Spectrosc.*, **12**, 1 (1977).
21. J. G. Jenkin, J. D. Riley, J. Liesegang and R. C. G. Leckey, *J. Electron Spectrosc.*, **14**, 477 (1978).
22. J. G. Jenkin, J. Liesegang, R. C. G. Leckey and J. D. Riley, *J. Electron Spectrosc.*, **15**, 307 (1979).
23. J. G. Jenkin, *J. Electron Spectrosc.*, **23**, 187 (1981).
24. P. D. Innes, *Proc. Roy. Soc.*, **A79**, 442 (1907).
25. E. Rutherford, *Phil. Mag.*, **28**, 305, (1914).
26. R. G. Steinhardt, *Anal. Chem.*, **23**, 1585 (1951).
27. K. Siegbahn, personal communication.
28. K. Siegbahn, C. N. Nordling, A. Fahlman, R. Nordberg, K. Hamrin, J. Hedman, G. Johansson, T. Bermark, S. E. Karlsson, I. Lindgren and B. Lindberg, *ESCA: Atomic, Molecular and Solid State Structure Studied by Means of Electron Spectroscopy*, Almqvist and Wiksells, Uppsala (1967).
29. C. R. Brundle and M. W. Roberts, *Proc. Roy. Soc.*, **A331**, 383 (1972).
30. M. P. Seah, *Surf. Sci.*, **32**, 703 (1972).
31. P. M. Williams, Chapter 9 in *Handbook of X-ray and Ultraviolet Photoelectron Spectroscopy* (Ed. D. Briggs), p. 313, Heyden and Son, London (1977).
32. W. E. Spicer, in *Electron and Ion Spectroscopy of Solids* (Eds L. Fiermans, J. Vennick and W. Dekeyser), p. 34, Plenum, New York (1978).
33. T. A. Carlson, *Surface and Interface Anal.*, **4**, 125 (1982).
34. D. C. Komingsberger and R. Prins, *Trends. Anal. Chem.*, **1**, 16 (1981).
35. A. Bianconi, *Appn. Surf. Sci.*, **6**, 392 (1980).
36. A. Benninghoven, *Z. Physik*, **230**, 403 (1970).
37. H. W. Werner, *Surface and Interface Anal.*, **2**, 56 (1980).
38. W. L. Baun, *Surface and Interface Anal.*, **3**, 243 (1981).
39. E. N. Haeussler, *Surface and Interface Anal.*, **2**, 134 (1980).
40. H. Oechsner and E. Stumpe, *Appl. Phys.*, **14**, 43 (1977).
41. I. V. Mitchell, *Phys. Bull.*, **30**, 23 (1979).
42. H. D. Hagstrum, in *Electron and Ion Spectroscopy of Solids* (Eds L. Fiermans, J. Vennick and W. Dekeyser), p. 273, Plenum, New York (1978).
43. J. W. Coburn and E. Kay, *CRC Critical Revs.*, in *Solid State Sciences*, **4**, 561 (1974).

44. A. W. Czanderna (Ed.), *Methods of Surface Analysis*, Elsevier, New York (1975).
45. R. E. Honig, *Thin Solid Films*, **31**, 89 (1975).
46. C. G. Pantano and T. E. Madey, *Appn. Surface Sci.*, **7**, 115 (1981).
47. M. P. Seah, in *Surface Analysis of High Temperature Materials—Chemistry and Topography* (Ed. G. Kemeny), Applied Science (to be published).
48. J. L. Campbell, B. H. Orr, A. W. Herman, L. A. McNelles, J. A. Thomson and W. B. Cook, *Anal. Chem.*, **47**, 1542 (1975).
49. D. J. Surman, J. A. Van der Berg and J. C. Vickerman, *Surf. Interface. Anal.*, **4**, 160 (1982).
50. M. K. Miller and G. D. W. Smith, *Metal Sci.*, **11**, 249 (1977).
51. G. Beamson, H. Q. Porter and D. W. Turner, *Nature*, **290**, 556 (1981).

Practical Surface Analysis
by Auger and X-ray Photoelectron Spectroscopy
Edited by D. Briggs and M. P. Seah

Chapter 2

Instrumentation

J. C. Rivière
UK Atomic Energy Research Establishment, Harwell, Didcot, Oxfordshire

2.1 Introduction

Without the means of measuring and recording electron energy distributions, electron spectroscopy would be merely speculation. The spectrometer is central to the whole subject, and it is therefore worth discussing spectrometers in some detail because of their importance. In this context the term 'spectrometer' is taken to include all aspects of an instrument that have any bearing on the process of the collection of electron spectroscopic data. Thus not only must the design and performance of the electron energy analyser itself be discussed, but also the design of electron and of X-ray photon sources, the nature of the ambient required in a spectrometer and how it is to be achieved, the ways in which a specimen for analysis is handled, i.e. transferred or manipulated, within the spectrometer and the methods of specimen surface preparation and treatment. It is the purpose of this chapter to describe current state of the art and practice in modern spectrometers used for XPS and AES in a way that will allow their operation and principles to be understood, and that will also provide some useful working relationships.

2.2 Vacuum Conditions

There are two reasons why electron spectrometers used in surface analysis must operate under vacuum. In the first place, electrons emitted from a specimen should meet as few gas molecules as possible on their way to the analyser so that they are not scattered and thereby lost from the analysis. Another way of expressing this requirement is that their mean free paths should be much greater than the dimensions of the spectrometer. This by itself does not impose much stringency on the working vacuum to be achieved, since vacua in the range 10^{-5}–10^{-6} torr would be adequate. However, the second reason *does* impose stringent requirements on the working vacuum. It has been established[1] in Chapter 5 (see Figure 5.5, in which the

observed relationship between electron kinetic energy and inelastic mean free path is plotted) that both XPS and AES are highly surface-specific techniques, with sampling depths typically of a few atom layers. Most of the electrons whose energies are analysed originate in fact in the first one or two atom layers. Such surface specificity, combined with elemental sensitivities of the order of 0.5 per cent of an atom layer, means that the techniques are very sensitive to surface contamination, from whatever source. Since in many experiments it is necessary to start with a well-characterized surface, either atomically clean or in some other stable condition, and since even very small amounts of contaminant can affect the course of an experiment drastically, it is clearly necessary to operate under conditions in which the rate of accumulation of contamination is negligible compared to the rate of change in the experiment. The principal source of contamination is from the residual gas in the vacuum system. Any textbook[2] on gas kinetic theory will provide the information that, to a good approximation, a monolayer of gas will accumulate on a surface in 1.5 s at a pressure of 10^{-6} torr and at room temperature, if every molecule hitting the surface stays there on impact (a sticking probability of unity). If one takes as typical requirements that no more than 0.05 atom layers of contaminant should accumulate in the space of 30 min, then, again for unity sticking probability, gas kinetics rule that the residual gas pressure should be 4×10^{-11} torr. In practice, sticking probabilities are not often quite as high as unity, and for the great majority of surface experiments it is found that a base pressure of 10^{-10} torr is adequate. Such a pressure falls into the regime of ultra-high vacuum (UHV).

2.2.1 UHV systems

2.2.1.1 Materials

Having established that XPS and AES should be operated at pressures in the UHV region, it is necessary to discuss briefly what implications that has for the materials of construction of spectrometers. Returning to gas kinetics again, but without going into detail, it is an easy matter to show that for practicable dimensions of chambers, pipework and pump, the ultimate vacuum attainable is always governed by the rate at which adsorbed gases leave the internal surfaces at ambient temperature. This outgassing rate, as it is called, varies considerably according to the chemical nature of the surface and its pre-treatment or history and according to the nature of the adsorbed gas. However, even in the best case, i.e. the lowest outgassing rate, if the temperature of the internal surfaces never goes above ambient, then UHV in the spectrometer will not be achieved, or at least not for a very long time indeed. In order to reach UHV in a reasonable time it is necessary to remove the adsorbed gases at a much faster rate than normal, and this is accomplished by

raising the temperature of the spectrometer, or baking it, to about 200 °C for a few hours. On cooling to ambient temperature again, the outgassing rate from the internal surfaces drops by several orders of magnitude and the pumps are then capable of producing and maintaining UHV.

The necessity of baking imposes restrictions on the materials that can be used in a spectrometer. Nothing must be used in construction that might disintegrate, outgas excessively or lose strength at the elevated temperature, and this applies throughout the life of the spectrometer, since bake-out is repeated at regular intervals. For these reasons it is not permissible, for example, to use elastomer materials such as Viton in seals between UHV and atmosphere since not only can elastomers oxidize and become brittle but at the baking temperature they also become porous and allow gases to pass through by diffusion. Thus the seals used in the joints between the components of the spectrometer must be of metal. The bake-out temperature precludes the use of indium (MP 156 °C), which is otherwise a useful sealing material, and in fact the two metals most commonly used for this purpose are gold and copper, with the emphasis heavily on the latter. Gold is used in the form of annealed wire compressed between flat flanges, but such seals, although completely reliable in use, suffer from the defects that it is difficult to position the wire correctly on a vertical flange, that the gold cannot always be removed from a flange after compression and that gold is expensive. The type of seal now employed almost universally is shown in Figure 2.1. A flat, annular, copper gasket fits exactly into the recess formed between two identical flanges when they are brought together. Each flange has a circular tapered knife-edge machined in its face. When the flanges are forced together by tightening the nuts on bolts placed through the bolt holes, the opposing knife-edges, which are of identical diameter, bite into the copper gasket and force the copper to flow away from them on each side. On the inside diameter there is no restriction, but on the outside the gasket is prevented from distorting by the walls of the recess into which it fits. Thus very large forces build up in the copper in that region and a very effective seal is formed between the gasket and the flange material on the outside of the knife-edge itself, capable of withstanding innumerable temperature cycles during bake-out.

The point of bringing in the above discussion about vacuum seals here is that it will be obvious by now that flanges of the type shown in Figure 2.1 (called, variously, 'knife-edge', 'copper-seal' or 'Conflat*' flanges) must be fabricated from a metal that is hard, retains its hardness indefinitely at baking temperatures and does not oxidize readily. Thus aluminium or its alloys, which are otherwise useful materials in UHV, cannot be used for UHV vacuum vessel construction since they are too soft for knife-edge flanges. As a result, the material used in the fabrication of the great majority of UHV

*Trade name of Varian Associates Ltd.

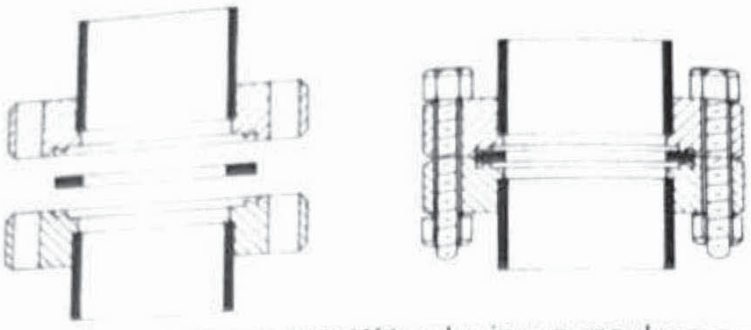


Figure 2.1 Knife-edge metal-to-metal UHV seal using an annular copper gasket. The left-hand diagram shows two identical flanges with circular knife-edges, and the gasket between them, before the seal is made. On the right the flanges have been clamped together and the gasket has been compressed against the inner walls of the recesses on the flanges by the action of the opposing knife-edges. The seal is made on the outer (sloping) surface of the knife-edge. This type of seal is also known as the Conflat® seal

vacuum vessels is stainless steel. The only exception occurs where magnetic screening is required, such as around energy analysers and in the vicinity of the specimen, in which case some manufacturers use mu-metal instead of stainless steel for the material of the vessel wall.

Although, as pointed out, the elastomer Viton should not be used between UHV and atmosphere, it is safe to use entirely within UHV if certain precautions are taken. These are that Viton rings should be de-greased and then pre-outgassed in an auxiliary vacuum system before insertion in the spectrometer and that those rings used in gate valve seats should never undergo bake-out with the valve in the closed position. Compressed Viton will take up a permanent set if baked repeatedly in that situation, leading to failure of the valve to close properly.

In addition to stainless steel and copper, there are several other materials that can safely be used between UHV and atmosphere, not as the principal constructional material but in components attached to the spectrometer for a variety of reasons. Before the advent of stainless steel vacuum systems, borosilicate glass was the constructional medium, and it is still used extensively in the form of windows mounted on flanges via graded glass-to-metal fused seals. For certain experiments quartz windows are also available. The other ceramic used widely is of course alumina; it is used in all places where electrical isolation is necessary to take current or voltage connections into the spectrometer.

Inside the spectrometer any material is permissible that does not cause an increase in the background pressure at either ambient or baking temperatures and that does not contain any constituents volatile at those temperatures that might settle on the inner surfaces and contaminate them. For instance, brass which contains a high proportion of volatile zinc must not be used, nor must any of the common plastic polymeric materials. However, it is permissible to

use PTFE-covered wire for instrumental connections in UHV if good quality material is used. The commonly used metals for small-scale fabrication in UHV, e.g. for special sample mounts, sample heating, electrical connections, etc., are copper, nickel, platinum, molybdenum, tantalum and tungsten, and most workers in electron spectroscopy would reckon to have a stock of at least some of those metals in the forms of wire and foil kept near the spectrometer.

2.2.1.2 Design and construction

These two aspects are interdependent; the design of any component, large or small, for UHV cannot be carried out without knowing exactly what the constructional constraints and requirements are. Constraints that have to be borne in mind continually are, for example, the relationship of flange positions to each other and to the external surface of a vacuum vessel, so that there is always access provided for the insertion and tightening of bolts and the allowance of adequate space both for weld preparation and for the welding operation itself. Normally such matters concern the manufacturer rather than the scientist, but if one wishes to design a special vacuum vessel oneself, or modify an existing one, then they should not be overlooked.

Another important part of vacuum system design, which applies not just to UHV, is the provision of adequate pumping speed at *all* points in the vessels to be evacuated. It is unfortunately still a not uncommon fault, even amongst reputable manufacturers who should know better, to design a perfectly good vacuum system for a spectrometer, good that is when the vessels are empty or carry no attachments, and then to place inside or attach to the vessels devices of such sizes and shapes that clearances are reduced to the point where the pumping speeds around the devices are severely restricted. This is the principal reason why the achievement of UHV by bake-out can take so long in some spectrometers. The relative arrangement of components inside the vacuum should be considered at the design stage not only from the electron spectroscopic point of view but also from the vacuum standpoint as well, and that applies with equal force to attachments that have to be accommodated on side ports. If necessary, to improve the pumping speeds in the latter case, additional by-pass pumping lines to the main pumps should be provided, or even auxiliary pumps. The estimation of conductances, and therefore of local pumping speeds, to adequate accuracy can be made from the various well-known formulae based on gas kinetics in the pressure region of molecular flow. These can be found in any good textbook[2] on vacuum physics and will not be repeated here.

Fabrication of vessels and other components in stainless steel involves inert gas arc welding, a technique which is now standard. For UHV, however, there are certain practices in such welding which it is essential to follow otherwise

the achievement of UHV will be difficult or impossible. These practices are aimed at avoiding trapped volumes with low-conductance paths to the vacuum side of the vessel wall; clearly such volumes could contain dirt or gases that would give rise to a constant source of contamination, i.e. act as 'virtual leaks'. To avoid their formation it is essential to specify that all joint welds must be internal *and* continuous. That is, enough filler rod must be on hand from the start to ensure that once the welding of, for example, a circular joint has begun it is continued without a break until the starting point is reached again. Only in this way can trapped volumes be avoided. Ideally, all such welding would be done by an electron beam in vacuum, but the latter method is slow and not suited to complex geometries such as are frequently encountered on the inside of UHV vessels.

The last operation to be carried out in the construction procedure is that of finishing and then cleaning the internal surfaces, and the effectiveness of this operation will govern the subsequent vacuum performance of the vessel or component. Nowadays the preferred finishing treatment is blasting with tiny glass beads, which not only removes scale, etc., from welds and surrounding areas but passivates the surface by introducing cold work into the surface region. in addition the topography of the surface is smoothed on a microscopic scale, so that potential traps for contamination are removed. Finishing by electropolishing, which was used widely not so long ago, is to be avoided at all costs, since the surface is physically removed to produce the polish and the weld regions are particularly susceptible to attack, allowing crevices and small holes to be uncovered. After glass-blasting it is essential to make sure that all the glass ball debris is removed, since if any remains it will eventually find its way into moving parts, with disastrous results.

De-greasing in clean pure solvent after finishing will produce a surface condition that would be usable without further treatment for high vacuum applications, but for UHV a final step is necessary, and that is pre-bake-out. By that is meant the baking of individual components on an auxiliary vacuum system, ideally to temperatures above those to be used during bake-out of the assembled spectrometer, in order to remove traces of the de-greasing solvent left on the internal surface. Following this final step, the components should be wrapped in clean plastic sheet or put into plastic bags, to remain clean until required for assembly. Needless to say, at no stage following glass-blasting is any component handled with bare hands; use of gloves is essential. In addition it is highly desirable that the pre-bake-out operation, the storage before use, and indeed the assembly, be carried out in dust-free enclosures, i.e. with filtered atmospheres.

The foregoing points of design and construction would not normally be ones that would be used by the individual research worker, since there are now very few laboratories that make their own vacuum equipment. Nor is it possible in the space available here to go into all the many other minor details

of instrument design that have to be taken into account. However, in assessing the comparative standards of fabrication of potential instrument suppliers, the practices described above form a useful general yardstick as to the likely performance and reliability.

2.2.1.3 Vacuum pumps

Only UHV pumps will be discussed under this heading, and then only briefly with regard to their respective advantages and disadvantages, since descriptions of their modes of operation can be found in many other places.

The four types of pump found in various combinations on commercially produced spectrometers are diffusion, sputter ion, turbomolecular and titanium sublimation, and it seems that each instrument manufacturer has his own preference. About the only area of agreement between manufacturers seems to be in the use of titanium sublimation pumps as auxiliary and not main pumps for UHV.

Diffusion pumps can be used over a wide range of required ultimate vacua, according to the type of oil used; for UHV they need oil of a vapour pressure less than 10^{-9} torr, of high resistance to degradation leading to volatile products and of reluctance to creep over surfaces. The polyphenyl ethers have been used with success for some years, but other oils with improved characteristics are now becoming available. Although diffusion pumps themselves are relatively cheap, they need efficient liquid nitrogen traps situated between them and UHV, and such traps in fact cost more than the pumps. When the costs of liquid nitrogen, of cooling water and of power over a long period are also taken into account, the cheapness is only apparent. Against this is the major advantage of diffusion pumps compared to those below—that they are well behaved and not temperamental. They are prepared to pump almost any gas that is not reactive towards the hot oil and will work for very long periods without needing attention if they are not abused.

Sputter ion pumps have many attractions. They do not use fluids, do not need cooling water or liquid nitrogen, nor indeed much power, and can be connected directly to the vessel to be evacuated without any intervening traps. They can be brought into operation at the throw of a switch rather than having a long warm-up time and can be switched off again just as easily. Although inherently expensive, since a control unit is needed with them, no additional components are necessary and there are no cooling costs. It is likely that over a long period of time they are cheaper to run than diffusion pumps. In general use, however, as the main UHV pump these advantages are more than outweighed by their disadvantages. Principally, they are fussy about what they will pump, and how much of it. As long as the gas load consists of the normal residual gas constituents, such as nitrogen, oxygen and carbon dioxide, there are no problems. With the modern improved designs

they will even pump noble gases with reasonable speeds, provided the noble gas pressure is kept below 10^{-6} torr and the pump is not exposed to it for long continuous periods. However, ion-pumping of helium is to be avoided, since helium diffuses into the titanium cathode and can cause cracking of the cathode if too great a volume of helium becomes incorporated. The same applies to hydrogen, with the additional complication that titanium forms a series of solid solutions and compounds with hydrogen, so that the end result of prolonged pumping of hydrogen could be disintegration of the cathodes. Furthermore, since water vapour is pumped by dissociation of the water molecule into hydrogen and hydroxyl ion in the pump discharge region, followed by reaction of the products with the titanium, excessive pumping of high pressures of water vapour will have the same eventual effect as pumping hydrogen. Another drawback of ion pumps is the memory effect. For example, since, as stated above, both helium and hydrogen are pumped by dissolution in the titanium, it follows that subsequent sputtering of the helium- or hydrogen-filled titanium will cause re-evolution of those gases. The problem can be overcome to a certain extent by a high temperature bake-out (~500°C) of the pump, but that operation is inconvenient, to say the least. In addition, although ion pumps will pump hydrocarbons effectively, the carbon remains associated with the titanium and can recombine with other gases to introduce contamination. Thus on an ion-pumped system there will always be molecules containing one and two, and often three, carbon atoms in the residual gas spectrum; methane in particular can be a principal component, along with hydrogen, at base pressures of less than 10^{-9} torr. For the same reason, admission of pure oxygen to an ion pump will cause an increase in the carbon monoxide partial pressure to the point where it interferes seriously with the course of an experiment. In summary, then, ion pumps are suitable, and highly convenient, for pumping atmospheric constituents, but cause many problems when their application differs from that.

Turbomolecular pumps are also attractive from the point of view of convenience. A turbomolecular pump is the only type of pump that can in principle take a vacuum system from atmospheric pressure to UHV (with bake-out, obviously), although in practice it is never used in that way since to do so regularly would impose too much strain on it. It will pump any gas, but the pumping efficiency decreases directly with the molecular weight of the gas, so that the efficiency for hydrogen and helium is not good; the residual gas at UHV will therefore be mainly hydrogen. The only limitations on the nature of the gas pumped will be either possible reaction with the lubricating oil on the low vacuum side or possible corrosive effects on the rotor blades. In normal practice neither of these possibilities causes any problems. Turbomolecular pumps do not need traps or baffles like diffusion pumps and, like ion pumps, are probably cheaper to run over a long period than diffusion pumps. There is some lingering doubt as to whether they are capable *on their own* of achieving pressures of less than 10^{-10} torr since they tend always to be used in conjunc-

tion with titanium sublimation pumps, but then the same can be said of diffusion pumps too. In fact most manufacturers will not guarantee ultimate vacua of 10^{-10} torr or better without the aid of a sublimation pump, whatever the type of principal pump used. Nowadays turbomolecular pumps run very quietly, one of the original objections being noise, and it seems that the only reasons that they are not used more often on spectrometers are the possibilities of vibration and an inherent distrust of their reliability, again based on past experience.

Titanium sublimation pumps are by far the simplest and cheapest of the UHV pumps and are used very widely indeed, but, as remarked earlier, never as the principal pump on a spectrometer. They can be, and are, operated at pressures as high as 10^{-6} torr, but in electron spectrometers and other UHV systems their main contribution is to achieve the last order of magnitude or so decrease in pressure necessary if ultimate vacua of 10^{-10} torr and below are to be reached. For maximum pumping speed the titanium should be evaporated onto a liquid nitrogen cooled surface, but no standard spectrometer offers such a facility, and indeed much of the benefit of their ease of operation would disappear if a liquid nitrogen feed had to be provided. At the least, however, they ought to be water cooled, but even that is unfortunately not as common as it should be. The first operation of a titanium filament for sublimation has to be carried out carefully, since considerable release of gas occurs on initial heating, but once outgassing is complete the heating current can be switched straight to the maximum on subsequent operations. Normally the pumping of a sublimation pump is timed by the control unit, so that the titanium is not used wastefully; as a rough guide the pump need be operated only every few hours for two minutes at a time, at pressures around 5×10^{-10} torr. Since sublimation pumps operate by 'gettering', i.e. by reaction of the evaporated titanium with reactive gases, they are excellent pumps for the major atmospheric constituents and for hydrogen and hydrocarbons, but do not pump noble gases at all. Furthermore, since the titanium film is not subsequently disturbed, as the cathode is in an ion pump, there are very few memory effects.

Based on the above remarks, the most trouble-free pumping systems would consist of diffusion or turbomolecular pumps acting in parallel with titanium sublimation pumps. Ion pumps are not recommended as the principal pumps, for the reasons given, although they should be used as appendage pumps for pumping small subsidiary volumes to which noble gases will not be admitted or occasionally as 'holding' pumps in the event of failure or maintenance shut-down of the principal pumps.

2.3 Sample Handling

Under this heading will be discussed all the steps necessary in the process of offering a sample to the energy analyser for analysis, from preparation of the

sample outside the spectrometer to insertion in, and positioning inside, the spectrometer, and to further preparation *in vacuo* that might be required.

2.3.1 Sample preparation

Most of the preparation that occurs before insertion in the spectrometer is devoted to mounting a sample on a probe or a manipulator, or on a stub, in such a way that the surface to be studied is presented for analysis with minimal contamination. The wearing of gloves to avoid direct skin contact with the sample or any other component to go into UHV must therefore be routine. Similarly, all tools used to help with sample preparation must be de-greased, handled only with gloves and reserved for the purpose in a clean area. Ideally the clean area should be a separate room or cubicle into which the air supply is filtered, but few laboratories have space to spare for that, and it is generally adequate to use an enclosed bench or a glove box for the purpose, with a positive flow of filtered air through it.

Methods of sample mounting vary from sample to sample and from one laboratory to another and are often the subject of considerable ingenuity. Where samples are to be analysed in a routine fashion at room temperature, i.e. do not require any special heating or cooling arrangements, modern spectrometers provide standard sample carriers, whose design varies with the manufacturer, but all of which have two things in common. Firstly, they have features such as recesses or grooves or spring clips that allow them to be transferred from the insertion system to the manipulator and, secondly, they have a flat area on which the sample is actually placed. The ingenuity arises in the ways in which awkwardly shaped samples can be fixed to this area in such a way that the materials used for fixing do not interfere with analysis of the sample surface of interest. Flat samples are always the easiest; if the reverse side is not of interest, they can be stuck to the carrier with a blob of methanol dag (*not* water dag), which dries rapidly and holds them firmly even after drying. Otherwise they can be held by being slipped under small spring clips of beryllium copper or some other springy metal fastened to the carrier.

The mounting of irregularly shaped samples is a function of individual sample geometry. In some cases the sample can be held sufficiently firmly against the carrier with a simple strap of some malleable metal such as nickel which can mould itself to the sample shape; the ends of the strap can either be spot-welded to the carrier or fixed under screw heads. In other cases it may be necessary to wrap the sample in metal foil, often of platinum, leaving the surface of interest exposed; the foil wrap acts as a cup in which the sample is held, the ends of the foil again being either spot-welded or held under screw-heads for fixing to the carrier. In whichever way it is mounted, care has always to be taken that the mounting material does not intrude into the analysed area or, if that is unavoidable, that the photo-electron spectrum of the mount does not interfere with that expected from the sample.

Powders pose some special problems. If they cannot be, or should not be, compressed or compacted, they can be placed loose in a recess either fixed to a carrier or part of a modified carrier, but there are obvious risks in so doing. Initial evacuation from atmospheric pressure, if carried out too quickly, will distribute much of the powder around the vacuum system; the powder is always likely to spill as the carrier is moved or transferred from one part of a spectrometer to another; and in some spectrometer configurations the analysed surface must be inclined to the horizontal at an angle of at least 30°, which must again lead to spillage into the system. A recommended way of 'fixing' a powder for analysis is to shake some of it over piece of indium foil to approximately uniform coverage and then to press gently down on the powder layer with a piece of clean, hard, metal such as molybdenum or tungsten. Enough of the powder then becomes embedded in the surface of the indium to give a continuous surface for analysis, and the excess powder can be shaken off. This method has the additional advantage that the proximity of the indium reduces surface charging effects, which can be severe in the analysis of loose powders. Another popular method of mounting powders is to use double-sided adhesive tape, but this is not recommended for AES.

In the more complicated situation where the sample has to be heated or cooled, either before or during analysis, the mounting of the sample needs greater precision. Generally, the experimental programmes involving such analyses will be of longer term than those in which samples are transferred on carriers, and samples are therefore mounted individually and directly on the manipulator rather than being transferred to it. Most manufacturers market either special probes or special manipulator attachments, for use with their spectrometers when heating or cooling is required, and an example of a combined stage is shown in Figure 2.2. Heating is provided indirectly by an insulated hot filament under the substrate supporting the sample and cooling by passing liquid nitrogen into a tank to which the substrate is connected by a thick copper braid. In general it is advisable to use a probe or attachment dedicated to one or the other, since the dual-purpose probe has inherent limitations on the temperatures it can achieve. Since the actual temperature reached by a sample during either heating or cooling depends very much on the extent of thermal contact of the sample and the substrate, it is always advisable, if it can be done without damage to the sample, to attach a thermocouple directly to it rather than rely on a thermocouple attached to the substrate. Alternatively, a dummy sample can be used at first, so that a calibration of the substrate temperature can be obtained in terms of the actual sample temperature.

One of the worrying aspects of the above method of sample heating, i.e. by conduction from a heated substrate, is that impurities present either on or in the substrate can become mobile at high temperature and contaminate the sample. Such an occurrence would clearly lead to spurious experimental results. It is therefore necessary, when the heating stage is first used, to run it

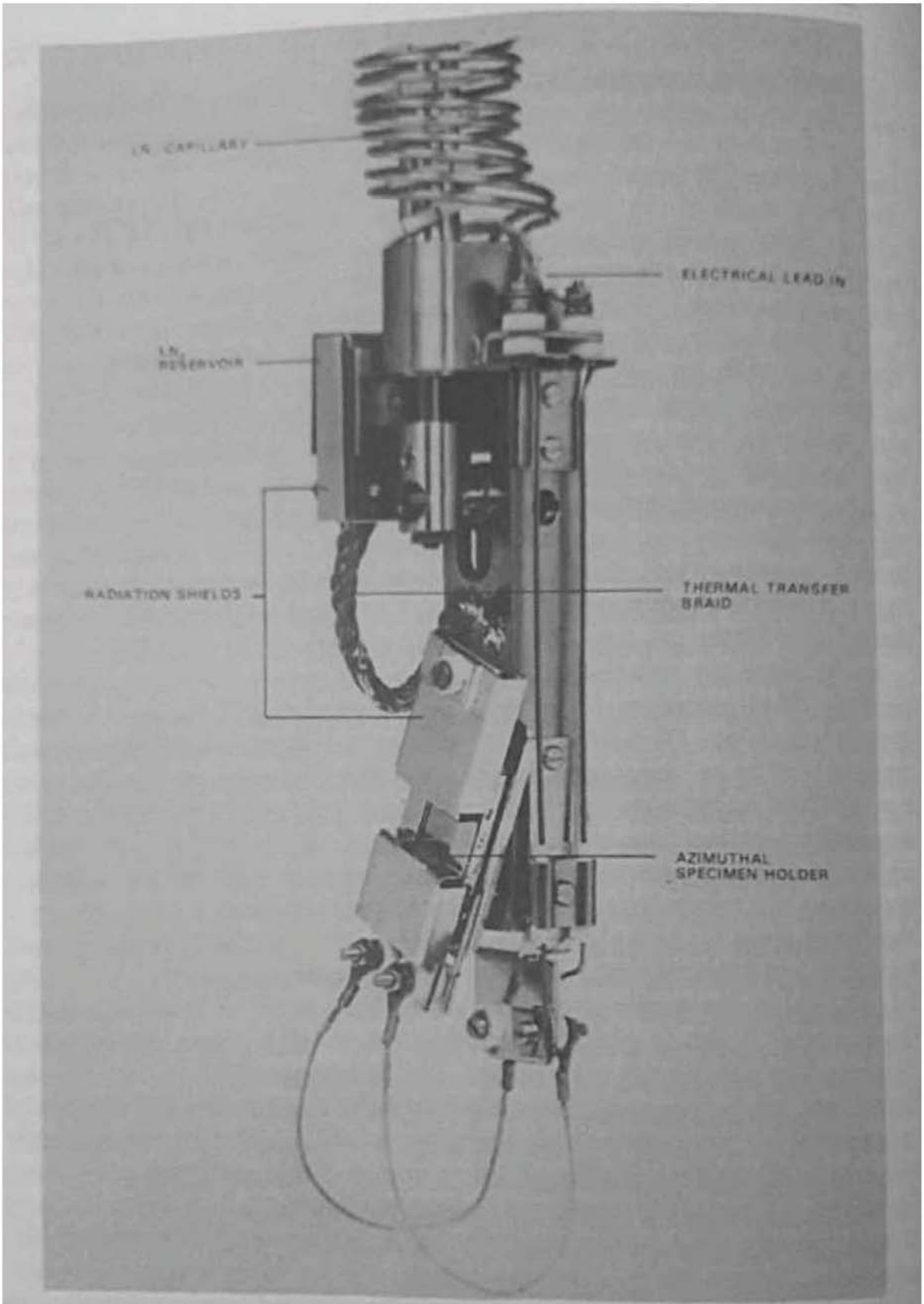


Figure 2.2 Photograph of a combined heating/cooling stage for specimen treatment. Heating is provided by conduction from an insulated hot filament under the substrate surface on which the specimen is mounted. Cooling is achieved by conduction along a copper braid fixed to a tank through which liquid nitrogen is passed. The stage is mounted on the shaft of a universal manipulator (see Figure 2.9). (Reproduced by permission of Vacuum Generators Ltd, Hastings)

for prolonged periods at the intended temperatures without any sample being present and to monitor the cleanliness of the substrate surface. If any contaminating species appears and persists, it should be removed by ion bombardment. The cycles of heating and ion bombardment should be continued until the contamination does not reappear on heating. In other words, during initial operation the substrate should be treated as if it were a sample, and cleaned in the same way.

In basic experimental work involving very pure materials, often in the form of single crystals, the above method of heating the sample is not regarded as adequate. There must be no possibility of contamination from any supporting material and the sample must be heated uniformly. According to the shape, thickness and nature of the sample, heating may be carried out by the passage of current through the sample, by conduction from heated supports, by radiation from an adjacent hot filament or by electron bombardment from a hot filament. Generally speaking, the supporting materials should either be the same and of the same purity as that of the sample or of a refractory metal such as tungsten or molybdenum that can be cleaned at a high temperature before being used as a support. For measurement of temperatures up to ~850 °C it is necessary to use a thermocouple either spot-welded very carefully to one edge of the sample or, if that is unacceptable, attached to a support as close to the sample as possible. Noble metal thermocouples, e.g. Pt–Pt/10% Rh, must always be preferred to the more common NiCr–NiAl thermocouples since the risk of contamination from the latter is much greater. For temperatures above 850 °C an optical pyrometer should be used, since at high temperatures there is an increased likelihood of reaction of the sample with thermocouple material; of course, it should be remembered that pyrometer readings are always subject to correction due to sample emissivity and to absorption in the glass window through which the sample is observed. With some mounting arrangements it may be possible to measure the sample temperature by using the temperature dependence of the resistance of wires supporting the sample.

2.3.2 Sample insertion

When a sample has had to be mounted individually on a manipulator, as discussed above, then clearly its insertion into the spectrometer involves bringing part of the spectrometer, usually the analysis chamber, to atmospheric pressure and bolting the manipulator directly to the chamber. This is of course a time-consuming procedure, since the chamber then has to be re-evacuated and the whole spectrometer baked in order to return to UHV. However, if the sample is one that is to be studied for a long period, then the time spent in regaining UHV is short in proportion. Until a few years ago this procedure was universal since no specimen transfer systems were available, which meant that for routine analyses at UHV sample turn-around was very

slow. Some improvement could be achieved in routine analysis by multiple-sample mounting, e.g. on a carousel which could be rotated via the rotary drive on the manipulator, but in general there was little flexibility in sample queueing for analysis.

One of the principal directions in which the sample turn-around has been greatly improved in recent years has been that of the method of inserting and transferring samples from ambient to UHV. This method at present takes two different forms, depending on the individual manufacturer's philosophy. In one form the carrier holding the sample is placed at the end of a long shaft, or probe, with a very highly polished cylindrical surface. The shaft slides, entirely grease-free, in a series of close-fitting sealing rings of either Viton or Teflon*, depending on the manufacturer, from ambient through a fore-chamber into the UHV chamber. The fore-chamber is pumped to at least 10^{-6} torr before the gate valve to UHV opens. Once inside the UHV chamber, the shaft is moved forward in a controlled manner until the carrier docks with, and locks into, an empty position on the sample carousel, when the shaft can be retracted. Removal of a sample simply involves the reverse procedure.

The advantages of the shaft method of sample insertion are that it is fast, that it can be made semi-automatic and that only one transfer operation is necessary. The disadvantage is that part of the shaft surface intruding into the UHV chamber has arrived there directly from the outside ambient atmosphere without any intervening bake-out or outgassing apart from initial pump-down in the fore-chamber. This, of course, is always true of the sample itself, and of its carrier, but their total surface area is much less than that of the exposed shaft surface, and in any case the carriers are subject to bake-out at some stage in their use. It is usual, then, during sample insertion by shaft for the pressure in the UHV chamber not only to rise out of the UHV region but to take some time to drop below 10^{-9} torr again; in fact, after a series of sample insertions and removals in quick succession, it may be impossible to regain 10^{-9} torr without bake-out. This method, then, is suited to those applications which require routine analysis of relatively inert surfaces and in which the maintenance of pressures in the analysis chamber near 10^{-10} torr at all times is sacrificed to the speed of sample turn-around.

The other form of specimen insertion could be termed the successive transfer method. In it, the carrier with the sample is placed on a holder in a small fore-chamber that is then evacuated from atmospheric pressure to $\sim 5 \times 10^{-3}$ torr with a trapped rotary pump. A gate valve between the fore-chamber and a second, or preparation, chamber is opened and the carrier is moved into the second chamber on a trolley enclosed entirely within the vacuum. After lifting the carrier onto another trolley in the preparation chamber, by manual operation via a bellows-sealed fork, the first trolley is returned to the fore-chamber

*Registered trade name.

and the gate valve is closed. Since the preparation chamber has a much greater volume than that of the fore-chamber and is pumped by a large diffusion pump, its pressure returns quickly to $\sim 10^{-9}$ torr. A second gate valve between the preparation chamber and the analysis chamber is then opened, the carrier moved through into the analysis chamber on the second trolley and transferred to the manipulator, again by manual operation. Once the transfer is complete, the trolley is retracted and the second gate valve closed. Since the pressure in the preparation chamber during transfer is $<10^{-9}$ torr, the UHV in the analysis chamber is hardly disturbed.

Clearly the above method is slower than the shaft insertion method, there is a not insignificant risk of dropping a carrier during transfer and it is difficult to see how it could be automated in any way. The overriding advantage from the experimental point of view is that the pressure in the analysis chamber always remains near the 10^{-10} torr achieved by bake-out, etc. If the sample itself were found to be a source of gas for any reason, it could be parked in the preparation chamber either until the outgassing had stopped or until it had been decided that it was unsuitable for analysis.

Not infrequently it is necessary to study samples that should not be exposed to atmospheric ambient at any stage, perhaps because the material of the sample itself would react too vigorously with the atmospheric constituent gases, or because there might be surface films of interest on the sample that would be destroyed or altered by reaction with the ambient, or even that the sample carries some chemical or radioactive contamination that should not be allowed to be dispersed. Various methods have been devised to minimize the exposure to the ambient during insertion of a sample into a spectrometer. The simplest consists merely of a large plastic bag surrounding either the insertion shaft or the entrance to the fore-chamber, as the case may be, in which is maintained a positive pressure of an inert gas such as argon. Sample mounting then takes place inside the bag by placing gloved hands through holes small enough to prevent back-streaming of atmospheric gases. A more elaborate, and more effective, version of this method is a glove box built to fit the end of the spectrometer into which insertion takes place, with proper sealed ports and sealed gloves. Again a positive pressure is maintained inside the box. Such a device has the advantages that many sensitive samples can be stored in it for future reference and that visibility of manipulation is very good.

More elaborate still for the protection of sensitive samples in transit to the spectrometer are the purpose-built transfer flasks, designed to pass through the standard port on a glove box and with integral flap valves. An example is shown in Figure 2.3. Inside the glove box, in an inert gas atmosphere, the valve is opened, the sample inserted and the valve closed again. The flask can then be removed, with the sample still surrounded by the inert gas, taken to the spectrometer and sealed to the fore-chamber. With the fore-chamber also filled with inert gas, the valve is opened and the flask and chamber pumped

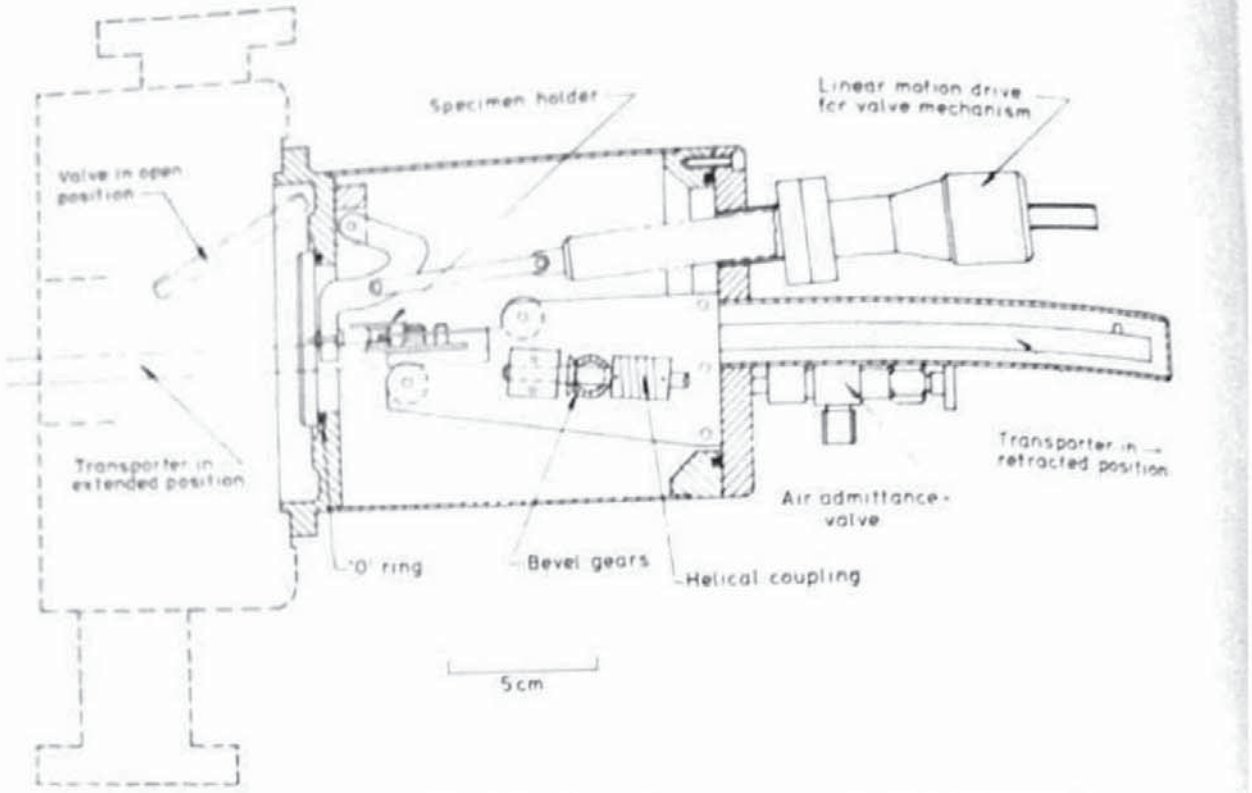


Figure 2.3 Specimen transfer flask for insertion of sensitive specimens into a spectrometer. The flask will pass through a standard port on a glove box and is designed to be attached to the entry port on ESCALAB. After the specimen has been loaded in an inert atmosphere, the valve is shut and the flask taken to the port, to which it is attached. The fore-chamber is then back-filled with inert gas, the valve opened and the chamber and flask evacuated together. Once a suitable vacuum has been achieved, the specimen is transported into the next chamber in the usual way. (Reproduced by permission of Vacuum Generators Scientific Ltd, East Grinstead)

out together to the normal vacuum required for sample insertion. After that, sample transfer follows the same procedure already described, using the transfer trolley in the flask rather than the one in the fore-chamber.

2.3.3 Sample treatment *in vacuo*

Treatments given to a sample after arrival in the spectrometer fall into three main categories, viz. preparation of a clean surface, depth profiling and surface reactions. Included in the latter are those treatments that consist of heating alone in order to produce surface segregation of impurities or to melt the sample if liquid metals are being studied.

2.3.3.1 Preparation of a clean surface

In the context of this volume, 'clean' is defined as the state of a surface in which the techniques of XPS and AES cannot detect characteristic spectral

features from any elements, except those constituting the bulk composition of the sample, above their limits of detection. Obviously if a surface analytical technique of greater sensitivity were to be applied to a surface judged clean by this definition, impurity elements would probably be detected. As always, cleanliness is a relative concept.

At first sight, the simplest cleaning technique is heat treatment. Several elemental materials, including silicon and the refractory metals, can be rendered clean according to the above definition by heating them to sufficiently high temperatures, generally a few hundred degrees Celsius below their melting points. That is, their surfaces will be clean while held *at* the high temperature. The problem is to maintain their cleanliness on cooling to room temperature, since as they cool their temperatures will pass through ranges in which impurities segregate quickly to the surface. The same problem arises with multielemental materials such as alloys and compounds whose temperatures cannot be raised sufficiently to achieve a clean surface while hot, but which have to undergo heat treatment as part of an experiment; in certain temperature ranges impurities present in the bulk at a low level can accumulate at the surface. The most common segregants are sulphur, oxygen and carbon, although in special cases other impurity elements such as nitrogen and phosphorus may appear in considerable amounts.

Heating a sample will thus hardly ever be effective by itself in producing a clean surface at room temperature. Another technique is needed to remove impurities physically from a surface, either those which have segregated as a result of heating or those present on the surface originally; by far the most commonly used is that of ion bombardment. In principle it is very simple. A beam of positive noble gas ions of energies between 500 eV and 5 keV, typically, is directed at the surface. As a result of the exchange of energy in the surface and subsurface regions, some atoms or clusters of atoms at the surface are given enough kinetic energy to leave the surface and be lost from it. In other words, the surface is eroded. The gas normally used is argon, chosen as a compromise between efficiency of removal of material, which increases with atomic weight, and vapour pressure at the temperature of liquid nitrogen traps, if in use; the heavier noble gases tend to condense at low temperatures and be difficult to pump away. The process of erosion of the surface is called sputtering and the source of ions is called an ion gun. Argon can be supplied to an ion gun in two ways: either by back-filling the system so that the pressure everywhere is that required to operate the gun or by passing the gas directly into the region of ion production so that only in that region is there a high pressure. The former method is a static one, in that it is used on ion pumped systems in which the ion pump must be valved off since it must not be allowed to pump argon continuously. The latter method is dynamic, since the exit apertures in the ion gun are sufficiently small for differential pumping to take place across them, allowing the pressure in the analysis

chamber to be maintained at a low level by either diffusion or turbomolecular pumps. Sample cleaning is much more likely to be successful in a dynamically pumped ion gun, where impurities released by erosion are swept away, than in a static system where recontamination can occur. Description of the various types of ion gun in use and their modes of operation will be found in Section 2.4.3.

As remarked above, the technique of cleaning by ion bombardment is in principle simple. In practice it can be straightforward, too, for there are many published examples where surfaces have been cleaned by ion bombardment alone. In the majority of cases, however, it is found that although ion bombardment will remove the bulk of a layer of contamination, there is always left behind a small amount, of the order of a few atom per cent, of contaminant atoms. Carbon and oxygen are generally the most difficult to remove completely. In many applications and analyses the residual low level of contaminant may be acceptable, if all that was expected of the bombardment was the removal of sufficient masking contamination to be able to carry out an analysis of the sample. In other applications it will be essential to produce a clean surface, as defined above, in which case the most common procedure is to use a combination of techniques, viz. alternate cycles of heating and of ion bombardment.

If the sample is thin, e.g. in the form of foil, and already reasonably pure, it should be possible, by appropriate choice of temperature and ion dose in the above procedure, to denude the sample completely of those impurities that segregate to the surface on heating. Once the sample is in that condition, it is then easy to renew the cleanliness of its surface when desired by a single cycle of bombardment and/or heating, since no further contamination of bulk origin will appear. For thick samples, it is clearly impractical to attempt to remove all segregating impurities from the entire sample, but in many cases careful adjustment of temperature and ion dose can again achieve a situation in which the surface region, extending to perhaps 1–2 μm, is denuded of impurities. Such a situation is obviously not stable under any subsequent heat treatment to temperatures near that used in the cleaning, but if there is to be no heat treatment, or only to a much lower temperature, then it represents a useful compromise.

Although ion bombardment is used so universally for surface cleaning, either by itself or in combination with heating, the potential user should be aware of the artefacts that can be introduced by it. These take the form of either or both chemical[3] or topographical[4] changes induced in the surface. Chemical changes here mean the alterations in the elemental composition or chemical state, or both together, of the major constituents that can occur at the surfaces of alloys and compounds during ion bombardment. Topographical changes are usually in the direction of increased surface roughness, or, occasionally, of statistically induced surface roughness. These effects are dis-

cussed in detail in Chapter 4, but it is useful to consider them briefly here in the context of surface cleaning.

