Application Note: 31014

Angle Resolved XPS

Key Words

- Surface Analysis
- Depth Profiles
- Thickness Measurement
- Relative Depth Plots

Introduction

Using angle resolved XPS (ARXPS), it is possible to characterize ultra-thin films without sputtering. In most cases, ARXPS can be considered to be a non-destructive technique. It therefore has the potential to probe sub-surface chemical states that would be destroyed by sputtering. From ARXPS data, it is possible to calculate the thickness and depth of ultra-thin layers and to construct a concentration depth profile.

The technique relies upon the fact that the information depth of XPS is < 10 nm, depending upon the material and the kinetic energy of the electron being measured.

This document provides a brief description of the models used by Thermo Fisher Scientific in the treatment of ARXPS data.

It begins with the mathematical basis for the measurement of layer thickness using ARXPS. The equations associated with the Beer-Lambert law and its application to the measurement of single and multiple overlayer thickness are described. Complicating features such as angular asymmetry and elastic scattering are also discussed.

Methods for the generation of concentration depth profiles are outlined and examples given. In this document all emission angles are measured with respect to the surface normal.

Inelastic Mean Free Path and Attenuation Length

There are two important quantities in ARXPS, the inelastic mean free path (IMFP) and the attenuation length (AL). These terms are related and, in the past, were used interchangeably but it is important that the difference between these terms should be understood.

Inelastic Mean Free Path

This is the average distance an electron travels between successive inelastic collisions. It should be understood that an electron may undergo elastic collisions between inelastic events and so the trajectory between the events may not be a straight line.

Attenuation Length

This general term can be applied to any radiation passing through any material. Essentially, it is a measure of the transparency of the material to the radiation. There is a quantity, μ, the attenuation coefficient, which describes the proportion of the radiation removed as it passes through a thickness Δx of the material. This is defined as:

$$\mu = \lim_{\Delta x \to 0} \frac{1-\left(I/I_0\right)}{\Delta x}$$

Where I_0 and I are the intensities of the incident and emergent radiation respectively.

The reciprocal of this term is known as the attenuation length, λ:

$$\lambda = 1/\mu$$

In many cases this leads to an equation of the general form:

$$I/I_0 = \exp(-\Delta x/\lambda)$$

This terminology can be applied to XPS and AES where electrons of a given energy are considered as they pass through a material.

If the intensity of the electron flux in a given material is measured in a particular direction, then the intensity is affected by both elastic and inelastic collisions. Only electrons which have not been scattered or which have been scattered elastically can contribute to the intensity of a photoelectron peak. Under these conditions, the IMFP is a constant term but AL can change with angle, as will be seen later. It is, however, the attenuation length that must be used in equations similar to that shown above.

Symbols

In the derivation of the equations required for calculating layer thickness, the following symbols are used:

I_A= The intensity of the photoelectron signal from element A in specific chemical state and from a specific transition.

I^{∞}_A = The intensity of the photoelectron signal from a element A in a specific chemical state and from a specific transition from a thick layer of material A.

λ = The attenuation length for a photoelectron emitted within a uniform material. The transition must be specified.

$\lambda_{B,A}$ = Attenuation length in material A of a photoelectron emitted from material B. Note that the element and its chemical state must be specified as well as the transition involved.

Thermo SCIENTIFIC

Beer-Lambert Law

Consider a layer of material of thickness dt at a depth t, Figure 1.

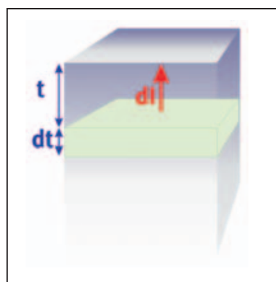

Figure 1: A layer of material, thickness dt, at a depth t emits photoelectrons with an intensity dI in a direction parallel to the surface normal.

The photoelectron intensity emitted from this layer (dI) in a direction parallel to the surface normal is

$$dI = Cdt$$

Where C is a constant involving X-ray flux density, sensitivity factors, geometric factors etc. Using the Beer-Lambert law the intensity reaching the surface from this thin layer parallel to the surface normal will be:

$$dI = C\exp(-t/\lambda)dt \ \ldots \ (1)$$

Integrating between zero and infinity, this becomes

$$I = C\lambda$$

To simplify, let $C\lambda = I^{\infty}$

Integrating equation (1) between t and infinity:

$$I = I^{\infty}\exp(-t/\lambda) \ \ldots \ (2)$$

where I is the photoelectron signal coming from all depths greater than t. This equation assumes emission at an angle normal to the surface.

The attenuation length for electrons varies with their kinetic energy, as shown in Figure 2.[1]

It will be seen from Figure 2 that the attenuation length increases with increasing kinetic energy in the energy range of interest in XPS (above a few tens of electron volts). For elemental silicon, the attenuation length of the Si 2p electron emitted due to Al Kα radiation is about 2.8 nm. Figure 3 shows a plot of intensity of a photoelectron peak as a function of depth for silicon.

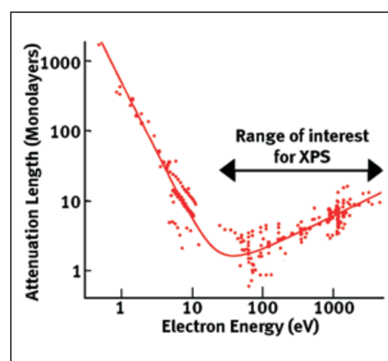

Figure 2: Attenuation length as a function of kinetic energy. Each data point represents a different element or transition.