Since the sputtering yields, i.e. the number of atoms sputtered per incident argon ion, vary substantially across the Periodic Table, as shown in Figure 2.4 from Seah,[5] it follows that in a material containing two or more elements the likelihood is that at least one of the elements will be removed preferentially due to its higher sputtering yield. Surface analysis after ion bombardment would therefore reveal a depletion in that element compared to the bulk or stoichiometric composition. Depletion would continue with further bombardment until an equilibrium situation was reached in which the relative atom populations of elements at the surface was balanced by the relative sputtering yields. This effect has been observed[6] frequently in alloy systems in which there is a significant disparity in sputtering yields between the constituents. If there are no matrix effects, and if the individual elemental sput-

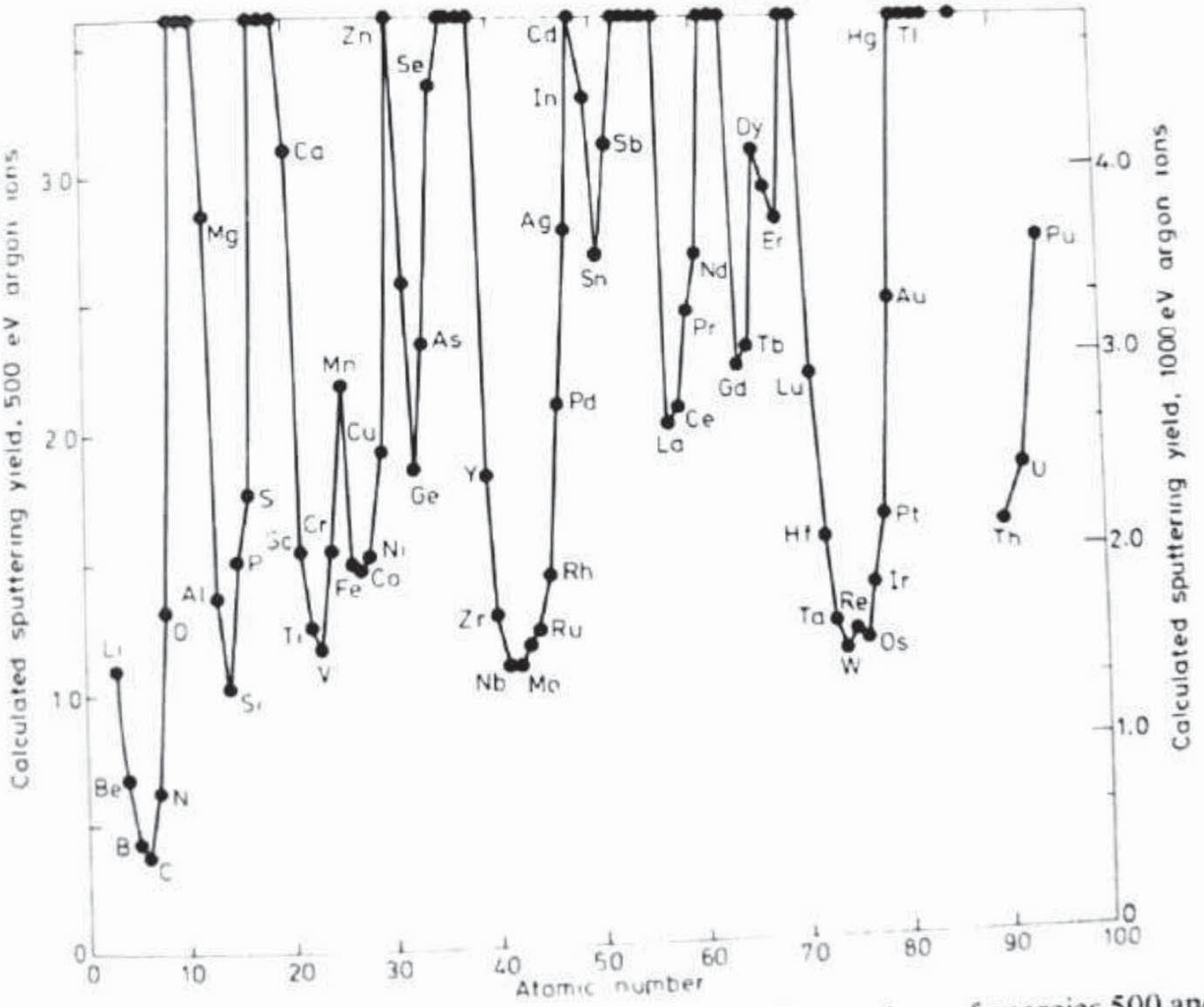


Figure 2.4 Predicted values of the sputtering yields of argon ions of energies 500 and 1000 eV across the Periodic Table. The yields of several elements with particularly high sputtering yields are not shown, for the sake of clarity, but may be found in the original paper by Seah[5]

tering yield are known, then it is possible to work back to a surface composition in the absence of sputtering; this will be discussed further in Chapter 5.

The other chemical changes induced by the bombardment itself can be equally serious. They include reduction of compounds[5] and persistence of elements due to 'knock-on'. In the former the more volatile elements such as oxygen in oxides, sulphur in sulphides, etc., are lost preferentially during bombardment even though their elemental sputtering yields may be similar to, or even lower than, those of the other constituents. Such loss is inevitably accompanied by a reduction in the oxidation state of, say, a metal in a metal oxide. In the effect of 'knock-on', as its name suggests, some elements, particularly light ones, can be driven further into the material by direct impact from a primary ion or by indirect impact via an energy cascade. The persistence of carbon in certain materials during prolonged bombardment, for example, is generally attributed to this effect.[7] It has not proved possible so far to take account in any systematic way of the artefacts of reduction and of 'knock-on' when attempting quantification in their presence.

The roughness of a surface can be increased by ion bombardment in several ways.[8] If the ions are directed at the sample surface from one direction only, i.e. from a single source, which is still the most common mode of operation, then in general any asperities will tend to be accentuated since sputtering efficiency is greater at higher incident angles. Eventually an equilibrium configuration will be reached in which the atom population in each different topographical position will be balanced by the sputtering yield in that position. A more serious situation can arise if there are impurity inclusions in the material whose sputtering yield is much less than that of the surrounding matrix, in which case the inclusion 'shadows' the material behind it from the incident ions and a conical pillar forms of progressively increasing height. Both these effects can be minimized, if not eliminated, by the use of two or more ion sources simultaneously from different directions, and an even better arrangement is to rotate the sample continuously during bombardment so that all angular effects are completely smoothed out. Unfortunately very few systems make provision for either of these arrangements.

Ion bombardment, then, although certainly effective in cleaning a surface, especially in conjunction with heating, brings in its train some highly undesirable side-effects, of which some are accountable but others as yet are not. Nevertheless, the technique is so widely used that it is clearly here to stay, and thus it is necessary to keep watch continually for intrusion of artefacts during the cleaning of any one sample.

In special cases other methods of producing clean surfaces are applicable. For example, soft metals such as lead, indium and tin have been cleaned[9] successfully by scraping their surfaces in vacuum with a tool, such as the edge of a razor blade, mounted on the end of a rod passing through a flexible

bellows. Cleaning in this case obviously corresponds to the creation of a new surface by mechanical removal of the original surface and its associated contamination. Another specialized method is that of fracture in vacuum. Some materials such as alkali halides, silicon, germanium, etc., will cleave easily along well-defined cleavage planes, so that all that is required to produce a clean surface with a composition close to that of the bulk is some mechanical ingenuity in the method of holding the material, cleaving it and offering it to the energy analyser for analysis. In other materials such as polycrystalline metals and alloys, which are too ductile at room temperature to cleave, fracture may still be possible by lowering the temperature to the point at which they become brittle. The fracture path may then either pass through the individual grains (transgranular) or follow the grain boundaries (intergranular); in either case a surface free from contaminants is produced, although the intergranular surface may contain certain impurities segregated from the bulk, as described in Chapter 7. An example[10] of a fracture stage is illustrated in Figure 2.5. All such stages possess the facility of cooling the sample to be fractured to temperatures near that of liquid nitrogen.

Methods of preparing atomically clean surfaces on bulk samples have recently been collected in a useful review by Musket *et al.*[11] The most widely used method of producing a clean surface without having to resort to ion bombardment or heating of the sample is that of deposition of pure material onto a substrate in ultra-high vacuum by evaporation either from a heated filament or crucible or by electron bombardment. It is possible in principle to deposit *any* solid element as a thin film in this way, as well as many compounds, but in practice the requirement of a pure and atomically clean surface imposes stringencies that prevent the use of some elements. The two overriding criteria are (a) that the evaporated film should be of the same composition as the starting material and (b) that there should be no release of gas during evaporation. Difficulties in complying with the first of these arise if the material to be evaporated has to be supported in any way during evaporation, since then there is always the possibility of reaction by alloying, or if the starting material is itself multicomponent, since it is then possible for differential evaporation rates to produce a deposit of different composition. The second criterion can usually be obeyed by careful attention to evaporation procedure, i.e. a thorough outgassing of the sample and its support before the film to be used is actually deposited.

The most straightforward evaporators are those made from the sample material without any support; they usually consist of a tightly wound coil of wire of the material through which a current is passed to raise the temperature by ohmic heating to the point at which the material sublimes freely. Obviously the properties of the material must be such that the coil does not sag at the temperature of sublimation, and for this reason the method is limited to a few of the more refractory metals. The advantages are that there

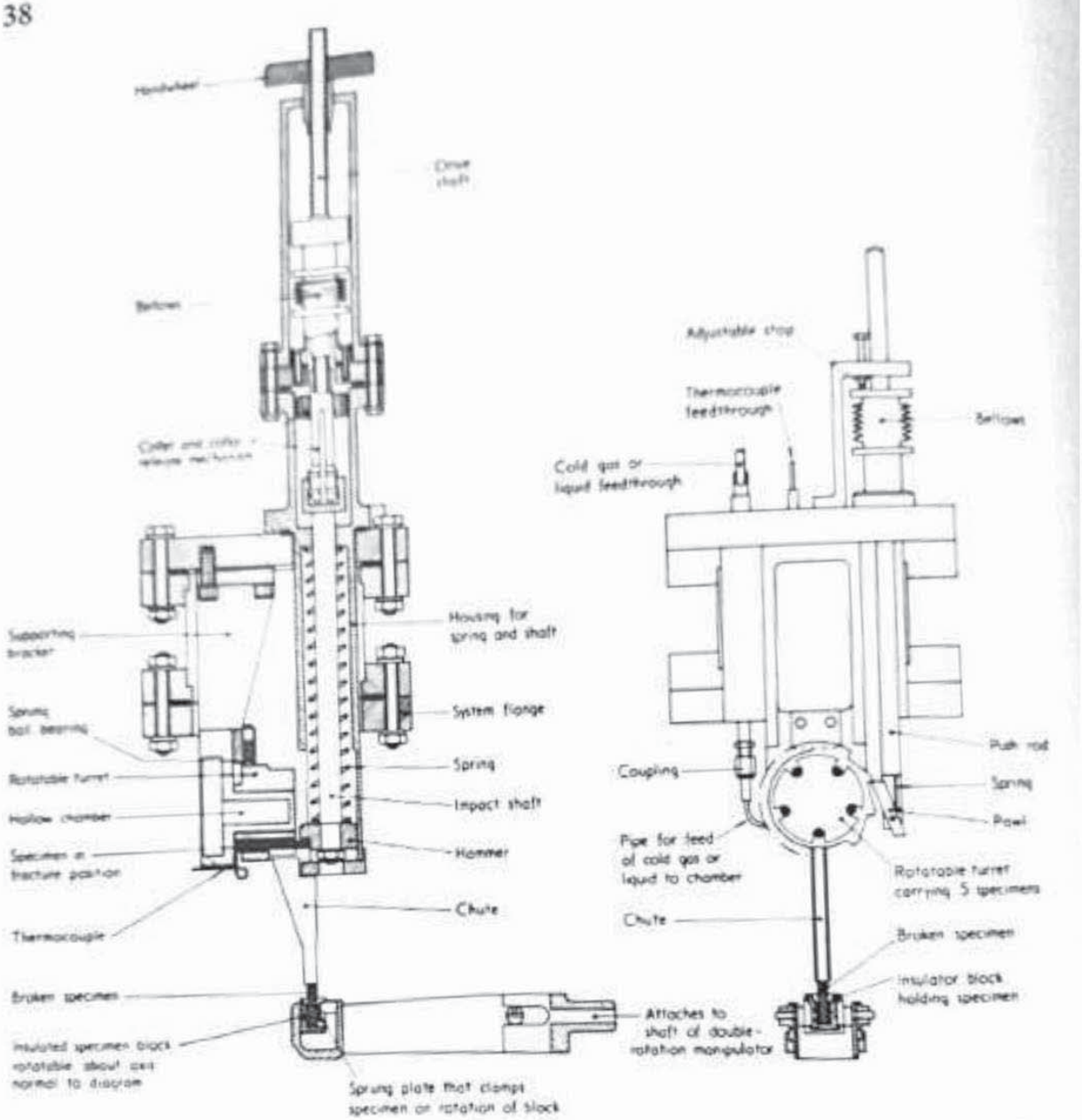


Figure 2.5 Impact fracture stage for fracture of brittle samples *in vacuo*. The rotating turret has holes for five specimens, which are clamped in place by grub screws. Rotation is achieved with a vertical push rod. To fracture a sample, the hammer is wound up against a powerful spring until at a predetermined point a release mechanism allows it to fly back and hit the free end of the sample. The broken piece then falls down a chute into a catcher with which it may be positioned for analysis. The sample may be cooled before fracture by passing liquid nitrogen into a tank on which the turret rotates. (After Coad *et al.*[10])

is no possibility of reaction with any support material and that the evaporator can be completely outgassed; the disadvantages are that the rate of deposition is slow and that the surroundings tend to become very hot unless proper radiation shielding is included.

Almost as straightforward are those evaporators that consist of a coiled filament supporting a sample material that wets but does not alloy with the filament material. Here the filament is invariably of ductile (i.e. annealed) tungsten or molybdenum. Metals that can be evaporated easily from such filaments include copper, silver and gold, typically, but others such as barium and magnesium are also possible. When using the noble metals it is desirable to pre-melt the sample onto the filament loops either in an auxiliary vacuum system or in a hydrogen stream, since the filament itself can then first be outgassed at a temperature much higher than that of subsequent evaporation. If that procedure is followed outgassing of the evaporator prior to evaporation merely involves re-melting the sample once or twice, following which evaporation can take place without any measurable rise in pressure. For more volatile non-alloying sample materials it may be difficult to outgas the evaporator to the same extent; the bare filament can always be pre-outgassed as above, but if attempts are made to melt the sample most of it may be lost prematurely. Instead it may be necessary to heat for long periods to temperatures at which the vapour pressure is still low. Some metals, such as cadmium, zinc, selenium, tellurium, the alkali metals and mercury, have vapour pressures at room or bake-out temperatures that are too high in any case to allow their presence in the bulk elemental form in an ultra-high vacuum system.

The majority of metals that one would like to be able to evaporate in order to produce a clean surface have the unfortunate property of alloying more or less rapidly with the filament material at the evaporation temperature. Included are all the transition metals, aluminium, the rare earths, platinum, zirconium, uranium and thorium, and some other less used metals. Clearly, if the sample alloys with the filament, with the formation of a liquid phase below the melting point of the filament material, not only will premature failure of the filament occur by breaking but if the vapour pressure of the alloy is also appreciable then the deposited film will be impure. In this situation it is necessary to design the filament-cum-sample configuration so that the possibility of disastrous alloying is minimized; what this means in practice is that the sample must be so distributed along the filament that at all points the local volume ratio of sample to filament materials is low—less then 1:5. If the evaporator can be arranged in that way, then it is possible to evaporate much of the sample in pure form before alloying causes breakage or contamination of the film. There are various ways that have been used to reduce the local material ratio, and ingenuity might well suggest others. One obvious one is to electroplate, or perhaps even ion plate, the sample material onto the filament as a uniform coating. The coating thickness, and therefore the local material ratio, can be controlled accurately. The disadvantage is the possibility of trapping impurities such as water in the plating, which might either be difficult to remove subsequently or cause poor adhesion to the filament. Another popular way of distributing the sample along the filament is to use

the coiled coil. In this arrangement the sample is in the form of a fine wire with a diameter much smaller than that of the filament material; before the filament itself is wound, the sample is wound around the filament wire in a spiral whose pitch maintains the desired low local material ratio. Yet another way is to form the filament itself from several strands of fine wire, so that when the sample material, normally hung on the filament in small loops, is melted, it runs into the interstices of the strands by capillary action and distributes itself fairly uniformly. To be avoided is the method of hanging loops of sample material on *single* filamentary wire; on melting, blobs are formed in the filament turns, with consequent rapid alloying and failure.

In all the methods just described for the evaporation of metals that alloy with the filament (which is almost invariably of tungsten) the outgassing of the evaporator before the film is finally deposited presents some problems. Since the temperature of useful evaporation is also likely to be that at which some alloying with the filament may occur, clearly the prior outgassing must take place at a lower temperature, with the consequence that prolonged periods of heating are necessary. This is usually a matter of trial and error, for it is difficult to predict in any one case just what procedure will be successful; the published literature should be consulted for some guide to the most likely method.

Some metals present particular difficulties either because they will not wet the refractory metals used for filaments or because they are sufficiently volatile for evaporation only at temperatures much higher than their melting points. Amongst these are gallium, indium, lead, silicon, thallium, calcium, lithium, arsenic, antimony and bismuth. Because of the impossibility of evaporation from filaments, they are generally evaporated from crucibles of either alumina, beryllia, boron nitride or carbon, as appropriate, heated indirectly by a surrounding element of refractory metal. Such evaporators can rarely be outgassed sufficiently thoroughly for use in UHV, even after very prolonged heating, and indeed preparation of clean surfaces of the above metals by the evaporation method is hardly ever attempted.

The other popular method of evaporation is by electron bombardment, and this can often be successful for materials that are too difficult for one reason or another for filament evaporation. Electrons are focused from a hot filament onto the material to be evaporated, heating it to a temperature that can be controlled by the power input. Normally the electron emitter is near earth potential and the material at a high positive potential, but in some configurations the polarities may be reversed. According to the focusing conditions and the power dissipated, all or only part of the material may be heated to the temperature of evaporation. Commercially available electron bombardment evaporators have water-cooled hearths on which the sample material is placed, which avoids the problems described above of alloying with the support material. On the other hand, such evaporators cannot be regarded as

UHV-compatible since they are not usually bakeable and cannot be outgassed properly before use. Electron bombardment evaporation has been used successfully in UHV, however, by the so-called pendant drop method, in which bombardment of the end of a wire or rod of the sample melts part of it, whereupon surface tension forces pull the molten volume into a blob that hangs from the end and acts as the evaporation source. Outgassing can be carried out by de-focusing the electron beam to heat the sample along its length without melting it.

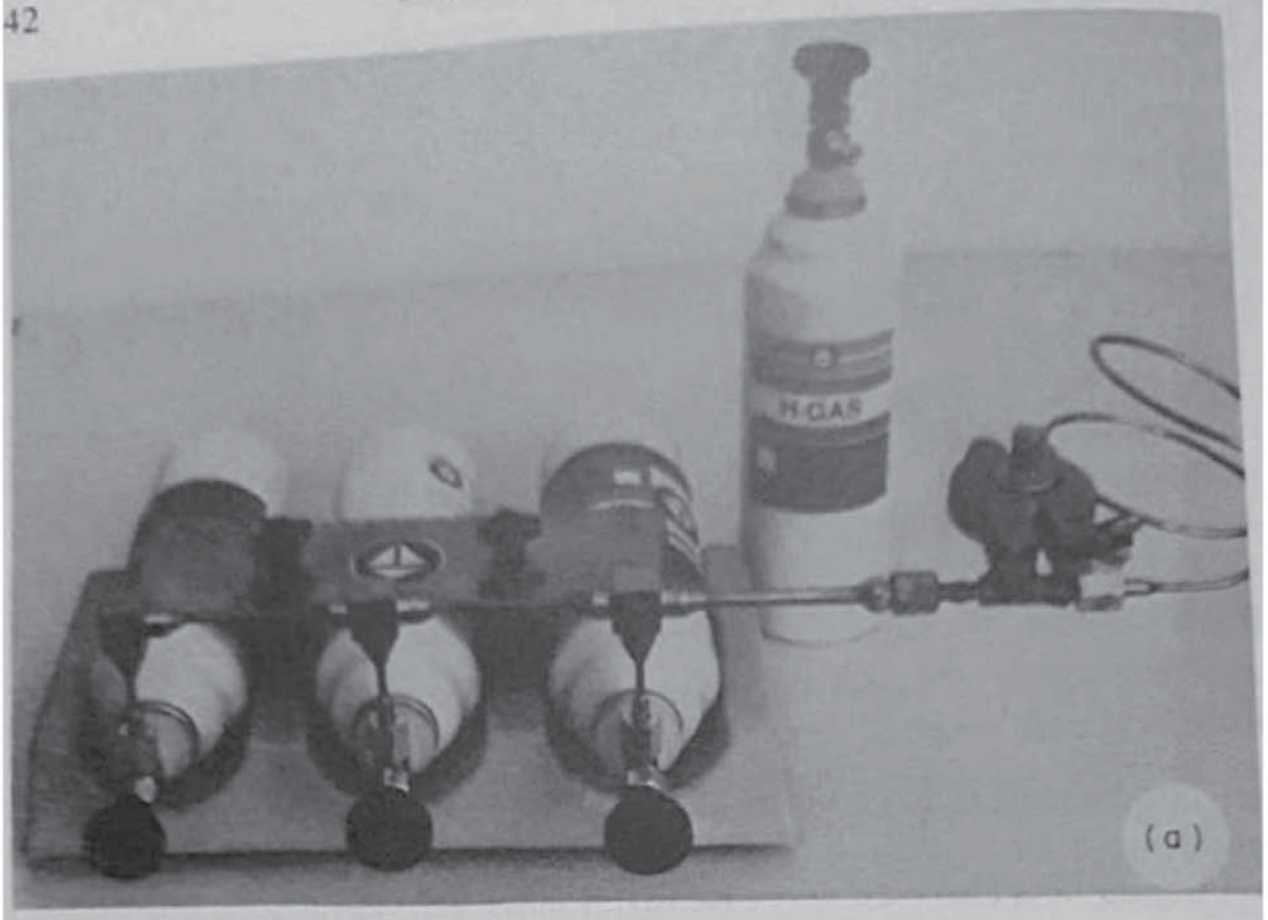
Whatever the method of evaporation used, it is always necessary either to remove the substrate from the line of sight of the evaporator or to swing a mask in front of it, during the preliminary outgassing period, to avoid the deposition of contaminated or impure material. This is particularly vital when a significant proportion of the sample is likely to be lost by premature evaporation during the outgassing procedure.

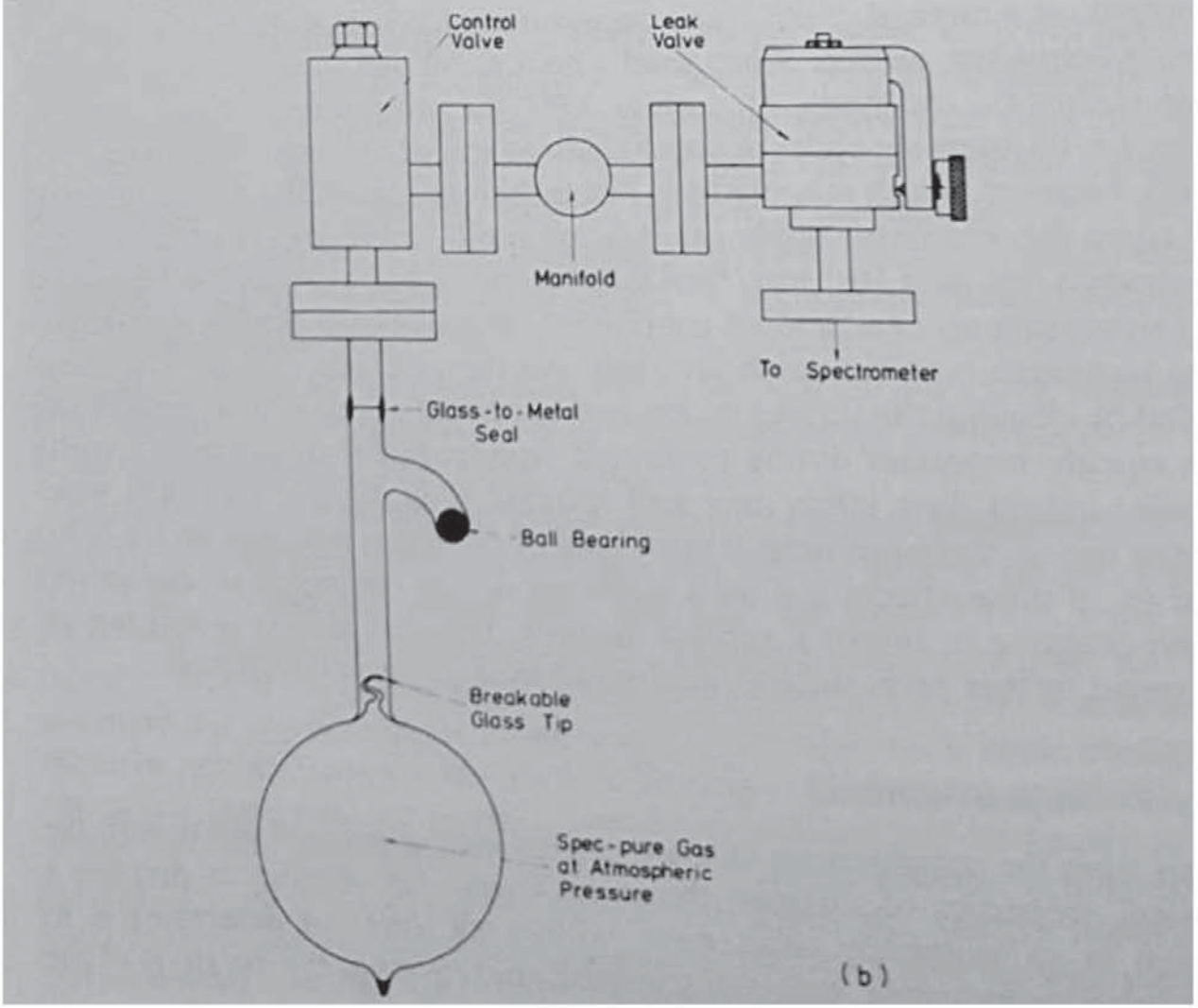
2.3.3.2 *Depth profiling*

Ion bombardment has already been mentioned as a potential method of producing a clean surface on a sample; some of its pitfalls have also been mentioned. It is also used very extensively indeed, in conjunction with surface analysis, as a method of obtaining compositional information as a function of depth below the surface. When used with a technique in which data acquisition is slow (at the moment), such as XPS, the information is built up stepwise, i.e. by alternate cycles of bombardment and of analysis. When used with AES, however, which is a relatively fast technique, and if the partial pressure of argon (or whichever noble gas is being used) in the region of the energy analyser is less than 10^{-5} torr, then bombardment and analysis can be carried out simultaneously and a more continuous compositional profile is obtained. The technique is called depth profiling. All the artefacts that may be introduced by cleaning the surface by ion bombardment are of course present and are equally important during continued bombardment to produce a depth profile; indeed their effect may well increase with depth. With such widespread use of depth profiling, a great deal of attention has been and is being paid to all these effects and their influence on the composition, and in fact depth profiling is almost a subject in itself. For that reason it will not be discussed further here, since a full discussion is given in Chapter 4.

2.3.3.3 *Surface reactions*

Apart from the requirements of some experiments concerned solely with the physical properties of surfaces, the usual reason for wishing to prepare a surface in an atomically clean state (within the limits of detection) is to establish a reproducible, well-characterized, starting point for the study of the


H-GAS
(a)


Control Valve
Leak Valve
Manifold
To Spectrometer
Glass-to-Metal Seal
Ball Bearing
Breakable Glass Tip
Spec-pure Gas at Atmospheric Pressure
(b)

reactivity of that surface. The reactivity may be towards gases and vapours, either singly or together, towards liquids, towards other solids, or towards impurity or other atoms diffusing to the surface from the bulk. Since gases are in general easy to handle in a vacuum system, it will come as no surprise that by far the greatest number of surface reactions studied have been those between gases and solids.[13] This is not the place to discuss the reactions themselves—only the means to bring them about.

On all commercially available spectrometers a gas admission and control system can be installed as an optional extra. Usually it is separate from the argon feed to the ion bombardment guns. A gas admission system must have the following features:

(a) A series of containers of spectroscopically pure gases, as many as the number of different gases needed for the experiment. These containers consist in some cases of glass flasks filled to atmospheric pressure, in others of metal cans filled to a few atmospheres pressure. In both cases the gas can be released by breaking a seal after the container is attached to the system. Figure 2.6 shows examples of each type.

(b) Fixing points, or stations, for the gas containers, each consisting of a flange and an associated control valve. For glass containers the mating flange carries a glass-to-metal seal ending in a length of glass tubing, to which the tube on top of the container is fused when the container is connected to the station. Pressurized metal can containers normally carry their own valves as an integral part of the seal puncturing mechanism, but because of the high pressures involved it is essential not to do away with the second valve at the fixing point, since additional control will be necessary. Again Figure 2.6 shows the arrangement in each case.

(c) A manifold to which all the stations are connected, which has a volume large enough to allow sensible mixing of one or more of the gases if the experiment requires a gas mixture, but not so large that the gases are used wastefully. Its volume may also be governed by the total pressure of the gas or gases required for the experiment in the reaction chamber, and

---

Figure 2.6 (a) An arrangement for the mixing and admission of pure gases to a spectrometer in a controlled way, using pressurized metal can containers. When the valve on a can is first screwed onto it, a hollow plunger breaks a seal, thereby releasing the gas. However, since the pressure in the cans is about eight atmospheres it is essential to have additional valves between the can and the system as a precautionary measure and for proper control. (By courtesy of Messer Griesheim Ltd) (b) A typical gas-handling arrangement using glass bottles containing one litre of spec-pure gas at one atmosphere pressure. The gas is released by lifting the soft iron slug with a magnet and dropping it onto the breakable tip. Care has to be taken in the process that powdered glass does not get carried into the seats of any valves, and it is usual to place a wad of clean glass wool in the neck of the bottle to prevent this happening

the volume of that chamber. Generally speaking, however, since the gas can be stored in the manifold at a relatively high pressure, e.g. several torr if necessary, and since the pressures used in experiments are rarely greater than 10^{-4} torr, the manifold volume can be kept small. Obviously the pressure in the manifold must be measurable and over quite a large range, from UHV to the torr region, which implies that provision must be made to accommodate two gauges with overlapping ranges.

(d) A high quality leak valve for the admission of gas from the manifold to the reaction chamber. The valve should be one whose design and precision in construction are such that the rate of admission of gas can be controlled finely and smoothly and that the setting for any particular chosen gas flow is reproducible. An example of the construction of such a valve is given in Figure 2.7. Needless to say, the valve should also be capable of shutting completely against a high pressure on the manifold side.

To avoid contamination of the gas or gases admitted to it, the manifold and all the connecting pipework must be fabricated to the same UHV standards as the rest of the spectrometer, so that it can also be subject to the same bake-out procedure. On the other hand, the glass or metal gas containers clearly cannot be baked, which means that between the manifold and the containers

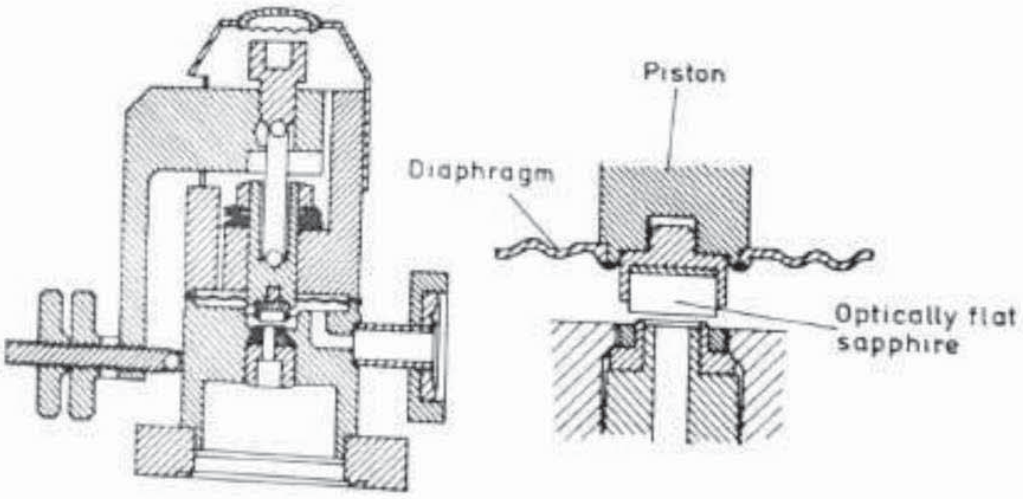


Figure 2.7 Leak valve for fine control of gas flow. The main construction of the valve is shown in the larger diagram and details of the seating mechanism in the smaller one. The valve seat consists of an optically flat sapphire disc pressing against a metal ring. The mechanical advantage of the drive mechanism is many thousand to one, so very fine control of the movement of the sapphire is possible. Provided the valve is not closed too tightly the setting for any particular flow rate is reproducible. (By courtesy of Varian Associates Ltd)

there is a region of temperature gradient during bake-out. Since the contain-
ers are positioned outside the bake-out zone, this region is accessible, and it is
good practice to warm it carefully and judiciously with a hand-held heat gun
once the bake-out temperature has been reached in the rest of the spectrome-
ter in order to drive off any condensable impurities.

Even though a single gas stored in a manifold may be considered pure, or a
gas mixture prepared in the manifold may have an accurately known composi-
tion, there is no guarantee that on admission to the reaction chamber the
purity or the composition will remain unchanged. Exchange can take place
with other gases adsorbed on the walls and other parts of the chamber, and
hot filaments can cause dissociation followed by recombination to form dif-
ferent molecules. If the experiment is a dynamic one, i.e. with the reaction
chamber open to the pumps so that the gases are flowing over the surface,
there is always the risk of other reactions taking place in the pumps them-
selves. This is not particularly likely with diffusion or turbomolecularly
pumped systems, but has certainly been observed in the presence of ion and
sublimation pumps. It is therefore very advisable to have on the reaction
chamber a means of checking the actual, as against the supposed, composition
of the gas being admitted to the clean surface, and this check is normally
performed by a small quadrupole mass spectrometer. Since such a device is
also invaluable for diagnostic purposes, i.e. during a fault condition, and for
leak testing, it can hardly be regarded as a luxury but as a necessity.

Reactions of clean surfaces with liquids have hardly ever been attempted in
UHV; the problems associated with transferring a liquid to a surface in a
controlled way in UHV seem too complex to be worth the effort. In recent
years the study of solid–solid reactions, on the other hand, has been increas-
ing steadily, stimulated by some of the requirements of the microelectronics
industry for the production of surface layers with certain electronic proper-
ties. These experiments are in general quite straightforward, and consist of
the deposition by evaporation of known quantities of one solid, usually a
metal, onto the surface of another solid, sometimes a metal[14] but more often a
semi-conductor.[15] The need for a completely hygienic deposition in terms of
purity and contamination, as discussed earlier, is particularly stringent here
since very small amounts of unwanted elements, either in the deposited film
or at the film/surface interface, can cause large changes in the electronic
properties of the layer. Reaction in some cases (e.g. palladium on silicon)[16]
occurs spontaneously at room temperature; in other cases it is necessary to
programme[17] the temperature of the substrate during the experiment either to
encourage reaction to take place during deposition or to prevent it doing so
until deposition is complete. Thus the basic surface preparation requirements
for the majority of solid–solid reaction studies are a properly designed
evaporator and a heating/cooling stage.

For the study of the adsorption at external surfaces of impurities segregat-

ing from the bulk the only basic surface preparation requirement is a heating stage capable of taking the material to the desired temperatures. If, however, it is the segregation to internal grain boundary surfaces that is to be studied, then the surfaces are prepared by fracture in UHV in one of the fracture stages already described and illustrated in Figure 2.5. In this case all heat treatments of the sample are carried out before insertion into UHV, although it will be necessary in most cases to *cool* the sample before fracture to increase the tendency to brittle fracture.

2.3.4 Sample positioning

There are several reasons why it is necessary to be able to manipulate the position of a sample in a finely controlled manner once it has been inserted into the spectrometer and brought into the proximity of the energy analyser.

Firstly, the shapes and sizes of samples can differ widely, so that the optimum position in terms of sensitivity and energy resolution for the surface analysis of one sample will not necessarily be optimum for the next. Secondly, it may be required during analysis to move the sample laterally in a reproducible way to allow different areas of it to be analysed; this of course applies more to XPS than AES since in the latter the electron beam can be deflected for the analysis of different areas and in any case the average area of analysis in AES is much smaller than in XPS. Thirdly, it is often desirable to be able to rotate the sample about an axis through its surface, again in a reproducible way, either to alter the angle of incidence of primary ions during ion bombardment or to alter the angle of take-off of electrons accepted by the analyser, for the purposes of profiling by variation of escape depth (see Chapters 3 and 4). Fourthly, if techniques other than electron spectroscopy are also available on the spectrometer, e.g. static SIMS that uses mass spectrometry rather than energy detection, it may be necessary for geometric reasons to have to move the sample into a position for analysis different from that for XPS and AES. Fifthly, as the reliability and degree of mechanical automation of surface analytical techniques improve, there will be increasing pressure from users for automatic sample change, which again involves accurate movement and positioning.

Accurate sample positioning is not an especially new requirement, or even one that is specific to XPS and AES. It has always, for example, been needed by LEED (low energy electron diffraction) from the earliest days of the study of surfaces, and many of the same reasons why it is needed by XPS and AES apply equally well to a whole range of other techniques. The considerations listed above suggest that at a minimum four degrees of freedom in movement should be available, viz. X, Y and Z linear motions and an axial rotational motion. For automatic sample change or for assuming analysis positions other than that for XPS and AES it would clearly be desirable, if not necessary, to

have azimuthal rotation as well in which the samples would be mounted on a carousel. The positioning movements actually provided by manufactuers in their *standard* production equipment for surface analysis vary somewhat and are dependent on their philosophy in respect of sample insertion. Where a sample carousel is used to accept a sample on its carrier from the end of a sliding shaft, the sheer bulk of the carousel makes the facility for axial rotation very difficult to incorporate in the positioning degrees of freedom, making any variation of the ion incidence angle or of the electron take-off angle impossible. On the other hand, of course, azimuthal rotation of the carousel lends itself to automatic sample changing and to making any one sample available for analysis by a different technique. Where a sample is inserted by the successive transfer method, it is of course placed finally on the end of the shaft of a sample manipulator, where it can be translated along any of the three orthogonal directions and also be axially rotated. If the manipulator is fitted with a double rotation attachment, usually an optional extra, then azimuthal rotation of rather limited radius can be added to the movements, but the mechanism is not sufficiently precise to be usable in automatic sample changing.

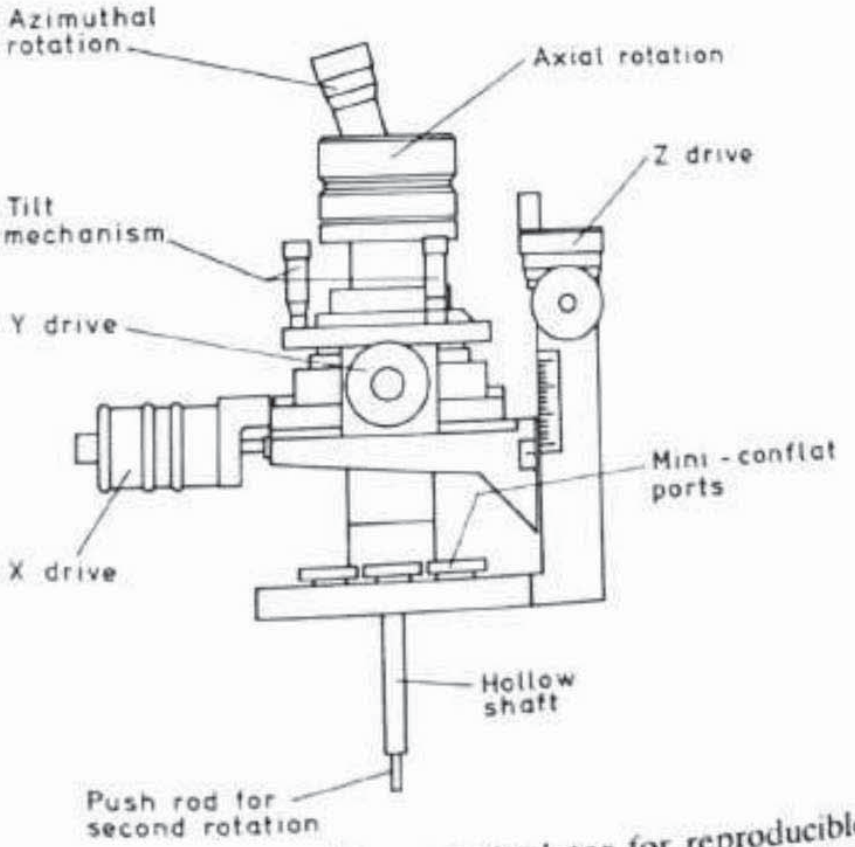


Figure 2.8 Diagram of high precision manipulator for reproducible positioning of specimens in a spectrometer. Movements available are shifts along the x, y and z axes, rotation about the central axis and about an axis at right angles to it (azimuthal rotation), and tilt. Electrical and other feedthroughs can be taken into the vacuum via a series of mini ports set into the main flange, and the whole device is bakeable to 250 °C. (Reproduced by permission of Vacuum Generators Ltd, Hastings)

For experiments of a basic nature, in which the sample is not inserted by either of the two methods already described but is mounted individually, all manufacturers will provide sample manipulators with precise, reproducible movements. An example of one of these is shown in Figure 2.8. The translational movements are taken up in a highly flexible all-welded bellows, whose length can be increased considerably if required to provide for additional Z translation over the 50 mm or so of the standard device. Axial rotation is derived from an offset knuckle joint that also operates through a flexible bellows. If azimuthal rotation is needed, then an additional mechanism is included that drives a cable or rod inside the hollow shaft on a push/pull basis, the cable being attached to a pulley that rotates the sample mount. Earlier versions of such universal manipulators had to be partially dismantled for bake-out, but all current models are of such construction that no dismantling is needed and their precision of movement and reproducibility is unaffected by repeated bake-out up to 250 °C. With most of them it is possible to obtain a variety of attachments for heating, cooling, azimuthal rotation, offset axial rotation, or almost any combination of these, so that their flexibility of application is considerable.

2.4 Sources

2.4.1 X-ray sources

Devices for the production of high fluxes of X-ray photons of characteristic energies were in use long before the advent of XPS, principally in X-ray diffraction (XRD) instruments. For XRD, however, the characteristic energies are normally required to be high, typically many scores of kiloelectronvolts whereas reference to the description of the basics of XPS in Chapter 3 shows that the surface-specific property of XPS depends on using soft X-rays of characteristic energies of only a few kiloelectronvolts. In addition, sources used in XRD operate traditionally by the bombardment of an anode at earth potential by electrons from a hot filament at a very high negative potential, so that any other surrounding surfaces also at ground potential would also be bombarded. Since in XPS the X-ray photons have to pass from the vacuum in the source to the vacuum in the analyser chamber through a thin window which acts as a barrier to electrons and to possible contamination, it is clear that if the traditional configuration were to be maintained the window would also receive electron bombardment. Thus in XPS sources the configuration is reversed in that the filament is near earth potential and the anode at a high *positive* potential. Figure 2.9 shows two commonly used configurations.

When one comes to consider the choice of material for the anode for an XPS source, it is necessary to consider first the matter of energy resolution.

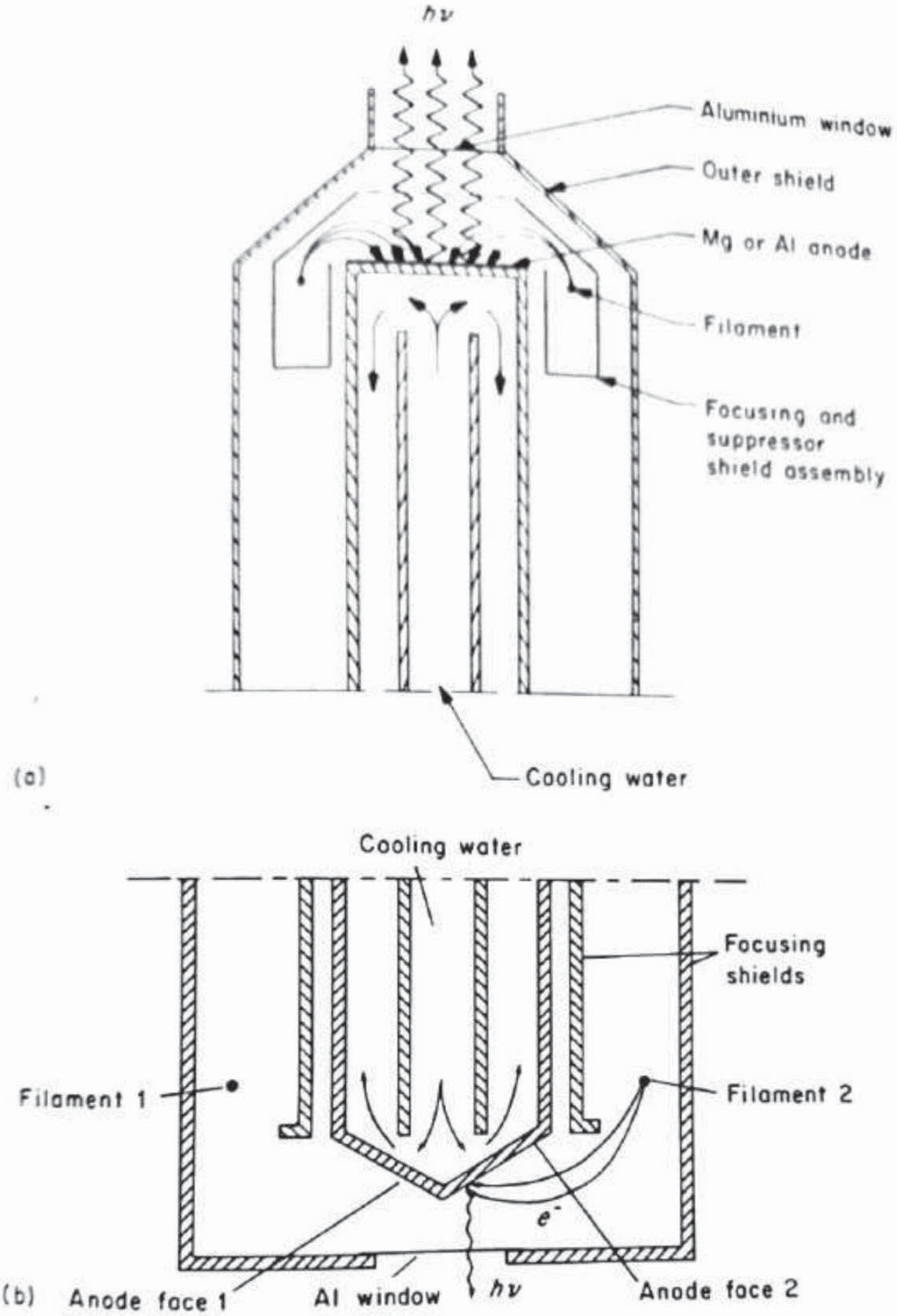


Figure 2.9 (a) Soft X-ray source with single anode of either magnesium or aluminium, deposited as a thick film on the flat end of a water-cooled copper block. The anode is surrounded by a cylindrical focusing shield at the same potential as the filament. An outer can acts as a radiation shield and carries the thin aluminium window that must be interposed between the target and sample. (Reproduced by permission of Perkin-Elmer, Physical Electronics Division) (b) Soft X-ray source with dual anode, allowing use of either magnesium or aluminium $K\alpha$ radiation by simple external switching without the need to break the vacuum in going from one to the other. The anode has a tapered end with two inclined faces on which films of magnesium and aluminium, respectively, are deposited. There are two semi-circular filaments, one for each face. The focusing arrangements are similar to those for the single anode of (a). (Reproduced from Barrie and Street[18] by permission of The Institute of Physics)

The Einstein relation governing the interaction of a photon with a core level,

$$E_{KE} = h\nu - E_B - e\phi \tag{2.1}$$

where E_{KE} = kinetic energy of ejected photo-electron
$h\nu$ = characteristic energy of incident X-ray photon
E_B = binding energy of core level electron
ϕ = work function term,

shows that the line width of E_{KE} will depend, amongst other factors, on the line width of $h\nu$. (Line width means here the full width at half-maximum height of the exciting X-ray lines or emitted photo-electron lines, as the case may be.) Since in XPS one is continually seeking chemical information by detailed analysis of individual elemental spectra, it follows that the most accurate information will be provided by working at the best instrumental energy resolution compatible with sensitivity, i.e. signal-to-noise, considerations. In practice, at the time of writing, this means using energy resolutions in the range 1.0–2.0 eV, mostly nearer the lower than the higher figure. Clearly, then, to avoid limitation of the achievable resolution by the line width of the X-ray source, it is necessary to use materials for the source anode whose line width is less than 1.0 eV. Table 2.1 lists X-ray line energies and line widths for various characteristic lines from some materials that would be suitable from other points of view, e.g. stability under prolonged electron bombardment. It can be seen that there are indeed very few materials whose

Table 2.1 Energies and widths of some characteristic soft X-ray lines

Line	Energy, eV	Width, eV
Y $M\zeta$	132.3	0.47
Zr $M\zeta$	151.4	0.77
Nb $M\zeta$	171.4	1.21
Mo $M\zeta$	192.3	1.53
Ti $L\alpha$	395.3	3.0
Cr $L\alpha$	572.8	3.0
Ni $L\alpha$	851.5	2.5
Cu $L\alpha$	929.7	3.8
Mg $K\alpha$	1253.6	0.7
Al $K\alpha$	1486.6	0.85
Si $K\alpha$	1739.5	1.0
Y $L\alpha$	1922.6	1.5
Zr $L\alpha$	2042.4	1.7
Ti $K\alpha$	4510.0	2.0
Cr $K\alpha$	5417.0	2.1
Cu $K\alpha$	8048.0	2.6

characteristic X-ray lines have sufficiently small widths. The yttrium and zirconium $M\zeta$ lines are narrow and have been used successfully in special applications, but their energies are much too low for general application since the number of photo-electron lines that could be excited would be insufficient for unambiguous analysis. The only other lines that fit the requirements are the $K\alpha$ lines of magnesium, aluminium and silicon. Although silicon has also been used[19] for special applications, it is not a metal, so that its heat transfer characteristics are poor and it is difficult to apply to an anode surface. Thus one is left with the magnesium and aluminium $K\alpha$ lines, and in fact it is these two that are used so universally in XPS.

Although the characteristic $K\alpha$ line energies of magnesium and aluminium are in the region of 1250–1500 eV, it is necessary to use exciting electron energies about an order of magnitude higher for efficient production of X-rays, according to the dependence[20] of emitted flux on bombarding energy shown in Figure 2.10. In all commercial spectrometers the maximum available accelerating potential is 15 kV, which is adequate. It is also necessary, for purposes of optimizing sensitivity at a given energy resolution, to be able to use as high an electron bombarding current as the source will stand since the photon flux will be directly proportional to that current. Against this design requirement must be set the necessity of making the physical size of the source near the sample analysing position as small as possible since the source must be placed as near as possible to the sample. The flux irradiating the sample will of course vary as the inverse square of the distance of the anode from the sample surface. All these requirements have led designers inexorably to the types of configuration shown in Figure 2.9, in which the necessary compactness of the source means that the maximum power dissipation that can be achieved is 1 kW, provided that adequate water cooling can be provided. The necessity of forced water cooling to remove heat from the anode before it melts also implies that the anode block must be of high heat conductance, which in turn means fabrication of the block and the integral water tubes from copper. Thus the anode material itself, i.e. the emitting surface, is normally deposited on the copper block as a thick film, typically ~10 μm in thickness, representing a compromise between being thick enough to exclude copper $L\alpha$ radiation and thin enough to allow adequate heat transfer. In most source control units the bombarding current can be selected from a set of switched fixed values, e.g. 5 mA, 10 mA, 20 mA, etc., the precise values varying with the manufacturer. The accelerating voltage, on the other hand, can be chosen anywhere in the continuous range from 0–15 kV, although in general the source will be unstable below ~2 kV.

Most commercially available X-ray sources are of the type shown in Figure 2.9(b), in which there are *two* filaments for electron bombardment and *two* anode surfaces, one of magnesium and one of aluminium. It is thus possible by simple external switching to go from Mg $K\alpha$ to Al $K\alpha$ radiation in a few

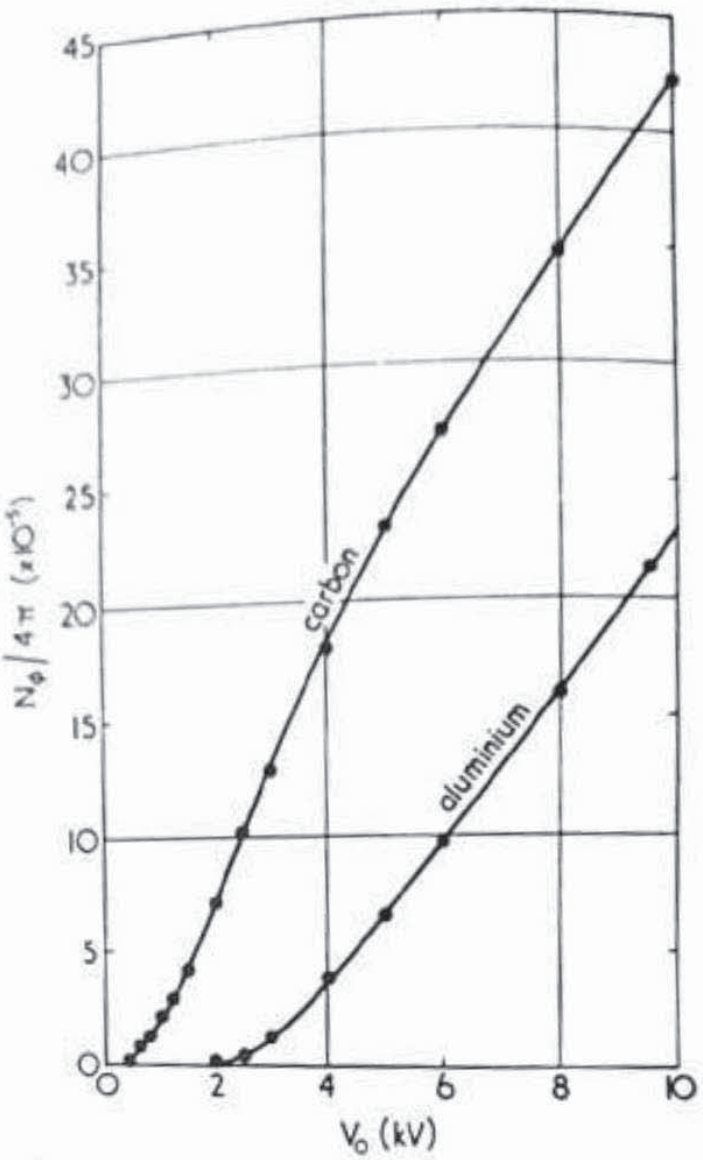


Figure 2.10 Dependence of efficiency of production of Al $K\alpha$ and C $K\alpha$ characteristic radiation on the energy of the bombarding electrons. (Reproduced from Dolby[20] by permission of The Institute of Physics)

seconds. There are two reasons why it is desirable to have the double-anode facility. One is that the two characteristic $K\alpha$ lines have different line widths, as can be seen by reference to Table 2.1, and although the Al $K\alpha$ line is of slightly greater general usefulness than that of magnesium, there are frequent occasions when the inherently better resolution obtained with the magnesium line is needed. The other reason is that in any X-ray-excited ejected electron spectrum both photo-electron and Auger peaks appear, and interferences can result. Since photo-electron energies are linked directly to the energy of the exciting photon, whereas Auger energies are fixed, a change in the X-ray line energy will resolve the interference. For special purposes, other combinations of anode material have been prepared, amongst which Mg/Zr and Al/Zr are popular, the Zr $L\alpha$ radiation being useful[21] when increased sensitivity for the detection of aluminium and silicon is needed.

As mentioned earlier, it is necessary to interpose a thin window between the anode and the sample to screen the latter from stray electrons, from heating effects and from any possible contamination originating in the source region. The material chosen for the window should obviously be reasonably transparent to the X-radiation being used, and for the Mg $K\alpha$ and Al $K\alpha$ lines it is both convenient and practicable to use aluminium foil about 2 μm thick. For that thickness the flux attenuation is about 24 per cent for Mg $K\alpha$ and 15 per cent for Al $K\alpha$. When other exciting radiation is used care must be taken that the appropriate window material is substituted for the standard aluminium. For example, if the $M\zeta$ radiation at 151.4 eV from a zirconium source were required, aluminium would be opaque and instead either beryllium or carbon windows would be suitable.

The description so far of X-ray sources for XPS refers for the most part to 'natural' or, more accurately, unmonochromatized radiation. It must be realized that the emission spectrum from *any* material is complex and consists of a broad continuous background (called Bremsstrahlung) on which are situated the more or less narrow characteristic lines. An example is shown in Figure 2.11 for aluminium. The Bremsstrahlung continuum is a function of the energy of the bombarding electrons, and after passing through the window will have a maximum between 20 and 40 per cent of that energy, according to the thickness of the window, so that it extends beyond the energy of the principal characteristic line and is therefore useful for exciting Auger transitions from atomic levels too deep to be ionized directly by the characteristic radiation. Complexities in the line spectrum arise from satellite emission associated with the principal lines and also from the multiplicity of lines as the Periodic Table is ascended, as discussed in Chapter 3.

2.4.1.1 X-ray monochromatisation

Removal of satellite interference, improvement of signal-to-background by eliminating the Bremsstrahlung continuum and selection of an individual line from the unresolved principal line doublet can be achieved by monochromatization of the emitted X-rays. There are several methods of doing this but they all depend on dispersion of X-ray energies by diffraction in a crystal, which is of course governed by the well-known Bragg relation:

$$n\lambda = 2d \sin \theta \tag{2.2}$$

where n = diffraction order
λ = wavelength of X-rays
d = crystal spacing
θ = Bragg angle

For first-order diffraction of Al $K\alpha$ X-rays, for which $\lambda = 0.83$ nm, it is found that quartz crystals are very suitable, since the crystal spacing of the

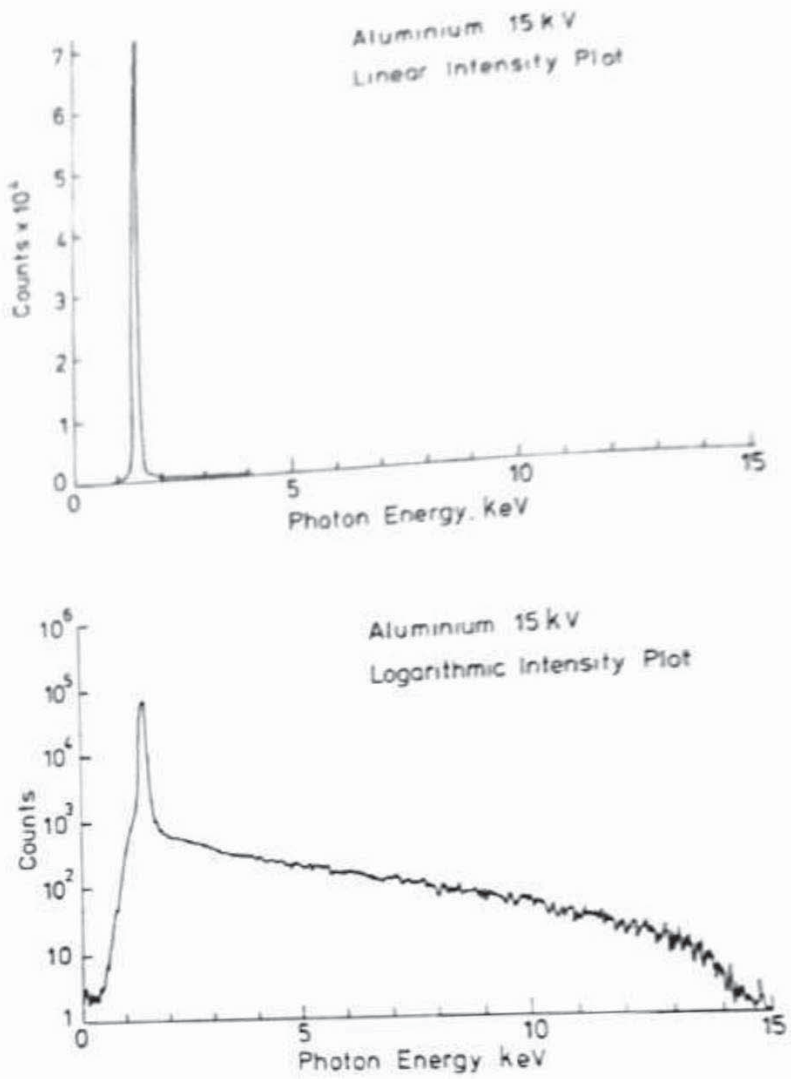


Figure 2.11 X-ray emission spectrum of an aluminium target under bombardment by 15 kV electrons, recorded by a lithium-drift detector through a beryllium window of thickness 7.5 μm. Upper curve, photon intensity plotted on a linear scale, on which little is evident except the intense characteristic $K\alpha$ line. Note that the energy broadening of the solid-state detector attenuates the peak by a factor of about 100. Lower curve, the same plotted on a logarithmic scale, that reveals more clearly the broad Bremsstrahlung background extending to energies much higher than the characteristic line. The background intensity at very low energies will have been reduced by absorption in the beryllium window. (Measurements by courtesy of Mr R. W. M. Hawes, Materials Development Division, Harwell)

$10\bar{1}0$ planes is 0.425 nm and the Bragg angle is thus 78.5°. Quartz has many advantages, since it can be obtained in perfect crystals of very large size which can easily be bent elastically or ground, and can be baked to high temperatures without damage or distortion.

The principle of the method of monochromatization used most estensively

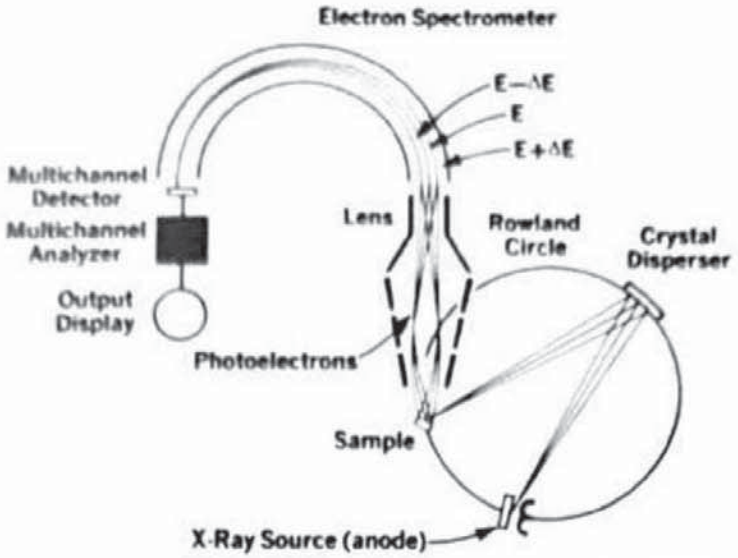


Figure 2.12 Principle of monochromatization of X-radiation. Selection of a particular wavelength is achieved by diffraction in a crystal of suitable lattice spacing, and for the Al $K\alpha$ lines quartz is a convenient crystal to use. The quartz crystal is placed on the surface of a Rowland circle, being accurately ground to match the curvature, and the X-ray source is placed at another point on the circle. Photons of the required wavelength will be focused at a third point on the circle, where the sample is placed. For the Al $K\alpha_1$ line the angle between incident and diffracted beams is 23°. A typical circle diameter would be 0.5 m. (From Kelly and Tyler[22])

in commercial instruments is that of fine focusing,[22] shown in Figure 2.12. The quartz crystal is placed on the surface of a Rowland or focusing sphere, having been ground accurately to the shape of the sphere, i.e. in both directions of curvature. With the anode itself lying on the Rowland circle, irradiated with a sharply focused electron beam, the $K\alpha$ X-rays are dispersed by diffraction and re-focused at another point on the Rowland circle where the sample is placed. If the $K\alpha_1$ component of the principal Al $K\alpha$ line is required, then the included angle between incident and diffracted rays will be 23°. According to the diameter chosen for the Rowland circle the dispersion around the circumference will vary; for the popular 0.5 m monochromator the dispersion is 1.6 mm/eV and for the larger 1 m instrument, 3.2 mm/eV. To achieve, therefore, line widths substantially less than the natural widths of the $K\alpha$ radiation, e.g. ~0.4 eV, involves focusing of the bombarding electron beam on the anode sufficiently finely that the width of the irradiated area on the sample in the dispersion plane is less than ~0.6 mm, using the 0.5 m monochromator.

Since the function of a monochromator is to allow only a selected small portion of the total $K\alpha$ emission to fall on the sample, it follows that the photon flux available for the same power dissipation is considerably less than with the unmonochromatized sources. For an aluminium anode, for example,

the flux will be lower by a factor of about 40 when using a l m monochromator compared to that from a standard source. Thus count rates will be much reduced, but against that must be set increased detectability in terms of signal-to-background since the background has been substantially removed, the absence of satellites resulting in a 'clean' spectrum, and excellent resolution. Most manufacturers can provide monochromators as optional extras, but they are inevitably both bulky and expensive. Very recently one manufacturer has also described the monochromatisation of Ag $K\alpha$ radiation.[23]

2.4.1.2 *Synchrotron radiation*

The soft X-ray sources discussed up to now have all been discrete line sources. Useful as these are, they have inherent limitations as far as XPS is concerned, the most obvious one being that of photo-ionization cross-section, since for any fixed line source the cross-section for one group of core levels may be near maximum, while for another group it may be near minimum. All that can be done to change the cross-sections in conventional equipment is to go to another fixed line source, which would not necessarily be of the right energy, and in any case the number of available line sources of any use has been seen to be very limited. Ideally one would like to have a continuously tunable source of high intensity at all the required energies. A close approximation to such a source exists in the form of synchrotron radiation.

It has been known for a long time that when a charged body is accelerated it emits radiation; the process is particularly efficient for an electron because of its very small mass. This effect is used in a synchrotron in which electrons are accelerated around an approximately circular torus (or 'doughnut') under the influence of pulsed magnetic fields. As they are accelerated around the curved path, the electrons emit light in a continuous spectrum whose intensity maximum is proportional to the radius of curvature and inversely proportional to the cube of the electron energy. Acceleration is taken to near-relativistic velocities, when the emitted radiation is concentrated in a very narrow cone tangential to the electron orbits, with an angular divergence of the order of 0.1 mrad (0.006°). It is therefore ideal for passing into a monochromator for energy selection. The general shape of the radiation spectrum of an electron moving in a curved orbit[24] is shown in Figure 2.13. Typically λ_c, the critical wavelength at the intensity maximum, is of the order of 0.1 to 0.4 nm, and the maximum energy would be between 1 and 10 keV. However, Figure 2.13 shows that towards longer wavelengths, i.e. lower energies, the intensity falls off quite slowly, so that there is still plenty of usable intensity at a few tens of electronvolts.

Typically a synchrotron radiation source would have a radius of many metres and the electron beam would have an energy of several GeV and current up to 1A, so it is not a device to be purchased ready-made from a

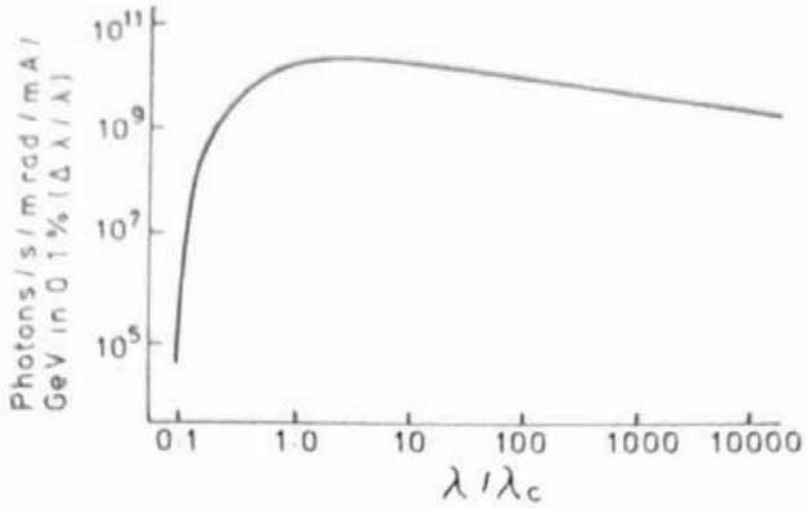


Figure 2.13 Shape of the radiation spectrum of an electron travelling in a curved orbit. The vertical scale of photon flux is a function only of the electron energy and current, while the horizontal scale is defined by λ_c, the so-called critical wavelength. If the maximum intensity is required to be at an energy near 1000 eV, then λ_c should be 1 °A or less. To achieve that the orbiting electrons must be accelerated to several gigaelectronvolts, and the radius of the orbit should be of the order of a few metres. (After Farge and Duke[24])

manufacturer and installed in one's laboratory. They are so large that they can only be nationally owned and would-be users have either to take their spectrometers to one or use the spectrometers available on site.

Despite the continuous nature of the spectrum of the emitted radiation, the synchrotron is not yet quite as continuously tunable as one would like because selection of the required energy or range of energies has to be performed by monochromators. So far monochromator design has enabled the energy region from about 20 eV to about 500 eV to be tunable continuously, but there is still quite a big gap to the higher energies of the discrete line sources. It is expected that this gap will be filled, but design constraints include operation in UHV and bake-out to reduce surface contamination of reflecting surfaces in the monochromator. Pressure to extend the range to 1000 eV or more comes from the fact that the X-ray photon flux from a typical synchrotron in that part of the spectrum is at least two orders of magnitude higher than from a conventional Al $K\alpha$ source. Even with losses in the monochromator, the flux will still be greater than available now.

2.4.2 Electron sources

In conventional AES the only function of the incident electron beam is to produce ionization in core levels in order to initiate the Auger transitions. For that purpose alone the energy in the beam is unimportant. If, however, elec-

tron loss spectra are to be studied then the line width, by analogy with X-ray sources, should be minimized. Without energy selection the width will be limited by the thermal spread, according to the Boltzmann distribution, to 0.3–0.5 eV. Such a width would be quite unacceptable in low energy electron loss spectroscopy (LEELS) in which vibrational spectra are studied, but a discussion of LEELS is outside the scope of this volume. For AES, then, no steps need be taken to reduce the primary beam width.

The two types of electron source used in AES are thermionic and field emission, and of these the former is by far the most common. All materials possess a work function, which is simply the energy necessary to remove an electron completely from the material. Work functions are usually[25] of several electronvolts magnitude and thus represent a barrier at the surface to the escape of electrons; at room temperature, and without any applied electric fields, electrons cannot escape and there is no emission. The two types of source use different physical methods to allow sufficient numbers of electrons to overcome the barrier and thus produce emission from the material. In the thermionic method the material is simply heated to a temperature high enough to give some of the electrons sufficient additional energy to pass over the barrier into the vacuum, so that the higher the temperature the greater is the thermionic current. The emission is governed by the Richardson equation:

$$J = A(1 - \bar{r})T^2 \exp(-e\phi/kT) \qquad \text{A/cm}^2 \tag{2.3}$$

where J = current density
A = Richardson constant, ideally equal to 120 A/(cm^2 deg^2) but varies with material
$\bar{r}$ = zero-field reflection coefficient for incident electrons
T = absolute temperature
ϕ = work function

Field emission operates rather differently. Basically it does not set out to give the electrons in the material additional energy, but to reduce the height of the work function barrier so that even at room temperature electrons can escape. Since most metallic work functions are between 4 and 5 eV[25] and the work function barrier extends over only a few nanometres, it can be calculated at once that the field at the surface due to the barrier is very large. To attempt to eliminate it completely by the application of an equal and opposite field would be quite impracticable, but that is not necessary since the application of even a moderate field will cause sufficient reduction in the barrier to produce adequate emission. This is because electrons have a finite probability of passing through the barrier by 'tunnelling'; the probability is too low at room temperature and under field-free conditions to be measurable, but when the height and shape of the barrier are altered sufficiently by the appli-

cation of a high electric field, then the probability becomes useful. To achieve the high field required the material is fashioned into a sharp point with a tip radius of typically 50nm, and a positive electrostatic field of many kilovolts is applied between it and an extraction electrode. The process is called field emission and is governed by the Fowler–Nordheim expression:

$$J = (1.55 \times 10^{-6} E^2/\phi)\exp[-6.86 \times 10^7 \phi^{3/2}\theta(x)/E] \quad \text{A/cm}^2 \quad (2.4)$$

where J = current density
E = field strength
$\theta(x)$ = Nordheim elliptic function
$x = 3.62 \times 10^{-4} E^{1/2}/\phi$

It will be noted from the expression (2.4) that the current density in field emission is strongly dependent on the value of the work function of the emitting surface, so that it is standard practice to use for the emitter a needle-shaped single crystal oriented with a crystal plane of low work function at the tip. Since the adsorption of residual gases in a vacuum system onto a clean metal surface almost always increases the work function, it is necessary, if the current density is not to decrease as a result, to maintain a particularly good level of UHV around the emitter. Before use any accumulated adsorption is removed by 'flashing' the emitter, i.e. raising its temperature briefly to a very high value. Because of the flashing necessity, and also because the emitter material must withstand the high electrostatic field without disintegrating, tungsten is always used—oriented so that a low work function plane such as the (310) is parallel to the tip.

The two types of source, thermionic and field emission, each have their respective advantages and disadvantages. Because the emitting area of a field emission source is so small and the emission is concentrated in a small angle, the current density per unit solid angle obtainable from it is much higher than from a thermionic source, and for this reason field emission sources are known as 'high brightness' sources. Figures quoted[26] are in excess of 10^7 A/(cm^2 sr) compared to a typical figure of 10^4 A/(cm^2 sr) for a tungsten thermionic filament. By using an emitter of lower work function, higher brightnesses can be obtained[26] for thermionic emitters, e.g. of the order of 10^6A/(cm^2 sr) for an LaB_6 crystal, but against these high brightness figures must be set both the greater difficulty of fabricating both field emission and LaB_6 sources and their variable reliability. A conventional tungsten thermionic filament may have a lower brightness, but it is relatively easy to make and to replace, its emission properties are entirely reproducible, it can be used over a wide range of emission currents and it is robust. With progressively improving technology, the emphasis will be placed more and more on the higher brightness sources, but at the time of writing the great majority of electron sources still use tungsten filaments in the thermionic mode.

The principal direction of development of electron sources for AES has always been, and still is, towards the production of ever-smaller irradiated areas, or spot sizes, on the sample from electron guns whose construction must be compatible with UHV. Until recently the latter requirement has meant that the electron optics of a gun have been entirely electrostatic, since there were difficulties in designing and fabricating the pole pieces and coils for electromagnetic optics that would withstand bake-out for UHV. However, these difficulties have been largely overcome in the high spatial resolution AES equipment available in the last few years, as will be seen below.

In purely electrostatic guns using tungsten thermionic filaments, the available spot size has decreased progressively from the original 400 μm at -20 μA and 2.5 keV obtained with a converted oscilloscope gun, to 40 μm at -5 μA and 5 keV in the guns used in the mid 1970s, to 5 μm at -10^{-7} A and 5 keV in the late 1970s and to 0.5 μm at -10^{-8} A and 10 keV typical of the present. With particularly careful alignment of filament and optics, with accurate fabrication of apertures and with attention to the elimination of stray magnetic fields, it is possible to achieve a spot size of $\sim$0.2 μm in current electrostatic guns, but for general purposes 0.5 μm is more usual. It is probable that for a simple electrostatic system operating at a voltage that does not need bulky insulators or expensive high-voltage supplies, the performance has now been virtually optimized.

One of the commercial designs for such an electrostatic gun is shown schematically in Figure 2.14. The assembly carrying the V-shaped tungsten filament and the first electrode, the Wehnelt cylinder, is mounted on a flexible bellows and has alignment screws to align the filament and Wehnelt aperture with the rest of the optics. Alignment of the filament itself with respect to the Wehnelt is achieved by a jigging arrangement, but it is difficult to make such an arrangement completely reproducible and optimum alignment is still a matter of experience.

Focusing and shaping of the electron beam is performed by two condenser lenses and an objective lens, and in addition any astigmatism (focusing in the X and Y directions not coincident) can be corrected by a set of stigmator poles before the objective lens. Finally, scanning of the beam is carried out by X and Y scanning poles at the exit from the gun. The gun is pumped through the final aperture and since that is small, with a consequently low pumping speed to the body of the gun, and particularly to the filament region, care has to be taken in outgassing a new filament so that an excessive pressure rise does not occur around it.

As remarked above, for the achievement of much smaller spot sizes, a change has to be made from electrostatic to electromagnetic optics. It is also necessary to operate at higher incident energies, in the range of 10–30 keV, since for electron optical reasons smaller spot sizes can be reached only if the energy in the electron beam is increased. For these reasons, reduction of a

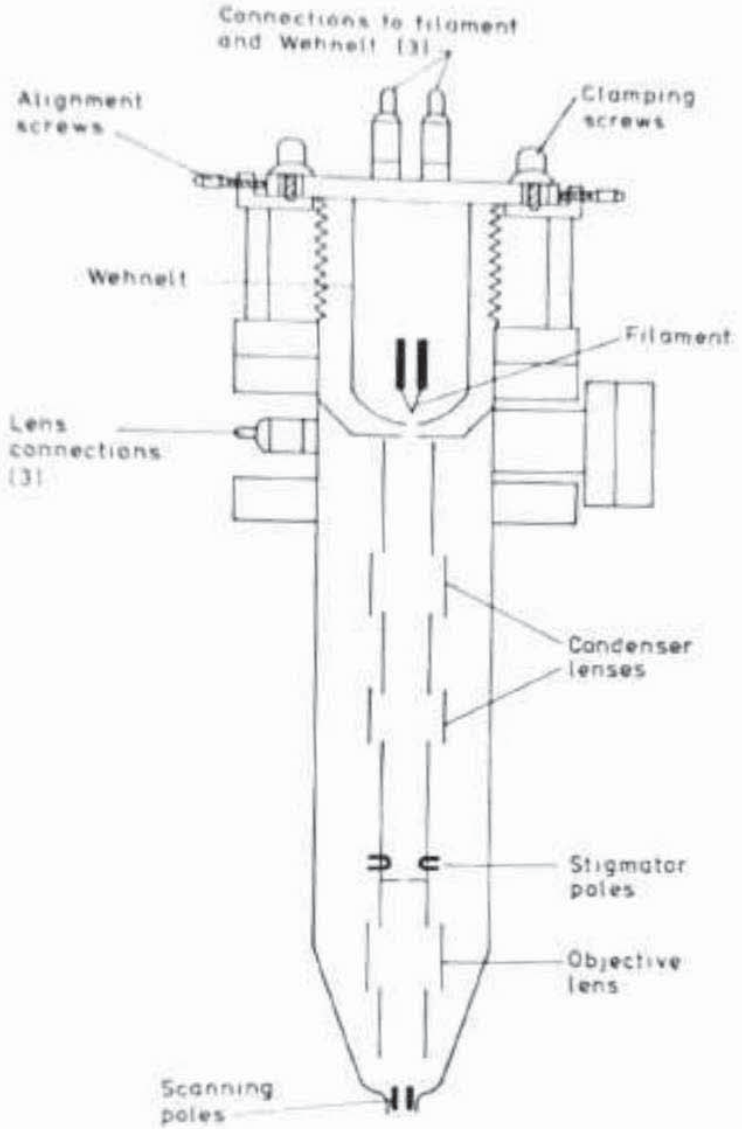


Figure 2.14 Electrostatically focused electron gun designed to produce beams of electrons of energies up to 10 keV. At beam currents between 10^{-9} and 10^{-8}A the optimum spot size on the specimen is about 0.5 μm. The filament, which is a tungsten hairpin, can be aligned with the Wehnelt cylinder by a jigging arrangement, while alignment with the optical axis of the gun is achieved by adjustment of the whole assembly via screws and a flexible bellows. (Reproduced by permission of Vacuum Generators Scientific Ltd, East Grinstead)

further order of magnitude in the spot size, from ~0.5 μm, involves a corresponding increase in the complexity of the electron gun and in the cost. One has to be quite sure that the additional information to be gained is worth the effort!

Just as for electrostatic guns, various designs of electromagnetic gun have appeared, but they are of necessity rather similar in basic layout. An example is given in Figure 2.15. There are two electromagnetic condenser lenses for 'squeezing' the electron beam to a small diameter, and an electromagnetic

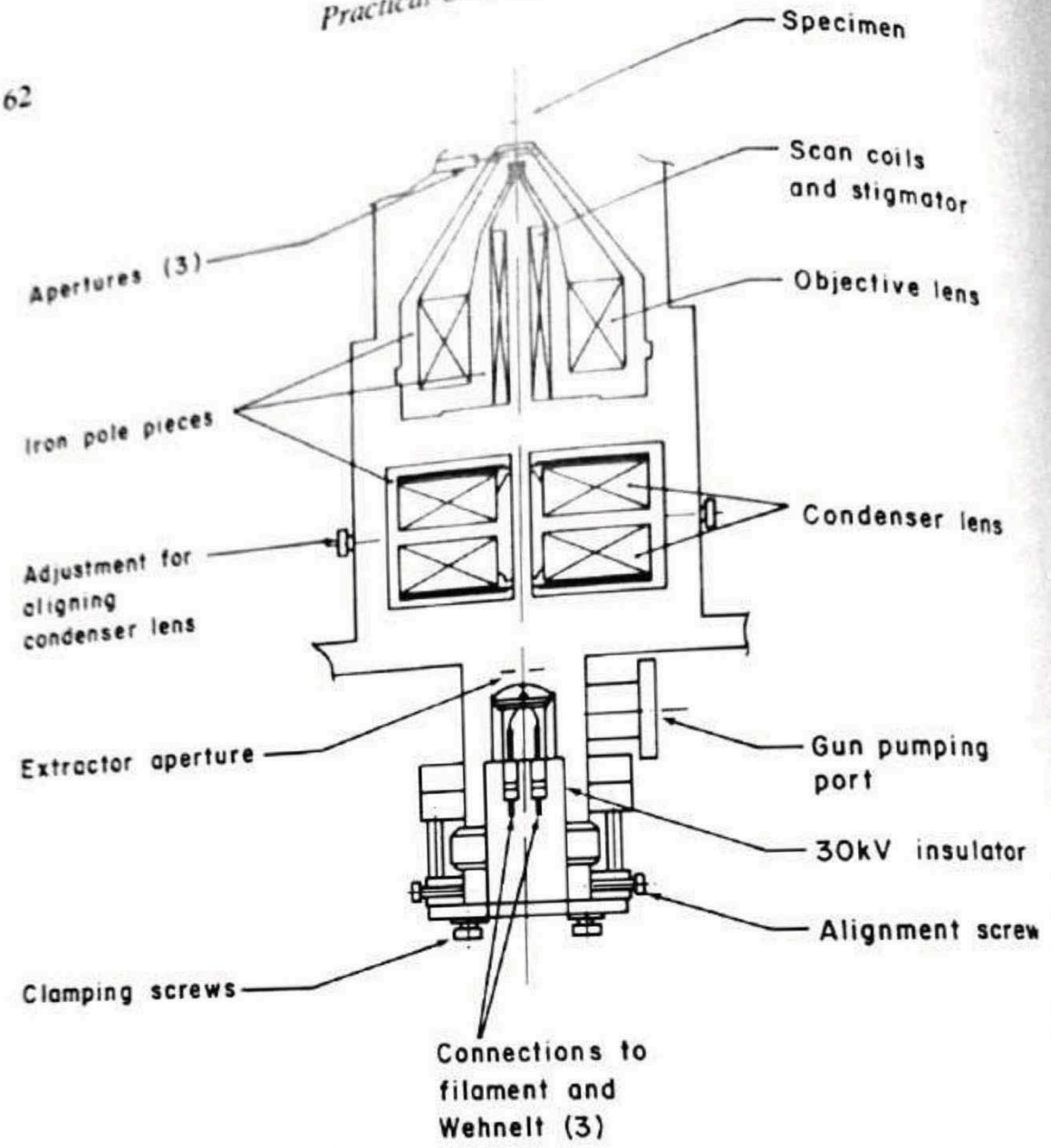


Figure 2.15 Electromagnetically focused electron gun designed to produce beams of electrons of energies up to 30 keV. At beam currents of a few nanoamperes the optimum spot size on the specimen is about 50 nm. As for the gun illustrated in Figure 2.14, the filament assembly can be aligned on the axis of the gun via alignment screws operating through a flexible bellows. Since the volume of the gun is much greater than that in Figure 2.14, a port is provided for separate pumping. The pole pieces and windings are bakeable to 160 °C. (Reproduced by permission of Vacuum Generators Scientific Ltd, East Grinstead)

objective lens for focusing. The windings for the lens are all totally enclosed, thus reducing the problems resulting from bake-out, although because of the lenses the maximum bake-out temperature has to be restricted to 160 °C. Filament mounting is similar to that in the electrostatic gun, and the assembly can be aligned as before with alignment screws operating on a bellows. Since the maximum voltage of operation is now 30 kV, however, the whole assembly is necessarily bulkier because of the additional insulation requirements.

Further alignment of the electron beam with the electron optical axis of the gun can be performed by adjusting the position of the condenser lenses themselves with set screws. Again the corrections for astigmatism and the scanning of the beam are carried out before the beam passes through the final aperture, and in this case the stigmator and the scanning coils are built into the objective lens, which is itself shaped in such a way as to allow reasonable access to the specimen. Instead of a fixed final aperture there is a sliding plate perforated with three apertures of different sizes, any of which can be selected by moving the plate into a predetermined position. Because of the sheer volume of the gun and the very low pumping speed through any of the final apertures, the gun is in this case separately pumped through a side port.

In the long run, the ability to achieve and to maintain over a useful length of time a spot size of the order of a few tens of nanometres depends not so much on the basic design, which tends inevitably to be fixed, but on the precision with which the lenses can be fabricated, particularly on the machining and alignment of the iron pole pieces. Also, it is important to construct the gun with very good mechanical stability and to be able to eliminate stray alternating magnetic fields completely from the gun and specimen region. The differences in performance between different manufacturers' guns will depend principally on those factors rather than on the possibility of producing a radically new design.

Field emission sources have also been used in electromagnet electron optics of the type illustrated in Figure 2.15, although rarely. The combination of the field emitter operated at high voltages (20-60 kV) and electromagnetic optics has pushed the progressive reduction in spot size to its present minimum of around 20 nm, but as yet very few spectrometers of that type are used in AES. However, field emission sources have been used more often with electrostatic optics, and this combination has achieved spot sizes comparable to those obtained with the thermionic source used with electromagnetic optics. A schematic example[27] is shown in Figure 2.16, along with the dimensions, from which it can be seen that the overall size is not large. The performance of this particular gun is that at a sample current of 10^{-9} A and a beam energy of 21 keV the spot size is 50 nm, at a working distance of 55 mm. An additional feature is the filament F located close to the first anode A1 used for outgassing the latter; one of the limitations in operation was found to be noise generated at the field emitter tip by ions released from the surface of A1 by electron bombardment.

Direct comparisons of the three emitters, viz. thermionic tungsten, LaB_6 and field emitting tungsten, indicate[26] that at beam currents above about 10^{-8} A, the LaB_6 source is superior to both thermionic and field emitting tungsten from the points of view of the spot size obtainable and the signal-to-noise ratio. At beam currents below 5×10^{-9} A, however, the field emission source is superior on both counts.

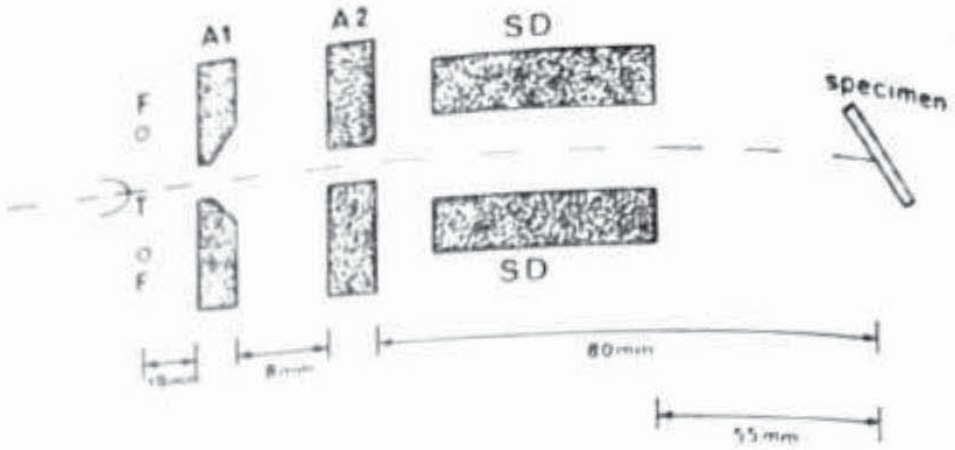


Figure 2.16 Schematic diagram of an electron gun using a field emission electron source. The field emitting tip T is that of a tungsten single crystal wire oriented [310] along its axis. F is a filament located close to the first anode A1, intended for outgassing the latter. A2 is the second anode and SD an eight-pole stigmator/deflector. At a beam energy of 21 keV and current of 10^{-9} A the gun focuses to 50 nm spot size at a distance of 55 mm from the SD. (After Prutton *et al.*[27])

2.4.3 Ion sources

Although it is certainly possible to excite Auger emission with primary ions, and there are several groups of workers whose principal study is ion-excited AES, the intention in this section is to discuss the devices used for ion erosion of surfaces, as described in Section 2.3.3. There are now several different types of ion source available which have been designed with surface cleaning and ion profiling in mind, each with its own characteristics and mode of operation, and the choice of type in any one situation is still partly a matter of personal preference and partly what is offered by the supplier of one's spectrometer.

The simplest type, now not used very much, is rather similar in construction and operation to a standard ionization gauge, and is shown schematically in Figure 2.17. As in the gauge a hot filament provides electron emission to a positively biased grid, the electrons making several oscillations within the grid space before being collected. Argon to pressures between 10^{-5} and 10^{-4} torr is admitted to the system, and the argon gas in the grid space is ionized. Unlike the ionization gauge the ion gun has no positive ion collector, but instead the positive argon ions are accelerated from the grid space towards the cylindrical electrode surrounding the filament and grid through a negative potential gradient of about 200 V. They pass through meshed apertures in that electrode and in an outer earthed screening can, the apertures being offset from the gun axis to avoid possible contamination effects from the hot filament. The entire assembly of filament, grid and accelerating electrode can be

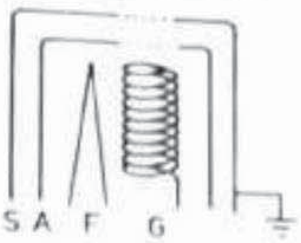


Figure 2.17 Design of an ion gun based on an ionization gauge geometry. The hot filament F emits electrons to a positively biased grid G, within which the electrons make several oscillations before being collected. With argon admitted to the gun at pressures between 10^{-5} and 10^{-4} torr, argon ions are produced by ionization in the grid space. The positive ions are then accelerated from the grid space towards a cylindrical electrode A surrounding the filament and grid, through a potential gradient of about 200 V. The ions pass out into the system through meshes in A and S, an outer screening can at earth potential. The mesh apertures are offset to avoid any contamination from the filament striking the specimen. By floating the filament–grid–accelerator with respect to earth, the ion energy can be varied from 200 to 500 eV. At an argon pressure of 10^{-4} torr, the ion current density on the specimen is about 1 μA/cm^2

floated on a positive potential with respect to the screen so that the energies of the emerging ions in the beam can be varied from 200 to 500 eV.

The 'ion gauge' type of ion source is not very powerful, producing a maximum ion current density of only about 1 μA/cm^2 at the maximum working pressure of 10^{-4} torr. Its advantages are cheapness, reliability and relative purity of ion beam, since few energetic neutral atoms are produced in the ionizing region.

Probably the most popular ion sources in current use employ gas discharges to produce ionization of the gas in the gun. The discharge can be initiated either by Penning discharge or by a high-voltage discharge. In the former the combined action of magnetic and electrostatic fields on argon at a high pressure in a confined space produces a cold-cathode discharge from which the positive ions are extracted. A typical arrangement is illustrated diagrammatically in Figure 2.18. Argon gas is piped directly into the back of the gun and thence to the discharge chamber rather than being present from back-filling of the entire analysis chamber, as happens with the 'ion gauge' type. After extraction the positive ion beam can be focused to a greater or lesser extent, depending on the sophistication of the lens system used. The simpler guns would be incapable of producing an ion spot size of less than 3 mm diameter; the more complex ones can focus to diameters of about 0.1 mm if the sample is at the correct distance from the final aperture. With the finer focused beams there are many advantages to be gained by rastering the beam on the surface, and so X and Y deflection plates are added after the focusing lens system. Guns of this type operate at accelerating voltages between 2 and 10 kV and, since the pumping speed through the gun to the discharge region is low, at

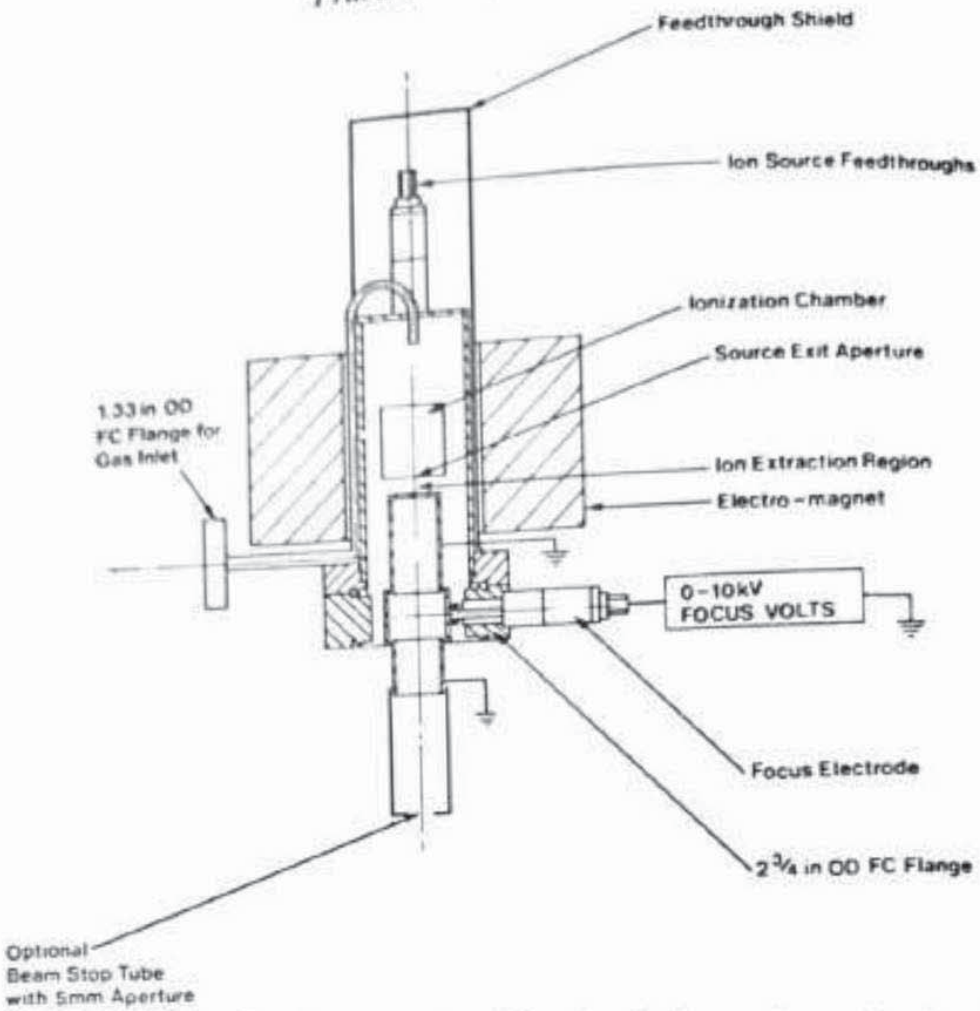


Figure 2.18 Design of an ion gun using a Penning discharge for production of positive ions. Argon gas is fed directly into the back of the gun in this case, rather than being admitted to the whole system, and since its exit speed is restricted the pressure in the interior of the gun is much higher than the pressure of argon in the system. During operation the pressure in the system is between 3×10^{-7} and 3×10^{-6} torr. A high voltage and an axial magnetic field are applied to an ionization chamber in which a discharge is induced and positive ions are extracted from the discharge into a focusing region. Ion energy can be varied betwen 500 eV and 10 keV, and the gun can deliver between 10 and 100 μA into an area of diameter about 5 mm, according to the pressure of argon in the source. (Reproduced by permission of Vacuum Generators Ltd, East Grinstead)

pressures in the analysis chamber between 3×10^{-7} and 3×10^{-6} torr. The ion current on the sample varies with both voltage and pressure, and as an example of performance it would provide about 20 μA into a spot of diameter 5 mm at an accelerating voltage of 5 kV and a chamber pressure of 4×10^{-7} torr. The focusing voltage would then be about 3.6 kV. Under the extreme conditions of 10 kV and 3×10^{-6} torr it would deliver 200 μA.

The greatly increased current density of such ion gun types over the 'ion

gauge' type is somewhat offset by one or two disadvantages. Ions inside the gun can strike other internal surfaces as well as being extracted into the beam, and hence sputtered material gradually accumulates; when the accumulation on insulating surfaces becomes sufficient breakdown or arcing occurs and the action of the gun becomes both unreliable and unreproducible. At that stage there is no option but to dismantle the gun and clean it according to the manufacturer's instructions, and it is generally desirable to keep a stock of spare insulating parts for replacement. The other disadvantage is that the beam emerging from the gun is not pure, in that as well as positive ions there is an appreciable proportion of energetic neutral atoms. Clearly it is only the current of positive ions striking the sample that can be measured, so the additional erosion effect of the neutrals is at the moment a little-known quantity. However, all the discharge-type ion guns suffer from this same drawback to a greater or lesser extent.

Ion sources that produce ion beams by high-voltage discharge depend on the geometry of their structure to increase the path lengths of the oscillating electrons and hence the ionization probability per electron. There are various possible geometries, but they all aim to shape the electrostatic field so that electrons are trapped for a significant length of time in one or more potential wells in the field. Because of the field configuration usually employed, this type of source is known as the 'saddle field' source, and a diagram of one of the smaller versions is shown in Figure 2.19. A pair of tungsten wires, surrounded for most of their length by shielding sleeves, protrude into the body of the gun, which in this case is spherical. A high positive voltage between 5 and 10 kV is applied to the wires, the body being earthed. Argon is admitted directly to the gun at a pressure in the region of 10^{-4} torr, whereupon a discharge is set up between the anode wires and the body, the latter acting as the cathode. The field configuration then constricts the electron paths to the 'figure-of-eight' shown in Figure 2.19, most of the ionization of the argon taking place in the loops at each end of the field. Positive ions are extracted through a slit cut in the body in the same plane as that of the field; the ion beam maintains the same width as that of the slit in the plane of the figure, but diverges progressively from the slit width in the directions above and below the plane. By adjusting the distance of the sample from the slit, the shape of the bombarded area can thus be altered from rectangular to square as desired. Since the only pumping path to the system is through the slit, pressures between 10^{-6} and 10^{-5} torr can be maintained in the system, with adequate pumping speed.

Saddle field sources are capable of providing very high ion current densities. Even a small one like that of Figure 2.19 can produce a density of 200 μA/cm^2 at 6–7 kV, while the larger ones are so powerful that they tend to be used not so much for ion bombardment in UHV as for what is called 'ion milling', in which, as the name suggests, large quantities of material are

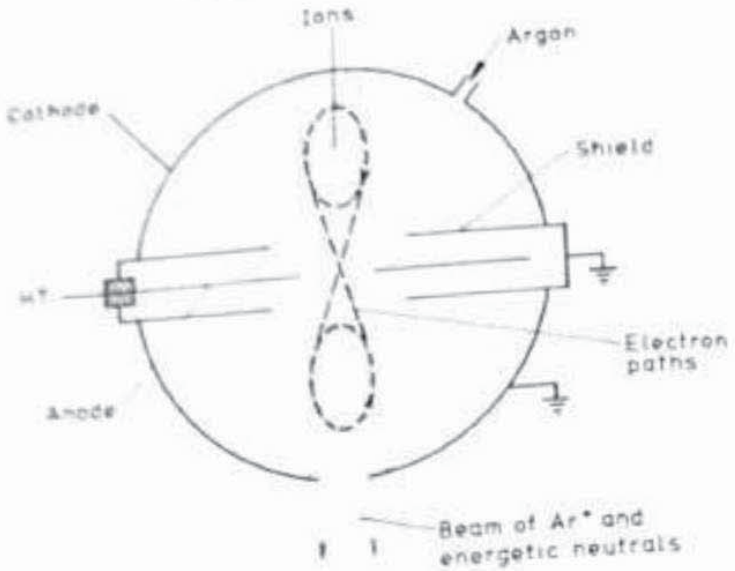


Figure 2.19 Design of an ion gun using a high-voltage discharge. The high voltage is applied to a pair of tungsten wires protruding into the body of the gun, which in this case is spherical. Argon gas is fed directly into the gun from a pipe at the back, and as shown in the design, the pumping speed for argon in the gun is low, so that the pressure inside is much higher than the pressure outside. A discharge is set up between the wires and the body and the electron paths are then constricted to the 'figure-of-eight' paths shown, with most of the ionization of the argon taking place in the loops at each end. Positive ions emerge through a slit cut in the body in the same plane as that of the field. Because of the shape of the field configuration this type of source is also known as the 'saddle field' source. At ion energies of 5–7 keV such a source can produce 100–200 μA of positive ions into an area of about 1 cm^2. (Reproduced by permission of Ion Tech Ltd, Teddington)

removed. As with the other sources, however, they have their disadvantages as well as their advantages. Unlike the sources using a Penning discharge, they tend to be unstable at voltages much below 5 kV and difficult to control, and are therefore restricted in general application to the high-voltage end of their operating regime. In addition, as can be seen from Figure 2.19, positive ions are not only extracted through the slit but bombard the interior of the body as well, so that it is possible for sputtered material also to pass through the slit and be deposited on the sample. It is therefore necessary either to fabricate the body from a metal that does not sputter easily, e.g. aluminium or titanium, or to insert an internal liner of such material. Finally, these sources do produce a rather high proportion of energetic neutral atoms in their beams, of the order of 50 per cent under maximum voltage operation, a higher proportion than from the types using Penning discharges.