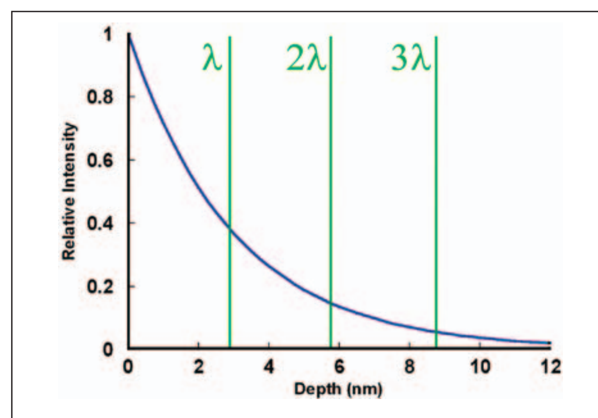

Figure 3: Relative intensity as a function of depth for Si 2p electrons emitted from silicon as a result of Al Kα radiation.

Figure 3 shows that, by considering electrons that emerge parallel to the sample normal, about 65% of the signal in electron spectroscopy will emanate from a depth of less than λ, 85% from a depth of less than 2λ, and 95% from a depth of less than 3λ.

Collection Angle

If electrons are collected at angles other than 0° with respect to the surface normal, these depths are decreased by a factor of cosθ, as can be seen in Figure 4. Equation (2) above now becomes

$$I = I^{\infty} \exp(-d/\lambda\cos\theta) \quad (3)$$

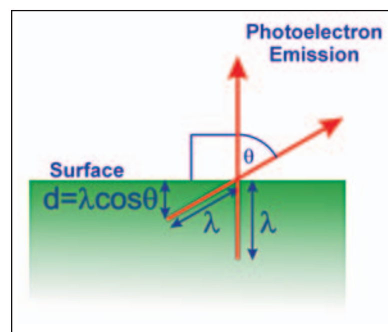

Figure 4: Greater surface specificity is achieved by collecting electrons at more grazing emission angles

If XPS signal is collected over a range of angles from near normal emission to near grazing emission then the analysis depth changes also. Figure 5 shows schematically the analysis of a thin metal oxide on a metal substrate. In this example, XPS data are collected from the metal at two angles, near normal (the “bulk angle”) and near grazing (the “surface angle”). Near normal emission produces a spectrum in which the metal peak dominates while the oxide peak dominates in the spectrum from the grazing emission.

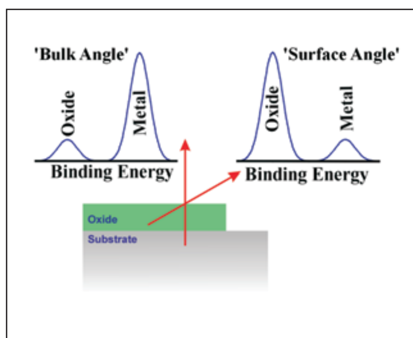

Figure 5: An illustration of the analysis of a thin metal oxide on a metal. The diagrammatic spectra show the effect of the collection angle on the elemental and oxide peaks of the metal.

This is the basis for angle resolved XPS measurements.

An Example of Angle Resolved XPS

In conventional ARXPS, the angular acceptance range of the spectrometer is selected by the user to provide the required angular resolution. Clearly, there must be a compromise between angular resolution and sensitivity (acquisition time). Spectra are then collected at each of a number of emission angles by tilting the specimen. An example of this type of experiment is shown in Figure 6 and Figure 7 for a sample of GaAs that has a thin oxide layer at its surface.

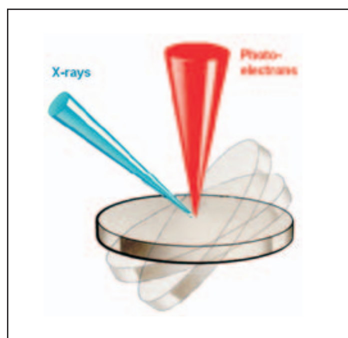

Figure 6: Conventional ARXPS, using sample tilting

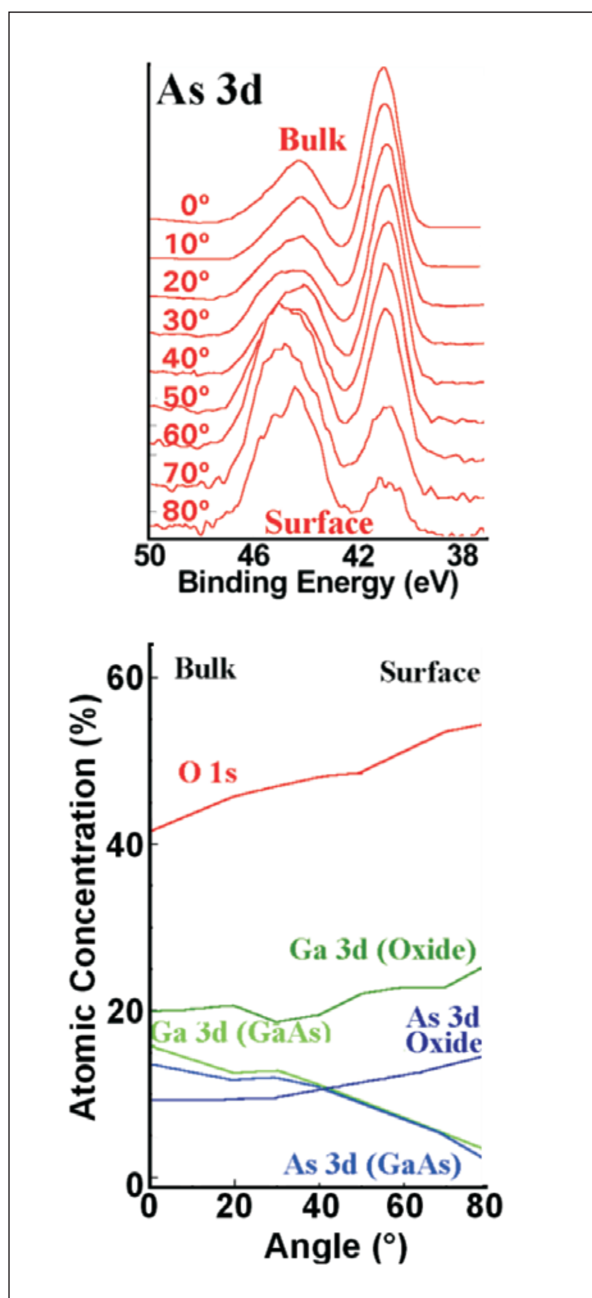

Figure 7: ARXPS experiment acquired by tilting the specimen. In this case, the specimen is gallium arsenide.