The most recently developed[29] type of ion source is the field emission liquid-metal ion source, whose principle is illustrated in Figure 2.20. A needle whose tip has a radius in the range 1–10 μm is placed with its shank passing

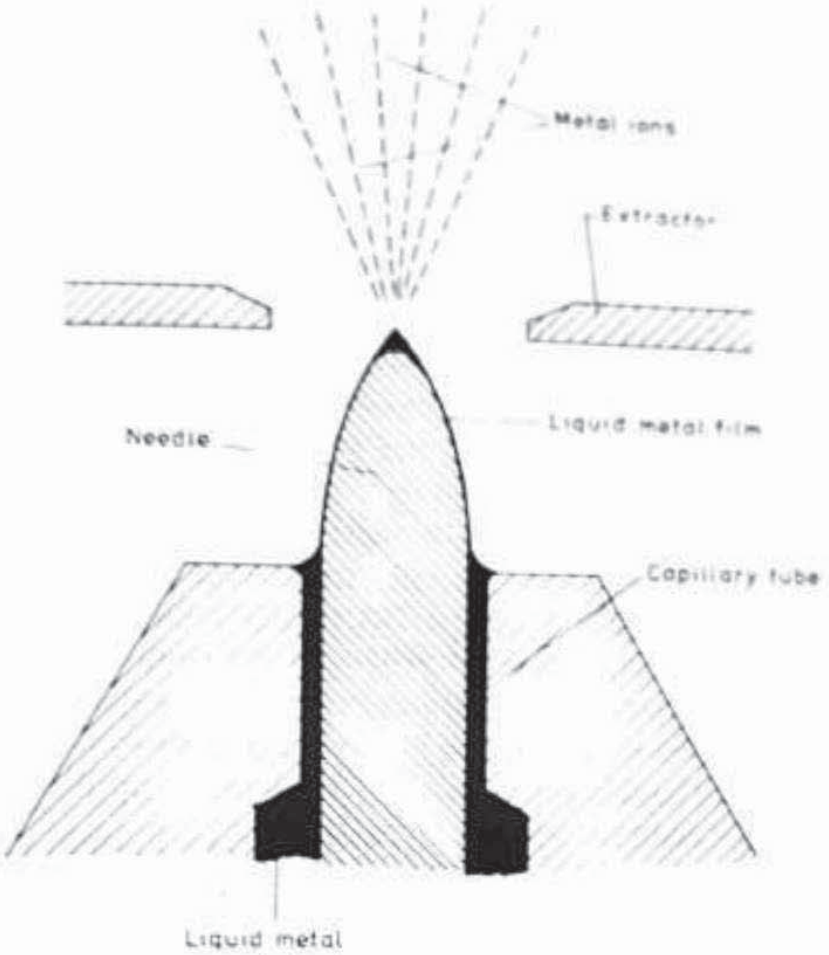


Figure 2.20 Schematic of the principle of operation of a liquid-metal field emission ion source. A needle with a tip of radius 1–10 μm dips into a reservoir of liquid metal through a close-fitting capillary tube. The liquid metal is drawn along the needle and over its tip by a capillary action, provided the liquid wets the material of the needle. A high voltage of between 4 and 10 kV is then applied to an extractor electrode placed a short distance in front of the tip, whereupon the liquid metal is drawn out into a cusp-like protrusion from which positive metal ions are extracted by field emission. Since high currents, of the order of 100 μA, can be drawn from the very small volume of the cusp, the source has a high brightness and the ion beam can thus be focused into a small spot while retaining usefully high currents. (Reproduced from Prewett and Jeffries[28] by permission of The Institute of Physics)

through a close-fitting capillary tube into a reservoir of liquid metal. The material of the needle is chosen so that the liquid wets it, but does not react with it. On wetting, the liquid metal is drawn up over the protruding needle and over its tip by capillary action, any loss being made good by flow from the reservoir. A high voltage between 4 and 10 kV is then applied between the needle and an extractor electrode positioned a short distance in front of the needle, whereupon the film of liquid metal covering the tip is distorted into a cusp-shaped protrusion. This protrusion is formed by the balance of forces due to electrostatic stress and surface tension. Due to the consequent very

high electric field at the tip of the cusp, a beam of positively charged metal ions is formed, which is then extracted through a circular aperture in the extractor. Although the source is described as a field emission source, it is in fact not yet clear whether the ionization mechanism includes field desorption or field ionization, or both, or indeed other contributing processes. Metals used as liquids in these sources have included to date cesium, indium, gallium, tin, bismuth, copper and gold; of these cesium has received most attention due to its potential applications in ion milling and in SIMS.

For reasons similar to those for which the field emission electron source was described as a 'high brightness' source, the liquid metal ion source is also a high brightness source. The high currents obtainable, ranging up to 100 μA, coupled with the very small emitting volume imply very high brightness values; unfortunately it is difficult to be precise because the volume is not known with any certainty. The high brightness, however, just as in the case of the electron source, leads to the possibility of producing useful currents in a finely focused beam, enabling ion erosion to be performed at much higher spatial resolution in principle than can be achieved with the gas discharge types. Preliminary experiments have shown[29] that ion spot sizes of less than 0.5 μm can be achieved without much effort, and it seems likely that resolutions of 0.1–0.2 μm should soon be possible.

There are at the moment several unknown quantities in respect of liquid-metal sources, although much research and development is being carried out and it will not be long before much more is known about them. They do seem to suffer from instabilities in certain regions of operating parameters, and this may be a function of the mechanisms of formation of the liquid cusp. The composition of the ion beam, in terms of proportions of singly and doubly charge ions, etc., varies with the metal being used and to a certain extent with the applied voltage. Molecular clusters may also be formed. There is a considerable divergence of the ion beam from the aperture, so that a carefully designed focusing lens system is required to make optimum use of the high brightness. However, these drawbacks are not necessarily insuperable and it is expected that liquid-metal ion sources will play an increasing part in surface analytical technology.

2.5 Electron Energy Analysers

The device which measures the energies (or, more correctly, the velocities) of electrons emitted or scattered from a surface is the heart of any spectrometer. In this section the device will be called the 'analyser' to distinguish it from the term 'spectrometer', which has been taken in this chapter to mean the complete instrument, as defined in Section 2.1. However, it should be noted that many authors use the term 'spectrometer' when they intend it to apply only to the analyser; neither usage is entirely wrong or entirely right, and it seems usually to be a matter of personal preference and definition.

Over the years of development of XPS and AES various types of analyser have been put forward as being eminently suited to the techniques, but there now seems to be a general consensus of agreement as to those that are most suitable. Because of the low kinetic energies of the electrons to be analysed in the techniques, and because magnetic fields are difficult to produce and handle in UHV, the analysers are all of the electrostatic type. The two principal ones are the cylindrical mirror analyser (CMA) and the concentric hemispherical analyser (CHA), which are both dispersive analysers, that is to say, in each the action of a deflecting electrostatic field disperses the electron energies so that for any given field only those energies in a certain narrow range are measured. Also still in use, and not always to be despised, is the retarding field analyser (RFA) based on the spherical electron optics used for low energy electron diffraction (LEED), in which a retarding potential is imposed in front of a collector allowing only those electrons of energies greater than the potential barrier to reach the collector. The descriptions of these three analysers will be given in the approximate chronological order in which they appeared on the surface analysis scene. Before that can be done, however, the requirements of XPS and AES in terms of energy resolution must be established.

2.5.1 Energy resolution requirements

There are two definitions of energy resolution in current use and it is necessary to be clear about the differences between them. The first is the *absolute* resolution, ΔE, usually measured as the full width at half-maximum (FWHM) height of a chosen observed peak. An alternative measure sometimes used is the width ΔE_B of the peak at its base, called the base width, and obviously in an ideal situation $\Delta E_B = 2\Delta E$. The second is the *relative* resolution defined as

$$R = \Delta E/E_0 \tag{2.5}$$

where E_0 is the kinetic energy at the peak position. The resolution R is usually expressed as a percentage, i.e. $(\Delta E/E_0) \times 100$. It is also common to find the relative resolution expressed as the *resolving power*, which is simply the reciprocal of R, i.e.

$$\rho = 1/R = E_0/\Delta E \tag{2.6}$$

Thus the absolute resolution can be specified independently of peak position in a spectrum, but the relative resolution can be specified only by reference to a particular kinetic energy.

Now in XPS it is necessary for identification of possible differences in chemical states of elements to be able to apply the same absolute resolution to any photo-electron peak in the spectrum, i.e. at any kinetic energy. As has been shown, the natural (i.e. unmonochromatized) line widths of the com-

2.5.2 Retarding field analyser (RFA)

Long before XPS and AES were established the technique of LEED was using electron optics of spherical symmetry to display surface diffraction patterns by acceleration of the diffracted electron beams onto a phosphor-coated collector or screen. The original arrangement consisted of two grids concentric with the screen, the inner grid being grounded,the second grid held at a negative potential close to that of the primary energy and the screen held at a

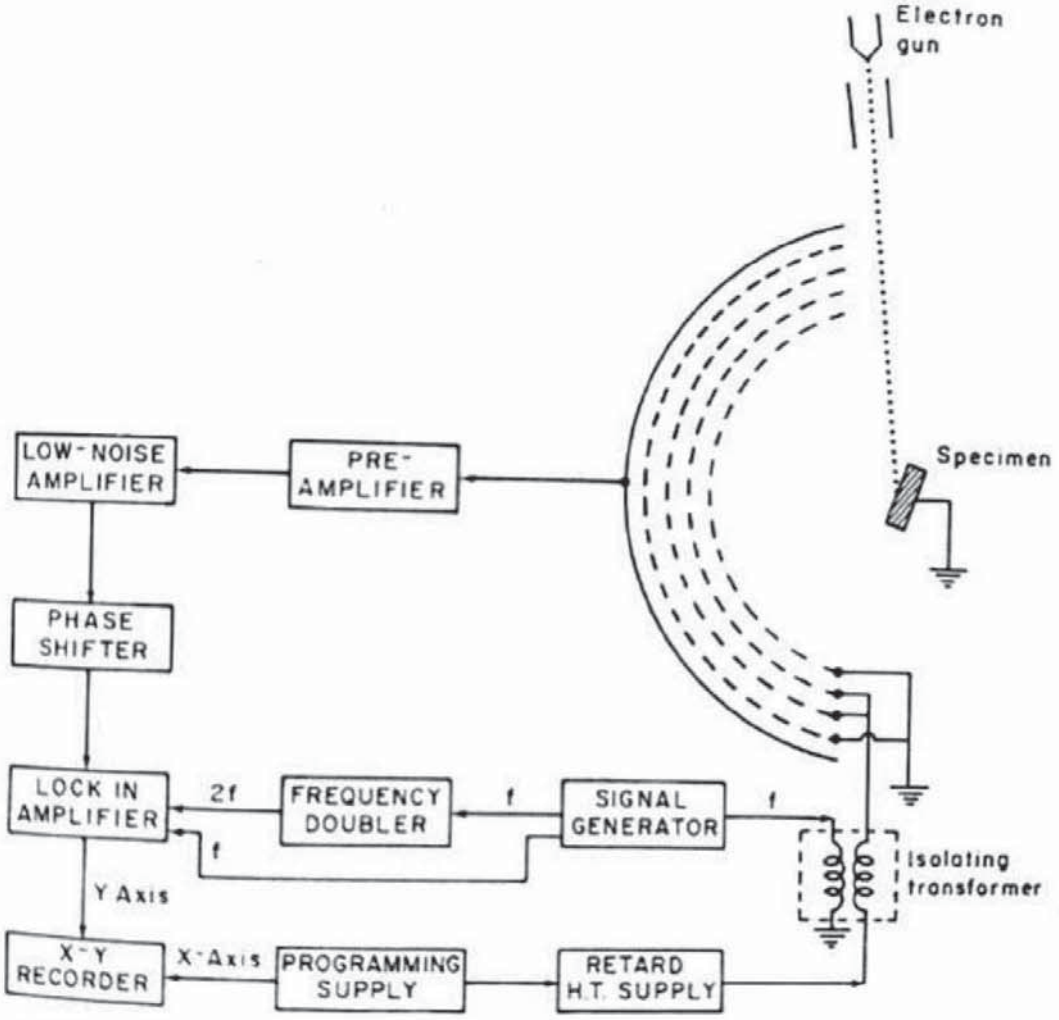


Figure 2.21 Schematic arrangement of a retarding field analyser (RFA) for AES based on four-grid LEED optics. The LEED electron gun would normally be shown protruding through the collector (screen) and grids, but is omitted here because for most analytical purposes it is unsuitable, since the electron energy range it can produce is too restricted and the spot size is too large. The retarding potential applied to the two inner grids is modulated with a small potential from a signal generator, via an isolating transformer, and the amplified signal from the collector is compared in the lock-in amplifier with either the fundamental or the first harmonic of the modulation frequency. The first comparison gives the energy distribution $N(E)$ according to expression (2.8) and the second the differential distribution $dN(E)/dE$ according to expression (2.9)

high positive potential. After it was realized[31] in about 1967 that the LEED optics could be used as an energy analyser the number of grids increased quickly to three and then four, with the final (and present) arrangement being that of Figure 2.21.

In the four-grid RFA of Figure 2.21 the incident electrons come either from the LEED gun projecting through the optics, to strike the sample at normal incidence, or from an auxiliary gun outside the optics, for grazing incidence excitation. The inner and outer grids are grounded, in order to provide a uniform field in the sample region and to shield the collector (i.e. the LEED screen) from capacitative coupling with the modulation on the other grids, respectively. The middle two grids are connected together and have the retarding potential applied to them; the reason for using two grids there is to reduce the serious lens effects suffered by electrons passing through the holes in a mesh. Superimposed on the retarding potential is a small modulating voltage from a signal source. The modulation is usually sinusoidal but need not be so.

The modulated collector current is then compared in a phase-sensitive detector (otherwise known as a lock-in amplifier or synchronous detector) with the reference signal from the signal source. Amplification of the component of collector current at the fundamental frequency ω then gives the energy distribution $N(E)$, while amplification of the component at the frequency 2ω gives the differential distribution $N'(E)$. These relationships can be derived quite easily by expansion of the signal as a Taylor series and algebraic rearrangement. Thus the collected current I is a function of the retarding potential V and, if the modulation is of the form $k \sin \omega t$, I can be written $I(V + k \sin \omega t)$. The Taylor expansion is then

$$I(V + k \sin \omega t) = I + k \sin \omega t(\mathrm{d}I/\mathrm{d}V) + \frac{k^2 \sin^2 \omega t}{2!} (\mathrm{d}^2I/\mathrm{d}V^2) + \cdots \quad (2.7)$$

When this is rearranged algebraically, the coefficient of $\sin \omega t$ is found to be

$$k(\mathrm{d}I/\mathrm{d}V) + (k^3/8) (\mathrm{d}^3I/\mathrm{d}V^3) + \ldots \quad (2.8)$$

and the coefficient of $\cos 2\omega t$ to be

$$(k^2/4) (\mathrm{d}^2I/\mathrm{d}V^2) + (k^4/48) (\mathrm{d}^4I/\mathrm{d}V^4) + \ldots \quad (2.9)$$

Sinse $(\mathrm{d}I/\mathrm{d}V)$ is the energy distribution $N(E)$ and $(\mathrm{d}^2I/\mathrm{d}V^2)$ the differential distribution $N'(E)$, it can be seen that the amplitude of the first harmonic is proportional to the former and of the second harmonic to the latter, *provided* the higher order derivatives can be neglected in each case. This will be so only if k is small enough. In practice this criterion will be obeyed if the modulation amplitude is less than half the width of an Auger peak; since the widths of Auger peaks vary, the modulation should obviously be chosen to resolve without distortion the narrowest peak in the spectrum.

Enhanced sensitivity can be obtained by using large modulating voltages, but at the cost of a distorted spectrum.

With sufficient care taken in construction and with some modifications to the grid spacings normally used for LEED, a relative resolution of about 0.3 per cent has been achieved for an RFA. For most RFA systems, though, the usually obtainable resolutions are in the range 0.5–1.0 per cent, which are just about adequate for AES. Where the RFA tends to score over the dispersive analysers is in the very low kinetic region, below 100 eV, since there a relative resolution of 0.5 per cent implies an absolute resolution of less than 0.5 eV, resulting in high resolution spectra.

Despite its inherent simplicity, the RFA is not suitable for general AES applications because of one overriding disadvantage, and that is that at any one retarding potential on the grids, *all* electrons with energies greater than that potential are allowed to reach the collector. Shot noise is thus generated by many electrons of energies other than those being analysed, with the result that the signal-to-noise characteristics are poor. There are also other problems such as uniformity of grid structure and the parallelism of the grids, for minor deviations from strict geometry can have devastating effects on the resolution obtainable. However, for those whose main research effort is devoted to the application of LEED to surface structural studies, the RFA presents a convenient means of checking the cleanliness of a surface at any stage in an experiment.

2.5.3 Cylindrical mirror analyser (CMA)

The drawbacks inherent in the RFA led the early users of AES to seek other types of energy analyser that would be better suited to the technique, and it was soon realized[32] that the CMA had many properties that made it almost ideal for the purpose. Over the years various commercial and home-made versions of the CMA have appeared, but in general the differences between the versions have been in details of construction, the basic form and mode of operation remaining unchanged.

The CMA is shown diagrammatically in Figure 2.22. Two cylinders of radii r_1 (inner) and r_2 (outer) are positioned accurately coaxially. The potential of the inner cylinder is at earth; that of the outer cylinder is $-V$. Electrons emitted from a source on the axis at an angle α pass through an aperture in the inner cylinder and those of a particular energy E_0 are deflected by the outer cylinder potential through another aperture to a focus on the axis. The general relationship between V and E_0 is

$$\frac{E_0}{eV} = \frac{K}{\ln(r_2/r_1)} \tag{2.10}$$

CMAs can and have been built for a variety of entrance angles α, but for

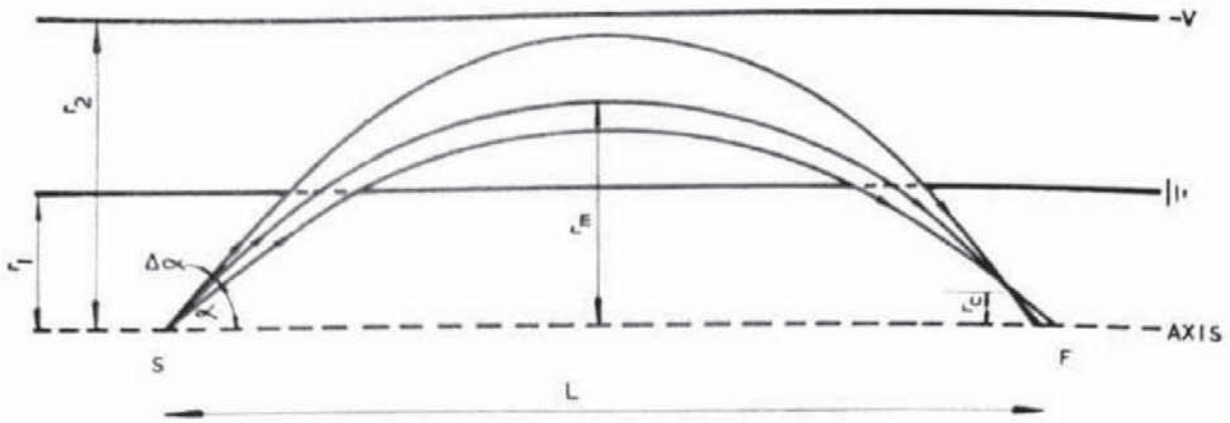


Figure 2.22 Diagrammatic arrangement of a cylindrical mirror analyser (CMA). The radii are r_1 for the inner cylinder and r_2 for the outer cylinder. The inner cylinder is earthed and a potential $-V$ is applied to the outer cylinder. Electrons emitted from a source S on the axis with a kinetic energy E_0 are re-focused at F according to the expression (2.10). The entrance angle α is chosen to be 42° 18′, since at that angle the CMA becomes a second-order focusing device. A typical angular aperture $\Delta\alpha$ would be 6°. L is the distance between S and F, r_c is the position of the minimum trace width and r_m the maximum distance off the axis for electrons entering the analyser at 42° 18′. (After Bishop, Coad and Rivière[36])

the special case where $\alpha = 42° 18'$, the CMA becomes a second-order focusing instrument.[33] For that case $K = 1.31$, and the above relationship between E_0 and V becomes

$$\frac{E_0}{eV} = \frac{1.31}{\ln(r_2/r_1)} \tag{2.11}$$

Apertures are in practice of finite width, and in any case it is desirable to achieve maximum sensitivity compatible with adequate resolution by operating the CMA with apertures that accept a spread of angles about α, typically with $\Delta\alpha = 6°$. Under these conditions the electron paths through the analyser, depicted in Figure 2.22, consist of annuli rather than an infinitely narrow trace, and it can be shown[33,34] that the minimum trace width ω_m is given, for $\Delta\alpha < 6°$, by

$$\frac{\omega_m}{r_1} = 7.76(\Delta\alpha)^3 \tag{2.12}$$

The important feature of the minimum trace width is its position, which is off-axis and slightly ahead of the focal position. Its distance r_c off the axis is given by

$$\frac{r_c}{r_1} = 5.28(\Delta\alpha)^2 \tag{2.13}$$

The importance lies in the fact that if an additional aperture is placed[35] at the position of minimum trace width, and made equal in size to it, a useful

improvement in resolution is obtained with little sacrifice in sensitivity. Some designs incorporate[36] the possibility of varying the width of the additional aperture, but this is not an operation that is easy to arrange for remote control, i.e. from outside the vacuum system, and therefore the tendency is to fix that aperture size at some convenient compromise and to allow variation instead of the sizes of the principal apertures.

For the unique angle 42° 18′ the distance on the axis between the source and focus is $6.1r_1$, and the other important constructional parameter is the maximum distance off the axis reached by electrons that are to pass through the exit apertures. This distance is given by[29]

$$\frac{r_m}{r_1} = \exp(1.3 \sin^2 \alpha) \tag{2.14}$$

Thus all the CMA parameters scale with the radius r_1 of the inner cylinder, and once that is chosen the rest follows more or less automatically.

Not shown in Figure 2.22 for reasons of clarity are some other important operational requirements, namely the fringe field plates. For practical reasons, i.e. access to the specimen position, the source of ejected electrons must lie *outside* the end of the CMA, and it can be seen at once that such an arrangement involves terminating the analyser rather abruptly. In fact, all commercial CMAs are cut back from their axes, so that their ends are cone shaped. However, the electrons passing through the analyser must see only uniform radial fields, otherwise there will be severe aberrations, so the field in the analyser must be terminated at each end without introducing field distortion in the aperture regions. In some CMAs this is achieved[36] by a series of ring-shaped electrodes whose potentials are adjusted via dividing resistors to ensure a smooth transition from the deflecting potential V to earth, in others by ceramic discs coated with a film whose radial resistance performs the same function as the dividing resistors, but in a more continuous fashion.

As in all dispersive analysers, simply scanning the potential $-V$ applied to the outer cylinder of a CMA gives the energy distribution of electrons passing through it directly, but it is important to note that because the transmission of the CMA varies as E_0 the distribution derived is not $N(E)$ but $EN(E)$. Thus modulation of V by a small sinusoidal signal, as described for the RFA, can still be used to extract a differential distribution, in this case by tuning the phase-sensitive detector to the fundamental frequency ω, and not to the second harmonic, but again the recorded distribution will be $EN'(E)$ rather than $N'(E)$.

The energy resolution of the CMA and CHA can be written[36,37] in terms of the base resolution ΔE_B as

$$\frac{\Delta E_B}{E_0} = Aw + B(\Delta\alpha)^n \tag{2.15}$$

where w = slit width (equal at the entrance and exit) and A, B and n are constants. A more approximate expression[38] involving the half-width ΔE is

$$\frac{\Delta E}{E_0} = \tfrac{1}{4}Aw + \tfrac{1}{4}B(\Delta\alpha)^n \tag{2.16}$$

which is valid only if $B(\Delta\alpha)^n = Aw/2$. For the CMA and CHA the values of the constants are given in Table 2.2.

Insertion of the values for the CMA into the two expressions above gives, for the base resolution,

$$\frac{\Delta E_B}{E_0} = \frac{0.36w}{r_1} + 5.55(\Delta\alpha)^3 \tag{2.17}$$

and, for the half-width resolution,

$$\frac{\Delta E}{E_0} = \frac{0.18w}{r_1} + 1.39(\Delta\alpha)^3 \tag{2.18}$$

In the case of the CMA the relatively wide angular apertures used can limit the resolution obtainable; in effect the source size replaces the slit widths in the above expressions. As an example, a CMA built with an inner radius r_1 of 3 cm and a semi-angular aperture $\Delta\alpha$ of 6° will have a theoretical half-width resolution $\Delta E/E_0$ of 0.15 per cent for a source size of 5 μm. However, for greater source sizes, the theoretical resolution degrades as follows: 10 μm, 0.16 per cent; 50 μm, 0.18 per cent; 100 μm, 0.21 per cent; 500 μm, 0.48 per cent; 1 mm, 0.80 per cent. The practical resolution obtainable will of course depend on the accuracy of fabrication and on stray magnetic fields, but in a reasonably well-constructed analyser, the degradation from these effects might cause the half-width resolution to deteriorate from 0.16 to 0.25 per cent for a 10 μm spot size.

The simple (single-pass) CMA as described up to now has been most effective in AES because of its very high transmission. This is the most important property for an AES analyser because the source area in AES using a highly focused electron beam is invariably much smaller than the acceptance area of the analyser. In other words, luminosity, which is the product of acceptance area and transmission, is not particularly relevant in AES. On the other hand, for XPS luminosity is all-important because, as described, standard unmono-

Table 2.2 Values of coefficients giving the base and half-width energy resolutions of the CMA and CHA

	A	B	n
CMA (42.3°)	$0.36/r_1$	5.55	3
CHA (180°)	$1/R_0$	1	2

chromatized X-ray sources are not focused but flood a large area of the sample with X-rays. In addition, XPS needs in general a greater absolute resolution than does AES. For these reasons, the single-pass CMA is not suitable for XPS. Another reason is that the precise energetic position of a photo-electron or Auger peak is so dependent on sample position in front of a single-pass CMA that reproducibility adequate for XPS would be totally lacking. (For example, Sickafus and Holloway[39] have shown that at 1500 eV a shift of 1 mm in the axial sample position causes a peak shift of 17.5 eV.)

The high transmission characteristics of the CMA are attractive for XPS, but higher luminosity should be achieved while maintaining adequate resolution. By using two CMAs in series, i.e. a two-stage or, as it is better known, a double-pass CMA, with pre-retarding grids before the entrance slit, an analyser suitable for XPS has been constructed.[40] It is shown diagrammatically in Figure 2.23. The two pre-retarding grids are spherical and centred on the source position of the specimen.

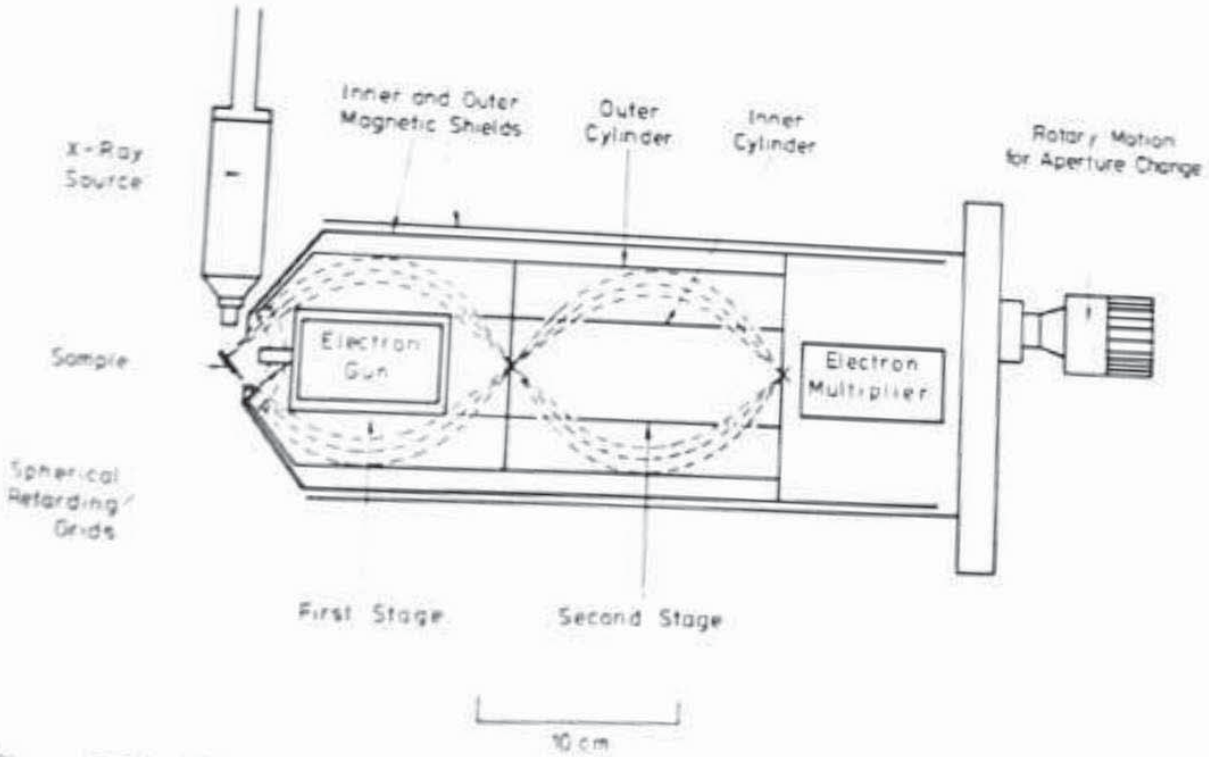


Figure 2.23 Diagrammatic arrangement of a double-pass CMA, used for both AES and XPS. The exit aperture from the first stage is the entrance aperture to the second stage. At the front end of the analyser are two spherical retarding grids centred on the source area of the sample that retard photo-electrons to a constant pass energy for XPS. For AES the grids are at earth potential, as is the inner cylinder. An externally operated rotary motion allows the entrance and exit apertures to the second stage to be changed remotely, from large sizes for XPS to small sizes for AES. The electron gun is situated on the axis of the CMA internally, but the X-ray source, of the type seen in detail in Figure 2.9(a), is external and positioned as close to the sample as the geometry will allow. (Reproduced from Palmberg[40] by permission of Elsevier Scientific Publishing Company)

The effect of pre-retardation is that by going from a kinetic energy of E_0 to a fixed analyser band-pass energy E_p, the energy resolution is improved (i.e. $\Delta E/E_p$ is reduced) by a factor (E_p/E_0). At the same time, the effective source volume is collapsed by a factor $(E_p/E_0)^{1/2}$. Thus, to maintain overall luminosity at a level suitable for XPS, the entrance and exit apertures to the *second* stage must be increased as far as possible, which means in practice to the point where the overall energy resolution is not quite degraded. To do this, the apertures can be altered by an external control so that either small apertures for AES or large apertures for XPS can be selected. In the double-pass CMA, therefore, the effective energy resolution is set by the first stage and the sensitivity by the second stage, although the two functions are clearly not inseparable.

For AES, the double-pass CMA is operated in the 'normal' mode, i.e. with the retarding grids grounded. If pre-retardation were attempted with AES, grid scattering would reduce the transmission undesirably. In XPS, however, although grid scattering reduces the transmission when using pre-retardation, the effective acceptance area is increased proportionately so that the luminosity is unaffected.

The opposing effects of collapse of the source area on retardation coupled with the effective source area increase due to grid scattering are functions both of kinetic energy and pass energy, and are somewhat difficult to estimate. The combined effect for any chosen pass energy will not necessarily be zero, or even linear with kinetic energy. In general, for absolute resolution ΔE worse than about 1.4 eV, the high luminosity of the double-pass CMA means that the signal intensity, for identical sample and operating conditions, is greater than for the CHA. However, as the pass energy is decreased to achieve absolute resolution of ≤ 1.0 eV, the signal intensity drops off more rapidly than in the CHA, and it is likely that the retardation grid effects mentioned above are largely to blame. Simply increasing the overall size of the CMA, i.e. scaling with the inner radius r_1, serves to reduce the effect for a given pass energy, and this increase is in fact the direction in which the manufacturers are currently going.

2.5.4 Concentric hemispherical analyser (CHA)

At roughly the same time (1969) as the CMA was established as highly suitable for AES, XPS was being recognized as a valuable surface-specific analytical technique, and the type of analyser used by its founder, Siegbahn,[41] was being adapted to operation in UHV. This was the concentric hemispherical analyser (CHA), also called the spherical deflector analyser, and it is still the most widely used energy analyser for XPS, being supplied in the spectrometers from at least four manufacturers, whereas the double-pass CMA is supplied from only one.

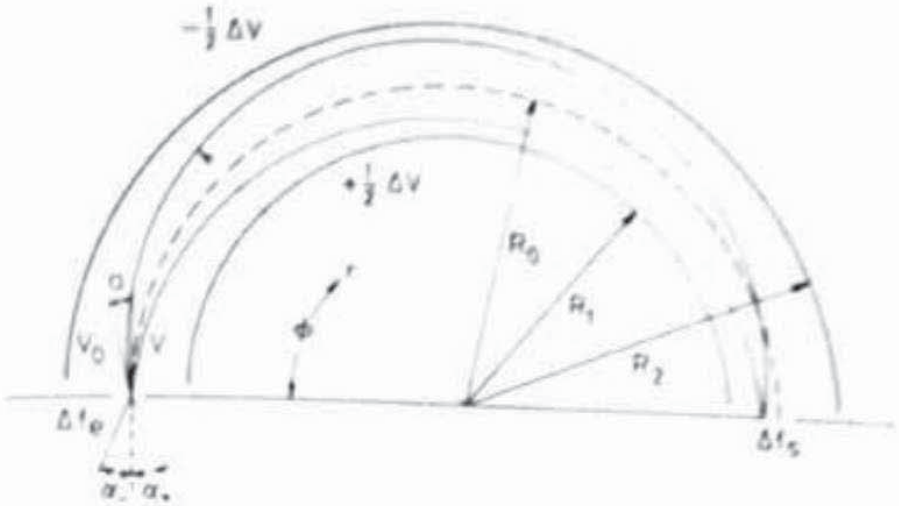


Figure 2.24 Principle of operation of the concentric hemispherical analyser. Two hemispherical surfaces of inner radius R_1 and outer radius R_2 are positioned concentrically. A potential ΔV is applied between the surfaces so that the outer is negative and the inner positive with respect to ΔV. R_0 is the median equipotential surface between the hemispheres, and the entrance and exit slits are both centred on R_0. If E is the kinetic energy of an electron travelling in an orbit of radius R_0, then the relationship between E and V is given by expression (2.21). ϕ and r are the angular and radial coordinates, respectively, of an electron of energy E_0 ($= e\Delta V_0$) entering the analyser at an angle α to the slit normal. If this electron is to pass through the exit slit, its path in the analyser is governed by the conditions of expression (2.22). (Reproduced from Roy and Carette[43] by permission of the National Research Council of Canada)

The CHA is shown schematically in Figure 2.24. Two hemispherical surfaces of inner radius R_1 and outer radius R_2 are positioned concentrically. A potential ΔV is applied between the surfaces so that the outer sphere is negative and the inner positive. There will be a median equipotential surface between the hemispheres of radius R_0, and the entrance and exit slits are each centred on a distance R_0 from the centre of curvature and lie on a diameter. In an ideal situation $R_0 = (R_1 + R_2)/2$, and that relationship is usually obeyed closely.

The coordinate position of an electron of initial kinetic energy E_0 travelling between the hemispheres is governed by two focusing conditions. Firstly,[42] if an electron enters the analyser at an angle α to the normal through the slit centre, its position in terms of r, its radial distance from the centre of curvature, and ϕ, its angular deflection, is given by

$$y = -y_0 \tan\alpha \sin\phi + (y_0 - c^2)\cos\phi + c^2 \qquad (2.19)$$

where $y = R_0/r$

y_0 = initial coordinate of electron

$c^2 = Ey_0^2/[E_0 \cos^2\alpha - 2E(1 - y_0)]$

Here E corresponds to the kinetic energy which an electron would have if

travelling on the circular orbit of radius R_0; for the 180° CHA, the first condition reduces to

$$y_f = 2c^2 - y_0 \tag{2.20}$$

where y_f is the final coordinate of the electron. The relationship between the deflecting potential ΔV and the ideal energy E is

$$e\Delta V = E(R_2/R_1 - R_1/R_2) \tag{2.21}$$

The second focusing condition arises[43] from the limits that must be imposed on the greatest and least radii of the electron trajectories. The maximum and minimum excursions can be expressed with respect to the median trajectory radius R_0 by

$$y_\pm = c^2 \pm (c^4 + 2k)^{1/2} \tag{2.22}$$

where $k = \frac{1}{2} y_0^2 (\tan^2 \alpha + 1) - c^2 y_0$

This second condition merely says that if the electron is to be transmitted by the analyser its final position y_f must be contained within the aperture of the exit slit and that the $y_\pm$ are not outside the limits imposed by the hemispherical surfaces.

In the earlier spectrometers, the mid-potential of the analyser was tied to the ground potential, and the scanning through the kinetic energy range was performed by varying the specimen potential. Nowadays, however, the situation is reversed in that the specimen is grounded and the whole analyser floats isolated from ground so that scanning is now carried out by varying the potential (or 'pedestal') on which the analyser sits.

The base and half-width resolutions can be written down from expressions (2.15) and (2.16), as well as from Table 2.2. They are

$$\frac{\Delta E_B}{E} = \frac{w}{R_0} + \alpha^2 \tag{2.23}$$

and

$$\frac{\Delta E}{E} = \frac{w}{2R_0} + \frac{\alpha^2}{4} \tag{2.24}$$

(Note that the entrance angle α into the CHA is equivalent here to the semi-angular aperture $\Delta\alpha$ of the CMA.) In general it is desirable to work with a large entrance angle for increased sensitivity, and since a compromise must be reached between sensitivity and resolution it is common to choose α so that $\alpha^2 \approx w/2R_0$. Then expressions (2.23) and (2.24) become

$$\frac{\Delta E_B}{E} = \frac{1.5w}{R_0} \tag{2.25}$$

and

$$\frac{\Delta E}{E} = \frac{0.63w}{R_0} \tag{2.26}$$

Thus, if R_0 has been chosen and a certain absolute resolution is required, there is an inverse relationship between the pass energy E and the slit width.

There are two retarding modes in which CHAs are currently used. In one, the electrons are decelerated by a constant factor, or ratio, from their initial kinetic energies. Clearly this mode operates at constant *relative* resolution; it is termed either the CRR or the FRR mode. In the other, the electrons are decelerated to a constant pass energy, and it is obvious that the mode is that of constant *absolute* resolution. It is termed the CAT or FAT mode, standing for constant or fixed analyser transmission. Each mode has its adherents and each its advantages and disadvantages. The CAT mode is generally easier for quantification, since the absolute resolution is the same at all parts of the spectrum, but the signal-to-noise characteristics become progressively worse towards low kinetic energies. The CRR mode is less easy to quantify, but small peaks at low kinetic energies can be detected without difficulty.

Retardation when used with a CHA is usually accomplished with planar grids across the entrance slit, but may also be achieved with a lens system. Lens systems are used in other ways, too, to improve the overall efficiency of analysis. By removing the sample from the entrance to the analyser, and instead projecting the analysed area of the sample onto the entrance slit, a much greater working distance is available below the lens, allowing closer approach of the X-ray source, electron gun and ion gun. In the cases of the two latter, this close approach is necessary to be able to achieve minimum beam spot sizes. According to whether either or both CRR and CAT modes of retardation are required, the lens is more or less complex, for in the CRR mode transmission between the sample and the analyser is constant, whereas in the CAT mode the transmission in the lens is varying. The former could be performed with a simple unipotential lens, but the latter needs variable ratios for which three or more elements are necessary.

Various criteria or figures of merit have been applied to the CMA and CHA to establish their respective ranking, including transmission,[44] luminosity and étendue[45] (the product of the entrance area and entrance solid angle). According to Roy and Carette,[46] who have summarized all these criteria for various analysers including these two, the CMA in the second-order focusing arrangement (i.e. $\alpha = 42.3°$) is consistently the best. However, they also make the very pertinent point that the criteria take no account of the way in which a particular analyser is to be applied, and that point is certainly relevant in the present context. The choice must depend on practical considerations. This is precisely the reason why the CMA has up to now been the prime choice for AES, because of its high transmission at moderate energy resolu-

tion, while the CHA is always preferred for XPS, because, particularly with the addition of a lens, it can maintain adequate luminosity at high-energy resolution. Also, of course, the CHA, with its narrow acceptance angle, is suited to angular dependence measurements, while the CMA is in general not so suited because it collects over 360° (although in the new models there is an internal shutter which can be used to select variable sectors; see Chapter 4.) Development of the CMA will probably go in the direction of increased size only, since it is difficult to see how it could accommodate a lens, whereas it is unlikely that the development of the input lens to the CHA has reached its optimum.

REFERENCES

1. M. P. Seah and W. A. Dench, *Surf. Interface Anal.*, **1**, 2 (1979).
2. For example, M. Pirani and J. Yarwood, *Principles of Vacuum Engineering*, Chapman and Hall, London (1961).
3. R. Kelly, *Surface Science*, **100**, 85 (1980).
4. G. W. Lewis, G. Kiriakides, G. Carter and M. J. Nobes, *Surf. Interface Anal.*, **4**, 141 (1982).
5. M. P. Seah, *Thin Solid Films*, **81**, 279 (1981).
6. P. H. Holloway and R. S. Bhattacharya, *J. Vac. Sci. Tech.*, **20**, 444 (1982).
7. W. O. Hofer and V. Littmark, *Phys. Lett.*, **71**, 6 (1979).
8. R. Kelly and O. Auciello, *Surface Science*, **100**, 135 (1980).
9. A. C. Parry-Jones, P. Weightman and P. T. Andrews, *J. Phys. C: Solid State Phys.*, **12**, 1587 (1979).
10. J. P. Coad, J. C. Rivière, M. Guttmann and P. R. Krahe, *Acta Met.*, **25**, 161 (1977).
11. R. G. Musket, W. McLean, C. A. Colmenares, D. M. Makowiecki and W. J. Siekhaus, *Applic. Surface Science*, **10**, 143 (1982).
12. C. Benndorf, H. Seidel and F. Thieme, *Rev. Sci. Inst.*, **47**, 778 (1976).
13. H. H. Madden, *J. Vac. Sci. Tech.*, **18**, 677 (1981).
14. M-G. Barthès and C. Pariset, *Thin Solid Films*, **77**, 305 (1981).
15. C. A. Crider and J. M. Poate, *Appl. Phys. Lett.*, **36**, 417 (1980).
16. S. Okada, K. Oura, T. Hanawa and K. Satoh, *Surface Science*, **97**, 88 (1980).
17. M. Eizenberg and R. Brener, *Thin Solid Films*, **88**, 41 (1982), and **89**, 355 (1982).
18. K. Yates, A. Barrie and F. J. Street, *J. Phys. E: Sci. Inst.*, **6**, 130 (1973).
19. J. E. Castle, L. B. Hazell and R. D. Whitehead, *J. Electron. Spec.*, **9**, 247 (1976).
20. R. M. Dolby, *Brit. J. Appl. Phys.*, **11**, 64 (1960).
21. J. E. Castle, L. B. Hazell and R. H. West, *J. Electron. Spec.*, **16**, 97 (1979).
22. M. A. Kelly and C. E. Tyler, *Hewlett-Packard Journal*, **24**, 2 (1972).
23. K. Yates and R. H. West, *Surf. Interface. Anal.* (in press).
24. Y. Farge and P. J. Duke (Eds), *European Synchrotron Radiation Facility*, Supplement I, European Science Foundation (1979).
25. J. C. Rivière, *Solid State Surface Science*, **1**, 79 (1969); H. B. Michaelson, *J. Appl. Phys.*, **48**, 4729 (1977).
26. A. Christou, *J. Appl. Phys.*, **47**, 5464 (1976).
27. M. Prutton, R. Browning, M. M. El Gomati and D. Peacock, *Vacuum*, **32**, 351 (1982).
28. P. D. Prewett and D. K. Jefferies, *J. Phys. D: Applied Phys.*, **13**, 1747 (1980).

29. P. D. Prewett and D. K. Jefferies, *Second International Conf. Low Energy Ion Beams*, Bath (1980).
30. J. C. Fuggle, in *Electron Spectroscopy* (Eds C. R. Brundle and A. J. Baker), Vol. IV, Academic Press, London (1982).
31. P. W. Palmberg and T. N. Rhodin, *J. Appl. Phys.*, **39**, 2425 (1968).
32. P. W. Palmberg, G. K. Bohn and J. C. Tracy, *Appl. Phys. Letters*, **15**, 254 (1969).
33. V. V. Zashkvara, M. I. Korsunskii and O. S. Kosmachev, *Sov. Phys. Tech. Phys.*, **11**, 96 (1966).
34. H. Z. Sar-El, *Rev. Sci. Inst.*, **41**, 561 (1970).
35. H. Hafner, J. A. Simpson and C. E. Kuyatt, *Rev. Sci. Inst.*, **39**, 33 (1968).
36. H. E. Bishop, J. P. Coad and J. C. Rivière, *J. Electron Spec.*, **1**, 389 (1972–1973).
37. J. A. Simpson, in *Methods of Experimental Physics* (Eds V. W. Hughes and H. L. Schultz), Vol. 4A, pp. 124–135, Academic Press, New York (1967).
38. M. E. Rudd, in *Low Energy Electron Spectrometry* (Ed. K. D. Sevier), pp. 17–32, Wiley-Interscience, New York (1972).
39. E. Sickafus and P. H. Holloway, *Surface Science*, **51**, 131 (1975).
40. P. W. Palmberg, *J. Electron. Spec.*, **5**, 691 (1974), and *J. Vac. Sci. Tech.*, **12**, 379 (1975).
41. K. Siegbahn, C. Nordling, A. Fahlman, R. Nordberg, K. Hamrin, J. Hedman, G. Johansson, T. Bergmark, S. Karlson, I. Lindgren and B. Lindberg, *ESCA, Atomic, Molecular and Solid State Structure Studied by Means of Electron Spectroscopy*, Almqvist and Wiksells Boktryckeri AB, Uppsala (1967).
42. E. M. Purcell, *Phys. Rev.*, **54**, 818 (1938).
43. D. Roy and J-D. Carette, *Canadian J. Phys.*, **49**, 2138 (1971).
44. M. P. Seah, *Surf. Interface Anal.*, **2**, 222 (1980).
45. D. W. O. Heddle, *J. Phys. E: Sci. Inst.*, **4**, 589 (1971).
46. D. Roy and J-D. Carette, *Topics in Current Physics*, Vol. 4, *Electron Spectroscopy for Chemical Analysis*, pp. 13–58, Springer-Verlag, Berlin (1972).

Practical Surface Analysis
by Auger and X-ray Photoelectron Spectroscopy
Edited by D. Briggs and M. P. Seah

Chapter 3
Spectral Interpretation

D. Briggs
ICI PLC, Petrochemicals and Plastics Division, Wilton, Middlesbrough, Cleveland, UK

J. C. Rivière
UK Atomic Energy Research Establishment, Harwell, Didcot, Oxfordshire

3.1 Introduction

In this chapter are considered the features which may be found in secondary electron and in X-ray photo-electron spectra, from the viewpoints both of their physical origin and of their information content. Since Auger features are contained in both spectra, while X-ray-excited spectra contain in addition photo-electron features, it is logical to describe the secondary electron spectrum first. Before doing either, however, it is necessary to discuss the derivation of the nomenclature used for the description of the two principal types of spectral feature.

3.2 Nomenclature

For a proper discussion of the nomenclature used in the two techniques being described here, one must go back to details of the momenta associated with the orbiting paths of electrons around atomic nuclei. Since an electron is a charged particle, its orbit around a nucleus induces a magnetic field whose intensity and direction depend on the electron velocity and on the radius of the orbit, respectively. Clearly the two latter quantities can be characterized by an angular momentum, called the *orbital* angular momentum, which of course is quantized since the electron can travel only in certain discrete orbitals. The characteristic quantum number is l, and l can take the values 0, 1, 2, 3, 4, . . . Another property of an orbiting electron is the electron spin, i.e. positive or negative, which also induces an inherent magnetic field; the latter

in turn has an associated *spin* momentum, characterized by a spin quantum number s, which can take either of the values $\pm\frac{1}{2}$. Thus the *total* electronic angular momentum is a combination of the orbital angular and spin momenta, and this combination is in fact simply the vector sum of the two momenta. However, it is most important in the present context to note that the vector summation can be carried out in *two ways*, rejoicing in the names of j–j coupling and L–S (also called Russell–Saunders) coupling, respectively.

3.2.1 j–j coupling

In this summation, the total angular momentum of a single isolated electron is obtained by summing vectorially the individual electronic spin and angular momenta. For the particular electron the *total* angular momentum is then characterized by the quantum number j, where $j = l + s$. Obviously j can take the values $\frac{1}{2}, \frac{3}{2}, \frac{5}{2}$, etc. To arrive at the total angular momentum for the whole atom, summation is then performed for all electrons, the result being the total atomic angular momentum with an associated quantum number J, where $J = \Sigma j$. This description of the summation is known as j–j coupling.

Strictly speaking, j–j coupling is the best description of electronic interaction in elements of high atomic number, i.e. $Z > \sim 75$, but in fact the nomenclature based on it has been used for both Auger and spectroscopic features for all parts of the Periodic Table. This does not matter for the features arising from photo-electron production, since the final state of the atom is singly ionized, but in the Auger process, where the final state is doubly ionized, interactions between the two holes in the final state can lead to situations where the j–j description is inadequate. This will be discussed below.

Under the j–j coupling scheme the nomenclature is based on the principal quantum number n and on the electronic quantum numbers l and j mentioned above. In the historical X-ray notation, states with $n = 1, 2, 3, 4, \ldots$ are designated K, L, M, N, ..., respectively, while states with various combinations of $l = 0, 1, 2, 3, \ldots$ and $j = \frac{1}{2}, \frac{3}{2}, \frac{5}{2}, \frac{7}{2}, \ldots$ are given conventional suffixes, 1, 2, 3, 4, ..., according to the listing in Table 3.1. The X-ray notation is almost always used for Auger transitions, so that, for example, in j–j coupling there would be six predicted KLL transitions, i.e. KL_1L_1, KL_1L_2, KL_1L_3, KL_2L_2, KL_2L_3 and KL_3L_3. This is still the most universally used nomenclature in AES.

The spectroscopic nomenclature is directly equivalent to the X-ray, and is more obviously related to the various quantum numbers. In it the principal quantum number appears first, then states with $l = 0, 1, 2, 3, \ldots$ are designated $s, p, d, f, \ldots$, respectively, and follow the first number, and finally the j values are appended as suffixes. Thus the state written L_3 in the X-ray notation, in which $n = 2$, $l = 1$ and $j = \frac{3}{2}$, would be written $2p_{3/2}$ in the spectro-

Table 3.1 X-ray and spectroscopic notation

Quantum numbers					
n	l	j	X-ray suffix	X-ray level	Spectroscopic level
1	0	$\frac{1}{2}$	1	K	$1s_{1/2}$
2	0	$\frac{1}{2}$	1	L_1	$2s_{1/2}$
2	1	$\frac{1}{2}$	2	L_2	$2p_{1/2}$
2	1	$\frac{3}{2}$	3	L_3	$2p_{3/2}$
3	0	$\frac{1}{2}$	1	M_1	$3s_{1/2}$
3	1	$\frac{1}{2}$	2	M_2	$3p_{1/2}$
3	1	$\frac{3}{2}$	3	M_3	$3p_{3/2}$
3	2	$\frac{3}{2}$	4	M_4	$3d_{3/2}$
3	2	$\frac{5}{2}$	5	M_5	$3d_{5/2}$
	etc.		etc.	etc.	etc.

scopic notation. In Table 3.1 the spectroscopic terms are listed opposite their X-ray equivalents. It is conventional to identify a photo-electron feature in terms of the spectroscopic name of the atomic level from which the photo-electron was ejected.