It is clear from the montage of As 3d spectra that the oxide peak (at higher binding energy) is dominant at the surface whereas the peak due to arsenic in the form of GaAs (at lower binding energy) is more dominant at near normal analysis angles. This phenomenon is repeated in the gallium spectra, as can be seen from the atomic concentration curves in Figure 7.

Parallel Acquisition ARXPS

On Thermo Scientific Theta Probe and Theta 300, angle resolved data may be collected over a 60° range of angles in parallel, without the need to tilt the sample. This is accomplished using the angle resolving lens and a 2-dimensional detector. A description of the lens operation is given in the application note AN31003.

Comparison of Conventional and Parallel Acquisition ARXPS

Conventional ARXPS, as illustrated in Figure 6, involves tilting the sample. Using parallel acquisition of angular data without tilting the sample, which is possible using Theta Probe or Theta 300 has a number of advantages.

1. ARXPS can be applied to large samples. It would be difficult to tilt a large sample in a typical XPS system, especially if data is required from a region near the edge of the sample. Theta 300 is therefore the only instrument capable of performing ARXPS measurements on complete semiconductor wafers.
2. Small area ARXPS is possible on Theta Probe or Theta 300 but would be very difficult using conventional methods:

   a) At all angles, the analysis position would have to be accurately aligned with the feature to be analyzed. This is difficult, especially if the analysis position is at some distance from the tilt axis and the required analysis area is very small.

   b) The analysis area changes as a function of angle, as can be seen in Figure 8. A worst case occurs when the transfer lens is used to define the analysis area. Using parallel angle acquisition, the analysis area and position is completely independent of the emission angle.
3. If an insulating sample is tilted, the required charge compensation conditions also change. Changes in peak position or shape may then be due to changes in the efficiency of charge compensation. Using parallel acquisition of angular data, the compensation conditions are the same for all angles and any changes in the spectra as a function of angle must reflect real chemical differences.

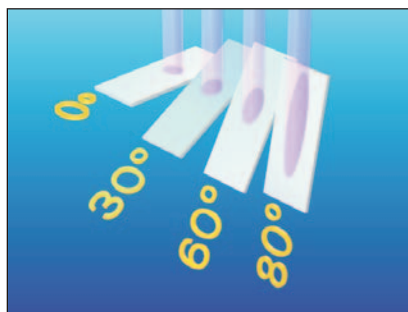

Figure 8: The analysis area changes as a function of angle, especially when using lens-defined small area analysis.

Determination of Structure and Composition

The Relative Depth Plot

This simple method is useful for the examination of samples having more than one layer on the substrate. The method produces a chart on which the order of the elements with depth is clearly shown. It does not provide a thickness or depth measurement but the positions of the species on the relative depth plot are related to the "average" depth of that species.

The plot is constructed for each species by taking the logarithm of the ratio of the peak area at near grazing emission angle to that at near normal emission. Figure 9 shows an example of a relative depth plot from silicon oxynitride on silicon. Carbon, a surface contaminant, can be seen nearest to the surface while the elemental silicon is deepest in the structure, as expected. In addition, it can be seen that nitrogen exists in two chemical states each having a different depth distribution.

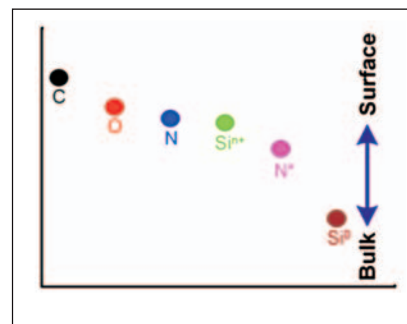

Figure 9: Relative depth plot from a silicon oxynitride film on silicon.

The relative depth plot has the advantage of being independent of any model and does not require the knowledge of the physical constants for the material. It can show, for example, the change in position of a species due to some form of treatment, such as annealing.

Two Layer Material

ARXPS can be used quantitatively to measure layer thickness; the background to the method will be described in this section.

Consider a thin layer, thickness d, of material A on a substrate B, Figure 10.

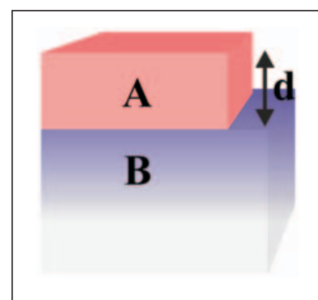

Figure 10: Thin layer, thickness d, of material A on a substrate B.

To obtain an expression for the signal from A, equation (1) must be integrated between 0 and d and becomes:

$$I_A = I^{\infty}_A \, [1\text{-}\exp(-d/\lambda_{A,A}\cos\theta)]$$

The signal from B arriving at the B-A interface is I^{∞}_B, assuming layer B is thick in comparison with λ_{BB}. This signal is then attenuated by passing though layer A. The signal emerging is therefore given by:

$$I_B = I^{\infty}_B \exp(-d/\lambda_{B,A}\cos\theta)$$

Note that the term $\lambda_{B,A}$ is the attenuation length in layer A for electrons emitted from layer B.

Taking the ratio of these signals:

$$\frac{I_A}{I_B} = R = R^{\infty}\frac{\left[1-\exp\left(-d/\lambda_{A,A}\cos\theta\right)\right]}{\exp\left(-d/\lambda_{B,A}\cos\theta\right)} \quad \text{... (4)}$$

Where $R^{\infty} = I^{\infty}_A / I^{\infty}_B$.

If $\lambda_{A,A} = \lambda_{B,A} = \lambda_A$ then this equation can be simplified to

$$R = R^{\infty}\left[\exp\left(\frac{d}{\lambda_A \cos\theta}\right) - 1\right]$$

Rearranging and taking the natural logarithm

$$\ln[1+R/R^{\infty}] = d/(\lambda_A \cos\theta) \quad \text{... (5)}$$

Plotting the left-hand side of this equation against $1/\cos\theta$ should then produce a straight line whose gradient is equal to d/λ_A.

The simple form of the equation (5) can only be used if $\lambda_A \approx \lambda_B$. This is true if the electrons detected from layers A and B have approximately the same energy (e.g. both are emitted from Si 2p). If this is not the case, then the more rigorous equation (4) should be used.

Figure 11 shows the required plot for a set of silicon dioxide layers of differing thickness on silicon. The thickness shown against each of the lines was derived from the ellipsometry measurements.

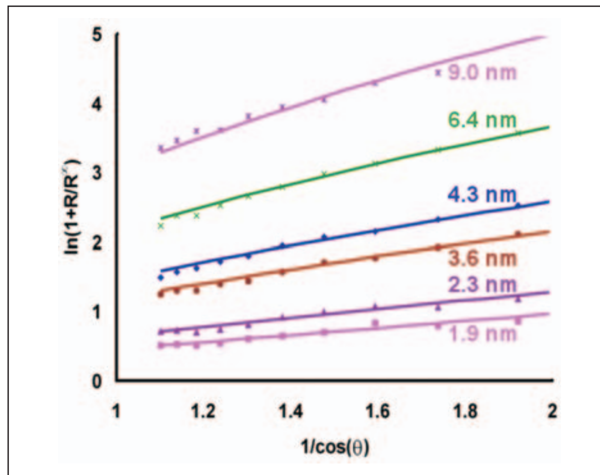

Figure 11: Calibration graphs for SiO_2 layers on Si.

In this Figure, the value for R^{∞} was taken to be 0.933 (see later) and, using $\lambda_{Si,SiO_2} = 3.6$ nm, the thickness of each layer can be calculated from the gradient of the lines. The thickness obtained for each layer is shown in Figure 12 where it is compared with the values obtained by ellipsometry.

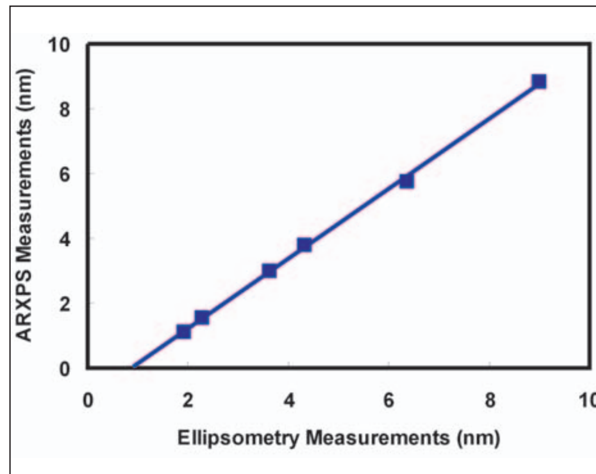

Figure 12: Comparison of the layer thickness measured by ARXPS with that measured using ellipsometry.

As can be seen from Figure 12, there is excellent linearity in this data and the gradient is close to unity (gradient = 1.08) but there is an intercept on the ellipsometry axis at 0.84 nm. This is due to the inclusion in the ellipsometry measurements of the airborne molecular contamination (AMC). The contamination layer, which is inevitably present, does not affect the thickness measured using ARXPS. Indeed, making certain assumptions, the thickness of the AMC layer can be determined using ARXPS.

The intercept shown in Figure 12 has also been observed by other workers.[2]

The Value for R^{∞}

The value for R^{∞} is the ratio of the intensities of the appropriate peaks from thick samples of the materials. In this context, 'thick' means greater than about 100 nm.

The values for the individual intensities will depend upon X-ray flux density, sensitivity factors, atomic number densities, asymmetry parameters, etc. For SiO_2 on Si everything should cancel (assuming the Si 2p peaks are used for both materials) except for the atomic number densities and the attenuation lengths in the two materials. This means that

$$R^{\infty} = \frac{N_{Si,SiO_2}\lambda_{Si,SiO_2}}{N_{Si,Si}\lambda_{Si,Si}}$$

Where $N_{X,Y}$ is the atomic number density (atoms per unit volume) of the element x in the material y. Note that R^{∞} contains an additional attenuation length term $\lambda_{Si,Si}$ (namely that, in elemental silicon, for electrons emitted from elemental silicon). The ratio of the atomic number densities is given by:

$$\frac{N_{Si,SiO_2}}{N_{Si,Si}} = \frac{\rho_{SiO_2} F_{Si}}{\rho_{Si} F_{SiO_2}} \quad \text{...(7)}$$

Where ρ_x is the density (mass per unit volume) of material x and F_x is the formula weight of x. When there is other than one atom of an element represented by the formula, then the formula weight should be divided by the number of those atoms present. For example, if the number density of oxygen atoms in silicon dioxide is required then F_{SiO2} should be divided by 2 to give the formula weight for 1 mole of O atoms. For SiO_2 on Si the values of these parameters are:

ρ_{Si} = 2.33 g cm^{-3}

ρ_{SiO_2}= 2.2 g cm^{-3}

F_{Si} = 28.09

F_{SiO_2}= 60.09

$\lambda_{Si,Si} \approx$ 2.9 nm (Si 2p)[3]

$\lambda_{Si,SiO_2} \approx$ 3.5 nm (Si 2p)

R^{∞} can be calculated for silicon dioxide on silicon:

R^{∞} = (2.2 x 28.09 x 3.5)/(2.33 x 60.09 x 2.9)= 0.53

R^{∞} can also be determined experimentally by measuring the intensities of thick (> 100 nm) layers of the materials. If this approach is used, however, care must be taken to ensure that the measurements are made under exactly the same conditions. An additional problem here is that a layer of contamination at the surface will attenuate the signal and unless the thickness of such a layer is the same for both materials, it will produce an erroneous value for R^{∞}. The contamination should not be removed by sputtering because this is known to reduce oxide layers by preferential sputtering and therefore provide an incorrect value for the intensity.