3.2.2 *L–S* coupling

The other way of carrying out the vectorial summation is first to sum all the individual electronic angular momenta and then all the individual electronic spin momenta. These two total momenta are then characterized by two quantum numbers, the total atomic *orbital* angular momentum quantum number L, which is equal to Σl, and the total atomic *spin* quantum number S, which is equal to Σs. Coupling of the two total momenta to give the total atomic angular momentum can then be characterized as before by the quantum number J, which is now, however, equal to $|L \pm S|$. Since L and S can take the values 0, 1, 2, 3, . . ., J can take any integral value between $|L-S|$ and $|L+S|$. The origin of the name 'L–S coupling' is obvious.

L–S coupling has been found to apply to elements of low atomic number, i.e. $Z < \sim 20$. In this coupling scheme the nomenclature is that of term symbols of the form $^{(2S+1)}L$ describing the electron distribution in the final state. By analogy with the spectroscopic notation of Table 3.1, states with $L = 0, 1, 2, 3, \ldots$ are designated in capitals $S, P, D, F, \ldots$, while the total spin quantum number S enters as the prefix $(2S + 1)$. (The state S corresponding to $L = 0$ should not here be confused with S, the total spin quantum number.) As for j–j coupling, L–S coupling predicts six possible components in the KLL series, listed in Table 3.2, but one of these is forbidden through the

Table 3.2 Notation in L–S coupling

Transition	Configuration	L	S	Term
KL_1L_1	$2s^02p^6$	0	0	^1S
$KL_1L_{2,3}$	$2s^12p^5$	1	0	^1P
		1	1	^3P
$KL_{2,3}L_{2,3}$	$2s^22p^4$	0	0	^1S
		[1	1	$^3P]^*$
		2	0	^1D

*Forbidden.

principle of conservation of parity[1]. Also shown in Table 3.2 is a frequently used way of describing the final-state configuration following an Auger transition ABC by writing down the electron populations in the levels B and C, e.g. $2s^12p^5$ for $KL_1L_{2,3}$.

The L–S classification and notation has been used mostly by those recording Auger spectra at high energy resolution in order to provide data for comparison with theoretical calculations, and is not normally encountered in everyday use. The same is true of the mixed, or intermediate, coupling scheme which must be used in the region of the Periodic Table where neither L–S nor j–j coupling is adequate to describe the final-state configuration. In intermediate coupling each L–S term is split into multiplets of different J values, so that the term symbols are now of the form $^{(2S+1)}L_J$. As can be seen from Table 3.3, ten possible final states are predicted in the KLL series, but one is forbidden for the same reasons as before. There have been several examples of experimental evidence for the existence of nine lines in a KLL spectrum.

Table 3.3 Notation in intermediate coupling

Transition	Configuration	L–S term	L	S	J	IC term
KL_1L_1	$2s^02p^6$	^1S	0	0	0	$^1S_0$
$KL_1L_{2,3}$	$2s^12p^5$	^1P	1	0	1	$^1P_1$
		^3P	1	1	0	$^3P_0$
			1	1	1	$^3P_1$
		^1S	1	1	2	$^3P_2$
$KL_{2,3}L_{2,3}$	$2s^22p^4$	^3P	0	0	0	$^1S_0$
			1	1	0	$^3P_0$
			[1	1	1	$^3P_1]^*$
		^1D	1	1	2	$^3P_2$
			2	0	2	$^1D_2$

*Forbidden.

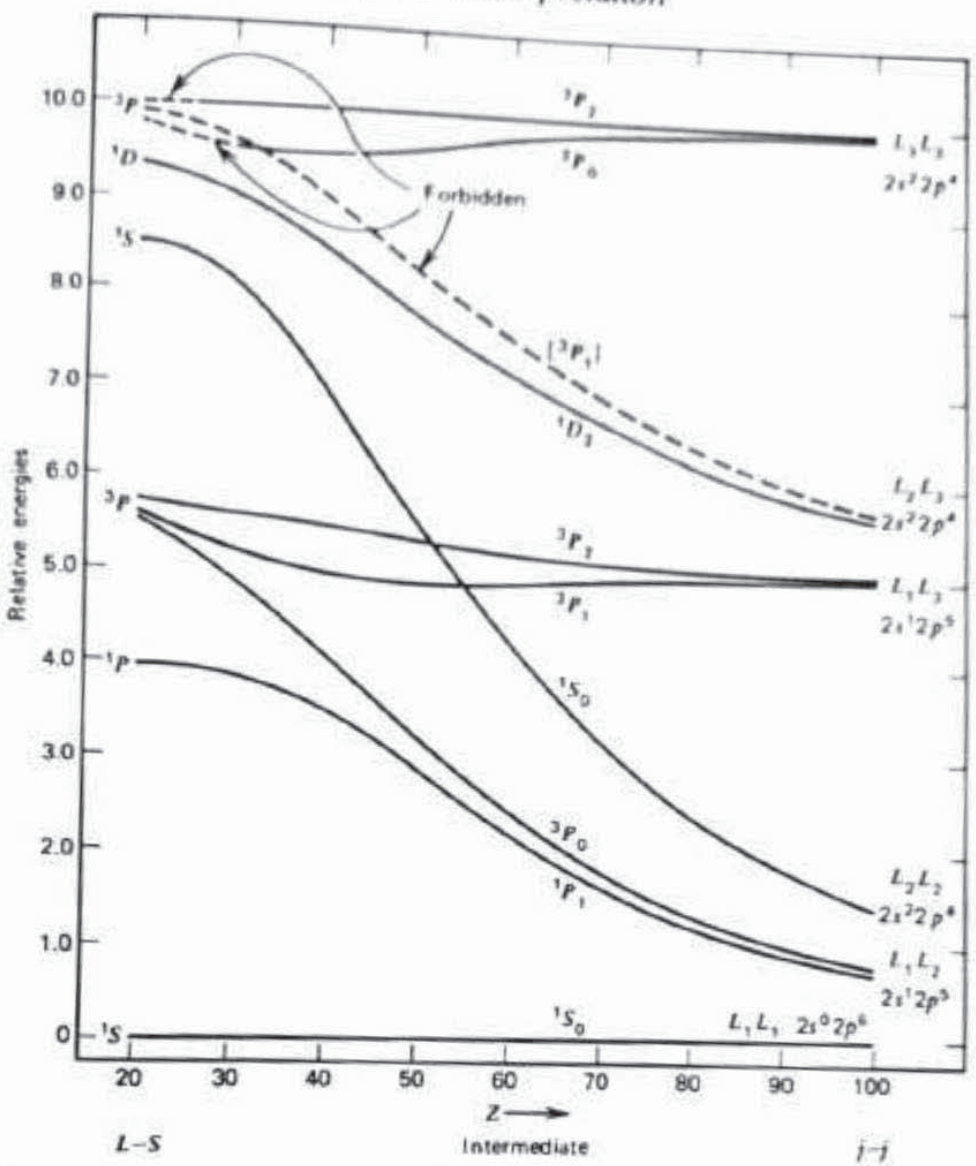


Figure 3.1 The transition from pure L–S through intermediate to pure j–j coupling across the Periodic Table, and its effects on relative energies in the KLL Auger series. (Reproduced from Sevier[2] by permission of John Wiley & Sons Inc.)

As Table 3.3 implies, it is customary to use a mixed notation, so that the KLL series, in approximate order of increasing energy, would be $KL_1L_1(^1S_0)$, $KL_1L_2(^1P_1)$, $KL_1L_2(^3P_0)$, $KL_2L_2(^1S_0)$, $KL_1L_3(^3P_1)$, $KL_1L_3(^3P_2)$, $KL_2L_3(^1D_2)$, $KL_2L_3(^1P_1)$ (forbidden), $KL_2L_3(^3P_0)$ and $KL_3L_3(^3P_2)$. The transition in the KLL series across the Periodic Table from pure L–S through intermediate to pure j–j coupling is shown in Figure 3.1.

3.3 The Electron-excited Secondary Electron Spectrum

When the energies of secondary electrons produced as a result of irradiation of a solid surface by an incident electron beam are analysed, the energy

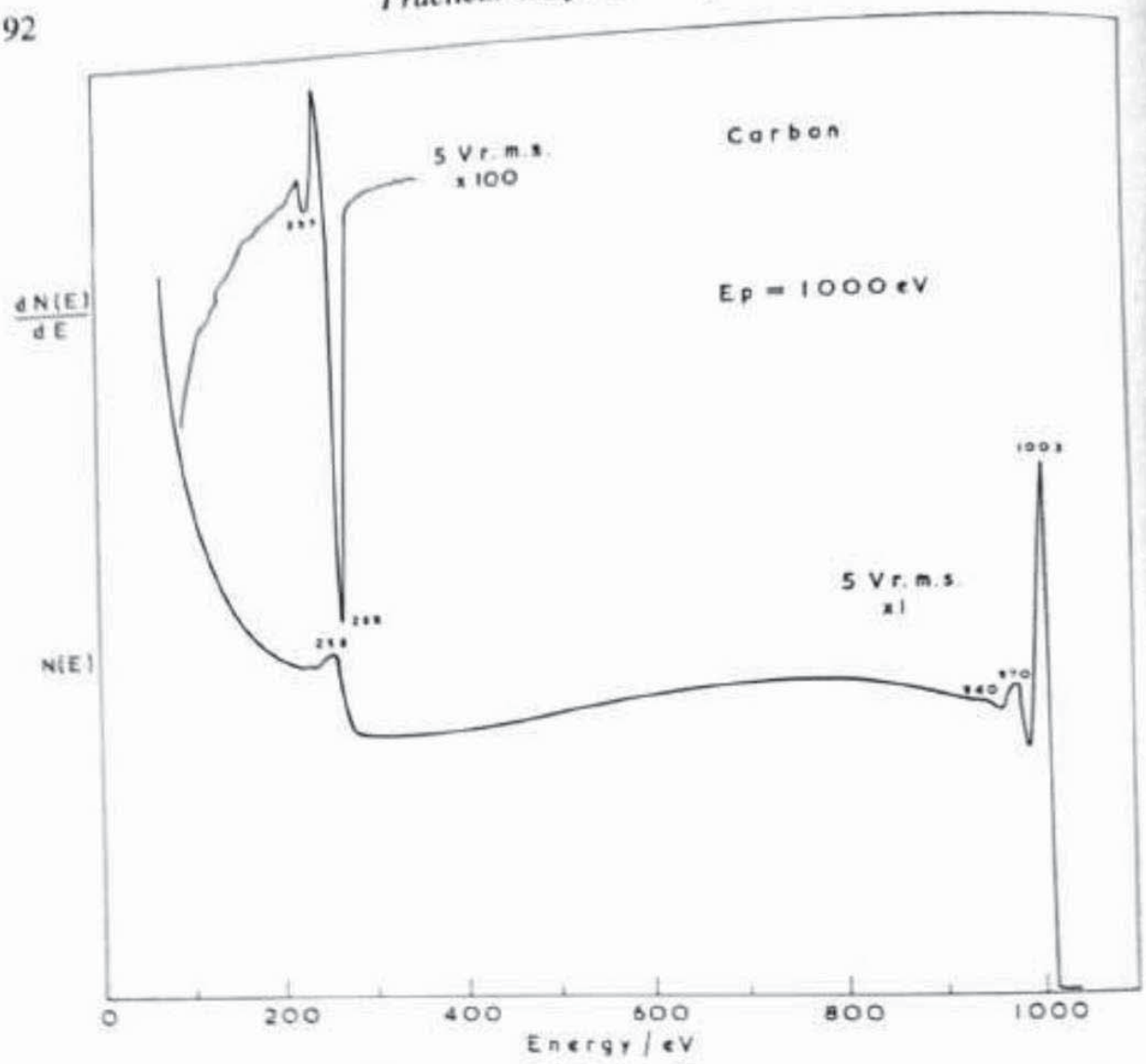


Figure 3.2 Lower curve: distribution of energies of secondary electrons ejected from a graphite surface by incident electrons of energy 1000 eV. Upper curve: differential distribution over the energy range containing the carbon *KLL* Auger peaks. In the differential distribution the peak 'position' is taken to be that of the high energy minimum, by convention

distribution of the ejected electrons looks typically like that in the lower part of Figure 3.2. At the energy of the primary electrons, there is a large, narrow, peak in intensity corresponding to those electrons that have been reflected from the surface with virtually no loss in energy, i.e. elastically scattered electrons. If the surface is that of a single crystal then these electrons contain structural information by virtue of diffraction from the regular atomic array, and are in fact used in LEED and RHEED. Associated with the elastic peak at lower energies is a series of smaller, broader, peaks whose intensity decreases successively away from the elastic peak. These are the plasmon loss peaks. At the other end of the energy distribution can be seen another large peak, very broad, peaking at a few electronvolts and extending many tens of electron-volts up the energy scale. In that peak are the so-called 'true' secondary

electrons; of course, *all* electrons ejected from a surface on primary electron irradiation are secondaries, but conventional terminology distinguishes the true secondaries from the others, mainly for historical reasons. The many investigations carried out under the heading of 'secondary electron emission' have all been concerned with the low-energy secondaries, which are produced by a cascade process.

Between the elastic and the true secondary peak, and often actually situated on the slope of the latter, are other features, some distinguishable easily as peaks, and some not, due mostly to Auger emission. Those few that are not Auger peaks are either associated plasmon loss peaks or ionization loss peaks. In conventional AES, where the highest spatial resolution is not sought and in which therefore the current in the incident electron beam can be maintained at 10^{-6}–10^{-7} A, it is normal to record the spectrum as the differential with respect to energy. Thus the energy distribution in the lower half of Figure 3.2 would appear as the differential distribution shown in the upper half of the figure. Not only does differentiation help to make weak features more readily identifiable, but it also removes the background to a large extent; it is particularly valuable in revealing more clearly features situated on the steep slopes of the 'true' secondary electron peak.

In the following section the features and their origins, arising from Auger emission and from various loss processes, will be described. Most of the description will be based on the differential spectrum, since it is still the one most commonly encountered.

3.3.1 Auger electron spectra

3.3.1.1 The Auger process

The left-hand side of Figure 3.3 shows a schematic energy level diagram of a solid, energy being measured downwards from an assumed zero of energy at the Fermi level. More formally, the zero should be taken as that at the vacuum level at an infinite distance, but in both XPS and AES it is normal to measure binding energies with respect to the Fermi level.

In the centre of Figure 3.3 is shown the sequence of events following ionization of a core level. For this example the K level is shown as being ionized by an incident electron, whose energy E_p must obviously be greater than the binding energy E_K of an electron in K. Due to the way in which the ionization cross-section depends on E_p it is in fact necessary in practice for E_p to be greater than about $5E_K$ for efficient ionization, as will be shown in Chapter 5. Following creation of a hole in the level K, the atom relaxes by filling the hole via a transition from an outer level, in this example shown as L_1. As a result of that transition the energy difference $(E_K - E_{L_1})$ becomes available as excess kinetic energy, and this excess energy can be used by the

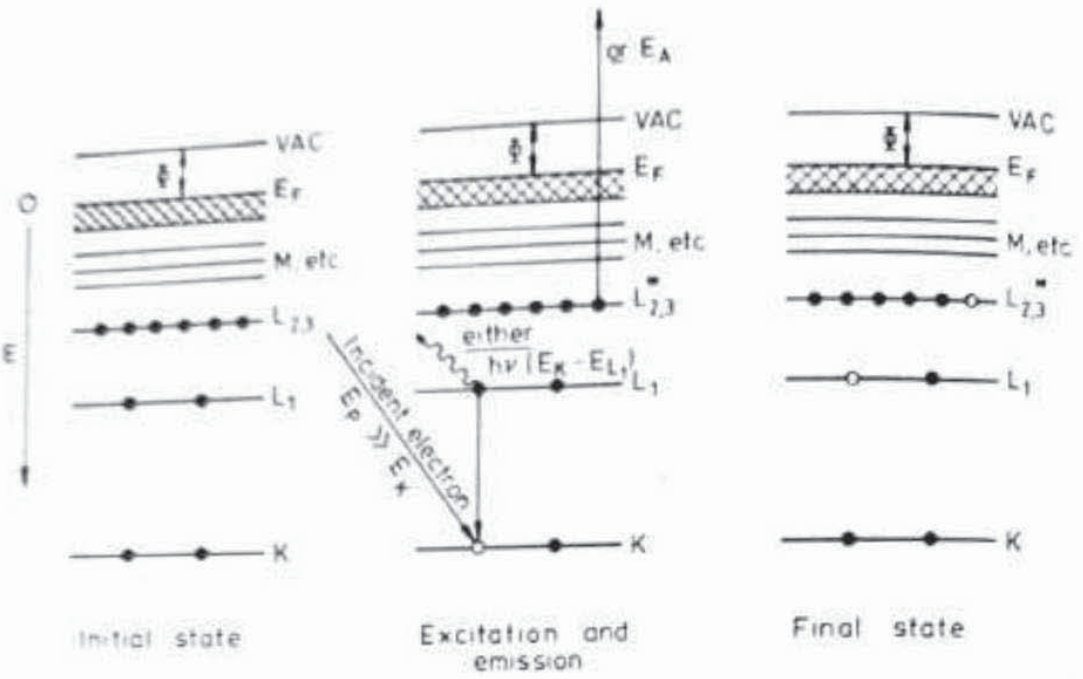


Figure 3.3 Schematic diagram of the process of Auger emission in a solid. The ground state of the system is shown at the left. In the centre an incident electron of energy E_p has created a hole in the core level K by ionization; for this to occur efficiently E_p should be $\geq \sim 5E_K$. The hole in the K shell is filled by an electron from L_1, releasing an amount of energy $(E_K - E_{L_1})$, which can appear as a photon of energy $h\nu = (E_K - E_{L_1})$ or can be given up to another electron. In this example the other electron is in the $L_{2,3}$ shell, and it is then ejected with energy $(E_K - E_{L_1} - E^*_{L_{2,3}})$; $E^*_{L_{2,3}}$ is starred because it is the binding energy not of $L_{2,3}$ in its ground state, but in the presence of a hole in L_1. The doubly ionized final state is shown on the right

atom in either of two ways. It can appear as a characteristic X-ray photon at that energy *or* it can be given to another electron either in the same level or in a more shallow level, whereupon the second electron is ejected. The first process is that of X-ray fluorescence, the second that of Auger emission. Clearly both cannot take place from the same initial core hole, so that they compete. However, as Figure 3.4 indicates for K shell ionization, the probability of relaxation by Auger emission is favoured overwhelmingly over that of X-ray fluorescence for relatively shallow core levels, i.e. with binding energies below about 2 keV. The same is true for the L, M, N, etc., atomic levels. If this were not so, signal strengths in AES would be much smaller and it would not therefore be as useful a technique as it is.

The Auger transition depicted in Figure 3.3 would be named in the conventionally used j–j coupling $KL_1L_{2,3}$. Obviously in the same notation other transitions are possible in the atom depicted in that figure, e.g. KL_1L_1, $KL_{2,3}L_{2,3}$, $L_1L_{2,3}L_{2,3}$, etc. The electrons taking part in the Auger process might also originate in the valence band of the solid, in which case the convention writes the transitions as, for example, $KL_{2,3}V$ if one electron comes from the valence band and, for example, KVV if both do.

The energy of the ejected Auger electron in the example of Figure 3.3 is

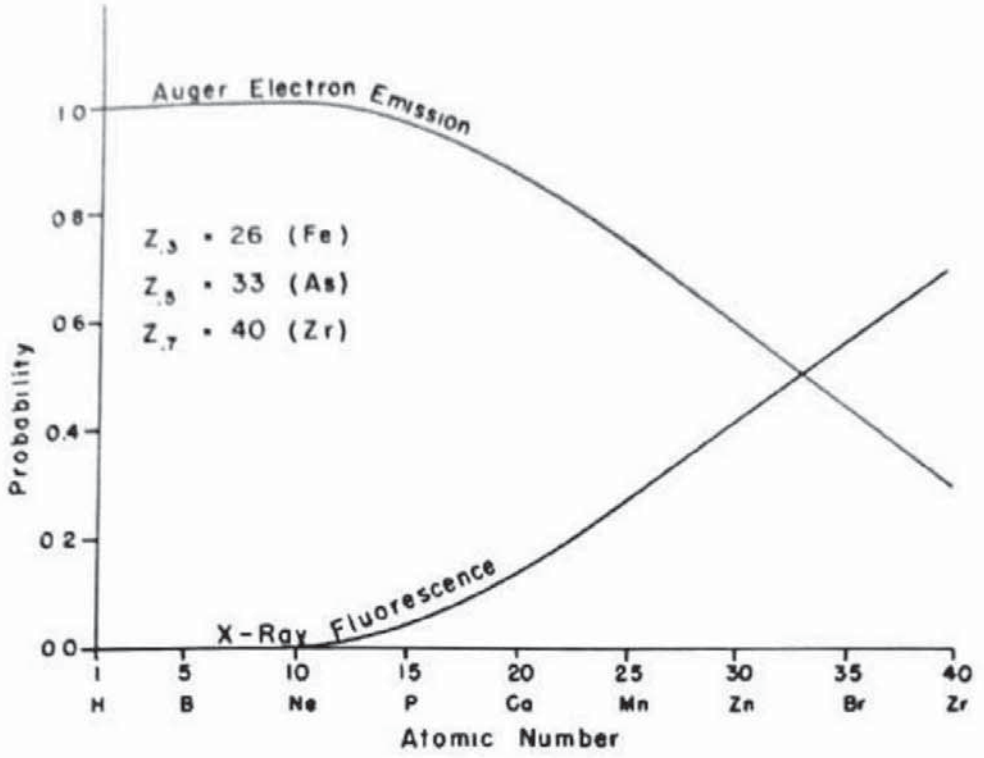


Figure 3.4 Relative probabilities of relaxation by emission of an Auger electron and by emission of an X-ray photon of characteristic energy, following creation of a core hole in the K shell

$$E_{KL_1L_{2,3}} = E_K - E_{L_1} - E^*_{L_{2,3}} \tag{3.1}$$

where E_i are the binding energies of the ith atomic energy levels. $E^*_{L_{2,3}}$ is starred because it is the binding energy of the $L_{2,3}$ level in the presence of a hole in level L_1, and is therefore different from $E_{L_{2,3}}$. The various contributions to the difference between E_i^* and E_i, in the general case, will be discussed below. It is sufficient for the moment to note that the Auger energy expressed by equation (3.1) is a function only of atomic energy levels, so that for each element in the Periodic table there is an unique set of Auger energies, there being no two elements with the same set of atomic binding energies. Thus analysis of Auger energies immediately leads to elemental identification. Even when either or both electrons in the Auger process originate in the valence band, the analysis provides elemental identification, since the dominant term in equation (3.1) is always the binding energy of the initially ionized core level.

For heavy elements it will be clear from equation (3.1) that the number of energetically possible Auger transitions, for a given primary electron energy, can become very large indeed, as the number of atomic levels proliferates. Fortunately for AES, the transition probabilities favour only a few of the very many, so that even for the heaviest elements the problem is not intractable.

3.3.1.2 Auger energies

An approximation to the energy E_{ABC} of the Auger transition ABC in an atom of atomic number Z, that has proved sufficiently accurate for most practical purposes, is the one derived empirically by Chung and Jenkins.[3] It is

$$E_{ABC}(Z) = E_A(Z) - \tfrac{1}{2}[E_B(Z) + E_B(Z+1)] - \tfrac{1}{2}[E_C(Z) + E_C(Z+1)] \quad (3.2)$$

$E_i(Z)$ are the binding energies of the ith levels in the element of atomic number Z and $E_i(Z+1)$ of the same levels in the next element up the periodic table, various compilations of E_{ABC}, such as that of Coghlan and Clausing[4], are available, based on equation (3.2).

A more physically acceptable expression for the Auger energy is

$$E_{ABC} = E_A - E_B - E_C - \mathcal{F}(BC:x) + R_x^{in} + R_x^{ex} \quad (3.3)$$

where $\mathcal{F}(BC{:}x)$ is the energy of interaction between the holes in B and C in the final atomic state x and the R_x are relaxation energies. The latter energy terms arise from the additional screening of the atomic core necessary when there is a hole in a core level, and it is achieved by an inward collapse or 'relaxation' of outer electronic orbitals towards the core. As the relaxation takes place, there is an increased interaction of the electrons in the outer orbitals with electrons in the inner orbitals, via Coulomb and exchange integrals. The term R_x^{in} is the *intra-atomic* relaxation energy, i.e. the relaxation energy appropriate to an isolated atom. In a molecule or in a solid, where additional screening electrons are available from the environment of the ionized atom, i.e. from other atoms or from the valence band, respectively, then the additional R_x^{ex} term appears, called the *extra-atomic* relaxation energy. The magnitudes of the $\mathcal{F}$ and R terms are often considerable. Table 3.4 gives examples calculated for the $L_3M_{4,5}M_{4,5}$: ^3P transitions of metallic nickel, copper and zinc by Kim, Gaarenstroom and Winograd.[5]

In the calculation of Auger energies using equation (3.3) it is customary to use experimentally determined binding energies for the E_i and calculated values for the other terms, so that the approach is semi-empirical. Various

Table 3.4 Calculated values of the $\mathcal{F}$ and R energies in equation (3.3) for the $L_3M_{4,5}M_{4,5}$: ^3P transition in Ni, Cu and Zn. (From Kim, Gaarenstroom and Winograd[5])

Metal	$\mathcal{F}(M_{4,5}M_{4,5}:{}^3P)$	$R^{in}_{^3P}$	$R^{ex}_{^3P}$
Ni	26.6	9.9	18.2
Cu	26.3	10.6	11.0
Zn	29.4	12.3	9.6

authors have provided tables of such semi-empirically determined Auger energies, the most extensive tabulation being that of Larkins,[6] who has covered the Auger series KLL, KLM, L_3MM, L_3MN, M_4NN, M_5NO, and N_5OO, all within the intermediate coupling system.

This identification of features in the secondary electron spectrum as arising from Auger emission could be made, in principle, by checking the observed kinetic energies against the energies calculated from expressions (3.2) or (3.3), the latter being more accurate and also providing energies of the individual final-state terms. On the other hand, there are by now several atlases of recorded Auger spectra available, in which the principle Auger peaks are labelled with their kinetic energies, making elemental identification easy in most cases. There are slight differences between these compilations, due to different recording conditions and different energy analysers used in each case, but these differences are not usually large enough to matter. It must be remembered that the spectra are in all cases shown in the differential $E\mathrm{d}N(E)/\mathrm{d}E$ mode, in which by *convention* the energetic position of the Auger peak is taken as that of the minimum in the high-energy negative excursion. Clearly such a conventional energy does not correspond to the actual peak energy in the undifferentiated, or direct, spectrum, and will differ from it by an amount of energy that depends on the peak width. Thus close agreement between Auger energies obtained from the differential distribution and those calculated should not be expected, but such a comparison is not usually attempted. In practise, the Auger atlases are used on a day-to-day basis in applied AES, in which the differential mode is still by far the more common, while the results of the detailed energy calculations are used by those working in the undifferentiated mode, generally derived from X-ray rather than electron excitation.

3.3.1.3 Auger intensities

Inspection of any one spectrum, or comparison of the spectra from any two adjacent elements in the above-mentioned atlases, reveals at once that the relative intensities of different Auger transitions within the same element, or of the same Auger transition in different elements, show large variations. The reasons for these variations are in general well known, but, unlike the situation for the calculation of Auger energies, that for the calculation of Auger intensities (or transition probabilities) is not nearly so good. There is no simple semi-empirical formula that will produce a set of usable relative intensities, and even the complex formulations that have to be used by the theoreticians have not yet been particularly successful in the accurate prediction of intensities, especially in the solid state. The relationship between the intensity of an Auger peak from a given element and its concentration, when present with other elements on a surface, will be discussed in Chapter 5, but

in the present context it is sufficient merely to note that intensities cannot be predicted with the same certainty as can energies. Where comparison of intensities is important, recourse must be had once again either to the published Auger atlases or to sets of elemental spectra compiled in one's own laboratory.

3.3.1.4 Characteristic Auger series

Since, as shown in Chapter 5, the dependence of the cross-section for ionization by electrons on the primary electron energy passes through a maximum at about 4–5 times the primary energy, and since a typical primary energy in current use would be 10 keV, there are necessarily limitations in any one part of the Periodic Table on the Auger transitions that can be excited with optimum efficiency. Thus Auger transitions based on K shell ionization are limited in usefulness to the range from lithium (in the solid state) to silicon, on L_3 shell ionization to the range from magnesium to rubidium, on M_5 ionization to the range from gallium to osmium, and so on. The practical effect of these limitations is that in each region of the Periodic Table there is a characteristic set, or series, of Auger transitions which is the most prominent under the standard experimental conditions normally employed. Some examples of

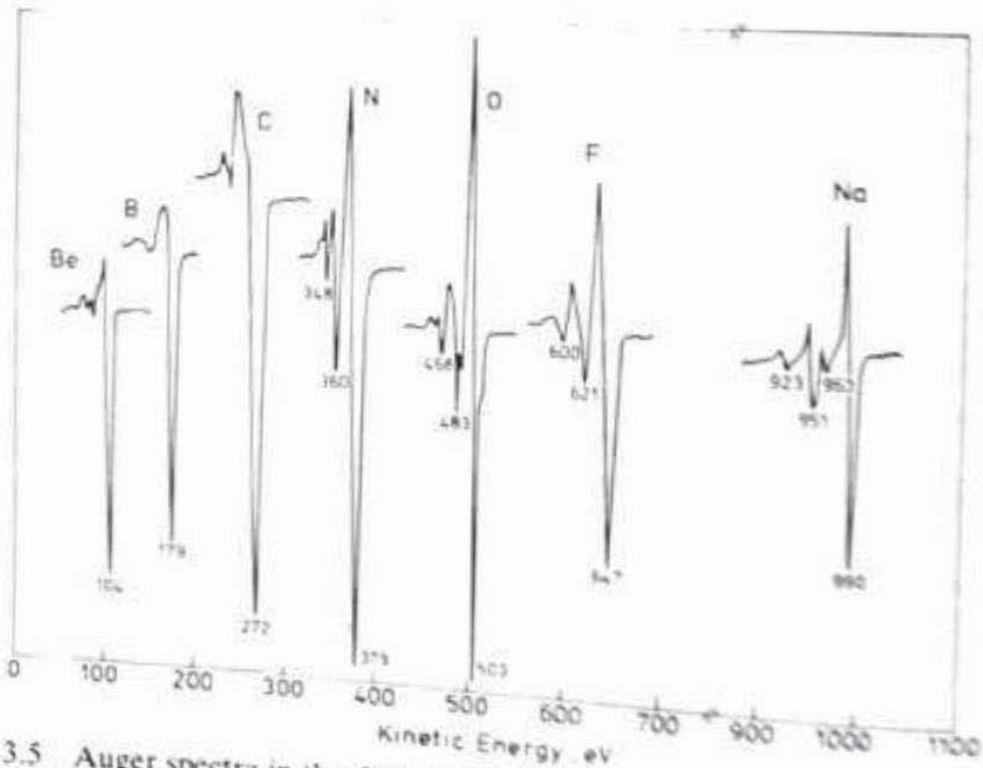


Figure 3.5 Auger spectra in the differential distribution characteristic of the lightest elements. The principal peak is the $KL_{2,3}L_{2,3}$. The relative intensities are not to scale. (Reproduced from Davis *et al.*[7] by permission of Perkin-Elmer Corporation)

these series are shown in Figures 3.5 to 3.9, taken from the *Handbook of Auger Electron Spectroscopy* by Davis *et al.*[7]

In Figure 3.5 it can be seen that for the light elements of the Auger transition $KL_{2,3}L_{2,3}$ (KL_1L_1 for beryllium) dominates for most of the range. For magnesium, aluminium and silicon, which are not illustrated, the low-energy $L_{2,3}VV$ transition also becomes strong and for the elements from phosphorus to calcium it is sufficiently intense to be used for identification, K shell ionization no longer being available as mentioned above. In the series of $3d$ transition metals, from scandium to copper, the characteristic 'fingerprint' is the LMM triplet, consisting, in order of decreasing energy, of the transitions L_3VV,

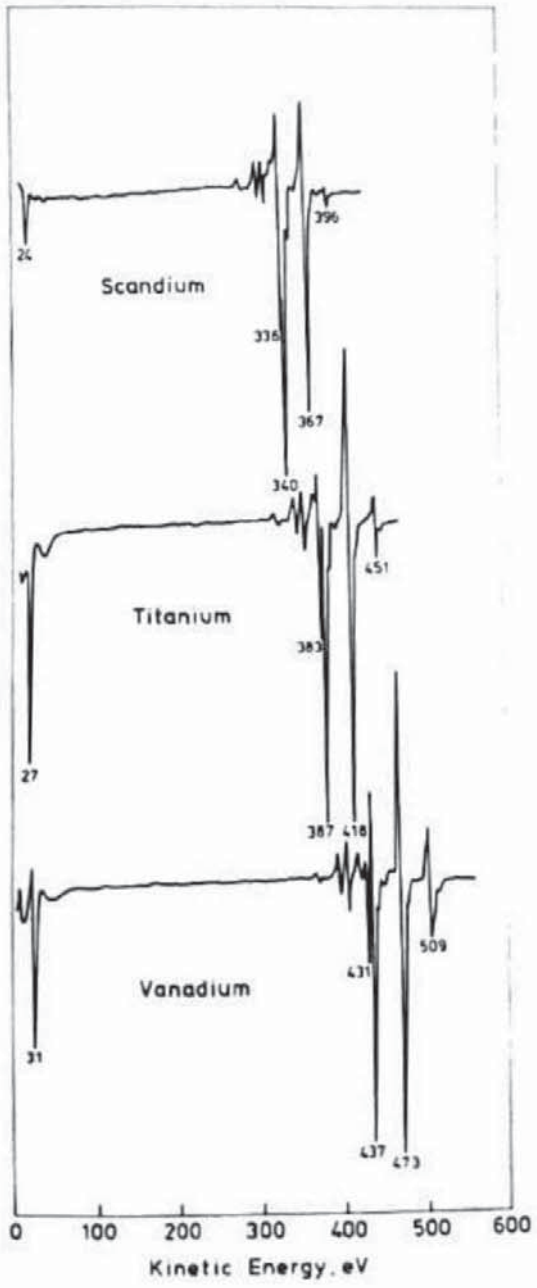


Figure 3.6 Auger spectra in the differential distribution for scandium, titanium and vanadium. The characteristic features are the LMM triplet and the low-energy $M_{2,3}VV$ peak. (Reproduced from Davis *et al.*[7] by permission of Perkin-Elmer Corporation)

$L_3M_{2,3}V$ and $L_3M_{2,3}M_{2,3}$. Figures 3.6 to 3.8 illustrate the progression of this triplet through the series, and it can be seen clearly how the relative magnitudes of the individual components change, particularly that of the L_3VV to the other two (note, also, Figure 5.7 in Chapter 5). The gradual increase in the intensity of the latter transition reflects, of course, the progressive filling of the valence band to completion at copper. Another striking and characteristic Auger series can be found amongst the 4*d* transition metals, and is shown in Figure 3.9. The principal feature there is the closely spaced $M_{4,5}N_{4,5}N_{4,5}$ doublet, the spacing being that between M_4 and M_5. In other regions of the

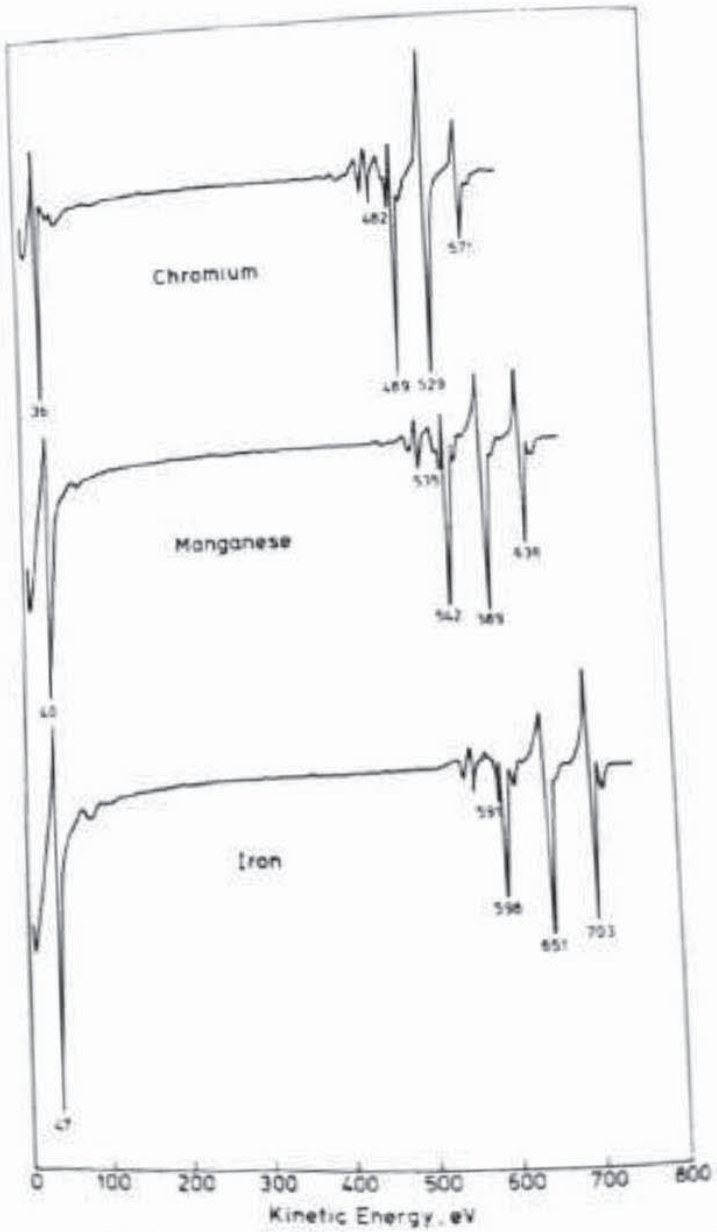


Figure 3.7 Auger spectra in the differential distribution for chromium, manganese and iron. The *LMM* triplet seen at the start of the first transition series in Figure 3.6 is now well developed, and the $L_{2,3}VV$ peak is also stronger. (Reproduced from Davis *et al.*[7] by permission of Perkin-Elmer Corporation)

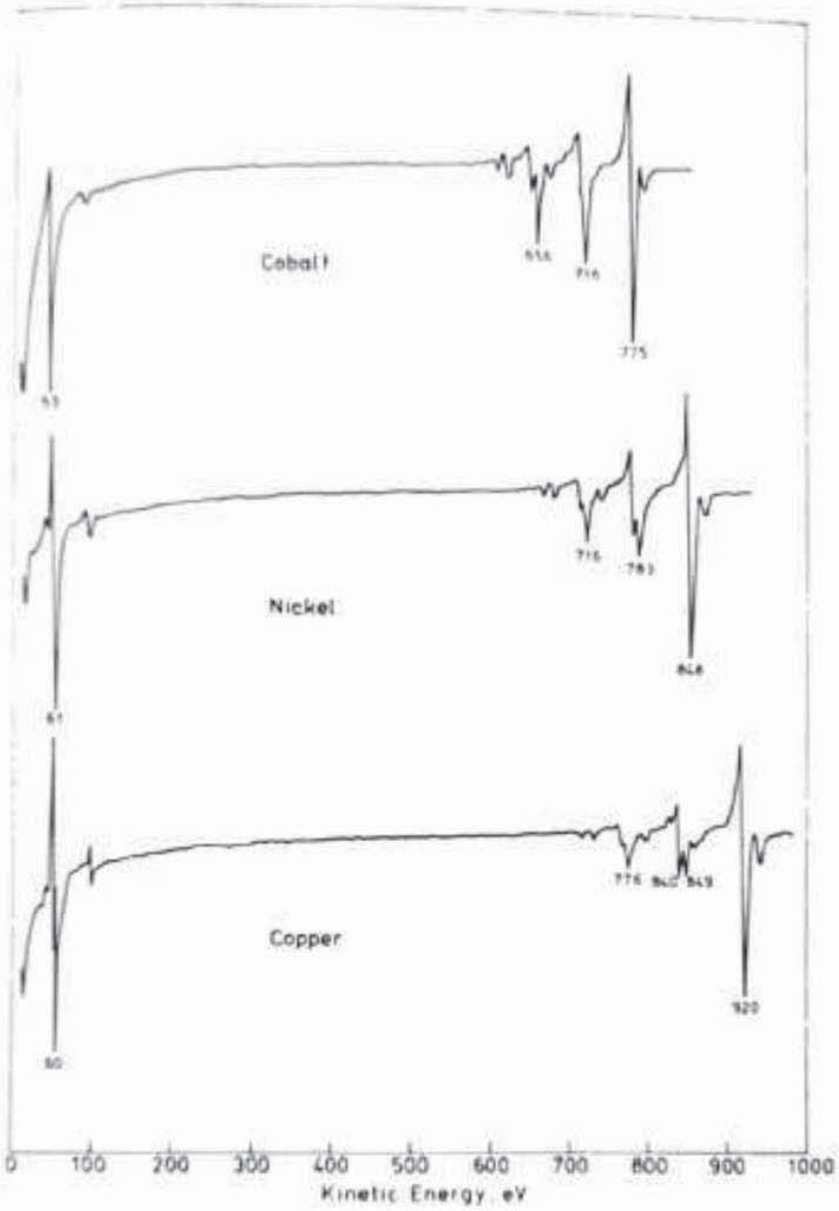


Figure 3.8 Auger spectra in the differential distribution for cobalt, nickel and copper. In the *LMM* triplet the L_3VV peak at the highest indicated energies is now predominant due to the filling of the metal *d* bands, the filling being complete at copper. (Reproduced from Davis *et al.*[7] by permission of Perkin-Elmer Corporation)

Periodic Table there are other series characteristic of the elements in those regions, but none is quite as immediately recognizable as any of those described above.

The Auger spectra in Figures 3.5 to 3.9 also reveal many minor features, mostly due to weak Auger transitions, and obviously in alloys or compounds or other multielement materials there is always present the possiblity of overlap of the Auger features of one or more of the constituent elements. Where, as for example amongst the 3*d* transition metals, there are several prominent transitions, it is always possible to find one that is clear of interference, but problems can arise where there is overlap between the position of the only Auger peak available for a light element and a prominent peak belonging to a

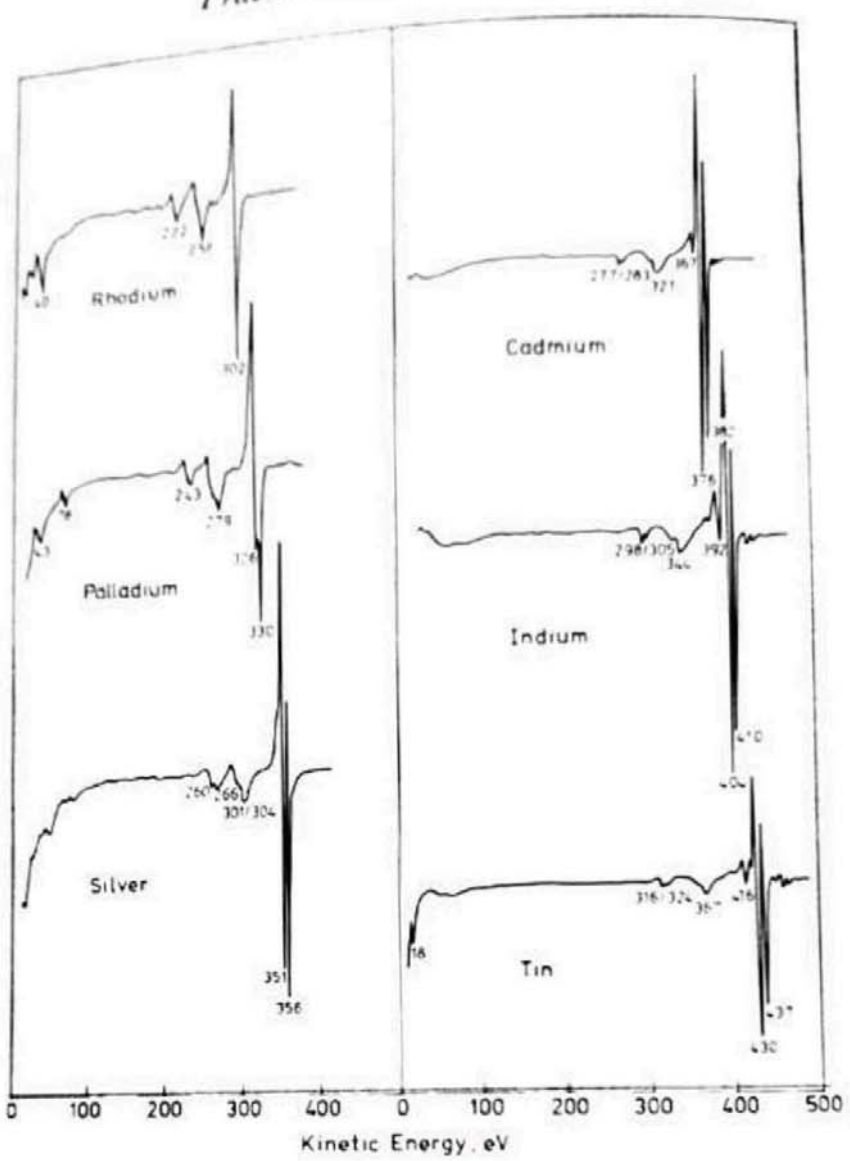


Figure 3.9 Auger spectra in the differential distribution for metals in the second transition series. Apart from rhodium and palladium, where it is poorly resolved, the characteristic feature is the sharp, closely spaced $M_{4,5}N_{4,5}N_{4,5}$ doublet, the separation being that between M_4 and M_5. (Reproduced from Davis *et al.*[7] by permission of Perkin-Elmer Corporation)

heavy element. Examples are chlorine and argon with molybdenum, boron with niobium, sulphur with zirconium and carbon with ruthenium. Where in doubt, it is best to use spectrum subtraction techniques, in which a standard 'clean' spectrum is subtracted from the suspect one.

3.3.1.5 Auger fine structure

Fine structure in Auger spectra can be seen frequently from both metals and non-metals, and originates either in chemical effects or in final-state effects. In general, structure due to chemical effects can usually be observed in conventional electron-excited AES, but that due to final-state or hole localization

effects must be recorded at much higher energy resolution, in the direct or undifferentiated mode, using X-ray excitation.

Chemical effects can arise from the same source as the so-called 'chemical shift' that is the strength of XPS, i.e. the binding energy of electrons in the core level being ionized can alter with chemical environment, leading to a

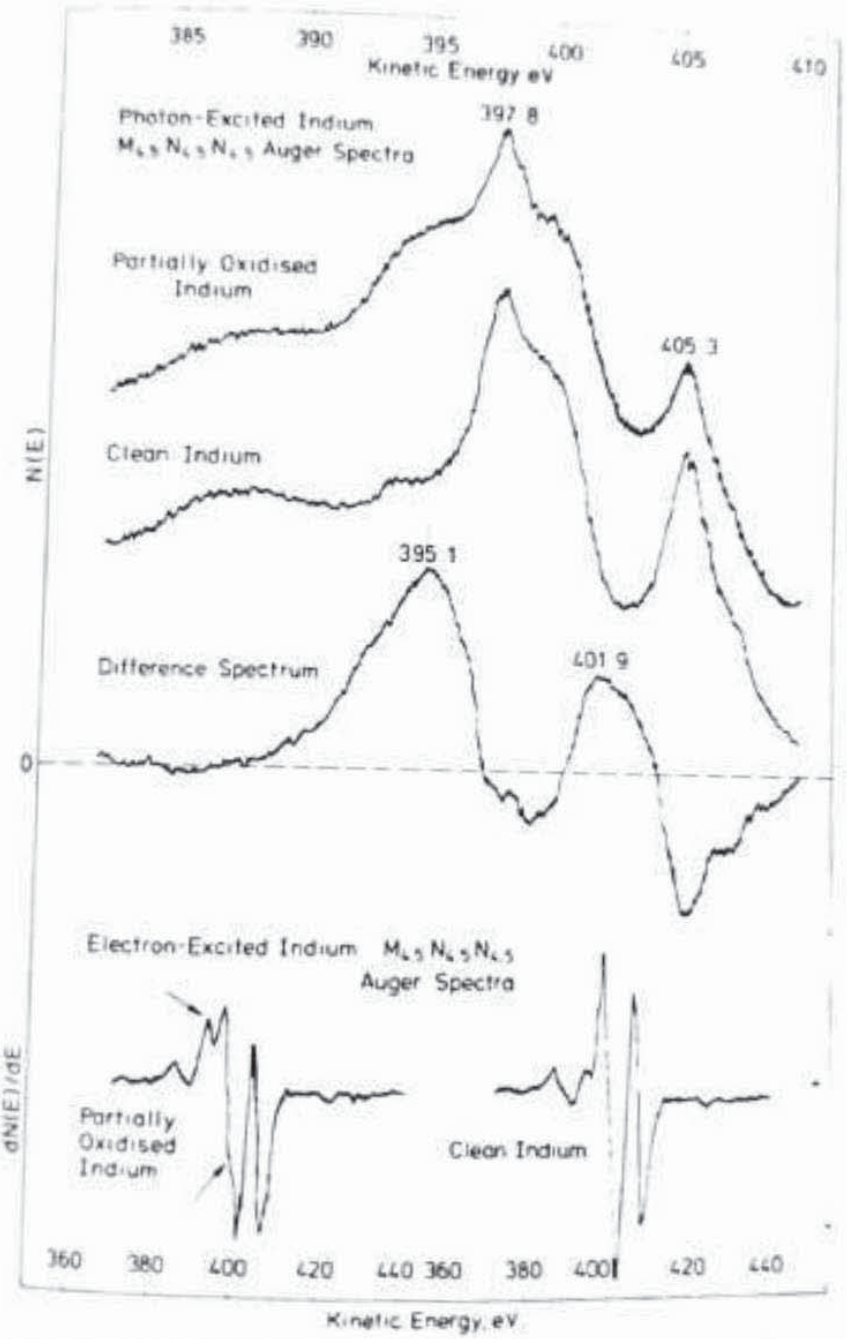


Figure 3.10 Example of fine structure in Auger spectra, due to a chemical shift in the binding energy of the ionized core level. The electron-excited differential Auger spectra of indium in the clean and partially oxidized states shown at the bottom reveal fine structure after partial oxidation, not seen in the clean state. Analysis of the photon-excited and undifferentiated Auger spectra in the top part shows by subtraction that the fine structure originates in shifts of 2.7 and 3.4 eV in the M_4 and M_5 levels on oxidation. The Auger spectrum in the partially oxidized state is thus a superposition of the clean and 'oxidized' spectra

similar shift in the consequent Auger transition. Where atoms of a particular element are present on a surface in more than one chemical state, then there is the consequent possibility of Auger fine structure by this mechanism. An example shown in Figure 3.10 is that of indium where the fine structure arrowed in the conventional differential spectrum at the bottom is seen to be due to a shifted Auger peak when the spectrum is integrated numerically, as shown at the top. In this case it is the shift of 2.8 and 3.4 eV in the M_4 and M_5 core levels, respectively, on oxidation that causes the shift in the $M_{4,5}N_{4,5}N_{4,5}$ Auger transitions and hence the appearance of the additional structure.