The experimentally determined values for R^{∞} are usually rather higher than the calculated, close to 0.93. The reason for this is not fully understood. It is partly because the intensity of the Si 2p peak from elemental silicon is reduced by plasmon excitation. This is an inelastic scattering process and occurs within elemental silicon to a greater extent than within silicon dioxide.[4]

Attenuation Length

If the attenuation length is known it can be used in the overlayer thickness calculator present as part of the *Avantage* data system.

If the value is unknown, the inelastic mean free path (IMFP) can be calculated using the method of Tanuma, Penn and Powell[5] (the same method that is used to calculate the IMFP data provided in reference 3). To convert the IMFP to attenuation length, the method described in reference 8 has been used.

Assumptions

A number of assumptions are made in the derivations above, the major ones are summarized here.

1. X-ray intensity is essentially unattenuated by passing through the analyzed volume.
2. Atom densities are constant within each of the materials considered.
3. Electron attenuation lengths are constant within each layer.
4. Attenuation lengths are independent of the direction of the analysis.
5. Elastic scattering is ignored.
6. X-ray photoelectron diffraction effects are ignored.
7. The specimen has a flat surface
8. Layers are continuous (i.e. there is no island structure).

Elastic Scattering

In the formulae derived above, it was assumed that there is no elastic scattering taking place and perturbing the results. This is not strictly valid under all conditions. The data in Figure 11 show only those measurements taken at angles of less than 60° from the surface normal. Figure 13 shows data from the same samples collected over a wider angular range.

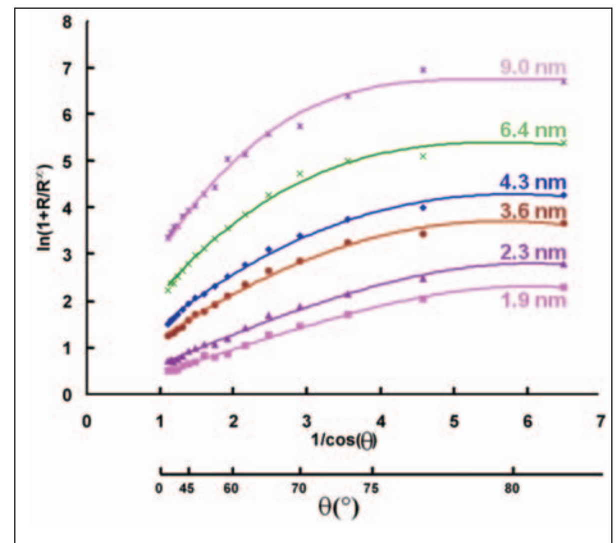

Figure 13: Data shown in Figure 11 but taken from a wider range of angles

In this diagram, it can be seen that the graphs become non-linear at large angles and the angle at which the non-linearity becomes apparent depends upon the thickness of the layer.

The non linearity is due to elastic scattering which causes the signal from the substrate to be larger than would be expected for a given layer thickness when the signal is collected at large angles. The origin of the effect is illustrated in Figure 14. This shows how electrons whose origin is deep within the sample can emerge from the surface at large emission angles.

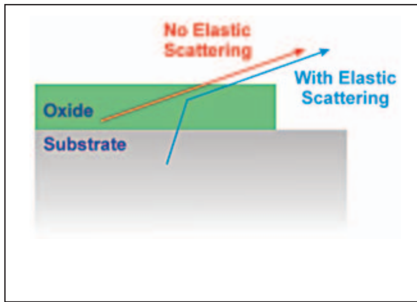

Figure 14: The mechanism by which the substrate signal is enhanced by elastic scattering.

Elastic scattering has the effect of increasing the attenuation length at large angles. These values have been calculated for Si 2p electrons in SiO_2 using the NIST Standard Reference Database 82 and are shown graphically in Figure 15.

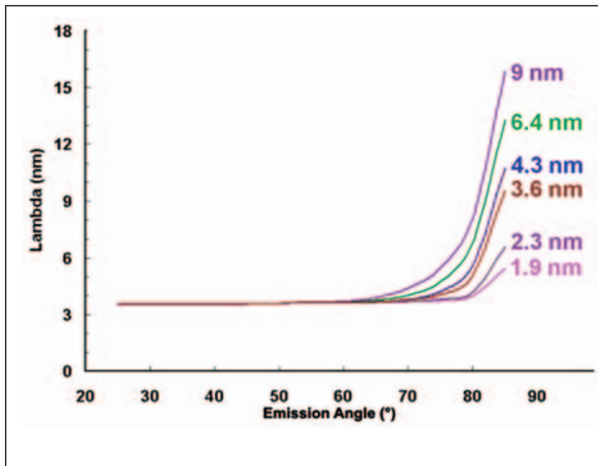

Figure 15: Values for the attenuation length of Si 2p electrons in SiO_2, as a function of layer thickness for a series of silicon oxide layers on silicon.

This is a very clear illustration of how attenuation lengths are affected by both layer thickness and angle of emission. For a thick layer, it can be seen from Figure 15, angles above 60° should be avoided while for thinner layers a larger angular range can be used in the calculation of layer thickness.

Asymmetry Factor

The asymmetry factor affects the angular dependence of the photoemission from a given orbital. It is an atomic effect, not a solid state effect, which means that its magnitude is independent of the chemical environment of the atom. The same correction is applied to Si 2p from elemental silicon and oxidized silicon. However, Si 2p would have a different asymmetry factor from both Si 2s and Al 2p.