The more commonly observed fine structure due to chemical effects, however, arises in the Auger spectra of non-metallic elements present on surfaces in one or more of various possible chemical states. When, for example, such an element is part of a chemisorbed layer on a metallic substrate, the most

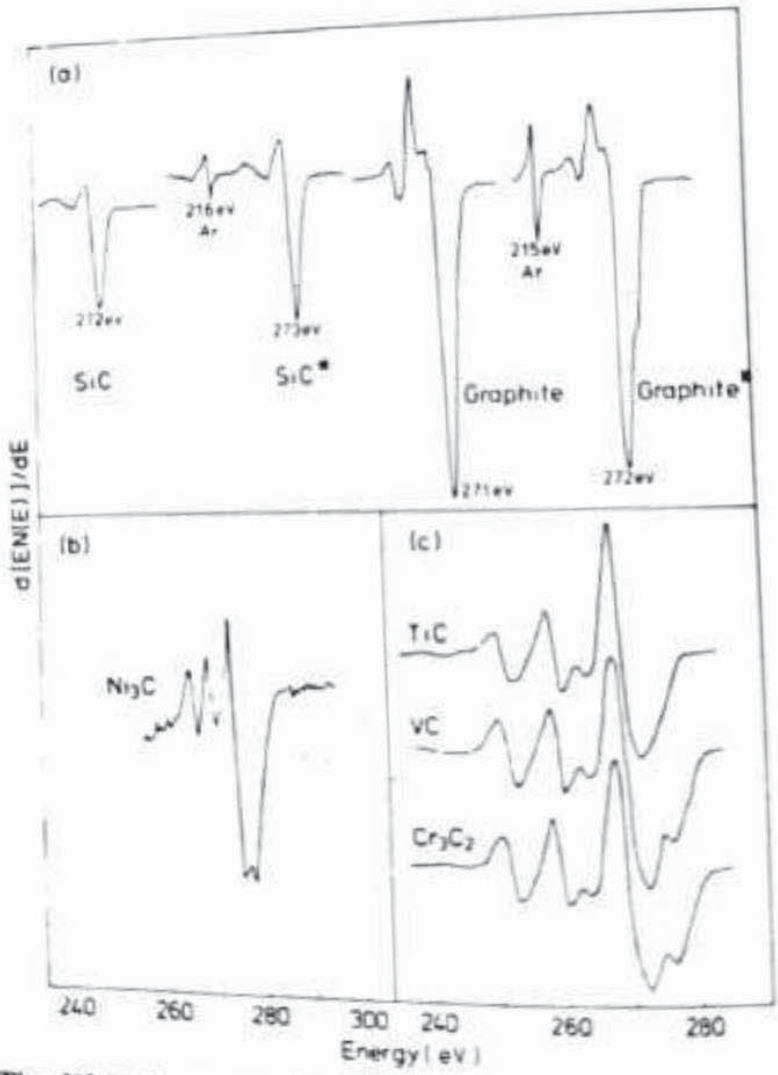


Figure 3.11 The *KLL* Auger spectra of carbon in different chemical situations. In the spectra in the upper part the asterisk * signifies recordings made after ion bombardment. (Reproduced from Kny[8] and Kleefeld[8] by permission of the American Institute of Physics and Elsevier Sequoia)

intense Auger transitions characteristic of the element will be those in which ionization of a core level in the element is followed by transfer of valence electrons *from the substrate* for the subsequent Auger relaxation. In other words, although the average energetic position of the resultant peak in the Auger spectrum will be characteristic of the element, which can thus still be identified, the shape of the peak, including the fine structure, will be related to the local density of the valence states adjacent to the chemisorption site. Since the local density of states (LDOS) will vary from one substrate to another, and from one type of site to another on any one substrate, it follows that variations in the shape and fine structure associated with the Auger peak(s) of the chemisorbed species will be observed.

The earliest example to be observed of these variations, and still one of the most striking, is that of the *KVV* Auger spectrum of carbon in different chemical situations. Figure 3.11, taken from various authors,[8] shows that both the overall appearance of the Auger peak shape, and the associated fine structure change markedly on going from graphite to various carbides, and even after ion bombardment of solid carbon. Note the particularly sharp fine structure present in the carbide spectra. Another example is given in Figure 3.12 for the sulphur $L_{2,3}VV$ Auger peak on three transition metals; the differences are not quite so great as for carbon, but are significant. Similar variations have been recorded for nitrogen, phosphorus and even oxygen,[10] which is more electronegative than the others mentioned here. Such valence spectra can be used either to derive information about the LDOS at the ionized surface site, by self-deconvolution, or, at a lower but more immediately useful level, to act as 'fingerprints' of the particular chemical form or compound present. In the present context of spectral interpretation it is sufficient to point out that the fine structure due to chemical effects is not uncommon and that it must therefore be recognized for what it is.

The other type of fine structure, originating in correlations between the two holes that are left in the atom at the end of the Auger process, is intrinsic to certain metals. In these metals Auger transitions of the type *CVV*, where C is a core level, produce line shapes whose widths even before deconvolution are much narrower than those of the valence bands *V*, and thus there would be no possibility whatsoever of deriving information about the valence density of states in the metals. The widths are in fact similar to those expected from the same transition in the free atom, allowing for solid-state broadening, and the Auger spectra are thus labelled 'quasi-atomic'. The physical explanation is that in such metals the effective interaction energy between the two holes in the final state is much larger than the width of the valence band, and therefore the holes cannot decay rapidly as they would be able to do if their interaction energy were lower. As a result the holes became localized near the ionization atom for a significant time, leading to a situation formally similar to that in an isolated atom. Under these circumstances the electronic correlation between

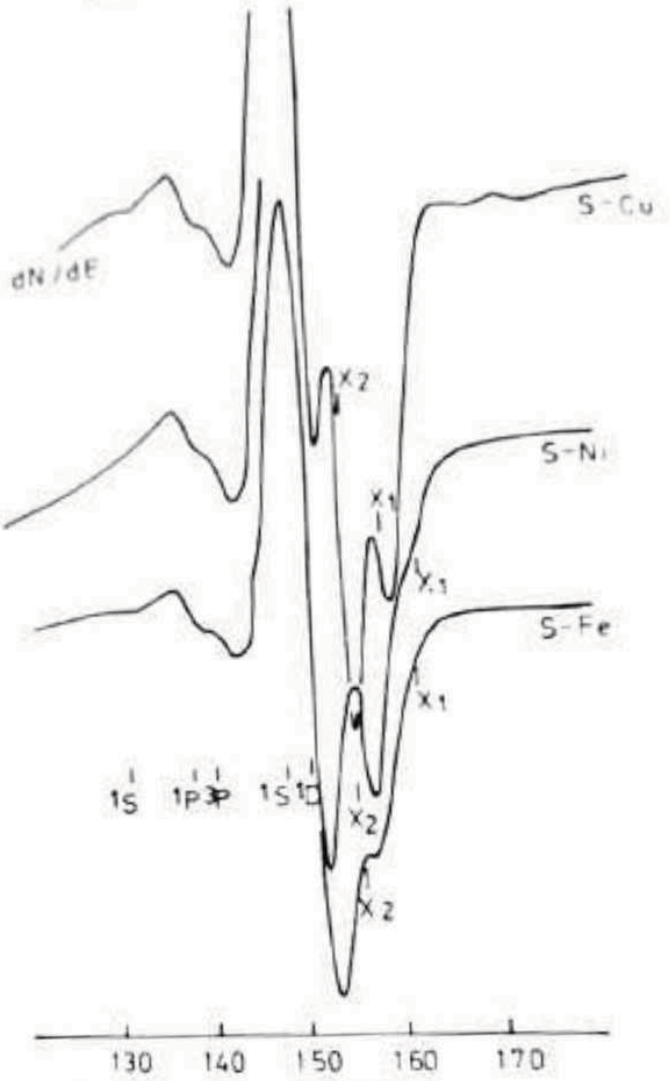


Figure 3.12 The $L_{2,3}VV$ Auger spectra of sulphur on the surfaces of three transition metals, showing chemical effects due to the differences in the bonding and therefore in the local densities of electronic states near the sulphur atoms. Calculated positions of spectral terms are indicated. The peaks labelled X_1 and X_2 are due to so-called 'inter-atomic' transitions. (After Salmeron, Baco and Rojo[9])

the holes gives rise to a splitting of the CVV spectrum into multiplets that can be described by the term symbols given in Section 3.2.2 on L–S and intermediate coupling.

The metals for which the 'quasi-atomic' type of behaviour has definitely been established are those from nickel to selenium, from silver to tellurium and from thallium to bismuth; it is possible there may be others. Some metals for which the 'quasi-atomic' Auger character is not normally observed can show that character when present in a dilute alloy.[11] Since some of the multiplet splittings are small, it is necessary to record the Auger spectra at higher energy resolution than that normally used; this is generally only possible by recording the undifferentiated spectrum by pulse counting in an XPS system, i.e. using X-ray rather than electron excitation. Examples of Auger spectra obtained in this way, showing well-developed fine structure arising from the

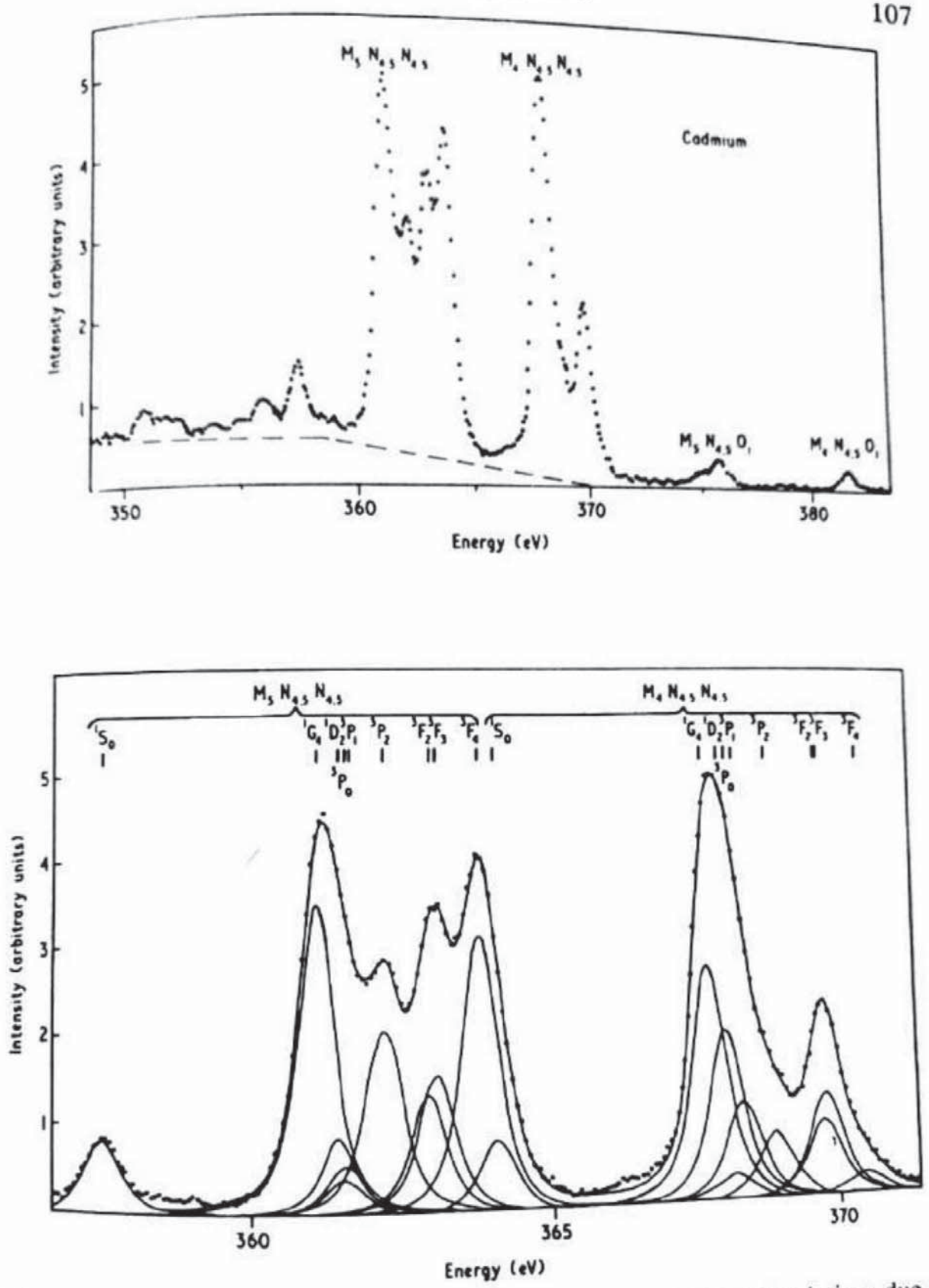


Figure 3.13 Fine structure in the $M_{4,5}N_{4,5}N_{4,5}$ Auger spectrum of cadmium due to multiplet splittings in the final state of the Auger process. The upper part is the observed spectrum with the assumed inelastic background shown as the dashed line. The lower part shows the calculated positions and intensities of the various multiplet components for each group, with the resultant theoretical envelope (solid line) compared with the experimental points. (Reproduced from Aksela and Aksela[12] by permission of the Institute of Physics)

above 'quasi-atomic' behaviour, are given in Figure 3.13. The nomenclature based on j–j coupling labels the two groups of peaks as $M_5N_{4,5}N_{4,5}$ and $M_4N_{4,5}N_{4,5}$ respectively, taking no account of the multiplet splittings. Underneath the experimental spectra are shown the calculated positions and relative intensities of the multiplets corresponding to the nine permitted final states; it is not easy to distinguish all nine since the energies of the $^1P_1$ and $^3P_1$ states are almost the same, but it can be seen that the energetic positions based on atomic structure calculations agree well with experiment.

3.3.1.6 *Plasmon loss features*

Any electron of sufficient energy passing through a solid can excite one or other of the modes of collective oscillation of the sea of conduction electrons. These oscillations have frequencies characteristic of the material of the solid, and therefore need characteristic energies for excitation. An electron that has given up an amount of energy equal to one of these characteristic energies, in the course of excitation, is said to have suffered a plasmon loss. Within the solid the loss is said to be that of a 'bulk' plasmon, and if the fundamental characteristic frequency of the bulk plasmon is ω_b, then the plasmon energy loss is clearly $h\omega_b$. Since electrons that have suffered a plasmon loss in energy can themselves suffer further losses of this kind in a sequential fashion, then a series of losses, all equally spaced by $h\omega_b$ but of decreasing intensity, will occur.[13]

At a surface the regular atomic lattice of the solid terminates and the conditions for the setting up of oscillations of frequency ω_b are no longer satisfied. Instead, a rather localized type of collective oscillation can be excited, of fundamental frequency ω_s, referring to the surface, where ω_s is less than ω_b. For a free electron metal ω_s can be shown[13] theoretically to be equal to $\omega_b/\sqrt{2}$, or more generally[14] to $\omega_b/(1 + \varepsilon)^{1/2}$, where ε is the dielectric constant.

When, therefore, the energies of electrons that have been ejected from a solid are analysed, there will be found associated with prominent peaks in the spectrum a series of plasmon loss peaks. The fundamental or 'first' plasmon loss will always be visible, and, dependent on the material and the experimental conditions, several multiple plasmon losses of decreasing size may also be visible. According to the surface condition of the material and to the kinetic energy of the parent peak, surface plasmon loss peaks may be discernible, but will always be of a lower intensity than the adjacent bulk plasmon peak.

A very clear example of an Auger spectrum containing all the above-mentioned features is given in Figure 3.14, in which the various losses are associated with the KLL peaks of aluminium. In the figure the KLL transitions are shown as excited by both electrons and soft X-ray photons, demonstrating that, for the same energy resolution, the shape of an Auger peak is the

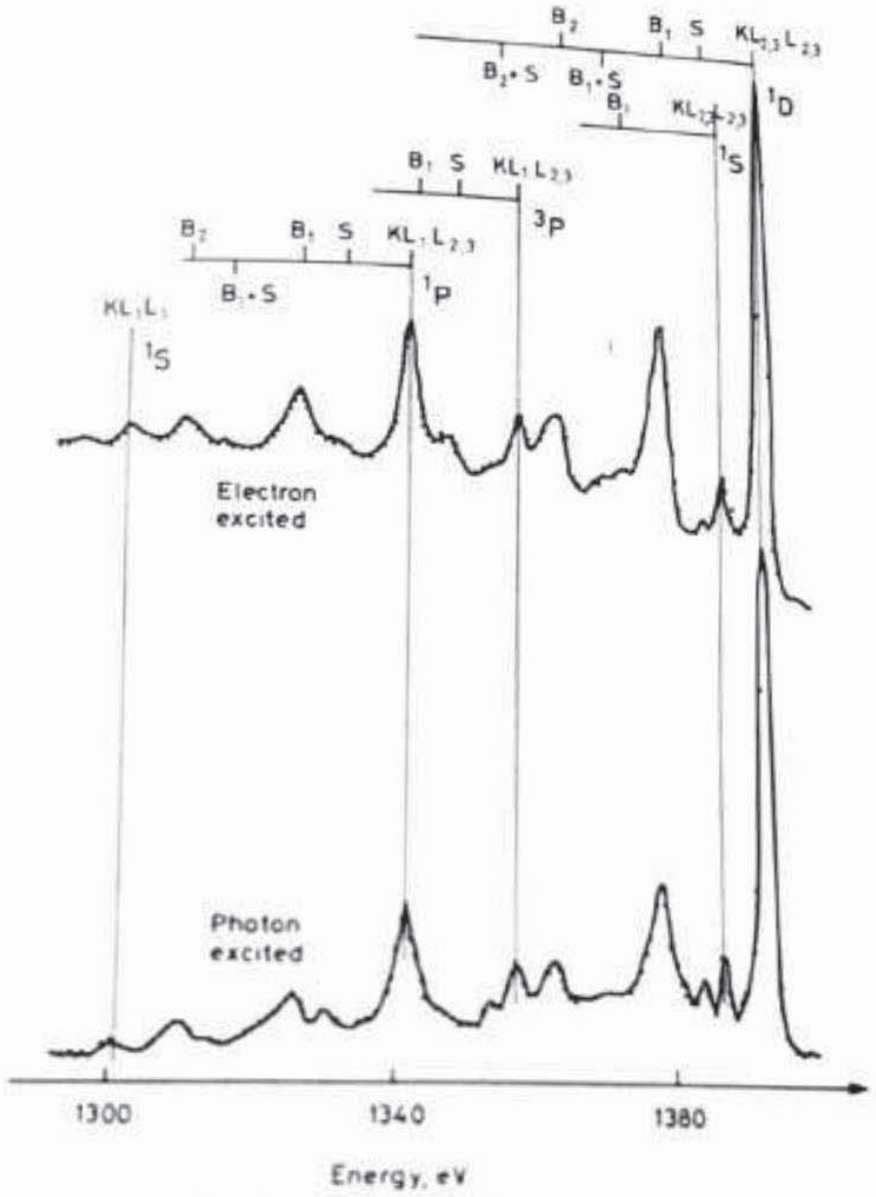


Figure 3.14 Electron-excited and photon-excited plasmon loss fine structure associated with the aluminium *KLL* Auger spectrum. The parent Auger peaks and their spectral terms are indicated as $KL_{2,3}L_{2,3}(^1D)$, $KL_{2,3}L_{2,3}(^1S)$, $KL_1L_{2,3}(^3P)$, etc. The surface plasmons are indicated as S and the successive bulk plasmons as B_1 and B_2; coupling between bulk and surface plasmons is thus $(B_1 + S)$ and $(B_2 + S)$. (Reproduced from Dufour *et al.*[16] by permission of The Royal Swedish Academy of Sciences)

same for either method of excitation and also that the same associated plasmon losses occur.

In an AES spectrum the largest peak is generally that of the elastically scattered electrons, and if the primary electron beam has a narrow spread in energy, then the elastic peak is also narrow when adequate energy resolution is used in the energy analyser. Under these conditions the plasmon loss structure associated with the elastic peak is both intense and well resolved, and the

surface plasmon loss $\hbar\omega_s$ in particular has proved useful as a surface cleanliness diagnostic. By reducing the primary energy to 50–300 eV, typically, where the penetration of the beam into the solid is very low and all loss processes are occurring at or near the surface, the surface plasmon loss can be considerably enhanced. The level of detectability of the resultant loss peak, coupled with the sensitivity of the surface plasmon loss to surface contamination and reaction, is then such that the state of cleanliness of a surface can be judged at least as sensitively from observation of that peak as that of the Auger peaks characteristic of contaminants.

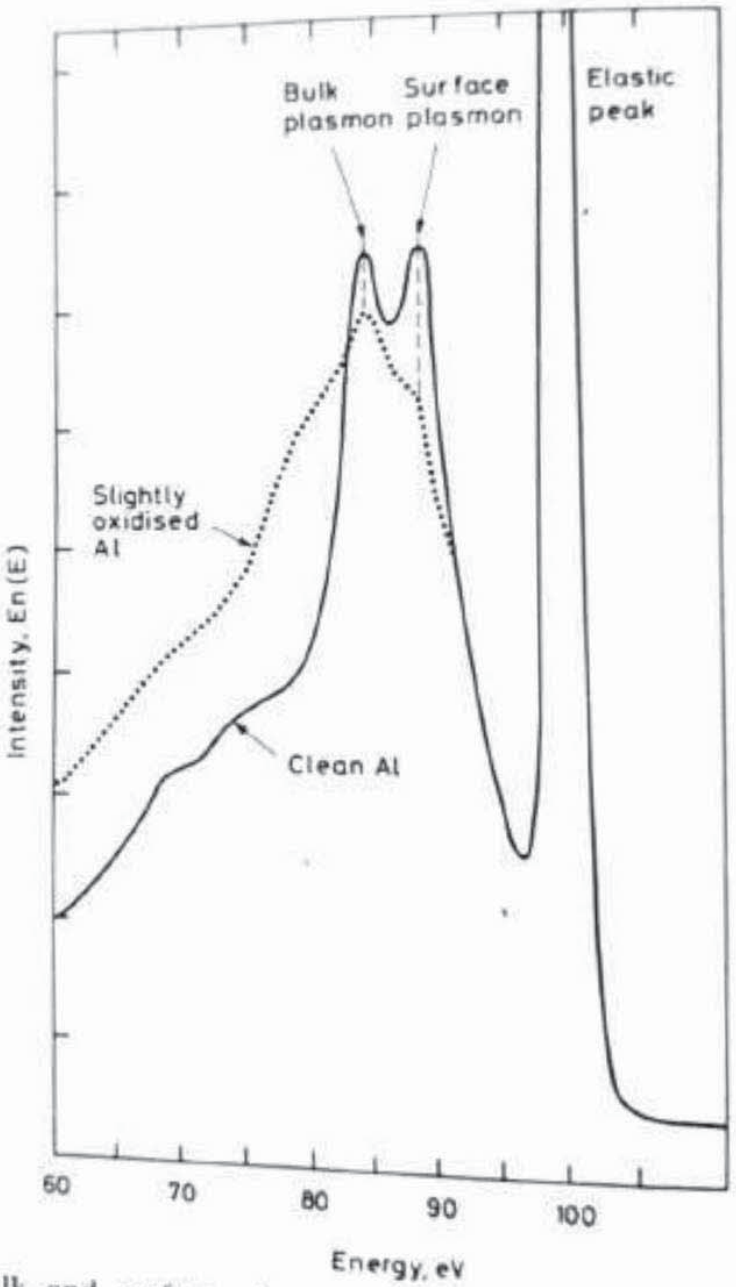


Figure 3.15 Bulk and surface plasmon losses from clean and slightly oxidized aluminium surfaces associated with the elastically scattered peak at a primary energy of 100 eV. The magnitude of the surface plasmon loss peak is a sensitive monitor of surface cleanliness. (Reproduced from Massignon *et al.*[16] by permission of the American Institute of Physics)

The surface plasmon loss peak has been used in this way by, for example, Massignon *et al.*[16] Figure 3.15, from their paper, shows how sharply the surface plasmon loss peak from aluminium is reduced due to slight oxygen contamination.

3.3.7 *Ionization loss features*

As depicted earlier in Figure 3.3, the first step in the electron-excited Auger process is the ionization of a core level by an incoming primary electron. Once an electron has entered a solid and lost energy by one or more loss processes any information about its origin is lost and it becomes indistinguishable from the other electrons in the solid. Thus on ionization two electrons are created in the solid, and since an electron can give up any fraction of its kinetic energy either of the two electrons can have energy from $(E_p - E_i)$ to zero. For ionization of E_i, kinetic energies above $(E_p - E_i)$ are not possible since all levels are filled up to E_f, from which E_i is measured, and there is thus nowhere for an excited electron to go except to empty, allowed, states beyond E_f.

In the undifferentiated secondary electron spectrum, therefore, the ionization loss feature would appear as a shallow step at a kinetic energy $(E_p - E_i)$, followed by a decreasing tail stretching all the way to zero energy. Such a step is very difficult to observe in the undifferentiated spectrum, but may be revealed more clearly, as for Auger features, by the usual differentiation with respect to energy.[17]

3.4 The X-ray Photoelectron Spectrum

3.4.1 Primary structure

Figure 3.16 depicts a *wide scan* spectrum of a clean silver surface obtained using Mg $K\alpha$ radiation and an analyser operating in the constant pass energy mode (constant ΔE). A series of peaks are observed on a background which generally increases to low kinetic energy (high binding energy) but which also shows step-like increases on the low kinetic energy side of each significant peak. When, as in this case, the X-ray source is non-monochromatic, the output consists of a broad continuous distribution (Bremsstrahlung radiation) upon which are superimposed lines characteristic of the target (anode) material. All these X-rays give rise to photo-electron emission according to the equation

$$E_K = h\upsilon - E_B - \phi \tag{3.4}$$

here E_K is the measured electron kinetic energy (KE), $h\upsilon$ the energy of the exciting radiation, E_B the binding energy (BE) of the electron in the solid and

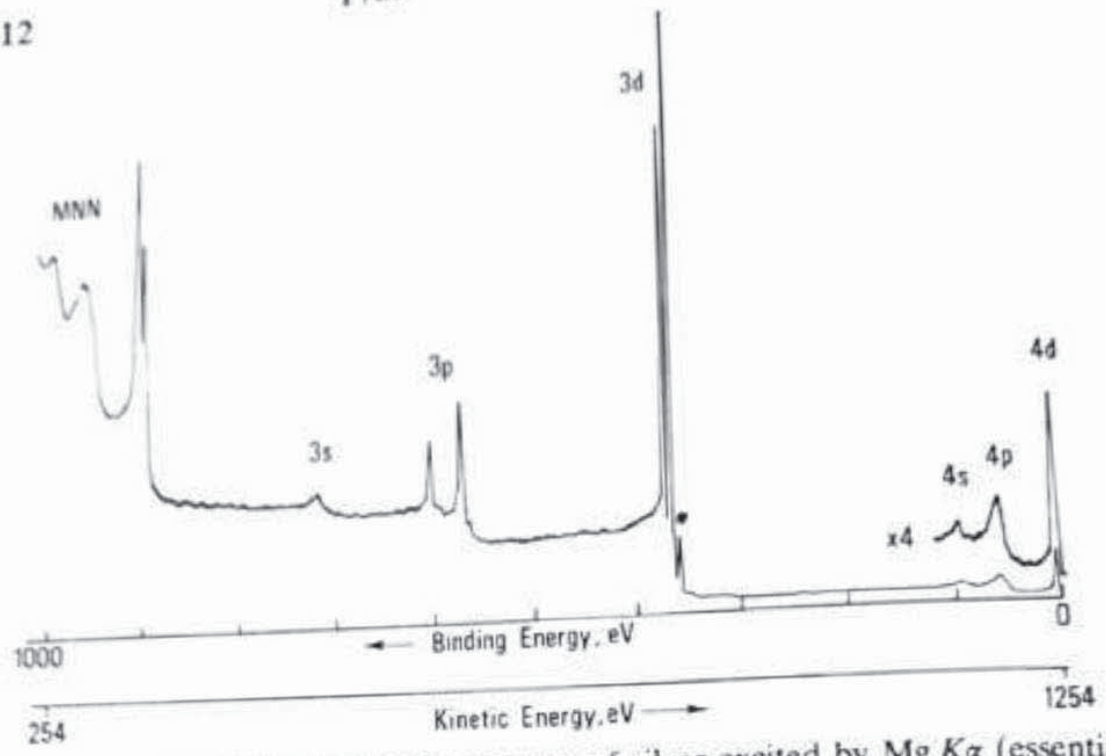


Figure 3.16 X-ray photo-electron spectrum of silver excited by Mg $K\alpha$ (essentially Mg $K\alpha_{1,2}$) and recorded with a constant analyser energy of 100 eV (*Ag 3*d* 'satellite' excited by Mg $K\alpha_{3,4}$)

ϕ the 'work function' (a catch-all term whose precise value depends on both sample and spectrometer, see Appendix I). Equation (3.4) assumes that the photo-emission process is *elastic*. Thus each characteristic X-ray (Mg $K\alpha_{1,2}$, etc., see below) will give rise to a series of photo-electron peaks which reflect the discrete binding energies of the electrons present in the solid.

The photo-emission process is *inelastic* if the photo-electron suffers an energy change (usually an energy loss) between photo-emission from an atom in the solid and detection in the spectrometer. Characteristic energy loss processes are discussed later; for the moment it is sufficient to note that the background 'step' to low kinetic energy of the photo-electron peaks is due to inelastic photo-emission (energy loss within the solid). Photo-emission by Bremsstrahlung radiation gives rise to a general background which is dominant in the low binding energy region of the spectrum. Secondary electrons resulting from inelastic photo-emission increasingly dominate the background at lower kinetic energy. The peaks observed in Figure 3.16 can be grouped into three basic types: peaks due to photo-emission from *core levels* and *valence levels* and peaks due to X-ray excited Auger emission (*Auger series*).

3.4.1.1 Core levels

The core-level structure in Figure 3.16 is a direct reflection of the electron structure of the silver atom. Mg $K\alpha$ radiation (1253.6 eV) is only energetic

Table 3.5 Spin-orbit splitting parameters

Subshell	j values	Area ratio
s	$\frac{1}{2}$	—
p	$\frac{1}{2}, \frac{3}{2}$	1 : 2
d	$\frac{3}{2}, \frac{5}{2}$	2 : 3
f	$\frac{5}{2}, \frac{7}{2}$	3 : 4

enough to probe the core levels of silver up to the 3s shell. It is immediately clear that the core levels have variable intensities and widths and that non-s levels are doublets.

The doublets arise through spin-orbit (j–j) coupling which has previously been discussed in Sections 3.2.1. Two possible states, characterized by the quantum number j ($j = l + s$) arise when $l > 0$ (see Table 3.1). The difference in energy of the two states, ΔE_j, reflects the 'parallel' or 'anti-parallel' nature of the spin and orbital angular momentum vectors of the remaining electron. The magnitude of this energy separation is proportional to the spin-orbit coupling constant ξ_{nl} which depends on the expectation value $\langle 1/r^3 \rangle$ for the particular orbital. The separation can be many electron volts. ΔE_j is therefore expected to increase as Z increases for a given subshell (constant n, l) or to increase as l decreases for constant n (e.g. Figure 3.16 shows for silver that the splitting of $3p > 3d$).

The relative intensities of the doublet peaks are given by the ratio of their respective degeneracies ($2j + 1$). Thus the area ratios and designations (nl_j) of spin-orbit doublets are given in Table 3.5.

Relative intensities The basic parameter which governs the relative intensities of core-level peaks is the atomic photo-emission cross-section, α. Values of α have been directly calculated and also derived from X-ray mass absorption coefficients. The calculated values of Scofield[15] are presented in Figure 5.14 of Chapter 5 to show the trends. Variation in the exciting energy also affects values of α, although for the commonly used Mg $K\alpha$ and Al $K\alpha$ radiations the effect on *relative* photo-emission cross-sections is very small. The transmission characteristics of the electron analyser, being a function of electron energy, apply an important modulation. These effects are thoroughly discussed in the context of quantification in Chapter 5.

Peak widths The peak width, defined as the full width at half-maximum (FWHM) ΔE, is a convolution of several contributions:

$$\Delta E = (\Delta E_n^2 + \Delta E_p^2 + \Delta E_a^2)^{1/2} \tag{3.5}$$

where ΔE_n is the natural or inherent width of the core level, ΔE_p is the width of the photon source (X-ray line) and ΔE_a the analyser resolution, all ex-

pressed as FWHM. Equation (3.5) assumes that all components have a Gaussian line shape.

The analyser contribution is discussed in detail in Chapters 2 and 5; it is the same for all peaks in the spectrum when the analyser is operated in the constant analyser energy (CAE) mode, but varies across the spectrum when the analyser is operated in the constant retard ratio (CRR) mode (since in this case $\Delta E/E$ is a constant).

The inherent line width of a core level, i.e. the range in KE of the emitted photo-electron, is a direct reflection of uncertainity in the lifetime of the ion state remaining after photo-emission. Thus from the uncertainity principle we obtain the line width (in energy units)

$$\Gamma = \frac{h}{\tau} = \frac{4.1 \times 10^{-15}}{\tau} \quad \text{eV} \tag{3.6}$$

with Planck's constant (h) expressed in electronvolt-seconds and the lifetime (τ) expressed in seconds. Clearly the narrowest core levels (e.g. Ag 3*d*) have lifetimes between 10^{-14} and 10^{-15} s whilst the broader core levels (e.g. Ag 3*s*) have lifetimes close to or even slightly less than 10^{-15} s.

Core-hole lifetimes are governed by the processes which follow photo-emission, in which the excess energy of the ion is dissipated or decays. Any of three mechanisms may be involved (Figure 3): emission of an X-ray photon (X-ray fluorescence), or emission of an electron either in an Auger process or in a Coster–Kronig process.

The Coster–Kronig process is a special type of Auger process in which the final doubly charged ion has one hole in a shell of the same principal quantum number (n) as that of the original ion. The relevant point at this juncture is that Coster–Kronig processes are heavily favoured for the *initial* decay of holes in core levels with low angular momentum quantum numbers (e.g. 2*s*, 3*s* and 3*p*, 4*s* and 4*p*, etc.) and in addition they are extremely fast. These core levels, therefore, are broader than those which decay by normal Auger processes.

It can be noted that the line widths of the principal light element core levels (1*s*, 2*p*) systematically increase with an increase in atomic number. The increase in valence electron density enhances the probability of the relevant Auger process (see below), decreasing the lifetime of the core hole. This effect can even be seen in high resolution spectra from one core level in different chemical environments. For example, the chemically distinct N atoms in the linear azide ion (N_3^-) give rise to two N 1*s* peaks (due to the chemical shift effect, see below), of which the peak due to the central atom, with the lowest electron density, is the narrower.

3.4.1.2 *Valence levels*

Valence levels are those occupied by electrons of low binding energy (say 0–20 eV) which are involved in de-localized or bonding orbitals. The spec-

trum in this region consists of many closely spaced levels giving rise to a *band* structure. Two situations can be distinguished as in Figure 3.17, viz. insulators and conductors (metals). Figure 3.17 illustrates the density of electron states (per unit energy in unit volume) in these two cases. In the case of an insulator the occupied *valence band* is separated from the empty *conduction band*, whilst in the case of a metal these bands overlap and the uppermost occupied state is termed the *Fermi level* (E_F). Note that E_F is not the true zero point of the electron energy scale, although BEs are often referenced to this point. The true zero is the vacuum level (E_v) and, to a first approximation, $E_F - E_v = \phi$, where ϕ is the work function of the material.

Figure 3.17 shows that at low energies the unoccupied states have structure. Consequently, if photo-electrons ejected from filled levels have KEs which fall within this structured region the *observed* intensity will be a convolution of the filled and empty density of states together with the matrix of transition probabilities. This is the case in ultra-violet photo-electron spectroscopy (UPS) and gives rise to the strong dependence of valence-level spectra on photon energy. In the case of XPS the KE of the valence photo-electrons is such that the final states are quite devoid of structure; thus the observed density of states closely reflects the initial filled density of states. In Figure 3.16 the peak labelled '4*d*' is in fact the conduction band spectrum which is dominated by 4*d* states. The complete spectrum at high resolution, Figure 3.18, shows the band structure and the sharp cut-off in electron density at E_F.

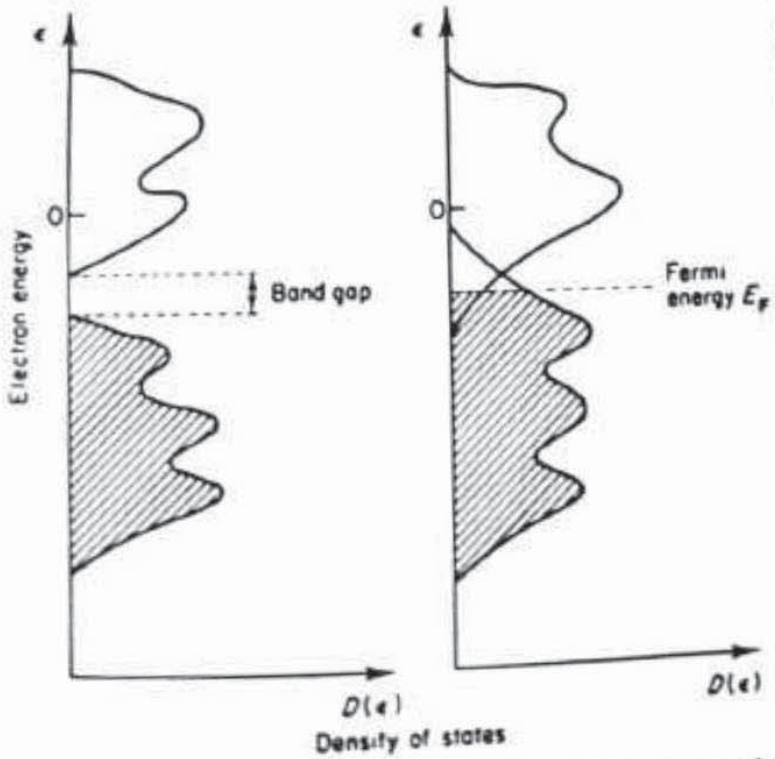


Figure 3.17 Schematic density of states for (a) an insulator and (b) a metal. The shading indicates the extent to which the energy levels are occupied. (After Orchard[19])

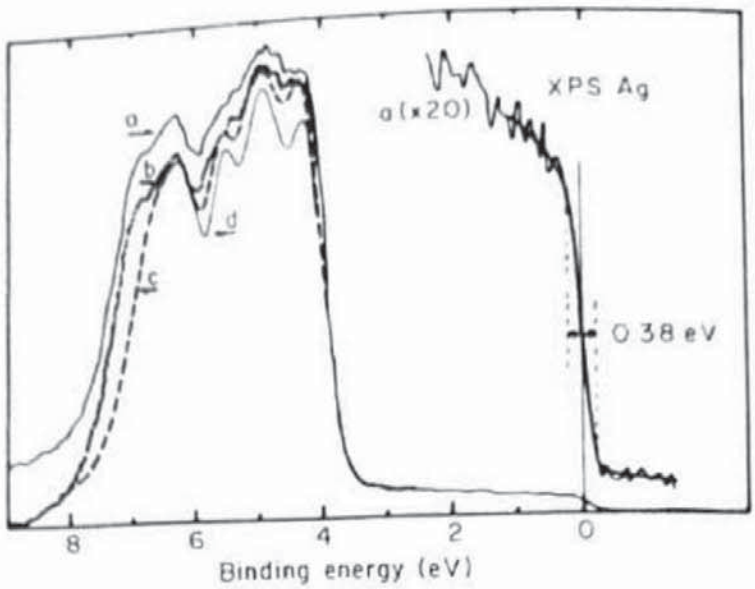


Figure 3.18 XPS valence spectrum for polycrystalline silver excited by monochromatic Al $K\alpha$ radiation in comparison with the theoretical density of states: (a) raw XPS data; (b) data after subtraction of smooth inelastic background; (c) and (d) total theoretical density of states after two different lineshape broadenings to include lifetime and shake-up effects. The Fermi edge has been expanded to reveal the instrumental resolution. (After Barrie and Christensen[20])

Cross-sections for photo-electron emission from valence levels are much lower than for core levels, giving rise to low intensities in general. However, valence-level spectra have their analytical value as discussed in Section 3.4.2.2.

3.4.1.3 Auger series

As explained in Section 3.3.1.1 the dominant decay mechanism for core-hole states is via radiationless transitions resulting in electron emission. The KEs of electrons resulting from Coster-Kronig transitions are very low and these are not normally detected; however, Auger electrons can give rise to dominant peaks in the 'photo-electron' spectrum as seen in Figure 3.16. The nomenclature for Auger peaks is described fully in Section 3.2.

Four principle Auger series can be observed in XPS. The KLL series is first observed for boron and can be exited through to sodium (with Mg $K\alpha$) or magnesium (Al $K\alpha$). With Ag $L\alpha$ (2984 eV) the whole series through to chlorine can be observed. The dominant line is KL_2L_3, as illustrated in Figure 3.19.

The LMM series is first observed with sulphur and can be followed through to germanium (Mg $K\alpha$), selenium (Al $K\alpha$) or ruthenium (Ag $K\alpha$). This series is very complex since lines due to initial ionization in L_2 or L_3 can then result

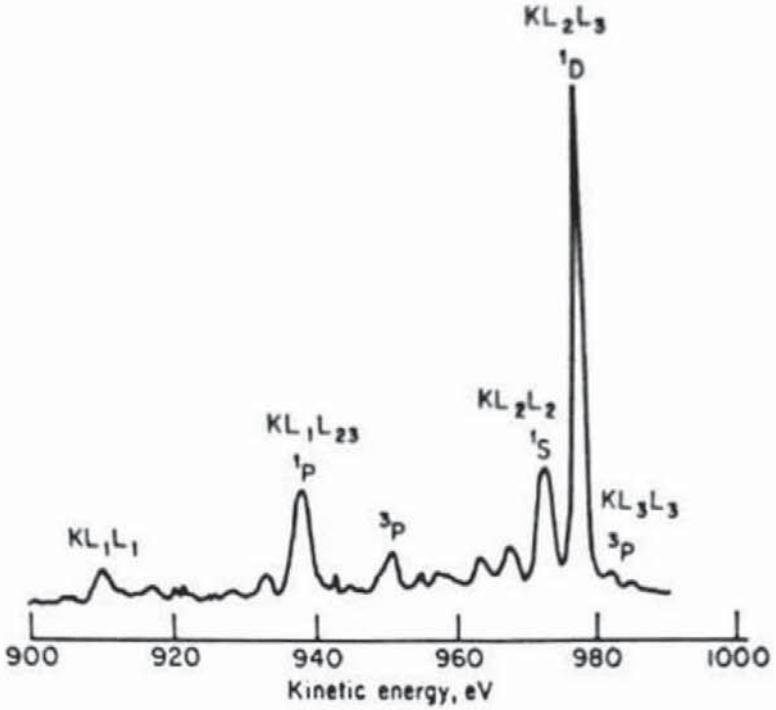


Figure 3.19 *KLL* Auger spectrum of sodium vapour. (After Hillig, Cliff and Mehlhorn[21])

in transitions involving M_{23} or M_{45}. The relative intensities of these groups changes with the atomic number (Z). An example is given in Figure 3.20.

The *MNN* series can be split into two types. The $M_{45}N_{45}N_{45}$ series probably first observed with molybdenum can be followed right through to neodymium with Mg $K\alpha$. The $M_{45}N_{67}N_{67}$ series is only accessible with higher energy photons (e.g. Ag $L\alpha$) because the theoretically observable lines from rare earth elements are weak and ill-defined until hafnium is reached (M_{45}

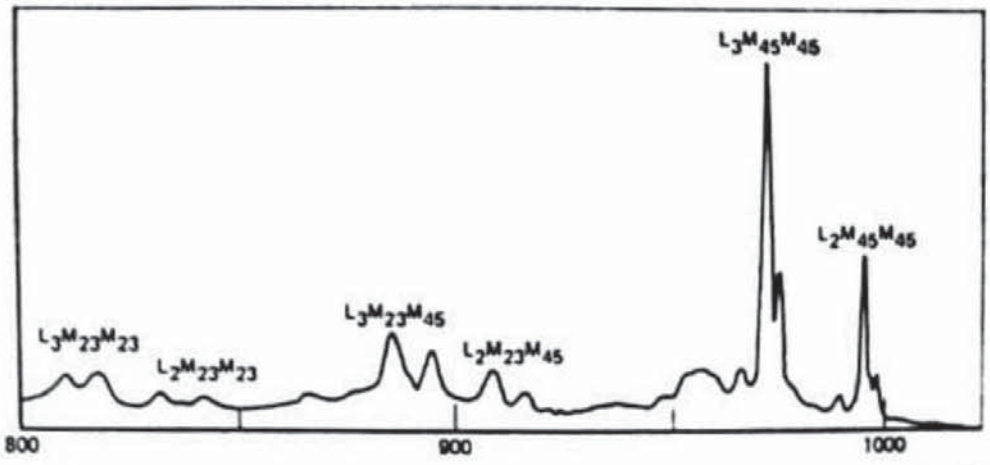


Figure 3.20 *LMM* Auger spectrum of zinc vapour. (After Aksela and Aksela[22])

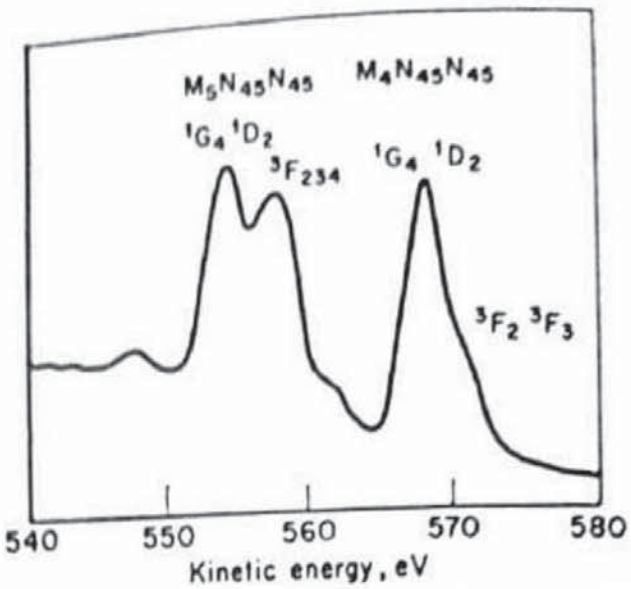


Figure 3.21 Cesium $M_{45}N_{45}N_{45}$ Auger spectrum from Cs_2SO_4. (After Wagner[23])

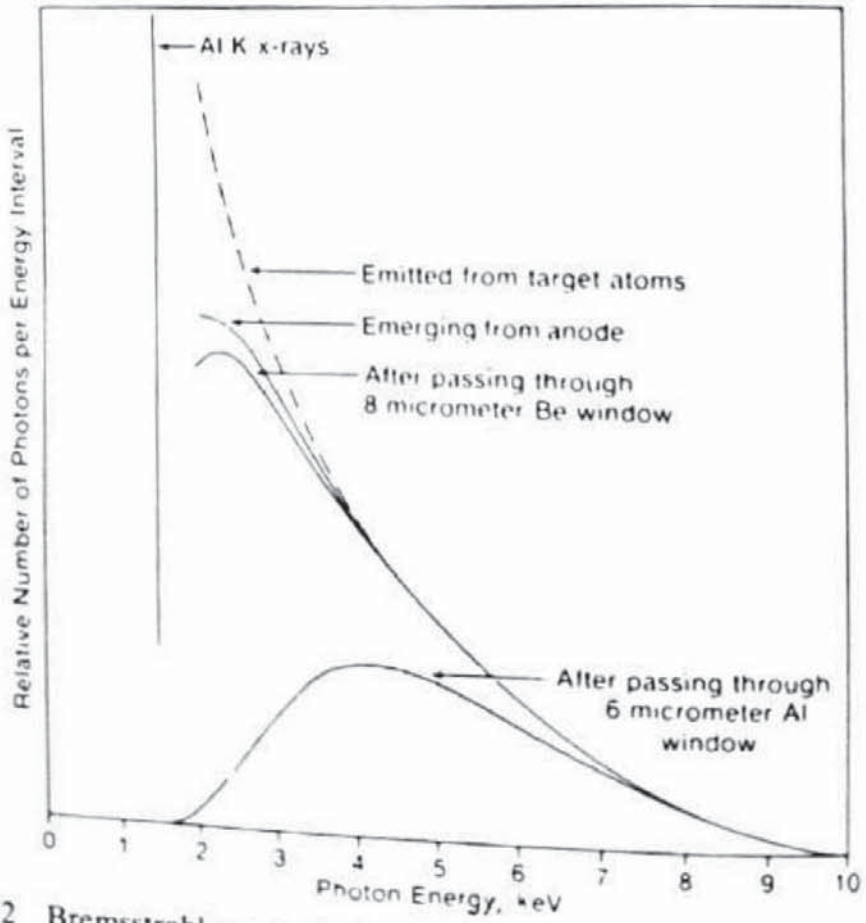


Figure 3.22 Bremsstrahlung radiation intensity from a conventional aluminium X-ray source. (Reproduced from Wagner and Taylor[24] by permission of Elsevier Scientific Publishing Co.). Note that use of a 2 μm aluminium window is roughly equivalent to using the beryllium window

shell = 1660 eV and therefore beyond Mg $K\alpha$ or Al $K\alpha$). A typical $M_{45}N_{45}N_{45}$ spectrum is shown in Figure 3.21.

The positions of the principal Auger peaks on the BE scale are included in Appendix 7. The KE of an Auger electron is given in equation (3.2); thus, in contrast to the photo-electron (equation 3.4) the Auger electron KE is independent of the exciting radiation. If, for example, the X-ray source is changed from Mg $K\alpha$ to Al $K\alpha$ all photo-electron peaks increase in KE by 233 eV whilst the Auger peaks remain unmoved. For spectra plotted directly in BE the reverse is true. This allows the two types of peak in the electron spectrum to be easily distinguished. Note also that the width of X-ray excited Auger peaks, as a consequence of equation (3.2), does not contain a contribution from the line width of the X-ray source.