The magnitude of the effect upon XPS quantification depends upon the angle between the X-ray source and the emitted electrons. With Theta Probe, because of the large angular range of collected photoelectrons, the sensitivity factor is dependent upon the emission angle from the sample. There is a difference between ARXPS performed on Theta Probe and ARXPS performed on other types of instrument. On Thermo Scientific ESCALAB 250, for example, ARXPS measurements are made by tilting the sample and making an analysis at each tilt angle. The angle between the X-ray source and the emitted electrons is constant for all tilt angles and so the data do not require a correction for the asymmetry factor.

The asymmetry factor is given by the equation[6]

$$L(\gamma) = 1 + 1/2\beta\ (^{3}/_{2}\sin^{2}\gamma\text{-}1)\quad (8)$$

Where β (the asymmetry parameter) is a constant for a given subshell of a given atom. Values of β vary from about 0.4 to 2 (the latter being the theoretical value for all s orbitals). For materials and transitions which are likely to be important for Theta Probe, the range of values will be between ~1 and 2. The form of the above equation is illustrated in Figure 16.

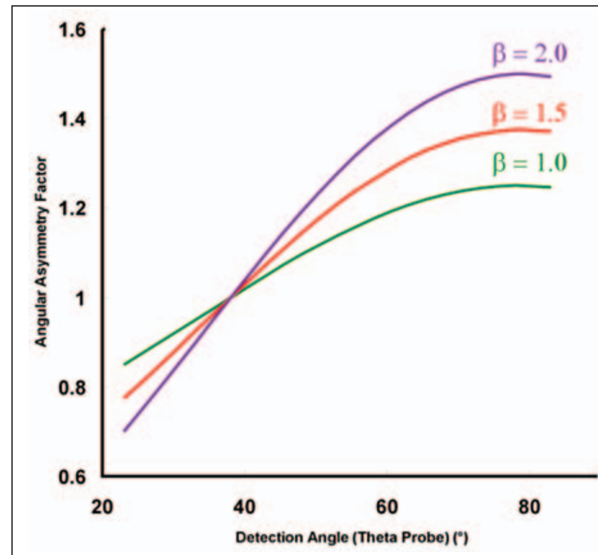

Figure 16: The angular asymmetry factor as a function of collection angle in Theta Probe for a range of β values

Strictly, the values for β are only applicable to photoelectron emission from gas phase atoms. In solids, the emission becomes more randomized by elastic scattering and β is therefore reduced. The extent of the reduction depends upon the average atomic number (Z) of the elements present in the material from which the electron is emitted.

According to one analysis, a form of the equation is[4,7]:

$$\beta^* = \beta(a - bZ + cZ^2) \quad ...(9)$$

(Where a = 0.781, b = 0.00514 and c = 0.000031)

Ratios of the intensities of the 2p:2s peaks from silicon and aluminum were plotted against the measured angle and compared with the calculated values. The results are shown in Figure 17.

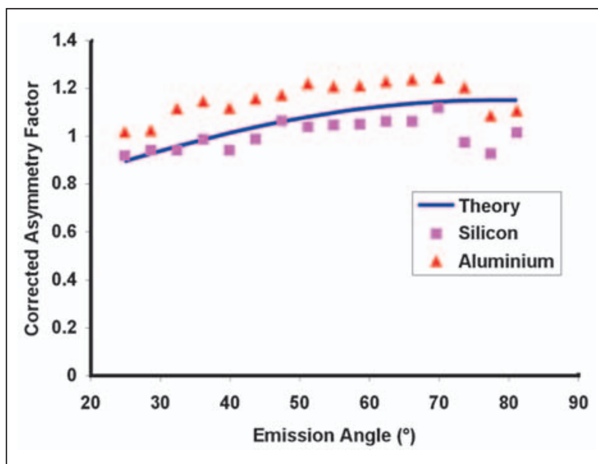

Figure 17: Theoretical and practical values for the ratio of peak intensities 2p/2s for aluminum and silicon. The theoretical values are calculated for measurements taken on Theta Probe.

The method used in the algorithms employed in the *Avantage* data system uses a more complex equation to make the correction to the asymmetry parameter. This uses an equation recommended by Seah and Gilmore,[8] based on detailed Monte-Carlo calculations by Jablonski.[9]

Effect of Asymmetry Factor on Measured Thickness

The following types of ARXPS measurements may be considered:

1. A metal oxide on its own metal (e.g. SiO_2 on Si). The thickness of such a layer is obtained by measuring ratio of the Si 2p signals from the oxide and from the metal. In this case, variations in sensitivity due to the asymmetry factor cancel out and no correction is required.
2. For other types of measurement, we determine the relative sensitivity, as a function of angle, from several different orbitals (2s, 2p, 3d etc) of each element in a standard sample. This may be accomplished using literature values for β, corrected using the equation of Seah and Gilmore.[8]

Multiple Overlayer Thickness Calculation

The single overlayer model cannot be used when more than one layer is present within the information depth of an ARXPS measurement. This is because some simplification has to be made regarding the substrate layer. For example, in the case of Al_2O_3 on SiO_2 on Si, when calculating the thickness of Al_2O_3 it is necessary to enter substrate material properties (e.g. density and band gap) into the single overlayer calculator. However, the substrate signal has a contribution from both SiO_2 and Si. If oxide properties are used in the calculation then the thickness of the layer will be underestimated but if Si data are used then the thickness will be overestimated. The magnitude of the error will depend upon the thickness of the silicon oxide layer.

For this reason, the Multi-overlayer calculator was developed for the *Avantage* data system. This calculator allows the thickness to be calculated of up to three layers on a substrate. Like the single overlayer calculator, it uses a database of the required physical properties and uses the TPP-2M[5] equation to calculate attenuation lengths.