In the above discussion it has been indicated which X-ray sources have possibilities for exciting the high energy Auger series (i.e. where the initial state energy is greater than that of Al $K\alpha$). However, for conventional single-target instruments the observation of principal photo-electron lines and these Auger lines in one experiment (see Section 3.4.2.3) is impossible, and attention has recently turned to the use of the Bremsstrahlung radiation for this purpose. As shown in Figure 3.22 useful intensities over a wide range of photon energy can be obtained from the normal aluminium target, particularly if a beryllium window is used in the X-ray source. Wagner and Taylor[24] have explored this alternative for twenty-four elements with Auger lines having KEs in the range 1400–2200 eV.

3.4.2 Information from primary structure

3.4.2.1 Core-level chemical shifts

The discovery, during the early days of XPS, that non-equivalent atoms of the same element in a solid gave rise to core-level peaks with measureably different BEs had a stimulating effect on the field. This BE differenee was dubbed the 'chemical shift' by analogy with nuclear magnetic resonance (n.m.r.) spectroscopy. Non-equivalence of atoms can arise in several ways: difference in formal oxidation state, difference in molecular environment, difference in lattice site and so on. The physical basis of the chemical shift effect is illustrated by a relatively simple model used with much success (in appropriate cases) to interpret chemical shift data—the charge potential model:[25]

$$E_i = E_i^0 + kq_i + \sum_{i \neq j} \frac{q_j}{r_{ij}} \tag{3.7}$$

where E_i is the BE of a particular core level on atom i, E_i^0 is an energy reference, q_i is the charge on atom i and the final term of equation (3.7) sums the potential at atom i due to 'point changes' on surrounding atoms j.

If the atom is considered to be an essentially hollow sphere on the surface of which the valence charge q_v resides, then the classical potential inside the sphere is the same at all points and equal to q_v/r_v, where r_v is the average valence orbital radius. A change in the valence electron charge (density) of Δq_v changes the potential inside the sphere by $\Delta q_v/r_v$. Thus the BE of all core levels will change by this amount. Moreover, as r_v increases, the BE shift for given Δq_v will decrease. In practice it is indeed found that for all cases where core–valence interaction is small the observable core levels from a given atom do suffer a similar BE shift and that BE shifts between equivalent compounds decrease as one descends a column of the periodic table (similar Δq_v, r_v increasing). If the summation term in equation (3.7) is abbreviated to V_i then the shift in BE for a given core level of atom i in two different environments is

$$E_i^{(1)} - E_i^{(2)} = k(q_i^{(1)} - q_i^{(2)}) + (V_i^{(1)} - V_i^{(2)}) \qquad (3.8)$$

The first term ($= k\,\Delta q_i$) clearly ensures that an increase in BE accompanies a decrease in valence electron density on atom i. The second term is not to be underestimated since it has the opposite sign to Δq_i. In molecular solids the atoms j are basically the atoms bonded to atom i, but in ionic solids the summation extends over the whole lattice; this is closely related to the Madelung energy of the solid and for this reason V is often referred to as a Madelung potential.

Equations (3.7) and (3.8) involve a number of simplifications. A major simplification of the charge potential model is that it neglects relaxation effects, i.e. no account is taken of the polarizing effect of the core hole on the surrounding electrons, both intra-atomic (on atom i) and extra-atomic (on atoms j). This is one example of a final-state effect, a term used to indicate the possibility that the electronic state of the photo-ionized atom may be different from that of the initial state (usually the atomic ground state). Other final-state effects are discussed later in this chapter.

In comparing equation (3.8) with equation (3.4) it will be realized that equation (3.8) is only valid if the work functions of the materials incorporating $i_{(1)}$ and $i_{(2)}$ are the same. When conductors are involved 'ϕ' is the spectrometer work function and the problem does not arise, but in the case of insulators $\phi_{(1)}$ and $\phi_{(2)}$ may not be identical. Moreover, insulators acquire a surface charge under X-irradiation and correction of non-identical charging is required before chemical shifts can be accurately measured. The problem of charge correction or BE referencing is dealt with in Appendix 2.

Figure 3.23 illustrates the chemical shift effect for three different types of nitrogen atom in the N 1*s* spectrum of a single compound. In this case the relative shifts are easily measured but the absolute N 1*s* binding energies require some calibration procedure to be adopted before they can be measured. The clear trend in this example is that the N 1*s* binding energy increases with an increase in the formal oxidation number of the nitrogen atom. Table 3.6 shows this to be generally true for nitrogen in both inorganic and organic

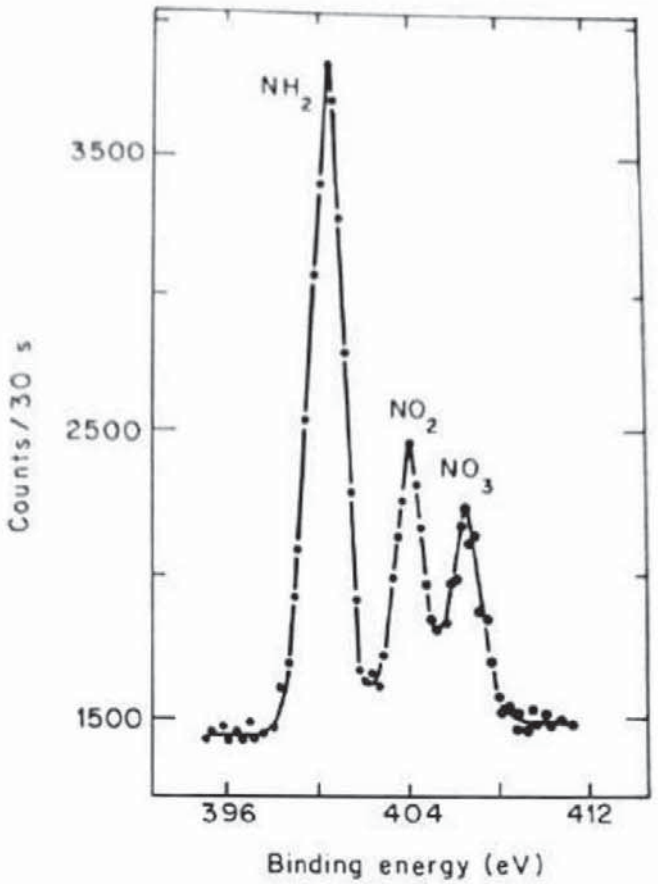


Figure 3.23 Nitrogen 1*s* spectrum from *trans*- $[Co(NH_2CH_2CH_2NH_2)_2(NO_2)_2]NO_3$. (Reprinted with permission from Hendrickson[26]. Copyright 1969 American Chemical Society)

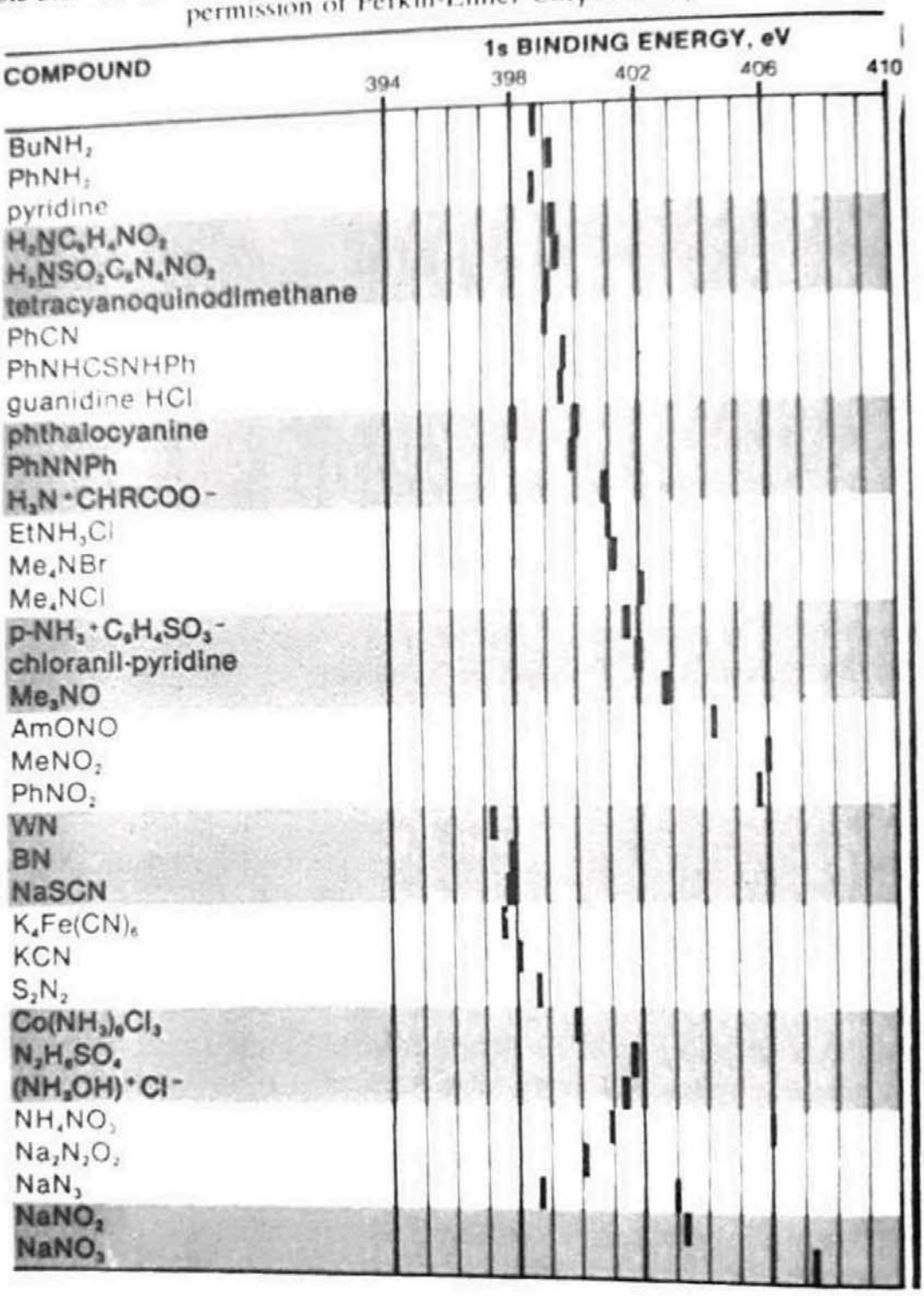
compounds. In situations where the formal oxidation state is the same the general rule is that the core-level BE of the central atom increases as the electro-negativity (electron withdrawing power) of attached atoms or groups increases. Many examples of core-level chemical shifts appear in later chapters and Appendix 4 gives an extensive compilation of reported values for most elements.

Many cases obviously arise in which distinct chemical states give rise to photo-electron peaks whose relative chemical shift is sufficiently small for the peaks to overlap considerably. Whilst in principle such a shift should be measurable if $\simeq 0.2$ eV this is impossible when the contributory peaks have a line width typically of 1–2 eV. However, if the profile of the contributory peaks is known then an attempt can be made to deconvolute the envelope. This procedure is discussed fully in Appendix 3.

3.4.2.2 Valence band structure

As mentioned in Section 3.4.1.2 valence band spectra relate closely to the occupied density of states structure. This is very useful in the study of the

Table 3.6 N 1s binding energies. (Reproduced from Wagner *et al.*[27] by permission of Perkin-Elmer Corporation)



electronic structure of materials, fundamental to many aspects of device applications, and in checking the accuracy of band structure calculations. As shown in Figure 3.18, very good agreement has been obtained in the case of silver and this is, in fact, quite typical of work on all types of metals. The applicability of various models for the electronic structure of alloys has been extensively studied in this way.

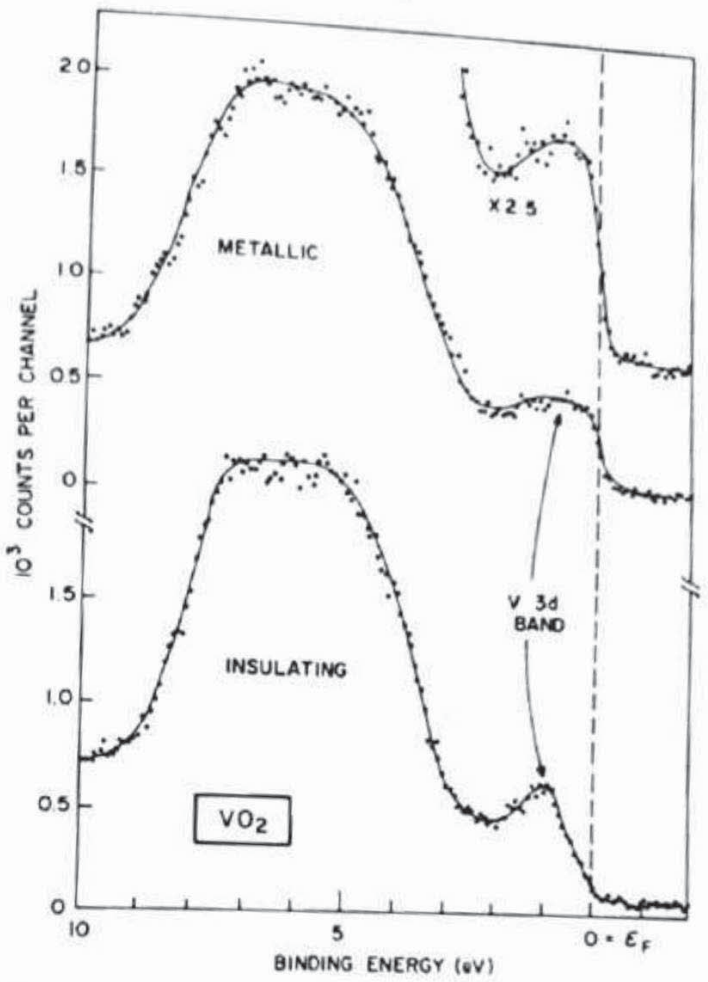


Figure 3.24 XPS valence bands of VO_2 at temperatures above and below the metal-insulator transition temperature. (Reproduced from Wertheim[28] by permission of Pergamon Press Ltd)

Further examples of the direct information obtained from valence band spectra are provided by the following. Vanadium dioxide (VO_2) is an insulator at room temperature but undergoes a structural change at 65 °C and becomes a metallic conductor. Figure 3.24 clearly shows the gap created at the lower temperature which results in the loss of conductivity. The tungsten bronzes are the family of materials with composition Na_xWO_3 ($0.32 < x < 0.93$). The metallic properties of these materials contrasts with the insulating behaviour of WO_3. Figure 3.25 shows that increasing the sodium content results in electron density increasing in a '*d*' band. This is absent in WO_3 (d^0) but is clearly seen in ReO_3 (the d^1 counterpart) which has metallic conductivity.

In the case of polymers, valence bands reveal structural information often unobtainable from core-level studies. This fingerprinting capacity of polymer valence band spectra is discussed in Chapter 9.

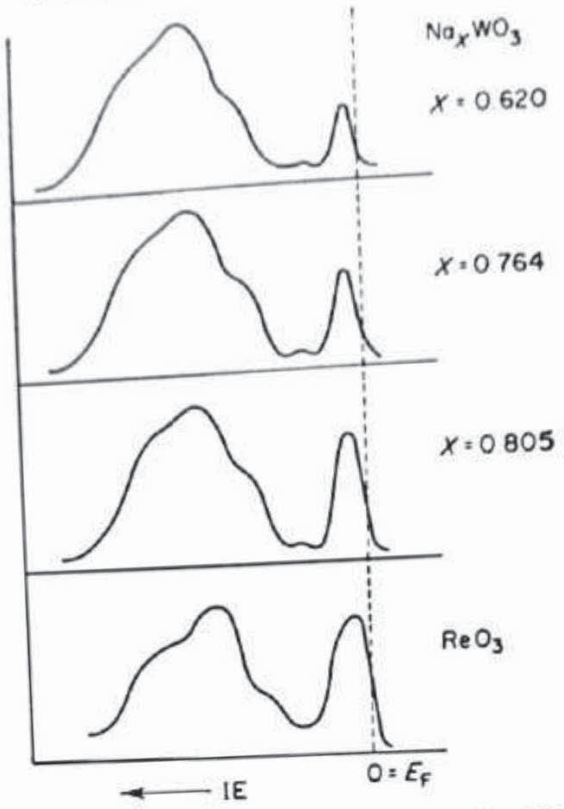


Figure 3.25 XPS valence bands of the tungsten bronzes Na_xWO_3 and ReO_3. (After Wertheim *et al.*[29])

3.4.2.3 *Auger chemical shifts*

For the Auger processes in which the final vacancies arise in core levels it is clear that a change in chemical state giving rise to a chemical shift in the photo-electron lines will also produce a chemical shift in the Auger lines. The magnitude of the Auger chemical shift is, however, often significantly greater than that of the photo-electron chemical shift. For instance, Wagner and Biloen,[30] in an early study of this effect, found the Auger chemical shift betwen metals and their oxides exceeded the photo-electron chemical shift by the following: Mg, 5.0; Zn, 3.7; Ga, 4.2; Ge, 3.5; As, 3.4; Cd, 4.9; In, 1.8; and Sn, 2.5 eV. Chemical shifts in Auger lines are discussed more fully in Section 3.3 but the reason for this effect can be appreciated from the following simplified argument.[31] As discussed in Section 3.4.2.2 the difference in BE between two chemical states depends on the change in core electron energy and the change in intra-atomic and extra-atomic relaxation energies. Thus, for a 1*s* (K) electron:

$$\Delta E_B(K) = \Delta\varepsilon(K) - \Delta R(K^+) \tag{3.9}$$

where $\Delta\varepsilon(K)$ is the 1*s* energy change and $\Delta R(K^+)$ is the change in relaxation energies (predominantly extra-atomic) in the final, singly ionized state. As

explained above (Section 3.4.2.1):

$$\Delta\varepsilon(K) \approx \Delta\varepsilon(L) \tag{3.10}$$

and the kinetic energy change is given by

$$\Delta E(K) = -\Delta E_B(K) = \Delta R(K^+) - \Delta\varepsilon(K) \tag{3.11}$$

For the KLL Auger process the kinetic energy change between chemical states is given by

$$\Delta E\ (KLL) = \Delta\varepsilon(K) - \Delta R(K^+) - 2\Delta\varepsilon(L) - \Delta R(L^+L^+) \tag{3.12}$$

where (L^+L^+) is the doubly ionized final state. It can be argued that $R(L^+L^+) \approx 4R(K^+)$ so from equations (3.10) and (3.12):

$$\Delta E\ (KLL) = 3\Delta R(K^+) - \Delta\varepsilon(K) \tag{3.13}$$

and combining equations (3.11) and (3.13):

$$\Delta E\ (KLL) - \Delta E(K) = 2\Delta R(K^+) \tag{3.14}$$

The difference between Auger and photo-electron chemical shifts therefore results from the difference in final-state relaxation energies between chemical states. It is possible to define a parameter, α, such that, for example:

$$\alpha = E(KLL) - E(K) \tag{3.15}$$

For insulators α is independent of any static charging and is a parameter, characteristic of a particular chemical state, which can be measured with greater accuracy than core-level BE or Auger peak KE. Wagner termed α the Auger parameter[32] and defined it practically as the difference between the KE of the most intense Auger peak and the KE of the most intense photo-electron line. The disadvantage of this procedure is that it can result in negative values. A 'modified' Auger parameter has been defined[33] as, for example:

$$\alpha' = \alpha + h\upsilon = E(KLL) + E_B(K) \tag{3.16}$$

which overcomes this problem. A good example of the use of this parameter comes from the work of West and Castle[34] on silicate minerals, in which a Zr $L\alpha$ source was used to excite the Si 1s and Si KLL levels. Many Auger parameters are given in Appendix 4.

3.4.2.4 *Auger lineshapes*

When an Auger process results in at least one vacancy in the valence levels then the intensity distribution of the lines in the Auger series can vary strongly from one compound to another. In effect the valence band or molecular orbital structure is convoluted into the Auger peak structure. This is nicely shown by the F KLL spectra in Figure 3.26. In fluorine the L_1 (2s) level is

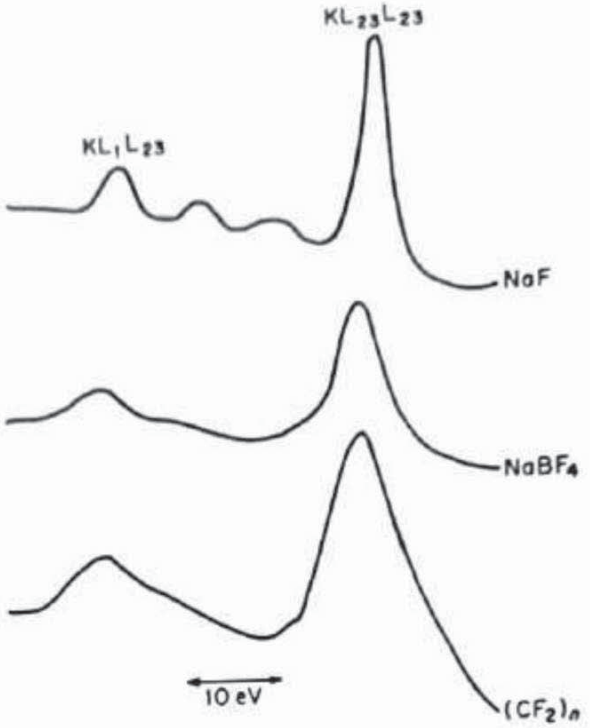


Figure 3.26 Fluorine KLL spectra for solids of different covalent character. (After Wagner[35])

essentially core-like, whilst the L_{23} ($2p$) levels are involved in bonding. In sodium fluoride the fluorine environment is similar to that of the free ion and the KLL structure appears to be similar to the Na KLL structure where all levels are core-like. In the complex ion and in the organic polymer, where the covalency of fluorine is increasing and the L_{23} levels are increasingly involved in bonding orbitals, the component lines in the KLL series become less sharply defined.

With oxygen the KLL group takes on a greater variety of forms, as shown in Figure 3.27. Note in particular how the separation between the KL_1L_{23} and $KL_{23}L_{23}$ lines varies by up to 5 eV and how the $KL_{23}L_{23}$ profile changes from an essentially single to a double peak. The extra variation in this peak is expected because both vacancies are now in the valence shell, whereas the L_1 shell is again semi-core-like. Once again the O KLL line shape shows characteristic variations from one type of compound to another.

3.4.3 Secondary structure

3.4.3.1 X-ray satellites and ghosts

Before discussing true secondary structure features in the spectrum it is necessary to identify the sources of 'spurious' low intensity peaks. These arise in two ways.

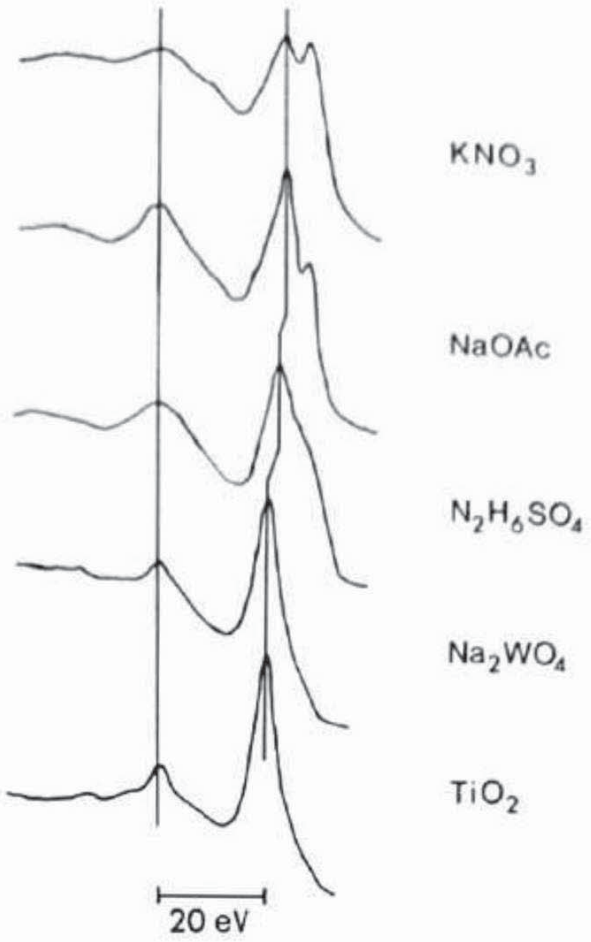


Figure 3.27 Oxygen *KLL* Auger spectra from several materials in the solid state. (Reprinted with permission from Wagner, Zatko and Raymond[36] Copyright 1980 American Chemical Society)

Standard X-ray sources are not monochromatic. Besides the Bremmsstrahling radiation mentioned in Section 3.4.1 and the principal $K\alpha_{1,2}$ line, magnesium and aluminium targets also produce a series of lower intensity lines, referred to as X-ray satellites. The transitions giving rise to $K\alpha_{1,2}$ (an unresolved doublet) are $2p_{3/2,1/2} \rightarrow 1s$. Satellites arise from less probable transitions (e.g. $K\beta$; valence band $\rightarrow 1s$) or transitions in a multiply ionized atom (e.g. $K\alpha_{3,4}$). The observed line positions and intensities are given in Table 3.7.

X-ray ghosts are due to excitations arising from impurity elements in the X-ray source. The most common ghost is Al $K\alpha_{1,2}$ from a Mg $K\alpha$ source. This arises from secondary electrons produced inside the source hitting the thin aluminium window (present to prevent these same electrons hitting the sample). This radiation will therefore produce weak ghost peaks 233.0 eV to higher KE of those excited by the dominant Mg $K\alpha_{1,2}$. Old or damaged targets can give rise to ghost peaks excited by Cu $L\alpha$ radiation, the main line from the exposed copper base of the target. These peaks appear 323.9 eV (556.9 eV) to lower KE of Mg $K\alpha_{1,2}$ (Al $K\alpha_{1,2}$) excited peaks. In dual-anode X-ray sources, which are commonly being used with one anode capable of yielding

Table 3.7 High-energy satellite lines from Mg and Al targets. (From data of Krause and Ferreira in Carlson[37])

X-ray line	Separation from $K\alpha_{1,2}$(eV) and relative intensity ($K\alpha_{1,2}$ = 100)	
	Mg	Al
$K\alpha'$	4.5 (1.0)	5.6 (1.0)
$K\alpha_3$	8.4 (9.2)	9.6 (7.8)
$K\alpha_4$	10.0 (5.1)	11.5 (3.3)
$K\alpha_5$	17.3 (0.8)	19.8 (0.4)
$K\alpha_6$	20.5 (0.5)	23.4 (0.3)
$K\beta$	48.0 (2.0)	70.0 (2.0)

high-energy X-rays (e.g. zirconium, silver and titanium), misalignments inside the source can lead to cross-talk between filaments and anodes which is a further source of ghost peaks.

A ghost source which has not previously received much attention is O $K\alpha$ arising from the oxide layers present on aluminium and magnesium target surfaces. The possible importance of this interference when recording weak signals has been pointed out by Koel and White[38] as a result of their study of X-ray excited C *KVV* and O *KVV* Auger spectra from adsorbed species. O $K\alpha$ peaks appear 728.7 eV (961.7 eV) to lower KE of Mg $K\alpha_{1,2}$ (Al $K\alpha_{1,2}$) excited peaks.

3.4.3.2 Multiplet splitting

Multiplet splitting (also referred to as exchange or electrostatic splitting) of core-level peaks can occur when the system has unpaired electrons in the valence levels. As an example, consider the case of the 3*s* electron in the Mn^{2+} ion. In the ground state the five 3*d* electrons are all unpaired and with parallel spins (denoted ^6S). After ejection of the 3*s* electron a further unpaired electron is present. If the spin of this electron is parallel to that of the 3*d* electrons (final state ^7S) then exchange interaction can occur, resulting in a lower energy than for the case of anti-parallel spin (final state ^5S). Thus the core level will be a doublet and the separation of the peaks is the exchange interaction energy:

$$E = (2S + 1)\, K_{3s3d} \tag{3.17}$$

where S is the total spin of the unpaired electrons in the valence levels ($\frac{5}{2}$ in this case) and K_{3s3d} is the 3*s*–3*d* exchange integral. The intensity ratio for the two peaks is given by

$$\frac{I(S + \frac{1}{2})}{I(S - \frac{1}{2})} = \frac{S + 1}{S} \tag{3.18}$$

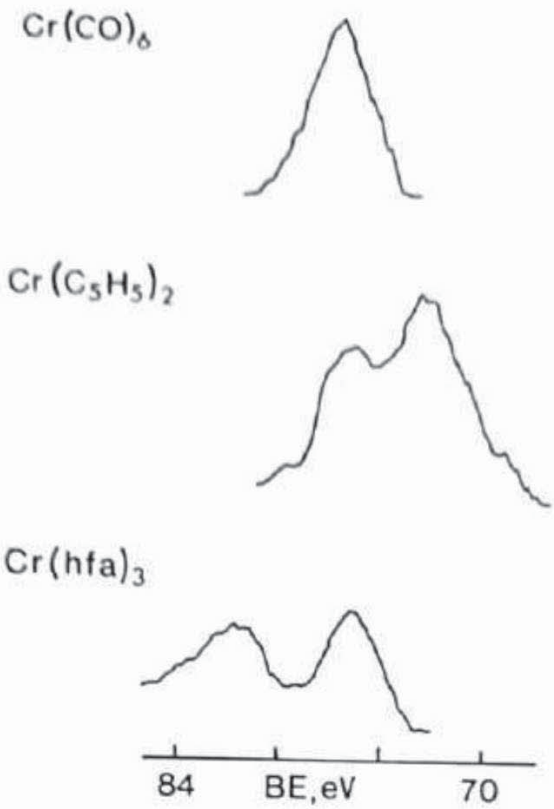


Figure 3.28 Multiplet splitting of the 3*s* peak from chromium compounds: $Cr(CO)_6$ is diamagnetic (no unpaired electrons); $Cr(C_5H_5)_2$ has two unpaired electrons in the e_{2g} level; $Cr(hfa)_3$ has three unpaired electrons in the t_{2g} level (hfa = CF_3COCH-$COCF_3$). Note also how the mean BE varies with the electronegativity of attached ligands. (Reproduced by Clark and Adams[41] by permission of North-Holland Publishing Co.)

For the 3*s* in Mn^{2+} this ratio is $(\frac{5}{2}+1)/\frac{5}{2}$ or 1.4 and E is calculated to be $\simeq$13 eV. In the real case of MnF_2 the intensity ratio is $\simeq$1.8 and the separation $\simeq$6.5 eV.[39] This discrepancy is due to relaxation and configuration interaction effects which complicate the simple model described above (see Ref. 40 for a detailed discussion). However, the data for chromium compounds shown in Figure 3.28 illustrates the fact that this model predicts the correct trends for changes in spin state.

Multiplet splitting for non-*s* levels is more complex because of the additional involvement of orbital–angular momentum coupling. Thus in Mn^{2+} the 3*p* level is split into four levels. Multiplet splitting is strongest when both levels involved are in the same shell (as with 3*s*, 3*p*–3*d*), but the effects are still apparent in transition metal systems for 2*p*–3*d* interaction. A good example is for cobalt which can exist in a variety of spin states. Multiplet splitting causes broadening (with asymmetry) in both $2p_{1/2}$ and $2p_{3/2}$ peaks, leading to apparent variation in the separation of the peak maxima. Table 3.8 gives some examples, showing that the doublet separation can be used diagnostically.[42] Similar variation has been reported for chromium compounds.[43]

Multiplet splitting is found strongly in the 4*s* levels of rare earth metals

Table 3.8 Variation in $2p$ spin-orbit splitting with spin state in Co complexes. (Reproduced from Briggs and Gibson[42] by permission of North-Holland Publishing Co.)

	Binding energy, eV			Unpaired electrons
	Co $2p_{1/2}$	Co $2p_{3/2}$	$2p_{1/2}$–$2p_{3/2}$	
$Co(acac)_2$	800.2	784.2	16.0	3
$Co(acac)_3$	799.0	784.0	15.0	0
$(PEt_2Ph)_2Co(C_6Cl_5)_2$	797.4	782.1	15.3	1
$(PEt_2Ph)_2Co(2\text{-methyl-1-naphthyl})_2$	797.0	781.5	15.5	1

(unpaired electrons in the $4f$ valence levels) and can also be seen in the spectra of organic radicals in which the free electron is relatively localized.

3.4.3.3 *Shake-up satellites*

To the valence electrons associated with an atom the loss of a core electron by photo-emission appears to increase the nuclear charge. This major perturbation gives rise to substantial reorganization of the valence electrons (referred to as relaxation) which may involve excitation of one of them to a higher unfilled level ('shake-up'). The energy required for this transition is not available to the primary photo-electron and thus the two-electron process leads to discrete structure on the low KE side of the photo-electron peak (shake-up satellite(s)).

Most detailed assignment work concerns atomic systems (e.g. neon gas) and shows that the transitions obey monopole selection rules. In the solid state the examples are found to be much more complex, except for organic systems. Conjugated and, especially, aromatic systems show shake-up satellites with intensities up to 5–10 per cent of the primary peak. In aromatic systems the satellite structure has been shown[44] to be due to $\pi \rightarrow \pi^*$ transitions involving the two highest filled orbitals and the lowest unfilled orbital (see Chapter 9).

Very strong satellites are observed for certain transition metal and rare earth compounds which have unpaired electrons in $3d$ or $4f$ shells respectively. The satellite intensity is usually much higher than would be expected for atomic-like shake-up (typically $\approx$10 per cent) and the reason for this has proved a great source of discussion.[40] A convincing explanation is that in the final state there is significant ligand–metal charge transfer such than an extra $3d$ or $4f$ electron is present compared with the initial state. This immediately explains why closed-shell systems (e.g. Cu^+, $3d^{10}$) do not exhibit shake-up satellites, whereas open-shell systems (e.g. Cu^{2+}, $3d^9$) do.[45] However, shake-up is a poor description of this process, as Fadley[40] has pointed out; a preferable

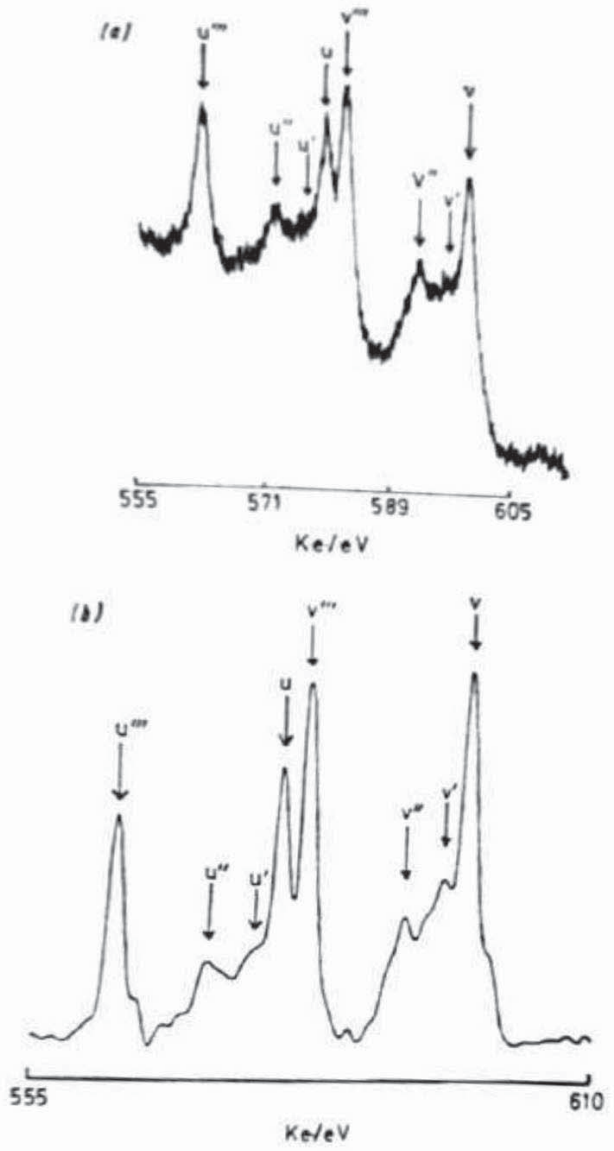


Figure 3.29 Cerium 3*d* signal from CeO_2: (a) raw Al $K\alpha$ excited spectrum, (b) deconvoluted spectrum. (After Burroughs *et al.*[46] by permission of the Royal Society of Chemistry)

description is that strong configuration interaction occurs in the final state due to relaxation. The complexity of such a process is illustrated in Figure 3.29 for cerium.

As in the case of multiplet splitting, shake-up satellites have diagnostic value. The distinguishing of copper (II) from copper (I) has already been mentioned. Tetrahedral nickel (II) gives satellites, square planar (diamagnetic) nickel (I) does not.[47] Cobalt (II) gives more intense satellites in the high spin (^4F) state than in the low spin (^2D) state, while cobalt (III) is diamagnetic and has no satellites.[42,48]

3.4.3.4 *Asymmetric metal-core levels*

In a solid metal there exists a distribution of unfilled one-electron levels above the Fermi energy (as shown in Figure 3.17) which are available for

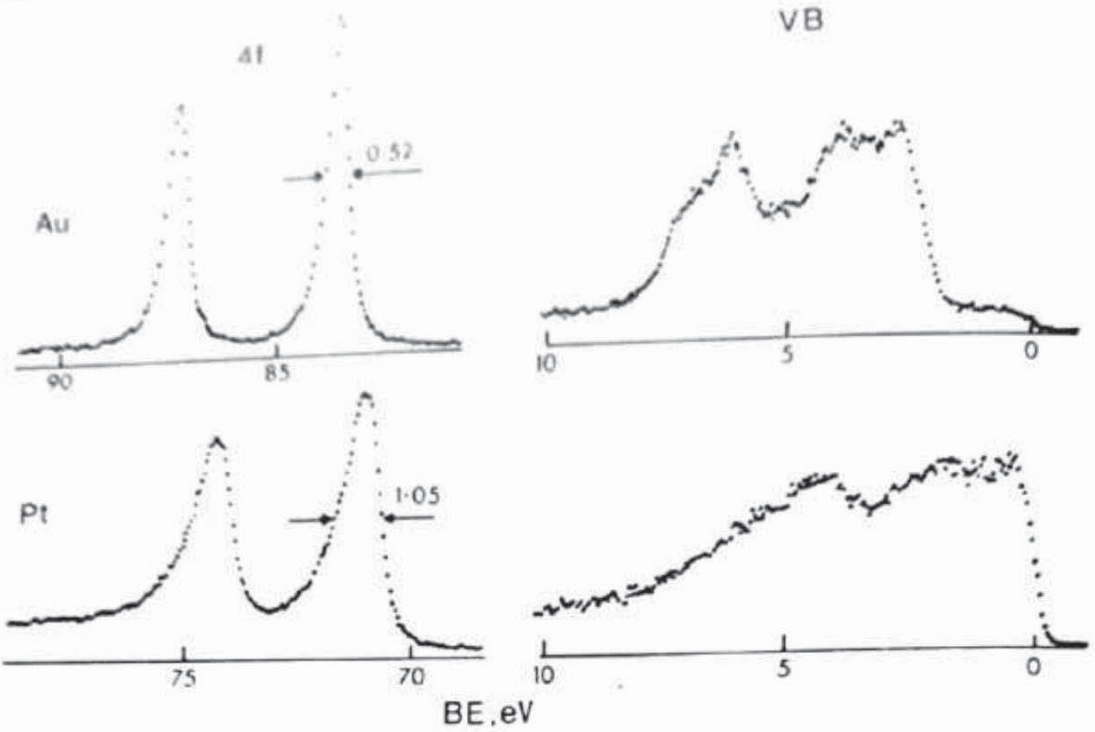


Figure 3.30 Core (4*f*) and valence (VB) photo-electron spectra of gold and platinum recorded using monochromatic Al $K\alpha$ radiation. Note the relationship between the degree of core-level asymmetry and the density of states at the Fermi level (BE = 0 eV). Reproduced from Barrie, Swift and Briggs[49])

shake-up-type events following core electron emission. In this case, instead of discrete satellites to low KE of the primary peak, the expectation is for the core level itself to be accompanied by a low KE tail. The higher the density of states at the Fermi level the more likely is this effect to be observed. This is confirmed by Figure 3.30. Such asymmetries must be borne in mind when attempting to deconvolute complex envelopes involving peaks from bulk metals (see Appendix 3).

3.4.3.5 Shake-off satellites

In a process similar to 'shake-up' valance electrons can be completely ionized, i.e. excited to an unbound continuum state. This process, referred to as 'shake-off', leaves an ion with vacancies in both the core level and a valence level. Discrete shake-off satellites are rarely discerned in the solid state because: (a) the energy separation from the primary photo-electron peak is greater than for shake-up satellites, which means the satellites tend to fall within the region of the broad inelastic tail, and (b) transitions from discrete levels to a continuum produce onsets of increased intensity (i.e. broad shoulders) rather than discrete peaks. However, a weak structure on the inelastic tail of C 1*s* peaks from polymers has been identified[50] (following double dif-

ferentiation) as a shake-off structure which images the polymer valence band structure.

Plasmon loss processes have already been discussed in Section 3.3.1.6. Loss peaks of exactly the type described are found associated with core levels in exactly the same way as with Auger peaks.[51]

3.4.4 Angular effects

Two types of angular effect are important in XPS. The first involves the increase in surface sensitivity obtained at low angles of electron exit to the surface, whilst the second is found in single-crystal systems giving rise to 'photo-electron diffraction'.

3.4.4.1 Enhancement of surface sensitivity

The reason for this effect is simply demonstrated by reference to Figure 3.31. If λ is the inelastic mean free path (IMFP) of the emerging electron then 95 per cent of the signal intensity is derived from a distance 3λ within the solid. However, the vertical depth sampled is clearly given by

$$d = 3\lambda \sin \alpha \tag{3.19}$$

and this is a maximum when $\alpha = 90°$. In the case of a substrate (s) with a uniform thin overlayer (o) the angular variation of intensities is given by

$$I_s^d = I_s e^{-d/\lambda \sin \alpha} \tag{3.20}$$

and

$$I_o^d = I_o(1 - e^{-d/\lambda \sin \alpha}) \tag{3.21}$$

where λ is the appropriate value for the observed photo-electron. In the ideal

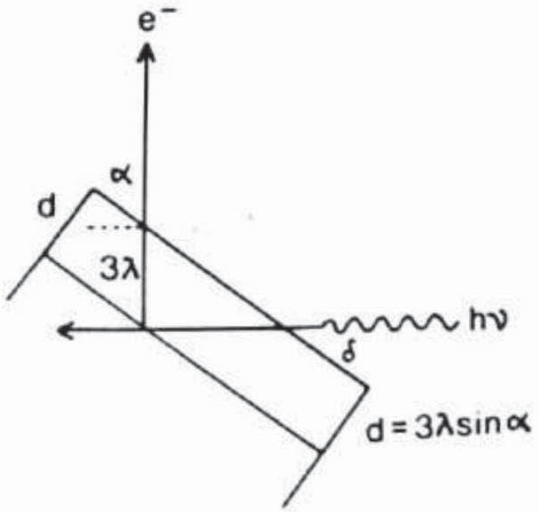


Figure 3.31 Surface sensitivity enhancement by variation of the electron 'take-off' angle

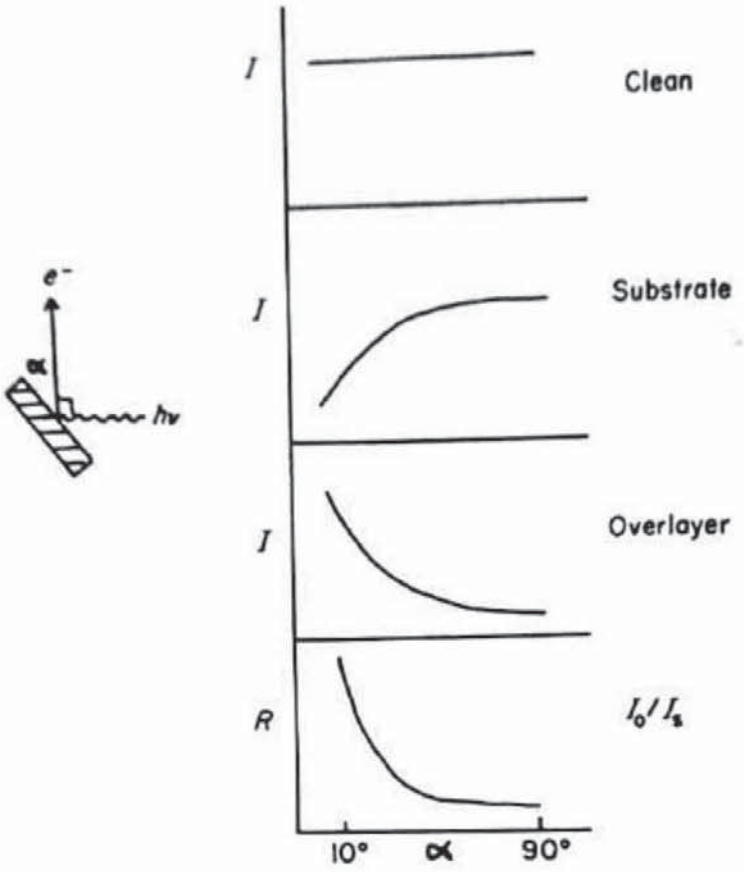


Figure 3.32 Theoretical angular dependent curves for a clean flat surface and a flat overlayer/substrate system

situation these equations lead to curves of the type shown in Figure 3.32. However, the real case is usually complicated by the fact that the system geometry imposes a response function also dependent on α. A useful summary of these effects for one particular spectrometer has been given by Dilks.[52] This complication is avoided by measurements of relative values of I_o/I_s so

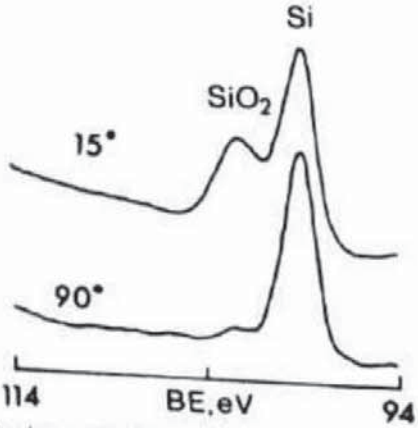


Figure 3.33 Effect of variation of take-off angle on the Si 2*p* spectrum from silicon with a passive oxide layer. Note the relative enchancement of the (surface) oxide signal at low angle (measured with respect to the surface). (After Wagner *et al.*[27])

that the instrument response function is cancelled. Thus, as shown in Figure 3.32, at low values of α, I_o/I_∞ increases significantly. An example of the effect is given in Figure 3.33.

The major requirement for surface sensitivity enhancement is that the surface is flat. Surface roughness leads to an averaging of electron exit angles and also to shadowing effects (both of the incident X-rays and emerging electrons) such that in most cases the surface enhancement effect cannot be observed.[53] In most commercial spectrometers the angle between the incident X-rays and detected electrons ($180 - \alpha - \delta$) is fixed. However, it has been shown that at *very* low δ there is also a surface sensitivity enhancement effect due to the rapid fall-off in X-ray penetration depth as δ tends to grazing incidence.[54]

3.4.4.2 *Photo-electron diffraction*

Measurement of core electron intensity from a single crystal surface as a function of electron take-off angle gives rise to plots of the type shown in Figure 3.34 (and for any such angle similar plots will arise by rotation in the plane of the surface, i.e. by variation of the azimuthal angle). The pronounced fine structure superimposed on the instrument response function is primarily

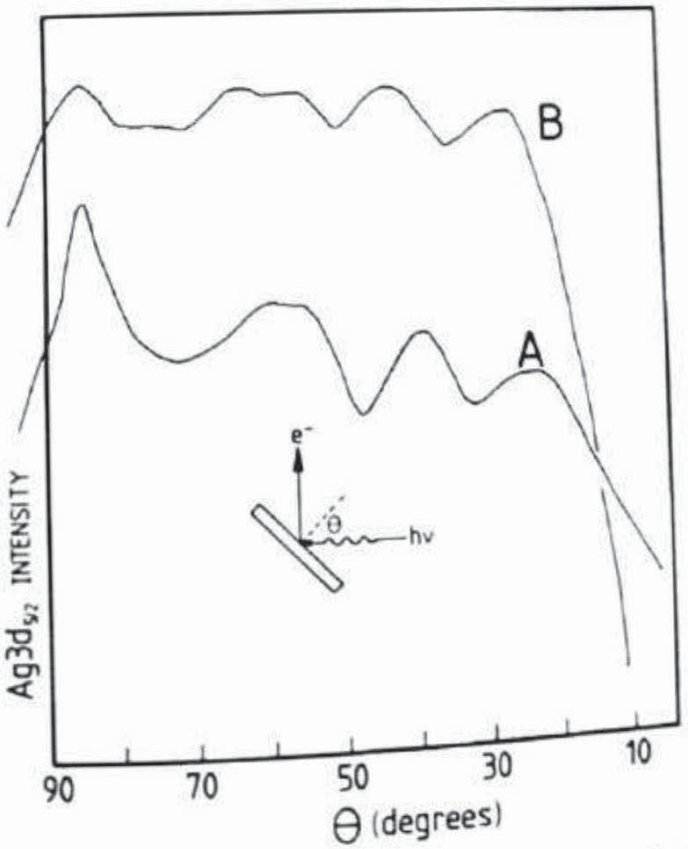


Figure 3.34 Angular variation of Ag $3d_{5/2}$ intensity from a clean (110) surface: (a) chemically polished, (b) not chemically polished after orientation. In both cases final cleaning was by argon ion etching and annealing cycles. (Reproduced from Briggs, Marbrow and Lambert[55] by permission of Pergamon Press Ltd.)

due to diffraction of photo-electrons from crystal planes. Studies of the related phenomena of high-energy electron channelling in crystals and Kikuchi band formation in LEED indicate[33] that photo-electron emission parallel with each set of planes (hkl) to within $\pm\theta_{hkl}$ (the first-order Bragg angle) will give rise to enhanced emission intensity. Here hkl are the usual Miller indices and θ_{hkl} is defined by

$$\lambda_e = 2d_{hkl} \sin \theta_{hkl} \tag{3.22}$$

where λ_e is the de Broglie wavelength of the core electron in question and d_{hkl} is the spacing between adjacent planes.

Photo-electron diffraction effects obviously complicate the quantification of peak intensity data from single crystal surfaces, but they have already been put to positive use analytically. Figure 3.35 shows how Evans and Scott[56] used X-ray photo-electron diffraction (XPD) patterns (the plot of relative core-

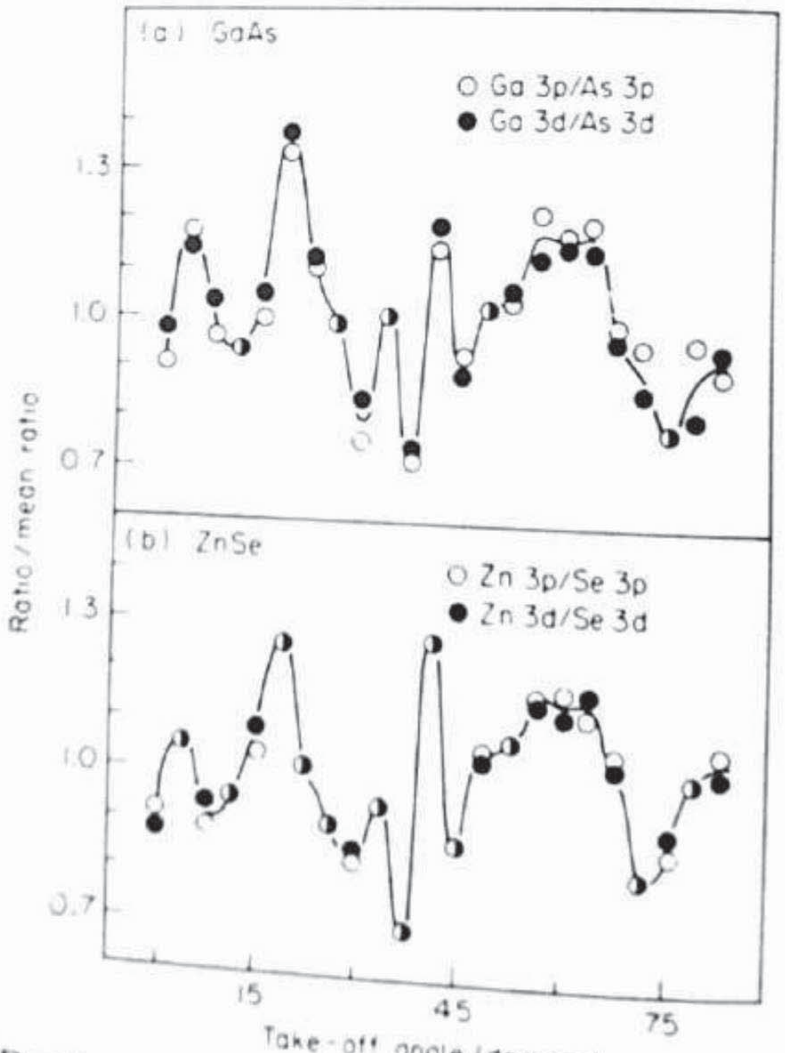


Figure 3.35 XPD patterns for (a) (100) GaAs wafer and (b) ZnSe layer grown on an identical GaAs substrate by hydrogen transport. The axis of rotation lay parallel to <110> and the take-off angle is given with respect to the surface normal. (After Evans and Scott[56])

level intensity as a function of take-off angle) to show that ZnSe grown onto (100) GaAs was expitaxial. The identical form of the plots immediately reveals identical crystal symmetry of substrate and overlayer. Evans, Adams and Thomas[57] have also carried out structural analyses of complex layer–silicate minerals using this technique.

3.4.5 Time-dependent spectra

Although XPS is generally regarded as a 'non-destructive' technique (particularly cf. AES and SIMS) this cannot be taken for granted. A fairly obvious problem concerns the exposure of materials to the high or ultra-high vacuum of the spectrometer. A recent systematic study[58] of several materials containing water of crystallization and some hydroxides has shown that their behaviour under vacuum can be predicted from the relationship between water vapour pressure and temperature (the Clausius–Clapeyron equation). Water loss is a particularly serious problem for intended studies of biological systems.

Less easy to predict is damage from the incident X-rays. In several cases the time dependence of XP spectra has been interpreted structurally and kinetically (e.g. reduction of platinum (IV) compounds, decomposition of $NaClO_4$).[59] Some radiation effects are essentially instantaneous at room temperature and therefore give rise to confusion. Perhaps the first example of this to be studied was the surface isomerization of the compound $(Ph_3P)_2Pt(C_2Cl_4)$ to $(Ph_3P)_2PtCl(C_2Cl_3)$.[60] Cooling the sample slows down these changes and often allows them to be followed *in situ*. X-radiation-induced changes are probably much more common than is often realized. A long literature argument about the appearance of a high BE (oxidized) S 2*p* peak in spectra of biologically important copper–sulphur proteins seems recently to have been resolved by the theory[61] that a radiation-induced redox reaction at surface copper–sulphur sites involving adsorbed water takes place, producing sulphone groups. Such examples indicate that radiation effects in XPS should be taken much more seriously than they have been hitherto.

References

1. See, for example, R. M. Eisberg, *Fundamentals of Modern Physics*, p. 436, John Wiley and Sons, New York (1963).
2. K. D. Sevier, *Low Energy Electron Spectrometry*, p. 55, John Wiley and Sons, New York (1972).
3. M. F. Chung and L. H. Jenkins, *Surface Science*, **21**, 253 (1970).
4. W. A. Coghlan and R. E. Clausing, *Atomic Data*, **5**, 317 (1973).
5. K. S. Kim, S. W. Gaarenstroom and N. Winograd, *Chem. Phys. Letters*, **41**, 503 (1976).
6. F. P. Larkins, *Atomic Data and Nuclear Tables*, **20**, 311 (1977).

7. L. E. Davies, N. C. MacDonald, P. W. Palmberg, G. E. Riach and R. E. Weber, *Handbook of Auger Electron Spectroscopy*, Physical Electronics Division, Perkin–Elmer Corporation, USA (1978). See also G. E. McGuire, *Auger Electron Spectroscopy Reference Manual*, Plenum, New York (1979).
8. E. Kny, *J. Vac. Sci. Tech.*, **17**, 658 (1980); J. Kleefeld and L. L. Levenson, *Thin Solid Films*, **64**, 389 (1979); M. A. Smith and L. L. Levenson, *Phys. Rev.*, **B16**, 1365 (1977).
9. M. Salmeron, A. M. Baro and J. M. Rojo, *Phys. Rev.*, **B13**, 4348 (1976).
10. See, for example, H. H. Madden, *J. Vac. Sci. Tech.* **18**, 677 (1981).
11. P. Weightman, *Rep. Prog. Phys.*, **45**, 753 (1982).
12. H. Aksela and S. Aksela, *J. Phys. B: Atom Mol. Phys.*, **7**, 1262 (1974).
13. R. H. Ritchie, *Phys. Rev.*, **106**, 874 (1957).
14. E. A. Stern and R. A. Ferrel, *Phys. Rev.*, **120**, 130 (1960).
15. G. Dutour, J. M. Mariot, P. E. Nilsson-Jatko and R. C. Karnatak, *Physica Scripta*, **13**, 370 (1976).
16. D. Massignon, F. Pellerin, J. M. Fontaine, C. le Gressus and T. Ichinokawa, *J. Appl. Phys.*, **51**, 808 (1980).
17. H. E. Bishop and J. C. Rivière (to be published).
18. J. H. Scofield, *J. Electron Spectrosc.*, **8**, 129 (1976).
19. A. F. Orchard, in *Handbook of X-ray and Ultraviolet Photoelectron Spectroscopy* (Ed. D. Briggs), p. 49, Heyden and Son, London (1977).
20. A. Barrie and N. E. Christensen, *Phys. Rev.*, **B14**, 2442 (1976).
21. H. Hillig, B. Cleff and W. Mehlhorn, *Z. Phys.*, **268**, 235 (1974).
22. S. Aksela and H. Aksela, *Phys. Lett.*, **A48**, 19 (1974).
23. C. D. Wagner, in *Handbook of X-ray and Ultraviolet Photoelectron Spectroscopy* (Ed. D. Briggs), p. 256, Heyden and Son, London (1977).
24. C. D. Wagner and J. A. Taylor, *J. Electron Spectrosc.*, **20**, 83 (1980).
25. K. Siegbahn *et al.*, *ESCA Applied to Free Molecules*, North Holland, Amsterdam (1969).
26. D. N. Hendrickson, J. M. Hollander and W. L. Jolly, *Inorg. Chem.*, **8**, 2642 (1969).
27. C. D. Wagner *et al.*, in *Handbook of X-ray Photoelectron Spectroscopy* (Ed. G.E. Muilenburg), Perkin–Elmer Corporation, USA (1979).
28. G. K. Wertheim, *J. Franklin Institute*, **298**, 289 (1974).
29. G. K. Wertheim, M. Campara, M. R. Shanks, P. Zuonsteg and E. Banks, *Phys. Rev. Lett.*, **34**, 738 (1975).
30. C. D. Wagner and P. Biloen, *Surf. Sci.*, **35**, 82 (1973).
31. C. D. Wagner, in *Handbook of X-ray and Ultraviolet Photoelectron Spectroscopy* (Ed. D. Briggs), p. 262, Heyden and Son, London (1977).
32. C. D. Wagner, *Farad. Discuss. Chem. Soc.*, **60**, 291 (1975).
33. C. D. Wagner, L. H. Gale and R. H. Raymond, *Anal. Chem.*, **51**, 466 (1979).
34. R. H. West and J. E. Castle, *Surf. Interface Anal.*, **4**, 68 (1982).
35. C. D. Wagner, in *Handbook of X-ray and Ultraviolet Photoelectron Spectroscopy* (Ed. D. Briggs), p. 266, Heyden and Son, London (1977).
36. C. D. Wagner, D. A. Zatko and R. H. Raymond, *Anal. Chem.*, **52**, 1445 (1980).
37. T. A. Carlson, *Photoelectron and Auger Spectroscopy*, Plenum, New York (1976).
38. B. E. Koel and J. M. White, *J. Electron Spectrosc.*, **22**, 237 (1981).
39. S. P. Kowalczyk, L. Ley, R. A. Pollak, M. R. McFeely and D. A. Shirley, *Phys. Rev.*, **B7**, 4009 (1973).
40. C. S. Fadley, in *Electron Spectroscopy: Theory, Techniques and Applications* (Eds C. R. Brundle and A. D. Baker), Vol. 2, Academic Press, London (1978).
41. D. T. Clark and D. B. Adams, *Chem. Phys. Lett.*, **10**, 121 (1971).

42. D. Briggs and V. A. Gibson, *Chem. Phys. Lett.*, **25**, 493 (1974).
43. G. C. Allen and P. Tucker, *Inorg. Chim. Acta*, **16**, 41 (1976).
44. D. T. Clark and A. Dilks, *J. Polym. Sci., Polym. Chem. Ed.*, **15**, 15 (1977).
45. D. C. Frost, C. A. McDowell and A. Ishitani, *Mol. Phys.*, **24**, 861 (1972).
46. P. Burroughs, A. Hamnett, A. F. Orchard and G. Thornton, *J. Chem. Soc. Dalton Trans.*, **1976**, 1686.
47. L. J. Matienzo, W. E. Swartz and S. O. Grim, *Inorg. Nucl. Chem. Lett.*, **8**, 1085 (1972).
48. D. C. Frost, C. A. McDowell and I. S. Woolsey, *Mol. Phys.*, **27**, 1473 (1974).
49. A. Barrie, P. Swift and D. Briggs, Preprints, 27th Pittsburgh Conference on Analytical Chemistry and Applied Spectroscopy (Cleveland, USA, 1976).
50. J. J. Pireaux, R. Caudano and J. Verbist, *J. Electron Spectrosc.*, **5**, 267 (1974).
51. See, for example, K. Siegbahn, *J. Electron Spectrosc.*, **5**, 3 (1974).
52. A. Dilks, in *Electron Spectroscopy: Theory, Techniques and Applications* (Eds C. R. Brundle and A. D. Baker), Vol 4, Academic Press, London (1981).
53. C. S. Fadley, *Prog. Solid State Chem.*, **11**, 265 (1976).
54. M. Mehta and C. S. Fadley, *Phys. Lett.*, **A55**, 59 (1975).
55. D. Briggs, R. A. Marbrow and R. M. Lambert, *Solid State Comm.*, **26**, 1 (1978).
56. S. Evans and M. D. Scott, *Surf. Interface Anal.*, **3**, 269 (1981).
57. S. Evans, J. M. Adams and J. M. Thomas, *Phil. Trans. Roy. Soc. Lond.*, **292**, 563 (1979).
58. K. Hirokawa and Y. Danzaki, *Surf. Interface Anal.*, **4** (1982).
59. R. G. Copperthwaite, *Surf. Interface Anal.*, **2**, 17 (1980).
60. D. T. Clark and D. Briggs, *Nature Phys. Sci.*, **237**, 15 (1972).
61. M. Thompson, R. B. Lennox and D. J. Zemon, *Anal. Chem.*, **51**, 2260 (1979).

Practical Surface Analysis
by Auger and X-ray Photoelectron Spectroscopy
Edited by D. Briggs and M. P. Seah

Chapter 4
Depth Profiling

S. Hofmann
Max-Planck-Institut für Metallforschung, Institut für Werkstoffwissenschaften, Seestrasse 92, D 7000 Stuttgart 1, FRG

4.1 Introduction

In-depth distribution analysis of chemical composition is a special case of micro local analysis, where the third dimension perpendicular to the surface of a sample is of primary interest. In principle, this task requires the compositional analysis of thin sections (in the ultimate dimension of a monoatomic layer) defined on a depth scale. It can be obtained either by (a) non-destructive or (b) destructive techniques.[1–7]

4.1.1 Non-destructive techniques

Non-destructive techniques are based on an analytical signal parameter (e.g. intensity, energy) which has a well-defined dependence on its depth of origin. For example, the energy loss of high-energy charged particles is proportional to their path length in a given material. This property is extensively used in RBS (Rutherford back-scattering of primary ions, typically He^+, 100 keV–5 MeV, 0.1–30 μm depth).[8,9] Here the atomic number dependent elastic scattering cross-section contains the analytical information whereas the energy loss yields the depth information. In NRA (nuclear reaction analysis), the resonance of nuclear reaction cross-sections with energy of the impinging particles is used for profiling: by increasing the primary beam energy the reaction products originate from progressively deeper regions in the sample. These techniques are excellent with respect to quantitative results of both concentration and depth scale, with sensitivities more or less comparable to electron spectroscopy. Because of low scattering cross-sections for light elements, RBS is only appropriate for elements of $Z > 10$ and is of limited use for light elements in a heavy element matrix. NRA is defined to the detection of low Z elements ($Z < 20$) due to the presently used accelerator energies ($\leq$5 MeV). The depth resolution obtainable with RBS and NRA, mainly due

to range straggling and useful energy resolution, is limited to between 5 and 200 nm depending on the material and total layer thickness.[8,9] For non-destructive techniques using X-ray analysis such as IIX (ion induced X-ray) analysis or EPMA (electron probe microanalysis), the depth resolution is even worse (typically fractions of 1 μm), which prevents their use in thin film depth profiling.

In electron spectroscopy, the intensity of a specific signal is dependent on the energy and the depth of origin of the Auger electrons or photo-electrons. The inelastic mean free path is the decisive parameter which determines the escape depth. For a given electron energy, a variation of the emission angle allows the variation of the effective escape depth between its full value perpendicular to the surface and almost zero.[4,10,11] This technique, which is limited to 50 nm total depth, is discussed in Section 4.5.1 in more detail.

4.1.2 Destructive techniques

Destructive techniques of depth profiling comprise the more classical methods of mechanical sectioning (tapered sectioning with recent improvements like ball cratering in conjunction with high lateral resolution techniques) and chemical sectioning with subsequent chemical analysis of the removed material. Apart from problems with specific materials (e.g. ceramics) depending on mechanical and chemical properties, these techniques suffer from the lack of control of surface reactions between subsequent steps, giving severe limitations in depth resolution, although recent progress in electrochemical polishing[12] has shown that depth resolution can be achieved in the 10 nm region.

A universally applicable 'sectioning' method is surface erosion by ion sputtering. The case of *in situ* combination with any surface analysis method has enhanced its now widespread application for depth profiling. A number of review articles have been published on this subject.[1-7]

Sputtering is a destructive method: the sample is bombarded with ions accelerated in an ion gun to an energy above 100 eV (typically 0.5–5 keV). A small fraction of the energy is transferred to surface atoms and causes them to leave the sample; they are sputtered away. Thus, the sample is successively decomposed in an abraded part which can be analysed, e.g. by SIMS, and in the residual surface which is analysed by AES and/or XPS.

The principle advantages of depth profiling by AES and XPS combined with ion sputtering are:

(a) The information depth is of the order of 1 nm.
(b) The analysis is independent of the sputtering yield.
(c) The influence of the matrix on the elemental detection sensitivity is small.

(d) The analysed area is small compared to the sputtered area, thus minimizing crater edge effects.

Point (d) is generally not fulfilled in XPS, as discussed later in Section 4.4.2.

4.2 Practice of Depth Profiling with AES and XPS

Since both AES and XPS analyse the residual surface left after a certain sputtering time, the basic implications with respect to depth profiling apply for both techniques. There are, however, a number of practical differences, e.g. detection speed, background and spatial resolution, which generally are more advantageous in Auger profiling. Therefore, we refer primarily to this technique if not otherwise stated. Any commercial equipment for AES/XPS comprises an electron energy analyser, an electron and/or X-ray source, an ion gun and the sample stage all mounted in a stainless steel chamber in which an ultra-high vacuum ($\leq 10^{-8}$ Pa) can be maintained.

4.2.1 Vacuum requirements

Considering the mean free path of electrons in gases, a vacuum better than 10^{-2} Pa would be sufficient for proper operation of the electron analysis. However, contamination of the surface during analysis has to be avoided.[13-15] Assuming a sticking coefficient of one, an adsorption monolayer is built up at a partial pressure of 10^{-4} Pa ($\approx 10^{-6}$ torr) in one second. To avoid an increase of 1 per cent of a monolayer for a typical measurement time (AES) of 100 seconds, the partial pressure of reactive gases such as CO, H_2O, C_xH_y should be below 10^{-8} Pa (10^{-10} torr), i.e. UHV. Inert gases (e.g. argon) which do not chemisorb can be tolerated up to a pressure several orders of magnitude higher.

4.2.2 Ion gun

The most frequently used ion guns in Auger spectrometers are simple electrostatic devices where the inert gas ions (Ar) are generated by collisional excitation with electrons of typically 100 eV energy from a hot filament. The positive ions are accelerated to between 0.5 and 5 keV and focused on the sample, creating a sputtering spot diameter of about 1–5 mm diameter. To achieve an ion current density of the order of 100 $\mu A/cm^2$, the pressure in the ion formation section should be about 5×10^{-3} Pa (5×10^{-5} torr). The general operation in ion pumped systems is to back-fill the whole chamber with argon with the ion pumps off. The sputtering gas purity required should meet the partial pressure ratio $10^{-8}/5 \times 10^{-3} = 2 \times 10^{-6}$. To avoid further contamination, a titanium sublimation pump with a cryopanel usually serves as an

efficient scavenger for the reactive gases. Compared to such 'static' systems, by using the more modern type of ion gun operating in a 'dynamic' mode with constant throughput of the sputtering gas (pumped away by the main chamber pump) the overall contamination is reduced. A further improvement is differentially pumping the ion gun by a getter pump or an added turbomolecular or diffusion pump. The decreased pressure in the main chamber ($< 10^{-2} \times$ ion gun pressure) which is due to the pressure stage action of the small opening (≈1 mm) at the end of the ion gun offers an additional advantage: the sputter damage of the electron gun cathode filament caused by accelerated positive ions is reduced and/or the X-ray source is prevented from arcing which is frequently encountered at a total pressure $>10^{-4}$ Pa.

Simple ion guns using a static spot must be aligned mechanically so that the analyser focus (together with the electron beam impinging on the sample) is close to the centre of the generally Gaussian ion current beam intensity distribution to avoid severe degradation of depth resolution.[14] The lack of uniformity of the ion beam intensity over millimetre dimensions precludes their application in XPS profiling.

For precise depth profiling, an ion gun with x/y beam deflection capability should be used. This enables an exact matching of analysed and sputtered areas to be made (e.g. by optimizing for the secondary electron emission monitored by a CMA). Rastering of the well-focused ion beam over a larger area (up to 10 × 10 mm) greatly improves the uniformity of the ion beam intensity,[15] leading to a flat bottom of the sputtering crater which is necessary for optimum depth resolution. Furthermore, at constant beam current the raster area is inversely proportional to the total primary ion density so that the sputtering rate is easily controlled. A rastered ion beam is a prerequisite for XPS profiling where the analysed area is some square millimetres.

To prevent shadowing effects in working with rough surfaces the angles between the ion beam , the electron beam and the electron take-off should be as small as possible. In particular, the ion beam should be directed close to the normal of the sample surface.[16] To avoid cone formation, sputtering with two guns at different angles is desirable.[17]

4.2.3 Analysing mode

Auger or photo-electron analysis can be done either discontinuously after subsequent sputtering or by continuous sputtering and simultaneous electron spectroscopy. In Auger profiling the latter technique is most frequently used together with multiplexing for different elemental peaks and a print-out of the respective Auger peak-to-peak heights (APPHS). In XPS, the strong increase in secondary electron emission during sputtering is unfavourable for simultaneous profiling. Furthermore, XPS peak area detection generally needs longer measurement times compared to AES, thus limiting the achievable

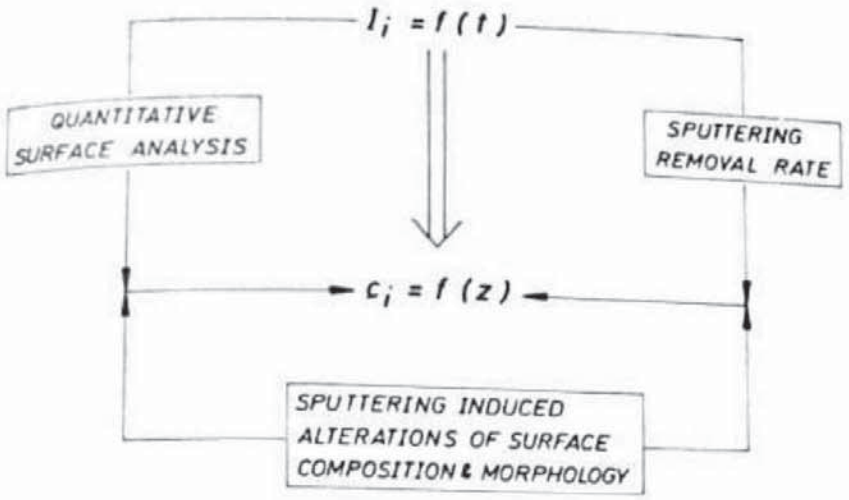


Figure 4.1 Principles of sputtering profile evaluation: conversion of a measured sputtering profile $I = f(t)$ to a true concentration profile $c = f(z)$[3]

depth resolution Δz_a for a given sputtering rate $\dot{z}$. If Δt_m is the time necessary for the measurement of a peak with a desired signal-to-noise ratio, $\Delta z_a = \dot{z}\Delta t_m$. Contrary to SIMS, where $\dot{z}$ and Δt_m are inversely proportional,[18] these quantities are independent in profiling by electron spectroscopy. Therefore Δz_a can be made negligibly small compared to the many other factors limiting depth resolution by reducing the primary ion current density or by discontinuous sputtering. In general, this procedure has to be used in XPS profiling because of the higher Δt_m. In AES profiling with a CMA, continuous sputter profiling with sputtering rates up to 5 nm/min is frequently applied. An additional advantage of a high sputtering rate is a reduced surface contamination (caused by adsorption from the residual gas atmosphere) which is due to increased sputter desorption. (cf. Section 4.4.2.1)

4.3 Quantification of Sputtering Profiles

The data primarily obtained during depth profiling experiment consist of signal intensities of the detected elements, I (e.g. Auger peak-to-peak heights), as a function of sputtering time, t, i.e. the 'measured sputtering profile', $I = f(t)$. The principal task is to obtain the original distribution of concentration c with depth z, $c = f(z)$, by an appropriate conversion of the measured data.[3–7]

The main steps of such an evaluation are visualized in Figure 4.1. First the sputtering time scale must be calibrated in terms of the main eroded depth, $z = f(t)$, and the intensity of the Auger signal must be calibrated in terms of local elemental concentration, $c = f(I)$. Thus a 'real' concentration profile is established which would be identical to the original profile if sputtering proceeded homogeneously in an ideal 'layer-by-layer' microsectioning. However,

profile distortions mainly due to sputtering-induced topographical and compositional changes of the instantaneous sample surface must be taken into account in a second step in order to reveal the true original profile, $c = f(z)$. The precision by which the shape of an evaluated profile resembles that of the original profile can be described by the concept of depth resolution,[1-7] which is the key parameter in profiling. A knowledge of the depth resolution or, more precisely, the depth resolution function is the prerequisite for a deconvolution of the measured concentration profiles.

The steps in profile evaluation are outlined in the following sections.

4.3.1 Calibration of the depth scale ($z = f(t)$)

The velocity of surface erosion is described by an instanteneous sputtering rate $\dot{z} = dz/dt$ which determines the mean eroded depth z as a function of sputtering time t according to

$$z(t) = \int_0^t \dot{z}\, dt \tag{4.1}$$

The sputtering rate $\dot{z}$ (in units of metres per second) is determined by

$$\dot{z} = \frac{M}{\rho N_A e} S j_p \tag{4.2}$$

where
M = atomic mass number
ρ = density (kg/m^3)
$N_A = 6.02 \times 10^{26}$ (Avogadro number)
$e = 1.6 \times 10^{-19}$ A s (electron charge)
S = sputtering yield (atom/ion)
j_p = primary ion current density (A/m^2)

For a constant sputtering rate, equation (4.1) shows a direct proportionality between sputtered depth and time:

$$z = \dot{z}t \tag{4.3}$$

In this case the z axis is linear with the sputtering time and (besides the zero point) only one point is necessary to determine the calibration factor $\dot{z}$. With equation (4.2), $\dot{z}$ can be calculated taking the sputtering rate S from the literature data[19-21] and measuring j_p with a Faraday cup. Figure 4.2 shows the orders of magnitude which may be achieved. However, the sputtering rate S depends on a variety of parameters including energy, mass and angle of the incident ions[19-21] and surface composition.[13] Therefore only a rough estimate can be expected from such a calculation.

A better method to obtain the sputtering rate $\dot{z}$ is to measure (for given ion gun operation parameters) the time required to sputter through a layer of

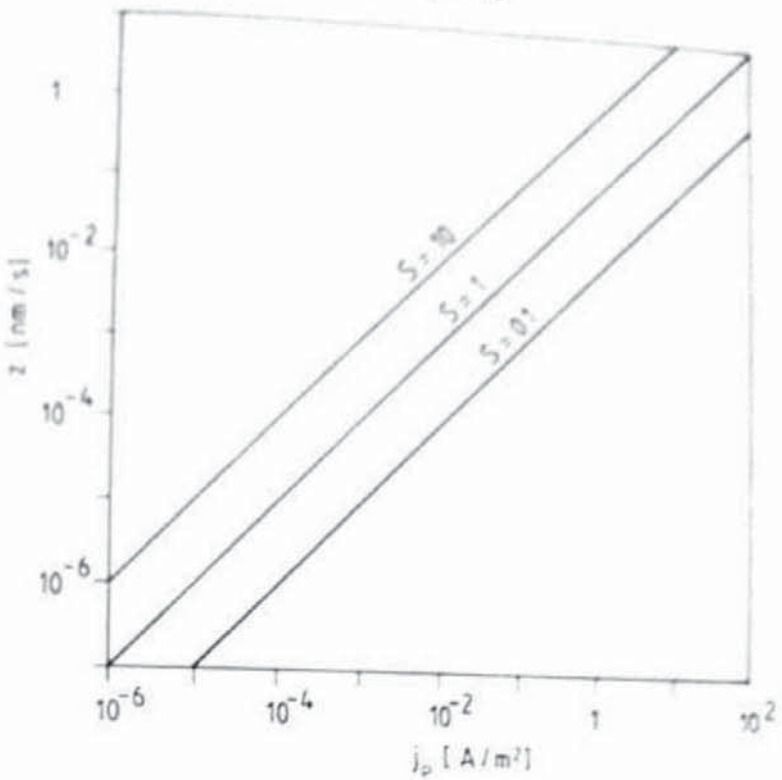


Figure 4.2 Sputtering rate $\dot{z}$ as a function of the primary ion density j_p for a mean $M/\rho = 10^{-2}$ m³/mol and three different sputtering yields S according to equation (4.2)

known thickness, e.g. a metallic evaporation layer or an oxide layer. Anodized tantalum foils[22,23] have proved very useful for this purpose, since the thickness is easily controlled by the formation voltage.[24] In addition, due to the sharp oxide–metal interface they provide a quick test method for the instrumental depth resolution (see Section 4.4). Thus $\dot{z}$ is determined with the knowledge of the Ta_2O_5 sputtering rate[23] as compared to the material under study.

Another possibility is to measure the crater depth z_0 after sputtering by conventional interferometry or stylus methods.[25,26]

In general, however, $\dot{z}$ varies with composition because according to equation (4.2) M, ρ and S are a function of composition, and the relation between sputtering time and depth becomes non-linear.[27,28] Even for a constant composition, the sputtering rate may change within the first few layers due to the gradual build-up of structural and compositional changes until a saturation value is obtained after the sputter depth is of the order of the mean projected range of the primary ions.[24]

For these reasons it is obvious that a precise determination of the true depth scale requires a measurement of the instantaneous sputtering rate during profiling. This can be done by *in situ* measurement of the residual thin film thickness, e.g. by direct mass measurement with a quartz microbalance,[30] by X-ray emission analysis[29] or by *in situ* laser optical interferometry,[31,32] as

developed recently by J. Kempf. Unfortunately, the latter methods are not available in standard equipment.

A first-order approximation of the correction for composition-dependent non-linearity in the time–depth relation has been demonstrated in the sputtering of Ni/Cr evaporation multilayers,[3,28] assuming a change of the sputtering rate proportional to composition. For a binary system A/B, the total instantaneous sputtering rate $\dot{z}$ is given by the mole fractions X_A, X_B and the sputtering rates of the pure components, $\dot{z}_A$, $\dot{z}_B$:

$$\dot{z} = X_A\dot{z}_A + X_B\dot{z}_B \tag{4.4}$$

According to equations (4.1) and (4.4):

$$z_1(t_1) = \int_0^{t_1} [X_A(t)\dot{z}_A + X_B(t)\dot{z}_B]\,dt \tag{4.5}$$

In this approximation, X_A and X_B may be taken from the normalized Auger signals I_A/I_A^0, I_B/I_B^0.

Limiting conditions and—in special cases—an estimation of the sputtering rate can be derived from the consideration of the electron spectroscopic signal quantification, as shown in the next section.

4.3.2 Calibration of the concentration scale ($c = f(I)$)

Quantitative AES and XPS are discussed in Chapter 5. With respect to depth profiling the important feature is the relation between intensity I_i and concentration c_i of an element i and can be described by an explicit function of the electron escape depth and—in the case of AES—an electron back-scattering factor:[33–38]

$$I_i = \frac{I_i^0}{\lambda_i}\int_0^{\infty} r_i(z)c_i(z)\exp\left(\frac{-z}{\lambda_i}\right)dz \tag{4.6}$$

where I_i^0 is the intensity for an elemental bulk standard, λ_i is the 'effective' escape depth of the Auger electrons or photo-electrons perpendicular to the surface, $c_i(z)$ is the local concentration (in mole fractions) at depth z and $r_i \geq 1$ is the back-scattering (cf. Chapter 5 of this book). The effective escape depth λ_i is determined by the inelastic mean free path[35] of the electrons for a given energy and material, λ_i^0, and by the angle of emission φ of the detected electrons with respect to the normal to the sample surface, according to:

$$\lambda_i = \lambda_i^0 \cos\varphi \tag{4.7}$$

Values for λ_i^0 are typically between 0.4 and 4 nm.[35] The angle φ is clearly defined for hemispherical analysers. For a cylindrical mirror analyser (CMA), a mean angle of emission, φ, can be defined.[33] Its dependence on the angle a

between the normal to the sample surface and the CMA axis is shown in Figure 4.3. For $\alpha > 47.7°$, part of the CMA acceptance cone is shadowed by the sample which gives a deviation from the cosine dependence, as shown in Figure 4.3.

The back-scattering factor is difficult to correct in the presence of concentration gradients. An approximation for binary systems has been given by Morabito and others.[6,17] In principle, the depth distribution of the excitation density must be known which is dependent on the concentration distribution. For a binary system, $r_{M,A} = (1 - r_B/r_A)c_A + r_B/r_A$. It has been shown[4,39] that for the intensity profile at a binary interface A/B a certain characteristic length for back-scattering (l_b) leads to an amplification factor for the A intensity given by $[1 + (r_B/r_A - 1)\exp - (z_0 - z)/l_b]$ (where z_0 is the position of the interface), which is in accordance with experimental results.[4] In general it is easier to apply the back-scattering correction after the escape depth correction. With the exception of extreme cases (sharp interfaces of elements with large difference in atomic number) the electron back-scattering in AES is

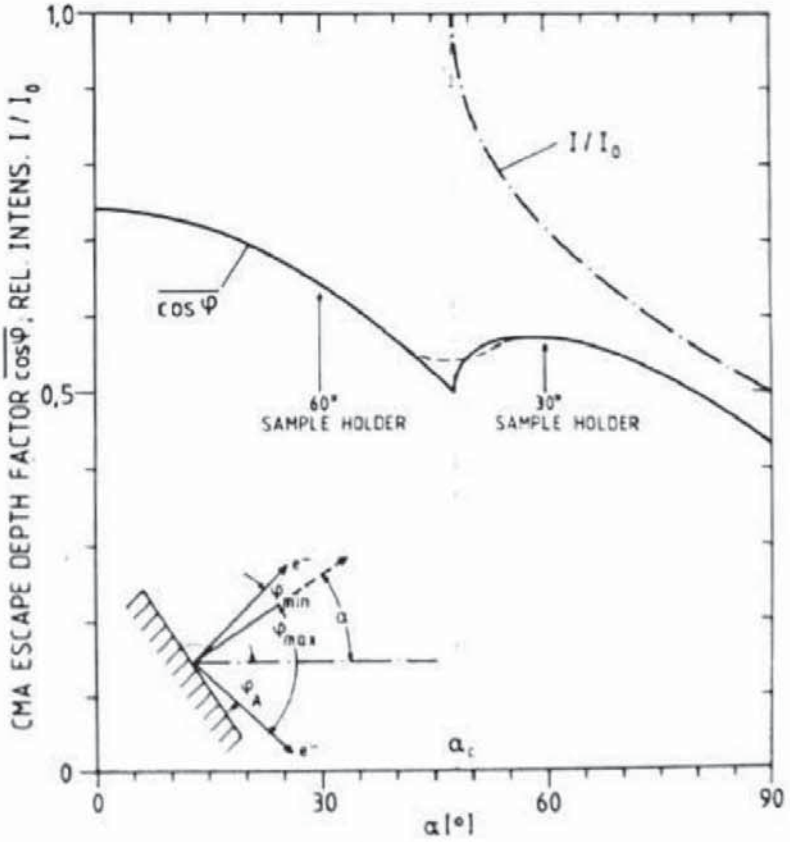


Figure 4.3 Mean effective escape depth factor ($\overline{\cos\varphi}$) as a function of the tilt angle α between the CMA axis and the sample normal. For $\alpha > \alpha_c = 47.7°$, the transmitted intensity I/I_0 decreases due to shadowing. The dashed line accounts for an angular acceptance $\Delta\varphi = 12°$

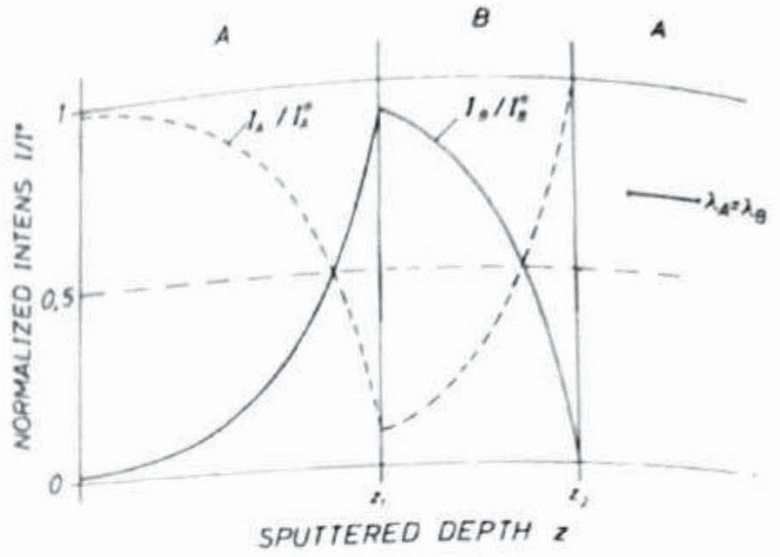


Figure 4.4 Shape of the measured profile of a thin sandwich layer A/B/A due to the escape depth influence after equations (4.8a,b), assuming ideal microsectioning[37]

only a second-order effect compared to the escape depth influence which is discussed in the following.

Neglecting r_i, equation (4.6) is valid for both AES and XPS. Depth profiling means a shift of the lower integration boundary from zero to an instantaneous sputter depth. Thus, for a given profile $c(z)$, integration of equation (4.6) as a function of the parameter z gives $I(z)$ for the case of ideal, continuous sputter erosion.[3,7,39] The expected intensity profile for the typical case of a thin sandwich layer of an element B in a matrix A is shown in Figure 4.4. Integration of equation (4.6) for this case gives

$$\left(\frac{I_B}{I_B^0}\right)_z = c_B\left[\exp\left(-\frac{z_1 - z}{\lambda_{B,A}}\right) - \exp\left(-\frac{z_2 - z}{\lambda_{B,B}}\right)\right] \tag{4.8}$$

for $z_1 - z \geq 0$ and $z_2 - z \geq 0$. Note that $\lambda_{B,A}$ is the electron escape depth of element B in A and $\lambda_{B,B}$ thatof B in B. For $\lambda_{BA} = \lambda_{BB} = \lambda$ the intensity–depth relations are, for $z \leq z_1$:

$$\left(\frac{I_B}{I_B^0}\right)_z = c_B\left[1 - \exp\left(-\frac{z_1 - z_2}{\lambda}\right)\right]\exp\left(-\frac{z_1 - z}{\lambda}\right) \tag{4.8a}$$

and for $z_1 < z < z_2$

$$\left(\frac{I_B}{I_B^0}\right)_z = c_B\left[1 - \exp\left(-\frac{z_2 - z}{\lambda}\right)\right] \tag{4.8b}$$

Since any concentration profile can be approximated by a successive number of thin layers of thickness $z_2 - z_1$, equation (4.8) provides a method

to receive the intensity at any depth z as a summation of these layers contributions.[7,39,42] In practice, it is sufficient to replace the upper integration limit (∞) by $z_u = z + 5\lambda$ (equation 4.6), because the maximum contribution (≤ 0.7 per cent) approaches the detection limit of AES or XPS.

A general solution of equation (4.6) for an arbitrary profile can be derived from equation (4.8) and is given by[7,43]

$$c_i(z) = \left(\frac{I_i}{I_i^0}\right) - \frac{\mathrm{d}(I_i/I_i^0)}{\mathrm{d}z}\lambda_i \tag{4.9}$$

Using equation (4.9), any measured sputtering profile can be corrected for the escape depth influence.[29,39,44,45] As seen from Figure 4.4, a shift of the real profile of 0.7λ to lower depth with respect to the measured profile is obtained, which is broadened by $\Delta z_\lambda \approx 1.6\lambda$.[5,7,39] It is interesting to note that a maximum obtainable gradient follows from equation (4.9):[47]

$$\left|\frac{\mathrm{d}I_i}{\mathrm{d}z}\right| \leq \frac{I_i}{\lambda_i} \tag{4.10}$$

This means a lower limit for the time–depth calibration factor $\dot{z}$, because $\mathrm{d}z = \dot{z}\,\mathrm{d}t$ gives

$$\dot{z} \geq \frac{\lambda_i}{I_i}\left|\left(\frac{\mathrm{d}I_i}{\mathrm{d}t}\right)\right| \tag{4.10a}$$

Frequently, calibration of the intensities in a depth profile is made using relative sensitivity factors[34,46] and taking the concentration to be proportional to the instantaneous intensity—that means neglecting the λ effect. In this case, the relative error $(\Delta c_i/c_i)_\lambda$ depends on the slope of the measured profile and is of the order of the relative change of the intensity within a depth λ:

$$\left(\frac{\Delta c_i}{c_i}\right)_\lambda \approx \frac{I_i(z) - I_i(z+\lambda)}{I_i(z)} \tag{4.11}$$

Applying equation (4.9) to the measured sputtering profile, a knowledge of the effective electron escape depth (and the back-scattering factor in AES) of the normalized intensities and of the instantaneous sputtering rate allows a quantitative evaluation of the measured depth profile (cf. Figure 4.1). For the special case of the detection of two peaks of the same element at different kinetic energies and therefore different escape depths λ_1, λ_2 (e.g. 60 and 920 eV copper AES peaks), equation (4.9) offers a possibility to derive the sputtering rate in terms of the two escape depths[47] if a measurable intensity slope with sputtering time exists. Since the right-hand side of equation (4.9) is the same for $I_1(\lambda_1)$ and $I_2(\lambda_2)$, it follows with $\mathrm{d}z = \dot{z}\,\mathrm{d}t$ that

$$\dot{z} = \frac{\lambda_1\left|\mathrm{d}(I_1/I_1^0)/\mathrm{d}t\right| - \lambda_2\left|\mathrm{d}(I_2/I_2^0)/\mathrm{d}t\right|}{(I_1/I_1^0) - (I_2/I_2^0)} \tag{4.12}$$

Although the determination of $\dot{z}$ by equation (4.12) is of limited accuracy, it allows a quick test of the order of magnitude of the sputtering rate.[47]

At this stage we have not yet considered the sputtering-induced distortions of the original concentration profile which must be taken into account for its evaluation.

4.4 Sputtering Profile Compared to the Original Concentration Profile: The Concept of Depth Resolution

4.4.1 Depth resolution

4.4.1.1 Definition of depth resolution

The influence of the mean escape depth of the Auger electrons or photoelectrons on the measured profile results in a certain broadening with respect to the true, original profile as discussed in Section 4.3.2. So far, we still have considered sputtering to proceed as an ideal microsectioning, continuous with eroded depth. Experimentally this cannot be realized because ion bombardment inevitably induces changes in the topography and in the composition of the surface region of a sample. These sputtering-induced effects are recognized in the measured profile as an additional apparent broadening which can be mathematically described by a resolution function $g(z,z')$,[48,49] as visualized in Figure 4.5. Any infinitesimal thin segment of the true concentration–depth distribution $c(z')$ is apparently broadened and the sum of all these contribu-

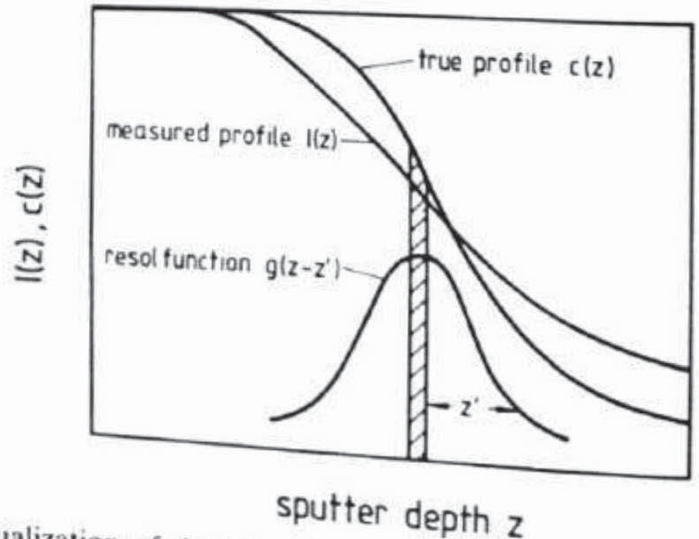


Figure 4.5 Visualization of the convolution process during sputter profiling: each thin layer of the true profile $c(z)$ is broadened into $g(z - z')$. The measured profile $I(z)$ is obtained by summation of the contributions from all such layers. (After Ho and Lewis[48])

tions at a sputtered depth z gives the normalized signal intensity at that depth when the resolution function is normalized:

$$\int_{-\infty}^{+\infty} g(z - z')\,dz' = 1 \tag{4.13}$$

The normalized intensity $I(z)$ can be expressed as a convolution integral:[3,48,49]

$$I_i(z) = \int_{-\infty}^{+\infty} c_i(z')g(z - z')\,dz' \tag{4.14}$$

If the resolution function $g(z - z')$ is known, the integral equation (4.14) can be deconvoluted and the true profile $c_i(z')$ is then determined. The procedures necessary for this step are outlined in Section 4.6.

It should be noted, that equation (4.14) also gives the convolution integral for quantitative AES/XPS, if $g(z - z') = \exp[-(z - z')/\lambda_i]$ (equations 4.5 and 4.6). In this case the solution is equation (4.9) and the shape of the resolution function for thin sections of a sample is similar to the curve depicted in Figure 4.4.

In many experiments of sputter profiling through sharp interfaces it has been observed that the measured profile can be approximated by an error function.[1-7,26,39,48-52] This gives a Gaussian function for $g(z - z')$, which is defined by a single parameter, the standard deviation σ. The most common definition of the depth resolution is $\Delta z = 2\sigma$, which corresponds to the difference of the depth coordinate z between 84 and 16 per cent of the intensity change at an interface.

The definition of depth resolution is shown in Figure 4.6(a), together with

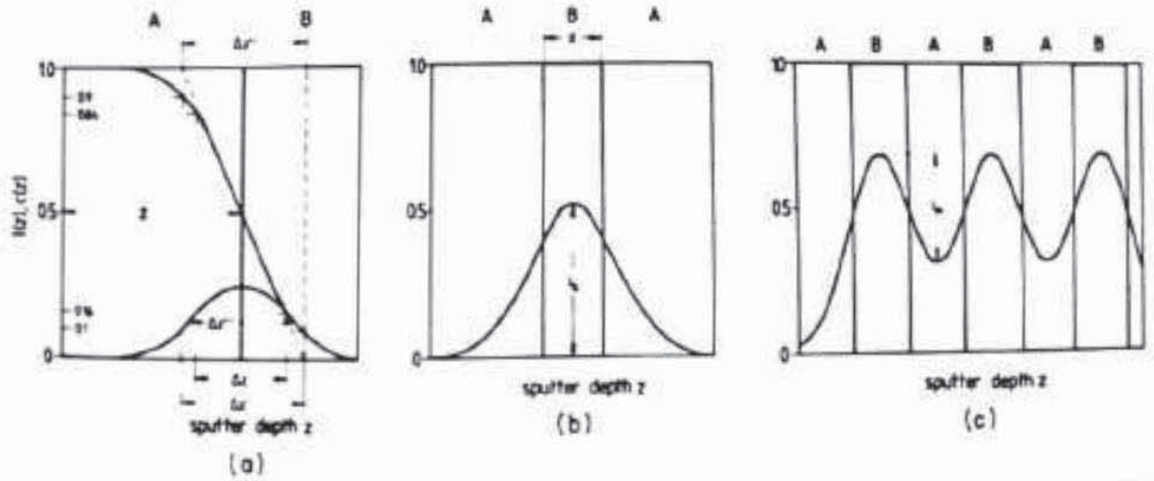


Figure 4.6 Determination of the depth resolution Δz for The Gaussian resolution function approximation: (a) step function true profile, $\Delta z = 2\sigma$ (0.84–0.16), $\Delta z'$ (0.9–0.1) = 2.564σ, $\Delta z''$ (inverse maximum slope) = 2.507σ. $\Delta z'''$ (FWHM of resolution function) = 2.355σ;[5,53] (b) single sandwich layer A/B/A where Δz is determined by I_i/I_0 according to equation (4.15a);[3,51] (c) multilayer sandwich structure A/B/A/B where Δz is determined by I_m/I_0 according to equation (4.15b)[3]

other definitions sometimes used. The relative depth resolution is defined by $\Delta z/z$ where z is the depth coordinate at the interface. For constant sputtering rate $\dot{z}$, the relative depth resolution corresponds to the respective value in sputtering time t coordinates, $\Delta z/z = \Delta t/t$.

The depth resolution can either be experimentally determined or estimated by calculations based on appropriate models.[3]

4.4.1.2 *Measurement of depth resolution*

The resolution function is directly obtained by sputtering through a very thin sandwich layer as visualized in Figure 4.5. A more frequently applied method is the sputtering through a rectangular profile and take the derivative.[3,48,49,56] In the Gaussian approximation a quantitative measure of the depth resolution Δz can be directly taken from the sputtering profile, as indicated in Figure 4.6(a).

For the case of a thin sandwich layer of thickness d, it has been shown[51] that the decrease of the maximum intensity I_0 (for ideal profiling) to a value I_s (Figure 4.6b) determines the depth resolution:

$$\frac{I_s}{I_0} = \operatorname{erf}\left(\frac{d}{\sqrt{2}\,\Delta z}\right) \tag{4.15a}$$

For multilayer sandwich structures (e.g. A/B/A/B/ . . .) with constant single layer thickness d a similar expression can be derived, which relates the amp-

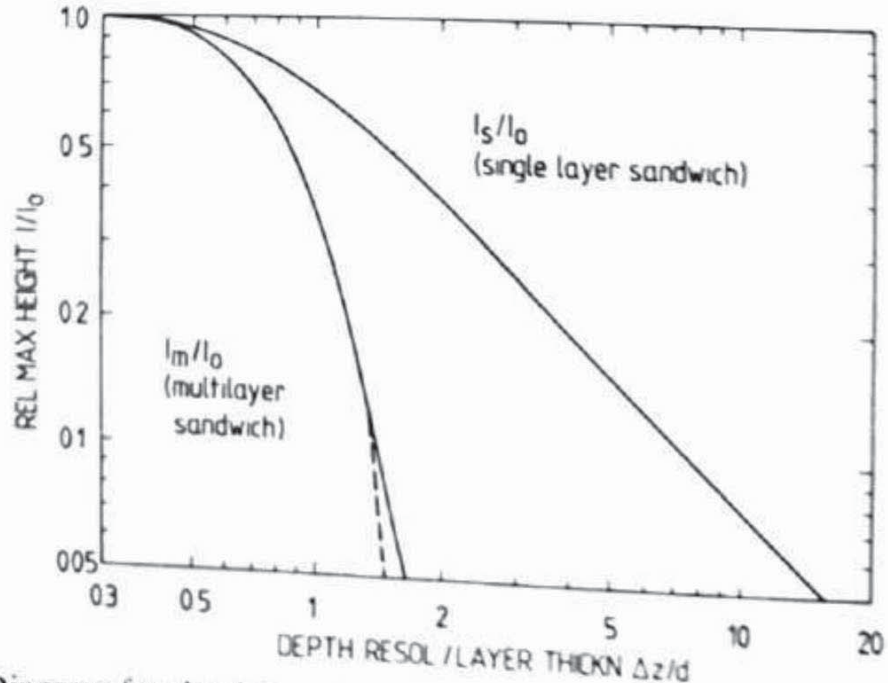


Figure 4.7 Diagram for the determination of Δz of single and multilayer structures of known thickness d shown in Figure 4.6(b) and (c) according to equations (4.15a,b). The approximations for the superposition of the two adjacent layers (——) and of four adjacent layers (– – –) is shown for the case of Figure 4.6(c)