The mathematics, again, rely on the manipulation of the Beer-Lambert equation. The ratio of the signal from layer i to that from the substrate is given by:

$$\frac{I_i}{I_s} = R_i^{\infty}\left(1 - \exp\frac{-d_i}{\lambda_{ii}\cos\theta}\right)\exp\left[\frac{1}{\cos\theta}\left(\sum_{j=1}^{j=n}\frac{d_j}{\lambda_{sj}} - \sum_{j=1}^{j=i-1}\frac{d_j}{\lambda_{ij}}\right)\right]$$

and the ratio of the signal from layer i to that from layer i+1 is given by:

$$\frac{I_i}{I_{i+1}} = \frac{R_i^{\infty}}{R_{i+1}^{0}}\frac{\left(1 - \exp\frac{-d_i}{\lambda_{ii}\cos\theta}\right)}{\left(1 - \exp\frac{-d_{i+1}}{\lambda_{i+1,i+1}\cos\theta}\right)}\exp\left[\frac{1}{\cos\theta}\left(\sum_{j=1}^{j=i}\frac{d_j}{\lambda_{i+1,j}} - \sum_{j=1}^{j=i-1}\frac{d_j}{\lambda_{ij}}\right)\right]$$

where d_i is the thickness of layer i and the other symbols have the definitions used earlier. These two equations are used for each of the layers in the stack and the thickness of each layer is calculated by least squares fitting of the experimental data to the equations.

Figure 18 shows the results of thickness measurements using the multi-overlayer thickness calculator on a series of samples of SiO_2 on Si samples which have had Al_2O_3 grown on them by atomic layer deposition (ALD). The thickness of each layer is shown as a function of the number of ALD cycles used in the preparation of the samples. The carbon layer is at the surface and is due to airborne molecular contamination (AMC).

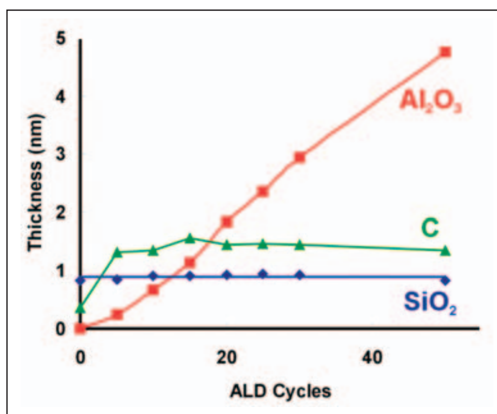

Figure 18: Thickness measurements using the multi-overlayer thickness calculator on a series of samples of SiO_2 on Si samples which have had Al_2O_3 grown on them by atomic layer deposition (ALD)

The graph shows that the growth rate of the Al_2O_3 layer is not uniform and that the thickness of the SiO_2 is independent of the number of ALD cycles. The thickness of the contamination layer depends only on the nature of the surface, it is thicker on Al_2O_3 than it is on SiO_2.

Reconstruction of Depth Profiles

When ARXPS data are collected from a number of angles, as with Theta Probe or Theta 300, it is, in principle, possible to use the data to obtain more than layer thickness measurements. The aim is to provide non-destructive depth profiles.

Unfortunately, there is no method for the direct transformation from a set of angle resolved XPS data to a concentration depth profile. The problem is compounded by the statistical noise on the data leading to uncertainty in the angle response curves. It is therefore necessary to use mathematical routines to produce the most probable depth profile. Various methods exist but the most promising for our purposes is based on Maximum Entropy (Max Ent) methods.

In this method, the sample is considered to consist of a number of "slabs" of material, each slab having uniform composition. Trial concentrations of each component material in each slab are then assumed and the ARXPS characteristics calculated. Further iterations then take place to find the "best fit". If least squares fitting to the data were used alone, noisy profiles would result due to over fitting of the angular data. Max Ent alone would produce profiles consisting of constant composition throughout. Both of these extremes would be unsatisfactory and some balance has to be found.

Maximum Entropy

The application of Maximum Entropy methods to ARXPS data has been described in detail elsewhere[10,11] and so only a brief outline of the method is described below.

Consider a sample from which a set of k angle resolved XPS spectra have been acquired. Each angle can be designated θ_k. If the sample consists of j slabs that contain i elements at concentration $c_{j,i}$ the integrated intensity can be expressed as follows:

$$I_i(\theta_k) = \sum_j I_{j,i}(\theta_k)$$

in which

$$I_{j,i}(\theta_k) = s_i c_{i,j} \exp\left[\frac{-dz\cos(\theta_k)}{\lambda_i}\right]^{j-1}$$

where dz is the thickness of each slab, s_i is the sensitivity factor for each element.

For each trial profile the sum of the squares of the errors (χ^2) is calculated according to the equation:

$$\chi^2 = \sum_k \frac{\left(I_k^{calc} - I_k^{obs}\right)^2}{\sigma_k^2}$$

Where σ_k is the standard deviation. This expression must now be minimized while maximizing the entropy term (S):

$$S = \sum_j \sum_i c_{j,i} - c_{j,i}^0 - c_{j,i}\log\left(\frac{c_{j,i}}{c_{j,i}^0}\right)$$

This can be done by maximizing the joint probability function (JPF):

$$Q = \alpha S - 0.5\chi^2$$

Where α is a regularizing constant. A large value for α would result in a profile which is too smooth while a small value for this term would result in the data being over fitted (i.e. fitted to the noise).

The iteration works by repeating this procedure for a range of different, trial profiles and searching for the largest value of Q.

With a good set of data from a suitable sample, this method can produce profiles that are consistent with the known structure of the material. It can do this without the imposition of any constraints.

Figure 18 shows an example of a profile calculated using the Maximum Entropy method.

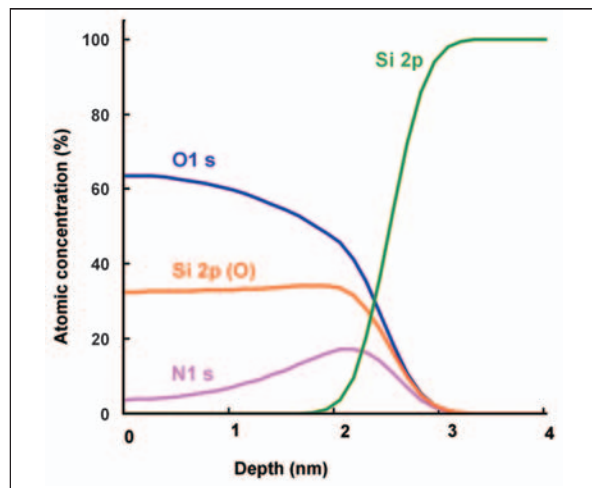

Figure 19: Example of a depth profile calculated using Maximum Entropy from ARXPS data. The sample was silicon oxynitride on silicon.

Profile Optimization

In the software we provide, the user has a choice of two methods for profile selection and optimization.

Method 1: A random profile is generated from which angle resolved XPS data is calculated. From this and the experimental data, the JPF calculated. Using the Powell method, the profile is adjusted to minimize the JPF, and improve the fit of the trial profile to the experimental data. The Powell method is an efficient iterative algorithm by which a local minimum in a multi-dimensional data set can be found.

It is likely that the minimum found by the procedure will be only a local minimum and so another trial profile is generated and optimized in the same way. This process continues until no further improvement in the profile is observed.

The disadvantage of this method is that it relies on the random selection of a profile that will, after Powell minimization, result in the global best fit to the data. There is no guarantee that this will occur nor any way of knowing when it has and thus automatically stopping the procedure

Method 2: This method uses a genetic algorithm (GA) to generate profiles. In this case, some number (normally 20) random profiles are generated and each is evaluated for the quality of its fit to the experimental data. Pairs of these profiles are allowed to 'breed' mathematically according to a set of rules. These rules determine the probability of a profile (or features from a given profile) appearing in a subsequent generation, depending upon the value of the JPF. Random "mutations" to these profiles are also introduced to help ensure that the global optimum profile is found. Several thousands of generations are typically required to achieve a satisfactory profile.

This method has the advantage that it provides continual improvement in every generation. Its disadvantage is that although it is good at finding the global rather than some local minimum, it does not necessarily find the optimum fit to this minimum, although it will be close. This disadvantage is eliminated by performing a Powell minimization after a sufficient number of GA generations.

Advantages of ARXPS

As can be seen from the above discussion, ARXPS is a technique that can provide information about near-surface layers to a depth of about 10 nm. As with conventional XPS, it also provides chemical state information. ARXPS can therefore provide information that is not available using other methods. Much of this falls into the category of non-destructive depth profiling.

When chemical state information is needed, sputtering should not be used because the required information may be destroyed. ARXPS allows measurements to be made on buried interfaces without the need to sputter the sample.

For quantitative information from near surface features, sputtering cannot be used. This is because there is a transient region of a few nanometers at the start of the sputtering process within which quantitative data is unreliable.

Atomic mixing causes depth resolution from sputtering methods to be insufficient, unless very low energy ions are used.

Techniques such as ellipsometry and spectroscopic ellipsometry are capable of providing film thickness information but they are incapable of providing chemical state information. Indeed, using these methods it is essential to know the composition of the films before their thickness can be determined.

In the past, ARXPS has not been used extensively because of the difficulty of performing the experiments. Now, using Theta Probe or Theta 300, most of the practical difficulties have been removed and ARXPS becomes viable as a routine analytical tool.

References:

1. P. Seah and W.A. Dench, Surface and Interface Analysis 1 (1979) 2
2. M.P. Seah et al, Surface and Interface Analysis, 36 (2004) 1269
3. Values for the attenuation lengths we calculated using the NIST Standard Reference Database 82. Practical values may differ slightly from those reported here.
4. M.P. Seah and S.J. Spencer, Surface and Interface Analysis 33 (2003) 515
5. S. Tanuma, C.J. Powell, and D.R. Penn, Surface and Interface Analysis, 1994, 21, 165
6. R.F. Reilman, A. Msezane and S.T. Mason, J. Electron Spectroscopy, 1976, 8, 389
7. A. Jablonski, Surface and Interface Analysis, 1989, 14, 659
8. M.P. Seah and I. Gilmore, Surface and Interface Analysis, 2001, 31, 835
9. A. Jablonski, Surface and Interface Analysis, 1995, 23, 29
10. A.K. Livesey and G.C. Smith, J. Electron Spectroscopy and Related Phenomena, 1994, 67, 439
11. R.L. Opila and J. Eng Jr., Progress in Surface Science, 2002, 69, 125

www.thermo.com/surfaceanalysis

In addition to these offices, Thermo Fisher Scientific maintains a network of representative organizations throughout the world.

Africa
+43 1 333 5034 127

Australia
+61 2 8844 9500

Austria
+43 1 333 50340

Belgium
+32 2 482 30 30

Canada
+1 800 530 8447

China
+86 10 8419 3588

Denmark
+45 70 23 62 60

Europe-Other
+43 1 333 5034 127

France
+33 1 60 92 48 00

Germany
+49 6103 408 1014

India
+91 22 6742 9434

Italy
+39 02 950 591

Japan
+81 45 453 9100

Latin America
+1 608 276 5659

Middle East
+43 1 333 5034 127

Netherlands
+31 76 579 55 55

South Africa
+27 11 570 1840

Spain
+34 914 845 965

Sweden/Norway/Finland
+46 8 556 468 00

Switzerland
+41 61 48784 00

UK
+44 1442 233555

USA
+1 800 532 4752

www.thermo.com

VG Systems Ltd. Trading as Thermo Fisher Scientific, East Grinstead, UK is ISO Certified.

AN31014_E 05/08M

©2008 Thermo Fisher Scientific Inc. All rights reserved. All trademarks are the property of Thermo Fisher Scientific Inc. and its subsidiaries.
Specifications, terms and pricing are subject to change. Not all products are available in all countries. Please consult your local sales representative for details.